

Supporting Information

Ishii et al. 10.1073/pnas.0905936106

SI Methods

Antibodies. α -B-crystallin was from M. Inomata (Tokyo, Japan); Band 4.1G was from M.C. Stankewich (New Haven, CT); Caspr-2 was from NeuroMab; CD47, CD9, and CDC42 were from BD Biosciences; CLIC4 was from S. Yuspa (Bethesda, MD); CNP was from Sternberger Biochemicals; DHPR was from Abnova; Fascin was from Millipore; GFAP was from Dako; MAG was from R.H. Quarles (Bethesda, MD); MBP was from E. Barbabrese (Farmington, CT); MOG (8–18C5) was from C. Linington (Aberdeen, United Kingdom); NDRG-1 was from U. Suter (Zurich, Switzerland); OSP was from Invitrogen; PLP was from M. Lees (Boston, MA); PSD-95 was described in Rasband et al. (1); Sec8 was from BD Biosciences; Sirtuin-2 was from Biomol; and Septin-4, HRP-conjugated goat anti-mouse and anti-rat, and donkey anti-rabbit and anti-goat IgGs were from Santa Cruz Biotechnology. The monoclonal antibody Caspr-2 was developed by and/or obtained from the University of California Davis/National Institute of Neurological Disorders and Stroke/National Institute of Mental Health NeuroMab Facility, supported by National Institutes of Health grant U24NS050606, and maintained by the Department of Pharmacology at the School of Medicine of the University of California, Davis, CA.

Human Brain Tissue. Brain tissues were obtained from 5 human subjects without neurological disease, between the ages of 53 and 77 years (details summarized in Fig. S1). All procurement of these tissues was approved by the Cleveland Clinic Institutional Review Board. The brains were removed by rapid autopsy and frozen at -70°C .

MS and Data Analysis. Myelin samples (150 μg protein) were solubilized in lysis buffer (25 mM Tris, 5 mM EDTA, protease inhibitor mixture) containing 1% amidosulfofetaine-14 (ICN Biochemical) at room temperature for 30 min, then precipitated with 2 volumes of 100% ethanol at -20°C overnight. The precipitate was solubilized in urea buffer (50 mM Tris-HCl, 5% SDS, 4 M Urea, 10 mM DTT, 2.5% glycerol, 0.01% bromophenol blue, pH 6.8). For the electrophoretic separation of each sample, 150 μg of protein was loaded in 3 combined wells on a 10% NuPAGE gel (Invitrogen). NuPAGE gels were Coomassie-stained and each gel lane was cut from top to bottom at approximately 2-mm intervals. The gel slices were digested with trypsin and analyzed as described previously (1). Digested proteins were analyzed using a linear ion trap mass spectrometer (Finnigan LTQ; Thermo Finnigan). Samples were loaded onto a 15-cm \times 100- μm capillary C18 reversed-phase column by a micro-autosampler (Famos; Dionex) and washed with Buffer A (5% acetonitrile, 0.4% acetic acid, 0.005% heptafluorobutyric acid), followed by LC-MS/MS analysis on the LTQ with 5% to 35% Buffer B (100% acetonitrile, 0.4% acetic acid, 0.005% heptafluorobutyric acid) linear gradient for 68 min. Each full MS scan was followed by 5 MS/MS scans of the 5 most intense peaks in the MS spectrum with dynamic exclusion enabled. The m/z scan range was 400 to 1,700 for full mass range. All data were searched against a non-redundant protein database (56,709 human entries, 42,910 mouse entries; Release 20041130; Advanced Biomedical Computing Center) using the SEQUEST algorithm (2). SEQUEST parameters were outlined as follows: mass tolerance of 1.0 Da for precursor ions and 0.0 Da for fragment ions, full tryptic constraint allowing 1 missed cleavage, allowing oxidation (+16 Da) of methionine. The database

search results were processed using INTERACT program and filtered with the following criteria: X_{corr} cutoff values are 1.9, 2.2, and 3.7 for 1+, 2+, and 3+ peptides, respectively; ΔC_n cutoff value is 0.1; proteins identified by at least 2 unique peptides in the same experimental fraction were accepted as high-confidence identification. To estimate the false-positive rate, the data sets were searched against a concatenated protein database with forward and reverse protein sequence (3). A peptide false-positive rate less than 0.75% was observed for samples. The probabilities of peptide identifications in our data set were computed using PeptideProphet software (4, 5). The complete INTERACT peptide sequence data and PeptideProphet scores for mouse and human subjects are tabulated in the Table S2. To enable comparison between human and mouse species National Cancer Institute identifiers (2004) were converted to the UniProt database (2007) entry using the protein ID. Proteins with more than one protein ID listed in column “Human Protein ID” and/or “Mouse Protein ID” represent redundant protein entries that have been combined into one record in the UniProt Database.

Certain proteins are precluded from MS detection because of their amino acid sequence. Specifically, tryptic proteolytic cleavage of proteins into smaller peptides is necessary for MS. Peptides that are too long or too short cannot be detected. Further transmembrane regions of the proteins cannot be analyzed by MS. For example, MAL/MVP17, a transmembrane myelin protein, could not be detected because it has only 3 tryptic cleavage sites, resulting in 4 peptides that are 1, 2, 29, and 119 aa long. In our study, the range for detectable peptide length was set at approximately 9 to 40 aa. Thus, only 1 peptide (29 aa) was potentially eligible for MS detection. However, this peptide forms part of the transmembrane region, and therefore could not be detected.

Microarray of Human CNS White and Gray Matter. Tissue blocks were visually inspected for the gray matter and white matter junction. Tissue for RNA isolation was obtained by separating gray and white matter with a scalpel and collecting 60- μm sections on a cryostat. Eight white matter and 8 gray matter samples from 3 non-neurological disease control brains (Fig. S1) were used for RNA isolation. We collected both gray matter and white matter from the same tissue blocks to minimize intra-sample variations. RNA was isolated according to Promega RNagents total RNA isolation system protocol (Promega), and integrity and 28S/18S ratios were determined with an Agilent 2100 Bioanalyzer (Agilent Technologies). Biotinylated cRNA probes (Affymetrix) made from 8 μg total RNA were hybridized to Affymetrix human genome U133A/U133B arrays containing 33,000 characterized genes (www.affymetrix.com/support/technical/datasheets/hgu133datasheet.pdf) according to manufacturer's instructions (Affymetrix). Microarrays were washed and stained with streptavidin-phycoerythrin conjugate (Molecular Probes) and biotinylated anti-streptavidin (Vector Laboratories), and scanned on a GS 450 gene array scanner (Hewlett Packard). Data retrieval was carried out using MAS 5.0 software (Affymetrix) and default normalizing and scaling options. The white matter and gray matter samples were processed identically. Intensity values were imported into Genespring (Agilent Technologies) and queried for the list of proteins. Average intensity values ($n = 8$) for each gene were used to determine W/G ratios (Fig. 3 in the main text). For genes reported not present (marked with an X in Fig. 3 in the main text), homologous gene names

were also searched to confirm its absence from the arrays. Transcripts that were below the threshold of detection (i.e., lower than background intensity) or called “absent” across all replicates by MAS5.0 gene flagging parameter was described as low (marked with an *X* in Fig. 3 in the main text). Using gene-mapping function (Genespring; Agilent Technologies), human IDs corresponding to mouse transcripts identified by Cahoy et al. (6) were identified. Intensity ratios of each transcript between white and gray matter was used to determine fold enrichment levels of different mRNAs in human white matter over gray matter.

Comparative Analysis of Our Proteome with Previous Transcriptome Database of Rodent Purified OLs, Neurons, and Astrocytes. To determine the cellular origin of the proteins identified in purified myelin fraction by NS, we compared our proteome database of 308 proteins to the transcriptome database reported by Cahoy et al. (6; supplemental data at <http://www.jneurosci.org/cgi/content/full/28/1/264/DC1>).

Gene IDs present in supplemental tables 4 and 6 of Cahoy et

al. (6) were used to select genes enriched in “astrocytes” and “neurons” among the proteins identified in our proteomic assay. To identify genes specifically enriched in OLs compared with OPCs, we performed a subtractive approach between genes in supplemental tables 5 (containing genes that were enriched in “OL-lineage cells” over “astrocytes” and “neurons”) and 17 (i.e., genes statistically enriched in OPCs) of Cahoy et al. (6). Transcripts that are present in only supplemental table 5 and not in supplemental table 17 (6) were selected as they represent genes associated with differentiated OLs. In addition, we also added transcripts from supplemental tables 18 (i.e., “genes statistically enriched in myelinating OLs compared to OPCs”) and 19 (i.e., “genes statistically enriched in MOG+ myelinating OLs compared to MOG negative OLs”) (6) to prepare the comprehensive list of 64 transcripts associated with mature OLs. It should be noted that, in addition to these, we have found another 47 proteins that have been previously reported in OLs (tan squares in column O in Fig. 3 in the main text). Together, this shows that at least 111 proteins identified in our myelin proteome are expressed by mature OLs, which strongly validates our MS identification of these proteins in myelin.

1. Rasband MN, et al. (2002) Clustering of neuronal potassium channels is independent of their interaction with PSD-95. *J Cell Biol* 159:663–672.
2. Eng JK, McCormack AL, Yates JR (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectrom* 5:976–989.
3. Peng J, et al. (2003) A proteomics approach to understanding protein ubiquitination. *Nat Biotechnol* 21:921–926.
4. Keller A, Nesvizhskii AI, Kolker E, Aebersold R (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal Chem* 74:5383–5392.
5. Nesvizhskii AI, Keller A, Kolker E, Aebersold R (2003) A statistical model for identifying proteins by tandem mass spectrometry. *Anal Chem* 75:4646–4658.
6. Cahoy JD, et al. (2008) A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function. *J Neurosci* 28:264–278.
7. Taylor CM, et al. (2004) Proteomic mapping provides powerful insights into functional myelin biology. *Proc Natl Acad Sci USA* 101:4643–4648.
8. Roth AD, Ivanova A, Colman DR (2005) New observations on the compact myelin proteome. *Neuron Glia Biol* 2:15–21.
9. Vanrobaeys F, Van Coster R, Dhondt G, Devreese B, Van Beeumen J (2005) Profiling of myelin proteins by 2D-gel electrophoresis and multidimensional liquid chromatography coupled to MALDI TOF-TOF mass spectrometry. *J Proteome Res* 4:2283–2293.
10. Werner HB, et al. (2007) Proteolipid protein is required for transport of sirtuin 2 into CNS myelin. *J Neurosci* 27:7717–7730.

Sample	Sex	Age (yr)	PMI (hr)	Cause of death	Brain region
Subject 1	F	53	36	Cardiopulmonary arrest	frontal lobe
Subject 2	M	65	8.5	Cardiac arrest	frontal lobe
Subject 3	M	53	12	Myocardial infarct	frontal lobe
Subject 4	F	65	7	Myocardial infarct	frontal lobe
Subject 5	M	77	5	Mesenteric bleeding	frontal lobe

Fig. S1. Summary of human subjects and brain tissues used in the study. Brain from subjects 3 and 4 was used for proteomic analysis, and brain from subjects 1, 3, and 5 was used for microarray analysis. PMI, postmortem interval.

Protein Name (Protein ID*)	Gene Symbol**	mRNA					
		W/G	p-value	O	N	A	U
1 Actin-binding protein anillin (ANLN)	ANLN	>>7.5	0.165				
2 Contactin-2 (CNTN2)	CNTN2	>>7.5	0.015				
3 Myelin-associated glycoprotein precursor (MAG)	MAG	7.4	0.000				
4 Tubulin polymerization-promoting protein family member 3 (TPPP3)	TPPP3	5.8	0.000				
5 Myelin basic protein (MBP)	MBP	5.3	0.009				
6 Myelin proteolipid protein (MYPR)	PLP1	5.1	0.000				
7 Protein S100-B (S100B)	S100B	4.3	0.001				
8 Myelin-oligodendrocyte glycoprotein (MOG)	MOG	3.8	0.007				
9 Septin-4 (SEPT4)	IB SEPT4	3.6	0.004				
10 Phosphatidylinositol-5-phosphate 4-kinase type-2 alpha (PI42A)	PIP4K2A	3.5	0.000				
11 Tight junction protein ZO-2 (ZO2)	TJP2	3.4	0.000				
12 Breast carcinoma amplified sequence 1 (BCAS1)	BCAS1	3.3	0.007				
13 Disheveled-associated activator of morphogenesis 2 (DAAM2)	DAAM2	3.3	0.004				
14 Differentiation-related gene 1 protein (NDRG1)	IB NDRG1	2.9	0.010				
15 Junctional adhesion molecule C (JAM3)	JAM3	2.9	0.188				
16 Chloride intracellular channel protein 4 (CLIC4)	IB CLIC4	2.8	0.016				
17 Myc box-dependent-interacting protein 1 (BIN1)	BIN1	2.8	0.000				
18 Tubulin polymerization-promoting protein (TPPP)	IB TPPP	2.7	0.003				
19 Dihydropteridine reductase (DHPR)	IB QDPR	2.6	0.004				
20 Dynamin-2 (DYN2)	DNM2	2.1	0.084				
21 Band 4.1-like protein 2 (E41L2)[Band 4.1 G]	IB EPB41L2	2.1	0.003				
22 Septin-7 (SEPT7)	IB SEPT7	1.9	0.016				
23 Septin-2 (SEPT2)	SEPT2	1.8	0.032				
24 NAD-dependent deacetylase sirtuin-2 (SIRT2)	IB SIRT2	1.7	0.003				
25 Phosphoserine aminotransferase (SERC)	PSAT1	1.6	0.003				
26 Annexin A5 (ANXA5)	ANXA5	1.6	0.071				
27 Oligodendrocyte-specific protein (CLD11)	CLDN11	1.6	0.007				
28 D-3-phosphoglycerate dehydrogenase (SERA)	PHGDH	1.5	0.044				
29 Dihydropyrimidinase-related protein 2 (DPYL2)	DPYSL2	1.5	0.130				
30 Ras-related C3 botulinum toxin substrate 1 (RAC1)	RAC1	1.5	0.043				
31 Potassium voltage-gated channel subfamily A member 1 (KCNA1)	KCNA1	1.5	0.173				
32 Synaptic vesicle membrane protein VAT-1 homolog (VAT1)	VAT1	1.4	0.160				
33 CD81 antigen (CD81)	CD81	1.3	0.395				
34 Growth-arrest-specific protein 7 (GAS7)	GAS7	1.2	0.209				
35 Adenylyl cyclase-associated protein 1 (CAP1)	CAP1	1.2	0.111				
36 Carbonic anhydrase 14 (CAH14)	CA14	1.1	0.375				
37 Neural cell adhesion molecule 2 (NCAM2)	NCAM2	1.1	0.732				
38 Glutathione S-transferase P (GSTP1)	GSTP1	1.1	0.173				
39 Basolateral Na-K-Cl symporter (S12A2)	SLC12A2	1.1	0.805				
40 Reticulon-4 (RTN4) [Nogo-A]	RTN4	1.0	0.512				
41 Engulfment and cell motility protein 1 (ELMO1)	ELMO1	1.0	0.900				
42 Spectrin beta chain, brain 1 (SPTB2)	SPTBN1	1.0	0.910				
43 Syntaxin-6 (STX6)	STX6	1.0	0.798				
44 cAMP-dependent protein kinase I-alpha regulatory subunit (KAPO)	PRKAR1A	1.0	0.981				
45 Lethal(2) giant larvae protein homolog 1 (L2GL1)	LLGL1	1.0	0.993				
46 Enoyl-CoA hydratase (ECHM)	ECH1	0.9	0.717				
47 B-cell receptor-associated protein 31 (BAP31)	BCAP31	0.9	0.066				
48 Heme oxygenase 2 (HMOX2)	HMOX2	0.9	0.319				
49 Rho GDP-dissociation inhibitor 1 (GDIR)	ARHGDI	0.9	0.547				
50 Rab GDP dissociation inhibitor alpha (GDIA)	ARHGDI	0.9	0.548				
51 Neurofascin precursor (NFASC)	NFASC	0.9	0.012				
52 2',3'-cyclic-nucleotide 3'-phosphodiesterase (CN37)	CNP	0.8	0.388				
53 ATP-citrate synthase (ACLY)	ACLY	0.8	0.223				
54 Receptor expression-enhancing protein 5 (REEP5)	REEP5	0.7	0.091				
55 Ras-related protein Rap-2a precursor (RAP2A)	RAP2A	0.7	0.066				
56 Septin-9 (SEPT9)	SEPT9	0.7	0.020				
57 Peroxiredoxin-2 (PRDX2)	PRDX2	0.7	0.272				
58 Synaptic vesicle glycoprotein 2A (SV2A)	SV2A	0.3	0.020				
59 p130Cas-associated protein (SNIP)	SNIP	x					
60 Cell adhesion molecule 4 precursor (CADM4)	CADM4	x					
61 Choline transporter-like protein 1 (CTL1)	SLC44A1	x					
62 Myosin-1d (MYO1D)	MYO1D	x					
63 Radixin (RADI)	RDX	x					
64 Immunoglobulin superfamily member 8 (IGSF8)	IGSF8	x					

Fig. S2. Short list of identified myelin proteins whose transcripts are enriched in mature oligodendrocytes. This list of 64 myelin proteins has been generated through analysis of multiple databases, combining proteomics and microarray analysis. The short-listing criteria are based on the enrichment of their transcripts in OLs (column O) but not in neurons (column N) and only rarely in astrocytes (column A), determined by analysis of previous microarray databases (6). The relative levels of mRNA signals of these proteins that were detected in white matter (W) over gray matter (G) are shown from highest to lowest, the top half being 7.4- to 1.5-fold enriched in white over gray matter (proteins above dotted line). Comparison with 4 other proteomes (7–10) shows that 31 of the proteins in this list are being reported for the first time in a myelin proteome (column U), 11 of which fall in the top half of the list. Eight of the proteins from the top half of the short list were selected for validation in multiple myelin samples by immunoblotting (IB), as shown in Fig. 4. *Human protein ID in the UniProt Knowledgebase; **Gene symbols as described in National Center for Biotechnology Information database. mRNA, Microarray analysis; X, transcripts below reliable detection threshold; P value, significance of mRNA enrichment between samples obtained from the intensity values of white and gray matter using a 2-sided t test assuming unequal variances.







Proteins	Detected by Mass Spec.		Detected by Immunoblotting	
	Mouse	Human	Mouse	Human
CD9	No	Yes		
CD47	No	Yes		
Fascin	No	Yes		
α B-Crystallin	No	Yes		
CDC42	Yes	No		
Sec8	Yes	No		

Fig. S3. Proteins that failed to be detected by MS in one species could be detected by immunoblot analysis in that species. A significant number of proteins were detected in this study by MS in either human or mouse only. In an effort to test if these proteins were truly species-specific in myelin, immunoblot analysis was performed on a select group of such proteins as an example. All 6 proteins that were examined that had failed to be detected by MS in that species could be detected by immunoblot analysis. This suggests that lack of detection of a given protein by MS in one species may result from technical limitations and may not signify true species differences.

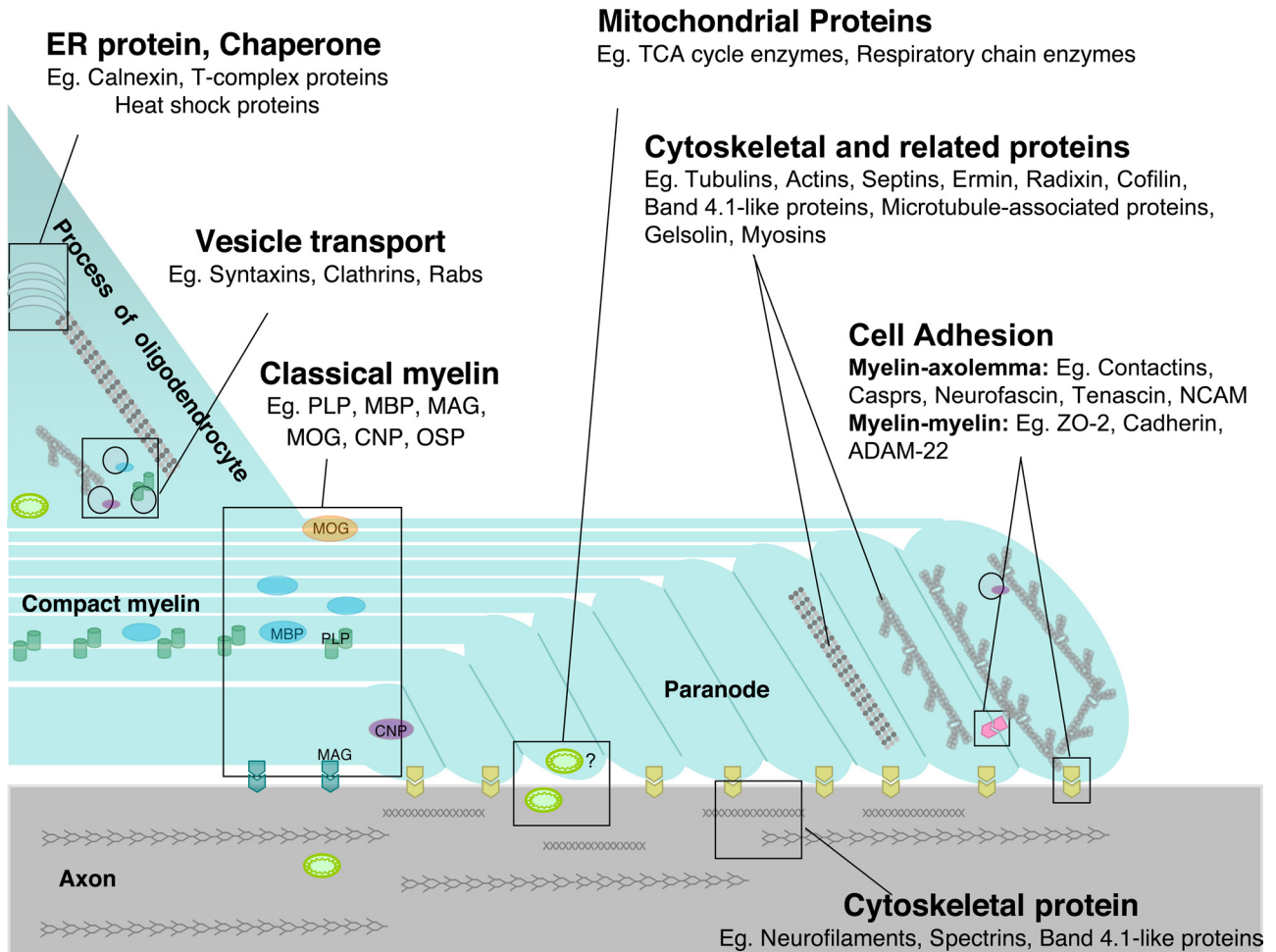


Fig. S4. Diagrammatic localization of identified myelin proteins from different functional categories. To place proteins that were identified in myelin fractions into specific biological contexts, they were categorized into different functional subgroups, including cytoskeletal, cell adhesion, vesicular transport, energy metabolism, endoplasmic reticulum, and chaperone proteins. Examples of proteins in these categories and their known/predicted localizations are presented diagrammatically. Because myelin and axolemma are tightly associated, some proteins that were identified in this study are likely to be of axonal origin or common to both myelin and axon, as depicted.

Other Supporting Information Files

[Table S1](#)

[Table S2](#)