NEUROSCIENCE RESEARCH NOTES

OPEN ACCESS | DATA NOTES ISSN: 2576-828X

Data on multiple post-translational modifications in Alzheimer's disease

Sayali Chandrashekhar Deolankar, Shashanka G. Koyangana, Arun H. Patil [#], Yashwanth Subbannayya, Prashant Kumar Modi, and T. S. Keshava Prasad *

Center for Systems Biology and Molecular Medicine, Yenepoya Research Centre, Yenepoya (Deemed to be University), Mangalore, India.

[#] Current address: Department of Pathology, Division of Cardiovascular Pathology, Johns Hopkins University School of Medicine, 720 Rutland Avenue, Ross Rm 626, Baltimore, MD.
* Correspondence: keshav@yenepoya.edu.in; Tel: +91-9972250102

Received: 9 February 2022; **Accepted:** 23 May 2022; **Published:** 10 June 2022
Edited by: Indranath Chatterjee (Tongmyong University, South Korea)
Reviewed by: Jhinuk Chatterjee (PES University, India); Lilach Soreq University College London, UK)
https://doi.org/10.31117/neuroscirn.v5i2.153

ABSTRACT: This article describes the data obtained for global post-translational modifications (PTMs) profiled for Alzheimer's Disease (AD) from two distinct human brain regions and one cerebrospinal fluid (CSF) sample. The PTM profiling was performed to identify phosphorylation, O-GluNAcetylation, methylation, acetylation and citrullination using three publicly available LC-MS/MS raw data sets (PRIDE ID: PXD004010, PXD002516, PXD004863). A total of 1,857 PTM harbouring proteins with 4,961 unique post-translationally modified peptides were identified. Among the modified peptides, 75 corresponded uniquely to proteins identified from CSF samples. The data is related to the research article "Dissecting Alzheimer's disease molecular substrates by proteomics and discovery of novel post-translational modifications (PTMs)".

Keywords: post-translational modification; Alzheimer's disease; multi-OMICS; proteomics; PTM biology; bioinformatics;

1.0 BACKGROUND

The data represents the discovery of several post-translational modifications (PTMs) in Alzheimer's disease (AD) by the reanalysis of previously acquired mass spectrometry data. These mass spectrometry data were profiled previously from AD brain tissues (Adav et al., 2016; Li et al., 2016) and CSF (Hansson et al., 2017) samples and submitted to a public repository (PXD004010, PXD002516, and PXD004863). Original articles by Adav et al., Li et al., and Hansson et al. have focussed only on describing expressed proteins identified from these samples. However, mass spectrometry data can also be mined further to identify PTMs of proteins. It is also known that PTMs play a major role in the functional activities of proteins. Therefore, a detailed reanalysis of the above-mentioned data sets was carried out to discover PTMs and performed scientific interpretation of PTMs data by Deolankar et al. (2019). A total of 4,045,380 MS/MS spectra were analyzed, which resulted in 5,417 unique peptides belonging to 2,309 proteins. When compared to literature total of 52 proteins were identified related to the AD pathway. Among these, 18 were unique to AD, and 28 shared pathways profile with Huntington's and

Parkinson's disease. The list of proteins unique to the AD pathway, with the modified peptides and the site of the modification, is described in **Table 1**. Supplementary **Tables S1** and **S2** list functional enrichment results for the proteins forming protein-protein interaction and co-expression. The identified clusters and the adopted methodology are depicted in **Figure 1**.

Table 1: List of modified peptides with PTMs for AD pathway

Gene Symbol (Master Protein Accessions)	Annotated Sequence	Modifications
GAPDH (NP_001276674.1)	LISWYDNEFGYSNr	R14(Methyl)
	VPTANVSVVDLTCr	R14(Methyl)
	FHGTVkAENGK	K6(Acetyl)
	VVDLMAHMASkE	M5(Oxidation); K11(Methyl)
	GALQNIIPASTGAAk	K15(Methyl)
	AVGkVIPELNGK	K4(Acetyl)
	VPTANVSVVDLtCrLEKPAK	T12(Phospho); R14(citru)
	TVDGPSGkLWR	K8(Acetyl)
	LVINGNPITIFQErDPSK	R14(citru)
	GALQNIIPAsTGAAK	T11(Phospho)
	IISNASCTTNCLAPLAk	K17(Methyl)
APOE (NP_001289617.1)	AATVGSLAGQPLQEr	R15(citru)
	LGADMEDVCGr	R11(citru)
	SELEEQLTPVAEETr	R15(Methyl)
APP (NP_958817.1; NP_000475.1)	VEsLEQEAANER	S3(Phospho)
	EVCSEQAETGPCr	R13(citru)
	LVFFAEDVGSNk	K12(Methyl)
ATP2A1(NP_775293.1)	FLEYETDLtFVGVVGMLDPPR	T9(Phospho); M16(Oxidation)
ATP2A2 (NP_733765.1)	rIGIFGQDEDVTSK	R1(citru)
	EFDELNPsAQR	S8(Phospho)
BAD (NP_004313.1)	RMsDEFVDSFK	S3(Phospho)
	SRsAPPNLWAAQR	S1(Phospho)
CALM1||*CALM2*||*CALM3* (XP_006720321.1; NP_001292553.1; NP_005176.1; XP_006720321.1; NP_001292553.1)	MKDTDSEEEIr	R11(Methyl)
	DGNGYIsAAELr	S7(Phospho)
	DGDGTITTk	K9(Methyl)
	DTDSEEEIrEAFr	R13(Methyl)
	MkDTDSEEEIR	M1(Oxidation); K2(Acetyl)
	EAFSLFDkDGDGTITTK	K8(Acetyl)
	MKDTDSEEEIr	R11(Methyl)
	ADQLtEEQIAEFK	T5(Phospho)
	DGNGYIsAAELr	S7(Phospho)
	DGDGTITTk	K9(Methyl)
	DTDSEEEIrEAFr	R13(Methyl)
	EAFSLFDk	K8(Methyl)
CAPN1(NP_001185797.1)	YLGQDYEQLr	R10(citru)

CDK5(NP_004926.1)	VRLDDDDEGVPSSALr	R2(citru)
	DLKPQNLLINr	R11(citru)
GRIN2A(XP_016878661.1)	SPDFNLTGsQSNMLK	S9(Phospho)
GSK3B||*GSK3A*(NP_002084.2; NP_063937.2; XP_006713673.1)	IQAAAStPTNATAASDANTGDR	T7(Phospho);
	LrYFFYSSGEK	R2(citru)
	GEPNVSyICsR	S10(Phospho)
IDE (NP_001309722.1)	FIIQSEKPPHYLESr	R15(citru)
MAPT (NP_058525.1; XP_005257428.1; XP_005257427.1; XP_005257419.1; XP_005257421.1)	SrLQTAPVPMPDLK	R2(citru)
	STPTAEAEEAGIGDTPsLEDEAAGHVTQAR	S1(Phospho)
	SGYSSPGSPGTPGSrSR	R17(citru)
	IATPrGAAPPGQK	R5(citru)
	kLDLSNVQSK	K1(Acetyl)
	TPPAPKtPPSSGEPPK	T7(Phospho)
	SPVVSGDTsPR	S9(Phospho)
	SGYssPGsPGTPGSR	S5(Phospho)
	SrLQTAPVPMPDLK	R2(citru); M10(Oxidation)
	AEEAGIGDTPSLEDEAAGHVTQEELr	R26(citru)
	SGDrSGYSSPGSPGTPGSR	R4(citru)
	KDQGGYtMHQDQEGDTDAGLK	Y6(Phospho); M8(Oxidation)
	TDHGAEIVYKsPVVsGDtSPR	S11(Phospho); S15(Phospho)
	TPPAPKtPPSSGEPPK	T7(Phospho)
	SkIGsTENLK	S5(Phospho)
	CGsLGNIHHKPGGGQVEVK	S3(Phospho)
	AEEAGIGDTPSLEDEAAGHVTQArMVSK	R24(citru)
	TPsLPtPPTREPK	S3(Phospho)
	IGsTENLK	S3(Phospho)
	AEEAGIGDTPsLEDEAAGHVTQAr	S11(Phospho);
	DQGGYTmHQDQEGDtDAGLKAEEAGIGDTPSLEDEAAGHVTQAR	M7(Oxidation); T15(Phospho)
	LQTAPVPMPDLk	K12(Methyl)
	STPTAEAEEAGIGDTPsLEDEAAGHVTQAr	R30(Methyl)
	VQIVYkPVDLSK	K6(Acetyl)
	DQGGYtMHQDQEGDTDAGLKEsPLQTPTEDGSEEPGSETSDAK	Y5(Phospho); M7(Oxidation);
	IGsLDNITHVPGGGNk	S3(Phospho)
	VAVVRtPPK	T6(Phospho)
	SPVVSGDTsPR	T8(Phospho)
	SGYssPGsPGTPGSR	S5(Phospho)
	IGSLDNITHVPGGGNkK	K16(Acetyl)
	TPPAPKtPPSSGEPPK	T7(Phospho)
PPP3CA (NP_001124163.1; NP_000935.1)	DAMPSDANLNSINk	K14(Methyl)
	rDAMPSDANLNSINK	R1(citru)
	LTAkEVFDNDGKPR	K4(Acetyl)
PPP3CB (NP_001135825.1)	QTLQSATVEAIEAEk	K15(Methyl)

MAPK3 (NP_002737.2)	rTEGVGPGVPGEVEMVK	R1(citru); M15(Oxidation)
	LKELIFQETAr	R11(citru)
	IADPEHDHTGFLTEYVAtr	T18(Phospho); R19(citru)
PSEN2 (NP_000438.2)	TSLMSAEsPTPR	S8(Phospho)
MAPK1 (NP_002736.3)	VADPDHDHTGFLTEyVATr	R19(citru)

Footnote: Phospho- Phosphorylation, Citu- Citrullination, Methyl- methylation, Acetyl -Acetylation.

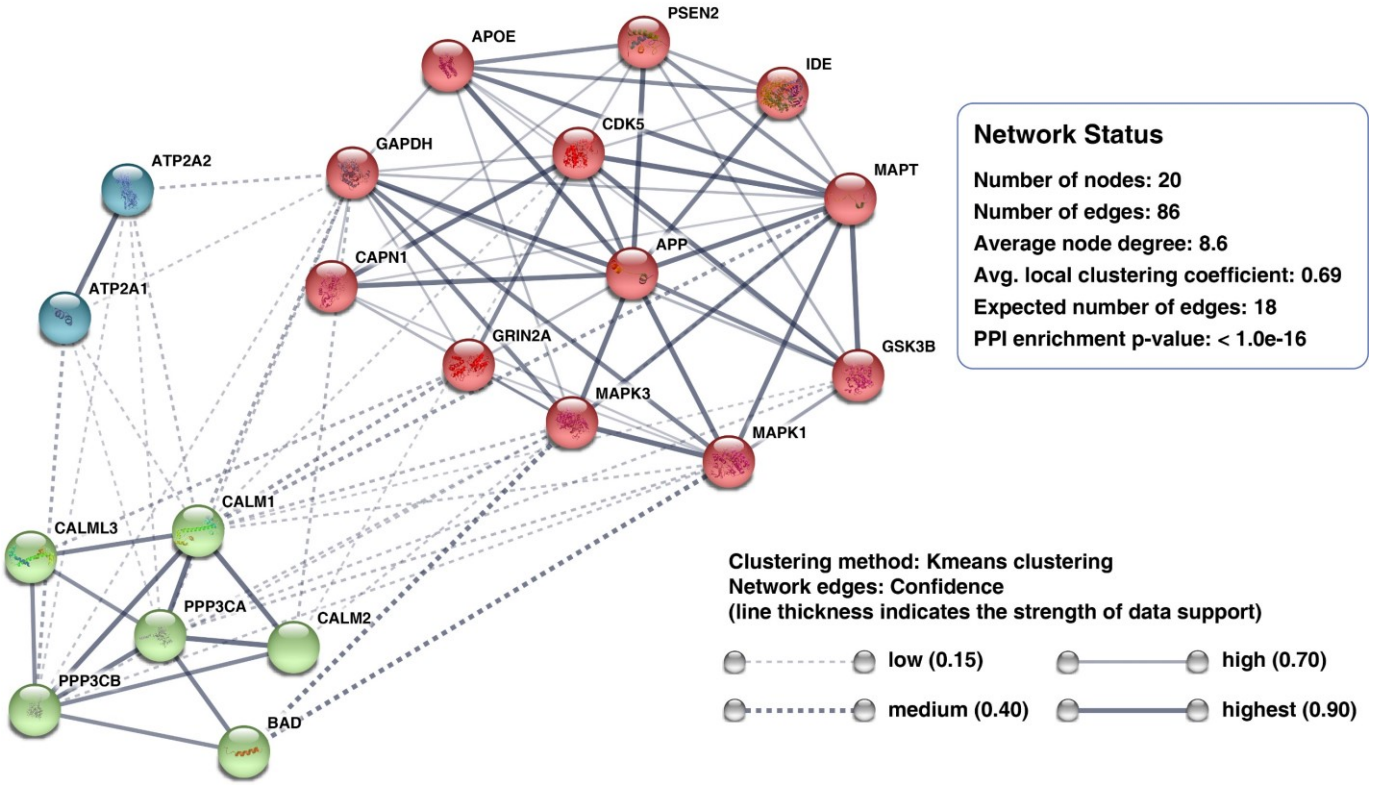

Figure 1. Protein-protein interaction and co-expression analysis network for proteins with PTMs from Alzheimer's disease pathway. The colours of the nodes represent the protein-protein interaction network for the selected proteins formed by k-means clustering methods with cluster limits set to 3 clusters. Each colour represents the group of proteins forming the clustered network.

2.0 MATERIALS AND METHODS

2.1 Selection of data source and database search

The high-resolution mass spectrometry-derived raw data were obtained from the PRIDE repository with the specific parameters as previously described (Deolankar et al., 2019). Global PTM search was performed against NCBI Human RefSeq (release 89) protein database as reference using SEQUEST-HT through Proteome Discoverer, Version 2.2 (PD 2.2) software suite (Thermo Scientific, Bremen, Germany). The detailed workflow for PD 2.2 search parameters and PTMs with localized amino acid residues was mentioned previously (Deolankar et al., 2019). A detailed methodology explaining the selection of data sets for search and parameters used for protein and peptide identification is mentioned in Supplementary methods.

2.2 PTM-Profiling

The select PTMs with ptmRS site probability greater than 75% for each PTM were profiled and summarized using the PTM-Pro tool (Patil et al., 2018). **Figure 2** provides a summary of PTMs from the selected studies.

2.3 Data analysis and comparison with available data resources

The functional enrichment for biological processes, molecular function and cellular components was carried out for PTM harbouring proteins from Alzheimer's

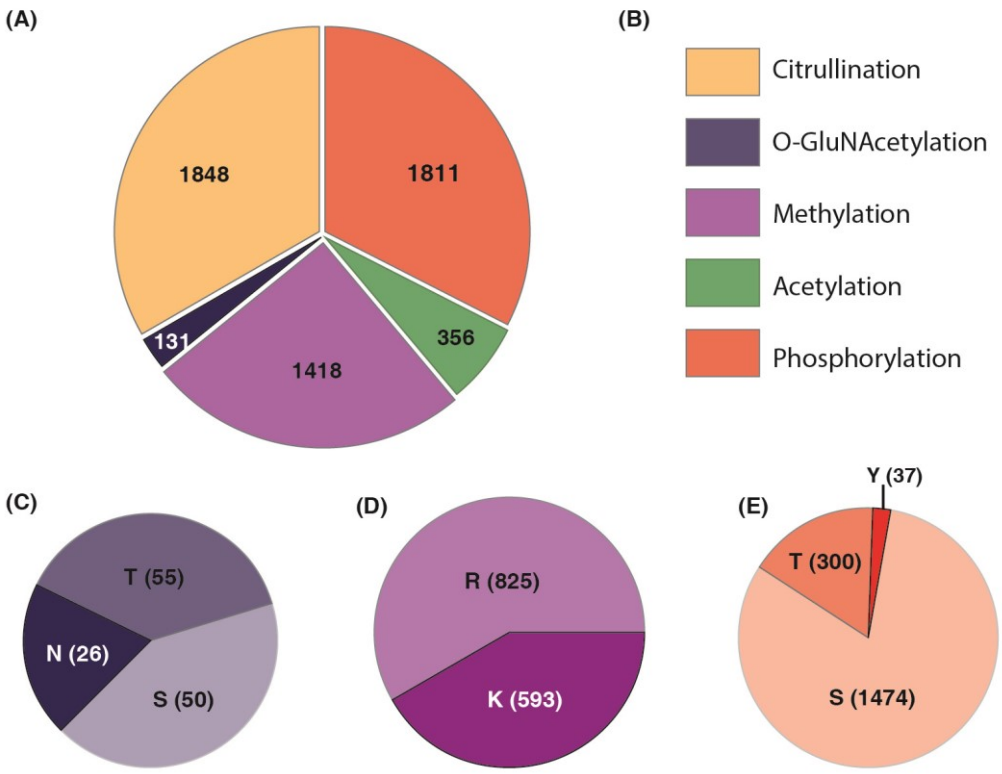

Figure 2. Distribution of PTMs from brain tissues and CSF. **(A)** The chart is describing the number of PTM identified in this study; **(B)** Colour code for describing the mentioned PTMs; and a summary of events identified as **(C)** O-GluNAcetylation of asparagine, serine and threonine, **(D)** methylation of lysine and arginine and **(E)** phosphorylation of serine, threonine and tyrosine.

disease pathway using STRING network analysis for multiple proteins (version 11.0) (Szklarczyk et al., 2019). Further, the analysis was extended to the identification of enriched pathways from the KEGG pathway (Kanehisa & Goto, 2000), Reactome databases, Protein family (Pfam) database and SMART domain identification analysis by using an inbuilt function from the STRING DB tool. Further, the resultant peptides (derived from PD searches) with select PTMs were compared with the PhosphoSitePlus database as described earlier (Deolankar et al., 2019).

3.0 DATA FORMAT & AVAILABILITY

3.1 Data accessibility

Analyzed supplementary data is appended to this article. The PD result files and MGF files for global PTM-profiling are available through the ProteomeXchange consortium with ID-PXD014042.

3.2 Related research article

The raw data presented here were previously generated and analyzed for a published article entitled "Dissecting Alzheimer's Disease Molecular Substrates by Proteomics and Discovery of Novel Post-translational Modifications." (Deolankar et al., 2019).

4.0 RELEVANCE

- This data provides a catalogue of the global post-translational modification profile from brain regions and CSF in Alzheimer's disease (AD).
- The data provides data on several PTMs on proteins associated with the AD pathway, which may serve as a resource for potential biomarkers or drug targets of AD.
- The most common modified peptides unique to Alzheimer's patients might be useful as prospective liquid biopsy biomarkers for the diagnosis of AD.

Supplementary Materials: The following files are available online: **Supplementary Table S1:** STRING DB analysis performed with a list of modified proteins from the Alzheimer's disease pathway; **Supplementary Table S2:** Protein-protein interaction and co-expression results from the STRING DB analysis performed with a list of modified proteins from the AD pathway; **Supplementary methods:** Provides a brief description of the methodology used for selection of data source, peptide and protein search using PD and PTM-profiling using the PTM-Pro tool.

Acknowledgements: The authors thank Karnataka Biotechnology and Information Technology Services (KBITS), Government of Karnataka, for the support to the Center for Systems Biology and Molecular Medicine (CSBMM) at Yenepoya (Deemed to be University), Mangalore under the Biotechnology Skill Enhancement Programme in Multiomics Technology (BiSEP GO ITD 02 MDA 2017). Sayali C. Deolankar is a recipient of a Senior Research Fellowship from the Indian Council of Medical Research (ICMR), Government of India. Arun H. Patil has been a faculty funded under the BiSEP program, Government of Karnataka.

Author Contributions: TSKP conceived and designed the experiments; SCD and PM performed the experiments; SCD, AKP and YS analyzed the data; AKP and SKG contributed to analysis tools; SCD wrote the paper, and TSKP revised the final version of the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

Adav, S. S., Gallart-Palau, X., Tan, K. H., Lim, S. K., Tam, J. P., & Sze, S. K. (2016). Dementia-linked amyloidosis is associated with brain protein deamidation as revealed by proteomic profiling of human brain tissues. *Molecular Brain, 9*, 20. https://doi.org/10.1186/s13041-016-0200-z

Deolankar, S. C., Patil, A. H., Koyangana, S. G., Subbannayya, Y., Prasad, T. S. K., & Modi, P. K. (2019). Dissecting Alzheimer's Disease Molecular Substrates by Proteomics and Discovery of Novel Post-translational Modifications. *OMICS, 23*(7), 350-361. https://doi.org/10.1089/omi.2019.0085

Hansson, K. T., Skillback, T., Pernevik, E., Kern, S., Portelius, E., Hoglund, K., . . . Gobom, J. (2017). Expanding the cerebrospinal fluid endopeptidome. *Proteomics, 17*(5), 1600384. https://doi.org/10.1002/pmic.201600384

Kanehisa, M., & Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research, 28*(1), 27-30. https://doi.org/10.1093/nar/28.1.27

Li, Y., Wang, X., Cho, J. H., Shaw, T. I., Wu, Z., Bai, B., . . . Peng, J. (2016). JUMPg: An Integrative Proteogenomics Pipeline Identifying Unannotated Proteins in Human Brain and Cancer Cells. *Journal of Proteome Research, 15*(7), 2309-2320. https://doi.org/10.1021/acs.jproteome.6b00344

Patil, A. H., Datta, K. K., Behera, S. K., Kasaragod, S., Pinto, S. M., Koyangana, S. G., . . . Prasad, T. S. K. (2018). Dissecting Candida Pathobiology: Post-Translational Modifications on the Candida tropicalis Proteome. *OMICS, 22*(8), 544-552. https://doi.org/10.1089/omi.2018.0093

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., . . . Mering, C. V. (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research, 47*(D1), D607-D613. https://doi.org/10.1093/nar/gky1131