



APPLICATION NOTE

Deep proteomic profiling of human cerebrospinal fluid by data-independent acquisition mass spectrometry coupled to bead-based protein enrichment

Authors: Amanda Weaver, Hayden Tomazin, and Joseph Crook

Highlights

- Bead-based protein enrichment coupled to DIA mass spectrometry applied to human CSF on the Orbitrap Excedion Pro / Vanquish Neo UHPLC platform.
- Average 1.85 – fold increase in protein group identifications versus a matched non-enriched preparation from the same single donor specimens, reproducible across all eight donors.
- Differentially abundant proteins between high- and low-glucose strata define functionally distinct extracellular matrix / humoral immunity and neural projection / cell adhesion signatures.

Introduction

Cerebrospinal fluid (CSF) is the clear, low-protein fluid that surrounds the brain and spinal cord. Continuously produced by the choroid plexus and turned over several times daily, it provides mechanical cushioning of the central nervous system (CNS), supports metabolic homeostasis, and clears soluble waste from the brain parenchyma¹. Because CSF exchanges directly with brain extracellular fluid, its molecular composition reflects the state of the underlying neural tissue with a fidelity plasma cannot match, making CSF analysis a cornerstone of laboratory diagnostics in neurodegenerative disease, CNS infection, intracranial hemorrhage, and CNS malignancy².

Despite this value, characterization of the CSF proteome by liquid chromatography-mass spectrometry (LC-MS) has historically been demanding. Total protein concentration in CSF is roughly 0.2–0.8 mg/mL, about one percent of plasma, and abundance spans many orders of magnitude, with a few highly abundant carrier proteins (albumin, immunoglobulins, transferrin) suppressing detection of the lower-abundance signatures usually of greatest interest³. Classical remedies—extensive offline fractionation, immunoaffinity depletion of high-abundance proteins, and pooling of large volumes—

improve depth but raise per-sample cost, reduce throughput, and limit applicability to single-donor specimens⁴.

Recent advances have improved the sensitivity, throughput, and reproducibility of single-shot CSF analysis. Faster-scanning Orbitrap instrumentation, data-independent acquisition (DIA), and advanced library search algorithms now permit quantitative profiling of complex matrices at low microgram-to-nanogram input⁵. In parallel, antibody-free bead-based enrichment strategies compress protein dynamic range through selective capture on physiochemically functionalized beads, enabling identification of low-abundance proteins from small, neat volumes of plasma and other fluids without immunodepletion⁶.

Here we apply a bead-based enrichment workflow coupled to DIA mass spectrometry to human CSF. Eight single-donor specimens from Alzheimer's patients, stratified by clinical glucose measurement, were processed in parallel using a standard in-solution iST and a bead-based ENRICH-iST workflow (both PreOmics). Peptides were analyzed in DIA mode on an Orbitrap Excedion Pro mass spectrometer coupled to a Vanquish Neo ultra-high performance liquid chromatography (UHPLC) system. Differential abundance between the two glucose strata was characterized in the enriched preparation and interpreted by gene ontology over-representation analysis.

Methods

Samples

Eight de-identified human Alzheimer's CSF specimens (Discovery Life Sciences) were used. Four had clinical glucose measurements below the laboratory reference range (50–80 mg/100 mL) and four above it. Glucose stratification was the sole selection criterion, applied regardless of clinical indication, though characteristics such as age, gender, and race were matched where possible.

Sample preparation

Each specimen was split into two aliquots. The first (10 μ L) was processed with the iST kit (PreOmics) per the manufacturer's protocol: single-step denaturation, reduction, and alkylation, followed by trypsin/LysC digestion and C18 spin-column cleanup. The second (200 μ L) was processed with the ENRICH-iST kit (PreOmics): incubation with EN-BIND buffer and pre-washed paramagnetic EN-BEADS, magnetic capture, three on-bead washes, on-bead denaturation/reduction/alkylation, on-bead trypsin/LysC digestion, and C18 cleanup. Peptides from both workflows were dried and stored at -80°C until LC-MS analysis.

LC-MS/MS analysis

Dried peptides were reconstituted in 16 μ L 4% ACN/0.1% formic acid, quantified using the Pierce Quantitative Fluorometric Peptide Assay (Thermo Fisher Scientific), and 200 ng of each was injected in technical duplicate, yielding 32 injections (8 specimens $\times$ 2 preparations $\times$ 2 technical replicates). Separation used a Vanquish Neo UHPLC system with a 75 $\mu\text{m} \times 500$ mm EASY-Spray PepMap Neo column (both Thermo Fisher Scientific) over a 75-minute gradient from 4% to 80% acetonitrile in 0.1% formic acid. Eluate was directed to an Orbitrap Excedion Pro mass spectrometer (Thermo Fisher Scientific) with the EasySpray NG

source, operated in data-independent acquisition (DIA) mode using collision energy of 28%, 30,000 resolution, and a 380–980 m/z scan range with 8 m/z isolation windows overlapping 1 m/z .

Data analysis

DIA raw files were processed in Spectronaut (Biognosys, version 20.5.260227.92449) and searched against a predicted Homo sapiens spectral library generated from the UniProt reference proteome FASTA (retrieved 30 January 2026) using DirectDIA+, a library-free search method, with an FDR of 0.01, fixed carbamidomethyl modification, and variable acetyl (N-term) and oxidation modifications. Protein group identification and label-free quantification were performed in Spectronaut, and downstream statistical comparisons in GraphPad Prism (Dotmatics, version 10.5.0). Differential abundance analysis between the high- and low-glucose cohorts was performed on the enriched preparation only. Gene ontology over-representation analysis for the Biological Process and Molecular Function categories used the Gene Ontology Enrichment Analysis tool (PANTHER 19.0)⁷. A schematic of the complete workflow is shown in **Figure 1**.

Results and Discussion

Bead-based enrichment increases protein group identifications across all donors

Across all eight donors, the bead-based ENRICH-iST workflow identified more protein groups than the parallel neat preparation from the same specimen (**Figure 2**). The neat preparation identified an average of 945 protein groups per sample and the enriched preparation 1732, an approximate 1.85-fold increase in proteome depth at a fixed 200 ng peptide loading, reproducible across all eight donors and both glucose strata. This gain is consistent with dynamic range compression: selective on-bead capture shifts a portion of the most abundant

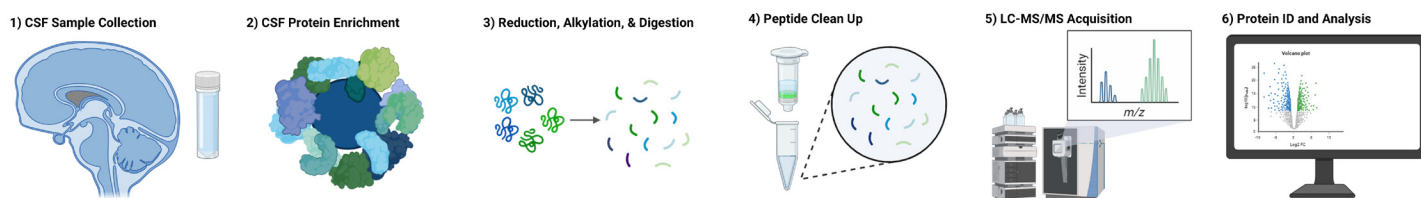


Figure 1. Schematic overview of the CSF discovery proteomics workflow. CSF specimens are collected and split into two aliquots. One aliquot is processed using a standard in-solution iST kit (PreOmics, neat preparation) and the other using a bead-based ENRICH-iST kit (PreOmics, enriched preparation). Both aliquots undergo reduction, alkylation, and trypsin/LysC digestion, followed by peptide clean-up. Peptides are analyzed by LC-MS/MS on a Vanquish Neo UHPLC coupled to an Orbitrap Excedion Pro mass spectrometer operating in DIA mode. DIA raw files are searched against a predicted Homo sapiens spectral library, and downstream statistical and functional analyses are performed in GraphPad Prism and the Gene Ontology Enrichment Analysis tool, respectively.

carrier proteins to the supernatant while retaining a broader population of lower-abundance proteins on the beads, letting the MS duty cycle sample more of the proteome at equivalent loading⁶. Its magnitude and consistency track with prior reports of bead-based enrichment in plasma and other low-protein fluids.

The depth was obtained from a single 75-minute gradient and one 200 ng injection per replicate, without offline fractionation or immunodepletion, with low replicate variability for both workflows, placing the per-sample timeframe within the range typical for cohort-scale studies. Unlike immunodepletion, bead-based enrichment preserves some carrier proteins such as albumin and the immunoglobulins, which can themselves carry information of interest and serve as the physiological vehicle for many low-abundance binders.

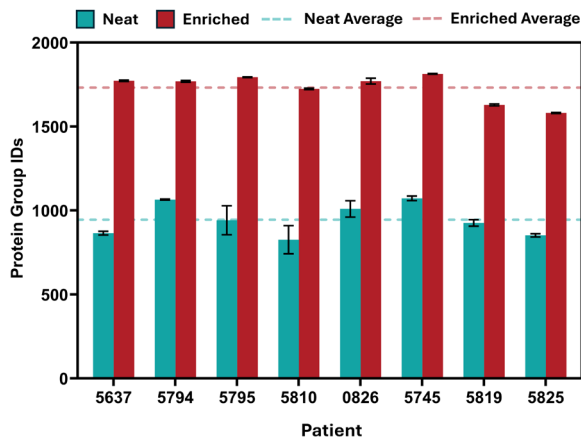


Figure 2. Protein group identifications by sample preparation method. Bars represent the mean number of protein group identifications across technical replicate injections ($n = 2$) for each donor sample ($n = 8$), prepared using either the neat iST workflow (teal) or the bead-based ENRICH-iST workflow (red). Error bars indicate the standard deviation between technical replicates. Dashed lines represent the cohort-wide average number of protein groups identified per sample for each preparation method (neat average = 945; enriched average = 1732). The enriched workflow yielded a 1.85 – fold improvement in protein group identifications on average and reproduced this gain in every donor in the cohort.

Differentially abundant proteins define distinct signatures between the glucose strata

Differential abundance analysis between the four high-glucose and four low-glucose donors was performed on the enriched preparation (Figure 3). 179 proteins were significantly elevated in the high-glucose cohort and 268 in the low-glucose cohort, spanning a broad fold-change range in both directions, with several candidates separated by more than two log2 units. Among the top hits by combined fold change and significance, ACAN, PEBP4, PAPLN, CKM, HRG, CRIPTO, GFRA3, KIF13A, and the immunoglobulin variable-region products IGLV8-61 and IGHV5-10-1 were elevated in the high-glucose group, while

PLA2G7, C1RL, HYAL1, MERTK, MRC1, MYL1, ACP1, TFRC, EHBP1L1, and EPHB2 were elevated in the low-glucose group.

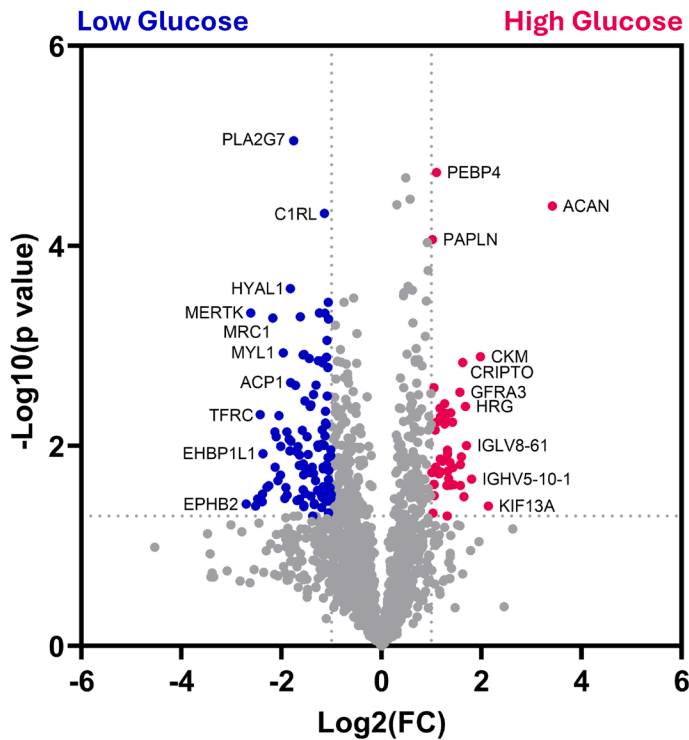


Figure 3. Differential abundance analysis between high-glucose and low-glucose CSF donors using the enriched preparation. Volcano plot of $-\log_{10}(\text{p value})$ versus $\log_2(\text{fold change})$ for protein groups quantified in both cohorts (high glucose $n = 4$, low glucose $n = 4$; enriched preparation only). Dotted lines mark the significance thresholds applied ($p \leq 0.05$ and $|\log_2 \text{FC}| \geq 1$). Proteins meeting both thresholds are colored pink (elevated in high-glucose) or blue (elevated in low-glucose); the top ten proteins by combined fold change and significance in each direction are labeled by gene symbol. The neat preparation was not included in this comparison.

Gene ontology enrichment points to distinct tissue origins for each stratum

Gene ontology over-representation analysis of the up-regulated set within each cohort (Figure 4) placed these proteins in coherent functional context. The high-glucose cohort was enriched for extracellular structure and humoral immunity terms, including extracellular matrix organization, external encapsulating structure, complement activation, extracellular matrix structural constituent, glycosaminoglycan binding, and heparin binding, mirroring its differentially abundant extracellular matrix constituents (ACAN, PAPLN), heparin- and glycosaminoglycan-binding proteins, immunoglobulin variable-region products, and the plasma protein HRG. This composition is consistent with a relative enrichment of plasma-derived and matrix-remodeling activity, a pattern associated in prior CSF studies with altered blood-brain barrier permeability and an attendant influx of plasma and complement proteins into

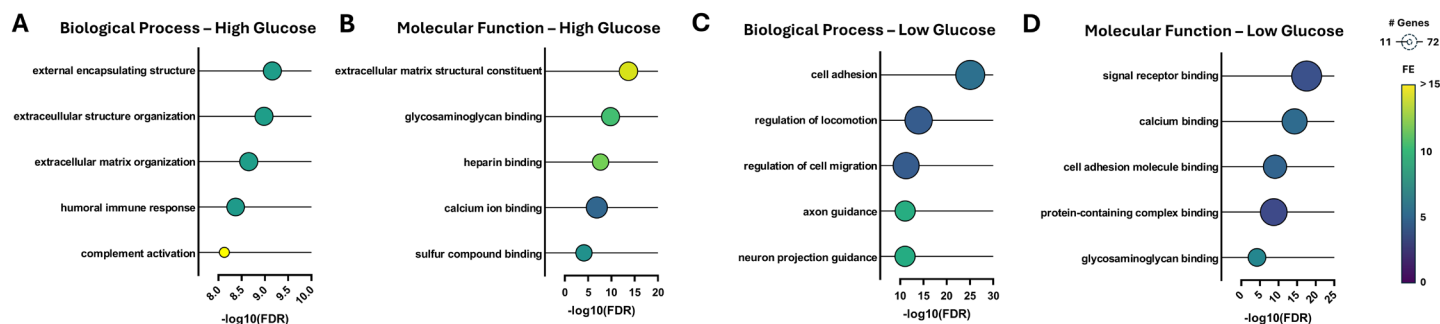


Figure 4. Gene ontology over-representation analysis for proteins differentially elevated in each glucose cohort. Top enriched terms for Biological Process (A, C) and Molecular Function (B, D) within the high-glucose (A, B) and low-glucose (C, D) cohorts. The x axis shows $-\log_{10}(\text{false discovery rate})$, marker diameter reflects the number of proteins contributing to each term, and marker color reflects the magnitude of fold enrichment (FE) according to the legend at right. Analyses were performed on the up-regulated protein set from each cohort identified in the enriched preparation (see Figure 3).

the CSF compartment⁸. The low-glucose cohort, by contrast, was enriched for cell adhesion, motility, and neural projection terms, including cell adhesion, regulation of cell migration, axon guidance, neuron projection guidance, signal receptor binding, and cell adhesion molecule binding, reflecting cell-adhesion molecules and ephrin and neurotrophic factor receptor family members (EPHB2, MERTK, GFRA3). This signature is more consistent with a parenchymal or cell-surface neural origin, and may reflect neural cell remodeling, synaptic turnover, or shedding of neuronal surface proteins into the CSF⁹.

Conclusion

We applied a bead-based enrichment workflow coupled to DIA mass spectrometry on the Orbitrap Excedion Pro platform to a

small set of human CSF specimens stratified by clinical glucose measurement. The enriched preparation produced an average 1.85-fold increase in protein group identifications relative to a matched non-enriched preparation, with low technical variability and a consistent gain across all eight donors. Differential abundance between the strata recovered functionally distinct signatures; extracellular matrix and humoral immunity in the high-glucose stratum, cell adhesion and neural projection in the low-glucose stratum, suggestive of differences in the relative contribution of plasma- and parenchyma-derived proteins. Larger disease-defined cohorts, alternative acquisition strategies and gradient lengths, paired plasma proteomics for cross-fluid validation, and targeted assays for proteins of interest are natural next steps.

REFERENCES

- Czarniak N, Kamińska J, Matowicka-Karna J, Koper-Lenkiewicz OM. Cerebrospinal Fluid-Basic Concepts Review. *Biomedicines*. 2023;11(5):1461. Published 2023 May 17. doi:10.3390/biomedicines11051461.
- Pingle SC, Lin F, Anekoji MS, et al. Exploring the role of cerebrospinal fluid as analyte in neurologic disorders. *Future Sci OA*. 2023;9(4):FSO851. Published 2023 Apr 4. doi:10.2144/fsoa-2023-0006.
- Schutzer SE, Liu T, Natelson BH, et al. Establishing the proteome of normal human cerebrospinal fluid. *PLoS One*. 2010;5(6):e10980. Published 2010 Jun 11. doi:10.1371/journal.pone.0010980.
- Schilde LM, Kösters S, Steinbach S, et al. Protein variability in cerebrospinal fluid and its possible implications for neurological protein biomarker research. *PLoS One*. 2018;13(11):e0206478. Published 2018 Nov 29. doi:10.1371/journal.pone.0206478.
- Kitata RB, Yang JC, Chen YJ. Advances in data-independent acquisition mass spectrometry towards comprehensive digital proteome landscape. *Mass Spectrom Rev*. 2023;42(6):2324-2348. doi:10.1002/mas.21781.
- Borràs E, Anastasi F, Pastor O, Suárez-Calvet M, Sabidó E. Enhanced proteome profiling of human cerebrospinal fluid using a commercial plasma enrichment strategy. *J Proteomics*. 2026;324:105575. doi:10.1016/j.jpro.2025.105575.
- Thomas PD, Ebert D, Muruganujan A, Mushayahama T, Albu LP, Mi H. PANTHER: Making genome-scale phylogenetics accessible to all. *Protein Sci*. 2022;31(1):8-22. doi:10.1002/pro.4218.
- Negro-Demontel L, Maleki AF, Reich DS, Kemper C. The complement system in neurodegenerative and inflammatory diseases of the central nervous system. *Front Neurol*. 2024;15:1396520. Published 2024 Jul 3. doi:10.3389/fneur.2024.1396520.
- Waldera-Lupa DM, Etemad-Parishanzadeh O, Brocksieper M, et al. Proteomic changes in cerebrospinal fluid from primary central nervous system lymphoma patients are associated with protein ectodomain shedding. *Oncotarget*. 2017;8(66):110118-110132. Published 2017 Nov 24. doi:10.18632/oncotarget.22654.