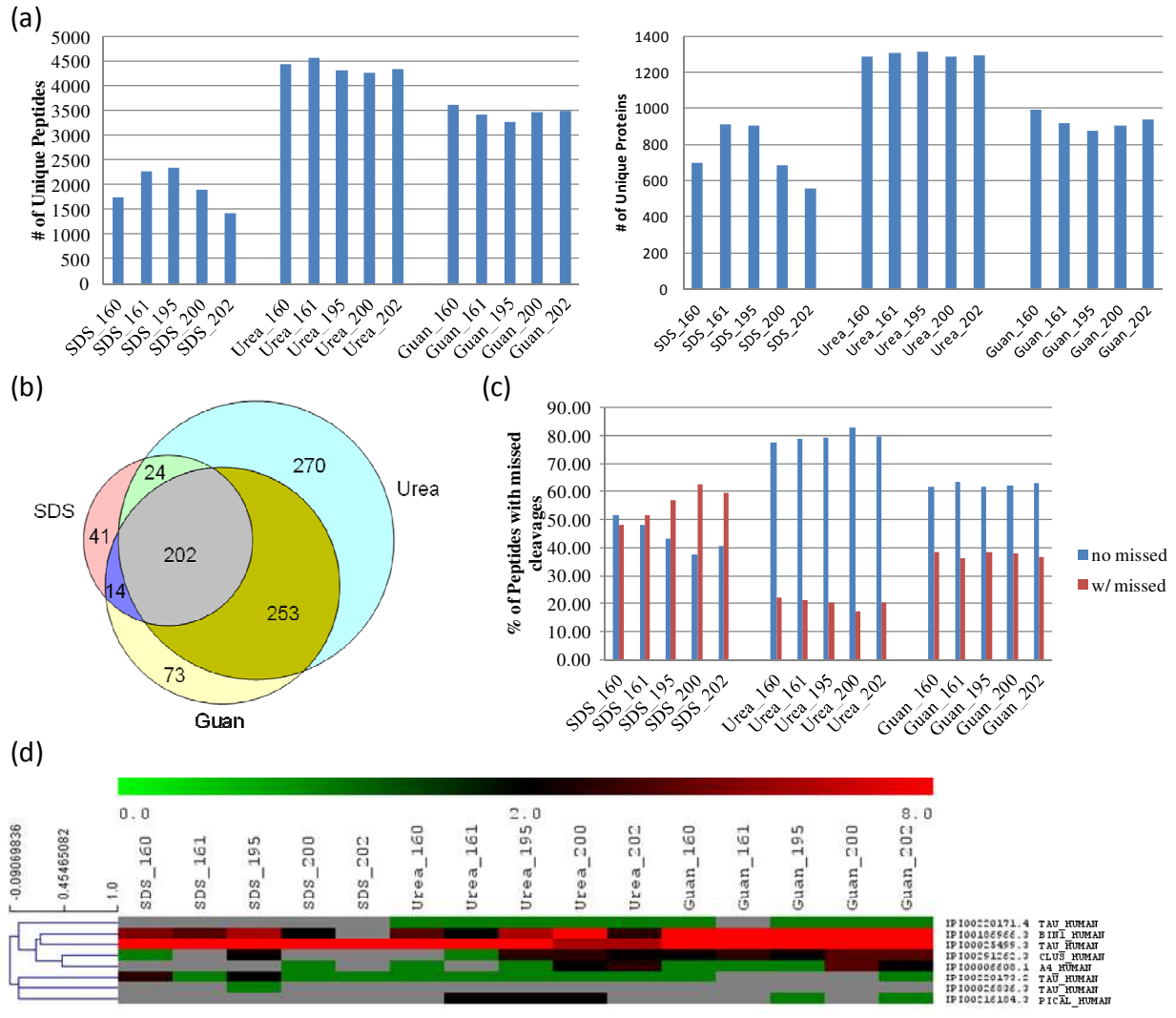
Supplementary Materials. Methods.

Comparison of Protein Extraction Protocols

It was important for this study to choose the most efficient protein extraction protocol for the Alzheimer's disease brain tissues from the currently-established proteomics sample preparation approaches. We compared three commonly used protein extraction protocols: (i) 8M urea with 2% SDS (labeled as SDS in Figure S1), (ii) 8M urea (labeled as Urea in Figure S1), (iii) 6M guanidine (labeled as Guan in Figure S1). Brain samples from five subjects were used to obtain statistically meaningful results for each protocol. The five subjects are fairly representative for the main study, including three Alzheimer's disease patients (labeled as 195, 200, and 202 in Figure S1) and two pathology-free controls (labeled as 160 and 161 in Figure S1). The derived tryptic peptide samples were adjusted to the same concentration before analyzed by reverse-phased LC coupled online with an LTQ Orbitrap mass spectrometer. Three criteria were used to assess the protocols. First, we evaluated the peptide and protein coverage in each protocol (see Figure S1.a). The number of identified peptides and proteins in the 8M urea extraction protocol is considerably higher than that of the other two protocols. About 1300 proteins are identified consistently across all five samples extracted by 8M urea. The 8M urea/2% SDS protocol yields the fewest peptide and protein identifications with the largest variation across the five samples. We also looked into the overlap of the proteins identified in each protocol to eliminate the possibility that a given protocol recovers quite a different portion of proteome from the others. Here, we only counted the proteins that were identified in all five samples to represent the protein recovery of each protocol. Figure S1.b clearly shows that these three protocols had good overlap while the urea protocol stands apart from the other two with the best recovery of the proteome. There are a total of 128 proteins that were not identified in

the urea protocol which account for 15% of all the 877 proteins. Thus, the 8 M urea protocol proves to be most efficient in terms of the proteome coverage. Second, we assessed how many identified peptides containing missed cleavages (see Figure S1.c). The missed cleavage refers to the presence of amino acid sequence “Lys-Xxx” or “Arg-Xxx” in the middle of a tryptic peptide while Xxx is neither Pro nor the *C*-terminus of a protein, and such peptides are less reliable to be used for protein quantitation. On average, peptides with missed cleavages account for 20% of all the peptides identified in a urea-extracted sample, 38% in a guanidine-extracted sample, and 56% in a urea/SDS-extracted sample. The 20% observed in the 8M urea protocol is in consensus with typical proteomics datasets acquired within the laboratory and it is suitable for the quantitation purpose, which makes the urea protocol again more favored than the other two. Lastly, we looked for the detection of proteins that are known to be closely related to Alzheimer’s disease (see Figure S1.d), including Amyloid precursor protein (A4) (Phiel, C. J., Wilson, C. A., Lee, V. M. Y. & Klein, P. S. GSK-3 regulates production of Alzheimer's disease amyloid- peptides. Nature 423, 435–439 (2003)), Bridging integrator (BIN1), Clusterin (CLUS), Phosphatidylinositol-binding clathrin assembly protein (PICAL) (Harold, D.; Abraham, R.; Hollingworth, P.; Sims, R.; Gerrish, A.; Hamshere, M.L.; et al. Genome-wide association study identifies variants at CLU and PICAL associated with Alzheimer's disease. Nat. Genet. 2009, 41, 1088-1093.), and Tau protein (TAU) (Avramopoulos, D. Genetics of Alzheimer’s disease: recent advances. Genome Med. 1, 34 (2009)). Figure S1.d uses the summed peptide spectral counts of the selected proteins in each sample to plot the heat map. The SDS protocol yields a poor coverage on A4, CLUS, PICAL and two isoforms of TAU; the guanidine protocol yields a poor coverage on PICAL and two isoforms of TAU; the urea protocol is the winner again with the best coverage of these selected proteins and only one isoform of TAU not detected. Taken all the aforementioned factors into consideration, the 8M urea protocol provides the most informative snapshot of the proteome for our quantitative proteomics study of Alzheimer’s disease, and was chosen for the main experiment presented in this paper.

Figure S1. Comparison of three protein extraction protocols. SDS refers to the protocol using 8M urea with 2% SDS and10 mM DTT in 50 mM Tris-HCl buffer (pH 7.6); Urea refers to the protocol using 8M urea with 10 mM DTT in 50 mM NH_4HCO_3 buffer (pH 7.8); Guan refers to the protocol using 6M Guanidine-HCl with 10 mM DTT in 50 mM NH_4HCO_3 buffer (pH 7.8). The numbers, 160, 161, 195, 200 and 202, indicate the sample ID. 160 and 161 are from pathogen-free subjects, while 195, 200, and 202 are from subjects with Alzheimer's disease. Panel (a) shows the number of unique peptides and proteins identified in each sample. Panel (b) illustrates the overlap of the identified proteins in each protocol. Please note that only proteins that were identified in all five samples prepared with a particular protocol are plotted in the Venn diagram to represent the result of that protocol. Panel (c) presents the percentage of identified peptides containing missed cleavages in each sample. Panel (d) depicts the recovery of the selected pathogen-associated proteins in each sample.



Some Details of Accurate Mass and Time Tag Technology

Creation of the database of the AMT tags of the tryptic peptides of the proteins present in the human brain cortex. 30 µg aliquots were pooled together into two separate pools (10 samples into control pool and 10 samples into AD pool) and subjected to LC fractionation by strong cation exchange (SCX) chromatography on a 200 mm × 2.1 mm Polysulfoethyl A column (PolyLC, Columbia, MD) preceded by a 10 mm × 2.1 mm guard column, using a flow rate of 0.2 mL/min. LC separations were performed using an Agilent 1100 series HPLC system (Agilent, Palo Alto, CA). Mobile phase solvents consisted of (A) 10 mM ammonium formate, 25% acetonitrile, pH 3.0 and (B) 500 mM ammonium formate, 25% acetonitrile, pH 6.8. Once loaded, isocratic conditions at 100% A were maintained for 10 min. Peptides were separated by using a gradient from 0-50% B over 40 min, followed by a gradient of 50-100% B over 10 min. The gradient was then held at 100% solvent B for another 10 min. Following lyophilization, all thirty fractions for each of the two pools collected during this gradient were dissolved in 25 mM ammonium bicarbonate and stored at -80^{O}C. Each SCX fraction was analyzed using LC-MS/MS platform described below. One survey MS scan was followed by MS/MS fragmentation spectra of top 10 most abundant ions.

MS/MS spectra were preprocessed using DeconMSn [53] and DtaRefinery [54] tools to assign the right monoisotopic peak and remove the systematic errors in mass measurement accuracy. The SEQUEST software (Thermo Fisher, San Jose, CA) was used to search the MS/MS data against the human UniProtKB/Swiss-Prot database release 15.10. Trypsin cleavage specificity was required for all of the considered peptides. The Xcorr, DeltaCn2 and dM criteria were optimized to give the maximum number of non-redundant peptide identifications, while maintain the FDR of non-redundant peptide identification estimated using reverse database approach [55] below 1%. The optimized criteria that were used to filter raw

SEQUEST results as follows: 1) deviation of parent ion mass (dM) < 3 ppm, Xcorr ≥ 1.4 and DeltaCn2 ≥ 0.1 for charge state +1; 2) dM < 2 ppm, Xcorr ≥ 1.5 and DeltaCn2 ≥ 0.11 for charge state +2; 3) dM < 2 ppm, Xcorr ≥ 2.0 and DeltaCn2 ≥ 0.12 for charge state +3; 4) dM < 2.5 ppm, Xcorr ≥ 1.4 and DeltaCn2 ≥ 0.17 for charge state +4. In total, the constructed database with identifiable peptides along their theoretical mass and observed elution time contained 36,965 unique peptide sequences corresponding to 5,369 proteins (72% identified with two or more peptides).

LC-MS and LC-MS/MS analyses. Each individual sample and each SCX fraction was analyzed with an automated custom-built capillary HPLC system coupled online to an LTQ Orbitrap mass spectrometer (Thermo Fisher, San Jose, CA) by using an electrospray ionization interface. The reversed phase capillary column was prepared by slurry packing 3-µm Jupiter C18 particles (Phenomenex, Torrence, CA) into a 75 µm i.d. × 65 cm fused silica capillary (Polymicro Technologies, Phoenix, AZ). The mobile phase solvents consisted of (A) 0.2% acetic acid and 0.05% TFA in water and (B) 0.1% TFA in 90% acetonitrile. An exponential gradient was used for the separation, which started with 100% A, and gradually increased to 60% B over 100 min. The instrument was operated in a data-dependent mode with an m/z range of 400-2000. One survey MS scan at 100K resolution was followed by MS/MS fragmentation spectra of top 10 most abundant ions for the datasets used for building the AMT database.

Peptide identification and quantitation. The analyses of quantitative LC-MS datasets were performed as previously described in details [56]. LC-MS feature detection was performed by deisotoping of individual MS scans using Decon2LS [57] followed by grouping monoisotopic peaks constituting LC profile and assignment of LC-MS features to peptides using VIPER [58], by matching their elution time and monoisotopic mass with 2% of normalized elution time and 2 ppm tolerances, respectively. False peptide identifications are likely to be present only in one or

few samples, thus to get a highly confident list of identifications we discarded all the peptides that are not fully present (10 out of 10) in either of two groups of subjects. False discovery rate of peptide identification for AMT approach was estimated using 11 Da shift approach [42] and was contained just below 1%. Peptide arbitrary abundance was measured using label-free approach as a sum of intensities of all monoisotopic ions constituting the LC-MS feature.