ARTICLE IN PRESS

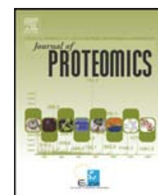
JPROT-02455; No of Pages 8

Journal of Proteomics xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Journal of Proteomics

journal homepage: www.elsevier.com/locate/jprot



Large-scale identification of membrane proteins based on analysis of trypsin-protected transmembrane segments

O. Vit [a], P. Man [b,c], A. Kadek [b,c], J. Hausner [b,c], J. Sklenar [d], K. Harant [e], P. Novak [b,c], M. Scigelova [f], G. Woffendin [g], J. Petrak [a]

[a] BIOCEV, First Faculty of Medicine, Charles University in Prague, Czech Republic
[b] Institute of Microbiology, Czech Academy of Sciences, Prague, Czech Republic
[c] Department of Biochemistry, Faculty of Science, Charles University in Prague, Czech Republic
[d] The Sainsbury Laboratory, Norwich Research Park, Norwich, UK
[e] Protemics Core Facility, Faculty of Science, Charles University in Prague, Czech Republic
[f] Thermo Fisher Scientific, Bremen, Germany
[g] Thermo Fisher Scientific, Hemel Hempstead, UK

ARTICLE INFO

Article history:
Received 30 November 2015
Received in revised form 3 February 2016
Accepted 4 March 2016
Available online xxxx

Keywords:
Integral membrane proteins
CNBr
Transmembrane
Missing proteins
Hydrophobicity
Lymphoma

ABSTRACT

Integral membrane proteins are generally under-represented in routine proteomic analyses, mostly because of their relatively low abundance, hydrophobicity and lack of trypsin-cleavage sites. To increase the coverage of membrane proteomes, various strategies have been developed, targeting mostly the extra-membrane segments of membrane proteins. We focused our attention to the rather overlooked hydrophobic transmembrane segments. Such peptides can be isolated after carbonate stripping and protease "shaving" of membranes isolated by simple centrifugation procedure. The treated membranes with embedded hydrophobic peptides can then be solubilized in organic solvents, re-digested with CNBr, delipidated and subjected to LC-MS/MS analysis. We modified the original "hppK" method, and applied it for the analysis of human lymphoma cells. We identified 1224 proteins of which two-thirds were IMPs with 1–16 transmembrane segments. This method allowed us to identify 13 "missing proteins" — proteins with no previous evidence on protein level.

Biological significance: Integral membrane proteins execute numerous essential functions and represent substantial part of eukaryotic proteomes. Our knowledge of their function and expression is, however, limited. Novel approaches extending our knowledge of membrane proteome are therefore highly desired. As we demonstrate here, a non-conventional method which targets rather overlooked hydrophobic transmembrane segments of integral membrane proteins has wide potential to provide the missing information on the membrane proteome. We show that it can deliver identification and potentially also quantification of hundreds of integral membrane proteins including the so called "missing proteins".

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Approximately one third of eukaryotic genes code for integral membrane proteins (IMPs). These molecules execute important functions — namely signal transduction, transmembrane transport, cell-cell communication, cell adhesion to extracellular matrix and many other processes [1]. Their roles make IMPs ideal pharmacological targets. In fact, membrane proteins are the targets of 50–60% of all currently approved drugs [2]. Yet, because of relatively low expression and, above all, their hydrophobicity, IMPs are generally under-represented in routine proteomic analyses [3]. Isolation of membrane-enriched fractions is often inefficient, being plagued by extensive contamination with major cytosolic proteins, components of cytoskeleton and other proteins. Carbonate washing only partially reduces the problem of nonspecific adhesion, as the percentage of IMPs of all identified proteins in isolated membrane fractions is typically only up to 40–50% [4–6]. Presence of both hydrophilic and hydrophobic domains makes IMPs amphipathic and therefore difficult to solubilize. Use of harsh conditions including high concentration of organic solvents or inclusion of detergents may interfere with protease digestion and/or MS analysis and often requires laborious removal [7,8]. Analysis is further complicated by the fact, that the hydrophobic (mostly alpha-helical) transmembrane domains of IMPs are resistant to digestion by conventional proteases due to inherently low solubility and the absence of specific cleavage sites.

To overcome some of these obstacles, methods targeting selectively either the hydrophilic (extra-membrane) or the hydrophobic (transmembrane, TM) sections of IMPs have been introduced. Extracellular hydrophilic sections of plasma membrane proteins can be targeted by surface biotinylation of amine residues or glycosylated aminoacids, followed by protease digestion and affinity purification of labeled extracellular peptides [9,10]. This strategy of "surfaceome" analysis has been applied to human leukemia cells [11,12], adipocytes [13], glioblastoma

http://dx.doi.org/10.1016/j.jprot.2016.03.016
1874-3919/© 2016 Elsevier B.V. All rights reserved.

Please cite this article as: O. Vit, et al., Large-scale identification of membrane proteins based on analysis of trypsin-protected transmembrane segments, J Prot (2016), http://dx.doi.org/10.1016/j.jprot.2016.03.016

ARTICLE IN PRESS

[14], stem cells [15,16] and other cell types. The surface capture method provides very high enrichment of IMPs, 45–90% of identified proteins in these analyses contain at least one predicted TM segment. The strategy is, however, limited to intact live cells and does not allow analysis of frozen samples, biopsies or post-mortem tissues, unless lectin-based glycocapture is used instead of surface biotinylation as recently demonstrated [17].

Alternative or rather complementary approach targets the overlooked "dark side" of IMPs — hydrophobic membrane-embedded segments. This unique approach has been developed by Adele Blackler and Christine Wu [18] and takes advantage of the fact, that the phosholipid bilayer protects the transmembrane peptides from enzymatic proteolysis. Isolated membrane vesicles can thus be opened at high pH and all non-protected non-membrane proteins can be "shaved" by a protease from both surfaces of the membranes. The treated membranes with embedded hydrophobic peptides can then be solubilized in organic solvents, re-digested with CNBr, delipidated and subjected to LC-MS. The method is not limited to plasma membrane and can also be applied to frozen samples or tissue biopsies. Despite its obvious potential to tap the valuable information hidden in the phospholipid bilayer, the strategy has not attracted many followers, possibly because of the laborious multi-step workflow and application of less common cleavage strategies (non-specific proteinase K and CNBr).

We followed the directions set up by Blackler and Wu [18], modified the method and applied it for a large-scale analysis of membrane proteins in human lymphoma cells. Our modification avoids ultracentrifugation step, includes on-column sample delipidation and, most importantly, replaces the non-specific proteinase K employed in the original method with trypsin. This critical modification prevents the analysis of numerous overlapping peptides generated by the non-specific proteinase K and thus makes the technique more sensitive and makes it compatible with quantitative analysis using SILAC labeling. Using this modified version of the technique, we identified 1224 proteins in the lymphoma cell line sample, of which two-thirds (802 proteins) were IMPs with 1–16 transmembrane segments according to the Tied Mixture Hidden Markov Model (TMHMM) prediction [19].

2. Methods

2.1. Isolation of membranes, carbonate stripping and trypsin digestion

Human mantle cell lymphoma Mino cells (CRL-3000, purchased from ATTC) were grown in 75 cm^2 cultivation flasks in Iscove's Modified Dulbecco's Medium (Lonza) and upon harvesting were washed in chilled Dulbecco's Phosphate Buffered Saline (Sigma). The harvested cells were stored at −80 °C until further processing. The cell pellets were resuspended in hypotonic lysis buffer (10 mM NaCl, 10 mM HEPES pH 7.5), kept on ice for 20 min, and homogenized by passing 20 times through a gauge 25 hypodermic needle. The homogenate was centrifuged (500 ×g, 5 min, 4 °C) to pellet unbroken cells and nuclei. The supernatant was treated with bovine deoxyribonuclease I (Sigma) (120 Kunitz units, 60 min incubation at 37 °C with 25 mM $MgCl_2$ and 5 mM $CaCl_2$ added) and centrifuged in a benchtop centrifuge (18,000 ×g, 30 min, 4 °C). The pellet was solubilized in ice-cold 100 mM Na_2CO_3, agitated for 30 min on ice and centrifuged again (18,000 ×g, 30 min, 4 °C). Resulting pellet was solubilized in 25 mM $(NH_4)HCO_3$ with sequencing grade modified porcine trypsin (Promega) and incubated at 37 °C overnight. After the tryptic digestion the suspension was again centrifuged at 18,000 ×g, the resulting pellet was solubilized in ice-cold 100 mM Na_2CO_3, agitated for 30 min on ice, and the suspension was frozen and thawed 3 times, before being finally centrifuged at 18,000 ×g.

2.2. Electron microscopy

Three aliquots were taken during the process of membrane isolation: A) after the first centrifugation at 18,000 ×g, B) after the first incubation in Na_2CO_3 and C) after digestion with trypsin. Upon centrifugation at 18,000 ×g, all four aliquots were resuspended in 250 mM sucrose at 1:1 v/v ratio. Fixation was done by overlaying samples with 2.5% glutaraldehyde in 0.1 M cacodylate buffer (pH 7.2), incubation for 30 min on ice and centrifugation at 18,000 ×g, 20 min, 0 °C. Postfixation was done in 2% OsO_4 in the same buffer. Fixed tissue was dehydrated through an ascending ethanol series and embedded in Araldite – Poly/Bed® 812 mixture (Polysciences). Thin sections were cut on a Reichert–Jung Ultracut E ultramicrotome and stained using uranyl acetate and lead citrate. Sections were examined and photographed using a JEOL JEM-1011 transmission electron microscope.

2.3. Preparation of standard membrane protein sample

Proton coupled chloride transporter from *Escherichia coli* (CLC-EC1, UniProt accession number P37019) was expressed and purified as described previously [20]. Protein was precipitated using acetone in the presence or absence of a model lipid, 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC, Avanti Polar Lipids).

2.3.1. Cyanogen bromide digestion

The pellet of isolated membranes (approx. 60 μl) was resuspended in 100 μl 70% trifluoroacetic acid (TFA) with 20 mg/ml cyanogen bromide (CNBr) and digested in dark at room temperature overnight. Digested peptides were then dried in speedvac and twice solubilized in 70% methanol and dried again to remove the remaining CNBr.

Standard protein was digested by CNBr according to the same protocol. Five micrograms of pure or DMPC-spiked CLC-EC1 were subjected to digestion overnight. DMPC alone was processed in the same manner as well. For method comparison purposes, CLC-EC1 was also digested by the original procedure as described previously by [18], where the samples were digested in 90% FA and no CNBr removal step followed the reaction.

2.3.2. Standard sample preparation for LC-MS analysis

To optimize the sample preparation method, the standard samples containing CLC-EC1, DMPC or their mixture were processed by three different protocols. In two of those, dried sample was reconstituted in 80% acetonitrile/10% H_2O/10% formic acid, sonicated and diluted with solvent A (0.5% formic acid in H_2O) to 10% acetonitrile concentration. The sample was then loaded on a peptide Macrotrap in an off-line holder (Optimize Technologies) and desalted using solvent A. Next, the trap was either subjected to direct elution with 150 μl solvent B (90% acetonitrile/5% water/5% formic acid) or to a modification of lipid removal procedure described previously [21,22]. In the latter, the trap was washed (sample delipidated) with dichloromethane/0.5% FA (up to 2500 μl). After delipidation the elution was done with 150 μl solvent B. In all cases the trap was finally cleaned by 40% acetonitrile/20% methanol/10% isopropanol/10% formic acid/10% water. Alternatively the sample was also processed by the procedure described by Blackler et al., where the supernatant as well as the pellet after lipid precipitation was desalted [18]. In all the three protocols, aliquots collected at every step of the preparation were analyzed by mass spectrometry.

Following the method optimization, the cyanogen bromide digested transmembrane peptides isolated from Mino cells were processed according to the procedure with dichloromethane delipidation on a peptide trap. Desalted and delipidated peptides were dried in speedvac and stored at −80 °C.

2.3.3. MS analysis of the standard sample

Development of the sample preparation procedure was monitored by MALDI-FT-ICR and ESI-FT-ICR on a 12T SolariX XR (Bruker Daltonics). Samples from the individual steps of offline clean-up and delipidation were either directly infused via an ESI source or spotted on a MALDI plate and overlaid with α-cyano-4-hydroxycinnamic acid. MALDI

Please cite this article as: O. Vit, et al., Large-scale identification of membrane proteins based on analysis of trypsin-protected transmembrane segments, J Prot (2016), http://dx.doi.org/10.1016/j.jprot.2016.03.016

ARTICLE IN PRESS

spectra were recorded with mass range from 400 to 6000 and ESI spectra from 200 to 3000.

2.4. LC-MS/MS analysis

For the trans-membrane peptide sample, the material was dissolved in 5 μl of 10% aqueous formic acid containing 80% (v/v) acetonitrile. The sample was further diluted by the addition of 4 volumes (20 μl) of water, and 5 μl of that solution were loaded onto a Thermo Scientific PepMap peptide trap (internal diameter 100 μm, length 2 cm, 5 μm particle size) using 0.05% aqueous formic acid containing 2% acetonitrile at flow rate 8 μl/min. Peptides were further separated on Thermo Scientific Acclaim EasySpray PepMap C18 RSLC column (internal diameter 75 μm, length 50 cm, 2 μm particle size, 100 Å pore size) maintained at a constant temperature (40 °C) and equilibrated with 2.4% (v/v) acetonitrile in 0.1% (v/v) aqueous formic acid (FA). Pumping was done by nano UHPLC (Easy-nLC 1000; Thermo Fisher Scientific) with a 90 min linear gradient of solvent B (80% acetonitrile in 0.1% aqueous formic acid) in solvent A (0.1% aqueous formic acid), according to the following profile: 0–5 min 10% B; 5–10 min ramping to 15% B; 10–95 min ramping to 55% B; 95–100 min ramping to 100% B; 100–104 min hold at 100% B; 104–105 min return to 10% B; 105–120 min equilibration at 10% B; flow rate of 300 μl/min. Total run time was 120 min.

Separated peptides were detected with the quadrupole-Orbitrap mass analyzer (Q Exactive; Thermo Fisher Scientific) using a data dependent acquisition in positive mode. Peptide parent ions were detected in a high resolution full scan (mass range 350–1500 *m/z*, 70,000 resolving power setting at *m/z* 200). The instrument was set so that 10 most intense ions of every full scan spectrum, meeting specific threshold criteria (minimum intensity 1.7×10^4 , charge state >1), were selected for MS/MS. Peptides were isolated with an isolation window of 3 Da, fragmented (HCD fragmentation with NCE 27 collision energy setting), and the resulting fragment ions were detected (17,500 resolving power setting). Other settings: target value 3×10^6 and 1×10^5 for full scan and MS/MS scan, respectively; maximum ion time 50 ms and 120 ms for full scan and MS/MS scan, respectively. Following their fragmentation the precursors were put on an exclusion mass list for 30 s (exclusion width ± 10 ppm from the parent ion).

Data processing: Proteome Discoverer v. 1.4 (Thermo Fisher Scientific) software package was used for protein identification. The spectra were searched using Thermo Fisher Scientific Sequest HT search engine against the human subset of Swiss-Prot database with added contaminant protein sequences (20,249 sequences in total). The search settings differed between the standard and membrane sample in the cleavage specificity, trypsin with 3 missed cleavage sites and cyanogen bromide (semi-specific) with 2 missed cleavage sites, respectively. Also, there was an extra dynamic modification, Met changing to homoserine lactone (Met −48.003) on any peptide C-terminus, defined for the membrane sample search. The other settings were the same for both samples: precursor mass tolerance 10 ppm; fragment mass tolerance 60 mDa; oxidation of Met residues (+15.995) set as dynamic modification; maximum 3 equal dynamic modifications per peptide allowed. The search results were validated with decoy database search strategy using Percolator [23] with target FDR 0.01 and validation based on q-value.

3. Results and discussion

The pioneering work of Adele Blackler and Christine Wu [18] tapped the previously unexplored information on transmembrane segments of membrane proteins buried in the lipid bilayer — it enabled identification of almost 500 IMPs. The use of proteinase K at high pH facilitates membrane stripping of associated proteins and promotes membrane vesicle opening resulting in efficient shaving of extramembrane protein material from both sides of membrane vesicles [18].

However, the low specificity of proteinase K produces numerous overlapping peptides leading to increased complexity of sample and potentially decreases the number of identified proteins. In our modified approach we used trypsin instead of proteinase K and verified its efficacy for the membrane shaving and analysis of integral membrane peptides. In the course of our experiments we also introduced several other modifications. We avoided the relatively laborious purification of membrane vesicles by ultracentrifugation steps, used different CNBr digestion protocol and performed on-column delipidation of solubilized membranes instead of lipid precipitation. In reference to the original hppK (high pH, proteinase K) strategy [18,24], we propose an acronym hpTC (high-pH-Trypsin-CNBr) for this modified method.

3.1. Isolation of membrane-enriched fraction

Instead of ultracentrifugation steps used in the original protocol we homogenized the lymphoma cells simply by passing through a hypodermic needle in a hypotonic buffer and isolated the crude membrane fraction as described previously [25] (Fig. 1A).

To cope with the viscous clump of DNA released from residual nuclear fragments during the subsequent steps, we treated the crude membrane sample with DNAse prior to the digestion. The crucial step – opening of membrane vesicles in 100 mM Na_2CO_3 at high pH – was performed essentially as described in the original protocol [18] (Fig. 1B). For the digestion of all non-membrane proteins and extramembrane sections of membrane proteins we chose trypsin instead of proteinase K employed by Blackler *et al.*

Since the peptides produced by non-specific proteinase K are highly overlapping, the resulting sample would be unnecessarily complex, compared to identical digest resulting from specific cleavage with trypsin. Also, since current state-of-the-art MS instrumentation provides better fragmentation spectra and also higher resolution and mass precision, it is no longer necessary to increase the protein identification confidence by fragmenting several overlapping peptides covering the same protein sequence. This in turn leads to higher individual signal intensities as the amount of each analyte protein is not split into several overlapping peptides concurrently occurring in the sample. Further, C-terminal arginine or lysine in each peptide generated by trypsin cleavage facilitate ionization and confident protein identification and validation. Last but not least, trypsin is compatible with SILAC labeling, enabling convenient semi-quantitative expression analysis of IMPs.

The successful digestion with trypsin in the re-formed membrane vesicles is visible in Fig. 1C as the vesicles are more sharply defined. To release the cleaved extramembrane peptide content from the vesicles we re-opened the vesicles again in 100 mM Na_2CO_3 at high pH and washed and pelleted the trypsin-treated membranes. To remove all potential soluble peptides, the vesicles were repeatedly freeze-thawed in the carbonate buffer.

3.2. Re-digestion of membrane peptides

Membrane-embedded peptides produced by the treatment of membranes with trypsin alone are rather large and hydrophobic for optimal LC-MS analysis. A typical vertebrate transmembrane alpha-helix consists of 20–30 amino acids depending on particular membrane type (ER, Golgi or plasma membrane) [18,26]. Furthermore a variable number of amino acid residues to the nearest trypsin cleavage site(s) or protein terminus must be added to such a theoretical calculation as well. In order to reduce the size of the peptides and lower their hydrophobicity, solubilization of the trypsin-treated membranes in strong acid (preferably FA or TFA) followed by chemical cleavage of hydrophobic peptides with CNBr can be used [27]. CNBr cleaves proteins or long peptides C-terminal to methionine residues and generates peptides with homoserine lactone at their C-termini. Also, as the cleavage is mediated only by a small chemical agent, it is well suited to digesting samples even in the membrane bilayer environment, where enzymatic digestion suffers from the very limited accessibility of large protein-based enzymes to the substrate. High toxicity of CNBr and its byproducts, however, needs to be considered and all

Please cite this article as: O. Vit, et al., Large-scale identification of membrane proteins based on analysis of trypsin-protected transmembrane segments, J Prot (2016), http://dx.doi.org/10.1016/j.jprot.2016.03.016

ARTICLE IN PRESS

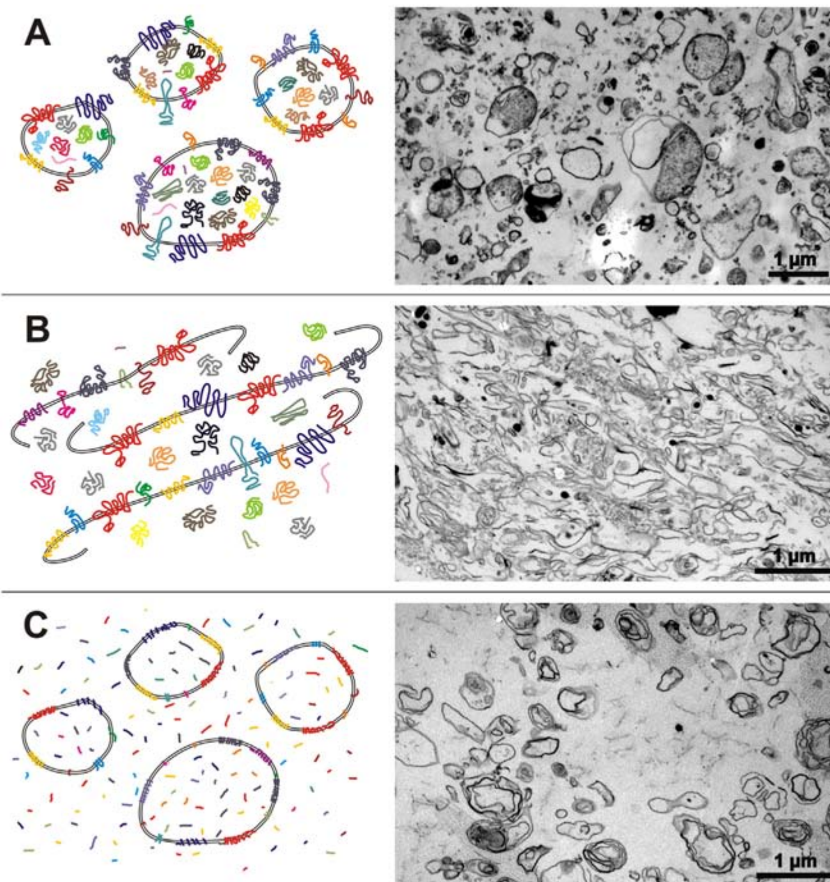


Fig. 1. Isolation, carbonate stripping and proteolytic shaving of membranes steps visualized by transmission electron microscopy (TEM). A) Upon homogenization and centrifugation steps, the isolated crude membrane fraction consists of variously sized membrane vesicles containing trapped proteins and other material. B) Carbonate stripping allows opening of the vesicles and release of trapped content. Due to the use of centrifugation during fixation for TEM, the opened membranes appear to be stacked. C) After tryptic digestion, the membrane outlines are more sharply defined and the vesicles appear empty compared to 1A. Due to previous opening and stacking, some of the resulting vesicles are multilayered.

work involving CNBr must be strictly performed in a fume hood. It should be mentioned that a combined cleavage with CNBr/trypsin has been used in membrane proteomics previously to analyze bacterial membrane, however, in a reverse order – i.e. CNBr cleavage followed by trypsin digestion and in different experimental design [28].

Initially, we performed a pilot experiment with a standard sample – membrane protein CLC-EC1. We performed the digestion according to the hppK-CNBr protocol which used solubilization and CNBr digestion in 90% formic acid. When applied to our model protein, CLC-EC1, substantial portion of the digestion products was found to be modified by formylation (Supplementary Fig. 1A). We therefore replaced the formic acid by TFA and despite the fact that formic acid was used for subsequent sample desalting, virtually no formyl adducts were found (Supplementary Fig. 1B).

3.3. Sample delipidation

After the chemical cleavage, the phospholipid-rich sample must be delipidated prior to the LC-MS/MS analysis. The original protocol employed lipid precipitation in aqueous-organic buffer [18]. Since we observed quite substantial amount of lipids in the supernatant of the precipitated and centrifuged sample, we replaced the precipitation step with on-column delipidation using dichloromethane [21,22]. The method was originally designed for non-ionic detergent removal from peptide digests but when applied to a detergent solubilized mitochondrial membrane protein ANT1, removal of the lipids was observed as well [22]. In a pilot experiment we tested the lipid removal procedure on pure CLC-EC1 digest and observed no peptide loss. In addition, even very hydrophobic peptides were detected in the final acetonitrile eluate from the trap column. Next we verified that CNBr treated lipid (DMPC) can be efficiently removed. As expected, the lipid eluted from the trap when dichloromethane was applied and no signal for the intact lipid or the acid-generated degradation product, lysoDMPC was found

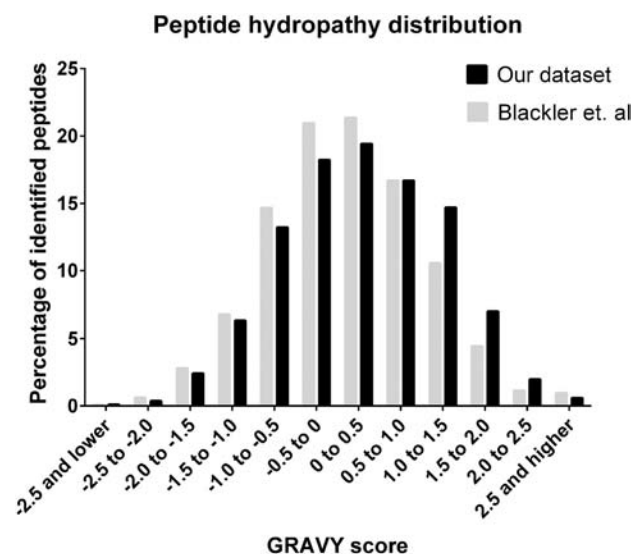


Fig. 2. Peptide hydrophobicity. Comparison of the distribution of peptide GRAVY score in our hpTC dataset with the original hppK-CNBr method. Distribution of peptides in our dataset is shifted toward higher GRAVY score (more hydrophobic peptides), presumably due to different method of sample delipidation.

Please cite this article as: O. Vit, et al., Large-scale identification of membrane proteins based on analysis of trypsin-protected transmembrane segments, J Prot (2016), http://dx.doi.org/10.1016/j.jprot.2016.03.016

ARTICLE IN PRESS

in the acetonitrile elution. Finally, we used DMPC-spiked CLC-EC1 for digestion and performed the purification procedure on this sample. The result is shown in Supplementary Fig. 2 and confirms that the on-column delipidation approach represents a feasible sample preparation for membrane proteomics samples. This is further supported by the hydrophobicity calculation (GRAVY score) [29] of the peptides identified in our dataset and those from the original dataset of Blackler et al. [18] obtained after lipid precipitation. The distribution can be seen in Fig. 2. Our dataset contained more peptides with high hydrophobicity (peptides with positive GRAVY score) probably due to their preferential capture on the hydrophobic stationary phase.

3.4. LC-MS/MS analysis

In the original paper, authors analyzed the delipidated membrane sample using MudPit approach [30] consisting of twelve 2-hour steps [18]. Our 1D LC-MS/MS analysis using a 90 min gradient on a 50 cm heated column connected to Orbitrap Q Exactive mass spectrometer resulted in identification of 1224 proteins with average sequence coverage 12.2%. Of the 1224 proteins 802 molecules (65.5%) were predicted to have at least one transmembrane domain by TMHMM [19]. For the complete list of the identified proteins see Supplementary Table 1.

3.5. Integral membrane proteins, transmembrane segments

Using the hpTC method we identified 802 IMPs with up to 16 transmembrane segments (except one outlying protein with 31 predicted transmembrane segments, Piezo1, Q92508) in the human lymphoma Mino cell sample, of which proteins with two or more TM segments represented roughly two thirds. (Fig. 3) The observed distribution is in agreement with genome-wide predictions for human proteins [1] suggesting that the hpTC method is not biased toward a specific group of transmembrane proteins. Furthermore a closer analysis showed that approximately half (1274/2561) of the unique peptides used for the identification of the IMPs overlapped with predicted transmembrane regions.

To illustrate the contribution of transmembrane peptides in the hPTC analysis several topological maps generated by Protter software [31] are shown in Fig. 4. Also, to determine the relative contribution of trypsin digestion and CNBr cleavage among the identified peptides, we calculated the distribution of terminal arginines and lysines and methionines converted by CNBr to homoserine lactones (Supplementary Fig. 3). Roughly 75% of all the peptides identified in our analysis were produced by CNBr or combined action of trypsin and CNBr. However, this proportion is significantly higher (97%) for the peptides which overlapped with predicted transmembrane segments. This underlines the benefit of CNBr use in the analysis of membrane proteins.

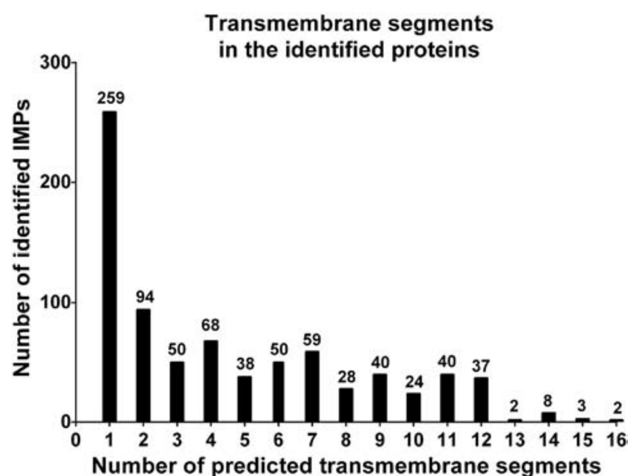


Fig. 3. Number of transmembrane segments in the identified IMPs. Number of predicted transmembrane regions (based on TMHMM prediction) in the 802 IMPs identified in our study.

3.6. Access to all cellular membrane compartments

Among the IMPs identified in the human lymphoma cells, there were numerous transporters, membrane enzymes, receptors, signal transduction proteins, and proteins with immunity-related activities and other functions. According to gene ontology annotations, the identified membrane proteins originate from various organelles, confirming the ability of the method to access not only plasmatic but all cellular membranes. Due to dynamic nature of cellular membranes (and according to their multiple G.O. annotations), many membrane proteins are associated with two or more cellular compartments. This complicates the evaluation and prevents any clear-cut calculation of protein distribution among the various membrane compartments. Nevertheless, we also observed proteins annotated exclusively to mitochondrial membrane (e.g. ADP/ATP translocases 1–3, numerous components of respiratory chain), to membranes of endoplasmic reticulum (e.g. GlcNAc-1-P transferase, UDP-glucuronic acid/UDP-*N*-acetylgalactosamine transporter) and Golgi (e.g. Alpha-(1,6)-fucosyltransferase, Beta-1,4-galactosyltransferase 3), proteins localized to nuclear envelope (e.g. lamin-B receptor, nucleoporin NDC1) and proteins known for their typical plasma membrane localization such as proteins belonging among the CD (cluster of differentiation) molecules. CD proteins expressed on the cell surface represent promising targets of modern anti-cancer drugs including therapeutic antibodies, as exemplified by rituximab targeting CD20 in B-cell lymphomas, trastuzumab targeting-HER2 in breast cancer patients or ipilimumab targeting CTLA-4 in patients with melanoma [32]. The information on expression of CD markers in cancer cells is therefore of tremendous importance. We identified 48 CD molecules in our analysis of human lymphoma cells (Table 1.)

3.7. Opening the black box

IMPs with large extra-membrane domains can provide several tryptic peptides and can be therefore identified in the conventional large scale proteomic analyses or using the surface-capture strategy. However, some membrane proteins possess only small extra-membrane segments offering no tryptic peptides of reasonable size. Such proteins thus inevitably escape detection by both surface capture methods and routine proteomic analyses. Here we show, that the CNBr cleavage of intra-membrane segments of such proteins may enable their identification as demonstrated by several proteins identified in our hpTC analysis. Such a typical difficulty to identify proteins with no extramembrane tryptic peptide can be exemplified by the identified Dolichyl-diphosphooligosaccharide-protein glycosyltransferase subunit 4 (OST4_HUMAN) or NADH-ubiquinone oxidoreductase chain 4L (NU4LM_HUMAN) in Fig. 4.

Recently, the Human Proteome Project sparked an interest in identification of “missing proteins” i.e. proteins having no evidence of existence on protein level [33]. The latest release of the neXtProt database (www.nextprot.org, September 2015) contains 3243 “missing proteins” [34]. In our dataset we identified 13 of such “previously unseen” proteins, namely O15342, O60478, Q14330, Q14656, Q5SWH9, Q6UWH6 (evidence on transcript level), A8MWL7, A2A368, C9J798, P69849 (based on homology) and O60361, Q5T1J5, Q92928 (classified as uncertain). Seven out of the thirteen proteins (O15342, O60478, Q14330, Q14656, Q5SWH9, Q6UWH6, A8MWL7) are IMPs with 2–7 transmembrane segments.

3.8. Proteins with no transmembrane domain. Contaminants?

The treatment of membranes with carbonate washing at high pH followed by trypsin digestion was in our study effective for removal of most but not all non-membrane contaminants. Among the 1224

Please cite this article as: O. Vit, et al., Large-scale identification of membrane proteins based on analysis of trypsin-protected transmembrane segments, J Prot (2016), http://dx.doi.org/10.1016/j.jprot.2016.03.016

ARTICLE IN PRESS

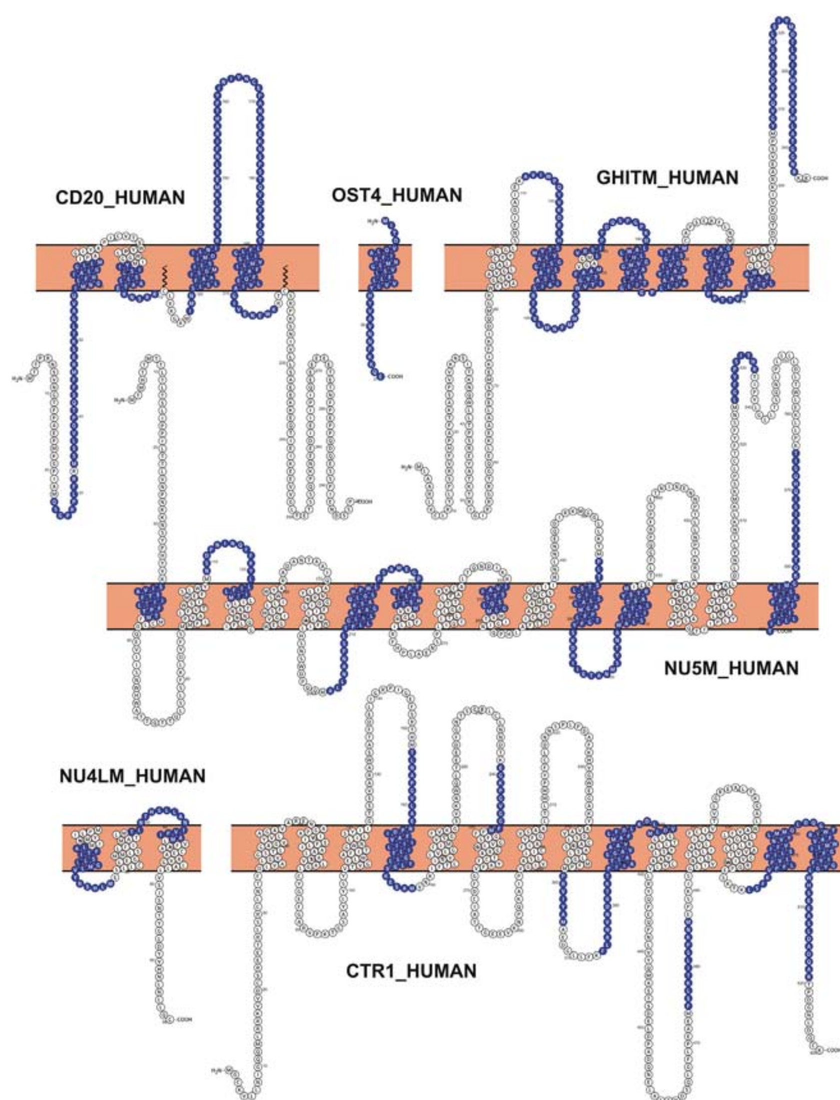


Fig. 4. Examples of sequence coverage of identified IMPs on topological prediction maps. Sequences of identified peptides used for the identification are labeled blue. Topology maps were generated using Protter (http://wlab.ethz.ch/protter). These examples illustrate that peptides overlapping with transmembrane segments significantly contribute to the protein identification. As demonstrated by OST4_HUMAN and NU4LM_HUMAN this method allows proteomic identification of small, highly hydrophobic IMPs with limited presence of trypsin cleavage sites.

identified proteins, 422 were predicted to have no transmembrane segment. Presence of these proteins in the hpTC isolate may be explained by their tight specific association with membrane lipids or proteins. For instance, we identified non-membrane protein beta-2-microglobulin, which is known to exist as a component of a large membrane-bound MHC class I complex and may be therefore shielded by the $\alpha1/\alpha2$ MHC heterodimer (also identified in our study) from the trypsin activity. Similarly, the presence of hydrophobic segments in the structure of non-membrane proteins may cause their non-specific sticking to the phospholipid bilayer during the sample preparation and may thus prevent their complete removal. In our lymphoma cell study we observed contamination by major cytosolic (alpha enolase, GAPDH, glucose-6-phosphate isomerase), cytoskeletal (actin, tubulin) and also nuclear (histones, small nuclear ribonucleoproteins) high-abundance proteins. However, as exemplified by histones, which may also be extra-nuclear and have been shown to directly interact with membrane [35,36], it is not always easy to decide where precisely to draw the line between proteins of interest and contaminants. Overall, despite the presence of some contaminating proteins in the results of our analyses, integral membrane proteins represented over 65% of all identified proteins, thus demonstrating high efficacy of the hpTC method.

3.9. Quantitative analysis of membrane proteome?

Analysis of membrane-embedded peptides clearly taps the invaluable information on membrane proteome. However, mere cataloguing of membrane proteome is not sufficient. Can we make such an analysis (semi)quantitative?

Implementation of trypsin instead of proteinase K to the method workflow potentially opens the door to quantitative membrane analysis using SILAC labeling [37]. To evaluate the theoretical potential of SILAC labeling for the hpTC method, we queried the content of canonic SILAC-compatible amino acids – arginine (R) and lysine (K) in the unique peptides identified in our lymphoma cell line analysis (Fig. 5). Out of the total number of 3884 peptides, 2055 (53%) contained at least one lysine, while arginine was present in 1965 (51%). Almost 83% of the peptides contained either K or R. Theoretically, an application of "double" (K and R) SILAC could provide semi-quantitative information on a significant

Please cite this article as: O. Vit, et al., Large-scale identification of membrane proteins based on analysis of trypsin-protected transmembrane segments, J Prot (2016), http://dx.doi.org/10.1016/j.jprot.2016.03.016

ARTICLE IN PRESS

Table 1
The list of identified CD molecules.

List of identified lymphoma membrane CD proteins			
CD No.	Accession	Entry name	Protein name
CD10	P08473	NEP_HUMAN	Neprilysin
CD11a	P20701	ITAL_HUMAN	Integrin alpha-L
CD18	P05107	ITB2_HUMAN	Integrin beta-2
CD19	P15391	CD19_HUMAN	B-lymphocyte antigen CD19
CD20	P11836	CD20_HUMAN	B-lymphocyte antigen CD20
CD21	P41597	CCR2_HUMAN	C-C chemokine receptor type 2
CD27	P26842	CD27_HUMAN	CD27 antigen
CD32	P31994	FCG2B_HUMAN	Low affinity immunoglobulin gamma Fc region receptor II-b
CD39	P49961	ENTP1_HUMAN	Ectonucleoside triphosphate diphosphohydrolase 1
CD40	P25942	TNR5_HUMAN	Tumor necrosis factor receptor superfamily member 5
CD43	P16150	LEUK_HUMAN	Leukosialin
CD45	P08575	PTPRC_HUMAN	Receptor-type tyrosine-protein phosphatase C
CD47	Q08722	CD47_HUMAN	Leukocyte surface antigen CD47
CD48	P09326	CD48_HUMAN	CD48 antigen
CD50	P32942	ICAM3_HUMAN	Intercellular adhesion molecule 3
CD53	P19397	CD53_HUMAN	Leukocyte surface antigen CD53
CD54	P05362	ICAM1_HUMAN	Intercellular adhesion molecule 1
CD63	P08962	CD63_HUMAN	CD63 antigen
CD70	P32970	CD70_HUMAN	CD70 antigen
CD71	P02786	TFR1_HUMAN	Transferrin receptor protein 1
CD72	P21854	CD72_HUMAN	B-cell differentiation antigen CD72
CD74	P04233	HG2A_HUMAN	HLA class II histocompatibility antigen gamma chain
CD79a	P11912	CD79A_HUMAN	B-cell antigen receptor complex-associated protein alpha chain
CD79b	P40259	CD79B_HUMAN	B-cell antigen receptor complex-associated protein beta chain
CD81	P60033	CD81_HUMAN	CD81 antigen
CD82	P27701	CD82_HUMAN	CD82 antigen
CD84	Q9UIB8	SLAF5_HUMAN	SLAM family member 5
CD92	Q8WWI5	CTL1_HUMAN	Choline transporter-like protein 1
CD97	P48960	CD97_HUMAN	CD97 antigen
CD98	P08195	4F2_HUMAN	4F2 cell-surface antigen heavy chain
CD99	P14209	CD99_HUMAN	CD99 antigen
CD102	P13598	ICAM2_HUMAN	Intercellular adhesion molecule 2
CD107a	P11279	LAMP1_HUMAN	Lysosome-associated membrane glycoprotein 1
CD132	P31785	IL2RG_HUMAN	Cytokine receptor common subunit gamma
CD147	P35613	BASI_HUMAN	Basigin
CD159a	P26715	NKG2A_HUMAN	NKG2-A/NKG2-B type II integral membrane protein
CD184	P61073	CXCR4_HUMAN	C-X-C chemokine receptor type 4
CD185	P32302	CXCR5_HUMAN	C-X-C chemokine receptor type 5
CD192	P41597	CCR2_HUMAN	C-C chemokine receptor type 2
CD197	P32248	CCR7_HUMAN	C-C chemokine receptor type 7
CD205	O60449	LY75_HUMAN	Lymphocyte antigen 75
CD217	Q96F46	I17RA_HUMAN	Interleukin-17 receptor A
CD222	P11717	MPRI_HUMAN	Cation-independent mannose-6-phosphate receptor
CD225	P13164	IFM1_HUMAN	Interferon-induced transmembrane protein 1
CD230	P04156	PRIO_HUMAN	Major prion protein
CD289	Q9NR96	TLR9_HUMAN	Toll-like receptor 9
CD298	P54709	AT1B3_HUMAN	Sodium/potassium-transporting ATPase subunit beta-3
CD361	P34910	EVI2B_HUMAN	Protein EVI2B

portion of proteins identified in our current hpTC analysis. However, we further evaluated the content of leucine, another essential amino acid compatible with SILAC labeling, which also occurs frequently in hydrophobic transmembrane segments. Leucine was found in 3213 (83%) of all the unique peptides. Taken together, vast majority of all identified peptides (97%) contained at least one lysine, arginine or leucine, suggesting that using a triple (K, R and L) SILAC labeling of the cells could provide semi-quantitative information on most proteins identified in our current hpTC analysis (Fig. 5).

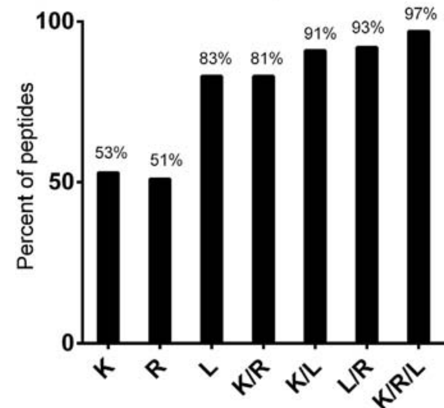


Fig. 5. Compatibility with SILAC. Percentage of unique peptides identified in our analysis containing SILAC-compatible amino acids.

4. Conclusions

In this study we demonstrate that the analysis of membrane-embedded peptides using the hpTC method enables very high enrichment of membrane proteins from all cellular compartments, including proteins with numerous transmembrane segments.

Identification of hundreds of integral membrane proteins in mammalian cells can be easily achieved in a single analysis and the whole experiment, including LC-MS/MS analysis, can be accomplished in approximately 50 h. The modified method avoids ultracentrifugation steps, employs more effective delipidation step and does not require any specialized instrumentation. Another important novelty of our modification to the original hppK method is the introduction of trypsin instead of non-specific proteinase K. This makes the method more sensitive and better suited for quantitative approaches, namely SILAC labeling and potentially opens the way toward a detailed semi-quantitative membrane proteome analysis using SILAC or super-SILAC strategy [37, 38]. However, the amphipathic nature of integral membrane proteins clearly calls for a combination of two complementary approaches. Joining forces of hydrophobic peptide-oriented hpTC approach with a soluble-peptide oriented method should provide the best results and enable deep expression analysis of membrane proteomes.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jprot.2016.03.016.

Transparency document

The Transparency document associated with this article can be found, in online version.

Acknowledgements

The work was supported by the Ministry of Health of the Czech Republic (www.mzcr.cz) via grants 15-32961A, NT12248-5, by the Czech Science Foundation (www.gacr.cz) via grant 15-14200S and by the Ministry of Education, Youth and Sports of the Czech Republic (www.msmt.cz) via the institutional grants PRVOUK P24/LF1/3 and SVV-2013-260033. Support from the Grant Agency of Charles University (389115) and European Regional Development Funds (CZ.1.05/1.1.00/02.0109 - BIOCEV) is gratefully acknowledged. Authors would like to thank Miroslav Hyliš for his guidance and assistance with electron microscopy.

Please cite this article as: O. Vit, et al., Large-scale identification of membrane proteins based on analysis of trypsin-protected transmembrane segments, J Prot (2016), http://dx.doi.org/10.1016/j.jprot.2016.03.016

ARTICLE IN PRESS

References

[1] L. Fagerberg, K. Jonasson, G. von Heijne, M. Uhlén, L. Berglund, Prediction of the human membrane proteome, Proteomics 10 (2010) 1141–1149.
[2] M.A. Yildirim, K.I. Goh, M.E. Cusick, A.L. Barabási, M. Vidal, Drug-target network, Nat. Biotechnol. 25 (2007) 1119–1126.
[3] D. Vuckovic, L.F. Dagley, A.W. Purcell, A. Emili, Membrane proteomics by high performance liquid chromatography-tandem mass spectrometry: analytical approaches and challenges, Proteomics 13 (2013) 404–423.
[4] J.R. Wiśniewski, A. Zougman, M. Mann, Combination of FASP and StageTip-based fractionation allows in-depth analysis of the hippocampal membrane proteome, J. Proteome Res. 8 (2009) 5674–5678.
[5] R.A. Mathias, Y.S. Chen, E.A. Kapp, D.W. Greening, S. Mathivanan, R.J. Simpson, Triton X-114 phase separation in the isolation and purification of mouse liver microsomal membrane proteins, Methods 54 (2011) 396–406.
[6] L. Cao, J.G. Clifton, W. Reutter, D. Josic, Mass spectrometry-based analysis of rat liver and hepatocellular carcinoma Morris hepatoma 7777 plasma membrane proteome, Anal. Chem. 85 (2013) 8112–8120.
[7] R.R. Loo, N. Dales, P.C. Andrews, The effect of detergents on proteins analyzed by electrospray ionization, Methods Mol. Biol. 61 (1996) 141–160.
[8] A.E. Speers, C.C. Wu, Proteomics of integral membrane proteins—theory and application, Chem. Rev. 107 (2007) 3687–3714.
[9] Y. Zhao, W. Zhang, Y. Kho, Y. Zhao, Proteomic analysis of integral plasma membrane proteins, Anal. Chem. 76 (2004) 1817–1823.
[10] B. Wollscheid, D. Bausch-Fluck, C. Henderson, R. O'Brien, M. Bibel, R. Schiess, R. Aebersold, J.D. Watts, Mass-spectrometric identification and relative quantification of N-linked cell surface glycoproteins, Nat. Biotechnol. 27 (2009) 378–386.
[11] P. Mirkowska, A. Hofmann, L. Sedek, L. Slamova, E. Mejstrikova, T. Szczepanski, M. Schmitz, G. Cario, M. Stanulla, M. Schrappe, V.H. van der Velden, B.C. Bornhauser, B. Wollscheid, J.P. Bourquin, Leukemia surfaceome analysis reveals new disease-associated features, Blood 121 (2013) e149–e159.
[12] A. Hofmann, B. Gerrits, A. Schmidt, T. Bock, D. Bausch-Fluck, R. Aebersold, B. Wollscheid, Proteomic cell surface phenotyping of differentiating acute myeloid leukemia cells, Blood 116 (2010) e26–e34.
[13] H. Moest, A.P. Frei, I. Bhattacharya, M. Geiger, B. Wollscheid, C. Wolfrum, Malfunctioning of adipocytes in obesity is linked to quantitative surfaceome changes, Biochim. Biophys. Acta 2013 (1831) 1208–1216.
[14] T. Bock, H. Moest, U. Omasits, S. Dolski, E. Lundberg, A. Frei, A. Hofmann, D. Bausch-Fluck, A. Jacobs, N. Krayenbuehl, M. Uhlen, R. Aebersold, K. Frei, B. Wollscheid, Proteomic analysis reveals drug accessible cell surface N-glycoproteins of primary and established glioblastoma cell lines, J. Proteome Res. 11 (2012) 4885–4893.
[15] R.L. Gundry, D.R. Riordon, Y. Tarasova, S. Chuppa, S. Bhattacharya, O. Juhasz, O. Wiedemeier, S. Milanovich, F.K. Noto, I. Tchernyshyov, K. Raginski, Tae H.J. Bausch-FluckD, S. Marshall, S.A. Duncan, B. Wollscheid, R.P. Wersto, S. Rao, J.E. Van Eyk, K.R. Boheler, A cell surfaceome map for immunophenotyping and sorting pluripotent stem cells, Mol. Cell. Proteomics 11 (2012) 303–316.
[16] B. DeVeale, D. Bausch-Fluck, R. Seaberg, S. Runciman, V. Akbarian, P. Karpowicz, C. Yoon, H. Song, R. Leeder, P.W. Zandstra, B. Wollscheid, D. van der Kooy, Surfaceome profiling reveals regulators of neural stem cell function, Stem Cells 32 (2014) 258–268.
[17] S.J. Deeb, J. Cox, M. Schmidt-Supprian, M. Mann, N-linked glycosylation enrichment for in-depth cell surface proteomics of diffuse large B-cell lymphoma subtypes, Mol. Cell. Proteomics 13 (2014) 240–251.
[18] A.R. Blackler, A.E. Speers, M.S. Ladinsky, C.C. Wu, A shotgun proteomic method for the identification of membrane-embedded proteins and peptides, J. Proteome Res. 7 (2008) 3028–3034.
[19] A. Krogh, B. Larsson, G. von Heijne, E.L. Sonnhammer, Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes, J. Mol. Biol. 305 (2001) 567–580.
[20] M. Maduke, D.J. Pheasant, C. Miller, High-level expression, functional reconstitution, and quaternary structure of a prokaryotic ClC-type chloride channel, J. Gen. Physiol. 114 (1999) 713–722.
[21] M. Rey, H. Mrázek, P. Pompach, P. Novák, L. Pelosi, G. Brandolin, E. Forest, V. Havlícek, P. Man, Effective removal of nonionic detergents in protein mass spectrometry, hydrogen/deuterium exchange, and proteomics, Anal. Chem. 82 (2010) 5107–5116.
[22] M. Rey, P. Man, B. Clémençon, V. Trézéguet, G. Brandolin, E. Forest, L. Pelosi, Conformational dynamics of the bovine mitochondrial ADP/ATP carrier isoform 1 revealed by hydrogen/deuterium exchange coupled to mass spectrometry, J. Biol. Chem. 285 (2010) 34981–34990.
[23] L. Käll, J.D. Canterbury, J. Weston, W.S. Noble, M.J. MacCoss, Semi-supervised learning for peptide identification from shotgun proteomics datasets, Nat. Methods 4 (2007) 923–925.
[24] C.C. Wu, M.J. MacCoss, K.E. Howell, J.R. Yates III, A method for the comprehensive proteomic analysis of membrane proteins, Nat. Biotechnol. 21 (2003) 532–538.
[25] Y.J. Jin, M.W. Albers, W.S. Lane, B.E. Bierer, S.L. Schreiber, S.J. Burakoff, Molecular cloning of a membrane-associated human FK506- and rapamycin-binding protein, FKBP-13, Proc. Natl. Acad. Sci. U. S. A. 88 (1991) 6677–6681.
[26] H.J. Sharpe, T.J. Stevens, S. Munro, A comprehensive comparison of transmembrane domains reveals organelle-specific properties, Cell 142 (2010) 158–169.
[27] W.A. Schroeder, J.B. Shelton, J.R. Shelton, An examination of conditions for the cleavage of polypeptide chains with cyanogen bromide: application to catalase, Arch. Biochem. Biophys. 130 (1969) 551–556.
[28] F. Fischer, D. Wolters, M. Rögner, A. Poetsch, Toward the complete membrane proteome: high coverage of integral membrane proteins through transmembrane peptide detection, Mol. Cell. Proteomics 5 (2006) 444–453.
[29] J. Kyte, R.F. Doolittle, A simple method for displaying the hydropathic character of a protein, J. Mol. Biol. 157 (1982) 105–132.
[30] M.P. Washburn, D. Wolters, J.R. Yates III, Large-scale analysis of the yeast proteome by multidimensional protein identification technology, Nat. Biotechnol. 19 (2001) 242–247.
[31] U. Omasits, C.H. Ahrens, S. Müller, B. Wollscheid, Protter: interactive protein feature visualization and integration with experimental proteomic data, Bioinformatics 30 (2014) 884–886.
[32] Y. Zhu, S.H. Choi, K. Shah, Multifunctional receptor-targeting antibodies for cancer therapy, Lancet Oncol. 16 (2015) e543–e554.
[33] L. Lane, A. Bairoch, R.C. Beavis, E.W. Deutsch, P. Gaudet, E. Lundberg, G.S. Omenn, Metrics for the Human Proteome Project 2013-2014 and strategies for finding missing proteins, J. Proteome Res. 13 (2014) 15–20.
[34] P. Gaudet, P.A. Michel, M. Zahn-Zabal, I. Cusin, P.D. Duek, O. Evalet, A. Gateau, A. Gleizes, M. Pereira, D. Teixeira, Y. Zhang, L. Lane, A. Bairoch, The neXtProt knowledgebase on human proteins: current status, Nucleic Acids Res. 43 (2015) D764–D770.
[35] E. Hariton-Gazal, J. Rosenbluh, A. Graessmann, C. Gilon, A. Loyter, Direct translocation of histone molecules across cell membranes, J. Cell Sci. 116 (2003) 4577–4586.
[36] K. Watson, N.J. Gooderham, D.S. Davies, R.J. Edwards, Nucleosomes bind to cell surface proteoglycans, J. Biol. Chem. 274 (1999) 21707–21713.
[37] S.E. Ong, B. Blagoev, I. Kratchmarova, D.B. Kristensen, H. Steen, A. Pandey, M. Mann, Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics, Mol. Cell. Proteomics 1 (2002) 376–386.
[38] T. Geiger, J. Cox, P. Ostasiewicz, J.R. Wisniewski, M. Mann, Super-SILAC mix for quantitative proteomics of human tumor tissue, Nat. Methods 7 (2010) 383–385.

Please cite this article as: O. Vit, et al., Large-scale identification of membrane proteins based on analysis of trypsin-protected transmembrane segments, J Prot (2016), http://dx.doi.org/10.1016/j.jprot.2016.03.016