

Eberhard Karls Universität Tübingen
Fakultät für Chemie und Pharmazie
Interfakultäres Institut für Biochemie

**QUANTITATIVE PROTEOMIC ANALYSIS
OF ENDOTHELIAL CELLS
DURING CAPILLARY MORPHOGENESIS**

Diplomarbeit

vorgelegt von
Marco Yannic Hein

Januar 2010

Diese Arbeit wurde im Zeitraum von Mai 2009 bis Januar 2010 unter der Leitung von Prof. Dr. Matthias Mann in der Abteilung für Proteomics und Signaltransduktion am Max-Planck-Institut für Biochemie in Martinsried durchgeführt.

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Martinsried, den 18. Januar 2010

Abstract

Angiogenesis, the growth of new blood vessels, is a fundamental physiological process. It is also of great significance during the development of tumours, which rely on blood vessels for supply of nutrients and oxygen and for the spreading of metastases. We studied the formation of capillaries using the established cell culture model of human endothelial cells plated on Matrigel, a three-dimensional extracellular matrix. Using quantitative mass spectrometry, we identified almost 8000 proteins and their regulation pattern during capillary morphogenesis.

Among the most upregulated proteins, we propose C-type lectin family 14 member A (CLEC14A) as a novel vessel marker. We characterized CLEC14A as a highly glycosylated transmembrane protein with a set of distinctly regulated phosphorylation sites on its cytoplasmic tail. In an attempt of elucidating the interaction and signaling network of CLEC14A, we identified LanC-like protein 1 (LANCL1) as specific interactor of the C-terminus of CLEC14A in its phosphorylated state.

Our findings will contribute to the understanding of the complex interplay of molecular mechanisms regulating vessel formation.

Zusammenfassung

Angiogenese, das Wachstum neuer Blutgefäße, ist ein fundamentaler physiologischer Prozess. Dieser ist unter anderem von großer Bedeutung bei der Tumorentwicklung, da Tumore auf Blutgefäße für ihre Versorgung mit Nährstoffen und Sauerstoff und für das Streuen von Metastasen angewiesen sind. Wir haben die Entstehung von Kapillaren anhand eines etablierten Zellkulturmodells untersucht, bei dem menschliche Endothelzellen auf Matrigel kultiviert werden, einer dreidimensionalen extrazellulären Matrix. Mit Hilfe von quantitativer Massenspektrometrie konnten wir annähernd 8000 Proteine und ihr Regulationsmuster während der Kapillarbildung identifizieren.

Unter den am stärksten hochregulierten Proteinen ist CLEC14A, ein C-Typ Lektin, welches wir als neuen Blutgefäßmarker vorschlagen. Wir haben CLEC14A als hochglycosyliertes Transmembranprotein charakterisiert, das in seinem cytoplasmatischen Teil einige spezifisch regulierte Phosphorylierungsstellen besitzt. Als Beitrag zur Entschlüsselung des Signal- und Interaktionsnetzwerks von CLEC14A konnten wir LANCL1 als Interaktionspartner von CLEC14A identifizieren, der spezifisch den phosphorylierten C-Terminus bindet.

Diese Erkenntnisse werden das Verständnis des komplexen Zusammenspiels der molekularen Mechanismen der Blutgefäßentstehung verbessern.

Acknowledgements

I hereby want to thank everybody who contributed to this work:

First of all many thanks to Matthias Mann for the privilege to work in your lab, for opening doors and creating a unique environment.

Likewise, many thanks to Sara Zanivan for the great supervision, for introducing me to the mysteries of angiogenesis and mass spectrometry, for your tireless dedication and support and for spreading your enthusiasm.

I am honored to be able to work with you two!

A special thank you goes to Doron Rapaport for making things possible in Tübingen and to Marc Schmidt-Supprian for help along the way.

Thanks to all the members of our department for sharing their knowledge and for keeping the spirits up.

Special thanks to Christian Eberl and Christian Kelstrup for valuable advice concerning the peptide pulldown. Thanks to Korbinian Mayr for shining the light at the end of the tunnel during seemingly endless maintenance hours.

Thanks to my office mates for the atmosphere, especially Tami Geiger for contemplative and exhilarating discussions and Marlis Zeiler for substantial amounts of computing time and not giving up hope.

Thanks to all my friends for your friendship, for appreciating my dedication to my work, for enduring scientific explanations and making me enjoy the world outside the lab.

Finally, my love and a huge thank you goes to my family, first of all to my parents for your unconditional support and your overwhelming faith in me, to my grandparents for your cheers and thoughts and to Ajla Hrle for your endless love.

I dedicate this work to you.

Contents

Abstract	v
Acknowledgements	vii
Abbreviations	x
1 Introduction	1
1.1 Angiogenesis	1
1.1.1 Morphology and Signalling of Angiogenesis	2
1.1.2 Angiogenesis <i>in vitro</i>	3
1.1.3 C-Type Lectins	4
1.1.4 Role of C-Type Lectin Family 14 Members in Angiogenesis	6
1.2 Mass Spectrometry Based Proteomics	7
1.2.1 General Workflow of Mass Spectrometry Based Proteomics	8
1.2.2 Quantitative Mass Spectrometry	9
1.2.3 Applications of Quantitative Mass Spectrometry	11
2 Results	13
2.1 The HUVEC Proteome	13
2.1.1 Dynamics of the HUVEC Proteome	16
2.1.2 Common Features of Clustered Proteins	18
2.2 Expression Pattern of CLEC14A	20
2.3 Posttranslational Modifications of CLEC14A	22
2.3.1 Glycosylation of CLEC14A	22
2.3.2 Phosphorylation of the CLEC14A Cytoplasmic Tail	23
2.4 Interaction Partners of CLEC14A	27
3 Discussion	31
4 Materials and Methods	35
4.1 Cell Culture	35
4.1.1 Isolation and Culture of Human Umbilical Vein Endothelial Cells	35
4.1.2 Cell Lysis	35
4.1.3 SILAC Labelling of HUVECs	36
4.1.4 Matrigel Assay	36
4.2 Pulldown Assays	37
4.2.1 Peptide Pulldown	37

4.2.2 Immunoprecipitation	37
4.3 Mass Spectrometry and Data Analysis.	38
4.3.1 1D-SDS-PAGE and In-Gel Digestion of Proteins	38
4.3.2 Nanoflow HPLC-MS/MS	39
4.3.3 Peptide Identification and Quantification	40
4.4 Protein Biochemistry	40
4.4.1 Deglycosylation Assay	40
4.4.2 Western Blotting and Immunodetection	40
A Appendix	42
A.1 SILAC Labelling of HUVECs	42
References	43

Abbreviations

ABC	ammonium bicarbonate	HUVECs	human umbilical vein endothelial cells
ACN	acetonitrile	ICAT	isotope-coded affinity tags
ANOVA	analysis of variance	iTRAQ	isobaric tag for relative and absolute quantitation
Arg10	L-arginine- $^{13}\text{C}_6$ $^{15}\text{N}_4$	KEGG	Kyoto Encyclopedia of Genes and Genomes
BSA	bovine serum albumin	Lam	laminin
CID	collision-induced dissociation	LANCL1	LanC-like protein 1
CLEC14A	C-type lectin domain family 14 member A	Lys8	L-lysine- $^{13}\text{C}_6$ $^{15}\text{N}_2$
CORUM	Comprehensive Resource of Mammalian Protein Complexes	MS/MS	tandem mass spectrometry, also MS ²
DTT	dithiothreitol	MS	mass spectrometry
ECM	extracellular matrix	PBS	Dulbecco's phosphate buffered saline
EHS sarcoma	Engelbreth-Holm-Swarm sarcoma in mice	PNGase	peptide N-glycosidase
ESI	electrospray ionization	PTM	posttranslational modification
FDR	false discovery rate	SILAC	stable isotope labelling by amino acids in cell culture
FN	fibronectin	TBS	Tris buffered saline
GFR	growth factor reduced Matrigel	TFA	trifluoroacetic acid
GO	Gene Ontology	VEGF	vascular endothelial growth factor
HIF1A	hypoxia inducible factor 1 α		
HPLC	high performance liquid chromatography		

1 Introduction

1.1 Angiogenesis

The cardiovascular system constitutes the principal network for distributing nutrients, gases and signalling molecules in the body. In compliance with this fundamental significance, it is the first functional organ which develops in the vertebrate body during ontogenesis [1]. The *de novo* formation of blood vessels in the embryo is defined as vasculogenesis [2]. It involves the differentiation of mesodermal stem cells into angioblasts and further into vascular endothelial cells. Endothelial cells are capable of producing a basal lamina and forming a lumen. They will later cover the entire inner surface of all blood vessels. During later stages of development, new vessels form by sprouting or intussusception from pre-existing ones, a process termed angiogenesis [3].

Angiogenesis occurs physiologically in growth and development and during wound healing. Pathologically, angiogenesis is of great significance for tumour progression: First, proliferating tumours rely on vascularization to meet their metabolic needs (fig. 1.1).

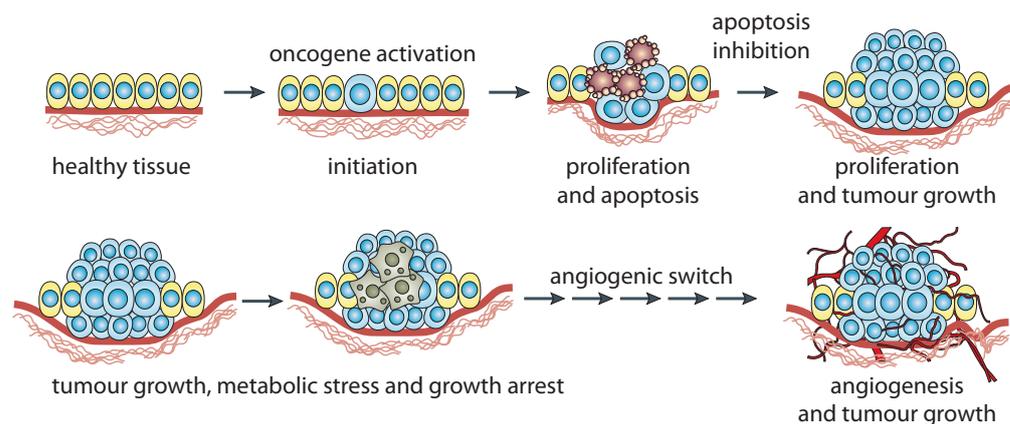


Figure 1.1: Significance of Tumour Angiogenesis.

Tumour development is initiated by mutations that cause oncogene activation. Apoptotic cell death is initially opposing tumour cell proliferation, until a subpopulation acquires defects in apoptosis. Proliferation progresses until the tumour reaches a critical size at which its centre is cut off the blood supply. A high percentage of humans at a certain age are thought to carry such *in situ* tumours without symptoms of disease [4]. Some tumours trigger the “angiogenic switch”, by which they succeed to trigger the invasive growth of newly formed blood vessels. This cures the tumour of hypoxia and metabolic stress and allows malignant progression. Adapted from ref. [5].

Second, angiogenesis allows the infiltration of the tumour by leukocytes, which provide chemokines and proteases, thereby assisting the detachment of tumour cells, their migration into the bloodstream and the formation of metastases [6]. Finally, also metastases need to become vascularized in order to grow [7].

The concept of tumourigenesis relying on angiogenesis was first proposed by Judah Folkman in 1971 [8] and has launched the field of anti-angiogenic cancer therapy.

1.1.1 Morphology and Signalling of Angiogenesis

Endothelial cells together with pericytes are the major players in angiogenesis. The sprouting of blood vessels comprises several consecutive steps, as outlined in fig. 1.2; these include the local degradation of the basement membrane, migration of endothelial cells, bipolar alignment, lumen formation and cell adhesion, and finally the formation of a new basement membrane [9].

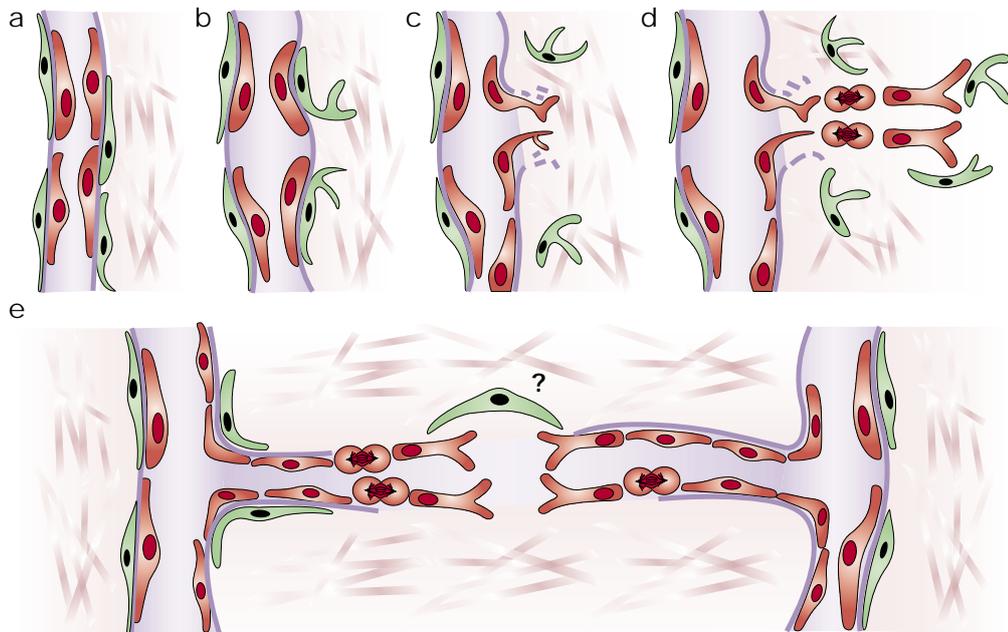


Figure 1.2: Morphological Changes During Angiogenesis.

a) A capillary lumen is lined by endothelial cells (red), which rest on a basement membrane (purple). Pericytes (green) are laid against the outside of the basement membrane. **b)** Angiogenic stimuli lead to the detachment of pericytes. **c)** This allows the degradation of the basement membrane and the migration of endothelial cells towards the chemotactic signal. **d)** Endothelial cells proliferate and stretch to form a capillary sprout. **e)** The sprouts form a lumen, accompanied by the formation of a new basement membrane and the attachment of pericytes. Finally, two sprouts fuse by a scarcely known mechanism. From ref. [10].

All of these processes are tightly regulated by activating and inhibiting molecules, which include mainly cytokines and growth factors, among them the members of the vascular endothelial growth factor (VEGF) and fibroblast growth factor (FGF) families [11].

The last two decades have revealed many proteins that are involved in the orchestration of the cellular events during angiogenesis: VEGF can be induced by a range of stimuli, e. g. hypoxia, and triggers tyrosine kinase pathways by binding to VEGF receptors, such as FLK1. Early responses to VEGF include an increased permeability of the vessels, mediated via vascular endothelial cadherin (VE-cadherin), the major adhesion molecule controlling cell to cell contacts in the endothelium [12, 13]. Matrix metalloproteases (MMPs) are then secreted in order to degrade the basement membrane [14]. Subsequent cell migration and remodelling events critically rely on integrin-mediated cell adhesion processes [15]. Next to VEGF and its receptors, the angiopoietin/TIE and ephrin-B/Eph-B systems of ligands and receptors mediate similar effects during later stages of angiogenesis [9].

Antiangiogenic molecules, such as angiostatin and thrombospondin-1, are known to counteract VEGF and other proangiogenic effectors.

Antiangiogenic therapy is promising to fight cancer and other diseases dependent on angiogenesis in a targeted fashion, thereby reducing adverse side effects. Compounds currently in clinical use mostly target VEGF signalling, either directly such as Bevacizumab, a monoclonal antibody against VEGF-A, or by blocking downstream receptor tyrosine kinases, such as Sunitinib, a small molecule inhibitor [16]. A detailed understanding of the interplay of molecular mechanisms regulating vessel formation is critical for the identification of new pharmacological targets and the development of new drugs.

1.1.2 Angiogenesis *in vitro*

Given the complexity of events during angiogenesis, it was long difficult to study this process. Early studies relied exclusively on animal models, involving e. g. the implantation of tumour tissue into laboratory animals and monitoring their vascularization. Later, an assay based on the chick embryo chorioallantoic membrane was developed, allowing for the screening of angiogenic regulators using chicken eggs. Only in 1980 it was discovered that angiogenesis can be studied using *in vitro* models of endothelial cells cultured in appropriate media and plated on special matrices. Endothelial cells were able to form tubular networks *in vitro*, which were almost identical, by light and electron microscopy, to capillary vascular beds *in vivo*. The morphological changes in these setups occurred slowly, generally requiring the cells to be kept in culture for many weeks [17].

In 1982, it was discovered that the Engelbreth-Holm-Swarm (EHS) sarcoma in mice had the unique property of producing extracellular matrix components that are distinct from cartilage but instead resemble the basement membrane [18]. Laminin I, type IV collagen, perlecan (heparan sulfate proteoglycan) and nidogen/entactin were identified as major constituents, next to growth factors and several matrix proteases. The mixture can be extracted in the cold and forms a gel when brought to body temperature.

A range of cell types show distinct growth or differentiation behavior when plated on EHS matrix [19]. Most significantly, endothelial cells were shown to form capillary-like structures. The process occurred within a few hours, compared to weeks in previous setups, and included distinct steps of cell adhesion, migration, alignment, protease secretion, and tubule formation. [20]. This underlined the significance on the interaction of endothelial cells with their surrounding matrix during angiogenesis.

Today, the EHS matrix is commercially available as “Matrigel” and is used in a plethora of cell biological assays [21, 22]. Among them are screens for angiogenic and anti-angiogenic molecules, which can be performed in high throughput and respond to almost all known regulatory factors [23]

1.1.3 C-Type Lectins

Lectins comprise a large family of carbohydrate-binding proteins, found ubiquitously in all kinds of taxa. They are usually large, multidomain proteins, yet the sugar-binding activity can be generally ascribed to a distinct carbohydrate-recognition domain (CRD) within the protein [24]. CRDs are structurally diverse and named only by their common function.

Animal lectins localize either within intracellular compartments, on the cell surface or are secreted. They function in (glyco)protein quality control, sorting and trafficking inside the cell, whereas they mediate cell-cell or cell-matrix recognition in the extracellular space.

One subgroup of lectins are the Ca^{2+} -dependent type (C-type) lectins, which are in turn divided into several families. All members share the conserved C-type lectin domain (CTLN), which is one class of CRDs. Next to their carbohydrate binding activity, CTLNs frequently serve in protein interaction and binding is not always Ca^{2+} -dependent [24]. Some CTLNs even lack the conserved motifs that are connected with sugar binding [25], among them all members of the family 14, or endosialin family, shown in fig. 1.3.

The endosialin family comprises four members: endosialin (CD248; tumour endothelial marker 1, TEM1), thrombomodulin (CD141), complement component C1q receptor (C1qR; CD93) and C-type lectin domain family 14 member A (CLEC14A;

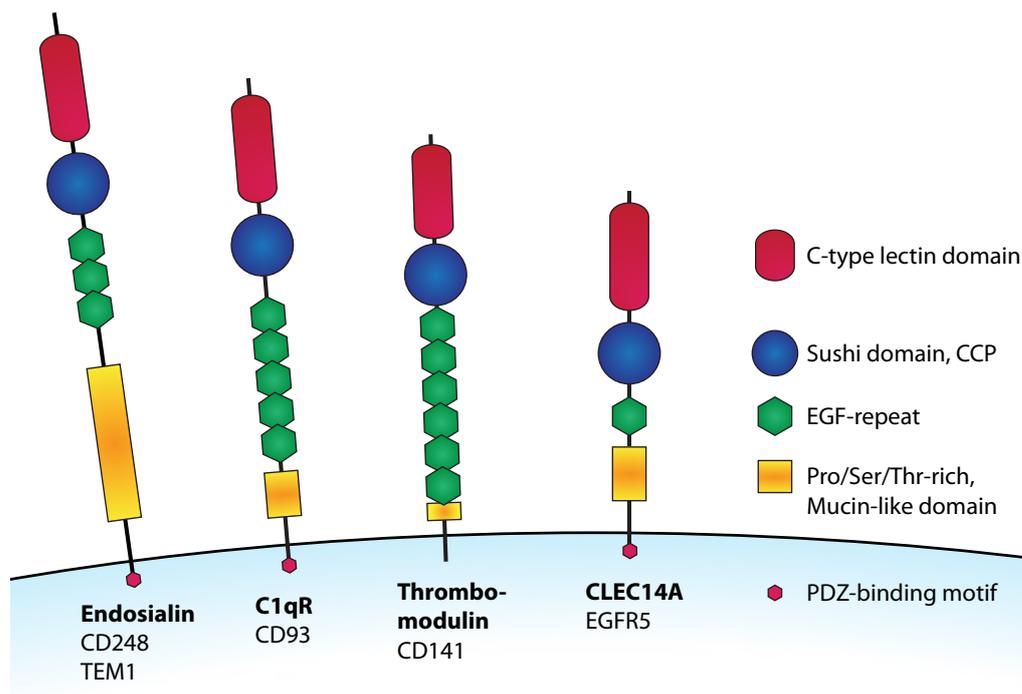


Figure 1.3: C-Type Lectin Family 14. All four members of the family 14 of C-type lectins share a set of conserved domains: the prototypical C-type lectin domain, the sushi domain, a number of EGF-like repeats and a proline, serine, threonine rich mucin-like domain. The C-terminal three amino acids of cytoplasmic act as a PDZ binding motif in some members.

epidermal growth factor receptor 5, EGFR5). All members of the Endosialin family are type I transmembrane proteins with sizes of the processed forms ranging from 469 to 740 amino acids. The large N-terminal extracellular domains all contain the eponymous C-type lectin domain, followed by a sushi domain (complement control protein domain, CCP domain), 2–6 epidermal growth factor (EGF)-like repeats and a loosely defined serine/threonine/proline-rich mucin-like domain of variable length. All mucin-like domains feature many potential O-glycosylation motifs and a strong O-glycosylation pattern was published in the case of endosialin [26]. The short cytoplasmic tails of 36–72 amino acids show little conservation among the paralogs.

The C-terminal 3 amino acids of C1qR were reported to function as a PDZ binding motif, which interacts with the PDZ domain of the protein GIPC1 [27]. The PDZ domain is named after the proteins PSD95, DlgA and zo-1, which were first shown to contain this domain.

By homology also endosialin contains the PDZ binding motif [26], yet no interaction partner(s) are known [28]. The extreme C-terminus of CLEC14A (amino acids

SDA) is identical to the PDZ binding motif in integrin alpha-5 and homologous to the sequence SEA in neuropilin-1, both of which were also shown to interact with GIPC1 [29, 30]. The respective sequence in thrombomodulin does not match the PDZ binding motif consensus.

Phosphorylation in the vicinity of PDZ binding motifs may influence the binding of PDZ domain containing proteins, and thereby serve as a regulatory mechanism. This was shown in the case of the C-terminus of the K⁺-channel Kir 3.2 binding to the PDZ domain protein PSD-95 [31].

1.1.4 Role of C-Type Lectin Family 14 Members in Angiogenesis

Endosialin was shown to be involved in tumour angiogenesis. It was discovered as tumour stromal antigen, expressed in vascular endothelial cells of malignant tumors [32]. Later, it was identified in a transcriptomic study that compared endothelia derived from tumours with ones derived from normal tissue. The protein was named tumour endothelial marker 1 (TEM1) for its distinct association with tumour-derived endothelia [33]. In another study based on the affinities of monoclonal antibodies, it was reported that endosialin is a cell surface glycoprotein which is predominantly expressed by fibroblasts and a subset of pericytes, that are associated with tumour vessels but not by tumour endothelium itself [34]. This discrepancy is matter of ongoing debate [28]. It is important to note that endosialin is not expressed in HUVECs *ex vivo* [26].

Endosialin knockout mice exhibit no difference to wild type animals in development, wound healing and tumour development [35]. However, there was a striking reduction in tumour growth, invasiveness, and metastasis after transplantation of tumours to abdominal sites in knockout mice. At the same time tumours were vascularized with a lot more small instead of medium and large vessels. These findings indicate a potential function for endosialin in vessel maturation and growth.

C-type lectin family member 14A is the fourth and largely uncharacterized member of the endosialin family [25]. There is very few experimental data about CLEC14A. One microarray study described CLEC14A expression to be endothelial cell restricted [36] (referred to as AW770514). Another study reported expression of mouse CLEC14A in the developing brain, and widely expressed in the adult body [37] (referred to as *Ceg1*).

Here, we report the first thorough characterization of CLEC14A as novel vessel marker, that is highly upregulated in HUVECs on Matrigel.

1.2 Mass Spectrometry Based Proteomics

Science has long sought to describe a given biological system, e. g. a cell, tissue or organism, in its totality. Fields of research that aim to provide information on such a global scale are referred to by the ending “-omics”.

A general concept of scientific discovery, especially in “omics” fields involves the relative comparison of different functional states of a biological system. Many approaches have been developed to deliver a comprehensive and unbiased comparison of two such states, but the task remains daunting.

Genomics research first was able to address biological questions comprehensively on the level of DNA. The field benefited greatly from a straightforward methodology based on oligonucleotide synthesis, hybridization and amplification by polymerase chain reaction [38]. Yet, biological effects are generally mediated by molecules that are located further downstream in the flow of genetic information: RNA species in some processes, but proteins in most cases.

The term “proteome” describes the entirety of proteins found in a given biological system. In contrast to the genome, the proteome is a highly dynamic ensemble: It follows internal (e. g. periodic) fluctuations and responds to external stimuli. Proteomic changes are achieved through altered rates of protein synthesis and protein degradation as well as through the activity of protein modifying enzymes.

More than a decade ago, the advent of DNA microarray technology, which harnesses the nucleic acid based methodology [39, 40], had paved the way to transcriptomics. The transcriptomic approach seeks to deduce functional information about protein expression from the levels of the corresponding mRNA transcripts. By using differentially labelled cDNA probes, microarrays allow the comparison of multiple functional states of a cell or tissue. However, it was shown that the levels mRNA and protein expression poorly correlate [41].

The recent development of ribosome profiling [42] allows a focused and quantitative view on those mRNA regions that are eventually transcribed. This approach yields expression data that correlate much better with the actual protein levels, but it is still far from being applicable on a general basis: It requires large sample amounts, involves many experimentally challenging steps and relies on expensive deep sequencing.

The chemical diversity of proteins has long hampered the development of methods for the comprehensive high-throughput analysis of the proteome. In recent years, mass spectrometry (MS) based approaches are emerging as the gold standard of proteomics research [43] and have proven capable of covering near-complete proteomes of eukaryotic cells [41].

This development was enabled by advances in sample preparation [44, 45], the establishment of quantitative methods [46], the advent of high-resolution mass spectrometers [47–49] and robust and automated workflows for bioinformatic analysis of the vast quantities of raw data [50].

1.2.1 General Workflow of Mass Spectrometry Based Proteomics

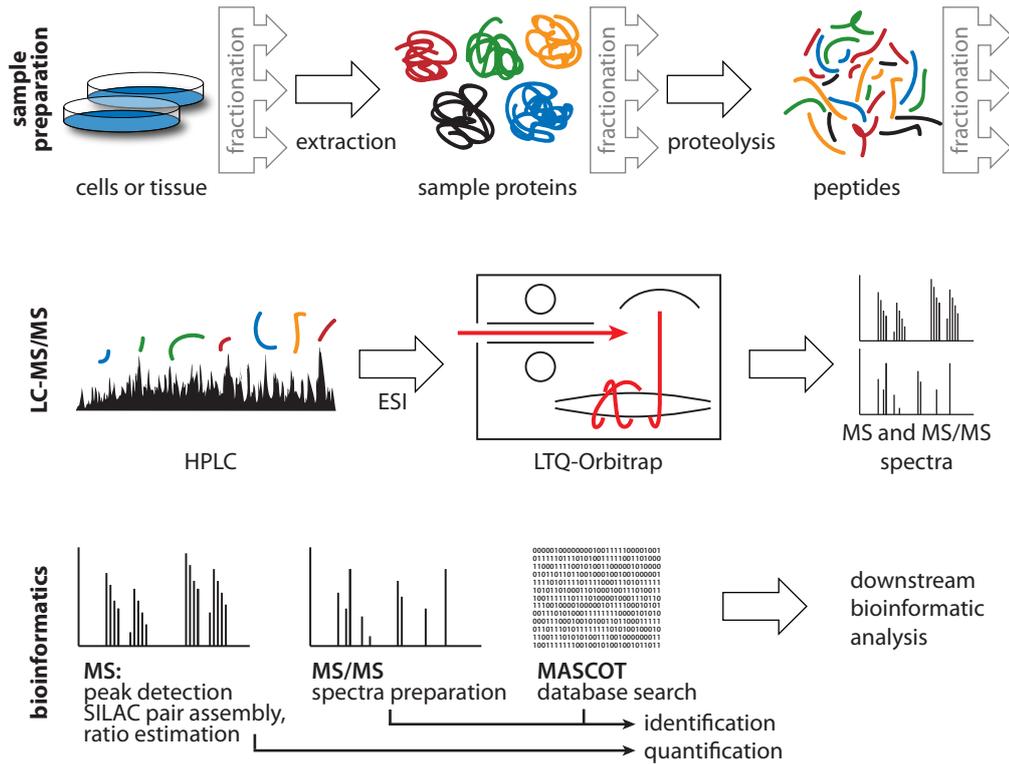


Figure 1.4: Basic Workflow of MS-Based Proteomics As Performed in Our Laboratory.
For details see main text.

A typical mass spectrometry based proteomic experiment follows the “bottom up” and “shotgun” principles, see fig. 1.4. Proteins are first extracted from the samples to be analysed, and digested into peptides, which are easily amenable to mass spectrometric analysis. Depending on the sample complexity, one- or multi-dimensional fractionation can be performed on the protein and/or peptide level. Each fraction is then further separated by high performance liquid chromatography (HPLC). The effluent of the HPLC column is subjected to electrospray ionization (ESI) and directly sprayed into the mass spectrometer. The mass spectrometer generally records full

MS spectra, alternating with fragmentation spectra (MS/MS) of the most abundant ions that are eluting at the given time.

“Shotgun” refers to the peptides being selected for fragmentation in a more or less non-targeted fashion. The identity of a peptide is deduced from its precursor mass and its characteristic fragment ion pattern. The proteins present in the mixture are later being reconstructed from the identified peptides (hence “bottom up”) during data analysis. This approach has its drawbacks, especially for the analysis of proteins with partial sequence identity, such as isoforms and splice variants, which need to be investigated based on peptides unique for one distinct protein species [51]. Yet, in terms of robustness, universality and throughput, it is far superior to the “top down” approach which starts from intact proteins.

1.2.2 Quantitative Mass Spectrometry

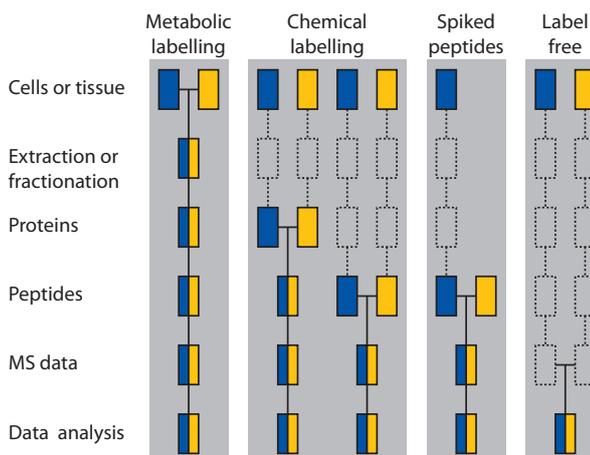
In many analytical setups, mass spectrometry is applied solely as a qualitative method used to identify proteins. However, quantitative information can be extracted from mass spectra and provides valuable information.

Mass spectrometry is not inherently quantitative, i. e. the intensity readout from the mass spectrometer is not directly correlated with the amount of peptide that was injected. This is largely a consequence of the different ionization propensities of different peptides, which in turn is caused by their chemical diversity.

A range of approaches have been developed to provide relative or absolute quantitative information. Many rely on labelling of peptide species with stable isotopes. This retains most of the chemical properties of the peptide while inducing a detectable mass shift, thereby allowing the relative quantification of the labelled versus the unlabelled form. Labelling can be performed chemically on the protein or peptide level, e. g. by using the ICAT [52] or iTRAQ reagents [53]. Alternatively, labels can be introduced metabolically e. g. by growing cells in media containing isotope-labelled compounds. The labelling method has consequences for the amenability of the samples and the error susceptibility: Metabolic labelling was first only applicable to cells and unicellular organisms, but has recently been extended to higher organisms such as the mouse [54]. Human tissue samples and clinical specimen obviously cannot be labelled metabolically and are only amenable to chemical labels, which can be applied on virtually any sample. Chemical labelling is performed rather late in the experimental workflow, thereby reducing the quantification accuracy due to variations caused by separate manipulation, whereas metabolically labelled samples can in principle already be mixed on the cell or tissue level (fig. 1.5).

Figure 1.5: Labelling Approaches and Quantitative Accuracy.

Blue and yellow boxes represent the different experimental conditions that are to be labelled and compared. After introduction of the label, the samples can be combined (horizontal lines) and from thereon be processed together. Dashed boxes and lines indicate stages of parallel handling, where uncompensated errors can occur that decrease the quantitative accuracy. Adapted from refs. [46, 55].



One popular and straightforward method of metabolic labelling is SILAC (stable isotope labelling by amino acids in cell culture), which was developed in our lab [56, 57]. Typically, cells are cultured in media containing either normal (“light”) or labelled (“heavy”) versions of essential amino acids such as arginine and lysine. In combination with the protease trypsin, which cleaves after exactly these amino acids [58], this results in one labelled amino acid per tryptic peptide, thereby introducing known mass shifts. By combining different versions of labelled amino acids, comparison of up to five states can be performed [59]. For a higher number of states, pairwise comparison can be done, using one common SILAC labelled reference as internal standard. Notably, labelled internal standards can be used for quantification of any unlabelled sample, provided that their composition overlaps sufficiently [60].

Despite advances in labelling methods, many experimental approaches would greatly benefit from the availability of robust label free quantification procedures. Label free methods are convenient and economical, omitting laborious or time-consuming labelling steps and the need for expensive isotopically labeled reagents.

Inherently, label free quantification is less accurate than label based methods, because the different samples are only compared at the final stage of the experiment (see fig. 1.5). Researchers have tried to deduce quantitative information from various properties of recorded mass spectra, e.g. the summed peak area of observed peptides or the number of acquired MS/MS spectra (“spectral counting”) [61]. Our laboratory exploits the relative intensities of identified peptides, together with recent advances in normalization algorithms for label free protein quantification; this approach has been shown to generate reliable data when combined with sophisticated statistical analysis of replicate experiments (J. Cox *et al.*, submitted to Nature Methods). Conceptual differences between label based and label free quantification methods are illustrated in fig. 1.6.

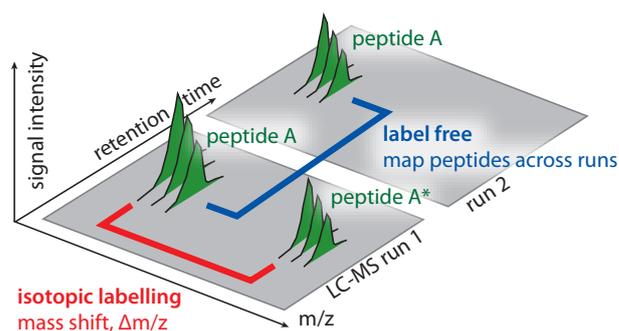


Figure 1.6: Label based vs Label Free Quantitative Mass Spectrometry. Isotopic labelling induces a mass shift between chemically equivalent peptide species (A and A*). Quantification is based on peak ratios within one spectrum. Label free quantification maps identical peptides across several LC-MS runs and compares the respective intensities. Adapted from ref. [62]

Many scientific questions, e. g. mathematical models of biological systems, require absolute quantification of proteins, while isotopic labelling provides only relative values. This issue can be solved by spiking in known amounts of synthetically produced labelled peptides or proteins [63]. This approach is generally limited to selected proteins of interest.

A recent study has claimed to deduce absolute quantitative information from unlabelled samples, using the intensity information from “high flyer” peptides [64], thereby challenging the dogma of mass spectrometry being inherently non-quantitative. It is evident that the margin of error of this method is rather wide, yet promising to provide some information where other methods are not applicable or high precision is not required.

1.2.3 Applications of Quantitative Mass Spectrometry

Quantitative mass spectrometry can monitor every property of a biological system that can be translated to a parameter observable in the m/z -intensity-retention time space, thereby making it a powerful tool in all fields of life science research [43, 57]; applications range from mapping the response of entire proteomes to certain stimuli to the identification of protein interaction networks.

Mass spectrometry is uniquely capable of detecting posttranslational modifications (PTMs) on proteins, which generally introduce a change in mass. With phosphorylation being among the most prominent PTMs, phosphoproteomics is one emerging field in recent years [65].

Analysis of protein interaction partners is another field boosted by MS based methods [66–68]. Novel interactors cannot only be identified, but also be quantified between controls and actual samples, allowing the discrimination of specific vs unspecific binding.

1 Introduction

With mass spectrometry being able to identify proteins *de novo*, based just on genomic information, the technology is highly valuable in the unbiased discovery of proteins in all kinds of experiments, offering alternative approaches for robustly established, but biased methods such as the western blot [69].

2 Results

2.1 The HUVEC Proteome

Mass spectrometry based proteomics offers powerful tools to study the complex processes during *in vitro* angiogenesis and opens up the prospect of identifying new angiogenic regulators with significance *in vivo*. The morphological changes that HUVECs undergo when plated on Matrigel are highly reminiscent of processes happening in the body during vessel formation and offer a unique amenability for examination.

HUVECs usually have an elliptic appearance and adopt a characteristic polygonal “cobblestone” morphology when grown to confluency. When seeded on Matrigel, HUVECs adhere within the first hour and stop dividing. Capillary formation begins after 1–2 h and is complete by 24 h [20]. Similar morphological changes occur on Matrigel depleted in growth factors (growth factor reduced, GFR).

Figure 2.1 illustrates the differences in HUVEC morphology.

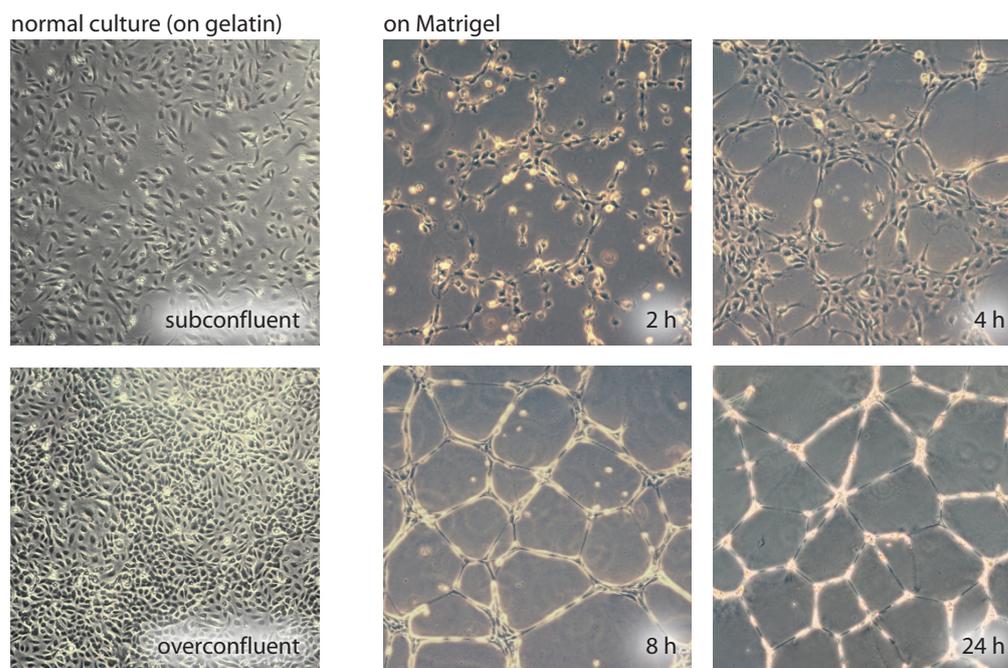


Figure 2.1: HUVEC morphology on different matrices. On gelatin coated plates, HUVECs grow in their typical “cobblestone” morphology (left panels). When seeded on Matrigel coated dishes, HUVECs undergo significant morphological changes and form capillary-like structures with a lumen, highly reminiscent of capillary vascular beds *in vivo*.

We designed an experiment which would allow to follow proteomic changes in endothelial cells during capillary formation and discern regulation patterns associated with morphological changes and changes triggered by extracellular matrices (ECMs) or growth factors. For this purpose, HUVECs were grown to confluency in normal culture on gelatin-coated dishes. Cells were then detached with trypsin and seeded onto dishes coated with either Matrigel or other ECMs, which did not induce capillary formation.

The following matrices were used:

Matrigel, basement membrane matrix extracted from the EHS sarcoma.

Matrigel GFR, prepared from EHS tumours by a modified protocol, which preserves the general composition, but eliminates most of the contained growth factors. Similarly to normal Matrigel, it triggers the formation of capillary-like structures.

Matrigel diluted 1:1000 contains all potentially involved factors, albeit in low concentrations. It is only able to trigger cell adhesion, but no capillary morphogenesis.

Laminin, extracted from the EHS sarcoma. Laminins are major basal lamina proteins and laminin 1 is the main component of Matrigel. Laminin 1 alone is not sufficient to trigger capillary morphogenesis of HUVECs, however in combination with a collagen I gel, laminin triggers elongation and anastomosing network formation and is therefore suggested to act as the principal factor during this process [20].

Fibronectin (FN), an ECM glycoprotein, is a general ligand for various integrins [70] and promotes rapid adhesion of many cell types, including HUVECs. Both fibronectin and integrins have been shown to function in embryonic cardiovascular development and tumour angiogenesis [15, 71].

Bovine serum albumin (BSA) serves as general negative control, since adhesion on BSA coated culture dishes relies on ECM proteins produced by the cells themselves.

Remaining cells in suspension were harvested as the time point zero. For quantification, lysates were mixed 2:1 with SILAC labelled internal standard (see section A.1) prior to in gel digestion and mass spectrometric analysis.

The experimental design is outlined in fig. 2.2

Given the kinetics of the process of capillary morphogenesis, cells on Matrigel were harvested after 12, 24 and 30 h. These time points were selected to cover the entire reorganization of the proteome, keeping in mind that proteomic changes lag behind their upstream stimuli.

Cells on ECMs without capillary formation, and on growth factor reduced Matrigel were collected after 24 h. All experiments were carried out in triplicate to allow statistical analysis.

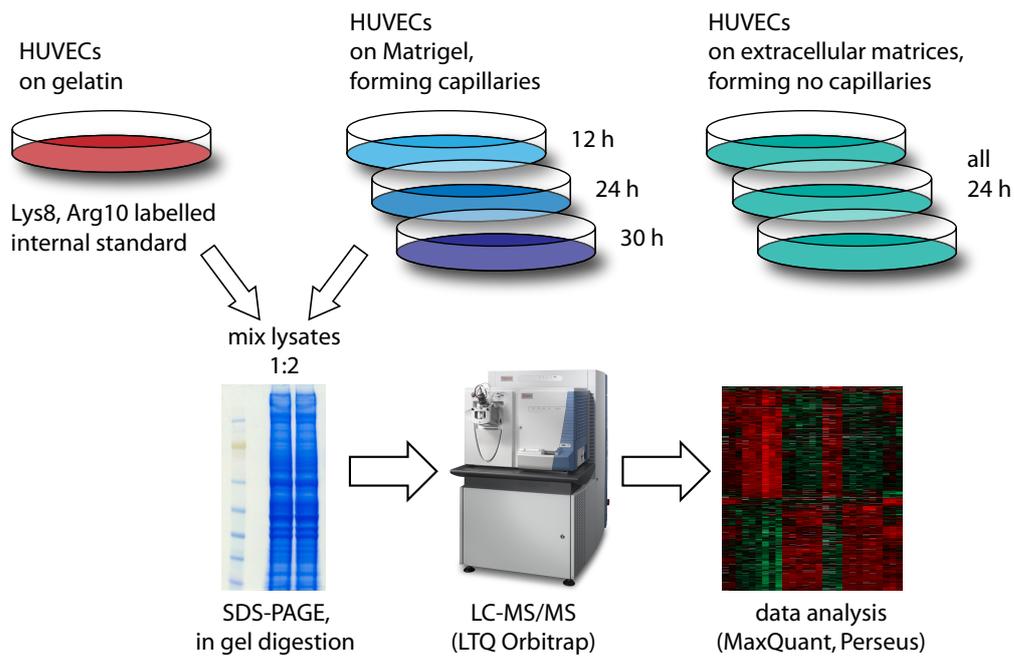


Figure 2.2: Experimental Design. HUVECs were seeded on Matrigel and the formation of capillary-like structures was followed over the course of 30 h. In parallel, HUVECs were cultured on other ECMs which do not trigger these morphological changes. These matrices included fibronectin, laminin, Matrigel diluted 1:1000, and BSA. After the indicated time points, the cells were lysed and mixed 2:1 with an internal standard consisting of the lysate of SILAC labelled HUVECs in normal culture. The mixtures were separated by 1D-SDS-PAGE, gels were cut into 18 slices and contained proteins digested in gel. Extracted peptides were subjected to LC-MS/MS. Acquired spectra were processed with MaxQuant for identification and relative quantification. Downstream bioinformatic analysis was done with Perseus.

The combined HUVEC proteome dataset contained 7959 proteins, that were identified with at least one unique peptide. Requiring at least two peptides per protein retained 7755 proteins. Among them were known endothelial cell markers [72] such as CD31, Cadherin-5, ICAM1, ICAM2, Neuropilin-1, Multimerin-2, Thrombomodulin, VCAM1, VEGFR1–3, and von Willebrand factor. The HUVEC proteome dataset, providing the qualitative information about proteins expressed in this cell type, together with the quantitative information about the regulation of these proteins in response to certain ECMs, provides a uniquely rich basis for follow-up research and discovery of novel marker proteins for capillary morphogenesis.

2 Results

To get an estimate of the reproducibility of the biological replicates we calculated pairwise correlation coefficients between log₂ transformed normalized SILAC ratios, see table 2.1. These numbers illustrate the general reproducibility of our experimental approach. Of note, correlation between experiments 1 and 2, which were carried out using cells from the same pool at the same passage, just 2 days apart, was higher than the correlations to experiment 3, which was performed with cells at one higher passage.

Table 2.1: Correlation of SILAC ratios between biological replicates

experiment	1 vs 2	1 vs 3	2 vs 3
Suspension	0.830	0.724	0.752
12 h	0.786	0.749	0.714
24 h	0.813	0.767	0.786
30 h	0.816	0.758	0.763
GFR	0.863	0.774	0.760
Matr. dil.	0.823	0.695	0.733
Laminin	0.758	0.818	0.757
FN	0.809	0.708	0.740
BSA	0.809	0.661	0.672
average	0.812	0.739	0.742

2.1.1 Dynamics of the HUVEC Proteome

To identify proteins that are significantly regulated on distinct matrices, we performed an analysis of variance (ANOVA) using our in-house developed Perseus software (J. Cox, unpublished). ANOVA requires independently acquired data, which are normally distributed (this is generally the case for logarithms of SILAC ratios) and homoscedastic (i. e. they possess the same variance within groups).

For this purpose, normalized SILAC ratios from all experiments were base-2 logarithmized and grouped into 9 experimental groups representing the different matrices or time points, each containing the 3 sets of data from the replicate measurements.

For each identified protein group, we then performed an F-test for the null hypothesis that the means of the ratios within all experimental groups are the same, and calculated significance values for the rejection of this hypothesis. Proteins were accepted in a list of significantly regulated proteins until a false discovery rate (FDR) of 0.01 was reached. The FDR was estimated by the method of random permutation, which compares the test statistics with statistics generated from data where

the group assignments of measured values were randomly swapped.

These criteria resulted in a list of 1419 proteins, or roughly 18 % of the observable HUVEC proteome, which are differentially expressed on at least one of the examined matrices or time points.

For visualization of these proteins, all logarithmized SILAC ratios were replaced by their corresponding standard scores $z = \frac{x-\mu}{\sigma}$ where x denotes the logarithmized SILAC ratio, μ the mean of all measured values for this protein and σ the standard deviation of these values. The proteins were then hierarchically clustered to receive a dendrogram of proteins with similar regulation patterns. Standard scoring allowed to display the regulation pattern of all proteins in one representation by colour coding (fig. 2.3).

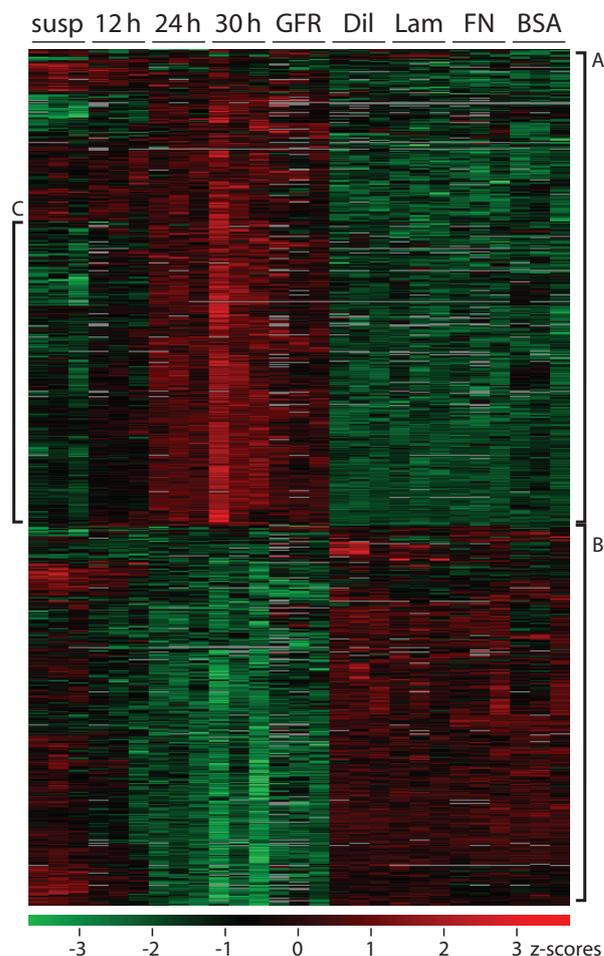


Figure 2.3: Dynamics of the HUVEC proteome.

Hierarchical clustering of differentially regulated HUVEC clusters proteins with similar regulation patterns. Each row represents one protein, the columns represent the different experiments, replicates are placed adjacent to each other. Colour coding is applied according to the z-score, which denotes the relative level of regulation of each protein between the experiments; green indicating down- and red upregulation. Grey colour indicated non determined values.

Two major clusters of proteins are recognizable and indicated with brackets labelled A and B. Further analysis was also done on subcluster C.

2.1.2 Common Features of Clustered Proteins

The analysis of regulated HUVEC proteins revealed two main groups of regulation patterns: one group of 788 proteins (labelled “cluster A” in fig. 2.3) which shows the general tendency to be upregulated on Matrigel, and a cluster of 631 proteins (“cluster B”), which are generally downregulated on Matrigel.

The highest peaks of regulation occurred in the 30 h Matrigel samples. The growth factor reduced Matrigel sample behaved similarly to the regular 24 h Matrigel sample, which is in accordance to the observable phenotype. The other matrices generally resulted in similar regulation patterns, except for some small subclusters of the B cluster, that exhibit upregulation on e. g. diluted Matrigel.

To assess the processes associated with the observed regulation patterns, we conducted analyses based on the Gene Ontology (GO) database, the Kyoto Encyclopedia of Genes and Genomes (KEGG) database of pathways and the Comprehensive Resource of Mammalian Protein Complexes (CORUM) database, whereby subclusters were screened for enrichment of certain annotations. Cluster C was enriched for membrane and extracellular proteins, for proteins of the respiratory chain and for proteins involved in the metabolism of N-glycans. On the other hand, cluster B was enriched for proteins containing pleckstrin homology (PH) domains, proteins involved in regulation of small G-protein signalling, mTOR signalling, protein dephosphorylation, mitosis and regulation of the cytoskeleton.

We were particularly interested in the identification of novel marker proteins for capillary morphogenesis. Therefore we took a closer look at cluster C, a subcluster of A, which contains proteins that are expressed at higher levels only in HUVECs on Matrigel over time, compared to the cells in suspension, while retaining low expression levels on matrices not triggering capillary morphogenesis. We calculated the relative levels of upregulation and gathered the proteins with the highest increases in table 2.2.

Among them were several histone subtypes, enzymes dealing with oxygen species (superoxide dismutase and heme oxygenase 1), some extracellular proteins (lactadherin and multimerin-2), molecules involved in cell adhesion (laminin γ -1, ICAM1 and cadherins) and some known players in angiogenesis or cancer (neuropilin-1, von Willebrand factor and ephrin B2). Interestingly, one previously uncharacterized member of the endosialin family of C-type lectins, CLEC14A, showed a strong pattern of upregulation.

Moreover, hierarchical clustering as shown in fig. 2.3 revealed that the pattern of CLEC14A across all experimental states closely resembled that of von Willebrand factor, a prototypical vessel marker [73], see fig. 2.4.

Table 2.2: Candidate Proteins for Capillary Morphogenesis of HUVECs on Matrigel

Gene Names	12 h	24 h	30 h	GFR	Matr. Dil.	Lam.	FN	BSA
CDH11	4.75	9.79	11.28	9.25	6.60	5.45	9.62	7.83
CDH13	1.23	3.86	5.08	2.74	1.73	1.30	1.68	2.37
CLEC14A	3.15	6.28	8.10	4.94	1.97	1.85	2.00	2.34
ICAM1	5.83	13.78	20.94	11.65	9.99	7.89	13.15	16.54
LAMC1	10.23	13.56	29.42	15.07	3.43	6.13	1.29	1.12
MFGE8	2.87	6.18	9.43	4.31	3.06	2.72	3.49	2.44
MMRN2	1.58	3.13	3.98	2.27	1.35	1.27	1.28	1.25
NRP1	3.56	4.81	5.30	3.20	3.86	2.49	2.10	3.57
VWF	1.30	2.75	4.00	2.26	1.42	1.22	1.24	1.21
EFNB2	7.99	10.74	13.99	11.21	N.D.	4.98	2.76	5.95
CLK3	N.D.	2.18	17.31	7.56	N.D.	1.34	N.D.	0.27
COL6A3	6.95	9.41	9.34	10.32	N.D.	0.81	N.D.	0.66
FGG	N.D.	4.92	24.76	8.64	N.D.	N.D.	0.04	N.D.
H2AFY	7.00	9.66	7.97	14.59	2.96	N.D.	4.38	N.D.
H2BFS	2.82	7.85	6.39	6.32	0.93	1.22	0.71	0.96
HIST1H4A	4.36	9.61	8.97	12.57	1.02	1.46	0.61	1.04
HIST2H2AA3	3.69	13.38	10.66	10.02	1.21	1.52	0.63	1.24
HIST2H2AB	1.61	10.42	10.32	7.26	0.91	1.26	0.94	1.03
HIST2H3A	2.91	6.94	8.23	8.04	2.08	1.68	0.87	1.35
HIST2H3PS2	3.66	15.63	15.73	20.30	1.09	0.17	2.11	2.85
HMOX1	2.49	7.53	11.46	8.68	0.57	0.92	0.50	0.60
SOD2	1.60	4.06	5.16	4.03	2.49	2.55	2.50	2.48

Candidate proteins were selected based on their regulation pattern (see fig. 2.3, cluster C) and on the relative level of upregulation compared to the cells in suspension. Given are the ratios of the medians of the individual normalized SILAC ratios in one experimental group and the cells in suspension. Proteins are indicated by the corresponding gene names.

N.D. = not determined.

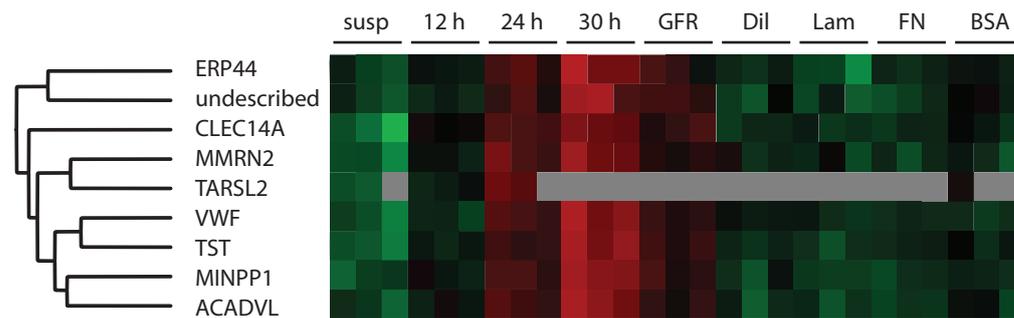


Figure 2.4: Closeup on Hierarchical Clustering. Figure 2.3 zoomed into the region showing CLEC14A and von Willebrand factor (VWF). Colour coding as in fig. 2.3.

2.2 Expression Pattern of CLEC14A

Human CLEC14A is a 490-residue, putative type I transmembrane protein, encoded by a single-exon gene located on chromosome 14. By homology, CLEC14A was categorized as fourth member of the C-type lectin family 14. Up to now, basically nothing is known about this protein. Its significant regulation pattern in HUVECs suggested it as a promising candidate for follow-up analysis. This pattern determined by our proteomic approach was confirmed by classical western blotting (fig. 2.5, A).

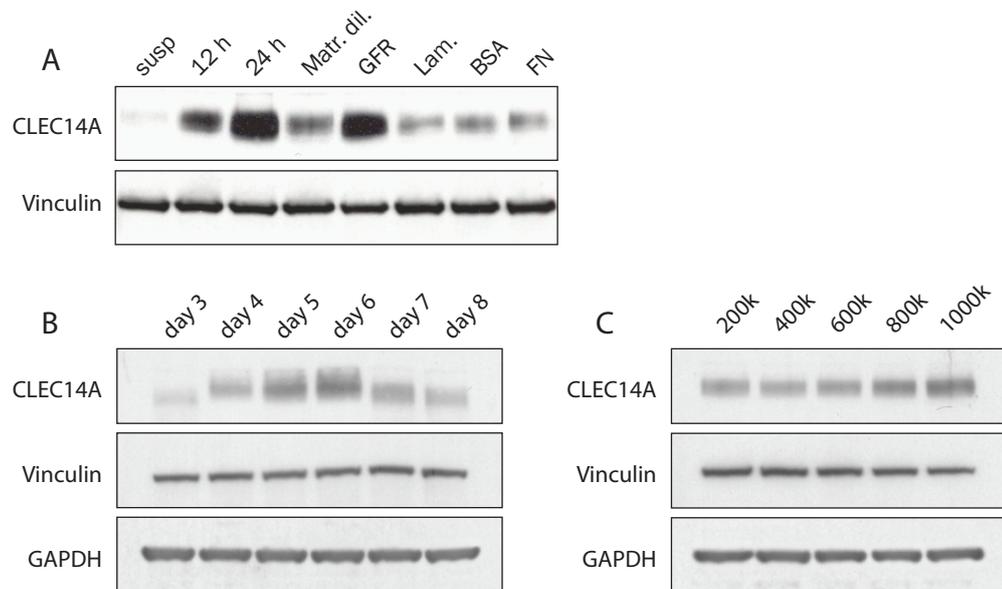


Figure 2.5: Regulation of CLEC14A under different cellular conditions.

A: Confirmation of proteomic data by western blotting. The samples were the same as for mass spectrometric analysis. B: HUVECs were seeded at a 1:2.4 ratio from a confluent dish, grown to confluency (day 3) and then harvested in the course of 5 days. C: HUVECs were seeded at densities of $2 \cdot 10^5$ – $1 \cdot 10^6$ cells per well of a 6-well plate and harvested 24 h later. Western blot for fig. A prepared by S. Zanivan.

Being a homolog of endosialin and being strongly upregulated on Matrigel proposed that this protein might play a role in angiogenesis. From its domain architecture, we suspected it to function in cell adhesion, cell-ECM or cell-cell interaction.

Adherens junctions are the main sites of cell to cell contacts in endothelia and were shown to be dynamically regulated by cell confluency [74]. Cadherins, some of which we could demonstrate to be regulated similarly to CLEC14A (see table 2.2) localize to these junctions. Moreover, the transition from subconfluent to overconfluent endothelial cell culture was suggested to recapitulate aspects of the transition

from the actively remodeling to the mature state of a vessel [74].

To probe a possible function of CLEC14A in this context, we analysed a dependency of CLEC14A expression on cell confluency. For this purpose, we seeded HUVECs subconfluently and harvested samples in the course of several days, while the cultures were growing from the just-confluent to the overconfluent state (fig. 2.5, B). For comparison, HUVECs were seeded at defined cell densities and harvested after 24 h (fig. 2.5, C). CLEC14A levels are increasing over time while the cell culture is growing overconfluent. After 6 days in culture, many cells die, thereby reducing the effective confluency of the culture. In contrast, in cells that are directly seeded at different densities, CLEC14A amounts increase only moderately with confluency. These findings suggest a role of CLEC14A in the maturation blood vessels, rather than during early events in cell adhesion.

Interestingly, besides increasing expression levels, we observed a band upshift correlating with the increase in amount over time. This gave the hint to examine posttranslational modifications of CLEC14A.

2.3 Posttranslational Modifications of CLEC14A

CLEC14A contains one of each of the prototypical C-type lectin, sushi, EGF-like repeat and mucin-like domains present in its paralogs. The mucin-like domain contains many predicted O-glycosylation motifs and one site matching the N-glycosylation consensus. Another possible N-glycosylation site is located in the sushi domain. For its paralog endosialin, O-glycosylation was experimentally shown, but no N-glycosylation could be detected [26].

The intracellular domain contains a series of serines matching kinase motifs. Sequence features of CLEC14A, including domains and possible sites of posttranslational modifications (PTMs) are depicted in fig. 2.6.

Figure 2.6: CLEC14A sequence features.

Colour coding of amino acid residues was done according to the domains in fig. 1.3. The predicted signal sequence and transmembrane region are printed in italics. All fragments that were identified by mass spectrometry throughout this study are printed in boldface. Residues matching the N-glycosylation motif NXS/T are highlighted in cyan, residues predicted to be O-glycosylated by the NetO-Glyc 3.1 Server [75] are indicated in purple and cytosolic residues matching kinase motifs according to phosida [76] are marked green.

```
MRPAFALCLL WQALWPGPGG GEHPTADRAG CSASGACYSL 40
HHATMKRQAA EEACILRGGG LSTVRAGAEI RAVLALLRAG 80
PGPGGSKDL LFWVALERRR SHCTLENEPL RGF5WLSSDP 120
GGLESRTLQW VEEPQRSCTA RRCAVLQATG GVEPAGWKEM 160
RCHLRANGYL CKYQFEVLCP APRPGAASNL SYRAPFLHS 200
AALDFSPPGT EVSALCRGQL PISVTCIADE IGARWDKLSG 240
DVLCPGPGRY LRAGKCAELP NCLDDLGGFA CECATGFELG 280
KDGRCVTSQ EGQPTLGGTG VPTRRPPATA TSPVPQRTWP 320
IRVDEKLGET PLVPEQDQNSV TSIPEIPRWG SQSTMSTLQM 360
SLQAESKATI TPSGSVSKF NSTSSATPQ AFDSSSAVVF 400
IFVSTAVVVL VILTMTVLGL VKLCFHESPS SQPRKESMGP 440
PGLSDPEPA ALGSSSAHCT NNGVKVGDCD LRDRAGALL 480
AESPLGSSDA 490
```

2.3.1 Glycosylation of CLEC14A

Differential protein glycosylation was recently appreciated as a regulatory mechanism for transmembrane receptors, particularly integrins, whose ligand affinities were suggested to vary with their attached glycan structures [77].

To validate possible glycosylation of CLEC14A, we treated HUVEC lysates with different deglycosylating enzymes and monitored the migration pattern of the protein on an SDS gel by western blotting (fig. 2.7). Untreated CLEC14A migrates in a broad band at an apparent molecular mass of roughly 65 kDa, compared to a theoretical mass of 51.6 kDa for the unmodified protein. N-deglycosylation results in a slight downshift of the band, whereas treatment with O-glycosidase has no apparent effect. However, treatment with sialidase (neuraminidase) or a combination of sialidase and O-glycosidase caused a significant downshift of the CLEC14A band. Take together, these results suggest that CLEC14A is both N-glycosylated and O-glycosylated, with the O-glycoside chains being highly substituted with sialic acid.

These substitutions are known to prevent cleavage by O-glycosidase alone. O- and N-glycosylations are likely to be confined to the mucin-like domain, which contains most of the predicted sites. Moreover, no peptide corresponding to this domain was found throughout this study. This can be explained by low accessibility of proteases, low ionizability of glycopeptides and unknown atomic compositions of PTMs, all of which complicate the identification of peptides via mass spectrometry.

Preliminary data indicate that the upshift of the CLEC14A band with increasing confluency (fig. 2.5, B) is not caused by differential N-glycosylation, as PNGase treatment of any sample resulted in an uniform downshift (data not shown).

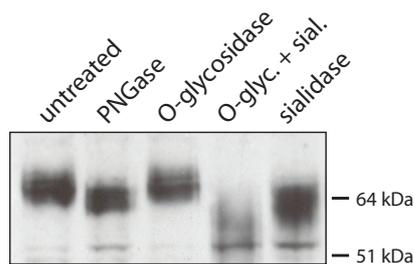


Figure 2.7: Deglycosylation of CLEC14A. HUVEC lysates were treated over night with deglycosylating enzymes and changes in the migration pattern of CLEC14A were analysed by SDS-PAGE and western blotting. PNGase = peptide N-glycosidase.

2.3.2 Phosphorylation of the CLEC14A Cytoplasmic Tail

Phosphorylation is a known mechanism for regulation of the binding of PDZ domain containing proteins to their recognition motifs, as shown for the K^+ -channel Kir 3.2 and the PDZ domain protein PSD-95 [31].

To further characterize CLEC14A, we investigated potential phosphorylation sites on the CLEC14A cytoplasmic tail. We enriched the protein from HUVEC lysate by immunoprecipitation and conducted a targeted mass spectrometric analysis of phosphorylated peptides. For this purpose, collision-induced dissociation in combination with multistage activation [78] was performed: in the case of phosphopeptides, the collisional energy applied to generate MS/MS spectra generally results in neutral loss of phosphoryl moieties, which are attached by rather weak bonds, with the consequential lack of further fragmentation. Multistage activation induces collisional activation of the product ions resulting from these neutral losses, thereby generating more structurally informative fragment ion species and increasing the confidence of peptide identification.

2 Results

Several possible phosphorylation sites were identified: serines 437, 483, 487 and 488. The latter three are located on the same tryptic peptide, which was found in the singly and doubly phosphorylated form. These three residues are in the direct vicinity of PDZ binding motif or on the motif itself, and therefore of special interest concerning potential interactors. Serine 483 could be unambiguously assigned as one location of the modification (fig. 2.8), whereas one could not definitely distinguish between serines 487 and 488 as the other modification site.

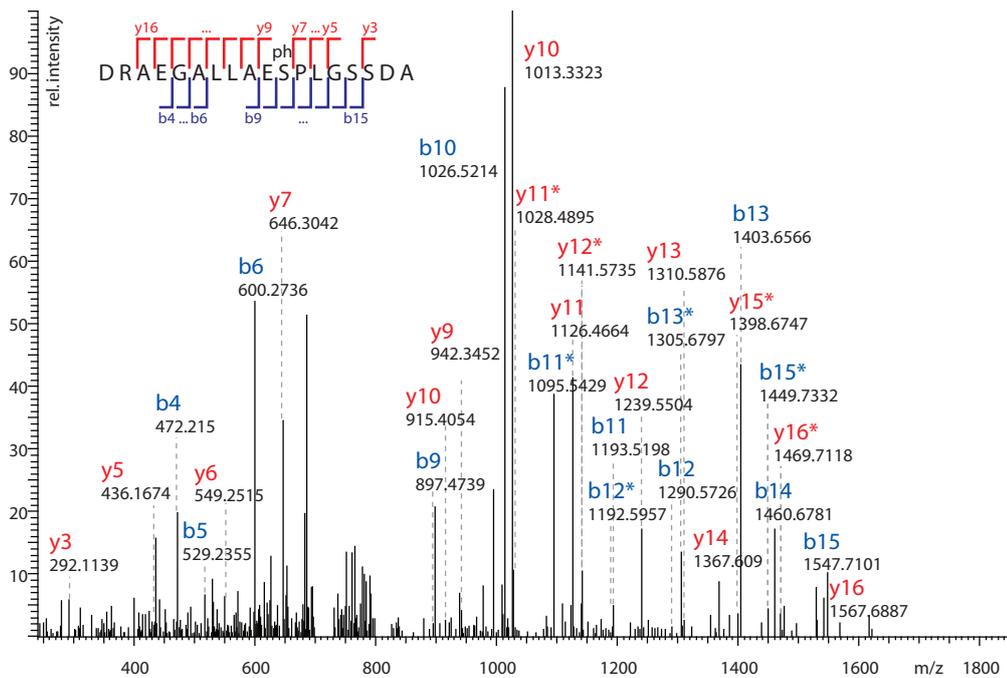


Figure 2.8: MS/MS spectrum of the CLEC14A peptide containing phospho-serine 483. Fragment ions are labelled according to the Roepstorff-Fohlman nomenclature [79]. Fragments with neutral losses of phosphoric acid (-97.977 Da) are marked with an asterisk.

To analyse whether the identified phosphorylation sites are regulated with the cellular state, we used lysates from just-confluent and 5 days overconfluent cultures, and from HUVECs after 20 h on Matrigel as starting material for the enrichment of the protein by immunoprecipitation. In any case, an equivalent amount of protein from SILAC labelled HUVECs was spiked in as internal standard.

Quantification was performed based on the SILAC ratios of the singly phosphorylated peptides KE(ph)SMGPPGLESDPEPAALGSSSAHCTNNGV, DRAEGALLAE(ph)SPLGSSDA, DRAEGALLAESPLG(ph)SSDA, and DRAEGALLAESPLG(ph)SSDA. The latter three contained one missed trypsin cleavage site. The corresponding clea-

ved peptide could not be used since it contains no lysine or arginine and therefore reports no SILAC ratio. Table 2.3 lists the determined protein ratio of CLEC14A compared to the internal standard and the SILAC ratios for the individual phosphopeptides. Data for peptides assigned to phosphoserines 487 and 488 were combined, since we assumed that they correspond to the same peptide. In case of multiple ratio counts, the averages are reported. The ratios were then normalized by the protein ratio, to yield the relative up- or downregulation of the specific site compared to the regulation of the protein. The protein ratio was calculated only from unmodified and methionine-oxidized peptides.

Table 2.3: Regulation of Identified Phosphosites of CLEC14A

examined species	cell state	SILAC ratio L/H	
		measured	normalized
entire protein	confluent	2.13	1.00
	overconfluent	3.69	1.00
	on Matrigel	3.39	1.00
pSer437-peptide	confluent	N.D.	N.D.
	overconfluent	5.30	2.49
	on Matrigel	10.79	5.07
pSer483-peptide	confluent	3.12	1.47
	overconfluent	4.88	1.32
	on Matrigel	2.27	0.67
pSer487/488-peptide	confluent	2.88	1.35
	overconfluent	5.19	1.41
	on Matrigel	5.89	1.74

Normalization was done by division by the ratio for the whole protein.

SILAC ratio L/H = ration of light vs heavy peak of a SILAC pair.

N.D. = not determined.

The numbers show that the relative phosphorylation of serines 487/488 remains virtually unchanged with cell confluency, and is slightly increased on Matrigel. In contrast, phosphorylation of serine 483 is reduced with confluency, and more significantly when the cells are plated on Matrigel, thereby reducing the absolute concentration of serine 483-phosphorylated CLEC14A. Serine 483 matches the recognition motif of glycogen synthase kinase-3 (GSK3), of which isoforms α and β have also been identified in the HUVEC proteome.

The most pronounced changes were detected for the phosphorylation of serine 437, which is increased 2.5-fold compared to the protein in overconfluent cells and more than 5-fold on Matrigel. Serine 437 matches the calcium/calmodulin-dependent protein kinase type II (CAMK2) consensus motif. CAMK2 chains α , δ and γ were found to be expressed in HUVECs.

2 Results

To investigate a possible evolutionary conservation of the identified phosphosites, we aligned the amino acid sequences of the CLEC14A orthologs from different mammals with available genome sequences (fig. 2.9). The cytoplasmic tail is generally highly conserved among mammals, with the notable exception of rodents, in which the C-terminal sequence diverges (in case of the guinea pig) or is truncated by about 20 residues (in case of mouse and rat). Yet, characteristics of the PDZ binding motif are found among all sequences, with the dog and horse sequences lacking the extreme C-terminal amino acid.

Serine 437 is only conserved among primates, whereas serines 483 and 487 are found in all but the mouse and rat sequences. Serine 488 exhibits a lower degree of conservation.

	430	440	450	460	470	480	
mouse	KLCFHKSRSSRTGK	GALDSPGVECD	AEATSLHHSST	QCTDIGVKSG	-----	-----	TVA
rat	KLCFHKSPSSRTRK	GALDSQVECD	AEATSLRHNSA	QCTDTGGK	-----	-----	GA
guinea pig	KLCFHKSPSSQPRK	GPLAPGGESE	AE-----	RSSSAHCTD	NGVKVRDSS	LQGRVEDCS	LAGSPHSSWVT
human	KLCFHESPSSQPRK	ESMGPPGLESD	PEPAALGSS	SAHCTNNGV	KVGD	CDLRDRAEG	ALLAESPLGSSDA
chimpanzee	KLCFHESPSSQPRK	ESMGPPGLESD	PEPTALSSSS	AHCTNNGV	KVGD	CDLRDRAEG	ALLAESPLGSSDA
orang utan	KLCFHESPSSQPRK	ESMGPPGLESD	PEPAALGSS	SAHCTNNGV	KVGD	CDLRDRAEG	ALLAESPLGSSDA
rhesus	KLCFRKSPSSQPRK	ESVGGPLESD	PEPAALGSS	SAHCTNNGV	KVGD	CGLRDRAEG	TLLAESPLGSSDA
marmoset	KLCFHKSPSSQPRK	GPMGPQGLESD	AEPALGSNS	AHCTNNGV	KVGD	CGLRDRAEG	ALLAESPLGSSDA
dog	KLCFHKSPSSQPRK	GPLAQPGVES	DTDAALRSS	SAQCTDYG	VKVGDCGL	RDRAEGASL	TESSLGSGD
horse	KLCFHKSPSSRPRK	GLAPPGVESD	AEAAALRST	SAHCTDNG	VNVGDRGL	RDRAERASL	TESSLGSSD
cat	KLCFHKSSSSQSRK	EPLAQPGMES	GTAAALRSS	SVHCTDNG	VKVGDCGL	RDRAEGASL	TGSSLGSGDT
cow	KLCFHKSPSPQQRK	GPLASPGTEG	DAEAAPLSS	SARSTDNG	VKAGDCGL	RDKVEGASL	TGSSLGSGDT
rabbit	KLCFHKDPSTRPRK	GPLALPGVES	NTEAN-ARSS	SAHCT-----	-----	DNGVRDRAEG	ASLTGSSLDGSDT

Figure 2.9: Alignment of the Cytoplasmic Tail of CLEC14A Orthologs. Serine residues found to be phosphorylated in human CLEC14A are marked in red. The amino acid numbering is that of the human sequence.

2.4 Interaction Partners of CLEC14A

The elucidation of the signalling and interaction network of CLEC14A will provide valuable information about its cellular function.

First, we sought to identify interaction partners of the cytoplasmic tail, especially of the PDZ binding motif, that respond to the phosphorylation state of the protein. For this purpose, we employed the established technique of peptide pulldowns using baits consisting of modified and unmodified synthetic peptides, corresponding to the C-terminal 16 amino acids of CLEC14A. Since the PDZ binding motif is located at the extreme C-terminus of the protein, it was reasonable to assume that physiological interactors would, at least in part, also interact with this relatively short peptide construct.

The mass spectrometry-based technique has its strength in identifying PTM-specific binders against a background of abundant contaminants. It was successfully applied in our lab for the identification of proteins binding to methylated histone tails [80] or tyrosine-phosphorylated peptides derived from insulin receptor substrates [81]. In contrast to these studies, which relied on SILAC labelling of cells, we followed a label free approach.

A specific PTM interactor will generally bind to the unmodified site with much less affinity, resulting in very high ratios in SILAC-based experiments. Unlike in proteomic approaches that seek to determine precise relative expression levels, numeric values are not critical for a qualitative assessment of the interaction. The label free approach therefore yields sufficient accuracy and supersedes the necessity for labelling cells.

Label free protein quantification using MaxQuant (*J. Cox et al.*, submitted to Nature Methods) determines ratios for individual peptides by comparing the corresponding integrated ion currents. Normalization is delayed until all peptide signals, which may be spread over several MS runs, have been acquired. Normalization requires a background of unspecific binders, that are pulled down irrespective of the different baits. Protein ratios are finally calculated by averaging peptide ratios.

Phosphorylation of serine 483 is promising to have physiological significance, since its levels are markedly reduced in cells on Matrigel, suggesting an inhibiting role during capillary morphogenesis, or a role during early events of this process. We performed peptide pulldowns using the peptides biotin-SGAEGALLAESPLGSSDA and biotin-SGAEGALLAE(ph)SPLGSSDA, coupled to streptavidin sepharose beads, as baits.

The experiment was carried out multiple times for the purpose of statistical analysis. We noticed that similar results could be obtained with a range of similar experimental procedures and therefore included all datasets from the optimization

2 Results

phase of the pulldown protocol into the final statistical analysis. The protocols differed slightly in the confluency of the HUVEC culture used, the number of washes, the concentration of Igepal in the washing buffers and the amount of beads, but generally followed the described protocol (section 4.2.1).

Dealing with protein-protein interaction, it is conceivable that some specific interactors are only identified in one sample, representing their physiological interactor, but not in the other. Given the functioning of the mass spectrometer, it will at the same time happen that other proteins are identified in just one of the samples by chance. To tackle this issue, all measured intensity values were log₂ transformed and missing values imputed with random numbers of a normal distribution centered at the detection limit of the mass spectrometer. Then, a t-test for equal group variance was applied and significant interactors identified by plotting the t-test p-values against the average protein ratios (volcano plot, fig. 2.10 B).

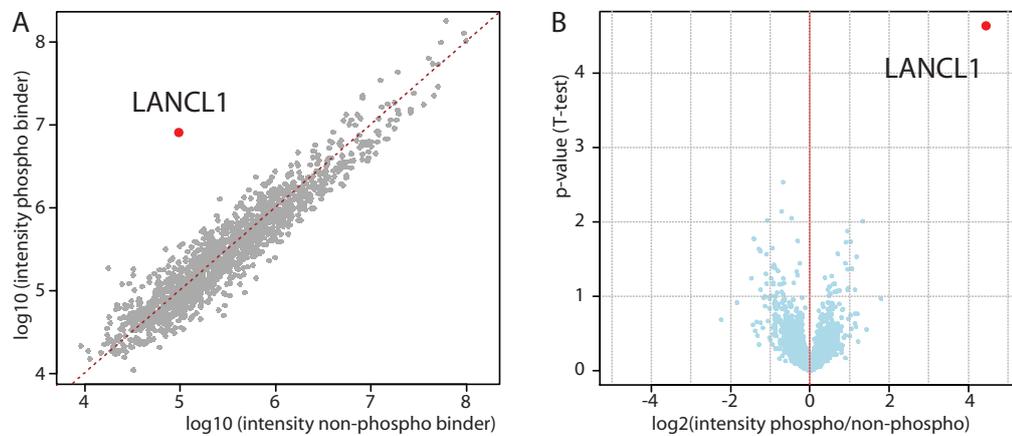


Figure 2.10: Peptide Pulldown of Phospho-Serine 483 Interactors.

A. Typical result of a single pulldown experiment. The normalized intensities resulting from the phospho- and non-phospho pulldown are plotted against each other. Unspecific binders appear on the bisecting line. Specific binders of the phosphorylated bait appear on the upper left, while binders of the unphosphorylated peptide appear on the lower right side of this line.

B. Volcano plot of a combination of 8 peptide pulldowns. The negative logarithmic P-values of the t-test are plotted against the average protein ratios. In this representation, significant interactors of the phosphorylated peptide appear on the right and separated from the cloud of proteins.

We identified LanC-like protein 1 (LANCL1) as specific interactor of the peptide carrying phosphorylated serine 483.

To evaluate a possible role of LANCL1 in capillary morphogenesis, we reassessed the regulation pattern of LANCL1 in our HUVEC proteome dataset. The protein was identified throughout all experimental conditions. Its regulation pattern, shown in table 2.4, did not pass the ANOVA test for significantly regulated proteins. However the tendency of slight downregulation in cells on Matrigel correlates with the regulation pattern of the phosphorylation of serine 483 of CLEC14A (table 2.3), thereby substantiating the significance of the CLEC14A-phSer483-LANCL1 interaction.

Table 2.4: Regulation Pattern of LANCL1

Gene Names	12 h	24 h	30 h	GFR	Matr. Dil.	Lam.	FN	BSA
LANCL1	1.04	0.79	0.74	0.88	1.06	1.26	0.92	0.87

Regulation pattern of LANCL1. Values are calculated as in table 2.2

3 Discussion

This study provides a comprehensive dataset of proteins expressed in human umbilical vein endothelial cells and their regulation pattern on various extracellular matrices. We identified almost 8000 proteins, thereby surpassing the number of previous identifications of HUVEC proteins [82, 83]. Our dataset allows in-depth bioinformatic, cell biological and biochemical follow-up at an extent which would by far exceed the scope of this thesis.

We identified C-type lectin family 14 member A as a promising candidate as novel vessel marker. CLEC14A is highly upregulated in HUVECs during the process of capillary formation on Matrigel. Similar levels of upregulation can be observed on a larger time scale while HUVEC cultures are growing overconfluent.

Retrospect analysis of available proteomic datasets from our lab and western blot detection indicate that CLEC14A is missing in commonly used cell lines such as HeLa and U2OS, and in a range of breast cancer-derived cells of non-endothelial origin, thereby consolidating CLEC14A as an endothelial marker.

Two recent large scale studies identified CLEC14A in the context of cancer. One recent report suggested CLEC14A (referred to as EGFR5) as a potential breast cancer biomarker, based on hypermethylation of CpG islands in several breast cancer cell lines [84]. Hypermethylation generally causes silencing of a gene, which is in accordance with our analysis of breast cancer cell lines.

The other mRNA based study found CLEC14A to be differentially expressed in malignant breast cancer samples compared to benign tumours or surrounding tissue, with CLEC14A being among the top downregulated genes in cancer [85]. The vessel marker CD31 (PECAM1) also appeared in this list of genes, suggesting that that different degrees of vascularization of the tumours compared to benign tissues are the reason for this finding.

Regulation of protein function by differential glycosylation was recently suggested for ECM receptors [77]. We showed that CLEC14A is a glycoprotein and we observed that PTMs might play a role during capillary formation, shown by a shift of the CLEC14A band in western blots. Our results indicate that N-glycosylation is not involved, while we need experimental confirmation for O-glycosylation.

Furthermore, we identified several phosphorylation sites on the short cytoplasmic tail of CLEC14A and could demonstrate that the extent of phosphorylation is differentially regulated depending on cell confluency and the capillary formation on Matrigel.

The extreme C-terminus of the protein is predicted to function as a PDZ binding motif. We identified LANCL1 to be a specific interactor of a peptide derived from the C-terminus of CLEC14A phosphorylated at serine 483. Further investigation will be necessary to characterize this interaction and substantiate whether it is dependent on the PDZ binding motif. The differences of the cytoplasmic tails of CLEC14A orthologs in primates and rodents indicate substantial differences in downstream signalling of CLEC14A and suggest to compare the interactomes of the human and mouse cytoplasmic domains.

Little is known about the physiological function of LANCL1. The protein belongs to the family of LanC-like proteins, which were suggested to catalyse modification reactions on peptides [86]. It was initially identified in an interaction experiment similar to our approach, using the C-terminus of the erythrocyte membrane protein stomatin as bait [87].

In 2007, LANCL1 was appreciated for its high affinity for glutathion [88] and could recently be cocrystallized with this molecule [89]. These findings call one study into question which reported the interaction of LANCL1 with the *Plasmodium* protein PfSBP1 during malaria infection [90]; the authors used GSH-coupled beads in combination with glutathion S-transferase (GST)-tagged PfSBP1.

Glutathion binding affinity links LANCL1 to the redox state of the cell, which is relevant for angiogenesis, as the induction of hypoxia inducible factor 1 α (HIF1A), which controls VEGF expression, is regulated in part by the balance of oxidized and reduced GSH [91].

The authors of the LANCL1 crystal structure identified the SRC homology 3 domain (SH3) of EPS8 as interaction partner of LANCL1 [89], which placed LANCL1 in the context of cell signalling. LANCL1 mutants defective in EPS8 interaction were shown to inhibit nerve growth factor-mediated neurite outgrowth.

So far, no PDZ domain containing proteins could be identified as interactors of the CLEC14A C-terminus. The protein GIPC1 in particular, which was reported to interact with homologous PDZ binding motifs of the endosialin family member C1qR and integrin α -5 [27, 29], and which is expressed in HUVECs, was not found among the interactors. Possibly the short peptide construct does not suffice to recapitulate all physiological interactions. In the case of C1qR, GIPC1 required a juxtamembrane region of the cytoplasmic tail for interaction, in addition to the actual PDZ binding motif [27].

We are working on the establishment of the interaction network of CLEC14A. In addition to intracellular interactors, we are searching for extracellular binding partners. Preliminary results from immunoprecipitation studies provide evidence for an interaction between CLEC14A and multimerin-2 and laminin subunit α -4 (S. Zanivan, unpublished data). Interestingly, both of these proteins are upregulated in HUVECs on Matrigel, with multimerin-2 appearing in the cluster of significantly regulated proteins in direct vicinity of CLEC14A (fig. 2.4).

To gain further insight into the molecular function of CLEC14A, we plan to analyse the phenotype of cells lacking CLEC14A expression. We could successfully knock down the protein by RNA interference and are currently exploring differences in cell adhesion and capillary morphogenesis by functional assays, fluorescence microscopy and the determination of the proteomic changes in CLEC14A negative cells.

We are currently performing experiments to assess the significance of CLEC14A *in vivo*. Preliminary immunohistochemical studies on tissue sections suggest that expression of CLEC14A is confined to blood vessels and possibly glands (S. Zanivan and P. Ostasiewicz, unpublished data) in both tumour and healthy tissue. Similar experiments are being performed using a mouse model of pancreatic islet carcinoma, which progresses in a series of defined states that allow to monitor the vascularization of the tumours (“Rip-Tag mouse”, [92, 93]). Preliminary data are in accordance with the general picture in human tissue (F. Maione, Institute for Cancer Research & Treatment, Candiolo, Italy; unpublished data) and we plan to explore if CLEC14A levels depend also on the stage of vessel development. We thereby hope to substantiate our *in vitro* finding of CLEC14A being upregulated during capillary morphogenesis and to endorse the protein as a novel regulator of vessel formation *in vivo*.

4 Materials and Methods

4.1 Cell Culture

4.1.1 Isolation and Culture of Human Umbilical Vein Endothelial Cells

Human umbilical vein endothelial cells (HUVECs) were cultured on polystyrene cell culture dishes (Corning, coated with 1 % gelatin) in medium 199 GlutaMax (Gibco Invitrogen), supplemented with 20 % fetal bovine serum (FBS, Gibco Invitrogen, heat inactivated at 56 °C for 30 min), 100 µg/ml heparin (sodium salt, Sigma-Aldrich), 0.2 % (v/v) bovine brain extract, and 100 U/ml Penicillin, 100 µg/ml Streptomycin (Invitrogen).

HUVECs were extracted by a method derived from the Jaffe protocol [94]. Umbilical cords were kindly supplied by W. Eiermann, (Rotkreuzklinikum München) and S. Pavan (Institute for Cancer Research and Treatment, Candiolo, Italy) and stored at 4 °C for up to 3 days prior to extraction. The umbilical cords were washed with PBS, then both ends were cut, blunt needles were inserted into the vein and clamped and the vessels were flushed with PBS to remove the blood. Endothelial cells were detached from the vessel walls by incubation with 0.2 % (w/v) Collagenase A from *Clostridium histolyticum* (Roche) in medium 199 for 20 min at room temperature. The cells were washed out with ~ 40 ml of PBS, collected by centrifugation at 1500 ×g for 5 min, resuspended in complete medium 199 (i. e. medium with all supplements) and seeded on gelatin-coated 15 cm dishes.

HUVECs were subcultured by standard protocols [95] seeding the cells at ratios of 1:2 to 1:6.

4.1.2 Cell Lysis

Cells were generally lysed in modified RIPA buffer (150 mM NaCl, 50 mM Tris·HCl pH 7.5, 10 % (v/v) glycerol, 1 mg/ml sodium deoxycholate, 1 % Igepal, protease inhibitor (Roche)) on ice for 30 min. Cellular debris was spun down in a cooled tabletop centrifuge at maximum speed and supernatants were collected and stored at –20 °C. Prior to further processing, protein amounts in cell lysates were quantified by the Bradford method [96] using Biorad's Quick-Start reagent with BSA as standard.

4.1.3 SILAC Labelling of HUVECs

For labelling HUVECs with stable amino acids in cell culture, 3–5 confluent primary P0 cultures were pooled and seeded at a ratio of 1:3–1:4 onto 10 cm dishes. After one day, the cells were switched to an analogous SILAC medium. This medium was composed from a custom-made 10× stock of medium 199 without arginine and lysine (Gibco Invitrogen), supplemented with 0.15 % v/w NaHCO₃ to adjust the pH, 2 mM L-glutamine and isotope labelled amino acids L-arginine-¹³C₆ ¹⁵N₄ hydrochloride (Sigma) and L-lysine-¹³C₆ ¹⁵N₂ hydrochloride (Sigma) at concentrations of 42 and 73 μg/ml, respectively. The cultures were grown to confluency, while changing the medium every other day, and then subcultured as before. After 2 days at passage 4, the cells were switched to a medium composed with dialysed FBS (dialysed against 150 mM NaCl using a 1 kDa cutoff dialysis tubing (Spectra/Por 7, Spectrum Labs)) and labelled amino acids at concentrations of 56 and 97 μg/ml for Arg10 and Lys8, respectively. At the end of each passage, one dish was lysed to check the level of incorporation of heavy amino acids.

4.1.4 Matrigel Assay

To analyse the formation of capillary-like structures, a confluent culture of HUVECs was starved over night in M199 supplemented with only 10 % FBS and antibiotics. The cells were then trypsinized and resuspended in starvation medium. Reference aliquots of cells in suspension were harvested by centrifugation and lysed immediately (*vide infra*). The remaining cells were seeded at a density of 1 200 000 cell per dish on 10 cm dishes coated with either Matrigel or several other matrices. Dishes were coated with 1 ml Matrigel Basement Membrane Matrix (BD Biosciences), diluted with 10 % M199. The matrix was applied beforehand at 4 °C and allowed to solidify for 45 min at 37 °C.

Additional coatings included 10 μg/ml fibronectin in PBS, 10 μg/ml laminin in PBS, Matrigel diluted 1:1000 in medium 199 and 3 % BSA in PBS. These coatings were applied over night at 37 °C, except for BSA, which was applied for 1 h at 37 °C.

Cells on control matrices were all harvested after 24 h by scraping the cells in lysis buffer. Cells on Matrigel were harvested after 12, 24 and 30 h by dissolving the matrix with MatriSpense cell recovery solution (BD Biosciences) according to the manufacturers recommendations. Cells were then pelleted by centrifugation and lysed.

In case of whole proteome samples intended for quantitative mass spectrometry, aliquots representing 100 μg of protein were mixed with 50 μg of SILAC-labelled internal standard, precipitated using the methanol-chloroform method [97] and resuspended in 1× LDS sample buffer (Invitrogen).

4.2 Pulldown Assays

4.2.1 Peptide Pulldown

Peptide pulldowns were carried out using synthetic, N-terminally biotin coupled peptides biotin-SGAEGALLAESPLGSSDA and biotin-SGAEGALLAE(ph)SPLGSSDA. Synthesis was performed by the core facility of the MPI for Biochemistry, Martinsried, Germany on Applied Biosystems 433 A automated peptide synthesizers using Fmoc chemistry. The peptides were lyophilized to remove organic solvent and then resuspended in 50 mM Tris·HCl, 150 mM NaCl, adjusting the pH to 7.5.

Per pulldown sample, 50 μ l of streptavidin sepharose beads (30 % (v/v) slurry, GE Healthcare) were washed twice in base buffer (150 mM NaCl, 50 mM Tris·HCl, pH 7.5, 10 % glycerol, 0.1 % Igepal), coupled with approximately 0.5 mg of peptide in 500 μ l base buffer for 30 min at room temperature, and washed again three times in base buffer. Between the washing steps, the beads were pelleted by centrifugation for 1 min at 1200 \times g.

Confluent HUVEC dishes were lysed in lysis buffer (base buffer, supplemented with 100 mM Na₃VO₄, protease and phosphatase inhibitors without EDTA (Roche), otherwise following the general cell lysis protocol (section 4.1.2). An equivalent of 0.5–3 mg of cellular protein was incubated with 50 μ l of washed, uncoupled streptavidin sepharose beads for 30 min at 4 °C on a rotating wheel. The precleared supernatant was then added to the peptide-coupled beads and incubated likewise for 2–3 h.

Finally, the supernatant was removed and the beads washed twice with base buffer supplemented with PhoStop tablets, and three times with base buffer containing only 50 mM NaCl. The beads were resuspended in 20 μ l of 2 \times LDS sample buffer and boiled for 5 min, prior to separation by SDS-PAGE.

4.2.2 Immunoprecipitation

Immunoprecipitation of CLEC14A was carried out for enrichment of the protein for targeted MS analysis of individual tryptic peptides. Per sample, 6 μ g of α CLEC14a antibody were coupled to 40 μ l of Dynabead slurry (Invitrogen). The beads were washed beforehand 2 \times in 100 mM sodium acetate, pH 5.0, 0.1 % Igepal, then mixed with the antibody solution plus 1 volume of sodium acetate buffer and shaken for 1 h at room temperature. After washing twice with sodium acetate buffer, the beads were resuspended in 200 mM sodium borate, pH 9.0, 0.1 % Igepal. For crosslinking, the beads were incubated with 25 mM dimethyl pimelimidate in sodium borate buffer for 45 min, and then washed again. To block remaining reactive groups, 200 mM ethanolamine in PBS, 0.1 % Igepal, pH 8.0 were added and the

beads shaken for 1 h. To remove residual uncoupled antibody, the beads were briefly washed with 50 mM glycine, pH 2.6, 0.1 % Igepal, then immediately neutralized with 50 mM Tris, 0.1 % Igepal, pH 7.5.

HUVECs were lysed as described, the lysate mixed with an equivalent amount of SILAC labelled lysate and diluted to a concentration of 1 mg/ml with lysis buffer. An equivalent of 3 mg of protein was precleared by incubation with 15 μ l of washed dynabeads for 30 min at 4 °C. Then, the precleared lysate was directly added to the antibody-coupled dynabeads and incubated for 1 h at 4 °C on a rotating wheel.

The beads were then washed twice with lysis buffer, twice with lysis buffer lacking detergents and inhibitors, and then resuspended in 20 μ l of 2 \times LDS sample buffer with 10 mM dithiothreitol (DTT) and boiled. The supernatant was loaded on an SDS gel (*vide infra*) and the band corresponding to CLEC14A excised for further in gel digestion and MS analysis.

Aliquots of the precleared and depleted lysates were taken for reference.

4.3 Mass Spectrometry and Data Analysis

4.3.1 1D-SDS-PAGE and In-Gel Digestion of Proteins

One-dimensional SDS polyacrylamide gel electrophoresis (1D SDS-PAGE) was used to reduce sample complexity, followed by tryptic in-gel digestion [98]. The protein samples were separated using NuPage Novex 4–12 % Bis-Tris gels and NuPage MOPS running buffer (Invitrogen) according to the manufacturer's instructions. The gels were stained with colloidal Coomassie (Invitrogen). For whole proteome samples, 18 slices were defined to yield equal protein amounts per slice according to visual observation of the staining intensity. Gels with samples from pulldown experiments were sliced to 4 slices, while eliminating highly abundant contaminant bands. Gel slices were cut into 1 mm³ cubes, washed thoroughly 2 \times with 25 mM ammonium bicarbonate (ABC), 50 % (v/v) acetonitrile (ACN) to remove the Coomassie staining and then dehydrated by washing with 100 % ACN. Remaining solvent was removed in a vacuum concentrator. The dry gel pieces were rehydrated with 10 mM DTT in 50 mM ABC and incubated for 1 h at 56 °C for reduction of disulfides. The resulting free thiol groups were subsequently alkylated by incubating the samples with 55 mM iodoacetamide in 50 mM ammonium bicarbonate for 45 min at 25 °C in the dark. The gel pieces were washed twice with 50 mM ammonium bicarbonate, again dehydrated with 100 % ACN, and dried in a vacuum concentrator. Then the gel pieces were rehydrated with 12.5 ng/ μ l of trypsin (sequencing grade, Promega) in 50 mM ammonium bicarbonate and incubated for 16–18 h at 37 °C for in-gel protein digestion.

Supernatants were transferred into fresh tubes, and the remaining peptides were extracted by shaking the gel pieces in 30 % ACN with 3 % TFA, followed by dehydration with 100 % ACN. The latter two steps were repeated. The extracts were combined, and organic solvent was removed in a vacuum concentrator. Desalting and concentration were carried out on RP-C₁₈ (Empore disks, 3M) StageTips [44]. The protocol by Rappsilber *et al.* was upscaled ~ 5-fold, using 2 layers of RP-C₁₈, adjusting to the sample amount. Samples were stored at 4 °C.

4.3.2 Nanoflow HPLC-MS/MS

Peptide mixtures were eluted from StageTips as described [44] and analysed by online reversed-phase nanoflow liquid chromatography tandem mass spectrometry (LC-MS/MS) on an EASY-nLC system (Proxeon Biosystems, Odense, Denmark) connected to the LTQ Orbitrap XL instrument (Thermo Fisher Scientific, Bremen, Germany) equipped with a Proxeon nanoelectrospray ion source. Chromatographic separation of the peptides took place in a 15 cm analytical column (75 µm inner diameter) in-house packed [99] with reversed-phase Repronil Pur C₁₈-AQ 3 µm resin (Dr. Maisch GmbH, Ammerbuch-Entringen, Germany). Peptide mixtures were autosampled onto the column with a flow of 500 nl/min and subsequently eluted with a flow of 250 nl/min from 5 % to 70 % ACN in 0.5 % acetic acid, in a 140 min gradient. The effluent from the HPLC column was directly electrosprayed into the mass spectrometer. The instrument was operated in data-dependent mode to automatically switch between full scan MS and MS/MS (MS²) acquisition. Survey full scan MS spectra (from m/z 300–1800) were acquired in the Orbitrap with a resolution of 60 000 at m/z 400 (after accumulation to a target value of 10⁶ charges in the linear ion trap) using the lock mass option for internal calibration of each spectrum [48]. The five most intense ions were sequentially isolated for fragmentation in the linear ion trap using collisionally induced dissociation (CID) with a normalized collision energy of 35 % at a minimum required signal of 1000. Target ions already selected for MS/MS were dynamically excluded for 90 s. The resulting fragment ions were recorded in the linear ion trap with unit resolution.

In case of targeted analysis of CLEC14A derived peptides, the instrument was programmed to sequence exclusively peptides on an inclusion list, representing all masses derived from the (optionally modified) C-terminal fragment. Additionally, multistage activation [78] was triggered when a neutral loss was detected, which corresponded to the mass of one phosphoryl group.

4.3.3 Peptide Identification and Quantification

Mass spectra were processed using the MaxQuant software (version 1.0.14.3) [100, 101]. Protein and peptide identification was performed using the Mascot search engine (version 2.2.2, Matrix Science [102]) by querying the concatenated forward and reversed International Protein Index protein sequence database (IPI version 3.62) plus 175 commonly observed contaminants. The minimum required peptide length was set to 6 amino acids, and trypsin was selected as specific enzyme [58] (cleavage at Arg-Pro and Lys-Pro was included); two missed cleavages were allowed. Carbamidomethylation (Cys) was set as fixed modification, whereas oxidation (Met) and N-acetylation were considered as variable modifications. Mass accuracies were set to 7 ppm for the parent ion and 0.5 Da for fragment ions.

For the relative quantification of the peptides against their SILAC-labeled counterparts, our in-house developed software MaxQuant was used [100]. For protein identification, the maximum protein and peptide false discovery rate (FDR) was set to 0.01. Proteins were considered identified with at one peptide uniquely assigned to the respective sequence.

4.4 Protein Biochemistry

4.4.1 Deglycosylation Assay

To assess the glycosylation state of CLEC14A, HUVEC lysates corresponding to 20 μ g of protein were diluted 2-fold in 150 mM sodium phosphate, pH 7.5, 25 mM EDTA. Then, 1 U of peptide N-glycosidase F (PNGase), 1 mU of O-glycosidase or 1 mU of Sialidase (all enzymes from Roche) were added and incubated for 18 h at 37 °C. LDS sample buffer was added and the samples were boiled prior to separation by SDS-PAGE and western blotting.

4.4.2 Western Blotting and Immunodetection

For western blot detection, protein samples corresponding to 20 μ g of total lysate were first separated by one-dimensional SDS polyacrylamide gel electrophoresis as described (section 4.3.1) and then electrophoretically transferred [103] onto PVDF membranes using wet blotting cells from BioRad at a constant voltage of 100 V for 1 h in transfer buffer (192 mM glycine, 20 mM Tris base, 20% (v/v) methanol). After blotting, the membranes were transiently stained with Ponceau S [104], destained in water and dried.

Prior to antibody detection, the membranes were rehydrated in Tris buffered saline with 1 % Tween-20 (TBS-T). Primary antibodies (see table 4.1) were diluted in TBS-T and incubated with the membrane for 1 h at room temperature. The membranes were then washed 3× with TBS-T for 5 min before adding the secondary antibody accordingly for 45 min. Finally, the membranes were washed thoroughly 3× in TBS-T. Western detection was performed using ECL detection reagents and films from GE Healthcare.

Table 4.1: Antibodies

Antibody	Type	Specificity	Dilution	Reference
α -CLEC14A	sheep IgG polyclonal	human, extracellular domain	1:3000	R&D Systems
α -GAPDH	rabbit IgG polyclonal	human	1:1000	Cell Signaling
α -vinculin	mouse IgG monoclonal	human	1:800	Sigma
α -sheep	donkey IgG HRP-coupled	sheep	1:5000	R&D Systems
α -rabbit	donkey IgG HRP-coupled	rabbit	1:3000	GE Healthcare
α -mouse	sheep IgG HRP-coupled	mouse	1:5000	GE Healthcare

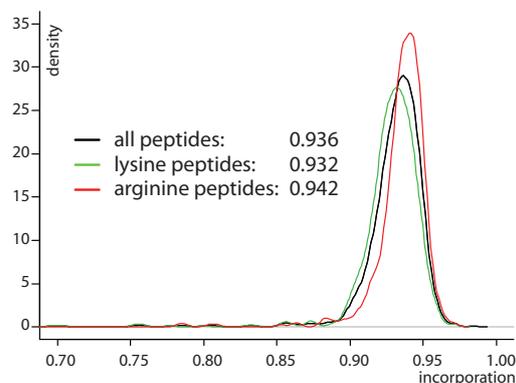
A Appendix

A.1 SILAC Labelling of HUVECs

Lysates from SILAC labelled HUVECs were used as internal standard throughout this study. The primary nature of HUVECs make complete labelling challenging: They cannot be kept in culture for an unlimited number of passages without losing their primary characteristics. Therefore, labelling had to be restricted to four passages. Moreover, cell growth was impaired when using dialysed serum. For this reason, cells were labelled in medium with regular FBS and only switched to an analogous medium containing FBS dialysed against a 1 kDa cutoff during the last passage.

Generally, 2–5 confluent primary cultures were pooled to reduce variability and switched to SILAC medium after one additional day. At the end of the fourth passage, incorporation of heavy amino acids generally reached ~93–95 % (fig. A.1).

Figure A.1: Incorporation of Heavy Amino Acids. Arg10, Lys8 labelled HUVEC lysate after 4 passages of labelling was analysed by mass spectrometry. Individual ratios from identified SILAC pairs were converted into a relative incorporation rate $1 - \frac{1}{\text{SILAC ratio} + 1}$ and plotted as density. Groups of peptides containing only lysine or only arginine show a similar incorporation as the entirety of peptides.



To reduce overestimation of the “light” peak due to residual unlabelled proteins in the internal standard lysate, sample and standard were mixed at a 2:1 ratio in case of whole proteome samples. Downstream bioinformatic analysis by the MaxQuant software would then normalize the observed ratios to a 1:1 median, thereby reducing this error.

Arginine to proline conversion, which occurs in some cell types and rises challenges in the calculation of SILAC ratios, could not be detected in HUVECs.

References

- [1] Coultas, L., Chawengsaksophak, K., and Rossant, J. *Nature* **438**, 937–945 (2005).
- [2] Risau, W. and Flamme, I. *Annu. Rev. Cell Dev. Biol.* **11**, 73–91 (1995).
- [3] Risau, W. *Nature* **386**, 671–674 (1997).
- [4] Folkman, J. and Kalluri, R. *Nature* **427**, 787 (2004).
- [5] Mathew, R., Karantza-Wadsworth, V., and White, E. *Nat. Rev. Cancer* **7**, 961–967 (2007).
- [6] Guo, W. and Giancotti, F. G. *Nat. Rev. Mol. Cell Biol.* **5**, 816–826 (2004).
- [7] Ellis, L. M. and Fidler, I. J. *Eur J Cancer* **32A**, 2451–2460 (1996).
- [8] Folkman, J. *N. Engl. J. Med.* **285**, 1182–1186 (1971).
- [9] Patan, S. *J. Neurooncol.* **50**, 1–15 (2000).
- [10] Bergers, G. and Benjamin, L. E. *Nat. Rev. Cancer* **3**, 401–410 (2003).
- [11] Yancopoulos, G. D., Davis, S., Gale, N. W., Rudge, J. S., Wiegand, S. J., and Holash, J. *Nature* **407**, 242–248 (2000).
- [12] Gavard, J. *FEBS Lett.* **583**, 1–6 (2009).
- [13] Vestweber, D. *Arterioscler Thromb Vasc Biol* **28**, 223–232 (2008).
- [14] Rundhaug, J. E. *J Cell Mol Med* **9**, 267–285 (2005).
- [15] Silva, R., D’Amico, G., Hodivala-Dilke, K. M., and Reynolds, L. E. *Arterioscler Thromb Vasc Biol* **28**, 1703–1713 (2008).
- [16] Eichhorn, M. E., Kleespies, A., Angele, M. K., Jauch, K.-W., and Bruns, C. J. *Langenbecks Arch Surg* **392**, 371–379 (2007).
- [17] Folkman, J. and Haudenschild, C. *Nature* **288**, 551–556 (1980).
- [18] Kleinman, H. K., McGarvey, M. L., Liotta, L. A., Robey, P. G., Tryggvason, K., and Martin, G. R. *Biochemistry* **21**, 6188–93 (1982).
- [19] Kleinman, H. K., McGarvey, M. L., Hassell, J. R., Star, V. L., Cannon, F. B., Laurie, G. W., and Martin, G. R. *Biochemistry* **25**, 312–318 (1986).
- [20] Kubota, Y., Kleinman, H. K., Martin, G. R., and Lawley, T. J. *J. Cell Biol.* **107**, 1589–1598 (1988).
- [21] Kleinman, H. K. and Martin, G. R. *Semin Cancer Biol* **15**, 378–386 (2005).
- [22] Arnaoutova, I., George, J., Kleinman, H. K., and Benton, G. *Angiogenesis* **12**, 267–274 (2009).
- [23] Auerbach, R., Lewis, R., Shinnars, B., Kubai, L., and Akhtar, N. *Clin Chem* **49**, 32–40 (2003).
- [24] Dodd, R. B. and Drickamer, K. *Glycobiology* **11**, 71R–79R (2001).
- [25] Drickamer, K. (2006). <http://www.imperial.ac.uk/research/animallectins/ctld/mammals/Groups/GroupXIV.html>.
- [26] Christian, S., Ahorn, H., Koehler, A., Eisenhaber, F., Rodi, H. P., Garin-Chesa, P., Park, J. E., Rettig, W. J., and Lenter, M. C. *J. Biol. Chem.* **276**, 7408–7414 (2001).
- [27] Bohlson, S. S., Zhang, M., Ortiz, C. E., and Tenner, A. J. *J Leukoc Biol* **77**, 80–89 (2005).
- [28] Greenlee, M. C., Sullivan, S. A., and Bohlson, S. S. *Curr. Drug Targets* **9**, 130–138 (2008).
- [29] Mourabit, H. E., Poinat, P., Koster, J., Sondermann, H., Wixler, V., Wegener, E., Laplantine, E., Geerts, D., Georges-Labouesse, E., Sonnenberg, A., and Aumailley, M. *Matrix Biol.* **21**, 207–214 (2002).

References

- [30] Valdeabri, D., Caswell, P. T., Anderson, K. I., Schwarz, J. P., König, I., Astanina, E., Caccavari, F., Norman, J. C., Humphries, M. J., Bussolino, F., and Serini, G. *PLoS Biol.* **7**, e25 (2009).
- [31] Cohen, N. A., Brenman, J. E., Snyder, S. H., and Bredt, D. S. *Neuron* **17**, 759–767 (1996).
- [32] Rettig, W. J., Garin-Chesa, P., Healey, J. H., Su, S. L., Jaffe, E. A., and Old, L. J. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 10832–10836 (1992).
- [33] St Croix, B., Rago, C., Velculescu, V., Traversono, G., Romans, K. E., Montgomery, E., Lal, A., Riggins, G. J., Lengauer, C., Vogelstein, B., and Kinzler, K. W. *Science* **289**, 1197–1202 (2000).
- [34] MacFadyen, J. R., Haworth, O., Robertson, D., Hardie, D., Webster, M.-T., Morris, H. R., Panico, M., Sutton-Smith, M., Dell, A., van der Geer, P., Wienke, D., Buckley, C. D., and Isacke, C. M. *FEBS Lett.* **579**, 2569–2575 (2005).
- [35] Nanda, A., Karim, B., Peng, Z., Liu, G., Qiu, W., Gan, C., Vogelstein, B., Croix, B. S., Kinzler, K. W., and Huso, D. L. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 3351–3356 (2006).
- [36] Ho, M., Yang, E., Matcuk, G., Deng, D., Sampas, N., Tsalenko, A., Tabibiazar, R., Zhang, Y., Chen, M., Talbi, S., Ho, Y. D., Wang, J., Tsao, P. S., Ben-Dor, A., Yakhini, Z., Bruhn, L., and Quertermous, T. *Physiol Genomics* **13**, 249–262 (2003).
- [37] Fu, J. Q., Wang, L., Chen, W., Ci, H. L., and Li, Y. P. *Shi Yan Sheng Wu Xue Bao* **37**, 409–417 (2004).
- [38] Bartlett, J. M. S. and Stirling, D. *Methods Mol Biol* **226**, 3–6 (2003).
- [39] Schena, M., Shalon, D., Davis, R. W., and Brown, P. O. *Science* **270**, 467–470 (1995).
- [40] Lashkari, D. A., DeRisi, J. L., McCusker, J. H., Namath, A. F., Gentile, C., Hwang, S. Y., Brown, P. O., and Davis, R. W. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 13057–13062 (1997).
- [41] de Godoy, L. M. F., Olsen, J. V., Cox, J., Nielsen, M. L., Hubner, N. C., Fröhlich, F., Walther, T. C., and Mann, M. *Nature* **455**, 1251–1254 (2008).
- [42] Ingolia, N. T., Ghaemmaghami, S., Newman, J. R. S., and Weissman, J. S. *Science* **324**, 218–223 (2009).
- [43] Aebersold, R. and Mann, M. *Nature* **422**, 198–207 (2003).
- [44] Rappsilber, J., Mann, M., and Ishihama, Y. *Nat. Protoc.* **2**, 1896–1906 (2007).
- [45] Wiśniewski, J. R., Zougman, A., Nagaraj, N., and Mann, M. *Nat. Methods* **6**, 359–362 (2009).
- [46] Ong, S.-E. and Mann, M. *Nat. Chem. Biol.* **1**, 252–262 (2005).
- [47] Makarov, A. *Anal. Chem.* **72**, 1156–1162 (2000).
- [48] Olsen, J. V., de Godoy, L. M. F., Li, G., Macek, B., Mortensen, P., Pesch, R., Makarov, A., Lange, O., Horning, S., and Mann, M. *Mol. Cell. Proteomics* **4**, 2010–2021 (2005).
- [49] Olsen, J. V., Schwartz, J. C., Griep-Raming, J., Nielsen, M. L., Damoc, E., Denisov, E., Lange, O., Remes, P., Taylor, D., Splendore, M., Wouters, E. R., Senko, M., Makarov, A., Mann, M., and Horning, S. *Mol. Cell. Proteomics* **8**, 2759–2769 (2009).
- [50] Kumar, C. and Mann, M. *FEBS Lett.* **583**, 1703–1712 (2009).
- [51] Tress, M. L., Bodenmiller, B., Aebersold, R., and Valencia, A. *Genome Biol* **9**, R162 (2008).
- [52] Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., and Aebersold, R. *Nat. Biotechnol.* **17**, 994–999 (1999).

- [53] Ross, P. L., Huang, Y. N., Marchese, J. N., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S., Purkayastha, S., Juhasz, P., Martin, S., Bartlet-Jones, M., He, F., Jacobson, A., and Pappin, D. J. *Mol. Cell. Proteomics* **3**, 1154–1169 (2004).
- [54] Krüger, M., Moser, M., Ussar, S., Thieversen, I., Luber, C. A., Forner, F., Schmidt, S., Zanivan, S., Fässler, R., and Mann, M. *Cell* **134**, 353–364 (2008).
- [55] Bantscheff, M., Schirle, M., Sweetman, G., Rick, J., and Kuster, B. *Anal. Bioanal. Chem.* **389**, 1017–1031 (2007).
- [56] Ong, S.-E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. *Mol. Cell. Proteomics* **1**, 376–386 (2002).
- [57] Mann, M. *Nat. Rev. Mol. Cell Biol.* **7**, 952–958 (2006).
- [58] Olsen, J. V., Ong, S.-E., and Mann, M. *Mol. Cell. Proteomics* **3**, 608–614 (2004).
- [59] Harsha, H. C., Molina, H., and Pandey, A. *Nat. Protoc.* **3**, 505–516 (2008).
- [60] Forner, F., Kumar, C., Luber, C. A., Fromme, T., Klingenspor, M., and Mann, M. *Cell Metab.* **10**, 324–335 (2009).
- [61] Liu, H., Sadygov, R. G., and Yates, J. R. *Anal. Chem.* **76**, 4193–4201 (2004).
- [62] Mueller, L. N., Brusniak, M.-Y., Mani, D. R., and Aebersold, R. *J. Proteome Res.* **7**, 51–61 (2008).
- [63] Hanke, S., Besir, H., Oesterheld, D., and Mann, M. *J. Proteome Res.* **7**, 1118–1130 (2008).
- [64] Malmström, J., Beck, M., Schmidt, A., Lange, V., Deutsch, E. W., and Aebersold, R. *Nature* **460**, 762–765 (2009).
- [65] Macek, B., Mann, M., and Olsen, J. V. *Annu. Rev. Pharmacol. Toxicol.* **49**, 199–221 (2009).
- [66] Vermeulen, M., Hubner, N. C., and Mann, M. *Curr. Opin. Biotechnol.* **19**, 331–337 (2008).
- [67] Mittler, G., Butter, F., and Mann, M. *Genome Res.* **19**, 284–293 (2009).
- [68] Butter, F., Scheibe, M., Mörl, M., and Mann, M. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 10626–10631 (2009).
- [69] Mann, M. *J. Proteome Res.* **7**, 3065 (2008).
- [70] Pankov, R. and Yamada, K. M. *J. Cell Sci.* **115**, 3861–3863 (2002).
- [71] Astrof, S. and Hynes, R. O. *Angiogenesis* **12**, 165–175 (2009).
- [72] Garlanda, C. and Dejana, E. *Arterioscler Thromb Vasc Biol* **17**, 1193–1202 (1997).
- [73] Zanetta, L., Marcus, S. G., Vasile, J., Dobryansky, M., Cohen, H., Eng, K., Shammian, P., and Mignatti, P. *Int. J. Cancer* **85**, 281–288 (2000).
- [74] Lampugnani, M. G., Corada, M., Andriopoulou, P., Esser, S., Risau, W., and Dejana, E. *J. Cell Sci.* **110** (Pt 17), 2065–2077 (1997).
- [75] Julenius, K., Mølgaard, A., Gupta, R., and Brunak, S. *Glycobiology* **15**, 153–164 (2005).
- [76] Gnad, F., Ren, S., Cox, J., Olsen, J. V., Macek, B., Oroshi, M., and Mann, M. *Genome Biol* **8**, R250 (2007).
- [77] Bellis, S. L. *Biochim. Biophys. Acta* **1663**, 52–60 (2004).
- [78] Schroeder, M. J., Shabanowitz, J., Schwartz, J. C., Hunt, D. F., and Coon, J. J. *Anal. Chem.* **76**, 3590–3598 (2004).
- [79] Roepstorff, P. and Fohlman, J. *Biomed Mass Spectrom* **11**, 601 (1984).
- [80] Vermeulen, M., Mulder, K. W., Denissov, S., Pijnappel, W. W. M. P., van Schaik, F. M. A., Varier, R. A., Baltissen, M. P. A., Stunnenberg, H. G., Mann, M., and Timmers, H. T. M. *Cell* **131**, 58–69 (2007).

References

- [81] Hanke, S. and Mann, M. *Mol. Cell. Proteomics* **8**, 519–534 (2009).
- [82] Bruneel, A., Labas, V., Mailloux, A., Sharma, S., Royer, N., Vinh, J., Pernet, P., Vaubourdolle, M., and Baudin, B. *Proteomics* **5**, 3876–3884 (2005).
- [83] Jorge, I., Navarro, P., Martínez-Acedo, P., nez, E. N., Serrano, H., Alfranca, A., Redondo, J. M., and Vázquez, J. *Mol. Cell. Proteomics* **8**, 1130–1149 (2009).
- [84] Chung, W., Kwabi-Addo, B., Ittmann, M., Jelinek, J., Shen, L., Yu, Y., and Issa, J.-P. J. *PLoS One* **3**, e2079 (2008).
- [85] Potapenko, I. O., Haakensen, V. D., Lüders, T., Helland, A., Bukholm, I., Sørli, T., Kristensen, V. N., Lingjærde, O. C., and Børresen-Dale, A.-L. *Mol Oncol* (2009).
- [86] Mayer, H., Bauer, H., Breuss, J., Ziegler, S., and Prohaska, R. *Gene* **269**, 73–80 (2001).
- [87] Mayer, H., Salzer, U., Breuss, J., Ziegler, S., Marchler-Bauer, A., and Prohaska, R. *Biochim. Biophys. Acta* **1395**, 301–308 (1998).
- [88] Chung, C. H. Y., Kurien, B. T., Mehta, P., Mhatre, M., Mou, S., Pye, Q. N., Stewart, C., West, M., Williamson, K. S., Post, J., Liu, L., Wang, R., and Hensley, K. *Biochemistry* **46**, 3262–3269 (2007).
- [89] Zhang, W., Wang, L., Liu, Y., Xu, J., Zhu, G., Cang, H., Li, X., Bartlam, M., Hensley, K., Li, G., Rao, Z., and Zhang, X. C. *Genes Dev.* **23**, 1387–1392 (2009).
- [90] Blisnick, T., Vincensini, L., Barale, J. C., Namane, A., and Breton, C. B. *Mol. Biochem. Parasitol.* **141**, 39–47 (2005).
- [91] Tajima, M., Kurashima, Y., Sugiyama, K., Ogura, T., and Sakagami, H. *Eur. J. Pharmacol.* **606**, 45–49 (2009).
- [92] Hanahan, D. *Nature* **315**, 115–122 (1985).
- [93] Hanahan, D. and Folkman, J. *Cell* **86**, 353–364 (1996).
- [94] Jaffe, E. A., Nachman, R. L., Becker, C. G., and Minick, C. R. *J. Clin. Invest.* **52**, 2745–2756 (1973).
- [95] Baudin, B., Bruneel, A., Bosselut, N., and Vaubourdolle, M. *Nat. Protoc.* **2**, 481–485 (2007).
- [96] Bradford, M. M. *Anal. Biochem.* **72**, 248–254 (1976).
- [97] Wessel, D. and Flügge, U. I. *Anal. Biochem.* **138**, 141–143 (1984).
- [98] Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V., and Mann, M. *Nat. Protoc.* **1**, 2856–2860 (2006).
- [99] Ishihama, Y., Rappsilber, J., Andersen, J. S., and Mann, M. *J. Chromatogr. A* **979**, 233–9 (2002).
- [100] Cox, J. and Mann, M. *Nat. Biotechnol.* **26**, 1367–1372 (2008).
- [101] Cox, J., Matic, I., Hilger, M., Nagaraj, N., Selbach, M., Olsen, J. V., and Mann, M. *Nat. Protoc.* **4**, 698–705 (2009).
- [102] Perkins, D. N., Pappin, D. J. C., Creasy, D. M., and Cottrell, J. S. *Electrophoresis* **20**, 3551–3567 (1999).
- [103] Towbin, H., Staehelin, T., and Gordon, J. *Proc. Natl. Acad. Sci. U. S. A.* **76**, 4350–4354 (1979).
- [104] Salinovich, O. and Montelaro, R. C. *Anal. Biochem.* **156**, 341–347 (1986).