

RESEARCH

Open Access



# The proteome of the blood–brain barrier in rat and mouse: highly specific identification of proteins on the luminal surface of brain microvessels by in vivo glyco-capture

Tammy-Lynn Tremblay<sup>1</sup>, Wael Alata<sup>1,2</sup>, Jacqueline Slinn<sup>1</sup>, Ewa Baumann<sup>1</sup>, Christie E. Delaney<sup>1</sup>, Maria Moreno<sup>1</sup>, Arsalan S. Haqqani<sup>1</sup>, Danica B. Stanimirovic<sup>1^</sup> and Jennifer J. Hill<sup>1\*</sup>

## Abstract

**Background** The active transport of molecules into the brain from blood is regulated by receptors, transporters, and other cell surface proteins that are present on the luminal surface of endothelial cells at the blood–brain barrier (BBB). However, proteomic profiling of proteins present on the luminal endothelial cell surface of the BBB has proven challenging due to difficulty in labelling these proteins in a way that allows efficient purification of these relatively low abundance cell surface proteins.

**Methods** Here we describe a novel perfusion-based labelling workflow: in vivo glyco-capture. This workflow relies on the oxidation of glycans present on the luminal vessel surface via perfusion of a mild oxidizing agent, followed by subsequent isolation of glycoproteins by covalent linkage of their oxidized glycans to hydrazide beads. Mass spectrometry-based identification of the isolated proteins enables high-confidence identification of endothelial cell surface proteins in rats and mice.

**Results** Using the developed workflow, 347 proteins were identified from the BBB in rat and 224 proteins in mouse, for a total of 395 proteins in both species combined. These proteins included many proteins with transporter activity (73 proteins), cell adhesion proteins (47 proteins), and transmembrane signal receptors (31 proteins). To identify proteins that are enriched in vessels relative to the entire brain, we established a vessel-enrichment score and showed that proteins with a high vessel-enrichment score are involved in vascular development functions, binding to integrins, and cell adhesion. Using publicly-available single-cell RNAseq data, we show that the proteins identified by in vivo glyco-capture were more likely to be detected by scRNAseq in endothelial cells than in any other cell type. Furthermore, nearly 50% of the genes encoding cell-surface proteins that were detected by scRNAseq in endothelial cells were also identified by in vivo glyco-capture.

**Conclusions** The proteins identified by in vivo glyco-capture in this work represent the most complete and specific profiling of proteins on the luminal BBB surface to date. The identified proteins reflect possible targets

<sup>^</sup>Danica Stanimirovic died on January 17, 2024. Dana is remembered as an extraordinary scientist, mentor, colleague, and friend.

\*Correspondence:

Jennifer J. Hill

Jennifer.Hill@nrc-cnrc.gc.ca

Full list of author information is available at the end of the article



for the development of antibodies to improve the crossing of therapeutic proteins into the brain and will contribute to our further understanding of BBB transport mechanisms.

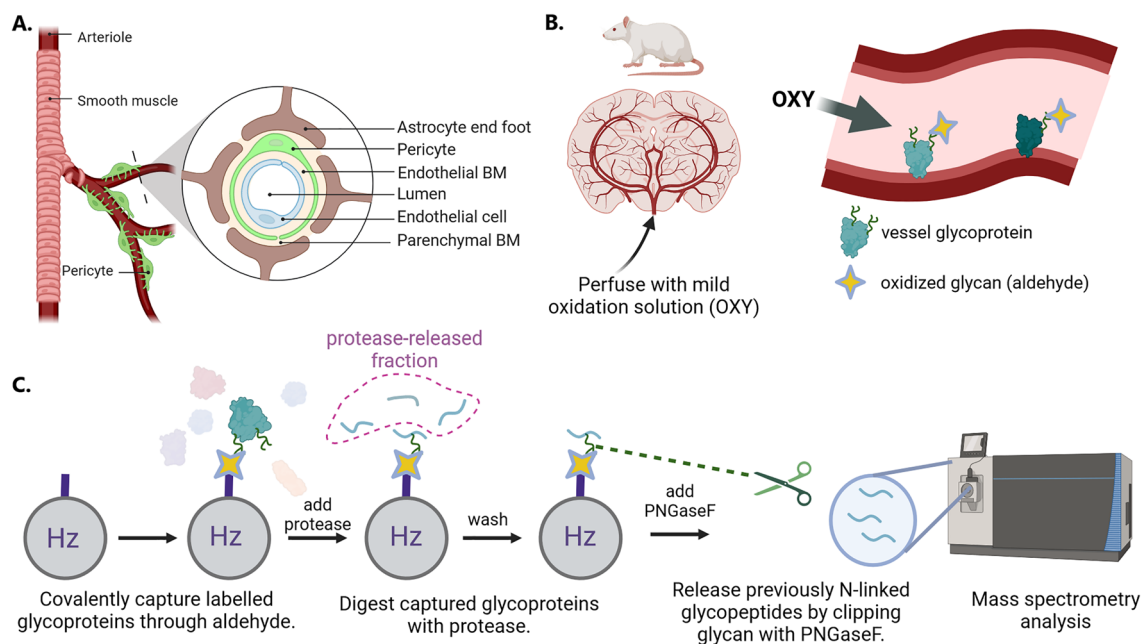
**Keywords** Blood–brain barrier, Proteomics, Luminal, Endothelial, Vessel

## Background

The blood–brain barrier (BBB) is composed of a tightly organized subset of cells, often referred to as the neurovascular unit (NVU) that controls the transport of molecules from the blood into the brain [1–4]. The prototypical NVU consists of endothelial cells connected by tight junctions to form microvessels, where the luminal surface of the polarized endothelial cells is in contact with blood and the basal surface is coated by an acellular basement membrane comprised mainly of extracellular matrix proteins such as collagens and laminins (Fig. 1A). This vascular basement membrane layer is further surrounded by pericyte cells and can merge with a parenchymal basement membrane in areas without pericytes [5]. Outside of the parenchymal basement membrane, astrocyte endfeet surround the vessel and provide a connection to neurons in the brain parenchyma. The NVU is thought to be the cellular basis for the tight regulation of transport of nutrients, metabolites, and immune cell

entry across the BBB and into the brain and central nervous system.

The luminal surface of endothelial cells is in direct contact with blood and includes a wide variety of transporter proteins that shuttle nutrients and metabolites across the BBB in a tightly-regulated manner via transcytosis [6]. The tight regulation of molecules crossing the BBB represents one of the largest challenges in the treatment of neurological diseases, such as Alzheimer’s and Parkinson’s disease, since many therapeutics are not able to enter the brain parenchyma due to their inability to cross the BBB efficiently, including nearly all protein-based drugs such as antibodies. One promising approach to facilitate trafficking of protein therapeutics into the brain is to use a “carrier” protein that binds to a transcytosing receptor present on the surface of the endothelial cells, allowing the therapeutic to piggyback on this receptor to gain access into the brain through receptor-mediated transcytosis (RMT) [7–10]. A thorough profiling of the



**Fig. 1** Overview of the in vivo glyco-capture method for labelling luminal glycoproteins at the blood–brain barrier. **A** Depiction of the typical neurovascular unit at the BBB, showing the lumen of the blood vessel and its surrounding cells and basement membrane (BM). **B** Perfusion-based labelling of glycoproteins in rats and mice using a mild oxidation solution to form aldehydes on luminal glycans, including glycoproteins. **C** Glyco-capture sample preparation workflow to specifically isolate peptides from proteins with oxidized glycans. A hydrazide (Hz) functionalized bead forms a covalent bond with aldehydes, capturing labelled glycoproteins and allowing unbound proteins to be washed away. Bound glycoproteins are then digested with a protease (trypsin, chymotrypsin) and unbound non-glycopeptides are washed away. Formerly N-glycosylated peptides are specifically released from the beads by the enzyme PNGase F and this collection of peptides is analyzed by mass spectrometry. Created with BioRender.com

proteins present at the BBB would improve our understanding of this unique physiological structure and the proteins responsible for trafficking through the BBB. Proteins present at the BBB also represent potential targets for development of therapeutics that can enable brain delivery.

Given the importance of the proteins present on the luminal surface of the BBB, several groups have focused on profiling these proteins using proteomic methods, as we have recently reviewed [11]. Proteomic analyses of isolated vessels have successfully identified many endothelial cell surface proteins, including known BBB transporter proteins [12–15]. However, these methods identify proteins from all cells present in the vessels and cannot determine the localization of these proteins, such as whether these proteins are present on the specialized luminal surface of the endothelial cells. To provide data on the proteins exposed to blood in brain vessels, chemical labelling has been used to specifically modify luminal proteins via perfusion of a reactive chemical label through animal brain capillaries [16–18]. All of these studies relied on perfusion of an N-hydroxysuccinimide (NHS)-ester modified biotin molecule that reacts with the N-terminus or exposed lysines on vessel-accessible proteins. Interestingly, significantly fewer proteins were identified using this approach in brain than from other organs, such as liver or kidney [11, 16, 18]. Furthermore, the specificity for cell surface proteins was generally low (~10–30%) in these approaches resulting in the identification of a very low number of cell surface membrane proteins. These findings suggest that labelling efficiency with biotin is much lower in the brain relative to other organs when relying on conjugation via NHS-esters.

One of the defining features of the BBB vasculature, relative to vasculature in other organs, is the particularly thick glycocalyx layer lining the luminal surface of the vessels [19]. As the name suggests, the glycocalyx is largely composed of glycans, in the form of glycoproteins, glycolipids, and proteoglycans. Given these two observations, we hypothesized that the glycan-based glycocalyx layer may be limiting the accessibility of luminal-facing proteins for reaction with NHS-esters. Therefore, since the majority of cell surface proteins are glycosylated, we decided to utilize a chemical labelling approach that targets glycans in the hopes of enriching luminal vessel proteins at the BBB in a more efficient and selective manner.

To this end, we modified a cell surface labelling approach which has been previously applied to isolated cells and tissues, called Cell Surface Capture (CSC) [20–22]. Unlike the NHS-biotinylation approach, which requires accessible primary amines (protein N-terminus or lysine residues), this cell surface labelling method specifically modifies glycans using mild oxidizing conditions.

This oxidation reaction produces an aldehyde group which can then be used as a chemical “handle” for specific reaction with a hydrazide or aminoxy group. While the original CSC method captured these aldehyde-containing molecules through subsequent reaction with a hydrazide-modified biotin molecule, we have found that the same selectivity for cell surface molecules can be maintained by directly capturing aldehyde-containing glycoproteins on a resin with an immobilized hydrazide group attached, a method referred to as Direct Cell Surface Capture (D-CSC) [23]. In this work, we adapt D-CSC to label luminal vessel proteins *in vivo* through terminal perfusion directly into the brain via the common/internal carotid. We demonstrate that this *in vivo* glyco-capture method is highly specific for cell surface proteins and is capable of identifying the majority of known BBB carrier targets in both rat and mouse animal models. The list of identified BBB luminal proteins presented here represent a valuable resource for future selection of RMT receptors to target for BBB carriers.

## Methods

### Animals

All experiments were performed in accordance with the Canadian Council on Animal Care and were approved by the Animal Care Committee of the National Research Council Canada. Wistar rats (male, 175–200 g) and four-week male Balb/C mice were purchased from Charles River Laboratories Inc. (Montreal, Canada). Upon arrival, animals were housed in the animal facility of the National Research Council Canada for at least 5 days on a 12-h light–dark cycle at  $22 \pm 3$  °C with free access to food and water. Details on animals used for each analysis is provided in Additional file 4: Table S1.

### Rat brain perfusion

Rats were deeply anesthetized under isoflurane gas, placed in supine position, and given a bolus *i.v.* injection of heparin (300 units/0.3 ml) via the right jugular vein. The left common carotid artery (CCA) was exposed and a perfusion line catheter made from polyethylene tubing (PE50) and filled with saline was inserted into the CCA and secured. Prior to perfusion both jugulars were cut, the chest cavity opened and the heart was cut. All animals were terminally-perfused via the CCA. Non-oxidized control rats (CTL) were perfused for 10 min at a rate of 2.5 ml/min with Live Cell Imaging Solution buffer (LCIS pH adjusted to 6.5; Life Science, cat# A14291DJ), containing 140 mM NaCl, 2.5 mM KCl, 1.8 mM CaCl<sub>2</sub>, 1.0 mM MgCl<sub>2</sub> and 20 mM Hepes). Oxidized/glyco-captured rats (OXY), were perfused for 8 min (2.5 ml/min) with freshly prepared Oxidation Solution (10 mM NaIO<sub>4</sub> dissolved in LCIS pH 6.5) followed by 2 min perfusion (2.5 ml/min) of

Quench Solution (30 mM Na<sub>2</sub>SO<sub>3</sub> dissolved in LCIS pH 6.5). At completion of perfusion, brains were dissected from the skull, weighed, and cut in half. The two hemispheres were frozen on dry ice and kept at – 80 °C until processed for vessel isolation.

#### Mouse brain perfusion

In situ brain perfusion was performed as described previously with slight modifications [24, 25]. Briefly, mice were deeply anesthetized by intraperitoneal injection of xylazine/ketamine (8/140 mg/kg). Then, the left common carotid artery was ligated at the heart side, and the external carotid artery at the bifurcation level to ensure that all the perfusate was directed straight to the brain. A polyethylene catheter filled with heparin (25 IU/mL) was then inserted in the left common carotid. The different solutions were perfused at a flow rate of 2.5 ml/min immediately after cutting the heart. The perfusion was performed in two subsequent steps: [1] eight minutes of Oxidation Solution, [2] two minutes of Quench Solution. In the control group, the first step was performed using only LCIS buffer with a pH 6.5, and the second step was performed as described above. After the perfusion, brains were extracted and the perfused hemispheres were kept at – 80 °C until processed for vessel isolation.

#### Vessel isolation and lysis

Vessels were isolated from either fresh- or snap-frozen brain tissues. Cerebellum was removed. In some cases, the perfused brain hemispheres from 2 animals were combined for vessel isolation (see Additional file 4: Table S1). Minced brains were resuspended in ice-cold PBS+20 mM sodium sulfite with 25 µL of protease inhibitor cocktail (Sigma, Cat# P8340) prior to homogenization with a Dounce homogenizer. The homogenate was filtered through a series of pluriStrainers® (pluriSelect, San Diego, CA) stacked in descending order from 300, 100, and 20 µm pore size. The pluriStrainers were fitted into a connector ring within a 50 ml conical tube and the homogenate was transferred onto the filter stack and filtered through the strainers using gentle suction. Ice-cold PBS (5 mL; Wisent, Saint-Jean-Baptiste, QC) was used to rinse the stacked strainers, then strainers were placed upside down in a new 50-mL conical tube and rinsed with 5 ml of ice-cold PBS per strainer to collect the vessels and capillaries. To release the vessels from the strainer, the buffer was forced through the filter by pipetting up and down with a 5 mL pipette. Microvessels and capillaries collected onto 100 µm and 20 µm strainers were collected in the same 50 mL conical tube and were centrifuged at 900×g for 5 min at 4 °C to pellet the vessels. The supernatant was discarded and proteins were extracted immediately from the vascular pellet by adding

1 mL (rat) or 0.5 mL (mouse) of Hydrazide Coupling Buffer ('Hz CB': 100 mM NaOAc, pH 5.5, 150 mM NaCl, 0.5% SDS, 1:200 v/v protease inhibitors and 1:1000 Benzodase) supplemented with 10 mM Na<sub>2</sub>SO<sub>3</sub>, and rotating the extract for 45 min at RT. Protein extracts were then frozen and stored at – 80 °C until glyco-capture. Protein concentration was determined using a DC protein assay (Bio-Rad).

#### BBB glyco-capture and digest

The protocol for hydrazide capture was based on previously published methods [26–28] with some modifications. Mouse (0.5–0.7 mg protein) or rat (2 mg protein) vessel extracts from perfused animals were diluted to between 0.5 and 1.1 mg protein/mL with Hz CB. Samples were incubated O/N at RT with 50 µL (mouse) or 100 µL (rat) of prewashed Affi-gel hydrazide beads (Bio-Rad). Beads were washed eight times with urea buffer (8 M urea, 0.4 M NH<sub>4</sub>HCO<sub>3</sub>) and four times with 50 mM NH<sub>4</sub>HCO<sub>3</sub>, followed by reduction (10 mM DTT for 1 h at 56°C) and alkylation (25 mM iodoacetamide for 1 h at RT in the dark). Beads were then washed four times with either 50 mM NH<sub>4</sub>HCO<sub>3</sub>/15% acetonitrile before incubation O/N at 37°C with 10µg of trypsin in 450 µL of the same buffer or, in rat samples only, four times with 50 mM NH<sub>4</sub>HCO<sub>3</sub> before incubation O/N at 25 °C with 10 µg of chymotrypsin in 450 µL of 50 mM NH<sub>4</sub>HCO<sub>3</sub>. The protein yield from the mouse vessel extract was too limited to allow use of both proteases. Trypsin-released or chymotrypsin-released peptides were collected and beads were washed three times with 1.5 M NaCl, three times with 60% acetonitrile-0.1% TFA, three times with 100% MeOH, and six times with 50 mM NH<sub>4</sub>HCO<sub>3</sub> before an incubation O/N at 37 °C with 4 units (mouse) or 6 units (rat) of PNGase F (Sigma) in 100 µL of 50mM NH<sub>4</sub>HCO<sub>3</sub>. N-linked peptides were collected by combining the supernatant with further elutions with 200 µL of 50 mM NH<sub>4</sub>HCO<sub>3</sub> and 200 µL of 50% acetonitrile/5% acetic acid. All eluates were pooled, dried under vacuum, and resuspended in 35 µL (mouse) or 50 µL (rat) of deionized water containing the glyco-captured peptide digests.

#### SV-ARBE cell surface and brain lysate glyco-capture

For brain lysate glyco-capture, whole brain tissue was thawed, cut into smaller pieces, and homogenized in Hz CB without SDS. After homogenization, SDS was added to a final concentration of 0.5% and samples were agitated for 1 h at RT and centrifuged at 16000 g for 10 min at RT to remove insoluble debris, and the protein concentration of the supernatant was determined using a DC protein assay (Bio-Rad). Glyco-capture was performed as described for the rat vessel extracts except that proteins

were oxidized in the whole brain lysate by using 15 mM Na-m-periodate for 1 h at RT in the dark, following the initial dilution into Hz CB. This reaction was then quenched by adding 30 mM Na<sub>2</sub>SO<sub>4</sub> for 10 min at RT before proceeding with overnight RT incubation with 50 μL of prewashed hydrazide beads and following the rest of the steps outlined above. For SV-ARBEC surface glyco-capture, cells were first oxidized on the plate prior to glyco-capture following the previously described protocol [23].

### Mass spectrometry

Peptide digests (10–30% of each sample) were analyzed by automated nanoLC-MS(/MS) on an Eclipse Tri-brid mass spectrometer coupled to a Dionex Ultimate 3000 UPLC system (Thermo Scientific). Peptides were trapped using an inline MicroTrap C18 (Phenomenex, 10×0.3 mm) and separated on a 10 cm×100 μm I.D. C18 column (Waters, 1.7 μm BEH130C18, 186003546) at 500 nL/minute using a 60 min gradient (solution A: 0.1% formic acid, solution B: 100% ACN/0.1% formic acid), followed by a 9 min equilibration at 1% solution B. MS spectra were acquired in the Orbitrap between 375 and 1800 Da m/z in profile mode at 120k resolution, while data-dependent rapid CID MS/MS scans were acquired in the ion trap in centroid mode with an intensity threshold of 1.5e4, a 1.6 Da isolation window, and normalized collision energy of 35%, and a 1 s cycle time. Dynamic exclusion was enabled (60s within 10 ppm), AGC was set to “custom”, and max inject time was set to “dynamic”.

### MS data analysis

Thermo raw files were converted to mzML using MSConvertGUI Version: 3.0.18278-ad334f4d5 (parameters: peakPicking cwt snr=0.1 msLevel=1-). MS2 data were searched against the mouse (11Aug2022-21986 entries) or rat (12Aug2022-22860 entries) Uniprot database with decoys added using the Pyteomics Python package in mode ‘shuffle’. The database search and scoring were done using MSFragger v3.5 and Philosopher v 4.4.0 as implemented in Fragpipe v18.0 (Parameters: 20 ppm precursor tolerance, 0.2 Da fragment tolerance with mass calibration and optimal parameter search, enzyme=trypsin (2 missed cleavage) or chymotrypsin (3 missed cleavage), fixed modification: C (carbamidomethyl), variable modifications: M (oxidation), N (deamidation), write decoys selected, file output=tsv file) [29, 30]. The tsv output files from MSFragger, which included decoy database hits, were subsequently filtered using a Python script to only keep peptide-spectrum-match (psm) rows that contained at least one consensus N-glycosylation sequence (N-X(!P)-S/T/C) and a number of deamidation on N modifications that was between 1 and x inclusive, where

x is the number of consensus N-glycosylation sites on the peptide. Next, the proteins represented by the remaining psm's were filtered to keep proteins with a Protein Prophet score  $\geq 0.75$  or  $> 0.8$  to keep the FDR less than 1%, based on remaining decoy hits. Supplementary data contains detailed protein identification lists for each individual experiment (Additional File 1: Data S1) and the peptide-spectrum match details (Additional File 2: Data S2).

### Data analysis

To build a list of mouse/rat cell surface proteins, we used the human cell surface list published in Sobotzki et al. [31] and converted the Gene Symbols to mouse using SynGo [32]. For proteins identified by in vivo glyco-capture that were not on this published list, proteins were added to the cell surface list manually only if the Uniprot subcellular location for the mouse or equivalent human gene was listed as “cell membrane” and other Uniprot information was consistent with extracellular plasma membrane expression. Functional classification and Gene Ontology (GO) enrichment analysis were implemented in Panther v 17.0 [33, 34] using Panther Go-Slim categories and Reactome pathway with the Binomial Test type and FDR correction for the statistical overrepresentation test. Single-cell RNAseq data was downloaded from the Allen Mouse Brain Atlas. <http://portal.brain-map.org/atlasses-and-data/rnaseq/mouse-whole-cortex-and-hippo-campus-10x>; file=“Gene Expression by Cluster, median (.csv)” [35]. To simplify the analysis, the number of cell clusters were minimized by combining all consecutive cell clusters with the same description name after the cell number keeping only the highest expression value for each gene. This simplification reduced the cell type clusters from 388 to 75. Outlier analysis of cell clusters were identified using ROUT (Q=0.1%) implemented in Prism 9.5.1.

## Results

### Identification of luminal BBB proteins in rats by in vivo glyco-capture

The workflow developed for the in vivo glyco-capture of proteins present on the blood-accessible surface of the BBB is described in Fig. 1. In brief, rodents were anesthetized and the brain perfused with a mild oxidizing solution to remove blood and to convert the cis-diols present in glycans to aldehydes, a reactive group that is rarely present in naturally occurring proteins (Fig. 1B). Animals in the control group were perfused only with buffer that did not contain any oxidation reagent. Brains were then extracted and homogenized on ice prior to enrichment of microvessels and capillaries via a filtration-based method [36–38]. The enriched microvessel fraction was lysed

and solubilized in an SDS detergent-containing buffer to ensure that all membrane proteins were fully solubilized. To isolate the proteins located on the luminal side of the brain microvasculature, the aldehyde-containing oxidized glycoproteins were covalently captured onto beads modified with a hydrazide group and extensively washed (Fig. 1C). Next, the immobilized glycoproteins were digested with a protease, either trypsin or chymotrypsin, releasing non-glycosylated peptides from the immobilized glycoproteins (“protease-released fraction”) and leaving only the glycosylated peptides attached to the beads. Lastly, PNGaseF, an enzyme that cleaves the bond between an N-linked glycan and asparagine, is used to release N-glycosylated peptides (“N-glycosylated fraction”).

Altogether, we collected samples from 10 individual rats using the above workflow, half of which were perfused with the mild oxidation solution (‘OXY’) and half perfused with buffer (‘CTL’), in order to confirm the specificity of the labelling approach. In theory, glycoproteins should only be labelled and captured after oxidation, so any proteins that are identified with the CTL perfusion are likely proteins with naturally-occurring aldehyde modifications or proteins that bind to the capture beads due to non-specific interactions. As can be seen in Additional file 3: Figure S1A (MSight), peptide signals from the N-glycosylated fraction are visible in both the CTL and OXY samples when the mass spectrometry data is visualized in a 2D representation using MSight [39]. As would be expected, there are significantly more peptide spots present in the OXY samples; these represent glycopeptides that have been specifically enriched through our *in vivo* glycocapture method. In contrast, the protease-released fraction showed similar peptide patterns between CTL and OXY (Additional file 3: Figure S1B) and thus was not analyzed further. In all cases, the peptide spot pattern is quite similar between individual animals receiving the same treatment showing reproducibility in the approach.

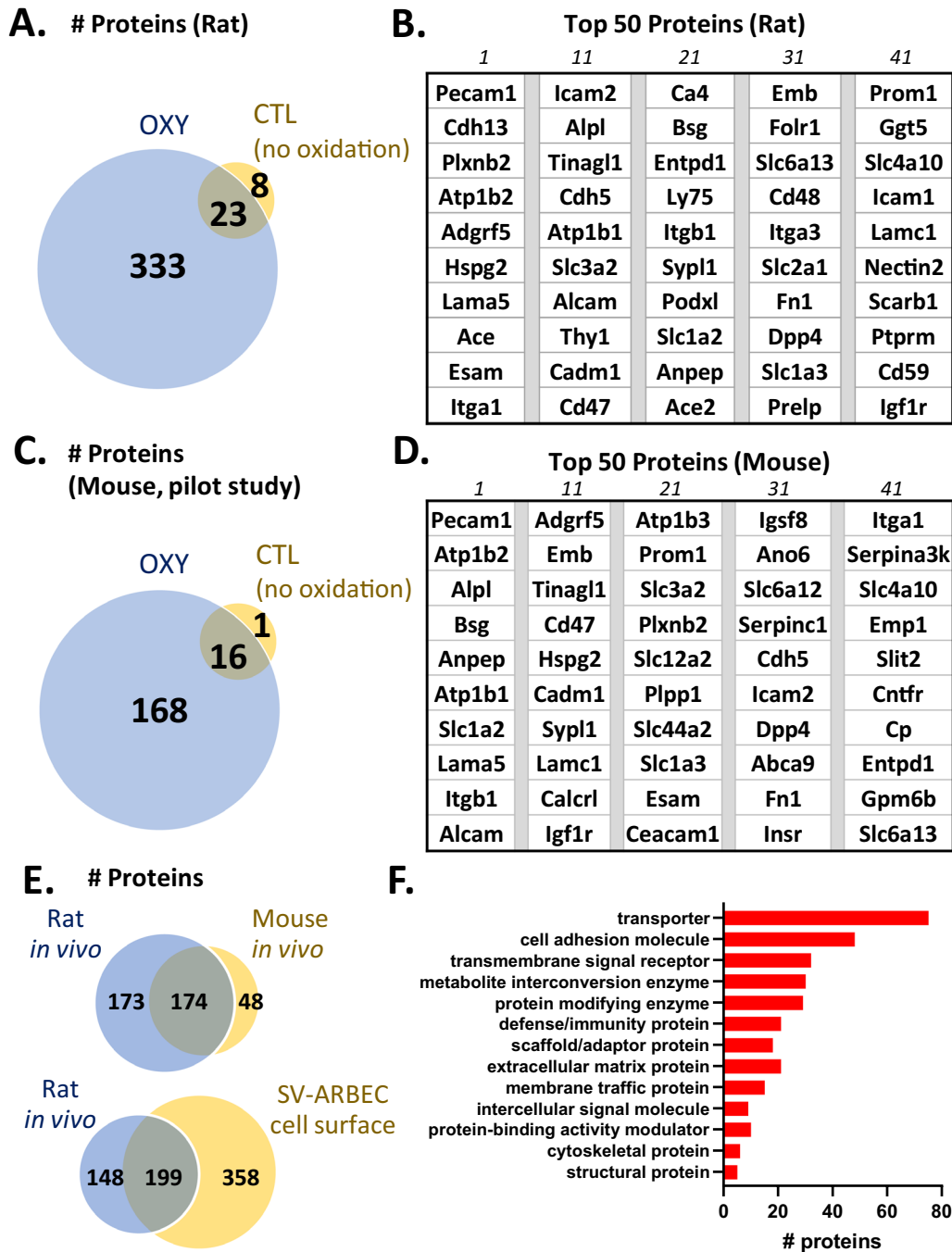
Using typical proteomic database search methods, the peptides present in the LC–MS data of the N-glycosylated fraction were identified and assigned to proteins to identify proteins present at the BBB. Following rigorous filtering of the rat data to maximize identifications while minimizing false positive identifications below 1% (see Methods), 356 proteins were identified in the OXY samples while only 31 proteins were identified in the CTL samples. Figure 2A shows that 23 of the 31 proteins identified in the CTL sample were also found in the OXY samples; of these 23 proteins, 14 were found at higher levels in the OXY fraction relative to the CTL fraction, with 1.5–64-fold more spectral counts assigned to them in the OXY fraction. Based on this observation, we

suspect that some of the CTL hits may represent carry-over in the nano-LC columns. In contrast, the other 9 proteins identified in both the CTL and OXY fractions had similar spectral counts in each sample, and likely represent background binding. These 9 proteins include common abundant proteins like histones and tubulins, together with a few predicted surface receptors (Lrp1, Atp1a3, and Slc6a11) which possibly contain naturally occurring aldehyde modifications, such as carbonylation. These 9 proteins were eliminated from our final list of rat proteins identified at the luminal surface of the rat BBB, resulting in the identification of 347 rat proteins (Additional file 4: Table S2). The 50 proteins with the highest number of spectral counts are highlighted in Fig. 2B.

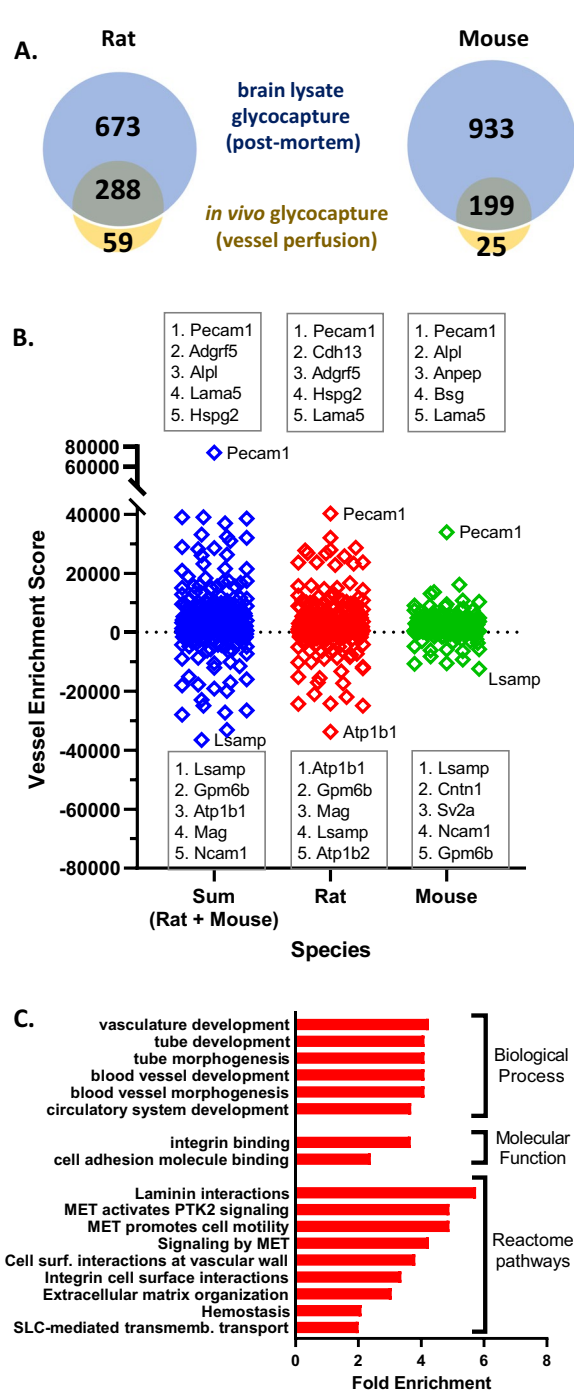
#### **Identification of luminal BBB proteins in mice by *in vivo* glycocapture**

A more complete understanding of the differences in protein composition at the BBB between species would be useful for the selection of animal models for various types of BBB studies. For this reason, we decided to extend our study into mouse, another common BBB model species. In some cases, two perfused hemispheres from different animals were combined prior to vessel enrichment (see Additional file 4: Table S1). Due to the smaller size of mouse brains, which resulted in a lower amount of total protein present in the enriched vessel lysate prior to glycocapture, we used only a single protease (trypsin) to avoid splitting this small sample into two separate digests. First, a pilot study using 8 mice was performed to confirm that the specificity of glycocapture was similar to the rat experiment. In this experiment vessels were pooled from two animals for each condition (CTL or OXY) prior to glycocapture, resulting in two CTL samples and two OXY samples in the LC–MS analysis. In this pilot study, 184 proteins were identified in the OXY samples while 17 proteins were identified in the CTL samples (Fig. 2C). Of the proteins identified in the CTL, 16 were also found in the OXY fraction and 12 of these proteins had higher spectral counts in the OXY sample (1.5–18-fold more than CTL samples), suggesting enrichment by oxidation labelling. To increase the number of proteins identified at the mouse BBB, we perfused an additional five mice with oxidation solution and compiled the proteins identified from all OXY perfused animals, removing the 4 proteins that were identified at equal levels in the CTL and OXY samples during our pilot study. This analysis led to the identification of 224 proteins present at the luminal BBB surface in mice (Additional file 4: Table S3). The 50 proteins with the highest number of spectral counts in mice are highlighted in Fig. 2D.

There is strong overlap between the proteins identified at the BBB in rat and in mouse (Fig. 2E, top panel).



**Fig. 2** In vivo glyco-capture enriches for BBB proteins with high specificity in both rat and mouse. **A** Number of proteins identified in rats perfused with oxidation solution (OXY) or a buffer control (CTL) after the glyco-capture workflow. **B** Top 50 proteins ranked by spectral counts that were identified in OXY rat samples. All proteins are listed by their corresponding HGNC gene symbol. **C** Number of proteins identified in a pilot mouse study perfused with oxidation solution (OXY) or a buffer control (CTL). **D** Top 50 proteins ranked by spectral counts identified in all OXY perfused mice. **E** Comparison of proteins identified by in vivo glyco-capture in rat and in mouse (top) and comparison of proteins identified by in vivo glyco-capture in rat with proteins identified by cell surface capture of the rat endothelial cell line, SV-ARBE (bottom). **F** Classification of all proteins identified by in vivo glyco-capture in either rat or mouse by Panther protein class



A combined list of proteins identified in either species contains 395 proteins, with 174 proteins found in both species (Additional file 4: Table S4). A large number of proteins (173) were identified only in rat and not in mouse, likely due to technical reasons, including the lower amount in input protein in mouse due to the smaller brain size and our subsequent decision to use

**Fig. 3** Comparison of in vivo glyco-capture to a whole brain lysate glyco-capture. **A** Number of proteins identified by in vivo glyco-capture and by brain lysate glyco-capture in rat and mouse. **B** Distribution of the vessel enrichment scores for all proteins identified by in vivo glyco-capture with the highest scoring and lowest scoring proteins indicated. **C** Gene ontology and Reactome pathway enrichment analysis for the 194 in vivo glyco-capture proteins with a vessel enrichment score > 1000. All listed categories had at least a twofold enrichment value with an FDR p-value < 0.05 when using a reference of the 1367 proteins identified in all glyco-capture experiments, including both in vivo glyco-capture and brain lysate glyco-capture

only one protease (trypsin) instead of two (trypsin and chymotrypsin). A high proportion of the proteins identified by in vivo glyco-capture in rat and mouse vessels are cell surface proteins, with 90% of the proteins predicted as cell surface and an additional 6% of protein predicted as secreted or extracellular matrix proteins that are often associated with the cell surface. These proteins contain a large number of proteins classified as transporters, cell adhesion molecules, and transmembrane signal receptors, as would be expected at the BBB (Fig. 2F) and contains several acknowledged endothelial cell markers, including PECAM1, VWF, CD34, Cldn5, ZO1 (Tjp1), and the majority of known BBB carriers, including insulin receptor (Insr), transferrin receptor (Tfrc), and Insulin-growth factor 1 receptor (Igf1r). The low number of proteins classified as cytoskeletal proteins highlights the specificity of our labelling and analysis method.

To obtain an initial indication of the similarity between our in vivo glyco-capture results and proteins expressed on endothelial cells, we compared the proteins identified by rat in vivo glyco-capture with those found on the cell surface of an immortalized adult rat brain microvascular endothelial cell line, SV-ARBE (Fig. 2E, bottom panel) [40]. In total, 557 proteins were identified on the surface of SV-ARBE cells by surface glyco-capture using similar labelling and analysis conditions as those used in the in vivo glyco-capture, as listed in Additional file 4: Table S5. Of these, 199 proteins were also found in the rat in vivo glyco-capture; therefore over 57% of the proteins identified by in vivo glyco-capture were also found in endothelial SV-ARBE cells.

#### Comparison of in vivo glyco-capture to brain lysate glyco-capture

To provide an indication of the proteins identified when glyco-capture labelling is not restricted to the perfused vessels, we also performed glyco-capture on a complete brain lysate by oxidizing all glycoproteins present within a solubilized brain lysate in both rat and mouse. More proteins were identified in the brain lysate glyco-capture



than by in vivo glyco-capture (Fig. 3A), likely due to the less restricted labeling procedure, and the higher amount of starting material for mouse (2 mg vs. 0.5–0.7 mg for the in vivo glyco-capture; 2 mg input was used in both cases for rat). In rat, 961 proteins were identified in the brain lysate glyco-capture, compared to the 347 that were identified in the in vivo glyco-capture. In mouse, this difference was even larger since the whole brain input was much less restricted by brain size than the in vivo glyco-capture, resulting in 1132 and 224 proteins identified in whole brain and in vivo glyco-capture respectively. The overlap in identifications between brain lysate glyco-capture and in vivo glyco-capture is high as highlighted in Fig. 3A. All identifications from the rat and mouse brain lysate glyco-capture are listed in Additional file 4: Table S6.

To identify proteins that are enriched in vessels, relative to the whole brain, we calculated a “Vessel Enrichment score” (VE score) for each identified protein. First, spectral counts associated with identified proteins in each dataset were converted to counts-per-million (CPM) to normalize between datasets. The VE score was then calculated for each protein by subtracting the CPM for that protein in the brain lysate glyco-capture from the CPM in the in vivo glyco-capture dataset. Thus, a positive VE score reflects a higher relative abundance for that protein in the vessel dataset relative to the whole brain dataset. The majority of the proteins in the in vivo glyco-capture dataset show a positive VE score (71% rat, 83% mouse, 76% combined species). To calculate VE scores for combined species data, the VE scores from each species were normalized to the median VE score for that species and then summed. The distribution of VE scores in the mouse and rat data, as well as the combined data set are shown in Fig. 3B with the 5 proteins with the highest and lowest VE scores in each dataset highlighted. Many of the proteins with the highest VE scores are well-known vessel marker proteins [41], such as *Pecam1* (CD31), *Anpep* (CD13), *Bsg* (CD147), and *Alpl* (Tissue nonspecific alkaline phosphatase; TNAP), with *Pecam1* having the highest score in all samples. In contrast, proteins with the lowest VE scores contain many proteins known to be widely expressed in neuronal cells, such as *Lsamp*, *Ncam1*, *Mag*, and *Sv2a*.

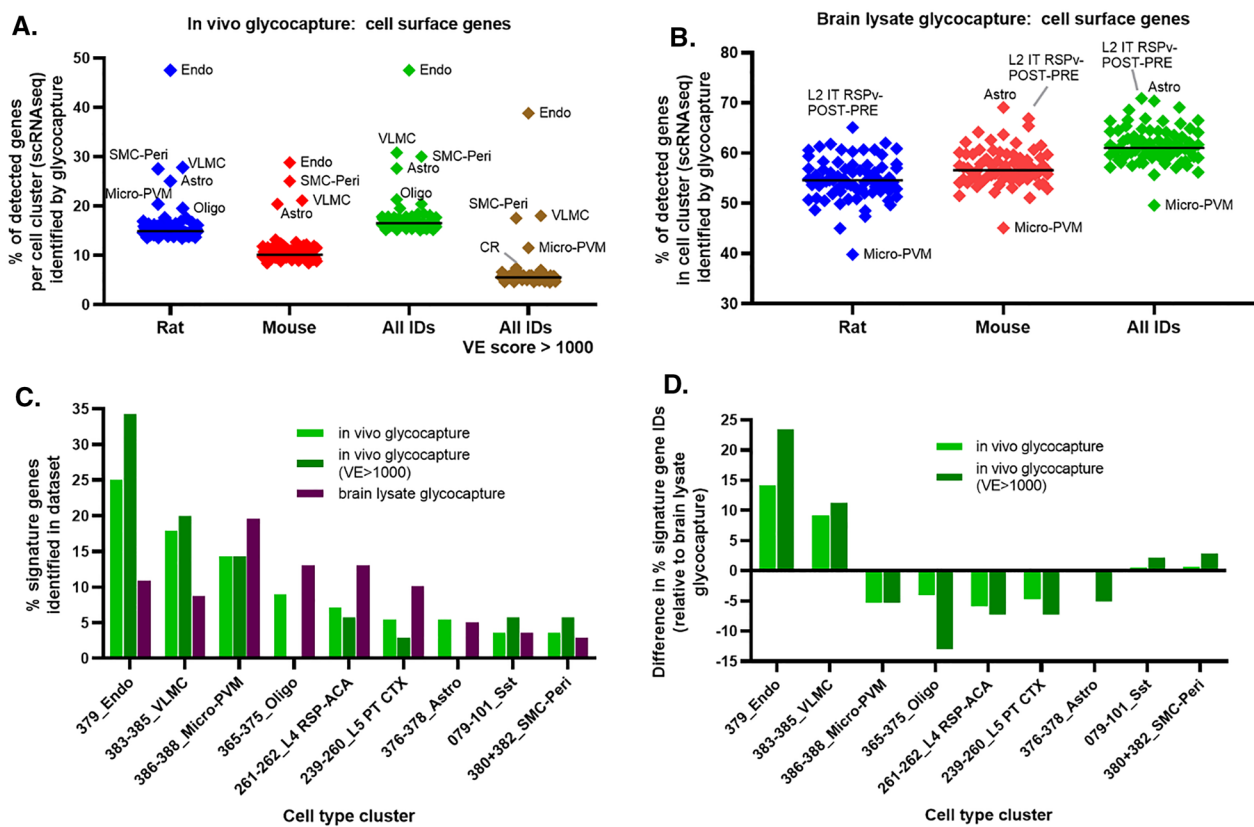
Next, we characterized the proteins identified with a high VE score (>1000) using a statistical overrepresentation test implemented by Panther [33]. As can be seen in Fig. 3C, several Gene Ontology (GO) terms were enriched with FDR-corrected p-values < 0.05, including several GO terms associated with vasculature and blood vessel development, as well as integrin and cell adhesion molecule binding. A number of Reactome pathways were also enriched in this set of proteins, including several

MET signaling pathways, cell surface interactions at the vascular wall, and SLC-mediated transmembrane transport. Taken together this enrichment analysis serves to further demonstrate the specificity of the proteins identified by in vivo glyco-capture.

#### Comparison of proteomic data to single-cell RNAseq data for cell-type analysis

To explore the cell types that were most likely to be labelled using our perfusion-based labelling method, we compared the proteins identified by in vivo glyco-capture to the clustered single-cell RNAseq (scRNAseq) data from mouse brain cells compiled by the Allen Brain Map project [35]. For each of the major brain cell type clusters, we calculated the percentage of cell surface genes detected by scRNAseq that were also detected by in vivo glyco-capture (Fig. 4A). The majority of the 75 major cell type clusters had a very similar percentage of detected genes identified within the in vivo glyco-capture dataset, as shown by the tight clustering of most cell types around the median value line. Only 4 cell-type clusters were found to be outliers in all in vivo glyco-capture datasets using ROUT (Q=0.1%) with all four of these cell-type clusters showing a higher percentage of detected genes in the in vivo glyco-capture datasets than would be expected based on the data from the other cell types. Endothelial cells (Endo) contained the highest percentage of RNAseq-detected genes found at the luminal BBB with up to 48% of the detected cell surface proteins found in the in vivo glyco-capture dataset. The other three cell types that were outliers in all data sets are also known to be associated with vessels. These include vascular leptomeningeal cells (VLMC), smooth muscle cells-pericytes (SMC-Peri), and micro-perivascular macrophages (Micro-PVM). A solid level of enrichment was seen for astrocytes (Astro) in both rat and mice, but interestingly astrocytes do not appear to show a high percentage of enriched genes when looking only at BBB proteins with a VE score > 1000, likely because these cells have a high proportion of glycoproteins even in the brain lysate glyco-capture since astrocytes are abundant in the brain (Fig. 4B). A very low level of enrichment in detection was also seen for oligodendrocytes (Oligo) and Cajal-Retzius cells (CR) in some datasets (see Fig. 4A). In contrast, no cells were found to be statistical outliers when the same analysis was completed with the brain lysate glyco-capture dataset (Fig. 4B) though a few cell types showed a trend toward increased or decreased identifications.

Next, we analyzed “signature genes” that were detected by single-cell RNAseq in ONLY one cell type cluster. Thus, these signature genes are more likely to be specific markers of that particular cell type. This analysis complements the cell-type analysis discussed above since it



**Fig. 4** Cell-type analysis of proteins identified by in vivo glyco-capture through comparison to single-cell RNAseq data. **A** For each in vivo glyco-capture dataset, the identified proteins were aligned to the genes that were detected by single-cell RNAseq in individual cell clusters using data from the Allen Brain Atlas. The percent of genes detected by scRNAseq (with a read count > 0) in each cell count that were identified in the glyco-capture dataset are shown. Statistical analysis comparing all cell types was performed to identify cell types that are outliers. These outlier cell types are labelled and include endothelial cells (Endo), smooth muscle cells-pericytes (SMC-Peri), vascular leptomeningeal cells (VLMC), micro-perivascular macrophages (Micro-PVM), and astrocytes (Astro), oligodendrocytes (Oligo), and Cajal-Retzius cells (CR). **B** The same analysis described in panel A was repeated using the brain lysate glyco-capture datasets. None of the cell types were tagged as outliers. A few cell types showing trends towards over or under representation are labelled. **C** The percentage of identified signature genes associated with each cell type in the listed proteomic dataset: in vivo glyco-capture (light green bar), in vivo glyco-capture with a vessel enrichment score > 1000 (dark green bar), or brain lysate glyco-capture (brown bar). Combined species data was used for the analysis and only cell types that had at least 5% identified signature genes in at least one dataset were included in the graph. **D** The difference in percent identified signature genes between the in vivo glyco-capture dataset and the brain lysate glyco-capture dataset for all identified in vivo glyco-capture proteins and identified in vivo glyco-capture proteins with a vessel enrichment score > 1000, for the cell types listed in panel C

does not rely on normalization to the depth of scRNAseq coverage for each cell type, which was a concern because endothelial cells and several of the others cell clusters described above show a relatively poor depth of coverage in the scRNAseq dataset, possibly due to the relative rarity of these cells in the brain. Specifically, we asked what percentage of signature genes identified in each glyco-capture dataset was associated with a particular cell type cluster. Again, the highest proportion of signature genes representing endothelial cells was found in the in vivo glyco-capture dataset (25%), with an even higher proportion of endothelial signature genes identified in the subset of proteins that have a VE score > 1000 (34%; Fig. 4C).

A significant portion of signature genes associated with VLMC (18%), Micro-PVM (14%), and Oligo (9%) cell clusters were also identified. Interestingly, while Endo, VLMC, Sst (GABAergic neurons), and SMC-Peri cell signatures were more likely to be represented in the proteins with a VE score > 1000, cell signature genes associated with Oligo and Astro clusters were completely absent in the high VE score subset. This likely reflects the high proportion of Oligo and Astro proteins represented in the brain lysate glyco-capture dataset, which was used in the VE score calculation. As a comparison, we also calculated the same values for the brain lysate glyco-capture dataset, shown in the brown column. Intriguingly, only Endo and

VLMC cell clusters showed more than 5% of associated signature genes in the *in vivo* dataset than in the whole brain dataset, as demonstrated by their positive values after correction based on the brain lysate glyco-capture identifications (Fig. 4D).

## Discussion

Here we describe a novel labelling and purification workflow to identify vessel-associated proteins at the luminal BBB surface, relying on terminal perfusion of a mild oxidation reagent to oxidize the glycans of glycoproteins at the luminal surface of vessels. This approach is based on our previously published D-CSC modification of the CSC cell surface protein method [22, 23]. Our modifications removed the necessity of a separate biotinylation reaction. The simplicity of the labeling in D-CSC, which requires only a single labelling step, lends itself well to labelling luminal vessel proteins *in vivo* through a perfusion approach. Previous work aimed at labelling proteins at the BBB through perfusion-based methods have been recently reviewed [11] and have generally relied on the direct biotinylation of proteins through reaction of an NHS-ester with lysine residues or the N-terminus. The results with this NHS-ester based labelling method are reported to be much poorer in brain than in other organs [16, 18], perhaps reflecting challenges in the accessibility of protein molecules due to the thicker glycocalyx. The *in vivo* glyco-capture approach described here demonstrated that direct labelling of glycans via oxidation allowed the identification of a large number of proteins while establishing a high level of specificity that has not been demonstrated with other workflows. Aldehyde-stability is a minor concern in this approach; reaction between the aldehyde on the oxidized glycan with other components in the lysate may result in loss of some labels. However, using *in vivo* glyco-capture, we identified more than twofold the number of proteins at the BBB than have previously been identified by other perfusion-based labelling methods based on NHS-ester proteins, while also identifying a much higher percentage of cell surface proteins [11, 16–18]. The perfusion methods used here may also have played a role in our results. We utilized intracarotid perfusion for rats and mice, both of which likely result in improved perfusion of the brain relative to the intracardiac perfusion that was used in the biotinylation studies. However, preliminary studies on the *in vivo* glyco-capture using intracardiac whole body perfusion showed similar results for brain as the ones found in this study (data not shown), suggesting that the labelling method may play a larger role in these improved results than the perfusion method does.

One advantage of the *in vivo* glyco-capture approach, as well as the Cell Surface Capture method that it is based

on, is that the specificity can be improved by filtering the peptide identifications to keep only peptides that contain a deamidated N-linked sequon as would be expected in a PNGaseF-released peptide. This extra filtering step eliminates many of the peptides that are visible in the CTL sample as well as many of the decoy hits, greatly improving the number of proteins identified while still effectively controlling the false-discovery rate (FDR). FDR control is especially challenging when spectral signal intensity is low, as it is in these experiments due to the small quantity of peptides isolated through this procedure. When the N-linked sequon filter was not used, the number of proteins successfully identified at a 1% FDR was considerably lower, mostly due to an increase in the decoy hits due to the large number of poor-quality MS2 spectra in this very low-level sample.

Many proven receptor-mediated transcytosis targets for BBB carriers were identified in this study, including insulin receptor (Insr) [42], Insulin-like growth factor receptor (Igf1r) [25], transferrin receptor (Tfrc) [43, 44], Glucose transporter type 1 (Slc2a1) [45], Low-density lipoprotein receptor-related protein 1 (Lrp1) [46], Cell cycle control protein 50A (Tmem30a) [7], 4F2 cell-surface antigen heavy chain (Slc3a2) [45, 47], Neonatal Fc receptor (Fcgrt) [48], and basigin (Bsg) [45, 49]. In contrast, a few known vessel proteins are notably absent from this list, including several of the efflux transporters, such as p-glycoprotein/MDR1 (Abcb1 in rat, Abcb1a/b in mouse) and Multidrug resistance-associated protein 2/MRP2 (Abcc2). There are several technical reasons why proteins might not be identified by *in vivo* glyco-capture. First, there is a limit to our ability to detect proteins expressed at low-levels on the cell surface, which highly biases our results to proteins that are present at reasonable abundance. This is likely the reason why many efflux transporters were not identified—of the 4 ABC transporters identified in this study (Abcc1, Abca9, Abcg2, and Abca8a), all were found with low spectral counts in only one of the two species—suggesting that they were near the limit of detection. In many ways, this bias towards identification of abundant proteins could be considered a positive for the development of RMT carriers for BBB crossing where an abundant receptor may be preferred. To this end, it is notable that most of the well-known RMT carriers are in the top 100 proteins as ranked by vessel enrichment score, which is highly weighted by abundance at the BBB. The second major reason why proteins may not be identified by *in vivo* glyco-capture is the absence of an N-linked glycosylation site that is located in a peptide of suitable size to be identified by the MS methods being used. *In vivo* glyco-capture will not identify proteins that are not glycosylated, or that only have o-linked glycosylation. In early

studies of in vivo glyco-capture, we attempted to improve our coverage of proteins by analyzing the peptides that are released from the hydrazide beads by trypsin or chymotrypsin in the protease-released fraction (see Fig. 1). A much larger number of proteins were identified in the protease-released fraction and these identifications were somewhat enriched in cell surface proteins. However, we also found a considerable number of proteins in this fraction from the CTL animals that were perfused only with buffer (see Additional file 3: Figure S1B), suggesting that the specificity for BBB proteins was not sufficient to lead to confident identification of BBB proteins. Based on this result, we decided not to analyze this fraction further. Because of these technical limitations, a negative result (non-identification) in this in vivo glyco-capture study is not necessarily indicative of the absence of a protein at the BBB.

Another major technical limitation of any perfusion-based labelling method is the difficulty in preventing transfer of the label through the vessel and into the surrounding tissue or cells. We specifically chose to use a perfusion rate that mimicked the pressure of normal blood flow to minimize any disruption of the BBB integrity and to keep microvessels completely intact. However, there is always the possibility that some labelling agent was able to leak outside the vessel at low levels. It is also possible that some of the oxidizing agent diffused into the brain through an active or passive process. We identified a number of proteins associated with the basement membrane (several laminins, nidogen 2, agrin), suggesting the possibility that the oxidation reagent may have penetrated somewhat deeper than just the luminal endothelial cell surface. However, it remains possible that these extracellular matrix proteins are present in lower amounts on the luminal cell surface. The cell-type analysis highlighted the possibility that some non-endothelial cells were also labelled by our method. While the highest enrichment was seen for genes that are known to be expressed in endothelial cells themselves, we also saw some enrichment in genes associated with SMC-Peri, VLMC, Micro-PVM, Astro, and Oligo cells suggesting the possibility that some of these cells may also be labelled in our method. Alternatively, the apparent enrichment in proteins associated with these non-endothelial cells may also be due to an artifact caused by the inherent similarity in gene expression between these other cell types and endothelial cells, coupled with the stochastic, random nature of detection of lower-level genes in scRNAseq data. In fact, the cell types that showed the strongest enrichment in our analysis (SMC-Peri, VLMC cells) loosely co-cluster with endothelial cells in the Allen brain map data [35]. Thus, it remains possible that only endothelial cells were labelled in our

experiment and that these other cell types were only highlighted due to their similar expression to endothelial cells. The relative low depth of coverage of these rarer cell types, as evidenced by the smaller number of genes detected, may also affect this analysis.

Similar to proteomic data, scRNAseq data also has an underrepresentation of cell surface proteins due to their relatively low expression levels compared to intracellular proteins involved in metabolism or the cytoskeleton. This leads to a higher percentage of non-detection events for cell surface proteins than more abundant protein types in scRNAseq data, which is one reason why we focused on cell surface genes/proteins in our comparative analysis. In fact, while 11.6% of all genes encode cell surface proteins, only  $7.1\pm 1.0\%$  of the genes detected by scRNAseq in each of the cell type clusters encode cell surface proteins (data not shown). In an outlier analysis, a larger percentage of cell surface protein-encoding genes were identified only in astrocytes, where 12.4% of detected genes encoded cell surface proteins, which likely is the cause of the higher percentage of brain lysate glyco-capture proteins represented in astrocytes relative to other cell types in Fig. 4A, B.

Many cell surface proteins were identified by in vivo glyco-capture. For further study as potential RMT carrier targets, the best options are likely to be proteins that were identified in both mouse and rat with high VE scores. These would include proteins such as Anpep, Esam, Icam2, and Alpl, several of which are known to play a role in endocytic processes including leukocyte extravasation [50–53]. However, it is important to recognize that the VE score is highly biased towards more abundant BBB proteins that contain larger extracellular domains and more accessible N-linked glycopeptides that happen to be well-suited to mass spectrometry analysis. Therefore, not all proteins with lower VE scores should be discounted and any protein with a positive VE score may be considered as a possible RMT candidate. For instance, one of the best-studied and most highly validated RMT targets, transferrin receptor (Tfrc) [54–56], was identified in this study with a VE score of 3617 (83rd percentile; rank 55 out of 347) in rat, but with a much lower VE score of 227 (20th percentile, rank 179 out of 224) in mouse. In this case, the apparent difference in VE scores between the two species appears to be due to technical reasons. Upon manual inspection of the peptide-spectrum-match data, two different unique N-glycopeptides were identified from Tfrc in rat, while only one unique N-glycopeptide was identified in mouse. Furthermore, this one N-glycopeptide had a very late elution time, which makes it more likely for this peptide to be “lost” due to minor variability in the nano-LC column performance between runs. In fact, a closer inspection of the data suggests this may

be the case; even in the rat data, one particular MS run showed a much higher identification of this Tfrc peptide relative to other runs. For this reason, the VE score should be considered as a useful tool for prioritizing possible RMT targets; however, it is also important to evaluate other criteria and additional datasets, as has been previously discussed in the literature [45, 57].

Another example of how technical details can affect the results is Cd59, an interesting potential novel BBB carrier candidate that was found in rat with a high enrichment score, but was not identified at all in mouse. Cd59 is known to be expressed in endothelial cells [58, 59], including in brain as shown by tissue IHC in Protein Atlas (<https://www.proteinatlas.org/>) [60]. In addition, Cd59 was calculated to be 68% luminal and 32% abluminal in HCBE/C3 cells based on apical and basolateral membrane fractionation by sucrose gradient [61] (data not shown). In the *in vivo* glyco-capture experiments, the observed species difference is likely due to differences in the sequence between mouse and human Cd59. Cd59 derives from only one gene in rat, while mouse contains two highly similar Cd59 proteins, Cd59a and Cd59b, with Cd59b showing restricted expression in testes while Cd59a is expressed more widely, including in endothelial cells, similar to Cd59 in rat and human [62]. All Cd59 proteins have only one N-linked glycosylation site. While Cd59b has an N-linked peptide similar to rat Cd59, the N-linked glycosylation site in Cd59a is present in an N-linked tryptic peptide that is too long to be easily identified by LC-MS under the conditions used in these experiments, explaining why this protein was not identified at the mouse BBB by *in vivo* glyco-capture.

A few proteins predicted to be present on the membranes of intracellular organelles or vesicles were identified in these experiments, with Synaptophysin-like 1 (Syp11) and the cation-dependent mannose-6-receptor (M6pr, also known as MPR46) showing the highest vessel enrichment score. The identification of these protein by *in vivo* glyco-capture with a relatively high vessel-enrichment score suggests that they may be expressed on the endothelial cell surface, or possibly in extracellular vesicles that are present at the BBB interface. However, it also remains possible that these proteins were identified due to local leakage of the labelling agent from the vessels or due to labelling of another cell-type interacting with the BBB (such as T cells) which may have been labelled through perfusion. It is intriguing that M6P-mediated uptake at the BBB has been shown to be very low in adult mice [63, 64], although we only identified M6pr in rat, not in mouse, so the possibility of a species difference remains. Interestingly, the cation-independent mannose-6-phosphate receptor (Igf2r, also known as MPR300 and MPRI), was also

identified in rats by *in vivo* glyco-capture and there is some evidence that this protein may be present in brain capillary endothelium in cows and pigs [65]. Lastly, it remains possible that M6pr could be present in mouse, but is held in an inactive state at the BBB. Additional experiments would be required to distinguish from these possibilities.

In addition to the integral cell membrane proteins that were identified in these experiments, a number of secreted proteins were identified as well. The highest scoring secreted proteins by vessel enrichment score included Tubulointerstitial nephritis antigen-like 1 (Tnagl1) and transferrin (Tf). These secreted proteins were likely labelled at the BBB due to their tight interaction with cell receptors that are present on endothelial cells, including integrins  $\alpha 5\beta 1$  and  $\alpha v\beta 1$  for Tnagl1 [66] or transferrin receptor (Tfrc) for Tf, all of which were also identified by *in vivo* glyco-capture.

In this work, a few proteins that would not be expected on the luminal endothelial cell surface were identified, such as Slc1a2 and Slc1a3, which are thought to be mainly expressed on astrocyte endfeet and are therefore associated with the abluminal surface [67]. Interestingly, Slc1a3 was identified in our rat SV-ARBE cell glyco-capture study showing that Slc1a3 is present in endothelial cells, at least in culture. As discussed above, it is possible that the mild oxidation reagent was able to penetrate the endothelial cell and label some abluminal proteins. Alternatively, this protein may be found at some level on the luminal surface, likely at a lower level than on the abluminal surface. Unfortunately, confirming luminal expression of the proteins identified in this work using orthogonal methods is extremely difficult. The luminal and abluminal membranes are separated by only 300–500 nm in human brain microvessels, making these two membranes difficult to resolve by typical light microscopy immunohistochemistry approaches [68, 69]. Electron microscopy studies can distinguish the two endothelial surfaces with high resolution, but require care in selecting the antibody and fixation conditions which can affect relative labelling of luminal and abluminal surfaces [70] while also requiring specialized expertise and significant time to scan vessels throughout the entire brain to identify potential differences in polarization in different brain areas or capillary structures. Interestingly, but perhaps not surprisingly, different endothelial cells have been shown to have different polarization of some transporters such as GLUT1 [68], suggesting that the same may be true for other BBB proteins, further complicating the issue. In a complementary approach, luminal and abluminal membranes can be isolated by fractionation [71, 72], but this appears to be an enrichment rather than a pure separation making interpretation somewhat difficult

and requiring correction for contamination [72]. Additional methods to determine the abluminal and luminal expression of proteins at the BBB is full of challenges and remains an area worthy of further method development.

### Next steps

In this work, we have demonstrated the use of *in vivo* glyco-capture to profile the proteins present at the blood–brain barrier in rat and mouse. In future work, we hope to implement relative quantification of these identified proteins to evaluate how protein expression changes at the BBB in various disease models, including Alzheimer’s and cancer. The application of *in vivo* glyco-capture could also be expanded beyond the brain by using whole body perfusion to identify vessel proteins present in peripheral organs such as kidney, lung, and liver. Lastly, extension of this work into other species, including non-human primates, would be of great benefit to the field.

### Conclusions

We have developed a novel perfusion-based labelling method that enabled the specific isolation and identification of proteins exposed to the luminal surface of microvessels in the brain of rats and mice. The majority of proteins identified were cell surface proteins and many known BBB receptors and transporters were found. The resulting list of identified proteins will serve as a valuable resource for future BBB studies.

### Abbreviations

|          |   |
|----------|---|
| BBB      | Blood–brain barrier                             |
| CCA      | Common carotid artery                           |
| CPM      | Counts per million                              |
| CSC      | Cell surface capture method                     |
| CTL      | Control/buffer-perfused animals                 |
| D-CSC    | Direct cell surface capture method              |
| GO       | Gene ontology                                   |
| Hz CB    | Hydrazide coupling buffer                       |
| LC–MS    | Liquid-chromatography coupled mass spectrometry |
| MS       | Mass spectrometry                               |
| NHS      | N-hydroxysuccinimide                            |
| NVU      | Neurovascular unit                              |
| OXY      | Oxidized/oxidation reagent perfused animals     |
| RMT      | Receptor-mediated transcytosis                  |
| scRNAseq | Single-cell RNAseq                              |
| VE       | Vessel enrichment                               |

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12987-024-00523-x>.

**Additional file 1: Data S1.** Individual protein identification lists for all datasets.

**Additional file 2: Data S2.** MSFragger peptide-spectrum match (psm) details after filtering for all datasets. Note that this table includes hits to the decoy database (if applicable).

**Additional file 3: Figure S1.** Example LC-MS data for control and oxidized rats.

**Additional file 4: Table S1.** Sample and MS run information. **Table S2.** Proteins identified by *in vivo* glyco-capture from rat brain. **Table S3.** Proteins identified by *in vivo* glyco-capture from mouse brain. **Table S4.** All proteins identified by *in vivo* glyco-capture from rat or mouse brain. **Table S5.** All proteins identified by cell surface glyco-capture in a rat endothelial cell line (SV-ARBEC). **Table S6.** All proteins identified by cell surface glyco-capture in a rat endothelial cell line (SV-ARBEC).

### Acknowledgements

The authors would like to thank Luc Tessier and Sam Williamson for their expert help ensuring that the mass spectrometers and LCs were working well prior to sample analysis.

### Author contributions

TLT, WA, JS, EB, CED, MM, and ASH contributed to the design and/or execution of experiments, DSB guided experimental execution and interpreted data, and JJH conceived of the concept, analyzed the data, interpreted data and drafted the manuscript. All authors revised the manuscript and approved the final version.

### Funding

This work was supported by the Beyond Brain Barriers and Multifunctional Antibody Therapeutics Programs at the National Research Council of Canada.

### Availability of data and materials

The full peptide identification results generated during the current study are available as Supplementary Data within this manuscript. All materials are available from commercial sources.

### Declarations

#### Ethics approval and consent to participate

All animal studies were performed in accordance with the Canadian Council on Animal Care and were approved by the Animal Care Committee of the National Research Council Canada.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>Human Health Therapeutics, National Research Council Canada, 100 Sussex Dr., Ottawa, ON K1A 0R6, Canada. <sup>2</sup>Present Address: Biology Program, New York University Abu Dhabi, Saadiyat Island Campus, P.O. Box 129188, Abu Dhabi, United Arab Emirates.

Received: 11 October 2023 Accepted: 19 February 2024

Published online: 04 March 2024

### References

- McConnell HL, Mishra A. Cells of the blood-brain barrier: an overview of the neurovascular unit in health and disease. *Methods Mol Biol.* 2022;2492:3–24.
- Iadecola C. The neurovascular unit coming of age: a journey through neurovascular coupling in health and disease. *Neuron.* 2017;96(1):17–42. <https://doi.org/10.1016/j.neuron.2017.07.030>.
- Stanimirovic DB, Friedman A. Pathophysiology of the neurovascular unit: disease cause or consequence. *J Cereb Blood Flow Metab.* 2012;32(7):1207–21.
- Pardridge WM. The isolated brain microvessel: a versatile experimental model of the blood-brain barrier. *Front Physiol.* 2020;11(May):1–27.

5. Thomsen MS, Routhé LJ, Moos T. The vascular basement membrane in the healthy and pathological brain. *J Cereb Blood Flow Metab.* 2017;37(10):3300–17.
6. Kadry H, Noorani B, Cucullo L. A blood–brain barrier overview on structure, function, impairment, and biomarkers of integrity. *Fluids Barriers CNS.* 2020;17(1):1–24. <https://doi.org/10.1186/s12987-020-00230-3>.
7. Haqqani AS, Delaney CE, Brunette E, Baumann E, Farrington GK, Sisk W, et al. Endosomal trafficking regulates receptor-mediated transcytosis of antibodies across the blood brain barrier. *J Cereb Blood Flow Metab.* 2018;38(4):727–40.
8. Pardridge WM. Blood-brain barrier and delivery of protein and gene therapeutics to brain. *Front Aging Neurosci.* 2020;11(January):1–27.
9. Pulgar VM. Transcytosis to cross the blood brain barrier, new advancements and challenges. *Front Neurosci.* 2019;13(JAN):1–9.
10. Pardridge WM. Receptor-mediated drug delivery of bispecific therapeutic antibodies through the blood-brain barrier. *Front Drug Deliv.* 2023;3(July):1–26.
11. Hill JJ, Haqqani AS, Stanimirovic DB. Proteome of the luminal surface of the blood–brain barrier. *Proteomes.* 2021. <https://doi.org/10.3390/proteomes9040045>.
12. Chun HB, Scott M, Niessen S, Hoover H, Baird A, Yates J, et al. The proteome of mouse brain microvessel membranes and basal lamina. *J Cereb Blood Flow Metab.* 2011;31(12):2267–81.
13. Al Feteisi H, Al-Majdoub ZM, Achour B, Couto N, Rostami-Hodjegan A, Barber J. Identification and quantification of blood–brain barrier transporters in isolated rat brain microvessels. *J Neurochem.* 2018;146(6):670–85.
14. Campeau A, Mills RH, Blanchette M, Bajc K, Malfavon M, Munji RN, et al. Multidimensional proteome profiling of blood-brain barrier perturbation by group B streptococcus. *mSystems.* 2020;5(4):e00368–20.
15. Zajec M, Kros JM, Dekker-Nijholt DAT, Dekker LJ, Stingl C, Van Der Weiden M, et al. Identification of blood-brain barrier-associated proteins in the human brain. *J Proteome Res.* 2021;20(1):531–7.
16. Roesli C, Fugmann T, Borgia B, Schliemann C, Neri D, Jucker M. The accessible cerebral vascular proteome in a mouse model of cerebral  $\beta$ -amyloidosis. *J Proteomics.* 2011;74(4):539–46. <https://doi.org/10.1016/j.jprot.2011.01.010>.
17. Simonian M, Shirasaki D, Lee VS, Bervini D, Grace M, Loo RRO, et al. Proteomics identification of radiation-induced changes of membrane proteins in the rat model of arteriovenous malformation in pursuit of targets for brain AVM molecular therapy. *Clin Proteomics.* 2018;15(1):1–8. <https://doi.org/10.1186/s12014-018-9217-x>.
18. Toledo AG, Golden G, Campos AR, Cuello H, Sorrentino J, Lewis N, et al. Proteomic atlas of organ vasculopathies triggered by *Staphylococcus aureus* sepsis. *Nat Commun.* 2019;1(10):4656.
19. Jin J, Fang F, Gao W, Chen H, Wen J, Wen X, et al. The structure and function of the glycocalyx and its connection with blood-brain barrier. *Front Cell Neurosci.* 2021;15(October):1–9.
20. van Oostrum M, Müller M, Klein F, Bruderer R, Zhang H, Pedrioli PGA, et al. Classification of mouse B cell types using surfaceome proteo-type maps. *Nat Commun.* 2019;10(1):1–9. <https://doi.org/10.1038/s41467-019-13418-5>.
21. Bausch-Fluck D, Hofmann A, Bock T, Frei AP, Cerciello F, Jacobs A, et al. A mass spectrometric-derived cell surface protein atlas. *PLoS ONE.* 2015;10(4):1–22.
22. Wollscheid B, Bausch-Fluck D, Henderson C, O'Brien R, Bibbel M, Schiess R, et al. Mass-spectrometric identification and relative quantification of N-linked cell surface glycoproteins. *Nat Biotechnol.* 2009;27(4):378–86.
23. Tremblay TL, Fauteux F, Callaghan D, Hill JJ. Cell surface profiling of cultured cells by direct hydrazide capture of oxidized glycoproteins. *MethodsX.* 2023;11(August):102349. <https://doi.org/10.1016/j.mex.2023.102349>.
24. Alata W, Paris-Robidas S, Emond V, Bourasset F, Calon F. Brain uptake of a fluorescent vector targeting the transferrin receptor: a novel application of in situ brain perfusion. *Mol Pharm.* 2014;11(1):243–53.
25. Alata W, Yogi A, Brunette E, Delaney CE, van Faassen H, Hussack G, et al. Targeting insulin-like growth factor-1 receptor (IGF1R) for brain delivery of biologics. *FASEB J.* 2022;36(3):1–19.
26. Zhang H, Li X-J, Martin DB, Aebersold R. Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat Biotechnol.* 2003;21(6):660–6.
27. Sun B, Ranish JA, Utleig AG, White JT, Yan X, Lin B, et al. Shotgun glycopeptide capture approach coupled with mass spectrometry for comprehensive glycoproteomics. *Mol Cell Proteomics.* 2007;6(1):141–9. <https://doi.org/10.1074/mcp.T600046-MCP200>.
28. Hill JJ, Tremblay TL, Fauteux F, Li J, Wang E, Aguilar-Mahecha A, et al. Glycoproteomic comparison of clinical triple-negative and luminal breast tumors. *J Proteome Res.* 2015;14(3):1376–88.
29. Kong AT, Leprevost FV, Avtonomov DM, Mellacheruvu D, Nesvizhskii AI. MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. *Nat Methods.* 2017;14(5):513–20.
30. Teo GC, Polasky DA, Yu F, Nesvizhskii AI. Fast deisotoping algorithm and its implementation in the MSFragger search engine. *J Proteome Res.* 2021;20(1):498–505.
31. Sobotzki N, Schafroth MA, Rudnicka A, Koetemann A, Marty F, Goetze S, et al. HATRIC-based identification of receptors for orphan ligands. *Nat Commun.* 2018;9(1):1–8. <https://doi.org/10.1038/s41467-018-03936-z>.
32. Koopmans F, van Nierop P, Andres-Alonso M, Byrnes A, Cijssouw T, Coba MP, et al. SynGO: an evidence-based, expert-curated knowledge base for the synapse. *Neuron.* 2019;103(2):217–234.e4.
33. Thomas PD, Ebert D, Muruganujan A, Mushayama T, Albou LP, Mi H. PANTHER: making genome-scale phylogenetics accessible to all. *Protein Sci.* 2022;31(1):8–22.
34. Mi H, Muruganujan A, Huang X, Ebert D, Mills C, Guo X, et al. Protocol Update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0). *Nat Protoc.* 2019;14(3):703–21. <https://doi.org/10.1038/s41596-019-0128-8>.
35. Yao Z, van Velthoven TJ, Nguyen TN, Goldy J, Sedeno-Cortes AE, Baftizadeh F, et al. A taxonomy of transcriptomic cell types across the isocortex and hippocampal formation. *Cell.* 2021;184(12):3222–3241.e26. <https://doi.org/10.1016/j.cell.2021.04.021>.
36. Betz AL, Firth JA, Goldstein GW. Polarity of the blood-brain barrier: distribution of enzymes between the luminal and antiluminal membranes of brain capillary endothelial cells. *Brain Res.* 1980;192(1):17–28.
37. Webster CI, Caram-Salas N, Haqqani AS, Thom G, Brown L, Rennie K, et al. Brain penetration, target engagement, and disposition of the blood-brain barrier-crossing bispecific antibody antagonist of metabotropic glutamate receptor type 1. *FASEB J.* 2016;30(5):1927–40.
38. Zhang W, Liu QY, Haqqani AS, Leclerc S, Liu Z, Fauteux F, et al. Differential expression of receptors mediating receptor-mediated transcytosis (RMT) in brain microvessels, brain parenchyma and peripheral tissues of the mouse and the human. *Fluids Barriers CNS.* 2020;17(1):1–17. <https://doi.org/10.1186/s12987-020-00209-0>.
39. Palagi PM, Walther D, Quadroni M, Catherinet S, Burgess J, Zimmermann-Ivol CG, et al. MSight: an image analysis software for liquid chromatography-mass spectrometry. *Proteomics.* 2005;5(9):2381–4.
40. Garberg P, Ball M, Borg N, Cecchelli R, Fenart L, Hurst RD, et al. In vitro models for the blood-brain barrier. *Toxicol In Vitro.* 2005;19(3):299–334. <https://doi.org/10.1016/j.tiv.2004.06.011>.
41. Goncharov NV, Nadeev AD, Jenkins RO, Avdonin P.V. Markers and biomarkers of endothelium: when something is rotten in the state. *Oxid Med Cell Longev.* 2017. <https://doi.org/10.1155/2017/9759735>.
42. Boado RJ, Zhang Y, Zhang Y, Pardridge WM. Humanization of anti-human insulin receptor antibody for drug targeting across the human blood-brain barrier. *Biotechnol Bioeng.* 2007;96(2):381–91.
43. Qian ZM, Li H, Sun H, Ho K. Targeted drug delivery via the transferrin receptor-mediated endocytosis pathway. *Pharmacol Rev.* 2002;54(4):561–87.
44. Yu YJ, Zhang Y, Kenrick M, Hoyte K, Luk W, Lu Y, et al. Boosting brain uptake of a therapeutic antibody by reducing its affinity for a transcytosis target. *Sci Transl Med.* 2011. <https://doi.org/10.1126/scitranslmed.3002230>.
45. Zuchero JYJ, Chen X, Bien-Ly N, Bumbaca D, Tong RK, Gao X, et al. Discovery of novel blood-brain barrier targets to enhance brain uptake of therapeutic antibodies. *Neuron.* 2016;89(1):70–82. <https://doi.org/10.1016/j.neuron.2015.11.024>.
46. Demeule M, Currie JC, Bertrand Y, Ché C, Nguyen T, Régina A, et al. Involvement of the low-density lipoprotein receptor-related protein in the transcytosis of the brain delivery vector Angiopep-2. *J Neurochem.* 2008;106(4):1534–44.

47. Chew KS, Wells RC, Moshkforoush A, Chan D, Lechtenberg KJ, Tran HL, et al. CD98hc is a target for brain delivery of biotherapeutics. *Nat Commun.* 2023;14(1):5053.
48. Pyzik M, Kozicky LK, Gandhi AK, Blumberg RS. The therapeutic age of the neonatal Fc receptor. *Nat Rev Immunol.* 2023;23(July):415–32.
49. Christensen SC, Krogh BO, Jensen A, Andersen CBF, Christensen S, Nielsen MS. Characterization of basigin monoclonal antibodies for receptor-mediated drug delivery to the brain. *Sci Rep.* 2020;10(1):1–13. <https://doi.org/10.1038/s41598-020-71286-2>.
50. Butz S, Vestweber D. The role of endothelial cell-selective adhesion molecule (ESAM) in neutrophil emigration into inflamed tissues. *Adhes Mol Funct Inhib.* 2007. [https://doi.org/10.1007/978-3-7643-7975-9\\_11](https://doi.org/10.1007/978-3-7643-7975-9_11).
51. He H, Guo J, Xu J, Wang J, Liu S, Xu B. Dynamic continuum of nanoscale peptide assemblies facilitates endocytosis and endosomal escape. *Nano Lett.* 2021;21(9):4078–85.
52. Tobys D, Kowalski LM, Czudaj E, Müller S, Zentis P, Pach E, et al. Inhibition of clathrin-mediated endocytosis by knockdown of AP-2 leads to alterations in the plasma membrane proteome. *Traffic.* 2021;22(1–2):6–22.
53. Ghosh M, Shapiro LH. CD13 regulation of membrane recycling: implications for cancer dissemination. *Mol Cell Oncol.* 2019. <https://doi.org/10.1080/23723556.2019.1648024>.
54. Johnsen KB, Burkhart A, Thomsen LB, Andresen TL, Moos T. Targeting the transferrin receptor for brain drug delivery. *Prog Neurobiol.* 2019;181(July):101665. <https://doi.org/10.1016/j.pneurobio.2019.101665>.
55. Paterson J, Webster CI. Exploiting transferrin receptor for delivering drugs across the blood-brain barrier. *Drug Discov Today Technol.* 2016;20:49–52. <https://doi.org/10.1016/j.ddtec.2016.07.009>.
56. Pardridge WM. Blood–brain barrier drug delivery of IgG fusion proteins with a transferrin receptor monoclonal antibody. *Expert Opin Drug Deliv.* 2014;12(2):207–22.
57. Haqqani AS. Prioritization of therapeutic targets of inflammation using proteomics, bioinformatics, and in silico cell–cell interactomics. *Methods Mol Biol.* 2019;2024:309–25.
58. Brooimans R, Van der Ark A, Tomita M, Van Es L, Daha M. CD59 expressed by human endothelial cells functions as a protective molecule against complement-mediated lysis. *Eur J Immunol.* 1992;22(3):791–7.
59. Ito S, Oishi M, Ogata S, Uemura T, Couraud PO, Masuda T, et al. Identification of cell-surface proteins endocytosed by human brain microvascular endothelial cells in vitro. *Pharmaceutics.* 2020;12(6):1–19.
60. Uhlén M, Björling E, Agaton C, Szigartyo CAK, Amini B, Andersen E, et al. A human protein atlas for normal and cancer tissues based on antibody proteomics. *Mol Cell Proteomics.* 2005;4(12):1920–32.
61. Zhang W, Bamji-Mirza M, Chang N, Haqqani AS, Stanimirovic DB. The expression and function of ABC transporters at the blood-brain barrier. In: Dorovini-Zis K, editor. *The blood-brain barrier in health and disease Volume 1, morphology, biology and immune function.* Boca Raton: CRC Press; 2015. p. 172–214.
62. Harris CL, Hanna SM, Mizuno M, Holt DS, Marchbank KJ, Morgan BP. Characterization of the mouse analogues of CD59 using novel monoclonal antibodies: tissue distribution and functional comparison. *Immunology.* 2003;109(1):117–26.
63. Urayama A, Grubb JH, Sly WS, Banks WA. Developmentally regulated mannose 6-phosphate receptor-mediated transport of a lysosomal enzyme across the blood-brain barrier. *Proc Natl Acad Sci U S A.* 2004;101(34):12658–63.
64. Pardridge WM. Blood-brain barrier delivery for lysosomal storage disorders with IgG-lysosomal enzyme fusion proteins. *Adv Drug Deliv Rev.* 2022;184:114234. <https://doi.org/10.1016/j.addr.2022.114234>.
65. Siupka P, Hersom MNS, Lykke-Hartmann K, Johnsen KB, Thomsen LB, Andresen TL, et al. Bidirectional apical–basal traffic of the cation-independent mannose-6-phosphate receptor in brain endothelial cells. *J Cereb Blood Flow Metab.* 2017;37(7):2598–613.
66. Shen M, Jiang YZ, Wei Y, Ell B, Sheng X, Esposito M, et al. Tinagl1 suppresses triple-negative breast cancer progression and metastasis by simultaneously inhibiting integrin/FAK and EGFR signaling. *Cancer Cell.* 2019;35(1):64–80.e7.
67. Danbolt NC. Glutamate uptake. *Prog Neurobiol.* 2001;2001(65):1–105.
68. Cornford EM, Hyman S. Localization of brain endothelial luminal and abluminal transporters with immunogold electron microscopy. *NeuroRx.* 2005;2(1):27–43.
69. Pardridge WM. The blood-brain barrier and neurotherapeutics. *NeuroRx.* 2005;2(1):1–2.
70. Devraj K, Klinger ME, Myers RL, Mokashi A, Hawkins RA, Simpson IA. GLUT-1 glucose transporters in the blood-brain barrier: Differential phosphorylation. *J Neurosci Res.* 2011;89(12):1913–25.
71. Hawkins RA, Viña JR, Mokashi A, Peterson DR, OKane R, Simpson IA, et al. Synergism between the two membranes of the blood-brain barrier: glucose and amino acid transport. *Am J Neurosci Res.* 2013;1(November 2017):1–25.
72. Kubo Y, Ohtsuki S, Uchida Y, Terasaki T. Quantitative determination of luminal and abluminal membrane distributions of transporters in porcine brain capillaries by plasma membrane fractionation and quantitative targeted proteomics. *J Pharm Sci.* 2015;104(9):3060–8.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.