Aβ42/Aβ40 Ratios of Presenilin 1 Mutations Correlate with Clinical Onset of Alzheimer’s Disease

Tang, Ning; Kepp, Kasper Planeta

Published in:
Journal of Alzheimer’s Disease

Link to article, DOI:
10.3233/JAD-180829

Publication date:
2018

Document Version
Peer reviewed version

Citation (APA):
$A\beta_{42}/A\beta_{40}$ ratios of presenilin 1 mutations correlate with clinical onset of Alzheimer’s disease

Ning Tang and Kasper P. Kepp*

*Technical University of Denmark, DTU Chemistry, Building 206, 2800 Kgs. Lyngby, DK – Denmark. *Phone: +045 45 25 24 09. E-mail: kpj@kemi.dtu.dk
Abstract
Missense mutations in presenilin 1 cause early onset familial Alzheimer’s disease in a way that is not understood. Increased $\text{A}\beta_{42}/\text{A}\beta_{40}$ ratios are the most consistent biochemical phenotype of such mutations in cultured cells and *in vivo*, and are thus considered central to the amyloid hypothesis. Previously, an inverse relation has been observed between the $\text{A}\beta_{42}/\text{A}\beta_{40}$ ratio of such mutants and the clinical age of symptom onset in patients carrying the mutation. However, a recent extensive study by Sun et al. of assayed presenilin 1 mutants concluded that such a relationship is not evident. To reconcile the disagreement, three different clinical datasets were compared directly with the data by Sun et al. After considering data noise and measurement uncertainty, we find a clear and highly significant inverse correlation between the $\text{A}\beta_{42}/\text{A}\beta_{40}$ ratio and the clinical age of onset in all three datasets even without removing noisy single- and double-patient data. With these data removed, the correlation coefficients increase further. The probability that these relationships are coincidental are approximately 0.1%.

**Key words:** Presenilin, Alzheimer’s disease, mutations, $\text{A}\beta_{42}/\text{A}\beta_{40}$ ratio, age of onset
Early-onset familial Alzheimer’s Disease (FAD) is a particularly severe form of Alzheimer’s disease (AD) caused by mutations in the genes coding for amyloid-β precursor protein (APP), presenilin 1 (PSEN1) or presenilin 2 (PSEN2) [1–3]. Despite constituting only a small fraction of all AD cases, these mutations are a cornerstone of efforts towards understanding the pathogenesis [4,5]. Variations in PSEN1 are the most common and also tend to cause earlier age of onset, with symptoms in some cases reported before the age of 35 [6–8].

Despite recent setbacks, the amyloid hypothesis still dominates current drug development efforts. The hypothesis argues that the Aβ peptides produced from APP upon cleavage by β- and γ-secretases gain a toxic function within neurons probably in the form of oligomers [9–13], and probably by direct interactions with cell membranes that provide the immediate environment of the newly formed peptides after cleavage by γ-secretase [14–17]. However, total measured levels of Aβ are commonly reduced in phenotypes of FAD mutants expressed in cultured cells, i.e. overload per se cannot be the cause of disease. Thus, the original “cascade” version of the hypothesis [12] focusing on toxicity by quantity has now been abandoned in favor of toxicity by quality [3]. The only consistent feature of most FAD mutations, and this is only partly true for PSEN1 and PSEN2, not APP [3], is the curious increase in the ratio of produced longer forms of Aβ, in particular Aβ42, vs. shorter isoforms, most notably Aβ40. This phenotype is a hallmark of most PSEN1 and PSEN2 mutations [18–21].

The increase in the Aβ42/Aβ40 ratio seems to support a pathogenic mechanism that puts little emphasis on total amounts of Aβ but instead focuses on the relative amount of the two peptide forms. Despite this, most drug discovery efforts seek to lower Aβ levels indiscriminately and have perhaps failed for this reason [10,22], since monomeric Aβ40 is well-known to serve important roles in the brain [23–26]. Conspicuously, not even the total amount of Aβ42 tends to increase consistently in PSEN1 phenotypes, only the ratio, since reduction of
both isoforms is common. As far as we know, only one plausible mechanism has been suggested that keeps the gain-of-function amyloid hypothesis alive from the viewpoint of these data, i.e. a competitive seeding mechanism where only the local critical amount of Aβ\textsubscript{42} vs. Aβ\textsubscript{40} is required to seed the formation of oligomers, even when global levels of Aβ\textsubscript{42} may be reduced [26].

The Aβ\textsubscript{42}/Aβ\textsubscript{40} ratio is the only approximately consistent biochemical phenotype of FAD variants, and it has accordingly been extensively explored [27–29]. If the ratio is pathogenic, it should manifest clinically. Thus, patients carrying PSEN1 mutations leading to higher Aβ\textsubscript{42}/Aβ\textsubscript{40} ratios would presumably also experience more severe disease on average, unless other modifiers make such relationships too noisy [29]. It is therefore of major importance that the Aβ\textsubscript{42}/Aβ\textsubscript{40} ratio has been reported to correlate inversely with the age of symptom onset in human mutant carriers, as seen in multiple studies [29–31]. In our view, this is the minimum necessary (but far from sufficient, as this phenotype could be a side effect of a loss of function mechanism) requirement for the gain-of-function hypothesis that dominates current drug development efforts.

Because of the issue with the noisy clinical data, which arises mainly from less trustworthy single- and double-patient observations, we previously argued that the age of onset should correlate more convincingly to Aβ\textsubscript{42}/Aβ\textsubscript{40} ratios if the noisy data are removed, and for the data where a mutation’s clinical age of onset has been confirmed three times or more, i.e. with the number of patients $N > 2$, the correlation is indeed strong [29].

Recently, Sun et al. performed the most complete assaying of PSEN1 mutants so far, providing a data set of Aβ\textsubscript{42}/Aβ\textsubscript{40} ratios that will be invaluable to the community for many years to come [32]. However, in addition to reporting these results, the authors also compared these assayed data to a clinical data set of age of onset. From this comparison they concluded that there is no significant correlation between the Aβ\textsubscript{42}/Aβ\textsubscript{40} ratio and the mean age of onset [32], in disagreement with our previous findings. It is important to determine whether there is
correlation, as its absence would substantially weaken the amyloid hypothesis (although correlation still does not guarantee causation). Sun et al. state directly in their abstract that there is no significant correlation, and this conclusion has been accommodated by e.g. Kelleher and Shen who favor a presenilin loss-of-function hypothesis of AD [33] and De Strooper’s group favoring a modified version of the gain-of-function mechanism modulated by multiple risk factors [34]. A third possibility is a broad loss-of-function mechanism with the impaired function being neuronal metal transport [26,35,36]. We want to dissect the relationship since a correlation for the full data set, if it exists, is truly remarkable and important to any etiology, regardless of its nature.

To further address this question, we re-examined the data reported by Sun et al. in relation to the various previous data sets and correlation studies that have been put forward [29–31]. The assayed data set arose from a single lab protocol and thus solves a major problem of data heterogeneity that undermines many genotype-phenotype analyses of this type [3,36]. However, the actual regression analysis (not the data produced) is affected by two issues that unfortunately change the main conclusions of the work. One is the use of clinical data taken directly from a single published webpage (https://www.alzforum.org/) without addressing clinical data noise arising from the many risk modifiers of the patients [3,29], and the other is the inclusion of an erroneous assayed data point during analysis.

As to the first point, the regression analysis by Sun et al. relies on age of onset data from (https://www.alzforum.org/) that are similar to but less complete than the data set that we used previously for the same purpose [29]. The data consist of single reported averages (without standard deviations) of the clinically reported values from sometimes many patients, sometimes a single patient. Because of the noise in the clinical data, it is necessary to study data set dependencies on conclusions and the effect of noise reduction [29]. The age of onset is highly uncertain when only reported in a single patient: To illustrate this, the variation in age of onset among multiple observations even within same families carrying the same mutant
commonly surpasses 5 years [3,29]. Accordingly, average values resulting from multiple observations are more significant. The fact that data points can be cherry-picked to have any desired correlation is not surprising as it is a hallmark of such noisy data. This does not mean that correlations are absent.

The second point is that the assayed data vary substantially in quality, and the assay covers the range to zero activity, where the calculated ratio becomes highly uncertain. Because of this, the standard errors reported in an experiment will not represent true errors due to small-size effects when handling the samples, and some measured Aβ_{42}/Aβ_{40} ratios will be much more uncertain than others. The largest problems arise in the limit where Aβ_{40} goes to zero.

To understand this, we first redid the analysis using the same data as Sun et al. (Data set 1, Supplementary Table S1). The most extreme outlier is G384A with a Aβ_{42}/Aβ_{40} ratio of 171, arising from the unique combination of a very large Aβ_{42} production (1.870) and an almost zero (0.011 vs. WT = 1) Aβ_{40} production. The reported ratio is almost ten times larger than the second largest value in the data set (A260V with Aβ_{42}/Aβ_{40} = 18.058) and is thus not a representative part of the overall data set. The remaining ratios range from 0-18 with reasonable spread and coverage, whereas the G384A data point is represented by an off-scale insert in Figure 5 of Sun et al.

If only this extreme outlier is removed from analysis, keeping all other data points directly, a highly significant correlation with p ~ 0.0011 is already found between Aβ_{42}/Aβ_{40} ratio and mean age of onset (Figure 1). Thus, the adequate removal of this single, extreme outlier out of the 138 mutations alone is enough to revert the main conclusion of Sun et al. We therefore conclude that even for the full data set of noisy clinical data, there is very significant correlation. This correlation is indeed remarkable and the p-value implies that the relationship is 0.1 % likely to be coincidental.

This correlation arose from the use of all noisy data from alz.org. As mentioned, another problem is the quality of these clinical data. In order to have a more systematic examination,
we also compared the assayed data to two other datasets (Dataset 2 and Dataset 3 in **Supplementary Tables S2 and S3, respectively**). Data set 2 includes consensus data from both [alzforum.org](http://alzforum.org) and the Alzheimer Disease & Frontotemporal Dementia Mutation Database [6], whereas Data set 3 was compiled by Ryman et al. using a distinct search approach [29,37]. Data set 1 and 2 are very similar but differ notably by the presence of five data points reported by Sun et al. (R35Q, G183V, N405S, L424V, I437V, and I439V), and in 11 other values of the age of onset that differ by more than 1 year between the two data sets, with the largest variation seen for M233L (60 vs. 42.8 years). Data set 3 by Ryman et al. is very distinct from the other data sets and thus serves as a strong external validation of the relationship discussed above. As seen from Figure 1, data sets 2 and 3 also show highly significant correlations with p = 0.0011 and 0.00023 using the ratios by Sun et al directly after removing G384A. This correlation is mostly determined by the amount of Aβ_{42} rather than enzyme activity or the amount of Aβ_{40}, in agreement with previous findings [30].
Figure 1. Linear regression study of average clinical age of symptom onset and enzyme activity, Aβ_{42}/Aβ_{40} ratio, amount of Aβ_{40} and amount of Aβ_{42} for three different datasets.

Despite the significance, the $R^2$ values from the study by Sun et al. are smaller than those of previous correlation studies. This relates partly to different assayed data for the Aβ peptides, and partly to the use clinical data with single-patient observations included. Previous correlation studies used a cell assay which simulated a membrane-like environment [29–31], whereas the detergent-based method was used by Sun et al. [32] Although this method is widely used and can produce reliable relative data, it may reduce the processing of γ-secretase, as
reported in many studies [38–42]. The presence and exact nature of the membrane may alter the production of Aβ, including the Aβ_{42}/Aβ_{40} ratio. In comparison with the cell-based assays, detergent-solubilized conditions tend to generate more longer Aβ peptides [40–45]. Sun et al. compared different methods using a small dataset containing 11 mutations [32]. The Aβ_{42}/Aβ_{40} ratios generated by the detergent-based assay were similar to those generated by the liposome-based or cell-based assay. However, some ratios did show significant difference, and even a 30% difference could explain some variations in the reported correlations studies. Considering the differences in protocol, it is remarkable that all data after reanalysis strongly support a correlation.

The correlations of Figure 1 are even more remarkable considering that they persist for the full data set, without any separation into subtypes of mutations. One could easily imagine that the correlation (and cause) prevails only for a certain chemical type of modifications relating to the protein structure and dynamics during C99 processing, as the ratio has been suggested to inversely relate to the stability and life time of the enzyme-C99 complex [46].

As mentioned, the noise in age of onset relates to the patient count \( N \). Figure 2 shows the results for all data where \( N > 2 \). As can be seen, the correlation remains highly significant between the Aβ_{42}/Aβ_{40} ratio and the mean age of onset, and the correlation coefficient has increased. This suggests that better refinement (i.e. noise reduction) of the clinical age of onset by the arrival of future clinical data will improve the correlation between Aβ_{42}/Aβ_{40} ratio and mean age of onset further.

It should be noted that the correlation is better for Dataset 3, which was compiled from Ryman et al. who specifically worked to compile additional, more complete clinical data [37]. This fully supports our notion that noise in the clinical data is a main reason for the lack of correlation and that correlation improves with data quality, strongly arguing for the authenticity of the relationships in Figure 2.
The main reason for the identified strong relationship is the quality of the clinical data of early onset FAD patients. In this context, it is essential that the data set covers well patients with early disease onset [29]. Severe early-onset mutations are less sensitive to risk modifiers than late-onset mutations[26], and thus it will be very hard to find similar relationships for sporadic AD, where younger patients may be less well sampled and where risk modulators will be very strong. To illustrate this, recent work by Nakamura et al.[47] suggests that the
Aβ$_{42}$/Aβ$_{40}$ ratio in plasma is a promising general biomarker: We plotted the reported age of the patients by Nakamura et al. against their reported Aβ$_{40}$/Aβ$_{42}$ ratio in plasma (Supplementary Figure S1). However interestingly, $P = 0.07$ and the direction of correlation is pathogenically meaningful as plasma Aβ$_{42}$ is relatively reduced opposite to what is seen in cell cultures and brains. Assuming that the ages reported by Nakamura et al. correlate with the individual ages of onset, which is not unlikely, this relationship may warrant future investigation as more data arrive.

Based on above analysis, we conclude that there is a highly significant correlation between the Aβ$_{42}$/Aβ$_{40}$ ratio and the clinical age of onset. The two features correlate also for the important new data obtained by Sun et al. once the single extreme outlier is removed. The correlation now persists regardless of protocol used, as it was already known for data assayed by other protocols, and already before reducing the inherent noise of the clinical data. Furthermore, the analysis indicates that improved data quality, as will result from noise reduction and the arrival of new clinical data in the future, will further strengthen this relationship, as best indicated by the strong correlation found when using the new compilation of clinical data by Ryman et al.

Acknowledgements

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

Conflict of interest

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


amyloid beta-protein similar to that in the senile plaques of Alzheimer’s disease is increased in vivo by the presenilin 1 and 2 and APP mutations linked to familial Alzheimer’s disease. *Nat. Med.* 2, 864–870.


