Molecular origins of tissue vulnerability to aberrant aggregation in protein misfolding diseases

Rosie Freer

Darwin College

This dissertation is submitted for the degree of Doctor of Philosophy

Preface

This dissertation is the result of my own work and includes nothings which is the outcome of work done in collaboration except as specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as specified in the text. It does not exceed the prescribed work limit for the relevant Degree Committee.

Acknowledgements

To comprehensively thank all who have helped and inspired me over the past three years would require a separate tome of equal length to my thesis, and it has been my pleasure to weave these interactions together into this final document. Most notably, I would like to thank my colleagues, for adopting a strange British girl into their clan, despite my terrible pasta and potentially worse office-singing. Pietro Sormanni has been a source of expertise and support from my very first day at the Vendruscolo lab. Thanks to his frank advice, Pietro has been instrumental to steering my PhD on the best possible course, as well as encouraging the exploration of new horizons. Gabi Heller has been the best scientific co-pilot one could wish for, her drive and enthusiasm for her field are infectious. I would like to thank Madeline MacNamara, Cambridge would not have been the same without you. Not many people would fly back from the Maldives to spend New Years with my family in Birmingham, you are crazy, and I am so grateful to have you in my life. Same goes for you Cameron Watson, Alex Théron-Grimaldi, Krystina Warrington, and Katrina Power. I would like to thank Chris Dobson, our stimulating conversations have generated more ideas than I've been able to explore over my PhD and I know your curiosity and insight will continue to spark my scientific exploits for many years to come. I'd like to thank Echo Wu Williamson, who went above and beyond to support my organisation of the group retreat. To; Jack, Joe, Holly, Katie, Lucy, Jon, Suzie, Emily, Sarah and Martin, thank you for a wonderful three years. You are the best family a girl could wish for. Special mention must be given to Jon, who crafted all the 'brain' figures in this thesis. A coding wizard, I know he has incredible things ahead.

And finally, I acknowledge Michele Vendruscolo. Once or twice in a lifetime, one meets an individual of such intellect that they are able to see beyond that gaze of all others. They will express views that will baffle you, or which sound wrong. And as time unfolds, you find yourself in awe as their vision becomes apparent to you, too. I have had the complete privilege of working for three years with Michele, and almost all that I have achieved over my PhD I owe to him. Michele has empowered me to accomplish more than I ever imagined to be possible - his unfaltering confidence in my scientific ability, generosity with his time, and belief in academic freedom have been instrumental in the development of my work. More broadly, Michele shares my curiosity in the world, and perhaps it was our kindred spirit in this sense which has made our conversations so stimulating. I will always be indebted to Professor Vendruscolo, and hope that our work together will transcend my time at Cambridge.

Abstract

Neurodegenerative disorders, including Alzheimer's disease (AD) and Parkinson's disease (PD), are increasingly common in our ageing society, are remain incurable. A major obstacle encountered by researchers in their attempts to find effective therapies is represented by the current lack of understanding of the molecular origins of these disorders. It is becoming clear that, although the aggregation of specific proteins, including amyloid β (Aβ) and tau in AD and α-synuclein in PD, hallmark these disorders, such behaviour is a consequence of a wider, system-level disruption of protein homeostasis. In order to identify the genetic factors contributing to such a disruption, the transcriptional changes that occur during neurodegenerative disease progression have received considerable scientific attention in recent years. In our approach, we considered another hallmark of these diseases - their characteristic patterns of spreading across the brain - to identify the nature of the transcriptional signature which underlies tissue vulnerability to protein aggregation. By understanding why tissues succumb in their characteristic sequential pattern in neurodegenerative diseases, and why some tissues remain almost completely resistant throughout, we hoped to obtain insight into the molecular origins of these disorders. Our results show that the AD progression can be predicted from a transcriptional signature in healthy brains related to the protein aggregation homeostasis of Aβ, tau, and the wider proteome. We highlight a relationship between a specific subproteome at high risk of aggregation (formed by supersaturated proteins), and the vulnerability to neurodegenerative diseases. We thus identify an AD-specific supersaturated set of proteins - termed the metastable subproteome, whose expression in normal brains recapitulates the staging of AD, with more vulnerable tissues having higher metastable subproteome expression. We find evidence of these vulnerability signatures transcending the tissue level of interrogation, with cellular and subcellular analysis also showing elevated levels of proteins known and predicted to predispose the aberrant aggregation of Aβ and tau. These results characterise the key protein homeostasis pathways in the inception and progression of AD, and establish an approach with the potential to be applied to other protein misfolding diseases, in the brain and beyond.

Table of Contents

List of figures

List of tables

1 Introduction

1.1 Protein Misfolding

1.1.1 Routes to protein misfolding

Protein folding is the process by which a protein acquires its native, biologically active conformation[1]. The stochastic nature and the complexity of this process can result in a failure to reach a correctly folded state, with proteins misfolding and aggregating [1-7].Given the relatively frequent failures in the folding process (Figure 1) [2], a complex cellular machinery has evolved to intervene with nascent, folded and intermediate states of proteins to ensure their correct behaviour [3, 4]. This 'protein homeostasis system' is crucial to determine the correct conformations, concentrations and trafficking of proteins.

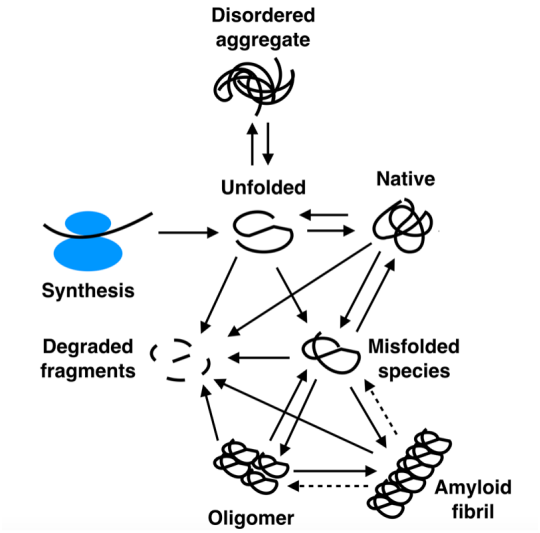

Figure 1.1: States accessible to proteins in the cellular environment. These states are complex and highly interconnected.

The hydrogen bonding propensity of the polypeptide backbone is common to all proteins and essential to enable their folding. However, this chemical property also affords polypeptide chains a ubiquitous potential to misfold and aggregate. There are many factors that influence the likelihood of protein misfolding, including those intrinsic to the amino acid sequences of proteins, their charge distribution [5] and the presence of specific hydrophilic-hydrophobic patterns [6]. Other traits that lead to an

elevated misfolding propensity can arise due to somatic or inherited mutations, or errors in the process of transcription or translation [7, 8]. These traits often destabilise the correct fold and act to favour the formation of misfolded and aggregated states. Misfolding can also occur due to factors associated with the cellular environment, including a failure of protein homeostasis machinery, errors in the post-translational modification or translation of proteins, or structural modifications due to environmental changes (such as pH) [7, 8].

1.1.2 Consequences of protein misfolding

In certain instances, a misfolded protein state can have a high tendency to self-associate. The amyloid state, an example of a highly organised protein assembly, results from the aggregation of misfolded proteins [9]. The kinetics of amyloid formation have been characterised to various extents for a number of proteins [10, 11]. This process is initiated by a nucleation step, which is generally energetically unfavourable and thus rate limiting [12]. Once critical nuclei have been generated, fibrillar extension contributes to the aggregation kinetics. This elongation step is typically kinetically and energetically favourable, because fibril ends act as a template for the addition of new monomeric units and fibrils are themselves more stable than the monomeric proteins at the typical cellular concentrations [13, 14]. Once formed, amyloid fibrils are more stable than even the native state [15, 16]. However, their formation is relatively rare because of high free energy barrier between the native and the aggregated states [17] and the presence of the protein homeostasis system. In addition to mature fibrils, a number of other structures have been described in association to protein aggregation, including oligomers, pores, annular structures, spherical micelles, and protofibrils [18, 19].

In a non-functional aggregated form, misfolded proteins can be resistant to clearance, and interfere with the capacity of affected cells to function normally. Recent studies have focused on identifying the aggregated protein species most damaging to cells, and there is now a consensus that soluble oligomers have the greatest in vivo toxicity [20, 21], although the corresponding mechanism has yet to be throughly elucidated. In

particular, Aβ oligomers have been shown to adversely affect synaptic structure and plasticity in neuronal cells [22]. A wide range of studies has indicated that a broadly similar process of cellular damage is elicited by soluble oligomers formed from other misfolded proteins. Both synthetic and and natural oligomers have been shown to induce apoptosis in cell cultures [23, 24], even at low concentrations, to block long term potentiation in brain cultures [25], and to impair synaptic plasticity and memory in animal models [26]. Protofibrils and other intermediate assemblies may also contribute to cell death [27], and mature fibrils can elicit toxicity in cultured cells [28], but only when at significantly higher concentrations than oligomers and protofibrils.

1.2 Protein misfolding regulation and dysregulation

1.2.1 Protein homeostasis

The conformations, concentrations, locations, and post-translational modifications of proteins are controlled by a complex cellular regulatory system, known as the protein homeostasis network (PHN) [29, 30]. In addition, this system ensures a balance between protein generation and degradation, allowing optimal cellular function whilst mitigating the risk of aggregation [30, 31]. Components of this network are expressed constitutively but at varying levels across tissues and cell types, and are often induced in response to a threat to protein homeostasis, for example the accumulation of misfolded proteins [32]. The PHN is composed of multiple molecular chaperone systems, quality control systems, and degradation pathways. Molecular chaperones are responsible for correct protein folding, for example by preventing or rescuing proteins (in particular nascent chains) from the formation of non-native interactions [4]. For example, during translation, highly hydrophobic portions of a polypeptide chain are exposed to the cellular environment and are vulnerable to misfolding or aggregation [33]. Molecular chaperones and co-chaperones, such as the heat shock proteins Hsp70 and Hsp40 act to ensure correct folding and avoidance of misfolding [4, 34]. The PHN performs many roles including the surveying of nascent chains for errors [35], and acting as folding checkpoints [36]. Degradation pathways often act in response to the quality control system to dispose of proteins which are damaged or

cannot be refolded into their native state [37]. These pathways have a key role in protein turnover and therefore protein levels in cells [38, 39].

1.2.2 Protein supersaturation

A protein is supersaturated when it is expressed beyond its limit of solubility. Under these conditions, the aggregated state is more stable than the native state, creating a high risk of aggregation.. It is increasingly recognised that proteins in the cellular environment are expressed at levels that exceed their critical concentrations [31]. These proteins are thus supersatured [40]. This aspect of protein behaviour is of central importance in understanding the molecular origins of neurodegenerative diseases, since it implies that protein aggregation poses an intrinsic and constant danger to cellular homeostasis [40-42].

A number of subproteomes have been suggested to influence tissue and brain specific risk to protein aberrant aggregation. These subproteomes include the primary proteins (protein aggregators) found in characteristic disease deposits [43], and the proteins in the protein homeostasis networks associated with protein aggregators [44]. In addition, one could consider the aggregation propensity of the wider proteome. An insoluble protein would provide a significant burden on the protein homeostasis system, allowing the most insoluble proteins to escape regulation and follow misfolding pathways towards aggregation [45]. Indeed this phenomena has been observed in *C. elegans* where the protein regulatory system was overwhelmed using aggregation-prone protein over-expression [46]. In response, discrete cytoplasmic aggregates were observed, and researchers further demonstrated that aggregate formation could be prevented using co-expression of the yeast heat shock chaperone Hsp104 [46]. This phenomenon has also been demonstrated in mammals; amyloidosis in mice can be induced by a diverse array of insoluble assemblies, including the yeast protein Sup35, and curli fibrils from the bacterium *E. coli* [47]. More generally, aggregation can affect a wide spectrum of proteins, as observed during ageing in *C. elegans,* where deposition was found to be widespread in the proteome [48]. These results can be rationalised by the observed anti-correlation between protein

aggregation rates and expression level, indicating that protein levels in the cell are at the limit of their solubility [31]. In this instance, one would expect a small perturbation in the PHN to result in widespread precipitation of proteins. This phenomenon is likely due to an evolutionary trade off between solubility and optimal functionality - as the vast majority of point mutations reduce protein stability [49].

In previous studies leading to the work that I present in this thesis, a supersaturation score was recently developed by researchers in our Centre to quantify the aggregation risk of specific proteins. This quantification allows the assessment of the burden on the protein homeostasis system from a subproteome, or the proteome as a whole [40]. This score is calculated using both experimentally measured and computationally predicted components - expression level and solubility respectively. The balance between these two factors determines the extent to which a protein is expressed beyond its solubility limit, and therefore provides a prediction of its risk of aberrant aggregation in the cell [40]. Proteins within the 5th percentile of this score are termed 'supersaturated'. Studies have found supersaturated proteins to be enriched in the co-aggregators found in amyloid plaques and neurofibrillary tangles, and the deposits found in ageing worms [40]. This result suggests that distinct stresses, including age and disease, can induce these proteins of high predicted risk to loose solubility in the cell when the aggregation threshold in a cell is reduced. It has also been suggested that since supersaturated proteins are highly vulnerable to aberrant aggregation [40], they could put a pressure on protein homeostasis and induce the misfolding and aggregation of other proteins that would otherwise be relatively soluble.

When considering protein homeostasis, supersaturation therefore presents another aspect of protein behaviour that needs to be highly regulated. By defining the most supersaturated proteins, we enable a more complete understanding of the factors which influence risk to aberrant aggregation.

1.3 Protein misfolding in disease

1.3.1 Characterising proteins which misfold in disease

Each protein misfolding disease is characterised by the deposition of a small specific subset of aggregating proteins [9]. The factors that determine which proteins will misfold and deposit in disease remain under intense investigation [9, 50, 51]. Although there is no known sequence or structural homology which unifies these molecules, an amyloidogenic propensity is common to all proteins. In addition, supersaturated proteins (see section 1.2.4) are enriched in the proteins which co-aggregate to form the disease-associated inclusions found in AD and PD [40].

Several theories exist for the pathogenic mechanism for protein misfolding diseases [52-55]. Some argue that protein aggregation is an artefact of neuronal death and not the cause, however there is much evidence to the contrary. Mutations in the gene that codes for tau are known to be associated with the incidence of various neurodegenerative diseases [56]. Transgenic animals expressing the mutant variants of misfolding proteins associated with the disease in humans develop neuropathological and clinical characteristics which reflect, at least in part, those found in the human pathology [57]. Perhaps the most compelling evidence of a crucial role for protein homeostasis in neurodegenerative diseases came from the discovery of a protective mutation for AD. The rare gene variant A673T has been found to decrease Aβ deposition in the brains of those who have developed AD. The mutation occurs close to the site of enzymatic cleavage necessary to form the aggregation prone polypeptide. Consequently, lower levels of the Aβ peptide were produced - giving indication that lower levels of key aggregation prone proteins have the potential to protect neural tissue against neurodegenerative disease. Given that another mutation at an identical site results in an Aβ variant more prone to misfolding and pathological aggregation, it is possible that this site also acts to influence the stability and solubility of Aβ, modulating pathological aggregation [58].

1.3.2 Mechanisms of misfolded protein toxicity

As discussed in section 1.1.2, oligomeric species have recently been shown to be the most toxic assemblies - several studies have suggested that levels of soluble Aβ oligomers show a better correlation to the presence and degree of cognitive deficits than the occurrence of plaque (large insoluble amyloid deposits) [59]. As is typical for the rapidly evolving field of protein misfolding diseases, the mechanisms by which this toxicity is conferred are under active study.

Evidence of oligomeric toxicity has been found in a number of neurodegenerative diseases, and in addition to much studied Aβ assemblies, there is evidence of the PD protein α-synuclein forming pathogenic oligomers [20]. Further, tau cytotoxicity has been associated to oligomeric structures [60], which arise prior to the paired helical filaments found in AD-characteristic tangles. However, it is unknown whether the mechanism of oligomer toxicity is common for different proteins and due to oligomer biophysical characteristics, or instead due to specific interactions, perhaps with a particular cell receptors or target proteins.

Researchers have suggested that specific neuronal or glial receptors could be conferring a response to Aβ oligomers. Work done by Malaplate-Armand et al. has suggested that extracellular oligomers could activate a signal transduction pathway that ultimately leads to cellular apoptosis [61]. A number of studies have also proposed that Aβ oligomers could stimulate a pro-inflammatory cascade through interaction with local astrocytes and glia [62]. Alternatively, oligomers may be indiscriminately affecting the function of diverse receptor channel proteins on neuronal cell membranes.

It has also been proposed that oligomers could be acting to drive aggregation of other disease-associated assemblies. This process could occur via a targeted and specific mechanism; for example Aβ oligomers could be inducing the hyperphosphorylation of tau [63], thereby driving its aggregation into paired helical filaments. However, the triggering of aberrant aggregation could be more widespread throughout the

proteome, damaging cells through a loss of function. The sequestration of proteins within neurons, in particular molecular chaperones, could perturb key cellular processes. These processes include protein synthesis, folding, trafficking and degradation, thereby inhibiting protein homeostasis and escalating the aggregation process [64]. In the protein misfolding disease amyotropic lateral sclerosis (ALS), gene regulatory proteins co-aggregate in the cytoplasm, preventing them from performing their designated roles in the nucleus [65]. In cystic fibrosis, pathology arises from the misfolding and defective trafficking of the primary disease-associated protein, CFTR [66-68]. Neurons could be particularly vulnerable to this threat, due to their limited renewal capacity [69].

Since oligomeric species are often highly hydrophobic, they could cause disruption to cell membranes by physically puncturing and permeabilising them [70]. It has been suggested that oligomers within a low molecular weight range would be most potent in causing damage [71]. The downstream consequences of oligomeric damage are somewhat clearer, as there is now a body of evidence to demonstrate that Aβ oligomers are able to inhibit the neuronal process of long term potentiation (LTP) long-term potentiation in vivo [72, 73], known to be crucial in memory formation. When derived from cells, these oligomers were able elicit a physiological impact at concentrations comparable to those found in human CSF [74]. Furthermore, LTP can be restored by treatment with treatment using anti-Aβ antibodies in vivo [74], when introduced using either passive diffusion or vaccination.

1.3.3 Factors of influence in protein misfolding inception in disease

Because misfolded states are permissive to pathways that often end in aggregation, aberrant aggregation in a cell can be indicative of a protein homeostasis failure [75]. In the event of protein self-association, oligomeric species may cause catastrophic damage in affected cells, and disease manifests [70]. Therefore, the likelihood that a protein misfolding disease will take hold in a given tissue is elevated by factors which promote the self-assembly of proteins specific to that disease.

In AD, plaques and tangles, primarily composed of the proteins Aβ and tau respectively, and have a specific spatiotemporal pattern of deposition in the brain.Indeed, tangle deposition has been found to closely follow the pattern of neuronal death [76]. Initiating in the parahippocampal gyrus, the disease spreads out to a number of cortical tissues, whilst for example the cerebellum is almost completely spared. This differential disease resistance suggests the existence an innate tissue vulnerability to the protein aggregation specific to AD. Differential vulnerability is also seen at the whole-brain level - AD onset can occur in individuals across a wide age range [77].

Two factors must be considered when determining brain tissue resistance to protein homeostasis failure. The first, being the contribution of brain-wide risk factors - more specifically genetic predisposition, age, and any lifestyle risk factors such as head injury. The second, a tissue intrinsic risk, is influenced by a tissue's specific proteomic signature, cellular resistance, and location relative to other tissues. The combination of these factors determine the fate of a tissue in the brain; whether it will succumb to protein misfolding damage, and if so, at what time point in an individual's lifetime disease inception will occur.

Environmental and genetic risks

Building on the identification Aβ gene mutations and their role in AD inheritance, subsequent genes whose variants have been identified which influence AD risk include those which relate to Aβ cleavage and clearance [78]. Both early and late onset AD have a genetic component, however the role of genetic factors is stronger for those who suffer from AD earlier in life - 30 to 60 years of age [79].

The strongest determinant of an individual's vulnerability to AD onset is age - whilst one in nine people 65 or older suffers from AD, one in three over 85s are living with AD [80]. Studies in model systems suggest that the influence of age on protein misfolding disease risk is due to a decline in protein homeostasis capacity over lifespan. In *C. elegans*, proteostatic capacity declines sharply after egg laying [81], and

in parallel it has been suggested that stem cells dedicate significantly more resource to proteostatic maintenance than differentiated cell types [82]. In mammals, the chaperone system decline is more gradual, but the end result - aberrant aggregation - is consistent across organisms. In addition, chemical modifications to Aβ which occur gradually over time could be responsible for the generation of a protein reservoir with a higher pathogenic potential. AD brains have a higher proportion of Aβ with N-terminal truncations and pyroglutamate modifications [83], both of which are known to occur over incrementally to proteins over considerable spans of time.

It has been suggested that neuronal cells have a particular vulnerability to protein misfolding disease, due to their longevity.- there is evidence that aggregates are reserved in the mother cell during mitotic division [84]. Therefore one could expect age-related declines in chaperone capacity to impact most strongly in this cell type.

A number of lifestyle factors have been linked to AD risk, some conclusively whilst others remain the subject of scientific debate. Head trauma carries a significant risk of AD [85], particularly upon repeated incidence and when involving a loss in consciousness. Traumatic brain injury (TBI) causes the deposition of Aβ in the injured brain [86], and studies of cortical impact injury in transgenic AD model mice found a significant increase in the levels of Aβ oligomers in the injured cortex [87].

These factors impact the vulnerability of tissues to protein homeostasis dysregulation throughout the brain, acting to reduce the time lag before disease onset in tissues (Figure 1.2). The greater this brain-wide risk, the shorter the time until disease inception, and until the disease reaches a given vulnerable tissue (Figure 1.2). It is possible that even for an individual who does not live to suffer from AD, their brain has a quantifiable lag time, and therefore the potential to succumb to the disease.

Tissue intrinsic risk

Not all brain tissues have equal vulnerability to AD, and their intrinsic susceptibility explains why some tissues remain resistant throughout disease progression (Figure

1.2). There are two circumstances under which a tissue can develop AD - through the spontaneous inception of the disease, or through the transmission of the disease from an adjacent tissue. Current theories for the origins of tissue intrinsic vulnerability are discussed in further detail in section 1.4. Results presented in this thesis will build an argument that a proteomic signature underlying vulnerability to the aberrant aggregation specific to AD is integral to determining the location in the brain where disease will first occur, and the subsequent patten of disease spread.

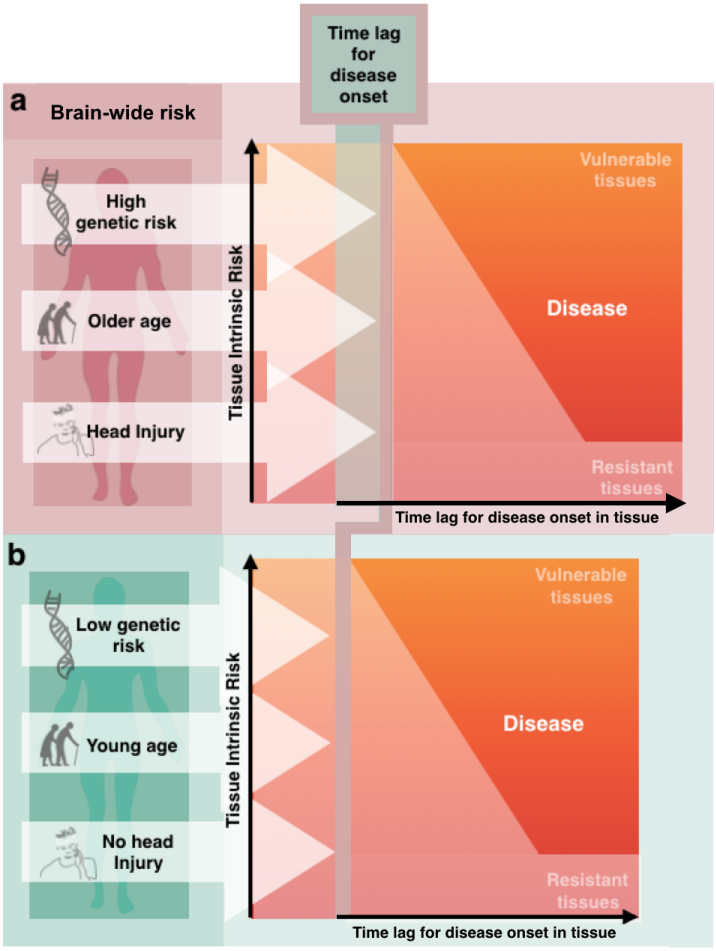

Figure 1.2: Interplay between whole brain and tissue-specific vulnerability to determine the time point of disease onset in AD.

A higher brain-wide risk **(a)** leads to a shorter lag time for both disease inception in the hippocampal region, and atrophy in subsequent tissues. Vulnerable tissues have a higher intrinsic risk and will therefore always succumb to disease first, regardless of brain-wide risks. Some tissues are below an intrinsic risk threshold and are resistant throughout. In **(b)**, a low brain-wide risk results in a longer lag time for disease onset. In summary, differential vulnerability to neurological protein misfolding diseases may result from the distance of neuronal populations to a catastrophic cliff ending in disease-specific protein homeostasis loss. Differential margins of resilience to disease are likely to depend on their function, genetic composition and history of stress exposure.

Although here AD specifically is discussed, this is a paradigm which can in theory be extended to all protein misfolding diseases. This model will be particularly relevant to protein misfolding diseases of the brain, where there are significant variations in connectivity and proteome composition between tissues, and a characteristic pattern of disease inception and spread.

1.3.4 Mechanisms of misfolding disease spread

For disease to spread to a secondary tissue, there are three key processes that render this brain region vulnerable (Figure 1.3). Protein misfolding diseases often incept within small specific neuronal populations, beyond which the disease progresses to affect a wider array of tissues, with clinically observable symptoms relating to the region affected by atrophy. It has be suggested by many groups in this field that the spatial progression of atrophy in tissues during protein misfolding disease occurs via the spread of the aggregated species [88]. This process parallels the onset of a cancer, where stochastic events at a genetic or protein homeostatic level may generate uncontrolled 'seeds' - small assemblies of misfolded proteins - catalysing further misfolding events.

For a misfolded protein to be capable of transmitting disease between tissues, the cells affected must have a number of key characteristics. Firstly, the tissue should be connected to an affected tissue. This may either by via direct proximity, or neuronal connections. Regions with synchronous baseline activity share anatomical connections, allowing researchers to deduce the brain's intrinsic connectivity networks [89]. Secondly, the secondary tissue must harbour an environment conducive to oligomeric amplification. There is currently limited evidence evaluating the relative roles played by these factors in influencing the time point at which a tissue succumbs to disease.

In some protein misfolding diseases, subsequent misfolding events can be seeded by proteins of the same species which have already achieved a pathogenic conformation, and can act as a template. This process is most commonly associated to prion disease,

a highly infectious neurodegenerative disease which can be transmitted through the exposure of a host organism to prion protein in it's pathogenic conformation [90]. Evidence is now emerging which suggests that other protein misfolding diseases including AD can incept via a similar mechanism [88], however this remains a field of active investigation and debate.

Finally, for disease to manifest in tissues with high oligomeric loads, neurons in the tissue must vulnerable to their toxicity. If seeds spread to a non-vulnerable region of the brain, oligomers may be controlled by an effective protein homeostasis system, or limited concentrations of the primary aggregating protein may deprive the aggregation process of it's necessary resource.

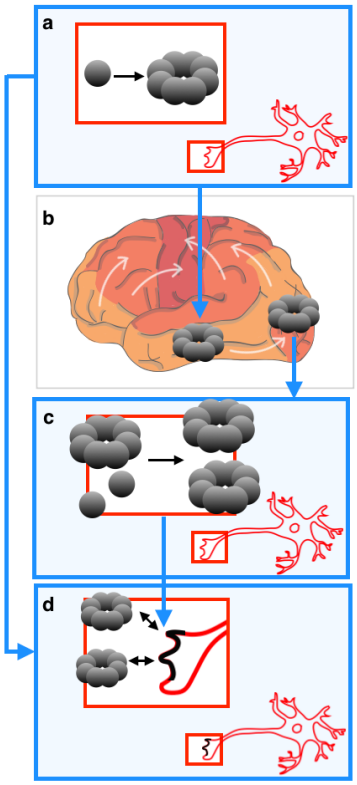

Figure 1.3: Possible mechanisms of misfolded protein spread. In the tissue where protein misfolding originates, an intrinsic environment favourable to the aggregation of Aβ into toxic oligomeric species (a) is present. Tissues to first succumb to a protein misfolding must be both in close proximity to the origin of aggregation, and contain neural tissue vulnerable to oligomeric damage (d). From the point of disease origin, aggregated species of misfolded protein spread to adjacent tissues directly, and non-adjacent tissues via neural connections (b). In secondary tissues which are exposed to misfolded protein aggregates, oligomeric amplification will occur at a rate determined by the proteomic composition and cellular characteristics (c). Secondary tissues with high oligomeric load will succumb to disease if they are vulnerable to oligomeric damage (d).

1.3.5 Introducing key protein misfolding diseases

Many of the most common protein misfolding diseases affect neural tissue, although other are systemic, or localised in specific non-neuronal tissues - a summary of notable examples in provided in Table 1.1.

Disease	Primary aggregating proteins	Most vulnerable tissues	Initial symptoms	Genes associated with inherited disease form
AD	Aβ, tau	Parahippocampal gyrus	Short term memory loss	APP, presenilin 1, presenilin 2 (most common)
PD	α synuclein	Substatia nigra pars compacta	Typically affects movement and coordination	SNCA (α syneuclin), PRKN, LRRK2, PINK1, DJ-1, ATP13A2
HD	HTT	Striatum	Problems with mood and mental abilities	HTT
Spinocerebellar ataxia	A SCA protein variant	Cerebellum	Dependent on protein variant, often eye movement affected and coordination suffers	The gene encoding the respective SCA protein variant
ALS	SOD, FUS, TDP43	Upper and lower motor neurons	Muscle stiffness, twitching, and weakness	SOD1 (most common)
Creutzfeldt–Jakob disease	Prion	Cortex	Rapidly progressing demetia	PRNP (prion protein)
AA amyloidosis	SAA	Liver, spleen and kidney	Weight loss, weakness, and edema	SAA

Table 1.1: Examples of protein misfolding diseases: a summary

Scientific attention first turned towards the role of protein aggregation in disease in the 20th century, when the first autopsies of AD brains were conducted. Deposits discovered in diseased brains were named ‘amyloid’. This reference to the latin word for starch; ‘amylum’, was made because these newly discovered deposits were stained by same dye that stained starch [91]. It has since been shown that the aggregated assemblies found in AD and other such diseases are comprised primarily of proteins. In each condition, each deposit type is composed of primarily composed of one protein, often assembled into an insoluble fibril-like structure [17].

When Alois Alzheimer described the first case of the condition that bears his name, the case was seen as a medical curiosity [92]. Over 100 years later, and AD and other protein misfolding diseases are seen by many as the next great healthcare crisis on the global horizon. 5 million people currently suffer from AD in the US, 40 million people worldwide [93]. The trend towards an escalating incidence, thought to be an unhappy consequence of rising life expectancies, is set to continue over this century. By 2050, 1-2% of people will be affected by AD, and a third to a half of those over 85s [94]. In the coming decades, this disease will also transition away from burdening only the most economically developed countries, by 2050 over 70% of cases will be in low or middle income nations [95] - protein misfolding diseases thus present a truly global healthcare challenge. In addition to the considerable emotional and social toll taken on sufferers and their support networks, there is also a considerable economic impact associated with the long term care required by many of those with AD. In the US, this cost was estimated to be $200 billion in 2015 [96], and is projected to exceed $1 trillion by 2050 [97]. There is currently no treatment available to prevent the progression of AD, with currently-approved pharmacotherapies providing only modest and transient symptomatic relief. Given that the predicted onset of the disease at a molecular level is decades before clinical symptom manifestation, the ability to achieve an earlier diagnosis will likely prove crucial to any effective treatment strategy. However, there are currently no fully validated biomarkers for the early diagnosis of this disease. It is therefore imperative that work is done to build understanding of the pathogenic mechanisms of these disease, and the origins of molecular tissue vulnerability, to provide those working to find a cure with clearer targets for prevention and early detection.

Early clinical manifestations of AD typically feature short term memory loss, most likely due to the high vulnerability of the hippocampal region to the disease. As the condition progresses, those affected can suffer from confusion, mood swings, apathy, and longer term memory loss. The aberrant assemblies found in AD brains on autopsy have been named plaques and tangles, and are primarily composed of Aβ and tau, respectively. Plaques are found in the extracellular space, whilst tangles are intracellular, and known to correlate well with the progression of cellular atrophy [98].

Familial AD can result from mutant variants of the Aβ gene, however most cases are later-onset and sporadic [99].

Aβ is produce by proteolytic cleavage of a longer membrane protein (see Section 2.1) encoded by the APP gene, which is expressed ubiquitously in neural and non-neural cells [100]. Tau, whose function is better understood, is a highly soluble cytoplasmic protein with an ability to stabilise microtubules [101]. Most disease-associated mutations currently identified within the APP gene occur either within the gene itself or in regions flanking the Aβ region. Mutations within flanking regions increase Aβ production, whilst mutations within the APP gene increase the risk of protein misfolding and subsequent oligomerisation [102, 103].

After AD, PD is the most common neurodegenerative disease. Symptoms are primarily associated with movement, although sleep pattern disruption and gastrointestinal problems can also be indicators of early disease onset. Dopaminergic neurons in the *substatia nigra pars compacta* (SNPC) are the first to succumb in disease progression, as evidenced by Lewy body deposition - a pathological hallmark of PD [104]. Lewy bodies are primarily composed of the protein α-synuclein, and accordingly disease-associated mutations have been found for the gene coding for this protein [105].

ALS, the most common form on motor neuron disease, results in progressive paralysis of skeletal muscles. Here, the most vulnerable neuronal sub-populations are spinal α-motor neurons and upper motor neurons in the brainstem and motor cortex [106]. In ALS, a number of species of characteristic deposits have been identified, primarily composed of proteins TDP-43, FUS, and SOD1. Genes encoding each of these proteins lead to familial forms of ALS when mutated [107].

Something of an outlier in the protein misfolding field, prion protein has a well elucidated mechanism of conformational-replication, spread, and toxicity. Interestingly, different perturbations in the prion protein result in an array phenotypes, suggesting intrinsic factors in human tissue confer differential vulnerability to each prion misfolded conformation [108].

Spanning cognitive and co-ordination deficits, Huntington's disease (HD) initially manifests in striatal GABA-ergic neurons [109]. Typical HD pathology involves deposits enriched in the protein huntingtin, whose risk of misfolding and eliciting disease is dependant on the number of CAG expansions in its polypeptide sequence [110]. Long spans of CAG repeats are also implicated in a further protein misfolding disease which manifests in GABAergic neurons - spinocerebellar ataxia (Sca1) [111]. The cell populations which succumb, however, differ; Sca1 develops in the giant Purkinje cells of the cerebellum [112], whereas HD manifests in neuronal populations in the striatum. It is interesting to note that the proteins that primarily aggregate in these diseases are both expressed in these vulnerable tissues, suggesting that a mutation associated to the CAG repeat will result in protein misfolding in different cellular environments dependant on the protein repeat. It is currently poorly understood why both regions are not affected in each case, although a number of theories for the origins of vulnerability to protein misfolding disease have been proposed, and are discussed in the following subchapter.

1.4 Origins of tissue vulnerability to neurodegenerative disease

1.4.1 Selective vulnerability

Selective vulnerability in protein misfolding conditions is manifested in the innate differential vulnerability of tissues, cell types, and subcellular localisations to disease [113]. This vulnerability is determined by many macroscopic and microscopic characteristics, including the location and connectivity of tissues, cellular function, and region-specific microenvironment harbouring particular combinations of subcellular components, including protein homeostasis machinery, and proteins critical to synaptic function [114-116].

As discussed in section 1.3.1, the concept of selective vulnerability has the potential to explain why differences in the spatiotemporal pattern of progression exist between neurodegenerative diseases. The genetic signature for neurodegenerative disease

predisposition is complex, incompletely understood, and suggests a multifactorial cause (Table 1.1). Two key barriers exist to the deconvolution of potential causative factors in protein misfolding diseases. The first relates to the many possible macroscopic and microscopic levels at which disease progression can be defined. It is possible to consider disease manifestation from the perspective of protein misfolding, soluble aggregate formation, insoluble aggregate deposition, neuronal atrophy, neuronal network breakdown, and clinical symptom manifestation. Without clarity on how the disease mechanism operates and spans across these levels of pathology, it is difficult to identify the root cause of the disease. Secondly, these disease-manifestations are most commonly studied in humans at post-mortem, at the end of the pathological cascade. The absence of a window into the presymptomatic events occurring in brains during protein misfolding disease, is perhaps why so much debate remains in the protein misfolding field around tissue vulnerability and disease origin.

By elucidating the factors which influence vulnerability to protein misfolding disease at a molecular level, there is potential to build a comprehensive understanding of the disease pathology, from proteins to symptoms. As a consequence, the development of targeted approaches which specifically protect neurons at risk, or intervene with critical protein homeostasis processes, will be possible.

1.4.2 Neural connectivity and activity

Connectivity studies suggest a particular vulnerability of highly connected 'hub' regions of the brain to AD pathology. Regions of the brain with marked Aβ deposition early in disease progression also show the highest level of neuronal connection to adjacent tissues [117]. These regions include, but are not exclusive to, the default mode network [118]. The construction of a network diffusion model based on the brain's connectivity network resulted in a prediction of AD topography in the brain, and studies have found that fibre pathways have a stronger statistical influence on the route of disease progression when compared to spatial proximity alone [119]. In vivo, exogenous seeding of metastable proteins in the hippocampus resulted in the development of diseased tissue in connected regions [120]. The mechanistic connection

between high connectivity and AD vulnerability is likely to be complex, with pathology spread and elevated activity levels both suggested to play a role.

Network hubs are more likely to be exposed to a 'prion-like' spread of Aβ assemblies. Evidence of prion-like spread of soluble aggregated assemblies has been identified in a number of protein misfolding diseases. Human transplanted tissue adjacent to endogenous, diseased tissue, has been observed to develop PD neuropathology [120]. Experimental transmission of AD-associated aggregates has also been observed in mouse models of the disease [120]. Whilst the mechanisms for this spread from cell to cell remain unclear, a number of possibilities have been proposed. These include the cellular release and free diffusion of early aggregated species through the extracellular space, transport via secreted vesicles, or trafficking through inter-cellular pathways.

The mechanistic connection between hub activity and disease vulnerability is yet to be elucidated, however elevated levels of local neural activity have been shown to be associated with elevated Aβ deposition [121]. To investigate the relationship between connectivity, level of activity, and protein misfolding disease risk, researchers have developed a computational model of the human brain. By predicting progressive atrophy based on activity level, it was possible to predict resultant perturbations in the structure and dynamics of the remaining brain network. Thus these results support a role for elevated intrinsic levels of neuronal activity in vulnerability to AD [122]. A further study sought to identify common phenotypic traits linking the neuronal groups most vulnerable to neurodegenerative disease. Recurring themes identified include intense firing patterns, and dynamic control of excitability relying on high cytosolic calcium fluxes [123]. For example, PD-vulnerable SNPC dopaminergic neurons are characterised by intense firing properties that rely on fast oscillations in cytosolic calcium levels and a high energy supply [123]. It is possible that high activity levels result directly in a proteomic environment at high risk of AD-specific aggregation: a recent study using in vivo micro-dialysis has demonstrated a correlation between interstitial fluid Aβ concentration and synaptic activity in APP transgenic mice [121].

The energy demand of a neuronal subpopulation is largely dependent on it's activity level, due to a frequent need to re-establish ion gradients [124]. Hippocampal CA1 neurons, which succumb early in the progression of AD [125], are characterised by particularly high levels of energy consumption. It has been proposed that an elevated metabolic demand leaves these cells vulnerable to hypoxia, and other metabolic stresses due to high production of free radicals [126]. At a higher rate of damage, cellular components would require more degradation and re-synthesis, escalating the metabolic demand on cells. The capacity of the brain to cope with this protein homeostatic burden is likely to decline with age, thereby enabling the escalating self-assembly of the most risky misfolding proteins. Indeed, mild cognitive dysfunction, which can often progress to AD, correlates with decreased glucose utilisation in the brain [127].

1.4.3 Gluscosaminoglycans

Glucosaminoglycans (GAGs) are long unbranched polysaccharides [128], found in abundance in the extracellular matrix, especially the basement membrane [129]. These macromolecules have a role in many cell processes, including cell adhesion and signalling [130]. GAG chains can attach to proteins to form a proteoglycan, heparan sulphate and chondroitin sulphate are the most common GAGs which form proteoglycans. The presence of GAGs in Aβ deposits was first identified using staining techniques on brain sections of AD patients [131]. Since, GAGs have been found associated with amyloid fibrils isolated from a number of neurodegenerative diseases [132]. Work investigating a possible role for GAGs in amyloidosis have found that sulphated GAGs accelerate the aggregation of a wide range of proteins, both metastable and non-aggregation prone [133-135]. The proposed mechanism centralises around the role of heparin sulphate in the formation of an aggregation seed, and subsequent impact of amyloid fibril stability [136, 137].

1.4.4 Metal ion homeostasis

Metal ions, which are required for a wide range of biochemical processes, have the potential to elicit cellular toxicity in their free forms [138]. Metal ion homeostasis across the brain could therefore have a role in determining the relative vulnerability of tissues to protein misfolding diseases. Metal ions may also influence neurodegenerative disease vulnerability through more specific mechanism, for instance it has been shown that zinc ions can stabilise Aβ oligomers [139].

1.4.5 Factors beyond neurons

Model systems of neurodegenerative disease suggest that there are factors in addition to the firing and neurochemical properties of brain regions which determine relative vulnerabilities. Knock-in mutations in mice modelling HD, Sca1 and FFI (another neurodegenerative disease) primarily affect the striatum, cerebellum, and thalamus respectively [140-142]. These are the earliest regions to exhibit disease onset in each case in the human disease counterpart, suggesting that these mouse models reflect these pathologies relatively well. However, murine thalamus contains a far lower proportion of GABAergic neurons than the human thalamus [143], suggesting that perhaps tissue location relative to other regions of the brain may play a role in vulnerability. In addition, non-neural cell composition could play a role - there is evidence for a role in glial and vascular cells in the disease process.

1.4.6 Inflammation

A characteristic feature of AD pathology, the inflammatory response is primarily driven by microglial cells, and has been suggested to be responsible for escalating disease progression [144]. Microglia are the primary immune cell in the brain, and work to prevent neural damage in the event of injury or pathogenic insult. During health, microglia work to support neuronal function, and have key functional roles including synaptic pruning, and the of promoting novel brain connectivity and development. The association between AD and mutations encoding the immune-related genes TREM2 and CD33 [145, 146] underscores the role of immune

dysregulation in disease vulnerability. However, there remains much debate about the stage in disease progression at which inflammation manifests - and crucially whether it is a causative factor or a consequence [144]. There are in the field two conflicting perspectives, the first being that immune system activation, mediated by microglia and astrocytes, follows as a consequence of Aβ deposition. The second perspective involves an earlier immune involvement. It is proposed that immune activity is sufficient to trigger AD pathology independently of Aβ deposition, and by promoting elevated Aβ levels initiates a self-reinforcing spiral towards neuronal atrophy [147].

In support of this second perspective, systemic inflammation from chronic diseases such as psoriasis is known to confer an elevated risk of AD [148]. A heightened AD incidence has also been observed in those who suffer from type II diabetes [149], a condition for which CNS inflammation and microglial activation have been described as important components, and for those with traumatic brain injury, known to be associated with neuroinflammation.

There is also evidence of Aβ aggregation acting to influence the inflammatory processes. It has been demonstrated that soluble Aβ oligomers and insoluble fibrils bind to a number of receptors expressed by microglia, including CD14, CD36, CD47 and α6β1 [150, 151]. There is evidence that in neurodegenerative disease, the supportive homeostatic roles played by microglia are lost, and instead they trigger chronic neuroinflammation [152].

1.4.7 Differential protein expression

A final potential explanation for selective vulnerability places focus on differences between the expression signature of tissues [153]. Expression levels of the protein which catalyses an aggregation cascade in the tissues first to succumb to the respective disease has, accordingly, been investigated. However studies chose a small number of sample tissues and lacked statistic rigour, and results were inconclusive [43].

1.4.8 Protein homeostasis

Beyond primary aggregators, the protein homeostasis system responsible for the aggregation kinetics of these metastable proteins should also be considered. An improved understanding for how protein control machinery components are distributed across the brain would allow an evaluation of their potential to influence regional vulnerability.

Much work has been done to understand the function of network of molecular chaperones and aggregation modulators who work to control protein misfolding in AD [154, 155]. To capitalise on this research, it is important to build understanding of the role played by the protein homeostasis networks of Aβ and tau in defining tissue resilience to AD. More broadly, by investigating the factors which underlie tissue vulnerability to AD, one can begin to interrogate the molecular mechanisms of disease pathogenicity. As yet, there is still no consensus in the field of neurodegenerative disease research about whether protein aggregation is a causative factor in disease inception and propagation. In the following chapter, we investigate levels of subproteomes responsible for the homeostatic control of two proteins key to AD aberrant aggregation; Aβ and tau, in normal brains. This approach allows one to ascertain whether the proteostatic environment in tissues most vulnerable to AD is conducive to aberrant disease-specific aggregation.

2 Aggregation Modulators

2.1 Introduction

2.1.1 AD-specific aggregation modulator networks

Much research effort has been dedicated to elucidating the molecular chaperone and post translational modification networks which modify Aβ and tau misfolding pathways [154, 155]. Figure 2.1 illustrates these proteins in the context of their role in these pathways; proteins in boxes have a scientific consensus regarding their influence on Aβ and tau aggregation, whilst unboxed proteins currently have conflicting or inconclusive results with respect to AD-specific aggregation. In the following chapter we discuss the roles played by the boxed proteins in Figure 2.1 in Aβ and tau homeostasis.

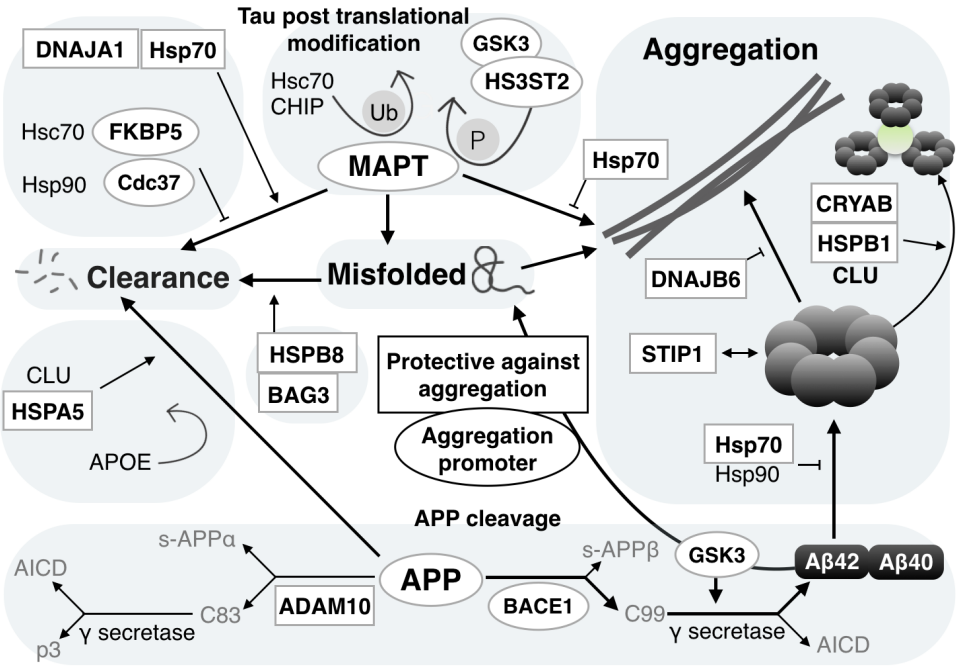

Figure 2.1: The protein homeostasis network responsible for modulating the aggregation dynamics of Aβ and tau

Aβ production via cleavage

Aβ is the cleaved product of the protein APP. Once expressed, three secretases can cleave the protein to produce two smaller fragments. If cleaved by α-secretase (encoded by ADAM10), the protein modification pathway is diverted away from the production of amyloidogenic Aβ [156-159], by producing the two APP fragments s-APPα and C83 [160]. By blocking synaptic trafficking of ADAM10, researchers have been able to generate a model of the initial phases of sporadic AD [161] by shifting APP metabolism towards Aβ production. Furthermore, the secreted ectodomain fragment aAPPα has been shown to have neuroprotective properties [162]. Such is the pivotal role of Aβ production in AD genesis, that two mutations in ADAM10 impairing the function of the secretase have been found to be associated with the familial late onset form of the disease [163]. If the levels of ADAM10 are elevated in mouse models of AD, Aβ plaque load is reduced [164].

If instead β secretase (encoded by BACE1) cleaves first, an on-pathway product for Aβ production, C99, results [165]. BACE1 is a membrane-bound aspartyl protease, and acts as the first and rate limiting step in the pathogenic APP spliced variant production. BACE1 cleavage is a crucial step in the production of Aβ; β secretase depletion almost completely abolishes Aβ production in mouse models [166-168]. The membrane bound stub C99 is next cleaved by γ secretase to form AICD (APP intracellular domain) and an Aβ variant [169]. This cleavage event occurs within the membrane, and liberates both polypeptides produced. This final cleavage site is variable and can occur at amino acids 37 to 43. Aβ42 is the most metastable variant of Aβ, with a strong propensity to oligomerise in vivo [170]. In addition, GSK3α, (but not GSK3β) has been shown to regulate APP cleavage, enhancing Aβ42 production [171, 172]. GSK3 is a constitutively active seriene/threonine kinase, known to have a wide range of functions in the cell, from glycogen metabolism to gene transcription. Therapeutic use of lithium, a GSK3 inhibitor, blocks the production of Aβ by interacting with APP cleavage at the γ-secretase step [172].

Aβ aggregation

Once produced, Aβ typically remains in it's soluble form for the majority of an individual's lifespan. If misfolding becomes dysregulated, this metastable peptide can begin to aggregate, forming soluble oligomers of low molecular weight, in addition to insoluble fibrillar deposits (Figure 1.1).

The heat shock proteins Hsp70 and Hsp90 have been found to inhibit the initial stages of Aβ aggregation. Studies suggest that these molecular chaperones cause structural changes to oligomers, but have little effect on fibrils [173]. Authors of a paper investigating the mechanism of this inhibition propose two possible models: 'holding' and 'refolding' [173]. In the former model, molecular chaperones bind the misfolded protein, probably via its substrate binding domain. In the latter, molecular chaperone binding results in a change of three dimensional structure of its target. The resultant configuration is more kinetically robust against a transfer to the aggregated state. Hsp70 production has been found to be induced in neuronal models in response to Aβ42 expression, and in models where this molecular chaperone is over-expressed, neurons are protected from the toxicity associated with aberrant Aβ accumulation [174]. These molecular chaperones have also been shown to be effective at more physiological levels; Hsp70 and Hsp90 have been found to inhibit Aβ self assembly at sub-stoichiometric concentrations [173].

A co-chaperone of Hsp70 and Hsp90, stress inducible phosphoprotein 1 (STIP1), has also been found to intervene early in the aggregation pathway of Aβ. Shown to protect neurons against Aβ oligomer damage, STIP1 reduces synaptic loss and cell death in cultured neurons, and preserves long term potentiation in hippocampal slices. Furthermore, haploinsufficient neurons were demonstrated to have increase vulnerability to Aβ oligomeric assault, a phenotype which could be rescued by treatment with recombinant STIP1 [175].

The action of the co-chaperone DNAJB6 has been analysed comprehensively by kinetic studies, and has been shown to retard the aggregation process of Aβ in a

concentration-dependant manner, extending to sub-stoichiometric levels [176]. Mechanistic investigations suggest that the high efficiency of DNAJB6 is due to its interaction with early aggregated species, rather than the Aβ monomer [176]. Oligomeric interaction, and often sequestration, is a mechanism of action suggested to be employed by a number of molecular chaperones to mitigate damage by this highly toxic aggregated species. One such example is αB-crystallin (CRYAB). Studies indicate a preferential binding of this molecular chaperone for early aggregated species, thereby inhibiting further self-association of Aβ to form fibrils [177]. Induction of CRYAB expression appears to be employed in vivo as a response to the pathogenic threat of Aβ - this small heat shock protein is found at elevated levels in brains affected by AD [178]. Accordingly, αB-crystallin has been shown to protect against Aβ toxicity in cell culture [179]. Another heat shock protein, HspB1, is also thought to sequester oligomers to produce large non-toxic aggregates in a neuroprotective mechanism [180]. As with DNAJB6, this mechanism of action appears to be highly effective; substoichiometric concentrations are sufficient to abolish Aβ oligomeric toxicity in mouse neuroblastoma cells [180].

Aβ clearance

In addition to attenuating the Aβ aggregation pathway, the protein homeostasis system also acts to prevent aggregation by targeting this highly amyloidogenic protein for clearance. HspA5 achieves this goal by inducing the production of cytokines, resulting in microglial phagocytosis of Aβ fragments [181]. In addition, experiments suggest that this molecular chaperone interferes with the maturation process of the APP protein to decrease Aβ40 and Aβ42 secretion [182]. This Hsp70 family member has furthermore been suggested to interact transiently and directly with the Aβ which is produced to ensure correct folding and limit misfolding risk [182]. The mechanisms of other proteins involved in the clearance and chaperoning of Aβ remain somewhat less clear. In AD model mice, the absence of clusterin (apoJ) and apoE resulted in both an earlier onset and higher prevalence of Aβ deposition [183]. ApoE and apoJ are lipophilic proteins with a role in lipid homeostasis, and are synthesised in neural tissue by astrocytes. The effects of these molecular chaperones on Aβ aggregation is

additive [183], however the precise mechanism of their action remains under active investigation.

Tau post translational modification and clearance

Whilst APP will only aggregate *in vivo* when cleaved to form Aβ, tau phosphorylation at number of key sites has been shown to predispose its assembly into fibrils. GSK3 is a ubiquitously expressed kinase, and has been shown to phosphorylate tau at most of the serine and threonine residues known to be hyperphosphorylated in aggregated tau filaments [184]. Tangle formation is known to be associated with neuronal atrophy, and accordingly mice with forebrain GSK3-β over-expression exhibit apoptotic neuronal death, in addition to tau hyperphosphorylation [185]. This model also recapitulates many other phenotypic aspects of AD, including an immune response, and impaired mental function [185]. Both GSK3-α and GSK3-β have been shown to induce the hyperphosphorylation of tau, in vitro and in cellular models [186-190]. It is thought that phosphorylation sites critical to aberrant aggregation are located around the microtubule binding domain, because an inability to bind microtubules is associated to tau's resilience to aggregation [191, 192]. Another enzyme known to induce pathogenic phosphorylation of tau, HS3ST2, is predominantly expressed in neurons. In a zebrafish tauopathy model where HS3ST2 was inhibited, there was a significant reduction in abnormally phosphorylated tau epitopes [193, 194].

Although chemical modification of tau can drive it towards aggregation, other processes, such as ubiquitination, can have a protective influence by signalling tau for clearance. Hyperphosphorylated tau bound to the molecular chaperone Hsc70 can be ubiquitinated by the E3 ubiquitin ligase CHIP (carboxy-terminus of the Hsc70-interacting protein) [195].

It is important to note that the role of Hsc70 in tau homeostasis is more nuanced and extends beyond this interaction alone. Hsc70 has many binding partners responsible for moderating it's activity, and has been suggested to have a role in mediating the folding of tau [196], and even slowing the process of tau cleavage relative to the

activity of other Hsp70 family members [197]. The ability of the molecular chaperone under study to accelerate tau clearance appears to depend on its capacity to recruit CHIP [198].

The protein DNAJA1 has also been suggested to act to stimulate tau clearance via the ubiquitin pathway [199]. When over-expressed in cell culture, DNAJA1 has been found to significantly reduce the levels of highly phosphorylated tau species found in intraneuronal tau aggregates [200].

Acting to effect the converse, FKBP5 forms a molecular chaperone complex with Hsp90 that prevents tau degradation [201]. In FKBP5 null mice, tau levels show a significant reduction, and when overexpressed in a murine model, tau is preserved [202]. In addition to functioning via the stabilisation of Hsp90, FKBP5 has also been found to influence the phosphorylation pattern of tau [201]. Another Hsp90 co-chaperone, Cdc37, appears to elicit the same effect on tau homeostasis. Studies have found that suppression of Cdc37 leads to tau clearance following Hsp90 inhibition, whilst Cdc37 over-expression preserves this amyloidogenic protein [203]. In parallel with the activity of FKBP5, there is evidence to suggest that Cdc37 is also able to alter tau kinase stability, altering the phosphorylation profile of tau and modulating its toxicity [203].

Degradation of misfolded proteins

In addition to clearance pathways specific to Aβ and tau, there exist a number of less specific pathways for targeting the degradation of misfolded proteins, also relevant to AD pathogenesis. HspB8 is a heat shock protein which is known to complex with its co-chaperone BAG3. When these molecular chaperones are over expressed in cell culture, autophagy is stimulated and the clearance of misfolded protein results. In primary neurons, autophagy activation resulted in a significant decrease in levels of highly phosphorylated tau [204].

Chapter 2.1.1 highlights the highly interconnected and varied functions of protein homeostasis components in the cell. Some proteins, such as β-secretase, have a clearly defined impact on the AD aggregation risk and a well understood mechanism of action. These proteins were taken forwards in this work to investigate the role played by aggregation modulators in determining intrinsic AD risk. For other proteins, such as Hsp90, known to inhibit both Aβ aggregation and tau clearance, it is currently difficult to discern their net impact on AD-specific aggregation risk. For this reason, these proteins have been excluded from consideration in this study, however as new research sheds light on the rich homeostasis network regulating Aβ, tau, and other relevant metastable proteins, we anticipate the genes under consideration could increase in number. Proteins with a well understood and significant influence on the aggregation risk of Aβ and tau are summarised in Table 2.1.

Gene name	Known function
Promote aggregation	
Cdc37	Suppression of Cdc37 destabilises tau, leading to its clearance, whereas Cdc37 overexpression stabilises tau.
GSK3A	Both GSK3β and GSK3α induce tau hyperphosphorylation. Also, GSK3α, but not GSK3β, has been shown to regulate APP cleavage resulting in the increased production of Aβ. Therapeutic use of lithium, a GSK-3 inhibitor, blocks the production of Aβ by interfering with APP cleavage at the γ-secretase step.
GSK3B	
HS3ST2	HS3ST2 expression is critical for the abnormal phosphorylation of tau in AD.
FKBP5	FKBP5 forms a mature molecular chaperone complex with Hsp90 that prevents tau degradation and promote neurotoxic tau aggregation.
BACE1 (β secretase)	BACE1 is the initiating and putatively rate-limiting enzyme in Aβ generation. BACE1 inhibition may thus block the production of Aβ and prevent the development of Aβ-associated pathologies.
APP	APP fragments aggregate to form amyloid plaques in AD.
MAPT	Aggregates of tau (the product of MAPT) form neurofibrillary tangles in AD.
Protect against aggregation	
HSPA1A (Hsp 72)	Molecular chaperones of the Hsp70 family have been implicated in the prevention of abnormal tau aggregation in adult neurons.
HSPA1B (Hsp 70)	
HSPA1L (Hsp 70)	
DNAJA1	DnaJA1 triages all tau species for ubiquitin-dependent clearance mechanisms.
HSPB8	The up-regulation of HSPB8 and BAG3 may enhance the ability of astrocytes to clear aggregated proteins released from neurons.
BAG3	
GRP78	Binds to APP and reduces Aβ secretion, and may stimulate Aβ clearance in microglia.
ADAM 10 (α secretase)	α-secretase, (the metalloprotease ADAM10), cleaves APP within the Aβ domain, thus preventing Aβ generation.
STIP1	May protect from Aβ oligomer toxicity.
CRYAB	Binds Aβ oligomers, and inhibits nucleation-dependent polymerization of amyloid fibrils.
HSPB1	HspB1 sequesters toxic Aβ oligomers and converts them into large nontoxic aggregates. Furthermore, the presence of HspB1 decreases the amount of Aβ42 released by cell lines.
DNAJB6	Interaction of the molecular chaperone DNAJB6 with growing Aβ42 aggregates leads to sub-stoichiometric inhibition of amyloid formation.

Table 2.1: Aggregation modulators of Aβ and tau

Molecular chaperones and post translational modifiers with a known net influence on the aggregation of Aβ and tau are listed, alongside their currently understood role in the homeostasis of these two key metastable proteins.

2.1.2 Assessing aggregation modulator impact on AD vulnerability

Although aggregation modulators have been studied at the individual level, or in small subsets, few researchers have taken a whole network perspective when considering their impact on AD-specific aggregation risk. We chose to look at disease risk from the perspective of intrinsic tissue vulnerability. This approach allowed for the consideration of both genetic and sporadic cases as one dataset, and for the study of normal brains before disease onset, thereby avoiding any conflicting signal arising from transcriptional changes present during disease progression. By contrast, comparing vulnerable from resistant individuals would either have required a comparison of those with a genetic predisposition to those without, or an analysis of brains after disease incidence. As discussed in section 1.4, specific tissues have an elevated intrinsic risk to AD, and will consistently succumb earlier in affected individuals.

2.1.3 Defining relative AD tissue vulnerability

This progression of AD through the brain was first identified by Braak in 1991 [125] using the deposition of tangles in brains studied at consecutive stages of the disease. Braak identified that brains of individuals who died whilst exhibiting clinical symptoms indicative of the early stages of the disease, exhibited very localised tangle deposition, initially in the parahippocampal gyrus. In contrast, the brains of those who died during much clinically later disease stages exhibited much more widespread tangle deposition, although regions such as the cerebellum were constantly spared even in the most severe cases. Braak noted that the pattern of tau spread was broadly consistent between brains, allowing the classification of a 'Braak staging'. Braak I corresponds to tissues which succumb in the earliest stages of the disease, and Braak 6 tissues are the last to be affected. Here, we refer to tissues who are completely resistant to AD and therefore without a Braak classification as 'non-Braak'. This classification corresponded symptomatic and histological events during the progression of AD, and allowed the identification of the most vulnerable, and the most resistant tissues.

2.1.4 Evaluating the association between subproteome expression and tissue vulnerability

In our approach, the Braak staging is used to test the power of subproteomes of interest to predict healthy tissue with intrinsic vulnerability to AD. To achieve this goal, a Δ score was defined. Δ scores are calculated by finding the expression of a given gene, in a given tissue, relative to expression of the corresponding gene in tissues completely resistant to AD. More details on the calculation process can be found in Methods, section 2.4. This approach allowed quantification of relative expression between tissues, and comparison between genes. A gene with a positive Δ score would therefore have elevated expression in the tissue of interest relative to resistant (non-Braak) tissues, and the inverse is true for a gene with a negative score. Data sources used in this study are summarised in Table 2.2, and discussed in further detail in Methods. Expression data was predominately sourced from the Allen Brain Atlas [205], a data resource unique for it's comprehensive and highly granular microarray analysis of normal human brain tissue. More details on the data used in this chapter can be found in Methods, section 2.4.1.

2.2 Results

2.2.1 Δ score analysis of Aβ and tau aggregation modulators

To investigate the possibility of an intrinsic tissue vulnerability to AD, determined in part by a protein complement which predisposes the aberrant aggregation of Aβ and tau, we employed Δ score analysis. Expression of aggregation modulators in tissues most vulnerable to AD was calculated, relative to expression in tissues completely resistant to AD ($\Delta_{BI\text{-}III}$). The results, (Figure 2.2), were striking. We found the Δ-Braak I-III score for aggregation promoters to be significantly elevated relative to that of aggregation protectors. Moreover, the median $\Delta_{BI\text{-}III}$ score for aggregation promoters was positive (0.48), indicating that aggregation promoters are expressed at elevated levels in the most AD-vulnerable tissues relative to resistant tissues. Conversely, the median $\Delta_{BI\text{-}III}$ score for protectors was negative (-0.47), and therefore expression of this

subproteome is on average expressed at a lower level in vulnerable tissues relative to those completely resistant to AD. We therefore find at a level of low granularity, an expression signature, which, if reflective of the proteomic environment in the cell, predisposes the aberrant aggregation in the tissues which succumb earliest to AD. The brains studied here are normal, and of ages decades before that typical of AD onset. We therefore suggest that this vulnerability signature is intrinsic to the tissue, and independent of any subsequent transcriptional changes which occur in the brain during disease progression.

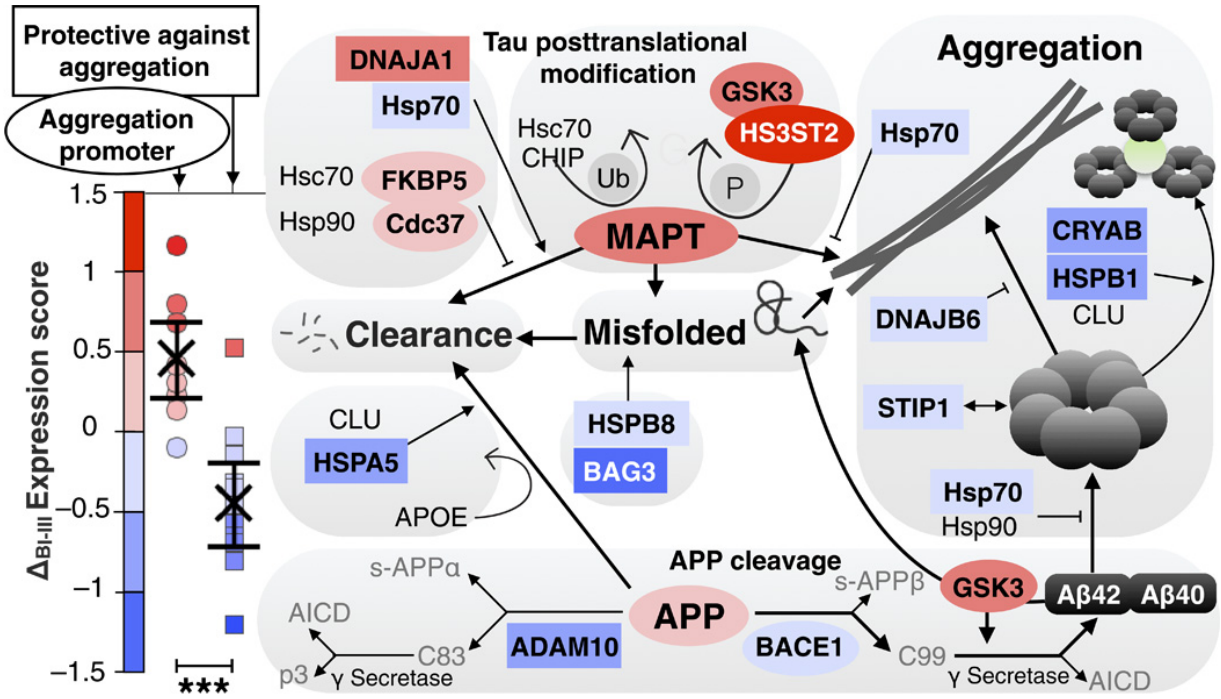

Figure 2.2: In vulnerable tissues in healthy brains, the expression patterns of specific molecular chaperones and post-translational modifiers predispose the aggregation of Aβ and tau. In vulnerable tissues in healthy brains, the expression patterns of specific molecular chaperones and post-translational modifiers predispose the aggregation of Aβ and tau. Proteins known to promote Aβ and tau aggregation (termed "promoters") are shown within ellipses, and proteins known to protect against it (termed "protectors") are shown within rectangular boxes (Table S4). Proteins whose roles in Aβ and tau aggregation are ambiguous in the literature are shown without frames and not considered in further analysis. Δ_{BI-III} scores are colour-coded according to the legend in the lower left corner. Error bars in the plot represent a 95% confidence interval on the mean (X symbol) calculated with 10^5 bootstrap cycles, ***$P < 0.001$ calculated with a two-tailed t test [206].

2.2.2 Evaluating the predictive power of aggregation modulator levels

To investigate the relationship between the levels of expression of protein homeostasis components and the tissue vulnerability to AD at a more granular level, we focused our analysis on tissues where an accurate alignment was possible between AD-staged tissues and the tissue parcellation used in microarray analysis (see Methods, Section 2.4). By focusing on tissues where we had the most accurate knowledge of the Braak

stage of the sample, we were able to distill a signal from the data set even when studying tissues from individual Braak stages (in contrast to section 2.2.1 where we compare Braak tissues I-III, to Braak tissues IV-VI and to non-Braak tissues)

We found that expression levels of protein homeostasis components that modulate the aggregation of Aβ and tau in healthy brains are good predictors of tissue vulnerability to AD and recapitulate the staging of the disease (Figure 2.3, full dataset in Appendix B1). This result indicates a role for AD-related protein homeostasis components not just in determining the region in which the disease will first incept, but also the route of spread through the brain. This result suggests that after the initiation of aberrant aggregation in a tissue, the disease will only spread to an adjacent tissue if the homeostasis of the disease-specific protein is vulnerable to collapse. Tissues with more a more robust capacity to control the misfolding of Aβ and tau, appear to be affected at correspondingly later stages of the the disease. An alternative hypothesis that cannot be excluded at this stage, is that tissues with shared function and neuronal connection also have highly similar proteomic composition. If pathogenic spread of AD were to occur along neuronal tracts between functionally-related regions of the brain, proteomic similarities between these regions could also contribute to the observed result. Indeed, AD progression has been shown to reflect to some extent the brain's connectome. It is likely that the relative contributions of connectivity and proteomic vulnerability to the route of AD spread will remain unclear until modelling work is completed taking both of these factors into consideration.

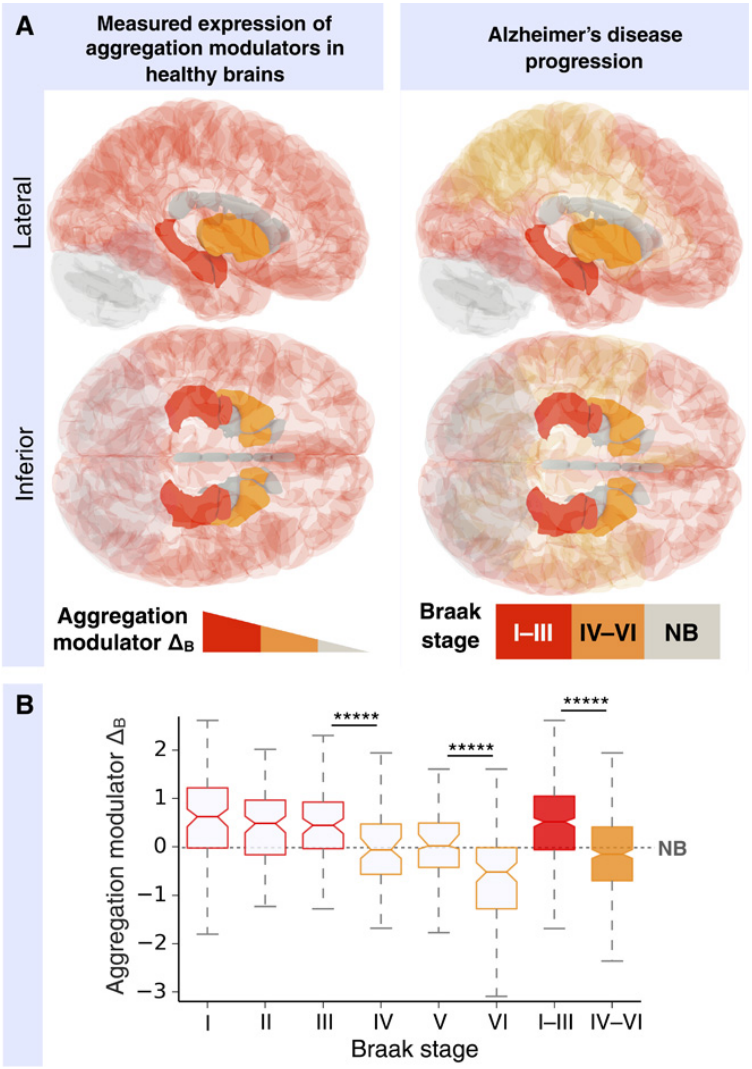

Figure 2.3: In healthy tissues, a protein homeostasis expression signature associated with Aβ and tau aggregation recapitulates the progression of AD well before the onset of the disease. In healthy tissues, a protein homeostasis expression signature associated with Aβ and tau aggregation recapitulates the progression of AD well before the onset of the disease. (A) Tissues are coloured according to the mean Δ score for expression in healthy brains of the aggregation modulators (protectors and promoters) of Aβ and tau aggregation (left) and to the Braak staging (right) The mean Δ score for aggregation modulator is calculated as the difference between the mean Δ scores for aggregation promoters and protectors in the region under scrutiny. (B) Box plot of the mean Δ scores for aggregation modulators [as calculated in (A)] in perfect-match tissues (see Methods) affected at progressive Braak stages (x axis). *****$P < 0.00001$; P values were calculated with Mann-Whitney U test with Benjamini-Hochberg multiple hypothesis testing correction [207]. NB, non-Braak.

At a physiological level, our interest in this study lies in understanding the relative levels of protein molecules in cells. However, research at present is limited in this field due to the lack of granular, high throughput proteomics data on normal human brain tissue [208]. Consequently, we have used mRNA expression data as a proxy for proteomic data in this study [205]. A natural progression from the central dogma, scientific opinion has generally held that there exists a direct correspondence between the level of mRNA transcripts and their protein product. However, some studies have indicated that in certain circumstances the correlation between mRNA and protein levels can be limited by factors including variance in half life and post-transcriptomic machinery. It was therefore imperative that we took steps to check the correlation between the expression data used in this study, and proteomic analysis of comparable brain tissues. For this control, we looked at protein levels from two perspectives; analysis with broad proteome coverage in two brain regions (Figure 2.4 a-b) , and a select number of proteins crucial to AD in a number of precise brain regions (Figure 2.4 c-d). Figures 2.4 a and b are based on immunohistochemistry studies of normal brain tissue, and demonstrate a statistically significant relationship between mRNA expression levels used in this study, and measured protein levels, in both the hippocampus and the cerebral cortex [208]. In a second study, immunoblotting was used to quantify the levels of tau in six different brain tissues; cerebellum, entorhinal cortex, occipital cortex, superior temporal gyrus, frontal cortex and hippocampus, relative to measured actin levels in each tissue [209]. We provide the Δ expression score for each region (relative to expression in non-AD tissues), for tau, to allow a comparison between protein levels as measured in this study and mRNA levels as used in this thesis. We see a good correspondence between levels in different tissues, with consistently depressed levels of tau in the cerebellum - a region of the brain known to be highly resistant to AD (Figures 2.4 c-d). We quantified this relationship by calculating the correlation coefficent between relative MAPT mRNA levels and relative tau protein levels as 0.9 ($p = 0.01$).

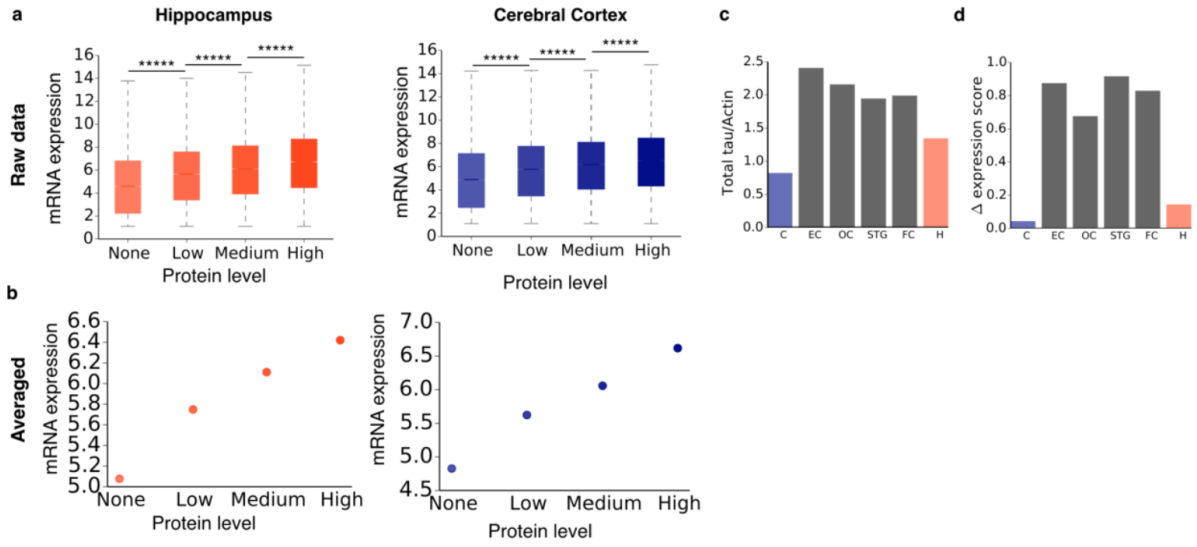

Figure 2.4: Analysis of the average correlation between mRNA and protein levels. (a) Correlation between mRNA and protein levels for two representative regions of the brain [208]. (b) Correlation between mRNA and protein level, averaged data; error bars are smaller than the data point symbols. (c) Measured levels of total tau protein [209], normalised to actin levels. The expression score for each region is relative to non-Braak tissues. C: cerebellum; EC: entorhinal cortex; OC: occipital cortex; STG: superior temporal gyrus; FC: frontal cortex and H: hippocampus. (d) Relative mRNA levels for the MAPT gene, for corresponding brain regions. The correlation coefficient between relative MAPT mRNA levels and relative tau protein levels is 0.9 ($p = 0.01$). Boxes represent the first and third quartiles of the distribution, whiskers the 1.5 inter-quartile range, and notches are standard errors on the median calculated with 104 bootstrap cycles. *****$p<0.00001$, calculated with a Mann-Whitney U test with Benjamini-Hochberg multiple hypothesis testing correction [207].

2.2.3 Testing the specificity of AD-aggregation modulators

Because of the generic role of many of the aggregation modulators studied in this analysis, such as the members of the Hsp70 family, we wanted to verify that the signal we observe in Figure 2.2 and 2.3 is present only for AD-specific aggregation. We therefore constructed lists of aggregation modulators associated with the protein homeostasis of TDP43 and SOD1, and a second gene set for the molecular chaperones and post translational modifiers known to influence α-synuclein aggregation. As can be seen in Figure 2.5a, the relatively high expression of aggregation promoters associated with ALS and PD in tissues most vulnerable to AD is less marked and not statically significant, the same is true of depressed relative aggregation protector expression. This result is observed despite a shared core of aggregation modulators between these diseases (Figure 2.5b), indicating that a small subset of disease-specific molecular chaperones have significantly different expression levels across brain tissue, and play an important role in determining disease specific vulnerability. Tissues defined as 'most vulnerable' for each alternative neurodegenerative disease are listed in Table 2.2.

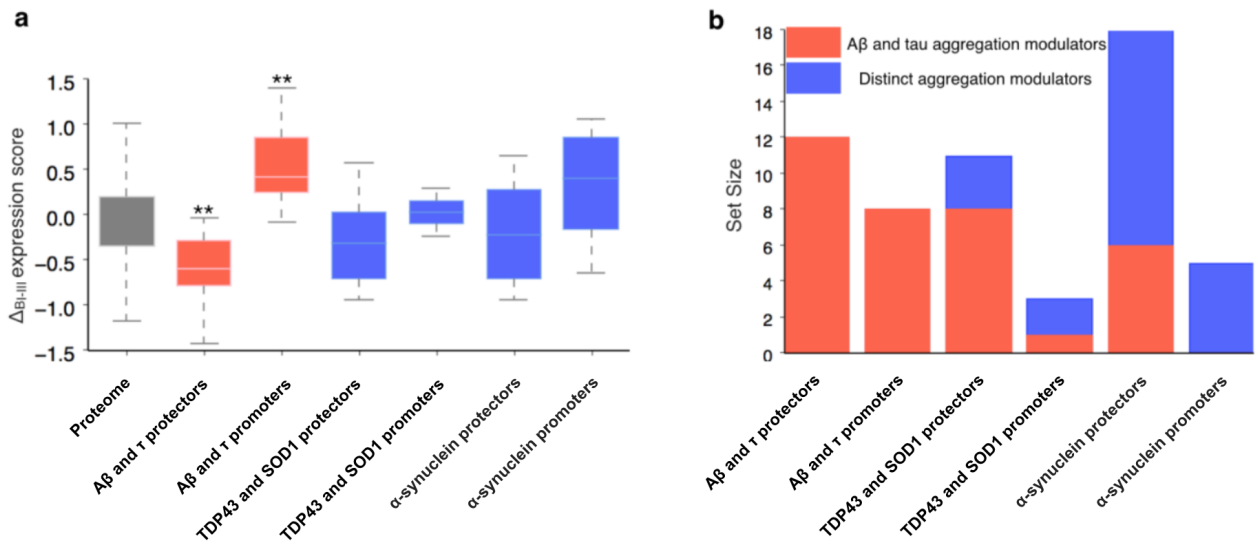

Figure 2.5: Δ_{BI-III} Scores for neurodegenerative disease–specific aggregation modulator sets. (a) Comparison of the distributions of mean Δ_{BI-III} scores for aggregation modulators of Aβ and tau (AD), α-synuclein (PD), and TDP43 and SOD1 (ALS). The only statistically significant differences (**$p < 0.01$), calculated with respect to ΔBI–III scores for the proteome using the Mann-Whitney U test with Benjamini-Hochberg multiple hypothesis testing correction [207], are found for AD-specific aggregation modulators. (b) Size of aggregation modulator sets described in panel a. The aggregation modulators common to the different diseases are shown in red, those distinct are shown in blue.

Region	Allen Brain Atlas code
FTD	
Frontal pole	FP-s, Fpi, FPm
Medial orbital gyrus	MOrG
Superior frontal gyrus	SFG-m, SFG-l
Middle frontal gyrus	MFG-s, MFG-i
Precentral gyrus	PrG-prc, PrG-sl, PrG-il, PrG-cs
Hippocampus	DG, CA1, CA2, CA3, CA4, IG, S
Amygdala	ATZ, AAA, BLA, BMA, CeA, COMA, IA, LA, NLOT, PLA
Cingulate gyrus, frontal part	CgGf-s, CgGf-i
Supramarginal gyrus	SMG-s, SMG-i
PD	
Substantia nigra pars compacta	SNC

Table 2.2: Vulnerable tissues for alternative neurodegenerative diseases
The name and Allen Brain Atlas code for tissues which succumb at the initial stages of FTD and PD are included

2.3 Conclusions

Although protein aggregation is the hallmark of neurodegenerative diseases, no conclusive evidence has been found to pinpoint the origins of disease toxicity to either the presence of insoluble aggregates, or intermediates in the aggregation process [21]. Through a series of studies aimed at investigating this perplexing situation, it is becoming increasingly clear that a defective protein homeostasis system is related to the pathogenic process of neurodegenerative diseases. This was initially evidenced by the appearance of of protein deposits in effected tissues. Building on these observations, we have adopted the strategy of trying to understand the factors which influence vulnerability to neuronal atrophy. We have thus collated the network of aggregation modulators related to the primary aggregators in AD, Aβ and tau. This network includes protein machinery important in peptide cleavage, post translational modification, oligomer sequestration, and misfolded protein clearance. By considering only aggregation modulators with a clear and additive influence on the aggregation process and toxicity potential of Aβ and tau, we were able to investigate a potential role for this subproteome in tissue vulnerability to AD. We have found that relative expression levels of aggregation modulators in normal brain tissue predispose AD-specific aggregation to extents relative vulnerability during disease progression. The contribution of other subproteomes to an environment vulnerable to AD-specific aggregation cannot be excluded (in fact, it could be expected). It is not know at this stage how the proteomic environment interacts with the connectivity of brain regions to determine vulnerability to disease spread one AD inception has occurred. These findings may inspire novel therapeutic approaches for AD, which, rather than trying to prevent a wide range of possible triggering events, could be based on the pharmacological enhancement of our natural defence mechanisms that maintain our proteome in a soluble state in the specific tissues where protein aggregation may take place more readily [29, 210, 211]. In summary, our results illustrate how the origins of variable tissue vulnerability to AD may lie within the proteome through the identification in vulnerable tissues of an intrinsic expression signature associated with protein aggregation, observed decades before the typical age of disease onset.

2.4 Methods

2.4.1 Data sources

The primary data source for this chapter is the Allen Brain Atlas [205]. The Allen Brain Atlas provided data from six healthy human brains from individuals aged 24 to 57 years. Samples were taken from more than 500 regions for each hemisphere, and more than 19,700 genes were studied with multiple probes. Microarray data were downloaded from the Allen Brain Atlas [205]. Data were downloaded from the Allen Brain Atlas on 19 December 2014.

2.4.2 Braak Staging

At progressive clinical stages of AD, conserved patterns of NFT deposition in neural tissues were observed, with increasingly large areas of the brain affected with advancing stages. In the Braak staging of AD [125], tissues were classified according to when, in the progression of AD, NFTs appear in constituent neurons. This method way employed because NFT formation is a pathological hallmark of AD and correlates well with cell atrophy [98]. In Braak stages I and II, NFT involvement is confined primarily to the transentorhinal region of the brain. In stages III and IV, limbic regions are also affected, with all regions of the hippocampus exhibiting AD pathology. In stages V and VI, there is extensive neocortical involvement. Even at late stages of AD, some regions of the brain, notably the cerebellum, remain unaffected; we classified these regions as “non-Braak.” In the original paper describing the Braak staging [125], disease stages were discussed sequentially, with the regions affected noted at each stage, in addition to the severity of the pathology in these regions.

2.4.3 Mapping with the Allen Brain Atlas

To assign the brain regions from the Allen Brain Atlas to the correct Braak stage, a rubric was developed. First, regions mentioned in the original paper [125], which we refer to as “Braak staging” regions, were assigned to the earliest Braak stage that they are associated with. Next, these regions were matched to the regions characterised in

the Allen Brain Atlas. When a region in the Allen Brain Atlas was larger than a Braak staging region, the whole of the Allen Brain Atlas region was allocated to the corresponding Braak stage. For instance, although only two isocortical layers were affected in Braak stage III, all isocortical tissues were assigned to Braak stage III because isocortical expression data were not distinguished by the layer they came from in the Allen Brain Atlas parcellation. For this reason, when investigating the relationship between Braak staging and expression signature, "perfect-match" regions provide the most accurate data. Perfect-match tissues have a high correspondence between their Braak and Allen Brain Atlas perimeters. Assignments from Braak to Allen Brain Atlas regions are listed in Table 2.3. Trends can be seen for perfect match and all tissues for both aggregation protectors are promoters comparing between Δ scores for Braak I-III and Braak IV-VI tissues, in Figure 2.6 a-b.

Of the two main types of tissue in the brain, white matter consists mostly of glial cells and myelinated axons, whereas grey matter has a more diverse composition, including neuronal cell bodies, dendrites, myelinated and unmyelinated axons, glial cells, synapses, and capillaries. Thus, because NFTs are not found in AD in the axon hillock or initial axon segment, one would not expect to see them in white matter in AD [212]. However, the volume of white matter does shrink in some regions during the progression of AD, where affected neurons have their cell bodies in grey matter and their axons in white matter. This fact implies that NFT appearance, and thus Braak staging, may not be ideal for describing vulnerability to AD in white matter tissues. However, the effect of this caveat on our study is limited because only 2 of the more than 500 regions studied in the Allen Brain Atlas include white matter and is accounted for in Figure 2.6c.

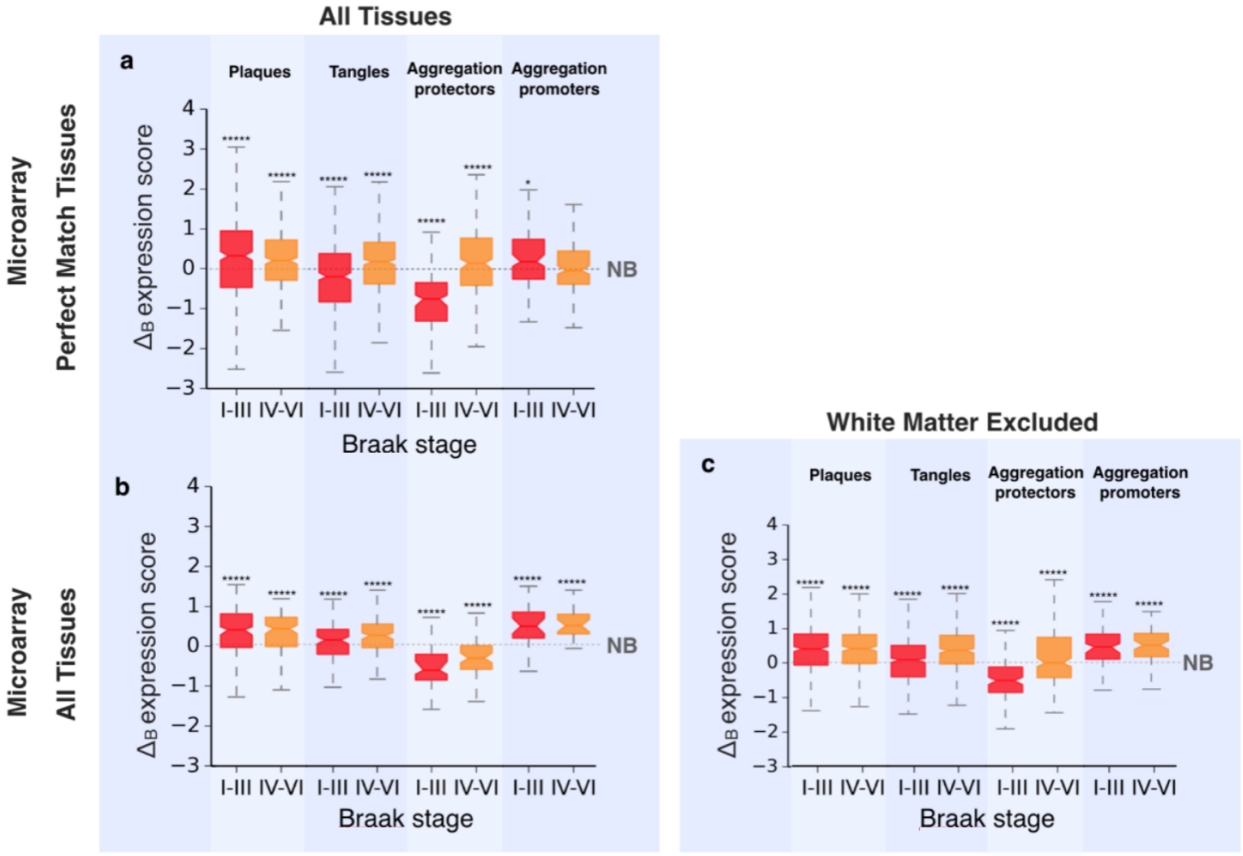

Figure 2.6: Distributions of ΔB scores for tissues affected at different Braak stages. (a) Boxplots of the ΔB scores for Aβ and tau aggregation modulators, and the co-aggregators in plaques and tangles calculated for 'perfect match' tissues (see table 2.3) progressively affected by AD (x-axis). (b) Same data as for panel a, but for all tissues, in this case the disease progression is less accurately mapped onto many the Allen Brain Atlas tissues. c) A comparison of ΔB score distributions with white matter data included and excluded from analysis. For each boxplot, the significance of the difference with the ΔB distribution for non-Braak (NB) tissues was calculated with a Mann-Whitney U test and Benjamini-Hochberg multiple hypothesis testing correction [207] *p < 0.05, ***p < 0.001, *****p<0.00001.

Tissue from Braak et al., 1991	Allen Brain Atlas region	Allen Brain Atlas code	Braak stage
CA1	CA1	*CA1*	I
Transentorhinal and entorhinal cortices	Parahippocampal gyrus	PHG-l, PHG-cos	I
Anterodorsal mucleus of the thalamus	Anterior group of nuclei	DTA	I
Magnocellular nuclei of the basal forebrain	Basal forebrain - includes septal nuclei, and nuclei within the substantia innominata	SptN, SI	I
Layers III and V in basal portions of frontal, temporal, and occipital asscoaition areas	Frontal, temporal, and occipital association areas.	AOrG, fro, FP-s, FP-l, Gre, trIFG, opIFG,orIFG, IRoG, LOrG, MOrG, MFG-s, MFG-i, PCLa-s, PCLa-i, PTG, POrG, SFG-m, SFG-l, SRoG, LiG-pest, LiG-str, Cun-pest, Cun-str, IOG, OP, SOG-s, SOG-i, OTG-s, OTG-i, FuG-l, FUG-its, FuG-cos, HG, ITG-mts, ITG-l, ITG-its, MTG-s, MTG-i, PLP, PLT, STG-i, STG-l, TP-s, TP-i, TP-m, TG	II
Subiculim	Subiculum	*S*	II
CA2	CA2	*CA2*	III
CA3	CA3	*CA3*	III
CA4	CA4	*CA4*	III
Fascia dentata	Dentate gyrus	*DG*	III
Cingulate gyrus	Cingulate gyrus	CgGp-s, CgGp-i, CgGf-s, CgGf-l, CgGr-i, CgGr-s, SCG	III
Amygdala	Amygdala	*ATZ, BLA, BMA, CeA, COMA, LA*	III
Reuniens nucleus of thalmus	Medial group of nuclei	DTM	IV
Hypothamic tuberomamillary nucleus	Tuberomamillary nucleus	TM	IV
Primary sensory areas of isocortex layer III	Primary sensory and motor cortices + any other isocortical regions that haven't been accounted for	PrG-sl, PrG-il, PrG-prc, PrG-cs, PoG-il, PoG-sl, PoG-cs, PoG-pcs, SMG-s, SMG-i, AnG-i, AnG-s , PCLp-l, PCLp-cs, Pcu-i, Pcu-s, SPL-s, SPL-i,	IV
Basal portions of the putamen	Putamen	*Pu*	IV
Accumbens nucleus	Nucleus accumbens	*Acb*	IV
Basal portions of the claustrum	Claustrum	Cl	IV
Substantia nigra pars compacta	Substantia nigra pars compacta	SNC	V
Lateral tuberal nucleus of the hypothalamus	Lateral tuberal nucleus	LTu	V
Antero-basal portions of the insula	Insula	LIG, SIG	V
Reticular nucleus of the thalamus	Reticular nucleus of the thalamus	R, ZI, FF	VI

Table 2.3: A correspondence between Allen Brain Atlas and Braak Stage tissues

All tissues assigned in the Braak staging of AD (Braak et al. 1991), and the corresponding regions in the Allen Brain Atlas (Hawrylycz et al. 2012) are listed. Codes are given to allow the identification of the specific regions in the Allen Brain Atlas database. Tissues which are not listed in this table, and therefore not noted as succumbing to AD in Braak's 1991 paper, are considered non-Braak in this study. Italics are used to denote 'perfect match' tissues.

2.4.4 Quantifying differential expression - the Δ score

Below we describe the Δ score methodology for analysis in this chapter, and chapter 3. Chapter 4 uses a number of different Δ score metrics, to study relative expression at a cellular and subcellular level, and to measure relative expression in the context of other types of neurodegenerative disease. More details of these other delta score calculation methodologies are provided in the methods section of chapter 4.

Because the expression of a given gene in the Allen Brain Atlas is measured by multiple probes [205], we first normalised the expression reading $E_{p,r}$ for each probe p in each region r in the Allen Brain Atlas as

$$E'_{p,r} = \frac{E_{p,r} - \mu_p}{\sigma_p} \quad (1)$$

where μ_p and σ_p are the average and SD of $E_{p,r}$ across all regions, respectively.

To quantify the variability of gene expression between tissues, we defined a Δ score (Figure 2.6) for a given probe p over a brain region R (which is typically made up of several Allen Brain Atlas regions) as

$$\Delta_{p,R} = \overline{E'_{p,R}} - \overline{E'_{p,\ \mathrm{NB}}} \tag{2}$$

where

$$\overline{E'_{p,\ \mathrm{NB}}} = \frac{1}{N_{\mathrm{NB}}} \sum_{r=1}^{N_{\mathrm{NB}}} E'_{p,r} \tag{3}$$

is the average of $E'_{p,r}$ for the non-Braak regions, that is, those N_{NB} regions that do not map onto any Braak staging regions [125], and

$$\overline{E'_{p,R}} = \frac{1}{N_R} \sum_{r=1}^{N_R} E'_{p,r} \tag{4}$$

represents the average of $E'_{p,r}$ over the set of Allen Brain Atlas regions mapped onto brain region R under scrutiny (for example, a Braak stage). Then, Δ scores for different probes measuring the same gene were averaged to give a gene Δ score

$$\Delta_{g,R} = \frac{1}{N_{p(g)}} \sum_{g=1}^{N_{p(g)}} \Delta_{R,p(g)} \tag{5}$$

where the average is over the $N_{p(g)}$ probes $p_{(g)}$ used to measure a gene g.

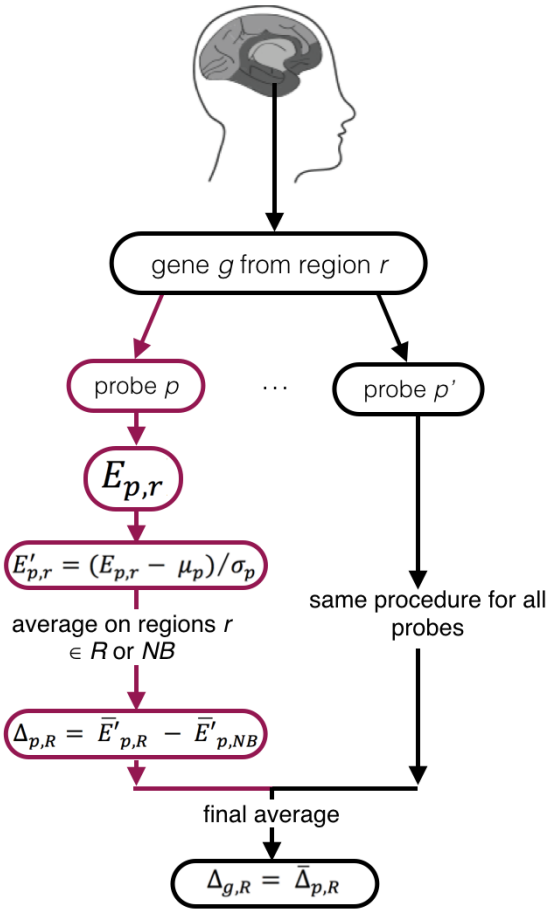

Figure 2.6: The Δ calculation process in more detail. Multiple probes are used to measure the expression of each gene. For each probe p, expression is measured for each region r, in the brain. Expression values for a probe are normalised across all tissues, by subtracting mean expression, and dividing by the standard deviation of expression. A probe Δ score is found by subtracting mean normalised expression in non-Braak regions (NB), from mean expression in a regional subset of interest. Finally, a Δ score for each gene is calculated by averaging Δ scores for the probes used to measure its expression.

2.4.5 Calculation of the relative expression for aggregation regulators

We undertook an unbiased literature search to identify all molecular chaperones and post-translational modifiers reported to affect the aggregation of Aβ or tau (Table 2.1). These aggregation regulators were sorted into three groups: (i) proteins that protect against aggregation (protectors), (ii) proteins that promote aggregation (promoters), and (iii) proteins whose cumulative role on aggregation was ambiguous in the current literature. We evaluated relative expression for proteins in groups 1 and 2 by calculating the Δ score distribution at consecutive Braak stages.

2.4.6 Evaluation of statistical significance

To assess the differences in the distributions of Δ scores between various data sets, we used the nonparametric Wilcoxon/Mann-Whitney U test [213], or a two-tailed t test [206], as specified in the figure captions. Because of the high number of data and hypotheses tested in this study, we adjusted the P values to reduce the false discovery rate (FDR). Specifically, for Figures 2.3-2.6 we used the Benjamini-Hochberg multiple hypothesis testing correction to control the FDR [207] because this method allows the cost paid for the control of multiplicity to be kept relatively low. More generally, from the analysis of the relationship between FDR, sensitivity, and study sample size [214, 215], it is known that microarray studies can be susceptible to large FDR, which, besides measurement variability, is primarily determined by the proportion of truly differentially expressed genes, the magnitude of the true differences, and sample size. Because our work relies on 3700 microarray studies (up to 900 samples from six brains), the FDR rate analysis was performed on a relatively large sample size, allowing for rather sensitive detection of truly differentially expressed genes. We further increased the statistical significance of the results and avoided a high false-negative rate by calculating the significance of the difference of Δ score distributions for groups of genes. In comparison to calculating the significance of the differences of Δ scores of individual genes, this approach greatly reduced the number of hypotheses in our study. These tests were performed using the SciPy and rpy2 modules for Python.

2.4.7 Shading of cortical and subcortical brain structures on three-dimensional representation

Figure 2.3 was created using a set of three-dimensional meshes of a human brain, which were constructed from 12 volumes acquired using magnetic resonance imaging [216-218]. Images were colored using the computer graphics software Blender.

3 The metastable subproteome

In the previous chapter, we have shown that tissues vulnerable to AD can be characterised by expression of the proteins which co-aggregate to form plaques and tangles, and the aggregation modulators of Aβ and tau. Together, the expression levels of these subproteomes provide a rationalisation of the specific proteins which aggregate during AD [153]. To build on this discovery, we wanted to understand if there were other subproteomes, whose expression level placed a burden on the protein homeostasis system, creating an environment vulnerable to disease-specific aggregation.

3.1 Introduction

3.1.1 A broader view on protein homeostasis

In this chapter, we build on the observation that vulnerability to disease-specific aggregation in neurodegenerative diseases such as AD could exist at a number of levels (Figure 3.1). We thus predicted that elevated levels of the primary aggregators in AD - Aβ and tau - would exist in the tissues which first succumb to AD. A number of proteins are found to co-aggregate in plaques and tangles alongside Aβ and tau respectively. We therefore selected AD co-aggregators as a secondary point of interrogation for investigating the molecular signature of AD vulnerability. Considering the protein aggregation process from a broader perspective, it is also possible that protein supersaturation plays a role in neuronal predisposition to aberrant aggregation. Supersaturated proteins (Section 1.2.2) are expressed beyond their solubility limit in tissues and have been found to be enriched in the deposits found in neurodegenerative disease brains, including plaques and tangles [40]. In addition to directly aggregating in disease, it is possible that high levels of supersaturated proteins put pressure of the protein homeostasis system of neurons, thereby enabling the unchecked misfolding of primary aggregators. Work by Morimoto et al. has experimentally demonstrated this process using a model of neurodegenerative disease in *C. elegans* and the over-expression of a protein known to be aggregation-prone [46].

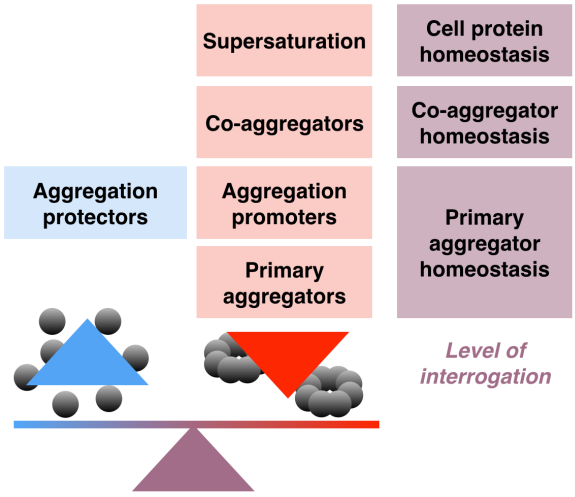

Figure 3.1: Levels of proteomic interrogation for investigation of vulnerability to disease-specific protein homeostasis loss.

3.1.2 Elucidation of a supersaturated subproteome specific to AD

Supersaturated levels were first calculated and later refined by Ciryam et al. [40]Expression levels were taken from experimental measurements across a number of tissues. We therefore sought to identify a subproteome of supersaturated proteins enriched in the peptides directly responsible to vulnerability to aberrant aggregation in AD. During AD progression in the brain, a robust down regulation of particular proteins has been observed. Strikingly, at the pathway level a correlation has been found between supersaturation level and extent of down regulation in AD [41]. It has been suggested that this relationship exists due to a protective response by diseased neurons to reduce the concentration of their most aggregation-prone proteins. A subset of the proteome is also down-regulated with age. Although a number of proteins are shared between down-regulated subsets, this is a distinct and robust process. The observation that proteins downregulated in age are enriched in supersaturated proteins has lead to the suggestion that this transcriptomic trend is also an innate cytoprotective mechanism to preserve protein solubility as molecular chaperone capacity declines. To identify the subset of supersaturated proteins most critical to AD-specific aggregation risk, we therefore found the intersection of supersaturated proteins, proteins which down-regulate in age, and proteins which down-regulate in AD. We will refer to the proteins in this intersection as the metastable subproteome .

3.2 Results

3.2.1 Relative expression of primary aggregators and co-aggregators in AD-vulnerable tissues

To determine if the brain tissues most vulnerable to neurodegenerative disease had an expression signature for the subproteomes identified in section 3.1, we utilised data from the Allen Brain Atlas, and the Δ score method to quantify differential expression. The Allen Brain Atlas is an online repository of microarray data for six normal human brains of ages 24-56. Data has a high level of granularity, over 500 measurements are taken for each individual, with 93% of known genes represented by at least two probes. The Δ score quantifies the expression of any given gene, in any given region, relative to expression in tissues completely resistant to the neurodegenerative disease under study (see section 2.4 for further details). The progression of AD through the brain was determined using Braak staging [125], which was developed through histopathological examination of the tangle profile in the brain during subsequent stages of AD, tangle deposition has since been shown to correlate with neuronal atrophy.

Our analysis revealed the presence of elevated expression levels of proteins that coaggregate in plaques and tangles in the tissues in which AD is first evident, as measured by the average ΔBI–III score (the Δ score for Braak regions I to III) (Figure 3.2 and Appendix B2). We tested the statistical significance of these results by calculating the ΔBI–III scores of 10^6 random sets of genes of equal size and comparing them to that of Aβ and tau aggregation modulators (Figure 3.3). We find that within the resultant distribution, the mean Δ score for co-aggregators in each case is at the upper tail of the distribution.

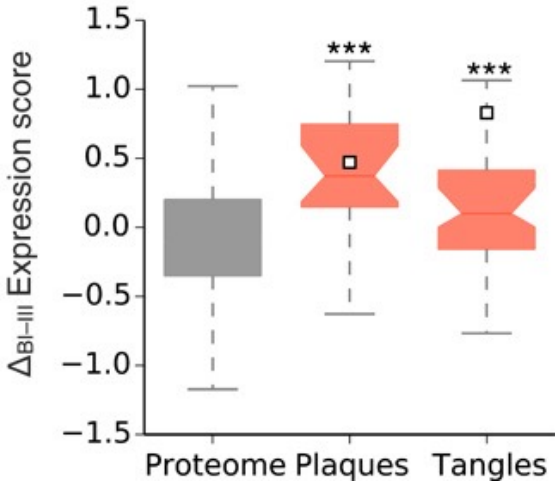

Figure 3.2: Tissue-specific transcriptional analysis of a subset of aggregation-prone proteins specific to AD.

Box plot of ΔBI–III (the Δ score for Braak regions I to III) for the whole proteome and the proteins that coaggregate with Aβ and tau in plaques and tangles. Aβ and tau are shown as square points in their respective distribution. ***$P < 0.001$; the statistical significance of the difference between the distributions of the coaggregators and that of the proteome was calculated using Mann-Whitney *U* test with Benjamini-Hochberg multiple hypothesis testing correction [207].

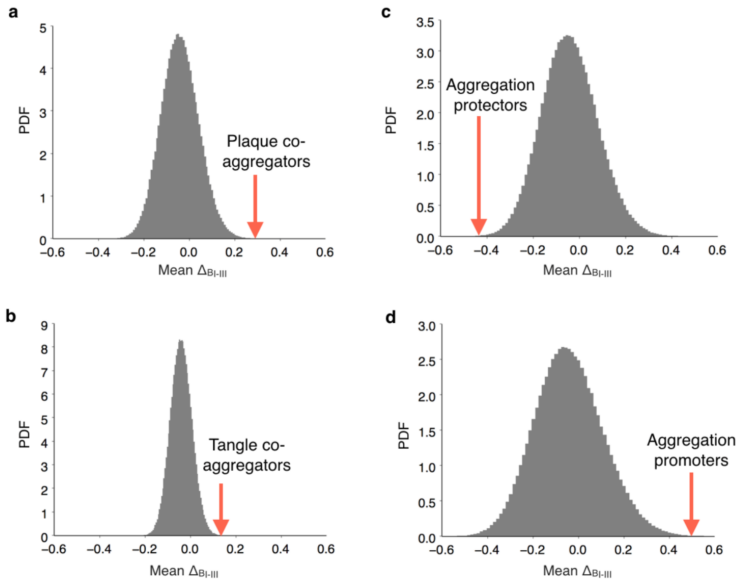

Figure 3.3: Analysis of ΔBI–III scores of random sets of genes. (a,b) Probability distribution functions (PDFs) of the ΔBI–III scores for 10^6 random gene sets of size equal to those corresponding to proteins that co-aggregate with Aβ in plaques (plaque co-aggregators) (a), and of that of proteins that co-aggregate with tau in tangles (tangle co-aggregators) (b); arrows indicate the mean ΔBI-III scores of plaque and tangle co-aggregators. (c,d) Probability distribution functions of the mean ΔBI–III scores for 10^6 random gene sets of size equal to that of Aβ and tau aggregation promoters (c), and protectors (d); arrows indicate the mean ΔBI-III scores of the aggregation modulators of Aβ and tau.

To test the ability of co-aggregator expression levels to differentiate tissues of increasing resistance to AD, we plotted the Δ scores for these subproteomes for tissues in Braak stages I-VI (Figure 3.4). For plaque co-aggregators, we find consistently elevated expression relative to non-AD tissues, and we find that tangle co-aggregator relative expression is not predictive of differential vulnerability between tissues known to succumb to the disease. This result indicates that whilst elevated co-aggregator expression (particularly with respect to plaques) is necessary for AD to develop in a given tissue, it is not sufficient to determine the primary site of disease onset, nor relative vulnerabilities between tissues.

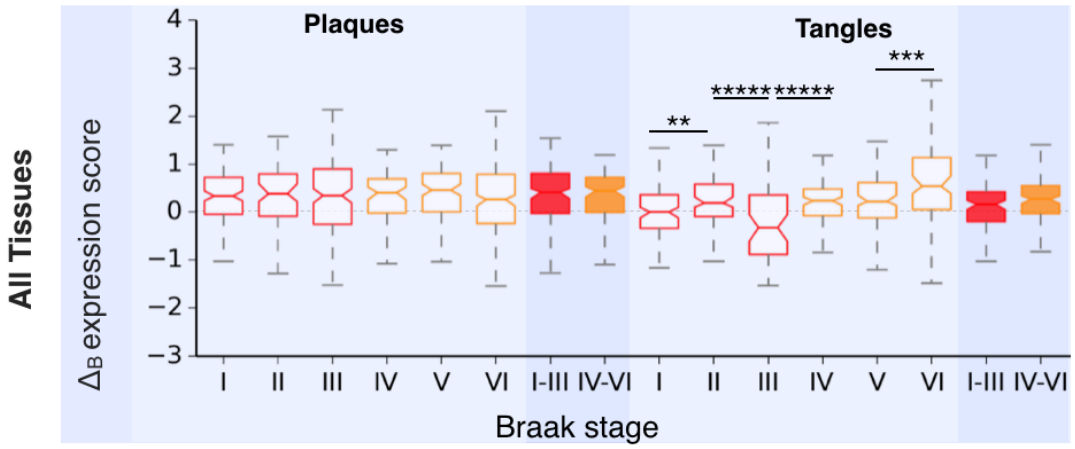

Figure 3.4: Distributions of ΔB scores for tissues affected at different Braak stages. Boxplots of the ΔB scores for the co-aggregators in plaques and tangles calculated for tissues progressively affected by AD (x-axis) For each boxplot, the significance of the difference with the prior ΔB distribution was calculated with a Mann-Whitney U test and Benjamini-Hochberg multiple hypothesis testing correction *p < 0.05, ***p < 0.001, *****p<0.00001 [207].

3.2.2 Investigating the metastable subproteome specific to AD

Having found an association between AD co-aggregators and tissue vulnerability to aberrant aggregation in disease, we next sought to identify a disease specific sub-proteome. We predicted that at high levels, this sub-proteome would create an environment conductive to AD-specific aggregation. We identified this metastable subproteome by finding the intersection between highly supersaturated proteins, proteins down-regulated in ageing, and proteins down-regulated in disease (Figure 3.5a). Using Δ score analysis, we find that this metastable subproteome has significantly elevated expression in the tissues most vulnerable to AD (Braak stage I-III tissues), (Figure 3.5b). In the previous chapter, we found that whilst co-aggregators found in plaques and tangles in AD brains also had significantly elevated expression in early AD tissues, they could not predict differential vulnerability between tissues. We assessed this

differential expression sensitivity by measuring the correlation for each gene in the subset between tissue AD vulnerability, and relative expression, as assessed using Δ score analysis. We find a significant correlation between tissue disease vulnerability to disease, and gene expression, for the metastable subproteome. The co-aggregators found in plaques and tangles show some correlation, but not at a level of significance. The co-aggregators found in Lewy bodies show no correlation, as expected (Figure 3.5c).

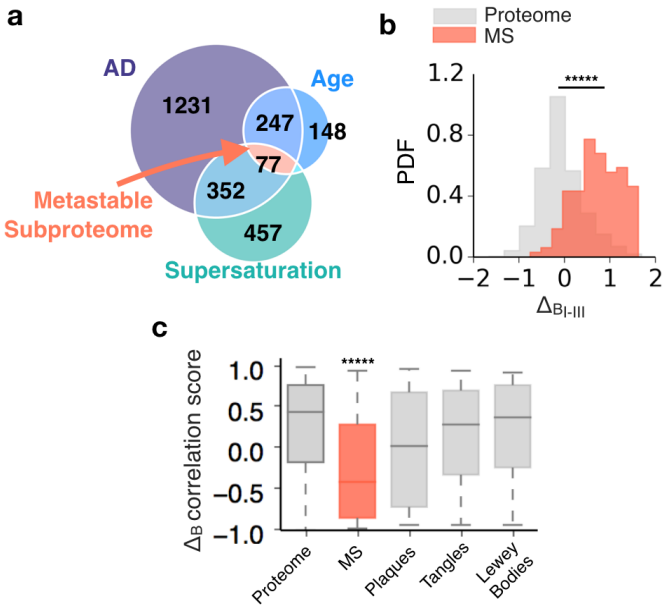

Figure 3.5: Expression of metastable subproteome specific to AD is elevated in AD-vulnerable tissues. (a) We defined an AD-specific metastable subproteome at risk of aggregation as the intersection of classes of proteins downregulated with age, downregulated in AD, and supersaturated. **(b)** Bar plot the Δ score for AD-vulnerable tissues averaged across all genes for all the possible intersections (x-axis) illustrated in panel a. **(c)** Distribution of the Δ score of genes comprising the metastable subproteome (MS, red) and the whole proteome (grey). ***** $p < 0.00001$, in c the statistical significance of the difference with $\Delta_{BI\text{-}III}$ averaged across the MS (first column) was estimated with a one-tailed t-test and in d the significance of the difference was estimated with a Mann Whitney U-test [207].

To check that the process by which the metastable subproteome was calculated was enriching for proteins relevant to AD vulnerability, we calculated the mean Δ score for all possible intersections illustrated in Figure 3.5a. The results are presented in Figure 3.6, and show that the metastable has the highest mean Δ score for tissues most vulnerable to AD.

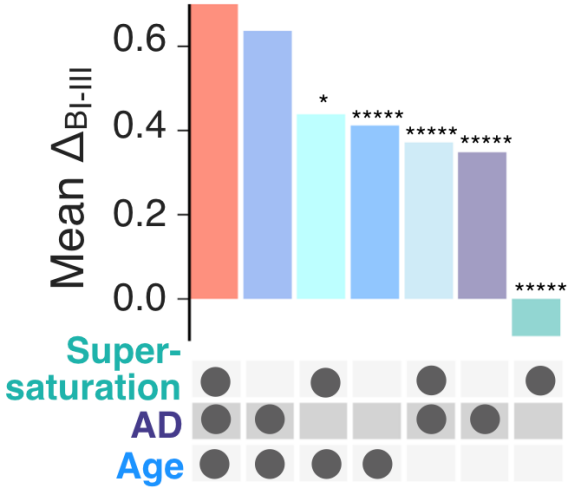

Figure 3.6: Mean Δ Braak I-III score for MS-component subproteome intersections. The mean Δ score for subproteomes comprised of proteins in different possible intersections of supersaturated proteins, and proteins downregulated in ageing and disease. These intersections are illustrated in figure 3.5a. *p < 0.05 and *****p<0.00001, 1 tailed t-test [206].

To test the robustness of these conclusions, we considered the overlap between highly supersaturated proteins and proteins with high relative expression in Braak I-III regions. By taking the intersection of the top 5% of these two subsets, we found a subset of proteins that are highly downregulated in AD (Figure 3.7). These results are fully consistent with those discussed above, and indicate that by analysing two properties that can be measured in health brains (supersaturation and high relative expression in early Braak regions), one can make predictions about proteins associated with the translational response to AD.To evaluate the extent of residual noise, we calculated the correlation between probe Δ scores. We found a good agreement between Δ scores for two probes measuring the same gene, the Pearson's correlation coefficient for this relationship is 0.7 (Figure 3.7d)

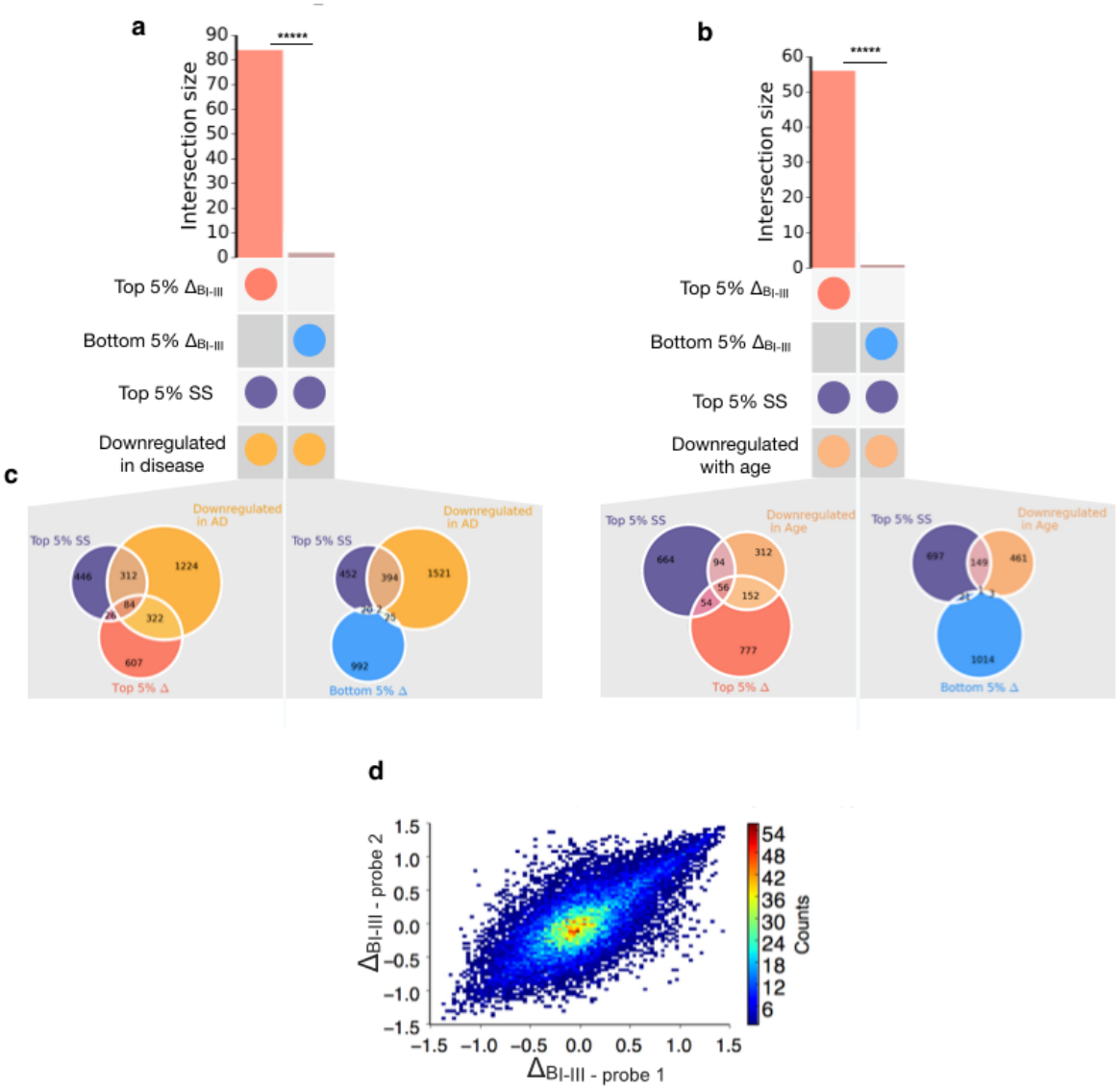

Figure 3.7: Testing Δ score signal robustness.

(a) Difference in overlap size between proteins downregulated in disease and supersaturated, and proteome members with top and bottom 5% Δ scores (b) Difference in overlap size between proteins downregulated in ageing and supersaturated, and proteome members with top and bottom 5% Δ scores (c) the sizes of all intersections related to (a) and (b). In (d) the correlation between Δ scores for two probes measuring the same gene was tested, resulting in a coefficient of correlation of 0.7.*****p<0.0001, Fisher exact test performed for panels (a) and (b).

3.2.3 Enrichment analysis of the metastable subproteome

We next sought to understand the physiological role of the metastable subproteome, in the cell. Using enrichment analysis of both subcellular localisations and biological processes, we find that the metastable subproteome is enriched for pathways known to be important for synaptic function; ion transport, mitochondrial energy metabolism, and synaptic transmission (Figures 3.8-3.9). It is interesting to note that synapses are the first subcellular location to degenerate during AD.

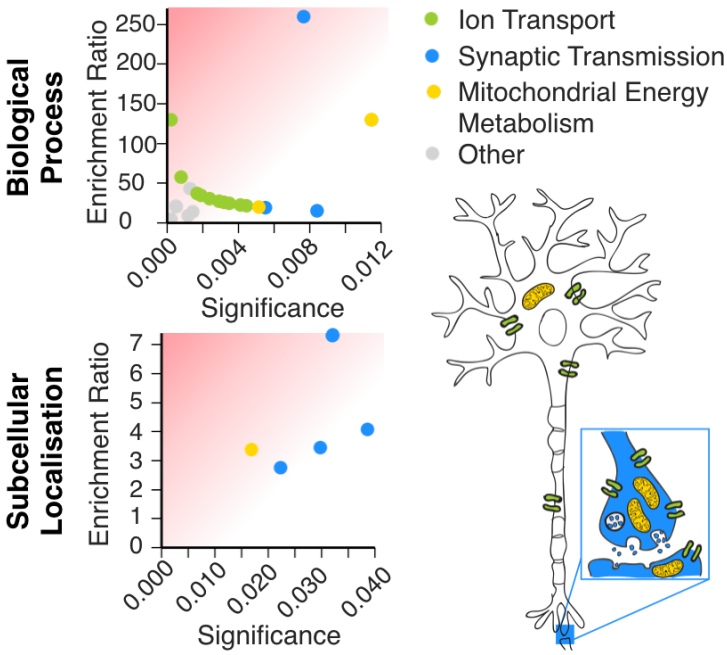

Figure 3.8: The metastable subproteome is enriched for synaptic localisation. (a) List of the eight most enriched biological processes. **(b)** List of the eight most enriched subcellular localisations. Enrichment ratios were calculated as the ratio between observed and expected values for each category (see Methods). All significantly enriched biological processes and subcellular localisations are reported in Figure 3.9.

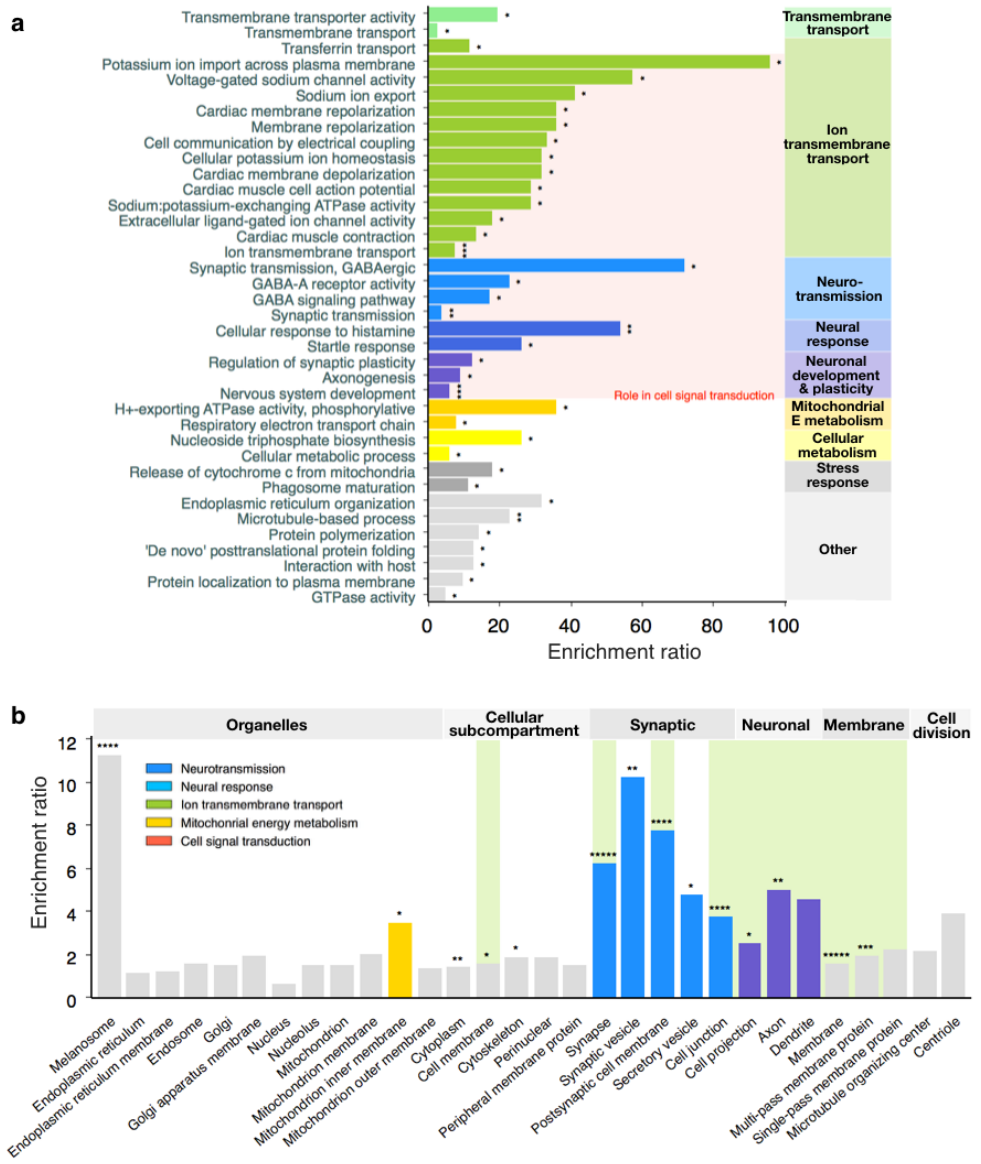

Figure 3.9: All significantly enriched metastable subproteome biological processes (a) and subcellular localisations (b).

Other common subcellular localisations are included for context.

*p < 0.05, **p < 0.01, ***p < 0.01, ****p < 0.001 and *****p<0.0001 one tailed Fischer exact test and Bonferroni multiple hypothesis correction [207].

3.2.4 Evaluating the predictive power of metastable subproteome levels with respect to AD progression in the brain

Finally, we sought to asses whether expression levels of the metastable subproteome in normals brains decades before disease onset, could predict the progression of AD. We find using Δ score analysis, that the most vulnerable tissues to AD (early Braak tissues), have significantly elevated relative expression levels of metastable subproteome proteins. In addition, we find that relative expression declines in tissues with progressively more AD resistance, who succumb at later stages of the disease (late Braak tissues), (Figures 3.10 a,b, full dataset in Appendix B3).

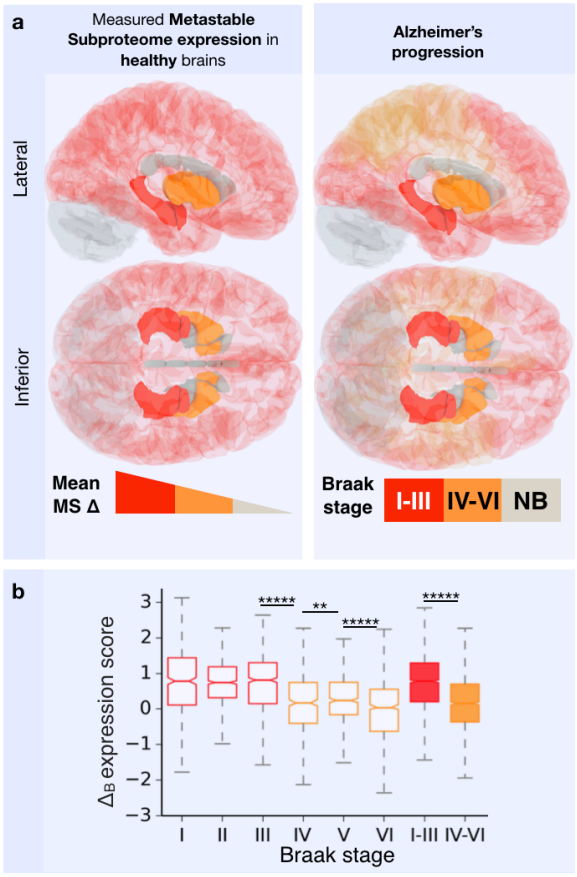

Figure 3.10: Metastable subproteome expression in normal tissues recapitulates AD progression. a) In the left panel, regions are coloured by the mean Δ score for metastable subproteome expression. In the right panel, regions are coloured by Braak stages. **b)** Bar plot with the mean Δ scores for the MS in 'perfect match' regions affected at the different Braak stages (x-axis). Perfect match tissues have a perfect correspondence between their Braak and Allen Brain Atlas perimeters. **p < 0.01 and *****p<0.00001 p-values for **b,** calculated with Mann-Whitney U test with Benjamini-Hochberg multiple hypothesis testing correction [207].

3.3 Conclusions

In this chapter we have identified a number of subproteomes in addition to the aggregation modulators discussed in Chapter 2, which underlie tissue vulnerability to AD. We have thus broadened our understanding of the molecular origins of susceptibility to aberrant aggregation in this disease. Although we have not quantified the relative influence of each subproteome on vulnerability, by looking at Δ score distributions at the granularity level of individual Braak stages, we have been able to compare the ability of these subproteomes to differentiate tissues of varying disease susceptibility. We have found that Aβ and tau, the primary aggregators in AD, both show elevated expression in the the most highly AD-vulnerable tissues. The small number of data points available limits our ability to examine this trend at the level of individual Braak stages or to meaningfully evaluate statistical significance. With the evolution of experimental techniques, we recommend this analysis is revisited when larger data sets are available. Looking beyond these primary aggregators, we evaluated the Δ scores for proteins which co-aggregate in smaller proportion with Aβ and tau in plaques and tangles respectively. We have found that these sub-proteomes do show significantly elevated expression in the most AD-vulnerable tissues, relative to those which are completely resistant. However, at an individual Braak stage level, Δ score distributions did not reflect the differential vulnerability of tissues to disease. This result indicates that whilst high levels of deposit co-aggregators create an environment conducive to AD-specific aggregation, they are not the critical factor which triggers disease onset in a given tissue. We next turned our attention to the aggregation propensity of the proteome as a whole in neurons, making use of a biophysical characteristic called supersaturation. We have defined an AD-specific metastable subset of proteins, which we found to be enriched for synaptic function. Crucially, we have found that the expression of this metastable subproteome in healthy tissue predicts the progression of AD through the brain.

Taken together, these results suggest a model for AD vulnerability at the molecular level. Here, the differential vulnerabilities of tissues to protein homeostasis loss are controlled by the availability of specific protein homeostasis components which control

primary aggregators, and the levels of disease-specific supersaturated proteins, which put pressure on the protein homeostasis system, enabling the escape and aggregation of misfolding primary aggregators. We also note that it is quite possible that we see an elevated Δ score signal for the MS specifically (see Figure 3.6) because the protein homeostasis networks that it interacts with are shared with those of the proteins which aggregate to form plaques and tangles in AD. This possibility would indicate the presence of disease-specific metastable subproteomes for other neurodegenerative diseases.

We will explore the possibility of a molecular signature for vulnerability in other neurodegenerative diseases associated with protein misfolding in the next chapter. We will also look at levels of analysis beyond the tissues level, analysing relative expression from a cellular and subcellular perspective.

3.4 Methods

3.4.1 Data sources

The primary data source for this chapter is the Allen Brain Atlas [205]. Details of the Allen Brain Atlas data and download date are given in Methods section 2.4.1. UniProt data [219] for subcellular localisation and biological process gene ontology assignments for most proteins were downloaded from www.uniprot.org/downloads on 21 May 2015. Protein IDs were converted between UniProt and Entrez ID (used by the Allen Brain Atlas) using the UniProt ID mapping service. With this procedure, expression data were assigned to about 90% of the human reference proteome.

3.4.2 Braak Staging

At progressive clinical stages of AD, conserved patterns of NFT deposition in neural tissues were observed, with increasingly large areas of the brain affected with advancing stages. In the Braak staging of AD [125], tissues were classified according to when, in the progression of AD, NFTs appear in constituent neurons because NFT formation is a pathological hallmark of AD and correlates well with cell atrophy. More

details of the use of Braak staging to map the progression of AD through brain tissue can be found in Methods section 2.4.2.

3.4.3 Mapping with the Allen Brain Atlas

To assign the brain regions from the Allen Brain Atlas to the correct Braak stage, a rubric was developed. More details on the mapping process can be found in Methods section 2.4.3. A summary of the final Allen Brain Atlas tissues allocated to each Braak stage can be found in Table 2.3. 'Perfect match' tissues denote tissues where the tissue studied in the Allen Brain Atlas analysis [205], was an exact correspondence to or smaller than the tissue allocated to a given Braak stage [125]. We include in Figure 3.11 an evaluation of the Δ score distribution for metastable subproteome proteins, considering all tissues, and the subset of prefect match tissues. We see an agreement in the trend in each case, with metastable subproteome expression most elevated relative to non-Braak tissues in tissues belonging to Braak stages I-III. Expression relative to non-Braak tissues is also elevated in tissues which succumb at a later stage in disease progression (Braak IV-VI tissues), but has a significantly lower median Δ score relative to Braak I-III tissues. The presence of a consistent trend across all brain tissues studied is therefore shown to exist both in the wider data subset and in perfect match tissues.

For technical reasons discussed in detail in Methods section 2.4, it was also important to ensure that the trend observed for metastable subproteome Δ scores was consistent when white matter tissues were excluded. The results of this control can be found in Figure 3.11c, and indeed there is significant and decreasing trend in relative expression as one examines consecutively less AD-vulnerable tissues.

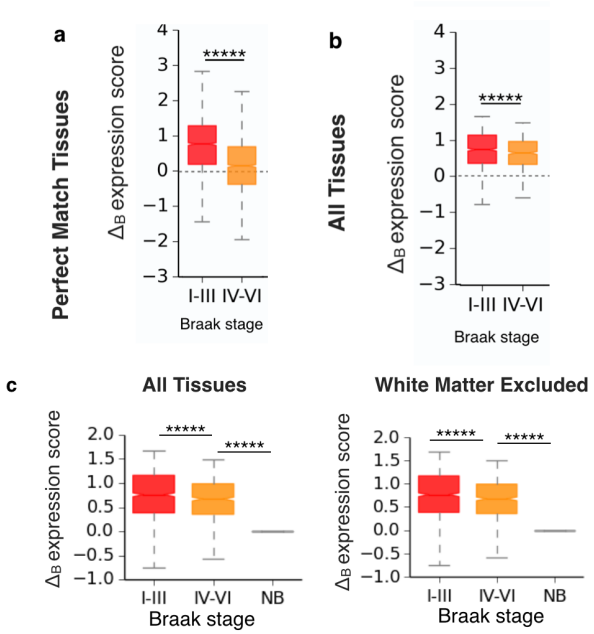

Figure 3.11: Distributions of ΔB scores for tissues affected at different Braak stages. (a) Boxplots of the ΔB scores for metastable subproteome calculated for 'perfect match' tissues (see table 2.3) progressively affected by AD (x-axis). (b) Same data as for panel a, but for all tissues, in this case the disease progression is less accurately mapped onto many the Allen Brain Atlas tissues. c) A comparison of ΔB score distributions with white matter data included and excluded from analysis. For each boxplot, the significance of the difference with the ΔB distribution for non- Braak (NB) tissues was calculated with a Mann-Whitney U test and Benjamini-Hochberg multiple hypothesis testing correction [207] *$p < 0.05$, ***$p < 0.001$, *****$p<0.00001$.

3.4.4 Quantifying differential expression - the Δ score

The Δ score used in analysis in this chapter quantifies expression of any given gene, in any given tissue, relative to expression in tissues resistant to AD. More detail on the Δ score calculation methodology can be found by referring to section 2.4.4 and Figure 2.6.

3.4.5 Evaluation of statistical significance

To assess the differences in the distributions of Δ scores between various data sets, we used the nonparametric Wilcoxon/Mann-Whitney U test, or a two-tailed t test, as specified in the figure captions. Because of the high number of data and hypotheses tested in this study, we adjusted the P values to reduce the false discovery rate (FDR). Specifically, for Figures 3.2, 3.4, and 3.9-3.11 we used the Benjamini-Hochberg multiple hypothesis testing correction to control the FDR because this method allows the cost paid for the control of multiplicity to be kept relatively low. More generally, from the analysis of the relationship between FDR, sensitivity, and study sample size, it is known that microarray studies can be susceptible to large FDR, which, besides measurement variability, is primarily determined by the proportion of truly differentially expressed genes, the magnitude of the true differences, and sample size. Because our work relies on 3700 microarray studies (up to 900 samples from six brains), the FDR rate analysis was performed on a relatively large sample size, allowing for rather sensitive detection of truly differentially expressed genes. We further increased the statistical significance of the results and avoided a high false-negative rate by calculating the significance of the difference of Δ score distributions for groups of genes. In comparison to calculating the significance of the differences of Δ scores of individual genes, this approach greatly reduced the number of hypotheses in our study. These tests were performed using the SciPy and rpy2 modules for Python. In additional to traditional statistical tests such as the t-test and Mann Whitney U-test, we tested the statistical significance of results in Figure 3.3 by calculating the $\Delta B_{I–III}$ scores of 10^6 random sets of genes of equal size and comparing them to that of Aβ and tau aggregation modulators (Figure 3.4). This allowed the verification of the accuracy of other statistical tests used in the paper, and to confirm that calculated p-values were not being distorted by the high number of data points used in the analysis.

3.4.6 Shading of cortical and subcortical brain structures on three-dimensional representation

Figure 3.10 was created using a set of three-dimensional meshes of a human brain, which were constructed from 12 volumes acquired using magnetic resonance imaging. Images were coloured using the computer graphics software Blender.

3.4.7 Enrichment analysis

To evaluate enrichment, we created an 'enrichment ratio'. This measure describes the ratio between observed and expected proteins in the metastable subproteome, for an assigned property under study

$$Enrichment\ ratio = \frac{MS_x/MS}{P_x/P} = \frac{observed}{expected}$$

where MS_x is the number of metastable subproteome members assigned to subcellular localisation or biological process x, MS is the total number of proteins in the metastable subproteome, P_x is the number of human proteome members assigned to subcellular localisation or biological process x, and P is total the number of proteins in the human proteome as a whole.

We evaluated the significance of our results using a one tailed Fisher exact test and Bonferroni multiple hypothesis correction. The background dataset was downloaded from uniprot.org, with all reviewed human proteins included (http://www.uniprot.org/downloads).
This dataset included subcellular localisation and biological process gene ontology assignments for each protein. Protein IDs were converted between UniProt and Entrez ID using the UniProt ID mapping service.

The gene ontology (GO) project provides a controlled vocabulary describing the biological processes a protein is involved in. Data available from the UniProt website lists selected terms derived from the GO project.

3.4.8 Defining a supersaturated subset

The supersaturation scores used in this study were taken from the latest publication by Ciraym et al. [41]. Supersaturation scores allow the evaluation of a protein's expression beyond its limit of solubility. Accordingly, these scores have two components - a score quantifying a protein's solubility, and the expression of the gene corresponding to the protein, in human tissue. The solubility score was calculated computationally, using the protein's polypeptide sequence as input. A high score indicated high insolubility. The supersaturated subset of proteins used in this study, are collated by extracting those genes within the top 5% of supersaturation scores [41].

3.4.9 Defining downregulated gene sets

Proteins downregulated during age and AD were calculated by Ciryam et al [41]. These subproteomes represent all found to be genes significantly downregulated. Corrections are made to exclude the influence of a co-correlation between transcriptional changes which occur during ageing and disease [41].

4 Looking beyond - further subproteomes in AD and other protein misfolding diseases

In Chapters 2 and 3, we found evidence of a transcriptional environment in tissues most vulnerable to AD conducive to the aggregation of Aβ and tau. We next sought to investigate the robustness of our hypothesis - that neuronal vulnerability to neurodegenerative disease is associated with an innate proteomic signature which predisposes disease-specific aberrant aggregation. In Chapter 2, we initially tested this hypothesis by looking for an AD transcriptional vulnerability signature associated to the protein homeostasis networks known to influence the aberrant aggregation of Aβ and tau. In Chapter 3, we tested this hypothesis further by analysis other subproteomes which we predict would influence the onset of AD-specific protein homeostasis loss. In this Chapter, instead of analysing new types of subproteome, we compare expression at different biological scales - cellular and subcellular. In addition, we expand our work to evaluate the presence of innate vulnerability to disease-specific aggregation in other neurodegenerative diseases. The combination of approaches used to investigate the vulnerability hypothesis throughout this thesis are summarised in Figure 4.1.

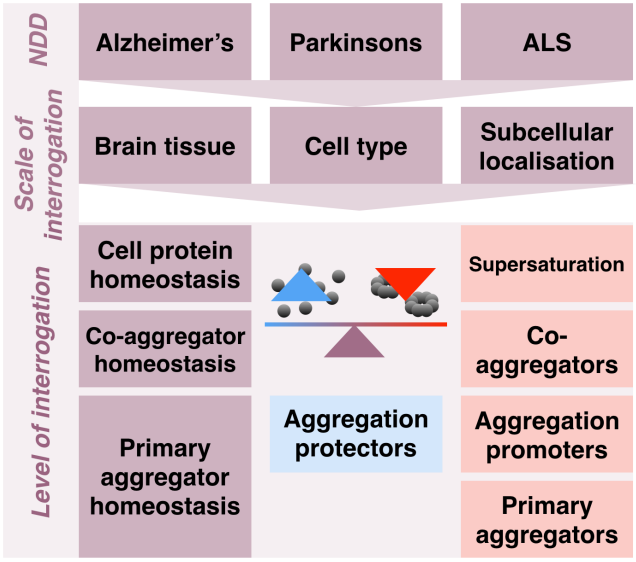

Figure 4.1: Levels of interrogation for investigation of vulnerability to disease-specific protein homeostasis loss. A number of neurodegenerative diseases associated with aberrant aggregation are studied in this Chapter. One can investigate a molecular signature associated to aberrant aggregation by comparing vulnerable and resistant tissue types, cell types, or subcellular localisations. The disease associated molecular signature can be studied from the perspective of cellular protein homeostasis, and of disease co-aggregators. In addition, one can interrogate the homeostatic regulation of the proteins most enriched in the deposits, which form the pathological hallmark of the neurodegenerative disease under study.

4.1 Introduction

4.1.1 Neurons are the most vulnerable cell type to neurodegenerative diseases

Although neurons are typically the primary cell of focus in neurodegenerative disease research, there are a diverse array of other cell types in neural tissue, making up significant proportion of the total cellular population. For example, the human brain has been found to contain almost 50% glial cells [220]. Neurons receive the lion's share of scientific attention for two reasons; their loss has the most direct impact on the manifestation of clinical symptoms, and they are the most vulnerable cell type to atrophy during neurodegenerative diseases progression. However, other cell types are affected by and implicated in neurodegenerative diseases pathogenesis - these cells

and their respective roles in neurodegenerative diseases will be introduced in the subsequent paragraphs.

Astrocytes in the brain form syncytium, into which neuronal networks embed themselves [221]. These glial cells perform a number of critical support functions; maintaining chemical homeostasis [222, 223], enabling neurogenesis, determining the microscopic architecture of grey matter [224-226] and protecting the brain against pathogenic insult. During AD, astrocytes respond via activation, triggering both morphological and molecular cell changes, and escalating the immunoinflammatory cascade characteristic to neurodegenerative diseases [227]. Astrocytes achieve this through the release of cytokines, pro-inflammatory factors, and neurotoxic reactive oxygen species [228]. In PD, activated astrocytes have been found in diseased tissue [229, 230]. The role of neuroinflammation in PD pathology remains under investigation, however this response is widely considered to be a downstream response to to the death of dopaminergic neurons [231]. It is likely that this inflammatory response results in further cell death - astrocytes are implicated in motor neurone death in ALS, and indeed, astrocytes derived from both familial and sporadic postmortem tissue are toxic to motor neurons [232].

Microglia, another glial cell type, are the primary immune effector cells in the brain. As with astrocytes, these cells have a crucial role in the maintenance of brain homeostasis and protection against infection [233]. In AD, a direct association has been found between plaque formation and microglial neuroinflammatory response. In AD brains, phenotypically activated microglia have been observed intimately associated with amyloid plaques, with extensions reaching into the deposit core [234, 235]. Microglia have been demonstrated to recognise and mount an immunological response to the Aβ peptide, and migrate towards aggregated Aβ species [233]. Unfortunately, despite their reactivity to amyloid aggregated species, microglia appear unable to clear them [233], consequently potentially resulting in more harm than help. In parallel to their role in AD, activated microglia have been found in close spatial association to areas of atrophy in PD [236-238]. Indeed it has been suggested that a chronic inflammation effected by microglia is a fundamental factor contributing to the

loss of highly PD-vulnerable dopamine-producing neurons [239, 240]. Neuroinflammation is also thought exacerbate motor neurone damage in ALS [241, 242]. Inflammatory cytokines released by microglia may result in glutamate excitotoxicity, resulting in neuronal excitotoxic death [243-245].

In the following of this chapter, we also investigate brain-derived endothelial cells. These cells are critical to the architecture of the blood brain barrier (BBB), forming an endothelial lining of neural vasculature. Endothelial cells have both metabolic and mechanical roles, working in concert with neuronal, glial, and smooth muscle cells to ensure optimal neural tissue function and homeostasis [246, 247]. Endothelial cells are primarily responsible for regulation of cerebral blood flow, and and chemical exchange across the blood-brain barrier. The ability of endothelial cells to fulfil their role is thought to be limited by the progression of AD pathogenesis: markers of endothelial dysfunction have been found to be elevated in the plasma of those with late-onset AD [248]. Endothelial cells also appear to participate in the inflammatory response broadly associated to neurodegenerative disease, and in AD brains express elevated levels of inflammatory cell adhesion molecules such as monocyte chemoattractant protein 1 (MCP-1) [249]. The overall impact of neurodegenerative diseases such as AD remains unclear, as despite evidence that dysfunction endothelial cells may also be proliferating in affected tissues: genome-wide expression profiling of AD tissue has identified the up-regulation of genes which promote angiogenesis [250].

Many of the impacts of neurodegenerative diseases on brain cells are likely to contribute to the atrophy of neurons observed in affected tissues - in particular immune activation and the loss of support functions carried out by cells such as astrocytes. Neurons are the primary site for the manifestation of many of the characteristic hallmarks of neurodegenerative diseases - for example tangle deposition in AD. Tangles develop intra-neuronally, progressing from early-stage inclusions, to mature neurofibrillary tangles (NFT), and finally observed as ghost tangles (the proteinaceous remnant of a dead neuron) [251]. Accordingly, neuron loss correlates with NFT deposition [76], and in turn the clinical manifestation of AD is considered to reflect the extent of neuron loss in the brain [252]. Lewy bodies, the histopathological

hallmark of PD, accumulate primarily in vulnerable subcortical neurons [253]. α-synuclein is the primary protein found in Lewy bodies. Aggregated αsynuclein species are thought to contribute to the pathogenic progression of PD [253]. Consequently, PD is characterised pathological by cell death in affected tissues - up to 70% of dopaminergic neurons of the SNPC have died by the end of life [254]. In parallel with AD and PD, ALS is characterised by the deposition of an aberrantly-aggregated proteins - primarily TDP43. Pathological TDP-43 is found mislocalised outside the nucleus, and typically aggregates in a neuron's cytoplasm or neurites [255].

In summary, it appears that neurons in particular have a vulnerability to the protein homeostasis failure characteristic of neurodegenerative diseases including AD, PD, and ALS. This observation leads one to consider the possibility that a predisposition to disease-specific protein aggregation is innate to neural cells, and that other cell types have proteomic environments with a greater robustness against this protein homeostasis loss.

4.1.2 Excitatory neurons are more vulnerable to AD than inhibitory neurons

Research published in early 2017 has provided insight into particular neuronal subclasses that appear to have elevated AD vulnerability. Investigators found that in the entorhinal cortex of transgenic mice over-expressing tau, excitatory but not inhibitory neurons were vulnerable to tau pathology [256]. The factors underlying this differential in vulnerability are currently unknown.

4.1.3 Synapses are the most vulnerable subcellular compartments to AD

In addition to a comparison of cell types, one can also consider vulnerability to neurodegenerative diseases at a subcellular level. In AD, there exists evidence that synapse loss occurs at an earlier stage in disease progression than whole-neuron loss [257]. Evidence of synaptic impairment has been found even at pre-symptomatic stages of the disease [258], suggested to be due to the presence of toxic oligomeric species of Aβ [72, 114, 258-262]. More extensive synaptic loss is found at later stages of

disease, with over 25% of synapses rendered dysfunctional in affected tissues at the onset of disease symptoms [263]. Furthermore, post mortem analysis of AD affected tissue indicates that synaptic loss correlates most robustly with aggregate deposition [263]. In AD patients, soluble Aβ levels show a significant correlation with the extent of synaptic loss [264].

Whilst there are a number of possible causes of elevated synaptic vulnerability in neurodegenerative diseases, for instance exocytotic impairment or perturbed intracellular trafficking [265], we take the view here that synapses may simply be the most fertile environment for the aggregation of disease-specific proteins. The subsequent presence of toxic aggregation intermediates would explain their particular vulnerability, even at the earliest stages of disease.

4.1.4 Commonalities and crucial differences between neurodegenerative diseases

The presence of aberrant aggregation hallmarking a number of neurodegenerative diseases strongly suggests a shared general pathogenic mechanism, involving a loss in protein homeostasis. In many neurodegenerative diseases, such as AD and PD, particular cell types and subcellular localisations show elevated vulnerability. However, at a tissue level, the site of disease inception and the spatio-temporal spread is characteristic to each condition, in addition to the primary proteins which form deposits. This indicates that there are proteomic factors which determine vulnerability to protein homeostasis loss in general, and different factors which influence the aberrant aggregation of particular protein subsets. If both of these factors coincide in a proteomic environment, it is likely that in that locality the disease with manifest.

4.2 Results

4.2.1 Elevated expression of supersaturated proteins characterises most vulnerable tissues in neurodegenerative diseases

Our first line of investigation aimed to establish the molecular factors responsible for creating an environment primed for the protein homeostasis dysfunction seen in protein misfolding diseases. We reasoned that these factors would place a generic

burden on the protein homeostasis system of neurons, and would exist at significantly elevated levels in the tissues which first succumb to neurodegenerative diseases-.

We therefore decided to investigate the role played by highly supersaturated proteins, in defining these tissues most vulnerable to aberrant aggregation. We have found that relative to the proteome, supersaturated proteins are significantly over-expressed in the tissues most vulnerable to AD, ALS, FTD, and PD (Figure 4.2). The agreement between regional vulnerability to neurodegenerative disease associated with aberrant aggregation, and tissue expression level of the most supersaturated proteins, is shown in Figure 4.2b.

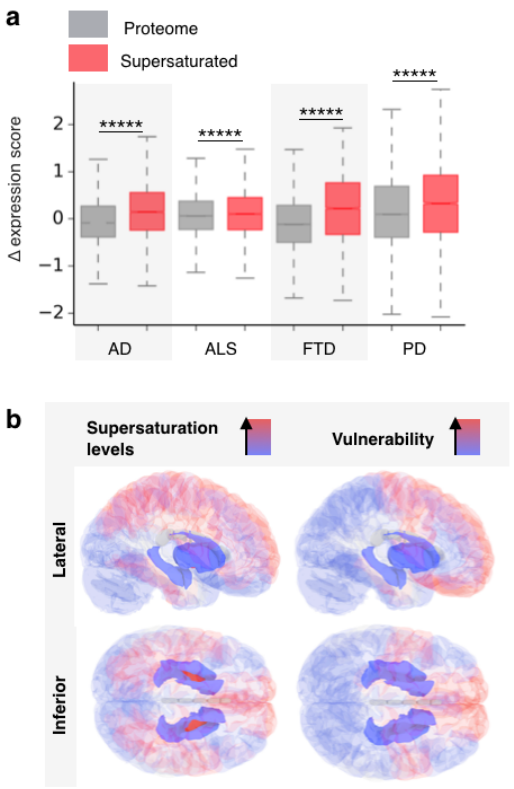

Figure 4.2: Supersaturated proteins are more highly expressed in vulnerable tissues in healthy brains relative to resistant tissues. a) Box plot of scores for supersaturated proteins. For each disease, Δ scores represent expression in vulnerable tissues, relative to expression in resistant tissues. ***** $p < 0.00001$; the statistical significance of the difference between the distributions of the protein set under study and that of the proteome was calculated with the Mann-Whitney U test with Benjamini-Hochberg multiple hypothesis testing correction [207]. **b)** Red regions on the brain image on the left indicate high relative expression of superstaurated proteins, blue indicates low levels. Red regions for the brain image on the right indicate the tissues most vulnerable to neurodegenerative disease onset

4.2.2 Neurons have the highest relative supersaturation burden relative to other brain cell types

The role of supersaturated proteins in defining vulnerability transcends tissues - the resilience of brain cell types to neurodegenerative disease can also be predicted in this way. To address this point, we used single-cell human mRNA sequencing data [266]. We thus found that neurons have significantly elevated levels of highly supersaturated proteins (Figure 4.3).

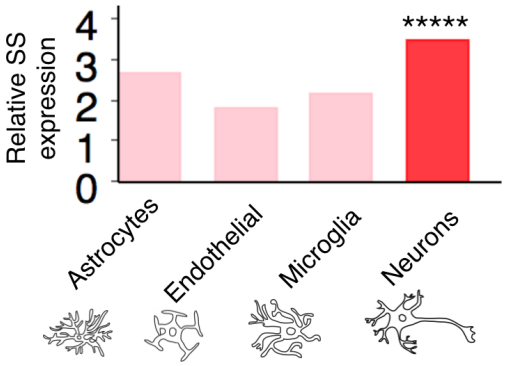

Figure 4.3: Relative expression of supersaturated proteins in different brain cell types. Expression analysis completed on human single cell RNAseq data. p-value calculated describes significance of expression in neurons, relative to that in other brain cell types plotted, calculated with the Mann-Whitney U test with Benjamini-Hochberg multiple hypothesis testing correction [207] *****$p < 0.0005$.

4.2.3 Neurons have an expression signature that predisposes them more to aggregation of Aβ and tau relative to other cell types

We next investigated whether Δ score analysis is capable of identifying neurons as the cell type most vulnerable to pathological aggregation in AD, from the perspective of AD-specific subproteomes. By calculating the relative expression of Aβ and tau for neurons, astrocytes, microglia, and endothelial cells, we found that their relative expression was elevated significantly in neurons (Figure 4.4). Simultaneously, we found that the relative expression of their aggregation protectors was the lowest and

that of their promoters was the highest. These results indicate that neurons also exhibit a cellular environment most conducive to Aβ and tau aggregation in comparison with the different brain cell types that we analyzed in this work.

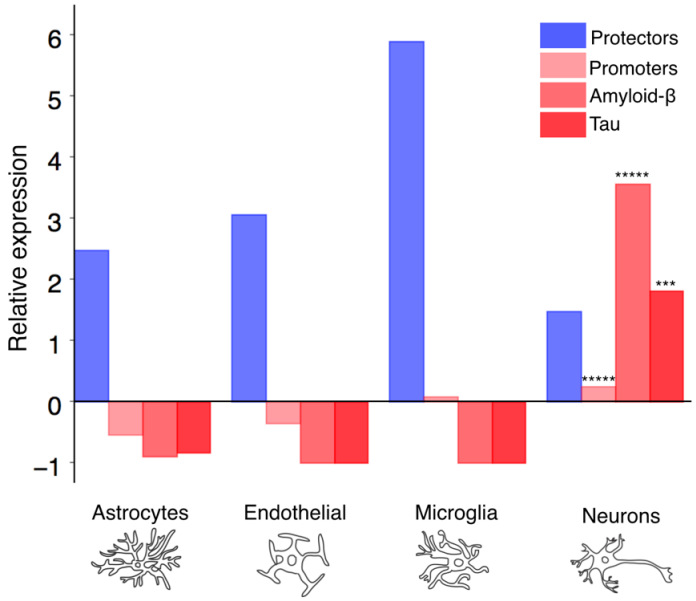

Figure 4.4: Expression of different components of Aβ and tau homeostasis in specific brain cell types. For different brain cell types, including neurons, astrocytes, microglia, and endothelial cells, we calculated the relative mRNA expression levels, as measured by the $E'_{g,c}$ score (see Methods), of genes corresponding to Aβ and tau, and the corresponding aggregation protectors and promoters. For each gene set in neurons, the significance of the difference with the expression distribution for all other brain cell types in combination was calculated using Mann-Whitney *U* test with Benjamini-Hochberg multiple hypothesis testing correction [207] ***$P <$ 0.001 and *****$P < 0.00001$.

Differential vulnerability can also be predicted by the metastable subproteome at the cellular level, again by utilising human single cell RNAseq data (Figure 4.5).

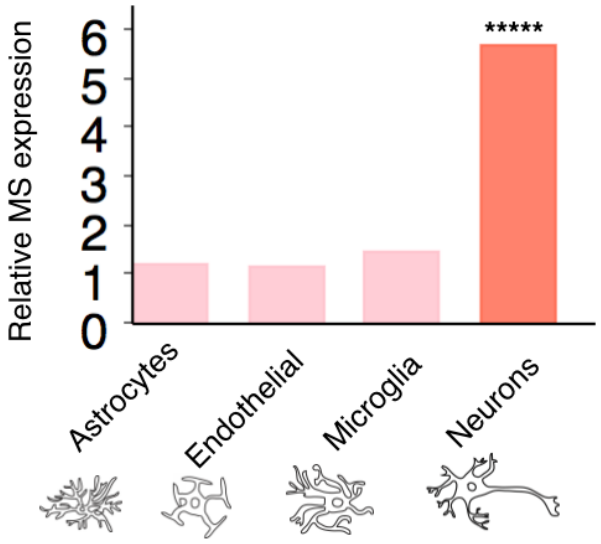

Figure 4.5: Relative expression of the metastable subproteome in a number of brain cell types. The significance level indicates relative expression of the MS for neurons is significantly higher when compared to relative MS expression in other cell types. For the metastable subproteome gene set in neurons, the significance of the difference with the expression distribution for all other brain cell types in combination was calculated using Mann-Whitney *U* test with Benjamini-Hochberg multiple hypothesis testing correction [207] ***$P < 0.001$ and *****$P < 0.00001$.

4.2.4 Excitatory neurons have an expression signature that predisposes the aggregation of Aβ and tau, relative to inhibitory neurons

Having found a molecular signature which aligns to elevated relative neural vulnerability to both AD and neurodegenerative diseases more generally, we next sought to investigate whether differential vulnerability between neurons mirrored previous results. To achieve this goal, we analysed differential expression from single cell mRNA data [267]. We categorised excitatory and inhibitory neurons in the data

set under study using a protocol established by the authors of the Allen Brain Atlas (for further details see Methods). We find that in both early and late stage AD tissues, levels of Aβ and tau aggregation promoters, the metastable subproteome, and plaque and tangle co-aggregators, are elevated in excitatory neurons relative to inhibitory neurons (Figure 4.6).

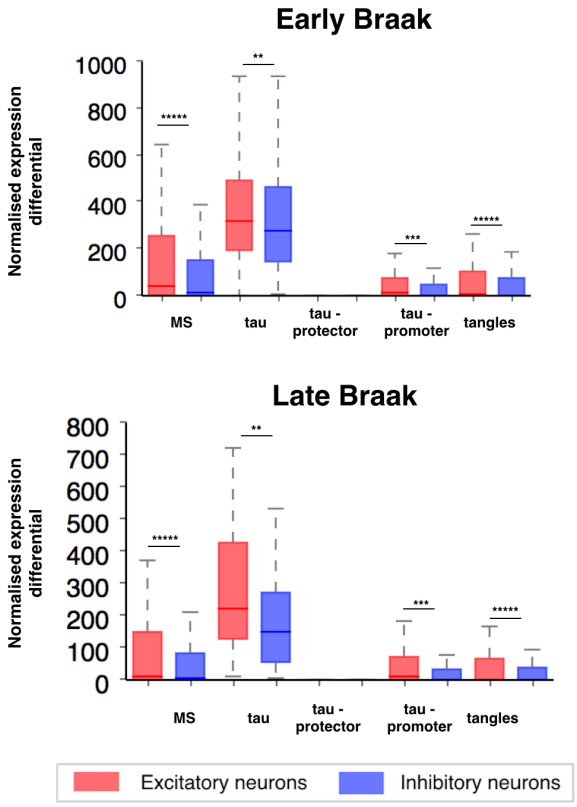

Figure 4.6: Differential expression of AD-specific aggregation modulators in excitatory and inhibitory neurons. Single cell expression data [267] were used to define neurons into excitatory and inhibitory subsets. Expression of gene sets specific to AD were assessed for cells derived from tissues affected in Braak stages III-IV (early Braak), and Braak stages V-VI (late Braak). The significance of the difference in normalised expression distributions from excitatory and inhibitory neurons was calculated using Mann-Whitney *U* test with Benjamini-Hochberg multiple hypothesis testing correction [207] ***$P < 0.001$ and *****$P < 0.00001$.

4.2.5 The synaptic proteome is enriched in both supersaturated proteins and in the metastable proteome

Next, we decided to investigate relative vulnerability at the subcellular level, using proteomics data describing protein levels in the synaptosome [268]. Focusing on the AD risk, we show how metastable proteins are prevalent in the synaptic environment (Figure 4.7a-c), and are enriched with respect to the most concentrated synaptic proteins (Figure 4.7a-b). We find that the metastable subproteome consistently has the most elevated enrichment for all metrics tested, relative to the other intersections possible between the subproteomes from which it is composed (Figure 4.7a-b). Synpatosomic enrichment of the highly supersaturated protein subset, the metastable subproteome, is significantly enriched in the synaptosomes relative to whole neurons (Figure 4.7e). This result suggests a relative elevated innate vulnerability to AD for the synapse as a subcellular localisation. Considering neurodegenerative diseases in a more general sense, we sought to investigate the enrichment of supersaturated proteins in the most and least prevalent sets of proteins in synaptosomes [268]. We found an enrichment of supersaturated proteins in synaptosomes (Figure 4.7c-d). These observations help rationalise why synapses are the first subcellular regions to degenerate during many neurodegenerative diseases.

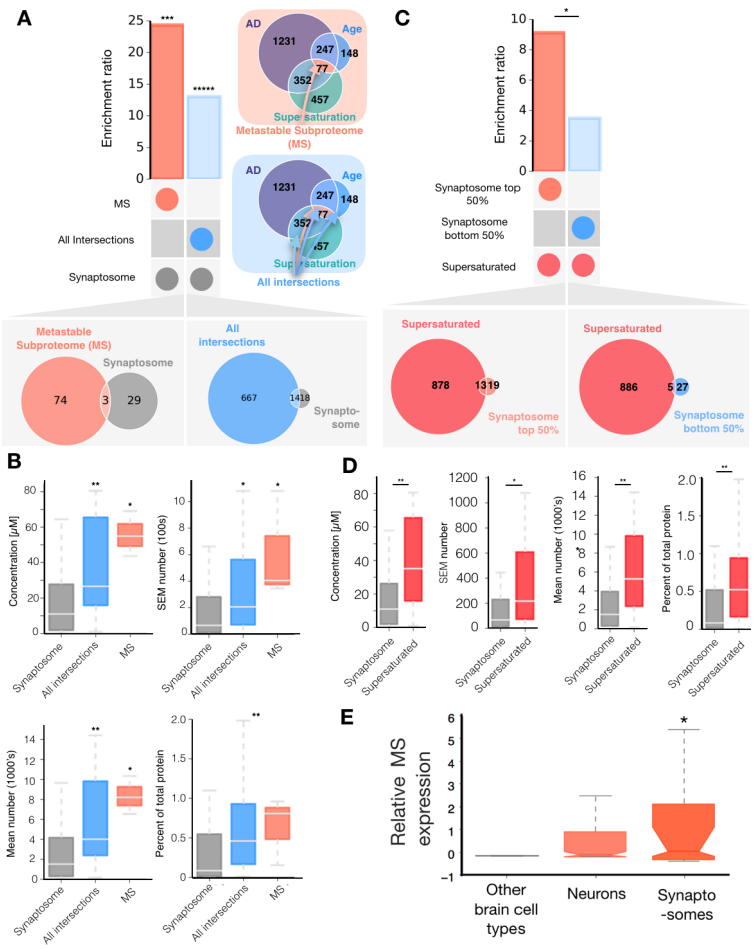

Figure 4.7: The synaptic environment is highly vulnerable to protein aggregation. a) Enrichment of the MS (left), and "intersectional" proteins (right), with respect to the most abundant synaptosomal proteins. Intersectional proteins are defined as those that are supersaturated, downregulated in AD, and downregulated in ageing, but not part of the metastable subproteome. b) Concentration, SEM number, protein percentage, and mean protein number for the metastable subproteome (red), intersectional proteins (blue), and all proteins measured in a synaptic buton (grey) (c) Enrichment of supersaturated proteins with respect to the most (left) and least (right) abundant synaptosomal proteins. (d) Concentration, SEM number, protein percentage, and mean protein number for supersaturated proteins (red), and all proteins measured in a synaptic buton (grey) are highlighted in red. e) Expression of the metastable subproteome, relative to expression of the proteome, in brain cell types, and synaptosomes. The significance of the result for neurons is evaluated relative to

other cell types, and for synaptosomes is evaluated relative to neurons. *p < 0.05, **p < 0.01, ***p < 0.01 and *****p<0.0001 Mann-Whitney U test [207].

4.2.6 Robustness testing of results

The data used in this study provided a number of opportunities to test the robustness of our hypothesis, using other neurodegenerative diseases to formulate predictions. In the first instance, we analysed the relative expression of proteins known to be vulnerable to misfolding and accumulation in PD in the tissues most vulnerable to the disease.

We found that genes corresponding to proteins that coaggregate within Lewy bodies [269] in PD (see Methods) [269] have elevated expression in the SNPC region] [270] that is highly vulnerable to this condition (Figure 4.8), a result that provides further support for the suggestion that the paradigm of tissue vulnerability that has been primarily investigated for AD her can applied to other neurodegenerative disorders.

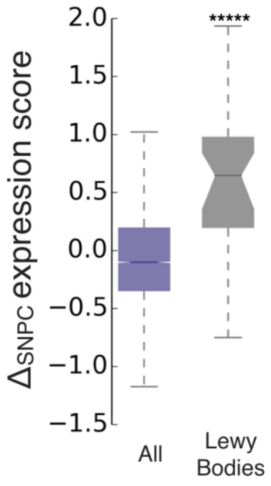

Figure 4.8 Δ_{SNPC} scores corresponding to aggregation-prone protein sets characteristic of PD. Distributions of the ΔSNPC scores corresponding to the whole proteome (All), and for proteins that co-aggregate in Lewy bodies. Δ scores are calculated for the SNPC region. Boxes represent the first and third quartiles of the distribution, whiskers the 1.5 inter-quartile range, and notches are standard errors on the median calculated with 104 bootstrap cycles. *****p<0.00001, calculated with a Mann-Whitney U test [207].

Because the key hypothesis in this work centralises on the concept of protein solubility, it is crucial for results to be meaningful that there is a strong

correspondence between mRNA levels analysed in this study and protein levels in the neurons within the respective samples.

Although a relatively weak correlation exists between mRNA and protein levels [208, 271], in this work, we considered average values across groups of genes, and hence, we expect stronger correlations to be present. To validate this type of approach, we verified that the patterns of gene expression analyzed here are consistent with the corresponding patterns of protein expression using two independent data sets (Figure 2.4) [209, 272], and in addition, we performed a control using proteome-level data, which were available for mouse tissues [273]. We found that the ranking of the tissue-specific risk using protein data is consistent with that using mRNA data (Figure 4.9).

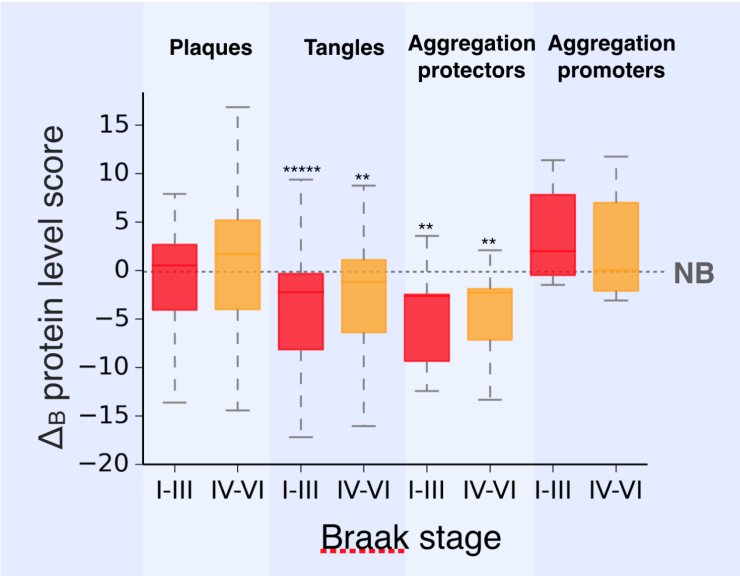

Figure 4.9: Proteome-based ΔB scores calculated using mice data for tissues affected at different Braak stages. Boxplots of relative levels of Aβ and tau aggregation modulators, calculated from mass spectrometry-based proteomics, for a subset of mouse tissues (36). Boxes represent the first and third quartiles of the distribution, whiskers the 1.5 inter-quartile range; notches are standard errors on the median calculated with 104 bootstrap cycles. For each boxplot, the significance of the difference with the Δ distribution for non-Braak (NB) tissues was calculated with a Mann-Whitney U test and Benjamini-Hochberg multiple hypothesis testing correction [207] **p < 0.01, *****p<0.00001.

4.2.7 Investigation of alternative hypotheses

Whilst this thesis provides evidence for the role played by disease-specific protein homeostasis in neurodegenerative disease pathogenesis, the possibility of other disease mechanisms cannot be discounted.

We therefore further investigated other possible causes of tissue vulnerability to AD, particularly the immune response (Figure 4.10) [274, 275]. Our analysis of $\Delta_{BI–III}$ scores of biochemical pathways listed in the Kyoto Encyclopedia of Genes and Genomes (KEGG) database [276] revealed that genes associated with inflammatory responses are expressed at elevated relative levels in healthy brains, whereas genes

involved in autoimmune responses are expressed at lower relative levels in AD-vulnerable tissues (Figure 4.10). Because no other immune pathway shows significant variation in expression (Figure 4.10), these results support previous suggestions of a role for inflammation in the pathogenesis of AD [274, 275]. Relative elevated expression of KEGG pathway proteins involved in the circadian cycle and heparin sulphate synthesis are also seen in early AD tissues (Figure 4.10). It is interesting to note that both of these processes have been implicated in AD pathogenesis (see section 1.4). Thus, the vulnerability of specific tissues in AD may result from the sum of a number of factors, including the expression levels of disease-specific, aggregation-prone proteins and their corresponding protein homeostasis complements, as well as the immune system.

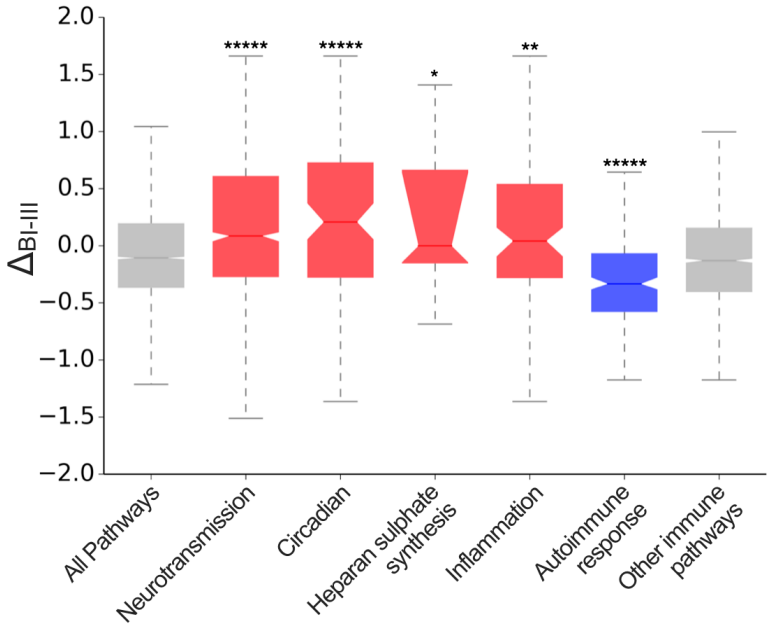

Figure 4.10: Distributions of Δ_{BI-III} scores for selected KEGG pathways.
Box plots of the Δ_{BI-III} score distributions for each pathway category (x axis) in the context of the whole proteome. "All pathways" is the distribution of the Δ_{BI-III} scores of all proteins in the human proteome with at least one KEGG pathway assigned. Boxes represent the first and third quartiles of the distribution, whiskers represent the 1.5 interquartile range, and notches are the standard errors on the median calculated with 10^4 bootstrap cycles. Significance values (**P < 0.01 and *****P < 0.00001) report the statistical significance of the difference with the first box plot (All pathways) calculated using Mann-Whitney U test with Benjamini-Hochberg multiple hypothesis testing correction [207]

4.3 Conclusions

In this chapter, we have investigated neurodegenerative disease vulnerability from perspectives beyond that of tissue-specific expression signatures relevant to AD. This approach had a two-fold objective. Firstly, to find evidence of relative vulnerability at smaller biological scales, and from the perspective of other diseases associated with protein homeostasis loss and other mechanisms suggested to be involved in disease progression. Secondly, to confirm the robustness of the robustness of results obtained in cChapters 2 and3.

Considering neurodegenerative diseases beyond AD, we have found that significantly elevated levels of supersaturated proteins were found in tissues most vulnerable to PD, ALS, and FTD. This result suggests a vulnerability to protein homeostasis loss underscores the inception of a number of neurodegenerative disease associated with aberrant protein deposition.

Moving from tissue-level analysis to cellular-level, we then compared expression signatures in neurons to those in other cell types, including astrocytes, microglia, and endothelial cells. We found significantly elevated levels of supersaturated proteins in neurons relative to other cell types. This higher burden on neuronal protein homeostasis, due to high concentrations of aggregation-prone proteins, could explain the notable vulnerability of neurones in many neurodegenerative diseases. Furthermore, when investigating AD-specific subsets, we find evidence of elevated relative vulnerability of neurons to AD-specific aggregation.

Neurons can be classified into a number of specialised groups. One of the broadest of these classifications is excitatory v. inhibitory. In accordance with experimental results revealing that excitatory neurons are most vulnerable to taupoathy, we find here that relative expression indicates a predisposition for AD-specific aggregation in these neurons relative to inhibitory neurons.

Moving to the subcellular level, we find evidence of enrichment of both supersaturated and metastable subproteome proteins in synaptosomes, relative to neurons as a whole.

This result aligns well with those discussed in chapter 3.2.3, where it was found that the metastable subproteome is enriched for proteins of synaptic function and subcellular localisation. Based on the results thus far in this chapter and the two previous, we suggest that an innate predisposition to aberrant aggregation in neurodegenerative disease exists in normal brains at the proteomic level.

Robustness checks confirmed that an expression level signature for disease-specific vulnerability was also present for PD, and that proteomic results reflected those found for the more widely available expression-level data.

Although this thesis has focused on the role of protein homeostasis in tissue vulnerability to neurodegenerative diseases, it is likely that other pathogenic mechanisms also contribute to disease inception and progression. From results in this chapter, and in literature, these mechanisms are likely to include inflammation and the presence of heparin sulphates.

4.4 Methods

4.4.1 Data sources

A full list of data sources for the results presented in this chapter is provided in table 4.1.

Figure	Data source	Reference
4.1	Allen Brain Atlas	M. J. Hawrylycz et al., An anatomically comprehensive atlas of the adult human brain transcriptome. Nature 489, 391-399 (2012).
4.3	RNAseq data from purified cell types of the human brain	S. Darmanis et al., A survey of human brain transcriptome diversity at the single cell level. Proceedings of the National Academy of Sciences 112, 7285 (2015).
4.4	RNAseq data from purified cell types of the human brain	S. Darmanis et al., A survey of human brain transcriptome diversity at the single cell level. Proceedings of the National Academy of Sciences 112, 7285 (2015).
4.5	RNAseq data from purified cell types of the human brain	S. Darmanis et al., A survey of human brain transcriptome diversity at the single cell level. Proceedings of the National Academy of Sciences 112, 7285 (2015).
4.6	RNA-seq dataset from adult human brain	Science. 2016 June 24; 352(6293): 1586–1590. doi:10.1126/science.aaf1204
4.7	Composition of isolated synaptic boutons	Wilhelm, Benjamin G., Sunit Mandad, Sven Truckenbrodt, Katharina Kröhnert, Christina Schäfer, Burkhard Rammner, Seong Joo Koo et al. "Composition of isolated synaptic boutons reveals the amounts of vesicle trafficking proteins." Science 344, no. 6187 (2014): 1023-1028.
4.8	Allen Brain Atlas	M. J. Hawrylycz et al., An anatomically comprehensive atlas of the adult human brain transcriptome. Nature 489, 391-399 (2012).
4.9	Proteomic analysis of the mouse brain	K. Sharma et al., Cell type-and brain region-resolved mouse brain proteome. Nature neuroscience 18, 1819 (2015).
4.10	Allen Brain Atlas	M. J. Hawrylycz et al., An anatomically comprehensive atlas of the adult human brain transcriptome. Nature 489, 391-399 (2012).

Table 4.1: Data sources for figures in chapter 4

4.4.2 Allen Brain Atlas data analysis

Methods regarding Braak staging are outlined in section (2.4.2), the mapping of Braak stages to Allen Brain Atlas regions are described in section (2.4.3). The equations detailing the process of Δ score calculation are given in section (2.4.4). A description of statistical techniques, which are consistent between chapters, are given in section 2.4.5. A description of the shading method used for cortical and subcortical brain structures on a three-dimensional representation, in figure 4.2, is given in section 2.4.6.

4.4.3 Key subproteome construction

In this chapter, both the metastable subproteome, and a subproteome of highly supersaturated proteins are used in analysis. Details of how these subproteomes are defined are given in sections 3.4.8 and 3.4.9.

4.4.4 Categorisation of neurons as excitatory or inhibitory

The cell categorisation method used in this chapter was taken from the one used in the Allen Brain Atlas to sort excitatory and inhibitory neurons from their expression signatures [277]. Cells were classified as either excitatory or inhibitory neurons or non-neuronal based on gene expression. Cells were classified as excitatory if the maximum expression of excitatory genes (Slc17a6, Slc17a7) was greater than the maximum expression of inhibitory (Gad1, Gad2, Slc32a1) or non-neuronal (Olig1, Gja1, Xdh, Ctss, Myl9) genes. Cells were classified as inhibitory if the maximum expression of inhibitory (Gad1, Gad2, Slc32a1) genes was greater than the maximum expression of excitatory (Slc17a6, Slc17a7) or non-neuronal (Olig1, Gja1, Xdh, Ctss, Myl9) genes. All remaining cells were classified as non-neuronal.

Prior to subsequent analysis, expression data were normalised. For each data set (excitatory or inhibitory), for each expression reading, the relative expression was calculated for each gene within each cell type (excitatory or inhibitory) as

$$E'_{g,c} = \frac{E_{g,c} - \mu_{g,c}}{\sigma_{g,c}}$$

where $E_{g,c}$ is the expression for each gene g in a given cell type c, $\mu_{g,c}$ is the mean expression of that gene in a given cell type c, and $\sigma g,c$ is the SD of expression of that gene in a given cell type c.

Tissue samples were limited in the dataset used for this figure, and therefore we limited tissue stage categorisations to early Braak (BA21, BA22, BA10, BA41), and late Braak (BA17).

4.4.5 Defining a vulnerability landscape in the brain

Regions which are described as the most common location for initial onset of four prevalent neurodegenerative diseases associated with aberrant aggregation, were collected and together define the most vulnerable tissues. Table 4.2 lists these regions for each disease (AD, ALS, FTD, and PD).

Disease	Allen Brain Atlas Code	Brain Regions
AD	PHG-l, PHG-cos, CA1, DTA, SptN, SI, AOrG, fro, FP-s, FP-l, Gre, trIFG, opIFG,orIFG, IRoG, LOrG, MOrG, MFG-s, MFG-i, PCLa-s, PCLa-i, PTG, POrG, SFG-m, SFG-l, SRoG, LiG-pest, LiG-str, Cun-pest, Cun-str, IOG, OP, SOG-s, SOG-i, OTG-s, OTG-i, FuG-l, FUG-its, FuG-cos, HG, ITG-mts, ITG-l, ITG-its, MTG-s, MTG-i, PLP, PLT, STG-i, STG-l, TP-s, TP-i, TP-m, TG, S, CA2, CA3, CA4, DG, CgGp-s, CgGp-i, CgGf-s, CgGf-l, CgGr-i, CgGr-s, SCG, ATZ, BLA, BMA, CeA, COMA, LA	Parahippocampal gyrus, CA1, Anterior group of nuclei, Basal forebrain - includes septal nuclei, and nuclei within the substantia innominata, Frontal, temporal, and occipital association areas, Subiculum, CA2, CA3, CA4, Dentate gyrus, Cingulate gyrus, Amygdala
PD	SNC	Substantia nigra pars compacta
FTD	FP-s, Fpi, FPm, MOrG, SFG-m, SFG-l, MFG-s, MFG-i, PrG-prc, PrG-sl, PrG-il, PrG-cs, SFG-m, SFG-l, MFG-s, MFG-i , DG, CA1, CA2, CA3, CA4, IG, S, ATZ, AAA, BLA, BMA, CeA, COMA, IA, LA, NLOT, PLA, CgGf-s, CgGf-i, SMG-s, SMG-i	Frontal pole, Medial orbital frontal gyrus, Superior frontal gyrus, Precentral gyrus, Dentate gyrus, hippocampus, Amygdala, Frontal cingulate gyrus, Supramarginal gyrus
ALS	PrG-prc, PrG-sl, PrG-il, PrG-cs, SFG-m, SFG-l, MFG-s, MFG-i , ACu, Amb, Arc, AP, CGS, CGMe, 8Co, CT, Cu, EL, EF, ECu, Gr, 12, IO, DIO, MIO, PrIO, BIO, IS, In, IFH, IPo, LPCu, MPCu, CMRt, GiRt, LMRt, PCRt, DPGi, LPGi, Gi, GiV, Gia, LMRt, EO, LRt, LRtPC, LRtS5, PCRt, Ro, Pa5, Pe5, PnB, Pr, Ramb, Sol, SolC, SolD, SolDL, SolG, SolI, SolIM, SolM, SolPaC, SolV, SolV, Sp5, Sp5C, Sp5Cg, SP5Cm, SP5Cz, Sp5I, Sp5O, Sge, SSp, 8Ve, ac, cc, hbc, hc, pc, smc, ar, arf, alv, amtg, agb, al, ap, bx, cgb, comb, cor, cbu, cbu-sc, cbu-ic, ec, emlgp, emlth, exc, fx, afx, bfx, fi, pfx, ilf, ithp, ic, imlgp, imlth, lls, lef, mp, mtg, mtt, mfb, mls, or-lp, milf, off , orpt, perf, ppf, ponb, rthp, sst, saf, szt, sm, st, scf, sthf, sprs, slf, thf, uf, 12, Amb, 5, 7	Agranular motor neocortex - Broadmann areas 4,6 Medulla oblongata at the level of N. XII - bulbar somatomotor neurons of N. XII

Table 4.2 Coding of regional vulnerability to four key neurodegenerative diseases

4.4.6 Relative expression for cell types

Data were obtained from a previous mRNA sequencing study of human brain tissue [266]. To evaluate the vulnerability of different brain cell types (figures 4.3-4.5), the relative expression was calculated for each gene within each cell type as

$$E'_{g,c} = \frac{E_{g,c} - \mu_{g,c}}{\sigma_{g,c}}$$

where $E_{g,c}$ is the expression for each gene g in a given cell type c, $\mu_{g,c}$ is the mean expression of that gene in a given cell type c, and $\sigma g,c$ is the SD of expression of that gene in a given cell type c.

5 Conclusions

In this thesis we have investigated the molecular origins of neurodegenerative diseases by an approach based on the analysis of the vulnerability of different cellular environments to protein aggregation. We have thus identified specific gene signatures in biological environments most vulnerable to neurodegenerative diseases which predispose to protein homeostasis loss.

For AD, the main focus of this work, we have shown that subproteomes related to AD-specific aggregation are found at elevated relative expression in vulnerable tissues, cell types, neuron types, and sub-neuronal localisations in normal brains. Quite generally, regions of the brain where a neurodegenerative disease first incepts appear to suffer from both elevated levels of a set of proteins known to be at high risk of aggregation (supersaturated proteins), and an expression signature which predisposes the protein homeostasis loss specific to the disease in question. In the case of AD, tissues that succumb early in disease progression have been found to have elevated expression of the primary aggregators found in plaques and tangles (Aβ and tau respectively), in addition to elevated relative expression of co-aggregators. Furthermore, levels of the proteins which regulate the aggregation of Aβ and tau are sufficient to predict the progression of AD through the brain, as is the metastable subproteome, an AD-specific supersaturated subset of proteins.

This analytical perspective has currently a number of limitations. Primarily, the analysis that we have reported is restricted by the available data, which are mostly at a transcriptional level. A relative lack of proteomic data limits at the moment our ability to directly interrogate the proteomic environment, which is clearly the one most directly implicated in protein aggregation. However, a homeostasis-associated vulnerability signature has been found in this work also at the proteomic level, at least for the limited data currently available at this level [273], indicating that our expression-level results may reflect the proteomic signature of the environment under study.

Data with greater breath and granularity, particularly at the tissue and the single-cell levels, would open up a number of new scientific opportunities. It would allow the evaluation of the relative weight which a number of factors, both within this study and external to it, have on tissue vulnerability. It would allow one to predict whether predisposition to Aβ or tau aggregation is more spatially related cell atrophy. It would allow one to untangle the relative roles of tissue vulnerability and connectivity in disease spread. Furthermore, these data would create an opportunity for models to be developed, where minimum solutions reveal the specific genes that correlate best with disease progression.

We should mention that the regulation of protein homeostasis is likely to involve many other cellular factors in addition to those considered here, including lipids, sugars, metabolites and nucleic acids. Although the impact of these factors on protein homeostasis is currently too limited to enable their systematic studies, we anticipate that advances in high-throughput experimental techniques will in the future enable the compilation of detailed lists of these cellular components, which in turn will make it possible to identify their specific roles in the avoidance of the toxic consequences of protein aggregation.

This unbiased approach could reveal new targets for the understanding and treatment of neurodegenerative diseases. With burgeoning scientific capabilities, we are just beginning to investigate the microenvironments within neuronal synapses, which may eventually lead to a mechanistic explanation of why these subcellular species are so vulnerable and why tau becomes so aggregation prone.

In summary, this work adds to a body of evidence that characterises protein homeostasis as an important component in the pathogenesis of neurodegenerative diseases. In a field where the factors influencing disease inception and progression still remain under debate, this work should help to add clarity, and promote new, helpful investigative avenues.

Bibliography

1. Onuchic, J.N., Z. Luthey-Schulten, and P.G. Wolynes, Theory of protein folding: the energy landscape perspective. Annual review of physical chemistry, 1997. **48**(1): p. 545-600.
2. Dobson, C.M., Protein folding and misfolding. Nature, 2003. **426**(6968): p. 884-890.
3. Bukau, B. and A.L. Horwich, The Hsp70 and Hsp60 chaperone machines. Cell, 1998. **92**(3): p. 351-366.
4. Hartl, F.U. and M. Hayer-Hartl, Molecular chaperones in the cytosol: from nascent chain to folded protein. Science, 2002. **295**(5561): p. 1852-1858.
5. Stigter, D. and K.A. Dill, Charge effects on folded and unfolded proteins. Biochemistry, 1990. **29**(5): p. 1262-1271.
6. Reynaud, E., Protein misfolding and degenerative diseases. Nature Education, 2010. **3**(9): p. 28.
7. Kelly, J.W., Alternative conformations of amyloidogenic proteins govern their behavior. Current opinion in structural biology, 1996. **6**(1): p. 11-17.
8. Soto, C., Protein misfolding and disease; protein refolding and therapy. FEBS letters, 2001. **498**(2-3): p. 204-207.
9. Chiti, F. and C.M. Dobson, Protein misfolding, functional amyloid, and human disease. Annu. Rev. Biochem., 2006. **75**: p. 333-366.
10. Knowles, T.P., et al., An analytical solution to the kinetics of breakable filament assembly. Science, 2009. **326**(5959): p. 1533-1537.
11. Iljina, M., et al., Kinetic model of the aggregation of alpha-synuclein provides insights into prion-like spreading. Proceedings of the National Academy of Sciences, 2016. **113**(9): p. E1206-E1215.
12. Lomakin, A., et al., On the nucleation and growth of amyloid beta-protein fibrils: detection of nuclei and quantitation of rate constants. Proceedings of the National Academy of Sciences, 1996. **93**(3): p. 1125-1129.
13. Esler, W.P., et al., Alzheimer's disease amyloid propagation by a template-dependent dock-lock mechanism. Biochemistry, 2000. **39**(21): p. 6288-6295.
14. Cohen, S.I., et al., Proliferation of amyloid-β42 aggregates occurs through a secondary nucleation mechanism. Proceedings of the National Academy of Sciences, 2013. **110**(24): p. 9758-9763.
15. Gazit, E., The “correctly folded” state of proteins: is it a metastable state? Angewandte Chemie International Edition, 2002. **41**(2): p. 257-259.
16. Baldwin, A.J., et al., Metastability of native proteins and the phenomenon of amyloid formation. Journal of the American Chemical Society, 2011. **133**(36): p. 14160-14163.
17. Knowles, T.P., M. Vendruscolo, and C.M. Dobson, The amyloid state and its association with protein misfolding diseases. Nature Reviews Molecular Cell Biology, 2014. **15**(6): p. 384-396.
18. Harper, J.D., C.M. Lieber, and P.T. Lansbury, Atomic force microscopic imaging of seeded fibril formation and fibril branching by the Alzheimer's disease amyloid-β protein. Chemistry & biology, 1997. **4**(12): p. 951-959.
19. Lashuel, H.A., et al., Mixtures of wild-type and a pathogenic (E22G) form of Aβ40 in vitro accumulate protofibrils, including amyloid pores. Journal of molecular biology, 2003. **332**(4): p. 795-808.
20. Winner, B., et al., In vivo demonstration that α-synuclein oligomers are toxic. Proceedings of the National Academy of Sciences, 2011. **108**(10): p. 4194-4199.
21. Sengupta, U., A.N. Nilson, and R. Kayed, The role of amyloid-β oligomers in toxicity, propagation, and immunotherapy. EBioMedicine, 2016. **6**: p. 42-49.
22. Selkoe, D.J., Soluble oligomers of the amyloid β-protein impair synaptic plasticity and behavior. Behavioural brain research, 2008. **192**(1): p. 106-113.
23. Gharibyan, A.L., et al., Lysozyme amyloid oligomers and fibrils induce cellular death via different apoptotic/necrotic pathways. Journal of molecular biology, 2007. **365**(5): p. 1337-1349.
24. Picone, P., et al., Aβ oligomers and fibrillar aggregates induce different apoptotic pathways in LAN5 neuroblastoma cell cultures. Biophysical journal, 2009. **96**(10): p. 4200-4211.
25. Townsend, M., et al., Effects of secreted oligomers of amyloid β-protein on hippocampal synaptic plasticity: a potent role for trimers. The Journal of physiology, 2006. **572**(2): p. 477-492.

26. Shankar, G.M., et al., Amyloid-β protein dimers isolated directly from Alzheimer's brains impair synaptic plasticity and memory. Nature medicine, 2008. **14**(8): p. 837-842.
27. Hartley, D.M., et al., Protofibrillar intermediates of amyloid β-protein induce acute electrophysiological changes and progressive neurotoxicity in cortical neurons. Journal of Neuroscience, 1999. **19**(20): p. 8876-8884.
28. Verma, M., A. Vats, and V. Taneja, Toxic species in amyloid disorders: Oligomers or mature fibrils. Annals of Indian Academy of Neurology, 2015. **18**(2): p. 138.
29. Balch, W.E., et al., Adapting proteostasis for disease intervention. science, 2008. **319**(5865): p. 916-919.
30. Powers, E.T., et al., Biological and chemical approaches to diseases of proteostasis deficiency. Annual review of biochemistry, 2009. **78**: p. 959-991.
31. Tartaglia, G.G., et al., Life on the edge: a link between gene expression levels and aggregation rates of human proteins. Trends in biochemical sciences, 2007. **32**(5): p. 204-205.
32. Walter, P. and D. Ron, The unfolded protein response: from stress pathway to homeostatic regulation. Science, 2011. **334**(6059): p. 1081-1086.
33. Frydman, J., Folding of newly translated proteins in vivo: the role of molecular chaperones. Annual review of biochemistry, 2001. **70**(1): p. 603-647.
34. Thulasiraman, V., C.F. Yang, and J. Frydman, In vivo newly translated polypeptides are sequestered in a protected folding environment. The EMBO journal, 1999. **18**(1): p. 85-95.
35. Ibba, M. and D. Söll, Quality control mechanisms during translation. Science, 1999. **286**(5446): p. 1893-1897.
36. Mandal, A.K., et al., Cdc37 has distinct roles in protein kinase quality control that protect nascent chains from degradation and promote posttranslational maturation. J Cell Biol, 2007. **176**(3): p. 319-328.
37. Connell, P., et al., The co-chaperone CHIP regulates protein triage decisions mediated by heat-shock proteins. Nature cell biology, 2001. **3**(1): p. 93-96.
38. Wickner, S., M.R. Maurizi, and S. Gottesman, Posttranslational quality control: folding, refolding, and degrading proteins. Science, 1999. **286**(5446): p. 1888-1893.
39. MacGurn, J.A., P.-C. Hsu, and S.D. Emr, Ubiquitin and membrane protein turnover: from cradle to grave. Annual review of biochemistry, 2012. **81**: p. 231-259.
40. Ciryam, P., et al., Widespread aggregation and neurodegenerative diseases are associated with supersaturated proteins. Cell reports, 2013. **5**(3): p. 781-790.
41. Ciryam, P., et al., Supersaturation is a major driving force for protein aggregation in neurodegenerative diseases. Trends in pharmacological sciences, 2015. **36**(2): p. 72-77.
42. Walther, D.M., et al., Widespread Proteome Remodeling and Aggregation in Aging C. elegans. Cell, 2015. **161**(4): p. 919-932.
43. Jackson, W.S., Selective vulnerability to neurodegenerative disease: the curious case of Prion Protein. Disease models & mechanisms, 2014. **7**(1): p. 21-29.
44. Voisine, C., J.S. Pedersen, and R.I. Morimoto, Chaperone networks: tipping the balance in protein folding diseases. Neurobiology of disease, 2010. **40**(1): p. 12-20.
45. Gidalevitz, T., et al., Progressive disruption of cellular protein folding in models of polyglutamine diseases. Science, 2006. **311**(5766): p. 1471-1474.
46. Satyal, S.H., et al., Polyglutamine aggregates alter protein folding homeostasis in Caenorhabditis elegans. Proceedings of the National Academy of Sciences, 2000. **97**(11): p. 5750-5755.
47. Westermark, P., K. Lundmark, and G.T. Westermark, Fibrils from designed non-amyloid-related synthetic peptides induce AA-amyloidosis during inflammation in an animal model. PLoS One, 2009. **4**(6): p. e6041.
48. David, D.C., et al., Widespread protein aggregation as an inherent part of aging in C. elegans. PLoS biology, 2010. **8**(8): p. 1925.
49. Reddy, B., S. Datta, and S. Tiwari, Use of propensities of amino acids to the local structural environments to understand effect of substitution mutations on protein stability. Protein engineering, 1998. **11**(12): p. 1137-1145.
50. Soto, C., Unfolding the role of protein misfolding in neurodegenerative diseases. Nature reviews. Neuroscience, 2003. **4**(1): p. 49.

51. Ross, C.A. and M.A. Poirier, Protein aggregation and neurodegenerative disease. Nature medicine, 2004. **10**(7): p. S10.
52. Allan Butterfield, D., Amyloid β-peptide (1-42)-induced oxidative stress and neurotoxicity: implications for neurodegeneration in Alzheimer's disease brain. A review. Free radical research, 2002. **36**(12): p. 1307-1313.
53. Barnham, K.J., C.L. Masters, and A.I. Bush, *Neurodegenerative diseases and oxidative stress.* Nature reviews. Drug discovery, 2004. **3**(3): p. 205.
54. Ballatore, C., V.M.-Y. Lee, and J.Q. Trojanowski, Tau-mediated neurodegeneration in Alzheimer's disease and related disorders. Nature reviews. Neuroscience, 2007. **8**(9): p. 663.
55. Perry, V.H., J.A. Nicoll, and C. Holmes, *Microglia in neurodegenerative disease.* Nature Reviews Neurology, 2010. **6**(4): p. 193-201.
56. Hutton, M., et al., Association of missense and 5′-splice-site mutations in tau with the inherited dementia FTDP-17. Nature, 1998. **393**(6686): p. 702-705.
57. Price, D.L. and S.S. Sisodia, Mutant genes in familial Alzheimer's disease and transgenic models. Annual review of neuroscience, 1998. **21**(1): p. 479-505.
58. De Strooper, B. and T. Voet, Alzheimer's disease: A protective mutation. Nature, 2012. **488**(7409): p. 38-39.
59. Näslund, J., et al., Correlation between elevated levels of amyloid β-peptide in the brain and cognitive decline. Jama, 2000. **283**(12): p. 1571-1577.
60. Maeda, S., et al., Increased levels of granular tau oligomers: an early sign of brain aging and Alzheimer's disease. Neuroscience research, 2006. **54**(3): p. 197-201.
61. Malaplate-Armand, C., et al., Soluble oligomers of amyloid-β peptide induce neuronal apoptosis by activating a cPLA 2-dependent sphingomyelinase-ceramide pathway. Neurobiology of disease, 2006. **23**(1): p. 178-189.
62. White, J.A., et al., Differential effects of oligomeric and fibrillar amyloid-β1–42 on astrocyte-mediated inflammation. Neurobiology of disease, 2005. **18**(3): p. 459-465.
63. Ma, Q.-L., et al., β-amyloid oligomers induce phosphorylation of tau and inactivation of insulin receptor substrate via c-Jun N-terminal kinase signaling: suppression by omega-3 fatty acids and curcumin. Journal of Neuroscience, 2009. **29**(28): p. 9078-9089.
64. Olzscha, H., et al., Amyloid-like aggregates sequester numerous metastable proteins with essential cellular functions. Cell, 2011. **144**(1): p. 67-78.
65. Mashiko, T., et al., Developmentally Regulated RNA-binding Protein 1 (Drb1)/RNA-binding Motif Protein 45 (RBM45), a Nuclear-Cytoplasmic Trafficking Protein, Forms TAR DNA-binding Protein 43 (TDP-43)-mediated Cytoplasmic Aggregates. Journal of Biological Chemistry, 2016. **291**(29): p. 14996-15007.
66. Kartner, N., 0. Augustinas, TJ Jensen, AL Naismith, and JR Riordan. 1992. Mislocalization of delta F508 CFTR in cystic fibrosis sweat gland. Nature Genetics. **1**: p. 321-327.
67. Cheng, S.H., et al., Defective intracellular transport and processing of CFTR is the molecular basis of most cystic fibrosis. Cell, 1990. **63**(4): p. 827-834.
68. Welsh, M.J. and A.E. Smith, Molecular mechanisms of CFTR chloride channel dysfunction in cystic fibrosis. Cell, 1993. **73**(7): p. 1251-1254.
69. Coelho, M. and I.M. Tolić, Asymmetric damage segregation at cell division via protein aggregate fusion and attachment to organelles. BioEssays, 2015. **37**(7): p. 740-747.
70. Glabe, C.G. and R. Kayed, Common structure and toxic function of amyloid oligomers implies a common mechanism of pathogenesis. Neurology, 2006. **66**(1 suppl 1): p. S74-S78.
71. Mannini, B., et al., Toxicity of protein oligomers is rationalized by a function combining size and surface hydrophobicity. ACS chemical biology, 2014. **9**(10): p. 2309-2317.
72. Walsh, D.M., et al., Naturally secreted oligomers of amyloid β protein potently inhibit hippocampal long-term potentiation in vivo. Nature, 2002. **416**(6880): p. 535-539.
73. Wang, H.-W., et al., Soluble oligomers of β amyloid (1-42) inhibit long-term potentiation but not long-term depression in rat dentate gyrus. Brain research, 2002. **924**(2): p. 133-140.
74. Cleary, J.P., et al., Natural oligomers of the amyloid-β protein specifically disrupt cognitive function. Nature neuroscience, 2005. **8**(1): p. 79-84.
75. Hipp, M.S., S.-H. Park, and F.U. Hartl, Proteostasis impairment in protein-misfolding and-aggregation diseases. Trends in cell biology, 2014. **24**(9): p. 506-514.

76. Whitwell, J.L., et al., MRI correlates of neurofibrillary tangle pathology at autopsy A voxel-based morphometry study. Neurology, 2008. **71**(10): p. 743-749.
77. Jacobs, D., et al., Age at onset of Alzheimer's disease Relation to pattern of cognitive dysfunction and rate of decline. Neurology, 1994. **44**(7): p. 1215-1215.
78. Tanzi, R.E. and L. Bertram, Twenty years of the Alzheimer's disease amyloid hypothesis: a genetic perspective. Cell, 2005. **120**(4): p. 545-555.
79. Pagon, R., et al., Early-Onset Familial Alzheimer Disease--GeneReviews (®).
80. Brookmeyer, R., et al., Forecasting the global burden of Alzheimer's disease. Alzheimer's & dementia, 2007. **3**(3): p. 186-191.
81. Shemesh, N., N. Shai, and A. Ben-Zvi, Germline stem cell arrest inhibits the collapse of somatic proteostasis early in Caenorhabditis elegans adulthood. Aging cell, 2013. **12**(5): p. 814-822.
82. Khodakarami, A., et al., Mediation of organismal aging and somatic proteostasis by the germline. Frontiers in Molecular Biosciences, 2015. **2**.
83. Schlenzig, D., et al., Pyroglutamate formation influences solubility and amyloidogenicity of amyloid peptides. Biochemistry, 2009. **48**(29): p. 7072-7078.
84. Lindner, A.B., et al., Asymmetric segregation of protein aggregates is associated with cellular aging and rejuvenation. Proceedings of the National Academy of Sciences, 2008. **105**(8): p. 3076-3081.
85. Mortimer, J.A., et al., Head injury as a risk factor for Alzheimer's disease. Neurology, 1985. **35**(2): p. 264-264.
86. Roberts, G., et al., Beta amyloid protein deposition in the brain after severe head injury: implications for the pathogenesis of Alzheimer's disease. Journal of Neurology, Neurosurgery & Psychiatry, 1994. **57**(4): p. 419-425.
87. Washington, P.M., et al., Experimental traumatic brain injury induces rapid aggregation and oligomerization of amyloid-beta in an Alzheimer's disease mouse model. Journal of neurotrauma, 2014. **31**(1): p. 125-134.
88. Goedert, M., F. Clavaguera, and M. Tolnay, The propagation of prion-like protein inclusions in neurodegenerative diseases. Trends in neurosciences, 2010. **33**(7): p. 317-325.
89. Damoiseaux, J.S. and M.D. Greicius, Greater than the sum of its parts: a review of studies combining structural connectivity and resting-state functional connectivity. Brain Structure and Function, 2009. **213**(6): p. 525-533.
90. Collinge, J., Prion diseases of humans and animals: their causes and molecular basis. Annual review of neuroscience, 2001. **24**(1): p. 519-550.
91. Kyle, R.A., *Amyloidosis: a convoluted story.* British journal of haematology, 2001. **114**(3): p. 529-538.
92. Mahendra, B., Dementia: a brief history of the concept, in Dementia. 1987, Springer. p. 1-18.
93. Prince, M.J., World Alzheimer Report 2015: the global impact of dementia: an analysis of prevalence, incidence, cost and trends. 2015: Alzheimer's Disease International.
94. Hebert, L.E., et al., Alzheimer disease in the United States (2010–2050) estimated using the 2010 census. Neurology, 2013. **80**(19): p. 1778-1783.
95. Sosa-Ortiz, A.L., I. Acosta-Castillo, and M.J. Prince, *Epidemiology of dementias and Alzheimer's disease.* Archives of medical research, 2012. **43**(8): p. 600-608.
96. Langa, K.M., Is the risk of Alzheimer's disease and dementia declining? Alzheimer's research & therapy, 2015. **7**(1): p. 34.
97. Okie, S., *Confronting Alzheimer's disease.* The New England journal of medicine, 2011. **365**(12): p. 1069.
98. Whitwell, J., et al., MRI correlates of neurofibrillary tangle pathology at autopsy A voxel-based morphometry study. Neurology, 2008. **71**(10): p. 743-749.
99. Henderson, A., et al., Environmental risk factors for Alzheimer's disease: their relationship to age of onset and to familial or sporadic types. Psychological medicine, 1992. **22**(02): p. 429-436.
100. Arking, R., Biology of aging: observations and principles. 2006: Oxford University Press.
101. Avila, J., et al., Role of tau protein in both physiological and pathological conditions. Physiological reviews, 2004. **84**(2): p. 361-384.

102. Citron, M., et al., Mutation of the beta-amyloid precursor protein in familial Alzheimer's disease increases beta-protein production. Nature, 1992. **360**(6405): p. 672.
103. Nilsberth, C., et al., The'Arctic'APP mutation (E693G) causes Alzheimer's disease by enhanced Aβ protofibril formation. Nature neuroscience, 2001. **4**(9): p. 887-893.
104. Braak, H., et al., Staging of brain pathology related to sporadic Parkinson's disease. Neurobiology of aging, 2003. **24**(2): p. 197-211.
105. Polymeropoulos, M.H., et al., Mutation in the α-synuclein gene identified in families with Parkinson's disease. science, 1997. **276**(5321): p. 2045-2047.
106. Ravits, J.M. and A.R. La Spada, ALS motor phenotype heterogeneity, focality, and spread Deconstructing motor neuron degeneration. Neurology, 2009. **73**(10): p. 805-811.
107. Sreedharan, J., et al., TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. Science, 2008. **319**(5870): p. 1668-1672.
108. Jackson, W.S., et al., Profoundly different prion diseases in knock-in mice carrying single PrP codon substitutions associated with human diseases. Proceedings of the National Academy of Sciences, 2013. **110**(36): p. 14759-14764.
109. Spokes, E., Neurochemical alterations in Huntington's chorea: a study of post-mortem brain tissue. Brain: a journal of neurology, 1980. **103**(1): p. 179-210.
110. Brinkman, R., et al., The likelihood of being affected with Huntington disease by a particular age, for a specific CAG size. American journal of human genetics, 1997. **60**(5): p. 1202.
111. Orr, H.T., et al., Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type 1. Nature genetics, 1993. **4**(3): p. 221-226.
112. Zoghbi, H.Y., *Spinocerebellar ataxias.* Neurobiology of disease, 2000. **7**(5): p. 523-527.
113. Saxena, S. and P. Caroni, Selective neuronal vulnerability in neurodegenerative diseases: from stressor thresholds to degeneration. Neuron, 2011. **71**(1): p. 35-48.
114. Selkoe, D.J., Alzheimer's disease is a synaptic failure. Science, 2002. **298**(5594): p. 789-791.
115. Yoshiyama, Y., et al., Synapse loss and microglial activation precede tangles in a P301S tauopathy mouse model. Neuron, 2007. **53**(3): p. 337-351.
116. Raj, A., A. Kuceyeski, and M. Weiner, A network diffusion model of disease progression in dementia. Neuron, 2012. **73**(6): p. 1204-1215.
117. Bero, A.W., et al., Bidirectional relationship between functional connectivity and amyloid-β deposition in mouse brain. Journal of Neuroscience, 2012. **32**(13): p. 4334-4340.
118. Greicius, M.D., et al., Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI. Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(13): p. 4637-4642.
119. Raj, A., et al., Network diffusion model of progression predicts longitudinal patterns of atrophy and metabolism in Alzheimer's disease. Cell reports, 2015. **10**(3): p. 359-369.
120. Ibrahim, T. and J. McLaurin, Protein seeding in Alzheimer's disease and Parkinson's disease: Similarities and differences. World J Neurol, 2014. **4**(4): p. 23-35.
121. Bero, A.W., et al., Neuronal activity regulates the regional vulnerability to amyloid-[beta] deposition. Nature neuroscience, 2011. **14**(6): p. 750-756.
122. Abdelnour, F., H.U. Voss, and A. Raj, Network diffusion accurately models the relationship between structural and functional brain connectivity networks. Neuroimage, 2014. **90**: p. 335-347.
123. Roselli, F. and P. Caroni, From intrinsic firing properties to selective neuronal vulnerability in neurodegenerative diseases. Neuron, 2015. **85**(5): p. 901-910.
124. Attwell, D. and S.B. Laughlin, An energy budget for signaling in the grey matter of the brain. Journal of Cerebral Blood Flow & Metabolism, 2001. **21**(10): p. 1133-1145.
125. Braak, H. and E. Braak, Neuropathological stageing of Alzheimer-related changes. Acta neuropathologica, 1991. **82**(4): p. 239-259.
126. Fehm, H., W. Kern, and A. Peters, The selfish brain: competition for energy resources. Progress in brain research, 2006. **153**: p. 129-140.
127. Arnaiz, E., et al., Impaired cerebral glucose metabolism and cognitive functioning predict deterioration in mild cognitive impairment. Neuroreport, 2001. **12**(4): p. 851-855.
128. Varki, A., et al., Essentials of glycobiology cold spring harbor laboratory press. New York, 1999: p. 66-84.

129. Hassell, J.R., et al., Isolation of a heparan sulfate-containing proteoglycan from basement membrane. Proceedings of the National Academy of Sciences, 1980. **77**(8): p. 4494-4498.
130. Rabenstein, D.L., Heparin and heparan sulfate: structure and function. Natural product reports, 2002. **19**(3): p. 312-331.
131. Snow, A.D., J.P. Willmer, and R. Kisilevsky, Sulfated glycosaminoglycans in Alzheimer's disease. Human pathology, 1987. **18**(5): p. 506-510.
132. Díaz-Nido, J., F. Wandosell, and J. Avila, Glycosaminoglycans and β-amyloid, prion and tau peptides in neurodegenerative diseases. Peptides, 2002. **23**(7): p. 1323-1332.
133. McLaurin, J., et al., A sulfated proteoglycan aggregation factor mediates amyloid-β peptide fibril formation and neurotoxicity. Amyloid, 1999. **6**(4): p. 233-243.
134. Suk, J.Y., et al., Heparin accelerates gelsolin amyloidogenesis. Biochemistry, 2006. **45**(7): p. 2234-2242.
135. Bourgault, S., et al., Sulfated glycosaminoglycans accelerate transthyretin amyloidogenesis by quaternary structural conversion. Biochemistry, 2011. **50**(6): p. 1001-1015.
136. McLaurin, J., et al., Interactions of Alzheimer amyloid-β peptides with glycosaminoglycans. European journal of biochemistry, 1999. **266**(3): p. 1101-1110.
137. Ancsin, J.B., Amyloidogenesis: historical and modern observations point to heparan sulfate proteoglycans as a major culprit. Amyloid, 2003. **10**(2): p. 67-79.
138. Stohs, S.J. and D. Bagchi, Oxidative mechanisms in the toxicity of metal ions. Free radical biology and medicine, 1995. **18**(2): p. 321-336.
139. Noy, D., et al., Zinc-amyloid β interactions on a millisecond time-scale stabilize non-fibrillar Alzheimer-related species. Journal of the American Chemical Society, 2008. **130**(4): p. 1376-1383.
140. Lin, C.-H., et al., Neurological abnormalities in a knock-in mouse model of Huntington's disease. Human molecular genetics, 2001. **10**(2): p. 137-144.
141. Watase, K., et al., A long CAG repeat in the mouse Sca1 locus replicates SCA1 features and reveals the impact of protein solubility on selective neurodegeneration. Neuron, 2002. **34**(6): p. 905-919.
142. Jackson, W.S., et al., Spontaneous generation of prion infectivity in fatal familial insomnia knockin mice. Neuron, 2009. **63**(4): p. 438-450.
143. Arcelli, P., et al., GABAergic neurons in mammalian thalamus: a marker of thalamic complexity? Brain research bulletin, 1997. **42**(1): p. 27-37.
144. Heneka, M.T. and M.K. O'Banion, Inflammatory processes in Alzheimer's disease. Journal of neuroimmunology, 2007. **184**(1): p. 69-91.
145. Guerreiro, R., et al., *TREM2 variants in Alzheimer's disease.* New England Journal of Medicine, 2013. **368**(2): p. 117-127.
146. Malik, M., et al., CD33 Alzheimer's risk-altering polymorphism, CD33 expression, and exon 2 splicing. Journal of Neuroscience, 2013. **33**(33): p. 13320-13325.
147. Wang, W.-Y., et al., Role of pro-inflammatory cytokines released from microglia in Alzheimer's disease. Annals of translational medicine, 2015. **3**(10).
148. Wotton, C.J. and M.J. Goldacre, Associations between specific autoimmune diseases and subsequent dementia: retrospective record-linkage cohort study, UK. J Epidemiol Community Health, 2017: p. jech-2016-207809.
149. Biessels, G. and L. Kappelle, Increased risk of Alzheimer's disease in Type II diabetes: insulin resistance of the brain or insulin-induced amyloid pathology?, 2005, Portland Press Limited.
150. Block, M.L., L. Zecca, and J.-S. Hong, Microglia-mediated neurotoxicity: uncovering the molecular mechanisms. Nature Reviews Neuroscience, 2007. **8**(1): p. 57-69.
151. Reed-Geaghan, E.G., et al., CD14 and toll-like receptors 2 and 4 are required for fibrillar Aβ-stimulated microglial activation. Journal of Neuroscience, 2009. **29**(38): p. 11982-11992.
152. Block, M.L. and J.-S. Hong, Microglia and inflammation-mediated neurodegeneration: multiple triggers with a common mechanism. Progress in neurobiology, 2005. **76**(2): p. 77-98.
153. Freer, R., et al., A protein homeostasis signature in healthy brains recapitulates tissue vulnerability to Alzheimer's disease. Science Advances, 2016. **2**(8): p. e1600947.
154. Muchowski, P.J. and J.L. Wacker, Modulation of neurodegeneration by molecular chaperones. Nature reviews. Neuroscience, 2005. **6**(1): p. 11.

155. Wilhelmus, M.M., R.M. De Waal, and M.M. Verbeek, Heat shock proteins and amateur chaperones in amyloid-Beta accumulation and clearance in Alzheimer's disease. Molecular neurobiology, 2007. **35**(3): p. 203-216.
156. Weidemann, A., et al., Identification, biogenesis, and localization of precursors of Alzheimer's disease A4 amyloid protein. Cell, 1989. **57**(1): p. 115-126.
157. Sisodia, S.S., Beta-amyloid precursor protein cleavage by a membrane-bound protease. Proceedings of the National Academy of Sciences, 1992. **89**(13): p. 6075-6079.
158. Haass, C. and D.J. Selkoe, Cellular processing of β-amyloid precursor protein and the genesis of amyloid β-peptide. Cell, 1993. **75**(6): p. 1039-1042.
159. Tomita, S., Y. Kirino, and T. Suzuki, Cleavage of Alzheimer's amyloid precursor protein (APP) by secretases occurs after O-glycosylation of APP in the protein secretory pathway Identification of intracellular compartments in which APP cleavage occurs without using toxic agents that interfere with protein metabolism. Journal of Biological Chemistry, 1998. **273**(11): p. 6277-6284.
160. Jäger, S., et al., α-secretase mediated conversion of the amyloid precursor protein derived membrane stub C99 to C83 limits Aβ generation. Journal of neurochemistry, 2009. **111**(6): p. 1369-1382.
161. Epis, R., et al., Blocking ADAM10 synaptic trafficking generates a model of sporadic Alzheimer's disease. Brain, 2010. **133**(11): p. 3323-3335.
162. Gralle, M., M.G. Botelho, and F.S. Wouters, Neuroprotective secreted amyloid precursor protein acts by disrupting amyloid precursor protein dimers. Journal of Biological Chemistry, 2009. **284**(22): p. 15016-15025.
163. Kim, M., et al., Potential late-onset Alzheimer's disease-associated mutations in the ADAM10 gene attenuate α-secretase activity. Human molecular genetics, 2009. **18**(20): p. 3987-3996.
164. Postina, R., et al., A disintegrin-metalloproteinase prevents amyloid plaque formation and hippocampal defects in an Alzheimer disease mouse model. Journal of Clinical Investigation, 2004. **113**(10): p. 1456.
165. Vassar, R., et al., β-Secretase Cleavage of Alzheimer9s Amyloid Precursor Protein by the Transmembrane Aspartic Protease BACE. science, 1999. **286**(5440): p. 735-741.
166. Cai, H., et al., BACE1 is the major [beta]-secretase for generation of A [beta] peptides by neurons. Nature neuroscience, 2001. **4**(3): p. 233.
167. Luo, Y., et al., Mice deficient in BACE1, the Alzheimer's [beta]-secretase, have normal phenotype and abolished [beta]-amyloid generation. Nature neuroscience, 2001. **4**(3): p. 231.
168. Roberds, S.L., et al., BACE knockout mice are healthy despite lacking the primary β-secretase activity in brain: implications for Alzheimer's disease therapeutics. Human molecular genetics, 2001. **10**(12): p. 1317-1324.
169. Small, S.A. and S. Gandy, Sorting through the cell biology of Alzheimer's disease: intracellular pathways to pathogenesis. Neuron, 2006. **52**(1): p. 15-31.
170. Jarrett, J.T., E.P. Berger, and P.T. Lansbury Jr, The carboxy terminus of the. beta. amyloid protein is critical for the seeding of amyloid formation: Implications for the pathogenesis of Alzheimer's disease. Biochemistry, 1993. **32**(18): p. 4693-4697.
171. Sun, X., et al., Lithium inhibits amyloid secretion in COS7 cells transfected with amyloid precursor protein C100. Neuroscience letters, 2002. **321**(1): p. 61-64.
172. Phiel, C.J., et al., GSK-3α regulates production of Alzheimer's disease amyloid-β peptides. Nature, 2003. **423**(6938): p. 435-439.
173. Evans, C.G., S. Wisén, and J.E. Gestwicki, Heat shock proteins 70 and 90 inhibit early stages of amyloid β-(1–42) aggregation in vitro. Journal of Biological Chemistry, 2006. **281**(44): p. 33182-33191.
174. Magrané, J., et al., Heat shock protein 70 participates in the neuroprotective response to intracellularly expressed β-amyloid in neurons. Journal of Neuroscience, 2004. **24**(7): p. 1700-1706.
175. Ostapchenko, V.G., et al., The prion protein ligand, stress-inducible phosphoprotein 1, regulates amyloid-β oligomer toxicity. The Journal of Neuroscience, 2013. **33**(42): p. 16552-16564.

176. Månsson, C., et al., Interaction of the molecular chaperone DNAJB6 with growing amyloid-beta 42 (Aβ42) aggregates leads to sub-stoichiometric inhibition of amyloid formation. Journal of Biological Chemistry, 2014. **289**(45): p. 31066-31076.
177. Bakthisaran, R., et al., αB-crystallin, a small heat-shock protein, prevents the amyloid fibril growth of an amyloid β-peptide and β2-microglobulin. Biochemical Journal, 2005. **392**(3): p. 573-581.
178. Lowe, J., et al., αB crystallin expression in nonlenticular tissues and selective presence in ubiquitinated inclusion bodies in human disease. The Journal of pathology, 1992. **166**(1): p. 61-68.
179. Wilhelmus, M.M., et al., Small heat shock proteins inhibit amyloid-β protein aggregation and cerebrovascular amyloid-β protein toxicity. Brain research, 2006. **1089**(1): p. 67-78.
180. Ojha, J., et al., Sequestration of toxic oligomers by HspB1 as a cytoprotective mechanism. Molecular and cellular biology, 2011. **31**(15): p. 3146-3157.
181. Kakimura, J., et al., Possible Involvement of ER Chaperone Grp78 on Reduced Formation of Amyloid-β Deposits. Annals of the New York Academy of Sciences, 2002. **977**(1): p. 327-332.
182. Yang, Y., R.S. Turner, and J.R. Gaut, The chaperone BiP/GRP78 binds to amyloid precursor protein and decreases Aβ40 and Aβ42 secretion. Journal of Biological Chemistry, 1998. **273**(40): p. 25552-25555.
183. Holtzman, D.M., In vivo effects of ApoE and clusterin on amyloid-β metabolism and neuropathology. Journal of Molecular Neuroscience, 2004. **23**(3): p. 247-254.
184. Hanger, D.P., B.H. Anderton, and W. Noble, Tau phosphorylation: the therapeutic challenge for neurodegenerative disease. Trends in molecular medicine, 2009. **15**(3): p. 112-119.
185. Hernandez, F., J.J. Lucas, and J. Avila, GSK3 and tau: two convergence points in Alzheimer's disease. Journal of Alzheimer's Disease, 2013. **33**(s1): p. S141-S144.
186. Hanger, D.P., et al., Glycogen synthase kinase-3 induces Alzheimer's disease-like phosphorylation of tau: generation of paired helical filament epitopes and neuronal localisation of the kinase. Neuroscience letters, 1992. **147**(1): p. 58-62.
187. Ishiguro, K., et al., Phosphorylation sites on tau by tau protein kinase I, a bovine derived kinase generating an epitope of paired helical filaments. Neuroscience letters, 1992. **148**(1): p. 202-206.
188. Lovestone, S., et al., Alzheimer's disease-like phosphorylation of the microtubule-associated protein tau by glycogen synthase kinase-3 in transfected mammalian cells. Current Biology, 1994. **4**(12): p. 1077-1086.
189. Cho, J.-H. and G.V. Johnson, Glycogen Synthase Kinase 3β Phosphorylates tau at both primed and unprimed sites differential impact on microtubule binding. Journal of Biological Chemistry, 2003. **278**(1): p. 187-193.
190. Asuni, A.A., et al., GSK3α exhibits β-catenin and tau directed kinase activities that are modulated by Wnt. European Journal of Neuroscience, 2006. **24**(12): p. 3387-3392.
191. Bramblett, G.T., et al., Abnormal tau phosphorylation at Ser 396 in Alzheimer's disease recapitulates development and contributes to reduced microtubule binding. Neuron, 1993. **10**(6): p. 1089-1099.
192. Yoshida, H. and Y. Ihara, τ in Paired Helical Filaments Is Functionally Distinct from Fetal τ: Assembly Incompetence of Paired Helical Filament-τ. Journal of neurochemistry, 1993. **61**(3): p. 1183-1186.
193. Sepulveda-Diaz, J.E., et al., HS3ST2 expression is critical for the abnormal phosphorylation of tau in Alzheimer's disease-related tau pathology. Brain, 2015. **138**(5): p. 1339-1354.
194. Sepulveda-Diaz, J.E., et al., HS3ST2 expression is critical for the abnormal phosphorylation of tau in Alzheimer's disease-related tau pathology. Brain, 2015. **138**(5): p. 1339-1354.
195. Smith, M.C., et al., The E3 ubiquitin ligase CHIP and the molecular chaperone Hsc70 form a dynamic, tethered complex. Biochemistry, 2013. **52**(32): p. 5354-5364.
196. Elliott, E., P. Tsvetkov, and I. Ginzburg, BAG-1 associates with Hsc70· Tau complex and regulates the proteasomal degradation of Tau protein. Journal of Biological Chemistry, 2007. **282**(51): p. 37276-37284.
197. Jinwal, U.K., et al., Imbalance of Hsp70 family variants fosters tau accumulation. The FASEB Journal, 2013. **27**(4): p. 1450-1459.
198. Shimura, H., et al., CHIP-Hsc70 complex ubiquitinates phosphorylated tau and enhances cell survival. Journal of Biological Chemistry, 2004. **279**(6): p. 4869-4876.

199. Abisambra, J.F., et al., DnaJA1 antagonizes constitutive Hsp70-mediated stabilization of tau. Journal of molecular biology, 2012. **421**(4): p. 653-661.
200. Arulselvam, K., DnaJ Molecular Chaperones Modulate Tau Levels. 2011.
201. Jinwal, U.K., et al., The Hsp90 cochaperone, FKBP51, increases Tau stability and polymerizes microtubules. The Journal of Neuroscience, 2010. **30**(2): p. 591-599.
202. Blair, L.J., et al., Accelerated neurodegeneration through chaperone-mediated oligomerization of tau. The Journal of clinical investigation, 2013. **123**(10): p. 4158.
203. Jinwal, U.K., et al., The Hsp90 kinase co-chaperone Cdc37 regulates tau stability and phosphorylation dynamics. Journal of Biological Chemistry, 2011. **286**(19): p. 16976-16983.
204. Lei, Z., C. Brizzee, and G.V. Johnson, BAG3 facilitates the clearance of endogenous tau in primary neurons. Neurobiology of aging, 2015. **36**(1): p. 241-248.
205. Hawrylycz, M.J., et al., An anatomically comprehensive atlas of the adult human brain transcriptome. Nature, 2012. **489**(7416): p. 391-399.
206. Fisher, R., Statistical methods for research workers, in Breakthroughs in Statistics. 1992, Springer. p. 66-70.
207. Benjamini, Y. and Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B (Methodological), 1995: p. 289-300.
208. Uhlén, M., et al., Tissue-based map of the human proteome. Science, 2015. **347**(6220): p. 1260419.
209. Čaušević, M., et al., β-Amyloid precursor protein and tau protein levels are differently regulated in human cerebellum compared to brain regions vulnerable to Alzheimer's type neurodegeneration. Neuroscience letters, 2010. **485**(3): p. 162-166.
210. Labbadia, J. and R.I. Morimoto, The Biology of Proteostasis in Aging and Disease. Annu Rev Biochem, 2015.
211. Habchi, J., et al., An anticancer drug suppresses the primary nucleation reaction that initiates the production of the toxic Aβ42 aggregates linked with Alzheimer's disease. Science advances, 2016. **2**(2): p. e1501244.
212. Kompoliti, K. and L. Verhagen, Encyclopedia of Movement Disorders, Three-Volume Set. Vol. 1. 2010: Academic Press.
213. Mann, H.B. and D.R. Whitney, On a test of whether one of two random variables is stochastically larger than the other. The annals of mathematical statistics, 1947: p. 50-60.
214. Storey, J.D. and R. Tibshirani, *Statistical significance for genomewide studies.* Proceedings of the National Academy of Sciences, 2003. **100**(16): p. 9440-9445.
215. Pawitan, Y., et al., False discovery rate, sensitivity and sample size for microarray studies. Bioinformatics, 2005. **21**(13): p. 3017-3024.
216. Desikan, R.S., et al., An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage, 2006. **31**(3): p. 968-980.
217. Destrieux, C., et al., Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. Neuroimage, 2010. **53**(1): p. 1-15.
218. Klein, A. and J. Tourville, 101 labeled brain images and a consistent human cortical labeling protocol. Front. Neurosci, 2012. **6**(171): p. 10.3389.
219. Consortium, U., *UniProt: a hub for protein information.* Nucleic acids research, 2014: p. gku989.
220. Azevedo, F.A., et al., Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. Journal of Comparative Neurology, 2009. **513**(5): p. 532-541.
221. Verkhratsky, A., et al., Astrocytes in Alzheimer's disease. Neurotherapeutics, 2010. **7**(4): p. 399-412.
222. Newman, E.A., D.A. Frambach, and L.L. Odette, Control of extracellular potassium levels by retinal glial cell K+ siphoning. Science (New York, NY), 1984. **225**(4667): p. 1174.
223. Kofuji, P. and E. Newman, Potassium buffering in the central nervous system. Neuroscience, 2004. **129**(4): p. 1043-1054.
224. Bushong, E.A., et al., Protoplasmic astrocytes in CA1 stratum radiatum occupy separate anatomical domains. Journal of Neuroscience, 2002. **22**(1): p. 183-192.

225. Nedergaard, M., B. Ransom, and S.A. Goldman, New roles for astrocytes: redefining the functional architecture of the brain. Trends in neurosciences, 2003. **26**(10): p. 523-530.
226. Bushong, E.A., M.E. Martone, and M.H. Ellisman, Maturation of astrocyte morphology and the establishment of astrocyte domains during postnatal hippocampal development. International Journal of Developmental Neuroscience, 2004. **22**(2): p. 73-86.
227. Rodriguez, J., et al., Astroglia in dementia and Alzheimer's disease. Cell death and differentiation, 2009. **16**(3): p. 378.
228. Heneka, M.T., J.J. Rodríguez, and A. Verkhratsky, *Neuroglia in neurodegeneration.* Brain research reviews, 2010. **63**(1): p. 189-211.
229. McGeer, P.L. and E.G. McGeer, Glial reactions in Parkinson's disease. Movement Disorders, 2008. **23**(4): p. 474-483.
230. Mena, M.A. and J. García de Yébenes, Glial cells as players in parkinsonism: the “good,” the “bad,” and the “mysterious” glia. The Neuroscientist, 2008. **14**(6): p. 544-560.
231. Booth, H.D., W.D. Hirst, and R. Wade-Martins, The Role of Astrocyte Dysfunction in Parkinson’s Disease Pathogenesis. Trends in Neurosciences, 2017.
232. Haidet-Phillips, A.M., et al., Astrocytes from familial and sporadic ALS patients are toxic to motor neurons. Nature biotechnology, 2011. **29**(9): p. 824-828.
233. Mandrekar-Colucci, S. and G.E. Landreth, *Microglia and inflammation in Alzheimer's disease.* CNS & Neurological Disorders-Drug Targets (Formerly Current Drug Targets-CNS & Neurological Disorders), 2010. **9**(2): p. 156-167.
234. Perlmutter, L.S., E. Barron, and H.C. Chui, Morphologic association between microglia and senile plaque amyloid in Alzheimer's disease. Neuroscience letters, 1990. **119**(1): p. 32-36.
235. Wisniewski, H., et al., Ultrastructural studies of the cells forming amyloid in the cortical vessel wall in Alzheimer's disease. Acta neuropathologica, 1992. **84**(2): p. 117-127.
236. Forno, L.S., *Neuropathology of Parkinson's disease.* Journal of Neuropathology & Experimental Neurology, 1996. **55**(3): p. 259-272.
237. Mirza, B., et al., The absence of reactive astrocytosis is indicative of a unique inflammatory process in Parkinson's disease. Neuroscience, 1999. **95**(2): p. 425-432.
238. Knott, C., G. Stern, and G. Wilkin, Inflammatory regulators in Parkinson's disease: iNOS, lipocortin-1, and cyclooxygenases-1 and-2. Molecular and Cellular Neuroscience, 2000. **16**(6): p. 724-739.
239. Liu, B., H.-M. Gao, and J.-S. Hong, Parkinson's disease and exposure to infectious agents and pesticides and the occurrence of brain injuries: role of neuroinflammation. Environmental health perspectives, 2003. **111**(8): p. 1065.
240. Herrera, A., et al., Inflammatory process as a determinant factor for the degeneration of substantia nigra dopaminergic neurons. Journal of neural transmission, 2005. **112**(1): p. 111-119.
241. Wang, Z., et al., Reduced expression of glutamate transporter EAAT2 and impaired glutamate transport in human primary astrocytes exposed to HIV-1 or gp120. Virology, 2003. **312**(1): p. 60-73.
242. Henkel, J.S., et al., Presence of dendritic cells, MCP-1, and activated microglia/ macrophages in amyotrophic lateral sclerosis spinal cord tissue. Annals of neurology, 2004. **55**(2): p. 221-235.
243. Pickering, M., D. Cumiskey, and J.J. O'Connor, Actions of TNF-α on glutamatergic synaptic transmission in the central nervous system. Experimental physiology, 2005. **90**(5): p. 663-670.
244. Tilleux, S. and E. Hermans, Neuroinflammation and regulation of glial glutamate uptake in neurological disorders. Journal of neuroscience research, 2007. **85**(10): p. 2059-2070.
245. Ilieva, H., M. Polymenidou, and D.W. Cleveland, Non–cell autonomous toxicity in neurodegenerative disorders: ALS and beyond. The Journal of cell biology, 2009. **187**(6): p. 761-772.
246. Woolsey, T.A., et al., Neuronal units linked to microvascular modules in cerebral cortex: response elements for imaging the brain. Cerebral Cortex, 1996. **6**(5): p. 647-660.
247. Lo, E.H., T. Dalkara, and M.A. Moskowitz, Mechanisms, challenges and opportunities in stroke. Nature reviews. Neuroscience, 2003. **4**(5): p. 399.
248. Zuliani, G., et al., Markers of endothelial dysfunction in older subjects with late onset Alzheimer's disease or vascular dementia. Journal of the neurological sciences, 2008. **272**(1): p. 164-170.

249. Grammas, P., Neurovascular dysfunction, inflammation and endothelial activation: implications for the pathogenesis of Alzheimer's disease. Journal of neuroinflammation, 2011. **8**(1): p. 26.
250. Pogue, A.I. and W.J. Lukiw, Angiogenic signaling in Alzheimer's disease. Neuroreport, 2004. **15**(9): p. 1507-1510.
251. Bancher, C., et al., Accumulation of abnormally phosphorylated τ precedes the formation of neurofibrillary tangles in Alzheimer's disease. Brain research, 1989. **477**(1): p. 90-99.
252. Niikura, T., H. Tajima, and Y. Kita, Neuronal cell death in Alzheimer's disease and a neuroprotective factor, humanin. Current neuropharmacology, 2006. **4**(2): p. 139-147.
253. Trojanowski, J.Q., et al., Fatal attractions: abnormal protein aggregation and neuron death in Parkinson's disease and Lewy body dementia. Cell Death & Differentiation, 1998. **5**(10).
254. Davie, C.A., *A review of Parkinson's disease.* British medical bulletin, 2008. **86**(1): p. 109-127.
255. Neumann, M., et al., Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Science, 2006. **314**(5796): p. 130-133.
256. Fu, H., et al., Tau pathology induces excitatory neuron loss, grid cell dysfunction, and spatial memory deficits reminiscent of early Alzheimer's disease. Neuron, 2017. **93**(3): p. 533-541. e5.
257. Small, D.H., S. San Mok, and J.C. Bornstein, OPINION: Alzheimer's disease and A [beta] toxicity: from top to bottom. Nature reviews. Neuroscience, 2001. **2**(8): p. 595.
258. Coleman, P., H. Federoff, and R. Kurlan, A focus on the synapse for neuroprotection in Alzheimer disease and other dementias. Neurology, 2004. **63**(7): p. 1155-1162.
259. Games, D., et al., Alzheimer-type neuropathology in transgenic mice overexpressing V717F β-amyloid precursor protein. Nature, 1995. **373**(6514): p. 523-527.
260. Mucke, L., et al., High-level neuronal expression of Aβ1–42 in wild-type human amyloid protein precursor transgenic mice: synaptotoxicity without plaque formation. Journal of Neuroscience, 2000. **20**(11): p. 4050-4058.
261. Masliah, E., et al., Altered expression of synaptic proteins occurs early during progression of Alzheimer's disease. Neurology, 2001. **56**(1): p. 127-129.
262. Buttini, M., et al., β-amyloid immunotherapy prevents synaptic degeneration in a mouse model of Alzheimer's disease. Journal of Neuroscience, 2005. **25**(40): p. 9096-9101.
263. Davies, C., et al., A quantitative morphometric analysis of the neuronal and synaptic content of the frontal and temporal cortex in patients with Alzheimer's disease. Journal of the neurological sciences, 1987. **78**(2): p. 151-164.
264. Lue, L.-F., et al., Soluble amyloid β peptide concentration as a predictor of synaptic change in Alzheimer's disease. The American journal of pathology, 1999. **155**(3): p. 853-862.
265. Keating, D.J., Mitochondrial dysfunction, oxidative stress, regulation of exocytosis and their relevance to neurodegenerative diseases. Journal of neurochemistry, 2008. **104**(2): p. 298-305.
266. Darmanis, S., et al., A survey of human brain transcriptome diversity at the single cell level. Proceedings of the National Academy of Sciences, 2015. **112**(23): p. 7285-7290.
267. Lake, B.B., et al., Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. Science, 2016. **352**(6293): p. 1586-1590.
268. Wilhelm, B.G., et al., Composition of isolated synaptic boutons reveals the amounts of vesicle trafficking proteins. Science, 2014. **344**(6187): p. 1023-1028.
269. Xia, Q., et al., Proteomic identification of novel proteins associated with Lewy bodies. Frontiers in bioscience: a journal and virtual library, 2008. **13**: p. 3850.
270. Fearnley, J.M. and A.J. Lees, Ageing and Parkinson's disease: substantia nigra regional selectivity. Brain, 1991. **114**(5): p. 2283-2301.
271. Vogel, C. and E.M. Marcotte, Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. Nature reviews. Genetics, 2012. **13**(4): p. 227.
272. Pontén, F. A Tissue-Based Map of the Human Proteome. in Scandinavian Journal of Immunology. 2015.
273. Sharma, K., et al., Cell type-and brain region-resolved mouse brain proteome. Nature neuroscience, 2015. **18**(12): p. 1819-1831.

274. Reed-Geaghan, E.G., et al., Deletion of CD14 attenuates Alzheimer's disease pathology by influencing the brain's inflammatory milieu. The Journal of Neuroscience, 2010. **30**(46): p. 15369-15373.
275. Gjoneska, E., et al., Conserved epigenomic signals in mice and humans reveal immune basis of Alzheimer/'s disease. Nature, 2015. **518**(7539): p. 365-369.
276. Kanehisa, M., et al., KEGG for representation and analysis of molecular networks involving diseases and drugs. Nucleic acids research, 2009. **38**(suppl_1): p. D355-D360.
277. Tasic, B., et al., Adult mouse cortical cell taxonomy by single cell transcriptomics. Nature neuroscience, 2016. **19**(2): p. 335.
278. Bartlett, A.I. and S.E. Radford, An expanding arsenal of experimental methods yields an explosion of insights into protein folding mechanisms. Nature structural & molecular biology, 2009. **16**(6): p. 582-588.
279. Bucciantini, M., et al., Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. nature, 2002. **416**(6880): p. 507-511.
280. Dobson, C.M., A. Šali, and M. Karplus, Protein folding: a perspective from theory and experiment. Angewandte Chemie International Edition, 1998. **37**(7): p. 868-893.
281. Lindquist, S.L. and J.W. Kelly, Chemical and biological approaches for adapting proteostasis to ameliorate protein misfolding and aggregation diseases–progress and prognosis. Cold Spring Harbor perspectives in biology, 2011. **3**(12): p. a004507.
282. Tsao, D. and N.V. Dokholyan, Macromolecular crowding induces polypeptide compaction and decreases folding cooperativity. Physical Chemistry Chemical Physics, 2010. **12**(14): p. 3491-3500.