Computing the Pathogenicity of Alzheimer’s Disease Presenilin 1 Mutations

Tang, Ning; Dehury, Budheswar; Kepp, Kasper Planeta

Published in:
Journal of Chemical Information and Modeling

Link to article, DOI:
10.1021/acs.jcim.8b00896

Publication date:
2019

Document Version
Peer reviewed version

Citation (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Computing the Pathogenicity of Alzheimer’s Disease Presenilin 1 Mutations

Ning Tang, BUDHESWAR DEHURY, and Kasper P. Kepp

J. Chem. Inf. Model., Just Accepted Manuscript • DOI: 10.1021/acs.jcim.8b00896 • Publication Date (Web): 25 Jan 2019

Downloaded from http://pubs.acs.org on January 28, 2019

Just Accepted

“Just Accepted” manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides “Just Accepted” as a service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. “Just Accepted” manuscripts appear in full in PDF format accompanied by an HTML abstract. “Just Accepted” manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are citable by the Digital Object Identifier (DOI®). “Just Accepted” is an optional service offered to authors. Therefore, the “Just Accepted” Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the “Just Accepted” Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these “Just Accepted” manuscripts.
Computing the Pathogenicity of Alzheimer’s Disease Presenilin 1 Mutations

Ning Tang†, Budheswar Dehury†, and Kasper P. Kepp‡*

†Department of Chemistry, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark.

* KPK: Phone: +045 45252409  E-mail: kpj@kemi.dtu.dk
Abstract

Alzheimer’s disease (AD) is one of the major global health challenges of the 21st Century. More than 200 distinct mutations in presenilin 1 (PSEN1) cause severe early-onset familial AD (FAD), and are thus of central interest to the etiology of AD. PSEN1 is the catalytic subunit of γ-secretase which produces Aβ, and the mutations tend to increase the produced Aβ_{42}/Aβ_{40} ratio. The molecular reasons for the pathogenesis of these mutations are unknown. We studied a close-to-complete dataset of PSEN1 mutations using 21 different computational methods hypothesized to reproduce pathogenesis, using both sequence- and structure-based methods with the full γ-secretase complex as input. First, we tested if pathogenicity can be estimated accurately using all possible mutations in PSEN1 as a direct control. Several methods predict the pathogenicity of the mutations (pathogenic vs. all other possible mutations) well, with accuracies approaching 90%. We then designed a stricter test for predicting the severity of the mutations estimated by the average clinical age of symptom onset for mutation carriers. Surprisingly, we can predict clinical age of symptom onset at 95% confidence or higher with several methods. Accordingly, our results show that simple biochemical properties of the amino acid changes rationalize an important part of the pathogenicity of FAD-causing PSEN1 mutations. Although pathogenic mutations generally destabilize γ-secretase, all tested protein stability methods failed to predict pathogenicity. Thus, either the static cryo-electron microscopy-derived molecular dynamics equilibrated structures used as input fail to capture the stability effect of mutated side chains, or protein stability is simply not a key factor in the pathogenicity. Our findings suggest that the chemical causes of FAD may be modelled and lend promise to the development of a semi-quantitative model predicting the age of onset of mutation carriers that could eventually become of care-strategic value.
Introduction

Alzheimer’s disease (AD) is the most common form of dementia (60−70% of all cases) and afflicts tens of millions of people worldwide.1,2 This chronic neurodegenerative disease is characterized by progressive cognitive and behavioral impairment leading eventually to loss of autonomy and death. The suffering of AD patients and close relatives is accompanied by a major burden on society health budgets due to the enormous cost of AD-related medical care.3 Currently, there is no effective therapy; available treatments only temporarily relieve symptoms.4 Accordingly, there is an urgent need to understand the biochemical origins of AD in order to develop new therapies.

AD comes in two types: Early-onset AD with clinical symptoms typically occurring between 30 and 65 years, and late-onset sporadic AD where clinical symptoms usually occur at +65 years. Although early-onset AD is a small minority it constitutes the main basis for research as disease can in some cases be linked to a mutation in a protein.3 In particular, some early-onset familial AD (FAD) is caused by inherited autosomal dominant gene mutations that are central to understanding the pathogenesis of AD.5,6 The genes encode for the amyloid-β precursor protein (APP),7,8 presenilin 1 (PSEN1)9–11 and presenilin 2 (PSEN2).12 Among these, mutations in PSEN1 are by far the most frequent.13 PSEN1 mutations also cause severe AD with clinical symptoms sometimes arising near 30 years of age, indicating their criticality and associated importance for understanding the pathogenesis of AD.14

PSEN1 is the catalytic subunit of the intramembranous aspartyl protease γ-secretase that cleaves C99 to produce the infamous β-amyloid peptide (Aβ) of variable length. FAD-causing mutations in PSEN1 commonly affect the cleavage by γ-secretase, leading to a change in the amount of each isoform of Aβ, typically increasing the Aβ42/Aβ40 ratio.15,16 To date, more than 200 PSEN1 FAD mutations have been identified.14

Despite extensive research, the pathogenic mechanism of PSEN1 mutations remains unknown. A plausible suggestion is that the mutations change the Aβ isoform distribution to produce
relatively more (but importantly not absolutely more) aggregation-prone toxic Aβ42 as argued by recent versions of the amyloid hypothesis emphasizing the Aβ42/Aβ40 ratio. It could also be caused by a loss of PSEN1 function, e.g. in relation to Notch cleavage, as this important substrate’s cleavage is likely to be also affected by the mutations. It was previously shown that the Aβ42/Aβ40 ratio measured in assays for a PSEN1 mutation (but not for APP mutations) correlates with the clinical severity of the mutation estimated from the reported average age of clinical symptom onset of the mutation carriers. A recent extensive assaying suggested that this correlation is weak. Re-analysis of these important new data, however, suggests that the relationship remains very strong, providing a firm and central basis for relating chemical PSEN1 modifications directly to AD pathogenesis.

We asked the question how a single change in an amino acid in this protein can be enough to cause disease. At the same time we want to know if the clinical age of AD onset can be estimated using basic chemical properties of the amino acid changes rather than the empirically observed effect on C99 processing. To address this, we studied the state-of-the-art computational tools argued to provide fast and efficient analysis of potential functional effects of missense mutations, and applied them to the missense PSEN1 mutations causing FAD, which form a cornerstone of current studies of AD etiology. The change in protein fold stability caused by mutation can potentially reduce the copy number of the functional protein (loss of function) or induce malicious misfolding (gain of function). Pathogenic missense mutations frequently impair native protein interactions and tend to reduce protein folding stability, but other properties such as the affinity of key substrates, catalytic function, solubility, compartmentalization, or signaling and processing of the protein can also be affected. The effect of mutations is examined either by structure-based methods using a protein 3D structure or using the amino acid sequence as input.

As structural input, we used the recently identified full-atom semi-open state of γ-secretase fully consistent with the cryo-electron-microscopy data but additionally including the missing...
residues and loops, equilibrated by molecular dynamics.\textsuperscript{45} We explored whether 21 different computational methods can qualitatively describe the pathogenicity of PSEN1 mutations using as a specifically designed ANOVA test all possible mutations in PSEN1 as a powerful control. This test should be useful to a range of other studies of pathogenic mutations. We also studied the more challenging task of quantitatively describing the pathogenic severity, as estimated by the average age of clinical symptom onset for patients carrying the mutations, as well-behaved data sets are available for this purpose.\textsuperscript{25,28,46} We show that some methods capture fundamental features of the mutations that correlate with the clinical AD data, the first such correlations for a large clinical data set\textsuperscript{46} beyond the empirically known $\text{A}\beta_{42}/\text{A}\beta_{40}$ ratio.\textsuperscript{28} Our findings lend promise to the development of semi-quantitative causal chemical models of FAD with predictive and care-strategic value.

**Computational Methods**

**Data collection.** Information on pathogenic and non-pathogenic PSEN1 mutations was obtained from the two widely used databases, Alzheimer Disease & Frontotemporal Dementia Mutation Database\textsuperscript{14} and the database of alzforum.org (www.alzforum.org/mutations). The insertions and deletions cannot be calculated because the methods generally require the site of mutation to remain intact, and thus the applied final data set contains only the missense single-point mutations in Table S1. In total, 235 known pathogenic mutations and four clinically confirmed non-pathogenic mutations are collected in Table S1. The average age of clinical symptom onset for each type of mutation was compiled from previously published systematic meta-analyses.\textsuperscript{25,28,46} The data set is the most complete for individual data on clinical symptom onset. For the calculations of clinical age of onset, we specifically used the values reported by Ryman et al.\textsuperscript{46}
Figure 1. (A) Alignment of the cryo-electron microscopy structure (PDB code 5A63, 3.4 Å)\(^47\) and our MD-equilibrated all-atom structure of the semi-open state of \(\gamma\)-secretase.\(^45\) (B) The sequence of PSEN1 (uniprot ID P49768, PSN1_HUMAN) used for calculating the mutation scores of the sequence-based methods. Known pathogenic mutations are shown in red.

**Protein stability calculations.** Computational estimates of the stability changes (\(\Delta\Delta G\)) upon mutation are widely used in protein design, but also in the evaluation of disease-causing missense mutations in particular for misfolding-prone proteins.\(^29\) The structure of \(\gamma\)-secretase used for our calculations is the previously established all-atom homology model equilibrated by 500 ns of molecular dynamics (MD).\(^45\) This model closely matches all coordinates of the cryo-electron microscopy structure of \(\gamma\)-secretase (PDB ID: 5A63)\(^47\) and contains the same experimental topology for all the transmembrane helices, with a total RMSD of 0.3 Å. The structural alignment is shown in Figure 1A with sites harboring pathogenic mutations marked in red. Figure 1B shows the corresponding sequence for PSEN1. Our structure contains all parts missing in the cryo-electron microscopy structure except the structurally elusive N-terminal of PSEN1 (70 amino acids). The missing parts were added using template modeling with subsequent 3x 500 ns MD equilibration.\(^45\) Accordingly, our in silico saturation mutagenesis only generates 7543
mutations for the structure-based methods, whereas for the sequence-based methods, all possible mutations also in this N-terminal part were computed. Our structural model represents the catalytically relevant mature auto-cleaved semi-open conformation state of γ-secretase, which we believe is the active state of the protein complex with optimal binding to C99 leading to longer retention time, maximal trimming, and shorter Aβ peptides. This "Fit-Stay-Trim mechanism" (FIST) implies that a fist-like tighter grab by γ-secretase makes C99 stay for longer time enabling C99 to be "squeezed" to smaller Aβ peptides by consecutive proteolytic cleavage.

For the FoldX (Version 4) ΔΔG calculations, the RepairPDB function was applied to the wild-type structure as recommended. The BuildModel function was employed to generate each individual single residue variant, and five independent runs were carried out. The ΔΔG values reported are the differences of the mean values of five independent runs. The typical prediction error of FoldX is ~1 kcal/mol. With Rosetta version 2017.36 (release 59679), the wild-type structure was cleaned and then relaxed using the REF2016 energy function. Then 20 structures were generated, and the structure with the lowest score was used for further ΔΔG calculation. The Cartesian version of the Rosetta ΔΔG application was used with three iterations. The final ΔΔG was calculated from the averaged energies. For DUET using the full γ-secretase as input, seven mutations (A192C, A192E, A192V, E184D, K187R, S104Y and V103W) could not be computed by the server and thus the all-else data set for DUET contains 7536 data points instead of 7543. The potential effect on the total average of < 0.01 kcal/mol will not affect any data or conclusions. For ΔΔG calculations using I-mutant, the secondary structure of the wild-type γ-secretase was first calculated by DSSP and then submitted to I-mutant 2.0 for ΔΔG calculation with default settings. For mCSM-Stability, SDM, DUET, POPMUSIC, HOTMUSIC, mCSM-PPI and EASE-MM ΔΔG calculations, the web servers were used with their default settings. As different methods have different sign conventions for stabilizing and destabilizing mutations, the sign was manually adjusted to our
present convention ($\Delta \Delta G < 0 = \text{stabilizing}$ and $\Delta \Delta G > 0 = \text{destabilizing}$) for all the methods to facilitate comparison.

Since the $\gamma$-secretase complex contains four subunits, we investigated the sensitivity of the computed outcomes to the presence of the additional subunits by also computing all data with the PSEN1 subunit alone. For the pathogenic mutations this gives similar results because most pathogenic mutations occur in PSEN1 parts with no close contact to other subunits, whereas a large part of all possible mutations in PSEN1 have contact with other subunits, and thus, the calculation for this control set is context-dependent. Full ANOVA significance data for all possible combinations of methods and structures are given in Table S2.

**Sequence-based disease prediction methods.** The amino acid sequence of PSEN1 used for the sequence-based methods (Figure 1B) calculations was obtained from UniProt (P49768 PSN1_HUMAN). The functional effect of the mutations was predicted using the individual web servers with default settings: PolyPhen-2,\textsuperscript{62} SIFT,\textsuperscript{63} PMUT,\textsuperscript{64} META-SNP,\textsuperscript{65} PHD-SNP,\textsuperscript{66} SNAP.2,\textsuperscript{67} SNAP,\textsuperscript{65} PANTHER,\textsuperscript{68} PROVEAN,\textsuperscript{69} FATHMM\textsuperscript{70} and SNPMUSIC.\textsuperscript{71} These methods generally use the amino-acid conservation obtained from multiple sequence alignments to predict the functional effect of a missense mutation. In addition, physicochemical properties of the involved amino acids, database annotations, and potential protein structural changes commonly associated with certain mutations are considered.\textsuperscript{29} The prediction scores represent the likelihood of a given mutation causing a deleterious functional change of the protein. A short summary of the studied methods is given in Table 1. In all cases, only the prediction scores were collected and used in the following analysis. The SNAP, SIFT, PANTHER and PHD-SNP score values were obtained from the meta-predictor META-SNP.

**Residue classification.** The relative solvent accessible surface area (RSA) of each residue was calculated using I-mutant 2.0 based on the wild type structure as input. The classification of buried/exposed properties was based on a RSA cutoff of 20%: Residues with RSA < 20% were classified as buried. Exposed residues were further divided into water-exposed and lipid-exposed.
Accordingly, mp_lipid_acc application\textsuperscript{72} from Rosetta was used to calculate the lipid accessibility of the residues. This method uses a concave-hull algorithm to classify lipid accessibility of residues based on the protein structures transformed in membrane coordinates and a defined membrane thickness. The mp_lipid_acc method achieves accuracies, specificities and sensitivities around 90% according to a manually curated benchmark dataset.\textsuperscript{72} The residues with RSA > 20% and predicted to be lipid accessible by Rosetta were labeled as lipid exposed, and the others were classified as water exposed.

<table>
<thead>
<tr>
<th>Method</th>
<th>Description</th>
<th>Input</th>
<th>Threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>FoldX</td>
<td>ΔΔG calibrated empirical force field using grid search</td>
<td>3D structure</td>
<td>ΔΔG &lt; 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Stabilizing</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ΔΔG &gt; 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Destabilizing</td>
</tr>
<tr>
<td>Rosetta</td>
<td>Knowledge-based potential using Cartesian space sampling</td>
<td>3D structure</td>
<td>ΔΔG &lt; 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Stabilizing</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ΔΔG &gt; 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Destabilizing</td>
</tr>
<tr>
<td>I-mutant 2.0</td>
<td>Support vector machine based protein stability predictor</td>
<td>3D structure</td>
<td>ΔΔG &lt; 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Stabilizing</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ΔΔG &gt; 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Destabilizing</td>
</tr>
<tr>
<td>mCSM-Stability</td>
<td>Graph-based structural signatures</td>
<td>3D structure</td>
<td>ΔΔG &lt; 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Stabilizing</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ΔΔG &gt; 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Destabilizing</td>
</tr>
<tr>
<td>SDM</td>
<td>Statistical potential energy function</td>
<td>3D structure</td>
<td>ΔΔG &lt; 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Stabilizing</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ΔΔG &gt; 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Destabilizing</td>
</tr>
<tr>
<td>DUET</td>
<td>Support vector machine based protein stability predictor, combines mCSM and SDM</td>
<td>3D structure</td>
<td>ΔΔG &lt; 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Stabilizing</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ΔΔG &gt; 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Destabilizing</td>
</tr>
<tr>
<td>POPMUSIC</td>
<td>Linear combination of statistical potentials</td>
<td>3D structure</td>
<td>ΔΔG &lt; 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Stabilizing</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ΔΔG &gt; 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Destabilizing</td>
</tr>
<tr>
<td>HOTMUSIC</td>
<td>Thermodynamics based and knowledge driven melting temperature (Δ\textsubscript{Tm}) predictor</td>
<td>3D structure</td>
<td>ΔT\textsubscript{m} &gt; 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Stabilizing</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ΔT\textsubscript{m} &lt; 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Destabilizing</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ΔΔG &gt; 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Stabilizing</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ΔΔG &lt; 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Destabilizing</td>
</tr>
<tr>
<td>EASE-MM</td>
<td>Sequence-based prediction with five support vector machine models</td>
<td>Sequence</td>
<td>ΔΔG &lt; 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Stabilizing</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ΔΔG &gt; 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Destabilizing</td>
</tr>
<tr>
<td>Algorithm</td>
<td>Type</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>----------------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>mCSM-PPI</td>
<td>Graph based structural signatures</td>
<td>3D structure</td>
<td></td>
</tr>
<tr>
<td>PolyPhen-2</td>
<td>Naive Bayes classifier trained by HumDiv and HumVar dataset</td>
<td>Sequence</td>
<td></td>
</tr>
<tr>
<td>SIFT</td>
<td>Uses sequence homology (PSI-BLAST) combined with known generic likelihood of amino acid substitutions</td>
<td>Sequence</td>
<td></td>
</tr>
<tr>
<td>PROVEAN</td>
<td>Uses pairwise sequence alignment scores (BLAST)</td>
<td>Sequence</td>
<td></td>
</tr>
<tr>
<td>PMUT</td>
<td>Neural network based classifier trained by a manually curated SwissProt dataset</td>
<td>Sequence</td>
<td></td>
</tr>
<tr>
<td>SNPMUSIC</td>
<td>Stability driven knowledge based classifier that uses statistical potentials</td>
<td>3D structure</td>
<td></td>
</tr>
<tr>
<td>PANTHER</td>
<td>A statistical method for scoring the functional likelihood of different amino acid substitutions</td>
<td>Sequence</td>
<td></td>
</tr>
<tr>
<td>PHD-SNP</td>
<td>Support vector machine based classifier trained by SwissProt dataset</td>
<td>Sequence</td>
<td></td>
</tr>
<tr>
<td>SNAP/SNAP.2</td>
<td>Neural network based classifier trained by experimentally obtained variant functional effect data</td>
<td>Sequence</td>
<td></td>
</tr>
<tr>
<td>META-SNP</td>
<td>Random forest based meta predictor combining SNAP, SIFT, PATHER and PHD-SNP</td>
<td>Sequence</td>
<td></td>
</tr>
<tr>
<td>FATHMM</td>
<td>Hidden Markov Models based classifier</td>
<td>Sequence</td>
<td></td>
</tr>
</tbody>
</table>
Results and Discussion

Computing the pathogenicity of PSEN1 mutations. The three main goals of this work are 1) to explore whether any computational method can describe the pathogenicity of the main genetic risk factor of FAD caused by PSEN1 mutations; 2) to test whether we can quantitatively rank the real clinical severity of the mutations using the reported average clinical age of symptom onset as observable; and 3) to identify the causes of failure or success, with a longer term aim of making new accurate disease-predicting methods specifically directed towards AD, which affects more than 30 million people worldwide.

To this end, we benchmarked 21 widely used protein-stability and sequence-based methods against the available clinical data. We designed a test using the complete hypothetical saturation mutagenesis as control, by computing all possible mutations at each position of the protein. After removing the mutations with confirmed pathogenicity, this test set can be compared to the pathogenic set using ANOVA analysis to explore the basis for qualitatively distinguishing pathogenic mutations from random mutations in PSEN1. We envision that this test will aid other studies of pathogenic mutations as it can be applied to any property suspected of relating to clinical pathogenicity as long as the property can be computed for all possible mutations within the protein.

The results of this analysis using ten widely used methods (FoldX, Rosetta, I-mutant, mCSM-Stability, SDM, DUET POPMUSIC, HOTMUSIC, mCSM-PPI and EASE-MM) are shown in Figure 2. Although the four confirmed non-pathogenic mutations do not constitute a statistically relevant control set, we added them to Figure 2 for completion; we hope that future establishing of non-pathogenic PSEN1 mutations will enable a test also using such data but our saturated mutagenesis test is sufficient for the purpose as shown below. Although the set of all other mutations includes both neutral as well as undiscovered pathogenic and non-pathogenic mutations, we believe that all possible mutations in the protein are the relevant background for testing any disease hypothesis.
As can be seen from Figure 2, the obtained ΔΔG values show a narrow range for all the methods, except FoldX and Rosetta. High values have been reported in previous studies using exhaustive mutagenesis with these two methods. Figure 2 consistently suggests that the pathogenic mutations generally tend to destabilize the protein. However, the non-pathogenic mutations do the same, and any random mutation is expected to be destabilizing, so the significance of this observation is not evident. Notably, as there were only four clinically confirmed non-pathogenic mutations reported in the database (and only three available for the structure-based protein stability methods as one is in the elusive N-terminal part), the non-pathogenic control is not significant and only serves as a comparison of interest. Instead, comparison to all other possible mutations by ANOVA enables a statistical test of the features that drive the pathogenicity of PSEN1 mutations.

To see whether any method captures the differences between the pathogenic mutations and all other possible mutations in the protein, the corresponding ANOVA data are shown in Table S2. There is no significant difference in average ΔΔG between the total group of all other mutations ("All.else") and the pathogenic group for any of the protein stability methods (p > 0.05). This is true regardless of using the full γ-secretase structure or only PSEN1 alone, even though the structural context affects the mutation sites that are located close to other subunits of the protein complex, albeit not so much the pathogenic mutations which are mainly located away from the other subunits.

The results for sequence-based methods are also shown in the lower part of Figure 2. Compared to protein stability methods, the mean scores for pathogenic and non-pathogenic mutations in some cases show significant differences, and several methods correctly classify the pathogenic mutations as pathogenic (Table S2). According to the applied ANOVA, all the sequence-based methods except SNPMUSIC produced significant differences between the control group and the pathogenic group at the 95% confidence level or higher (p < 0.05).
Figure 2. Distribution of scores from protein stability methods (FoldX, Rosetta, I-mutant, mCSM-Stability, SDM, DUET, POPMUSIC, HOTMUSIC, mCSM-PPI and EASE-MM) and sequence-based disease prediction methods (PolyPhen-2, SIFT, PROVEAN, PMUT, SNPMUSIC, PANTHER, PHD-SNP, SNAP, META-SNP, SNAP.2) for all non-pathogenic, pathogenic, and all other possible ("All.else") mutations. For the protein stability methods (FoldX, Rosetta, I-mutant, mCSM-Stability, SDM, DUET, POPMUSIC, HOTMUSIC and mCSM-PPI), the dot represents using the full γ-secretase structure as input and the triangle represents using only PSEN1 with different color in the panel. The black dot is the mean value, and the bar shows the standard error. Data points have been randomly separated on the horizontal axis for clarity, with the black shapes indicating the density of points.
For the structure-based stability methods (FoldX, Rosetta, I-mutant, mCSM-Stability, SDM, DUET, POPMUSIC, HOTMUSIC and mCSM-PPI), the obtained ΔΔG values using only PSEN1 as input are also shown in Figure 2. The distributions are very similar. However, the mean values were significantly different for the group of all possible mutations when using the two protein structures as input (Table S2). Thus, we conclude that other protein subunits in some cases contribute energetically to the ΔΔG effects of PSEN1 mutations, but not very much for the pathogenic mutations which tend to be oriented away from the other subunits. Using PSEN1 alone, there is still no significant difference between pathogenic and all other possible mutations, in agreement with the results obtained using the full γ-secretase structure. Accordingly, the structural sensitivity does not affect our conclusion that ΔΔG is not significantly explaining the pathogenicity of PSEN1 mutations.

Distinct behavior of buried, water-exposed, and membrane-exposed residues. Previous studies have revealed that most pathogenic mutations occur in buried sites of proteins or in surface sites involved in molecular interactions, which have a potentially larger impact on stability and function.\textsuperscript{34–39} Accordingly, we want to know how the methods work for different types of sites, as site-specific differences could in principle relate to FAD pathogenicity. Therefore, the PSEN1 residues were divided into buried, water-exposed and membrane-exposed based on their relative solvent accessibility (RSA) values (cutoff 20% in RSA). Then the ΔΔG values and scores were calculated for each residue category, and ANOVA was again performed. The results are shown in Figure 3 (additional data in Figure S1 and Table S3). The confirmed non-pathogenic mutations are all exposed, potentially consistent with a neutral effect on protein function although more data are required for any strong conclusion on this. ΔΔG values still exhibited no ability to discriminate between pathogenic mutations and all other possible mutations in the protein (Table S3), except for I-mutant applied to water-exposed residues. Therefore, in general, we conclude that the loss of protein stability is not a main driver of PSEN1 FAD pathogenicity, as far as the structural models can be trusted to not miss a central
physiological context, notably perhaps the membrane environment. We cannot rule out that
stability or membrane packing conformation can contribute to pathogenicity, but membrane
models cannot currently be run by the programs. Thus as used here, the results mainly serve as
relevant negative control examples of protocols that do not capture the pathogenicity.

As seen from Figure 3 and Table S3, the sequence-based disease prediction methods cannot
capture the differences between pathogenic and all other possible mutations for all residue
categories; they work mainly for buried and water-exposed residues. However, compared to
protein stability methods, many of these methods show a very promising and much better ability
to identify pathogenic mutations. We thus decided to study these methods further.

**Accuracy and sensitivity of the methods.** To further test the ability of the methods to
distinguish pathogenic mutations, receiver operating characteristic (ROC) analysis was
performed using the area under the curve (AUC) as evaluation metric; the results are shown in
Figure 4A. Other evaluation metrics and the optimal threshold of all the tested methods are
shown in Table S4. Based on the AUC values, Figure 4A lists the top-10 methods among all the
benchmarked methods. All these methods use evolutionary information from multiple sequence
alignments, in agreement with the above-described ANOVA results. Interestingly, as shown in
Figure 4A, most of these methods perform reasonably well for distinguishing pathogenic and
non-pathogenic mutations, indicating their potential use in predicting the pathogenicity of newly
identified PSEN1 mutations, which could be of diagnostic interest. The META-SNP method
exhibited the best distinguishing power with AUC of 0.92, followed by PolyPhen (0.91) and
PMUT (0.91).
Figure 3. Scores and score distribution obtained from bioinformatics methods (PolyPhen-2, SIFT, PROVEAN, PMUT, SNPMUSIC, PANTHER, PHD-SNP, SNAP, META-SNP, SNAP.2) for buried, water exposed and lipid exposed residues. Each method has a unique color. The black dot is the mean score value for each group and the error bar is the standard error. Data points have been randomly separated on the horizontal axis for clarity, with the black shapes indicating the density of points.
Figure 4. (A) Receiver operating characteristics curves for the top-10 methods (META-SNP, PolyPhen-2, PMUT, PHD-SNP, SIFT, SNAP.2, SNAP, PANTHER, PROVEAN and FATHMM) according to the area under curve values. (B) Accuracy of the sequence-based methods (PolyPhen-2, SIFT, PROVEAN, PMUT, SNPMUSIC, PANTHER, PHD-SNP, SNAP, META-SNP, SNAP.2 and FATHMM) for predicting the pathogenicity of a mutation. The inset figure shows the prediction accuracy for buried (red color), water exposed (blue color) and lipid exposed (green color) residues.

Figure 4B shows the prediction accuracy. Notably, FATHMM has a prediction accuracy of 1, classifying all the PSEN1 mutations as pathogenic but failing to identify non-pathogenic mutations. PolyPhen exhibited the second-highest prediction accuracy (0.95), followed by SNAP.2 (0.89) and PROVEAN (0.87). The prediction accuracy for each residue category is also shown in the insert panel of Figure 4B. In general, the methods have better prediction accuracy for buried residues.
Amino acid conservation patterns central to the successful tested methods seem to be important descriptors of real clinical FAD caused by PSEN1 mutations. The three most accurate methods all use amino acid conservation information for prediction.\textsuperscript{62,67,69} To understand this finding in more detail, we calculated the conservation score for each residue of PSEN1 as shown in Figure S2. As can be seen, most of the pathogenic mutations occur in conserved regions indicating that the residues could be structurally and functionally critical to the enzymatic activity of \(\gamma\)-secretase. This agrees well with the FIST mechanism of \(\gamma\)-secretase that involves perturbation of the compactness of the protein complex during C99 processing, leading to a change in the A\(\beta\)42/ A\(\beta\)40 ratio,\textsuperscript{45,77} but more studies are needed to confirm this relationship between C99 processing and the structure-disrupting effect of PSEN1 mutations.

**Correlation between allele frequency, age of onset, and method scores.** The recent increase in population sequencing data makes the use of computational screening methods of major importance in rationalizing variant phenotypes. Such data allow us to measure the effect of the natural selection on missense mutations across the allele frequency spectrum.\textsuperscript{78} Allele frequency is often used as a key criterion for clinical interpretation of sequence variants. Therefore we analyzed the predicted pathogenicity using the four methods with highest prediction accuracy for the PSEN1 missense mutations found in the database of the Exome Aggregation Consortium (ExAC) and the Genome Aggregation Database (gnomAD).\textsuperscript{79} The results are shown in Figure 5. As the gnomAD has more data for missense mutations in PSEN1, the final allele frequency used in the analysis was compiled from gnomAD. As seen from Figure 5, mutations with high frequency are more likely to be benign and accurately predicted by the methods, as implied by the clinical diagnostic filtration criteria. However, we also noticed that a few confirmed pathogenic mutations have relatively high allele frequency so this relationship is not capable of predicting pathogenicity in all cases. Still, our data suggest that the allele frequency provides information that should be incorporated in future prediction studies of FAD.
Figure 5. The distribution of scores by PolyPhen-2, PROVEAN, PMUT, and SNAP.2 vs. the allele frequency of PSEN1 mutations according to the Genome Aggregation Database. The red, blue, and green points represent random, pathogenic, and non-pathogenic mutations, respectively. The inset plots are the allele frequencies after removing the high allele frequencies. The three horizontal lines in the left panel are the score quantiles for all mutations, whereas the lines in the right panel are the score quantiles for mutations found in gnomAD.

We have shown that some computational methods with good accuracy qualitatively captures the pathogenicity of FAD-causing PSEN1 mutations. We want to consolidate this finding by also testing the methods quantitatively against actual clinical data reflecting the severity of each mutation. We used the clinical age of onset data from Ryman’s meta-analysis\(^46\) as these are the least noisy available. The results of the linear regression analysis are shown in Figure 6. There is no correlation between the age of onset and ΔΔG, but there are several significant relationships for the disease prediction methods. When separating into residue categories it is clear that some residue types are much better described than others (Figure S3), although all the stability methods fail to produce significant correlation for all residue types except three methods, where one of the three residue types is significantly described (Figure S4).
Figure 6. The linear relationship between scores of sequence- and structure-based disease prediction methods and clinical age of symptom onset for all pathogenic mutations in total.
Figure 6 reveals significant correlations for the sequence-based disease prediction methods, with $p < 0.05$. Best correlations were found for water-exposed residues. The small $R^2$ values in Figure 6 arise from the noise in the clinical data due to many risk modifiers; thus a high $R^2$ could never be expected for these relationships, as the noise in clinical data will always reduce any correlation. What is notable is however that a significant linear relationship is obtained, and the probability that these relationships are coincidental is smaller than 5% for several of these methods. This quantitative relationship is the first of its kind and the only one to describe clinical severity of FAD mutations aside from the experimentally observed $\alpha\beta_{42}/\alpha\beta_{40}$ ratio.

Based on all above analysis, several methods (notably PolyPhen, PANTHER, PMUT) showed good prediction power for identifying the pathogenic PSEN1 mutations not only qualitatively but also quantitatively. We note that the scores of Polyphen tend to cluster near 1 as a feature of a binary classification, and thus other methods such as PANTHER and PMUT with comparable significant $p$-values show better correlation structure across the data range. To understand the reasons for the success of these methods, we analyzed the PolyPhen-2 scores in more detail. The 20 amino acids were divided into eight groups: aliphatic amino acids (AVILMC), aromatic amino acids (FWY), negative charged (DENQ), positive charged (HKR), phosphorylatable (STY), small (AGST), and proline (P) and glycine (G). The PolyPhen-2 scores were recalculated based on these groups with results shown in Figure 7. For lipid-exposed residues, the mutation of any glycine, proline or charged amino acids is expected to be a severe perturbation of the $\gamma$-secretase structure and function. From our pathogenic mutation database (Table S1) we find 125 pathogenic mutations associated with glycine, proline and charged amino acids, which account for 53% of all confirmed pathogenic mutations. Moreover, these mutations were also associated with early age of onset (average ~ 42 years) (Table S1). For other mutation categories, we did not find systematic patterns. We consider this finding significant as it suggests that disruption of $\gamma$-secretase conformation state rather than fold stability per se is a likely driver of FAD caused by PSEN1 mutation, in agreement with our general FIST model of $\gamma$-secretase-related FAD.
Figure 7. The PolyPhen-2 scores and score distribution for buried, water-exposed, and lipid-exposed residues according to different mutation categories. The black dot is the mean value for each category and the bar shows the standard error. Three red horizontal lines represent the quantiles of score value. The black horizontal line represents the mean score of the pathogenic mutations. G and P represent glycine and proline, respectively.

Conclusions

The pathogenicity of PSEN1 mutations is accurately described by several methods as indicated by the present study. The good performance of methods such as PolyPhen-2, SNAP.2, and PROVEAN points to the importance of residue conservation in FAD caused by PSEN1 mutations. All the protein stability methods failed to capture the pathogenicity of the PSEN1 mutations suggesting protein stability may not a key factor in PSEN1-mutation caused FAD. Compared to water soluble proteins, the prediction accuracy for membrane proteins is probably much lower. The experimentally based, MD-equilibrated protein structure used in the present
study may not correctly capture the γ-secretase stability changes within a real membrane. Alternatively, loss of protein stability is not a key driver for FAD caused by PSEN1 mutation. We speculate that PSEN1 mutations will change the conformation of the protein-membrane system in a way that is beyond study by any of these methods, and we are currently pursuing this problem by explicit large-scale MD simulations of select, prominent mutants.

Our main result is that we have identified several computational methods that recapitulate the pathogenicity of PSEN1 mutations, which are the most important genetic risk factors of early-onset FAD, not just qualitatively (pathogenic vs. all other possible mutations) but also semi-quantitatively (by correlating with clinical age of onset of each type of pathogenic mutation). The evolutionary amino acid conservation patterns are central to this success. Mutations involving proline, glycine, and charged residues contribute particularly to pathogenicity, supporting a model where conformational integrity, rather than fold stability per se, is a molecular driver of disease. The "Fit-Stay-Trim mechanism" (FIST) implies that conformational integrity of γ-secretase is central to FAD, as a tighter grab by γ-secretase increases the residence time of C99 within the catalytic pocket, leading to enhanced trimming and smaller Aβ peptides by consecutive proteolytic cleavage. In contrast, PSEN1 mutations tend to disturb this conformational integrity, most likely by favoring the open state of γ-secretase, which leads to shorter residence time, less trimming, and a relative increase in the produced toxic Aβ42.

We hope that our findings will be of value in the development of more accurate quantitative models of disease that can help as a tool of personalized medicine, most notably the diagnosis and strategic planning of patients with FAD.

Acknowledgements

The Danish Council for Independent Research | Natural Sciences (DFF), grant case 7016-00079B, and the Novo Nordisk Foundation, grant NNF17OC0028860, are gratefully acknowledged for supporting this work.
Supporting Information Available

The supporting information file contains the clinical data sets (Table S1), details of the ANOVA (Table S2 and S3), properties of the optimization procedures (Table S4), distributions of computed ΔΔG values (Figure S1), conservation scores (Figure S2), and linear relationships with clinical data divided into water-exposed, lipid-exposed, and buried residues (Figure S3 and S4). This material is available free of charge via the Internet at http://pubs.acs.org.

Conflict of interest

The authors declare that they have no conflict of interest associated with this work.

References


(6) Bateman, R. J.; Aisen, P. S.; De Strooper, B.; Fox, N. C.; Lemere, C. A.; Ringman, J.


(17) Kepp, K. P. Alzheimer’s Disease Due to Loss of Function: A New Synthesis of the


(34) Teng, S.; Madej, T.; Panchenko, A.; Alexov, E. Modeling Effects of Human Single


(50) Schymkowitz, J.; Borg, J.; Stricher, F.; Nys, R.; Rousseau, F.; Serrano, L. The FoldX


TOC graphic
TOC graphic

87x34mm (600 x 600 DPI)
Figure 1

108x60mm (300 x 300 DPI)
Figure 2

101x101mm (300 x 300 DPI)
Figure 3

101x114mm (300 x 300 DPI)
Figure 4

194x109mm (300 x 300 DPI)
Figure 5

685x330mm (300 x 300 DPI)
Figure 6

304x431mm (300 x 300 DPI)
Figure 7

127x93mm (300 x 300 DPI)