Journal of Equine Veterinary Science 34 (2014) 451–458

Contents lists available at ScienceDirect

Journal of Equine Veterinary Science

journal homepage: www.j-evs.com

Short Communication

Proteomic Characterization of Equine Cerebrospinal Fluid

Carolyn J. Broccardo PhD [a], Gisela Soboll Hussey DVM, PhD [b], Lutz Goehring DVM, PhD [c], Paul Lunn DVM, PhD [d], Jessica E. Prenni PhD [a,*]

[a] Proteomics and Metabolomics Facility, Colorado State University, Fort Collins, CO
[b] Department of Pathobiology and Diagnostic Investigation, Michigan State University, Lansing, MI
[c] Department of Clinical Sciences, Colorado State University, Fort Collins, CO
[d] College of Veterinary Medicine, North Carolina State University, Raleigh, NC

ARTICLE INFO

Article history:
Received 17 May 2013
Received in revised form 16 July 2013
Accepted 26 July 2013
Available online 4 October 2013

Keywords:
Equine
Cerebrospinal fluid
Proteomics
Normalized spectral abundance factor
Gene ontology

ABSTRACT

Cerebrospinal fluid (CSF) is a biofluid that is reflective of overall health. Although proteomic profiling of human CSF has been performed in the context of a variety of disease states, this report represents the first comprehensive proteomic analysis of equine CSF. A total of 320 proteins were confidently identified across six healthy horses, and these proteins were further characterized by gene ontology terms mapped in UniProt, and normalized spectral abundance factors were calculated as a measure of relative abundance. Theses results provide an optimized protocol for analysis of equine CSF and lay the groundwork for future studies involving the study of equine CSF in the context of pathogenic disease states.

1. Introduction

Cerebrospinal fluid (CSF) is a highly informative biofluid that can be mined for candidate biomarkers of disease [1,2]. However, studies of CSF are challenging because of the invasiveness of sample collection in humans and animals and the difficulty in obtaining sufficient sample amount from small animal models. Large animal models present a unique opportunity to investigate CSF due to the ease of obtaining large sample volumes and the ability for time course collection. Despite this advantage, there have been no large scale proteomic investigations of equine CSF. The aim of this study was to comprehensively characterize the proteome of equine CSF. Results of this work provide the groundwork for future hypothesis-driven research with far reaching applications in the study of equine pathology and infectious disease research.

2. Materials and Methods

All CSF collection procedures were approved by the Institutional Animal Care and Use Committee at Colorado State University. Briefly, CSF was collected from six healthy horses via subarachnoid catheter or manual draw [3]; protease inhibitors were added (Pierce, Rockford, IL, USA), and samples were frozen at −80°C (Fig. 1). Protein concentrations were determined via Bradford assay (Thermo Scientific, Rockford, IL, USA) [4], and 30 μg of each sample underwent in-solution digestion using Protease Max (Promega, Madison, WI, USA) and urea. Samples were solubilized in a solution of 8 M urea and 0.2% protease max and then reduced, alkylated, and digested with 1% protease max and trypsin at 37°C for 3 hours. Samples were dried in a vacuum centrifuge (Speed Vac; Savant), desalted using PepClean C18 spin columns (Pierce), dried, and resuspended in 30 μL of 3% acetonitrile (ACN) and 0.1% formic acid. All solvents, water, and acid were liquid

Conflict of interest: none.

* Corresponding author at: Dr. Jessica E. Prenni, PhD, Proteomics and Metabolomics Facility, Colorado State University, 2021 Campus Delivery, Fort Collins, Colorado 80523.
E-mail address: Jessica.Prenni@colostate.edu (J.E. Prenni).

0737-0806/$ – see front matter © 2014 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jevs.2013.07.013

chromatography tandem mass spectrometry (LC-MS/MS) grade (Sigma, St. Louis, MO). Online two-dimensional LC-MS/MS (2D-LC-MS/MS) with strong cation exchange (SCX) and subsequent reverse-phase chromatography were performed as follows. Ten-microgram aliquots of digested peptides from each biological sample ($n = 6$) were loaded onto a 3.5-μm 50 × 0.8-mm column (Zorbax BIO-SCX II; Agilent Technologies, Santa Clara, CA, USA) for individual analysis. Peptides were eluted from the SCX column in a step-wise procedure using increasing concentrations of NaCl in 0.3% ACN and 0.1% formic acid (20-μL NaCl injections at 15, 30, 45, 60, 75, 90, 120, 150, 300, and 500 mM). Peptides from each individual salt injection were then purified and concentrated using an online enrichment column (Zorbax C18, 5 μm, 5 × 0.3 mm; Agilent). Subsequent chromatographic separation was performed on a reverse-phase nanospray column (model 1100 nano-high performance liquid chromatography, Zorbax C18, 5 μm, 75-μm inside diameter × 150 mm column; Agilent) using a 60-minute linear gradient from 25%-55% buffer B (90% ACN, 0.1% formic acid) at a flow rate of 300 nL/min. Peptides were eluted directly into the mass spectrometer (LTQ linear ion trap; Thermo Scientific), and spectra were collected over an m/z range of 200-2,000 by using a dynamic exclusion limit of two MS/MS spectra of a given peptide mass for 30 seconds (exclusion duration of 90 seconds). Compound lists of resulting spectra were generated using Bioworks version 3.0 software (Thermo Scientific) with an intensity threshold of 5,000 AU and one scan per group. This workflow generated 10 raw data files per sample. Tandem mass spectrometry spectra results were searched against the National Center for Biotechnology Information nonredundant (NCBInr) protein database by using a taxonomy filter for "equine" concatenated to a reverse decoy database (14,473 total entries), using both the Mascot database search engine (version 2.3.02; Matrix Science, Boston, MA) and the Sequest engine (version 27, revision 11; Sorcerer, Sage-N Research, San Jose, CA). The search parameters used were as follows: average mass; peptide mass tolerance of 2.5 Da; fragment ion mass tolerance of 1.0 Da; complete tryptic digestion allowing one missed cleavage; variable modification of methionine oxidation; and a fixed modification of cysteine carbamidomethylation. Peptide identifications from both of the search engines were combined using protein identification algorithms in Scaffold 3 (version 3.6.4; Proteome Software, Portland, OR, USA) [5,6]. All data files from all SCX fractions and six biological replicates were then combined using the "mudpit" option in Scaffold 3, generating a composite listing for all proteins identified across all runs and animals. Peptide and protein probability thresholds of 99% and 95%, respectively, and a minimum of two unique peptides were applied and resulted in a peptide false discovery rate of 0.2% based on peptides matching to the decoy database [7,8]. Proteins containing shared peptides were grouped by Scaffold to satisfy the laws of parsimony. Manual validation of MS/MS spectra was performed for all protein identifications above the probability thresholds that were based on only two unique peptides. Criteria for manual validation included 1) a minimum of at least five theoretical y or b ions in consecutive order greater than 5% of the maximum peak intensity; 2) an absence of prominent unassigned peaks greater than 5% of the maximum peak intensity; and 3) indicative residue-specific fragmentation, such as intense ions N-terminal to proline and immediately C-terminal to aspartate and glutamate. NCBI gi accession numbers were converted to UniProt accession numbers using ID mapping in UniProt. Gene ontology (GO) terms were then collected in UniProt, using the Retrieve gene ontology tool for the 145 proteins whose NCBI gi accession numbers could be converted to UniProt accession numbers.

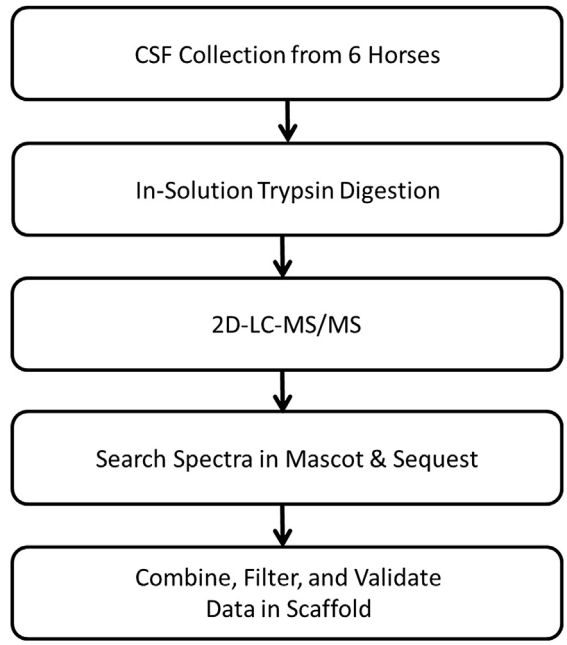

Fig. 1. Experimental workflow. Cerebrospinal fluid was collected from 6 healthy horses, digested, and analyzed via 2D-LC-MS/MS. Spectra were searched against the NCBInr equine protein database using both Mascot and Sequest. Results were combined, filtered, and validated in Scaffold 3. 2D-LC-MS/MS two-dimensional liquid chromatography tandem mass spectrometry; NCBInr, National Center for Biotechnology Information non-redundant.

3. Results and Discussion

The objective of this study was to comprehensively characterize the CSF proteome of the healthy horse. To enable the detection of as many proteins as possible, CSF from each animal was analyzed separately by 2D-LC-MS/MS. The resulting data from each animal were combined to generate an inclusive list of identified proteins that formed a representative healthy equine CSF. A total of 320 proteins (Table 1) were identified that met the filtering criteria; this list is publically available in the proteomics identifications (PRIDE) repository (dataset identifier PXD000133) [9]. The overall abundance of proteins in healthy animals was assessed by calculation of normalized spectral abundance factors (NSAF) [10]. Proteins in Table 1 are listed in order of their NSAF value, where NSAF = (unweighted spectrum count/molecular weight in kilodaltons (kDA))/sum of spectral counts for the overall experiment. NSAF provide a measurement of relative abundance within a sample by accounting for overrepresentation of peptides from large molecules compared to proteins with lower molecular weight. The most abundant proteins based on NSAF values