Chemical Tools for Isolation of Endogenous Claudins from Carcinoma Cells

A thesis submitted in partial fulfilment of the requirements for the degree of

Master of Research

by

Payal Barua

MS (Biochemistry and Molecular Biology)

Department of Molecular Science

Macquarie University

Sydney, Australia

October, 2017

List of Contents

Abstract

Claudins are a family of tetraspan transmembrane proteins that regulate the barrier function and ion selective paracellular permeability of the tight junction. Claudins have diverse roles in normal tissues as well as different cancers, and it is necessary to understand the regulation of endogenous claudins in different cellular environments. As low-abundance and hydrophobic proteins, detection and analysis of claudins has been a major challenge. Here a chemical method for tagging and enriching endogenous surface claudins is investigated for analysis by mass spectrometry. Model human colorectal carcinoma, HCT-116 cells, were surface labelled with sulfo-NHS-S-S-biotin, followed by NeutrAvidin affinity purification and then LC/MS/MS analysis. Various cell-lysis, labelling and elution conditions for optimized recovery of claudins were examined by western-blot analysis. Following the optimized labeling and affinity pull-down protocol, the biotinylated elute fraction was analysed in 1D liquid LC/MS/MS analysis. A total of 389 unique protein groups were identified including 58 proteins found in ≥2 samples. Although peptides of claudins still remain to be detected, 19 plasma membrane proteins were identified with the optimized protocol. This sets the stage for a future advanced workflow that combines membrane fractionation with surface biotinylation for MS based proteome-wide analysis of claudins.

Declaration

I certify that the work in this title entitled "Chemical Tools for Isolation of Endogenous Claudins from Carcinoma Cells" has not been previously submitted as part of requirements for a degree to any other university or institution other than Macquarie University. I also certify that the thesis is an original piece of research and it has been written by me. Any help and assistance that I have received in my research work and the preparation of the thesis itself have been appropriately acknowledged. Finally, I certify that all information sources and literature used are indicated in the thesis.

The research presented in this thesis was approved by Macquarie University biosafety review committee, reference number 5201700491, on 22th May, 2017.

Acknowledgement

I would like to acknowledge the following people, without whom this thesis would have not been possible:

My supervisor, Dr. Fei Liu, for giving me the opportunity to explore a fascinating area of research. Thanks for your endless support and guidance all through the journey.

Collaborator Dr Charlie Ahn and Abidali Mohammedali for all the training and support provided for experimental procedure and evaluation.

Subash Adhikari and Matthew Fitzhenry for helping with LC/MS/MS analysis.

Dr Andrew Piggot and APAF staff for instrumental support.

Macquarie University for a research scholarship.

My loving husband and friend, Kawsar for his unconditional love and support.

List of abbreviations

°C	Degree celsius
TJ	Tight junction
ZO	Zonula occludens
Ocln	Occludin
JAM	Junctional adhesion molecule
CAR	Coxsackie and adenovirus receptor
ESAM	Endothelial cell-selective adhesion molecule
CLD	Claudin
MDCK	Madin-Darby canine kidney
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
EMPs	Epithelial membrane proteins
ECL	Extracellular loop
ICL	Intracellular loop
CPE	*Clostridium perfringens* enterotoxin
PTM	Post translational modifications
MS	Mass spectrometry
GST	Glutathione S-transferase
SILAC	Stable isotope labeling with amino acids in cell culture
Gαi2	G-protein subunit αi2
LC	Liquid chromatography
LC-MS	Liquid chromatography mass spectroscopy
aa	Amino acid
Kda	Kilo dalton
CHAPS	3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate hydrate
hESCs	Human embryonic stem cells
BBB	Blood-brain barrier
DMEM	Dulbecco's modified eagle medium
FBS	Fetal bovine serum
PBS	Phosphate buffer saline
RT	Room temperature
TBS	Tris-buffered saline
BSA	Bovine serum albumin
HEPES	4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid, N-(2-Hydroxyethyl)piperazine-N′-(2-ethanesulfonic acid)
EDTA	Ethylenediaminetetraacetic acid
SDS	Sodium dodecyl sulfate
PAGE	Polyacrylamide gel electrophoresis
LDS	Lithium dodecyl sulfate
EGFR	Epidermal growth factor receptor
HRP	Horseradish peroxidase
MS/MS	Tandem mass spectrometry

ACN	Acetonitrile
FA	Formic acid
CID	Collision induced dissociation
GO	Gene ontology
DNA	Deoxyribonucleic acid
FT	Flow through
DTT	Dithiothreitol
GO.CC	Gene ontology cellular component
HSP	Heat shock protein

List of Figures

List of Tables

Chapter 3

1. Introduction

The transmembrane protein claudin family is the key group protein of the tight junction (TJ) in epithelia and endothelia. This family of proteins regulates the barrier functions of TJ and ion selective paracellular permeability by forming pores. Their tissue specific expression and function specific to subcellular localization facilitate their involvement with complex signaling processes inside the cell. After the discovery in 1998 by the Tsukita group, claudins are being investigated for their physiological and pathological roles in cells and tissues [1]. Relentless efforts are given to comprehend their role in maintaining cell polarity, integrity as well as in neoplastic growth and progression. However these low abundant and hydrophobic proteins are very resistant to capture and characterization. Existing approaches and facilities provide only a discreet view of the role and organization of claudins. For a systematic view of claudins, integration of different approaches and new methodologies are needed.

1.1. Tight junction and claudins

Tight junctions (TJ) are subcellular compartments made by strands of proteins at the lateral portion of the epithelial sheets. Tight junctions mostly regulate the paracellular permeability forming close contacts between the adjacent epithelial cells. TJ proteins are frequently reported to control the signals in the epithelial milieu, cell proliferation as well as cell transformation.

In 1963, Farquhar and Palade described TJ as a barrier forming structure between the epithelial cells [2]. Later on, fridge fracture studies described these TJs forming fused or kissing points with the adjacent cells at the lateral epithelial cell membranes (Fig.1. a, b). Molecular organization of the TJ is complex, involving many transmembrane and cytoplasmic plaque protein components [3]. The transmembrane TJ proteins, including the claudins, TAMPs or TJ-associated marvel domain-containing proteins (occludin, tricellulin and marvelD3) and the immunoglobulin-like domain containing adhesion proteins (JAM and CAR), are considered as major barrier proteins. Various cytoplasmic plaque proteins beneath them interact with the transmembrane TJ proteins as well as cytoskeletal proteins. Among the cytoplasmic plaque proteins, scaffolding proteins zonula occludens (ZO-1, ZO-2 and ZO-3), cingulin and cingulin-like protein 1, are the peripheral

1

proteins that structurally support the transmembrane barrier proteins by attaching them with cytoskeletal proteins. Also, other signaling components of cytoplasmic plaque, such as– GTP-binding proteins, protein kinases and phosphatases as well as transcriptional factors are found to associate with TJ [3]. Till date, the number of identified TJ associated proteins are around 40, although more are yet to be reported [4].

Furuse *et al.* first described claudins as key molecules of TJ structure and functions [1]. This group of researchers provided a series of extensive research on physiological role of claudins in TJ. They showed that claudins can actually generate TJ strands in occludin deficient cells and combinations of claudin-1 /-3 and claudin- 2/-3 can also form heterotypic TJ strands [5-6]. Later on, Van Itallie *et al.* described the role of the claudin family of proteins in charge selectivity and pore formation and regulation of overall barrier functions in epithelial tissues [7].
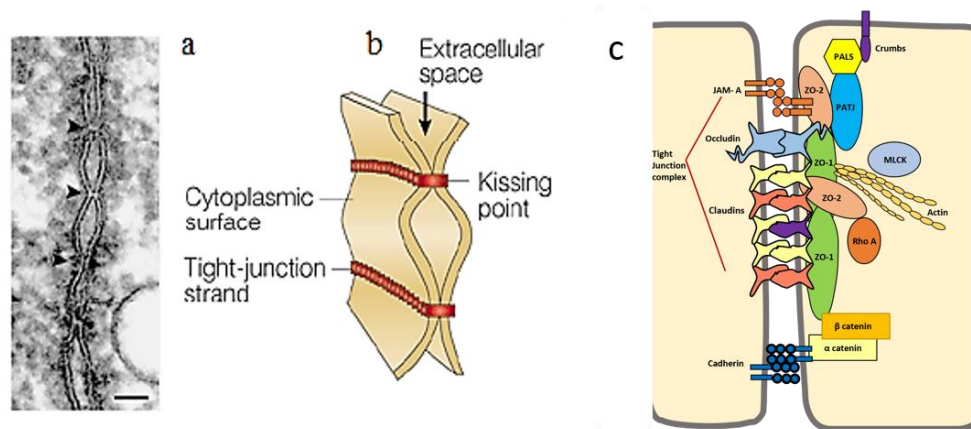
Fig. 1.1 (a) Freeze-fracture image of ultrathin section of TJ between epithelial cells (b) Schematic of 3D structure of TJ (c) Molecular organization of TJ complex

Nevertheless, the type and combination of claudins are a determinant of TJ functionality and diversity in various tissues. Apparently, CLD-1, 3, 4, 5 and 8 are found mainly to seal the paracellular clefts [8-10], on the contrary, CLD- 2, 10, 15, 17 and 21 from ion selective paracellular channels and increase solute permeability through the cleft [11-13]. Also, claudins mingling with themselves, as well as other proteins, have found to fine tune the barrier permeability. In Madin-Darby canine kidney II (MDCK II) cells, Epidermal Growth Factor (EGF) altered the claudin

profiles resulting in an augmentation of barrier function [14-15]. An activation of EGF receptor suppressed CLD-2 expression, but facilitated the distribution and higher expression of CLD-1, -3, and -4 at the same time [15]. Until now, the role and assembly of claudins in cellular signaling is not fully understood as the endogenous claudins are very difficult to isolate.

1.2. Human claudin protein family

Claudin proteins comprised of a multigene family have been extensively analyzed in many organism including human. Till date more than 23 genes for claudins are reported for human where two transcript variants are included [16-17]. These genes are located through 13 chromosomes- 1, 3, 4, 6, 7, 8, 11, 13, 16, 17, 21, 22, and X in human [16]. On the basis of sequence similarity, the members of claudin gene family are classified as classic (claudin- 1-10, 14, 15, 17 and 19) and non-classic (claudin- 11-13, 16, 18, 20-24) claudins. The classic claudins share conserved motif at the C-terminal binding domain, while the non-classics have great variety at C-terminal [18].
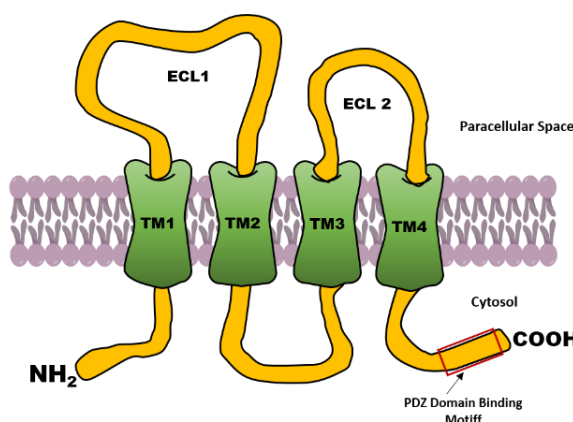
Fig. 1.2 General scheme of claudin protein structure. TM 1-4 -transmembrane domain 1-4, ECL-1, 2- Extracellular loop 1, 2.

The claudin multigene family encodes at least 23 proteins in human. As per the structural as well as functional similarities, claudin protein family is included in the pfam00822 superfamily characterized by tetraspan membrane domains [19]. Being a member of this family, claudins share

sequence similarity among themselves and other members of this superfamily such as Epithelial Membrane Proteins (EMPs), MP-20, CACNGs etc. Human claudins are in the range of 200−300 amino acids (20−35 kDa) and share the common structural features; such as transmembrane regions, extracellular loop (ECL) and intracellular loop (ICL)[18, 20]. The four transmembrane domains are comprised of around 24 aa while the ECL1 and ECL2 have around 50 aa and 22 aa with consensus W-X(17-22)-W-X(2)-C-X(8-10)-C sequence on ECL1 (Fig. 2) [21]. The cytosolic portion of claudins comprised of N-terminus (~ 10 a a), the ICL (~12 a a) and the C-terminus (25–55 a a), is the interaction sites for the other junctional and signaling proteins. Especially, the PDZ domain binding motif YV at the C-terminus (Fig. 2), which is conserved among the classic claudins and also common in non-classic ones, is well known for binding ZO proteins.
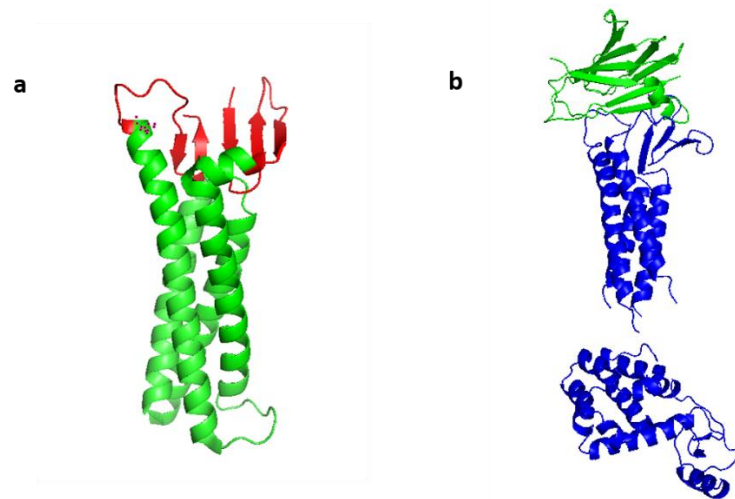
Fig. 1.3 (a) X-ray crystal structure of mouse claudin-15. The four transmembrane domains are in green, the extracellular domains are in red colour, and the C-terminal of the protein is not showed here. (b) X-ray crystal structure of *Clostridium perfringens* enterotoxin (CPE) bound human claudin-4 with C-terminal. The CPE is in green and the claudins-4 is in blue colour.

The first X-ray crystal structure of mouse claudin-15, determined by Suzuki and colleagues gave us the opportunity to get a closer view of this protein (Fig.3.a) [22].The crystal structure showed the first extracellular loop (ECL-1) comprised by three characteristic β-sheets, where the other one (ECL-2) have only one β-sheet anchored on four α helical transmembrane domains. The charged residues distributed along the β-sheet of the two ECL loops contribute to the charge selectivity of this claudin [22]. Recently the crystal structures of mouse and human claudins were studied using

Clostridium perfringens enterotoxin (CPE), which is a well characterized binding partner of claudins. The C-terminal domain of the CPE bound to mouse claudin 19 and human claudin-4 has provided a closer view of the structural regulation and assembly of the claudins in TJ strands [23-24]. CPE binds tightly to the ECL loop of human claudin-4 and changes the native β-sheet conformation of this protein, which enables the disassembly of claudins from TJ strands (Fig. 3. B) [24]. The surface accessible residues on the ECL loops can serve as good molecular targets for studying further assembly and regulation of claudins.

1.3. Claudin expression and localization

Claudins are expressed in numerous tissue types in accordance with their role to form different barriers as well as paracellular pores. For example, claudins-1, 2, 3, 4, 5, 7, 11, 12, and 15 are widely expressed in a large variety of tissues, whereas others (claudins-14, 16, 17, 20, and 22) are rare and restricted in specific tissues. Again among the common claudins, higher expression of claudins-3, 4 and 7 are found in epithelial tissues, while their expression lowers in other tissues such as in brain tissues[25]. The blood brain barrier (BBB) in the brain tissues are uniquely constituted by claudin-5[26]. However, the regulation of this differential and spatial expression of this proteins still remained to understand.

Claudin localization also varies in a function specific manner. Most claudins are found on membrane at the junctional points. Although, some claudins are found at cellular locations other than TJ, playing a diverse role in regulating cell motility, division, and migration [27]. The non-junctional as well as cytoplasmic claudins are generally observed to be found in the vesicles along the basolateral plasma membrane. Evidence of cytoplasmic vesicular localization of claudin 3, 4 and 7 has been found in regular epithelial cells[28]. Also localization of non-junctional claudins were also observed on cell surface. Studies with CPE showed binding with claudin-3 on HEK cell surfaces, but not on TJ [29]. On the contrary, nuclear localization of claudins are rare in normal tissues but often reported in cancer cells [30]. Nevertheless, isolation and extractability of claudins could vary with their diverse location. These differential pools of claudins are of particular importance for understanding overall claudin regulation in normal and diseased cells.

1.4. Claudins in cancer

The claudin family proteins are frequently reported to be involved with cancer cells. Although in some cancers some claudins are thought to be diagnostic and prognostic markers, recent studies suggest a pivotal role of claudins in metastasis and tumor progression [30-31]. As most of the cancers originate from the epithelial layers, it is within reason that claudins' involvement could be pivotal in the progression of cancers. Overexpression of claudins have been reported in a range of cancers, including breast, colorectal, pancreas, cervical, squamous cell, stomach, nasopharyngeal, ovary, and thyroid [32-34]. Overexpression of claudins also resulted in alteration of their cellular localization and function in certain cancers. For example, claudin-7 overexpression in breast cancer cell lines plays a functional role in cancer progression [35]. On the other hand, loss of junctional claudins, resulting into disruption of cell-cell junction and communication, is often observed in neoplasia. Downregulation of some claudins were found to be related with tumor cell motility and metastatic progression. For example, relapse in breast cancer is known to involve decreased claudin-1 expression, [36] and apoptosis was induced upon re-expression of claudin-1 in metastatic MDA-MB 361 cells [37].

The role of claudins in cancer is complex, and very little is understood regarding the mechanisms of claudin regulation and their implication in cancers. The role of claudins in tumorigenesis and invasive progression has been extensively studied using knock down and overexpressed cancer cell lines as well as with clinical tumor samples. Gene expression analysis revealed effects of varying expression levels of certain claudins in stage specific growth and metastatic progression in various cancer cells. Effects of claudins in various stages and subtypes were elaborately studied in different colorectal cancer cells and tissues. For instance, claudin-1 was found to be overexpressed in most of the primary and metastatic colorectal tumors and crucial for tumor growth and progression. Forced expression of claudin-2 in SW480 and HCT116 cell lines showed increased tumorigenesis and progression [38]. Also, aberrant localization of claudins were seen in different colorectal cancer cells, such as SW480, SW620 cells.

Aberrant expression profiles of claudins often influence other cellular signaling cascades of growth and survival[30, 39]. Recent studies with colorectal cancer cells demonstrated a role of

claudins in the epithelial mesenchymal transition process that enables cancer cells to acquire invasive and metastatic properties. It is well documented that, external epigenetic and environmental factors induce mutation in the APC (adenomatous polyposis coli) gene, which in turn upregulate the expression of claudin-1 in colorectal carcinomas[30]. Overexpression claudin-1 suppresses the expression of claudin-7 in some colorectal cancer cells [40]. Loss of certain claudins (e.g.-claudin-7) result in disruption of claudin-claudin and claudin-β integrin association of the epithelial barrier. It is hypothesized that, with the epithelial barrier compromised, the 'outside-in' and 'inside-out' signals in the epithelia as well as intra-cellular signaling in individual cells are dysregulated. In colorectal cancer cells, the altered signaling induces loss of cell polarity and triggers epithelial mesenchymal transition [30]. However, other colorectal cancer cells such as the HCT-116 cells have very low or no expression of claudin-1. Despite of extensive gene and RNA expression investigation, the study of endogenous claudins is scanty. Poor detection and extraction difficulties of claudins posed major obstacles in understanding the mechanistic role of this family of proteins.

1.5. Detection of endogenous claudins

Detection of claudins in various epithelial tissues and cell lines have mostly been possible by using the immunodetection method, although alternative methods such as small molecular tracer and fluorescent labeled peptides have also been reported. Immunofluorescent imaging has also been effective in profiling localization specific expression of claudin. For example, immunofluorescent staining of human colon and jejunum showed claudin-2 expression is restricted in colon epithelia, while claudins-4 had observed in both of the tissues [41]. In a very recent study immunostaining of clinical tissues of colorectal cancer patients showed altered localization of caudins-4 and 7 to cytoplasm from plasma membrane location, than the surrounding healthy tissues [42].

Polyclonal and monoclonal antibodies of some claudins are available and widely used in the investigation of relative expression of individual claudins. However, these antibodies are frequently reported to bind multiple epitopes across different members of claudin family, thus not always reliable to detect specific claudin members [43]. Besides cross-reactivity, it is difficult to use

immunodetection method to estimate the relative stoichiometry of claudins and their post translational modifications (PTM) in tissues. Several efforts has been made to detect phosphorylation by immunoaffinity enrichment. For example- tyrosin phosphorylated claudin-5 have been enriched by a general phosphor-Tyr antibody (Shen, W., 2011). However, that lacks proteome-wide study of major PTMs of claudins, including phosphorylation, palmitoylation and ubiquitination. PTM profiling of endogenous claudins still remains challenging due to their low abundance and lack of effective enrichment method. Hence, for a systematic and precise expression and PTM profiling of these low abundant membrane proteins, a more reliable and rigorous technique is now of high priority.

In recent days, mass spectrometry (MS) based approach have been used for characterization and quantification of endogenous claudins. Till date, human claudins-1, 2, 3, 4, 5, 6, 7, 10, 11, 12, and 18, are identified by means of more than one unique peptides with good log (*E*) scores. While the others, human claudin- 8, 9, 14, 17, 20, 22, 24, and 25 are yet to be detected with certainty [44]. Also, there are no reliable MS data available for the splice variant of human claudin-7, 10, 11, 18, and 19. Mass spectrometry of claudins is limited by sample preparation since these tetraspan membrane proteins are resistant to enrichment by standard isolation techniques [45].

1.6. Enrichment of endogenous claudins

Although low abundance membrane proteins such as claudins are difficult to enrich and characterize, diverse enrichment methods have been investigated to detect these proteins. In a seminal study, affinity enrichment method using a synthetic *Clostridium perfringens* enterotoxin, CPE fused with glutathione S-transferase (GST) enabled enrichment of claudin-1, 3, 4, 5 and 7 from rat cholangiocytes [46]. Several claudins, namely- claudin-3 and 4 are well known as the receptor of CPE [47]. The C-terminal of CPE binds with the second extracellular loop, ECL2 of claudin 3 and 4. In the study mentioned above, they tried to isolate the tight junction complex by using stable isotope labeling with amino acids in cell culture (SILAC) in conjunction with GST-CPE affinity enrichment [46]. Numerous proteins of nuclear and mitochondrial origin, such as DNA-directed RNA polymerase II, Ran-binding protein-2 and two other claudins, claudin-1 and 5 were co-enriched with claudin-3 and 4. However, other than claudin-1 and 5, all the proteins co-enriched and identified by MS, were not tight junction -associated protein and considered as a

false positive result in that study. Clearly, CPE based enrichment was useful for isolating specific claudin members, yet failed to isolate the junctional complex attached with that. Recently published X-ray crystallographic structural of CPE bound human claudin-4 showed that binding of this enterotoxin changes the native conformation of ECL loops of claudin and disrupts the claudin-claudin association of tight junction[23-24]. Possibly CPE also affect claudins' interaction with other junctional proteins and thus remained unsuccessful in co-enriching the junctional complex with claudins. Though CPE is an excellent tool for the enrichment of several claudins, its application is very limited for endogenous tight junction proteins and other claudins.

Monoclonal antibodies designed for immunoprecipitation have also been used to enrich junctional proteins for proteomic studies. Tang *et al.* had enriched the junctional complex with a junction targeted anti-PKC zeta antibody from the T84 epithelial cell line [48]. From the enriched fraction, the purity of the junction isolated was confirmed by western blot and over 900 proteins had been identified by MS. However, among the 912 proteins identified by MS, only 20 were known tight junction proteins. The enriched claudins, namely claudin-1 was detected by western blot in that study, but the peptide identified by MS was not sufficient to be enlisted with certainty. The major cause of a low number of tight junction protein profiled might be the differential expression pattern of claudins and their PTM. Not all of the claudins are expressed in one tissues at a time. Also, the hydrophobic membrane spanning domains and potential PTM sites in claudins may have limited the identification of the peptides by MS. Nevertheless, being the very first attempt in proteomic profiling of tight junction proteins, this study helped to identify the key problems regarding these protein's enrichment and identification by MS. Later on, 12 constitutively associated proteins with claudin-5 were isolated by co-immunoprecipitation using anti claudin-5 antibody from brain endothelial cells [49]. In this study, G-protein subunit αi2 (Gαi2) was discovered as a new interacting partner of claudin-5 by co-immunoprecipitation and identified by liquid chromatography-mass spectrometry (LC/MS). Consequently, the immunoprecipitation technique has allowed claudins' enrichment for individual investigation, but of very limited use for proteome-wide investigations with MS, mostly because of non-specific protein binding.

1.6.1. Subcellular fractionation of claudins

Beside antibody-based enrichment, subcellular fractionation or membrane partition had also proved useful to spatially segregate and subsequently isolating hydrophobic membrane proteins. The low abundance hydrophobic proteins in the unfractionated lysates are broadly miscalculated in MS-based proteomic quantification due to their low copy numbers and poor solubility. On the contrary, the hydrophilic and high abundant proteins are largely over-estimated. Some membrane proteins such as- claudins and kinases are extremely low in copy number but have a very specific subcellular localization. Thus subcellular fractionation increases the probability of detecting the low copy number and extremely hydrophobic transmembrane proteins by MS [50]. This method was almost universally applied to various tissues and could be adjusted and optimized for different proteomic studies.

The basic workflow of subcellular fractionation (Fig-1.4) comprised of two major steps- disruption of the cellular arrangement by homogenization and then fractionation and separation of the intact organelles from the homogenate based on their physical property [50]. For the first step, cells or tissue are collected and mechanically homogenized following isolation of larger organelles such as nucleus by slow centrifugation. Then for the second step clarified post-nuclear supernatant with smaller organelles can be fractionated by various means. The membrane portion can be fractionated by ultracentrifugation or density gradient centrifugation (Fig-1.4 step 6). The ultracentrifugation is very robust and can provide a crude membrane fraction containing all type of membrane proteins [51]. Alternatively, different membrane fraction having diverse lipid-protein ratios can be separated by a discontinuous sucrose gradient centrifugation or in 35% sucrose cushion [52-53].
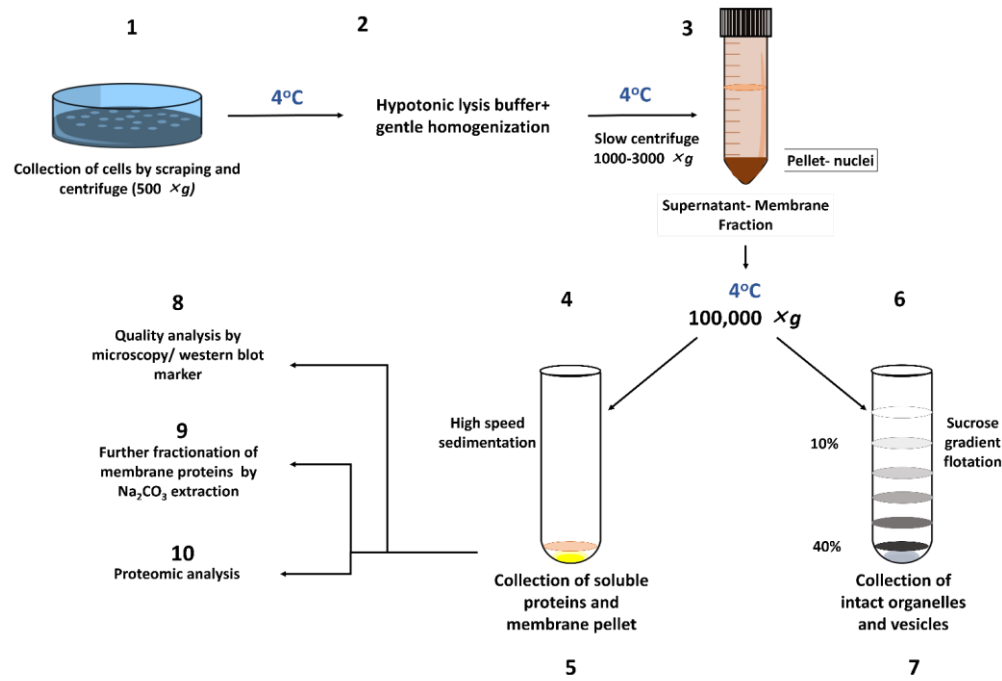
Fig. 1.4 Workflow of subcellular fractionation for membrane protein enrichment.

However, fractionation is difficult to employ for cultured cells because of altered cytoskeletal arrangement and cytoplasmic aggregates that pose a major obstacle in getting homogenized subcellular fractions. Isolation of organelle fractions by ultracentrifugation also have a high risk of contamination across the fractions thus result in weak and unreliable enrichment [54], Nevertheless some limitations such as cytoplasmic aggregation in lysate, can be minimized by using mechanical homogenization and subsequent precipitation of nuclei by slow centrifugation [55]. By using a hypotonic buffer along with gentle mechanical homogenization intact fractions of organelles could be isolated which can substantially reduce contamination of fractions at the subsequent ultracentrifugation step. For example, Tang *et al* had fractionated the junctional complex using a hypotonic buffer and gentle homogenization of T84 cells by a 20G needle [48]. Beside this, use of detergents such as Triton X-100, CHAPS and sodium caprate to solubilize the hydrophobic proteins like claudins, have improved the enrichment [56-57]. For instance, claudin-4 and 5 showed increased solubility in Triton X-100 with sodium caprate [56]. Also, zwitterionic detergent CHAPS was effective for the enrichment of occludin, ZO-1, 2 and claudins by immunoprecipitation [57]. Nonetheless, both mechanical homogenization and detergent use is highly cell type specific and need rigorous methodological optimization [45, 58].

11

In a recent study, membrane fractionation had applied to investigate the membrane proteome in human embryonic stem cells (hESCs) and successfully identified 2292 membrane proteins including CLD- 3, 6 and 7 [59]. They have improved the fractionation of membrane proteins by ultracentrifugation and subsequent sodium carbonate treatment. While, previous attempt of large-scale proteome-wide studies in hESCs had failed to detect claudins without fractionation [60]. Although this physical approach has some limitations, it could be partially useful for claudins with further developments. Alternative membrane protein specific approaches coupled with subcellular fractionation can be a key to get spatial segregation and tissue type specific enrichment of claudins.

1.6.2. Chemical probes for claudin enrichment:

While physical approach can provide partial segregation of membrane proteins from the cytosolic proteins, the chemical approach can go deeper by selectively capturing the relevant proteins. In chemical proteomics, tagging the desired proteins with small molecular chemical probes is frequently used to identify as well as quantify membrane proteins [61].

There are two types of covalent labeling probes available- activity based and affinity based probes. Both of which are widely used in the functional and structural characterization of proteins as well as enzymes [62]. While the activity based probes were proved effective for profiling the enzymes, the affinity based probes have labeled mostly the endogenous proteins irrespective of their enzymatic activities [62-63]. Thus for a structural protein like claudin, affinity based probes preserve the protein structure would be a promising way for isolation and enrichment [64]. These chemical probes are tailored with an affinity ligand which forms a covalent bond with target proteins while carrying a fluorophore for reporting. These type of affinity based probes have been used successfully in membrane proteomics for investigating protein-protein interaction [64]. Thiol or amine-reactive group containing probes have been reported to capture membrane proteins including claudins [65]. For example, in cysteine mutated claudin-2, thiol-reactive reagent was used to map the paracellular pathway in MDCK cells. As most of the claudins has the surface accessible cysteines, thiol-reactive probes had been used widely for investigating claudin- protein

interactions. For instance, cysteine labeling has been successfully used to identify CLD-1 as a partner of tetraspanin CD-9 in human A431 epidermoid carcinoma cells [66].
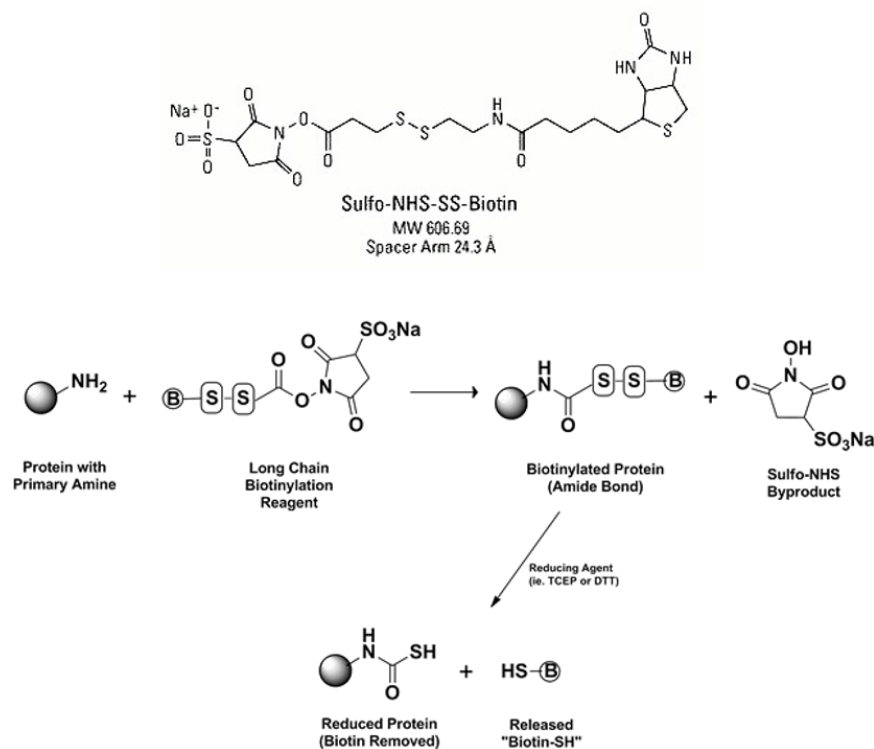
Fig. 1.5 Mechanism of biotinylation of protein by sulfo-NHS-SS-biotin probes

Amine-reactive probes have also been used to label and detect proteins as well as study the tight junction barrier function [8, 67]. Primary amine-reactive labels such as sulfo-NHS-SS-biotin, comprising a thiol-cleavable disulfide bridge are commonly used to label the cell surface exposed lysine residues of the membrane proteins. These biotinylated probes are non-cell permeable and only label the cell surface proteins *in vivo* [68]. To achieve a reasonable enrichment and to avoid cross contamination among the organelle fractions, cell surface labeling by biotin probes and subsequent affinity purification have been proposed recently [54, 69]. The use of cell surface labeling by biotinylation has been reviewed for enrichment in integral membrane proteins in tissues [70]. In a novel method reported in 2015, biotinylated probes were used to isolate and enrich almost 7% of the total cellular protein in plasma membrane fractions from mouse brain tissues [71]. They used

sulfo-NHS-SS-biotin to label the membrane proteins of brain tissue slices and enriched the biotinylated proteins by streptavidin pull-down. Although many other hydrophobic membrane proteins, including tight junction protein ZO-2 were tagged and enriched, this method was not as successful for enriching claudins in their plasma membrane fractions. Only claudin-11 and isoform b of claudin-10 were found in the flow through fraction. However, this could also be due to claudins' localization in the cytoplasm. Nevertheless, the surface labeling method can be optimized for targeting claudins.

A combination of cell surface biotinylation and subcellular fractionation was shown to be effective for membrane protein isolation in breast cancer cell lines [72]. Here membrane proteome of MDA-MB-231 and osteotropic B02 metastatic variants were compared using biotinylation and streptavidin affinity enrichment. In this study, they have used the amine-reactive sulfo-NHS-SS-biotin to label the cell surface proteins followed by membrane fractionation by the traditional ultracentrifugation method. After that, the biotinylated proteins were affinity enriched in streptavidin resin and eluted by reducing the biotin probe. The enriched proteins were then analyzed by western blotting and MS. However, claudins were not specifically investigated in that study[72]. To further increase the enrichment of claudins and related junctional proteins for MS investigation, new probes conferring both specificity and spatial segregation is required for claudins. To be more precise in this respect, a combination of improved covalent probes and physical fractionation approach could be a way to overcome the enrichment obstacles.

1.7. Aims and objectives

Claudins are the regulators of paracellular permeability and integrity of epithelia and endothelia. Being a key structural component of tight junction strands, this family of proteins forms homo and heterotypic complex with other claudins as well as other structural proteins. With a variety of tissue specific expression profiles, claudins play an important role in coordinating signaling at cell-cell contact. As such, claudin dysregulation is involved in cellular transformation and metastatic progression. Claudins were studied as markers of a range of cancers because of their altered expression and altered cellular localization in number of carcinoma cells. Important progress has been made in structural characterization and understanding the mechanism of barrier regulation of claudins, with much more to be investigated regarding their role and regulations in

14

various cancers. To get a better understanding of regulation and mechanistic role of this proteins in cancer, profiling of this family of proteins as well as relevant partner proteins are necessary. To overcome the major problems of claudin detection- antibody cross-reactivity, heavy post translational modification and poor extractability, a more reliable techniques should be applied. A mass spectrometry based detection could provide us with exact identification as well as quantification of these proteins. A proteomic technology platform that allows molecular characterization of claudins with localization specificity, in conjunction with existing immuno-detection methods, would greatly facilitate current efforts in understanding endogenous claudin biology and pathology.

Proteomic profiling of endogenous claudins is challenging due to general difficulty in extracting low-abundance membrane proteins. As discussed earlier, diverse enrichment methods and technologies are known for profiling endogenous claudins and interacting proteins, however they suffer from nonspecificity issue and poor enrichment for MS based analysis. Membrane fractionation by ultracentrifugation have been tried for plasma membrane proteins enrichment, while needs cell type specific optimization. Affinity based chemical probes were also employed and were encouraging to detect of several claudins, although lacks exact subcellular localization specificity. Although a number of chemical enrichment method with probes were applied for membrane protein enrichment, a combination approach of membrane fractionation and chemical enrichment has not been demonstrated for endogenous claudin proteomics and can be established as an advanced workflow for claudin analysis.

Here, the aim of this study is to examine efficient methods for capturing surface proteins with chemical probes and to construct a reproducible workflow for proteomic detection of claudins by mass spectrometry as the first step towards this advanced workflow. This optimized workflow would then be combined with a specifically tailored membrane enrichment method for claudins to increase the proteomics coverage and detection limits. A chemical proteomic approach of cell surface tagging by biotinylation, followed by a NeutrAvidin pull-down and affinity purification was investigated here using model HCT-116 cells. Step by step optimization for isolation and purification of the targeted membrane proteins were done for subsequent detection by western blot and one-dimensional liquid chromatography mass spectrometry (LC/MS/MS) analysis. Due

to the fact that the HCT-116 cells do not express the common claudins, such as claudin-1 and 2, only claudin-4 was used for optimizing the immunodetection method. Along with claudin-4, EGFR was used as a marker of enrichment method. Considering the very limited timeline of this training project, reproducibility of the workflow remains to be confirmed, yet the complete workflow has been explored for surface protein detection. Future work on this workflow will include optimization of the protein capture step with a better immobilization surface with less background and a more streamlined sample preparation protocol to minimize protein loss before mass spectrometry. An efficient chemical capture protocol will then be combined with membrane partition methods to achieve claudin detection and analysis by mass spectrometry with better localization specificity. Which will greatly facilitate the understanding of biology as well as pathology of claudins.

2. Experimental methods

2.1. Cell culture

HCT 116 wild type colorectal cancer cell line was provide by Charlie Ann from the Department of Biomedical Sciences at Macquarie University. Cells were seeded in 15 cm culture plates in Dulbecco's Modified Eagle Medium (DMEM, ThermoFisher Scientific), supplemented with 10% fetal bovine serum (FBS, ThermoFisher Scientific). Cells were washed with phosphate-buffered saline (PBS, pH 7.4) and collected with a plastic scraper. Cell pellets were reconstituted in PBS and centrifuged at 1000×g for 5 minutes at 4 ºC. The collected cell pallets were stored at -80 ºC for further analysis.

2.2. Biotinylation of Cell surface and Plasma Membrane Proteins

Live HCT-116 cells (2-3×10^7 cells/plate) were washed twice with room temperature (RT) PBS and then washed twice with ice-cold PBS. Cell surface biotinylation was carried out by adding 10 mL of Sulfo-NHS-SS-biotin (ThermoFisher Scientific) solution (0.25 mg/mL in ice-cold PBS) to each plate. The plates were incubated at 4 °C with gentle agitation on a platform rocker for 15 min. For control plates, cells were treated with either 10 mL of quenched Sulfo-NHS-SS-biotin solution (0.25 mg/mL Sulfo-NS-SS-biotin quenched by treating with tris-buffered saline (TBS, pH-7.6) at 50 °C for 15 min) or ice-cold PBS at 4 °C for 15 min. The labelling reaction was quenched by washing with 10 mL of ice-cold TBS once. Then the cells were gently scraped from the plate with a plastic scrapper and reconstituted in TBS. The cell pellets were collected after centrifuging at 400×g for 10 mins, at 4 °C. The washing process was repeated twice to remove unreacted biotin. The cell pellets were either used immediately or stored at -80 °C for further analysis.

2.3. Biotinylation of bovine serum albumin (BSA) as a control

Biotinylation of BSA (Sigma Aldrich) was performed according to the supplier's technical protocol. A solution of 2 mg/mL BSA was incubated for 30 min at RT with 10 mM solution of

sulfo-NHS-S-S biotin (20 fold molar excess of biotin). Unreacted biotin was removed by repeated washing with PBS using a Vivaspin filter unit (Sartorius-Stedim) with 5 KDa molecular weight cut off. After washing, the recovered biotinylated BSA solution was quantified by BCA assay and stored at -20 °C.

2.4. Cell lysis

Cell pellets ($2\text{-}3\times10^7$ per pellet) from storage were thawed on ice and dissolved in 1 mL of hypotonic lysis buffer (10 mM HEPES, 1.5 mM $MgCl_2$, 10 mM KCl and 1 mM EDTA; pH 7.4) with 1x mix of protease inhibitors (Roche) prior to lysis.

2.4.1. Lysis by Probe Sonication

Cells were suspended in 1 mL of hypotonic lysis buffer in a 15 mL falcon tube and sonicated by 10 bursts of a probe sonicator (Branson Sonifier 450, John Morris) at duty cycle 50. The sonication was repeated twice, resting the cells on ice for 2 min in between each sonication cycle. The resultant cell lysates were then centrifuged at either $400\times g$ for 10 min at 4 °C or $3000\times g$ for 10 min at 4 °C for comparison. The clarified supernatant was transferred to a fresh 1 mL microfuge tube for analysis.

2.4.2. Lysis by bead beating

Equal amount (0.35 g) of ceramic and glass beads were mixed in a 2 mL Precellys soft tissue homogenizing tube (Bertin Technologies). Cells were suspended in 1 mL of lysis buffer and transferred to the homogenizing tube with ceramic and glass beads. The homogenizing tubes were loaded into Precellys 24-tissue homogenizer (Bertin Technologies) and homogenized by vortexing at 5500- 1×20 speed for 30s. Vortexing was repeated twice, resting the tubes on ice for 2 min in between each cycle. The resultant cell homogenate was then centrifuged at the same speed and condition used for homogenization. The clarified supernatant was transferred to a fresh 1 mL microfuge tube for analysis.

2.4.3. Cell Lysis by syringe homogenization

Cells were suspended in 500 µL of lysis buffer and taken in a 1 mL syringe. Cells were slowly passed through a 26 G needle for 10 times. The process was repeated after resting the cells on ice for 1 min. The resultant cell lysates were then centrifuged as indicated above in section 2.4.1. The clarified supernatants were collected and normalized with lysis buffer to 1 mL volume.

2.4.4. Lysis by Dounce Homogenizer

Cells were suspended in 1 mL of hypotonic lysis buffer and transferred to the glass-tube of Dounce homogenizer. The cell suspensions were homogenized by either 50 strokes of the tight fitted pestle operated with a drill or with 100 strokes by hand. The process was carried out at 4 ^{o}C in a cold room. The resultant cell lysates were then centrifuged as mentioned above on the probe sonication section (section 2.4.1). The clarified supernatants were collected for analysis.

2.5. Affinity Pull-down and Elution of Plasma Membrane Proteins

Biotinylated plasma membrane proteins were affinity enriched by binding with Pierce™ NeutrAvidin agarose resin (ThermoFisher Scientific). Settled beads (160 µL) were taken in Pierce™ snapcap spin column (ThermoFisher Scientific) and washed thrice with 200 µL of lysis buffer. 1mg/mL clarified cell lysates (800 µL) (biotinylated/control) were incubated with the conditioned NeutrAvidin beads in spin columns at 4 ^{o}C in an end-over-end rotator for overnight. The next day the columns were centrifuged at 1000×g for 1 min, and the unbound flow through and cytosolic fractions were collected. Three subsequent washes of the beads were carried out with lysis buffer to wash all the unbound cytosolic proteins. The washed beads with biotinylated proteins were then incubated with elution buffer (160 µL) (50mM DTT in 1% SDS) at 90 ^{o}C for 30 min. After incubation, the eluted plasma membrane fractions were collected by centrifuging the columns at 1000× g, for 2 min at RT. The eluted fractions were stored at -20 ^{o}C with previously collected lysates, flow through fractions, washed fractions and beads.

2.6. Protein Quantification

The protein content in the clarified cell lysates, collected flow through fractions and recovered biotinylated BSA were quantified using Pierce™ BCA protein assay kit (ThermoFisher

Scientific). BSA, supplied with the kit, was used as standard. The absorbance was taken at 562 nm using a FLUOstar OPTIMA microplate reader (BMG Labtech). The relative protein quantity of the eluted plasma membrane fractions of the beads was calculated from the relative density the known standard BSA from the SDS PAGE gels.

2.7. SDS-PAGE Analysis

All the collected lysates, flow through and eluted fractions were subjected to SDS PAGE analysis using TruPAGE™ precast gel system (Sigma-Aldrich). Samples were mixed with TruPAGE™ 4X LDS sample loading buffer (Sigma-Aldrich) with 100mM DTT and heated at 90ºC for 5 min before running into TruPAGE™ 4–20%, 10 x 10cm, 12-well precast gel. The gels were run with 1x TruPAGE™ Tris-MOPS SDS express running buffer (Sigma-Aldrich) for 56 minute at 180 V in an XCell SureLock™ Mini-Cell Electrophoresis System (ThermoFisher Scientific). After the run, the gels were subjected to either western blot analysis or stained for visualization by scanning. For staining procedure, the protein bands in the gels were fixed by immerging the gel in coomassie fixing solution (24.8% Methanol, 2.48% Orthophosphoric acid (v/v)) for 30 min. Then the gels were stained overnight with coomassie blue solution (25% Methanol, 3% orthophosphoric acid, 1.87 g/L Brilliant Blue G-250). The gels were destained with coomassie destaining solution (40% methanol, 10% acetic acid (v/v)) and scanned using a normal precision scanner (Hp).

2.8. Western Blot Analysis

The gels were blotted using the iBlot® Dry Blotting System (ThermoFisher Scientific) on iBlot® PVDF transfer stack (ThermoFisher Scientific) and blotted for 7 min. The efficiency of the dry blot transfer was checked using Ponceau S solution (Sigma-Aldrich). The membranes were washed in TBS-Tween20 (TBST, pH-7.6) and blocked with 5% skimmed milk for 1 hour at RT. After blocking, the membranes were incubated overnight in Rabbit polyclonal anti-EGFR primary antibody (Sigma-Aldrich, 1:10000 dilution) and Rabbit monoclonal anti-Claudin-4 primary antibody (Abcam, 1:1000 dilution) at 4 ºC. The next day, the membrane was washed in TBST and incubated for 1 hour in Sheep anti-Rabbit IgG (H+L), HRP secondary antibody (1:8000 dilution) at RT. The membranes were visualized using Novex™ ECL Chemiluminescent

Substrate Reagent Kit (ThermoFisher Scientific) in a LAS-3000 (Fujifilm Life Science) imaging system.

2.9. Sample preparation for mass spectrometry analysis

Samples were prepared using two different digestion methods. The whole cell lysate of HCT-116 (no probe) was processed by in-solution digestion method. The eluted fraction was digested using tube-gel digestion method[73].

2.9.1. In-solution digestion of whole cell lysate

Approximately 50 µg of clarified cell lysates (untreated) were taken for in-solution digestion. The volume of each of the lysate samples was adjusted to 100 µL by adding 50 mM ammonium bicarbonate (Sigma-Aldrich). The samples were then reduced with 10 mM dithiotheritol (DTT, Sigma-Aldrich) for 30 min at 55 °C. After reduction, samples were alkylated with 20 mM iodoacetamide (IAA, Sigma-Aldrich) for 30 min in dark at RT. Then 20 µL of 0.2 µg/µL trypsin (Sigma-Aldrich) solution in 100 mM ammonium bicarbonate was added in 25:1 protein-trypsin ratio and incubated at 37 °C for 16 hours.

The digested peptides were then dried and reconstituted with 0.1% FA. The acidified peptides were desalted by binding with 100 µL OMIX C18 tip (Varian) and eluted with 0.1% FA and 70% ACN. Finally, the samples were dried in vacuum concentrator and stored at -20 °C for MS analysis.

2.9.2. Tube gel digestion of membrane enriched fraction

For tube gel digestion, 50 µL of elute with approximately 30 µg of proteins in each biotinylated and 13 µg in control (no probe) were taken. Samples were alkylated with 150mM IAA for 30 min in dark. Then 18.5µL of acrylamide/bis acrylamide (40%, 37.5:1, Bio Rad), 2.5 µL of 10% APS (Ammonium persulfate; Sigma-Aldrich) and 1µL of 100% TEMED (N, N, N', N' - Tetramethylethylenediamine; BioRad) were added to each tube to transform the solution into 40% tube-gels.

Proteins in the tube-gels were digested by in-gel trypsin digestion. After 30 min, when the gels were set, they were cut into small pieces and washed thrice with 1 mL of 50mM ammonium bicarbonate in 50% ACN. Then the gel pieces were further dehydrated with 100% ACN and dried using a vacuum concentrator. The dried gel pieces were incubated with 83 ng/ µL trypsin (Sigma) in 50mM ammonium bicarbonate (pH 8) in 10:1 protein-trypsin ratio at 37 °C for 16 hours.

The next day, the digested peptides were extracted twice with 200 µL of 0.1% FA in 50% ACN, then fully dehydrated with 200 µL 100% ACN. Later, the peptide extracts were vacuum dried and reconstituted with 0.1% FA. Then the peptides were then desalted similarly with 100 µL 100 µL OMIX C18 tip and eluted with 0.1% FA and 70% ACN. Finally, the samples were dried in vacuum concentrator and stored at -20 °C for MS analysis.

2.10. LC/MS/MS analysis

Liquid chromatography mass spectrometry (LC-MS) analysis was performed on an Easy nLC system connected to a Q-Exactive™ Hybrid Quadrupole-Orbitrap™ mass spectrometer (ThermoFisher Scientific). The equivalent of 2 µg of peptide was dissolved in 10 µL of 0.1% FA and loaded to an in-house packed 100 µm × 3.5 cm reversed phase peptide trap (Solid core Halo® 2.7 µm 160 Å ES-C18, Advanced Materials Technology), peptide separation was carried out using a self-packed 75 µm × 10 cm (Solid core Halo® 2.7 µm160 Å ES-C18, Advanced Materials Technology) column. Separation for cell lysate samples were conducted using a linear gradient of 90% ACN in 0.1% FA (Buffer B) reaching from 0% to 50% in 110 min, 85% in 112 min, 85% in 120 min. Biotinylated elute samples were separated using linear gradient of the same buffer reaching from 0% to 50% in 50 min, 85% in 52 min, 85% in 60 min. Flow rate was set at 300 nL/min. The Q-Exactive Quadrupole-Orbitrap (ThermoFisher Scientific) was set up in a data dependent MS/MS mode where a full spectrum scan (350–1600 m/z, resolution 35,000) was followed by a selection of maximum of ten Collision induced dissociation (CID) tandem mass spectrum (100 to 2,000 m/z). Peptide ions were selected as the 10 most intense peaks of the MS1

scan. The normalized collision energy used was 35% in CID. We applied a dynamic exclusion list of 45 s.

2.11. Proteomic data analysis

Raw data were analyzed in the MaxQuant[74] version 1.6.0.16. The MS/MS spectra were matched against the human Uniprot FASTA database. Enzyme specificity was set to trypsin, search cysteine carbamidomethylation as a fixed modification and N-acetylation of protein, oxidation of methionine as variable modifications. Up to two missed cleavages were allowed for protease digestion, and peptides had to be fully tryptic. A common list of contaminants, reverse protein IDs and proteins with 0 intensity were removed prior to further analysis. Respective Gene ontology-cellular compartment information for each protein were retrieved from UniProt.

3.   Results and discussion

Isolation and characterization of plasma membrane proteins are recognized as a very challenging task. Due to their hydrophobicity, poor solubility and lower abundance in comparison to cytoplasmic proteins, plasma membrane proteins are generally underrepresented in standard proteomic studies[75-76]. Our target proteins in this study, the claudins, are low abundant proteins and distributed in various locations in cells[44]. In the current study, the aim is to investigate methods that can integrate into an optimal and compatible workflow (Fig. 3.1) for chemically capturing and enriching surface proteins followed by proteomic detection and analysis. An advanced workflow such as this would then incorporate future chemical tagging of claudin proteins specifically to increase the proteomics coverage and detection limits. Model HCT-116 cells were used to investigate the methods that would constitute the workflow, starting with cell surface tagging by biotinylation, followed by a NeutrAvidin pull-down and affinity purification for subsequent detection by western blot and one-dimensional liquid chromatography mass spectrometry (LC/MS/MS) analysis. Each stage of this workflow was investigated and optimized in the study here, and the results and discussion are structured accordingly.

In brief, a number of common cell lysis methods were investigated to select a cell homogenization method with better recovery of membrane proteins from HCT-116 cells (Fig. 3.1 (b), section 3.1). The biotinylation condition was optimized for tagging claudin as well as other plasma membrane proteins (Fig. 3.1 (a), section 3.2). For the affinity purification stage, different ratio of NeutrAvidin bead to protein quantity were trialed along with numerous elution conditions, as analyzed by western blots (Fig. 3.1 (d-f), section 3.3). For 1D LC/MS/MS analysis of the enriched surface membrane protein fraction, tube- gel assisted digestion methods were tested to increase the peptide coverage of the membrane protein fraction (Fig. 3.1 (g-h), section 3.4). The proteomic analysis of HCT-116 whole cell lysates was also conducted for comparison.
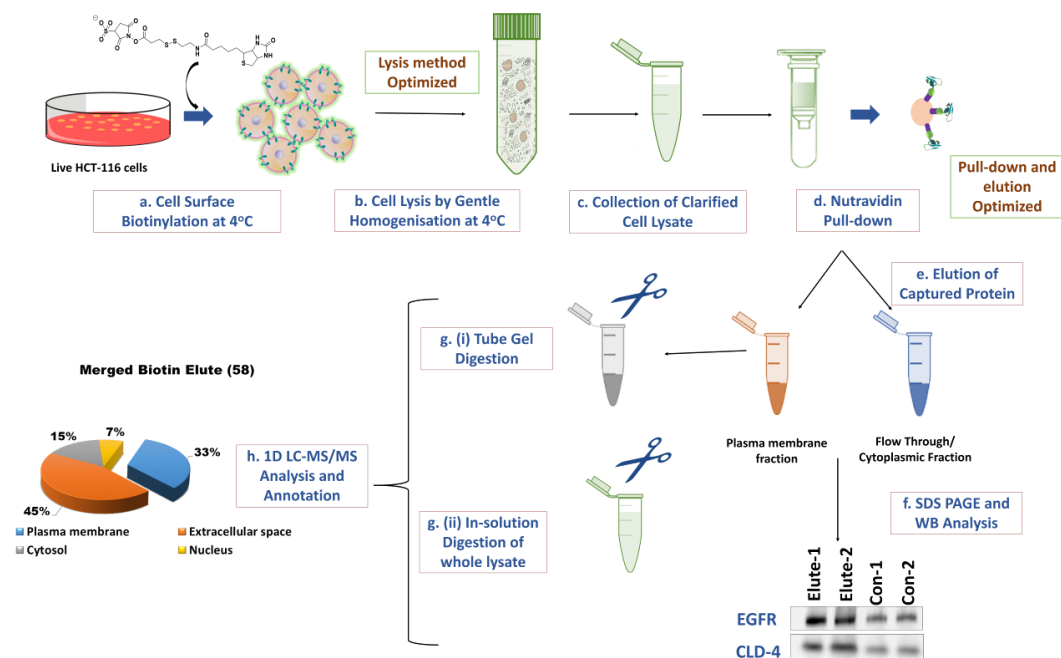
Fig. 3.1 Overview of the workflow investigated. (a–e) Biotinylation and pull-down of membrane proteins. (f) SDS-PAGE and western blot analysis of the eluted fraction. (g–h) LC/MS/MS analysis of elute fraction and crude lysate and peptide annotation.

3.1. Mechanical cell lysis method optimisation for membrane protein enrichment in HCT-116 cells

Different cell lines are known to have different cytoskeletal organizations that require suitably developed protein extraction methods. Here cell lysis methods were investigated for the model cells (HCT-116) to find the best protein extraction conditions. HCT-116 cells have been investigated for plasma membrane protein extraction by Nagano, K. *et al.*, where they applied probe sonication for cell lysis and a combination of low (1000×g for 10 min) and high speed centrifuge (100,000×g) for pelleting the membrane fraction[77]. Use of tight fitted Dounce homogenization followed by sucrose gradient cellular fractionation was also reported for extracting membrane fraction in transfected HCT-116 cells for analyzing signaling proteins [78].

In this study, five different mechanical lysis methods, namely, probe sonication, bead beating, syringe homogenization using 26G needle, Dounce homogenization using drill, Dounce homogenization using hand strokes were investigated. All of these methods were frequently

25

reported for membrane protein enrichment from various cultured epithelial like cells [48, 59, 77, 79]. Gentle mechanical homogenization such as syringe lysis or Dounce homogenization, has been reported to be effective for increasing membrane protein ratio in the clarified supernatant by keeping the larger organelles intact, which can be removed by centrifugation at lower speeds. Removal of the nucleus greatly reduces formation of cytoskeletal aggregates and prevents release of DNA in the lysate[55]. Also junctional proteins, such as claudins and other membrane proteins, have unique complex formed beneath the plasma membrane. Gentler methods could keep those membrane protein complexes intact which in turn increases the ratio of membrane or membrane associated proteins in the lysate.

Along with gentle mechanical lysis, a hypotonic lysis buffer (10 mM HEPES, 1.5 mM $MgCl_2$, 10 mM KCl and 1 mM EDTA; pH 7.4), reported as effective for membrane protein extraction was chosen [79]. Hypotonic lysis buffer swells the cells by osmotic shock and facilitates the breakage of plasma membrane upon mechanical homogenization. In addition, factors such as temperature and centrifuge speed were also investigated. As shown below, protein recovery was analyzed first by protein quantification to estimate the total amount recovered by each method with different centrifuge speed, and then by western blot to evaluate recovery of specific membrane proteins.

3.1.1. Effect of cell lysis methods on protein recovery quantified by BCA method

The amount of total protein in cell lysates by different methods shows significant variation. The probe sonication method, that disrupts the cells with ultrasound waves, produced lysates with the highest quantity of protein (1.41–1.70 mg/mL) than all other four methods. The bead beating method (0.83–1.19 mg/mL) and syringe homogenization method (0.94–1.18 mg/mL) recovered a moderate amount of proteins. While the Dounce homogenization with drill showed a comparably moderate level of protein recovery (1.03–1.05 mg/mL) in the lysate, Dounce homogenization with hands gave slightly better protein recovery (1.10–1.23 mg/mL) then those with drill. Given that sonication is the most disruptive amongst all the method used, proteins could be released from organelles due to the heat generated in microscopic compartments of the cells. Also, foaming in the lysate was observed for the bead beating and syringe homogenization method, which could also cause significant protein loss.

For all of the five lysis methods, centrifuge speed was investigated and analyzed by protein quantification. For each method, centrifuging was performed at 400×g for 10 min or 3000×g for 10 min, and the former was found to be slightly more effective for higher lysate protein quantity. For example, for the two Dounce-based methods, the samples centrifuged at 400×g showed higher protein concentration (1.23 mg/mL) than those centrifuged at 3000×g (1.10 mg/mL).

3.1.2. Effect of cell lysis methods on protein recovery by western blot analysis

Western blot analysis of the lysate samples was conducted using two membrane proteins as markers: EGFR and claudin-4. Cell lysate from MDA-MB-468 breast cancer cells was used as positive control for EGFR and Caco-2 gastric adenocarcinoma cell lysate was used as positive control for claudin-4. The lysates from four different methods and two different centrifuge speed, showed distinguishable variation in EGFR and claudin-4 by western blot analysis. As shown in Fig 3.2, lysates from all five methods centrifuged at 400×g (Fig 3.2 (a)) were compared to those centrifuged at 3000×g (Fig 3.2(b)). Lysates from the Dounce method by hand showed relatively higher amount of protein for both EGFR and claudin-4. On the contrary, sonicated samples showed much less recovery for EGFR, but a good recovery for claudin-4. It is possible that, sonication disrupts the nucleus that results in aggregation of larger proteins (i.e. EGFR). Dounce homogenization is known as a gentle method to rupture the plasma membrane only, leaving the nucleus intact for removal by centrifuge. However, the Dounce method fitted with a drill provided relatively less recovery compared to that by hand. It is possible that the rotating motion from the drill is not as effective compared to the up-down thrusting motion by hand. The Dounce homogenization method by hand was found to be the most suitable mechanical homogenization method for the HCT-116 cells.
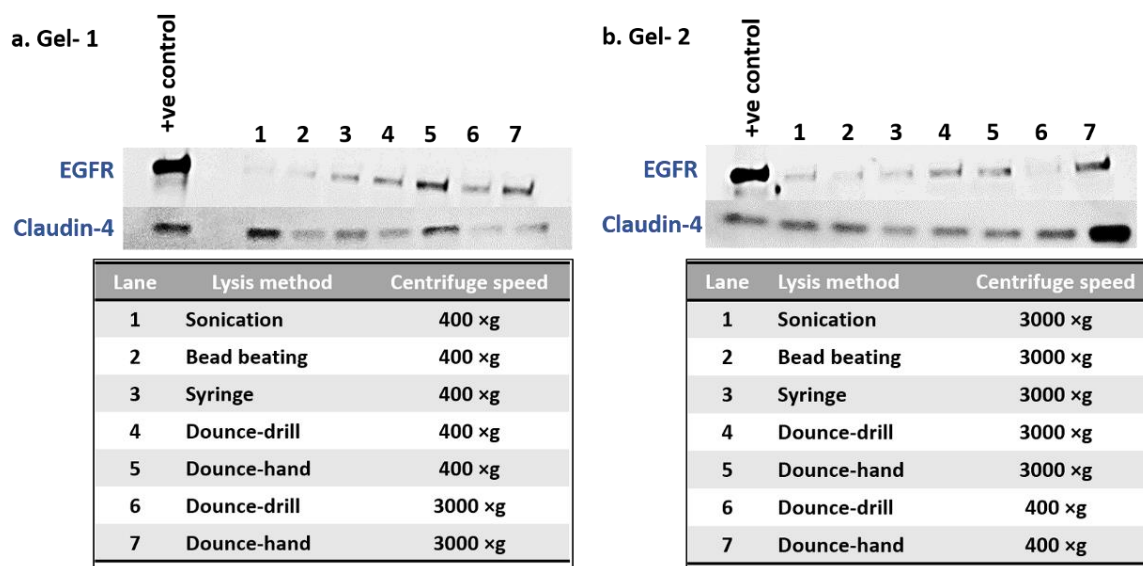
Fig. 3.2 Western blot of the lysates from different lysis methods using EGFR and claudin-4 as markers. (a) Gel- 1 Five lysates from five methods at 400×g compared with Dounce lysates at 3000×g. (b) Gel- 2: Five lysates from five methods at 3000×g compared with sonicated and Dounce lysates at 400×g.

3.2.   Biotinyation of cell surface and plasma membrane proteins in HCT-116 cells

Water soluble biotin probes have been widely reported to tag and purify cell surface and plasma membrane proteins in various tissues and cell lines [71-72, 75, 80]. Biotinylation of HCT-116 cells using sulfo-NHS-SS-biotin was not reported previously, although another noncleavable biotin label was used but resulted in poor protein recovery[77]. Here, commercially available sulfo-NHS-SS-biotin probe was used for HCT-116 cells with membrane impermeability due to the negatively charged sulfo group. Cleavability of the S-S bond with DTT would allow subsequent release of captured proteins [80]. As discussed below, the efficiency of the biotinylation and affinity pull-down were evaluated by western blot analysis.

3.2.1. Efficiency of biotinylation and elution of surface proteins

The biotinylation reaction was performed under typical literature conditions at 4 $^{\circ}$C for 15 min to minimize cell loss upon prolonged biotin treatment and to restrict the labeling of cytosolic proteins. Cell death and loss during washes was observed upon prolonged biotin treatment, which in turn resulted in poor amount of proteins in whole cell lysate (data not shown). For subsequent purification of the biotin labeled proteins, various types of avidin-based affinity purification methods are well established[75, 81]. Due to the extraordinary stability of biotin-avidin complex, elution of the biotinylated proteins from avidin remained challenging[80-81]. As discussed below, following biotinylation, optimization of the elution step is critical for recovery of tagged proteins for subsequent LC/MS/MS analysis.

Charge neutral NeutrAvidin agarose beads were used to capture the biotinylated membrane proteins. These beads were reported to exhibit a near neutral isoelectric point and less nonspecific binding properties than other avidin beads such as streptavidin. The biotinylated cell lysate was incubated overnight in a NeutrAvidin packed column. Two negative controls, either quenched or no probe were used to check the extent of the non-specific binding of the sticky proteins with NeutrAvidin beads. The quenched probes were previously heated with primary amine containing buffer (TBS; pH- 7.6) to quench the active amine reactive site of the probe. Cells treated with this quenched probe will not labeled on the cell surface, but will have lysates that occupy biotin binding sites of the beads showing background in terms of non-specific binding. The negative control with no probes will demonstrate the extent of non-specific binding of the charged or sticky components to the beads.
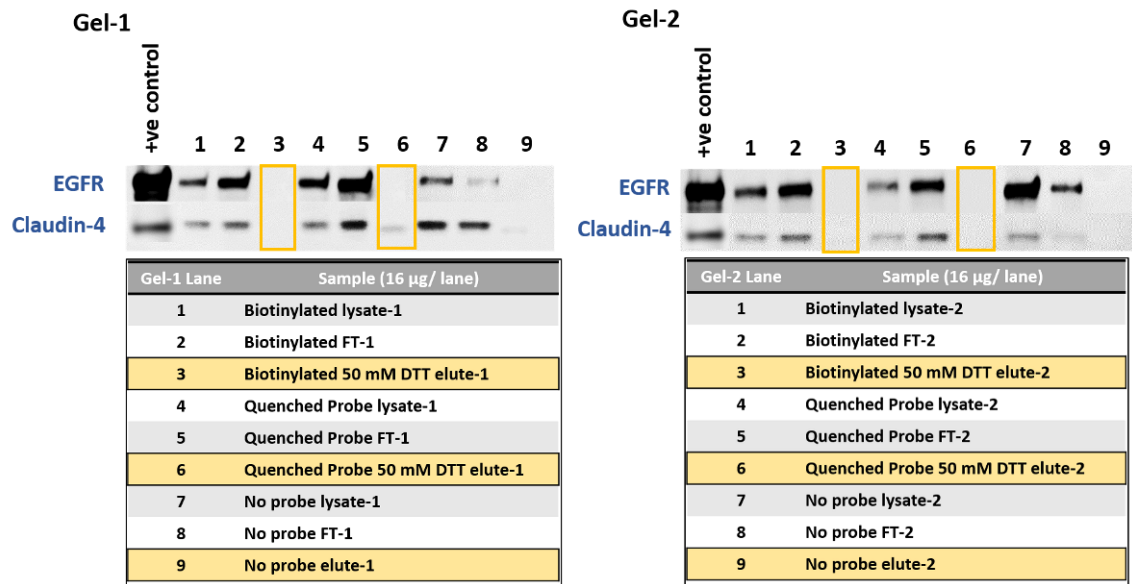
Fig: 3.3 Efficiency of biotinylation and elution was compared Western blot analysis. Three fractions from biologically duplicated samples- biotinylated, quenched probe treated and no probe treated were compared in Gel-1 and Gel-2.

A mild elution buffer (50 mM DTT in lysis buffer, pH 7.4 and 1x mixture of protease inhibitor) was first used to elute the captured proteins at RT. The eluted fraction (elute), flow through (FT) and whole cell lysate were analysed by western blot with claudin-4 and EGFR antibodies. The EGFR and claudin-4 were present in both the lysate and flow through fractions (Fig 3.3, (a) lane 1 and 2, 4 and 5, 7 and 8 (b) lane 1 and 2, 4 and 5, 7 and 8) in all three groups of samples but the elute fraction in either the biotin-labeled or control samples showed no detectable proteins in the western blot analysis (Fig (a) lane 3, 6 and 9 (b) lane 3, 6 and 9). The presence of claudin-4 the flow through fractions is possibly an indicator of its cytosolic location in HCT-116 cells, which corroborates with the recently reported data of altered membrane to cytosolic localization of claudin-4 in colon carcinomas [42].

It appeared that the elution condition likely did not cleave the S-S bond with the 50 mM DTT in it. Elution with 50-100 mM DTT with or without 1% SDS at 50 °C was previously reported to cleave the S-S bond of the Sulfo-NHS-S-S biotin probe[72, 75-76]. As the lysis buffer was not supplemented with any other reducing agent or detergent, possibly cellular lipids forming a layer around the beads were impeding the reduction of S-S bond by DTT.
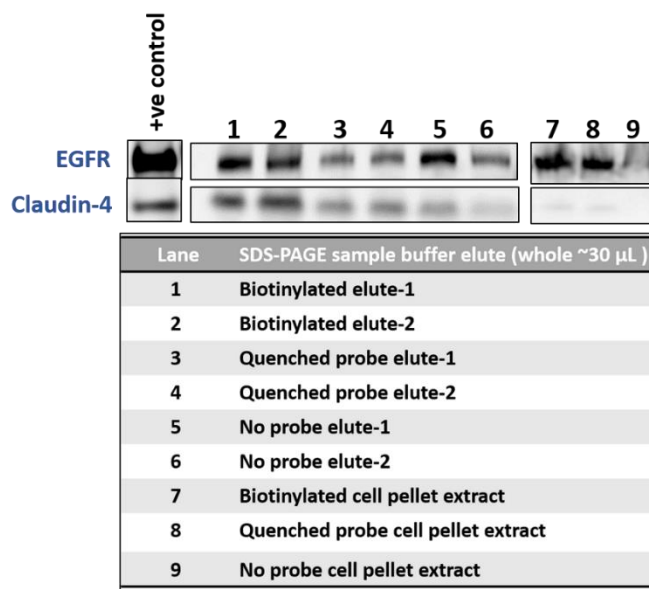
Fig. 3.4 Western blot analysis of eluted fractions after biotinylation using SDS-PAGE sample buffer.

A harsher condition for eluting the biotinylated proteins from the NeutrAvidin beads was tested using the SDS PAGE sample buffer as the base buffer (40% Glycerol, 4% Lithium Dodecyl Sulfate (LDS), 4% Ficoll 400, 0.025% Phenol Red, 0.025% Brilliant Blue G250, 2 mM EDTA) with 50 mM DTT. While the mild elution condition of using the lysis buffer with 50 mM DTT remained inefficient for protein elution, boiling with this base buffer and 50 mM DTT, released the proteins from the beads. The western blot analysis of the total biotin labeled extracts showed about twice the amount of EGFR and claudin-4 recovered compared to the negative control fractions (quenched probe treated or no probe) (Fig 3.4). Western blot analysis of enriched claudin-4 was reported by Lohrberg, D. from caco-2 cell lines, using a GST-CPE fusion column for affinity enrichment[46]. Considering the lower expression of claudin-4 in HCT-116 cells then that in Caco-2 cells, enrichment of claudin-4 by cell surface biotinylation was encouraging.

However, nonspecific sticking of these two membrane proteins were also observed in both group of control elute fractions (Fig 3.4 lane 3–6). As discussed earlier, negative controls used here (quenched probe or no probe) were designed to examine the level of nonspecific binding to the beads. The levels of nonspecific sticking were consistent with the controls of either with no probes or with quenched probes. Clearly the beads used, were not as efficient as the suppliers claimed. Concentrating the eluted fractions provided a better detection. Also, alternative

immobilization surface material other than NeutrAvidin agarose beads can be tested for future analysis.

To ensure that labeled surface proteins were not collected in the remaining cell pellets after lysis, proteins were also extracted from the pellets by boiling with SDS PAGE loading buffer and analyzed by western blot (Fig 3.4 lane 7–9). Absence of claudin-4 in the pellet extraction confirmed the efficiency of the optimized Dounce homogenization method described in the previous section. On the contrary, a moderate amount of EGFR from cell pellet extraction indicated possibly altered nuclear and vesicular EGFR trapped with the nucleus and cellular debris. Considering the heavy glycosylation of membrane EGFR, it also indicates the extraction difficulties of this protein by biotinylation[82].

3.2.2. Optimisation of elution condition for protein recovery

To evaluate the optimal elution condition for recovering the captured proteins, different elution buffer in combination with different temperature and incubation time was trialed using a fixed amount (50 µg) of biotinylated BSA (1mg/mL) as a positive control was incubated with 50 µL (dry volume as recommended ratio by suppliers) of beads. Altogether 11 different conditions were tested and subjected to SDS PAGE analysis (Fig 3.5 (b)). Here, Buffer-1, SDS-PAGE loading buffer with 50 mM DTT showed higher efficiency at 55 °C to recover more than 70% of BSA from bead (Fig 3.5 (a) lane 6). However, prolonged and high temperature incubation with SDS-PAGE loading buffer found to reduce protein recovery up to 10% (Fig 3.5 (a) lane 2–5). A milder elution buffer Buffer-2 (1% SDS and 50 mM DTT) solution recovered the highest (83%) amount of BSA upon 30 min incubation at 90 °C (Fig 3.5 (a) lane 9). On the contrary, the mildest buffer, Buffer-3 (50 mM DTT in PBS) remained inefficient to elute most of the BSA (Fig 3.5 (a) lane 13–14). Being simple in composition and milder than SDS-PAGE loading buffer, 1% SDS and 50 mM DTT offers the most suitable elution condition for NeutrAvidin beads.
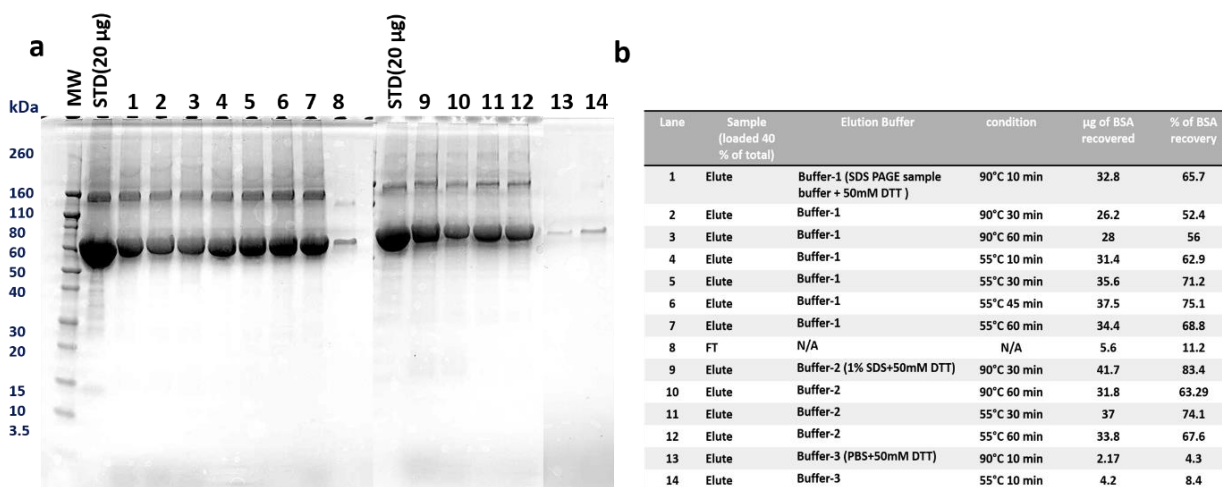
Fig. 3.5 SDS-PAGE analysis of different elution condition compared (a) Gel image of biotinylated BSA recovered from different elution condition. (b) Sample description of SDS-PAGE and percentage of proteins recovered calculated by relative densitometry.

The table in panel (b):

Lane	Sample (loaded 40 % of total)	Elution Buffer	condition	µg of BSA recovered	% of BSA recovery
1	Elute	Buffer-1 (SDS PAGE sample buffer + 50mM DTT)	90°C 10 min	32.8	65.7
2	Elute	Buffer-1	90°C 30 min	26.2	52.4
3	Elute	Buffer-1	90°C 60 min	28	56
4	Elute	Buffer-1	55°C 10 min	31.4	62.9
5	Elute	Buffer-1	55°C 30 min	35.6	71.2
6	Elute	Buffer-1	55°C 45 min	37.5	75.1
7	Elute	Buffer-1	55°C 60 min	34.4	68.8
8	FT	N/A	N/A	5.6	11.2
9	Elute	Buffer-2 (1% SDS+50mM DTT)	90°C 30 min	41.7	83.4
10	Elute	Buffer-2	90°C 60 min	31.8	63.29
11	Elute	Buffer-2	55°C 30 min	37	74.1
12	Elute	Buffer-2	55°C 60 min	33.8	67.6
13	Elute	Buffer-3 (PBS+50mM DTT)	90°C 10 min	2.17	4.3
14	Elute	Buffer-3	55°C 10 min	4.2	8.4

3.2.3. Optimisation of bead-protein ratio for protein capture

The recommended ratio from the suppliers of NeutrAvidin resin (ThermoFisher Scientific) was to incubate 1–2 µg of pure biotinylated protein in 1 µL of dry/settled bead. With the surface biotinylated cell lysate, it is assumed that 5–10% of the total protein could be biotinylated. The optimal bead-protein ratio for HCT-cells is not known but critical to understand the extent of protein recovery, therefore optimization of the bead-protein ratio was carried out with sulfo-NHS-SS-biotin and NeutrAvidin beads.

To evaluate capturing capacity of the beads and optimal recovery of the captured proteins, various amounts (500, 250, 100, 75, 50 and 25 µg) of biotinylated BSA (1 mg/mL solution) were incubated overnight with 50 µl of settled NeutrAvidin beads at 4 ºC. After overnight incubation, the bead bound proteins were recovered by boiling the beads in 50 µL of 1X SDS PAGE sample buffer and 50 mM DTT (as mentioned in the previous section for 10 minutes at 90 ºC. 20µl of the eluted proteins were loaded in the SDS PAGE gel wells. Gel bands were analysed by relative densitometry of the standard BSA (15 µg) loaded (Fig 3.5 (a). lane 2). Higher bead to protein

ratio showed higher levels of labelled proteins recovered. A ratio of 10:1, protein (µg)–bead (µL) (Fig. 3.5. (a) Lane 3 and 4) failed to capture most of the biotinylated BSA that remained in the FT fraction. Ratios of 1:1 or 1:2 improved the capture rate to 50–60% of labeled BSA.

Lane	Sample	Protein (µg) to bead (µL) ratio	% of total sample loaded	µg of BSA recovered	% of recovery
1	Elute	10:1	40	12.8	2.6
2	FT	10:1	5	271.6	54
3	Elute	5:1	40	13.3	5.3
4	FT	5:1	8	153.7	61.5
5	Elute	2:1	40	22.2	22.2
6	FT	2:1	20	50.15	50.15
7	Elute	1.5:1	40	20	26.7
8	FT	1.5:1	27	32.5	43.29
9	Elute	1:1	40	23.6	47.3
10	FT	1:1	40	8.7	17.4
11	Elute	1:2	20	30.1	60.2
12	FT	1:2	40	3.3	6.6
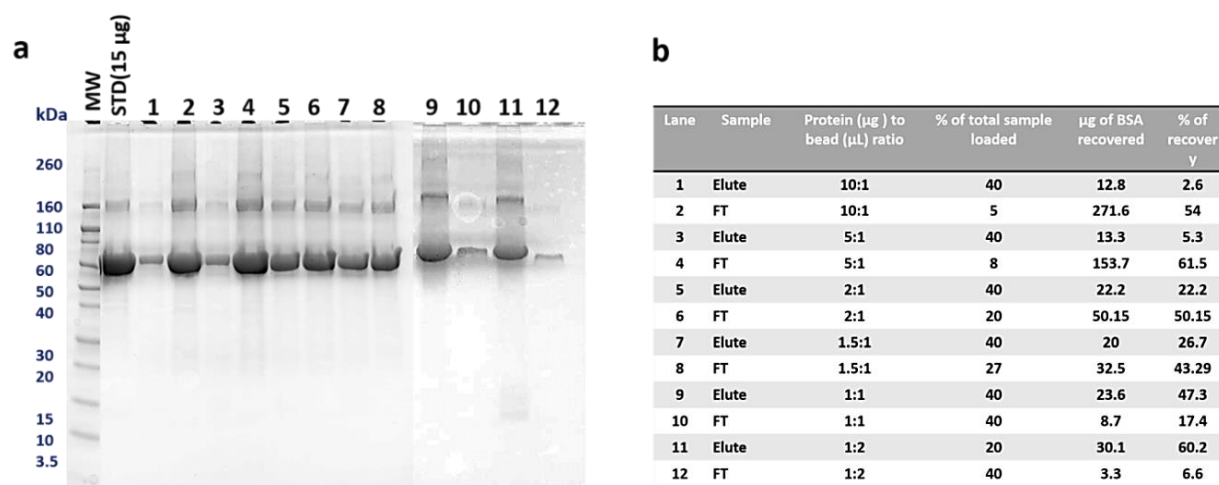
Fig. 3.6 SDS-PAGE analysis of different protein-bead ratio compared with biotinylated BSA (a) Gel image of biotinylated BSA recovered from incubation with of different protein-bead ratio. (b) Protein-bead ratio of SDS-PAGE samples and percentage of proteins recovered in elute and FT.

3.2.4. Optimal protein-bead ratio and elution condition for capturing spiked biotinylated BSA in cell lysate

To optimize the capturing and elution condition in cell lysates, biotinylated BSA was spiked into HCT-116 cell lysates. A solution of 50 µg of biotinylated BSA was spiked in 500 µL of cell lysate (normalized to 1 mg/mL) and incubated in three different protein (µg)-bead (µL, dry volume) ratio (2:1, 1:1, 1:2). After eluting with the optimized elution buffer (1% SDS, 50 mM DTT), all three elutes and FT fractions were subjected to SDS PAGE analysis. Highest recovery (50%) was found with the 1:2 protein (µg)–bead (µL) ratio (Fig 3.7(a) lane 5). However, the protein recovery rate in the lysate was 10% less compared to the recovery found in the previous section (section-3.2.3).
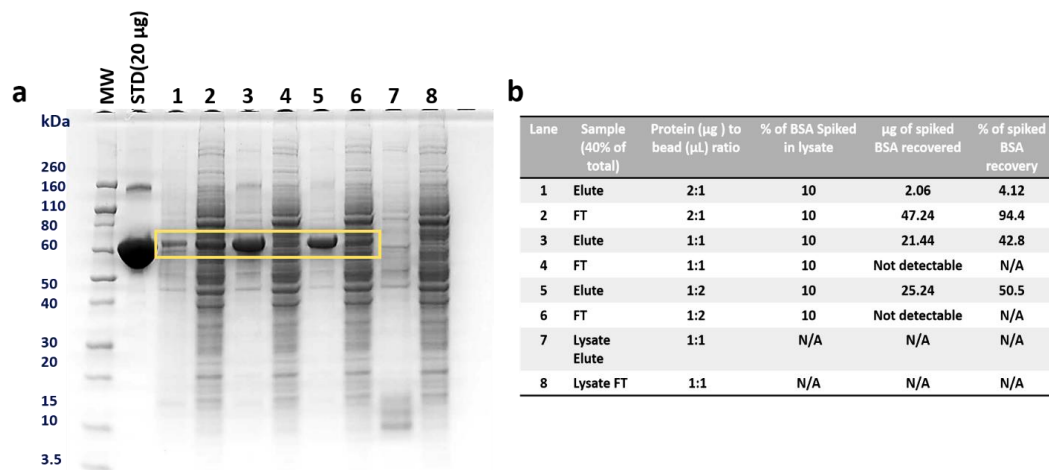
Fig: 3.7 (a) SDS PAGE analysis of effect on capturing spiked biotinylated BSA compared with different protein (µg)–bead (µL) ratio. (b) Percentage of BSA recovered in elute and FT calculated by relative densitometry with corresponding protein–bead ratio.

Lane	Sample (40% of total)	Protein (µg) to bead (µL) ratio	% of BSA Spiked in lysate	µg of spiked BSA recovered	% of spiked BSA recovery
1	Elute	2:1	10	2.06	4.12
2	FT	2:1	10	47.24	94.4
3	Elute	1:1	10	21.44	42.8
4	FT	1:1	10	Not detectable	N/A
5	Elute	1:2	10	25.24	50.5
6	FT	1:2	10	Not detectable	N/A
7	Lysate Elute	1:1	N/A	N/A	N/A
8	Lysate FT	1:1	N/A	N/A	N/A

To evaluate recovery limits, samples with different percentages of spiked BSA in cell lysates and normal buffer (PBS) were compared. A range of 10–50 µg of biotinylated BSA (2–10%) was spiked into either 500 µL cell lysate (1 mg/mL) or 500 µL PBS for incubation with a fixed amount (50 µL) of beads. The protein (µg)–bead (µL) bead ratio was in the range of 1:5 to 1:1. The captured proteins were eluted by incubating at our optimized elution condition and buffer (1% SDS and 50 mM DTT, at 90 ºC for 30 min) and subjected to SDS PAGE analysis.

The recovered BSA in the eluted fraction in both of the gels confirmed the efficiency of the optimized elution condition for both matrices (Fig. 3.8 (a) (b)), biotinylated BSA could still be recovered when spiked at 2%, or 20 µg/mL BSA concentration, although 100% recovery of the spiked proteins remained elusive. However, it appears that nonspecific binding from lysate proteins to the beads would become competitive if labeled protein is less than 1 µg/mL in concentration. For elute with 10 and 25 µg of spiked protein (20–50 µg/mL), clear recovery of BSA was possible in PBS but not as clear in the lysate (Fig. 3.8 (a) lane- 5, 7).

a. Gel-1

Gel-1 Lane	Samples (40% of total)	% of BSA Spiked in lysate
1	FT	10
2	Elute	10
3	Elute	7
4	FT	7
5	Elute	5
6	FT	5
7	Elute	2
8	FT	2
9	Lysate Elute	none
10	Lysate FT	none

b. Gel-2

Gel-2 Lane	Samples (40% of total)	% of BSA in PBS
1	Elute	10
2	FT	10
3	Elute	7
4	FT	7
5	Elute	5
6	FT	5
7	Elute	2
8	FT	2
9	Lysate Elute	none
10	Lysate FT	none
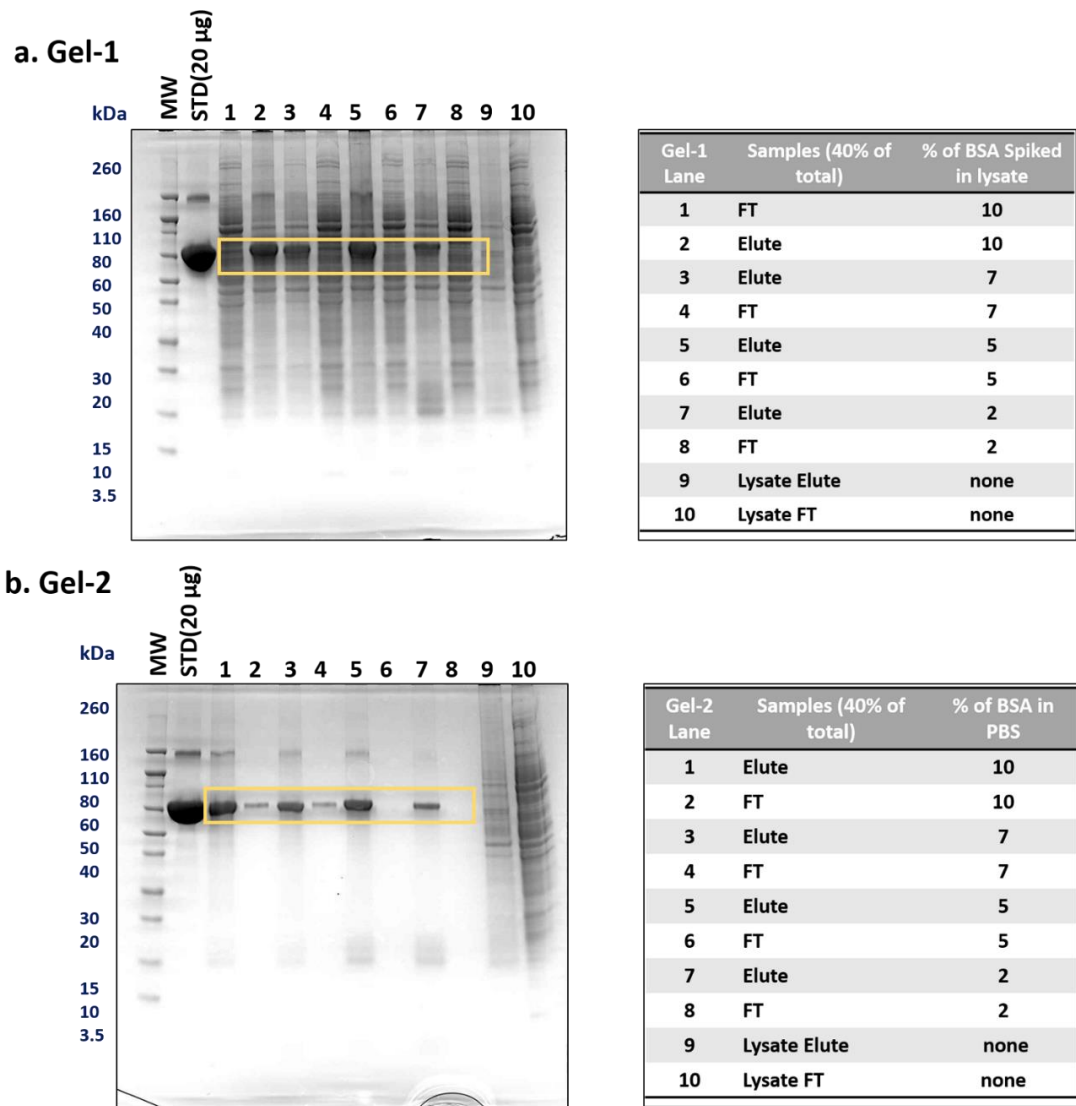
Fig. 3.8 SDS PAGE of matrix effect comparison with spiked BSA in cell lysates. (a) Gel-1- recovered BSA from 10–2% spiked biotinylated BSA in cell lysate. (b) Gel-2- recovered BSA from 10–2% biotinylated BSA in PBS.

3.3.   Surface biotinylation and membrane protein isolation from HCT-116 cells

3.3.1. Surface biotinylation of plasma membrane protein

Cells were biotinylated for 15 min at 4º C as described in the section 3.2 in triplicates (~2.5×10^7 cells/plate) with a slight modifications in the washing steps. Cell loss upon biotin treatment and washing steps was minimized by very slow rocking in a platform rocker during biotinylation following a gentle wash with TBS on plate. Subsequent washes were carried out after harvesting the cells by scrapping. The control cells (no probes used) were similarly washed with PBS. The optimized Dounce homogenization method provided lysates with good protein yields with minimal cell loss (biotinylated, 2–2.4 mg/mL and untreated control lysate, 2.7 mg/mL) (data not shown).

The protein (µg)–bead (µL) ratio investigated was 1:1 or 1:2, using 800 µL of biotinylated or control/untreated cell lysate (1 mg/mL concentration) with 80 or 160 µL of dry/settled NeutrAvidin beads. This was with the assumption that around 10% of the total proteins in lysate could be biotinylated. The captured proteins were eluted with the optimized elution condition (1% SDS and 50 mM DTT, at 90 ºC, for 30 min). A small fraction (10 µL) of each of the eluted PM fractions was subjected to SDS-PAGE analysis (Fig 3.9). The total protein content in the PM fractions was estimated from relative densitometry of the gel bands.

Densitometry data indicated that incubating in 1:2 protein (µg)–bead (µL) ratio (Fig 3.9 (b)) captured almost 1.5 times more protein than that from 1:1 protein (µg)–bead (µL) ratio (Fig 3.9 (a)). However, an increase in bead amount also increased the elution volume yet kept the protein concentration higher than the 1:1 protein (µg)–bead (µL) ratio. Although a moderate avidin elution contamination is prominent in all the elute lanes (Fig 3.9 (a) (b)), considering the total amount of protein recovery, the 1:2 protein (µg)–bead (µL) ratio appeared to be most suitable for preparation for proteomic analysis.
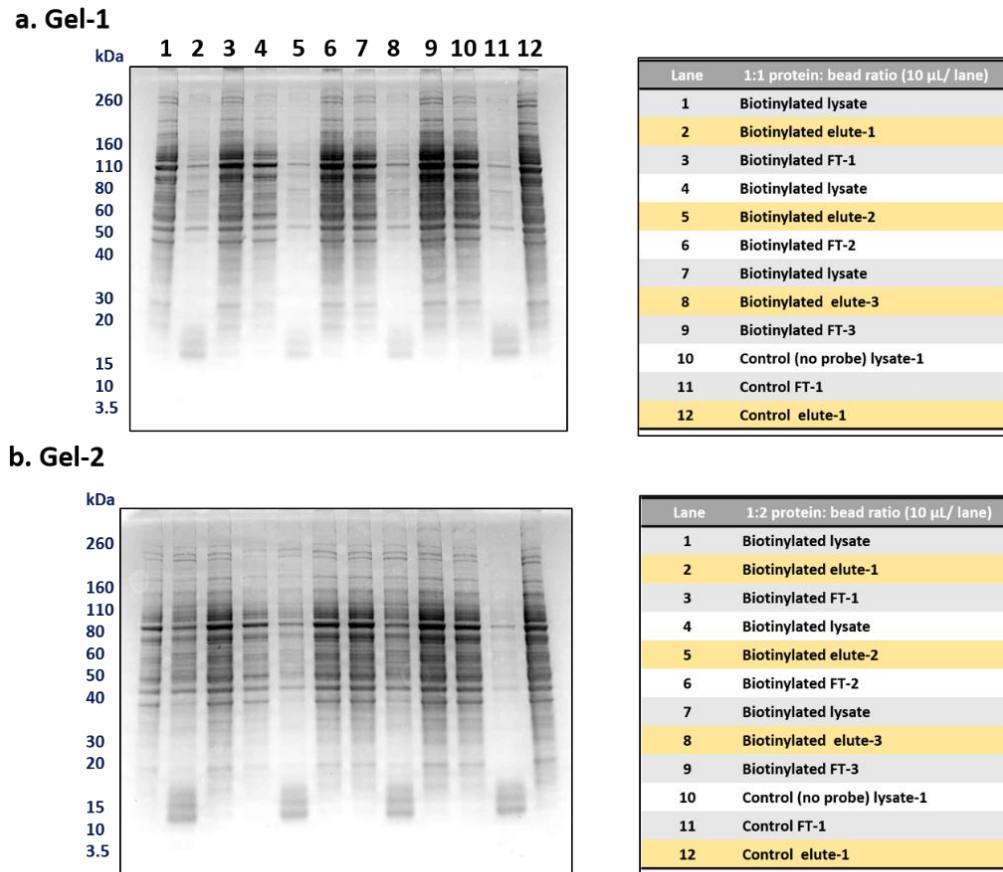
Fig. 3.9 SDS PAGE analysis of two different protein (µg)–bead (µL) ratio compared for protein recovery in elute. (a) Gel-1: protein capture and recovery in elute compared with 1:1 protein (µg)–bead (µL) ratio (b) protein capture and recovery in elute compared with 1:2 protein (µg)–bead (µL) ratio.

3.3.2. Western blot analysis of protein enrichment

The elute fractions were analyzed by western blot analysis prior to LC/MS/MS analysis. Along with the lysate and FT fractions, ~ 7 µg of proteins eluted were subjected to western blot analysis with EGFR and claudin-4 antibodies. Along with MDA-MB-468 breast cancer cell lysate was used as a positive control for EGFR, and the Caco-2 gastric adenocarcinoma cell lysate for claudin-4.
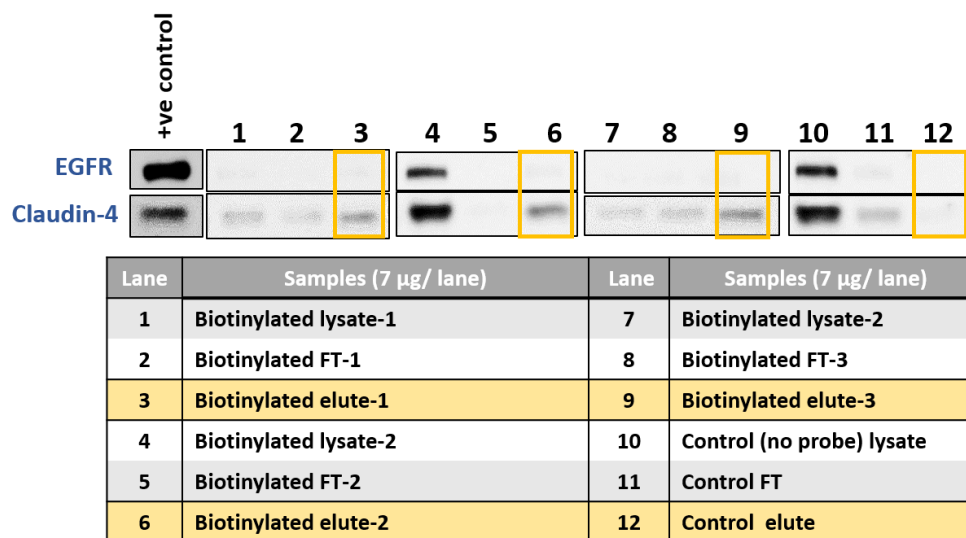
Fig. 3.10. Western blot analysis of the three fractions (lysate, FT and elute) from triplicated biotinylated and no probe control samples. Elute lanes are highlighted using orange boxes

The presence of claudin-4 in all the biotin elute fractions (Fig 3.10 lane 3, 6 and 9) confirmed us efficiency of the biotinylation method for tagging claudins. The absence of claudin-4 in the elute fraction of the control with no probes is consistent with the presence of claudin-4 in the FT fraction (Fig. 3.10 lane 12). Considering the fact that only about 15% of elute were loaded for western blot analysis and diluted condition of this elute, a lower band intensity than the previous data (section 3.2, Fig 3.4) was expected. Although, the missing bands in lysate lanes of biotinylated lysate-1 and 3 (Fig 3.10 lane 1 and 7) was not consistent with the lysate bands in biotinylated lysate-2 and no probe control lysate (Fig 3.10 lane 4 and 10). A possible cause of this loading variation could be over estimation of protein quantities in the lysate by the BCA protein quantification method. Other methods such as Bradford's method could be used to in the future to verify the protein quantities and resolve potential technical issues.

The issue of EGFR protein not being observed in any of the elute fractions (Fig. 3.10 lane 3, 6 and 9) is inconsistent with previous observations (Fig. 3.4 lane 1, 2 (top)). This in part confirms the potential issue of protein quantification as mentioned earlier. Also, it is possible that heavy glycosylation of surface EGFR in colorectal carcinomas could impede the surface labeling of this protein [82]. As we have observed in previous sections (section 3.1 and 3.2), significant quantities

of EGFR remained in the cell pellet after extraction, indicating much less enrichment for this protein, compared to claudins, by the surface biotinylation approach.

3.4. Proteomic analysis of biotinylated enriched fraction

The plasma membrane, cell surface, junctional and extracellular matrix proteins are the surface exposed proteins available for biotinylation. As our biotinylated cell surface proteins contains 1% SDS and 50 mM DTT after elution, removal of detergent was necessary prior to analysis by LC/MS/MS. For all of the biotinylated elute fractions, a gel-assisted digestion approach was selected for detergent removal. Two digestion methods, tube gel digestion and SDS-PAGE in-gel digestion, were commonly reported in membrane proteomic analysis. The tube gel digestion method originally reported by Lu, X. and Zhu, H. for high throughput analysis of membrane proteins, remove various types of detergents without performing any electrophoresis [73]. Recently, Smolders, K. *et al.* applied this method for proteomic analysis of biotinylated membrane proteins of mouse brain tissues [71]. On the contrary, traditional SDS-PAGE method removes detergent by electrophoresis but can result in substantial peptide loss.

As discussed below, the effectiveness of our biotinylation and enrichment method was evaluated in terms of the reproducibility and percentage of membrane proteins from the elute fraction as compared to that from the whole cell lysate. The proteomic data was also compared with existing literature for an evaluation of the reported workflow.

3.4.1. Comparative membrane proteomic coverage in untreated cell lysate and biotinylated elute

Proteomic analysis of the eluted fraction (~2.5×10^7 cells in each) identified a range of 126–277 protein groups for each of the biological triplicate. After removal of false positives (zero intensity) and duplicate protein IDs from the control (no probe) elute fractions, a total of 389 unique protein groups were identified from triplicated biotinylated elute fraction. Gene ontology cellular component (GO.CC) analysis annotated the subcellular distribution for all 389 biotin

enriched proteins. As a comparison, triplicated cell lysates (~2.5× 10^7 cells in each) were also evaluated by LC/MS/MS to reveal a total of 604 protein groups. A total of 742 proteins were reported by a previous study of HCT 116 membrane proteomics with biotinylation, in which 200–300 proteins were found from samples each containing 10^8 HCT-116 cells [77]. In this study, despite of using at least four times less cells (2.5× 10^7), the number of proteins identified was not significantly lower (126–277).

Table. 3.1 Subcellular distribution of merged proteins identified in ≥2 biotinylated elute and lysate samples.

Subcellular distribution	Merged biotin elute	Merged cell lysate
Plasma membrane	19	41
Extracellular Space	26	52
Cytosol	9	27
Nucleus	4	53
Total	58	173
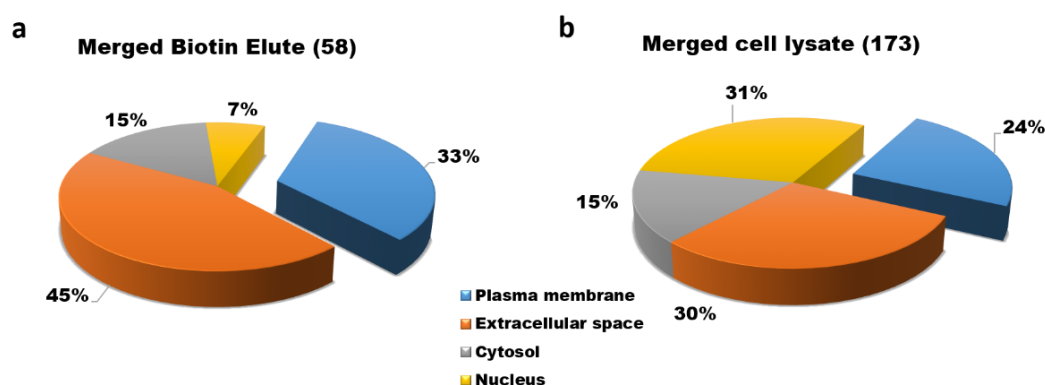
Fig 3.11 Relative subpopulation distribution of the proteins identified in ≥2 samples in biotinylated elute and cell lysate. (a) Proteins identified in ≥2 biotinylated elute samples (b) Proteins identified in ≥2 untreated cell lysate samples.

To get a better comparative view, two protein lists, one from biotinylated elute fractions (biological triplicate) and the other whole cell lysate samples, were analysed for protein identification. Out of the 389 proteins, 58 proteins (Supplementary Table 1), identified in at least

two samples out of the biological triplicate, were taken into account for this analysis. On the other hand, 173 proteins (Supplementary Table 2) in the whole cell lysate list were identified from two or more lysates. The subcellular distribution data of the two lists showed significant differences in terms of protein identified from various subcellular locations (Table 3.1, Fig. 3.11). However, none of the targeted claudin family members was detected, even though detection by western blot analysis was observed (section 3.3.2, Fig. 3.10). A relatively higher percentage of plasma membrane proteins is seen in the elute fraction. Abundant and secreted in the extracellular space, many of the extracellular matrix proteins can be available for biotin labeling. The minimal percentage of nuclear proteins along with higher percentage plasma membrane proteins is an indicator of effective labeling as well as enrichment method.

A total of 27 proteins were found in common from the two lists shown in Figure 3.11a and 3.11b. Among this common set of proteins, about 60% were found to be distributed in extracellular space as well as cytoplasm. Some of the nuclear and cytosolic proteins can be secreted into the extracellular space and cell surface to be biotinylated. Commonly identified proteins, such as members of the heat shock protein family, namely HSP 90-beta, HSP 90-alfa, HSP 60 and HSP 27, could be examples of that. These proteins are abundant cellular stress specific marker proteins that can have cytosolic and cell surface localizations, and therefore can be detected in both the lysate and enriched fractions. Two proteins, namely annexin A2 and focal adhesion protein copine-3, are also commonly identified in the elute fraction and whole cell lysate samples. Membrane adhesion proteins such as annexin A2 are well known as a metastatic marker of colorectal cancer and abundant in colorectal cancer cells[83]. Among the 148 unique proteins in the cell lysate list, 34 were plasma membrane proteins, including 8 junction-localized proteins (Supplementary 2). Detection of junctional proteins in the lysate is an indicator of effective junctional complex preservation during cell lysis. Protein loss due to tube-gel digestion method could be a reason for not observing these junctional proteins in the biotin elute fraction.

Table. 3.2 List of proteins identified from biotinylated elute fraction previously reported by Nagano's study [77].

Entry no	Protein id	Protein name	Subcellular distribution
1	P05023	Sodium/potassium-transporting ATPase subunit alpha-1	Plasma membrane
2	P05141	ADP/ATP translocase 2	Plasma Membrane
3	P43243	Matrin-3	Nucleus
4	Q99832	T-complex protein 1 subunit eta	Cytosol
5	P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	Nuclear/ Extracellular space
6	P16401	Histone H1.5	Nuclear/ Extracellular space
7	P62805	Histone H4	Nuclear/ Extracellular space
8	P62906	60S ribosomal protein L10a	Nuclear/ Extracellular Exosome
9	P48643	T-complex protein 1 subunit epsilon	Cytosol/ Extracellular Exosome
10	P09651	Heterogeneous nuclear ribonucleoprotein A1	Nuclear/ Extracellular space
11	P10412	Histone H1.4	Nuclear/ Extracellular space
12	P62857	40S ribosomal protein S28	Cytosol/ Extracellular space

A total of 33 unique proteins were common to the lysate and elute samples, with nine proteins annotated as plasma membrane proteins. A total of 12 proteins from the list (Table 3.2) were previously reported by Nagano's study that also utilized surface biotinylation of HCT 116 cells [77]. The remaining 18 proteins from the list (Table 3.3) of 33 unique proteins were not previously reported by them [77]. For example, brain-specific angiogenesis inhibitor 1 (BAI1), a seven transmembrane domain containing G protein coupled junctional complex protein, was identified. Protein BAI1 was reported to have lower expression in colorectal cells [84]. Although a number of cytosolic and nuclear proteins were in the list, several integral components of plasma membrane proteins, such as mitochondrial 2-oxoglutarate/malate carrier protein (OGCP) and scavenger receptor class F member 2 (SRECRP-1), were also identified.

Table. 3.3. List of proteins identified from biotinylated elute fraction, previously not reported by Nagano's study [77].

Entry	Protein ID	Protein name	Subcellular distribution
1	O14514	Brain-specific angiogenesis inhibitor 1	Cell-cell junction
2	Q02978	Mitochondrial 2-oxoglutarate/malate carrier protein	Plasma membrane
3	O14745	Na(+)/H(+) exchange regulatory cofactor 1	Plasma membrane
4	Q96GP6	Scavenger receptor class F member 2	Plasma membrane
5	P0C7V7	Putative signal peptidase complex catalytic subunit SEC11B	Membrane
6	P02545	Prelamin-A/C	Cytosol/Extracellular matrix
7	P08670	Vimentin	Cytoskeleton/
8	P0DN76	Splicing factor U2AF 35 kDa subunit-like protein	Nucleus
9	Q9UI15	Transgelin-3	Nucleus
10	Q8TEX9	Importin-4	Nucleus
11	Q8NCF5	NFATC2-interacting protein	Nucleus
12	Q8NHX4	Spermatogenesis-associated protein 3	unknown
13	Q9NR09	Baculoviral IAP repeat-containing protein 6	Cytosolic
14	Q13151	Heterogeneous nuclear ribonucleoprotein A0	Nucleus
15	Q9NQT5	Exosome complex component RRP40	Cytosolic
16	Q13428	Treacle protein	Cytosolic
17	O76003	Glutaredoxin-3	Cytosolic
18	P11586	C-1-tetrahydrofolate synthase	Cytosolic

In addition to use of non-cleavable sulfo-NHS-LC-biotin for labeling in the previous study [77], a combination of high speed membrane protein sedimentation, peptide pull-down by monomeric avidin, and in-solution digestion was used, containing significant points of differentiation from the study here. Our biotinylation method, using cleavable sulfo-NHS-SS-biotin probe and protein pull down by NeutrAvidin beads, did not provide the expected ease of sample manipulation, due to the fact the chemical cleavage did not proceed nearly as readily as claimed by the supplier. In future the chemical cleavage can be improved by using possibly a different immobilization surface or changing to a different probe that can be cleaved by light. This will increase enrichment efficiency and reduce the level of background as well.

The total number of proteins identified in each biotinylated elute fractions showed high variability in terms of number of proteins identified. Several technical factors possibly contributed to this lower reproducibility of protein identification form the biotin elute fractions. Instrumental errors occurred during the LC/MS/MS analysis that resulted in gaps in the MS spectra of two of the three samples. Thus, some peptides from the two elute fractions were missing from the identification process, which in turn resulted in data variation. The tube gel digestion method, followed for eluted sample processing, requires rigorous cleaning and washing steps to remove detergent. Along with that, a laborious peptide extraction due to bigger gel volume, may have possibly caused some peptide loss. An alternative of gel based detergent removal could minimize the sample processing step and can boost the number of more membrane proteins identified. The filter-aided digestion method reported [85] can be tested for reproducible sample preparation step.

4.   Conclusion and future direction

Extraction and purification of endogenous claudin family of proteins are challenging. In the current study, methods were investigated for establishing chemical labelling and affinity enrichment protocols for claudin isolation and detection by mass spectrometry using HCT-116 cells. Important steps such as cell lysis methods, labelling, capture and elution conditions were optimized to increase cell surface protein as well as claudin capture and recovery for proteomic analysis. An optimized cell surface labeling and enrichment method was applied to triplicated HCT-116 cells, and moderate amount of enrichment of claudin-4 was confirmed by western blot analysis. Analysis of the biotinylated elute fraction (~2.5× 10^7 cells in each) by LC/MS/MS identified 126–277 protein groups for each of the biological triplicate. Comparative analysis of proteomic data showed relatively higher percentage of plasma membrane proteins in the biotinylated elute fraction. Reproducibility issues were observed due to instrumental error and peptide loss upon gel assisted digestion. Due to the time limitation on this training project, these technical issues will be addressed at a later time.

The optimized workflow has showed a promising result for claudin isolation in the elute fraction. In future, an alternative immobilization surface or changing to a photo-cleavable chemical probe will be considered for improving elution efficiency, reducing elute fraction complexity, and streamlining sample preparation for mass spectrometry. Further improvement of MS sample preparation will be sought with an alternative filter aided in-solution digestion method. This chemical-capture module with then be combined with optimized membrane partition protocols to establish an advanced workflow for endogenous claudin analysis by mass spectrometry. This new technological platform in hand, will expedite the understanding of claudin biology and pathology.

5.   References

1.   Furuse, M.; Fujita, K.; Hiiragi, T.; Fujimoto, K.; Tsukita, S., Claudin-1 and -2: Novel integral membrane proteins localizing at tight junctions with no sequence similarity to occludin. *J. Cell Biol.* **1998,** *141* (7), 1539-1550.
2.   Farquhar, M. G.; Palade, G. E., Junctional complexes in various epithelia. *J. Cell Biol.* **1963,** *17* (2), 375-412.
3.   Balda, M. S.; Matter, K., Tight junctions as regulators of tissue remodelling. *Curr. Opin. Cell Biol.* **2016,** *42*, 94-101.
4.   Furuse, M., Molecular basis of the core structure of tight junctions. *Cold Spring Harbor perspectives in biology* **2010,** *2* (1), a002907.
5.   Furuse, M.; Sasaki, H.; Fujimoto, K.; Tsukita, S., A single gene product, claudin-1 or -2, reconstitutes tight junction strands and recruits occludin in fibroblasts. *J. Cell Biol.* **1998,** *143* (2), 391-401.
6.   Furuse, M.; Sasaki, H.; Tsukita, S., Manner of interaction of heterogeneous claudin species within and between tight junction strands. *J. Cell Biol.* **1999,** *147* (4), 891-903.
7.   Van Itallie, C. M.; Anderson, J. M., The role of claudins in determining paracellular charge selectivity. *Proc. Am. Thorac. Soc.* **2004,** *1* (1), 38-41.
8.   Furuse, M.; Hata, M.; Furuse, K.; Yoshida, Y.; Haratake, A.; Sugitani, Y.; Noda, T.; Kubo, A.; Tsukita, S., Claudin-based tight junctions are crucial for the mammalian epidermal barrier. *J. Cell Biol.* **2002,** *156* (6), 1099-1111.
9.   Van Itallie, C.; Rahner, C.; Anderson, J. M., Regulated expression of claudin-4 decreases paracellular conductance through a selective decrease in sodium permeability. *J. Clin. Invest.* **2001,** *107* (10), 1319-1327.
10.   Angelow, S.; Kim, K. J.; Yu, A. S., Claudin-8 modulates paracellular permeability to acidic and basic ions in MDCK II cells. *J. Physiol.* **2006,** *571* (1), 15-26.
11.   Colegio, O. R.; Van Itallie, C. M.; McCrea, H. J.; Rahner, C.; Anderson, J. M., Claudins create charge-selective channels in the paracellular pathway between epithelial cells. *Am. J. Physiol.: Cell Physiol.* **2002,** *283* (1), C142-C147.
12.   Van Itallie, C. M.; Fanning, A. S.; Anderson, J. M., Reversal of charge selectivity in cation or anion-selective epithelial lines by expression of different claudins. *Am. J. Physiol.: Renal Physiol.* **2003,** *285* (6), F1078-F1084.
13.   Amasheh, S.; Meiri, N.; Gitter, A. H.; Schöneberg, T.; Mankertz, J.; Schulzke, J. D.; Fromm, M., Claudin-2 expression induces cation-selective channels in tight junctions of epithelial cells. *J. Cell Sci.* **2002,** *115* (24), 4969-4976.
14.   Chen, S. P.; Zhou, B.; Willis, B. C.; Sandoval, A. J.; Liebler, J. M.; Kim, K.-J.; Ann, D. K.; Crandall, E. D.; Borok, Z., Effects of transdifferentiation and EGF on claudin isoform expression in alveolar epithelial cells. *J. Appl. Physiol.* **2005,** *98* (1), 322-328.
15.   Singh, A. B.; Harris, R. C., Epidermal growth factor receptor activation differentially regulates claudin expression and enhances transepithelial resistance in Madin-Darby canine kidney cells. *J. Biol. Chem.* **2004,** *279* (5), 3543-3552.
16.   Baltzegar, D. A.; Reading, B. J.; Brune, E. S.; Borski, R. J., Phylogenetic revision of the claudin gene family. *Mar. Genomics* **2013,** *11*, 17-26.
17.   Morita, K.; Furuse, M.; Fujimoto, K.; Tsukita, S., Claudin multigene family encoding four-transmembrane domain protein components of tight junction strands. *Proc. Natl. Acad. Sci.* **1999,** *96* (2), 511-516.
18.   Krause, G.; Winkler, L.; Mueller, S. L.; Haseloff, R. F.; Piontek, J.; Blasig, I. E., Structure and function of claudins. *Biochim. Biophys. Acta, Biomembr.* **2008,** *1778* (3), 631-645.

19.   Anderson, J. M.; Van Itallie, C. M., Physiology and function of the tight junction. *Cold Spring Harbor perspectives in biology* **2009,** *1* (2), a002584.

20.   Van Itallie, C. M.; Anderson, J. M., Claudins and epithelial paracellular transport. *Annu. Rev. Physiol.* **2006,** *68*, 403-429.

21.   Turksen, K.; Troy, T.-C., Barriers built on claudins. *J. Cell Sci.* **2004,** *117* (12), 2435-2447.

22.   Suzuki, H.; Nishizawa, T.; Tani, K.; Yamazaki, Y.; Tamura, A.; Ishitani, R.; Dohmae, N.; Tsukita, S.; Nureki, O.; Fujiyoshi, Y., Crystal structure of a claudin provides insight into the architecture of tight junctions. *Science* **2014,** *344* (6181), 304-307.

23.   Saitoh, Y.; Suzuki, H.; Tani, K.; Nishikawa, K.; Irie, K.; Ogura, Y.; Tamura, A.; Tsukita, S.; Fujiyoshi, Y., Structural insight into tight junction disassembly by Clostridium perfringens enterotoxin. *Science* **2015,** *347* (6223), 775-778.

24.   Shinoda, T.; Shinya, N.; Ito, K.; Ohsawa, N.; Terada, T.; Hirata, K.; Kawano, Y.; Yamamoto, M.; Kimura-Someya, T.; Yokoyama, S.; Shirouzu, M., Structural basis for disruption of claudin assembly in tight junctions by an enterotoxin. *Sci. Rep.* **2016,** *6*, 33632.

25.   Hewitt, K. J.; Agarwal, R.; Morin, P. J., The claudin gene family: expression in normal and neoplastic tissues. *BMC cancer* **2006,** *6* (1), 186.

26.   Irudayanathan, F. J.; Trasatti, J. P.; Karande, P.; Nangia, S., Molecular architecture of the blood brain barrier tight junction proteins–a synergistic computational and in vitro approach. *J. Phys. Chem. B* **2015,** *120* (1), 77-88.

27.   Tsukita, S.; Furuse, M.; Itoh, M., Multifunctional strands in tight junctions. *Nat. Rev. Mol. Cell Biol.* **2001,** *2* (4), 285-293.

28.   Matsuda, M.; Kubo, A.; Furuse, M.; Tsukita, S., A peculiar internalization of claudins, tight junction-specific adhesion molecules, during the intercellular movement of epithelial cells. *J. Cell Sci.* **2004,** *117* (7), 1247-1257.

29.   Winkler, L.; Gehring, C.; Wenzel, A.; Müller, S. L.; Piehl, C.; Krause, G.; Blasig, I. E.; Piontek, J., Molecular determinants of the interaction between Clostridium perfringens enterotoxin fragments and claudin-3. *J. Biol. Chem.* **2009,** *284* (28), 18863-18872.

30.   Singh, A. B.; Uppada, S. B.; Dhawan, P., Claudin proteins, outside-in signaling, and carcinogenesis. *Pflugers Arch. - Eur. J. Physio.* **2017,** *469* (1), 69-75.

31.   Webb, P. G.; Spillman, M. A.; Baumgartner, H. K., Claudins play a role in normal and tumor cell motility. *BMC Cell Biol.* **2013,** *14* (1), 19.

32.   Blanchard, A. A.; Skliris, G. P.; Watson, P. H.; Murphy, L. C.; Penner, C.; Tomes, L.; Young, T. L.; Leygue, E.; Myal, Y., Claudins 1, 3, and 4 protein expression in ER negative breast cancer correlates with markers of the basal phenotype. *Virchows Archiv.* **2009,** *454* (6), 647-656.

33.   Tőkés, A.-M.; Kulka, J.; Paku, S.; Szik, Á.; Páska, C.; Novák, P. K.; Szilák, L.; Kiss, A.; Bögi, K.; Schaff, Z., Claudin-1,-3 and-4 proteins and mRNA expression in benign and malignant breast lesions: a research study. *Breast Cancer Res.* **2005,** *7* (2), R296.

34.   Swisshelm, K.; Macek, R.; Kubbies, M., Role of claudins in tumorigenesis. *Adv. Drug Delivery Rev.* **2005,** *57* (6), 919-928.

35.   Kominsky, S. L.; Argani, P.; Korz, D.; Evron, E.; Raman, V.; Garrett, E.; Rein, A.; Sauter, G.; Kallioniemi, O.-P.; Sukumar, S., Loss of the tight junction protein claudin-7 correlates with histological grade in both ductal carcinoma in situ and invasive ductal carcinoma of the breast. *Oncogene* **2003,** *22* (13), 2021-2033.

36.   Morohashi, S.; Kusumi, T.; Sato, F.; Odagiri, H.; Chiba, H.; Yoshihara, S.; Hakamada, K.; Sasaki, M.; Kijima, H., Decreased expression of claudin-1 correlates with recurrence status in breast cancer. *Int. J. Mol. Med.* **2007,** *20* (2), 139-144.

37.   Hoevel, T.; Macek, R.; Swisshelm, K.; Kubbies, M., Reexpression of the TJ protein CLDN1 induces apoptosis in breast tumor spheroids. *Int. J. Cancer* **2004,** *108* (3), 374-383.

38.   Dhawan, P.; Ahmad, R.; Chaturvedi, R.; Smith, J. J.; Midha, R.; Mittal, M. K.; Krishnan, M.; Chen, X.; Eschrich, S.; Yeatman, T. J.; Harris, R. C.; Washington, M. K.; Wilson, K. T.; Beauchamp, R. D.; Singh, A. B., Claudin-2 expression increases tumorigenicity of colon cancer cells: role of epidermal growth factor receptor activation. *Oncogene* **2011,** *30* (29), 3234-3247.

39.   Tabariès, S.; Siegel, P., The role of claudins in cancer metastasis. *Oncogene* **2016**.

40.   Dhawan, P.; Singh, A. B.; Deane, N. G.; No, Y.; Shiou, S.-R.; Schmidt, C.; Neff, J.; Washington, M. K.; Beauchamp, R. D., Claudin-1 regulates cellular transformation and metastatic behavior in colon cancer. *J. Clin. Invest.* **2005,** *115* (7), 1765-1776.

41.   Escaffit, F.; Boudreau, F.; Beaulieu, J. F., Differential expression of claudin-2 along the human intestine: Implication of GATA-4 in the maintenance of claudin-2 in differentiating cells. *J. Cell. Physiol.* **2005,** *203* (1), 15-26.

42.   Hahn-Strömberg, V.; Askari, S.; Ahmad, A.; Befekadu, R.; Nilsson, T. K., Expression of claudin 1, claudin 4, and claudin 7 in colorectal cancer and its relation with CLDN DNA methylation patterns. *Tumor Biol.* **2017,** *39* (4), 1010428317697569.

43.   Tsukita, S.; Yamazaki, Y.; Katsuno, T.; Tamura, A., Tight junction-based epithelial microenvironment and cell proliferation. *Oncogene* **2008,** *27* (55), 6930-6938.

44.   Liu, F.; Koval, M.; Ranganathan, S.; Fanayan, S.; Hancock, W. S.; Lundberg, E. K.; Beavis, R. C.; Lane, L.; Duek, P.; McQuade, L., Systems Proteomics View of the Endogenous Human Claudin Protein Family. *J. Proteome Res.* **2016,** *15* (2), 339-359.

45.   Mitic, L. L.; Unger, V. M.; Anderson, J. M., Expression, solubilization, and biochemical characterization of the tight junction transmembrane protein claudin-4. *Protein Sci.* **2003,** *12* (2), 218-227.

46.   Lohrberg, D.; Krause, E.; Schümann, M.; Piontek, J.; Winkler, L.; Blasig, I. E.; Haseloff, R. F., A strategy for enrichment of claudins based on their affinity to Clostridium perfringens enterotoxin. *BMC Mol. Biol.* **2009,** *10* (1), 61.

47.   Katahira, J.; Inoue, N.; Horiguchi, Y.; Matsuda, M.; Sugimoto, N., Molecular cloning and functional characterization of the receptor for Clostridium perfringens enterotoxin. *J. Cell Biol.* **1997,** *136* (6), 1239-1247.

48.   Tang, V. W., Proteomic and bioinformatic analysis of epithelial tight junction reveals an unexpected cluster of synaptic molecules. *Biol Direct* **2006,** *1*.

49.   Luissint, A.-C.; Federici, C.; Guillonneau, F.; Chrétien, F.; Camoin, L.; Ganeshamoorthy, K.; Couraud, P.-O., Guanine nucleotide-binding protein Gαi2: a new partner of claudin-5 that regulates tight junction integrity in human brain endothelial cells. *J. Cereb. Blood Flow Metab.* **2012,** *32* (5), 860-873.

50.   Huber, L. A.; Pfaller, K.; Vietor, I., Organelle proteomics. *Circ. Res.* **2003,** *92* (9), 962-968.

51.   Lund, R.; Leth-Larsen, R.; Jensen, O. N.; Ditzel, H. J., Efficient isolation and quantitative proteomic analysis of cancer cell plasma membrane proteins for identification of metastasis-associated cell surface markers. *J. Proteome Res.* **2009,** *8* (6), 3078-3090.

52.   Mannová, P.; Fang, R.; Wang, H.; Deng, B.; McIntosh, M. W.; Hanash, S. M.; Beretta, L., Modification of host lipid raft proteome upon hepatitis C virus replication. *Mol. Cell. Proteomics* **2006,** *5* (12), 2319-2325.

53.   Patwardhan, A. J.; Strittmatter, E. F.; Camp, D. G.; Smith, R. D.; Pallavicini, M. G., Comparison of normal and breast cancer cell lines using proteome, genome, and interactome data. *J. Proteome Res.* **2005,** *4* (6), 1952-1960.

54.   Elschenbroich, S.; Kim, Y.; Medin, J. A.; Kislinger, T., Isolation of cell surface proteins for mass spectrometry-based proteomics. *Expert Rev. Proteomics* **2010,** *7* (1), 141-154.

55.   Cox, B.; Emili, A., Tissue subcellular fractionation and protein extraction for use in mass-spectrometry-based proteomics. *Nat. Protoc.* **2006,** *1* (4), 1872-1878.

56.   Sugibayashi, K.; Onuki, Y.; Takayama, K., Displacement of tight junction proteins from detergent-resistant membrane domains by treatment with sodium caprate. *Eur. J. Pharm. Sci.* **2009,** *36* (2), 246-253.

57.   Lynch, R. D.; Francis, S. A.; McCarthy, K. M.; Casas, E.; Thiele, C.; Schneeberger, E. E., Cholesterol depletion alters detergent-specific solubility profiles of selected tight junction proteins and the phosphorylation of occludin. *Exp. Cell Res.* **2007,** *313* (12), 2597-2610.

58.   Schuck, S.; Honsho, M.; Ekroos, K.; Shevchenko, A.; Simons, K., Resistance of cell membranes to different detergents. *Proc. Natl. Acad. Sci.* **2003,** *100* (10), 5795-5800.

59.   McQuade, L. R.; Schmidt, U.; Pascovici, D.; Stojanov, T.; Baker, M. S., Improved Membrane Proteomics Coverage of Human Embryonic Stem Cells by Peptide IPG-IEF. *J. Proteome Res.* **2009,** *8* (12), 5642-5649.

60.   Schulz, T. C.; Swistowska, A. M.; Liu, Y.; Swistowski, A.; Palmarini, G.; Brimble, S. N.; Sherrer, E.; Robins, A. J.; Rao, M. S.; Zeng, X., A large-scale proteomic analysis of human embryonic stem cells. *BMC genomics* **2007,** *8* (1), 478.

61.   Leitner, A.; Lindner, W., Chemistry meets proteomics: The use of chemical tagging reactions for MS-based proteomics. *Proteomics* **2006,** *6* (20), 5418-5434.

62.   Jing, C.; Cornish, V. W., Chemical tags for labeling proteins inside living cells. *Acc. Chem. Res.* **2011,** *44* (9), 784.

63.   Heal, W. P.; Dang, T. T.; Tate, E. W., Activity-based probes: discovering new biology and new drug targets. *Chem. Soc. Rev.* **2011,** *40* (1), 246-257.

64.   Hayashi, T.; Hamachi, I., Traceless affinity labeling of endogenous proteins for functional analysis in living cells. *Acc. Chem. Res.* **2012,** *45* (9), 1460-1469.

65.   Giron, P.; Dayon, L.; Sanchez, J. C., Cysteine tagging for MS-based proteomics. *Mass Spectrom. Rev.* **2011,** *30* (3), 366-395.

66.   Kovalenko, O. V.; Yang, X. H.; Hemler, M. E., A novel cysteine cross-linking method reveals a direct association between claudin-1 and tetraspanin CD9. *Mol. Cell. Proteomics* **2007,** *6* (11), 1855-1867.

67.   Nitta, T.; Hata, M.; Gotoh, S.; Seo, Y.; Sasaki, H.; Hashimoto, N.; Furuse, M.; Tsukita, S., Size-selective loosening of the blood-brain barrier in claudin-5–deficient mice. *J. Cell Biol.* **2003,** *161* (3), 653-660.

68.   Ding, L.; Zhang, Y.; Tatum, R.; Chen, Y.-H., Detection of tight junction barrier function in vivo by biotin. *Claudins: Methods and Protocols* **2011**, 91-100.

69.   Weekes, M. P.; Antrobus, R.; R Lill, J.; Duncan, L. M.; Hör, S.; Lehner, P. J., Comparative analysis of techniques to purify plasma membrane proteins. *Journal of Biomolecular Techniques* **2010,** *21* (3).

70.   Vit, O.; Petrak, J., Integral membrane proteins in proteomics. How to break open the black box? *J. Proteomics* **2017,** *153*, 8-20.

71.   Smolders, K.; Lombaert, N.; Valkenborg, D.; Baggerman, G.; Arckens, L., An effective plasma membrane proteomics approach for small tissue samples. *Sci. Rep.* **2015,** *5*.

72.   Kischel, P.; Guillonneau, F.; Dumont, B.; Bellahcène, A.; Stresing, V.; Clézardin, P.; Pauw, E. A. D.; Castronovo, V., Cell Membrane Proteomic Analysis Identifies Proteins Differentially Expressed in Osteotropic Human Breast Cancer Cells. *Neoplasia* **2008,** *10* (9), 1014-IN11.

73.   Lu, X.; Zhu, H., Tube-Gel Digestion A Novel Proteomic Approach for High Throughput Analysis of Membrane Proteins. *Mol. Cell. Proteomics* **2005,** *4* (12), 1948-1958.

74.   Cox, J.; Mann, M., MaxQuant enables high peptide identification rates, individualized ppb-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **2008,** *26* (12), 1367-1372.

75.   Hörmann, K.; Stukalov, A.; Müller, A. C.; Heinz, L. X.; Superti-Furga, G.; Colinge, J.; Bennett, K. L., A surface biotinylation strategy for reproducible plasma membrane protein purification and tracking of genetic and drug-induced alterations. *J. Proteome Res.* **2016,** *15* (2), 647-658.

76.   Karhemo, P.-R.; Ravela, S.; Laakso, M.; Ritamo, I.; Tatti, O.; Mäkinen, S.; Goodison, S.; Stenman, U.-H.; Hölttä, E.; Hautaniemi, S., An optimized isolation of biotinylated cell surface proteins reveals novel players in cancer metastasis. *J. Proteomics* **2012,** *77*, 87-100.

77.   Nagano, K.; Shinkawa, T.; Kato, K.; Inomata, N.; Yabuki, N.; Haramura, M., Distinct cell surface proteome profiling by biotin labeling and glycoprotein capturing. *J. Proteomics* **2011,** *74* (10), 1985-1993.

78.   Henkhaus, R. S.; Roy, U. K. B.; Cavallo-Medved, D.; Sloane, B. F.; Gerner, E. W.; Ignatenko, N. A., Caveolin-1-Mediated Expression and Secretion of Kallikrein 6 in Colon Cancer Cells. *Neoplasia (New York, N.Y.)* **2008,** *10* (2), 140-148.

79.   Han, C.-L.; Chien, C.-W.; Chen, W.-C.; Chen, Y.-R.; Wu, C.-P.; Li, H.; Chen, Y.-J., A multiplexed quantitative strategy for membrane proteomics opportunities for mining therapeutic targets for autosomal dominant polycystic kidney disease. *Mol. Cell. Proteomics* **2008,** *7* (10), 1983-1997.

80.   Elia, G., Biotinylation reagents for the study of cell surface proteins. *Proteomics* **2008,** *8* (19), 4012-4024.

81.   Rybak, J. N.; Scheurer, S. B.; Neri, D.; Elia, G., Purification of biotinylated proteins on streptavidin resin: a protocol for quantitative elution. *Proteomics* **2004,** *4* (8), 2296-2299.

82.   Sethi, M. K.; Kim, H.; Park, C. K.; Baker, M. S.; Paik, Y.-K.; Packer, N. H.; Hancock, W. S.; Fanayan, S.; Thaysen-Andersen, M., In-depth N-glycome profiling of paired colorectal cancer and non-tumorigenic tissues reveals cancer-, stage-and EGFR-specific protein N-glycosylation. *Glycobiology* **2015,** *25* (10), 1064-1078.

83.   Xiu, D.; Liu, L.; Qiao, F.; Yang, H.; Cui, L.; Liu, G., Annexin A2 Coordinates STAT3 to Regulate the Invasion and Migration of Colorectal Cancer Cells In Vitro. *Gastroent. Res. Pract.* **2016,** *2016*, 10.

84.   Fukushima, Y.; Oshika, Y.; Tsuchida, T.; Tokunaga, T.; Hatanaka, H.; Kijima, H.; Yamazaki, H.; Ueyama, Y.; Tamaoki, N.; Nakamura, M., Brain-specific angiogenesis inhibitor 1 expression is inversely correlated with vascularity and distant metastasis of colorectal cancer. *Int. J. Oncol.* **1998,** *13* (5), 967-1037.

85.   Wiśniewski, J. R.; Zougman, A.; Nagaraj, N.; Mann, M., Universal sample preparation method for proteome analysis. *Nat. Methods* **2009,** *6* (5), 359-362.

Appendices

Supplementary Table-1. List of proteins found in ≥2 biotin elute fractions

Entry no	Protein ID	Protein names	Identified in samples
		Plasma Membrane Proteins	
1	P07355	Annexin A2	2
2	O14514	Brain-specific angiogenesis inhibitor 1	2
3	P05783	Keratin, type I cytoskeletal 18	3
4	P05141	ADP/ATP translocase 2	2
5	P09211	Glutathione S-transferase P	2
6	O75369	Filamin-B	2
7	P08865	40S ribosomal protein SA	2
8	P06733	Alpha-enolase	3
9	P07900	Heat shock protein HSP 90-alpha	3
10	P08670	Vimentin	3
11	P08238	Heat shock protein HSP 90-beta	3
12	P07237	Protein disulfide-isomerase (PDI)	3
13	P04406	Glyceraldehyde-3-phosphate dehydrogenase (GAPDH)	3
14	P04792	Heat shock protein beta-1	2
15	P35232	Prohibitin	2
16	Q02978	Mitochondrial 2-oxoglutarate	2
17	P05023	Sodium/potassium-transporting ATPase subunit alpha-1	2
18	O14745	Na(+)/H(+) exchange regulatory cofactor NHE-RF1 (NHERF-1)	2
		Nuclear Proteins	
19	P43243	Matrin-3	2
20	P0DN76	Splicing factor U2AF 35 kDa subunit-like protein (U2 small nuclear RNA	2
21	Q9UI15	Transgelin-3	2
22	Q13151	Heterogeneous nuclear ribonucleoprotein A0 (hnRNP A0)	2
		Cytoplasmic Proteins	
23	Q8TEX9	Importin-4 (Imp4)	3
24	Q8NCF5	NFATC2-interacting protein	2
25	Q8NHX4	Spermatogenesis-associated protein 3	2
26	P0C7V7	Putative signal peptidase complex catalytic subunit SEC11B	2
27	Q9NR09	Baculoviral IAP repeat-containing protein 6	2
28	Q96GP6	Scavenger receptor class F member 2	2
29	P06748	Nucleophosmin (NPM)	2
30	Q9NQT5	Exosome complex component RRP40	2

31	Q13428	Treacle protein (Treacher Collins syndrome protein)	2
		Extracellular Matrix Proteins	
32	Q99832	T-complex protein 1 subunit eta (TCP-1-eta)	2
33	P22392	Nucleoside diphosphate kinase B (NDK B)	2
34	P22626	Heterogeneous nuclear ribonucleoproteins A2/B1 (hnRNP A2/B1)	2
35	P62826	GTP-binding nuclear protein Ran	2
36	P61604	10 kDa heat shock protein	2
37	P16401	Histone H1.5	2
38	P62805	Histone H4	2
39	P62906	60S ribosomal protein L10a	2
40	P62942	Peptidyl-prolyl cis-trans isomerase FKBP1A	2
41	P60660	Myosin light polypeptide 6	2
42	P04075	Fructose-bisphosphate aldolase A	2
43	O76003	Glutaredoxin-3	3
44	P06454	Prothymosin alpha	3
45	P78371	T-complex protein 1 subunit beta (TCP-1-beta) (CCT-beta)	3
46	P19338	Nucleolin	2
47	P61978	Heterogeneous nuclear ribonucleoprotein K	3
48	P48643	T-complex protein 1 subunit epsilon	3
49	P60174	Triosephosphate isomerase	3
50	Q06830	Peroxiredoxin-1	2
51	P09651	Heterogeneous nuclear ribonucleoprotein A1	2
52	P10412	Histone H1.4	2
53	P11586	C-1-tetrahydrofolate synthase	2
54	P02545	Prelamin-A/C	2
55	P07195	L-lactate dehydrogenase B chain	2
56	Q99880	Histone H2B type 1-L	2
57	P38646	Stress-70 protein	2
58	P62857	40S ribosomal protein S28	2

Supplementary Table-2. List of protein identified in ≥2 of the cell lysate samples.

Entry no	Protein ID	Protein name	Identified in samples
		Plasma membrane	3
1	P07900	Heat shock protein HSP 90-alpha	3
2	P63261	Actin, cytoplasmic 2	3
3	P08238	Heat shock protein HSP 90-beta	3
4	P06733	Alpha-enolase	3
5	P04406	Glyceraldehyde-3-phosphate dehydrogenase	3
6	P11142	Heat shock cognate 71 kDa protein	3
7	P10809	60 kDa heat shock protein, mitochondrial	3
8	P13639	Elongation factor 2	3
9	P06576	ATP synthase subunit beta, mitochondrial	3
10	P04083	Annexin A1	3
11	P09211	Glutathione S-transferase P	3
12	P08865	40S ribosomal protein SA	3
13	P11021	78 kDa glucose-regulated protein	3
14	P12429	Annexin A3	3
15	P15311	Ezrin	3
16	P07355	Annexin A2	3
17	P08758	Annexin A5	3
18	P04792	Heat shock protein beta-1	3
19	P06744	Glucose-6-phosphate isomerase	3
20	P11413	Glucose-6-phosphate 1-dehydrogenase	3
21	P14625	Endoplasmin	3
22	P27797	Calreticulin	3
23	P05783	Keratin, type I cytoskeletal 18	3
24	P23528	Cofilin-1	3
25	P68402	Platelet-activating factor acetylhydrolase IB subunit beta	2
26	P61586	Transforming protein RhoA	2
27	P07602	Prosaposin	2
28	Q9H4A4	Aminopeptidase B	2
29	Q15046	Lysine--tRNA ligase	2
30	P52272	Heterogeneous nuclear ribonucleoprotein M	2
31	Q16555	Dihydropyrimidinase-related protein 2	2
32	Q9ULV4	Coronin-1C	2
33	P60709	Actin, cytoplasmic 1	2
34	P62834	Ras-related protein Rap-1A	2

35	Q5TZA2	Rootletin (Ciliary rootlet coiled-coil protein)	2
36	Q99497	Protein/nucleic acid deglycase DJ-1	2
37	P16070	CD44 antigen	2
38	P08195	4F2 cell-surface antigen heavy chain	2
39	Q00610	Clathrin heavy chain 1	2
40	P07237	Protein disulfide-isomerase	2
41	O75369	Filamin-B	2
		Extracellular matrix	
42	P38646	Stress-70 protein, mitochondrial	3
43	P46783	40S ribosomal protein S10	3
44	P78371	T-complex protein 1 subunit beta	3
45	P61604	10 kDa heat shock protein, mitochondrial	3
46	P18669	Phosphoglycerate mutase 1	3
47	P07195	L-lactate dehydrogenase B chain	3
48	P05386	60S acidic ribosomal protein P1	3
49	P00558	Phosphoglycerate kinase 1	3
50	P22234	Multifunctional protein ADE2	3
51	P16949	Stathmin	3
52	P34932	Heat shock 70 kDa protein 4	3
53	P50990	T-complex protein 1 subunit theta	3
54	P17066	Heat shock 70 kDa protein 6	3
55	P32119	Peroxiredoxin-2	3
56	P40925	Malate dehydrogenase	3
57	P24752	Acetyl-CoA acetyltransferase, mitochondrial	3
58	Q14974	Importin subunit beta-1	3
59	Q16658	Fascin	3
60	P29401	Transketolase	3
61	P13667	Protein disulfide-isomerase A4	3
62	P60660	Myosin light polypeptide 6	2
63	P60842	Eukaryotic initiation factor 4A-I	2
64	P30050	60S ribosomal protein L12	2
65	P40939	Trifunctional enzyme subunit alpha,	2
66	Q02790	Peptidyl-prolyl cis-trans isomerase FKBP4	2
67	P68363	Tubulin alpha-1B chain	2
68	O43175	D-3-phosphoglycerate dehydrogenase	2
69	Q58FF8	Putative heat shock protein HSP 90-beta 2	2
70	Q9NTK5	Obg-like ATPase 1	2
71	P10768	S-formylglutathione hydrolase	2
72	P05387	60S acidic ribosomal protein P2	2
73	P31939	Bifunctional purine biosynthesis protein PURH	2
74	P17987	T-complex protein 1 subunit alpha	2
75	P50502	Hsc70-interacting protein	2
76	O75347	Tubulin-specific chaperone A	2
77	P49588	Alanine--tRNA ligase, cytoplasmic	2

78	Q9Y617	Phosphoserine aminotransferase	2
79	P30048	Thioredoxin-dependent peroxide reductase	2
80	Q9H299	SH3 domain-binding glutamic acid-rich-like protein 3	2
81	P07954	Fumarate hydratase	2
82	P27348	14-3-3 protein theta	2
83	P78417	Glutathione S-transferase omega-1	2
84	P62942	Peptidyl-prolyl cis-trans isomerase FKBP1A	2
85	Q15102	Platelet-activating factor acetylhydrolase IB subunit gamma	2
86	P46779	60S ribosomal protein L28	2
87	P07741	Adenine phosphoribosyltransferase	2
88	P41250	Glycine-tRNA ligase	2
89	P61916	Epididymal secretory protein E1	2
90	P50395	Rab GDP dissociation inhibitor beta	2
91	Q96KP4	Cytosolic non-specific dipeptidase	2
92	O75874	Isocitrate dehydrogenase	2
93	P0DP25	Calmodulin-3	2
		Cytoplasmic protein	
94	P25398	40S ribosomal protein S12	3
95	P63220	40S ribosomal protein S21	3
96	O75607	Nucleoplasmin-3	3
97	P06748	Nucleophosmin	3
98	P62917	60S ribosomal protein L8	3
99	Q9Y2Z0	Protein SGT1 homolog	3
100	P26640	Valine--tRNA ligase	2
101	P24534	Elongation factor 1-beta	2
102	Q01105	Protein SET	2
103	P29692	Elongation factor 1-delta	2
104	P43487	Ran-specific GTPase-activating protein	2
105	P31948	Stress-induced-phosphoprotein 1	2
106	Q99873	Protein arginine N-methyltransferase 1	2
107	P47914	60S ribosomal protein L29	2
108	Q07020	60S ribosomal protein L18	2
109	P36551	Oxygen-dependent coproporphyrinogen-III oxidase	2
110	P12956	X-ray repair cross-complementing protein 6	2
111	P31153	S-adenosylmethionine synthase isoform type-2	2
112	P0DP23	Calmodulin-1	2
113	P0DP24	Calmodulin-2	2
114	Q3BBV1	Neuroblastoma breakpoint family member 20	2
115	P63313	Thymosin beta-10	2
116	Q8IYT4	Katanin p60 ATPase-containing subunit A-like 2	2
117	P80297	Metallothionein-1X	2
118	O15355	Protein phosphatase 1G	2
119	P39687	Acidic leucine-rich nuclear phosphoprotein 32 family member A	2

120	Q99733	Nucleosome assembly protein 1-like 4	2
		Nuclear protein	
121	P14618	Pyruvate kinase PKM	3
122	Q06830	Peroxiredoxin-1	3
123	P62826	GTP-binding nuclear protein Ran	3
124	P68371	Tubulin beta-4B chain	3
125	P09382	Galectin-1	3
126	P62937	Peptidyl-prolyl cis-trans isomerase A	3
127	P60174	Triosephosphate isomerase	3
128	P06454	Prothymosin alpha	3
129	P12277	Creatine kinase B-type	3
130	P40926	Malate dehydrogenase, mitochondrial	3
131	P0DMV9	Heat shock 70 kDa protein 1B	3
132	P00338	L-lactate dehydrogenase A chain	3
133	Q5VTE0	Putative elongation factor 1-alpha-like 3	3
134	P22392	Nucleoside diphosphate kinase B	3
135	P63104	14-3-3 protein zeta/delta	3
136	P26447	Protein S100-A4	3
137	P07737	Profilin-1	3
138	P04075	Fructose-bisphosphate aldolase A	3
139	P22314	Ubiquitin-like modifier-activating enzyme 1	3
140	P26641	Elongation factor 1-gamma	3
141	P46777	60S ribosomal protein L5	3
142	P31947	14-3-3 protein sigma	3
143	P36578	60S ribosomal protein L4	3
144	O43707	Alpha-actinin-4	3
145	Q15185	Prostaglandin E synthase 3	3
146	P31949	Protein S100-A11	3
147	P20962	Parathymosin	3
148	P61956	Small ubiquitin-related modifier 2	3
149	P00441	Superoxide dismutase	2
150	P61978	Heterogeneous nuclear ribonucleoprotein K	2
151	Q00839	Heterogeneous nuclear ribonucleoprotein U	2
152	P07437	Tubulin beta chain	2
153	P12004	Proliferating cell nuclear antigen	2
154	P37837	Transaldolase	2
155	P30044	Peroxiredoxin-5, mitochondrial	2
156	P30086	Phosphatidylethanolamine-binding protein 1	2
157	P23526	Adenosylhomocysteinase	2
158	P19338	Nucleolin (Protein C23)	2
159	P30041	Peroxiredoxin-6	2
160	P55786	Puromycin-sensitive aminopeptidase	2
161	P63244	Receptor of activated protein C kinase 1	2
162	P30101	Protein disulfide-isomerase A3	2

163	P55072	Transitional endoplasmic reticulum ATPase	2
164	Q16531	DNA damage-binding protein 1	2
165	P09960	Leukotriene A-4 hydrolase	2
166	Q9UHV9	Prefoldin subunit 2	2
167	Q13126	S-methyl-5'-thioadenosine phosphorylase	2
168	Q99880	Histone H2B type 1-L	2
169	Q13404	Ubiquitin-conjugating enzyme E2 variant 1	2
170	P83916	Chromobox protein homolog 1	2
171	Q8NHW5	60S acidic ribosomal protein P0-like	2
172	Q15233	Non-POU domain-containing octamer-binding protein	2
173	O76021	Ribosomal L1 domain-containing protein 1	2