

SCIENTIFIC REPORTS



OPEN

Diversity of Amyloid-beta Proteoforms in the Alzheimer's Disease Brain

Norelle C. Wildburger¹, Thomas J. Esparza¹, Richard D. LeDuc², Ryan T. Fellers², Paul M. Thomas², Nigel J. Cairns^{1,3,4,5}, Neil L. Kelleher^{2,6,7}, Randall J. Bateman^{1,3,5} & David L. Brody^{1,5}

Amyloid-beta (A β) plays a key role in the pathogenesis of Alzheimer's disease (AD), but little is known about the proteoforms present in AD brain. We used high-resolution mass spectrometry to analyze intact A β from soluble aggregates and insoluble material in brains of six cases with severe dementia and pathologically confirmed AD. The soluble aggregates are especially relevant because they are believed to be the most toxic form of A β . We found a diversity of A β peptides, with 26 unique proteoforms including various N- and C-terminal truncations. N- and C-terminal truncations comprised 73% and 30%, respectively, of the total A β proteoforms detected. The A β proteoforms segregated between the soluble and more insoluble aggregates with N-terminal truncations predominating in the insoluble material and C-terminal truncations segregating into the soluble aggregates. In contrast, canonical A β comprised the minority of the identified proteoforms (15.3%) and did not distinguish between the soluble and more insoluble aggregates. The relative abundance of many truncated A β proteoforms did not correlate with post-mortem interval, suggesting they are not artefacts. This heterogeneity of A β proteoforms deepens our understanding of AD and offers many new avenues for investigation into pathological mechanisms of the disease, with implications for therapeutic development.

A β has been the major target for disease-modifying therapeutic development in AD, driven in large part by studies of Down syndrome and studies of autosomal dominant AD and sporadic, late-onset AD, which implicate increased production, decreased clearance, and A β aggregation^{1,2}. However, little is known about the proteoforms (*i.e.*, all protein variants of a single gene including post-translational modifications and sequence variants)³ of A β in human AD brain. Previous studies have demonstrated some sequence heterogeneity and post-translational modifications (PTMs) of the A β peptide in amyloid-beta plaques⁴⁻⁹. Yet amyloid-beta plaques, one of the pathological hallmark of AD, correlate only moderately with dementia¹⁰⁻¹². As a result, focus has shifted in recent years to the most toxic forms of A β , soluble aggregates previously termed 'oligomers' and other appellations¹³, as they demonstrate a strong correlation with dementia¹⁴⁻¹⁹.

Toward this end, we used high-resolution mass spectrometry to obtain highly accurate measurements and peptide-sequencing information of intact, undigested A β from soluble and insoluble A β aggregates isolated from human AD brain (Table 1). Soluble A β aggregates were isolated with our fully quantitative, sensitive, specific, and efficient purification strategy²⁰ (Fig. 1). For comparison, more insoluble A β aggregates were analyzed from the same brain samples. Analysis of undigested (*i.e.*, top-down proteomics) peptides by nano-liquid chromatography tandem mass spectrometry (nLC-MS/MS) allows unambiguous identification of A β and fully characterizes the composition of individual proteoforms. The advantages of nLC-MS/MS over other techniques like matrix-assisted laser desorption/ionization (MALDI) include: *i*) chromatographic separation of various proteoforms before mass

¹Department of Neurology, Washington University School of Medicine, 660 South Euclid Avenue, St. Louis, MO, 63110, United States. ²Proteomics Center of Excellence, Northwestern University, Evanston, IL, United States.

³Department of Neurology, Knight Alzheimer's Disease Research Center, 4488 Forest Park Pkwy, St. Louis, MO, 63112, United States. ⁴Department of Pathology and Immunology, 660 South Euclid Avenue, St. Louis, MO, 63110, United States. ⁵Department of Neurology, Hope Center for Neurological Disorders, 660 South Euclid Avenue, St.

Louis, MO, 63110, United States. ⁶Department of Molecular Biosciences, Northwestern University, Evanston, IL, United States. ⁷Department of Chemistry, Northwestern University, Evanston, IL, United States. Correspondence and requests for materials should be addressed to N.C.W. (email: n.wildburger@wustl.edu) or D.L.B. (email: brodyd@neuro.wustl.edu)

Pt No.	Status	Age, yr	PMI, hr	Gender
1	CDR 3	73	12	male
2	CDR 3	83	6.3	female
3	CDR 3	92	18	female
4	CDR 3	86	15.5	female
5	CDR 3	89	6	female
6	CDR 3	82	10.75	male

Table 1. Characteristics of Human Brain Frontal Cortex Samples.

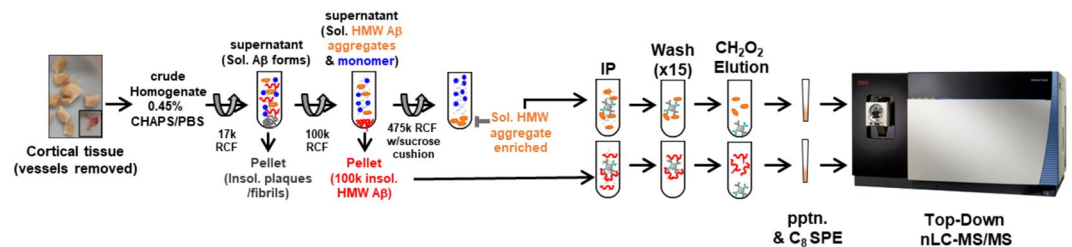


Figure 1. Isolation and comparison of A β proteoforms from human AD brain. Workflow for isolation of endogenous A β proteoforms. Cortical tissue was dounce homogenized with sub-critical micelle concentration of CHAPS followed by differential ultracentrifugation, anti-A β dual antibody immunoprecipitation, and elution in neat formic acid. Immunoprecipitated A β proteoforms in the HMW soluble aggregates fractions and more insoluble fractions were precipitated from undigested non-A β proteins using C $_8$ TopTips and analyzed by nLC-MS/MS. RCF, relative centrifugal force; Sol., soluble; LMW, low molecular weight; HMW, high molecular weight; IP, immunoprecipitation; pptn., precipitation; SPE, solid phase extraction.

spectrometry analysis, which offers better sample separation to effectively reduce sample complexity and improve resolution, *ii*) higher mass accuracy for the intact proteoforms, and *iii*) identification of fragment ions from direct *de novo* sequencing. Purification of soluble A β aggregates from frozen human AD brain by differential ultracentrifugation and immunoprecipitation resulted in \sim 10,000-fold enrichment and \sim 60% recovery of starting material²⁰. Even so, A β comprised approximately 0.1% of the total protein content of the soluble high molecular weight (HMW) aggregates (Supplementary Table 1), requiring further enrichment before mass spectrometry analysis. We achieved substantial additional separation of A β from the other, as-yet-uncharacterized protein components of the soluble A β aggregates through an optimized C $_8$ solid phase extraction protocol adaptable to varying amounts of total A β input (Supplementary Fig. 1; Supplementary Table 2). Using these methods, we analyzed A β extracted from brain tissue of six cases of AD each of whom had clinically severe AD dementia at expiration in an untargeted mass spectrometry approach (Supplementary Figs 2–4).

We identified 26 A β proteoforms in AD brain (Fig. 2). The proteoforms were identified with a high degree of mass accuracy, as were fragment ions that allowed unambiguous determination of their primary structures (Fig. 3, Supplementary Fig. 5). Contained within our dataset were numerous forms of A β with N-terminal truncations (A β x-40/42), C-terminal truncations (A β 1-x), and in one case (A β 11-34), both N- and C-terminal truncations (Fig. 2). The canonical forms of A β traditionally studied, A β 1–42 and A β 1–40, were present throughout all participants in all fractions. A β 1–42, in both its oxidized and un-oxidized forms had one of the highest relative abundances in the mass spectrometry datasets. A β 1–40 was also abundant, but with more variability amongst participants in both fractions, particularly in its oxidized form. Interestingly, no A β 1–38, or any 38- or 37-terminating proteoforms, were identified in our dataset even though this species is commonly detected in cerebrospinal fluid (CSF)^{21, 22}.

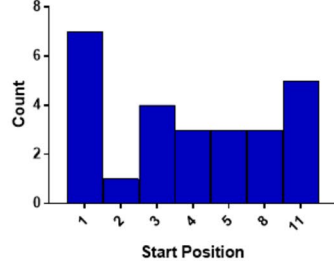
The N-terminal truncations identified include A β 2-x, A β 3-x, A β 4-x, A β 5-x, A β 8-x, and A β 11-x. C-terminal truncated forms of A β include A β x-20, A β x-31, and A β x-34 (Fig. 2A). The most frequent start and end positions were at 1 (D; aspartic acid) and 42 (A; alanine), respectively (Fig. 2B,C), like canonical A β 1–42. However closer inspection revealed that the most frequent proteoform length was 40 amino acids followed by 32 (Fig. 2D) and not a full 42 amino acid sequence. Only 4 out of 26 proteoforms were full-length (15.3%). The heterogeneity in sequence length can be largely attributed to the N-terminal heterogeneity (Fig. 2B) as 19 out of the 26 (73%) proteoforms contained N-terminal truncations as has been previously observed^{22, 23}. While less frequent by comparison, 8 out of 26 (30%) proteoforms contained C-terminal truncations. Overall, the majority of proteoform sequences terminated at alanine 42 (18 out of 26, 69.2%; Fig. 2C) compared to 9.5% terminating in valine 40 (4 out of 26) in keeping with previous studies demonstrating that A β x-42 aggregation rates were higher in amyloidosis²⁴. Traditionally used A β 1–40, A β 1–42, or A β 1-x (*i.e.*, canonical N-terminus) enzyme-linked immunosorbent assays (ELISAs) would not detect the majority of these forms, which fits well with our previous report that much of the A β measured by ELISA in human brain interstitial fluid is neither A β 1–40 nor A β 1–42²¹.

The post-translationally modified forms of A β comprised N-terminal pyro-glutamate, oxidation of methionine, and in one instance an ammonium ion adduct at the C-terminus, which to our knowledge has not been

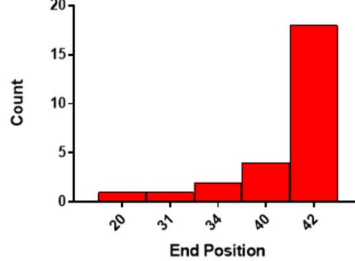
A



B



C



D

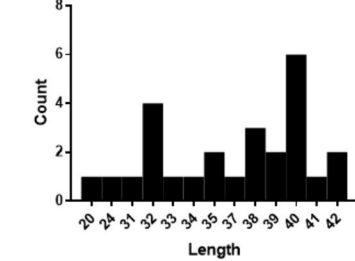


Figure 2. Heterogeneity of A β proteoforms in the human Alzheimer's disease brain. **(A) Top:** full protein sequence of amyloid precursor protein with A β 1–42 sequence highlighted (grey). **Left:** Identified A β proteoform sequence names, with blue bars representing the alignment with canonical A β . **Right:** Heatmap and quantification of the log₂ relative intensity of each A β proteoform for each of the six severe Alzheimer's disease brain samples in the soluble aggregate and more insoluble fractions. **Far right:** Peptide score, indicated confidence of peptide identification³⁸. **(B)** Frequency of A β proteoform amino acid start positions identified in CDR3 cohort; 1 = D, aspartic acid. **(C)** Frequency of A β proteoform amino acid end positions identified in CDR3 cohort; 42 = A, alanine. **(D)** Frequency of A β proteoform lengths identified in CDR3 cohort. Pt., Participant; HMW, high molecular weight; N.D., not detected; Ox., Oxidation; Pyro., Pyro-glutamate; Amm., ammonium ion.

previously reported. No reagents in our isolation protocol contained ammonia. Many of the other forms, including pyro-glutamate modification of amino acid 11 and truncated A β 4–42, were of high relative abundance similar to canonical A β 1–42. One caveat to cross-proteoform comparison is that the relative abundance in the mass spectrometry datasets is largely dependent on the ionization efficiency of various A β proteoforms. The ionization efficiency can be altered by sequence truncations and/or PTMs. Thus, the most reliable quantitative comparisons can only be made for a given proteoform between patients or fractions in this study.

Differential mass spectrometry (dMS) analysis revealed no statistically significant differences in A β proteoform abundance of the 26 A β proteoforms identified using two independent platforms, between soluble aggregates and the more insoluble material (Fig. 4; Supplementary Table 3)²⁵. However, it did uncover trends in regard to the types of proteoforms in each fraction. N-terminally truncated proteoforms, where the hydrophilic region of the peptide is removed leaving the mid-domain and hydrophobic C-terminus, were more enriched in the insoluble fraction relative to the soluble aggregates²⁶. In contrast, enriched species in the soluble A β aggregate fraction were proteoforms with nearly the entire hydrophobic C-terminus truncated. One proteoform, A β 11–34, contained both N- and C-terminal truncations, which has not been described previously. A β 1–34 was 5.6-fold more abundant in the soluble aggregates fraction than in the more insoluble material (Fig. 4, Supplementary Fig. 6). A β 1–20 and A β 11–34, both previously reported in the CSF^{22,27}, were 4- and 2.9-fold more abundant in the more soluble material, respectively (Supplementary Figs 7 and 8). Notably, A β 1–42 and A β 1–40 appeared to be the least significantly differentially expressed between the two fractions, remaining well within the ± 2 -fold change threshold (Fig. 4).

Next we sought to determine if the post-mortem interval (PMI) could affect A β proteoform truncations and PTMs (Supplementary Fig. 9). The signal intensity (*i.e.*, relative abundance) of A β 2–42 and A β 3–42 showed a negative correlation in the soluble aggregates as a function of PMI and no significant effect in the more insoluble aggregates. This is the opposite of the increase that would be expected if these truncations were an artefact of PMI.

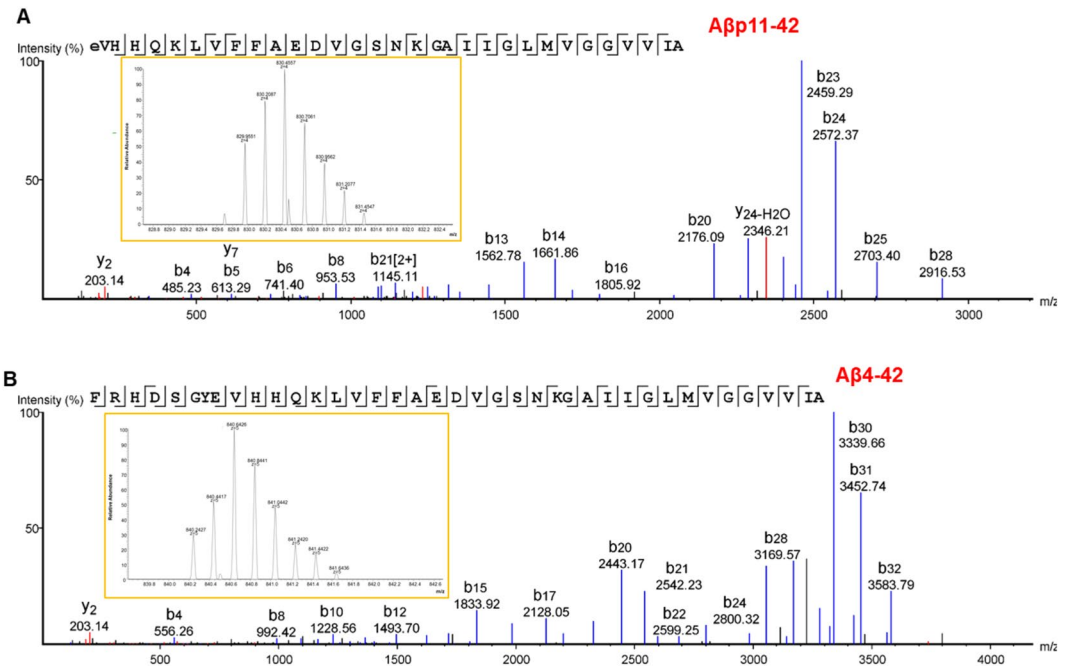


Figure 3. Mass spectrometry of A β proteoforms from human AD brain. **(A)** nano-Liquid chromatography-tandem mass spectrometry (nLC-MS/MS) spectrum for undigested, full-length A β p11–42. Each peak represents an ion fragmented from A β p11–42, with peaks labeled ‘b’ representing N-terminal fragment ions and peaks labeled ‘y’ representing C-terminal fragment ions. The numbers indicate measured mass-charge ratio (m/z). The single letter amino acid code across the top indicates the *de novo* sequence identified by mass spectrometry, which matches the amyloid precursor protein sequence corresponding to A β p11–42. The line breaks between amino acids indicate a cleavage of the amide bond between two adjacent amino acids producing fragment ions. The lines below each amino acid indicate a detected ‘b’ ion, and lines above indicate a detected ‘y’ ion. *Inset:* isotopic envelope for the +4 charged, full-length A β p11–42; the peaks are spaced 0.25 daltons apart at $z = +4$ because the naturally occurring isotopes (e.g. ^{13}C and ^{15}N) differ by 1 dalton. For the +4 ion, the observed m/z was 829.9551 (theoretical $m/z = 829.9523$), which was 3.4 parts-per-million (ppm) error from the theoretical mass of A β p11–42. **(B)** Spectrum for full length A β 4–42. For the +5 ion, the observed m/z was 840.2427 (theoretical $m/z = 840.2399$), which was 3.3 ppm error from the theoretical mass of A β 4–42.

The N-terminal truncations A β 4–42, A β 4–40, A β 5–42 Amm., A β 5–42, A β 8–42, and A β 8–40 all demonstrated no significant correlation with PMI in either the soluble or more insoluble aggregates (Supplementary Fig. 9N–U). The C-terminally truncated proteoforms found with high relative expression in the soluble aggregates – A β 1–34, A β 1–20, and A β 11–34 – demonstrated no significant correlation with PMI (Supplementary Fig. 9A,G,V). Proteoform A β 1–31 also showed no significant correlation. Intriguingly, the truncated proteoform A β 11–42, in both the un-oxidized and oxidized state, and as well as A β p11–42 showed a high positive correlation in the more insoluble aggregates with PMI, but a high to moderate negative trend in the soluble aggregates (Supplementary Fig. 9Y,Z). Given that the trend for increased relative abundance of A β 11–42 occurs irrespective of the oxidation status of methionine 35, we postulate that this may be due to the N-terminal truncation rather than oxidation. This may reflect residual BACE1 activity in the brain even after death²⁸ for which neither our standard inhibitors (aprotinin and leupeptin) nor the protease cocktail, would block or inactivate (Supplementary Fig. 2). However, these results should be interpreted with caution due to the small sample size; only two participants had PMI of ≥ 12 hours.

To summarize, we have demonstrated that human brain soluble A β aggregates consist of an extraordinarily heterogeneous population of A β proteoforms. Other human brain proteins such as alpha-synuclein have not been reported to contain such substantial heterogeneity²⁹, though tau post-translational modifications are protean in transgenic mouse brains³⁰. Our results provide the first direct information regarding the proteoform constituents of soluble A β aggregates in human AD brain of which we are aware. Soluble A β aggregates may be important therapeutic targets, as they are believed to be the most toxic form of A β . Notably, human brain soluble A β aggregates have been reported to be substantially more toxic than similarly sized synthetic A β aggregates^{31,32}. These results suggest the hypothesis that previously unrecognized A β proteoforms may contribute to the modulation of aggregate toxicity in AD. However, these results should be interpreted with caution as some of these A β proteoforms have also been detected in cognitively intact individuals with A β and tau pathology³³. The lack of control (*i.e.*, unaffected) samples in this study limits the interpretation of these findings. The specificity of these A β proteoforms cannot be determined through comparison to similarly prepared non-AD brain or early stage AD samples. Clearly, a great deal of additional investigation with a larger cohort will be required to determine the clinical relevance, aggregation properties, and toxicity of the A β proteoforms described here. Cross-sectional study of A β proteoforms to map disease specificity (control *versus* AD), disease progression specificity (control

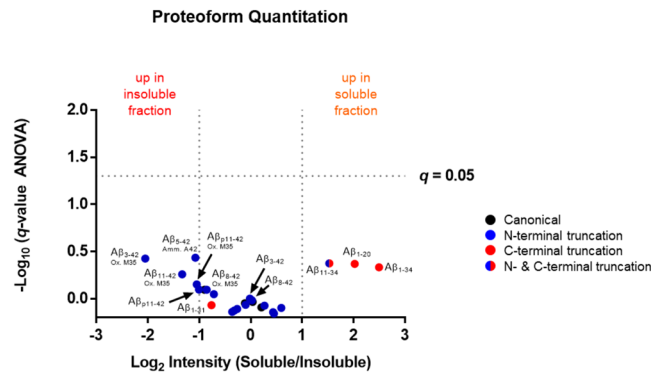


Figure 4. Determination of significant differentially expressed A β proteoforms. Logarithmic ratios of soluble versus insoluble fractions are plotted against the negative logarithmic q -values (q is a false discovery rate corrected p -value, where $q \leq 0.05$ is considered statistically significant; above horizontal line) of the hierarchical linear model with replicates nested within participants and participants treated as random effects to test the fixed effect difference between soluble and insoluble aggregates. Vertical lines designate the logarithmic ratio indicating ± 2 -fold change. All A β proteoforms are listed in Supplementary Table 3 with their respective q -values and \log_2 fold change. Ox., Oxidation; Amm, ammonium ion; M, methionine; A, alanine.

versus mild versus severe AD) is the next logical step, as this will inform future efforts in animal models of the disease. The extent to which the most relevant A β proteoforms are recapitulated in animal models of AD-related pathology remains to be determined and may be of great importance for preclinical therapeutic development.

Methods and Materials

Regulatory Compliance.

All protocols were carried out in accordance with the Charles F. and Joanne Knight Alzheimer's Disease Research Center and Washington University guidelines. This specific study was approved by the Knight Alzheimer's Disease Research Center tissue committee. All donors or their surrogates gave informed consent for their brains to be used for research studies.

Human Tissue.

Clinically and neuropathologically well-characterized human brain tissue samples were obtained from the Charles F. and Joanne Knight Alzheimer's Disease Research Center (knight ADRC), Washington University School of Medicine, Saint Louis, Missouri. At the time of death, informed consent was obtained from the next of kin in accordance with the local Institutional Review Board. Cognitive status at expiration was determined using a validated retrospective post-mortem interview with an informant to establish the Clinical Dementia Rating (CDR)³⁴. We used frozen tissue from the frontal lobe (Brodmann areas 8/9) of severely demented participants with Alzheimer's disease dementia (CDR3, mean age at death = 84 ± 6.6 yrs; $n = 6$; mean post-mortem interval = 11.4 ± 4.8 hours (Table 1). In addition, parietal lobe was used for method development in addition to frontal lobe. Routinely, the right cerebral hemisphere was coronally sliced at 1 cm intervals and frozen by contact with pre-cooled Teflon[®]-coated aluminum plates and temperature equilibrated by immersion into liquid nitrogen vapor in a cryo-vessel. Following freezing, tissues were placed in Ziploc[®] storage bags and stored in freezer at -80°C ³⁵.

Amyloid-beta (A β) Extraction and Separation of Soluble Aggregates from Insoluble Material.

Complete methods are described in Esparza *et al.*²⁰. Briefly, 1–2 g of frozen CDR3 (severe AD) frontal cortical samples, including both gray and white matter, were weighed, stripped of pia mater, leptomeningeal, and intraparenchymal vessels to the fullest extent possible, and dounce homogenized at a 10:1 buffer volume:tissue weight ratio using a constant 25–40 manual strokes. Homogenization buffer consisted of ice-cold 1X phosphate buffered saline (PBS) (137 mM sodium chloride, 7.76 mM sodium phosphate dibasic, 2.17 mM monopotassium phosphate, 2.7 mM potassium chloride) with 0.45% (w/v) (3-((3-cholamidopropyl) dimethylammonio)-1-propanesulfonate) (CHAPS) and 1X protease inhibitor (2 $\mu\text{g}/\text{mL}$ aprotinin and 1 $\mu\text{g}/\text{mL}$ leupeptin). The resulting homogenate was rotated for 15 min at 4°C before centrifugation at $17,000 \times g$ for 30 min in a Sorvall RC 5B centrifuge with a SS-34 rotor. The supernatant was centrifuged in Beckman Optima XPN-100 centrifuge with a Ti70.1 rotor for 60 min at $100,000 \times g$ at 4°C , and the pellet (insoluble material; *see below*) was resolubilized in 5 M guanidine hydrochloride. The supernatant of the $100,000 \times g$ spin was centrifuged for 60 min at $475,000 \times g$ at 4°C with a 70% sucrose cushion. The bottom layers (~ 2 mL) atop of the sucrose cushion (*see Fig. 1a*) were used for analysis of soluble A β aggregates. These methods have been demonstrated to preserve the size distribution of soluble A β aggregates present in PBS lysates, and do not result in artefactual aggregation of monomeric A β ²⁰. All pipet tips and tubes were blocked with 2% bovine serum albumin to significantly reduce non-specific loss of A β .

Purification of Soluble A β Aggregates.

The layers atop the 70% sucrose cushion of the $475,000 \times g$ spin were immunoprecipitated with 100 $\mu\text{L}/\text{mL}$ of a 50% slurry of beads conjugated to the monoclonal antibodies HJ3.4 and HJ5.1 overnight (22 hrs) at 4°C . HJ3.4 binds the N-terminus of canonical A β , and HJ5.1 binds a mid-domain epitope. Beads were washed 15 times in (1 mL each) 1X PBS and eluted with formic acid at room

temperature for 15 min. Total protein and A β concentrations were determined by NanoOrange (Molecular Probes, Eugene, OR) and ELISA, respectively. 5 ng of total A β from each participant sample was dried to completeness *in vacuo*. The 2D Clean-Up Kit (GE Healthcare, Piscataway, NJ) was used to desalt and delipidate the samples. Precipitated samples were subjected to C $_8$ TopTips (Glygen, Columbia, MD) to separate full-length proteins from A β peptides (Supplementary Fig. 1, Supplementary Tables 1 and 2). C $_8$ TopTips were conditioned in 60% ACN/ 0.05% TFA (w/v) and equilibrated in 0.05% TFA (*aq.*, v/v) three times each for 1 min at 2,000 rotations-per-minute (rpm). Next, precipitated samples were resuspended in 80 μ L of neat formic acid and applied to the C $_8$ TopTips and spun for 1 min at 2,000 rpm. The flowthrough was collected and dried to completeness *in vacuo* and stored at -80°C until analysis by nLC-MS/MS.

Purification of Insoluble A β Aggregates. The pellet of the $100,000 \times g$ spin was resolubilized in 5 M guanidine hydrochloride (pH 8.0) overnight at 4°C . The resulting guanidine solubilized 100 k pellet was centrifuged in a MicroCL 17 R centrifuge at $17,000 \times g$ to remove any guanidine insoluble material. Next, the supernatant was diluted 1:10 (0.5 M guanidine, final concentration) in 1X PBS and immunoprecipitated with 100 μ L/mL of a 50% slurry of immobilized monoclonal antibodies HJ3.4 and HJ5.1 overnight (22 hrs) at 4°C . Beads were washed 15 times in 1X PBS (1 mL each) and eluted 3X with 100 μ L formic acid at room temperature for 5 min each. Total protein and A β concentrations were determined by NanoOrange (Molecular Probes, Eugene, OR) and ELISA, respectively²⁰. 5 ng of total A β from each patient sample was dried to completeness *in vacuo*. The 2D Clean-Up Kit (GE Healthcare, Piscataway, NJ) was used to desalt and delipidate the samples. Precipitated samples were subjected to C $_8$ TopTips (Glygen, Columbia, MD) to separate full-length proteins from A β peptides (Supplementary Fig. 1, Supplementary Tables 1 and 2). C $_8$ TopTips were condition in 60% ACN/0.05% TFA (v/v) and equilibrated in 0.05% TFA (*aq.*, v/v) three times each for 1 min at 2,000 rpm. Next, precipitated samples were resuspended in 80 μ L of neat formic acid and applied to the C $_8$ TopTip and spun for 1 min at 2,000 rpm. The flowthrough was collected and dried to completeness *in vacuo* and stored at -80°C until analysis by nLC-MS/MS.

A β ELISA. The A β 1-x ELISA was performed as previously described²⁰. Briefly, mouse monoclonal HJ5.1 (mid-domain) was used to coat 96-well Nunc MaxiSorp plates (464718, Nalge Nunc, Rochester, NY) at 20 μ g/mL in a carbonate buffer (35 mM sodium bicarbonate, 16 mM sodium carbonate, 3 mM sodium azide, pH 9.6) using 100 μ L/well overnight at 4°C . After washing 5x and blocking in 2% bovine serum albumin (BSA, A7030, Sigma-Aldrich, St. Louis, MO) in 1X PBS for 30 min at room temperature, samples and standards (A β 1-40 on a 8-point standard curve) were plated in dilution buffer²⁰ and incubated overnight at 4°C . Samples and standards were developed by incubating in sequence with biotinylated HJ3.4 (canonical N-terminus) at 100 ng/mL for 1 hour at room temperature, streptavidin HRP-20 (65R-S103PHRP, Fitzgerald, Acton, MA) for 30 min at room temperature, and finally 3, 3', 5, 5' - tetramethylbenzidine (TME) (T5569, Sigma-Aldrich, St. Louis, MO) for measurement on a BioTek Synergy 2 plate reader at 650 nm as previously described²⁰.

Protease Inhibitor Control. While the truncation seen is not likely a post-mortem artifact, we used only two protease inhibitors in our homogenization buffer instead of a cocktail of inhibitors, which would protect against a broader spectrum of protease activities. To evaluate the possible impact of using only two protease inhibitors in the truncation profile of A β , we analyzed a representative sample (Pt1) using our standard inhibitors (aprotinin and leupeptin) or Halt Protease and Phosphatase Inhibitor Cocktail (78443, Thermo Fisher Scientific) at 1X (final concentration). Three grams of frontal tissue was diced (on ice) and split evenly for homogenization in either the standard buffer or the cocktail buffer and processed for insoluble and soluble fractions as described above (Supplementary Fig. 2). Mass spectrometry analysis we performed as described below. Spider search results with a score ($-10\log P$) of 31.7 or higher (an estimated FDR value of 2.3% at the peptide level). Only truncated proteoforms were considered for further analysis.

nLC-MS/MS. Samples from each individual CDR3 participant ($n = 6$, biological replicates) across the HMW soluble and insoluble A β fractions were split into two technical replicates (of 5 ng total A β as measured by ELISA) with the exception of Participant 4 (Pt4), which did not have enough material for a second technical replicate. All samples were prepared (precipitation and C $_8$ SPE) in two block-randomized sets (<https://www.random.org>)³⁶ for each set of participant replicates across two days. Samples were resuspended in 1%/10%/5% FA/ACN/MeOH (v/v) and analyzed by non-stop nLC-MS/MS in a block-randomized fashion³⁶ - to eliminate systemic bias due to run order - for a total of 22 Thermo.raw files on a LTQ-Orbitrap Fusion (Thermo Fisher Scientific). Throughout data acquisition, quality assurance/quality control (QA/QC) standard of 20 fmoles BSA (#P8108S, New England BioLabs, Ipswich, MA) was injected every 12-18 hrs to monitor instrument drift and variability over time. 29 BSA peptides from each interwoven QA/QC run were analyzed with AutoQC for instrument performance metrics³⁷ (Supplementary Fig. 3).

Separations were performed using an online NanoAcquity UPLC (Waters). The chromatographic separation was performed on an ACQUITY UPLC HSS T3 (360 μ m OD \times 75 μ m ID) column packed with 10 cm C $_{18}$ (1.8 μ m, 100 \AA , Waters) at 300 nL/min and heated to 60°C . Mobile phases were 0.1% FA in water (A) and 0.1% FA in ACN (B). Samples were eluted from the column with the gradient ramped to 35% B over 65 min and further increased to 95% B over 8 min and held for an additional 6 min. Total run time, including column equilibration, sample loading, and analysis was 89 min. The mass spectrometer was operated in data-dependent mode to automatically switch between MS and MS/MS acquisition. The survey scans at mass-charge ratio (m/z) 400-2000 (MS) were acquired in the Orbitrap at high resolution (60,000 at m/z 400) in profile mode, and the MS/MS spectra were acquired in the Orbitrap (15,000 at m/z 400) in centroid mode using XCalibur, version 3.0 (Thermo Fisher Scientific). Ion injection times for the MS and MS/MS scans were 500 ms each. The automatic gain control targets were set as 2×10^5 for MS and MS/MS in the Orbitrap. The most abundant precursor ions from each MS scan

were sequentially isolated and fragmented in the Orbitrap using HCD (isolation width 2.0 Da, normalized collision energy 30%, activation Q 0.250, and activation time 10 ms) within a 3 sec duty cycle (TopSpeed method). Dynamic exclusion (± 10 ppm relative to precursor ion m/z) was enabled with a repeat count of one, a maximal exclusion list size of 500, and an exclusion duration of 60 s. Monoisotopic precursor selection (MIPS) was enabled and unassigned ions were rejected.

Mass Spectrometry Data Processing. MS files (.raw) were imported into PEAKS (version 8, Bioinformatics Solutions Inc., Waterloo, ON) and searched against a UniprotKB/SwissProt Human database of reviewed, canonical sequences (October 2015; 20,204 entries) appended with the cRAP contaminant database (January 2015 version, The Global Proteome Machine, www.thegpm.org/cRAP/index.html). Precursor ion mass tolerance was set to 10 ppm, and fragment mass tolerance was 0.1 Da with no enzyme specificity. All modifications in the UniMod database (<http://www.unimod.org>) were considered in the PEAKS search. PEAKS automatically generates a decoy-fusion database, which appends a decoy sequence to each protein identification for the calculation of FDR³⁸. The Spider search results with a score ($-10\log P$) of 31.9 or higher (an estimated FDR value of 2.9% at the peptide level) for the CDR3 cohort data. Spider is an algorithm tool within PEAKS, which we utilized to search peptide spectrum matches not identified by the database search by altering the amino acids systematically at each residue until a new, better peptide sequence is constructed from the MS/MS data³⁹. All non-A β peptides and those A β proteoforms containing formylation were removed. Immunoprecipitation elution was in neat formic acid (Sigma #94318); thus, formylation occurring endogenously or by the elution conditions is indistinguishable. Further, samples were resuspended in MS buffer containing 5% methanol (v/v) to maintain solubility of A β , which in combination with 1% FA could lead to artefactual methylation via Fischer esterification.

In the CDR3 cohort, 27 A β proteoforms were identified meeting the criteria described above. However, we sought reproducible identification of the 27 A β proteoforms on an independent platform. The filtered list of A β proteoforms were *i*) assigned identification numbers by a third party who had no part in the initial analysis and *ii*) transformed into chemical formulas. This chemical formula list was queried against each .raw file with National Resource for Top-Down Proteomics (NRTDP) pipeline version 1.3 at ± 10 ppm precursor ion mass tolerance and a 4 min retention time window alignment (Supplementary Fig. 4). If a given proteoform was not identified by both platforms (PEAKS and NRTDP), it was removed from analysis; 26 out of 27 were reproducibly identified on both platforms. MS intensity was calculated across all peaks within an isotopic cluster extracted with Skyline for each proteoform identified in Fig. 2 and exported as an Excel file.

Differential Mass Spectrometry (dMS). In this method relative quantitation is performed from the full-scan MS. The high-resolution MS provides m/z , retention time, charge state, and relative abundance (intensity) of precursor ion across multiple samples for comparative analysis^{40–43}. A table of proteoform intensities was imported into a custom SAS script for analysis. To detect differential (label-free) abundance from the intensities data⁴⁰, the MS intensity measures (calculated as described above) for each proteoform were standardized to Z-scores across all measures of that proteoform. Different charge states were treated as independent measures of the same proteoform as reflected in the increased degrees of freedom for certain tests (Supplementary Table 3).

ANOVA based on a hierarchical linear model (HLM) with replicates nested within patients, and patients treated as random effects, was used to test the fixed effect difference between soluble aggregate A β and more insoluble A β fractions. All calculations were done using SAS PROC MIXED with restricted maximum likelihood estimations (SAS Institute, Cary, NC) and type 3 sums of squares (where appropriate). The HLM was used to test for differences in mean intensity between CDR3 soluble and insoluble A β fractions, while allowing each biological replicate to have its own overall mean. Each p -value of the resulting 26 F tests was corrected for multiple testing (q -value), and those with an FDR of ≤ 0.05 were considered significant⁴⁴. Next, the same model was run on the \log_2 -converted raw intensities. The difference in estimated mean between aggregate A β and more insoluble A β fractions in these tests was taken as an estimate of the overall fold change within the fractions. Thus two separate ANOVA analyses are run, the first to test the statistical significance of abundances of proteoforms between fractions, and the second to estimate effect size^{25, 45}. The mass spectrometric data have been deposited in ProteomeXchange (<http://proteomecentral.proteomexchange.org>) via the PRIDE partner repository⁴⁶ with the data set identifier PXD005119.

Data Visualization. Analyses were performed and visualized with Excel 2013, PRISM (version 7) for correlation analysis of signal intensity and PMI and graphing volcano plots or SAS (version 9.4) for hierarchical linear modeling.

References

- Hardy, J. *et al.* Pathways to Alzheimer's disease. *Journal of internal medicine* **275**, 296–303, doi:10.1111/joim.12192 (2014).
- Hardy, J. A. & Higgins, G. A. Alzheimer's disease: the amyloid cascade hypothesis. *Science* **256**, 184–185 (1992).
- Smith, L. M., Kelleher, N. L. & Proteomics, C. F. T. D. Proteoform: a single term describing protein complexity. *Nat Methods* **10**, 186–187, doi:10.1038/nmeth.2369 (2013).
- Mori, H., Takio, K., Ogawara, M. & Selkoe, D. J. Mass spectrometry of purified amyloid beta protein in Alzheimer's disease. *J Biol Chem* **267**, 17082–17086 (1992).
- Saido, T. C. *et al.* Dominant and differential deposition of distinct beta-amyloid peptide species, A beta N3(pE), in senile plaques. *Neuron* **14**, 457–466 (1995).
- Kuo, Y. M., Emmerling, M. R., Woods, A. S., Cotter, R. J. & Roher, A. E. Isolation, chemical characterization, and quantitation of A beta 3-pyroglyutamyl peptide from neuritic plaques and vascular amyloid deposits. *Biochem Biophys Res Commun* **237**, 188–191, doi:10.1006/bbrc.1997.7083 (1997).
- Roher, A. E., Palmer, K. C., Yurewicz, E. C., Ball, M. J. & Greenberg, B. D. Morphological and biochemical analyses of amyloid plaque core proteins purified from Alzheimer disease brain tissue. *J Neurochem* **61**, 1916–1926 (1993).

8. Roher, A. E. *et al.* Structural alterations in the peptide backbone of beta-amyloid core protein may account for its deposition and stability in Alzheimer's disease. *J Biol Chem* **268**, 3072–3083 (1993).
9. Portelius, E. *et al.* Mass spectrometric characterization of brain amyloid beta isoform signatures in familial and sporadic Alzheimer's disease. *Acta Neuropathol* **120**, 185–193, doi:10.1007/s00401-010-0690-1 (2010).
10. Selkoe, D. J. Alzheimer's disease: a central role for amyloid. *J Neuropathol Exp Neurol* **53**, 438–447 (1994).
11. Terry, R. D. *et al.* Physical basis of cognitive alterations in Alzheimer's disease: synapse loss is the major correlate of cognitive impairment. *Ann Neurol* **30**, 572–580, doi:10.1002/ana.410300410 (1991).
12. Dickson, D. W. *et al.* Correlations of synaptic and pathological markers with cognition of the elderly. *Neurobiol Aging* **16**, 285–298; discussion 298–304 (1995).
13. Selkoe, D. J. & Hardy, J. The amyloid hypothesis of Alzheimer's disease at 25 years. *EMBO Mol Med* **8**, 595–608, doi:10.15252/emmm.201606210 (2016).
14. McLean, C. A. *et al.* Soluble pool of Abeta amyloid as a determinant of severity of neurodegeneration in Alzheimer's disease. *Ann Neurol* **46**, 860–866 (1999).
15. Lue, L. F. *et al.* Soluble amyloid beta peptide concentration as a predictor of synaptic change in Alzheimer's disease. *Am J Pathol* **155**, 853–862 (1999).
16. Shankar, G. M. *et al.* Amyloid-beta protein dimers isolated directly from Alzheimer's brains impair synaptic plasticity and memory. *Nat Med* **14**, 837–842, doi:10.1038/nm1782 (2008).
17. Walsh, D. M. & Selkoe, D. J. A beta oligomers - a decade of discovery. *J Neurochem* **101**, 1172–1184, doi:10.1111/j.1471-4159.2006.04426.x (2007).
18. Tomic, J. L., Pensalfini, A., Head, E. & Glabe, C. G. Soluble fibrillar oligomer levels are elevated in Alzheimer's disease brain and correlate with cognitive dysfunction. *Neurobiol Dis* **35**, 352–358, doi:10.1016/j.nbd.2009.05.024 (2009).
19. Esparza, T. J. *et al.* Amyloid- β oligomerization in Alzheimer dementia versus high-pathology controls. *Ann Neurol* **73**, 104–119, doi:10.1002/ana.23748 (2013).
20. Esparza, T. J. *et al.* Purification and Quantitative Characterization of Soluble HMW Amyloid-beta from Alzheimer's disease Brain Lysates. *Manuscript in Preparation Scientific Reports* **6**, 38187, doi:10.1038/srep38187 (2016).
21. Brody, D. L. *et al.* Amyloid-beta dynamics correlate with neurological status in the injured human brain. *Science* **321**, 1221–1224, doi:10.1126/science.1161591 (2008).
22. Kummer, M. P. & Heneka, M. T. Truncated and modified amyloid-beta species. *Alzheimers Res Ther* **6**, 28, doi:10.1186/alzrt58 (2014).
23. Schieb, H. *et al.* Beta-amyloid peptide variants in brains and cerebrospinal fluid from amyloid precursor protein (APP) transgenic mice: comparison with human Alzheimer amyloid. *J Biol Chem* **286**, 33747–33758, doi:10.1074/jbc.M111.246561 (2011).
24. Patterson, B. W. *et al.* Age and amyloid effects on human central nervous system amyloid-beta kinetics. *Ann Neurol*. doi:10.1002/ana.24454 (2015).
25. Ntai, I. *et al.* Applying label-free quantitation to top down proteomics. *Anal Chem* **86**, 4961–4968, doi:10.1021/ac500395k (2014).
26. Pike, C. J., Overman, M. J. & Cotman, C. W. Amino-terminal deletions enhance aggregation of beta-amyloid peptides *in vitro*. *J Biol Chem* **270**, 23895–23898 (1995).
27. Portelius, E., Westman-Brinkmalm, A., Zetterberg, H. & Blennow, K. Determination of beta-amyloid peptide signatures in cerebrospinal fluid using immunoprecipitation-mass spectrometry. *J Proteome Res* **5**, 1010–1016, doi:10.1021/pr050475v (2006).
28. Vassar, R. *et al.* Beta-secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE. *Science* **286**, 735–741 (1999).
29. Kellie, J. F. *et al.* Quantitative measurement of intact alpha-synuclein proteoforms from post-mortem control and Parkinson's disease brain tissue by intact protein mass spectrometry. *Sci Rep* **4**, 5797, doi:10.1038/srep05797 (2014).
30. Morris, M. *et al.* Tau post-translational modifications in wild-type and human amyloid precursor protein transgenic mice. *Nat Neurosci* **18**, 1183–1189, doi:10.1038/nn.4067 (2015).
31. Noguchi, A. *et al.* Isolation and characterization of patient-derived, toxic, high mass amyloid beta-protein (Abeta) assembly from Alzheimer disease brains. *The Journal of biological chemistry* **284**, 32895–32905, doi:10.1074/jbc.M109.000208 (2009).
32. Jin, M. *et al.* Soluble amyloid beta-protein dimers isolated from Alzheimer cortex directly induce Tau hyperphosphorylation and neuritic degeneration. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 5819–5824, doi:10.1073/pnas.1017033108 (2011).
33. Sergeant, N. *et al.* Truncated beta-amyloid peptide species in pre-clinical Alzheimer's disease as new targets for the vaccination approach. *J Neurochem* **85**, 1581–1591 (2003).
34. Berg, L. Clinical Dementia Rating (CDR). *Psychopharmacol Bull* **24**, 637–639 (1988).
35. Cairns, N. J. *et al.* Neuropathologic assessment of participants in two multi-center longitudinal observational studies: the Alzheimer Disease Neuroimaging Initiative (ADNI) and the Dominantly Inherited Alzheimer Network (DIAN). *Neuropathology* **35**, 390–400, doi:10.1111/neup.12205 (2015).
36. Oberg, A. L. & Vitek, O. Statistical design of quantitative mass spectrometry-based proteomic experiments. *J Proteome Res* **8**, 2144–2156, doi:10.1021/pr801009y (2009).
37. Bereman, M. S. *et al.* An Automated Pipeline to Monitor System Performance in Liquid Chromatography-Tandem Mass Spectrometry Proteomic Experiments. *J Proteome Res*, doi:10.1021/acs.jproteome.6b00744 (2016).
38. Zhang, J. *et al.* PEAKS DB: de novo sequencing assisted database search for sensitive and accurate peptide identification. *Mol Cell Proteomics* **11**(M111), 010587, doi:10.1074/mcp.M111.010587 (2012).
39. Han, Y., Ma, B. & Zhang, K. SPIDER: software for protein identification from sequence tags with de novo sequencing error. *J Bioinform Comput Biol* **3**, 697–716 (2005).
40. Wiener, M. C., Sachs, J. R., Deyanova, E. G. & Yates, N. A. Differential mass spectrometry: a label-free LC-MS method for finding significant differences in complex peptide and protein mixtures. *Anal Chem* **76**, 6085–6096, doi:10.1021/ac0493875 (2004).
41. Mazur, M. T. *et al.* Quantitative analysis of intact apolipoproteins in human HDL by top-down differential mass spectrometry. *Proc Natl Acad Sci USA* **107**, 7728–7733, doi:10.1073/pnas.0910776107 (2010).
42. Bildl, W. *et al.* Extending the dynamic range of label-free mass spectrometric quantification of affinity purifications. *Mol Cell Proteomics* **11**(M111), 007955, doi:10.1074/mcp.M111.007955 (2012).
43. Ntai, I., Toby, T. K., LeDuc, R. D. & Kelleher, N. L. A Method for Label-Free, Differential Top-Down Proteomics. *Methods Mol Biol* **1410**, 121–133, doi:10.1007/978-1-4939-3524-6_8 (2016).
44. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)* **57**, 289–300 (1995).
45. Wildburger, N. C. *et al.* Quantitative proteomics reveals protein-protein interactions with fibroblast growth factor 12 as a component of the voltage-gated sodium channel 1.2 (nav1.2) macromolecular complex in Mammalian brain. *Mol Cell Proteomics* **14**, 1288–1300, doi:10.1074/mcp.M114.040055 (2015).
46. Vizcaino, J. A. *et al.* The PRoteomics IDentifications (PRIDE) database and associated tools: status in 2013. *Nucleic Acids Res* **41**, D1063–1069, doi:10.1093/nar/gks1262 (2013).

Acknowledgements

We would like to thank the participants in the Knight ADRC who donated their brains for the study, David Holtzman and Hong Jiang for antibodies, and Tom Ellenberger for allowing us to use his ultracentrifuge. This work was supported by the Jeane B. Kempner PostDoctoral Fellowship (NCW), NIH R01 NS065069 (DLB), the Cure Alzheimer's Fund (DLB), Coins for Alzheimer's Research Trust (NJC), Alzheimer's Drug Discovery Foundation (NJC), BrightFocus Foundation (DLB), NIH R01 NS065667 (RJB), Washington University School of Medicine (RJB), and the Hope Center for Neurological Diseases. This research was carried out in collaboration with the National Resource for Translational and Developmental Proteomics under Grant P41 GM108569 from the National Institute of General Medical Sciences, National Institutes of Health. We thank Michael Gross, John Cirrito, David Holtzman, Amos Heckendorf, Cheryl Lichti, and Nicolas Barthélemy for advice and helpful discussions. Human brain samples were provided by the Washington University Knight Alzheimer's Disease Research Center (P50 AG05681 and P01 AG3991).

Author Contributions

N.C.W. and D.L.B. designed research. T.J.E. performed all fractionations and I.P.s. N.C.W. prepared all samples for top-down mass spectrometry, performed all mass spectrometry experiments, and mass spectrometry data analysis. N.C.W., R.D.L., and R.T.F. performed mass spectrometry data analysis. N.J.C. provided brain tissue, pathological assessments, and critically revised the manuscript. N.C.W. prepared the figures. N.C.W. and D.L.B. wrote the paper with intellectual contributions from all authors. R.J.B. and D.L.B. supervised the work and critically revised the manuscript. N.C.W., R.J.B. and D.L.B. obtained funding. N.L.K. and P.M.T. provided top-down expertise and R.J.B. provided the Orbitrap FusionTM. All authors read and approved the final manuscript.

Additional Information

Supplementary information accompanies this paper at doi:[10.1038/s41598-017-10422-x](https://doi.org/10.1038/s41598-017-10422-x)

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017