

Using in vivo intact structure for system-wide quantitative analysis of changes in proteins

Received: 19 February 2024

Accepted: 16 October 2024

Published online: 29 October 2024

Ahrum Son¹, Hyunsoo Kim^{1,2}, Jolene K. Diedrich¹, Casimir Bamberger¹,
Daniel B. McClatchy¹, Stuart A. Lipton^{1,3,4} & John R. Yates III¹ ✉

Mass spectrometry-based methods can provide a global expression profile and structural readout of proteins in complex systems. Preserving the in vivo conformation of proteins in their innate state is challenging during proteomic experiments. Here, we introduce a whole animal in vivo protein footprinting method using perfusion of reagents to add dimethyl labels to exposed lysine residues on intact proteins which provides information about protein conformation. When this approach is used to measure dynamic structural changes during Alzheimer's disease (AD) progression in a mouse model, we detect 433 proteins that undergo structural changes attributed to AD, independent of aging, across 7 tissues. We identify structural changes of co-expressed proteins and link the communities of these proteins to their biological functions. Our findings show that structural alterations of proteins precede changes in expression, thereby demonstrating the value of in vivo protein conformation measurement. Our method represents a strategy for untangling mechanisms of proteostasis dysfunction caused by protein misfolding. In vivo whole-animal footprinting should have broad applicability for discovering conformational changes in systemic diseases and for the design of therapeutic interventions.

The proper functioning of cellular machinery depends on the ability to maintain the functional structures of proteins. Proper folding of proteins is necessary to engage with partners in complexes and to perform catalytic activities. Protein folds or shapes can be measured by powerful, high-resolution ex vivo techniques such as X-ray crystallography, NMR, and Cryo-electron microscopy (Cryo-EM)^{1–5}. Cryo-EM can be used to analyze large protein complexes if they are extracted from cells or are produced recombinantly prior to deposition on the grid and frozen^{6–8}. Multiplexed Ion Beam imaging (MIBI) and ion beam tomography are capable of imaging cells and tissues, but they are not explicitly used to study the structure of proteins and protein complexes. Modeling algorithms can generate protein structures from ex-vivo protein cross-linking data, while in vivo cross-linking analyses generate protein-protein interaction data. Thus, because no methods are available to determine the high-resolution structures of multiple

proteins in vivo, we are still limited in our ability to elucidate the structures of proteins at a proteome-wide scale in the cellular milieu.

Protein “footprinting” methods were developed to probe the folding and interactions of proteins (such as epitope sites in antigens) using protease restriction or covalent labeling to identify exposed regions of proteins^{9,10}. The data generated in protein footprinting experiments is often low resolution, but the potential scale of experiments has made it an attractive method. In 2010 West et al. used proteome-scale footprinting in *S. cerevisiae* to determine off target binding of cyclosporin¹¹. A variety of protein labeling methods have been developed that provide low resolution ex vivo structural information about proteins^{6,12–15}. Picotti and colleagues developed a limited proteolysis method to map ligand binding and protein folding in cell lysates and biofluids^{6,16–18}. In 2015 Espino et al. used lasers to activate hydroxy radicals in vivo to label proteins, providing the first attempt to

¹Department of Molecular Medicine, The Scripps Research Institute, La Jolla, California, USA. ²Department of Convergent Bioscience and Informatics, Chungnam National University, 99 Daehak-ro, Yuseong-gu, Daejeon, Republic of Korea. ³Neurodegeneration New Medicines Center, The Scripps Research Institute, La Jolla, CA, USA. ⁴Department of Neurosciences School of Medicine University of California, San Diego, La Jolla, CA, USA.

✉ e-mail: jyates@scripps.edu

footprint an intact cell¹⁹. Their approach has now been extended to the transparent worm *C. elegans*, which was chosen so the laser beam could penetrate the worm²⁰. Additionally, there are various amino acid labeling methods available for measuring protein structural changes^{21–23}.

Bamberger et al. developed Covalent Protein Painting (CPP), a chemical approach for quantitative protein footprinting to measure in vivo changes to protein conformations on a proteome scale²⁴. In CPP, formaldehyde, a chemical that rapidly permeates cells and tissues, is used to label proteins by forming a Schiff's base at solvent exposed lysine residues. These unstable intermediates are converted to dimethyl labels by reduction with cyanoborohydride. After lysis of cells or tissue, denaturation and digestion of proteins, a second labeling with a different "weight" reagent is performed to label inaccessible amino acid residues. By using heavy and light isotope versions of the reagents, a quantitative measure of lysine accessibility can be obtained. Using this method, Bamberger et al. probed the conformational changes of a proteome from postmortem brain tissue to reveal structural changes and altered protein-protein interactions in the brain tissue of AD patients²⁴. In another study, Bamberger et al. measured the altered conformations of proteins in 60 cancer cell lines (NCI60)²⁵. Because the CPP protein labeling method begins with the widely used formaldehyde fixation step for in vivo dimethyl labeling, it should be extensible to whole animal labeling to study models of disease.

Methods to measure alterations of protein conformations in vivo are needed to study diseases caused by protein misfoldings, which create loss or gain of function disruptions to biological processes, including Alzheimer's disease, a common misfolding disease that is characterized by plaques of amyloid-beta and tangles of tau proteins. As observed by Bamberger et al., late-stage neurodegenerative diseases in humans are characterized by the misfolding of many additional proteins, suggesting that there is a generalized failing of proteostasis²⁴. Techniques that allow in vivo measurement of protein folding would be a powerful tool for the study of these misfolding diseases.

Here, we use AD as a model to test our hypothesis that the global measurement of structural changes of proteins in tissues can be used to understand changes in their biological functionality during progression of protein misfolding diseases. We reason that it is important to capture proteins in their innate states to preserve the complex cellular milieu without the protein degradation that might occur during extraction and homogenization of organs. In this study, we introduce the use of CPP in an animal model to probe dynamic changes in protein structures in vivo. This technique is applied in mouse tissue to study proteome-wide structural changes of proteins in progressing AD. We identified proteins whose structures were altered in co-expressed protein communities across 7 types of mouse tissue, which helps us understand the role of spatially altered proteins in various biological processes.

Results

A comprehensive structural information readout platform for proteome across mouse organs

Our primary goal was to reveal how in vivo alterations of protein conformations across tissues were associated with physiological disturbances that characterize Alzheimer's disease (AD). To prevent protein degradation in response to external stimuli, it was necessary to capture the innate folding state in vivo. Thus, we expanded and applied the CPP method previously used for living cells²⁴ to in vivo dimethyl label proteins in a mouse model by incorporating the strategy of perfusion. We validated the methodology's capacity to discern structural changes in proteins by evaluating the presence of potential confounding factors that might influence the interpretation of the results. These factors include the impact of age-related changes in vascular structure on the permeability of reagents, the potential

influence of protein abundance on labeling efficiency, and the correlation between accessibility obtained from tissue blocks and tissue lysates post-organ extraction and accessibility obtained from labeling via perfusion. Additionally, the agreement between the protein structures measured by our method and the crystal structures was determined. [See Supplementary Notes under subheading 'Validation for the utility of CPP via perfusion' and Supplementary Figs. 1–5] Here, we presented a method to in vivo dimethyl label the body-wide proteome of an intact mouse prior to organ harvest. Our method involved sequentially diffusing labeling reagents through blood vessels throughout whole body so that proteins surfaces were light-dimethylated [$[(CHD_2)_2]$ in vivo. After harvest and homogenization of each tissue and lysis of the cells, proteins were denatured and proteolyzed with chymotrypsin followed by labeling of the newly exposed lysine sites with heavy-dimethyl [$(^{13}CD_3)_2$] tags (Fig. 1).

The accessibility of each lysine site was determined from the ratio of the intensity of light-labeled peptide vs. the sum of intensities of the light- and heavy-labeled peptides. Prior to presenting the results, it is important to outline the scope and objectives of the CPP method to prevent potential misinterpretation. Dimethylation by the CPP method does not occur on lysines that are already modified (i.g. acetylation or ubiquitination). Even for lysines at the same position, the accessibility can vary depending on whether they have undergone prior modifications. However, it is not possible to determine which specific modification hinders the dimethyl labeling. The goal of this method is simply to measure surface accessibility of lysine residues and the labeling process can be effected by PTMs, mutations, or protein-protein interactions that block a site of labeling. It is important to note that while this may influence accessibility results, the overall purpose of the method remains to detect structural changes in proteins, rather than pinpointing the specific causes of these structural changes. Furthermore, since non-natural isotopic dimethylation was used in both initial and secondary dimethylation, naturally occurring dimethylation can be distinguished from dimethylation that occurred during the experiment. This approach aims to provide a quantification of protein structural alterations, regardless of the specific type of PTM or modification involved.

We systematically investigated the structural changes in the proteomes of 7 tissues in an AD mouse model ranging in age from 6 to 15 months, as well as in normal control (NC) mice to exclude the effects of aging. A total of 43,014 dimethyl-labeled peptides that mapped to 5,217 proteins across all tissues were identified at a peptide false discovery rate (FDR) of <1%. Among the labeled peptides, 1219 labeled peptides that mapped to 498 proteins were identified in all 7 tissues, whereas 24,026 labeled peptides that mapped to 4,952 proteins were tissue-specific (Fig. 2A). Likewise, the highest proportion of labeled proteins were tissue-specific (37.3% ($n = 1,947$)), whereas labeled proteins that were identified in all 7 tissues comprised 11.1% ($n = 578$) of all labeled proteins (Fig. 2B).

We next sought to examine the reproducibility of the in vivo CPP method for each mouse tissue. We evaluated the correlations of the accessibility between biological replicates for each age, then averaged these correlations. We found the highest R values for NC (0.842) and AD (0.782) in brain, whereas thymus showed the lowest correlation, 0.549 for NC and 0.505 for AD (Fig. 2C–F, Supplementary Table 1). Correlations averaged across all ages in all 7 tissues showed R values of 0.668 and 0.649 for NC and AD, respectively. Because we observed strong correlations of the biological replicates regardless of anatomical source or pathological conditions of the tissue, we concluded that perfusion-based dimethyl-labeling was reliable. Also, we noticed that even in the presence of different pathological conditions, correlations of accessibility were higher within the same tissues compared to those observed between different tissues at the same age. The labeling efficiency after perfusion labeling ranged from 91.5 to 97.5% depending on the tissue labeled. We conducted an empirical comparison of

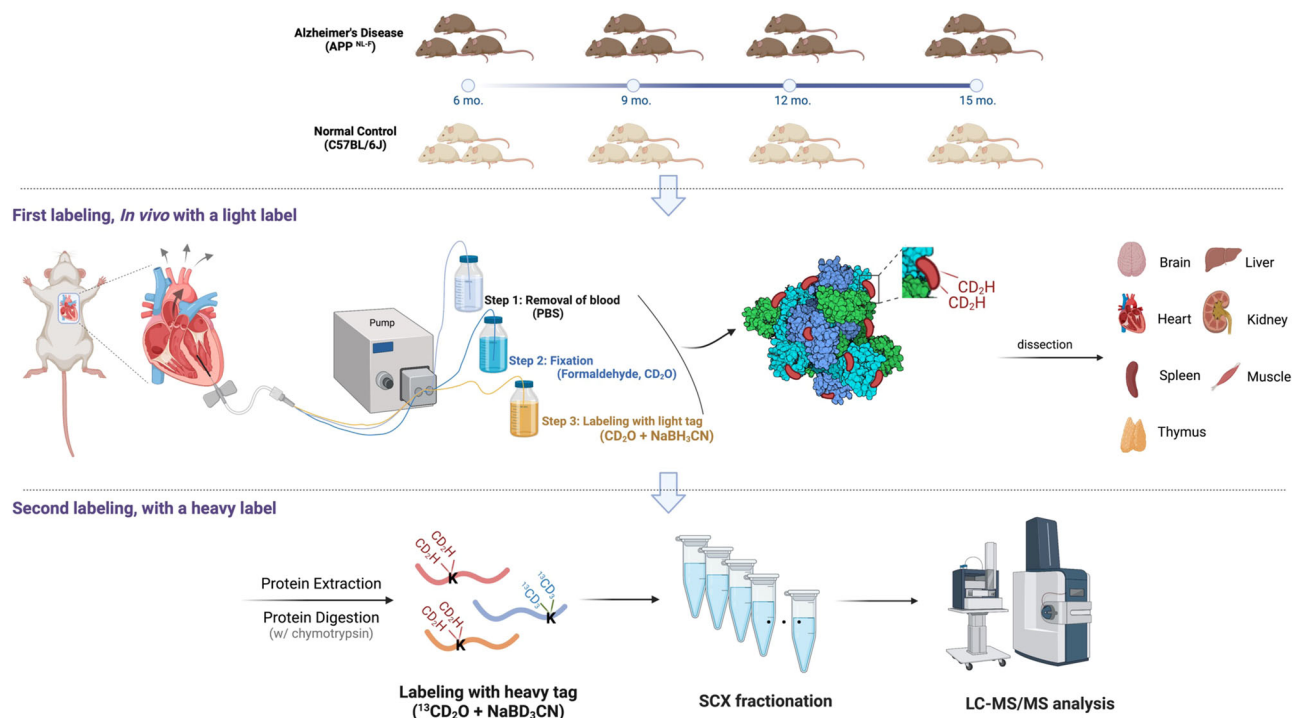


Fig. 1 | Strategy for identifying dimethyl-labeled peptides. Three mice per each age group (ranging 6 months to 15 months) were used for AD (APP^{NL-F}) and NC (C57BL/6J). The first step of the CPP workflow consists of three sub-steps that were conducted via perfusion: (i) blood was washed by PBS, (ii) tissue was fixed by formaldehyde, and (iii) exposed lysine sites of the native proteins were labeled with

light-dimethylation ($[CD_2H]_2$). Proteins from each of the seven organs were extracted and digested separately with chymotrypsin, after which the newly exposed lysine sites were labeled with heavy dimethylation ($[C^{13}D_3]_2$). Created in BioRender. Kim, H. (2022) BioRender.com/d78w669.

dimethylation patterns across tissues and determined that the dimethyl labeling method yielded tissue-specific information, revealing distinct patterns unique to each tissue type examined. [See Supplementary Notes under subheading ‘Dimethylation pattern encoding tissue-specific protein conformation’ and Supplementary Figs. 6–7].

Variability of the conformational changes among 7 tissues

We used labeled peptides that were detected in all seven of the tissues we analyzed to quantitatively measure the structural differences in proteins. To minimize the effect of tissue-biased accessibility, we used quantile-normalized values for comparison across tissues (Supplementary Fig. 8). Each tissue was analyzed separately by mass spectrometry, but since these samples were extracted from the same mouse, quantile normalization was conducted to facilitate comparison between the tissues. The normalized values are used exclusively for tissue comparative analysis. The accessibility value can only be used to compare accessibility under different conditions. Therefore, normalized values were used to analyze long-term changes in accessibility. We sought to identify patterns of changes in accessibility that occur specifically for AD in proteins from 6mo to 15mo. To achieve this, we tested for differences between NC and AD over time using a spline model from 6 to 15 months for each labeled peptide (Fig. 3A). Brain tissue was the most structurally affected by AD, with 686 peptides in AD brain showing significantly different patterns of accessibility compared to NC from 6 to 15 months (Fig. 3B). In heart, kidney, and thymus, fewer than 25 peptides showed significantly different patterns of change in AD relative to NC. There were no labeled peptides that exhibited significant different patterns in accessibility across all tissues from 6 months to 15 months. However, a total of 10 labeled peptides consistently showed significantly different patterns in accessibility across up to four different organs each in the AD model, and there were no peptides showing a common significant change in accessibility in more than five tissues. To examine the conformational

changes of proteins specifically impacted by AD, we corrected for the confounding effect of age by dividing the individual accessibility of AD by the accessibility of its corresponding sequence of NC, resulting in a metric we refer to as “fold-change”. We evaluated the variability in structural changes depending on tissues relative to the brain using 10 peptides that showed distinct patterns of change in accessibility in AD from those in NC in four tissues (Fig. 3C–F, Supplementary Table 2). At the early stage of AD (6 months), the biggest conformational discrepancy due to AD was observed between the liver and brain. The lysine site of TAKGLF (Eno1) was not only more accessible in AD liver than in AD brain, but also more accessible in AD liver than in NC liver. As AD progressed, AD-specific structural changes in muscle and spleen were greater than those in liver when the fold-change between each tissue and the brain was compared. (Fig. 3G–H). A functional enrichment analysis of the 10 proteins retrieved KEGG pathways associated with metabolism, glycolysis, and the TCA cycle (Fig. 3I). For only 5 (51–67 amino acid (AA) of Atp5f1d, 469–479 AA of Dpysl2, 48–57 AA of Eno1, 1472–1483 AA of Flna, and 277–284 AA of Ppp2cb) out of 1219 peptides commonly detected in the 7 tissues, the accessibility patterns in all 7 tissues during AD progression were not significantly different from those during normal aging, and no significant difference in accessibility was observed across the 7 tissues under each condition (4-ages, disease). This indicates that the regions corresponding to these 5 peptides were not affected by AD in any of the 7 tissues. As expected, this investigation confirms that the effects of AD are most significantly observed in brain tissue, but by quantifying conformational changes of proteins across all tissues, we found tissue-specific variations associated with AD.

Conformations of proteins in the brain are changing as AD progresses

Proteomic investigations into AD pathology have primarily relied on the analysis of differential protein expression. Bai et al. profiled the

differentially expressed proteins and identified the protein networks that are affected during AD progression²⁶. Co-expressed proteins and altered protein expression in human brain tissue of asymptomatic and

symptomatic AD patients were reported by Seyfried et al.²⁷. Savas et al. measured protein expression in several mouse models of AD using quantitative mass spectrometry²⁸. Despite extensive proteomic studies

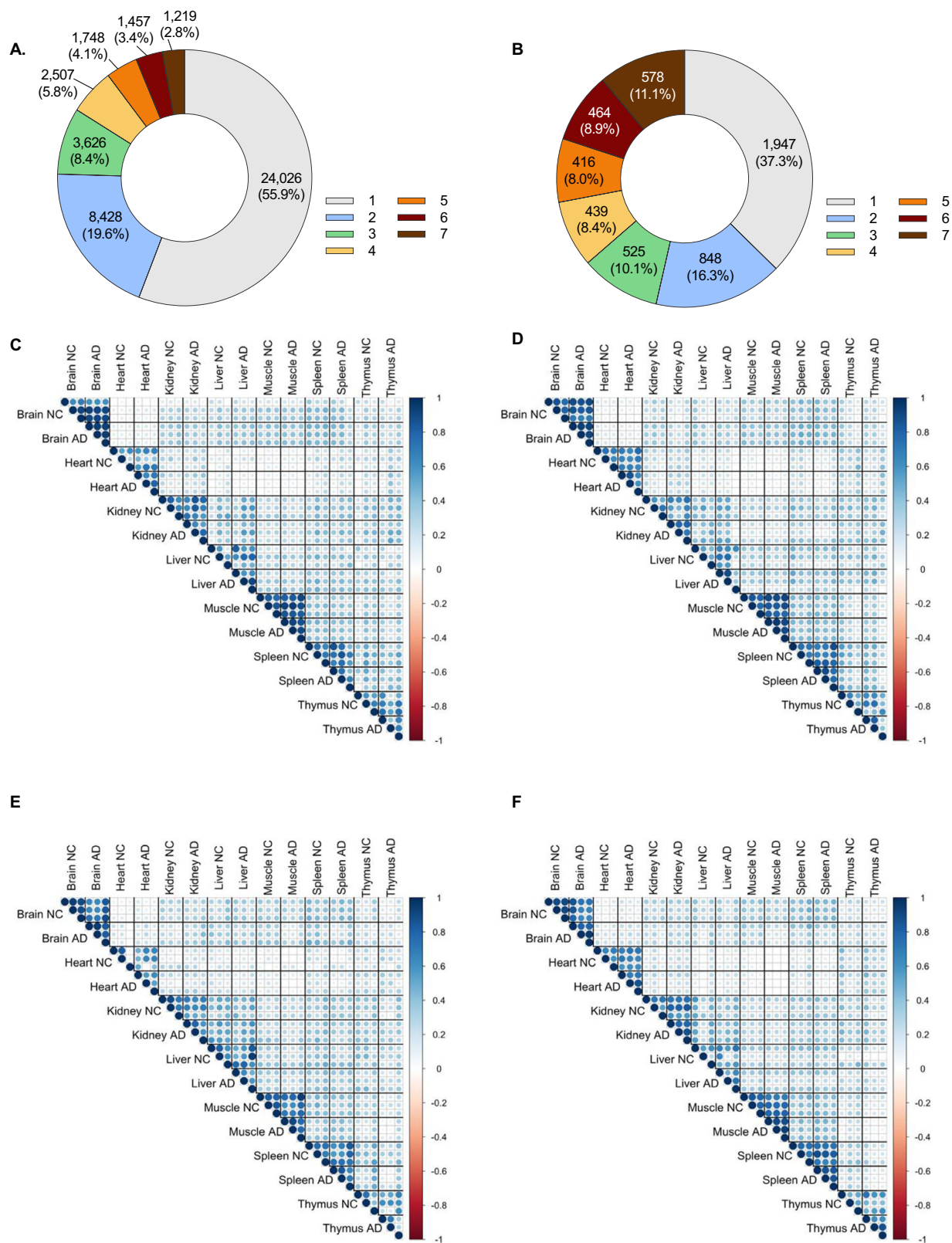
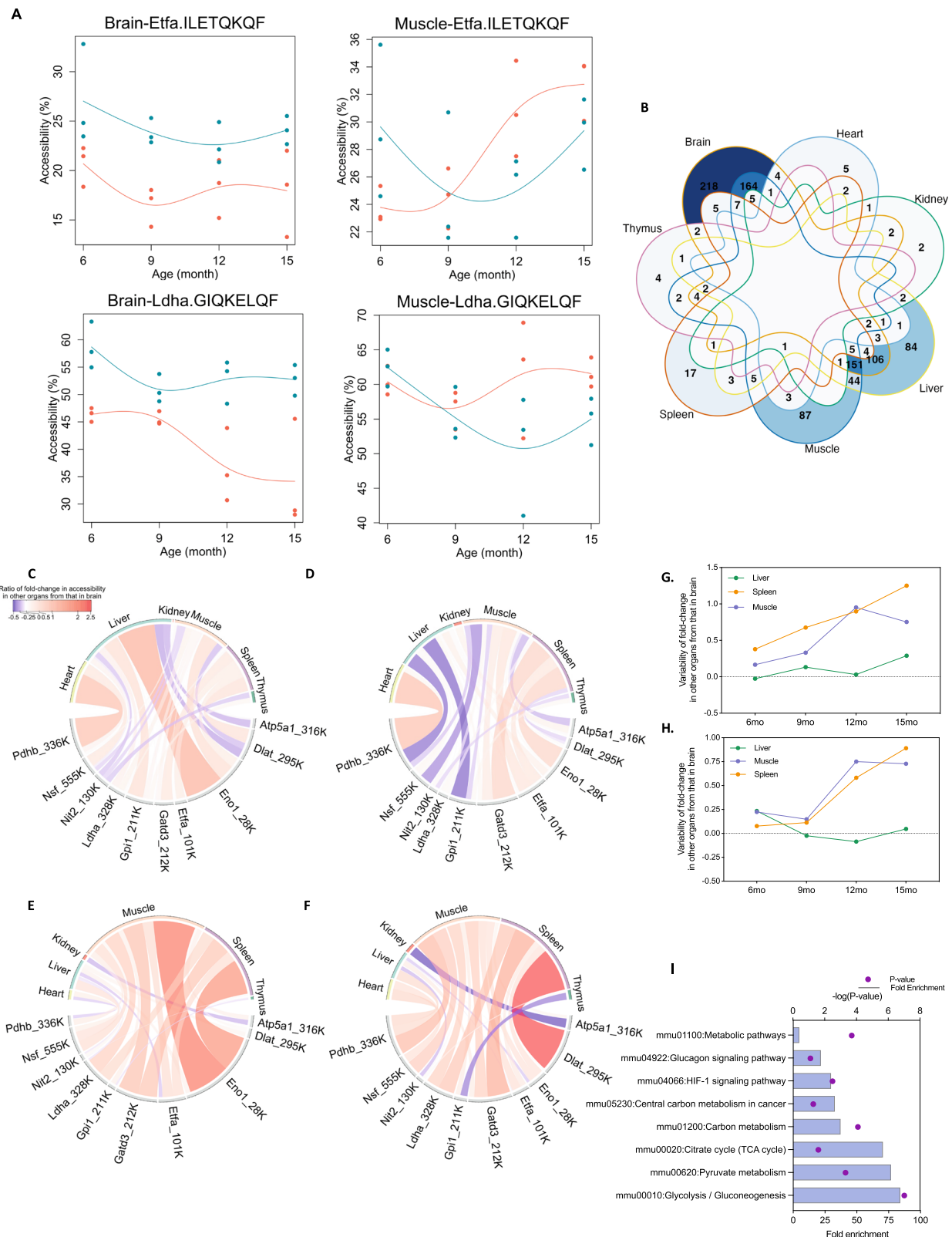


Fig. 2 | Pattern of labeling across tissues. **A** More than half of the total labeled peptides were tissue-specific. Less than 3% of a total of labeled peptides were peptides common to all seven tissues. **B** The proportions of labeled proteins were determined by assigning proteins to the labeled peptides. Unlabeled proteins were

not counted. The largest portion of labeled proteins was tissue-specific proteins, and the portion of proteins common to all 7 tissues was the third largest portion. **C–F** Biological triplicates were correlated across 7 tissues at 6 months (**C**), 9 months (**D**), 12 months (**E**) and 15 months (**F**).



on AD and some footprinting studies on AD related proteins, there have been no comprehensive in vivo studies of protein structures as AD progresses^{24,29–31}. To uncover the changes in the 3D structure of brain proteins corresponding human homologs, we focused on 780 proteins known to be expressed in the brain based on the Human Protein Atlas database (<https://www.proteinatlas.org>) and also found to be dimethyl-labeled in our brain dataset. To identify lysine sites that change significantly in accessibility during the progression of AD and

which also differ from NC, we used statistical methods to test the accessibility of 3,456 peptides corresponding to 780 proteins. Given the limited sample size (3) in our study, it is challenging to assume or test “normality”³². As a result, we chose to employ a non-parametric test to minimize the likelihood of Type I errors, or false positive results³³. Non-parametric tests apply relatively flexible assumptions to small sample sizes, potentially yielding more reliable results³⁴. We first used a Kruskal-Wallis test to compare lysine site accessibility among

Fig. 3 | Variability of the conformational changes depending on the tissues.

A Two representative peptides (ILETQKQF and GIQKELQF) were shown representatively. The normalized values were utilized to fit spline models. The accessibility of each peptide for both NC (blue) and AD (red) exhibited significantly distinct patterns from 6 to 15 months (Benjamin Hochberg, adjusted P -values = 0.0036 (brain), 0.0282 (muscle) for ILETQKQF, adjusted P -values = 0.0029 (brain), 0.0028 (muscle) for GIQKELQF). **B** Venn diagram shows the number of peptides exhibiting significant differences in the trend of accessibility changes between NC and AD. There were no peptides from AD that showed a significant difference in accessibility changes compared to NC in all 7 tissues during the period from 6 to 15 months. The value of zero was not indicated. **C–H** During AD progression, 10 common peptides

exhibited distinct patterns in accessibility between NC and AD in different four tissues. The variabilities for the structural changes in each tissue were calculated based on the value of brain using the formula: (fold-change of other tissue - fold-change of brain) / fold-change of brain at 6 mo (**C**), 9 mo (**D**), 12 mo (**E**), and 15 mo (**F**). Only the first three amino acids were shown. AGTAEAIKAL of Gatd3 (**G**) and GIQKELQF of Ldha (**H**) showed a difference in the magnitude of accessibility change in muscle and spleen compared to that in the brain as AD progressed. The number indicates the position of the labeled lysine site within the sequence. **I** Enriched KEGG pathways with 10 proteins. P -value were corrected with Bonferroni. Source data are provided as a Source Data file.

the four age groups (6, 9, 12, and 15 months) of AD (P -value ≤ 0.05), and then we used a Mann-Whitney test to compare lysine site accessibility between AD and NC per age (P -value ≤ 0.05). This approach allowed us to select labeled peptides that showed significant changes in the accessibility throughout disease progression (6–15 months), while also exhibiting significant differences from NC at each age point. Of the 3,456 peptides that were tested, 83 peptides corresponding to 62 proteins showed a significant difference in accessibility in all tests. Among 83 peptides, the lysine position in four pairs of peptides was the same: APVISAEEKAY and APVISAEEKAYHEQL for Tuba1, HPEQLITG-KEDAANNY and ITGKEDAANNY for Tuba1, QVVLVEPKTAW and QYQVVLVEPKTAW for Cnp, and RYLSEVASGENKQTTVSNSQQAY and SEVASGENKQTTVSNSQQAY for Ywhab. Most peptides exhibited a tendency for decreased accessibility with aging. Accessibility in the peptides from the AD group demonstrated a distinctive pattern characterized by a precipitous decline at the 9-month time point (Fig. 4A, B). These results suggest that the structural changes in these proteins are more influenced by AD pathology than by the normal aging process. Furthermore, it suggests that AD may not only alter the structures of individual proteins, it may also produce disease-related physiological changes that modify communities of multiple proteins acting as networks.

We examined 62 mouse proteins that have human homologs known to be highly enriched in the human brain to confirm that they are more abundant in the mouse brain than in the other six organs. The relative abundance of protein in the brain was determined by dividing the protein abundance in brain tissue by the average protein abundance of six other tissues for each age group. Of the 62 proteins, 60 (excluding two brain-specific proteins) exhibited a range of enrichment factor from 0.13-fold (Eef2) to 1,232-fold (Tuba1b) when compared to their expression in other tissues (Fig. 4C). On average, 51 of these proteins were expressed at higher levels in the brain than in other tissues, while 9 proteins were expressed at lower levels in the brain than in other tissues. Expression of a total 60 proteins was also enriched 21.7-fold in NC and 20-fold in AD on average. No tendency in expression was observed in either aging or AD status (Supplementary Fig. 9A). The abundances of most of the 62 proteins in brain was not significantly different between NC and AD nor did they show a consistent pattern that correlated with AD. For example, 2',3'-cyclic-nucleotide 3'-phosphodiesterase (Cnp), which is associated with neuronal cells and glial cells³⁵, was found to be enriched in the brain. Cnp in the brain was relatively highly, with levels ranging from 13.4- to 23.9-fold in the NC and from 17.0- to 27.7-fold in the AD across four ages (Supplementary Fig. 9B). The accessibility of 2 peptides of Cnp decreased in progressing AD but remained unchanged in the aging NC (Fig. 4D, E). Additionally, differences in expression levels of Cnp in brain were not statistically significant between NC and AD, and they did not change significantly with AD progression (Supplementary Fig. 9C).

These results suggest that the structural information obtained through the accessibility of lysine sites can complement protein expression data by revealing changes that are not fully characterized by protein expression levels alone. This method provides an approach

to collect both protein expression and structural information to gain a more comprehensive understanding of the changes that occur in a proteome during the progression of a disease.

Differential structural changes of tightly regulated proteins

We hypothesized that the changes in protein expression could accompany structural changes to proteins, which may suggest alterations to their physiological function. To evaluate our hypothesis, we initially analyzed the abundance of proteins in tightly co-regulated protein networks by modularizing them into protein communities using a WGCNA algorithm³⁶. WGCNA was applied to the dataset of each tissue; 4,295 proteins in brain, 3,074 proteins in heart, 4,186 proteins in kidney, 4,298 proteins in liver, 1,721 proteins in muscle, 4,800 proteins in spleen, 4,159 proteins in thymus were used to build protein co-expression networks. No outliers were detected after all samples were hierarchically clustered using average distance and Pearson's method. For brain, the lowest soft threshold power was 22, with an R^2 of more than 0.75. This network consisted of 17 modules of proteins related by their co-expression across control and disease tissues based on the TOM-based dissimilarity, after merging the modules with dissimilarity (Supplementary Fig. 10A, B). The WGCNA analysis also divided the protein data sets into 14, 17, 13, 10, 12 and 10 modules for heart, kidney, liver, muscle, spleen and thymus, respectively (Supplementary Fig. 11A–G).

We evaluated the correlation of the co-expressed proteins with AD by comparing the co-expression of proteins in AD and NC within each module, irrespective of age. Correlation of coefficient $R > 0.4$ and P -value < 0.05 were set as the criteria for significant correlation for the AD-related expression. We found that only a limited number of protein communities were significantly correlated with AD-related characteristics, with one module (M3) in the brain showing a correlation of 0.41, two modules (M11 and M14) in the kidney showing correlations of 0.43 and 0.44, two modules (M2 and M5) in the muscle showing correlations of 0.44 and 0.46, and one module (M4) in the spleen showing a correlation of 0.44. No modules in other tissues showed a significant correlation related with AD (Fig. 5A). We also assessed whether the direction (positive or negative) and strength of the association between the protein community expression in each module and AD remained consistent following the subdivision of the samples into four age groups (Supplementary Fig. 11A–G). Strong correlations were observed at a certain age in a few modules, while in most modules, the direction of correlation was inconsistent across the four age groups. For example, when samples from all age groups were included, a high correlation was observed for module 3 (M3) of brain ($R = 0.41$) (Fig. 5A, B). When the samples were divided by age, M3 proteins from the 9-month samples were found to be highly negatively correlated with a value of $R = 0.82$. At 12 and 15 months, samples were also negatively correlated, with R values of 0.42 and 0.72, respectively. In contrast, M3 proteins from the 6-month samples showed a positive correlation with a value of $R = 0.4$ (Supplementary Fig. 11). The direction of AD-related association for the protein abundance-based was found to fluctuate during the progression of AD. Therefore, we focused on the proteins in modules displaying significant related with AD across all samples to investigate the conformational changes of the co-regulated proteins.

We investigated the conformational changes of the co-regulated proteins from the 4 types of tissues (brain, kidney, muscle, and spleen) that showed significant correlation based on WGCNA. To examine

whether the conformational changes of proteins were specifically impacted by AD, the labeled peptides were clustered based on the fold-change of the accessibility during the progression of AD using the

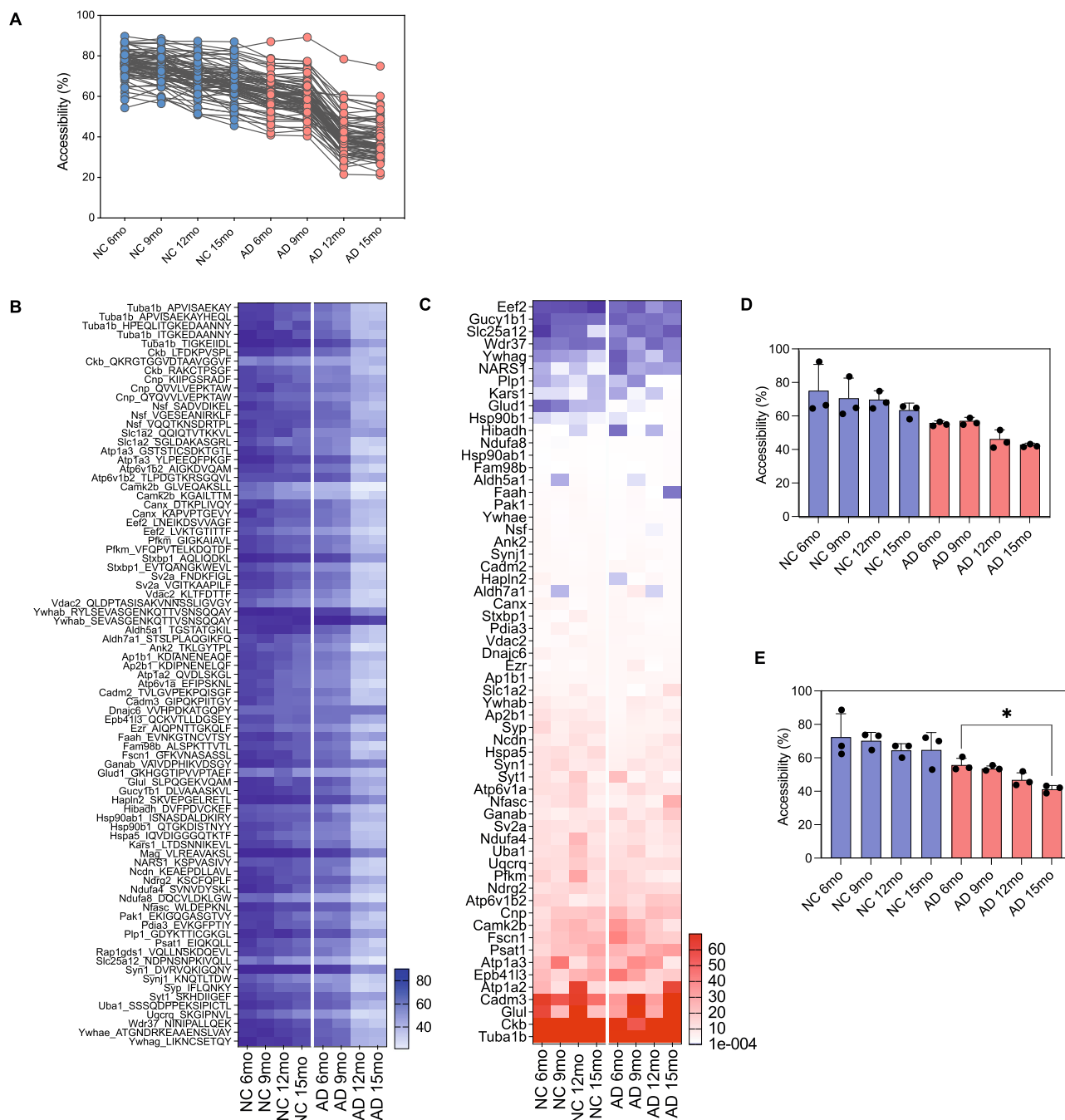


Fig. 4 | Changes in the structures of proteins that are known to be associated with brain. A, B Accessibility of 83 labeled peptides that mapped to 62 proteins were significantly changed. The accessibility values in AD groups decreased more steeply than those in NC groups (A). The proteins and peptides corresponding to each change in the accessibility are indicated B. **C** Expression of 60 brain proteins were compared to averaged expression in six other tissues, as enrichment factor (C). The minimum enrichment factor was 0.13 for Eef2 in NC at 15 months and the maximum enrichment factor was 1232 for Tuba1b in NC at 12 months. Expression of 6 proteins (Eef2, Gucy1b1, NARS1, Slc25a12, Wdr37, Ywhag) was lower in brain than the expression of the corresponding proteins in other tissues at all ages in NC and AD. The bar indicates the enrichment factor, with red indicating an enrichment factor more than 1, and blue indicating an enrichment factor less than 1. Proteins enriched more than 70-fold are marked in red. **D, E** Two of the three peptides of

Cnp share one lysine site, and variations in accessibility for these two peptides are represented. KIIPGSRADF (D) is located at 87-96 amino acid of Cnp and QYQVVL-VEPKTAW (E) is located at 141-153 amino acid of Cnp. While they exhibited a decreasing trend in AD, the accessibility values were lower in AD compared to NC, and the trend in AD was steeper than in NC. In the AD group, both peptides showed P-values below 0.05 (Kruskal-Wallis test, $p = 0.0084$ (D), $p = 0.0008$ (E)), indicating statistically significant changes in peptide accessibility during AD progression. Experiments were performed in biological triplicates per group, with each dot representing an individual mouse. The significance levels shown in the graph represent P-value from post-hoc analysis (Dunn's test, adjusted $p = 0.0327$ for AD 6mo vs. AD 15mo). Blue indicates NC groups, pink indicates AD groups. Asterisk (*) denotes P-value < 0.05 from Dunn's test. Error bars indicate the mean $\pm$ standard deviation. Source data are provided as a Source Data file.

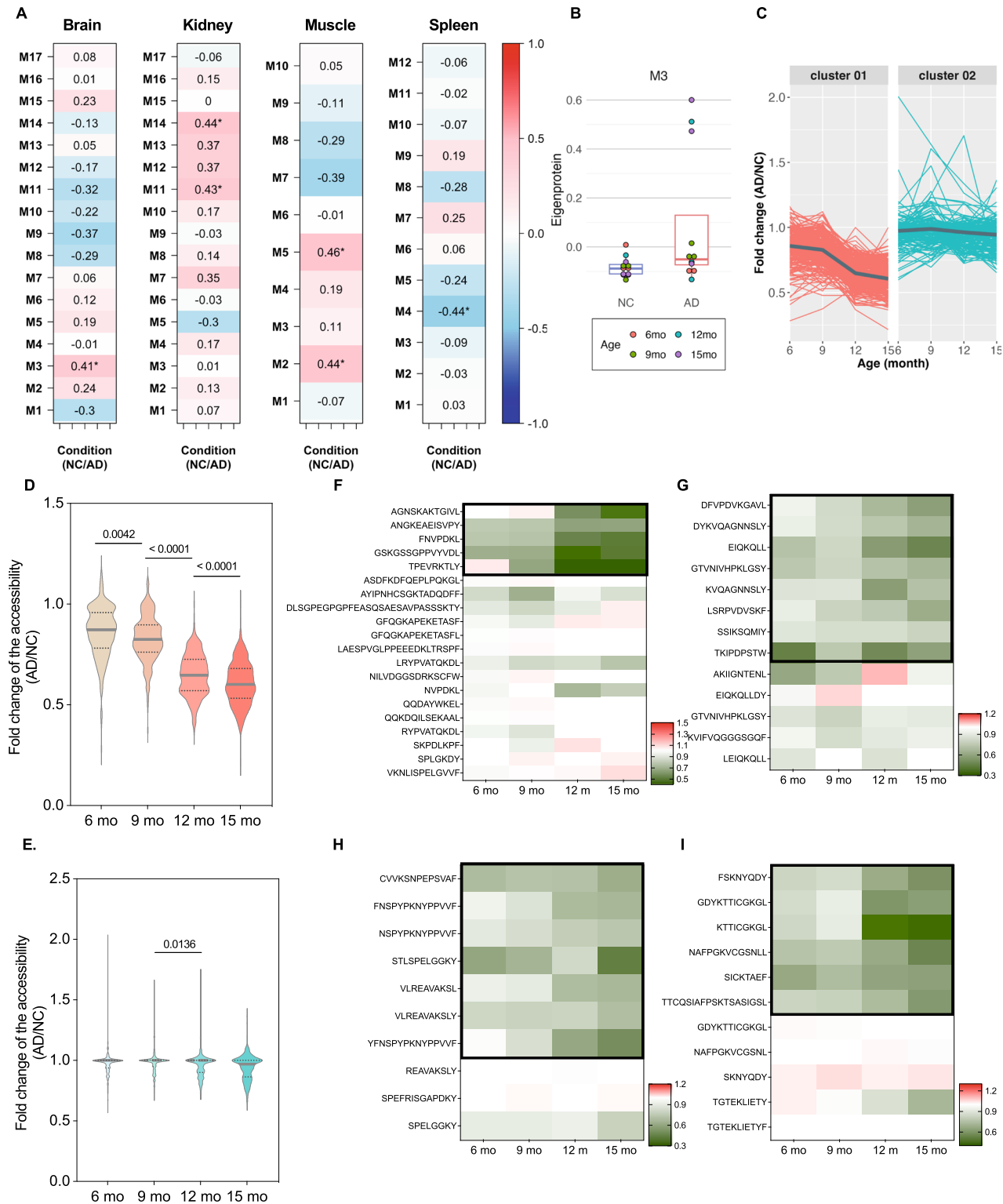


Fig. 5 | The structural changes of the co-expressed proteins by WGCNA. A Four of the seven tissues (brain, kidney, muscle and spleen) showed significant correlated modules ($R^2 > 0.4$, Two-sided Student's t-test, P -value < 0.05). **B** In module 3 (M3) of brain, the eigenprotein level between AD and NC was assessed, with dot colors indicating the age of the mice. Three mice per age were used for both NC and AD. The minima/maxima are the lowest/highest data point. Center line denotes median, box edges indicate the 25th and 75th percentiles, and whiskers extend to ± 1.5 interquartile range (IQR). **C–E** The labeled peptides of M3 proteins were clustered based on the fold-change of the accessibility (**C**). Distribution of the fold-

change of the accessibility of cluster 1 (**D**) and cluster 2 (**E**). Of 481 labeled peptides that mapped onto 174 proteins in M3, fold-change of accessibility of 268 peptides showed a consistent decrease in progressing AD in cluster 1 and fold-change of accessibility of 213 peptides in cluster 2 did not significantly change in progressing AD. P -values from two-sided Student's t-test are shown between the indicated groups. **F–I** The fold changes of the accessibility for Map1a (**F**), Psat (**G**), Mag (**H**), and Plp1 (**I**) are presented. The peptides in the bold box were included in cluster 1. Peptides in the bold box clearly decreased. Source data are provided as a Source Data file.

K-means clustering algorithm. The number of clusters was determined via an optimization algorithm (Supplementary Fig. 12). For the brain, the 469 labeled peptides that were mapped onto 174 proteins constituting M3 were clustered into two clusters (Fig. 5C). Accessibility of the lysine sites of 265 peptides (123 proteins) in cluster 1 showed a steadily decreasing pattern, with a slight decrease during early AD development (6–9 months), a dramatic decrease from 9 to 12 months, and a slight decrease again during the late stage of AD development from 12 to 15 months (Fig. 5D). This suggests that beginning at 9 months, the lysine sites included in cluster 1 became significantly sterically inaccessible due to AD. On the other hand, lysine sites of 204 peptides (113 proteins) in cluster 2 showed relatively stable accessibility during AD development (Fig. 5E). Of the 174 proteins in M3, 62 are also present in both cluster 1 and cluster 2 (represented by 2 or more peptides), and the labeled peptides of 61 and 51 proteins were exclusively included in cluster 1 and cluster 2, respectively (Supplementary Fig. 13). Twenty peptides that mapped to Map1a (primarily a neuronal protein) were most frequently identified in M3 proteins (Supplementary Fig. 14), and 5 of 20 peptides were hidden as AD progressed, but 15 lysine sites were spatially stable (Fig. 5F). Subsequently, 19- and 14-labeled peptides were mapped to Pkg1 and Mdh1, respectively, and the peptides included in cluster 1 exhibited a consistently decreasing pattern of accessibility. The lysine sites of the peptides of Psat1, Mag, and Plp1, which are highly expressed in oligodendrocytes and were included in the 62 brain proteins that constituted M3, became inaccessible during progression of AD (Fig. 5G–I). The datasets from kidney, muscle, and spleen were processed separately to reveal the AD-induced spatial changes of proteins (Supplementary Fig. 15). Collectively, our findings suggest that the steric changes of proteins occur concurrently with changes in co-expression of proteins as AD progresses.

Biological functions of protein communities whose conformational changes precede expression changes

We sought to uncover how the network undergoing conformational changes was related to biological functions, particularly those involved in neurodegenerative diseases. To achieve this, we examined the interactions of proteins, which had shown altered patterns of accessibility fold-change via K-means clustering, as well as their functional implications. For instance, the proteins that were included in cluster 1 were used for the brain dataset.

We investigated how changes in the structure and expression of proteins could affect the progression of AD by examining the pathways and biological functions they are involved in. We performed the analysis for the enrichment network on Metascape (metascape.org), considering the inter-term similarity and intra-term redundancy in the enriched terms³⁷. A total of 113 significantly enriched terms were grouped into 20 clusters based on their similarities and redundancies in the brain dataset (Supplementary Fig. 16 and Supplementary Table 3). Fifteen proteins (Gnas, Mdh1, Ogdh, Pkg1, Ppp1cb, Slc1a3, Sod2, Taldo1, SdhA, Oxct1, Ndufa8, Epm2aip1, Aldh1l1, EtfA, and Ugp2) were enriched in “generation of precursor metabolites and energy” (node-107) with the most significantly enriched having a *P*-value of 1.78×10^{-9} . “Energy derivation by oxidation of organic compounds” (node-1) was also enriched significantly, but was similar to node 107 with kappa score of more than 0.3. It has been demonstrated that an abnormality of carbon and energy metabolism occurs in neurodegenerative disease since neurons require large amounts of energy to maintain their normal activity, and metabolic decline of the brain contributes to cognitive impairment^{38,39}. Interestingly, we discovered that structural changes preceded the expression changes in the protein communities “generation of precursor metabolites and energy” and “carbon metabolism” (node-96) (Fig. 6A, B, D, E and Supplementary Fig. 17). To isolate protein expression changes specifically attributed to AD progression while excluding effects of normal aging, we

calculated the fold-change of abundance by dividing the protein abundance in the AD group by that in the NC group. This approach allowed us to analyze changes in protein expression levels as time progressed from 6 to 15 months. Protein expressions in node-107 and node-96 increased in AD but remained stable in NC during aging. Expression fold-change increased significantly at 15 months, whereas accessibility fold-change decreased significantly after 9 months. These findings further support the arguments in an earlier study that reported that in the mouse model of APP^{NL-F}, synaptic loss become noticeable at 9–12 months⁴⁰. The breakdown of the blood-brain barrier (BBB) is subsequently accompanied by a decline in synaptic integrity⁴¹. This disruption induces alterations in blood flow, decreased expression of glucose transporter proteins⁴², and modifications in signaling pathways that regulate the use of neuronal glucose⁴³. The decreased energy levels ultimately trigger a cascade effect, culminating in the downstream failure of energy-dependent processes and resulting in synaptic loss⁴⁴. We also noted that “metal ion homeostasis” (node-108) was significantly enriched with 12 proteins (Ank1, Calb1, Calb2, Gnas, Itpr1, Prkcb, Slc12a4, Slc1a3, Sod2, Vapb, Fis1, Immt). The homeostasis of metal ions is also known to be essential to maintain the normal function of brain, and abnormally elevated iron in brain is recognized to induce cell death and to be a cause of several neurodegenerative diseases including AD^{45,46}. Zinc also has an essential role in protein binding for enzymatic activity or to modulate synaptic transmission, and abnormal levels of zinc have been reported to be implicated in AD^{47,48}. The enriched proteins in metal ion homeostasis showed no significant change in expression but a significant decrease in accessibility from 9 months (Fig. 6C, F and Supplementary Fig. 16).

Next, we used the STRING database to identify 45 out of 123 proteins included in cluster 1 that interacted directly with each other through 30 edges. Of these, 39 proteins were associated with either brain-related terms or the enriched terms in which proteins showed conformational changes preceding the expression change (node-97, node-107, and node-108) (Fig. 7A, Supplementary Table 4). We noted that Plp1 and Mag, which were associated with central nervous system and abnormal nervous system and were also known to be located in extracellular space, directly interacting with each other. Alphafold2-Multimer^{49–51} was utilized to predict the complex structure of Mag and Plp1. It showed that the two alpha carbons of the lysine residues FSKNYQDY of Plp1 and YFNSPYKKNYPVVF of Mag were located within 13.9 Å (Fig. 7B) of each other, thus allowing them to interact⁵² and making it possible that these adjacent peptides bind with each other. FSKNYQDY showed a greater decrease in accessibility in AD than in NC (Fig. 7C), and YFNSPYKKNYPVVF exhibited a similar accessibility in AD and NC at 6 months, but the lysine sites became inaccessible during AD progression, whereas no difference was observed in NC (Fig. 7D).

In addition to the brain, it has been reported that the peripheral system plays a role in amyloid-β clearance. Approximately 40–60% of brain-derived amyloid-β is transported across the blood-brain-barrier into the peripheral system for clearance, although the mechanisms of this process remain unclear⁵³. Spleen is composed of a variety of immune cells (with 7–8% of all cells being monocyte/macrophage) and has a role in blood filtering and immunological functions. In addition, the spleen monocytes/macrophages are reported to be involved in clearing amyloid beta⁵⁴. Still, the physiological mechanisms underlying the association between the peripheral organs and AD remain unknown. In the ontology network of the spleen dataset, we noted “carbon metabolism” (node-147) and “neutrophil degranulation” (node-151). In these communities of proteins, the lysine sites in AD became exposed during progressing AD, while the accessibility in NC remained stable from 6 months to 15 months (Supplementary Fig. 18A, B). No significant change in the expression of proteins was observed for proteins of carbon metabolism (Supplementary Fig. 18D), but by 15 months the expression of proteins in neutrophil

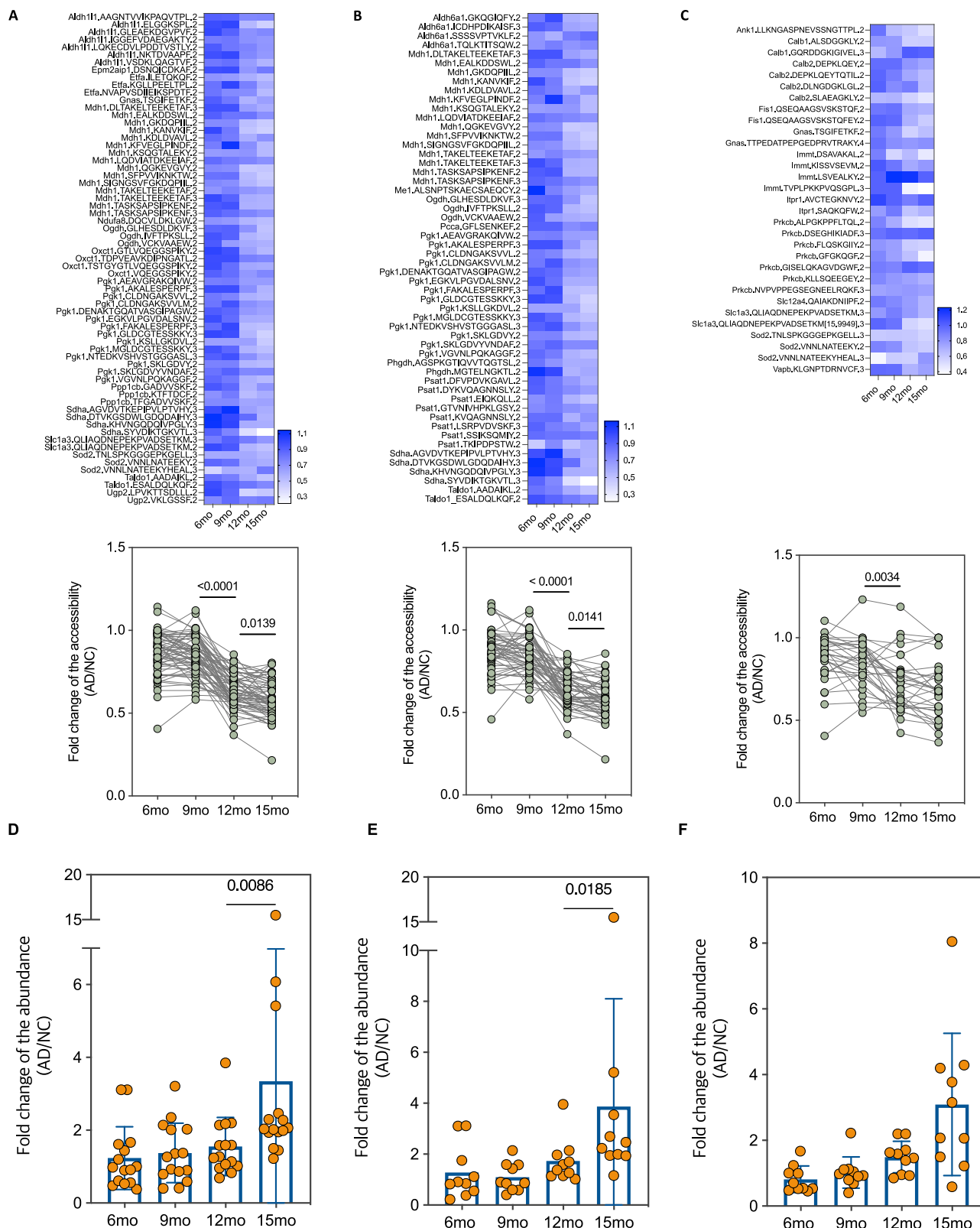


Fig. 6 | The biological functions affected by structural changes of composed proteins. **A–C** Structural changes of proteins enriched in generation of precursor metabolite and energy (**A**), carbon metabolism (**B**), and metal ion homeostasis (**C**) are shown. The heatmap (top) showed variations in the fold-change of the accessibility based on peptide sequence. The scatter plots (bottom) were plotted irrespective of peptide sequences. The p-value was calculated using Student's t-test.

D–F Expression change of proteins enriched in generation of precursor metabolite and energy (**D**), carbon metabolism (**E**), and metal ion homeostasis (**F**). The fold-changes of the expression level are presented. Three mice per age were used for both NC and AD. Error bars indicate the mean $\pm$ standard deviation. P-values from two-sided Mann-Whitney test are shown between the indicated groups. Source data are provided as a Source Data file.

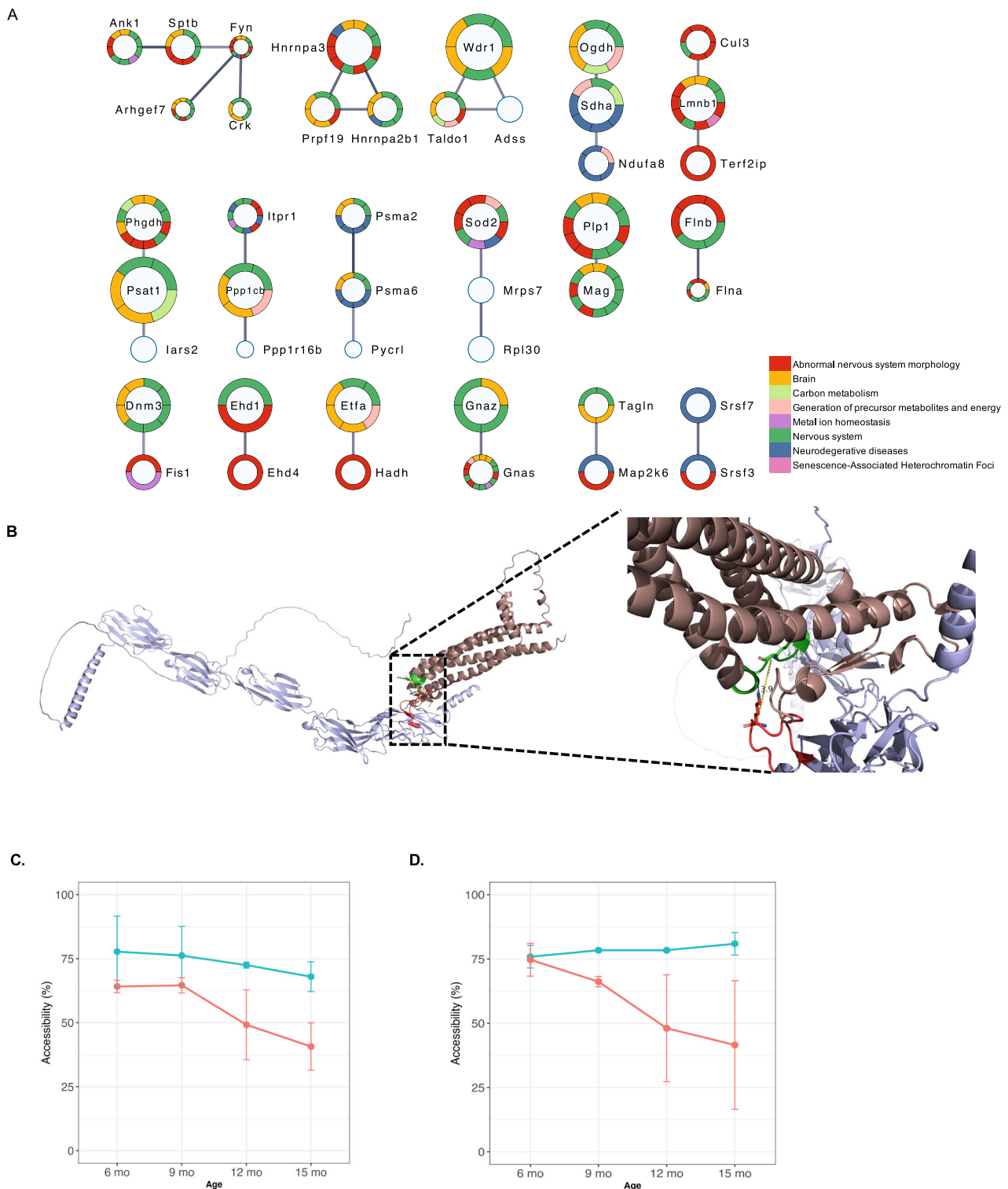


Fig. 7 | The physical interactions of the proteins in M3 of brain. **A** Forty-five proteins in the brain dataset were physically interacted. Each node indicates a protein. The ring color of the node indicates the terms that the protein is associated with. The size of node represents the number of significantly changed lysine sites, with very small nodes indicating no significantly changed peptides, small nodes indicating one significantly changed peptide, medium nodes indicating 2–4 significantly changed peptides, and large nodes indicating more than 4 significantly changed peptides. **B** The structure of the Plp1-Mag complex was predicted using AlphaFold-Multimer. The structure in dark pink is Plp1 and the

structure in light purple is Mag. FSKNYQDY of Plp1 and YFNSPYPKNYPPVVF of Mag were presented in green and red, respectively. The right panel is an enlarged view of the complex on the left. The distance between alpha-carbons of two lysine sites was 13.9 Å. **C, D** Structural changes in adjacent peptide regions with potential binding, with variation of the accessibility of site in AD (pink) and NC (green) for FSKNYQDY (**C**) and YFNSPYPKNYPPVVF (**D**). Three mice per age were used for both NC and AD. Error bars indicate the mean $\pm$ standard deviation. Source data are provided as a Source Data file.

degranulation differed significantly from the expression at 12 months (Supplementary Fig. 18C). The results from the other tissues are shown in the supplementary data (Supplementary Figs. 19–20 and Supplementary Tables 5–8). Collectively, the results presented here provide compelling evidence of a relationship between conformational changes and protein expression, thereby highlighting the significant impact of organ-specific alterations of biological function during the progression of AD.

Discussion

This study elucidated AD-associated conformational changes in the proteomes of seven tissues in mice. We used an AD mouse model (APP^{NL-F}), which expresses APP at wild-type levels while producing elevated pathogenic amyloid- β through an APP knock-in approach. This model reduces the risk of artificial phenomena that might be observed with an APP overexpressed mouse model and enhances the interpretability of the results⁴⁰. We found co-regulated proteins whose accessibility changed in 4 of 7 tissues, and we connected the structural differences of proteins between progressing AD and normal aging of unaffected mice to possible alterations in their biological functions. Whole animal perfusion was used to deliver reagents to comprehensively dimethyl-label mouse organs with minimal intervention (i.e., organ excision, tissue homogenization, or protein extraction) that could denature or alter proteins. Our method for quantitatively measuring accessibility determines the relative fraction of inaccessible over accessible for each lysine site. Changes in accessibility can be interpreted as a change in protein folding or a change in interaction with another molecule and thus can be a surrogate for protein conformation changes between different conditions.

The methodology used in this study is referred to as “in vivo labeling,” as we argue that during the postmortem interval immediately after the cessation of mouse heartbeats, the properties of the “living system” of cells do not immediately change, and, at the moment proteins are labeled, the living system maintains consistency, even if the mouse is in the process of ‘dying’ at that time. This is somewhat of a philosophical question—is organismal death the same as cellular death? Organs are removed from brain dead patients and used in transplants. For example, hearts and lungs must be transplanted within approximately four hours after being removed from the donor. Livers can be preserved between 12 and 18 h; a pancreas can be preserved 8–12 h; intestines can be preserved approximately 8 hours; kidneys can be preserved 24–48 h⁵⁵.

In this study, we focused on changes in proteins of the brain. The amyloid- β protein serves as a marker of AD and is known to accumulate as plaques in the brain, but we were not able to detect structural changes of amyloid- β . We were able to quantify the accessibility of three peptides from APP in only one 15-month mouse. Although it has been known that the amyloid beta plaque accumulates primarily in the cortex, hippocampus, basal ganglia, thalamus, and basal forebrain⁵⁶, we used whole brain tissue sample in this study because the purpose of this study was to globally investigate the change of the proteome. The structural changes we encountered in Psat1 (phosphoserine aminotransferase 1), Mag (myelin-associated glycoprotein), and Plp1 (proteolipid protein 1) are particularly intriguing because they are associated with oligodendrocytes in the brain, which have recently been found to produce amyloid- β in addition to neurons and thus contribute to amyloid- β plaque burden⁵⁷. The normal structure and function of Plp1 is also indispensable to the function of myelin in supporting neuronal axon activity and thus the conduction of normal brain electrical impulses.

It has been suggested that kidney function is linked to brain activity⁵⁸, and changes in kidney function may play a role in the development and progression of AD. Studies have shown that the MRI images of the brains of AD patients were similar to that of patients with kidney disease⁵⁹ and a systematic meta-analysis demonstrated that cognitive impairment is significantly related to malfunction of

kidney⁶⁰. Despite persuasive evidence of the link between the kidney and AD, the exact physiological mechanisms underlying this relationship are not fully understood. From our WGCNA analysis, we found two modules (M11 and M14) to be significantly co-expressed in kidney and we investigated the enriched functions of the structurally altered proteins using GO enrichment analysis. In both modules, purine-related functions were most significantly enriched. Our findings from the kidney are supported by the results of metabolomic studies showing that guanosine monophosphate (which is derived from purine guanine and associated with purine metabolism) was dysregulated in the brain of an AD mouse model based on APOE4 allele mutant mice⁵⁵. Therefore, it can be inferred that purine metabolism may play a role in the link between kidney function and AD.

In our previous study, we observed structural changes in mitochondrial succinate dehydrogenase B in brain lysates of AD patients⁶¹. We highlighted that lysine #137 of succinate dehydrogenase B was less accessible in AD patients than in control subjects. Interestingly, in this study we have found that mitochondrial succinate dehydrogenase B in mice is also affected by AD (Supplementary Fig. 21). In AD mouse samples, the accessibility of lysine #139 in succinate dehydrogenase B changed significantly from 6 to 15 months of age, with a decrease in exposure beginning at 9 months. In contrast, the accessibility of this residue remained consistent across all four time points in the NC group. The high sequence homology of these proteins (92%) between mouse and human species, coupled with the observed decrease in accessibility of lysine residues at positions 139 and 137, respectively, as AD progresses, suggests that this protein is structurally and functionally conserved across species. Furthermore, this result demonstrates that the dimethyl labeling approach combined with perfusion can be used as a viable alternative to tissue lysate-labeling. Interestingly, changes in the activity of tricarboxylic acid (TCA) cycle enzymes and the electron transport chain (ETC), have recently been reported to affect energy production and thus synapse loss in AD models⁶². Given that succinate dehydrogenase represents both a critical enzyme in the TCA cycle and complex II of the ETC, our findings further support the presence of structural changes in the proteins underlying mitochondrial energy metabolism in AD. Furthermore, we observed that our findings from the APP^{NL-F} model were reproducible using 5XFAD model (See Supplementary Notes under subheading ‘Validation with different AD mouse model’ and Supplementary Fig. 22).

Limitations of this study include the lack of a standard to monitor the distribution of reagent solution to organs in the body during the first labeling step. Although some signs such as body twitching, tail flicking, and head moving in the anesthetized animals was observed, a reliable quantitative standard to assess the extent of labeling in each organ would be useful. We defined labeling efficiency as the ratio of initially labeled peptides over identified lysine-containing peptides per tissue. An additional limitation of this method is its reliance on lysine residues as conformational reporters. Lysines typically comprise only 5–7% of amino acids in proteins, which constrains the amount of structural information that can be obtained. The quality and comprehensiveness of the analysis therefore scales with the sequence coverage of individual proteins. Higher sequence coverage generates more MS/MS data from lysine-containing peptides, thereby increasing the completeness of the structural analysis. However, this selective focus on lysine residues means that significant portions of the protein structure remain unexamined, potentially limiting the overall structural insights.

It is likely that changes in protein structure are a result of failing proteostasis and that expression changes are a function of alterations in protein synthesis and/or degradation. Mutations of DNA in somatic cells accumulate as we age^{63–66} and can result in changes to protein sequences, necessitating more effort to keep proteins properly folded. Our method provides a means to measure protein surface accessibility as a surrogate for protein conformation in vivo and in animal models to

study of the role of protein folding in aging and AD. Our proteomic analysis showed changes in protein structures in multiple tissues during the progression of AD. Even though the patterns of change in non-brain tissue did not correspond exactly with those in the brain tissue (and it is not clear they should), our analysis showed changes in protein structure and expression in other tissues in the AD mouse model. In conclusion, this method to measure in vivo alterations to protein surface accessibility in animal models of disease provides a means to measure a previously unexplored characteristic of proteins to provide insights into how physiological systems are perturbed.

Methods

Ethical Statement

Our research complies with all relevant ethical regulations. Animal facilities were AAALAC (Association for Assessment and Accreditation of Laboratory Animal Care) approved, and protocols (07-0083) were in accordance with the IACUC (Institutional Animal Care and Use Committee).

Animals/Tissue collection

Female mice (APP^(NL-F)⁴⁰) were purchased from RIKEN Brain Science Institute and female C57BL/6 were obtained from The Scripps Research Institute breeding colony. Mice were housed in plastic cages located inside a temperature- and humidity-controlled animal colony and were maintained on a standard cycle (a 12 h day/night cycle). Mice were sacrificed at 6, 9, 12, and 15 months of age.

First dimethyl-labeling and tissue collections

Mice were anesthetized by inhalation of 1% isoflurane. Chests of the anesthetized mice were opened by cutting the ribcage. The left heart ventricle was punctured with a perfusion needle and a small cut was made in the right atrium to allow outflow of the perfusion solutions. Blood components were washed away with prewarmed pH 7.4 phosphate-buffered saline (PBS) for 10 min. The mice were perfused with 20 mL of fixation solution (1% CD₂O) at a flow rate of 2.0 mL/min. Immediately afterward, 40 mL of the solution for the first light-dimethylation reaction (0.3 mM NaBH₃CN, 1% CD₂O in pH 7.4 PBS) was added at a flow rate of 2.0 mL/min. Organs were quickly excised and cut into 50 mg of tissue blocks. The tissue blocks were incubated in the same labeling solution (0.3 mM NaBH₃CN, 1% CD₂O in pH 7.4 PBS) for 10 min, and then the reaction was quenched by immersing the tissue blocks in 50 mM ammonium bicarbonate (ABC) solution for 5 min.

Tissue homogenization and protein extraction

Tissue blocks were placed in 100 μ L of 20 mM 2-[4-(2-hydroxyethyl)piperazin-1-yl] ethanesulfonic acid (HEPES) pH 7.4 and were homogenized with a pestle until no chunks were visible. The tissue samples were sonicated for 10 cycles (pulse-on 5 sec, pulse-off 3 s, amplify 30%) and then homogenates were clarified by centrifugation at 8,000 g at 4 °C for 30 min. Protein precipitation was performed by adding 400 μ L of 100% methanol, 100 μ L of 100% chloroform and 300 μ L of water to the sample. After vortexing vigorously, the samples were centrifuged at 8,000 g at 4 °C for 30 min. The large aqueous layer was discarded. The samples were washed by adding 800 μ L of 100% methanol and vortexing vigorously. After centrifugation at 8,000 g at 4 °C for 30 min, the supernatant was removed. The methanol washing step was repeated 3 times. Methanol was removed and the pellet was air-dried. The pellet was dissolved in 100 μ L of 1% sodium deoxycholate (SDC) in 20 mM HEPES pH 7.4. The protein concentration was determined with a BCA protein Assay kit following the instructions from the vendor (23225, Thermo Scientific).

Proteolysis of labeled proteins with chymotrypsin

Aliquots of tissue samples that contained 200 μ g of proteins were adjusted to 80 μ L with 1% SDC in 20 mM HEPES pH 7.4. The proteins

were reduced with 10 mM TCEP (Tris(2-carboxyethyl)phosphine hydrochloride) and 1% SDC in 20 mM HEPES pH 7.4 at 60 °C for 60 min on a shaker. Reduced proteins were alkylated with 20 mM IAA (iodoacetamide) for 30 min at 25 °C in the dark. Denatured proteins were digested with chymotrypsin (Promega) at 1:100 (enzyme:substrate(w:w)) at 37 °C for 16 h. Samples were acidified with formic acid to a final concentration of 1%. The sample was centrifuged at 8,000 g at 4 °C for 30 min and the supernatant was transferred to a new tube. The sample was centrifuged again at 8,000 g at 4 °C for 30 min to collect the clean sample and the supernatant was transferred to a new tube.

Second dimethyl-labeling and desalting

Pierce C18 spin tips (87784, ThermoFisher) were used for the second dimethyl-labeling step and desalting. A multipipette and a 96-well plate were used to prepare multiple samples in one batch. The C18 tips were activated by aspirating and dispensing 100 μ L each of 100% methanol and 100% acetonitrile (ACN). After the C18 tips were washed with 100 μ L of 0.1% formic acid, the samples were loaded onto the C18 tips by aspiration. To clean the samples bound to C18 tips, 100 μ L of 0.1% formic acid was aspirated and dispensed, and the pH was adjusted by aspirating 20 mM HEPES pH 7.4. Peptides bound to the C18 tips were dimethyl-labeled by aspirating 1% formaldehyde (¹³CD₃O), 0.3 mM Sodium cyanoborodeuteride (NaBD₃CN) and the saturated tips were incubated for 15 min at 25 °C. The reaction was quenched by aspirating 50 mM ABC and incubating for 10 min at 25 °C. After washing C18 tips with 0.1% formic acid, the labeled peptides were eluted with 100 μ L of 40% ACN in 0.1% formic acid followed by 100 μ L of 60% ACN in 0.1% formic acid. The eluted peptides were lyophilized.

Strong cation exchange (SCX) fractionation of peptides

SCX fractionation was conducted with commercial spin columns (90008, ThermoFisher Scientific). The pH of the sample was reduced by adding 800 μ L of 30% ACN in 0.1% formic acid. The spin column was equilibrated by adding 400 μ L of 30% ACN in 0.1% formic acid. It was then centrifuged at 1,000 g for 5 min, and the flow-through solution was discarded. The sample was loaded on the spin column and was centrifuged at 1,000 g for 3 min. Flow-through was stored for LC-MS/MS analysis. The peptides were eluted with consecutive 200 μ L aliquots of elution buffer containing 10 mM, 30 mM, 50 mM, 70 mM, 100 mM, 150 mM and 300 mM of ammonium acetate. All elution buffer aliquots contained 0.1% formic acid and 30% ACN.

LC-MS/MS analysis

Samples were loaded onto EvoTips in accordance with the manufacturer's protocol. Liquid chromatography-mass spectrometry (LC-MS) analysis was performed using an Evosep One system (Evosep) coupled to a timsTOF Pro mass spectrometer (Bruker Daltonics). Chromatographic separation was achieved using a 15 cm $\times$ 150 μ m inner diameter column packed with BEH C18 1.7 μ m particles (Waters) and featuring an integrated, in-house pulled tip. The 30 samples per day (SPD) method was employed for sample separation. Mobile phase A consisted of 0.1% formic acid in water, and mobile phase B comprised 0.1% formic acid in acetonitrile. Mass spectrometric data acquisition was conducted in parallel accumulation serial fragmentation (PASEF) mode. Each 1.1-second acquisition cycle included one MS1 survey trapped ion mobility spectrometry (TIMS)-MS scan followed by PASEF MS/MS scans. The dual TIMS analyzer was configured with an ion accumulation and ramp time of 100 ms each, covering an ion mobility range of $1/K_0 = 0.6\text{--}1.6$ Vs/cm². Precursor ions for MS/MS analysis were isolated using a 2 Th window for $m/z < 700$ and a 3 Th window for $m/z > 700$, spanning a total m/z range of 100–1700. Collision energy was applied as a linear function of increasing mobility, ranging from 59 eV at $1/K_0 = 1.6$ VS/cm² to 20 eV at $1/K_0 = 0.6$ Vs/cm². Singly charged precursor ions were excluded using a polygon filter. Precursor selection for tandem mass spectrometry was based on an

intensity threshold of 2,500 and a target value of 20,000, with an active exclusion period of 24 s.

Peptide identification and quantification

Raw files were searched against mouse proteins from Swiss-Prot-Uniprot database (retrieved 03/13/2022, 51,076 entries) containing canonical and isoform sequences, using MSFragger (version 17.1) in the FragPipe pipeline with mass calibration and parameter optimization enabled⁶⁷. Philosopher was used to filter all peptide-spectrum matches. Quantification analysis was performed with IonQuant. The parameter setting of chymotrypsin allowed for two missed cleavage sites and the minimal required peptide length was set to six amino acids. Dimethyl peptide pairs were identified using variable modifications of light (Δ mass: 32.0564) and heavy labeling (Δ mass: 36.0757) of lysine, oxidation of methionine (Δ mass: 15.9949), fixed modification of heavy dimethylation (Δ mass: 36.0757) on N-terminus and the carbamidomethylation of cysteine (Δ mass 57.0214). Precursor tolerance was set to 50 ppm and fragment tolerance was set to 50 ppm. Isotope error was set to 0/1/2. The minimum number of fragment peaks required to include a PSM (peptide-spectrum match) in modeling was set to two, and the minimum number required to report the match was four. The top 150 most intense peaks were considered, and a minimum of 15 fragment peaks were required to search a spectrum. The data were also searched against a decoy database and protein identifications were accepted at 1% peptide false discovery rate (FDR). All identified peptides were heavy-dimethylated on N-terminus.

Determination of accessibility of lysine sites and protein expression

Each peptide with a lysine site should be either light- or heavy-dimethylated, depending on the accessibility of lysine site. The difference in intensity of the peptides labeled in the first isobaric labeling step versus the second yields a relative abundance ratio R^{68} . The R value represents the proportion of the peptide in which a specific lysine site was accessible for dimethylation and is independent of the overall protein amount in the sample²⁴. The relative accessibility of a lysine residue for dimethylatizon is assessed by the value of accessibility; Accessibility (%) = $R/(1 + R) \times 100$. Accessibility was used as an indicator of structural changes and is based on the extent of dimethylation induced by exposure to formaldehyde and cyanoborohydride. Therefore, peptides that were unlabeled in both heavy and light forms, as well as those naturally modified by dimethylation, were excluded from the accessibility calculations. The level of protein abundance was determined by summing the intensities of the top 3 unique peptides. Peptide abundances were summed together regardless of their labeling status.

k-Nearest Neighbor (kNN) Imputation

Missing values were imputed by kNN method, a machine learning method through the VIM package in R^{69,70}. The kNN approach is based on the assumption that there are correlations between spot volume patterns across different variables. This method addresses missing data by identifying spots that exhibit volume patterns similar to the spot under investigation⁷¹. To estimate the missing value, it calculates a weighted average using data from the k spots determined to be most similar. The similarity is quantified using Euclidean distance, which is then used to weight each spot's contribution to the final estimate. The optimal number of k -neighbors must be determined through empirical testing for each specific dataset.

Weighted correlation network analysis (WGCNA)

A weighted protein co-expression network was built using the value of protein abundance from blockwiseModules WGCNA function⁷². Construction of weighted gene co-expression networks was conducted independently for each of 7 tissue datasets. The soft thresholding

powers were determined with the R function pickSoftThreshold³⁶. To pick an appropriate soft-thresholding power for network construction, the value of power was raised to 50. The chosen values were the smallest threshold that resulted in a scale-free R^2 fit of 0.75 and the networks were created by calculating the component-wise minimum values for topologic overlap. Soft threshold powers varied across seven tissues as follows: 22 for brain, 16 for heart, kidney, 12 for liver, muscle and spleen, and 26 for thymus. BlockwiseModule function was run with the following parameters: TOMType = "signed", maxBlockSize = 5000, mergeCutHeight = 0.1, verbose = 3. Module eigenprotiens (MEs) were calculated the correlation between the traits of AD (AD vs. NC). Multiple comparisons were accounted for by FDR correction across modules, and the P -values for the modules were reported.

Statistical analysis

Differentially expressed proteins and altered accessibilities (%) between pairs of different age groups (6, 9, 12, and 15 months) in AD or between different pathological conditions (AD and NC) per age were found using Mann-Whitney tests independently. Kruskal-Wallis was used to simultaneously compare the accessibilities among four age groups (6, 9, 12, and 15 months) or among three age groups (6, 9, and 12 months or 9, 12, and 15 months) in AD. These comparisons were tested with Kruskal-Wallis followed by Bonferroni's comparison *post hoc* test independently. The criterion for significance was a P -value less than 0.05.

Enrichment of Gene ontology (GO) and protein-protein interaction analysis

GO analysis was performed with ClueGO (a plug-in in Cytoscape) to identify the significant biological functions of the proteins in the WGCNA module⁷³. Protein-protein networks were detected using Metascape and the following databases: STRING, BioGrid, InWeb_IM, OmniPath^{37,74,75}. The resultant network contained the subset of proteins that form physical interactions with at least one other list member, the confidence cutoff of physical interaction was set to medium (0.5) or strong (0.7). Visualization of the protein-protein interaction network was performed on the Cytoscape combining STRING.

Complex modeling with AlphaFold2-multimer

We used AlphaFold2-multimer to predict the protein-protein interaction motif of each complex. AlphaFold2-multimer modeling was performed with ColabFold⁷⁶. Input multiple sequence alignment (MSA) features were generated by local ColabFold using the "MMseqs2 (Uniref + Environmental)" MSA mode. By default, the constructed MSAs contain both unpaired (per-chain) and paired sequences. AlphaFold2-multimer was run with one or several options from the following list: model type = alphafold2_multimer v3, num recycles = 3, recycle early stop tolerance = 0.5, max msa = auto, num seeds = 1. The models were ranked by confidence score and rank 1 was selected as the most accurate model. The distance between two lysine residues was calculated using PyMOL2 version 2.5 (Schrödinger, LLC).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All mass spectrometry raw data in this study have been deposited to MassIVE repository with identifier MSV000091970 [<https://doi.org/10.25345/CSVTIH07H>]. Source data are provided with this paper.

References

1. Jones, S. & Thornton, J. M. Principles of protein-protein interactions. *Proc. Natl. Acad. Sci. USA* **93**, 13–20 (1996).

2. Keskin, O., Gursoy, A., Ma, B. & Nussinov, R. Principles of protein-protein interactions: what are the preferred ways for proteins to interact? *Chem. Rev.* **108**, 1225–1244 (2008).
3. Kalheim, L. F., Fladby, T., Coello, C., Bjornerud, A. & Selnes, P. [18F]-Flutemetamol uptake in cortex and white matter: comparison with cerebrospinal fluid biomarkers and [18F]-Fludeoxyglucose. *J. Alzheimers Dis.* **62**, 1595–1607 (2018).
4. Sakakibara, D. et al. Protein structure determination in living cells by in-cell NMR spectroscopy. *Nature* **458**, 102–105 (2009).
5. Oikonomou, C. M. & Jensen, G. J. Cellular electron cryotomography: toward structural biology in situ. *Annu. Rev. Biochem.* **86**, 873–896 (2017).
6. de Souza, N. & Picotti, P. Mass spectrometry analysis of the structural proteome. *Curr. Opin. Struct. Biol.* **60**, 57–65 (2020).
7. Benjin, X. & Ling, L. Developments, applications, and prospects of cryo-electron microscopy. *Protein Sci.* **29**, 872–882 (2020).
8. Renaud, J. P. et al. Cryo-EM in drug discovery: achievements, limitations and prospects. *Nat. Rev. Drug Discov.* **17**, 471–492 (2018).
9. Sheshberadaran, H. & Payne, L. G. Protein antigen-monoclonal antibody contact sites investigated by limited proteolysis of monoclonal antibody-bound antigen: protein “footprinting”. *Proc. Natl. Acad. Sci. USA* **85**, 1–5 (1988).
10. Slavov, N. Measuring protein shapes in living cells. *J. Proteome Res.* **20**, 3017 (2021).
11. West, G. M. et al. Quantitative proteomics approach for identifying protein-drug interactions in complex mixtures using protein stability measurements. *Proc. Natl. Acad. Sci. USA* **107**, 9078–9082 (2010).
12. Xu, G. & Chance, M. R. Hydroxyl radical-mediated modification of proteins as probes for structural proteomics. *Chem. Rev.* **107**, 3514–3543 (2007).
13. Mendoza, V. L. & Vachet, R. W. Probing protein structure by amino acid-specific covalent labeling and mass spectrometry. *Mass Spectrom. Rev.* **28**, 785–815 (2009).
14. Geer, M. A. & Fitzgerald, M. C. Energetics-based methods for protein folding and stability measurements. *Annu. Rev. Anal. Chem. (Palo Alto Calif.)* **7**, 209–228 (2014).
15. Li, K. S., Shi, L. & Gross, M. L. Mass spectrometry-based fast photochemical oxidation of proteins (FPOP) for higher order structure characterization. *Acc. Chem. Res.* **51**, 736–744 (2018).
16. Pepelnjak, M., de Souza, N. & Picotti, P. Detecting protein-small molecule interactions using limited proteolysis-mass spectrometry (LiP-MS). *Trends Biochem. Sci.* **45**, 919–920 (2020).
17. Cappelletti, V. et al. Dynamic 3D proteomes reveal protein functional alterations at high resolution in situ. *Cell* **184**, 545–559 e522 (2021).
18. Mackmull, M. T. et al. Global, in situ analysis of the structural proteome in individuals with Parkinson’s disease to identify a new class of biomarker. *Nat. Struct. Mol. Biol.* **29**, 978–989 (2022).
19. Espino, J. A., Mali, V. S. & Jones, L. M. In cell footprinting coupled with mass spectrometry for the structural analysis of proteins in live cells. *Anal. Chem.* **87**, 7971–7978 (2015).
20. Espino, J. A. & Jones, L. M. In vivo hydroxyl radical protein footprinting for the study of protein interactions in *Caenorhabditis elegans*. *J. Vis. Exp.* (2020).
21. Yu, K. et al. Global profiling of lysine accessibility to evaluate protein structure changes in Alzheimer’s disease. *J. Am. Soc. Mass Spectrom.* **32**, 936–945 (2021).
22. Lu, H. et al. DiLeu isobaric labeling coupled with limited proteolysis mass spectrometry for high-throughput profiling of protein structural changes in Alzheimer’s disease. *Anal. Chem.* **95**, 9746–9753 (2023).
23. Yin, K., Tong, M., Sun, F. & Wu, R. Quantitative structural proteomics unveils the conformational changes of proteins under the endoplasmic reticulum stress. *Anal. Chem.* **94**, 13250–13260 (2022).
24. Bamberger, C. et al. Protein footprinting via covalent protein painting reveals structural changes of the proteome in Alzheimer’s disease. *J. Proteome Res.* **20**, 2762–2771 (2021).
25. Bamberger, C., Diedrich, J., Martinez-Bartholome, S. & Yates, J. R. Cancer conformational landscape shapes tumorigenesis. *J. Proteome Res.* **21**, 1017–1028 (2022).
26. Bai, B. et al. Deep multilayer brain proteomics identifies molecular networks in Alzheimer’s disease progression. *Neuron* **105**, 975–991 e977 (2020).
27. Seyfried, N. T. et al. A multi-network approach identifies protein-specific co-expression in asymptomatic and symptomatic Alzheimer’s disease. *Cell Syst.* **4**, 60–72 e64 (2017).
28. Savas, J. N. et al. Amyloid accumulation drives proteome-wide alterations in mouse models of Alzheimer’s disease-like pathology. *Cell Rep.* **21**, 2614–2627 (2017).
29. Gau, B., Garai, K., Frieden, C. & Gross, M. L. Mass spectrometry-based protein footprinting characterizes the structures of oligomeric apolipoprotein E2, E3, and E4. *Biochemistry* **50**, 8117–8126 (2011).
30. Klinger, A. L. et al. A synchrotron-based hydroxyl radical footprinting analysis of amyloid fibrils and prefibrillar intermediates with residue-specific resolution. *Biochemistry* **53**, 7724–7734 (2014).
31. Li, K. S., Rempel, D. L. & Gross, M. L. Conformational-Sensitive Fast Photochemical Oxidation of Proteins and Mass Spectrometry Characterize Amyloid Beta 1-42 Aggregation. *J. Am. Chem. Soc.* **138**, 12090–12098 (2016).
32. Ghasemi, A. & Zahediasl, S. Normality tests for statistical analysis: a guide for non-statisticians. *Int. J. Endocrinol. Metab.* **10**, 486–489 (2012).
33. ÖZTUNA, D., ELHAN, A. H. & TÜCCAR, E. Investigation of four different normality tests in terms of Type 1 error rate and power under different distributions. *Turkish J. Med. Sci.* **36**.
34. Fagerland, M. W. t-tests, non-parametric tests, and large studies—a paradox of statistical practice? *BMC Med Res Methodol.* **12**, 78 (2012).
35. Sprinkle, T. J. 2’,3’-cyclic nucleotide 3’-phosphodiesterase, an oligodendrocyte-Schwann cell and myelin-associated enzyme of the nervous system. *Crit. Rev. Neurobiol.* **4**, 235–301 (1989).
36. Zhang, B. & Horvath, S. A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol. Biol.* **4**, Article17 (2005).
37. Zhou, Y. et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, 1523 (2019).
38. Knobloch, M. & Paolicelli, R. C. Myeloid Metabolism as a New Target for Rejuvenation?—Comments on Restoring Metabolism of Myeloid Cells Reverses Cognitive Decline in Ageing. *Nature*. 2021 Feb;590(7844):122-128. *Immunometabolism* **3**, e210034 (2021).
39. Navarro, A. & Boveris, A. Mitochondrial nitric oxide synthase, mitochondrial brain dysfunction in aging, and mitochondria-targeted antioxidants. *Adv. Drug Deliv. Rev.* **60**, 1534–1544 (2008).
40. Saito, T. et al. Single App knock-in mouse models of Alzheimer’s disease. *Nat. Neurosci.* **17**, 661–663 (2014).
41. Nation, D. A. et al. Blood-brain barrier breakdown is an early biomarker of human cognitive dysfunction. *Nat. Med.* **25**, 270–276 (2019).
42. Simpson, I. A., Chundu, K. R., Davies-Hill, T., Honer, W. G. & Davies, P. Decreased concentrations of GLUT1 and GLUT3 glucose transporters in the brains of patients with Alzheimer’s disease. *Ann. Neurol.* **35**, 546–551 (1994).
43. Talbot, K. et al. Demonstrated brain insulin resistance in Alzheimer’s disease patients is associated with IGF-1 resistance, IRS-1 dysregulation, and cognitive decline. *J. Clin. Invest.* **122**, 1316–1338 (2012).
44. Krukowski, K. et al. Traumatic brain injury in aged mice induces chronic microglia activation, synapse loss, and complement-dependent memory deficits. *Int. J. Mol. Sci.* **19**, (2018).

45. Drayer, B. et al. MRI of brain iron. *AJR. Am. J. Roentgenol.* **147**, 103–110 (1986).
46. Levi, S. & Finazzi, D. Neurodegeneration with brain iron accumulation: update on pathogenic mechanisms. *Front. Pharm.* **5**, 99 (2014).
47. Paoletti, P., Vergnano, A. M., Barbour, B. & Casado, M. Zinc at glutamatergic synapses. *Neuroscience* **158**, 126–136 (2009).
48. Sensi, S. L., Paoletti, P., Bush, A. I. & Sekler, I. Zinc in the physiology and pathology of the CNS. *Nat. Rev. Neurosci.* **10**, 780–791 (2009).
49. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
50. Tunyasuvunakool, K. et al. Highly accurate protein structure prediction for the human proteome. *Nature* **596**, 590–596 (2021).
51. Evans R. et al. Protein complex prediction with AlphaFold-Multimer. *bioRxiv*, (2022).
52. Vangone, A. & Bonvin, A. M. Contacts-based prediction of binding affinity in protein-protein complexes. *Elife* **4**, e07454 (2015).
53. Marques, M. A. et al. Peripheral amyloid-beta levels regulate amyloid-beta clearance from the central nervous system. *J. Alzheimer's Dis.* **16**, 325–329 (2009).
54. Yu, Z. Y. et al. Physiological clearance of Aβ by spleen and splenectomy aggravates Alzheimer-type pathogenesis. *Aging Cell* **21**, e13533 (2022).
55. Nuriel, T. et al. Neuronal hyperactivity due to loss of inhibitory tone in APOE4 mice lacking Alzheimer's disease-like pathology. *Nat. Commun.* **8**, 1464 (2017).
56. Thal, D. R., Rub, U., Orantes, M. & Braak, H. Phases of Aβ deposition in the human brain and its relevance for the development of AD. *Neurology* **58**, 1791–1800 (2002).
57. Sasmita, A. O. et al. Oligodendrocytes produce amyloid-beta and contribute to plaque formation alongside neurons in Alzheimer's disease model mice. *Nat. Neurosci.* (2024).
58. Tang, X. et al. Association of kidney function and brain health: a systematic review and meta-analysis of cohort studies. *Ageing Res. Rev.* **82**, 101762 (2022).
59. Guo, Z. et al. Common 1H-MRS characteristics in patients with Alzheimer's disease and vascular dementia diagnosed with kidney essence deficiency syndrome: a preliminary study. *Alter. Ther. Health Med.* **23**, 12–18 (2017).
60. Etgen, T., Chonchol, M., Forstl, H. & Sander, D. Chronic kidney disease and cognitive impairment: a systematic review and meta-analysis. *Am. J. Nephrol.* **35**, 474–482 (2012).
61. Pankow, S., Bamberger, C., Martínez-Bartolomé, S., Park, S.-K., Yates, J. R. Identification of in vivo CFTR conformations during biogenesis and upon misfolding by covalent protein painting (CPP). *bioRxiv*, (2021).
62. Andreyev, A. Y. et al. Metabolic bypass rescues aberrant S-nitrosylation-induced TCA cycle inhibition and synapse loss in Alzheimer's disease human neurons. *Adv. Sci. (Weinh.)* **11**, e2306469 (2024).
63. Miller, M. B., Reed, H. C. & Walsh, C. A. Brain somatic mutation in aging and Alzheimer's disease. *Annu. Rev. Genomics Hum. Genet.* **22**, 239–256 (2021).
64. Choudhury, S. et al. Somatic mutations in single human cardiomyocytes reveal age-associated DNA damage and widespread oxidative genotoxicity. *Nat. Aging* **2**, 714–725 (2022).
65. Kim, J. et al. Prevalence and mechanisms of somatic deletions in single human neurons during normal aging and in DNA repair disorders. *Nat. Commun.* **13**, 5918 (2022).
66. Luquette, L. J. et al. Single-cell genome sequencing of human neurons identifies somatic point mutation and indel enrichment in regulatory elements. *Nat. Genet.* **54**, 1564–1571 (2022).
67. Kong, A. T., Leprevost, F. V., Avtonomov, D. M., Mellacheruvu, D. & Nesvizhskii, A. I. MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. *Nat. Methods* **14**, 513–520 (2017).
68. Bamberger, C., Pankow, S., Park, S. K. & Yates, J. R. 3rd. Interference-free proteome quantification with MS/MS-based isobaric isotopologue detection. *J. Proteome Res.* **13**, 1494–1501 (2014).
69. Templ, M., Kowarik, A. & Filzmoser, P. Iterative stepwise regression imputation using standard and robust methods. *Comput. Stat. Data* **55**, 2793–2806 (2011).
70. Templ, M., Alfons, A. & Filzmoser, P. Exploring incomplete data using visualization techniques. *Adv. Data Anal. Cl.* **6**, 29–47 (2012).
71. Jung, K. et al. Stastical evaluation of methos for the analysis of dynamic protein expression data from a tumor study. *Rev. Stat. J.* **4**, (2006).
72. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinf.* **9**, 559 (2008).
73. Bindea, G. et al. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* **25**, 1091–1093 (2009).
74. Doncheva, N. T., Morris, J. H., Gorodkin, J. & Jensen, L. J. Cytoscape StringApp: network analysis and visualization of proteomics data. *J. Proteome Res.* **18**, 623–632 (2019).
75. Doncheva, N. T. et al. Cytoscape stringApp 2.0: analysis and visualization of heterogeneous biological networks. *J. Proteome Res.* (2022).
76. Mirdita, M. et al. ColabFold: making protein folding accessible to all. *Nat. Methods* **19**, 679–682 (2022).

Acknowledgements

We thank Dr. Claire Delahunty for critical reading the manuscript. Figure1 was created with the help of BioRender (BioRender.com/d78w669). National Institute on Aging - National Institutes of Health (RF1AGO61846-01, J.R.Y.) National Institute on Aging - National Institutes of Health (5R01AGO75862, J.R.Y.).

Author contributions

A.S., C.B., and J.R.Y. conceived the project. A.S. and H.K. performed the experiments. J.D. measured the samples on the mass spectrometer. A.S. and H.K. analyzed the data and the results. C.B. and D.M. provided the critical feedback. S.A.L. provided 5XFAD mouse model and feedback. J.R.Y. supervised the project. A.S. wrote the manuscript and prepared the figures with help from all authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-53582-x>.

Correspondence and requests for materials should be addressed to John R. Yates.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

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-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024