AIJR Preprints
Section: Preprints
Article Id: 540, Version: 1, 2024
DOI: https://doi.org/10.21467/preprints.540
URL: https://preprints.aijr.org/index.php/ap/preprint/view/540

{Click on above link or DOI to see the latest available version of this article}

NOT PEER-REVIEWED

Version 1: Received: 10 April 2024 / Approved: 14 April 2024 / Online: 15 April 2024

In Silico Analysis of Mutations Along the Amyloidogenic Pathway in Alzheimer's Disease

Wenqi Zhao*

Research Intern, Cambridge University Centre of Misfolding Diseases

zhao.j@etoncollege.org.uk

ABSTRACT

Through in silico simulation of mutations and their effect on protein structure, we conclusively examine the impact of mutations along the amyloidogenic pathway in three steps: as factors which undermine the suppression of Aβ production from BACE-1; the inhibition of amyloid breakdown by neprilysin; and the aggregation of Aβ monomers through oligomeric and fibril stages. We verified the significance of mutations in miRNA that particularly complement with BACE1. We discovered novel mutations that impede most significantly on neprilysin function. And we examined the importance of mutations on the propensity of Aβ to aggregate. The results are significant: the framework and algorithm of the paper can be employed to make accurate predictions for patients from simple and widely accessible genetic data. Beyond that, given the ubiquity of proteins within our body, the functions for modelling miRNA suppression, predicting protein function and calculating protein aggregation also have widespread uses in all areas of human biology and medicine.

1 Introduction

Alzheimer's disease (AD) is the most common neurodegenerative disease, affecting 55 million people globally (WHO). Its prevalence increases, with ageing being the most common risk factor among others. Approximately 73% of people in the US aged 75 or older are living with a neurodegenerative condition or at a high risk of developing one in coming years[1]. Victims of AD suffer memory loss, communication barriers and hallucination, leading to a gradual descent and regression in health, resulting in death. Since the first discovery of neuritic plaques by Dr Alois Alzheimer in 1904, the amyloid pathway as the pathogenic cause of Alzheimer's Disease has been extensively studied both in vivo and in vitro. Similarly, there has been an effort to thoroughly analyse the genetics of AD. Yet this is often capped at correlative biomarkers such as PSEN1/2, without the linkage to the underlying biological processes. Yet, treatment of AD is severely limited and only provide symptomatic relief. The combination of genetic and environmental factors - be that socioeconomic, exposure to pollution or a stressful lifestyle - contributes towards the pathogenesis of AD. The constant increase in patients necessitates research to better the understanding of the backbone of Alzheimer's as a prerequisite to discovering novel treatment methodologies.

How to Cite:

Wenqi Zhao, "In Silico Analysis of Mutations Along the Amyloidogenic Pathway in Alzheimer's Disease" *AIJR Preprints*, 540, Version 1, 2024.

There are two different pathways for the cleavage of Amyloid Precursor Protein (APP): amyloidogenic and non-amyloidogenic (Figure 1). Taking the latter first, we have the cleavage of APP by α-secretase, producing sAPPα (Aβ42), which is soluble and released into the intracellular space. This is followed by γ-secretase cleavage that forms a P3 protein, whose significance is barely known, and AICD, the intracellular component. On the other hand, the amyloidogenic pathway involves the sequential cleavage by β [8] and γ-secretase. β-secretase produces sAPPβ, a shorter variation of the product of the non-amyloidogenic pathway, Aβ40[2]. Therefore, the activity of β-secretase (BACE-1) sets off the amyloidogenic pathway and is thus paramount in the onset of Alzheimer's, in particular, Early Onset AD.

Secretases are mostly functional upon attaining the quaternary structure which is tetrameric[13], involving subunits coded for by presenilin 1 and presenilin 2 genes – whereby the mutations in these genes are associated with EOAD. It is believed that the driving factor behind the propensity of aggregation is the solubility of the Aβ fractions. In the monomeric stage, Aβ42 is more prone to aggregation than Aβ40, as it is less soluble[5]. Modifications to Aβ as a monomer impact its solubility; for instance, glycation, the addition of carbohydrates to the end chain of peptides, can affect the N-terminal structure of the protein.

Neprilysin (NEP) is a metalloprotease, involving a singular zinc atom associated with the HExxH (His-Glu-x-x-His) binding motif which is critical for the cleavage action of amyloid-beta. It is 749 amino acid residues long and transmembrane, with the extracellular domain extending from Thr49-Trp749[15]. Specifically, they cleave Aβ at the residues Glu3-Phe4, Gly9-Trp10, Phe19-Phe20, Ala30-Ile31, and Gly33-Leu34[11]. It has been reported previously that it is one of the crucial enzymes in cleaving the APP protein, thus showing a mitigatory effect on the propensity of APP to aggregate and form toxic oligomers and fibrils. There is sound evidence that NEP is the rate-limiting enzyme for Aβ degradation, and the inhibition of which, using thiorphan, increases Aβ levels.

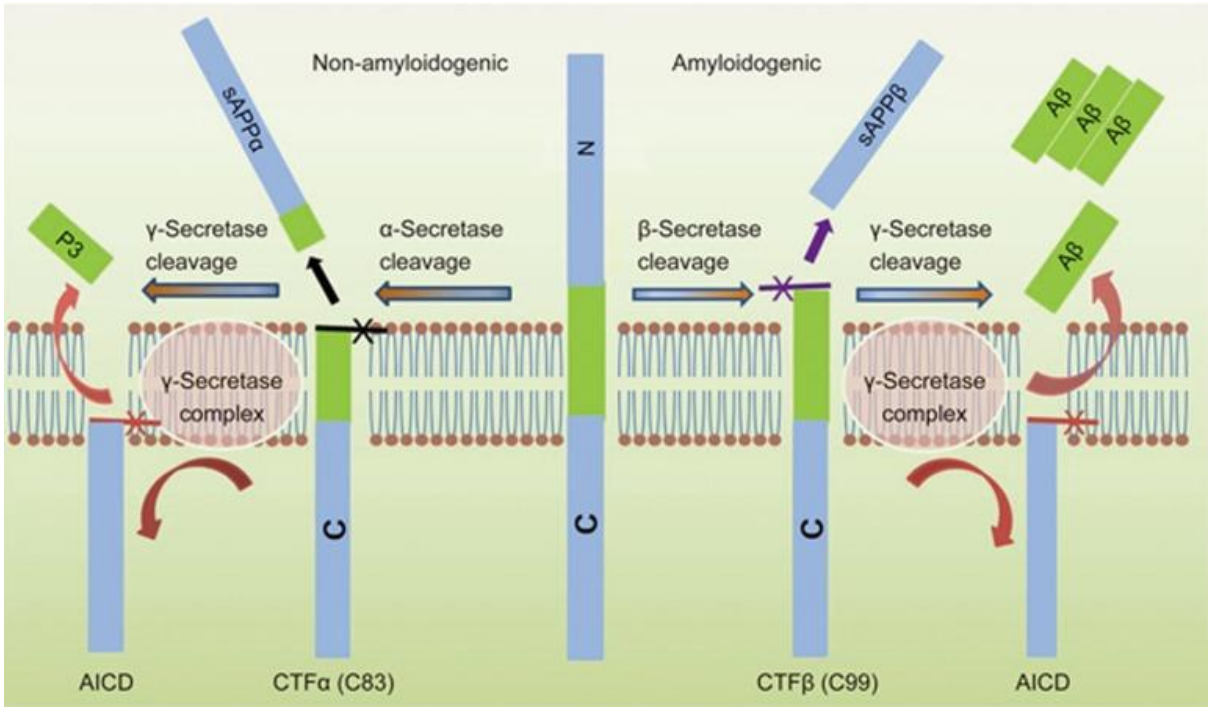


Figure 1: Illustration of amyloidogenic and non-amyloidogenic pathway

2 Methods

2.1 Production of Aβ

In this section, we examine the significance of mutations on miRNA fragments which prevent their ability to regulate BACE1 expression. We found significant differences in protein production and post-transcriptional regulation as a result of mutations on the complementary sequence in particular. The production of Aβ is hinged upon the dominance of the amyloidogenic pathway over the non-amyloidogenic counterpart. It is therefore the increase in β-secretase activity (BACE-1) that directly related with greater Aβ production. MicroRNAs (miRNA) have a crucial role in the post-transcriptional silencing and regulation of protein synthesis. There is a many-to-many mapping relationship between miRNA's and the mRNAs being regulated. Oftentimes, there are 6-8 base pairs which complement with bases in the 3' (or less commonly 5') untranslated region (UTR) of the mRNA. There are 3 major steps within translation as we shall model, for which miRNA have the potency to undermine. Indeed, even Single Nucleotide Polymorphisms (SNPs) [3] can severely impact the ability for miRNA to carry out its function of, in this case, regulating BACE1 expression and mitigating the amyloidogenic pathway.

2.1.1 Kinetic Model of Protein Production

The first step, with constant rate k_1, involves the 40S ribosomal subunit binding to the mRNA. This is followed by the recognition of the AUG motif by the subunit (complementary to the initiator tRNA in ribosomes). The rate of this is modelled as k_2. Finally, we have the recruitment of the final initiation factor, the 60S subunit, after which translation initiation is complete, with reaction rate k_3. The initial kinetic model is as follows:

$$\frac{dA}{dt} = -k_1 A$$

$$\frac{dB}{dt} = -k_2 B + k_1 A$$

$$\frac{dC}{dt} = -k_3 C + k_2 B$$

$$\frac{dD}{dt} = k_3 C$$

A is a protein in stage 1, B in stage 2, C in stage 3 and D ready for elongation. With rates of reaction k_1, k_2 and k_3 as 0.5, 0.2 and 0.2 [17], the expected reaction is shown in Figure 4. Previous computational models [17] validate and suggest that miRNA represses translation predominantly on the k3 reaction rate. This is because it is largely repressing translation at a late step, independent of the presence of a 5' cap. Furthermore, as mentioned, mRNA carries out deadenylation of the Poly(A) tail and protein degradation, thereby undercutting the maximum yield of protein for the elongation process in addition to slowing down the rate.

Where A is a protein in stage 1, B in stage 2, C in stage 3 and D ready for elongation. With rates of reaction k_1, k_2 and k_3 as 0.5, 0.2 and 0.2 [17], the expected reaction is depicted in Figure 4. Previous computational models [17] validate and suggest that miRNA represses translation predominantly during the recruitment of the 60s ribosomal subunit modelled by the reaction rate k_3. This is because the miRNA largely represses translation at a later step, independent of the presence of a 5' cap on the mRNA. Furthermore, the mRNA carries out deadenylation of the Poly(A) tail and protein degradation, thereby reducing the maximum yield of protein for the elongation process in addition to slowing down the rate.

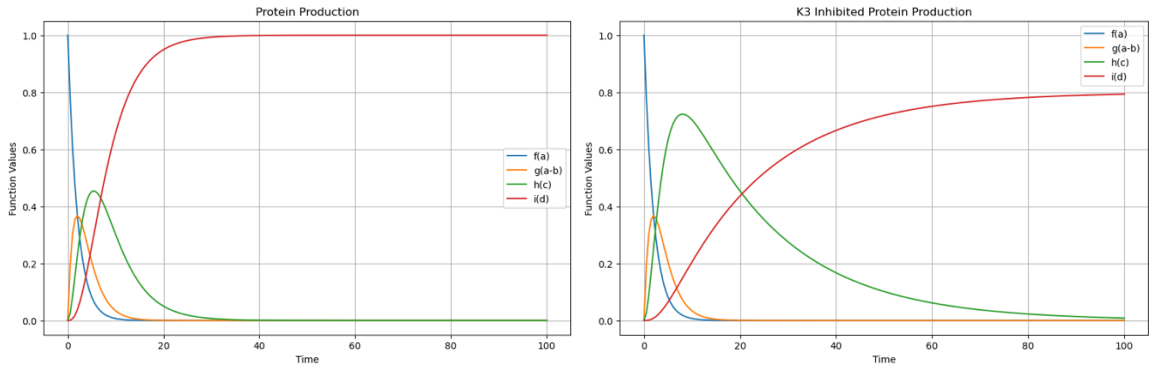


Figure 2: Expected rate of protein initiation in the ribosome in the absence of mRNA silencing vs the actual rate of protein initiation with mRNA silencing affecting the recruitment of the 60S subunit

2.2 Cleavage of Aβ with Neprilysin

2.2.1 Active Site

Based on the simple "induced fit" model of enzyme function, the significance of a mutation's impact on an enzyme's binding affinity and thus its functional capacity is heavily related to the proximity of the conformational change to the HExxH motif at the heart of the active site. Therefore, as the first step, we create a simple mathematical abstraction for the significance of a mutation.

$$f(x_1, x_0) = D(x_1, H) \times \delta(x_1, x_0) \times P(x_1, x_0)$$

Where the function $f(x_1, x_0)$ calculates the significance of the mutation, given parameters (x_1, x_0), the mutant amino acid the wild-type amino acid, respectively.

2.2.2 Distance to Histidine Amino Acid

$$D(x_1, H)$$

Given that the sequence of paramount importance in metalloprotease neprilysin is the HExxH motif, mutations closer to this active site are more likely to impact its ability in cleaving Aβ. Therefore, we calculate the distance in ˚Awithin the Protein Data Bank (PDB) file between the target (x_1) and H residue. We used string searching for the HExxH binding motif, finding the first Histidine residue at the index 521. The visualisation is in Figure 7. Therefore, we examine the influence of mutations on APP that influences factors affecting their aggregation: alignment

similarity, net charge, solvent accessibility and β-sheet. We find significant effects that mutations can have on this process.

2.2.3 Amino Acid Comparison

$$\delta(x_1, x_0)$$

The second metric we have in analysing the effect of mutant amino acids on the overall structure, and hence function, of neprilysin is the difference between the mutant and wild- type AAs on account of several metrics. First, we obtained hydrophobicity values from the Wimley-White scale [20]. This is significant, since hydrophobic side-chains oftentimes force these particular amino acids into the interior of the molecule, where the hydrophilic ones cluster on the outside, as it is favourable for them to form hydrogen bonds with water and polar molecules. Indeed, vast conformational changes have arisen from even single AA polymorphisms, such as sickle cell anaemia where glutamic acid is substituted by valine. The latter is far more hydrophobic and hence reduces the overall solubility of haemoglobin, causing disease. Second, we gather the volume of the amino acids. Importantly, this influences the bending and folding of amino acid chains. Specifically, amino acids which are too large disrupt the positions of the groups that form bonds, be that disulfide bridges, hydrophobic or ionic bonds, with amino acids down the polypeptide chain as it folds. Third, we measure the isoelectric point (pI value) of the amino acid. This is because their charge at a given pH influences solubility and binding patterns. After gathering numerical values, we normalised them to the range −1 to 1, with the maximum and minimum being at each end.

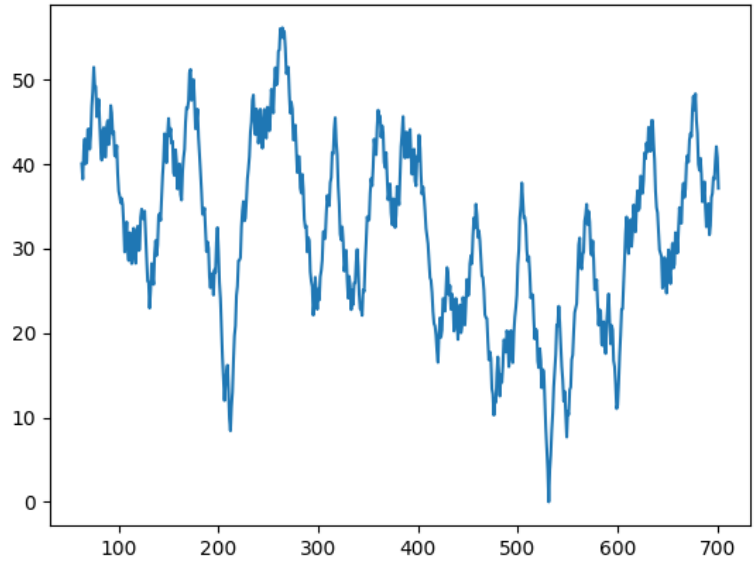


Figure 3: Distance between residues within Neprilysin and the Histidine Atom on residue 521 - where the active site is located

$$f(x) = \frac{2(1 - min(x))}{max(x) - min(x)}$$

With that for each column, we plotted a 3-dimensional graph of each amino acid (Figure 4). To calculate the potential for structural difference, we calculate the Euclidean distance between the wild-type and mutant amino acids.

2.2.4 Mutation Rate

Finally, we need to calculate the Probability that a given mutation does occur (since the simulation for mutations is done at a random rate, we need to reverse engineer the process). This is estimated based on two parameters. First, the number of nucleotide substitutions that needs to occur. That is, a mutation from Phenylalanine (coded for by UUU) to Leucine (coded for by UUA) is much more likely to happen than from Phenylalanine to Alanine (coded for by GCU). Second, there is an observed and intuitive difference between purine or pyrimidine substitutions: from A to G and vice versa, or from C to U and vice versa (deemed a transition mutation), and purine to pyrimidine substitutions (called transversions) [14]. This is simply because of the larger structural difference as purines have 2 nitrogenous rings while pyrimidine only have 1. Indeed, it is observed that transitions are 1.697 times more frequent than transversions. Thus, a simple formula for calculating the ease of mutation is done through

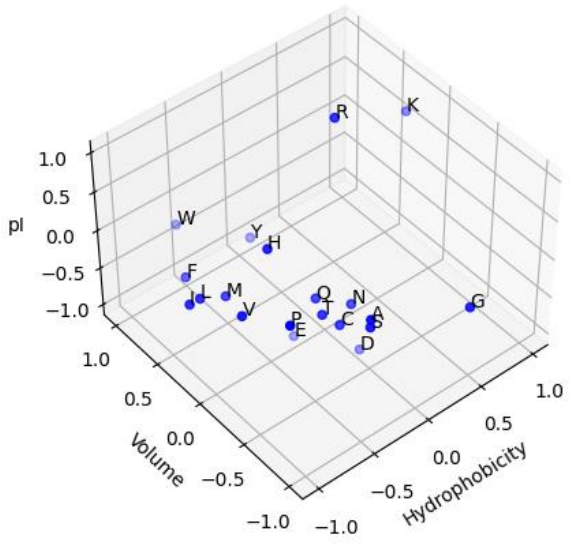


Figure 4: 3D visualisation of similarities of Amino Acids

$$M = (M_{tv} + 1.697 \times M_{ts})$$

Where M = number of mutations, M_{tv} = number of transversions and M_{ts} = number of transitions. The visualisation of the ease of mutation is in Figure 5.

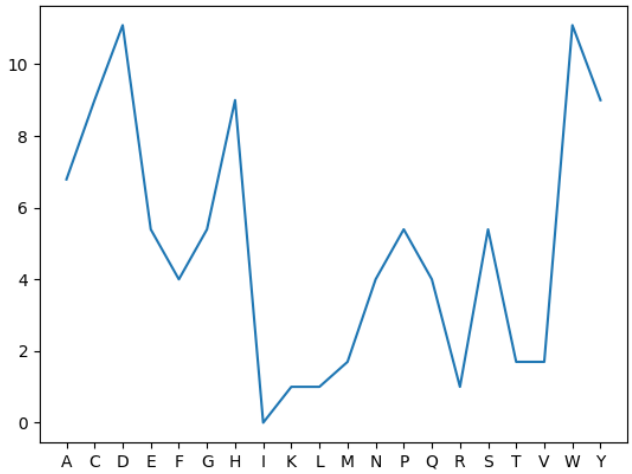


Figure 5: With reference amino acid as Isoleucine (I), the ease of mutation between amino acids based on the number and type of nucleotide substitutions needed in the base code.

2.3 Aggregation of Aβ

Mutations are a rare, random change in DNA, that alters the nucleotide sequence of a gene, which in turn translates to changes further down the transcriptional and translational process. Theoretically, a singular mutation has the potential to create catastrophic changes. For example, sickle cell anaemia occurs when glutamic acid, a polar amino acid at the surface of the β−chain, is replaced by valine, which is non-polar. This makes the hydrophilic outside of the globular haemoglobin much less soluble, resulting in clots and other severe symptoms.

2.3.1 Data Acquisition and Processing

1. We obtain files of the desired genes in FASTA format, from the National Centre for Biotechnology Information (NCBI) [16]
2. Using the Biopython Library [4], we perform transcription of the DNA into mRNA, based on the A-U, C-G, T-A complementary base-pair rules.
3. We create functions that simulate the three major mutations: substitution (the changing of one nucleotide to the other), insertion and deletion. The function for substitution takes in an mRNA sequence and a mutation rate as parameters. The function for insertion takes in the base mRNA sequence, the number of insertions as well as the maximum length of inserted nucleotide sequence. The function for deletion takes in the base mRNA sequence, as well as the maximum length and number of deletions.
4. A selection of random mutations are simulated on the mRNA, with differing magnitudes in rate. The resulting sequence is translated into a peptide chain based on the genetic code for homo sapiens.
5. We output the resulting protein sequence to Robetta, software developed by the University of Washington that uses comparative modelling to infer a protein's tertiary structure (in a Protein Data Bank (PDB) file format) from its amino acid sequence.

2.3.2 Comparing Structural Similarities

In order to interpret the effect mutations at the nucleic level has on the macro-molecular structure and function of proteins, we compare these similarities of mutated proteins based on several simplifying assumptions

1. The most basic assumption is known as Anfinsen's dogma, where for small globular proteins, the secondary structure is almost entirely dependent on the primary amino acid structure[6]. Therefore, we employ a modified BLAST function that compares the percentage alignment of two sequences. First, we align and superimpose the sequences, and then calculate the number of amino acids at each position which are the same.
2. To examine the differential likelihood of aggregation, we calculate the Half Sphere Exposure of proteins. This measures the solubility of the protein through the number of Cα atoms within a given radius of a Cα residue (split into a top and bottom half). This explains the accessibility of solvent molecules, such as water, which impacts its solubility [7].
3. Furthermore, we calculate the net charge of a protein using the Henderson-Hasselbach Equation [19]. For proteins Histidine (H), Arginine (R) and Lysine (K) which can become protonated to form NH3+, the concentration of which that would be positively charged at a given pH (HA) is

$$pH = pKa + \log \frac{1 - HA}{HA}$$

   For acidic proteins Aspartic Acid (D) and Glutamic Acid (E), the proportion that is negatively charged is calculated by the same formula, except we are interested in 1−HA. Thus, the function for net charge calculates the difference between the positive charge, a sum of the number of H, R, K multiplied by their respective protonated rates, and the negative charge, a sum of the number of D, E multiplied by their respective deprotonated rates.

4. Finally, we examine the concentration of β-pleated sheets. These are heavily associated with a higher likelihood of protein aggregation [10] since the edge sheets form H bonds with neighbouring edge sheets easily. Therefore, a higher concentration of β−sheet structures is an important factor in the likelihood of aggregation. Here, we use the DSSP (Dictionary of Secondary Structure of Proteins) library to analyse the percentage of this secondary structure.

3 Results and Discussion

3.1 The role of miRNA mutations in the Production of Aβ

We first simulated mutations in the miRNA (mi-R29b-1 [9]) sequence. For both the mutated and original miRNA sequence, we calculate the maximum complementary alignment with the BACE1 enzyme - where the miRNA would bind to in order to inhibit mRNA expression. The percentage difference in the alignment sequences were then compared. This investigated the potency of mutations to disrupt the site of mRNA inhibition. This difference, visualised in Figure 5, will therefore impact miRNA's ability to decrease k3 and repress BACE1 production.

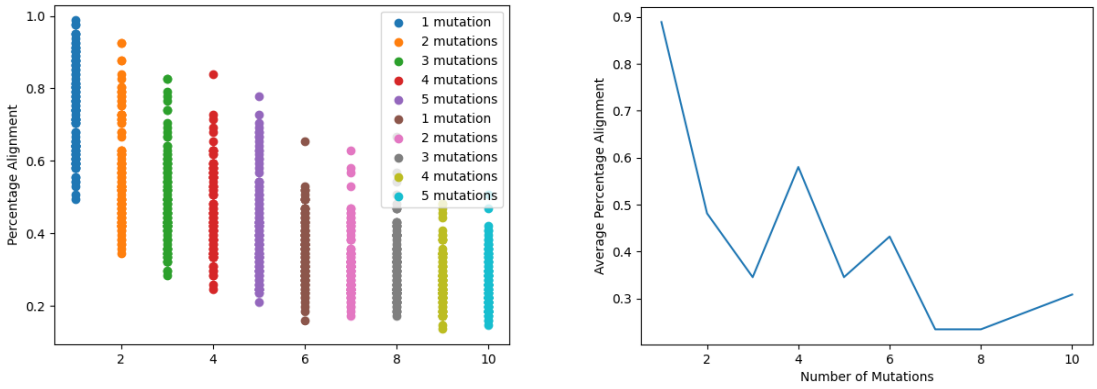


Figure 6: Right Graph: Scatter plot of percentage alignment (x axis) versus number of mutations (y axis). Left Graph: Line graph of median percentage alignment (x axis) versus number of mutations (y axis).

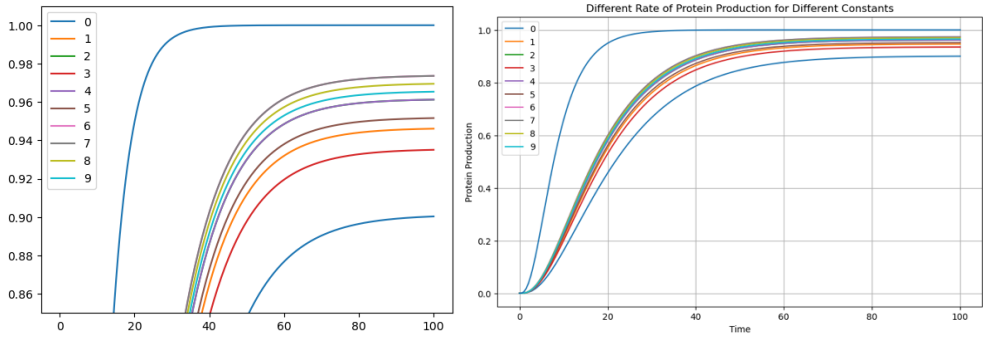


Figure 7: Effect of miRNA mutation on the rate of protein production. The figure is zoomed in on the right for more clarity

Therefore, we can suggest based on our observations that the binding site for miRNA is impacted significantly as the number of mutations increase. Based on that, we model the kinetics of protein production with updated constants. That is, the percentage 'un- alignment' (1 - percentage alignment) is equal to the percentage decrease in k3 and k4. Thus, a lower percentage accuracy because of mutations will lead to less of an ability to reduce the constants in the model.

Ultimately, we can see a significant degree of impact of miRNA mutations on its potency to regulate protein expression in Figure 6. All the mutated forms of miRNA leads to a greater alignment with the original kinetic model, with more mutations broadly leading to more similarity, although he statistical significance of that is reasonably low.

3.2 Significant mutations for Neprilysin Cleavage

We normalised the values to between 0 and 1, where any value closer to 1 is equivalent to a more impacting mutation (smaller distance to H residue, larger Euclidean distance on AA visualisation, easier mutation). We generated a set of 100 mutations at random indices with random final amino acids. Out of this list, the most substantial mutations, which align with our expectations from the start, are illustrated in. They are:

Index and AA	Distance to H	AA similarity	Ease of Mutation	Final Value
547 D → W	0.975	0.792	**0.917**	**0.708**
419 K → L	0.879	0.978	**0.800**	**0.688**
497 R → E	0.786	1.000	**0.872**	**0.685**
511 F → K	0.634	0.965	**0.900**	**0.551**
131 A → K	0.767	0.808	**0.844**	**0.522**

We then simulated each of those mutations using Robetta, a protein structure prediction algorithm, and indeed found significant (for the expectation of SNPs) conformational differences to the original Neprilysin PDB file, justifying our conclusion.

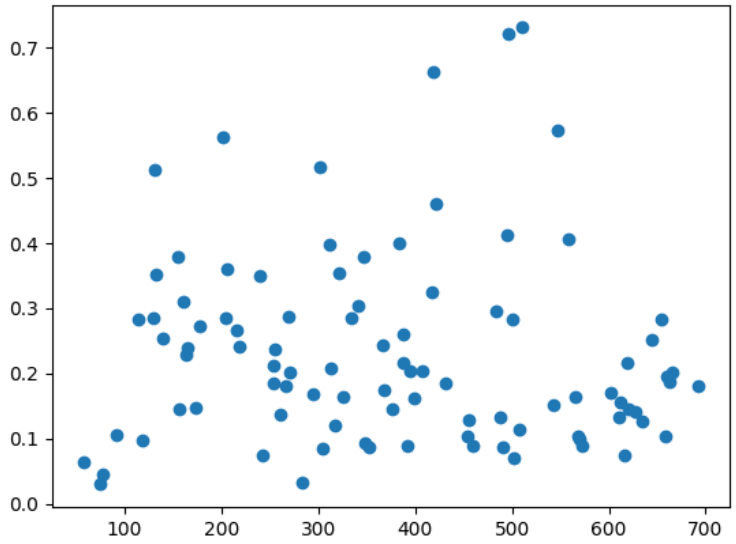


Figure 8: Significance of Final Mutations

3.3 Effect of mutations on aggregation of Aβ

3.3.1 Mutation vs Alignment Similarity

From the above graphs we see a strong negative correlation between the rate of mutation (calculated as the rate of substitution, number of insertions and number of deletions), supporting the claim that mutations are a significant cause in disrupting protein structure and causing aggregation. Furthermore, the production of short peptide chains can bind and aggregate with each other. However, there is a strong degree of unpredictability. First, mutations in the DNA which change the reading frame are likely to generate far more structural disparity. Hence, we spikes in the graphs which are similar to the original APP protein, because the mutation only changed singular amino acids instead of the structure as a whole. Furthermore, mutations which introduce premature stop codons (UGA, UAA, UAG) significantly alter the structural conformation of the proteins.

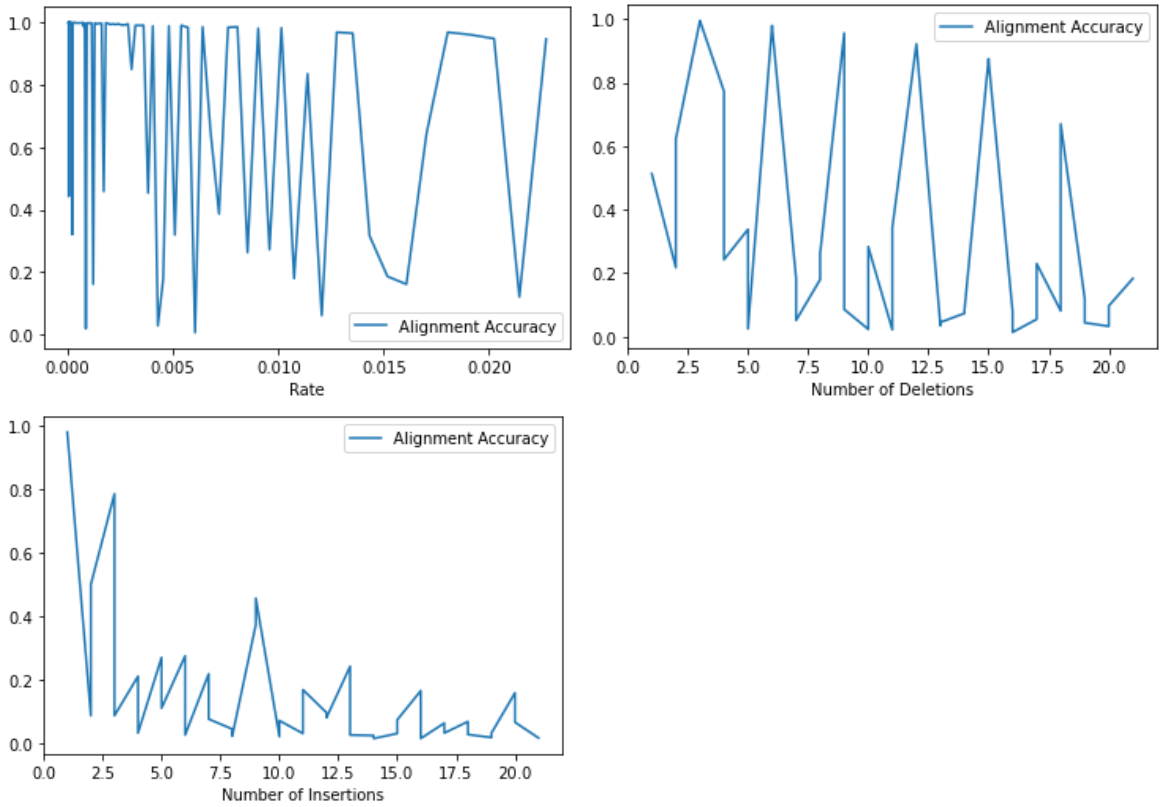


Figure 9: Simulation of mutation rate vs alignment accuracy

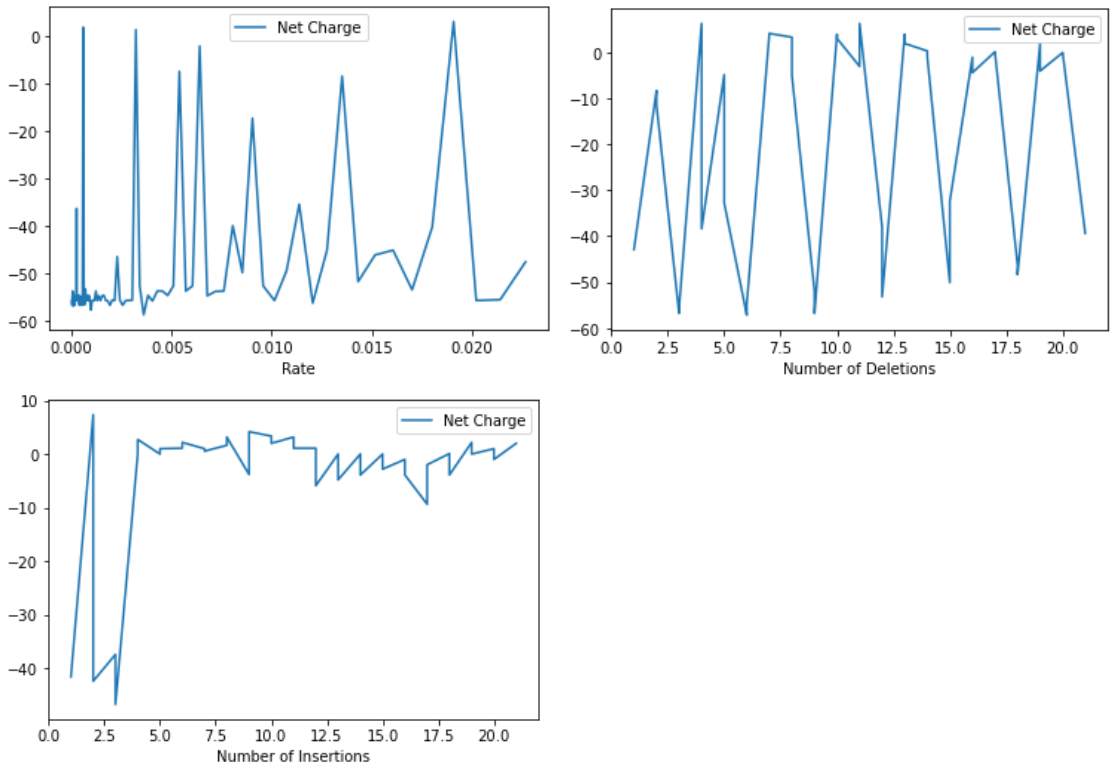


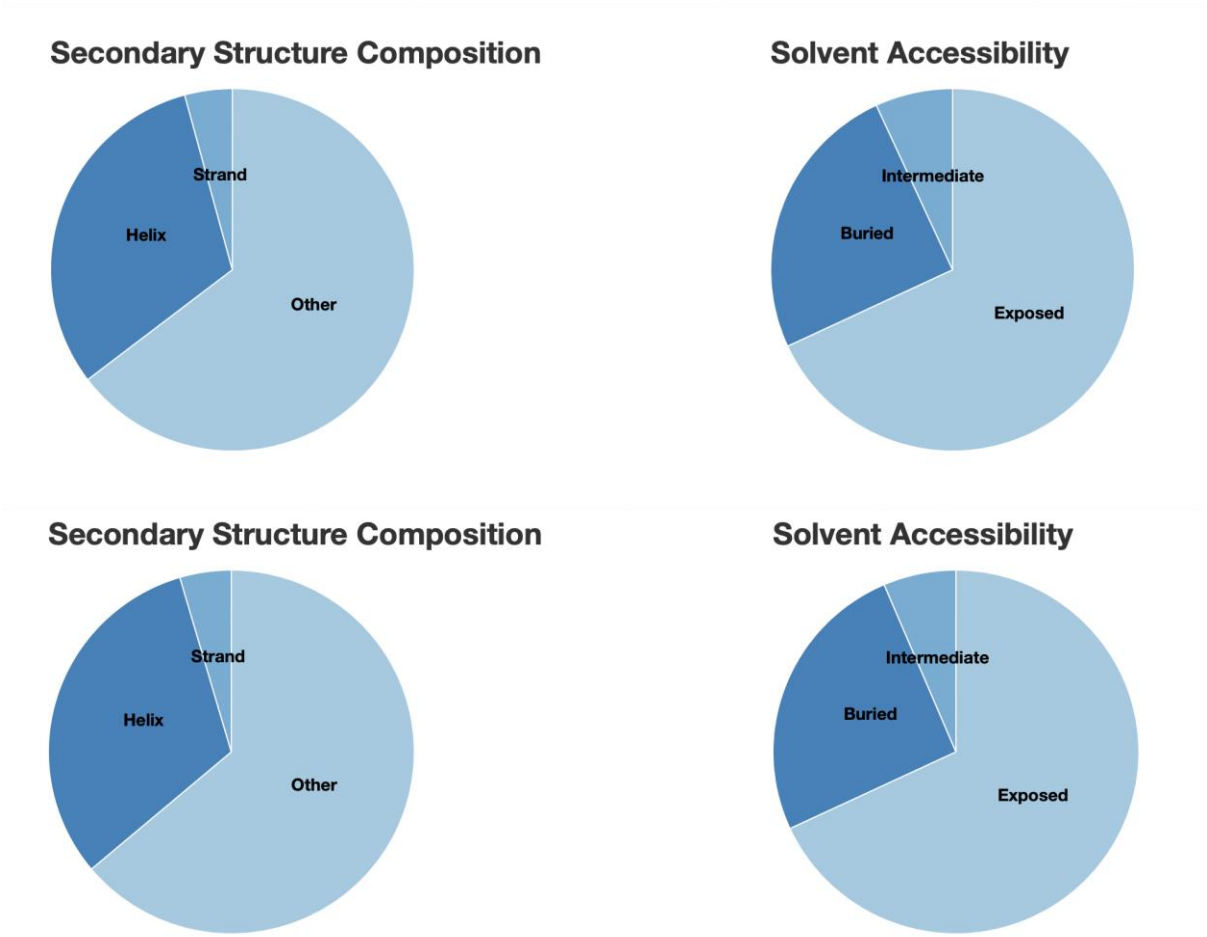
Figure 10: Simulation of mutations rate against net charge

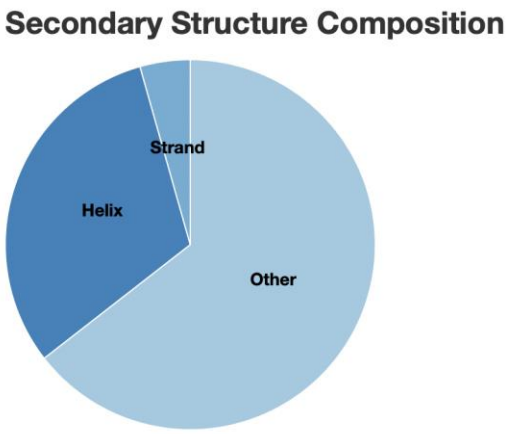
3.3.2 Mutation vs Net Charge

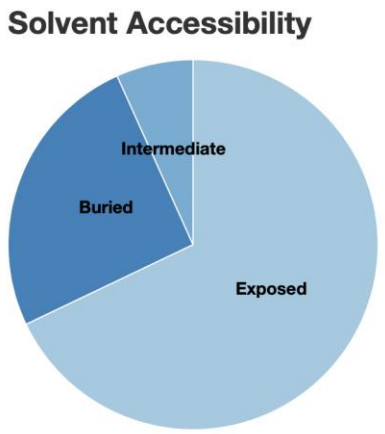
A low net charge leads to a decrease in solubility (for example, ionic compounds are far more soluble as the charged components can interact with the polar water molecules) [12]. Therefore, a small net charge is characteristic of a protein's likelihood of aggregation. From the simulation, we see differing results. When the number of insertions increase, the net charge quickly concentrates to being very close to 0. For the number of deletions, we see the same pattern of dips that occur every time a frame shift does not happen. For substitution mutations, the net charge spikes upwards to 0 every so often, without a clear pattern. Therefore, although a huge degree of uncertainty and fortuity exists, there is a significant likelihood mutations of all kinds (they do not even need to happen at a large rate, as indicated by the cluster of spikes in the first graph), can lead to the production of clusters of protein fragments that are insoluble, which clump and aggregate.

3.3.3 Mutation vs Solvent Accessibility and β-sheets

Here, it is very unfair and biologically insignificant to simulate mutations to the same extent as before; therefore, we minimise the quantity of mutations from 1 to 10. Using prediction software Predict Protein [18], predictions based on the mutated amino acid sequence were made. However, given the small mutation size, no severe differences or effects were observed. The images are the secondary structure composition (α-helices, β-strands and other) and solvent accessibility of sequences with 1, 2 and 10 mutations, respectively.







4 Conclusion

4.1 Significance

Mutations along the amyloidogenic pathway are important on the multitude of processes that affect Alzheimer's disease. Our research not only validates this, but, more importantly, provide frameworks of pre-emptive risk assessments. Patients' genetic data can be analysed by the formulae derived in this paper to provide insight onto the likelihood of pathogenesis. Furthermore, the genetic data of entire demographics can be screened for the mutations we've outlined to be the most harmful.

4.2 Limitations and Further Development

The novelty of mutations means that the work here is theoretical - there simply are not enough people with or without such mutations to experimentally verify the level of Aβ biomarkers in their brain because of these mutations. Furthermore, we made many simpli- fying assumptions. In the examination of miRNA and BACE1, we assumed that the impact on protein production is directly proportional to the number of mutations within the com- plementary strand: this may overlook nuanced scenarios (consecutive mutations, types of mutations etc.) that could have differing effects. In the second section, there were perhaps more metrics we could have examined to model the importance of amino acid mutations more accurately. Furthermore, perhaps different weights on the three metrics (which are currently assumed to be equally important) would be more representative.

5 Declarations

5.1 Availability of data and material

The full code is available upon reasonable request to the author(s).

5.2 Acknowledgements

With sincerest gratitude to Samuel Dada, Rahul Arora and James Tomkins from Cambridge University in their critical review of the manuscript, and Professor Michele Vendruscolo for the direction and guidance throughout.

5.3 Competing Interests

The author declares no competing interests.

References

[1] Alzheimer's Association. Alzheimer's Disease Facts and Figures. url: https://www.alz.org/alzheimers-dementia/facts-figures
[2] Marco Calabr`o et al. "The biological pathways of Alzheimer disease: A review". In: AIMS neuroscience 8.1 (2021), p. 86.
[3] Lena J Chin et al. "A SNP in a let-7 microRNA complementary site in the KRAS 3 untranslated region increases non–small cell lung cancer risk". In: Cancer research 68.20 (2008), pp. 8535–8540.
[4] Peter J. A. Cock et al. "Biopython: freely available Python tools for computational molecular biology and bioinformatics". In: Bioinformatics 25.11 (Mar. 2009), pp. 1422–1423. issn: 1367-4803. doi: 10.1093/bioinformatics/btp163
[5] Audrey Gabelle et al. "Correlations between soluble α/β forms of amyloid precursor protein and Aβ38, 40, and 42 in human cerebrospinal fluid". In: Brain research 1357 (2010), pp. 175–183.
[6] Giorgia Gambardella et al. "The anfinsen dogma: Intriguing details sixty-five years later". In: Inter- national journal of molecular sciences 23.14 (2022), p. 7759.
[7] Thomas Hamelryck. "An amino acid has two sides: a new 2D measure provides a different view of solvent exposure". In: Proteins: Structure, Function, and Bioinformatics 59.1 (2005), pp. 38–48.
[8] Harald Hampel et al. "The β-secretase BACE1 in Alzheimer's disease". In: Biological psychiatry 89.8 (2021), pp. 745–756.
[9] S´ebastien S H´ebert et al. "Loss of microRNA cluster miR-29a/b-1 in sporadic Alzheimer's disease correlates with increased BACE1/β-secretase expression". In: Proceedings of the National Academy of Sciences 105.17 (2008), pp. 6415–6420.
[10] Joelle AJ Housmans et al. "A guide to studying protein aggregation". In: The FEBS Journal 290.3 (2023), pp. 554–583.
[11] Steven Howell, Josephine Nalbantoglu, and Philippe Crine. "Neutral endopeptidase can hydrolyze - amyloid(1–40) but shows no effect on -amyloid precursor protein metabolism". In: Peptides 16.4 (1995), pp. 647–652. doi: 10.1016/0196-9781(95)00021-b.
[12] Ryan M Kramer et al. "Toward a molecular understanding of protein solubility: increased negative surface charge correlates with increased solubility". en. In: Biophys. J. 102.8 (Apr. 2012), pp. 1907– 1915.
[13] Ji Young Lee et al. "Allosteric modulation of intact γ-secretase structural dynamics". In: Biophysical journal 113.12 (2017), pp. 2634–2649.
[14] Michael Lynch. "Rate, molecular spectrum, and consequences of human mutation". In: Proceedings of the National Academy of Sciences 107.3 (2010), pp. 961–968.
[15] NN Nalivaeva, IA Zhuravin, and AJ Turner. "NEPRILYSIN expression and functions in development, ageing and disease". In: Mechanisms of Ageing and Development 192 (Dec. 2020), p. 111363. doi: 10.1016/j.mad.2020.111363.
[16] National Center for Biotechnology Information (NCBI). url: https : / / www . ncbi . nlm . nih . gov / gene?Cmd=DetailsSearch&Term=351.
[17] Tracy Nissan and Roy Parker. "Computational analysis of miRNA-mediated repression of translation: implications for models of translation initiation inhibition". In: Rna 14.8 (2008), pp. 1480–1491.
[18] "PredictProtein - Predicting Protein Structure and Function for 29 Years". In: Nucleic Acids Research 49.W1 (May 2021), W535–W540. issn: 0305-1048. doi: https://doi.org/10.1093/nar/gkab354
[19] Ronald J. Tallarida and Rodney B. Murray. "Henderson—Hasselbalch Equation". In: Manual of Phar- macologic Calculations: With Computer Programs. New York, NY: Springer New York, 1987, pp. 74– 75. isbn: 978-1-4612-4974-0. doi: https://doi.org/10.1007/978-1-4612-4974-0_24
[20] William C Wimley and Stephen H White. "Experimentally determined hydrophobicity scale for pro- teins at membrane interfaces". In: Nature structural biology 3.10 (1996), pp. 842–848.