

Protein level analyses in tumor cell metastasis

Dissertation

with the aim of achieving a doctoral degree (Dr. rer. nat.)

at the Faculty of Mathematics, Informatics and Natural Sciences

Departement of Chemistry

University of Hamburg

submitted by

Sebastian Gärtner

Hamburg, 2019

1st reviewer: Prof. Dr. Klaus Pantel

2nd reviewer: Prof. Dr. Horst Weller

Date of the disputation: 5th July, 2019

Members of the examination committee of the disputation:

Prof. Dr. Hartmut Schlüter (chairman)

Prof. Dr. Klaus Pantel

Dr. Maria Riedner

**The dissertation with the title „Protein level analyses in tumor cell metastasis“ was prepared at the
Institute of Tumor Biology at the University Medical Center Hamburg-Eppendorf, Center for
Experimental Medicine, Hamburg from November 1st, 2015 until July 5th, 2019.**

I. List of publications

This thesis is not part of a publication at the moment. A part of the thesis is in the process of registration for patent protection at the University Medical Center Hamburg-Eppendorf, Hamburg, Germany.

II. Table of contents

I.	List of publications	1
II.	Table of contents	2
III.	Table of tables.....	5
IV.	Table of figures	6
V.	List of abbreviations	9
1.	Zusammenfassung.....	13
2.	Abstract	15
3.	Introduction	17
3.1.	Breast cancer subtypes	19
3.1.1.	Luminal subtype	19
3.1.2.	ErbB-2 ⁺ subtype	20
3.1.3.	Basal-like subtype	20
3.2.	CTCs/DTCs/CSCs.....	21
3.3.	Epithelial-to-mesenchymal transition (EMT)	22
3.4.	Hypoxia.....	24
3.5.	CUB domain-containing protein 1 (CDCP1).....	27
3.6.	Neuropilins	28
3.7.	Discoidin, CUB and LCCL domain-containing protein 2	29
3.8.	Chimeric antigen receptor T-cell technology (CAR-T).....	30
3.9.	Magnetic activated cell sorting (MACS)	30
3.10.	Therapy.....	32
3.10.1.	Surgery	33
3.10.2.	Radiation	33
3.10.3.	Chemotherapy	33
3.10.4.	Systemic therapy for non-metastatic breast cancer	34
4.	Aim of the thesis	35
5.	Results	36
5.1.	Cell line characterization by relevant marker proteins.....	36
5.2.	Validation of MS-data by Western Blot analysis.....	42
5.3.	Potential biomarker candidates from SILAC-MS-proteome analysis	55
5.4.	Validation of biomarker candidates by Western Blot analysis	63
5.5.	Expression of potential biomarker candidates on CTC cell lines.....	65
5.6.	Potential biomarker candidate expression on PBMC of healthy donors	69

5.7.	Potential biomarker candidate expression on erythrocytes of healthy donors	73
5.8.	CDCP1 expression on cell lines from other tumor entities	75
5.9.	Immunofluorescence staining of CDCP1 in breast cancer and DTC cell lines	76
5.10.	Immunofluorescence staining of Neuropilin-1 in breast cancer and DTC cell lines ..	80
5.11.	Single cell isolation from tumor cell lines via CDCP1-MACS-system	82
5.12.	Detection of CDCP1 on breast cancer tumor cell lines	86
5.13.	Detection of CDCP1 on circulating tumor cells (CTCs) of breast cancer patients	89
5.14.	Detection of soluble CDCP1-variant in the blood plasma of breast cancer patients via ELISA	92
5.15.	Analysis of CDCP1 in the whole breast cancer patient cohort	100
5.15.1.	Analysis of CDCP1 for cT1-cases in the whole breast cancer patient cohort	101
5.15.2.	Analysis of CDCP1 for DCIS-cases in the whole breast cancer patient cohort.....	103
5.16.	Analysis of CDCP1 in the “early detection” breast cancer patient cohort.....	107
5.16.1.	Analysis of CDCP1 in cT1 cases from the “early detection” breast cancer patient cohort.....	109
5.16.2.	Analysis of CDCP1 in DCIS cases from the “early detection” breast cancer patient cohort.....	111
5.17.	Investigation of the influence of microenvironmental changes on the expression levels of the biomarker candidates.....	115
5.17.1.	Investigation of the influence of hypoxia on the expression levels of the biomarker candidates.....	116
5.17.2.	Investigation of the influence of cobalt on the expression levels of the biomarker candidates.....	123
6.	Discussion	133
6.1.	Conclusion.....	150
7.	Outlook	152
8.	Methods and Materials	154
8.1.	Methods	154
8.1.1.	Patients.....	154
8.1.2.	Cultivation conditions of cell lines.....	155
8.1.3.	Colloidal coomassie staining of polyacrylamide gels (Neuhoff staining)	157
8.1.4.	Cell harvest, SDS-Page, Western Blots and membrane stripping	158
8.1.5.	Culture conditions of cell lines under hypoxia (1% O ₂)	159
8.1.6.	Culture conditions of cell lines under cobalt-chloride (CoCl ₂)	160
8.1.7.	Antibodies used for Western blot analysis.....	162
8.1.8.	Isolation of mononuclear cells from peripheral blood	164
8.1.9.	Generation of cytopspins from cell lines	164

8.1.10.	Antibodies used for immunocytochemical staining analysis.....	164
8.1.11.	Immunocytochemical staining of cell lines.....	165
8.1.12.	Cultivation and immunofluorescent analysis of cell lines in chamber slides	166
8.1.13.	Stable isotope labeling by amino acids in cell culture (SILAC).....	166
8.1.14.	Sample preparation for mass spectrometry.....	169
8.1.15.	LC-MS/MS analysis	170
8.1.16.	Data Analysis.....	177
8.1.17.	Isolation and enumeration of tumor cells via CellSearch® System	178
8.1.18.	Fluorescence-activated cell sorting (FACS).....	179
8.1.19.	Quantitative <i>Sandwich</i> -ELISA (Enzyme-linked Immunosorbent Assay).....	180
8.1.20.	Magnetic-activated cell sorting (MACS)	182
8.2.	Materials.....	183
8.2.1.	Chemicals and solutions	183
8.2.3.	Commercial kits and solutions	184
9.	References	185
9.	Supplement.....	I
9.1.	List of used hazardous substances (chemicals and solutions) with hazard statements according to GHS (Globally Harmonised System) classification.....	I
9.2.	MS2-spectra (large scale)	VI
9.3.	Research reports for the analysis of CDCP1 with the CellSearch system.....	XIV
10.	Danksagung	XXI
11.	Eidesstattliche Erklärung.....	XXII
12.	Declaration on oath	XXIII

III. Table of tables

Table 1: Result of MS analysis for chosen validation proteins.	42
Table 2: Results of MS analysis for chosen biomarker candidates.	55
Table 3: Breast cancer patients analyzed for CDCP1 with the CellSearch System.	91
Table 4: Clinicopathological parameters for the whole breast cancer cohort.	105
Table 5: Clinicopathological parameters for the “early detection” cohort.	113
Table 6: Levels of significance for the biomarker candidates under hypoxia.	122
Table 7: Levels of significance for the the biomarker candidates under exposure to cobalt.	131
Table 8: Primary antibodies for Western Blot analysis.	162
Table 9: Overview of the used chemicals and solutions.	183
Table 10: Overview of the used commercially purchased solutions and kits.	184

IV. Table of figures

Figure 1: Overview of the different molecular subtypes in breast cancer.....	19
Figure 2: Degradation of HIF-1 α under normal oxygen conditions	25
Figure 3: Stabilization of HIF-1 α under hypoxic oxygen conditions.....	26
Figure 4: Cell surface molecule-dependent MACS isolation procedure	32
Figure 5: Characterization of different breast cancer cell lines and DTC cell lines I.....	37
Figure 6: Characterization of different breast cancer cell lines and DTC cell lines II.	38
Figure 7: MS1-spectrum of human-Connexin-43	44
Figure 8: MS/MS (MS2)-spectrum of human-Connexin-43.....	45
Figure 9: MS1-spectrum of human-78 kDa glucose-regulated protein (Grp78)	46
Figure 10: MS/MS (MS2)-spectrum of human-78 kDa glucose-regulated protein (Grp78).....	47
Figure 11: MS1-spectrum of human-CD44 antigen (CD44)	48
Figure 12: MS/MS (MS2)-spectrum of human-CD44 antigen (CD44).....	49
Figure 13: MS1-spectrum of human-Integrin alpha 5.....	50
Figure 14: MS/MS (MS2)-spectrum of human-Integrin alpha 5.....	51
Figure 15: MS1-spectrum of human- epidermal growth factor receptor (EGFR)	52
Figure 16: MS/MS (MS2)-spectrum of human-epidermal growth factor receptor (EGFR).....	53
Figure 17: Western Blot analysis of the validation proteins.....	54
Figure 18: MS1-spectrum of human- CUB domain-containing protein 1 (CDCP1)	57
Figure 19: MS/MS (MS2)-spectrum of human-CUB domain-containing protein 1 (CDCP1).....	58
Figure 20: MS1-spectrum of human- Neuropilin-1 (NRP1).....	59
Figure 21: MS/MS (MS2)-spectrum of Neuropilin-1 (NRP1)	60
Figure 22: MS1-spectrum of human-DCBLD2	61
Figure 23: MS/MS (MS2)-spectrum of human-DCBLD2.....	62
Figure 24: Western blot analysis of potential biomarker candidates	63
Figure 25: Characterization of the CTC cell lines	66
Figure 26: Candidate expression on the CTC cell lines	67
Figure 27: Western blot analysis of CDCP1 on PBMC.....	69
Figure 28: Western blot analysis of NRP1 on PBMC	70
Figure 29: Western blot analysis of NRP2 on PBMC	71
Figure 30: Western blot analysis of DCBLD2 on PBMC	72
Figure 31: Western blot analysis of CDCP1 on erythrocytes.....	73
Figure 32: Western blot analysis of NRP1 on erythrocytes	74
Figure 33: Western blot analysis of CDCP1 and NRP1 on cell lines of different tumor entities	75
Figure 34: Immunofluorescence staining of CDCP1	77
Figure 35: Chamber slide CDCP1-immunofluorescence staining.....	79
Figure 36: Immunofluorescent staining of NRP-1.....	80
Figure 37: Chamber slide NRP-1-immunofluorescence staining	81

Figure 38: FACS analysis for anti-human-CDCP1 antibody (R&D AF2666).....	82
Figure 39: Immunofluorescence staining of CDCP1 after MACS isolation	84
Figure 40: FACS analysis for anti-human-CDCP1 antibody (NBP2-60224AF488)	86
Figure 41: CellSearch analysis of breast cancer cell lines for CDCP1	87
Figure 42: CellSearch-CTC-analysis of a breast cancer patient for CDCP1	89
Figure 43: Western Blot analysis of recombinant human-CDCP1 (Novoprotein) I.....	92
Figure 44: Western Blot analysis of recombinant human-CDCP1 (Novoprotein) II	93
Figure 45: Analysis of CDCP1 in cell culture supernatant	94
Figure 46: Analysis of PBMC of healthy female donors by SDS PAGE and Coomassie staining	96
Figure 47: Western Blot analysis of PBMC of healthy female donors for CDCP1	97
Figure 48: Stability test for CDCP1	98
Figure 49: Reliability determination for CDCP1-ELISA.....	99
Figure 50: ROC analysis for CDCP1 in the whole breast cancer patient cohort.....	100
Figure 51: ROC analysis for CDCP1 in the whole breast cancer patient cohort for cT1.....	101
Figure 52: ROC analysis for CDCP1 in the whole breast cancer patient cohort for DCIS.....	103
Figure 53: ROC analysis for CDCP1 in the “early detection” cohort	107
Figure 54: ROC analysis for CDCP1 in the “early detection” cohort for cT1	109
Figure 55: ROC analysis for CDCP1 in the “early detection” cohort for DCIS	111
Figure 56: Western blot analysis of the biomarker candidates under hypoxia	116
Figure 57: Quantification of CDCP1 under hypoxia.....	118
Figure 58: Quantification of NRP1 under hypoxia	120
Figure 59: Quantification of NRP2 under hypoxia	121
Figure 60: Western blot analysis of the biomarker candidates under cobalt	123
Figure 61: Quantification of Hif-1 α under cobalt	125
Figure 62: Quantification of CDCP1 under cobalt.....	127
Figure 63: Quantification of NRP1 under cobalt	129
Figure 64: Quantification of NRP2 under cobalt	130
Figure 65: Morphological analysis cell lines lines under cobalt	132
Figure 66: Suggested hypothetical model for the CDCP1 ⁺ /CK ⁺ /CD45 ⁻ CTC subpopulation	141
Figure 67: Suggested model for influence of cobalt on Hif-1 α stability	160
Figure 68: Overview of the workflow for SILAC-labeling.....	168
Figure 69: Setup of the Quadrupole Orbitrap mass spectrometer	170
Figure 70: Principle of electrospray ionization.....	172
Figure 71: Typical arrangement of a Quadrupole mass analyzer	174
Figure 72: Cross section of the Orbitrap mass analyzer and the connected C-trap	175
Figure 73: MS2-spectrum of human-Connexin-43 (large scale)	VI
Figure 74: MS2-spectrum of human-78 kDa glucose-regulated protein (Grp78) (large scale)	VII
Figure 75: MS2-spectrum of human-CD44 antigen (CD44) (large scale)	VIII
Figure 76: MS2-spectrum of human-Integrin alpha 5 (large scale).....	IX
Figure 77: MS2-spectrum of human- epidermal growth factor receptor (EGFR) (large scale)	X
Figure 78: MS2-spectrum of human- CUB domain-containing protein 1 (CDCP1) (large scale)	XI

Figure 79: MS2-spectrum of human- Neuropilin-1 (NRP1). (large scale)	XII
Figure 80: MS2-spectrum of human- DCBLD2 (large scale)	XIII
Figure 81: Research report for the breast cancer sample Heid-716.....	XIV
Figure 82: Research report for the breast cancer sample Heid-796.....	XV
Figure 83: Research report for the breast cancer sample Heid-805.....	XVI
Figure 84: Research report for the breast cancer sample Heid-786.....	XVII
Figure 85: Research report for the breast cancer sample Heid-791.....	XVIII
Figure 86: Research report for the breast cancer sample Heid-nachL11/172	XIX
Figure 87: Research report for the breast cancer sample Heid-790.....	XX

V. List of abbreviations

aa	Amino acid
AC	Alternating current
ADP	Adenosine diphosphate
ANOVA	Analysis of variance
ATP	Adenosine triphosphate
AUC	Area under curve
BC	Breast cancer
BCA	Bicinchoninic acid
BL1	Basal-like 1
BL2	Basal-like 2
BMP-1	Bone morphogenetic protein-1
BRCA1	Breast cancer, early onset 1
BSA	Bovine serum albumine
c	Concentration
CAR-T	Chimeric antigen receptor T-cell
CD	Cluster of differentiation
CI	Confidence interval
CID	Collision-induced dissociation
CK	Cytokeratin
CO ₂	Carbon dioxide
CSC	Cancer stem cell
CTC	Circulating tumor cell
C1r/C1s	Complement C1r/C1s subcomponent
d	Days
DNA	Desoxyribonucleid acid
DAPI	4',6'-diamidino-2-phenylindole
DC	Direct current
DCIS	Ductal carcinoma <i>in situ</i>
DPBS	Dulbecco's Phosphate-Buffered Saline
DTC	Disseminated tumor cell
DTT	Dithiothreitol

E	Epithelial
e	elementary charge
ECM	Extracellular matrix
eCTCs	epithelial circulating tumor cells
EDTA	Ethylenediaminetetraacetic acid
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
E_{kin}	Kinetic energy
ELISA	Enzyme-linked Immunosorbent Assay
emCTCs	Partial EMT circulating tumor cells
EMT	Epithelial-to-mesenchymal transition
ER	Estrogen receptor
ESI	Electrospray ionization
FACS	Fluorescence-activated cell sorting
FCS	Fetal calf serum
FDA	Federal Drug Administration
FDR	False discovery rate
FGF	Fibroblast growth factor
FIH-1	Factor inhibiting HIF-1 α
FITC	Fluorescein isothiocyanate
FSC	Forward scattering
FT	Fourier transformation
G	Grading
g	Gram
h	Hour
HCD	Higher-energy collisional dissociation
HCl	Hydrochloric acid
HER2/ErbB-2	Human epidermal growth factor receptor 2
HPLC	High-performance liquid chromatography
HR	Hormone receptor
HRE	Hypoxia response elements
IF	Immunofluorescence
IEF	Isoelectric focusing

IPG	Immobilized pH-gradient
kDa	Kilo Dalton
kVh	Kilo volt per hour
L	flight distance
LC	Lung cancer
m	Mass
M	Mesenchymal
MACS	Magnetic-activated cell sorting
mc	Monoclonal
mCTCS	Mesenchymal circulating tumor cells
MET	Mesenchymal-to-epithelial transition
mg	Milligram
mg/L	Milligram per liter
μ A	Micro ampere
μ L	Microliter
μ M	Micromolar
mL	Milliliter
mmol/L	Millimol per liter
MS	Mass spectrometry
MW	Molecular weight
m/z	Mass-to-charge ratio
nm	Nanometer
OD	Optical density
ODD	Oxygen-dependent degradation domain
PaCa	Pancreatic cancer
PAGE	Polyacrylamide gel electrophoresis
PBMC	Peripheral blood mononuclear cells
PC	Prostate cancer
pc	Polyclonal
PE	Phycoerythrin
pH	Concentration hydrogen ions
PHD	Prolyl hydroxylase domain proteins
PR	Progesteron receptor

PVDF	Polyvinylidene fluoride
RF	Radio frequency
ROC	Receiver operating characteristics
RPC	Reversed phase chromatography
rpm	Rounds per minute
RT	Room temperature
RTK	receptor tyrosine kinase
s	Second
scFv	Single chain variable fragment
SDS	Sodium dodecyl sulfate
SILAC	Stable isotope labeling by amino acids in cell culture
SSC	Side scattering
t	Flight time
TBST	Tris-buffered Saline with Tween 20
TNBC	Triple negative breast cancer
TOF	Time of flight
U	Acceleration Voltage
UV	Ultraviolet
v	Velocity
V	Volt
VHL	Von Hippel-Lindau protein
z	charge
°C	Degree Celsius
%	Percent

1. Zusammenfassung

Brustkrebs war die am häufigsten vorkommende Tumorentität bei Frauen weltweit im Jahr 2018 und war verantwortlich für etwa 15 % der Tumor-assoziierten Todesfälle bei Frauen. Der metastatische Rückfall mit Ausbreitung in Leber, Gehirn, Knochenmark und andere Organe führt dabei in den meisten Fällen zum Tod. Die unterschiedlichen Brustkrebs-Subtypen weisen stark variierende Therapiemöglichkeiten sowie verschieden stark ausgeprägtes Metastasierungspotenzial auf. Der Östrogen-Rezeptor-positive Subtyp, der etwa zwei Drittel aller Brustkrebsfälle ausmacht, ist hormonell sehr gut therapierbar. Der Östrogen-Rezeptor-, Progesteron-Rezeptor- sowie der Her2 (ErbB-2)- negative Subtyp (triple negativer Brustkrebs) hingegen korreliert mit schlechter Prognose, hohem Anteil an metastatischem Rückfall und schlechter Therapierbarkeit. Die Metastasierung beginnt mit den sogenannten zirkulierenden Tumorzellen (englisch: circulating tumor cell, CTC). Dies sind Tumorzellen, die vom Primärtumor abgegeben werden, sich über das Blutssystem verteilen und das Sekundärorgan infiltrieren. Nach Austritt aus dem Blutssystem in Sekundärorgane werden diese Tumorzellen disseminierte Tumorzellen (englisch: disseminated tumor cells, DTC) genannt. CTCs und DTCs spielen eine große Rolle bei der Detektion, Therapieentscheidung, Therapieüberwachung und bei der Prognose. Zirkulierenden Tumorzellen mit mesenchyalem oder epithelalem/mesenchyalem Phänotyp wird ein hohes Metastasierungspotenzial nachgesagt. Die gegenwärtig etablierten Marker (EpCAM, Cytokeratine) zur Detektion und Anreicherung von Tumorzellen aus dem Blut von Patienten werden stark bei Zellen mit dem epithelialen Phänotyp exprimiert, sind aber bei mesenchymalen Tumorzellen herunterreguliert. Das Ziel dieser Arbeit war es neue Oberflächenprotein-Biomarker für Tumorzellen mit mesenchymalen Eigenschaften zu identifizieren und diese „Detektionslücke“ zu schließen. Dazu wurde ein SILAC-Massenspektrometrie-basierter Proteomvergleich zwischen der Knochenmark-metastatischen Brustkrebs Zelllinie MDA-231 BO2 und der disseminierten Tumorzelllinie BC-M1, isoliert von den Metastasen einer Brustkrebs-Patientin, durchgeführt. Dabei wurden die Oberflächen-Proteine CDCP1, NRP-1 (und NRP-2) und DCBLD2 als mögliche Biomarker identifiziert und auf DTC-Zelllinien von anderen Entitäten und anderen Brustkrebszelllinien nachgewiesen. Für die Proteine CDCP1 und NRP-1 konnte ein Immunfluoreszenz-Nachweis für Brustkrebs-Zelllinien etabliert werden. Für das Protein CDCP1 konnte sowohl ein durchflusszytometrisches (FACS) Nachweisverfahren sowie ein magnet-basiertes Zellseparationsverfahren (MACS) zur Isolation von Tumorzellen über CDCP1 aus dem Blut etabliert werden. Des Weiteren konnte für die lösliche Variante von CDCP1, welche durch proteolytische Spaltung entsteht, ein immunologisches

Antikörper-basiertes Nachweisverfahren (ELISA) für den Nachweis aus dem Blutplasma von Brustkrebspatienten entwickelt werden. Lösliches CDCP1 konnte im Blutplasma von Brustkrebspatienten nachgewiesen werden (Sensitivität: 29.1 %, Spezifität: 96.5 %). Weiterhin konnte für das Protein CDCP1 der Nachweis mittels des CellSearch-Systems, welches Tumorzellen über das epitheliale Oberflächenmolekül EpCAM anreichert, etabliert werden. CDCP1 konnte auf CTCs von Brustkrebspatienten nachgewiesen werden. Die Verwendung von CDCP1 im CellSearch-System sowie der Immunfluoreszenz-basierte Nachweis dienen zur Detektion von schwer zu identifizierenden CTCs und DTCs mit mesenchymalen oder epithelial/mesenchymalen Eigenschaften. Die Verwendung des MACS-basierten Isolations-Systems dient hingegen zur Isolation von Tumorzellen mit mesenchymalen oder epithelial/mesenchymalen Eigenschaften über CDCP1. Dadurch ergibt sich die Möglichkeit solche Tumorzellen mit anderen Verfahren weiter zu charakterisieren und so die Prozesse der Metastasierung besser verstehen zu können.

2. Abstract

Breast cancer was the leading cancer entity for women worldwide in 2018 and is responsible for 15 % of cancer-related deaths. The relapse followed by metastatic spread into liver, brain, bone marrow and other distant organs is associated with cancer-related deaths. The therapy outcome varies between different breast cancer subtypes. Furthermore, the different subtypes differ in their metastatic potential. For example, for the estrogen-receptor positive subtype, which accounts for two thirds of all breast cancer cases, is good treatable by hormonal therapy. On the other hand, the estrogen receptor-, progesterone receptor- and Her2 (ErbB-2)- negative Subtype (triple negative breast cancer) correlates with bad prognosis, high probability for a metastatic relapse and ineffective treatment possibilities. The properties of metastasis formation are connected to so called circulating tumor cells (CTC). These are tumor cells released by the primary tumor followed by distribution via the blood circulation and infiltration of distant organs. After extravasation to secondary organs, these tumor cells are called disseminated tumor cells (DTC). CTCs and DTCs play a crucial role in detection, therapy decision, therapy monitoring and prognosis. CTCs with mesenchymal or epithelia/mesenchymal phenotype are thought to be major driver in metastasis formation. The current established and used biomarkers (EpCAM, cytokeratins) for detection and enrichment of tumor cells from the blood of patients are strongly limited to epithelial properties and downregulated in mesenchymal tumor cells. The aim of this work was the closing of this “detection gap” for tumor cells with mesenchymal attributes and the identification of new cell surface biomarker proteins for such cells. To address this question, a SILAC-mass spectrometry based proteome analysis between the bone marrow metastatic breast cancer cell line MDA-231 BO2 and the DTC cell line BC-M1 which was isolated from bone marrow metastasis of a breast cancer patient was performed. The cell surface proteins CDCP1, NRP-1, (NRP-2) and DCBLD2 were identified as potential new biomarkers. The expression of these proteins was detected on DTC cell lines from other entities and on different breast cancer cell lines. For CDCP1 and NRP-1 an immunofluorescent analysis for breast cancer cell lines was established. Additionally, a flow cytometry based analysis (FACS) as well as a magnetic activated cell sorting (MACS) based analysis was established for breast cancer cell lines and CDCP1. Furthermore, an immunological antibody based analysis (ELISA) was established for the detection of the soluble variant of CDCP1 (generated by proteolytic cleavage) in the blood of breast cancer patients. Soluble CDCP1 was detected in the blood of breast cancer patients (sensitivity: 29.1 %, specificity: 96.5 %). In addition, the detection of CDCP1 on CTCs of breast cancer patients with the CellSearch system (tumor cell enrichment via epithelial cell surface molecule EpCAM)

was established. CDCP1 was detected on CTCs of breast cancer patients. The application of the CDCP1 detection with the Cell Search System and the immunofluorescent analysis serve as detection methods for badly detectable CTCs and DTCs with mesenchymal or epithelial/mesenchymal attributes in breast cancer patients. However, the application of the MACS-based isolation method provides a potential enrichment technique for badly characterized tumor cells with mesenchymal or epithelial/mesenchymal properties via CDCP1. Hence, these isolated tumor cells can be further analyzed which might lead to a better understanding of the process of metastasis formation.

3. Introduction

Development of tumors (tumorigenesis) in humans is a process with several steps. Thereby genetic changes cause the stepwise transformation of healthy human cells into malignant descendants. The diagnosis of several cancer types correlate with advanced age and tumorigenesis requires several genetic alterations like mutations in the p53 suppressor gene or the gene encoding for retinoblastoma protein [1, 2] resulting in the development of invasive tumors by traversing several intermediate precancerous states [3]. These genetic alterations result in growth advantages and thereby leading to tumor formation [1, 3, 4]. Defects in regulatory mechanisms cause abnormal cell proliferation and homeostasis in cancer cells. In 2000, Hannahan and Weinberg proposed the six “hallmarks of cancer” which contribute to a better understanding of the diverse mechanisms leading to neoplastic diseases [1]. Independence of growth signals, impassivity to growth-inhibitory signals, apoptosis avoidance, unlimited replicative abilities, angiogenic properties, tissue invasion and metastasis formation arbitrate malignancy. Normal cells need stimulatory signals otherwise they are not able to proliferate whereas in cancer cells some oncogenes are thought to mimic normal growth-signal molecules. Additionally, in normal cells antiproliferative signal molecules mediate quiescence and homeostasis whereas in tumor cells the signal cascade is disturbed due to downregulation or mutation of corresponding growth receptors. The expansion of cancer cells is not only caused by the unlimited proliferation but also by the disturbed cell turnover due to inefficacy of programmed cell death. These three mentioned conditions lead to the independence of cancer cell growth from environmental influences and signals. The fast increasing number of tumor cells in a primary tumor causes the increased demand of energy, oxygen and nutrients which is provided by blood supply. Therefore tumor cells upregulate the expression of angiogenic inducers or downregulation of angiogenic inhibitors [1, 5]. The invasive properties and the metastatic potential of cancer cells are promoted by the five other hallmark abilities. Both attributes include alterations in the process of cell-cell- and cell-ECM- interactions as well as the use of extracellular proteases which are necessary for tissue infiltration and distant seeding of metastases. These derivatives of the primary tumor are responsible for 90 % of cancer-related deaths in humans [1, 6]. Another property of cancer cells is the genomic instability and arising mutations which provide these mutated cells with advantages [1]. The progress in cancer research has led to the expansion of the hallmarks of cancer by three additional characteristics that favor tumor progression. These are the adaption of the cellular energy metabolism, the escape of the surveillance and elimination by the immune system as well as the support of the acquired hallmark properties by innate immune cell mediated

inflammation [7]. The relevance of this disease caused by the existence of more than 100 cancer types with different subtypes is reflected in the number and statistics of cancer cases in the worldwide population.

In 2018, 18.1 million new cancers cases were diagnosed worldwide followed by 9.6 million cancer-related deaths [8]. Published data by the WHO (world health organization) demonstrated that lung cancer has occupied the sixth position in global causes of deaths for 2016 whereas in the year 2000 lung cancer ranked at position nine. Interestingly, cancer was not listed in the first ten positions of global causes of deaths in low-income and lower-middle-income countries for 2016. However, lung-, liver- and stomach cancer have occupied position four, nine and ten in upper-middle-income countries in 2016. Additionally, in high-income countries lung cancer, colon and rectum cancer and breast cancer were listed at the positions four, seven and ten [9]. The ageing of the population has contributed markedly to the increasing number of new cancer cases worldwide in the past years. For example, 6.7 million new cancer cases were observed in adults with the age of 65 years or older in 2012. This number incorporates almost half of the cases diagnosed worldwide in 2012 while the portion of adults aged 65 years or older was 8 % of the world population. For males aged 65 years or older, prostate cancer was the leading cancer entity whereas for females aged 65 years or older breast cancer was the leading cancer type in 2012. For 2035, the total number of new cancer cases (combination of all cancer types) was estimated to increase up to 14 million among adults aged over 65 years and older [10].

Breast cancer was the second leading entity with 11.6% (2.1 million) of all 18.1 million newly diagnosed cases in 2018 among all ages and both sexes. Breast cancer contributed with a portion of 6.6 % (0.6 million) to all cancer deaths (9.6 million) in 2018 among all ages and both sexes. This correlates with the fifth position for cancer deaths behind lung (18 %), colorectum (9 %), stomach (8 %) and liver cancer (8 %) among all ages and both sexes. In detail, breast cancer was the leading cancer entity for women worldwide in 2018 regarding incidence and mortality. 24.2 % (2.1 million) of all new cancer cases (9.6 million) were caused by breast cancer and 15 % (0.6 million) of all cancer-related cases of death (4.2 million) were provoked by this entity [8].

3.1. Breast cancer subtypes [11, 12]

There are several clinical and prognostic factors that classify different types and stages of breast cancer. Tumor size, lymph node metastasis status, cT-stage, status of differentiation (grading), histologic origin and expression level of progesterone receptor (PR), estrogen receptor (ER) or human epidermal growth factor receptor 2 (HER2, ErbB-2) are important clinical and prognostic factors for the classification of breast cancer at primary diagnosis [13, 14]. These criteria underline the heterogeneity of this disease. Breast cancer can be subdivided into four major subtypes (**figure 1**): “luminal A”, “luminal B”, “HER2/ErbB-2 overexpressed” and “basal-like” [15, 16]. These subtypes differ in clinical outcome, prognosis, infiltrated distant organs, metastatic relapse, therapy possibilities and response [13, 17].

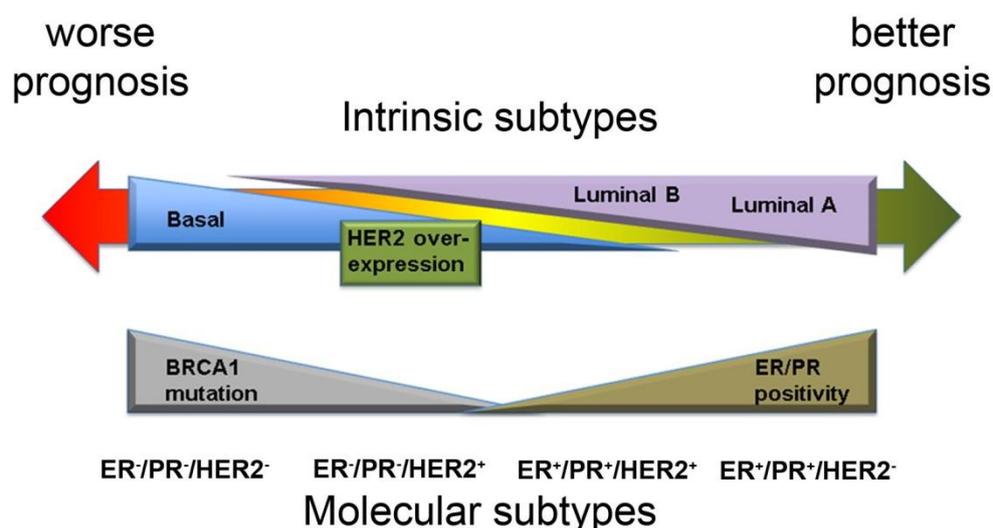


Figure 1: Overview of the different molecular subtypes in breast cancer. The image was edited according to the template figure 2 from “Breast cancer intrinsic subtype classification, clinical use and future trends” by Dai *et al.* [15].

3.1.1. Luminal subtype

More than two thirds of all breast cancer cases are hormone receptor-positive (HR: ER/PR) tumors with a better prognosis than HR-negative breast cancers [13]. Their expression pattern is similar to their underlying luminal epithelial component of the breast including the expression of luminal cytokearatin 8/18 [12, 15, 18]. More than 90% of HR⁺/HER2⁻ tumors can be classified into luminal A (ER⁺/PR⁺/HER2⁺) or luminal B (ER⁺/PR⁺/HER2⁻) subtype [13, 15, 19] in which luminal A tumors represent the largest part [15]. Tumors of luminal B subtype have a higher proliferation rate due to

increased expression of genes responsible for proliferation like p53, BRCA1 or ErbB-2 [11, 13, 20] whereas tumors of luminal A subtype show higher expression levels of ER-related genes like Cathepsin [11, 15, 21]. Hormone therapy is an effective treatment strategy for luminal breast tumors whereas chemotherapy is less successful [15, 22]. Normally, luminal subtypes of breast cancer exhibit a good prognosis but with a worse prognosis for luminal B subtypes compared to luminal A tumors [15, 21].

3.1.2. ErbB-2⁺ subtype

ErbB-2 -overexpressing breast tumors represent 15% to 25% of invasive breast cancers [13] and they exhibit an ER⁻/PR⁻/ErbB-2⁺ expression pattern. This subtype is very heterogeneous because differences in ErbB-2-transcription levels are not observed in all clinically ErbB-2⁺ tumors [15] which causes varying responses to ErbB-2-targeted therapies [13, 23]. Additionally, 40% to 80% of ErbB-2-overexpressing breast cancers show a tp53 mutation and are usually grade 3 [15]. ErbB-2-overexpressing breast cancers have a worse prognosis but they show a good response to therapies that address ErbB-2 (trastuzumab) [13, 24]. But not all ErbB-2-overexpressing tumors show a satisfying response to trastuzumab. Particularly, an upregulation of the chemokine receptor CXCR4 is associated with a resistance against trastuzumab [25]. Additionally, they respond well to anthracycline and taxane-based neoadjuvant therapy with an increased complete response than luminal breast cancers. The higher probability of an early relapse causes the bad prognosis [22].

3.1.3. Basal-like subtype

Breast tumors of basal-like subtype are a special subtype among breast cancers [13] with an ER⁻/PR⁻/ErbB-2⁻ expression status which is referred to as triple-negative breast cancer (TNBC). TNBC accounts for 10 % to 20 % of all breast cancer cases [13, 26] and can be subdivided in six molecular subtypes: two basal-like (BL1/BL2), an immunomodulatory subtype (IM), a mesenchymal subtype (M), a mesenchymal stem-like subtype and a luminal androgen receptor subtype [27]. The expression levels of specific proteins are similar to the basal epithelial cells of other organs and breast myoepithelial cells [12, 15]. These tumors are characterized by absence or low expression of hormone receptors and ErbB-2, high expression of proteins related to basal origin like EGFR, CK5/6/14/17 and genes connected with proliferation [12, 18]. The

expression of basal cytokeratins is thought to correlate with reduced expression of BRCA1 [28] and TP53 mutation [11, 29]. 60% to 90% of TBNC are of basal origin [15, 30] and they have the worst prognosis of all tumors of the breast with high aggressivity and the lack of systemic therapy [13]. Additionally, this subtype correlates with a younger age of the patients [31]. Chemotherapy with anthracycline and taxane is the only treatment option due to the lack of receptors ER/PR/ErB-2 and the bad prognosis is caused by less therapy options and not by chemoresistance [22]. As a consequence of missing receptors, several FDA-approved therapies fail [32]. TNBC incorporates an additional molecular subtype called claudin-low tumors [33] which lack the expression of ER, PR, ErB-2, Claudin-3, Claudin-4, Claudin-7 and E-Cadherin [34]. These claudin-low triple-negative tumors are thought to contain tumor cells exhibiting properties similar to those of stem cells and attributes of epithelial-to-mesenchymal transition (EMT) [33].

3.2. CTCs/DTCs/CSCs

Primary tumors are capable of releasing single tumor cells or tumor cell clusters which infiltrate the surrounding tissue followed by the entrance into the blood circulation. These tumor cells are called circulating tumor cells (CTCs) [35]. The model of dissemination was long time accepted as a late event in cancer progression but present research revealed that cancer cells also disseminate in early precancerous lesions [36]. After leaving the environment of the primary tumor, cancer cells have to conquer several difficulties to survive and reach a distant organ. Tumor cells in the blood circulation have to protect themselves from anoikis-mediated cell death due to loss of cell-cell and cell-ECM interactions. Additionally CTCs have to escape the attack of the immune system and their survival depends on the durability to shear forces in the blood stream [35]. Experimental models suggest that millions of CTCs are present in the whole body but only a small portion is able to escape the immune system and therapy and enter a distant organ (liver, bone, lungs). After extravasation in secondary organs, these cells are called disseminated tumor cells (DTCs) [36] and either reach a dormant state or cause metastatic outgrowth [35-37]. The state of quiescence after entering secondary organs is called clinical cancer dormancy and is characterized by the termination of cell division. It can take up to ten years until the dormant tumor cells start to proliferate again due to potential signals from the surrounding microenvironment [35, 36, 38]. Osteoclasts are thought to be involved in the osteolytic process which favors the formation of bone metastases [35, 39, 40]. The plasticity or the ability of cancer cells to switch between different phenotypes represents a necessary

property. Without this property the efficiency of the metastatic cascade would be markedly reduced. Without acquiring mesenchymal attributes epithelial tumor cells would not be able to migrate by invasion of the surrounding tissue otherwise mesenchymal CTCs/DTCs are capable of seeding distant metastasis without regaining epithelial properties [35, 41-43]. Another rare subpopulation of cancer cells with tumor-initiating potential are called cancer stem cells (CSC). These cells exhibit the ability of self-renewal as well as the potential of differentiation like normal hematopoietic stem cells. A small subset of CTCs/DTCs shows a breast cancer stem-cell phenotype ($CD44^+/CD24^{-/low}$) [44]. Additionally, most of the CTCs/DTCs are non-proliferative and chemotherapy-resistant which are also supposed attributes of cancer stem cells [44-47]. There are two possible explanations why the dysregulation of pathways that control self-renewal might cause tumorigenesis and lead to the transformation of stem cells into different types of cancer. The ability of self-renewal is activated in stem cells but not in differentiated cells, thereby the probability might be higher to sustain this property in stem cells than the complete activation of this program in differentiated cells which would require less mutations. Furthermore, due to the longer existence of stem cells the probability of the accumulation of mutations is increased compared to differentiated cells with a shorter lifetime [48]. CTCs, DTCs and especially CSCs seem to exhibit an increased potential to perform EMT and to disseminate to secondary organs [44, 49].

3.3. Epithelial-to-mesenchymal transition (EMT)

EMT is a normal process in embryonic development, organogenesis, gastrulation and development of the peripheral nervous system [50, 51] which causes epithelia-mesenchymal plasticity and enables these cells to migrate and form different tissues and organs [35, 52, 53]. The process of EMT plays a relevant role in wound healing and tissue regeneration [35, 54]. Cancer cells from the primary tumor also take advantage of this program to start invasion and migration. As a consequence metastases occur which are primarily responsible for cancer-related deaths [35]. Motility and invasiveness are essential requirements for metastatic spread of malignant cells [7, 35]. Before detaching from primary tumor, the number of cell-cell contacts of the tumor cells decrease by downregulation of proteins responsible for adherence (Cadherins) which in turn results in the loss of the apical-basal polarity [35, 51]. After numerous genetic changes the tumor cells are released from the primary tumor, penetrate the basement membrane [51] and intravasate actively or passively the blood circulation or the lymphatic vessels. The circulating tumor cells (CTC) are

exposed to physical stress like shear forces in the blood stream or collision with blood cells. Additionally they have to avoid anoikis, a process which normally leads to apoptosis by the lack of cell-ECM contacts and escape the immune system [35, 55]. After reaching a distant organ, tumor cells extravasate the blood circulation by binding to the endothelium after reaching small capillaries [56]. After extravasation, disseminated tumor cells (DTC) die, remain in tumor dormancy or activate the reverse process of EMT, the mesenchymal-to-epithelial transition (MET), to gain the ability to proliferate and to found the outgrowth of macrometastasis [35, 57, 58]. The bone marrow seems to be the favorite organ for the accumulation of DTC [36]. In breast cancer the outgrowth of metastasis can take several years after removal of the primary tumor [38]. One reason for the late metastatic relapse is the dormancy state of DTC with low proliferation rates [35, 36] resulting in weak therapy effects in this period followed by resuming signal-triggered proliferation [59, 60]. CTC with mesenchymal attributes after passing through EMT have advantages in motility but not in proliferation [42]. In general, after complete EMT, the expression of epithelial proteins like EpCAM or E-Cadherin is downregulated due to the loss of cell-cell connections, cytoskeletal alterations and changes in keratin expression patterns followed by the final upregulation of mesenchymal proteins like Vimentin and N-Cadherin [35, 61]. Additionally, it was demonstrated that a subpopulation of CTCs, especially clusters, exhibit epithelial as well as mesenchymal properties [35]. This phenotype is referred to as partial EMT or hybrid EMT/MET and is thought to be involved in early dissemination, dormancy and metastatic outgrowth [41, 62]. Circulating tumor cells with dynamic changes between epithelial and mesenchymal phenotype [63] or with a partial EMT phenotype can make use of the advantages of both phenotypes which includes the migratory properties of mesenchymal cells and the cell-cell connections of epithelial cells [35, 63]. A complete epithelial or mesenchymal phenotype does not favor the metastatic outgrowth. Properties of both phenotypes have to be present otherwise tumor cells would not be able to migrate without mesenchymal attributes or to proliferate, grow and form metastasis without properties of an epithelial phenotype [35, 42]. The same applies to DTCs, DTC with mesenchymal attributes remain in dormancy after extravasation at secondary sites [41] whereas the retrieval of the epithelial character by MET is necessary for metastatic outgrowth [41, 43]. The activation of EMT can be caused by several different stimuli. TGF- β , Wnt or Interleukin-6 can provoke EMT but also nicotine, alcohol or UV-light are possible inducers [35, 43, 57]. These signals activate transcription factors like Snail, Twist and ZEB that favor the mesenchymal phenotype [57]. Hypoxia is also known as activator of EMT in breast cancer [64, 65].

3.4. Hypoxia

Oxygen is the second most frequent element appearing in the atmosphere of the earth [66]. The oxygen metabolism is a strictly regulated process and is necessary for the existence of mammalian cells [67]. The aerobic respiration enables mammalian cells to metabolize ADP to ATP. The energy required for this conversion is generated by the reduction of oxygen to water in the presence of hydrogen by the mitochondrial electron transport chain [66]. The oxygen concentration in the air is 21 % under normal atmospheric pressure whereas the partial oxygen concentration in alveoli is 13.5 % and in the arterial blood circulation the oxygen concentration is 9.5 %. The physiological oxygen concentration in peripheral tissue is between 4 % and 7.5 % and is referred to as physoxia depending on the tissue. The range between 1 % and 5 % oxygen is called physiological hypoxia and oxygen concentrations below 1 % are referred to as pathological hypoxia [68]. Hypoxia is also thought to support tumor cell dissemination [69-71] at which the bone marrow represents the favorite secondary site for DTCs in breast cancer [69, 72-75]. To gain the ability to found metastasis after extravasation DTC have to adapt to new microenvironmental conditions [76]. DTC have to deal with low oxygen concentrations of approximately 1 % O₂ (hypoxia) in the hematopoietic stem cell niche [69, 77]. Hypoxia promotes angiogenesis, metastasis formation, therapy resistance [76] and is able to activate EMT [64], thus generating tumor cells with mesenchymal attributes which contribute to metastatic spread because of their stem cell properties and resistance to therapies [57, 69]. Tumor cells react to hypoxia by activation of hypoxia-inducible factor family of transcription factors (HIF-1 and HIF-2), especially Hif-1 α [65]. Additionally, the HIF-independent signaling pathways mediated by mTOR (mammalian target of rapamycin) kinase and the UPR (unfolded protein response) contribute to the adaptation of tumor cells to low oxygen concentrations [76]. The heterodimer HIF1 is a transcription factor composed of the subunit HIF-1 β and strongly regulated subunit HIF-1 α . Both subunits comprise domains which participate in DNA binding and dimerization [65, 78]. The HIF-1 α subunit additionally comprises an oxygen-dependent degradation domain (ODD). Under normal oxygen concentrations, two proline residues (P402 and P564) in the ODD of HIF-1 α become hydroxylated by prolyl hydroxylase domain proteins (PHD). The hydroxylation is followed by binding of the von Hippel-Lindau protein (VHL) to the hydroxylated proline residues of HIF-1 α . This in turn leads to proteasomal degradation through transferring a poly-ubiquitin chain mediated by an E3 ubiquitin ligase (**figure 2**). An

additional hydroxylation of an asparagine (N803) by FIH-1 (factor inhibiting HIF-1 α) inhibits the HIF-1 α binding to the co-activators p300 and CREB binding protein (CBP) which is followed by a reduced activity of HIF-1 α under normal oxygen conditions (normoxia) [65, 79, 80].

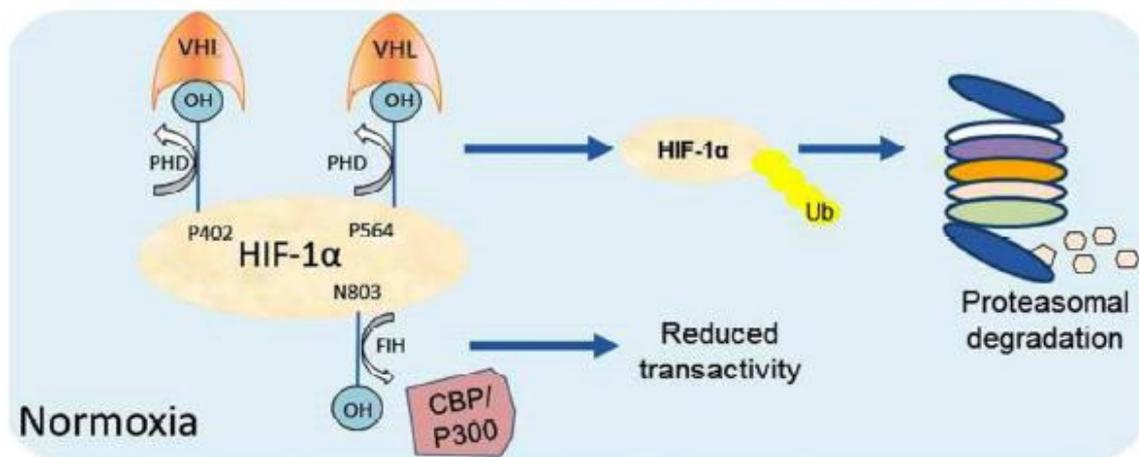


Figure 2: Degradation of HIF-1 α under normal oxygen conditions (normoxia). The image was edited according to the template figure 1 from “Hypoxia-inducible factor 1 and breast cancer metastasis by Liu *et al.* [65].

Low level oxygen concentrations prevent the hydroxylation of both proline residues and the asparagine residue which leads to HIF-1 α stabilization, accumulation followed by dimerization with HIF-1 β in the nucleus. This heterodimer in turn binds to a specific DNA sequence located in hypoxia response elements (HRE) within the target gene region and triggers their activation [65] (**figure 3**). The expression of more than 1000 genes is regulated by HIF-1 and HIF-2 as a cause of decreased oxygen supply [81].

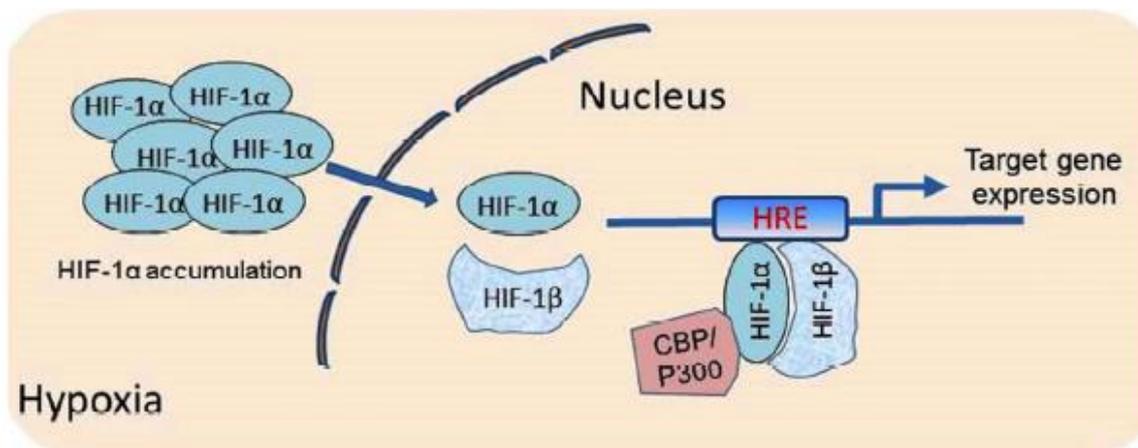


Figure 3: Stabilization of HIF-1 α under hypoxic oxygen conditions (hypoxia). The image was edited according to the template figure 1 from “Hypoxia-inducible factor 1 and breast cancer metastasis by Liu *et al.* [65].

Hypoxic regions in tumors [82] are not the only reason for the activation of the HIF-1 pathway. This pathway can also be activated by oxygen-independent growth factors. For example, the binding of the growth factors insulin-like growth factor-2 (IGF2) to its receptor insulin-like growth factor 1 receptor (IGF1R) or binding of transforming growth factor- α to its receptor epidermal growth-factor receptor (EGFR) mediate signal-transduced activation of HIF-1 α expression and proliferation of tumor cells. The stimulation of the activation of HIF-1 α by oxygen-independent growth-factors is different from the activation due to hypoxic response. The growth-factor mediated expression of HIF-1 α is dependent on the cell type. Additionally, growth-factors excite the expression of HIF-1 α by activating the phosphatidylinositol 3-kinase (PI3K) or the mitogen-activated protein kinase (MAPK) pathway compared to hypoxia which induces the stabilization of HIF-1 α [83].

3.5. CUB domain-containing protein 1 (CDCP1)

The protein CUB domain-containing protein 1 (CDCP1, UniProtKB: Q9H5V8) is a single-pass type I membrane glycoprotein (also known as *substractive immunization associated 135 kDa* (SIMA135) [84], gp140 [85], *transmembrane and associated with Src kinases* (Trask) [86] or CD318) with an intracellular, a membrane and an extracellular domain. The extracellular domain consists of three CUB (C1r/C1s, Uegf, Bmp1) domains and the complete sequence contains 14 sites for N-glycosylations as well as 20 cysteine residues potentially involved in formation of disulfide bonds [84-87]. Structural features comprise a full sequence length of 836 amino acids consisting of 29 amino acid signal peptide sequence, a cytoplasmic domain of 150 amino acids, a transmembrane domain of 21 amino acids and an extracellular domain of 636 amino acids [88]. The calculated molecular weight of this protein is around 90.1 kDa, however the observed molecular weight in Western Blot analysis is around 135 kDa due to N-Glycosylations[84]. CDCP1 (cleaved form of CDCP1) was first mentioned in 1996 [89]. The proof of the expression of CDCP1 in leukemia, metastatic colon cancer and breast cancer was first mentioned by Scherl-Mostageer *et al.* in 2001 [87] and by Hooper *et al.* in 2003 [84]. In addition, the detection of CDCP1-expression on a subset of different hematopoietic cells as well as colon cancer cells and mamma carcinoma cells while the surrounding tissue showed no expression of CDCP1 was reported [90]. In 2006, Ikeda *et al.* could demonstrate that high-level expressing breast tumors show higher rates of proliferation than tumor with low expression levels of CDCP1 [91]. In 2008, Uekita *et al.* revealed that expression as well as phosphorylation levels of CDCP1 correspond with the ability of invasion of scirrhous gastric cancer cells. Furthermore they showed that migration and anchorage dependent survival are responsible for invasion and peritoneal dissemination [92]. These results highlight the important role of CDCP1 in tumor biology and thus contributing with high probability to the formation of metastases at distant sites. This thesis is supported by the determination of CDCP1 on CTCs of metastatic breast cancer patients as well as the association of CDCP1 with mesenchymal or hybrid epithelial/mesenchymal breast cancer or DTC cell lines shown in this work.

3.6. Neuropilins

The neuropilins are single-pass type I transmembrane glycoproteins with a size between 120 kDa and 140 kDa. The existence of Neuropilin-1 (NRP-1, Vascular endothelial cell growth factor 165 receptor (VEGF15R), CD304, BCDA-4, UniProtKB: O14786) and Neuropilin-2 (NRP-2, Vascular endothelial cell growth factor 165 receptor 2 (VEGF15R2), UniProtKB: O60462) have been determined in vertebrates [93]. Neuropilin-1 is known to be involved in several different processes including the immune system, development of cardiovascular system, neuron-guidance, migration and angiogenesis. Furthermore, NRP-1 plays an important role in tumor pathogenesis [93-96] and is expressed on growth cones of Semaphorin III-sensitive neurons mediating chemorepulsive activity [97]. NRP-1 is composed of a small cytoplasmic intracellular domain (44 aa), a short transmembrane domain (23 aa) and a large N-terminal extracellular domain (835 aa) [98]. The extracellular (N-Terminus) part contains three subdomains: a complement-binding CUB domain (a1/a2), coagulation factor V/VIII (b1/b2) domain and a meprin or MAM domain [93, 99]. The co-receptor activity of NRP-1 is ensured by dimerization mediated by the transmembrane domain and the meprin domain [99, 100]. The cytoplasmic domain mediates binding to several proteins involved in signaling as well as in promoting angiogenic activity [93, 101, 102]. NRP-1 represents an important co-receptor for a wide range of extracellular molecules acting as ligands, respectively class III and class IV semaphorins [93, 103], as well as vascular endothelial growth factor 165 (VEGF165) [93, 104, 105] and transforming growth factor beta (TGF- β) [93, 106, 107]. Furthermore, several secreted isoforms of NRP-1, described as sNRP-1 (soluble NRP-1) exist lacking cytoplasmic and transmembrane domains but still keeping the ability to bind ligands [93, 108].

Neuropilin-2 is also a single-pass transmembrane glycoprotein like Neuropilin-1 a small cytoplasmic intracellular domain (42 aa), a short transmembrane domain (25 aa) and a large N-terminal extracellular domain (844 aa) [98]. NRP-2 plays an important role in angiogenesis as well as in supporting breast cancer cell migration. There is evidence that this cell surface receptor mediates endothelial cell migration which might result in angiogenesis [109]. Additionally, NRP-2 is a co-receptor for binding of vascular endothelial growth factor C (VEGF-C) which is involved in lymphangiogenesis and thereby promoting lymph node metastasis in breast cancer [110] and several other human cancers [109, 111]. NRP-2 binds with high affinity to semaphorin E and semaphorin IV and not semaphorin III like NRP-1 [112]. The repulsive activity is mediated by the binding of the CUB and coagulation factor domains of NRP-2 to semaphorin

IV. In addition, NRP-1 and NRP-2 are able to form heterodimers and the composition of the structure of NRP-2 is very similar to the structure of NRP-1 [97] but NRP-2 exhibits several alternative splice variants [97, 112].

3.7. Discoidin, CUB and LCCL domain-containing protein 2

Discoidin, CUB and LCCL domain-containing protein 2 (DCBLD2; UniProtKB: Q96PD2) is a single-pass type I transmembrane glycoprotein with a molecular weight of 130 kDa [113, 114] which is also known as CUB and LCCL-homology, coagulation factor V/VIII homology domains protein 1 (CLCP1) and endothelial and smooth muscle cell-derived neuropilin-like protein (ESDN) [115]. The neuropilin-like protein DCBLD2 was first identified in vascular injury [114, 115] and has a regulatory function in vascular cell growth [114]. DCBLD2 is involved in the negative regulation of cell proliferation [114, 116] but seems to play a crucial role in proliferation of tumor cells and metastasis formation [116]. Furthermore, DCBLD2 might be involved in the EGFR stimulation of cancer cells [115]. The extracellular region of DCBLD2 is composed of a signal sequence, a CUB domain, a LCCL homology domain and a coagulation factor V/VIII-homology domain. The structure shows high similarity to the neuropilins but also distinct differences. One difference is the presence of a LCCL domain between the CUB domain and the coagulation factor V/VIII-homology domain. Furthermore, DCBLD2 consists of a single-pass transmembrane region and an intracellular C-terminal scaffolding domain [113]. Interestingly, DCBLD2 exhibits the longest cleavable secretion signal among eukaryotes [114]. Semaphorin 4B might be a potential ligand for DCBLD2 leading to clustering of DCBLD2 and resulting in activation of bound SFKs (Src family kinases) but it was also observed that binding of semaphorin 4B to DCBLD2 leads to ubiquitinylation of DCBLD2 followed by degradation. The receptor tyrosine kinases EGFR and VEGFR mediated signaling seems to be affected by DCBL2 [113].

3.8. Chimeric antigen receptor T-cell technology (CAR-T)

These properties highlight the possible important role of CDCP1, NRP-1, NRP-2 and DCBLD2 in tumor progression, cancer cell dissemination and thus contributing with high probability to the formation of metastases at distant sites. All biomarkers might also function as potential immune target for personalized therapies. Highly personalized approaches, like the chimeric antigen receptor T-cell (CAR-T) technology, are an option to specifically attack cancer cells while reducing unwanted side effects. CARs usually contain specificity-conferring extracellular antibody single chain variable fragment (scFv), a CD3z-domain and intracellular costimulatory domains. The procedure begins with the isolation of T-cells from the blood of a patient, followed by a genetic modification to recognize the surface protein of interest and re-injection into the donor [117-119].

3.9. Magnetic activated cell sorting (MACS)

One part of this work was to acquire a better understanding of cancer progression, especially breast cancer. CTCs and DTCs released from the primary tumor play an important role in this process. Mainly, CTCs with mesenchymal (mCTCs) or hybrid epithelial/mesenchymal phenotype (emCTCs) are in the main focus. The enrichment of CTCs or DTCs from the blood or the bone marrow of cancer patients by surface molecules is a crucial step for analyzing these cells regarding their biological properties and behavior in the context of disease progression, metastatic potential and therapy decisions. Currently, immunomagnetic methods (MACS) are widely-used to isolate tumor cells. For example, the immunomagnetic isolation via the extracellular domain of EpCAM is used to capture EpCAM-positive CTC in metastatic breast cancer [120]. This approach enables the isolation of CTC with an epithelial phenotype. The disadvantage of this approach is the loss of EpCAM-negative CTCs. EpCAM-negative tumor cells might have undergone a complete EMT which leads to a mesenchymal phenotype or a partial EMT which leads to a hybrid epithelial/mesenchymal phenotype. Mesenchymal CTCs are thought to be more invasive than the epithelial counterparts [36] but the hybrid epithelial/mesenchymal phenotype is linked with metastasis formation and metastatic relapse [36, 74]. These findings indicate that there is need for a cell surface molecule which can be used for a MACS-based isolation of CTC with mesenchymal or hybrid epithelial/mesenchymal attributes. Hence, the large extracellular domain of CDCP1 might be applicable

for MACS-based isolation of CTCs/DTCs with mesenchymal or hybrid epithelial/mesenchymal phenotype. Due to the fact that different biological variants of CDCP1 correlate with dissemination and invasion of tumor cells [121], CDCP1-positive tumor cells might provide information about metastasis formation-potential and metastasis development.

For the labeling procedure an indirect and a direct variant can be distinguished. The antigen of interest on the cell surface can be hybridized with an antibody or a biotinylated antibody followed by the binding of the super-paramagnetic particles (MACS-nanobeads, size range: 20-100 nm [122]) which are conjugated to a secondary antibody (recognizing the primary antibody) or to Streptavidin (recognizing the biotinylated antibody). This approach represents the indirect technique. The faster direct approach is represented by the direct binding of MACS bead-conjugated antibodies to the cell surface antigen of interest. For the isolation procedure a positive and a negative selection can be differentiated. The positive selection enriches the cells harboring the antigen of interest on their cell surface and the cells lacking the cell surface antigen are discarded. Whereas the negative selection isolates all cells lacking the cell surface antigen of interest while the cells expressing the cell surface antigen are retained on the magnetic column [123]. After successful labeling with particles the cells are applied to a ferromagnetic iron-column positioned in a magnetic field. The procedure is described for the positive selection. Nanoparticle-labeled cells retain in the column and unlabeled or antigen-negative cells pass the column and will be discarded. Thereafter, the magnetic column is removed from the magnet and the labeled cells are eluted and further analysis can be performed [122, 124] (**figure 4**).

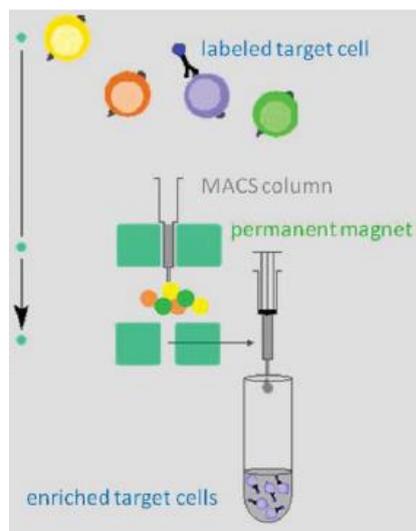


Figure 4: Cell surface molecule-dependent MACS isolation procedure. The complex composed of the biotinylated antibody and the Streptavidin-conjugated magnetic beads (black) recognize and bind to the target cell (purple). The cell-antibody-bead complex binds to the MACS column placed into a permanent magnet. The unlabeled cells or cells negative for the surface protein are discarded after passing the column without binding. After removal of the column from the magnet the labeled cells are collected and represent an enriched population. The image was edited according to the template figure 1 from “Small But Mighty: How the MACS1-Technology Based on Nanosized Superparamagnetic Particles has Helped to Analyze the Immune System Within the Last 20 Years” by Grutzkau *et al.* [122].

3.10. Therapy

More than a half of the tumors in the breast are diagnosed by mammogram screening procedures. 62 % of the tumors in the breast in the United States are restricted to the breast at the time point of indication whereas 31 % have already infiltrated the lymph nodes and 6 % exhibit metastatic spread to distant organs and regional lymph node infiltration [125, 126]. The diagnosis is based on histopathology according to standardized criteria. The most prevalent histologic patterns of breast tumors are the invasive ductal carcinoma which occurs in 50-75 % of the patients and the invasive lobular carcinoma which is diagnosed in 5-15 % of all breast cancer patients. The other breast cancer cases are composed of carcinomas with mixed lobular/ductal carcinomas and other infrequent histologic patterns [125]. Different therapy options are available for the diverse histological patterns and subtypes. Surgical resection, chemotherapy and radiation therapy as well as different directed therapies against target molecules are applied.

3.10.1. Surgery

The local therapies for non-metastatic breast tumors are characterized by a total mastectomy, the whole resection of the breast or a lumpectomy, the resection of the tumor while conserving the whole breast followed by radiation. Neoadjuvant therapy prior to surgery might increase the eligibility for a breast-conserving lumpectomy [125, 127, 128].

3.10.2. Radiation

Another local therapy option for non-metastatic breast cancer is the radiation therapy. Radiation is applied either to the whole breast or a part of the breast after lumpectomy. The chest wall after mastectomy as well as the regional lymph nodes might also be treated by radiation. For example, radiation after lumpectomy correlated with a decreased recurrence of breast cancer. Postlumpectomy radiation is performed with a dose of 42.5 Gray (Gy) over 16 fractions. Ionizing radiation causes destruction of bases as well as single-strand DNA breaks and double-strand DNA-breaks [129]. For metastatic breast cancer surgical resection as well as radiation are used for palliative purposes [125, 130-132].

3.10.3. Chemotherapy

In general, chemotherapy is a strategy to prevent recurrence in stage I-III breast cancer patients and the only therapy option for triple-negative breast cancer cases. Normally, the breast cancer subtypes define the therapy strategy. Patients with hormone receptor (HR) positive tumors receive endocrine therapy and some patients additionally chemotherapy as well. All patients with ErbB-2 positive tumors receive trastuzumab-based therapy directed against ErbB-2 and chemotherapy as well as additional endocrine therapy if the tumor is HR-positive as well. Triple-negative breast cancer patients receive chemotherapy alone. The therapeutic aim for metastatic cancer patients is the prolongation of life as well as the alleviation of occurring symptoms [125].

The additional chemotherapy options for HR-positive tumors are adriamycin/cyclophosphamide (AC), adriamycin/cyclophosphamide/paclitaxel (AC-T) or docetaxel/cyclophosphamide (TC). Patients with

ErbB-2 positive tumors receive additional chemotherapy combined with ErbB-2 targeted therapies which are paclitaxel/trastuzumab (TH), Adriamycin/cyclophosphamide/paclitaxel/trastuzumab \pm pertuzumab (AC-TH \pm P), docetaxel/carboplatin/trastuzumab \pm pertuzumab (TCH \pm P) and additional endocrine therapy if the tumor is HR-positive. Patients with triple-negative breast cancer receive the chemotherapies adriamycin/cyclophosphamide (AC), adriamycin/cyclophosphamide/paclitaxel (AC-T) or docetaxel/cyclophosphamide (TC) [125].

3.10.4. Systemic therapy for non-metastatic breast cancer

Here, the subtypes of breast cancer define the applied directed therapy. The cell surface receptors estrogen receptor alpha ($ER\alpha$) with an incidence of 70 % and the progesterone receptor (PR), together referred as hormone receptors [125, 133] as well as the epidermal growth factor receptor 2 (ErbB-2) with an incidence of 20 % represent direct therapy targets in breast cancer [125, 134]. The systemic therapy for HR-positive breast tumors is composed of endocrine agents (tamoxifen, letrozole, anastrozole, exemestane) disrupting the ER-signaling. Patients with tumors overexpressing ErbB-2 might receive a therapy with anti-ErbB-2 antibodies (trastuzumab, pertuzumab) or small-molecule tyrosine kinase inhibitors (lapatinib, neratinib) [125]. For example, tamoxifen competitively inhibits the binding of estrogen to $ER\alpha$ by selectively modulating this receptor whereas the aromatase inhibitors (letrozole, anastrozole and exemestane) reduce the concentration of circulating estrogen by preventing the modification from androgens to estrogen [125]. The application of tamoxifen for five years decreases the probability of recurrence by 50 % compared without application of endocrine therapy [125, 135].

4. Aim of the thesis

The aim of this thesis is to acquire a better understanding of the bone metastasis formation in cancer patients, especially caused by breast cancer. CTCs and DTCs released from the primary tumor play an important role in this process. Mainly, CTCs with mesenchymal (mCTCs) or hybrid epithelial/mesenchymal phenotype (emCTCs) are in the main focus of cancer researchers. Their biological properties provide these cells with advantages for the survival in the blood circulation and in seeding metastasis in distant tissues. Furthermore mCTCs and emCTCs are connected with therapy resistance and thereby they act as originator for late disease relapse and bad therapy results. On the contrary, less aggressive epithelial CTCs are more easily to detect and to isolate because of the existence of epithelial marker and surface proteins (e.g. EpCAM, cytokeratins). Especially, the CellSearch system is a very effective tool to isolate epithelial CTCs by enrichment via EpCAM. However, mCTCs and emCTCs lacking these epithelial proteins are difficult to detect and to isolate in the background of hematopoietic cells which also exhibit a mesenchymal phenotype [136]. Therefore our goal was the determination of cell surface marker proteins expressed on mesenchymal or hybrid epithelial/mesenchymal tumor cells to close this “detection-gap”. This achievement improves effectiveness of isolation, detection, analysis of biological characteristics and understanding of the involvement in metastasis formation of these types of tumor cells. The current CTC isolation or detection methods use epithelial markers and therefore they do not respect the potential metastasis-founding tumor cells with mesenchymal or partial EMT/MET phenotype [36, 137]. The identified markers CDCP1, NRP-1, NRP-2 and DCBLD2 could serve as isolation molecules for these types of tumor cells by antibody-dependent enrichment techniques like MACS.

5. Results

To gain deeper insights into the understanding of bone metastasis formation of breast cancer and to improve the diagnostics of metastasis-promoting mesenchymal CTCs/DTCs a cell line model was established. For this purpose, several breast cancer cell lines and DTC cell lines were compared regarding their status of receptor tyrosine kinases and epithelial or mesenchymal properties to identify the most suitable cell lines for the model.

5.1. Cell line characterization by relevant marker proteins

Different proteins were used to characterize the cell lines by Western Blot analysis. The set of cell lines composed of MDA-231, MDA-231 SA, MDA-231 BO2 and BC-M1 was used as a validation set for the MS-data. The other set of cell lines composed of MDA-231, Hs578t, MCF-7, MDA-468, BC-M1, PC-E1 and LC-M1 was used to identify the expression status of potential biomarker proteins among different breast cancer cell lines and DTC cell lines isolated from other entities. **Figure 5** shows the expression status of CD44, CD24, Vimentin, N-Cadherin, E-Cadherin, EpCAM, CK5/8/18/19 and cytokeratins (detected by A45/BB3 or AE1/AE3) for the different cell lines. **Figure 6** shows the expression levels of EGFR, ErbB-2, AKT, pAKT, Erk 1/2, pErk 1/2, STAT3, pSTAT3, Hif-1 α and CXCR4. These proteins were used to classify these cell lines related to their degree of differentiation, origin, phenotype or metastatic potential.

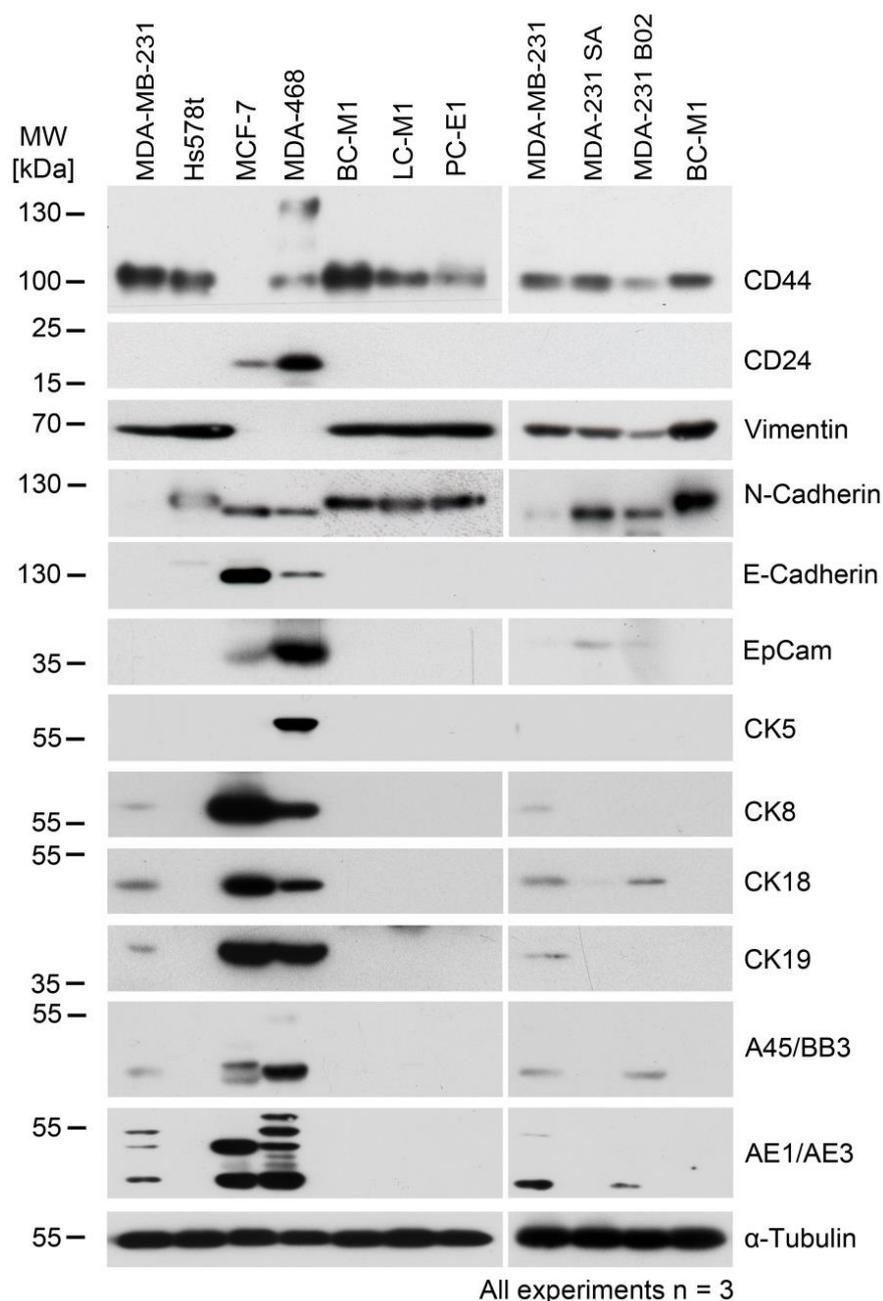


Figure 5: Western blot analysis of different breast cancer cell lines and DTC cell lines. Proteins for the characterization of the subtype, phenotype as well as stem cell properties of the cell lines are shown. α -Tubulin served as a loading control. All experiments were performed as biological triplicates. The cell line panel (figure 5, right side) with MDA-231, MDA-231 SA, MDA-231 B02 and BC-M1 was used for the validation of the MS-data set. The cell line panel (figure 5, left side) with MDA-231, Hs578t, MCF-7, MDA-468, BC-M1, LC-M1 and PC-E1 was used for the evaluation of the biological relevance of these proteins.

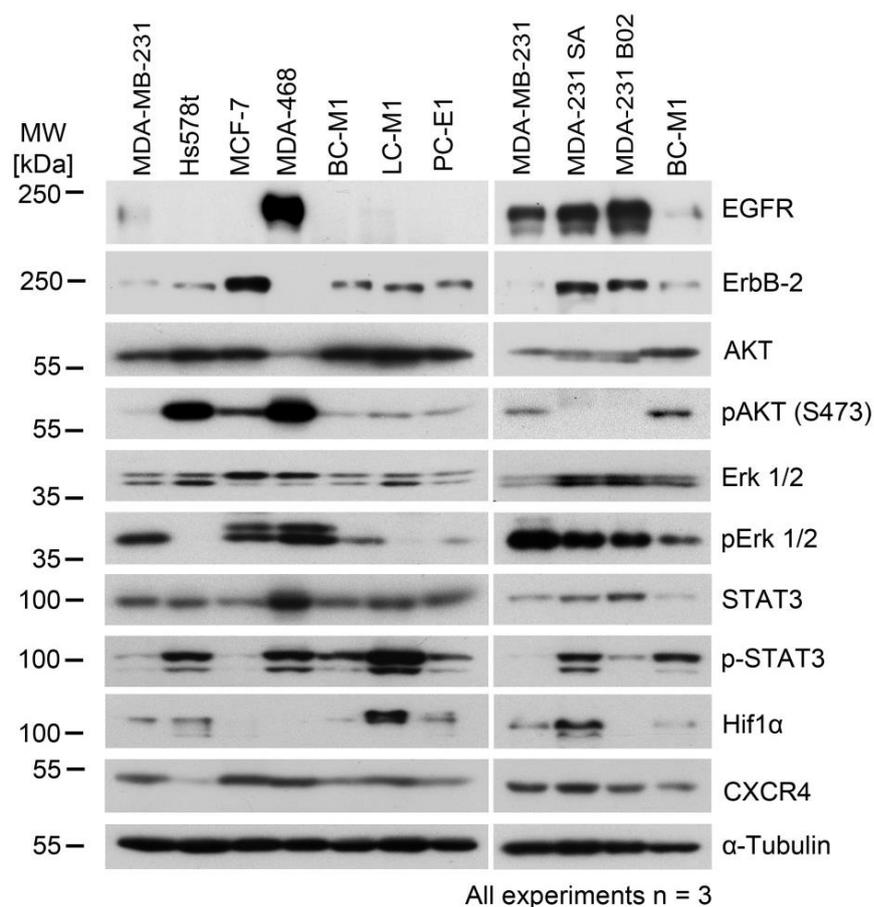


Figure 6: Western blot analysis of different breast cancer cell lines and DTC cell lines. Shown are metastasis relevant proteins like CXCR4 or the receptor tyrosine kinases EGFR and ErbB-2 and signaling relevant proteins as well as a cell stress correlated protein (Hif1 α). α -Tubulin served as a loading control. All experiments were performed as biological triplicates. The cell line panel (**figure 6**, right side) with MDA-231, MDA-231 SA, MDA-231 B02 and BC-M1 was used for the validation of the MS-data set. The cell line panel (**figure 6**, left side) with MDA-231, Hs578t, MCF-7, MDA-468, BC-M1, LC-M1 and PC-E1 was used for the evaluation of the biological relevance of the analyzed proteins.

MCF-7 and MDA-468 are characterized by the expression of the epithelial cell adhesion molecule (EpCam), E-Cadherin, CD24 and the cytokeratins CK8, CK18 and CK19. MDA-468 additionally expresses CK5 and CD44 with a double band (MW:100 kDa and 130 kDa). The expression of the cytokeratins was further proven by the pan-antibody-cocktails A45/BB3 and AE1/AE3. A45/BB3 recognizes the cytokeratin heterodimers CK8/18 and CK9/19, whereas AE1 detects cytokeratins 10,14,14,16 and AE3 detects the expression of cytokeratins 1,2,3,4,5,6,7,8. Additionally, MCF-7 and MDA-468 are positive for CXCR4, the receptor tyrosine kinases EGFR (MDA-468) or ErbB-2 (MCF-7), but negative for Vimentin, Hif1 α and N-Cadherin. MCF-7 is positive for AKT_{total}, whereas MDA-468 has a very weak expression. In contrast, MDA-468 exhibits a strong expression of phosphorylated AKT (p-AKT S473). MCF-7 is weakly positive for p-AKT (S473). MCF-7 and MDA-468 show a strong expression of mitogen-activated protein kinase 3 (MAPK 3, Erk1) and a very weak expression of mitogen-activated

protein kinase 1 (MAPK 1, Erk2). Both cell lines exhibit a strong expression of the corresponding phosphorylated proteins (pErk1/2), at which slightly higher signals are detected for MDA-468. For STAT3, MCF-7 shows a weak one-band signal, whereas MDA-468 exhibits a strong expression of STAT3. The phosphorylated STAT3 (p-STAT3) is difficult to detect in MCF-7, whereas MDA-468 shows a strong double-band signal.

Hs578t is positive for CD44, Vimentin, N-Cadherin and negative for CD24, EGFR, E-Cadherin, EpCam and cytokeratins 5, 8, 18, 19. The absence of cytokeratin expression was also proven by pan-antibody cocktails AE1/AE3 and A45/BB3. ErbB-2, Hif1 α and CXCR4 exhibit weak signals. Hs578t has a strong expression of AKT_{total} and an even higher expression of AKT phosphorylated at position S473. Erk1/2 is detectable in this cell line with a slightly higher expression level of mitogen-activated protein kinase 1 (MAPK 1, Erk2). The phosphorylated variant (pErk1/2) is not detectable. Hs578t exhibits a moderate signal for STAT3 and a strong signal for the phosphorylated variant (p-STAT3).

MDA-231 is characterized by a strong expression of CD44, Vimentin and CXCR4. This cell line is negative for CD24, E-Cadherin, EpCam and CK5 whereas N-Cadherin, ErbB-2, EGFR, Hif1 α and the cytokeratins 8,18,19 show a weak expression pattern. The weak cytokeratin expression pattern was also proven by pan-antibody cocktails AE1/AE3 and A45/BB3. MDA-231 exhibits a weak expression of p-AKT (S473) and a slightly higher expression of the unphosphorylated protein AKT. Both, mitogen-activated protein kinase 3 (MAPK 3, Erk1) and mitogen-activated protein kinase 1 (MAPK 1, Erk2) are detectable in MDA-231. The phosphorylated proteins (pErk1/2) show a higher expression level. STAT3 is detectable with a moderate signal, whereas p-STAT3 shows a weak signal.

MDA-468 [8, 138-140], MDA-231 [138, 139, 141] and Hs578t [138, 142] belong to the triple-negative cell lines lacking the expression of estrogen-, progesterone- and ErbB-2 receptor. The cancer stem cell phenotype of MDA-231 and Hs578t was determined by the expression status of CD44⁺/CD24⁻ by Western Blot analysis (**figure 5**). Additionally, it was shown that these two cell lines are Vimentin-positive. Hs578t was negative for all analyzed cytokeratins whereas MDA-231 exhibits a weak expression of CK8/18/19. These results suggest that the triple-negative cell lines MDA-231 and Hs578t exhibit stem cell characteristics with mesenchymal attributes. MDA-468 is positive for CD44 and strongly positive for CD24. Additionally, this cell line is negative for Vimentin and positive for epithelial proteins EpCam, E-Cadherin and cytokeratins CK5/8/18/19. This expression pattern suggests that MDA-468 has no

properties of stem cells but originates from epithelial tissues. These findings indicate that the origin of the triple-negative cell line MDA-468 differs from the origin of the triple-negative cell lines MDA-231 and Hs578t.

The according bone metastatic sublines MDA-231 SA and MDA-231 BO2 of the parental cell line MDA-231 are positive for CD44 (MDA-231 BO2 has a lower expression), Vimentin (MDA-231 BO2 has a lower expression), CXCR4 (MDA-231 BO2 has a lower expression), ErbB-2 and EGFR (MDA-231 SA has a lower expression). EpCam is detectable with a very weak signal in MDA-231 SA, whereas MDA-231 BO2 lacks EpCam. CK18 is detectable only in MDA-231 BO2. The pan-antibody cocktails AE1/AE3 and A45/BB3 detect a weak signal for MDA-231 BO2. Hif1 α is only detectable in MDA-231 SA. Both cell lines are negative for CD24, N-Cadherin, E-Cadherin and the cytokeratins 5, 8, 19. MDA-231 SA as well as MDA-231 BO2 are positive for AKT_{total} but show no detectable levels of phosphorylated AKT^{S473}. Both, mitogen-activated protein kinase 3 (MAPK 3, Erk1) and mitogen-activated protein kinase 1 (MAPK 1, Erk2) are strongly expressed by both bone metastatic cell lines. The phosphorylated forms (pErk1/2) also show a strong expression. STAT3 is detectable in both cell lines, with a slightly higher expression level in MDA-231 BO2, whereas p-STAT3 has a higher expression level in MDA-231 SA and a very weak signal in MDA-231 BO2.

The bone marrow DTC cell lines BC-M1 (derived from a breast cancer patient), LC-M1 (derived from a lung cancer patient) and PC-E1 (derived from a prostate cancer patient) are positive for CD44, Vimentin, N-Cadherin and have a weak expression of ErbB-2 and CXCR4. The DTC cell lines are negative for EGFR, CD24, E-Cadherin, EpCam, cytokeratins 5,8,18 and 19. Additionally, there were no detectable signals for the pan-cytokeratin antibodies AE1/AE3 and A45/BB3. The DTC cell lines differ in their Hif1 α expression levels. BC-M1 is negative, whereas PC-E1 is weakly positive and LC-M1 exhibits the strongest expression levels. BC-M1, LC-M1 and PC-E1 exhibit strong signals for AKT_{total} whereas the phosphorylated variant (pAKT S473) has very low expression levels. Erk1/2 is detectable with a moderate signal in all three DTC cell lines, whereas the phosphorylated protein (pErk1/2) is only detectable in BC-M1. STAT3 is detectable in the DTC cell lines with moderate signals for BC-M1 and PC-E1 and a higher expression level for LC-M1. The phosphorylated protein (p-STAT3) is detectable in the DTC cell lines with moderate expression levels for BC-M1 and PC-E1 and with the strongest signal for LC-M1.

The DTC cell line BC-M1 and the bone metastasis cell line MDA-231 BO2 were chosen for the cell line model representing tumor cell dissemination (BC-M1) and outgrowth of macroscopic bone metastasis (MDA-231 BO2). The protein expression profiles were compared by a SILAC-mass spectrometry based approach. The goal was to identify new biomarkers for detection of mesenchymal or metastasis promoting tumor cells to improve the detection and the understanding of dissemination and bone metastasis outgrowth.

5.2. Validation of MS-data by Western Blot analysis

Before selection of potential biomarker candidates, the MS data was validated by Western Blot analysis to verify the reliability of this data set. For this purpose, metastasis relevant proteins were analyzed by Western Blot. The proteins Connexin-43, Grp78, CD44, Integrin alpha 5, and EGFR were used as validation proteins. **Table 4** shows the results of the MS proteome analysis of the validation proteins.

Table 1: Determination of the differential expression of validation proteins Connexin-43 (Cx43), 78 kDa glucose-regulated protein (Grp78), CD44 antigen (CD44), Integrin alpha 5 (ITGA5), Integrin beta 3 (ITGB3) and epidermal growth factor receptor (EGFR) by SILAC-LC-MS/MS proteomic analysis. The proteome comparison between BC-M1 and MDA-231 BO2 was performed in 4 biological replicates. Heat shock 70 kDa protein 14 was shown as a reference protein which is not differentially expressed between both cell lines.

Swiss-Prot acc no.	recommended protein name by UniProtKB/Swiss-Prot (short name)	total number of peptides analyzed	number of unique peptides analyzed	number of biological replicates	average value (signal ratio) ^a	standard deviation	p-value ^c
P17302	Connexin-43 (Cx43)	9	3	4	-15.8 ^b	2.435	1.299 x 10 ⁻²⁹
P11021	78 kDa glucose-regulated protein (Grp78)	179	21	4	-2.35 ^b	0.333	6.787 x 10 ⁻²⁶³
P16070	CD44 antigen (CD44)	10	3	4	-2.88 ^b	0.712	2.125 x 10 ⁻¹⁴
P08648	Integrin alpha 5 (ITGA5)	8	4	4	-15.38 ^b	4.024	1.080 x 10 ⁻²⁴
P00533	Epidermal growth factor receptor (EGFR)	23	9	4	10.47	2.512	6.573 x 10 ⁻²²
Q0VDF9	Heat shock 70 kDa protein 14 (HSPA14)	25	7	4	1.049	0.361	0.254

^aA negative value of the average signal ratio correlates with a protein overexpression in BC-M1 whereas a positive value correlates with a protein overexpression in MDA-231 BO2. ^bThe quantification was performed manually. ^cStudent's test, with $p < 0.005$ was considered as a significant difference.

Connexin-43 was analyzed by 9 peptides and 3 unique peptides were used for the quantification. The average expression value^a shows a 15.8 fold-change higher expression in the cell line BC-M1 compared with MDA-231 BO2. Grp78 was analyzed by 179 peptides and 21 unique peptides were used for the quantification. The average expression value^a shows a 2.35 fold-change higher expression in the cell line BC-M1. CD44 was analyzed by 10 peptides and 3 unique peptides were used for the quantification. The average expression value^a shows a 2.88 fold-change higher expression in the cell line BC-M1. Integrin alpha 5 was analyzed by 8 peptides and 4 unique peptides were used for the quantification. The average expression value^a shows a 15.38 fold-change higher expression in the cell line BC-M1. EGFR was analyzed by 23 peptides and 9 unique peptides were used for the quantification. The average expression value^a shows a 10.47 fold-change higher expression in the cell line MDA-231 BO2. HSPA14 was analyzed by 25

peptides and 7 unique peptides were used for the quantification. The average expression value^a (1.049) indicates that this protein is equally expressed between the cell lines BC-M1 and MDA-231 BO2. The p-values of the validation proteins underlines the significantly different expression of these proteins between the cell lines. The p-value of HSPA14 (0.254) indicates that the expression of this protein does not differ significantly between the compared cell lines. For each validation protein (**table 1**) a characteristic MS1- (**figure 7, 9, 11, 13, 15**) and MS2 (**figure 8, 10, 12, 14, 16**) spectrum is shown. For the interpretation of the MS1-spectra it is necessary to respect that ¹²C₆-labeled L-Lysin and L-Arginine were incorporated into the proteins of the cell line BC-M1. ¹³C₆-labeled L-lysine and L-arginine were incorporated into the proteins of the cell line MDA-231 BO2. Due to the fact that L-Lysin and L-Arginine contain six carbon atoms, the mass (m) between ¹²C-labeled lysin/arginine and ¹³C-labeled lysin/arginine differs in six daltons (Da) but the difference of the mass to charge-ratio (m/z) between the light- and heavy-labeled peptide depends on the charge (z) of the analyzed peptides.

The annotated MS/MS fragment ion (MS2)-spectra were used for the identification of the corresponding protein. Peptide fragments occur by bond-breaking of the bond between the carbon atom and the nitrogen atom of the peptide bond after collision-induced dissociation (CID) followed by higher-energy C-trap dissociation (HCD) [143]. The peptide belongs to the y-ion series if the charge is kept at the carboxyl-terminus (C-terminus) of the fragment. Accordingly, if the charge is kept at the amino-terminus (N-terminus), the peptide fragment belongs to b-ion series. The m/z-values of the peptide fragments provide information about the contained amino acids. Thus, the primary structure (amino acid sequence) of a protein can be reconstructed from the y-ion or b-ion series [144-146].

In total, 6674 proteins were identified by the LC-MS/MS based proteome analysis between BC-M1 and MDA-231 BO2. A quantification could not be performed for 2145 proteins. A quantification could be performed for 4529 proteins resulting in a signal ratio value. 785 proteins with a signal ratio between either 0 and 0.5 or 2 and infinity were considered as differentially expressed proteins. 3744 proteins with an expression value between 0.5 and 2 were considered as not differentially expressed proteins. All 785 differentially expressed proteins were analyzed regarding their localization in the cell and their expression levels on PBMCs, immune and other blood cells.

Figure 7 shows the MS1 spectrum of a peptide of Connexin-43. SDPYHATSGALSPAK was one of the peptides used for the quantification of the different expression levels (table 1, average value (signal ratio)^a) of the protein Connexin-43 between the cell line BC-M1 and MDA-231 BO2. The higher intensity of the $[M+2H]^{2+}$ light-peaks is caused by the $^{12}\text{C}_6$ -labeled peptide and indicates a higher expression level of Connexin-43 in the cell line BC-M1. The difference between the $[M+2H]^{2+}$ Light-peaks is 0.5 m/z because the isotopes differ in 1 Da and the peptide exhibits a two-fold charge. The same applies to the differences between the $[M+2H]^{2+}$ heavy-peaks. The corresponding peaks (751.3670 / 754.3772; 751.8685 / 754.8784; 752.3700 / 755.3804) between the $^{12}\text{C}_6$ -labeled peptide and the $^{13}\text{C}_6$ -labeled peptide differ in 3 m/z. This is caused by the difference of six daltons between $^{12}\text{C}_6$ -labeled and $^{13}\text{C}_6$ -labeled lysin/arginine and the two-fold charge of the peptide.

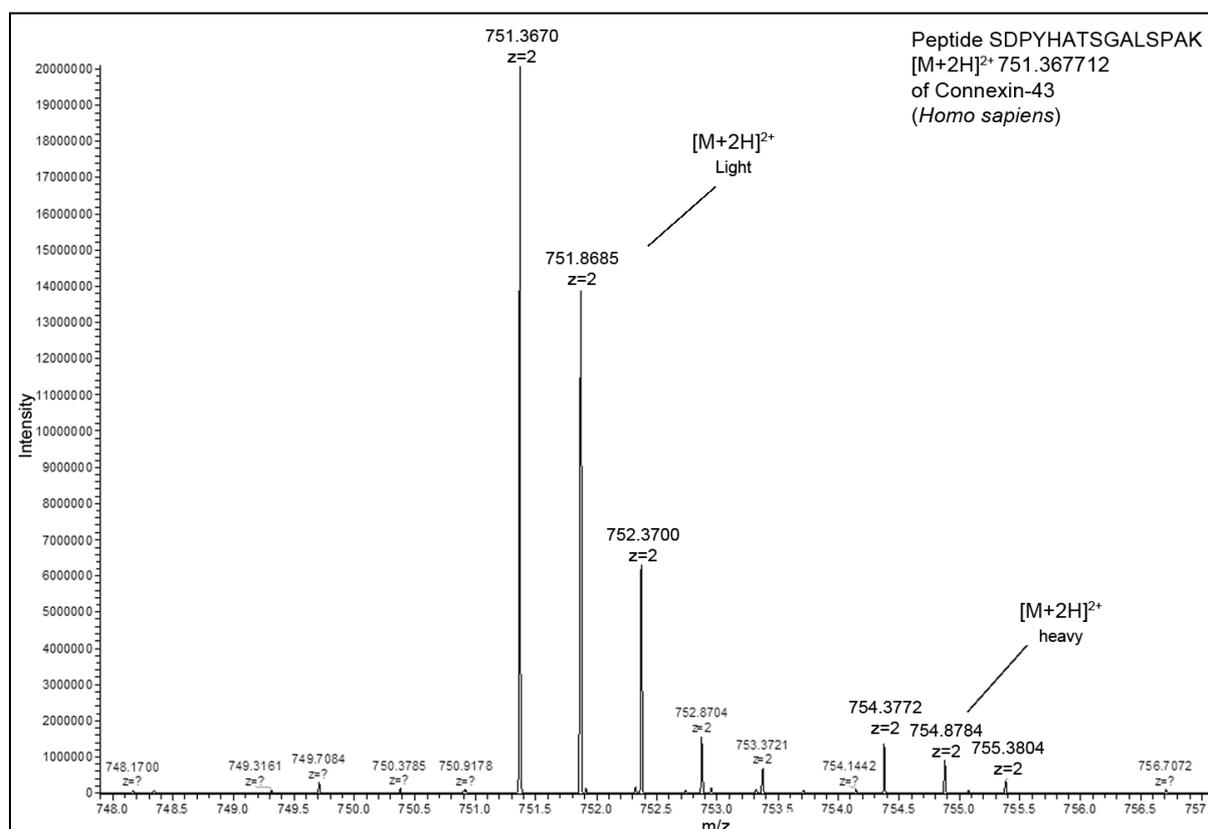


Figure 7: MS1-spectrum for the peptide SDPYHATSGALSPAK of human-Connexin-43. This spectrum was measured in positive ion mode and the peptide has a two-fold charge. The mass to charge-ratio (m/z) is plotted on the x-axis and signal intensity is plotted on the y-axis. The light-labeled peptide was derived from the cell line BC-M1 and the heavy-labeled peptide is derived from the cell line MDA-231 BO2. A SILAC-peptide-pair with three isotope peaks ($[M+2H]^{2+}$ light) for the $^{12}\text{C}_6$ -labeled peptide SDPYHATSGALSPAK (751.3670, 751.8685, 752.3700) and three peaks ($[M+2H]^{2+}$ heavy) for the $^{13}\text{C}_6$ C-labeled peptide SDPYHATSGALSPAK (754.3772, 754.8784, 755.3804) is shown. A relevant mass range is displayed.

The fragment ion spectrum (MS2) of the $^{12}\text{C}_6$ -labeled peptide SDPYHATSGALSPAK (**figure 8**) was one of the peptides used for the identification of the corresponding protein Connexin-43. The spectrum contains the y- and b-ion series. For the analyzed peptide, the complete y-ion series (except of y_{14}^+ and y_{15}^+) was detected and automatically annotated. The C-terminal $^{12}\text{C}_6$ -labeled lysine is characterized by the y_1^+ at 147.16 m/z.

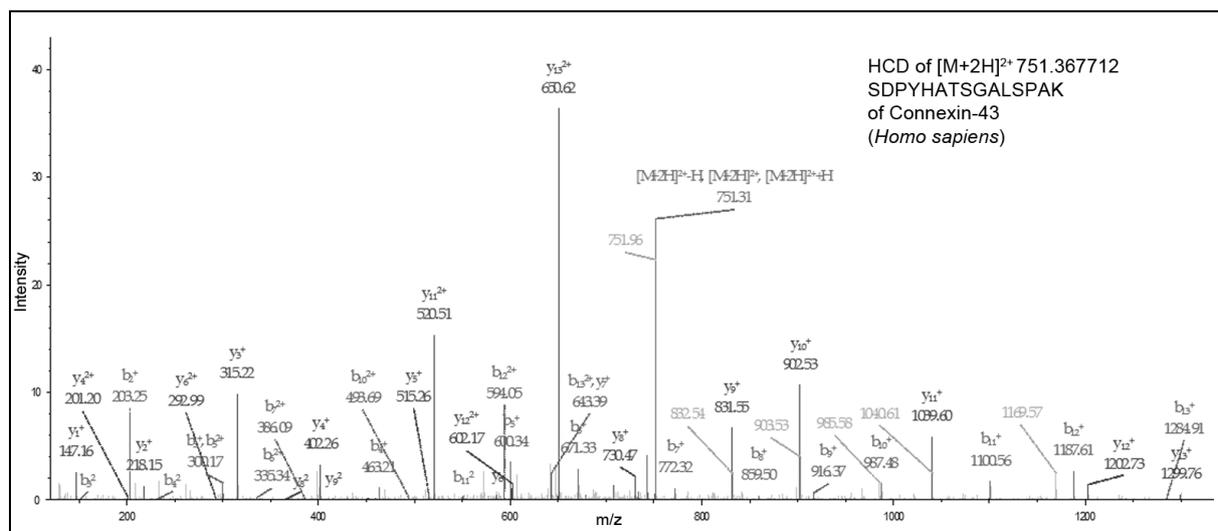


Figure 8: MS/MS (MS2)-spectrum for the $^{12}\text{C}_6$ -labeled peptide SDPYHATSGALSPAK of human-Connexin-43. This fragment-ion spectrum was generated by higher energy collisional dissociation (HCD). The mass to charge-ratio (m/z) is plotted on the x-axis and signal intensity is plotted on the y-axis. A relevant mass range is displayed. A larger scale image of this spectrum is shown in the supplement.

Figure 9 shows the MS1 spectrum of a peptide of Grp78. ELEEIVQPIISK was one of the peptides used for the quantification of the different expression levels (**table 1**, average value (signal ratio)^a) of the protein Grp78 between the cell line BC-M1 and MDA-231 BO2. The higher intensity of the $[M+2H]^{2+}$ Light-peaks is caused by the $^{12}\text{C}_6$ -labeled peptide and indicates a higher expression level of Grp78 in the cell line BC-M1. The difference between the $[M+2H]^{2+}$ Light-peaks is 0.5 m/z because the isotopes differ in 1 Da and the peptide exhibits a two-fold charge. The same applies to the differences between the $[M+2H]^{2+}$ heavy-peaks. The corresponding peaks (699.3973 / 702.4069; 699.8987 / 702.9084; 700.3998 / 703.4098) between the $^{12}\text{C}_6$ -labeled peptide and the $^{13}\text{C}_6$ -labeled peptide differ in 3 m/z. This is caused by the difference of six daltons between $^{12}\text{C}_6$ -labeled and $^{13}\text{C}_6$ -labeled lysin/arginine and the two-fold charge of the peptide.

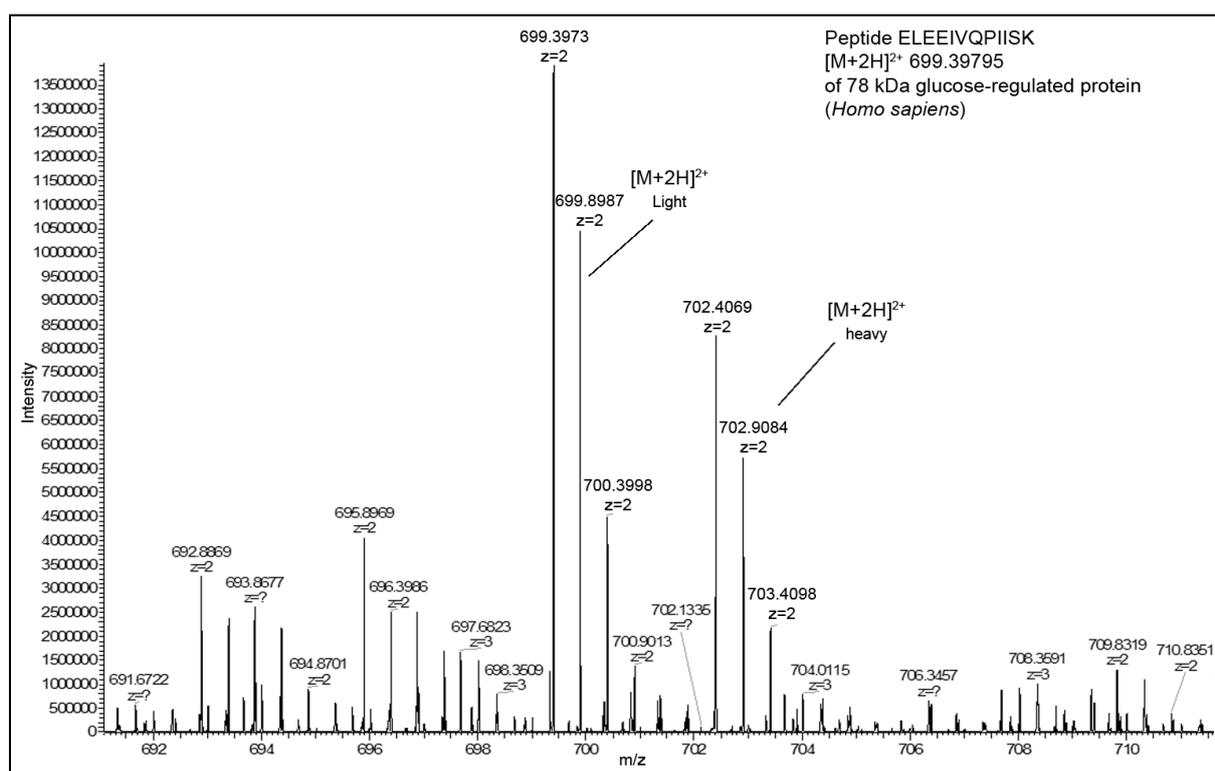


Figure 9: MS1-spectrum for the peptide ELEEIVQPIISK of human-78 kDa glucose-regulated protein (Grp78). This spectrum was measured in positive ion mode and the peptide has a two-fold charge. The mass to charge-ratio (m/z) is plotted on the x-axis and intensity is plotted on the y-axis. The light-labeled peptide is derived from the cell line BC-M1 and the heavy-labeled peptide is derived from the cell line MDA-231 BO2. A SILAC-peptide-pair with three peaks ($[M+2H]^{2+}$ light) for the $^{12}\text{C}_6$ -labeled peptide ELEEIVQPIISK (699.3973, 699.8987, 700.3998) and three peaks ($[M+2H]^{2+}$ heavy) for the $^{13}\text{C}_6$ -labeled peptide ELEEIVQPIISK (702.4069, 702.9084, 703.4098) is shown. A relevant mass range is displayed.

The fragment ion spectrum (MS2) of the $^{12}\text{C}_6$ -labeled peptide ELEEIVQPIISK (**figure 10**) was one the peptides used for the identification of the corresponding protein Grp78. The spectrum contains the y- and b-ion series. For the analyzed peptide, the complete y-ion series (except of y_{12}^+) was detected and computationally annotated. The C-terminal $^{12}\text{C}_6$ -labeled lysine is characterized by the y_1^+ at 147.17 m/z.

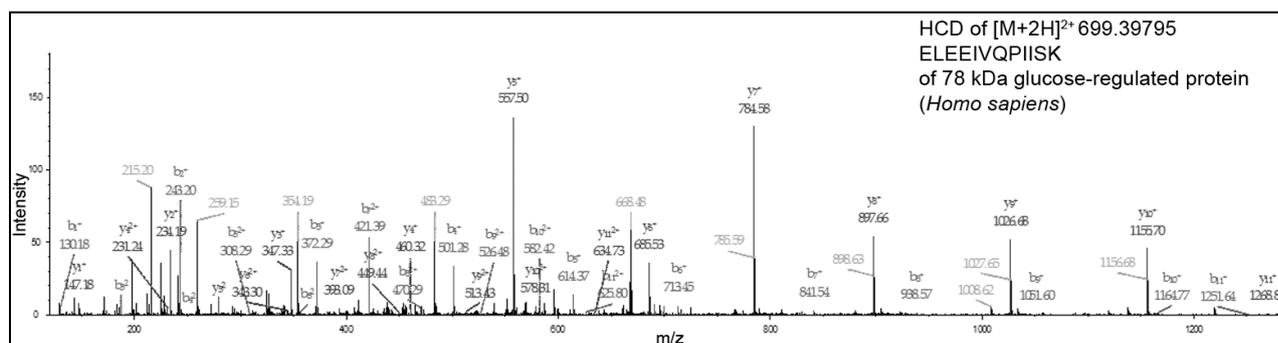


Figure 10: MS/MS (MS2)-spectrum for the $^{12}\text{C}_6$ -labeled peptide ELEEIVQPIISK of human-78 kDa glucose-regulated protein (Grp78). This fragment-ion spectrum was generated by higher energy collisional dissociation (HCD). The mass to charge-ratio (m/z) is plotted on the x-axis and intensity is plotted on the y-axis. A relevant mass range is displayed. A larger scale image of this spectrum is shown in the supplement.

Figure 11 shows the MS1 spectrum of a peptide of CD44 antigen. FAGVFHVEK was one of the peptides used for the quantification of the different expression levels (**table 1**, average value (signal ratio)^a) of the protein CD44 between the cell line BC-M1 and MDA-231 BO2. The higher intensity of the $[M+2H]^{2+}$ light-peaks is caused by the $^{12}\text{C}_6$ -labeled peptide and indicates a higher expression level of CD44 in the cell line BC-M1. The difference between the $[M+2H]^{2+}$ light-peaks is 0.5 m/z because the isotopes differ in 1 Da and the peptide exhibits a two-fold charge. The same applies to the differences between the $[M+2H]^{2+}$ heavy-peaks. The corresponding peaks (517.2773 / 520.2872; 517.7788 / 520.7887; 518.2798 / 521.2902) between the $^{12}\text{C}_6$ -labeled peptide and the $^{13}\text{C}_6$ -labeled peptide differ in 3 m/z. This is caused by the difference of six daltons between $^{12}\text{C}_6$ -labeled and $^{13}\text{C}_6$ -labeled lysin/arginine and the two-fold charge of the peptide.

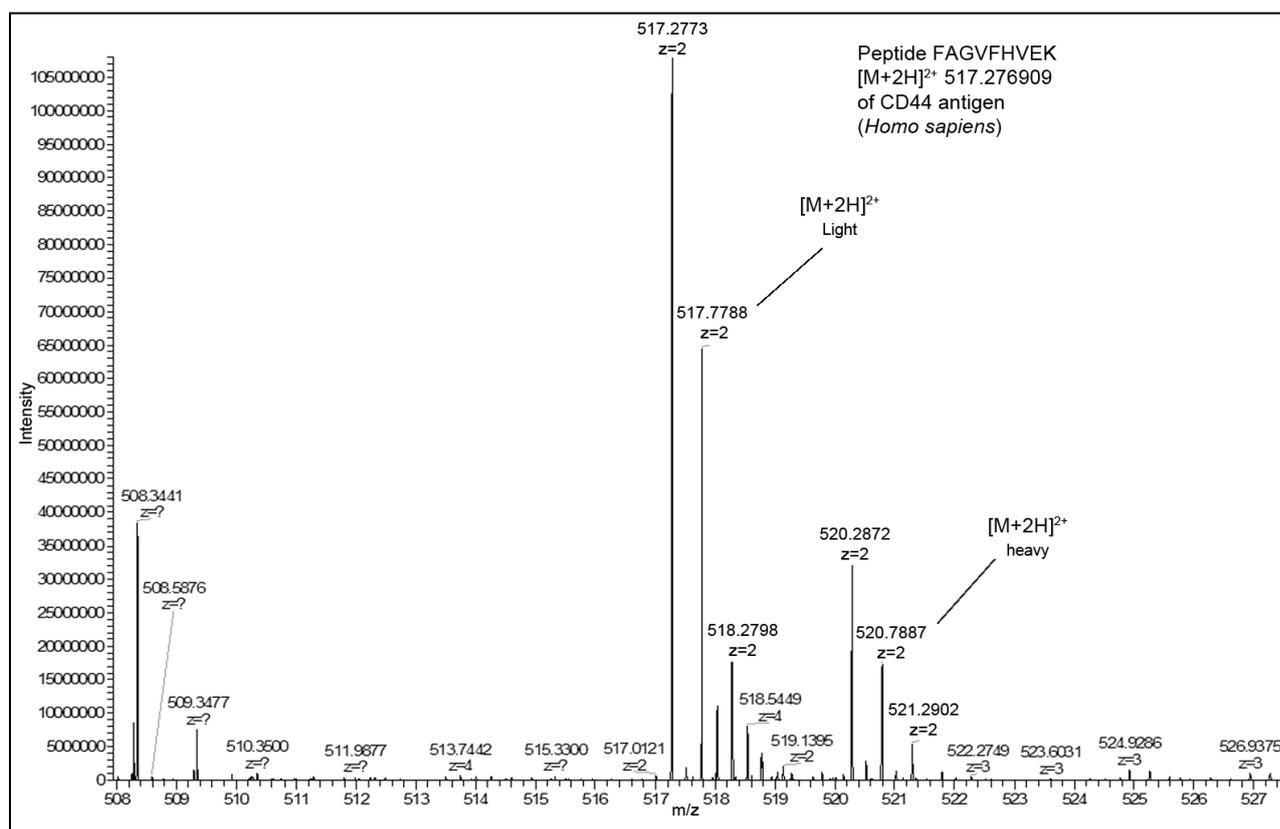


Figure 11: MS1-spectrum for the peptide FAGVFHVEK of human-CD44 antigen (CD44). This spectrum was measured in positive ion mode and the peptide has a two-fold charge. The mass to charge-ratio (m/z) is plotted on the x-axis and signal intensity is plotted on the y-axis. The light-labeled peptide was derived from the cell line BC-M1 and the heavy-labeled peptide is derived from the cell line MDA-231 BO2. A SILAC-peptide-pair with three peaks ($[M+2H]^{2+}$ Light) for the $^{12}\text{C}_6$ -labeled peptide FAGVFHVEK (517.2773, 517.7788, 518.2798) and three peaks ($[M+2H]^{2+}$ Heavy) for the $^{13}\text{C}_6$ -labeled peptide FAGVFHVEK (520.2872, 520.7887, 521.2902) is shown. A relevant mass range is displayed.

The fragment ion spectrum (MS2) of the $^{13}\text{C}_6$ -labeled peptide FAGVFHVEK (**figure 12**) was one the peptides used for the identification of the corresponding protein CD44. The spectrum contains the y- and b-ion series. For the analyzed peptide, the complete y-ion series (except of y_9^+) was detected and automatically annotated. The C-terminal $^{13}\text{C}_6$ -labeled lysine is characterized by the y_1^+ at 153.13292 m/z.

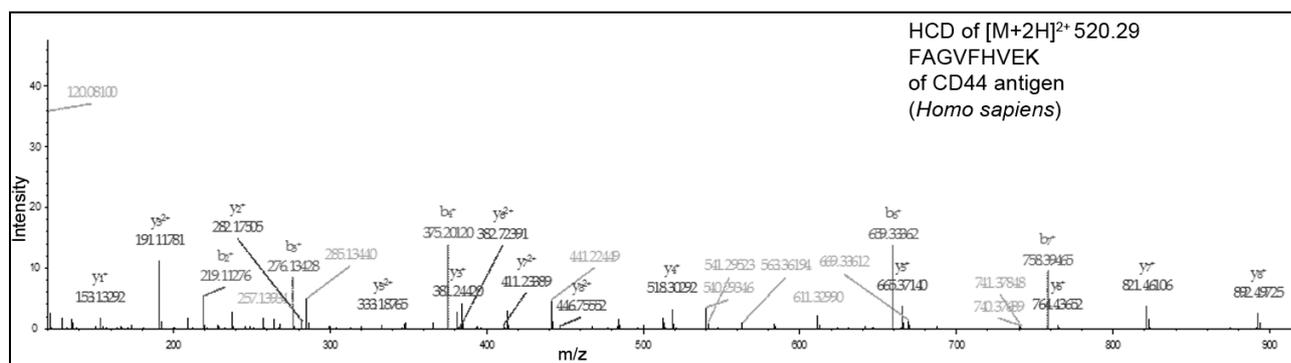


Figure 12: MS/MS (MS2)-spectrum for the $^{13}\text{C}_6$ -labeled peptide FAGVFHVEK of human-CD44 antigen (CD44). This fragment-ion spectrum was generated by higher energy collisional dissociation (HCD). The mass to charge-ratio (m/z) is plotted on the x-axis and intensity is plotted on the y-axis. A relevant mass range is displayed. A larger scale image of this spectrum is shown in the supplement.

Figure 13 shows the MS1 spectrum of a peptide of Integrin alpha 5. LLESSLSSEGEPEYK was one of the peptides used for the quantification of the different expression levels (**table 1**, average value (signal ratio)^a) of the protein Integrin alpha 5 between the cell line BC-M1 and MDA-231 BO2. The higher intensity of the $[M+2H]^{2+}$ light-peaks is caused by the $^{12}C_6$ -labeled peptide and indicates a higher expression level of Integrin alpha 5 in the cell line BC-M1. The difference between the $[M+2H]^{2+}$ light-peaks is 0.5 m/z because the isotopes differ in 1 Da and the peptide exhibits a two-fold charge. The same applies to the differences between the $[M+2H]^{2+}$ heavy-peaks. The corresponding peaks (991.9775 / 994.9870; 992.4785 / 995.4861; 992.9785 / 995.9881) between the $^{12}C_6$ -labeled peptide and the $^{13}C_6$ -labeled peptide differ in 3 m/z. This is caused by the difference of six daltons between $^{12}C_6$ -labeled and $^{13}C_6$ -labeled lysin/arginine and the two-fold charge of the peptide.

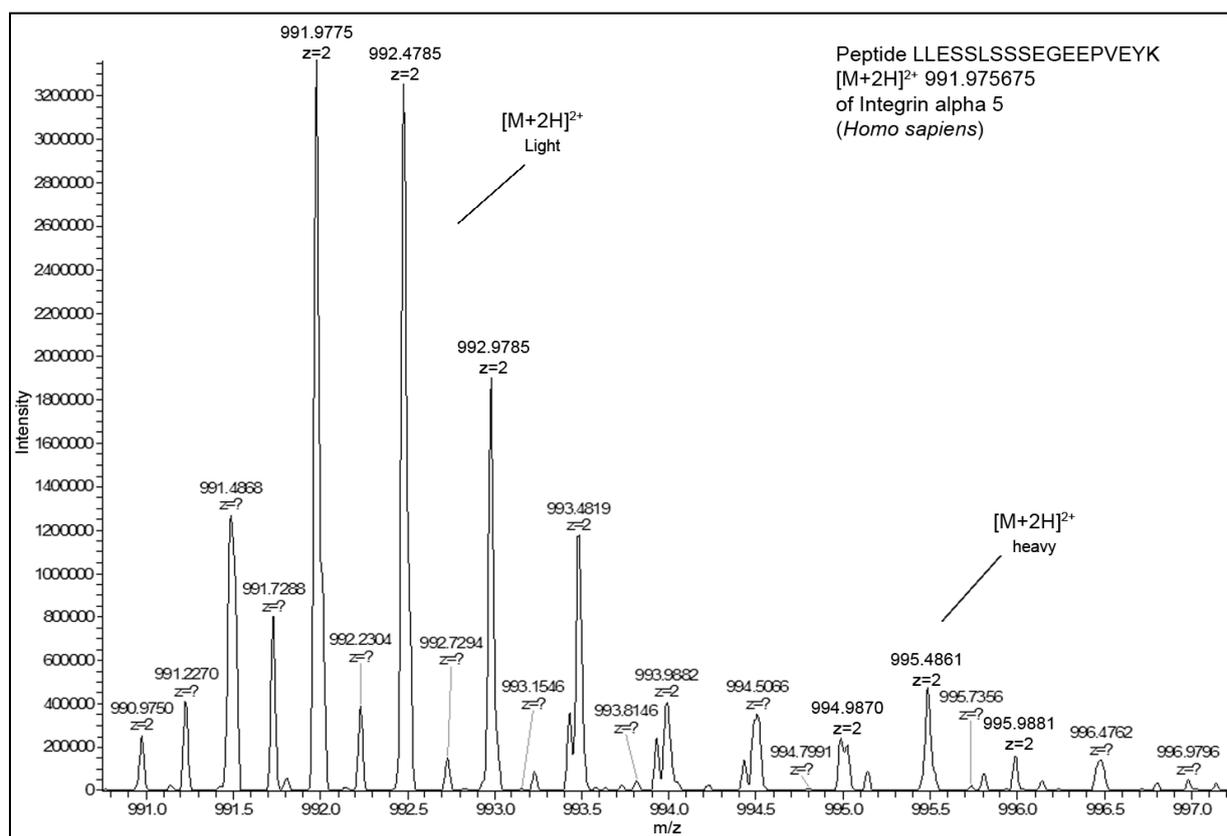


Figure 13: MS1-spectrum for the peptide LLESSLSSEGEPEYK of human-Integrin alpha 5. This spectrum was measured in positive ion mode and the peptide has a two-fold charge. The mass to charge-ratio (m/z) is plotted on the x-axis and intensity is plotted on the y-axis. The light-labeled peptide was derived from the cell line BC-M1 and the heavy-labeled peptide is derived from the cell line MDA-231 BO2. A SILAC-peptide-pair with three peaks ($[M+2H]^{2+}$ light) for the $^{12}C_6$ -labeled peptide LLESSLSSEGEPEYK (991.9775, 992.4785, 992.9785) and three peaks ($[M+2H]^{2+}$ heavy) for the $^{13}C_6$ -labeled peptide LLESSLSSEGEPEYK (994.9870, 995.4861, 995.9881) is shown. A relevant mass range is displayed.

The fragment ion spectrum (MS2) of the $^{12}\text{C}_6$ -labeled peptide LLESSLSSEGEPEYK (**figure 14**) was one of the peptides used for the identification of the corresponding protein Integrin alpha 5. The spectrum contains the y- and b-ion series. For the analyzed peptide, the complete y-ion series (except of y_{17}^+ and y_{18}^+) was detected and automatically annotated.

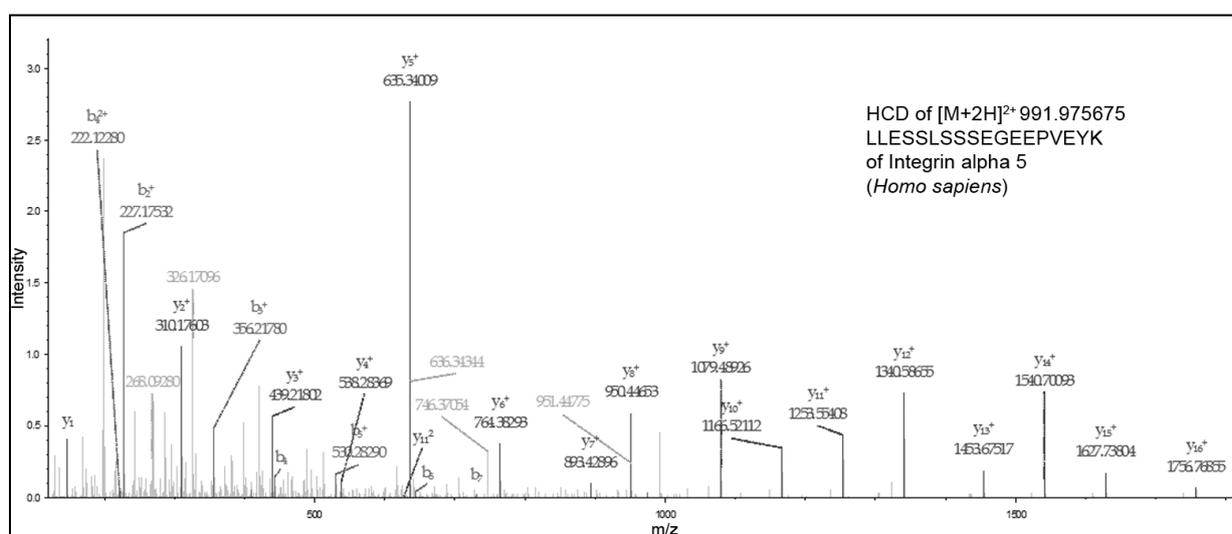


Figure 14: MS/MS (MS2)-spectrum for $^{12}\text{C}_6$ -labeled the peptide LLESSLSSEGEPEYK of human-Integrin alpha 5. This fragment-ion spectrum was generated by higher energy collisional dissociation (HCD). The mass to charge-ratio (m/z) is plotted on the x-axis and intensity is plotted on the y-axis. A relevant mass range is displayed. A larger scale image of this spectrum is shown in the supplement.

Figure 15 shows the MS1 spectrum of a peptide of EGFR. EISDGDVIISGNK was one of the peptides used for the quantification of the different expression levels (**table 1**, average value (signal ratio)^a) of the protein EGFR between the cell line BC-M1 and MDA-231 BO2. The higher intensity of the $[M+2H]^{2+}$ heavy-peaks is caused by the $^{13}C_6$ -labeled peptide and indicates a higher expression level of EGFR in the cell line MDA-231 BO2. The difference between the $[M+2H]^{2+}$ light-peaks is 0.5 m/z because the isotopes differ in 1 Da and the peptide exhibits a two-fold charge. The same applies to the differences between the $[M+2H]^{2+}$ heavy-peaks. The corresponding peaks (673.8449 / 676.8542; 674.3461 / 677.3558; 674.8482 / 677.8568) between the $^{12}C_6$ -labeled peptide and the $^{13}C_6$ -labeled peptide differ in 3 m/z. This is caused by the difference of six daltons between $^{12}C_6$ -labeled and $^{13}C_6$ -labeled lysin/arginine and the two-fold charge of the peptide.

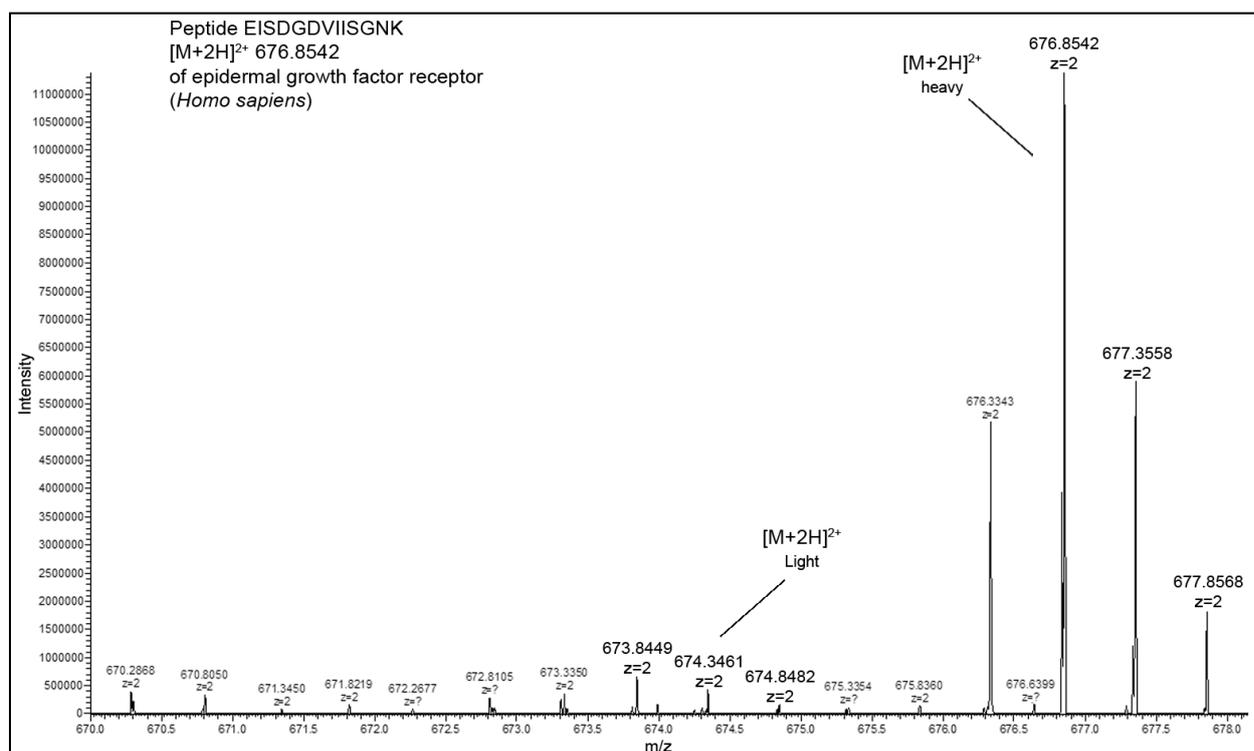


Figure 15: MS1-spectrum for the peptide EISDGDVIISGNK of human- epidermal growth factor receptor (EGFR). This spectrum was measured in positive ion mode and the peptide has a two-fold charge. The mass to charge-ratio (m/z) is plotted on the x-axis and intensity is plotted on the y-axis. The light-labeled peptide was derived from the cell line BC-M1 and the heavy-labeled peptide is derived from the cell line MDA-231 BO2. A SILAC-peptide-pair with three peaks ($[M+2H]^{2+}$ light) for the $^{12}C_6$ -labeled peptide EISDGDVIISGNK (673.8449, 674.3461, 674.8482) and three peaks ($[M+2H]^{2+}$ heavy) for the $^{13}C_6$ -labeled peptide EISDGDVIISGNK (676.8542, 677.3558, 677.8568) is shown. A relevant mass range is displayed.

The fragment ion spectrum (MS2) of the $^{13}\text{C}_6$ -labeled peptide EISDGDVIISGNK (**figure 16**) was one the peptides used for the identification of the corresponding protein EGFR. The spectrum contains the y- and b-ion series. For the analyzed peptide, the complete y-ion series (except of y_{13}^+) was detected and automatically annotated. The C-terminal $^{13}\text{C}_6$ -labeled lysine is characterized by the y_1^+ at 153.10 m/z.

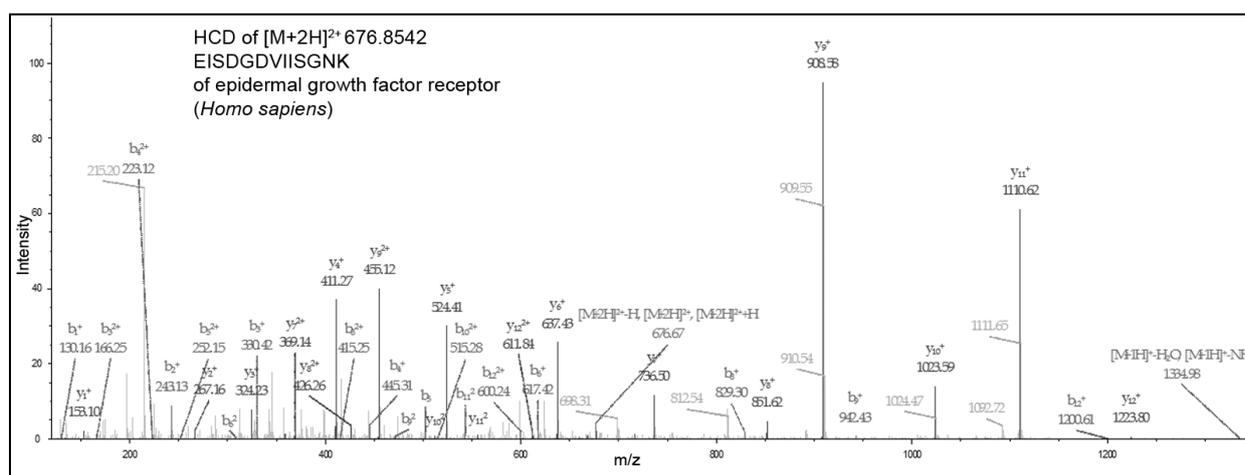


Figure 16: MS/MS (MS2)-spectrum for the $^{13}\text{C}_6$ -labeled peptide EISDGDVIISGNK of human-epidermal growth factor receptor (EGFR). This fragment-ion spectrum was generated by higher energy collisional dissociation (HCD). The mass to charge-ratio (m/z) is plotted on the x-axis and intensity is plotted on the y-axis. A relevant mass range is displayed. A larger scale image of this spectrum is shown in the supplement.

The MS2 spectra of the validation proteins (**figure 8, 10, 12, 14, 16**) indicate that the chosen validation proteins (**table 1**) have been identified correctly by the fragment ion spectra (y-ion series). To check the reliability of the MS data, the correctly identified proteins were validated by Western Blot analysis on a validation cell line panel composed of MDA-231, MDA-231 SA, MDA-231 BO2 and BC-M1 (**figure 17**).

The Western Blot analysis (**figure 17**) indicates a strong overexpression of Connexin-43 and Integrin alpha 5 and a moderate overexpression of Grp78 and CD44 in the cell line BC-M1. A strong differential expression of EGFR was detected for the cell line MDA-231 BO2. These results correlate with the signal ratios from the SILAC-MS analysis listed in **table 1**.

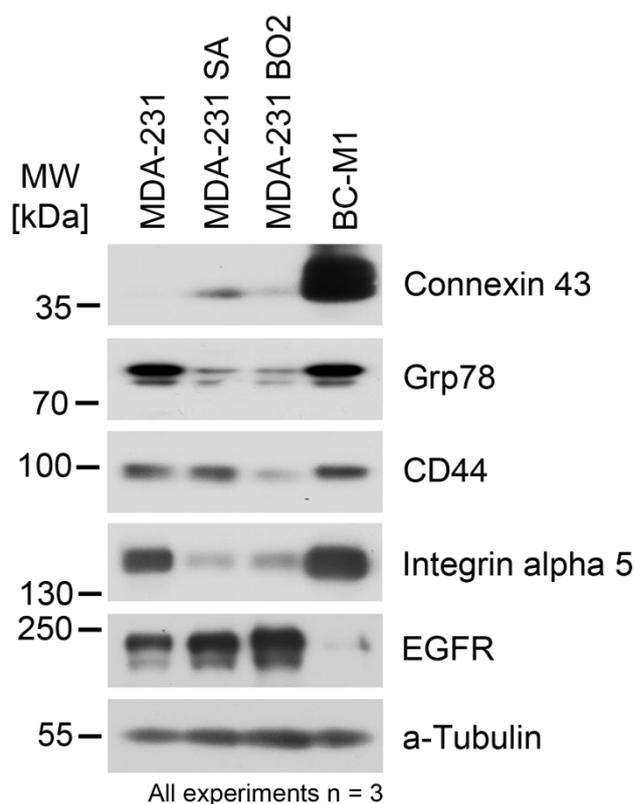


Figure 17: Western Blot analysis of proteins used for validation of the MS data from the proteome comparison analysis between BC-M1 and MDA-231 BO2. α -Tubulin serves as a loading control. All experiments were performed as biological triplicates.

The comparison of the Western Blot analysis (**figure 17**) with the MS1-spectra (**figure 7, 9, 11, 13, 15**) and the signal ratios (**table 1**) of the validation proteins indicate a reliable MS-data set and a trustable basis for further analysis.

5.3. Potential biomarker candidates from SILAC-MS-proteome analysis

After MS-data validation by Western Blot analysis, potential biomarker candidates were chosen. These candidates had to fulfill different criteria. The biomarkers should be overexpressed in the cell line BC-M1. The proteins should be present on the plasma membrane with a large extracellular domain. Furthermore, protein expression databases should predict the absence of the potential biomarker candidates on blood cells, including red blood cells, PBMC and other immune cells. The proteins CUB domain-containing protein 1 (CDCP1), Neuropilin-1 (NRP1) and Discoidin, CUB and LCCL domain-containing protein 2 (DCBLD2) fulfilled all criteria and were considered as potential biomarkers. The chosen biomarker candidates of the SILAC-MS analysis are shown in **table 2**.

Table 2: Determination of the differential expression of the biomarker candidates CUB domain-containing protein 1 (CDCP1), Neuropilin-1 (NRP1) and Discoidin, CUB and LCCL domain-containing protein 2 (DCBLD2) by SILAC-MS/MS proteomic analysis. The proteome comparison between BC-M1 and MDA-231 BO2 was performed in 4 biological replicates. Heat shock 70 kDa protein 14 was shown as a reference protein which is not differentially expressed between both cell lines.

Swiss-Prot acc no.	recommended protein name by UniProtKB/Swiss-Prot (short name)	total number of peptides analyzed ^e	number of unique peptides analyzed	number of biological replicates ^d	average value (signal ratio) ^a	standard deviation	p-value ^c
Q9H5V8	CUB domain-containing protein 1 (CDCP1)	6	4	4	-7.267 ^b	3.126	2.864 x 10 ⁻¹¹
O14786	Neuropilin-1 (NRP1)	24	8	4	-2.881 ^b	1.093	3.098 x 10 ⁻¹⁹
Q96PD2	Discoidin, CUB and LCCL domain-containing protein 2 (DCBLD2)	3	2	4	-2.837 ^b	0.868	1.713 x 10 ⁻³
Q0VDF9	Heat shock 70 kDa protein 14 (HSPA14)	25	7	4	1.049	0.361	0.254

^aA negative value of the average signal ratio correlates with a protein overexpression in BC-M1 whereas a positive value correlates with a protein overexpression in MDA-231 BO2. ^bQuantification was performed manually. ^cStudent's test, ^dtotal number of analyzed biological replicates, ^eFor CDCP1 and NRP-1: peptides were determined in 3 out of 4 replicates, for DCBLD2: peptides were determined in 2 out of 4 replicates, with $p < 0.005$ was considered as a significant difference.

The protein CDCP1 was quantified by 4 unique peptides with a 7.27-fold higher expression in BC-M1 compared with MDA-231 BO2 by LC-MS/MS analysis. NRP1 was quantified by 8 unique peptides with a 2.88-fold higher expression in BC-M1. The protein BCBLD2 was quantified by 2 unique peptides with a 2.84-fold higher expression in BC-M1. In the following figures two mass spectra for a chosen peptide of CDCP1 (**figure 18, 19**), NRP1 (**figure 20, 21**) and DCBLD2 (**figure 22, 23**) are shown. These spectra were used for quantification (MS1-spectrum) and identification (MS/MS (MS2)- spectrum) of the chosen proteins. For Neuropilin-2, the software was not able to calculate signal ratio because NRP-2 was not detected in the cell line MDA-231 BO2. After database and literature research, NRP2 was also considered as a potential biomarker candidate and therefore included into the candidate list, because NRP2, like NRP1, fulfilled all criteria which were defined before the start of the project.

Figure 18 shows the MS1 spectrum of a peptide of CDCP1. LSLVLVPAQK was one of the peptides used for the quantification of the different expression levels (**table 2**, average value (signal ratio)^a) of the protein CDCP1 between the cell line BC-M1 and MDA-231 BO2. The higher intensity of the $[M+2H]^{2+}$ light-peaks is caused by the $^{12}\text{C}_6$ -labeled peptide and indicates a higher expression level of CDCP1 in the cell line BC-M1. The difference between the $[M+2H]^{2+}$ light-peaks is 0.5 m/z because the isotopes differ in 1 Da and the peptide exhibits a two-fold charge. The same applies to the differences between the $[M+2H]^{2+}$ heavy-peaks. The corresponding peaks (534.3130 / 537.3460; 534.8143 / 537.8057; 535.3056 / 538.3446) between the $^{12}\text{C}_6$ -labeled peptide and the $^{13}\text{C}_6$ -labeled peptide differ in 3 m/z. This is caused by the difference of six daltons between $^{12}\text{C}_6$ -labeled and $^{13}\text{C}_6$ -labeled lysin/arginine and the two-fold charge of the peptide.

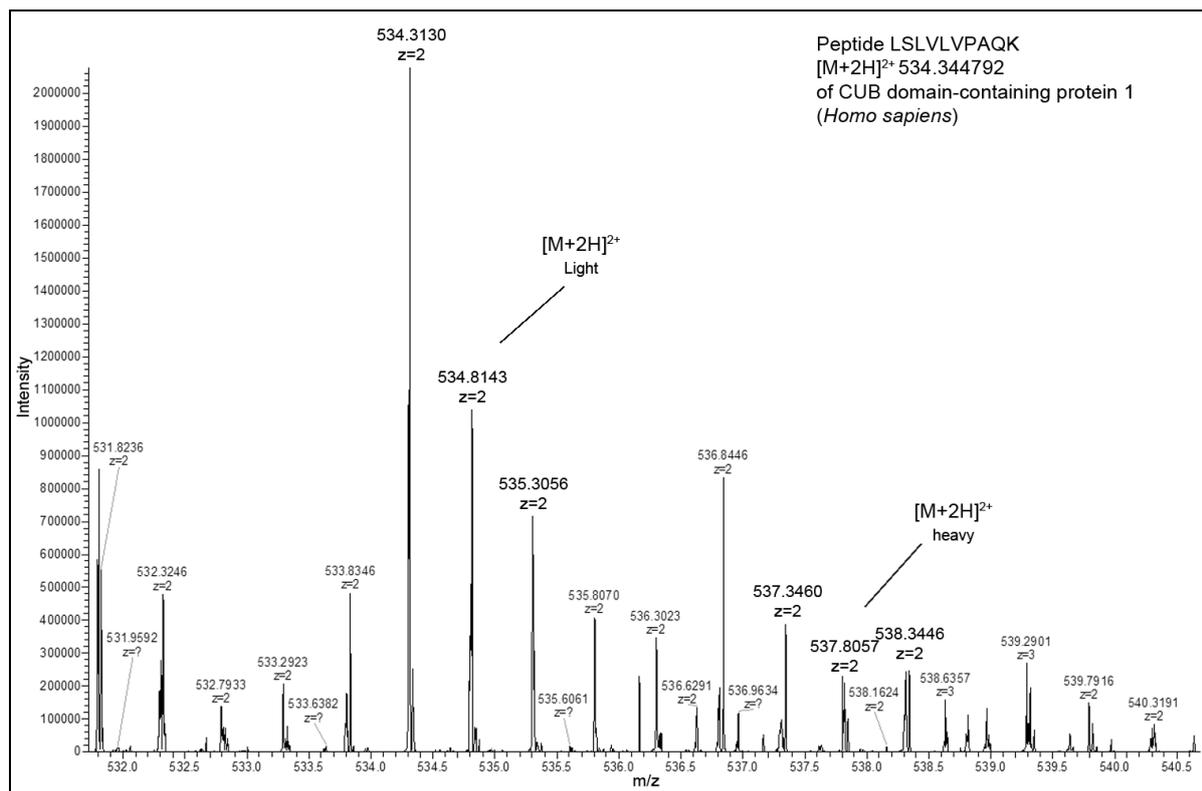


Figure 18: MS1-spectrum for the peptide LSLVLVPAQK of human- CUB domain-containing protein 1 (CDCP1). This spectrum was measured in positive ion mode and the peptide has a two-fold charge. The mass to charge-ratio (m/z) is plotted on the x-axis and intensity is plotted on the y-axis. The light-labeled peptide was derived from the cell line BC-M1 and the heavy-labeled peptide is derived from the cell line MDA-231 BO2. A SILAC-peptide-pair with three peaks ($[M+2H]^{2+}$ Light) for the $^{12}\text{C}_6$ -labeled peptide LSLVLVPAQK (534.3130, 534.8143, 535.3056) and three peaks ($[M+2H]^{2+}$ Heavy) for the $^{13}\text{C}_6$ -labeled peptide LSLVLVPAQK (537.3460, 537.8057, 538.3446) is shown. A relevant mass range is displayed.

The fragment ion spectrum (MS2) of the $^{12}\text{C}_6$ -labeled peptide LSLVLVPAQK (**figure 19**) was one the peptides used for the identification of the corresponding protein CDCP1. The spectrum contains the y- and b-ion series. For the analyzed peptide, the complete y-ion series (except of y_{10}^+) was detected and automatically annotated. The C-terminal $^{12}\text{C}_6$ -labeled lysine is characterized by the y_1^+ at 147.19 m/z.

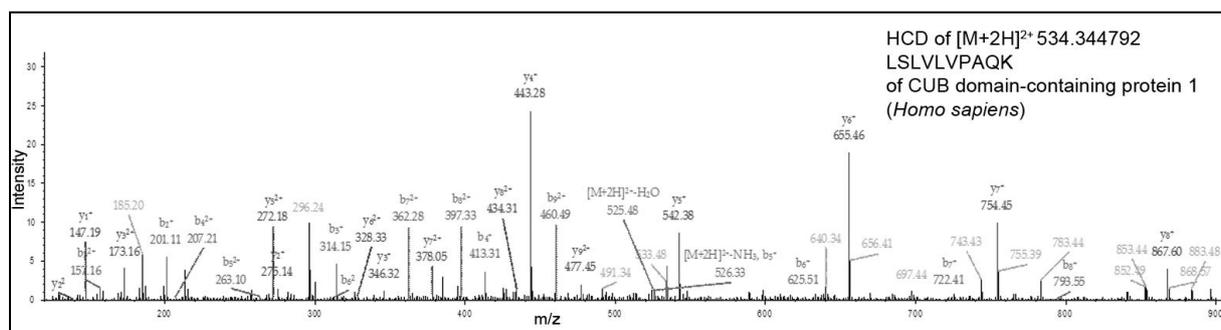


Figure 19: MS/MS (MS2)-spectrum for the $^{12}\text{C}_6$ -labeled peptide LSLVLVPAQK of human-CUB domain-containing protein 1 (CDCP1). This fragment-ion spectrum was generated by higher energy collisional dissociation (HCD). The mass to charge-ratio (m/z) is plotted on the x-axis and intensity is plotted on the y-axis. A relevant mass range is displayed. A larger scale image of this spectrum is shown in the supplement.

Figure 20 shows the MS1 spectrum of a peptide of NRP-1. IESPGYLTSPGYPHSYHPSEK was one of the peptides used for the quantification of the different expression levels (**table 2**, average value (signal ratio)^a) of the protein NRP-1 between the cell line BC-M1 and MDA-231 BO2. The higher intensity of the $[M+4H]^{4+}$ light-peaks is caused by the $^{12}C_6$ -labeled peptide and indicates a higher expression level of NRP-1 in the cell line BC-M1. The difference between the $[M+4H]^{4+}$ light-peaks is 0.25 m/z because the isotopes differ in 1 Da and the peptide exhibits a four-fold charge. The same applies to the differences between the $[M+4H]^{4+}$ heavy-peaks. The corresponding peaks (587.2829 / 588.7875; 587.5334 / 589.0382; 587.7839 / 589.2891; 588.0346 / 589.5395) between the $^{12}C_6$ -labeled peptide and the $^{13}C_6$ -labeled peptide differ in 1.5 m/z. This is caused by the difference of six daltons between $^{12}C_6$ -labeled and $^{13}C_6$ -labeled lysin/arginine and the four-fold charge of the peptide.

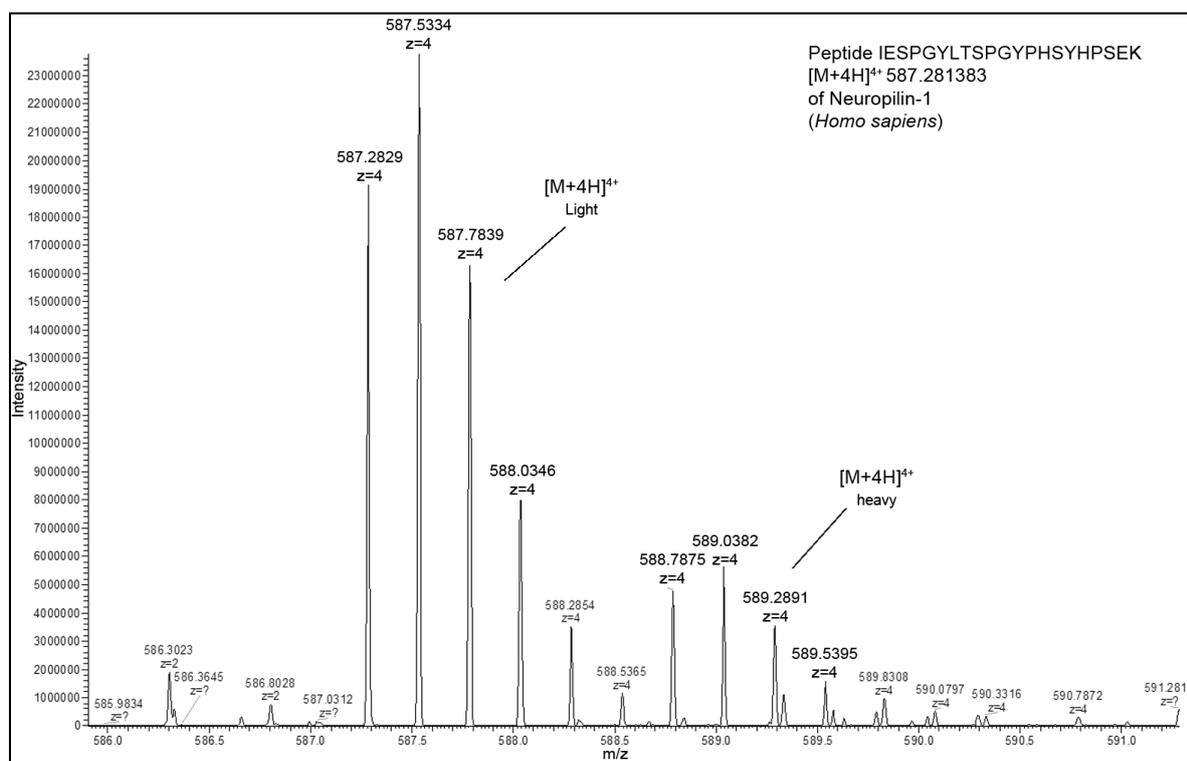


Figure 20: MS1-spectrum for the peptide IESPGYLTSPGYPHSYHPSEK of human- Neuropilin-1 (NRP1). This spectrum was measured in positive ion mode and the peptide has a two-fold charge. The mass to charge-ratio (m/z) is plotted on the x-axis and intensity is plotted on the y-axis. The light-labeled peptide was derived from the cell line BC-M1 and the heavy-labeled peptide is derived from the cell line MDA-231 BO2. A SILAC-peptide-pair with four peaks ($[M+4H]^{4+}$ light) for the $^{12}C_6$ -labeled peptide IESPGYLTSPGYPHSYHPSEK (587.2829, 587.5334, 587.7839, 588.0346) and four peaks ($[M+4H]^{4+}$ heavy) for the $^{13}C_6$ -labeled peptide IESPGYLTSPGYPHSYHPSEK (588.7875, 589.0382, 589.2891, 589.5395) is shown. A relevant mass range is displayed.

The fragment ion spectrum (MS2) of the $^{12}\text{C}_6$ -labeled peptide IESPGYLTSPGYPHSYHPSEK (**figure 21**) was one the peptides used for the identification of the corresponding protein NRP-1. The spectrum contains the y- and b-ion series. For the analyzed peptide, a partial y-ion series (except of y_{13}^+ to y_{21}^+) and a partial b-ion series (except of b_1^+ and b_7^+ to b_{21}^+) was detected and automatically annotated.

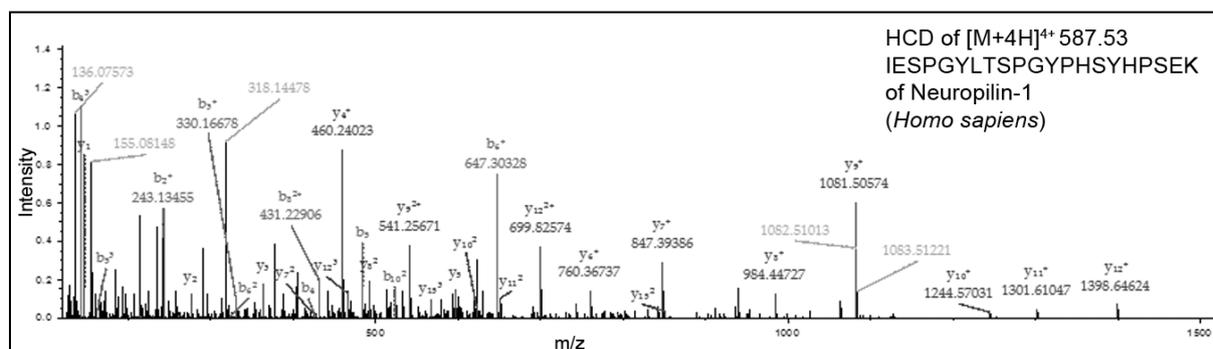


Figure 21: MS/MS (MS2)-spectrum for the $^{12}\text{C}_6$ -labeled peptide IESPGYLTSPGYPHSYHPSEK of Neuropilin-1 (NRP1). This fragment-ion spectrum was generated by higher energy collisional dissociation (HCD). The mass to charge-ratio (m/z) is plotted on the x-axis and intensity is plotted on the y-axis. A relevant mass range is displayed. A larger scale image of this spectrum is shown in the supplement.

Figure 22 shows the MS1 spectrum of a peptide of DCBLD2. DGECDVFK was one of the peptides used for the quantification of the different expression levels (table 2, average value (signal ratio)a) of the protein DCBLD2 between the cell line BC-M1 and MDA-231 BO2. The higher intensity of the $[M+2H]^{2+}$ light-peaks is caused by the $^{12}\text{C}_6$ -labeled peptide and indicates a higher expression level of DCBLD2 in the cell line BC-M1. The difference between the $[M+2H]^{2+}$ light-peaks is 0.5 m/z because the isotopes differ in 1 Da and the peptide exhibits a two-fold charge. The same applies to the differences between the $[M+2H]^{2+}$ heavy-peaks. The corresponding peaks (485.2041 / 488.2144; 485.7058 / 488.7145; 486.2069 / 489.2431) between the $^{12}\text{C}_6$ -labeled peptide and the $^{13}\text{C}_6$ -labeled peptide differ in 3 m/z. This is caused by the difference of six daltons between $^{12}\text{C}_6$ -labeled and $^{13}\text{C}_6$ -labeled lysin/arginine and the two-fold charge of the peptide.

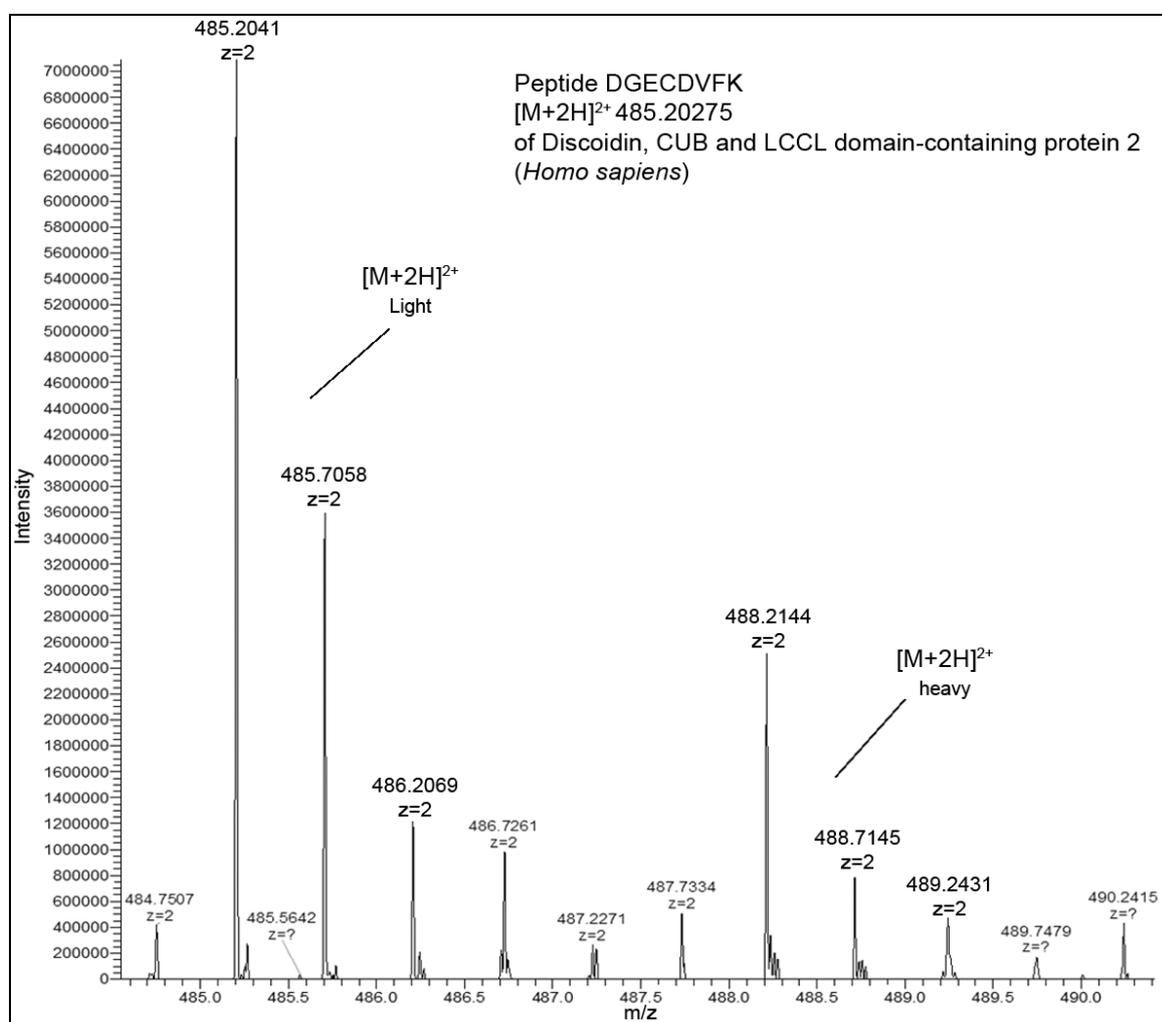


Figure 22: MS1-spectrum for the peptide DGECDVFK of human- Discoidin, CUB and LCCL domain-containing protein 2 (DCBLD2). This spectrum was measured in positive ion mode and the peptide has a two-fold charge. The mass to charge-ratio (m/z) is plotted on the x-axis and intensity is plotted on the y-axis. The light-labeled peptide was derived from the cell line BC-M1 and the heavy-labeled peptide is derived from the cell line MDA-231 BO2. A SILAC-peptide-pair with three peaks ($[M+2H]^{2+}$ light) for the $^{12}\text{C}_6$ -labeled peptide DGECDVFK (485.2041, 485.7058, 486.2069) and three peaks ($[M+2H]^{2+}$ heavy) for the $^{13}\text{C}_6$ -labeled peptide DGECDVFK (488.2144, 488.7145, 489.2431) is shown. A relevant mass range is displayed.

The fragment ion spectrum (MS2) of the $^{12}\text{C}_6$ -labeled peptide DGECDVFK (**figure 23**) was one the peptides used for the identification of the corresponding protein (DCBLD2). The spectrum contains the y- and b-ion series. For the analyzed peptide, the complete y-ion series (except of y_8^+) and partial b-ion series (b_2^+ to b_4^+) was detected and automatically annotated. The C-terminal $^{12}\text{C}_6$ -labeled lysine is characterized by the y_1^+ at 147.11279 m/z.

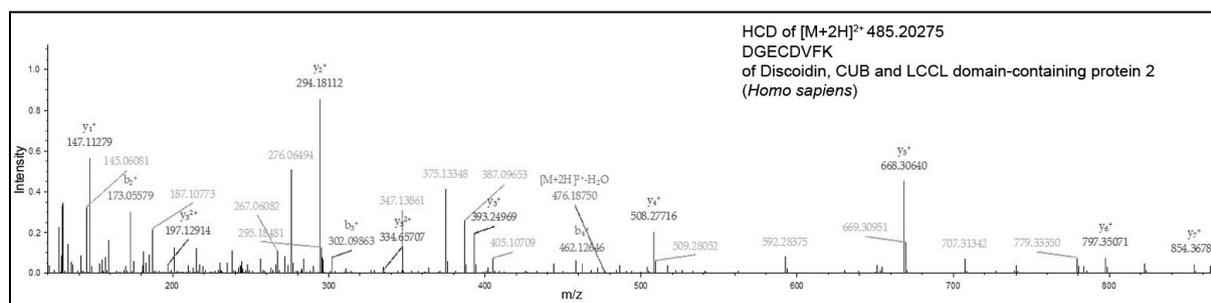


Figure 23: MS/MS (MS2)-spectrum for the $^{12}\text{C}_6$ -labeled peptide DGECDVFK of human- Discoidin, CUB and LCCL domain-containing protein 2 (DCBLD2). This fragment-ion spectrum was generated by higher energy collisional dissociation (HCD). The mass to charge-ratio (m/z) is plotted on the x-axis and intensity is plotted on the y-axis. A relevant mass range is displayed. A larger scale image of this spectrum is shown in the supplement.

The MS2 spectra (**figure 19, 21, 23**) indicate that the chosen biomarker candidates (**table 2**) have been identified correctly by the fragment ion spectra (y-ion series or partial b-ion series). The MS1 spectra (**figure 18, 20, 22**) show that all chosen candidates are overexpressed in the cell line BC-M1. To check the reliability of the MS data, the potential biomarker candidates were validated by Western Blot analysis on a validation cell line panel composed of MDA-231, MDA-231 SA, MDA-231 BO2 and BC-M1 (**figure 24**). This was a necessary control to reproduce the signal ratios determined MS-analysis (MS1-spectra, **figure 18, 20, 22**).

5.4. Validation of biomarker candidates by Western Blot analysis

The findings of the SILAC-MS analysis (**table 2**) have been validated for the proteins CDCP1, NRP1, NRP2 and DCBLD2 by Western Blot analysis (**figure 24**) for different breast cancer cell lines and dissemination cells lines from breast cancer metastasis patient (BC-M1), lung cancer metastasis patient (LC-M1) and prostate cancer metastasis patient (PC-E1). Neuropilin-2 (NRP-2) was additionally considered as potential biomarker. NRP-2 fulfilled all criteria for the choice of the biomarker candidates from SILAC-MS analysis.

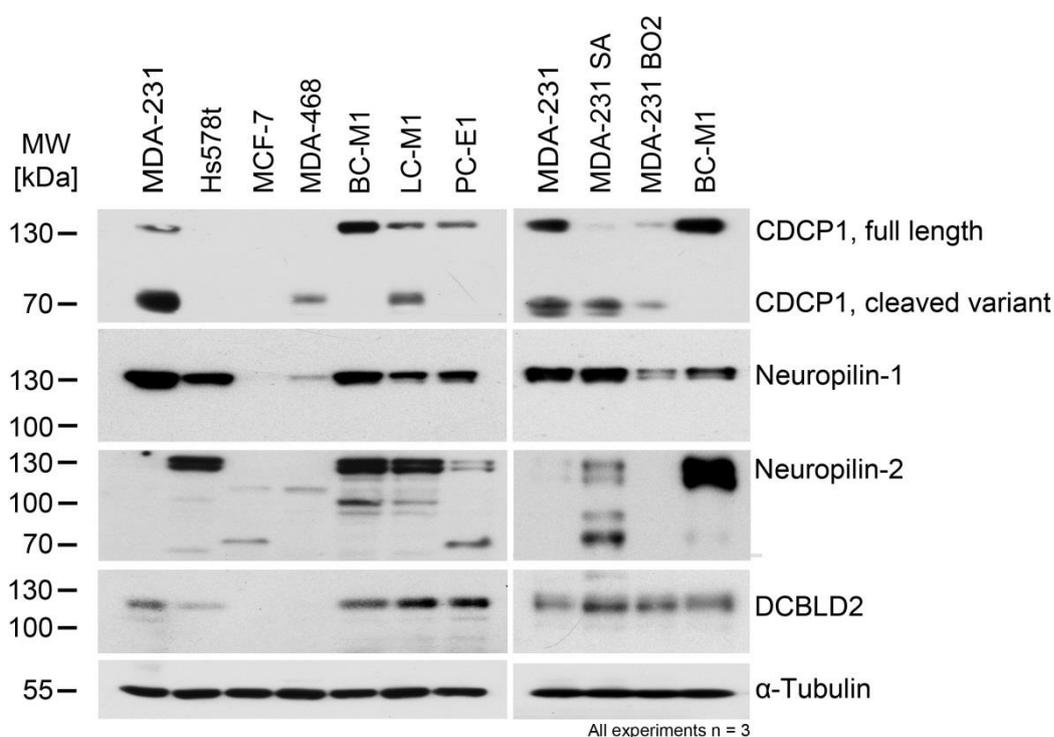


Figure 24: Western blot analysis of potential biomarkers CDCP1, NRP1, NRP2 and DCBLD2 on different breast cancer cell lines and DTC cell lines. α -Tubulin served as a loading control. All experiments were performed as biological triplicates.

The DTC cell lines (BC-M1, LC-M1, PC-E1) show the highest expression of full-length CDCP1 (130 kDa) whereas the cleaved variant of CDCP1 (70 kDa) was only detected in LC-M1 within the DTC cell lines. The bone marrow metastatic cell lines MDA-231 BO2 and MDA-231 SA contain low amounts of full-length CDCP1 but a higher amount of the cleaved variant. MDA-231 SA has a higher expression of the cleaved membrane-fragment of CDCP1 than MDA-231 BO2. MDA-231 shows the highest expression level of the cleaved variant of CDCP1 but displays a low amount of full length CDCP1. HS578t and MCF-7 are

negative for both variants of CDCP1. MCF-7 will serve as a negative control for CDCP1 in further experiments. In general, cytokeratin-low expressing or cytokeratin-negative and Vimentin-positive cell lines show the highest expression of CDCP1. The analysis of CDCP1 in BC-M1 by Western Blot confirms the dependability of the mass spectrometry data. Interestingly, the detection of CDCP1 in other breast cancer cell lines and DTC cell lines (LC-M1 and PC-E1) derived from other tumor entities support the reliability of our BC-M1/MDA-231 BO2 model. In addition, the detection of both CDCP1 variants in the highly invasive breast cancer cell line MDA-231 is a further confirmation of the relevance of CDCP1 in context of metastasis formation. Both triple negative (ER-/PR-/ErbB-2-) cell lines MDA-231 (positive for both CDCP1 variants, positive for Vimentin and weakly positive for cytokeratins) and MDA-468 (positive for cleaved variant of CDCP1, but negative for Vimentin and positive for cytokeratins) are a good example for the use of CDCP1 as a biomarker for triple-negative tumor cells. In contrast, the cell line Hs578t (CDCP1-negative, but positive for Vimentin and negative for cytokeratins) represents an exception regarding the correlation between CDCP1 and mesenchymal phenotype.

The DTC cell lines (BC-M1, LC-M1, PC-E1), MDA-231, MDA-231 SA and Hs578t exhibit the highest expression levels of Neuropilin-1. MDA-231 BO2 expresses a lower amount of NRP-1 compared to the strong positive cell lines. MDA-468 expresses a low amount of this protein and MCF-7 is negative for NRP-1.

The highest expression of Neuropilin-2 (MW 130 kDa) is detectable in the cell lines Hs578t, BC-M1 and LC-M1. MDA-231 SA and PC-E1 express lower amounts of NRP-2. The cell lines MDA-231, MDA-231 BO2, MCF-7 and MDA-468 are negative for the expression of the 130 kDa variant of NRP-2. The origin of lower molecular weight bands is not clear and needs to be further investigated.

For Discoidin, CUB and LCCL domain-containing protein 2 (BCBLD2), the DTC cell lines (BC-M1, LC-M1, PC-E1), MDA-231, MDA-231 SA and MDA-231 BO2 exhibit the highest expression levels. Hs578t produces lower amounts of this protein. MCF-7 and MDA-468 are negative for DCBLD2.

The results of the Western Blot analysis (**figure 24**) correlate with the MS1-spectra (**figure 18, 20, 22**) and the signal ratios (**table 2**) of the potential biomarker candidates and provide a reliable basis for further analysis.

5.5. Expression of potential biomarker candidates on CTC cell lines

The cell line CTC-ITB001 was established at the Institute of Tumor Biology at the University Medical Center Hamburg-Eppendorf by Andra Kuske [147]. This cell line was generated from peripheral blood (7.5 mL) of a metastatic breast cancer patient with more than 100 CTCs/7.5 mL blood. The first diagnosis of this patient described a lymphatic and bone metastatic bilateral (invasive-lobular/invasive-ductal) mamma carcinoma with ER⁺/PR⁺/ErbB-2⁻ receptor status. The cell line CTC-ITB001-MIND is the corresponding mouse-xenograft model (see methods and materials). The CTC cell line CTC-ITB001 represents a suitable model for the CTC-biomarker establishment because it was generated from CTCs of a breast cancer patient. The corresponding xenograft cell line CTC-ITB001-MIND is a model for metastasis caused by CTCs. Both models are important for the establishment and validation of CDCP1 as a biomarker. These CTC cell lines were characterized for relevant proteins by Western Blot analysis (**figure 25**) to classify them into biological context.

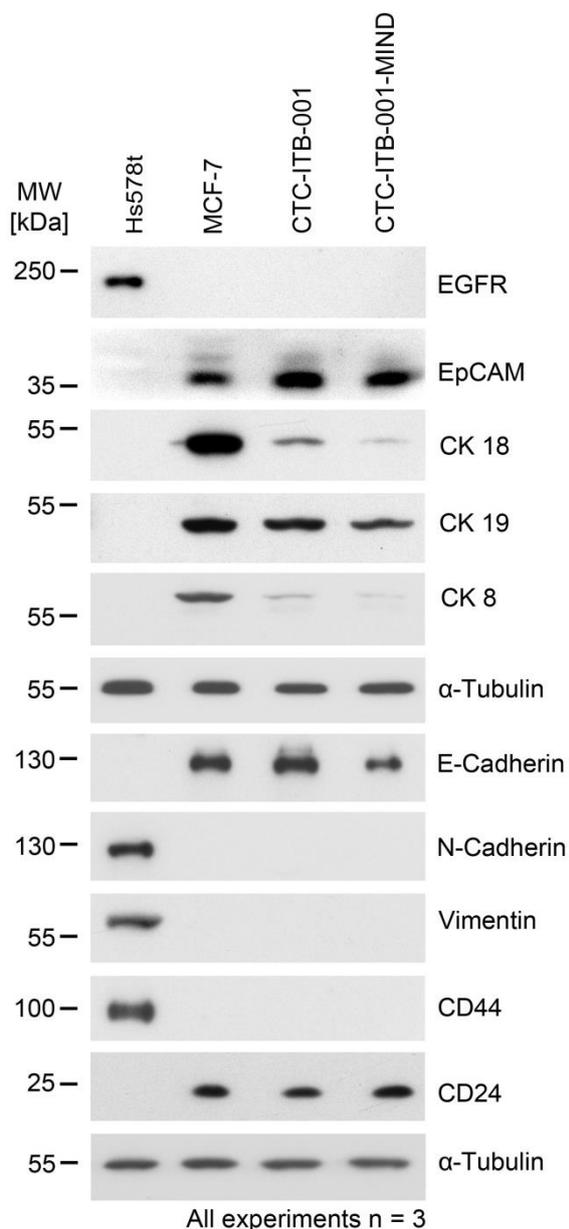


Figure 25: Western blot analysis of the CTC cell lines CTC-ITB001 and the corresponding mouse xenograft model CTC-ITB001-MIND. Hs578t served as a control cell line with mesenchymal properties, MCF-7 served as control cell line with epithelial origin. α -Tubulin served as a loading control.

The cell lines Hs578t and MCF-7 were characterized in **figure 5 and 6**. Hs578t serves as a control cell line with mesenchymal properties whereas MCF-7 serves as a control cell line with epithelial attributes. Both CTC cell lines are negative for EGFR, N-Cadherin, Vimentin and CD44. Additionally, both cell lines are positive for EpCAM, CK19, E-Cadherin and CD24. They exhibit a weak expression of CK18 and CK8.

The mouse xenograft CTC cell line CTC-ITB001-MIND shows a slightly weaker expression of CK18, CK8 and E-Cadherin and a slightly stronger expression of CD24 than the parental CTC cell line CTC-ITB001.

The CTC cell lines were analyzed for the expression level of the biomarker candidates CDCP1, NRP-1, NRP-2 and DCBLD2 (**figure 26**).

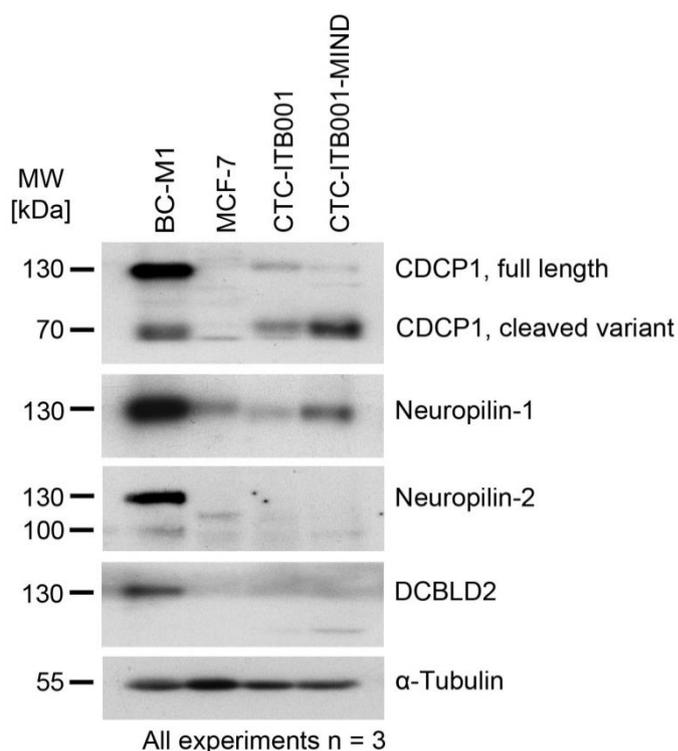


Figure 26: Western blot analysis of CDCP1 on the circulating tumor cell line CTC-ITB001 and the corresponding mouse xenograft model CTC-ITB001-MIND. BC-M1 served as a positive control, MCF-7 served as negative control and α -Tubulin served as a loading control.

The expression of the full length variant of CDCP1 (130 kDa) has a low signal intensity and the cleaved product of CDCP1 (70 kDa) shows a slightly higher expression in the CTC cell line CTC-ITB001. The signal intensity of the full length variant in the corresponding mouse xenograft model CTC-ITB001-MIND shows a weaker expression than the full length variant in the originating cell line CTC-ITB001. On the contrary the expression of the cleaved variant of CDCP1 in the cell line CTC-ITB001-MIND is stronger than the expression of the 70 kDa-variant in CTC-ITB001.

CTC-ITB001 shows a weak expression of Neuropilin-1 whereas the corresponding mouse xenograft cell line exhibits a stronger expression of NRP-1. On the contrary, NRP-2 as well as DCBLD2 is not detectable in both CTC cell lines by Western Blot analysis.

5.6. Potential biomarker candidate expression on PBMC of healthy donors

Considering the biomarker candidates as potential marker for CTC/DTC detection of breast cancer patients it is necessary to prove that the proteins CDCP1, NRP-1, NRP-2 and DCBLD2 are not expressed on peripheral blood mononuclear cells (PBMC) of healthy female donors. That is very important to guarantee that the marker candidates are only expressed on tumor cells. Otherwise the biomarker expression on PBMCs would lead to high rates of false positive non-tumor cells and the analysis would be ambiguous and time consuming. Therefore the expression of these proteins on PBMC of healthy female donors was analyzed by Western Blot (**figure 27 – 30**).

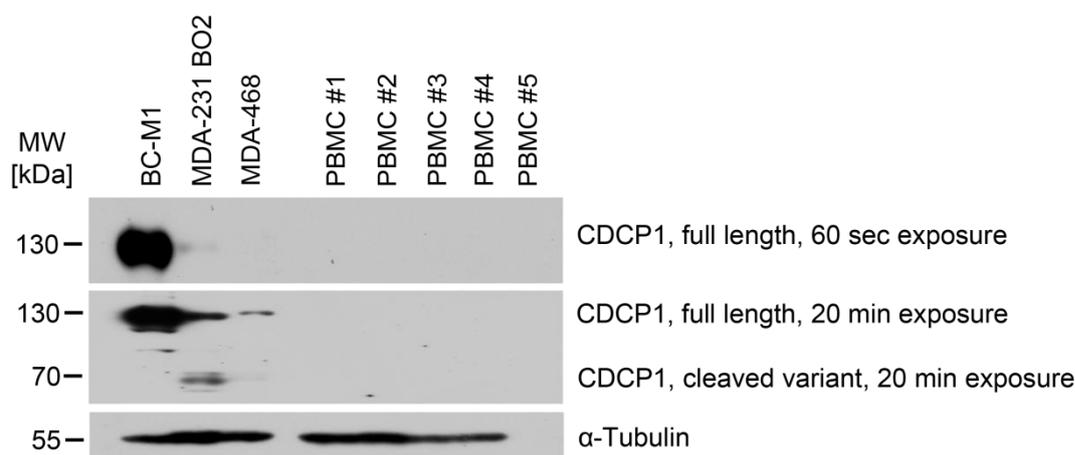


Figure 27: Western blot analysis of CDCP1 on peripheral blood mononuclear cells of five healthy women (PBMC#1-PBMC#5). BC-M1 and MDA-231 BO2 served as positive controls and α -Tubulin serves as a loading control.

Both, the full length variant (130 kDa) and the cleaved membrane fragment (70 kDa) of CDCP1 are not detectable in peripheral blood mononuclear cells (PBMC #1 to PBMC #5) of healthy female donors, even after long exposure times of 20 minutes (**figure 27**). Whereas the full length variant produces a strong signal after 60 seconds in the cell line BC-M1 and a very weak signal in the cell line MDA-231 BO2. After 20 minutes exposure time the full length variant is detectable in MDA-231 BO2 and MDA-468 as well as the cleaved membrane variant in MDA-231 BO2.

The biomarker candidate Neuropilin-1 is not detectable in peripheral blood mononuclear cells (PBMC #1 to PBMC #5) of healthy female donors, even after longer exposure times of 15 minutes (**figure 28**). In contrary, NRP-1 is detectable after three minutes exposure time in the control cell lines MDA-231 BO2 and BC-M1.

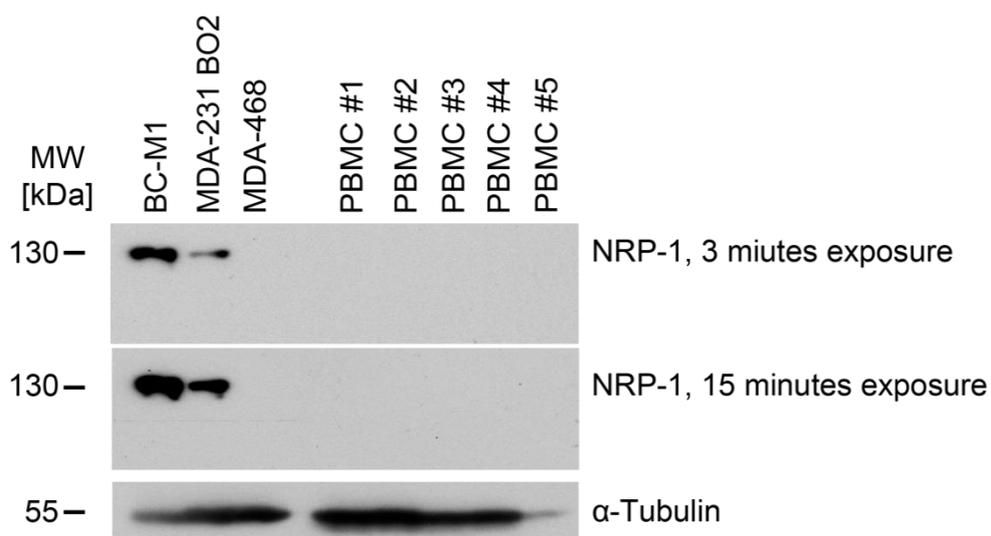


Figure 28: Western blot analysis of NRP1 on peripheral blood mononuclear cells of five healthy women (PBMC #1-PBMC #5). BC-M1 and MDA-231 BO2 served as positive controls and α -Tubulin serves as a loading control.

The biomarker candidate Neuropilin-2 is detectable with a weak signal between 100 kDa and 130 kDa in peripheral blood mononuclear cells (PBMC #1 to PBMC #5) of healthy female donors (**figure 29**). Additionally, a band of unknown origin with a strong signal at about 55 kDa is detectable in the samples PBMC #1 to PBMC #5. The bands in the PBMC samples are not detectable in the control cell lines BC-M1, MDA-231 BO2 and MDA-468.

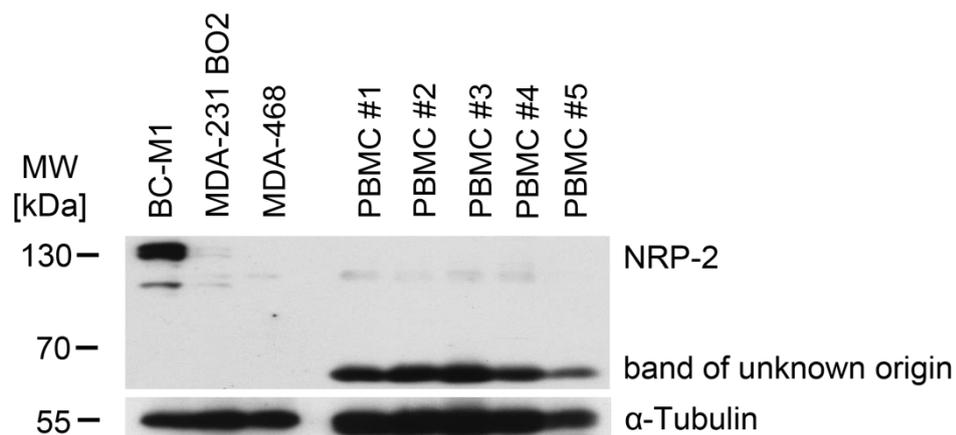


Figure 29: Western blot analysis of NRP2 on peripheral blood mononuclear cells of five healthy women (PBMC#1-PBMC#5). BC-M1 and MDA-231 BO2 served as positive controls and α -Tubulin serves as a loading control.

The potential marker protein Discoidin, CUB and LCCL domain-containing protein 2 (DCBLD2) is not detectable in peripheral blood mononuclear cells (PBMC #1 to PBMC #5) of healthy female donors, even after long exposure times of 20 minutes (**figure 30**). In contrary, the protein is detectable in the control cell lines BC-M1, MDA-231 BO2 and MDA-468 even after short exposure times of 60 seconds.

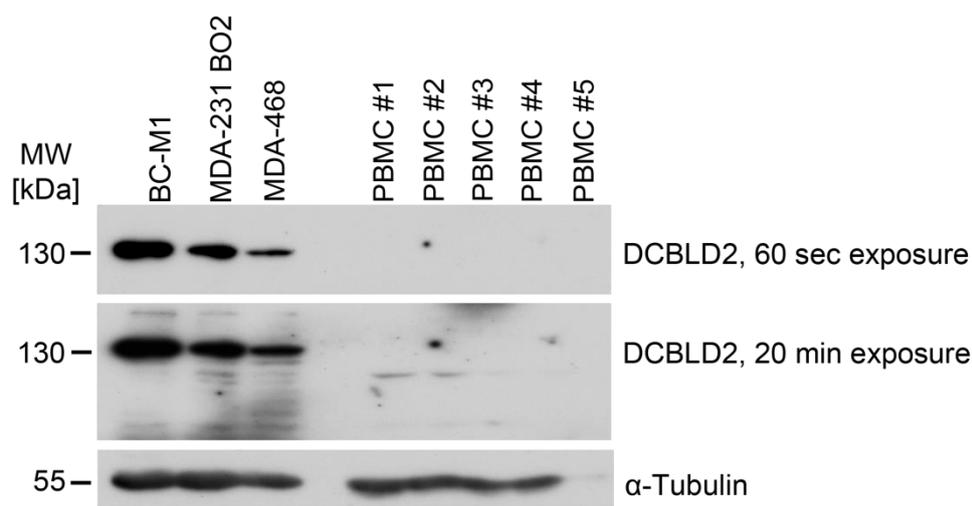


Figure 30: Western blot analysis of DCBLD2 on peripheral blood mononuclear cells of five healthy women (PBMC#1-PBMC#5). BC-M1 and MDA-231 BO2 serve as positive controls and α -Tubulin serves as a loading control.

Unlike tumor cells, there is no detectable expression of CDCP1 (**figure 27**), NRP-1 (**figure 28**) and DCBLD2 (**figure 30**) on peripheral blood mononuclear cells (PBMC#1 to PBMC#5) of healthy female donors in western blot analysis, even after longer exposure times (20 minutes for DCBLD2 and CDCP1, 15 minutes for NRP-1). However, NRP-2 is detectable in the peripheral blood mononuclear cells (PBMC#1 to PBMC#5) of healthy female donors with a strong signal at about 55 kDa and a weak signal between 100 kDa and 130 kDa (**figure 29**). These bands are of unknown origin and they are not detectable in the control cell lines BC-M1, MDA-231 BO2 and MDA-468.

The potential marker proteins CDCP1, NRP-1 and DCBLD2 fulfill all criteria for CTC/DTC detection markers at the moment and will be further investigated. The strong signals of unknown origin of NRP-2 in PBMC of healthy donors results in the exclusion of this marker candidate. NRP-2 will not be further investigated at this point. DCBLD2 will not be further investigated because of the lack of a reliable and executable antibody. The initially used antibody for DCBLD2 as well as the new batch did not produce reliable results and there were no other usable replacements available.

5.7. Potential biomarker candidate expression on erythrocytes of healthy donors

Considering the extracellular domain of the cell surface proteins CDCP1 and NRP-1 as a potential target for immune cells (e.g. CAR-T cells) or cytotoxic treatment (e.g. with nanoparticles) it is essential that red blood cells are negative for the extracellular domains of CDCP1 or NRP-1. That is very important to prevent attacks of immune cells or cytotoxic treatment against healthy non-tumor cells. Therefore the expression of NRP-1 (**figure 31**) and CDCP1 (**figure 32**) in red blood cells of healthy female donors was investigated by Western Blot analysis.

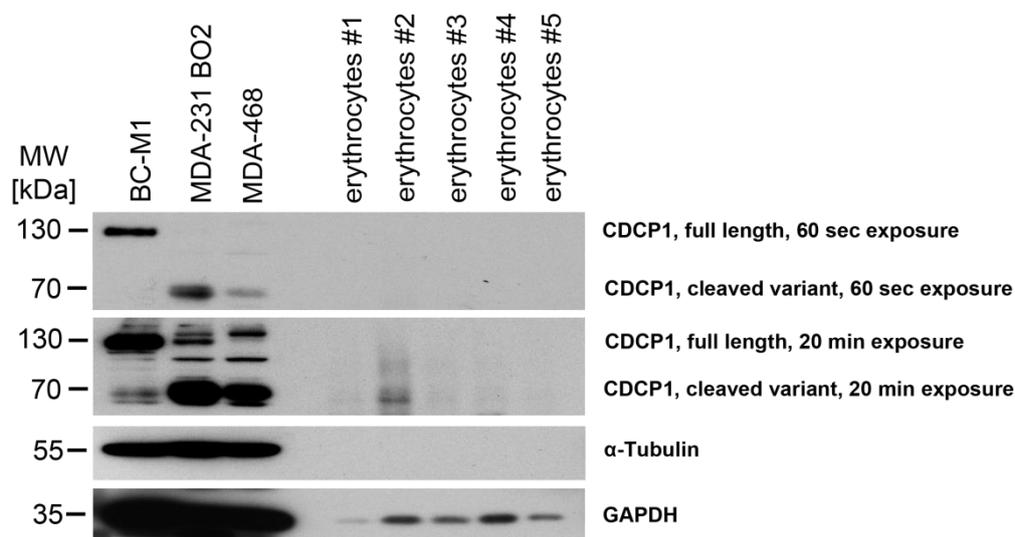


Figure 31: Western blot analysis of CDCP1 on erythrocytes of five healthy women (erythrocytes#1-erythrocytes#5). BC-M1 and MDA-231 BO2 serve as positive controls, α -Tubulin serves as a loading control for the three cell lines (BC-M1, MDA-231 BO2 and MDA-468) and GAPDH served as a loading control for the erythrocytes.

The full length variant of CDCP1 was not detectable on erythrocytes of healthy female donors (erythrocytes#1-erythrocytes#5), even after longer exposure times of 20 minutes (**figure 31**). In contrast, the full length variant was detectable in BC-M1 after an exposure time of 60 seconds and after 20 minutes exposure time in MDA-231 BO2 and MDA-468. However, there was a weak signal detectable at about 70 kDa in the sample erythrocytes #2 after an exposure time of 20 minutes. For an exposure time of 60 seconds there was no detectable signal in the erythrocytes for the cleaved variant of CDCP1. However, the

cleaved fragment was detectable after an exposure time of 60 seconds in the control cell lines MDA-231 BO2, MDA-468 and after 20 minutes exposure time in BC-M1.

NRP-1 was not detectable in the erythrocytes (erythrocytes#1-erythrocytes#5) of healthy female donors, even after longer exposure times of 20 minutes (**figure 32**). However, NRP-1 (130 kDa) was detectable with strong signals in the control cell lines BC-M1, MDA-231 BO2 and MDA-468 after exposure times of 60 seconds and 20 minutes. The signal at about 100 kDa in the control cell lines is caused by a reaction of unknown origin. This band is caused by an anti-NRP-1 antibody (CST, D62C6) which was later replaced by the anti-NRP-1 antibody (Abcam) because the Abcam-NRP-1 antibody did not generate a band of unknown origin. This experiment (**figure 32**) was not repeated with the Abcam-NRP-1 antibody.

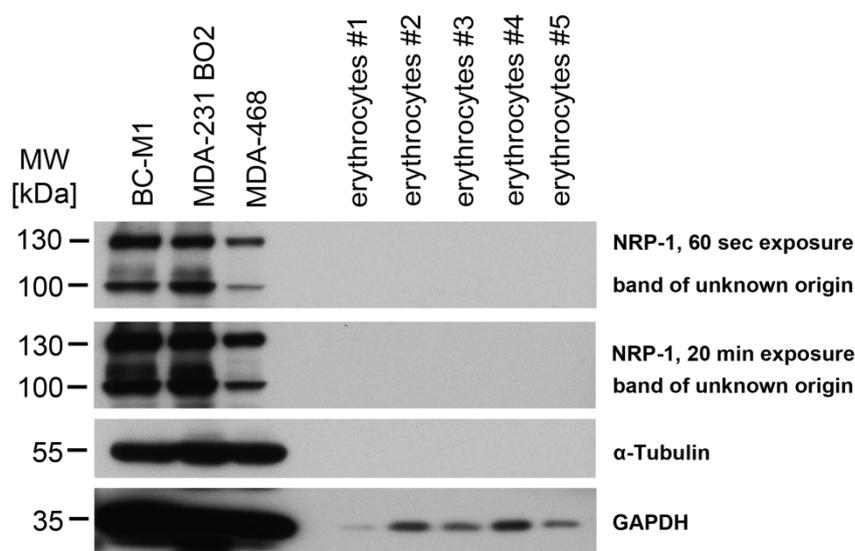


Figure 32: Western blot analysis of NRP1 on erythrocytes of five healthy women (erythrocytes#1-erythrocytes#5). BC-M1 and MDA-231 BO2 serve as positive controls, α -Tubulin serves as a loading control for the three cells lines and GAPDH serves as a loading control for the erythrocytes.

5.8. CDCP1 expression on cell lines from other tumor entities

To determine the presence of CDCP1 in other tumor entities Western Blot analysis for CDCP1 of pancreas-, prostate-, head and neck- and lung cancer cell lines was performed (**figure 33**).

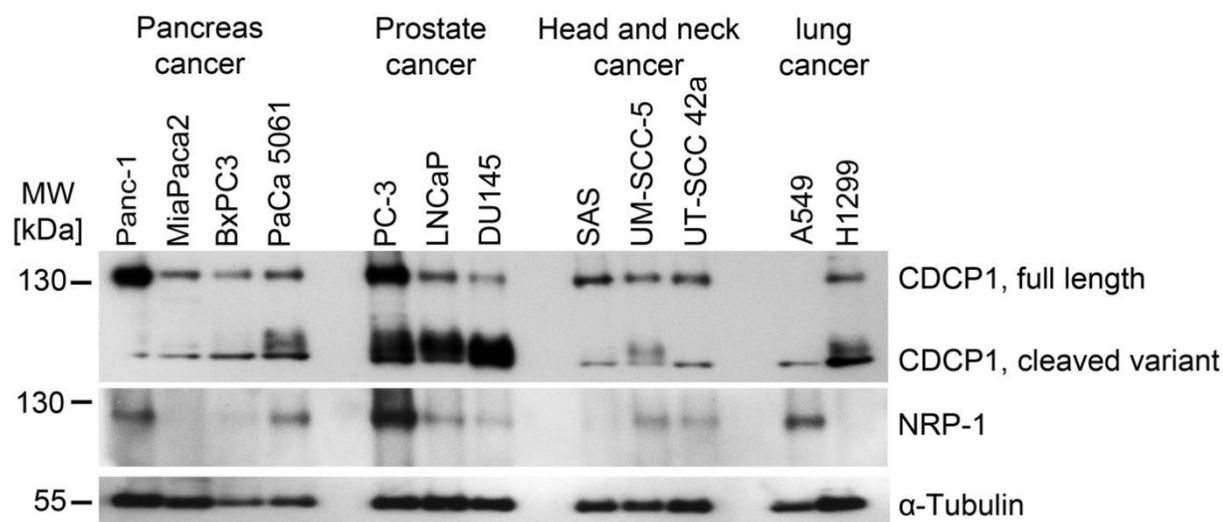


Figure 33: Western blot analysis of the cleaved and full length variant of CDCP1 on cell lines of different tumor entities. α -Tubulin served as a loading control.

Figure 33 shows that the expression of CDCP1 and NRP-1 is not restricted to breast cancer cell lines. CDCP1 is detectable in pancreas-, prostate-, head and neck- and lung- cancer. The expression level of the full length variant of CDCP1 exhibits a strong signal in the pancreatic cell line Panc1 as well as the prostate cancer cell line PC-3. All other cell lines have a moderate expression of the full length variant. The pancreatic cell line 5061 as well as the head and neck cancer cell line UT SCC 5 and the lung cancer cell line A549 have a strong signal for the cleaved variant of CDCP1. The prostate cancer cell lines PC-3, LNCaP and DU145 have the highest expression level for the cleaved CDCP1-fragment. All other cell lines exhibit a low expression level for the cleaved variant.

NRP-1 is detectable in the pancreas cancer cell lines Panc1 and 5061 with a moderate expression level whereas Panc2 and BxPC3 are negative for NRP-1. The highest expression level of NRP-1 is detectable in the prostate cancer cell line PC-3. LNCaP and DU145 have a weak expression. The head and neck cancer cell line SAS is negative for NRP-1 whereas UT SCC 5 and UT SCC 42a exhibit a weak expression level. The lung cancer cell line A549 is strongly positive for NRP-1 whereas the cell line H1299 is negative.

5.9. Immunofluorescence staining of CDCP1 in breast cancer and CTC cell lines

To determine the ability of an antibody for efficient isolation and specific diagnostic purposes of CTCs it has to be tested for specificity and sensitivity by immunofluorescent staining. Thereby it can be verified if the antibody is able to recognize the extracellular domain of endogenous CDCP1 on fixed tumor cells. Otherwise the antibody is not capable of recognizing CDCP1 on CTCs for isolation or detection methods. The immunofluorescent staining was established with the CDCP1⁺/CK⁺ cell line MDA-231 as a positive control and the CDCP1⁻/CK⁺ cell line MCF-7 as a negative control (**figure 5 and 24**). It is also necessary to verify the absence of CDCP1 expression on CD45⁺ PBMC by immunofluorescent staining which was already shown by Western Blot (**figure 27**). CDCP1 was analyzed by immunofluorescence staining for different breast cancer and CTC cell lines (**figure 34**).

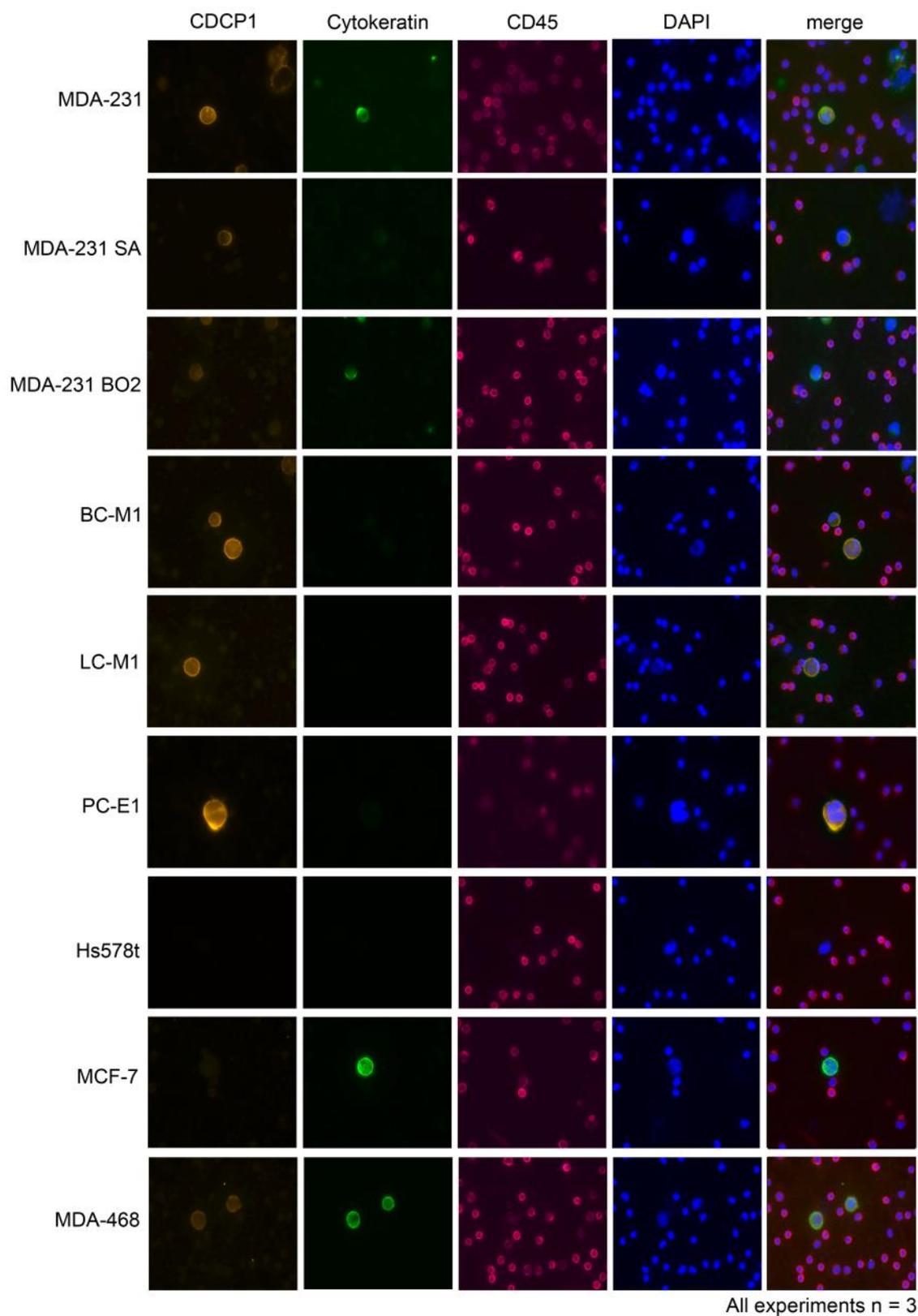


Figure 34: Immunofluorescence staining of CDCP1 in different cell lines and PBMC of healthy female donors. Staining for CDCP1 is shown in orange in the channel “CDCP1”, staining for pan-cytokeratin is shown in green in the channel “cytokeratin”, staining for the leukocyte exclusion marker CD45 is shown in magenta in the channel “CD45”, staining for the nucleus is shown in blue in the channel “DAPI” and the overlay of all before mentioned images is shown in the channel “merge”.

The DTC cell lines BC-M1, LC-M1 and PC-E1 exhibit the strongest immunofluorescent staining for CDCP1 and no detectable signal for cytokeratin. The cell line MDA-231 also exhibits a strong signal for CDCP1 and cytokeratin-pattern typical for this cell line. Cell line MDA-231 SA and MDA-231 BO2 have a weak signal for CDCP1 and very weak signals for cytokeratin. MDA-468 exhibits a weak staining signal for CDCP1 and a strong signal for cytokeratin. In contrast, MCF-7 with a strong signal for cytokeratin has no detectable signals for CDCP1 and Hs578t has neither detectable signals for CDCP1 nor signals for cytokeratin. These results are comparable to Western Blot analysis (**figure 5 and 24**). It is shown that CD45-positive PBMCs of healthy female donors are negative for CDCP1-immunofluorescence staining as shown in western blot analysis of PBMC (**figure 27**). The signal intensity for CDCP1⁺/CK^{-/weak} tumor cell lines (MDA-231 MDA-231 SA, MDA-231 BO2, BC-M1, LC-M1, PC-E1) is high enough to classify these cells as CDCP1-positive. Even the CK⁺/CDCP1^{weak} cell line MDA-468 is confirmed as CDCP1-weakly positive like shown in Western blot analysis (**figure 24**). All of the CDCP1⁺ tumor cell lines can be easily distinguished from CDCP1⁻ cell lines (MCF-7 and Hs578t). The used antibody for CDCP1 immunofluorescence staining is not able to discriminate between the full length variant and the cleaved membrane-anchored variant because this antibody was generated by immunization with almost the whole extracellular domain (Phe30 – Leu666). Due to the polyclonality the antibody recognizes full length CDCP1 as well as membrane-bound cleaved CDCP1.

Additionally, CDCP1 was analyzed by immunofluorescence staining in the CDCP1⁺/CK⁺ breast cancer cell line MDA-231 grown in chamber slides. To gain information about the distribution of CDCP1 expression regarding the natural morphology, MDA-231 was cultivated in chamber slides and stained for CDCP1 (**figure 35**).

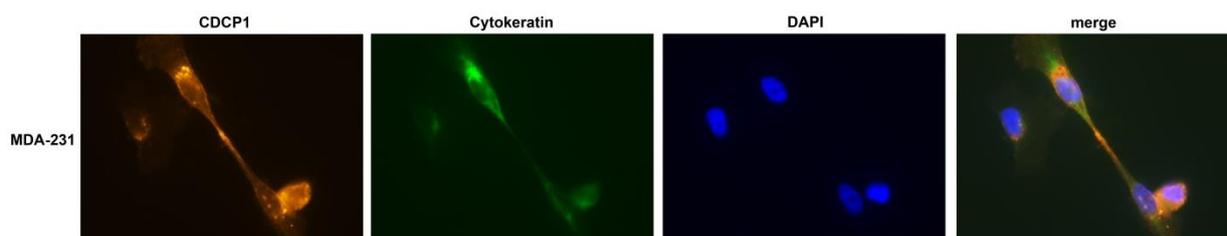


Figure 35: Chamber slide CDCP1-immunofluorescence staining of MDA-231. CDCP1 is shown in orange in the channel “CDCP1”, staining for pan-cytokeratin (AE1/AE3) is shown in green in the channel “cytokeratin”, staining for the nucleus is shown in blue in the channel “DAPI” and the overlay of all before mentioned images is shown in the channel “merge”.

Figure 35 shows four MDA-231 cells. The staining for CDCP1 is visible as a membrane staining and as speckles near the nucleus. The highest intensity of the CDCP1-expression is located in the membrane filaments at the point of contact to the connected cells. The moderate expression of cytokeratins is located near the nucleus and exhibits the typical cytokeratin staining pattern of MDA-231.

5.10. Immunofluorescence staining of Neuropilin-1 in breast cancer and DTC cell lines

To determine the ability of an antibody for efficient isolation and specific diagnostic of CTCs it has to be tested for specificity and sensitivity by immunofluorescent staining. Thereby it can be verified if the antibody is able to recognize the endogenous NRP-1 on fixed tumor cells. Otherwise the antibody is not capable of recognizing NRP-1 on CTCs for isolation or detection methods. The immunofluorescent staining was established with the NRP-1⁺/CK⁻ cell line BC-M1 as a positive control and the CDCP1⁻/CK⁺ cell line MCF-7 as a negative control (**figure 5 and 24**). It is also necessary to verify the absence of NRP-1 expression on CD45⁺ PBMC by immunofluorescence which was already shown by Western Blot (**figure 28**). NRP-1 was analyzed by immunofluorescent staining for BC-M1 and MCF-7 (**figure 36**).

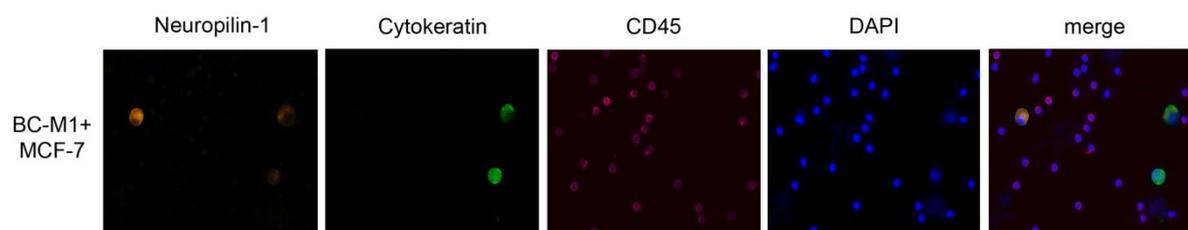


Figure 36: Immunofluorescent staining of NRP-1 for the DTC cell line BC-M1, breast cancer cell line MCF-7 and PBMC of healthy female donors. Staining for NRP-1 is shown in orange in the channel “Neuropilin-1”, staining for pan-cytokeratin is shown in green in the channel “cytokeratin”, staining for the leukocyte exclusion marker CD45 is shown in magenta in the channel “CD45”, staining for the nucleus is shown in blue in the channel “DAPI” and the overlay of all before mentioned images is shown in the channel “merge”. BC-M1 serves as a NRP-1 positive control cell line whereas MCF-7 serves as NRP-1 negative control cell line.

The DTC cell line BC-M1 exhibits a strong expression of NRP-1 which is detectable as membrane staining whereas NRP-1 is detectable in MCF-7 with very weak signals. This correlates with the Western Blot analysis (**figure 24**). Additionally, cytokeratin is not detectable in BC-M1 whereas MCF-7 exhibits a strong cytokeratin expression. These results correlate with Western Blot analysis (**figure 5 and 24**). For PBMC of healthy female donors there is no expression of NRP-1 detectable which correlates with the Western blot analysis (**figure 28**). The NRP-1 immunofluorescence staining shows that the anti-NRP-1 antibody is able to recognize endogenous NRP-1 on tumor cells. The potential biomarker NRP-1 is able to distinguish between breast cancer tumor cells and PBMC of healthy donors.

Figure 37 shows a MDA-231 cell in the middle of the image with a NRP-1 membrane staining and small speckles of NRP-1 located in the cytoplasm after cultivation of the cells in chamber slides. The highest intensity of the NRP-1-expression is located in the membrane area. The moderate expression of cytokeratins is located near the nucleus and exhibits the typical cytokeratin staining pattern of MDA-231.

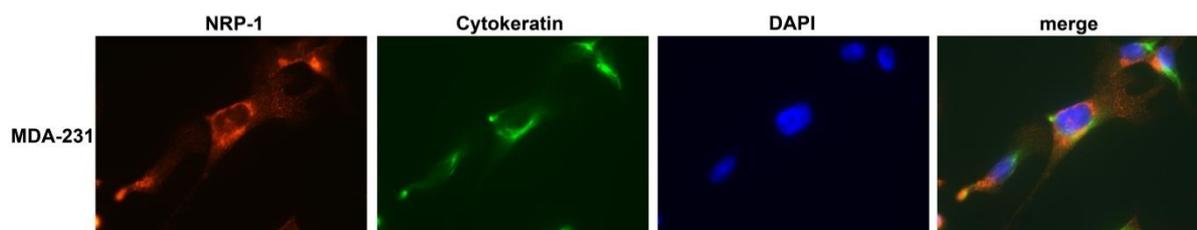


Figure 37: Chamber slide NRP-1-immunofluorescence staining of MDA-231. NRP-1 is shown in orange in the channel “NRP-1”, staining for pan-cytokeratin is shown in green in the channel “cytokeratin”, staining for the nucleus is shown in blue in the channel “DAPI” and the overlay of all before mentioned images is shown in the channel “merge”.

5.11. Single cell isolation from tumor cell lines via CDCP1-MACS-system

The commonly used EpCam-dependent isolation of tumor cells provides the enrichment of CTC with epithelial properties. To increase the probability of the enrichment of CTC with a mesenchymal or hybrid epithelial/mesenchymal phenotype, CDCP1-dependent MACS-isolation will be performed. Therefore, Streptavidin-conjugated magnetic beads are conjugated with the CDCP1-antibody which has to be biotinylated for this purpose. To test the suitability of the biotinylated anti-CDCP1 antibody (Novus Biologicals, BAF2666) with the MACS isolation method this antibody was previously validated by FACS. This step was necessary to confirm the ability of this antibody to bind the extracellular domain of CDCP1 on viable and unfixed cells. For this purpose, FACS experiments were performed with unconjugated anti-CDCP1 antibody (R&D AF2666) to simulate the secondary reaction with the streptavidin-conjugated magnetic beads in the MACS approach (**figure 38**).

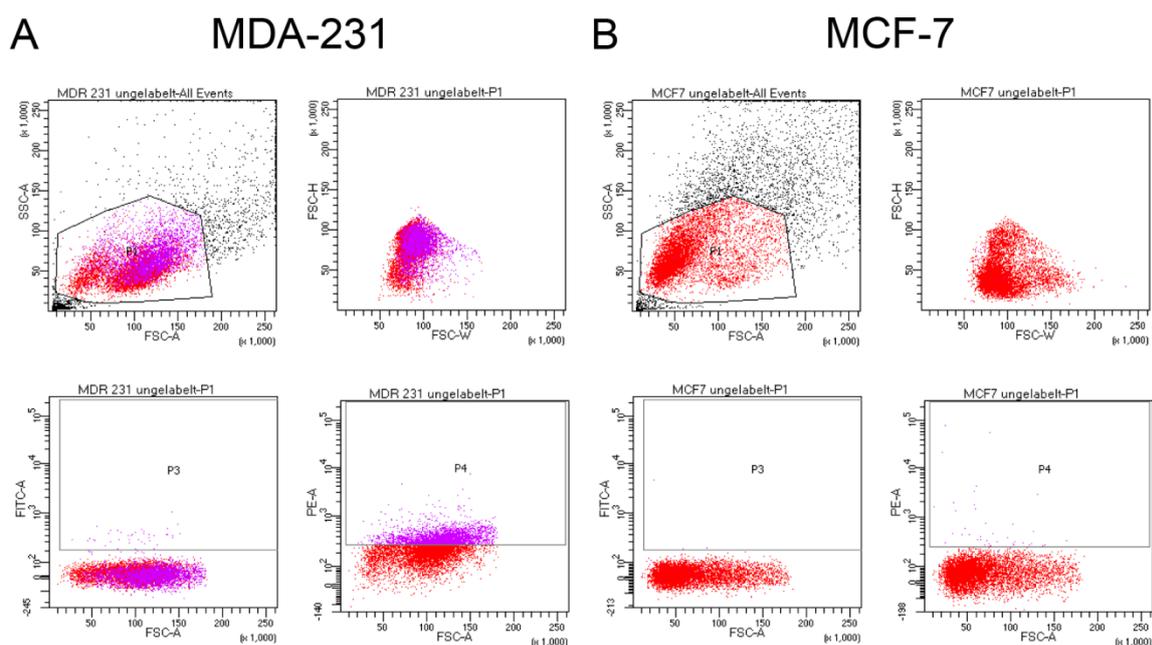


Figure 38: FACS analysis of breast cancer cell lines MDA-231 and MCF-7 with unconjugated anti-human-CDCP1 antibody (R&D AF2666). (A) MDA-231 was used as a positive control. (B) MCF-7 was used as a negative control. The x-axis in the plot “P1” represents the forward light scattering (FCS-A) whereas the y-axis represents the side-scattered light (SSC-A). FCS correlates with the size or the volume of the cells [148] and SSC gives insights into the internal structures [149]. The plot “P2” is used for the exclusion of cell-doublings but was not used for these experiments. The graph “P3” shows the correlation between forward scattered light (FSC-A) and the fluorophore FITC. This fluorophore was not used but served as internal negative control. The plot “P4” shows the correlation between the forward scattered light (FCS-A) and the fluorophore PE. The PE-conjugated secondary antibody was used to recognize the unconjugated CDCP1-antibody.

For MDA-231 (**figure 38, A**) 8.095 out of 10.000 cell events (**A, P1, top left**) were further analyzed and 3.056 (37.8 %) MDA-231 cells were detected as CDCP1- positive (**A, P4, bottom right**). For MCF-7 (**figure 38, B**) 6.765 out of 10.000 cell events (**B, P1, top left**) were further analysed and 37 (0.5 %) MCF-7 cells were detected as CDCP1-positive (**B, P4, bottom right**). These results correlate with the results for the above mentioned cell lines for immunofluorescence staining (**figure 34**) and Western Blot analysis (**figure 24**). These results confirm that this anti-CDCP1 antibody is able to specifically recognize viable and CDCP1-positive tumor cells. Due to these facts MACS-based tumor cell isolation can be performed with this antibody.

A MACS-based isolation method using the cell surface protein CDCP1 was first established with the CDCP1-positive and cytokeratin-positive breast cancer tumor cell line MDA-231. Different parameters were tested to generate an isolation method for CDCP1-positive CTCs/DTCs of cancer patients with a suitable yield. The results of the optimal protocol are shown here:

The average recovery rate of 50 MDA-231 cells spiked into a PBMC fraction of a healthy donor was 55 %, which means that approximately 27 out of 50 cells could be positively selected and isolated with the CDCP1-MACS-based isolation method (**figure 39**).

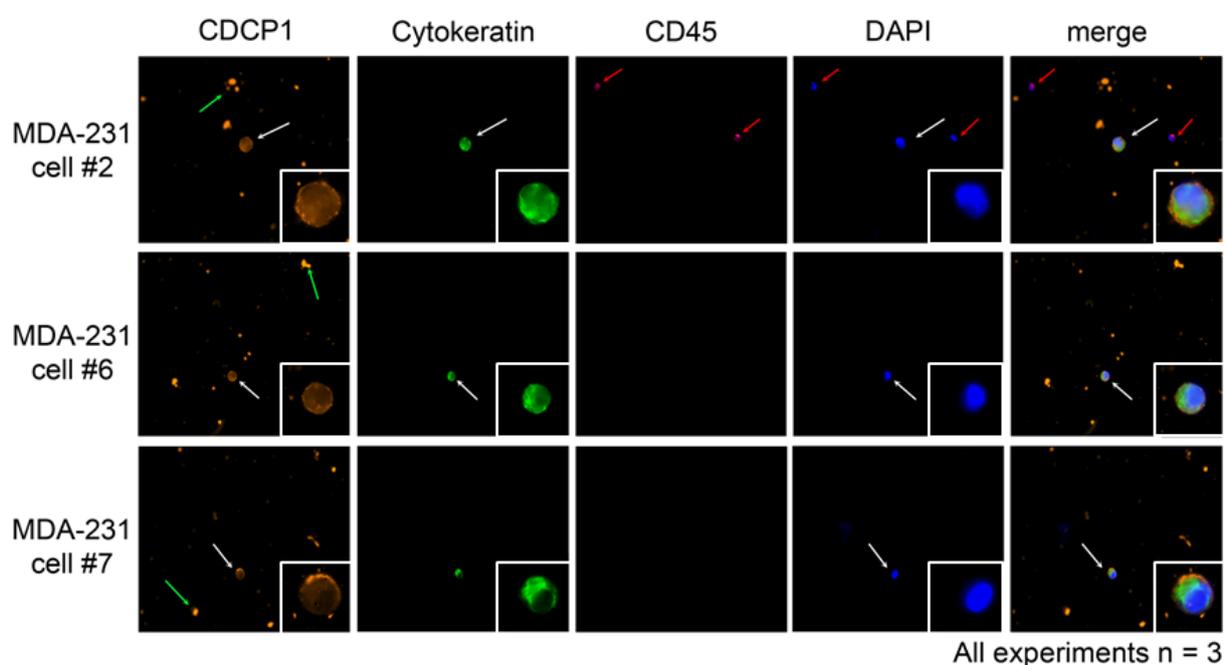


Figure 39: This figure shows three selected MDA-231 cells enriched by positive selection with a MACS-based isolation approach using the cell surface protein CDCP1. The channel “CDCP1” shows the detection of the CDCP1-capturing antibody by a Fluorophor-conjugated secondary antibody, channel “cytokeratin” represents the cytokeratin staining, channel “CD45” shows the leukocyte exclusion marker CD45, DAPI staining for nucleic acids is shown in the channel “DAPI” and the merge image combines an overlay of all channels. CDCP1⁺/CK⁺/CD45⁻/DAPI⁺ MDA-231 cells are indicated by white arrows and a magnification is shown in the white square. CDCP1⁻/CK⁻/CD45⁺/DAPI⁺ PBMCs are marked by red arrows. Green arrows highlight CK⁻/CD45⁻/DAPI⁻ unspecific CDCP1-signals.

The MDA-231 cell line was chosen as a reference cell line to determine the parameters to establish a system which is able to isolate CDCP1-positive CTCs/DTCs from cancer patients. The staining pattern of the CDCP1-capturing antibody correlates with the CDCP1-staining for the tumor cell lines (**figure 34**) and represents the typical membrane staining-pattern of a predominantly expressed cell surface protein. The cytokeratin staining shows the typical cytokeratin pattern typical for MDA-231 similar to the staining-pattern for the analyzed tumor cell lines (**figure 34**). CDCP1/cytokeratin-positive MDA-231 cells are

negative for CD45 and the number of unspecific selected CD45-positive cells is very low (See cell 6/7 in **figure 39**). Except of the image “CD45/cell 2” there are two CD45-positive but CPDP1-and cytokeratin-negative cells which were probably isolated by unspecific binding to the column and not by positive selection via CDCP1-capturing. The CK-/DAPI speckles in the channel “CDCP1” appearing not circularly and stronger/brighter than the represented MDA-231 cells seem to be caused by unspecific binding of biotinylated antibody or by binding of Streptavidin-conjugated beads to the surface of the microscope slid. The appearance of these signals in a different layer than the cells could be an indication for the before mentioned explanation. The number of MCF-7 cells enriched by positive CDCP1 selection was below four cells out of 50 spiked cells. This corresponds with an unspecific isolation rate of CDCP1-negative cells below 8%. These results show that this method is suitable for the isolation of tumor cells with MDA-231-like phenotype out of the blood of cancer patients.

5.12. Detection of CDCP1 on breast cancer tumor cell lines

For the possibility of detecting CDCP1 on circulating tumor cells (CTCs) or disseminating tumor cells (DTCs) of breast cancer patients with the CellSearch system it is necessary to validate and test the applied antibody for specificity and sensitivity. For this purposes the Alexa488-conjugated anti-human-CDCP1 antibody (NBP2-60224AF488) was investigated in fluorescence-activated cell sorting (FACS) (**figure 40**).

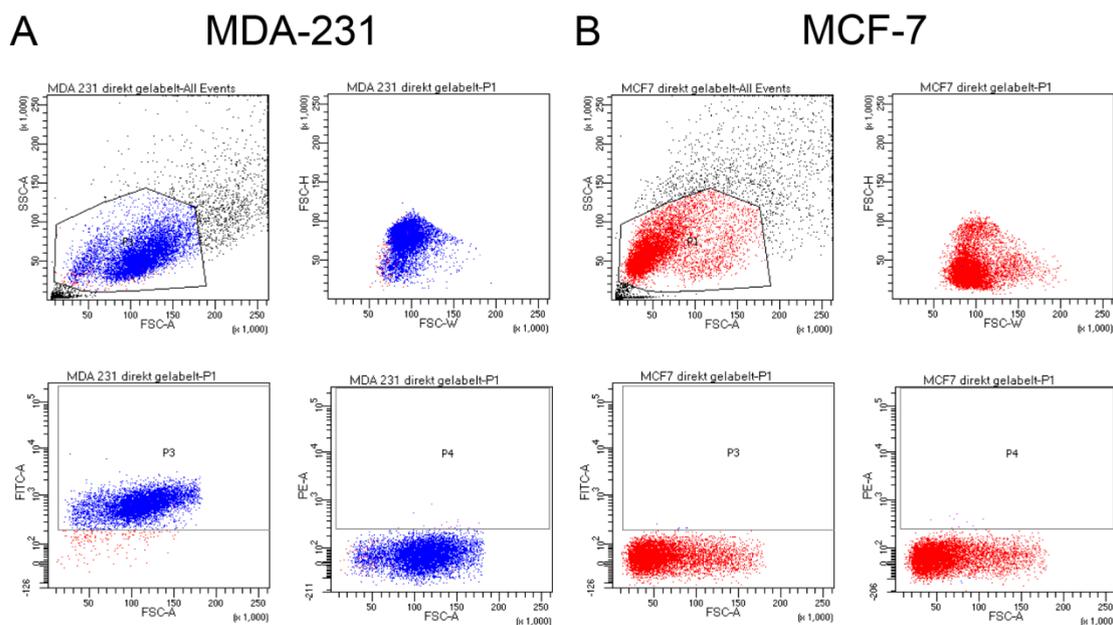


Figure 40: FACS analysis of breast cancer cell lines with Alexa488-conjugated anti-human-CDCP1 antibody (NBP2-60224AF488). (A) MDA-231 was used as a positive control. (B) MCF-7 was used as a negative control. The x-axis in the plot “P1” represents the forward light scattering (FCS-A) whereas the y-axis represents the side-scattered light (SSC-A). FCS correlates with the size or the volume of the cells [148] and SSC gives insights into the internal structures [149]. The plot “P2” is used for the exclusion of cell-doublets but was not used for these experiments. The graph “P3” shows the correlation between forward scattered light (FSC-A) and the fluorophore FITC. The FITC-conjugated CDCP1-antibody was used to recognize CDCP1 on the cell surface of the analyzed cells. The plot “P4” shows the correlation between the forward scattered light (FCS-A) and the fluorophore PE but this fluorophore was not considered in this experiment and served as internal negative control.

For MDA-231 (**figure 40, A**) 7.485 out of 10.000 cell events (**A, P1, top left**) were further analysed and 7.300 (97.5%) MDA-231 cells were detected as CDCP-1 positive (**A, P3, bottom left**). For MCF-7 (**figure 40, B**) 7.474 out of 10.000 cell events (**B, P1, top left**) were further analysed and 5 (0.07%) MCF-7 cells were detected as CDCP1-positive (**B, P3, bottom left**). These results correlate with the results for the above mentioned cell lines for immunofluorescence staining (**figure 34**) and Western Blot analysis (**figure 24**). These results confirm that the anti-CDCP1 antibody is able to specifically recognize viable and CDCP1-positive tumor cells. Due to these facts CTC analysis (CellSearch System) can be performed with this antibody.

Furthermore, the parameters for the use in the CellSearch System of the Alexa488-conjugated anti-human-CDCP1 antibody (NBP2-60224AF488) had to be determined. The parameters were identified by the analysis of a CDCP1 strong positive cell line (MDA-231), a CDCP1 weak positive cell line (MDA-468) and a CDCP1 negative cell line (MCF-7). The results after establishment of the suitable parameters are shown in **figure 41**.

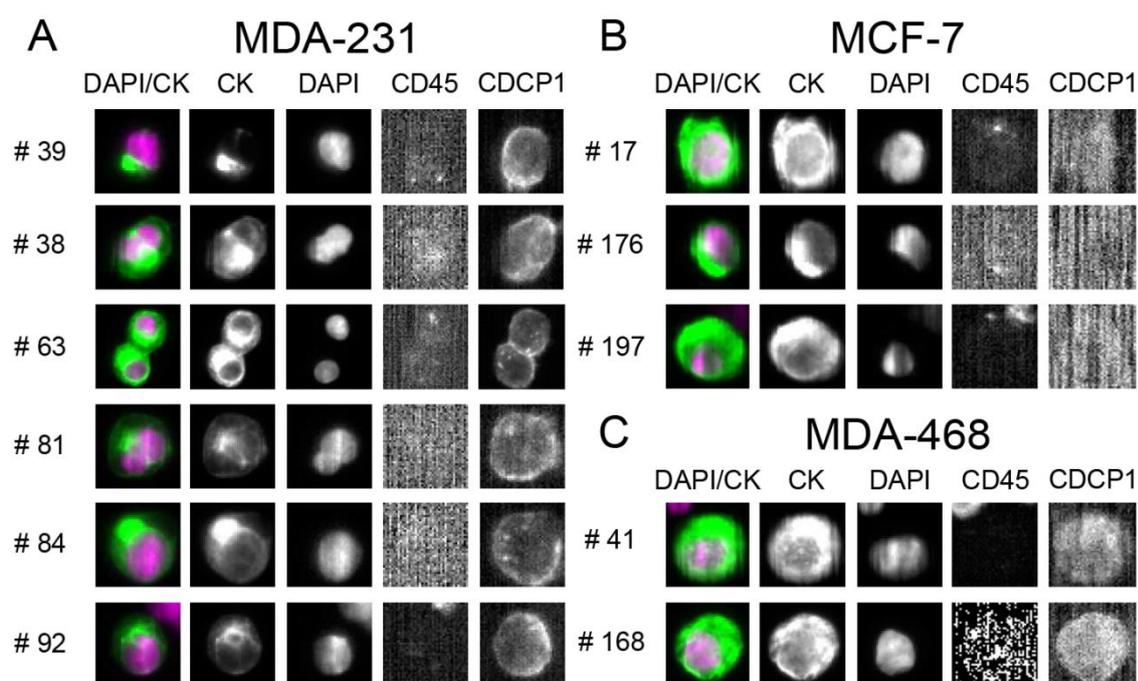


Figure 41: CellSearch analysis of breast cancer cell lines with Alexa488-conjugated anti-human-CDCP1 antibody (NBP2-60224AF488). cytokeratin staining is shown in the channel “CK”, staining of the nucleus is shown in the channel “DAPI”, staining of CD45 is shown in the channel “CD45”, CDCP1 staining is shown in the channel “CDCP1” and overlay images of the channels “CK” and “DAPI” are shown in the channel “DAPI/CK”. (A) Six chosen images of cell line MDA-231 are shown. (B) Three chosen images of cell line MCF-7 are shown. (C) Two chosen images of cell line MDA-468 are shown.

As a positive control, every of the shown MDA-231 cell (**figure 41 A**) exhibits a nucleus (shown in the channel “DAPI”) and a cytokeratin-pattern (shown in the channel “CK”) specific for the cell line MDA-231. The staining for the leukocyte exclusion marker CD45 is negative for this cell line, while the staining for the membrane protein CDCP1 is positive and the highest fluorescence intensity originates from areas of the cell membrane as already shown for the immunofluorescence staining (**figure 34**). The results for the cell line MDA-231 show that this antibody is able to recognize the extracellular domain of CDCP1. This leads to a membrane staining with a typical stronger signal around the cell and a weaker cytoplasmic staining.

As a negative control, every of the shown MCF-7 cells (**figure 41 B**) exhibit a nucleus (shown in the channel “DAPI”) and a cytokeratin-pattern (shown in the channel “CK”) specific for the cell line MCF-7. The staining for the leukocyte exclusion marker CD45 and the membrane protein CDCP1 are negative for this cell line. These results correlate with immunofluorescence staining (**figure 34**) and Western Blot analysis (**figure 24**). The MDA-468 cell line was chosen because of a moderate expression of CDCP1 (**figure 24 and 34**). Every of the shown MDA-468 cells (**figure 41 C**) exhibit a nucleus (shown in the channel “DAPI”) and a cytokeratin-pattern (shown in the channel “CK”) specific for the cell line MDA-468. The staining for the leukocyte exclusion marker CD45 is negative for this cell line, while the staining for the membrane protein CDCP1 is positive but weaker as the CDCP1-staining for MDA-231(**figure 7**). The intensity of the CDCP1-staining correlates with immunofluorescence staining (**figure 34**) and Western Blot analysis (**figure 24**).

5.13. Detection of CDCP1 on circulating tumor cells (CTCs) of breast cancer patients

The determined optimal parameters were used for the analysis of blood of breast cancer patients. The blood samples (7.5 mL) of metastatic breast cancer patients were analyzed for the expression of the cell surface protein CDCP1 by the CellSearch system (**figure 42**).

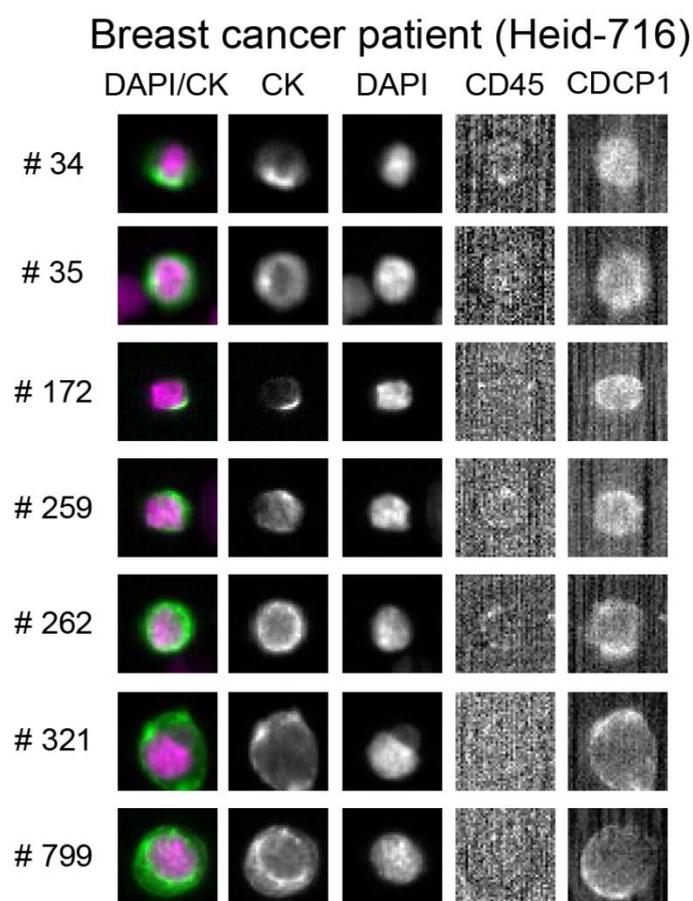


Figure 42: CellSearch-CTC-analysis of a breast cancer patient for CDCP1. Cytokeratin staining is shown in the channel “CK”, staining of the nucleus is shown in the channel “DAPI”, staining of CD45 is shown in the channel “CD45”, CDCP1 staining is shown in the channel “CDCP1” and overlay images of the channels “CK” and “DAPI” are shown in the channel “DAPI/CK”.

In the breast cancer patient (Heid-716) shown in **figure 42**, 1460 tumor cells were determined as intact CK⁺/CD45⁻ CTCs after EpCAM-dependent enrichment with the CellSearch system. 18 of 1460 CTCs (1.2 %) were detected as CDCP1⁺/CK⁺/CD45⁻ CTCs. Two of the demonstrated cells (**figure 42**, #321, #799) have a membrane staining pattern for CDCP1 comparable with MDA-231 (**figure 41 A**). The signal intensity of the large number of analyzed CDCP1⁺/CK⁺ CTCs from this breast cancer patient is

comparable with the signal intensity of CDCP1 in the cell line MDA-468 (**figure 41 C**) analyzed by immunofluorescent staining. MDA-468 exhibits a weak expression of CDCP1. **Table 3** shows an overview of the determination of CDCP1 in all analyzed breast cancer patients. In total, 41 breast cancer patients were analyzed. Seven (17 %) of the analyzed 41 patients contained CDCP1⁺/CK⁺/CD45⁻ CTCs whereas 17 (41 %) of 41 analyzed patients contained CK⁺/CD45⁻ CTCs. The average percentage of CDCP1⁺/CK⁺/CD45⁻ CTCs in the population of CK⁺/CD45⁻ CTCs was at about 4 %. The low probability for the isolation of CDCP1⁺/CK⁻ CTCs with the CellSearch by the fact that EpCam-negative cells are usually negative for cytokeratins as shown for Western Blot analysis (**figure 5**, exception: MDA-231).

Table 3: Overview of the breast cancer patients analyzed for CDCP1 with the CellSearch System.

sample	date	CTC	CDCP1	comment
Heid 716	31.07.2018	1890 CTCs	18 CDCP1+	1460 intact CTCs, 430 apoptotic
Heid 717	02.08.2018	negativ	Negativ	xxxxxx
Heid 718	06.08.2018	3 CTC	Negativ	2 relativ intact CTC, 1 apoptotic CTC
Heid 719	06.08.2018	1 CTC	Negativ	unclear (CD45+)
Heid 720	14.08.2018	negativ	Negativ	xxxxxx
Heid 767	14.01.2019	8 CTC	Negativ	4 relativ intact CTC, 4 apoptotic CTC
Heid 768	16.01.2019	negativ	Negativ	xxxxxx
Heid 769	16.01.2019	1 CTC	Negativ	1 apoptotic CTC
Heid 770	21.01.2019	8 CTC	Negativ	xxxxxx
Heid 771	23.01.2019	negativ	Negativ	xxxxxx
Heid 776	07.02.2019	negativ	Negativ	xxxxxx
Heid 777	12.02.2019	negativ	Negativ	xxxxxx
Heid 778	19.02.2019	49 CTC	Negativ	3 relativ intact CTC, 46 apoptotic CTC
Heid 779	19.02.2019	xxxx	Xxxx	CellSearch run aborted
Heid 780	21.02.2019	xxxx	Xxxx	CellSearch run aborted
Heid 781	21.02.2019	850 CTC	13 very weak CDCP1+	xxxxxx
Heid 782	27.02.2019	284 CTC	Negativ	CellSearch run aborted
Heid 784	28.02.2019	1 CTC	Negativ	1 apoptotic CTC
Heid 785	04.03.2019	14 CTC	Negativ	2 relativ intact CTC, 12 apoptotic CTC
Heid 786	06.03.2019	587 CTC	ca 10% very weak CDCP1+	450 relativ intact CTC, 140 apoptotic CTC
Heid 787	06.03.2019	1 CTC	Negativ	xxxxxx
Heid 788	06.03.2019	negativ	Negativ	xxxxxx
Heid 789	08.03.2019	1 CTC	Negativ	1 CTC apoptotic
Heid 101	08.03.2019	542 CTC	Negativ	434 relativ intact CTC, 108 apoptotic CTC
Heid L11/172	08.03.2019	368 CTC	12 very weak CDCP1+	250 intact CTC, 120 apoptotic CTC
Heid 790	13.03.2019	20 CTC	2 CDCP1+	4 relativ intact CTC, 16 apoptotic CTC
Heid 791	15.03.2019	negativ	Negativ	xxxxxx
Heid 792	15.03.2019	2 CTC	Negativ	1 relativ intact CTC, 1 apoptotic CTC
Heid 793	21.03.2019	negativ	Negativ	xxxxxx
Heid 794	25.03.2019	2 CTC	Negativ	apoptotic
Heid 795	25.03.2019	negativ	negativ	xxxxxx
Heid 796	26.03.2019	168 CTC	a few very weak CDCP1+	25 relativ intact CTC, 143 apoptotic
Heid 797	26.03.2019	6 CTC	negativ	all apoptotic
Heid 799	02.04.2019	4 CTC	negativ	xxxxxx
Heid 800	05.04.2019	negativ	negativ	xxxxxx
Heid 801	08.04.2019	1 CTC	negativ	xxxxxx
Heid 802	08.04.2019	28 CTC	negativ	10 relativ intact, 18 apoptotic CTC
Heid 803	10.04.2019	1 CTC	negativ	xxxxxx
Heid 804	11.04.2019	2000 CTC	negativ	1700 relativ intact, 300 apoptotic CTC
Heid 805	11.04.2019	26 CTC	1 very weak CDCP1+	15 relativ intact, 11 apoptotic CTC
Heid 806	12.04.2019	negativ	negativ	xxxxxx

5.14. Detection of soluble CDCP1-variant in the blood plasma of breast cancer patients via ELISA

The proteolytic cleavage of the full length variant of CDCP1 (MW: 130 kDa) by serine proteases leads to the generation of a membrane-bound CDCP1 variant (MW: 70 kDa) and a soluble CDCP1-fragment (MW: 65 kDa). It is possible that the soluble fragment is released by tumor cells *in vivo* into the blood circulation. To analyze this, the soluble fragment of CDCP1 will be used for the generation of a blood plasma-based detection method (ELISA) for breast cancer patients.

As a first step, it is necessary to establish an internal control for the Sandwich-ELISA. Recombinant human-CDCP1 is used as a protein standard to convert the measured optical density into protein concentrations of CDCP1. The recombinant CDCP1 was tested with both Sandwich-ELISA antibodies by Western Blot analysis (**figure 43 and 44**) to determine specificity and sensitivity of the antibodies to recognize the recombinant protein.

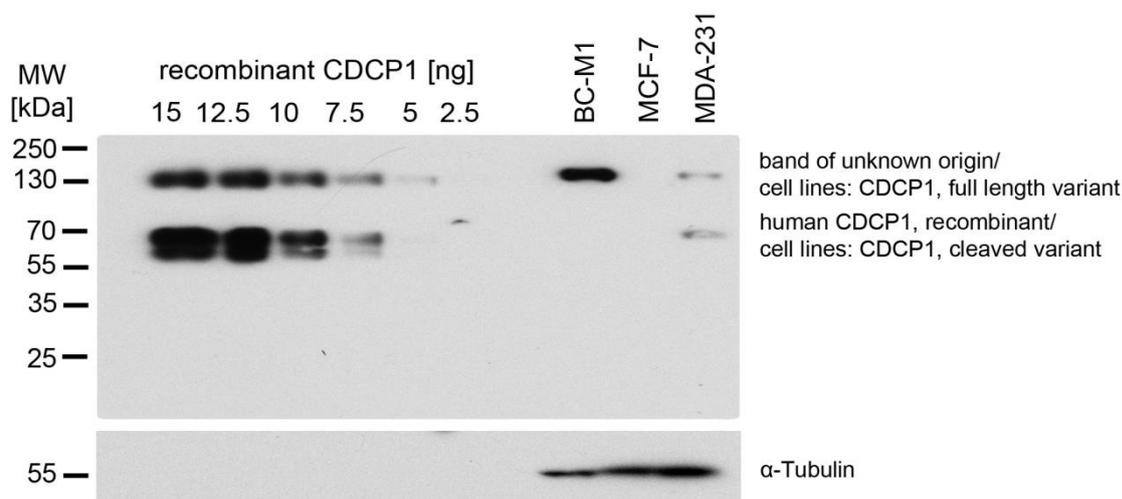


Figure 43: Western Blot analysis of recombinant human-CDCP1 (Novoprotein) with anti-human CDCP1 antibody (Thermo Fisher). Different amounts (15 ng, 12.5 ng, 10 ng, 7.5 ng, 5 ng and 2.5 ng) of recombinant CDCP1 were loaded on the gel. The cell lines (15 µg) were used as controls. BC-M1 was used as a positive control for the full length CDCP1, MCF-7 was used as a CDCP1-negative cell line and MDA-231 was used as positive control for the full length variant and the cleaved fragment of CDCP1. α-Tubulin served as a loading control.

Figure 43 shows the Western Blot analysis of recombinant human-CDCP1 with the Sandwich-ELISA-capturing antibody (epitope: Ala33-Leu666). The recombinant CDCP1 was applied in descending concentrations from 15 ng to 2.5 ng in 2.5 ng-steps. The recombinant protein (Novoprotein, ci80, Phe30-

Ser341) has a calculated molecular mass of 36 kDa without glycosylation. The expected molecular weight is around 60 kDa due to glycosylation. The CDCP1-capturing antibody recognizes recombinant CDCP1 at a mass below 70 kDa. The signal intensity decreases with descending concentration. Five ng of recombinant CDCP1 is detectable with a very weak signal. The lowest concentration (2.5 ng) is not detectable by Western Blot analysis. This antibody detects another signal at a size of 130 kDa for the recombinant protein. For the control cell lines, full length CDCP1 is detectable in BC-M1 and both variants are detectable in MDA-231. MCF-7 serves as a negative control and CDCP1 is not detectable.

Figure 44 shows the Western Blot analysis of recombinant human-CDCP1 with the Sandwich-ELISA-detection antibody. The recombinant CDCP1 (Novoprotein, ci80, Phe30-Ser341) was applied in descending concentrations from 15 ng to 2.5 ng in 2.5 ng-steps. The CDCP1-detection antibody recognizes recombinant CDCP1 (epitope: Phe30-Leu666) at a mass below 70 kDa. The signal intensity decreases with descending concentration. Five ng of recombinant CDCP1 is still detectable with a moderate signal intensity. The lowest concentration (2.5 ng) is not detectable by Western Blot analysis. For the control cell lines, full length CDCP1 is detectable in BC-M1 and both variants are detectable in MDA-231. MCF-7 serves as a negative control and CDCP1 is not detectable.

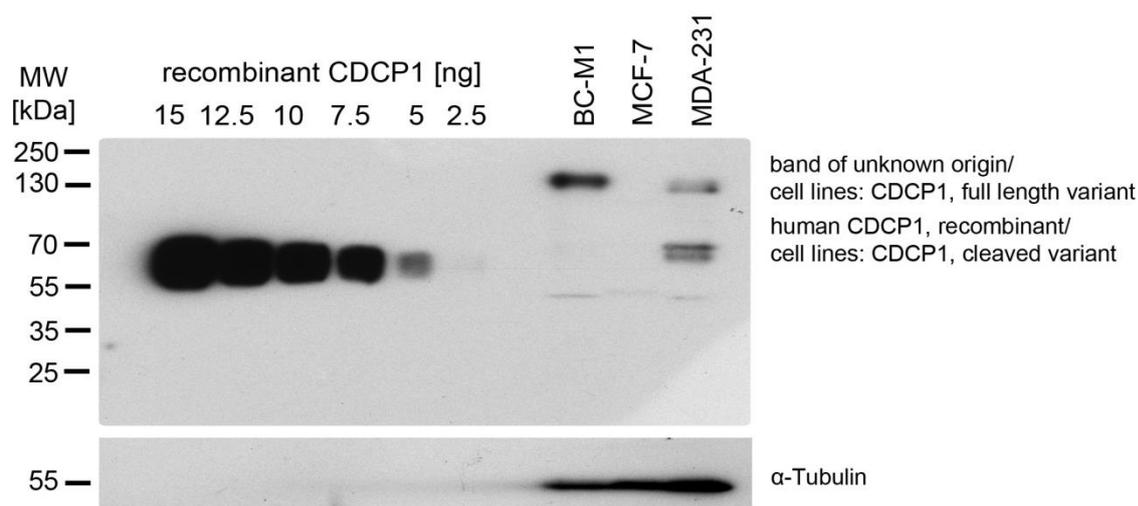


Figure 44: Western Blot analysis of recombinant human-CDCP1 (Novoprotein) with anti-human CDCP1 antibody (R&D). Different amounts (15 ng, 12.5 ng, 10 ng, 7.5 ng, 5 ng and 2.5 ng) of recombinant CDCP1 were loaded on the gel. The cell lines (30 μ g) were used as controls. BC-M1 was used as a positive control for the full length CDCP1, MCF was used as a CDCP1-negative cell line and MDA-231 was used as positive control for the full length variant and the cleaved fragment of CDCP1. α -Tubulin served as a loading control.

The cell culture supernatant of tumor cell lines was used to establish and validate an ELISA (**figure 45**) which is able to determine the concentration of soluble CDCP1 in the blood plasma of breast cancer patients. After determination of the different parameters (concentration, blocking reagent and incubation times) anti-human CDCP1 antibody (Thermo Fisher, dilution 1:150) was used as the coating/capturing antibody. The anti-human CDCP1 antibody (R&D systems, dilution 1:150) was used as the detection antibody (second primary antibody).

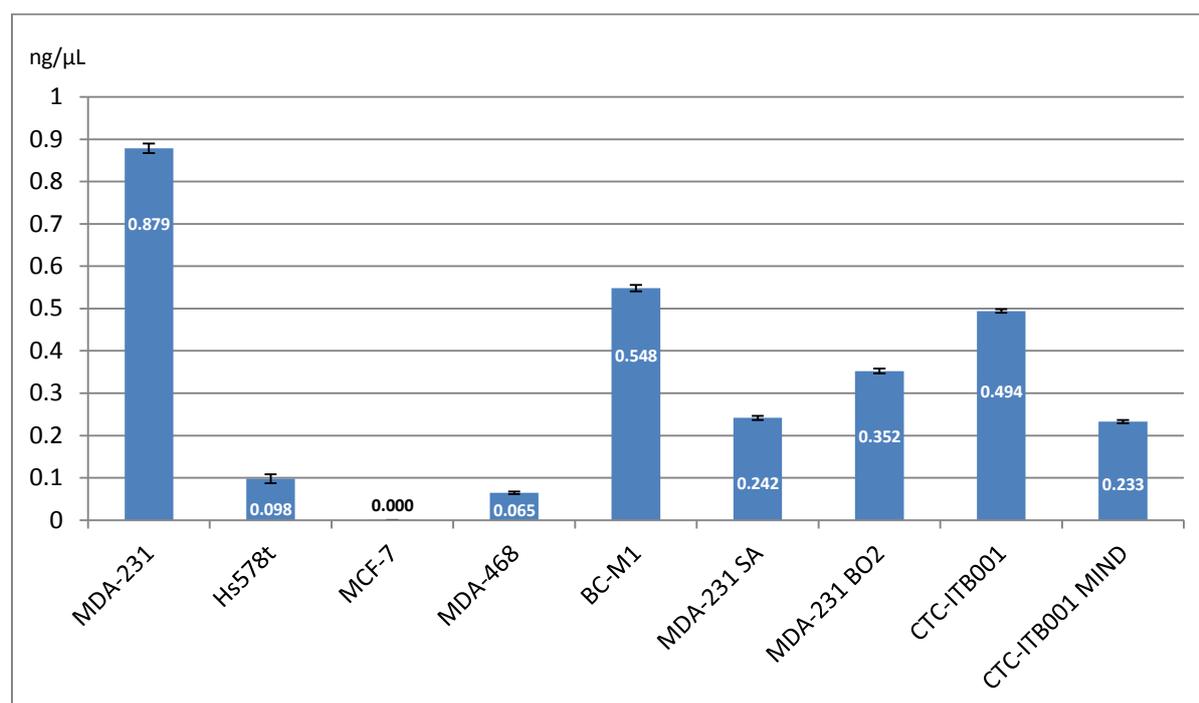


Figure 45: Cell culture supernatant (1 μ L) of different cell lines was diluted in DMEM+10% FCS (99 μ L) and the concentration (ng/ μ L) of CDCP1 was measured with anti-CDCP1 Sandwich ELISA. The values were generated in biological triplicates and shown as a mean value (Blue bar) with errors bars (black lines) showing the standard deviation.

The cell line MDA-231 has the highest concentration of CDCP1 in the supernatant, whereas CDCP1 is not detectable in MCF-7 by ELISA. The supernatant of the cell line Hs578t contains a low amount of CDCP1 as well as MDA-468. The supernatant of the cell line BC-M1 contains the second highest concentration of CDCP1. The cell line CTC-ITB001 shows the third highest concentration. Similar concentrations of CDCP1 are measured in the supernatant of MDA-231 SA, MDA-231 BO2 and CTC-ITB001-MIND. Due to the use of different antibodies with different epitopes for Western Blot analysis and Sandwich-ELISA for the detection of CDCP1 the correlation between the results of both methods is not explicit. The high amount of CDCP1 in cell culture supernatant of MDA-231 (**figure 45**) correlates very well with the signal for the cleaved variant of CDCP1 in Western Blot analysis (**figure 24**). Also, CDCP1 is not detectable in

MCF-7 neither by Sandwich-ELISA nor by Western Blot analysis. In Hs578t a low amount of CDCP1 is detectable in cell culture supernatant whereas in Western Blot analysis CDCP1 is not detectable. The cleaved variant of CDCP1 is detectable in MDA-468 in Western Blot analysis but not in cell culture supernatant by Sandwich-ELISA. CDCP1 in cell culture supernatant is detectable in DTC cell line BC-M1 (0.548 ng/ μ L) by Sandwich ELISA and the full length can be seen in Western Blot analysis. The Sandwich-ELISA is able to prove the existence of CDCP1 in cell culture supernatant of MDA-231 SA and MDA-231 BO2 (**figure 45**) and the cleaved variant is also detectable by Western Blot analysis (**figure 24**). In addition, CDCP1 is also detectable in cell culture supernatant of the CTC cell lines CTC-ITB001 and CTC-ITB001-MIND by Sandwich-ELISA (**figure 45**) as well as in Western Blot analysis (**figure 26**).

In summary, CDCP1 is detectable in cell culture supernatant of cell lines with mesenchymal phenotype like MDA-231, BC-M1 and the bone metastatic sublines MDA-231 SA and MDA-231 BO2. These results correlate very well with the Western Blot analysis (**figure 24**). However, CDCP1 is not detectable in the cell culture supernatant of cell lines with epithelial phenotype (MCF-7/MDA-468) whereas CDCP1 is detectable in MDA-468 in Western Blot analysis (**figure 24**). CDCP1 is detectable in the epithelial CTC cell lines CTC-ITB001 and CTC-ITB001-MIND by Western Blot analysis (**figure 26**) and even in the cell culture supernatant by Sandwich-Elisa (**figure 45**). Due to these facts the analysis of CDCP1 in the blood of breast cancer patients appears to be promising.

The analysis of blood plasma of breast cancer patients requires the determination of unspecific reactions of the Sandwich-ELISA antibodies with components of the blood plasma. Therefore the blood plasma of five healthy female donors by was analyzed by Western Blot analysis (**figure 47**). The protein separation with the sample set of BC-M1, MDA-231 BO2 and MDA-468 was performed three times. One gel was used for a Coomassie staining to verify loading of plasma samples to confirm equal protein loading of each sample (**figure 46**). The other two polyacrylamide-gels were transferred to PVDF membranes. The membranes were incubated with either anti-human CDCP1 antibody (Thermo Fisher) or anti-human CDCP1 antibody (R&D) and analyzed by Western Blot (**figure 47**).

Figure 46 shows the Coomassie stained protein separation gel for the control cell lines BC-M1, MDA-231 BO2 and MDA-468 and five blood plasma samples of healthy female donors of an age over 50 years. The lanes were loaded with same amount of total protein (20 μ g). The band pattern of the plasma samples has a similar intensity and confirms the quality of sample purification, preparation and protein concentration determination.

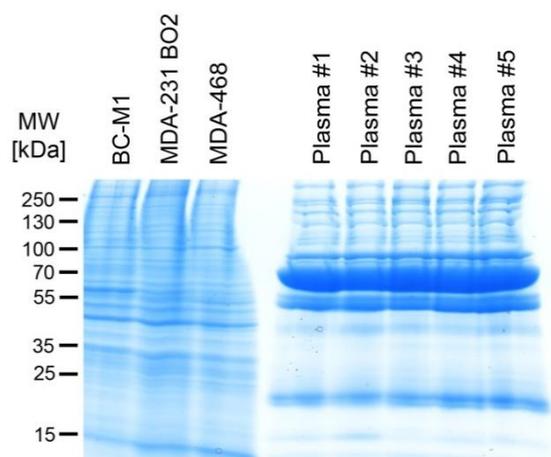


Figure 46: Analysis of PBMC of healthy female donors by SDS PAGE and Coomassie staining. BC-M1, MDA-231 BO2 and MDA-468 serve as control cell lines.

The two PVDF-membranes were incubated with either anti-human CDCP1 antibody (Thermo Fisher, middle image, **figure 47**) or anti-human CDCP1 antibody (R&D, upper image, **figure 47**). CDCP1 is not detectable in the blood plasma of healthy female donors with the capturing-antibody. The full length variant for BC-M1 and the cleaved variant for MDA-231 BO2 are detectable. In contrast, CDCP1 is detectable with a signal at about 130 kDa in the blood plasma of the healthy individuals with the detection antibody. Furthermore there is a weak signal below 55 kDa detectable in plasma sample #1 to #4 and a strong signal at the same size in plasma sample #5. Additionally, the full length variant of CDCP1 is detectable in BC-M1 and MDA-231 BO2.

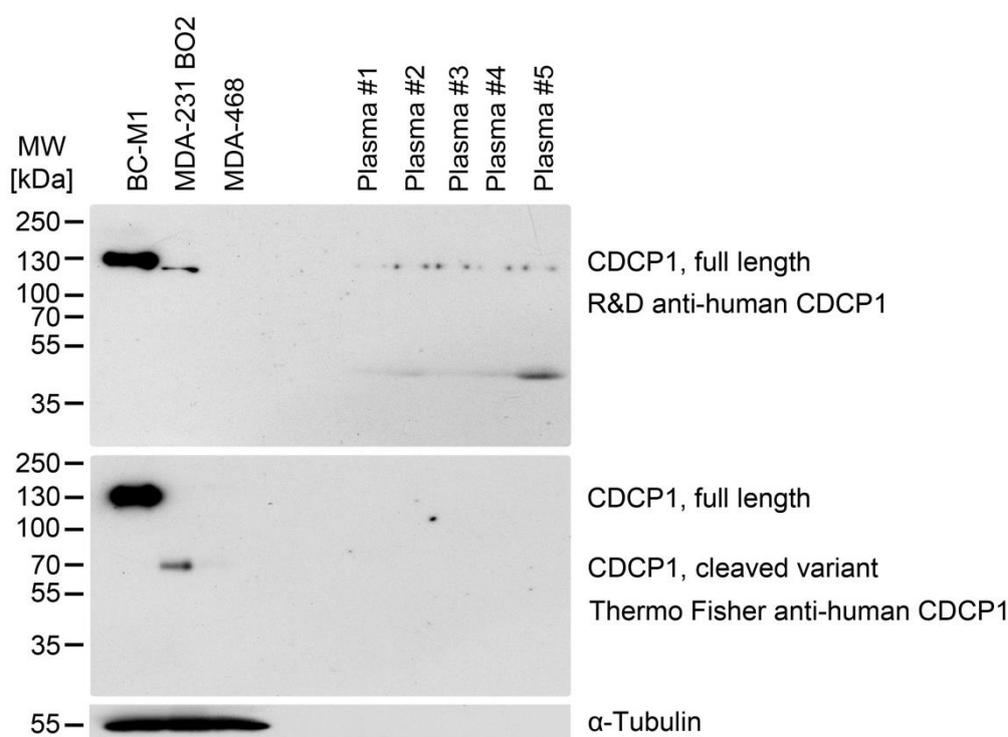


Figure 47: Western Blot analysis of PBMC of healthy female donors with the Sandwich-ELISA antibodies (R&D anti-human CDCP1: upper image, Thermo Fisher anti-human CDCP1: middle image). BC-M1, MDA-231 BO2 and MDA-468 serve as control cell lines. α -Tubulin served as a loading control.

Initial analysis of a small sample set of randomly chosen blood plasma samples of breast cancer patients revealed the presence of CDCP1 in 5 of 12 metastatic breast cancer patients.

Another important parameter is the stability of CDCP1 in the blood plasma of breast cancer patients. The blood plasma of one patient (Gyn17.047) was chosen for stability testing of CDCP1 under different temperatures and time periods by Sandwich-ELISA (**figure 48**). This sample was stored at different temperatures for different time periods.

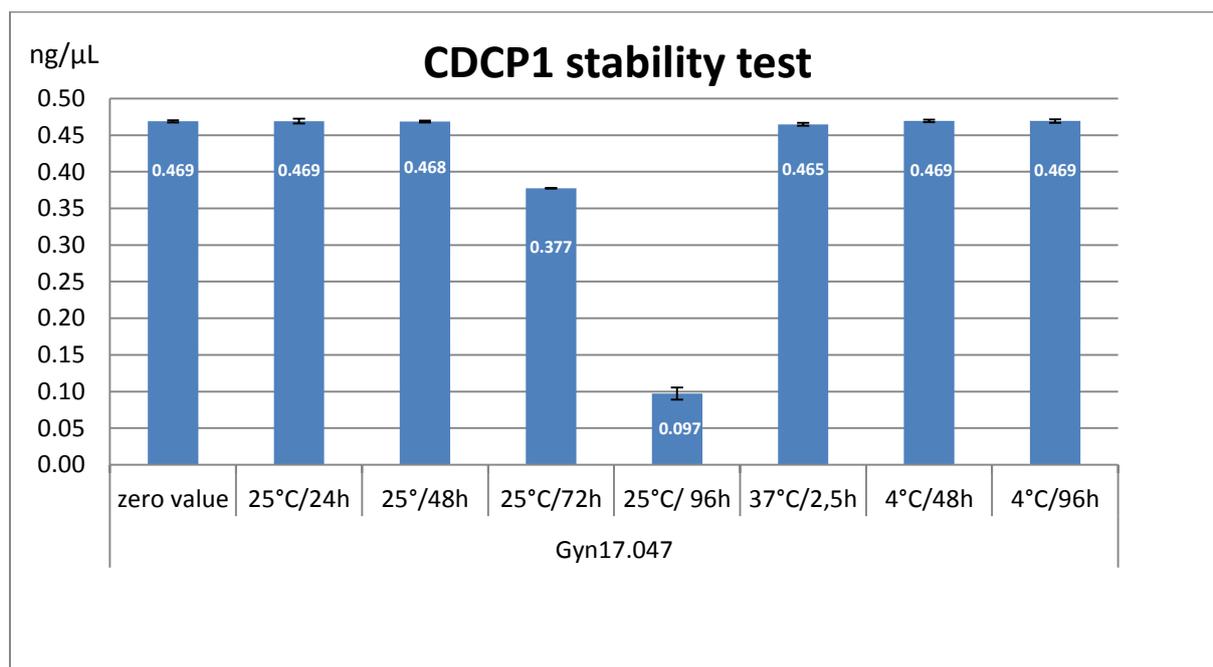


Figure 48: Determination of the stability of CDCP1 in blood plasma of a breast cancer patient. The analysis was performed by CDCP1-Sandwich ELISA. The zero value was taken after initial thawing.

The CDCP1 concentration did not decrease after storage for 48 h at 25°C compared to the zero value. The concentration decreased after 72 h at 25°C by 20 % and major changes were detectable for the storage for 96 h at 25°C by 79 %. For the storage at 37°C for 2.5 h and for the storage at 4°C for 96 h there were no changes in the CDCP1 concentration detectable.

The following chapter examines the analysis of CDCP1 in the blood plasma of a large breast cancer patient cohort.

To prove the reliability of the CDCP1 Sandwich ELISA and to verify the correctness of the determined CDCP1 concentrations in the blood plasma of breast cancer patients each sample was analyzed independently by two researchers (**figure 49**). This was necessary to avoid confusion of samples due to the large number of samples ($n = 771$) and to confirm the reproducibility of the CDCP1 concentrations in the blood plasma of breast cancer patients determined by Sandwich ELISA.

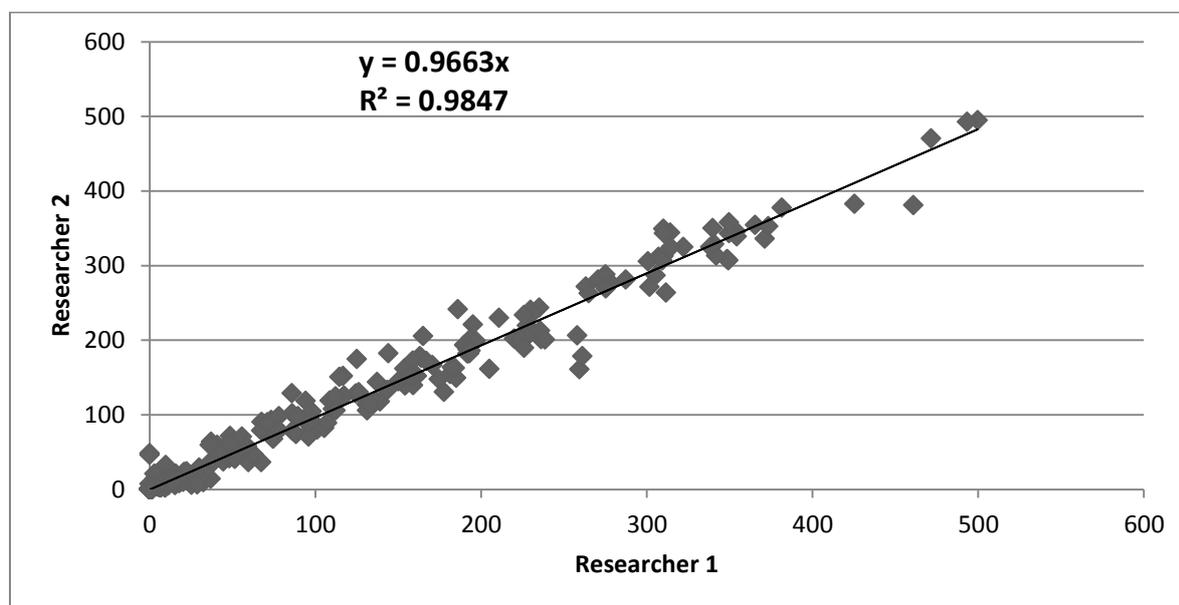


Figure 49: Determination of CDCP1 in the blood plasma of breast cancer patients. All samples from the whole cohort ($n = 771$) were analyzed independently by two researchers (researcher 1 and researcher 2) by CDCP1-Sandwich ELISA. The two determined CDCP1-concentrations of each sample were plotted against each other. The black line signifies the linear regression fit (y) and R^2 describes the determination coefficient.

The slope of the straight line (**figure 49**, black line, y) is 0.9663 and the determination coefficient (**figure 49**, R^2) of 0.9847 signifies a good accordance and reproducibility with minor differences for most of the samples.

5.15. Analysis of CDCP1 in the whole breast cancer patient cohort

CDCP1 was determined in the blood plasma of 771 breast cancer patients (whole patient cohort) and 426 controls (healthy female donors with an age of over 50 years) by Sandwich-ELISA. In total, 1197 samples were analyzed and characterized by ROC analysis (**figure 50**) for specificity and sensitivity.

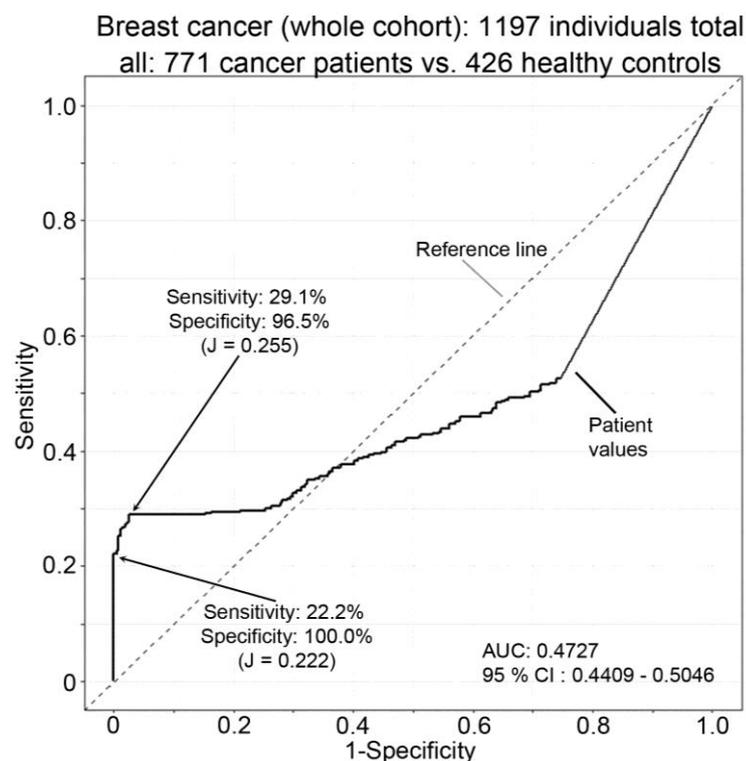


Figure 50: ROC (receiver operating characteristics) analysis for the determination of CDCP1 in the whole breast cancer patient cohort (771 patients and 426 healthy controls) by Sandwich-ELISA. 1-Specificity is plotted on the x-axis, whereas the sensitivity is plotted on the y-axis. The diagonal line (grey) is the reference line and signifies a random process. The black line represents the patient values. AUC: area under the curve, CI: confidence interval, J: Youden-index (degree which indicates the performance of a test to distinguish between two groups) [150].

The analysis of CDCP1 in the blood plasma of 771 breast cancer patients and 426 healthy controls by Sandwich-ELISA can be used to quantify the amount of CDCP1. The results were analyzed with a ROC analysis (**figure 50**) and correlated with different clinicopathological parameters of the patients data (**table 3**). For the complete cohort (771 patient samples) the CDCP1-analysis is able to discriminate between healthy and not healthy with a sensitivity of 22.2 % and specificity of 100 % (Youden-index: 0.222). If the specificity is decreased to 96.5 % the sensitivity reaches a value of 29.1 % (maximal Youden-index: 0.255). The area under curve has a value of 0.4727. The 95 % confidence interval ranges from 0.4409 (bottom level) to 0.5046 (top level). CDCP1 was measured with a mean value of 37.9 ng/mL \pm 87.7 ng/mL

(standard deviation) in the whole cohort (771 breast cancer patients). CDCP1 was detected with a mean value of $0.89 \text{ ng/mL} \pm 1.81 \text{ ng/mL}$ (standard deviation) in the healthy control group (426 individuals).

5.15.1. Analysis of CDCP1 for cT1-cases in the whole breast cancer patient cohort

To assess the utility of CDCP1 for the early detection of breast cancer, the samples with cT1-staging were selected (**figure 51**) from the group of all patients (**figure 50**). The cT1 cohort consisted of 379 patients and was compared with 426 healthy controls. . In total, 805 samples were analyzed and characterized by ROC analysis for specificity and sensitivity.

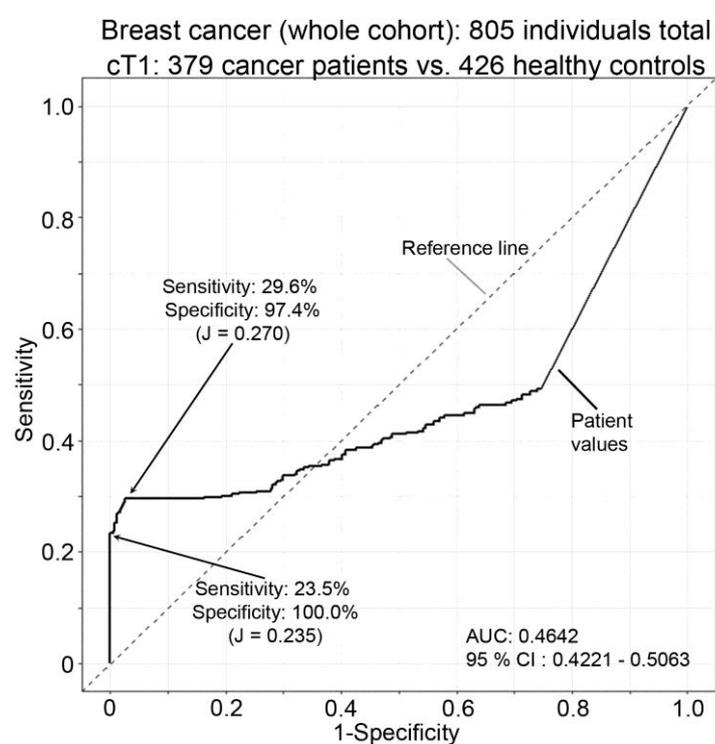


Figure 51: ROC (receiver operating characteristics) analysis for the determination of CDCP1 for tumor staging cT1 of the whole breast cancer patient cohort (379 patients and 426 healthy controls) by Sandwich-ELISA. 1-Specificity is plotted on the x-axis, whereas the sensitivity is plotted on the y-axis. The diagonal line (grey) is the reference line and signifies a random process. The black line represents the patient values. AUC: area under the curve, CI: confidence interval, J: Youden-index.

The analysis of CDCP1 in the blood plasma of 379 breast cancer patients and 426 healthy controls by Sandwich-ELISA can be used to quantify the amount of CDCP1. The results were analyzed with a ROC analysis (**figure 51**) and correlated with different clinicopathological parameters of the patients data (**table**

3). For the cT1-cases of the whole cohort (379 patient samples) the CDCP1-analysis is able to discriminate between healthy and disease with a sensitivity of 23.5 % and specificity of 100 % (Youden-index: 0.235). If the specificity is decreased to 97.4 % the sensitivity reaches a value of 29.6 % (maximal Youden-index: 0.270). The area under curve has a value of 0.4642. The 95 % confidence interval ranges from 0.4221 (bottom level) to 0.5063 (top level). CDCP1 was measured with a mean value of 41.6 ng/mL \pm 93.8 ng/mL (standard deviation) in the whole cohort only considering cT1 cases (379 breast cancer patients). CDCP1 was detected with a mean value of 0.89 ng/mL \pm 1.81 ng/mL (standard deviation) in the healthy control group (426 individuals).

5.15.2. Analysis of CDCP1 for DCIS-cases in the whole breast cancer patient cohort

To determine the usability of CDCP1 for the early detection of precancerous lesions like DCIS, the samples with DCIS were selected (**figure 52**) from the group of all patients (**figure 50**). The DCIS cohort consisted of 71 patients and was compared with 426 healthy controls by Sandwich-ELISA. In total, 497 samples were analyzed and characterized by ROC analysis for specificity and sensitivity. The cohort was manually reduce to 71 patients by exclusion of 700 patients so that only patients with DCIS were taken into consideration.

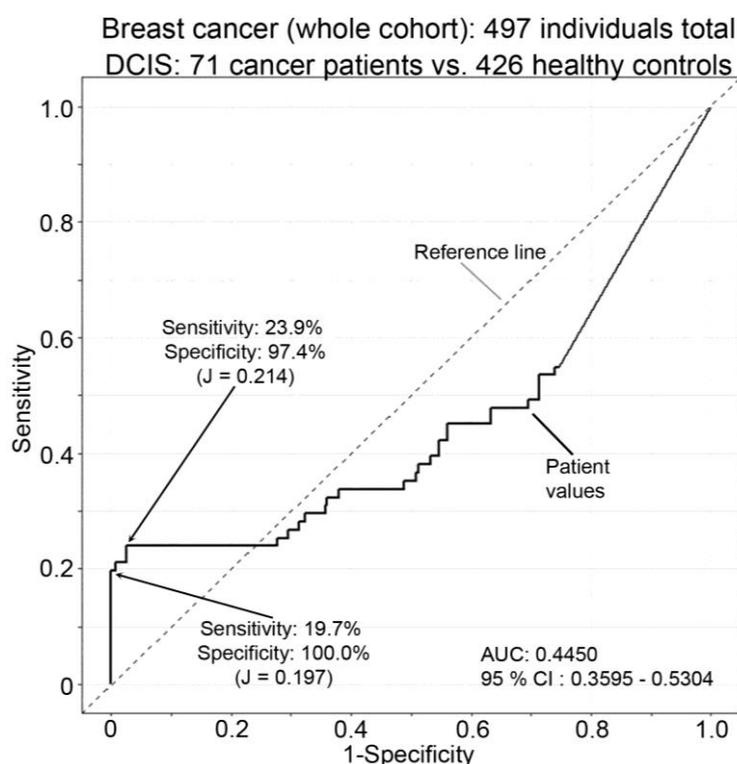


Figure 52: ROC (receiver operating characteristics) analysis for the determination of CDCP1 for diagnosis DCIS of the whole breast cancer patient cohort (71 patients and 426 healthy controls) by Sandwich-ELISA. Specificity is plotted on the x-axis, whereas the sensitivity is plotted on the y-axis. The diagonal line (grey) is the reference line and signifies a random process. The black line represents the patient values. AUC: area under the curve, CI: confidence interval, J: Youden-index.

The analysis of CDCP1 in the blood plasma of 71 breast cancer patients (DCIS) and 426 healthy controls by Sandwich-ELISA can be used to quantify the amount of CDCP1. The results were analyzed with a ROC analysis (**figure 52**) and correlated with different clinicopathological parameters of the patients data

(table 3). For the diagnosis “DCIS” of the whole cohort (71 patient samples) the CDCP1-analysis is able to discriminate between healthy and not healthy with a sensitivity of 19.7 % and specificity of 100 % (Youden-index: 0.197). If the specificity is decreased to 97.4 % the sensitivity reaches a value of 23.9 % (maximal Youden-index: 0.214). The area under curve has a value of 0.445. The 95 % confidence interval ranges from 0.3595 (bottom level) to 0.5304 (top level). CDCP1 was measured with a mean value of 28.0 ng/mL \pm 68.8 ng/mL (standard deviation) in the whole cohort only considering DCIS-cases (71 breast cancer patients). CDCP1 was detected with a mean value of 0.89 ng/mL \pm 1.81 ng/mL (standard deviation) in the healthy control group (426 individuals).

The blood plasma CDCP1 concentration of the 771 breast cancer patients from the whole cohort was correlated with different clinicopathological parameters (**table 7**). The CDCP1 concentration was separated into two groups below and above a cutoff value (18.41 ng/mL). This cutoff value was determined by ROC analysis. The correlated parameters are diagnosis, detection, stage, grading, histology, lymph node invasion, lymphatic invasion, vascular invasion, subtyping and therapy.

Table 4: Correlation of clinicopathological parameters with CDCP1 concentrations (mean value and standard deviation) in the blood plasma of breast cancer patients for the whole cohort. The table shows the number of cases, the corresponding mean value and the standard deviation as well as the percentage of the parameter-related cases for the determined CDCP1 concentration [ng/mL]. Additionally, the table provides the number and the percentage of the parameter-related cases with values below the cutoff value as well as above the cutoff value. The first column on the right-hand side shows the level of significance (p-value calculated by ANOVA analysis) for each subgroup of the clinicopathological parameters ($p < 0.05$ was considered significant).

	n	c (CDCP1) [ng/mL]		c (CDCP1) < 18.41 ng/mL		c (CDCP1) > 18.41 ng/mL		p-value
		n	%	n	%	n	%	
All	771	37.9 +- 87.7	100.0	600	77.8	171	22.2	-
Diagnosis	767	37.7+-87.6	100.0	597	77.8	170	22.2	0.283
Carcinoma	692	38.8 +- 89.4	90.2	536	77.5	156	22.5	
Ductal carcinoma <i>in situ</i>	75	27.4 +-68.1	9.8	61	81.3	14	18.7	
Detection	761	37.4+-87.2	100.0	593	77.9	168	22.1	0.146
Primary detection	683	35.8 + -84.7	89.8	535	78.3	148	21.7	
Recurrence	78	51.0 + - 106.1	10.2	58	74.4	20	25.6	
Stage	741	37.7+-87.9	100.0	579	78.1	162	21.9	0.582
cT1	379	41.6 +-93.8	51.1	290	76.5	89	23.5	
cT2	236	35.1+-83.8	31.8	188	79.7	48	20.3	
cT3	44	39.4+-93.3	5.9	34	77.3	10	22.7	
cT4	11	10.8+-35.1	1.5	10	90.9	1	9.1	
pTis (DCIS)	71	28.0+-68.8	9.6	57	80.3	14	19.7	
Grading	727	37.8+-88.3	100.0	569	78.3	158	21.7	0.023
G1	111	49.7+-96.5	15.3	83	74.8	28	25.2	
G2	330	43.2+-96.4	45.4	251	76.1	79	23.9	
G3	286	27.0+-88.3	39.3	235	82.2	51	17.8	
Histology	598	37.0+-86.3	100.0	470	78.6	128	21.4	0.573
Infiltrating ductal	176	48.5+-100.8	29.4	131	74.4	45	25.6	
Infiltrating lobular	87	36.1+-94.9	14.5	70	80.5	17	19.5	
papillary	8	24.8+-69.5	1.3	7	87.5	1	12.5	
no special type	292	30.6+-74.5	48.8	234	80.1	58	19.9	
others	35	31.9+-79.2	5.9	28	80.0	7	20.0	
Lymph node status	689	38.2+-87.9	100.0	535	77.6	154	22.4	0.999

Positive	218	38.2+-92.5	31.6	170	78.0	48	22.0	
Negative	471	38.2+-85.8	68.4	365	77.5	106	22.5	
Lymphatic invasion	615	39.0+-89.8	100.0	477	77.6	138	22.4	0.211
Positive	147	47.1+-106.1	23.9	113	76.9	34	23.1	
Negative	468	36.5+-84.0	76.1	364	77.8	104	22.2	
Vascular invasion	610	38.7+-89.6	100.0	474	77.7	136	22.3	0.208
Positive	10	3.3+6.7	1.6	9	90.0	1	10.0	
Negative	600	39.3+-90.3	98.4	465	77.5	135	22.5	
Subtyping	665	38.2+-88.9	100.0	516	77.6	149	22.4	0.358
ER+PR positive	477	40.8+-93.7	71.7	369	77.4	108	22.6	
ErbB-2 positive	82	36.8+-74.3	12.3	59	72.0	23	28.0	
Triple negative	106	27.2+-76.0	15.9	88	83.0	18	17.0	
Therapy	723	37.3+-87.4	100.0	564	78.0	159	22.0	0.049
Neoadjuvant	43	7.3+-24.7	5.9	40	93.0	3	7.0	
Adjuvant	540	36.7+-86.5	74.7	423	78.3	117	21.7	
Radiation	66	53.2+-96.0	9.1	45	68.2	21	31.8	
others	74	45.0+-104.4	10.2	56	75.7	18	24.3	

^amean value \pm standard deviation

^bANOVA analysis

The p-value for the grading (0.023) shows a significant correlation with CDCP1. A lower concentration signifies the grading status “G3”. Additionally, the p-value for therapy represents a significant correlation with CDCP1 in the blood plasma of breast cancer patients. Lower concentrations of CDCP1 are detectable in patients who received a neoadjuvant therapy prior to surgery. No other clinicopathological parameters correlated significantly with CDCP1.

5.16. Analysis of CDCP1 in the “early detection” breast cancer patient cohort

The large number of samples of breast cancer patients allowed the restriction of the initial whole cohort to an “early detection” cohort with remaining statistical power. Exclusion of cases like having received neoadjuvant therapy (see method section) yielded a collective of breast cancer cases as found at the time of primary breast cancer diagnosis. Additionally all cases with a recurrence were excluded. Therefore it was possible to analyze the quality of CDCP1-ELISA in detecting breast cancer prior to any medical intervention or influences through disease relapse.

CDCP1 was measured in the blood plasma of 440 breast cancer patients and 426 controls (healthy female donors) by Sandwich-ELISA. In total, 866 samples were analyzed by a ROC curve (**figure 53**).

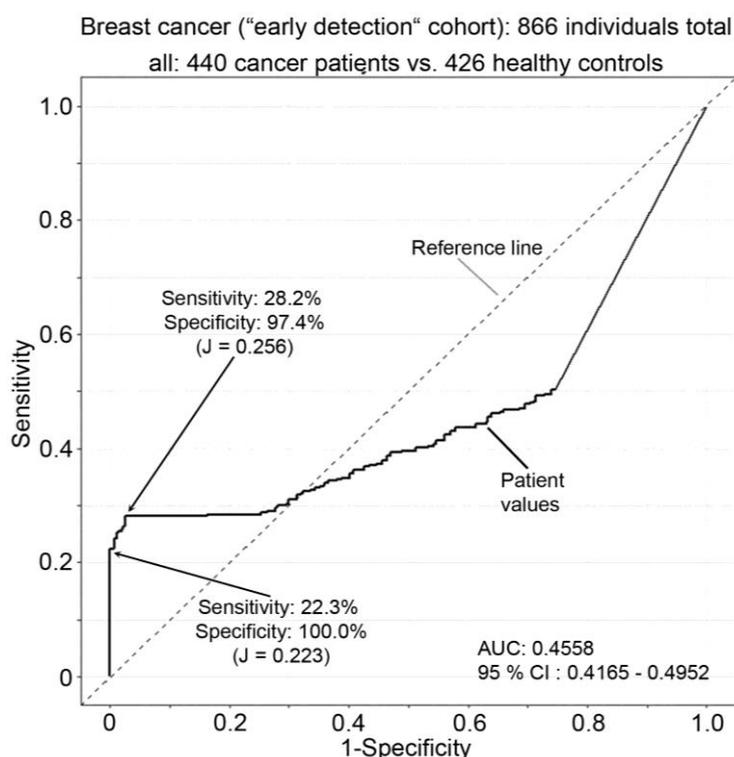


Figure 53: ROC (receiver operating characteristics) analysis for the determination of CDCP1 in the breast cancer patient “early detection cohort” (440 patients and 426 healthy controls) by Sandwich-ELISA. Specificity is plotted on the x-axis, whereas the sensitivity is plotted on the y-axis. The diagonal line (grey) is the reference line and signifies a random process. The black line represents the patient values. AUC: area under the curve, CI: confidence interval, J: Youden-index.

The analysis of CDCP1 in the blood plasma of 440 breast cancer patients and 426 healthy controls by Sandwich-ELISA can be used to quantify the amount of CDCP1. The results were analyzed with a ROC analysis (**figure 53**) and correlated with different parameters of the patient data (**table 4**). For the “early

detection"-cohort (440 patient samples) the CDCP1-analysis is able to discriminate between healthy and not healthy with a sensitivity of 22.3 % and specificity of 100 % (Youden-index: 0.223). If the specificity is decreased to 97.4 % the sensitivity reaches a value of 28.2 % (maximal Youden-index: 0.256). The area under curve has a value of 0.4558. The 95 % confidence interval ranges from 0.4165 (bottom level) to 0.4952 (top level). CDCP1 was measured with a mean value of 37.8 ng/mL \pm 89.1 ng/mL (standard deviation) in the "early detection" cohort (440 breast cancer patients). CDCP1 was detected with a mean value 0.89 ng/mL \pm 1.81 ng/mL (standard deviation) in the healthy control group (426 individuals).

5.16.1. Analysis of CDCP1 in cT1 cases from the “early detection” breast cancer patient cohort

Furthermore, it was examined if CDCP1 can be used as a biomarker for the detection of early stages of breast cancer (cT1 stadium) or precursor diseases like the ductal carcinoma in situ (DCIS). The cT1 cases were selected from “early detection” cohort (**figure 53**). Initially, all cases with tumor stage I (cT1 stadium) were analyzed. Thereby the sample number was reduced to 241 patient samples. These were compared with 426 healthy controls by ROC analysis (**figure 54**). In total, 667 samples were analyzed.

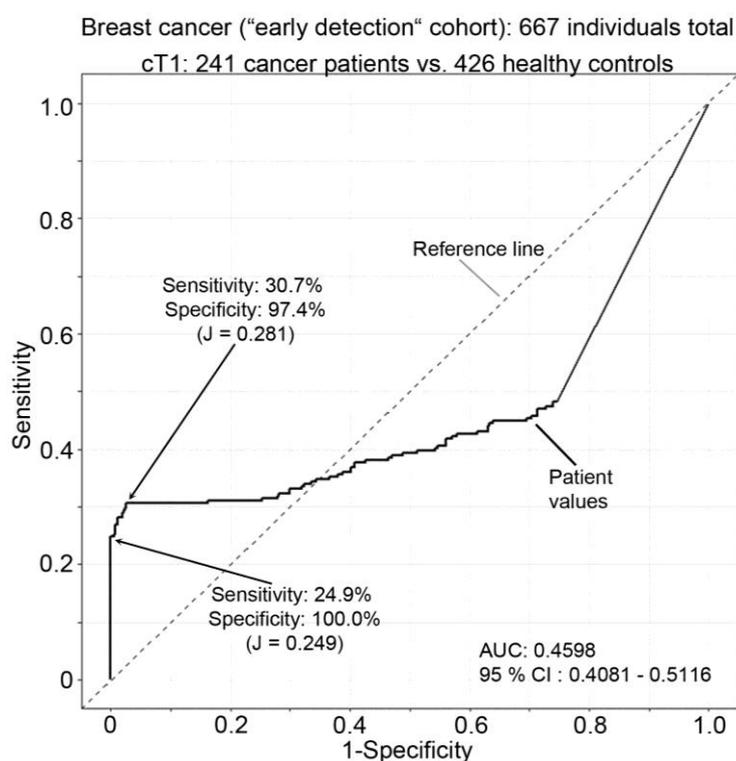


Figure 54: ROC (receiver operating characteristics) analysis for the determination of CDCP1 in the breast cancer patient “early detection” cohort for cT1-cases (241 patients and 426 healthy controls) by Sandwich-ELISA. Specificity is plotted on the x-axis, whereas the sensitivity is plotted on the y-axis. The diagonal line (grey) is the reference line and signifies a random process. The black line represents the patient values. AUC: area under the curve, CI: confidence interval, J: Youden-index. Only cases with tumor stage cT1 were analyzed.

The concentration of CDCP1 was determined in the blood plasma of 241 breast cancer patients with tumor stage cT1 and 426 healthy controls by Sandwich-ELISA. The results were analyzed with a ROC analysis (**figure 54**). The test is able to detect 24.9% of stage cT 1 cases in the breast cancer “early detection cohort” with a specificity of 100% (Youden-index: 0.249). If the specificity is decreased to 97.4% the sensitivity reaches a value of 30.7% (maximal Youden-index: 0.281) which is a promising result. The area under curve has a value of 0.4598. The 95% confidence interval ranges from 0.4081 (bottom level) to 0.5116 (top level).

CDCP1 was measured with a mean value of 44.7 ng/mL \pm 98.4 ng/mL (standard deviation) in stage cT1 breast cancer patients in the “early detection” cohort (241 breast cancer patients). CDCP1 was detected with a mean value 0.89 ng/mL \pm 1.81 ng/mL (standard deviation) in the healthy control group (426 individuals).

5.16.2. Analysis of CDCP1 in DCIS cases from the “early detection” breast cancer patient cohort

For the detection of precancerous stages like DCIS, samples with DCIS-diagnosis were isolated from the “early detection” cohort and analyzed by ROC analysis. Thereby the sample number contained 45 patient samples. These were compared with 426 healthy controls by ROC analysis. In total, 471 samples were analyzed (**figure 55**).

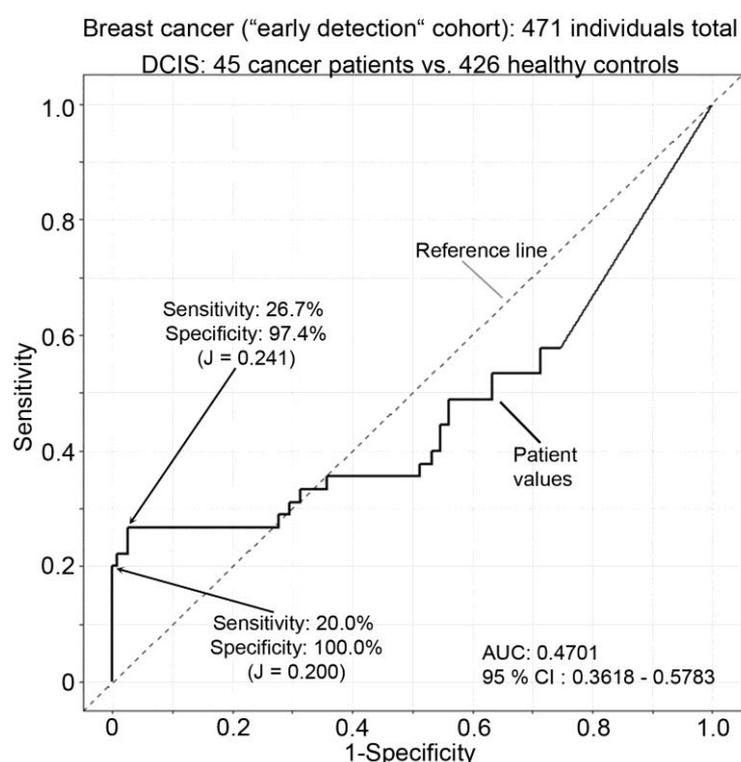


Figure 55: ROC (receiver operating characteristics) analysis for the determination of CDCP1 in the breast cancer patient “early detection” cohort for DCIS-diagnosis(45 patients and 426 healthy controls) by Sandwich-ELISA. Specificity is plotted on the x-axis, whereas the sensitivity is plotted on the y-axis. The diagonal line (grey) is the reference line and signifies a random process. The black line represents the patient values. AUC: area under the curve, CI: confidence interval, J: Youden-index. Only cases with a DCIS-diagnosis were analyzed.

CDCP1 was detected with a sensitivity of 20.00% and a specificity of 100% (Youden index: 0.200) in all cases with a DCIS diagnosis in the “early detection” cohort (**figure 55**). By decreasing the specificity to 97.4% the sensitivity would increase to 26.7% (maximal Youden index: 0.241). The area under curve has a value of 0.4701. The 95% confidence interval ranges from 0.3618 (bottom level) to 0.5783 (top level). CDCP1 was measured with a mean value of 26.4 ng/mL \pm 69.7 ng/mL (standard deviation) in patients

with DCIS diagnosis in the “early detection” cohort (45 patients). CDCP1 was detected with a mean value $0.89 \text{ ng/mL} \pm 1.81 \text{ ng/mL}$ (standard deviation) in the healthy control group (426 individuals).

The blood plasma CDCP1 concentration of the 440 breast cancer patients from the “early detection”-cohort was correlated with clinicopathological parameters (**table 8**). The CDCP1 concentration was separated into two groups below and above the cutoff value (18.622 ng/mL). This cutoff value was determined by ROC analysis. The correlated parameters are diagnosis, detection, stage, grading, histology, lymph node invasion, lymphatic invasion, vascular invasion, subtyping and therapy.

Table 5: Correlation of clinicopathological parameters with CDCP1 concentrations (mean value and standard deviation) in the blood plasma of breast cancer patients for the “early detection” cohort. The table shows the number of cases, the corresponding mean value and the standard deviation as well as the percentage of the parameter-related cases for the determined CDCP1 concentration [ng/mL]. Additionally, the table provides the number and the percentage of the parameter-related cases with values below the cutoff value as well as above the cutoff value. The first column on the right-hand side shows the level of significance (p-value calculated by ANOVA analysis) for each subgroup of the clinicopathological parameters (p < 0.05 was considered significant).

	n	c (CDCP1) [ng/mL]		c (CDCP1) < 18.622 ng/mL		c (CDCP1) > 18.622 ng/mL		p-value ^b
		n	%	n	%	n	%	
All	440	37.8+-88.9	100.0	342	77.7	98	22.3	-
Diagnosis	439	37.9+-89.1	100.0	341	77.7	98	22.3	0.461
Carcinoma	393	39.0+-91.0	89.5	305	77.6	88	22.4	
Ductal carcinoma <i>in situ</i>	46	28.7+-70.6	10.5	36	78.3	10	21.7	
Detection	440	37.8+-89.0	100.0	342	77.7	98	22.3	N/A
Primary detection	440	37.8+-89.0	100.0	342	77.7	98	22.3	
Recurrence	0	0	0.0		0.0		0.0	
Stage	440	37.8+-89.0	100.0	342	77.7	98	22.3	0.354
cT1	241	44.7+-98.4	54.8	181	75.1	60	24.9	
cT2	124	33.4+-82.6	28.2	99	79.8	25	20.2	
cT3	22	24.2+-55.7	5.0	18	81.8	4	18.2	
cT4	8	0.25+-0.35	1.8	8	100.0	0	0.0	
pTis (DCIS)	45	26.4+-69.7	10.2	36	80.0	9	20.0	
Grading	429	38.1+-89.9	100.0	335	78.1	94	21.9	0.127
G1	82	54.3+-100.5	19.1	59	72.0	23	28.0	
G2	188	38.3+-93.8	43.8	147	78.2	41	21.8	
G3	159	29.5+-78.1	37.1	129	81.1	30	18.9	
Histology	336	35.9+-85.4	100.0	265	78.9	71	21.1	0.266
Infiltrating ductal	140	46.6+-99.4	41.7	105	75.0	35	25.0	
Infiltrating lobular	53	13.2+-48.0	15.8	47	88.7	6	11.3	
papillary	5	39.6+-87.9	1.5	4	80.0	1	20.0	
no special type	112	33.0+-78.7	33.3	88	78.6	24	21.4	
others	26	36.1+-88.1	7.7	21	80.8	5	19.2	
Lymph node status	403	39.0+-90.6	100.0	1487	369.0	432	107.2	0.394

Positive	136	33.6+-87.8	33.7	107	78.7	29	21.3	
Negative	267	41.7+-92.0	66.3	205	76.8	62	23.2	
Lymphatic invasion	369	39.3+-91.6	100.0	1107	300.0	320	86.7	0.421
Positive	89	46.1+-106.1	24.1	68	76.4	21	23.6	
Negative	280	37.1+-86.6	75.9	218	77.9	62	22.1	
Vascular invasion	368	39.4+-91.7	100.0	814	221.2	237	64.4	0.255
Positive	7	0.23+-0.18	1.9	7	100.0	0	0.0	
Negative	361	40.1+-92.4	98.1	278	77.0	83	23.0	
Subtyping	389	38.3+-90.6	100.0	302	77.6	87	22.4	0.777
ER+PR positive	294	40.1+-94.1	75.6	227	77.2	67	22.8	
ErbB-2 positive	47	30.6+-69.1	12.1	34	72.3	13	27.7	
Triple negative	48	35.2+-88.5	12.3	41	85.4	7	14.6	
Therapy	440	37.8+-89.0	100.0	342	77.7	98	22.3	0.559
Neoadjuvant	0	0	0.0	0	0.0	0	0.0	
Adjuvant	353	36.0+-88.3	80.2	278	78.8	75	21.2	
Radiation	34	53.1+-97.2	7.7	24	70.6	10	29.4	
others	53	39.7+-88.9	12.0	40	75.5	13	24.5	

^amean value \pm standard deviation

^bANOVA analysis

The blood plasma CDCP1 concentration of the 440 breast cancer patients from the early detection cohort was correlated with different clinicopathological parameters. The CDCP1 concentration was separated into two groups below and above a cutoff value. This cutoff value was determined by ROC analysis. The correlated parameters are diagnosis, detection, stage, grading, histology, lymph node invasion, lymphatic invasion, vascular invasion, subtyping and therapy. There was no significant correlation between CDCP1 and clinical parameters detectable.

5.17. Investigation of the influence of microenvironmental changes on the expression levels of the biomarker candidates

Cancer cells inside of primary tumor can be exposed to hypoxia due to the restricted vascularization with blood vessels. Cancer cells can be exposed to cell stress due to the disturbed angiogenesis and resulting ischemia. The ability for adapting to such crucial changes might lead to an evolutionary advantage [151, 152]. The undersupply with oxygen, hypoxia, is one of the states causing cell stress and an adaptive reaction of the cancer cell to avoid cell death. Furthermore, circulating tumor cells can be subjected to hypoxia at secondary sites like the bone marrow [72, 77]. Therefore the changes of the biomarker candidate expression levels after cultivation under hypoxic conditions was investigated. The cell lines MCF-7, MDA-478, MDA-231 and BC-M1 were selected as representatives for the respective cancer subtypes. The exposure to one percent oxygen for 14 days should simulate the situation inside of a primary tumor with low oxygen supply [68] or the lower oxygen concentration in the hematopoietic stem cell niche [77, 153]. The following exposure to ten percent oxygen should simulate the situation after leaving the area with low oxygen supply inside the primary tumor and entering well-oxygenated blood circulation. Additionally, this model could simulate the situation when dormant cancer cells leave the bone marrow and enter the blood circulation again.

5.17.1. Investigation of the influence of hypoxia on the expression levels of the biomarker candidates

Figure 56 shows the changes of the biomarker expression levels under hypoxia (1% O₂, 14 days) and following reoxygenation (10% O₂, four hours) by Western Blot analysis

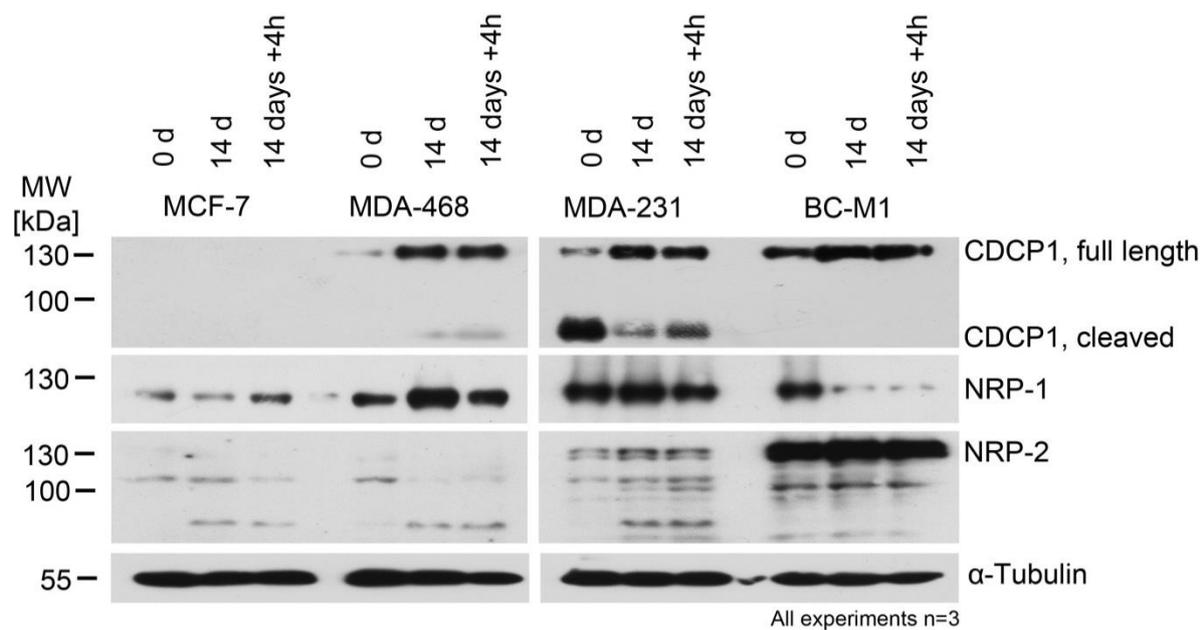


Figure 56: Western blot analysis of the cleaved and full length variant of CDCP1, NRP-1 and NRP-2 on breast cancer cell lines after exposure to hypoxia (14 d) and following reoxygenation (14 d +4h) 0d are standard cell culture conditions. α -Tubulin served as a loading control.

The comparison of the biomarker expression levels between standard oxygen conditions (0d) and hypoxia followed by reoxygenation (14d + 4 h) was performed to simulate the event where CTCs leave hypoxic regions of the primary tumor and enter the well-oxygenated blood circulation. The time period of four hours of reoxygenation was chosen because an estimated half-life for CTCs of 2.4 hours was observed but CTCs were also detected between three and eight hours [154]. This analysis provides important information about the ability of the biomarkers to detect CTCs developed in hypoxic regions. Due to the fact that hypoxia is a possible activator of EMT [64, 65] which is known to play a crucial role in invasiveness and metastatic spread [7, 35] it is necessary to generate this type of CTCs.

The cell line MCF-7 is negative for all variants of CDCP1 under all analyzed culture conditions. MDA-468 has a weak expression of the full length variant of CDCP1 (**figure 57A**) but the cleaved variant is not detectable under normal culture conditions by Western Blot analysis (**figure 56**) and has a very low signal intensity determined by quantification (**figure 57 B**). The signal intensity of the full length variant increases after 14 days of hypoxia (14d) by 226 % and by 193 % after reoxygenation (14 days+ 4h) compared to the standard cultivation conditions (0d). The cleaved variant is detectable with an increase by 2313 % after 14 days of hypoxia and reoxygenation for four hours (14 days+ 4h) compared to the standard cultivation conditions (0d) but this increase is not significant.

MDA-231 has a moderate expression of the full length variant of CDCP1 (**figure 57 C**) under normal culture conditions (0d), whereas the cleaved variant exhibits a strong signal (**figure 57 D**) under normal culture conditions (0d). After 14 days of hypoxia (14d) the expression level of the full length variant slightly increases by 41 % and keeps stable with minor changes after reoxygenation (14d + 4h). The expression level of the cleaved variant decreases significantly by 57 % after 14 days of hypoxia (14d) and keeps stable at this reduced expression level with minor changes after reoxygenation (14 days+ 4h).

BC-M1 has the strongest expression level of full length CDCP1 (**figure 57 E**) under normal culture conditions (0d) for the analyzed cell lines (**figure 56**). With the exposure to an O₂ concentration of one percent for 14 days (14d) the expression level of the full length variant does not change significantly but increases by 35 % after hypoxia and reoxygenation (14d + 4h) compared with the standard cultivation conditions (0d).

A significant reduction of CDCP1 is only detectable for the cleaved variant of CDCP1 in MDA-231 between normal cultivation conditions (0d) and 14 days hypoxia (14d) or 14 days of hypoxia followed by reoxygenation (14d +4h).

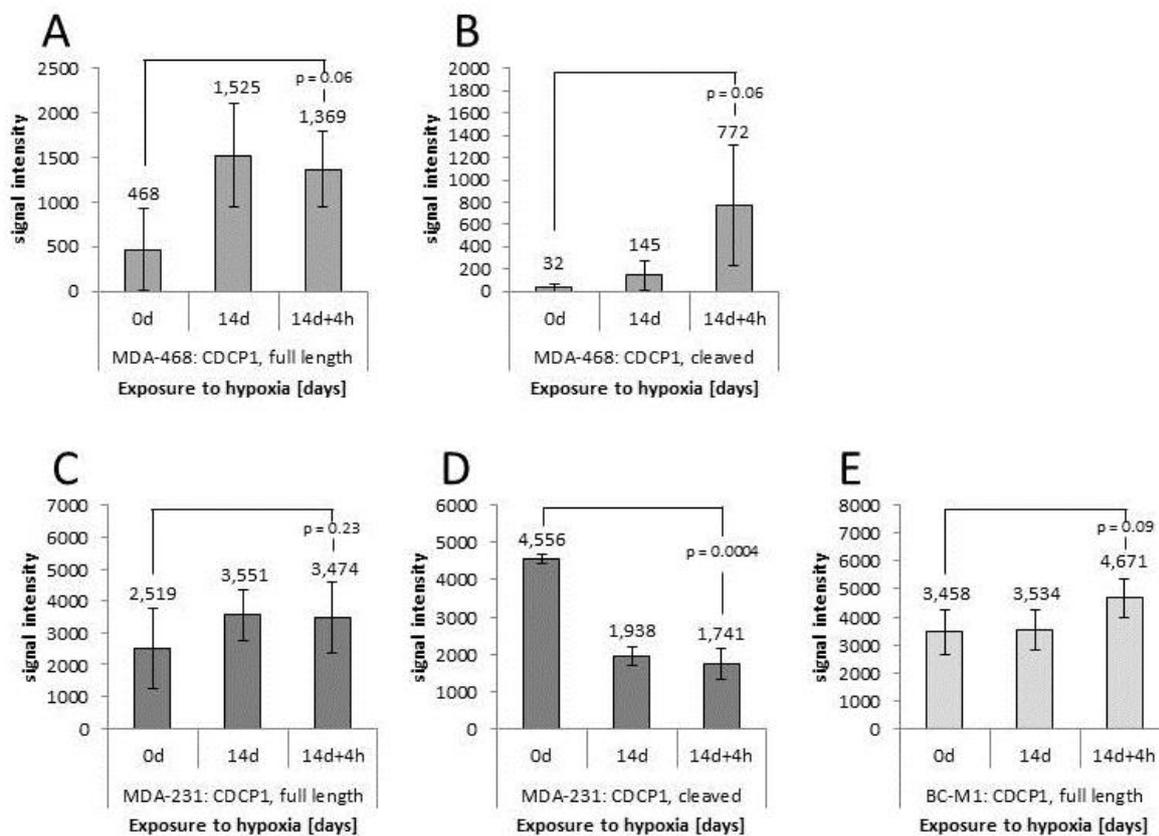


Figure 57: Quantification of the Western Blot analysis for the changes of CDCP1 under exposure to hypoxia (14d) and reoxygenation (14d +4h) for the cell lines MDA-468 (A,B), MDA-231 (C,D) and BC-M1 (E). The level of significance (p) is shown for the difference between the state without atmospheric oxygen concentration (0d) and the exposure one percent oxygen for 14 days followed by exposure to ten percent oxygen for four hours (14d+ 4h). The signal intensity is represented in arbitrary units which were generate by the quantification software Quantity One.

MCF-7 has a weak expression of NRP-1 (**figure 58 A**) under normal culture conditions (0d). After cultivation under one percent of oxygen for 14 days (14d) NRP-1 expression slightly decreases by 59 % whereas the signal intensity significantly increases by 175 % after following reoxygenation for four hours (14 days +4h) compared to 14 days of hypoxia.

NRP-1 is detectable with a weak expression level (**figure 58 B**) under normal culture conditions (0d) in MDA-468. The expression level of NRP-1 is significantly elevated by 407 % after cultivation under hypoxia for 14 days and after reoxygenation by 367 % compared to the expression level under standard cultivation (0d).

MDA-231 exhibits the strongest signal of NRP-1 (**figure 58 C**) under normal conditions (0d) of all analyzed cell lines. The expression level of NRP-1 is slightly decreased by 28 % after 14 days under hypoxia without significance (14d) and keeps stable at this reduced expression level with minor changes after reoxygenation (14d + 4h).

BC-M1 has also a strong expression level of NRP-1 (**figure 58 D**) under normal culture conditions (0d). Compared to standard cultivation (0d), the expression levels of NRP-1 are significantly reduced by 67 % after 14 days of hypoxia (14d) and by 78 % after 14 days of hypoxia and reoxygenation for four hours (14d + 4h).

A significant change of NRP-1 is only detectable in MDA-468 (increase) and BC-M1 (decrease) between atmospheric oxygen concentration (0d) and 14 days of hypoxia followed by reoxygenation (14d +4h).

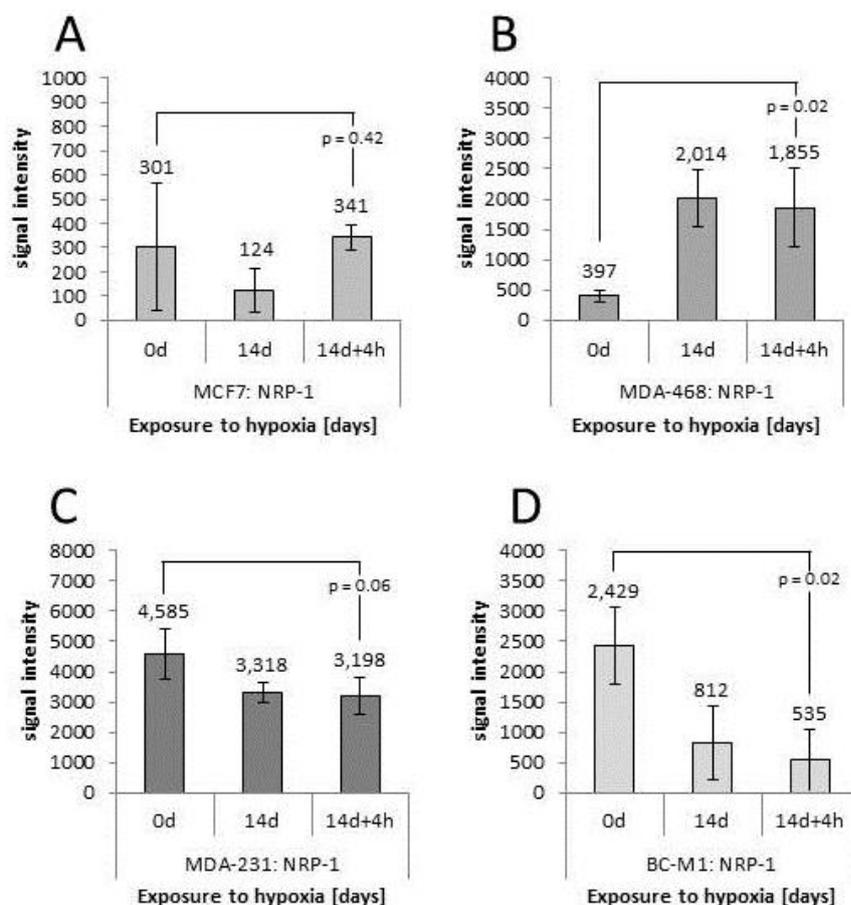


Figure 58: Quantification of the Western Blot analysis for the changes of CDCP1 under exposure to hypoxia (14d) and reoxygenation (14d +4h) for the cell lines MDA-468 (A,B), MDA-231 (C,D) and BC-M1 (E). The level of significance (p) is shown for the difference between the state without atmospheric oxygen concentration (0d) and the exposure one percent oxygen for 14 days followed by exposure to ten percent oxygen for four hours (14d+ 4h). The signal intensity is represented in arbitrary units which were generate by the quantification software Quantity One.

MCF-7 and MDA-468 do not exhibit a signal at a molecular weight of 130 kDa for NRP-2. The detectable signals cannot be related accurately to specific signals of NRP-2 without fail. Therefore a quantification was not performed.

For NRP-2, there are moderate signals of a double band (**figure 59 A**) detectable at a molecular weight of 130 kDa under normal culture conditions (0d) in MDA-231. The expression level slightly increases by 44 % after 14 days under hypoxia (14d) and a decrease by 22 % was detected after reoxygenation (14d + 4h) compared to standard cultivation (0d) . The signals at about 100 kDa and 70 kDa are the same detectable signals for MCF-7 and MDA-468 but their origin remains unclear.

BC-M1 has the strongest expression level NRP-2 (**figure 59 B**) under normal culture conditions (0d) for the analyzed cell lines (**figure 56**). The expression level slightly decreases by 23 % after 14 days of hypoxia (14d). NRP-2 increases significantly (**table 5**) by 33 % after reoxygenation (14d + 4h) compared to the expression level after 14 days of hypoxia (14d). The expression levels of NRP-2 do not differ significantly between standard cultivation conditions (0d) and cultivation under hypoxia for 14 days and following reoxygenation (14d + 4h).

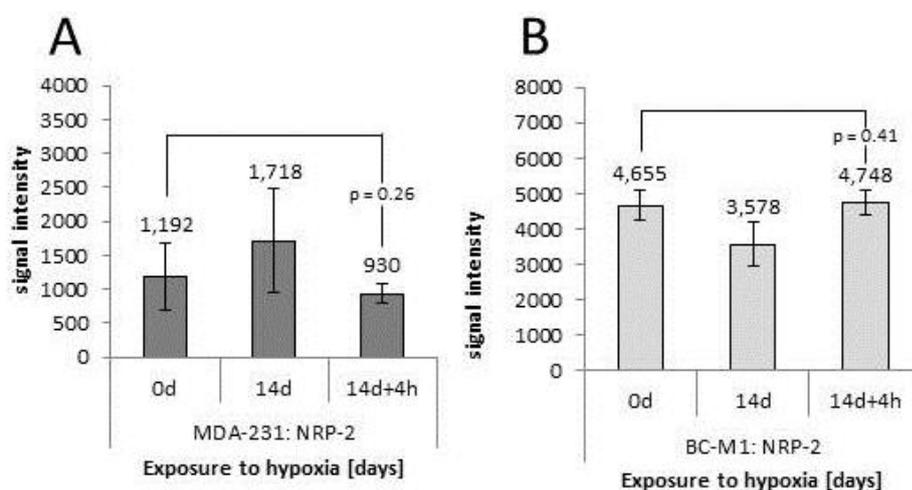


Figure 59: Quantification of the Western Blot analysis for the changes of NRP-2 under exposure to hypoxia (14d) and reoxygenation (14d +4h) for the cell lines MDA-231 (A) and BC-M1 (B). The level of significance (p) is shown for the difference between the state without atmospheric oxygen concentration (0d) and the exposure one percent oxygen for 14 days followed by exposure to ten percent oxygen for four hours (14d+ 4h). The signal intensity is represented in arbitrary units which were generate by the quantification software Quantity One.

Table 6: Levels of significance (p-value) for the expression level differences of the biomarkers under exposure to hypoxia (14d) and following reoxygenation (14d + 4h) for the combination of the different time points. 0d/4d signifies the difference between cultivation under normal oxygen concentration (0d) 14 days of hypoxia (14d).

MCF-7	0d/14d ^a	0d/14d +4h ^a	14d/14d +4h ^a
NRP-1	0.2103	0.4224	0.0202
MDA-468	0d/14d	0d/14d +4h	14d/14d +4h
CDCP1, full length	0.0572	0.0553	0.3868
CDCP1, cleaved	0.1557	0.0641	0.0950
NRP-1	0.0046	0.0174	0.4367
MDA-231	0d/14d	0d/14d +4h	14d/14d +4h
CDCP1, full length	0.1941	0.2337	0.4697
CDCP1, cleaved	0.0001	0.0004	0.2957
NRP-1	0.0592	0.0648	0.4072
NRP-2	0.2323	0.2573	0.1147
BC-M1	0d/14d	0d/14d +4h	14d/14d +4h
CDCP1, full length	0.4634	0.0898	0.0909
NRP-1	0.0310	0.0156	0.3239
NRP-2	0.0552	0.4094	0.0384

^a p-values calculated by Student *t* test

The biomarkers CDCP1, NRP-1 and NRP-2 exhibit detectable expression levels after 14 days of hypoxia and even after 14 days of hypoxia and following reoxygenation for Western Blot analysis. However, NRP-1 was significantly reduced only in BC-M1 after 14 days of hypoxia and 14 days of hypoxia with following reoxygenation but was still detectable.

5.17.2. Investigation of the influence of cobalt on the expression levels of the biomarker candidates

The cultivation of cancer cells with cobalt-supplemented cell culture medium can simulate hypoxia by exclusively stabilizing the hypoxia-inducible factor 1 α (HIF-1 α). The cells can be then cultivated under normal oxygen- and nitrogen concentrations and the possible changes in expression levels can be related to the stabilization HIF-1 α different to hypoxia where different pathway can be activated and thereby generating a stimulus [76]. Therefore specific cancer cell lines were cultivated with cobalt-containing cell culture medium and the changes in the expression levels of the biomarker candidates by Western Blot analysis (**figure 60**) were analyzed. The cell lines were chosen because of their undetectable or weak expression level of HIF-1 α under cultivation without cobalt-chloride (S).

Figure 60 shows the Western Blot analysis of the biomarker expression level changes under exposure to cobalt-chloride (CoCl₂) for different time periods.

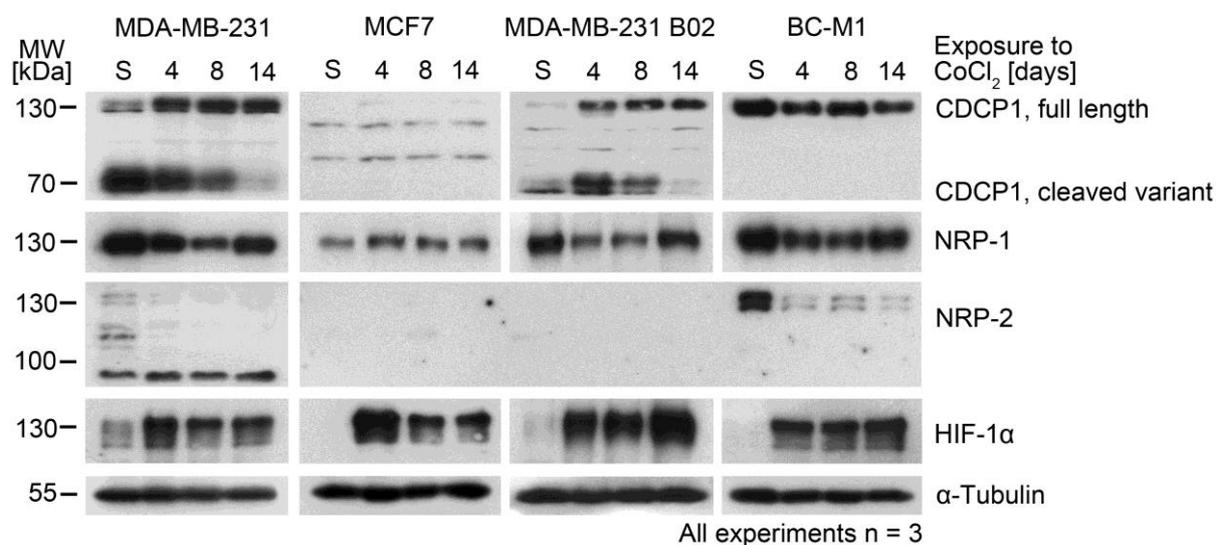


Figure 60: Western blot analysis of the cleaved and full length variant of CDCP1, NRP-1 and NRP-2 on breast cancer cell lines after exposure to cobalt-chloride (CoCl₂) for different time periods. The cells were cultivated under normal conditions without CoCl₂ (S), with CoCl₂ for four days (4), with CoCl₂ for 8 days (8) and with CoCl₂ for 14 days (14). α -Tubulin serves as a loading control.

Figure 61 shows the quantitative analysis of the HIF-1 α expression levels determined by Western Blot analysis (**figure 60**) under exposure to cobalt for the different time points. There is an almost undetectable expression of HIF-1 α in the cell lines MCF-7 (**figure 61 B**) and BC-M1 (**figure 61 D**) and a weak expression in MDA-231 (**figure 61 A**) and MDA-231 BO2 (**figure 61 C**) under normal culture conditions without exposure to cobalt-chloride (0(S)).

Four days of exposure to CoCl₂ (4) induces the stabilization of HIF-1 α resulting in an elevation of the expression levels in MDA-231 by 168 %, in MCF-7 by 5765 %, in MDA-231 BO2 by 485 % and in BC-M1 by 1446 %. The expression level of HIF-1 α keeps stable for 8 days (8) and 14 days (14) with minor changes of the intensity for all analyzed cell lines except of MDA-231. After stabilization of HIF-1 α after four days (4) in MDA-231 the expression level decreases significantly (**table 6**) by 54 % after exposure for eight days (8) and by 38 % after 14 days (14) compared to the HIF-1 α expression levels after four days exposure to cobalt.

There is a significant change in the expression level of HIF-1 α between cultivation without cobalt (0 (S)) and 14 days of exposure to cobalt (14) for MCF-7, MDA-231 BO2 and BC-M1. For MDA-231 the expression difference does not change significantly between cultivation without cobalt (0(S)) and eight days or 14 days of exposure to cobalt. The stabilization of HIF-1 α mediated by CoCl₂ was successful.

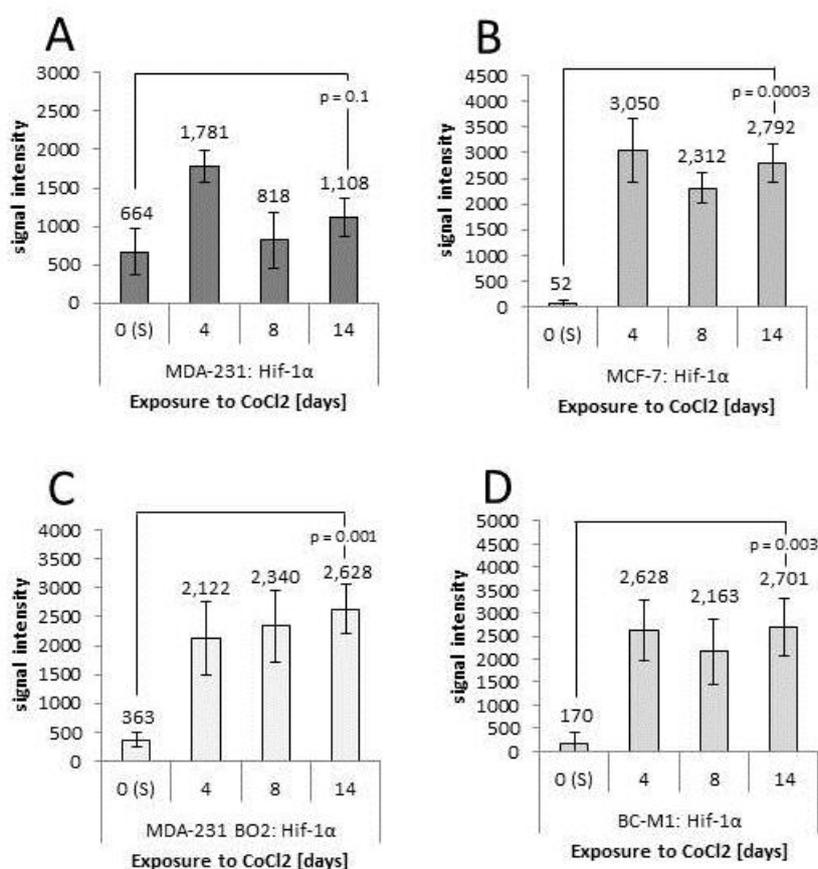


Figure 61: Quantification of the Western Blot analysis for the changes of Hif-1 α under exposure to CoCl₂ for the cell lines MDA-231 (A), MCF-7 (B), MDA-231 BO2 (C) and BC-M1 (D) and for different time periods. The level of significance (p) is shown for the difference between the state without CoCl₂ (0 (S)) and the exposure to CoCl₂ for 14 days (14). The signal intensity is represented in arbitrary units which were generated by the quantification software Quantity One.

Figure 62 shows the quantitative analysis of the CDCP1 expression levels determined by Western Blot analysis (**figure 60**) under exposure to cobalt for the different time points. MDA-231 exhibits a moderate expression of the full length variant of CDCP1 (**figure 62A**) and a strong expression of the cleaved variant of CDCP1 (**figure 62B**) under normal culture conditions (S). The expression level of the full length variant increases significantly (**table 6**) by 45 % after four days (4), by 64 % after eight (8) days and by 79 % after 14 days of exposure to cobalt (14) compared to standard cultivation. Whereas the cleaved variant of CDCP1 decreases by 7 % after four days, by 39 % after eight days and by 81 % after 14 days exposure to cobalt-chloride until it is hardly detectable by Western Blot analysis (**figure 60**). CDCP1 is not detectable in MCF-7 independent of the exposure to cobalt-chloride.

The full length variant of CDCP1 is detectable with a very weak signal in MDA-231 BO2 (**figure 62 C**) under normal culture conditions (S) whereas the cleaved variant is detectable with a moderate signal (**figure 62 D**). The expression level of the full length variant increases significantly for the time points of eight days (2746 %) and 14 days (8036 %) under exposure to cobalt compared to the initial state. The expression level of the cleaved variant of CDCP1 increases after four days by 49 % and decreases after eight days by 58 % and by 93 % after 14 days compared to the cultivation without cobalt.

BC-M1 exhibits the highest expression level of full length CDCP1 (**figure 62 E**) under normal culture conditions (S). BC-M1 does not express the cleaved variant of CDCP1, independent of the exposure to cobalt-chloride. The expression of the full length variant decreases after stabilization of HIF-1 α after four days by 51 %, by 21 % after eight days and by 58 % after 14 days compared to standard culture conditions. There is a significant change in the expression levels of the full length variant of CDCP1 for MDA-231 (increase), MDA-231 BO2 (increase) and BC-M1 (decrease) detectable between cultivation without cobalt (0 (S)) and 14 days of exposure to cobalt (14). The cleaved variant of CDCP1 shows a significant reduction in MDA-231 and MDA-231 BO2 after 14 days of exposure to cobalt compared to standard culture conditions.

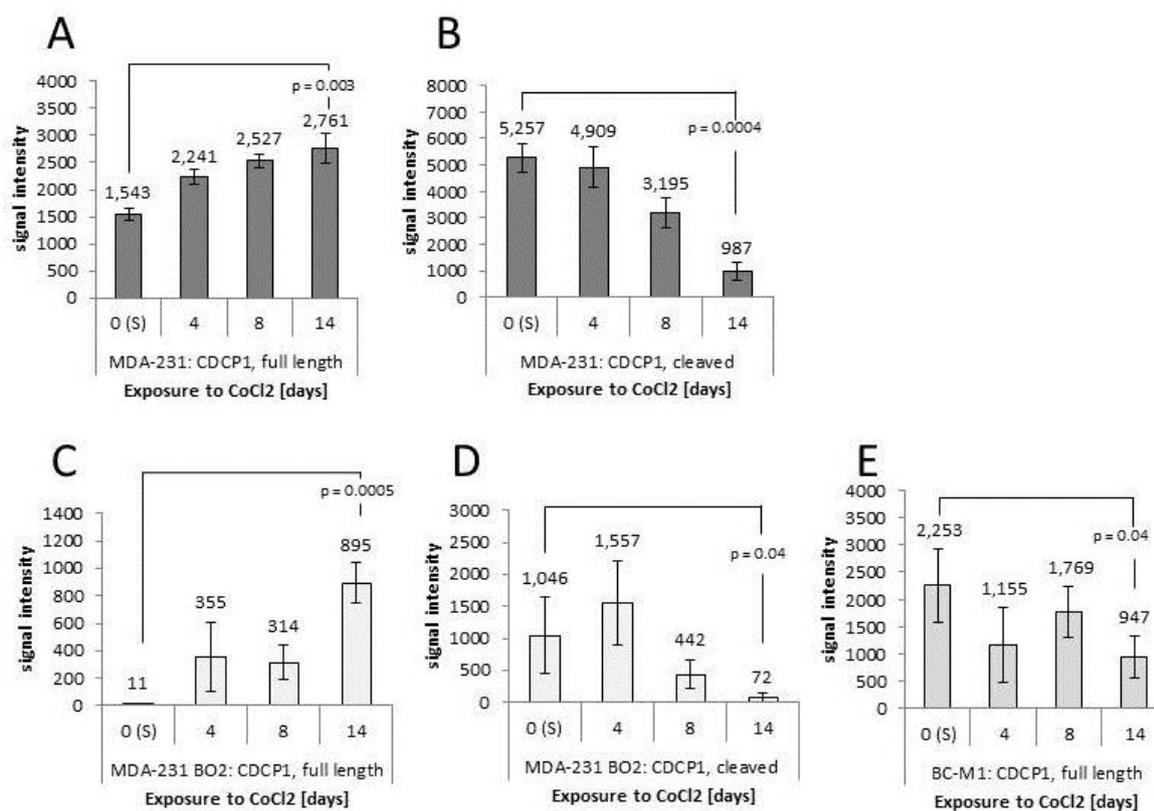


Figure 62: Quantification of the Western Blot analysis for the changes of the full length variant and the cleaved variant (cleaved) of CDCP1 under exposure to CoCl₂ for the cell lines MDA-231 (A), MCF-7 (B), MDA-231 BO2 (C) and BC-M1 (D) and for different time periods. The level of significance (p) is shown for the difference between the state without CoCl₂ (0 (S)) and the exposure to CoCl₂ for 14 days (14). The signal intensity is represented in arbitrary units which were generated by the quantification software Quantity One.

Figure 63 shows the quantitative analysis of the NRP-1 expression levels determined by Western Blot analysis (**figure 60**) under exposure to cobalt for the different time points. NRP-1 is detectable with a strong signal (**figure 63 A**) under normal culture conditions (S) in MDA-231. The expression level decreases by 23 % after four days, by 44 % after eight days and by 24 % after 14 days of exposure to cobalt compared with the initial state (0 (S)). All of these changes are significant.

NRP-1 is detectable with a weak signal (**figure 63 B**) in MCF-7 under normal culture conditions (0(S)). The NRP-1 expression levels slightly increase with increasing time of exposure to cobalt but these changes are not significant (**table 6**).

NRP-1 is detectable in MDA-231 BO2 with a strong signal (**figure 63 C**) under normal culture conditions (0 (S)). The expression decreases significantly by 39 % after four days and by 34 % after eight days of exposure to CoCl_2 . After 14 days the expression is slightly increased by 27 % compared to the cultivation without cobalt but the difference is not significant.

BC-M1 exhibits a high expression of NRP-1 (**figure 63 D**) under normal culture conditions (0(S)). The expression level decreases by 26 % after four days, by 39 % after eight days and by 27 % after 14 days of exposure to cobalt compared with the initial state (0 (S)). All of these changes are significant.

The expression differences of NRP-1 between the initial state (0 (S)) and cobalt-exposure for 14 days (14) are significant for the cell lines MDA-231 and BC-M1.

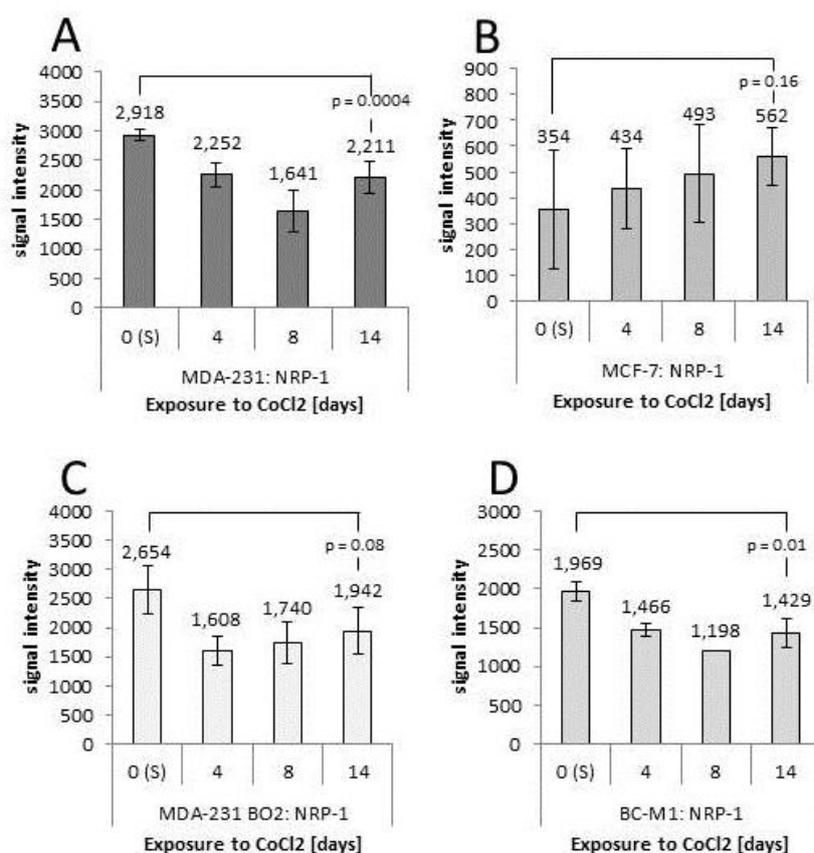


Figure 63: Quantification of the Western Blot analysis for the changes of the full length variant and the cleaved variant (cleaved) of NRP-1 under exposure to CoCl₂ for the cell lines MDA-231 (A), MCF-7 (B), MDA-231 BO2 (C) and BC-M1 (D) and for different time periods. The level of significance (p) is shown for the difference between the state without CoCl₂ (0 (S)) and the exposure to CoCl₂ for 14 days (14). The signal intensity is represented in arbitrary units which were generated by the quantification software Quantity One.

Figure 64 shows the quantitative analysis of the NRP-2 expression levels determined by Western Blot analysis (**figure 60**) under exposure to cobalt for the different time points. NRP-2 is not detectable in MCF-7 and MDA-231 BO2.

NRP-2 is detectable with a weak signal (**figure 64 A**) in MDA-231 under normal culture conditions (0(S)). The expression level decreases by 28 % after four and eight days and by 32 % after 14 days of exposure to cobalt compared to the initial state (0 (S)) but these changes are not significant (**table 6**).

BC-M1 exhibits a high expression of NRP-1 (**figure 64 B**) under normal culture conditions (0(S)). The expression level decreases by 55 % after four days, by 67 % after eight days and by 84 % after 14 days of exposure to cobalt compared with the initial state (0 (S)). All of these changes are significant.

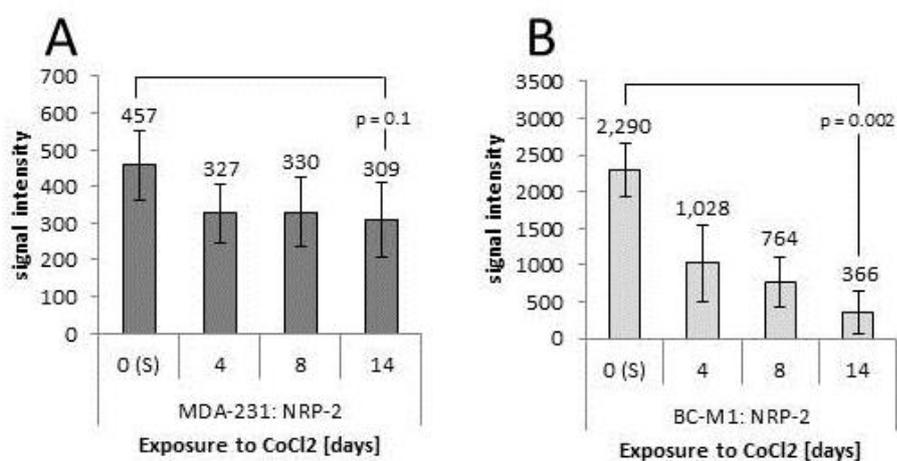


Figure 64: Quantification of the Western Blot analysis for the changes of NRP-2 under exposure to CoCl₂ for the cell lines MDA-231 (A) and BC-M1 (B) for different time periods. The level of significance (*p*) is shown for the difference between the state without CoCl₂ (0 (S)) and the exposure to CoCl₂ for 14 days (14). The signal intensity is represented in arbitrary units which were generated by the quantification software Quantity One.

Table 7: Levels of significance (p-value) for the expression level differences of the biomarkers under exposure to cobalt-chloride for the combination of the different time points. 0 (S)/4 signifies the difference between the exposure without CoCl₂ (0 (S)) and four days exposure to CoCl₂ (4).

MDA-231	0 (S)/4 ^a	0 (S)/8 ^a	0 (S)/14 ^a	4/8 ^a	4/14 ^a	8/14 ^a
CDCP1, fl	0.0031	0.0006	0.0026	0.0495	0.0339	0.1728
CDCP1, cleaved	0.3126	0.0104	0.0004	0.0302	0.0012	0.0043
NRP-1	0.0076	0.0077	0.0004	0.1680	0.4032	0.1644
NRP-2	0.1034	0.1221	0.0996	0.4851	0.4274	0.4189
Hif-1	0.0065	0.3334	0.0925	0.0154	0.0222	0.2001
MCF-7	0 (S)/4	0 (S)/8	0 (S)/14	4/8	4/14	8/14
NRP-1	0.3544	0.2752	0.1594	0.3760	0.1996	0.3409
Hif-1	0.0012	0.0002	0.0003	0.1006	0.3203	0.2408
MDA-231 BO2	0 (S)/4	0 (S)/8	0 (S)/14	4/8	4/14	8/14
CDCP1, fl	0.0605	0.0139	0.0005	0.4218	0.0283	0.0065
CDCP1, cleaved	0.2295	0.1257	0.0415	0.0424	0.0165	0.0479
NRP-1	0.0193	0.0398	0.0799	0.3930	0.1879	0.3135
Hif-1	0.0092	0.0057	0.0010	0.3729	0.2013	0.3076
BC-M1	0 (S)/4	0 (S)/8	0 (S)/14	4/8	4/14	8/14
CDCP1, fl	0.0919	0.2284	0.0390	0.1789	0.3637	0.0656
NRP-1	0.0048	0.0006	0.0139	0.0040	0.4022	0.0749
NRP-2	0.0233	0.0062	0.0021	0.2896	0.1403	0.1403
Hif-1	0.0038	0.0097	0.0028	0.2662	0.4573	0.2313

^a p-values calculated by Student *t* test

The biomarker candidates CDCP1, NRP-1 and NRP-2 exhibit detectable expression levels after all analyzed time periods for the exposure to cobalt determined by Western Blot analysis. However, NRP-2 was significantly reduced only in BC-M1 after the exposure to cobalt but remains detectable by Western Blot analysis. The expression level of NRP-2 in MDA-231 was hardly detectable after exposure to cobalt for 14 days.

The cell lines exposed to CoCl_2 were morphologically analyzed by brightfield microscopy (**figure 65**).

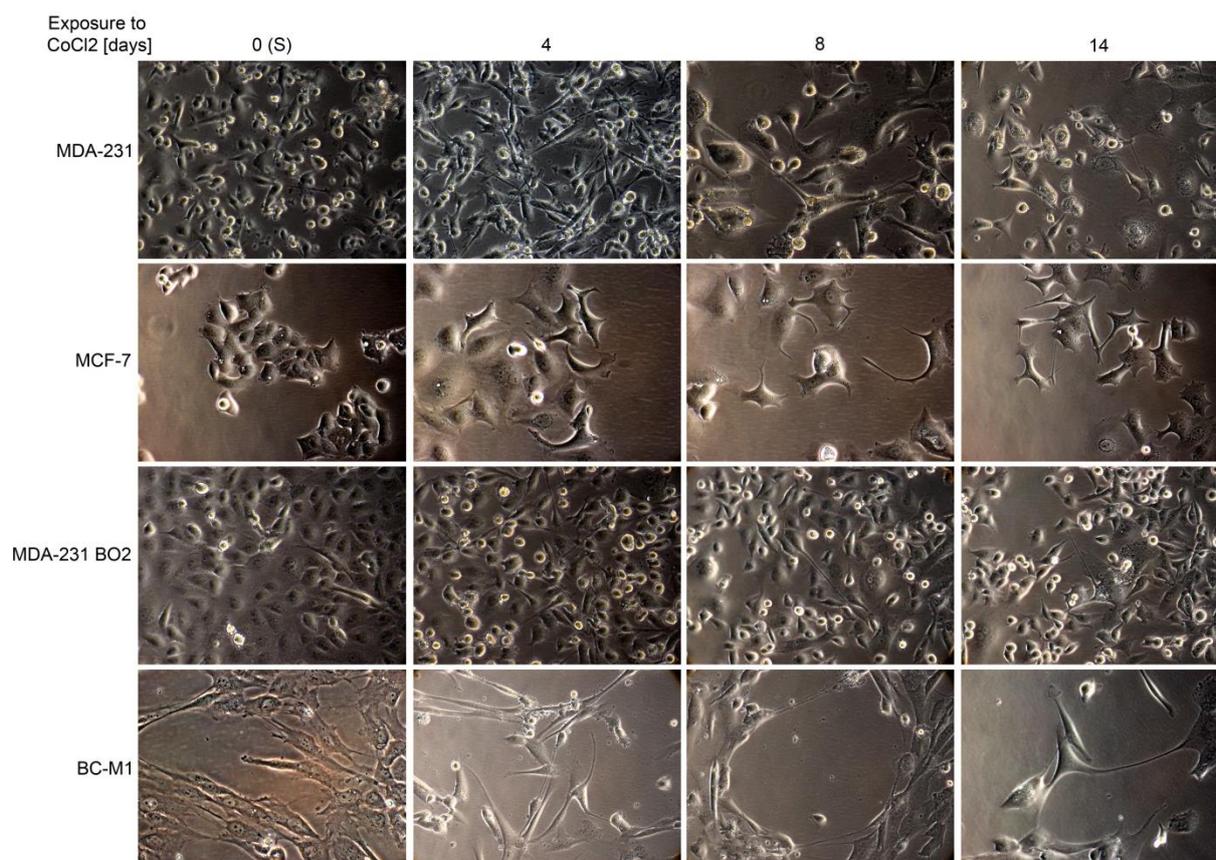


Figure 65: Morphological analysis of the cell lines MDA-231, MCF-7, MDA-231 BO2 and BC-M1 under standard cultivation conditions (0 (S)) and different time periods under exposure to CoCl_2 by brightfield microscopy. The images were taken with 320-fold magnification.

Each cell lines show morphological changes under the exposure to CoCl_2 . The spindle-shaped form of MDA-231 under cultivation without cobalt becomes more pronounced under the exposure to cobalt for four days (4). Additionally, the cells form straight and long protrusions. After eight (8) and 14 days (14) of exposure to cobalt the confluency decreases and the length of the protrusions increases. Furthermore these protrusions mediate cell-cell connections.

MCF-7 cells grow in close cell-cell contact with a round to polygonal shape under cultivation without cobalt (0 (S)). After four days of exposure to cobalt (4) the amount of cell-cell contacts decreases and the morphology becomes more spindle-shaped-like with short cell protrusions. After eight days (8) and 14 days of exposure to cobalt (14) the confluency decreases. Additionally the amount of cell-cell contacts decreases and there are more single cells visible. The spindle-shaped form becomes more pronounced and the protrusions become longer and mediate cell-cell connections.

MDA-231 BO2 grows in close cell-cell contact with a shape between round and polygonal under cultivation without cobalt (0 (S)). The morphology is heterogeneous. After four days of exposure to cobalt (4) the amount of cell-cell contacts decreases and the cells develop short cell protrusions. After eight days (8) and 14 days of exposure to cobalt (14) the confluency and the cell-cell contacts decrease and the morphology becomes more spindle-shaped-like with longer cell protrusions mediating cell-cell connections.

BC-M1 grows with the lowest confluency compared to the cell lines MDA-231, MDA-231 BO2 and MCF-7. BC-M1 cells exhibit a spindle-like and filamentous shape under cultivation without cobalt (0 (S)). They grow in small groups with areas without cells between the cell groups and have short protrusions. After exposure to cobalt the amount of cell-cell contacts decreases and there are more single cells with a more spindle-like shape visible. The amount and the length of the protrusions mediating cell-cell contacts increases.

CDCP1, NRP-1, NRP-2 and DCBLD2 were identified as potential biomarker candidates by SILAC-MS based proteome analysis. The expression of these proteins on breast cancer cell lines correlated with mesenchymal or hybrid epithelial/mesenchymal properties. For the use as diagnostic biomarker for breast cancer, the lack of these proteins (except for NRP-2) on PBMC of healthy donors was determined by Western Blot analysis and an immunofluorescence staining was established for CDCP1. Additionally, a MACS-based isolation method was established for the isolation of CDCP1⁺ breast cancer tumor cell line MDA-231 and CDCP1 was detected on CTCs from breast cancer patients with the CellSearch System. Furthermore, a Sandwich-ELISA was established for detection of CDCP1 in the blood plasma of breast cancer patients and CDCP1 was detected in the blood plasma 22 % of the analyzed breast cancer patients. Additionally, it was shown that the expression of the marker candidates remains stable in most of the analyzed cell lines after exposure to hypoxia or after hypoxia-simulation by cobalt-chloride. The following chapter deals with the evaluation of the results and the classification into the current state of knowledge.

6. Discussion

Even though early detection of breast cancer (BC) made progress in the past, it is still challenging to establish a simple and reliable test that is applicable to the clinic [155]. These cancer-related deaths are often caused by metastatic expansion into distant organs (e.g. the bone marrow) [44, 155-159]. A study by Braun *et al.* [160] showed that the presence of micrometastasis in breast cancer was correlated with the cancer-related death in 25 % of the patients. Otherwise some patients did not survive while metastatic cells (micrometastasis) could not have been detected [74, 157]. Moreover, a chemotherapy-resistant DTC subpopulation with stem cell attributes might play a crucial role in the formation of metastasis [57]. These cells exhibit mesenchymal properties with an altered epithelial protein expression profile [35, 51, 57, 69]. The downregulated expression of epithelial proteins like EpCAM or the modified expression pattern of cytokeratins results in the hindered detection and isolation of these mesenchymal cells [136] and leads to a “detection gap” for this type of cells. Thereby these cells lack epithelial proteins like cytokeratins and EpCAM [161] which are necessary for conventional detection and enrichment techniques.

First, a suitable model for the proteome comparison analysis had to be defined. Therefore a set of different breast cancer cell lines and DTC cell lines from different entities were investigated for their grade of epithelial differentiation as well as for their expression status of receptor tyrosine kinases EGFR/ErbB-2 and the corresponding signaling pathways which are thought to contribute to malignant transformation [162]. A suitable model is represented by a macrometastatic cell line compared with a cell line which originates from a micrometastatic state. The cell lines were characterized for a set of cancer-related proteins.

The cell lines were investigated for their CD44/CD24 phenotype. The expression pattern CD44⁺/CD24 is connected with properties of breast cancer stem cells/progenitor cells (BCSC) like the ability of self renewal, formation of new tumors and therapy resistance [163]. CD44⁺/CD24^{-/low} usually correlate with mesenchymal phenotype [164]. CD24 is thought to be involved in the transition between EMT and MET [164]. CD44 plays an important role in tumor progression, metastasis formation [165] and adhesion to the extracellular matrix via binding to its ligand hyaluronic acid [166]. The unusual band at 130 kDa of CD44 in MDA-468 might be caused by an isoform CD44v [167]. The DTC cell lines as well as Hs578t, MDA-231

and the corresponding metastatic sublines are CD44⁺/CD24⁻ whereas MCF-7 exhibits a CD44⁻/CD24⁺ phenotype and MDA-468 CD44⁺/CD24⁺.

The DTC cell lines as well as Hs578t, MDA-231 and the corresponding metastatic sublines are positive for Vimentin and for N-Cadherin except for MDA-231 which is negative for Vimentin. The cell lines MCF-7 and MDA-468 are negative for the expression of Vimentin and N-Cadherin. Vimentin plays an important role in cell-ECM-contacts [168], whereas N-Cadherin seems to be involved in mediating cell-cell-connections and associated with a fibroblast-like phenotype [169]. The N-Cadherin positive cell lines show a fibroblast-like or spindle-shaped morphology. In general, the expression of Vimentin and N-Cadherin is connected to the transition from epithelial to mesenchymal state (EMT) [170].

The DTC cell lines as well as Hs578t and MDA-231 are negative for EpCam and E-Cadherin. A very weak expression of EpCAM and no detectable expression of E-Cadherin was determined in MDA-231 SA and MDA-231 BO2. The expression of these proteins signifies an epithelial phenotype, especially E-Cadherin mediates connections between cells and their environment [1]. These findings correlate with the EpCam/E-Cadherin expression profile and the growth properties of MCF-7/MDA-468 (tight cell-cell contacts) and less cell connections in MDA-231 and the DTC cell lines.

With regard to the cytokeratins as epithelial markers, CK5 is expressed in cells of basal epithelial differentiation [171] and basal-like triple negative breast cancer cells like MDA-468 [172] which could be verified by Western Blot analysis (**figure 5**). The cytokeratins 8/18/19 characterize cells of luminal (glandular) differentiated origin [173]. The cell lines MDA-231 and MCF-7 show a luminal differentiation state (CK8/18/19⁺, CK5⁻) whereas MDA-468 exhibits a basal-like differentiation state because of the expression of CK5 (basal) and additional expression of the luminal cytokeratins 8/18/19. The absence of EpCam, CK5/8/18/19 underlines the low degree of differentiation and cancer stem/progenitor cell properties of Hs578t and the DTC cell lines BC-M1, LC-M1 and PC-E1 [174]. In summary, BC-M1 and MDA-231 BO2 have a low degree of epithelial differentiation with high expression levels of mesenchymal marker proteins. Additionally, these two cell lines exhibit a CD44⁺/CD24⁻ expression pattern which is thought to be typical for cancer stem cells [163].

In addition to the degree of epithelial differentiation, the presence of receptor tyrosine kinases allow a characterization of the investigated cell lines [175, 176]. An EGFR expression was detected in the cell line MDA-468 with the strongest signal. The overexpression is caused by EGFR gene amplification [177].

MDA-231 BO2 exhibits a moderate EGFR expression level which is elevated compared with its parental cell line MDA-231. Moreover, ErbB-2 was not detected in BO2, so that MDA-231 BO2 can be considered as an EGFR-driven phenotype. In contrast, BC-M1 exhibits an ErbB-2/ErbB-3-driven phenotype [178] and was determined as EGFR⁺ by Western Blot analysis.

For the evaluation of the phenotypic behaviour, the analyzed cell lines were investigated for three essential signaling pathways.

The Akt-pathway is a complex signaling pathway that plays an important role in breast cancer cell growth and proliferation and is thought to mediate resistance to several therapies. Among other activators, receptor tyrosine kinases (RTK) EGFR and ErbB-2 mediate the downstream signaling after stimulation with growth factors. [179]. The phosphoinositide 3 kinase (PI3K) which is downstream of ErbB-2 [180] is significantly involved in the activation of protein kinase B (Akt) by phosphorylation [179, 181]. Moderate levels of activated (phosphorylated at Ser473) Akt (pAkt) were detected in BC-M1 and the DTC cell lines, whereas the bone metastatic sublines MDA-231 SA and MDA-231 BO2 were negative for pAKT in Western Blot analysis.

The Erk1/2 (Mitogen-activated protein kinase 3/ Mitogen-activated protein kinase 1) including MAP Kinase pathway is the most relevant in breast cancer and ends in the control of cell proliferation and apoptosis [182]. This pathway can be activated by growth factors like EGF, IGF, Insulin and TGF- β mediated by receptor tyrosine kinases [182]. Both kinases Erk1/2 are activated through phosphorylation at threonine-202 and tyrosine-304 residue by other kinases [183]. Activated Erk2 (pErk2) was detectable in every cell line with the exception of Hs578t and LC-M1. Phosphorylated Erk2 was detectable in MDA-231, MDA-231 SA, MDA-231 BO2, BC-M1 and PC-E1 with decreasing intensity in this order.

STAT3 belongs to a family called signal transducers and transcription factors and is activated and present in a large number of breast cancers [184]. Transcriptional activity like cellular transformation, tumor initiation and tumor progression is mediated by STAT3 due to external signals from cytokines and growth factors to their corresponding receptors [185]. STAT 3 is activated through phosphorylation of a tyrosine-705 and serine-727 residue but the unphosphorylated STAT3 might have transcriptional and regulatory influence as well [184]. Phosphorylated STAT3 (pSTAT3) was detectable in MDA-231, MDA-231 BO2 and MCF-7 with a very low intensity. The highest expression levels of pSTAT3 were detected in Hs578t, MDA-468 and LC-M1. MDA-231 SA, BC-M1 and PC-E1 exhibit moderate expression levels of pSTAT3.

HIF-1 α (hypoxia inducible factor 1 α) is an oxygen-regulated transcriptional activator which is stabilized under hypoxic conditions and activates target genes for angiogenesis, cell proliferation, survival and glucose metabolism [65, 186, 187]. High expression levels of HIF-1 α under atmospheric oxygen concentrations were detected in the DTC cell line LC-M1 and MDA-231 SA whereas MDA-231 BO2 was negative and BC-M1 exhibits a weak expression of HIF-1 α . Usually HIF-1 α stabilization is caused by low oxygen concentrations [186] but hypoxia is not the only impulse for HIF-1 α induction [186, 188] in breast cancer. The synthesis of HIF-1 α can also be stimulated by growth-factors and cytokines in an oxygen-independent manner [83]. Additionally, HIF-1 α was associated with the involvement in resistance in prostate cancer under normoxic conditions [189] and reactive oxygen species (ROS) might play a role in HIF-1 α regulation. These observations might explain the determined expression levels of the breast cancer and DTC cell lines by Western Blot analysis under normal oxygen conditions.

The C-X-C chemokine receptor type (CXCR4) is a chemokine receptor specific for the ligand stromal-derived-factor-1 (SDF-1) [190]. This receptor-ligand complex plays a crucial role in cell survival, proliferation, chemotaxis, migration, adhesion and CXCR4 is expressed in most breast cancers [190, 191]. The most important property of CXCR4 is the potential involvement in metastatic spread to distant organs (liver, lung, bone marrow) in breast cancer because the ligand SDF-1 is expressed in the infiltrated organs [190, 192]. CXCR4 was detected in every analyzed breast cancer cell line and DTC cell line with the lowest amount in Hs578t suggesting that all analyzed cell lines might seed metastasis in distant organs with a lower probability for the cell line Hs578t.

To identify novel marker proteins for the early detection of breast cancer a SILAC-MS/MS based proteome analysis between BC-M1 and MDA-231 BO2 was performed to identify potential marker proteins. BC-M1 was isolated from the bone marrow of a metastatic breast cancer patient and was detected as an EGFR⁻/cytokeratin⁻/EpCam⁻/ErbB-2^{low}/CXCR4⁺/Vimentin⁺ phenotype [193, 194]. MDA-231 BO2 is a bone metastatic subline of the parental cell line MDA-231 and was determined as an EpCam^{low}/cytokeratin^{low}/CXCR4⁺/Vimentin⁺/ErbB-2⁺/EGFR⁺⁺ phenotype [195-197].

Before potential marker candidates were chosen, in a first step, the data of the mass spectrometric analysis was validated. Therefore five cancer related proteins (**table 1**) were validated by Western Blot analysis (**figure 17**) and the expression levels were compared with the signal ratios (**table 1**) determined by the MS1 spectra for Connexin-43 (**figure 7**), Grp78 (**figure 9**), CD44 (**figure 11**), Integrin alpha 5 (**figure 13**) and

EGFR (**figure 15**). The comparison of the signal ratios and the expression levels determined by Western Blot analysis confirmed the reliability of the MS data which is the basis for further analysis. The requirements for the choice of potential marker candidates are characterized by the expression in BC-M1, the existence of an extracellular domain and the identification by at least two unique peptides.

The mass spectrometric based analysis of the two compared proteomes of the bone marrow DTC cell line and the bone metastatic cell line MDA-231 BO2 revealed 785 differentially expressed proteins. Among the proteins overexpressed in BC-M1, CDCP1, NRP-1, NRP-2 and DCBLD2 were the most promising marker candidates after database research and verification of defined criteria. NRP-2 was additionally chosen because this protein fulfilled the same requirements and belongs to the neuropilin-family. The potential marker candidates were identified by MS2-spectra (CDCP1: **figure 19**, NRP-1: **figure 21** and DCBLD2: **figure 23**). The four candidates were validated by Western Blot analysis (**figure 24**) and the expression levels of CDCP1, NRP-1 and DCBLD2 were compared with the signal ratios (**table 2**) determined by the MS1 spectra for CDCP1 (**figure 18**), NRP-1 (**figure 20**) and DCBLD2 (**figure 22**). The determined signal ratios (**table 2**) of the MS analysis correlated with the expression levels determined by Western Blot analysis (**figure 24**).

All marker candidates were detected on the DTC cell lines of patients from breast-, lung- and prostate-cancer (BC-M1, LC-M1, PC-E1) by Western Blot analysis. This highlights the potential role of the marker candidates in the process of tumor cell dissemination.

The expression of CDCP1 was observed in MDA-231, MDA-468, MDA-231 SA and MDA-231 BO2 as well as in the circulating tumor cell lines CTC-ITB001 und CTC-ITB001-MIND (**figure 24 and 26**). The absence of CDCP1 in the CD44⁺/CD24⁻ CSC cell line Hs578t shows that there might be no direct connection between cancer stem cell properties and CDCP1. The strong expression of CDCP1 in MDA-231 might support the role of CDCP1 in the process of cellular migration, adhesion as well as anchorage-independent growth [198, 199] and thereby suggesting a potential role in the process of EMT. The presence of CDCP1 in the triple-negative breast cancer cell lines (MDA-468, MDA-231) is a promising result because CDCP1 might function as a therapy target for badly treatable triple-negative tumors lacking therapeutic targets like estrogen-, progesterone and ErbB-2 receptor [32, 200]. The determination of the expression of CDCP1 in the highly invasive breast cancer cell line MDA-231 is a promising finding supporting the connection between aggressiveness and invasiveness of cancer cells expressing the cleaved

variant of CDCP1 [92, 121]. The clinical relevance in the detection of circulating tumor cells might be supported by the determination of CDCP1 in the breast cancer patient CTC cell line CTC-ITB001 and CTC-ITB001-MIND. Additionally, the cleaved membrane-bound variant of CDCP1 was detected on CTC-ITB001 as well as in CTC-ITB001-MIND with a slightly higher expression. This fact supports the hypothesis that cleaved membrane-bound CDCP1 fragment might correlate with the property of tumor-initiating potential and invasiveness [92, 121].

The cell surface protein NRP-1 was detected in every cell line except of MCF-7. The determination of the expression in the DTC cell lines highlights a possible involvement in the dissemination of tumor cells. The determination of NRP-1 in the CSC cell line Hs578t supports a feasible connection between this protein and cancer stem cell attributes [201, 202]. The high expression levels of NRP-1 in the hybrid EMT cell line MDA-231 could propose an influence in the complex process of EMT [203].

The highest expression levels of NRP-2 were detected in the DTC cell lines BC-M1, LC-M1, in the breast cancer cell line Hs578t and in the bone metastatic cell line MDA-231 SA. Moderate levels were determined in the DTC cell line PC-E1. The strong expression in BC-M1 might suggest a potential function in tumor cell dissemination and the additional strong expression in the CSC cell line Hs578t suggests a feasible involvement in promotion of stem cell like attributes in breast cancer cells [204].

The presence of DCBLD2 was verified in all analyzed cell lines except of MCF-7 and MDA-468. The strong expression in the DTC cell lines might correlate with an involvement in tumor cell dissemination in tumors from different entities. The expression status of DCBLD2 was reported to correlate with the enhancement of cell motility in lung cancer cells [205]. Cell motility is an important property of the EMT process which could explain the expression of DCBLD2 in the hybrid EMT cell line MDA-231.

Next, the utility of CDCP1, NRP-1, NRP-2 and DCBLD2 as CTC/DTC detection marker proteins was investigated. Therefore the expression of the candidate markers was investigated on PBMC of healthy female donors by Western Blot analysis. There was no expression of CDCP1, NRP-1 or DCBLD2 on PBMC of healthy female donors detectable. Furthermore, CD45⁺/CDCP1⁺ PBMCs or CD45⁺/NRP-1⁺ PBMCs could not be detected with immunofluorescence analysis which supports the finding of the absence of CDCP1, NRP-1 and DCBLD2 on PBMC analyzed by Western Blot. These are necessary requirements for the use as a diagnostic biomarker to distinguish CDCP1⁺ tumor cells and NRP-1⁺ tumor cells from normal PBMCs. A typical membrane staining pattern for CDCP1 was observed in particular for the BC-

M1 and MDA-231 BO2. The immunofluorescent analysis revealed a membrane staining pattern for NRP-1 in the DTC cell line BC-M1 as well as in MDA-231. The results of the immunofluorescence show that CDCP1 and NRP-1 are suitable for the detection of CTC/DTC with mesenchymal attributes of breast cancer patients.

NRP-2 was not further investigated as a detection marker after the determination of a strong signal (below 70 kDa) in PBMC of healthy female donors. The origin of this band might be caused by the expression of NRP-2 on PBMC of healthy female donors by cross-reaction of the anti-NRP-2 antibody with another protein expressed in PBMC. It was shown that NRP-2 is expressed during the differentiation of macrophages which might explain the signal of NRP-2 in PBMC of healthy donors [206]. It was shown that the marker candidates CDCP1 and NRP-1 could serve as CTC/DTC detection markers due to the lack of their expression on PBMC of healthy donors and the established immunofluorescent staining. DCBLD2 might also function as a detection marker for CTC/DTC but an immunofluorescent staining could not be established for this antibody. However, the lack of the expression of DCBLD2 on PBMC of healthy donors shows the suitability of this protein for detection applications.

For the improvement of isolation techniques for mCTCs, emCTCs or tumor or cancer cells with tumor initiating potential (CSC) a MACS-based isolation system to capture CDCP1-expressing tumor cells was established. This approach might be suitable for the detection of early disseminating tumor cells with increased metastatic potential [41, 62]. For previous testing of the suitability for the MACS-system, the unconjugated CDCP1-antibody was validated with MDA-231 cells by FACS analysis. The applied antibody detected 37.8% of the viable and unfixed MDA-231 cells by FACS analysis. The moderate recovery rate could be explained by the losses during secondary antibody reaction. The negative control, MCF-7, was detected with 0.5 % which could be explained by unspecific reactions or by the ability of the antibody to bind cells expressing a low number of antigen molecules.

After successful validation of the unconjugated CDCP1-antibody, CDCP1-dependent MACS isolation of MDA-231 cells was performed. Improvements of the protocol resulted in the increase of an average recovery rate of 55%. That means that more than half of the hybrid EMT MDA-231 cells could be isolated. This is a promising result compared to the average recovery rate of 41 % of MDA-231 cells with the CellSearch System, the diagnostic standard for CTC detection [207]. A directly magnetic-bead labeled CDCP1-antibody could potentially increase the average recovery rate by avoiding the secondary reaction

between biotinylated CDCP1-antibody and Streptavidin-conjugated magnetic beads. Similar EpCAM-based isolation methods were less effective. For example, MDA-231 cells were isolated by EpCAM-based immunomagnetic enrichment with a recovery rate of 0.23 % [208] whereas the FDA approved EpCAM-dependent enrichment system CellSearch yielded a recovery rate of MDA-231 from 12 % [209] to 40.6 % [207]. These findings indicate that CDCP1-isolation of hybrid EMT breast cancer cells is more efficient than EpCAM-dependent isolation regardless of the applied method.

For the detection of CDCP1-expressing CTCs in breast cancer patients using the CellSearch System, the CDCP1-antibody was proved to function in a specific and sensitive manner prior to the use in the CellSearch device. Therefore the directly fluorophore (Alexa588)-conjugated CDCP1-antibody was first established in the FACS analysis to determine the optimal parameters. The previous testing was performed with MDA-231 cells as a positive control and MCF-7 cells as a negative control. The recovery rate for CDCP1-positive MDA-231 cells was 97.5 % whereas the recovery rate for CDCP1-negative MCF-7 cells was 0.07 %. The high recovery rate for MDA-231 indicates the efficiency of the fluorophore-conjugated CDCP1 antibody to recognize the epitope in unfixed cells. This high recovery rate also supports the thesis that the CDCP1-based MACS isolation could be improved by the application of a directly magnetic-bead conjugated CDCP1-antibody.

After validation of the fluorophore-conjugated CDCP1-antibody by FACS analysis, the parameters for the use in the CellSearch system were determined by the analysis of a strong CDCP1-positive cell line (MDA-231), a weakly CDCP1-positive cell line (MDA-468) and a CDCP1-negative cell line (MCF-7). After optimization of the protocol, breast cancer patients were analyzed for the presence of CDCP1-positive CTCs with the CellSearch system. CTCs were detected by cytokeratin staining and application of CD45 as an exclusion marker after EpCam-dependent enrichment and CTCs were analyzed for CDCP1. 41 % of the analyzed patients (17 out of 41 patients) contained CK⁺/CD45⁻ CTCs. CDCP1⁺/CK⁺/CD45⁻ CTCs were detected in four percent of the patients containing CK⁺/CD45⁻ CTCs after EpCAM-dependent enrichment. The low detection rate of CDCP1⁺/CK⁺/CD45⁻ CTCs (4 %) might be caused by the EpCAM-based enrichment. It was shown that the highest expression levels of CDCP1 (**figure 24**) were detected in the EpCAM-negative DTC cell lines and the hybrid EMT cell line MDA-231 (**figure 5**). These findings and the fact that CSC cell line Hs578t is negative for CDCP1 suggest a correlation between CDCP1 and a mesenchymal or hybrid EMT phenotype and to a lesser extend with cancer stem cell features. The determination of CDCP1 in mesenchymal tumor cells or tumor cells that have undergone EMT could

explain the low detection rate of CDCP1⁺/CK⁺/CD45⁻ CTCs [210] and the decreased isolation efficiency of EpCAM^{low/-} CTCs [211]. An analysis of metastatic breast cancer patients revealed that 16 % of detected CTCs expressed epithelial markers and 22 % of these epithelial CTCs co-expressed mesenchymal markers [212]. That means that the percentage of this epithelial/mesenchymal CTC subpopulation is 3.5 % of all detected CTCs. These findings suggest that the CDCP1⁺/CK⁺/CD45⁻ CTCs detected by the CellSearch System might belong to the rare subgroup of CTCs with a partial EMT phenotype (**figure 66**). The mesenchymal CTC-subpopulation exhibits too low expression levels of EpCAM/CK which might lead to the loss of the detection possibility for this CTC subgroup [213]. A CDCP1-dependent enrichment with the CellSearch System or a CDCP1-dependent MACS isolation might improve the detection possibilities of the CTC subgroups which are not detectable with an EpCAM-dependent isolation method.

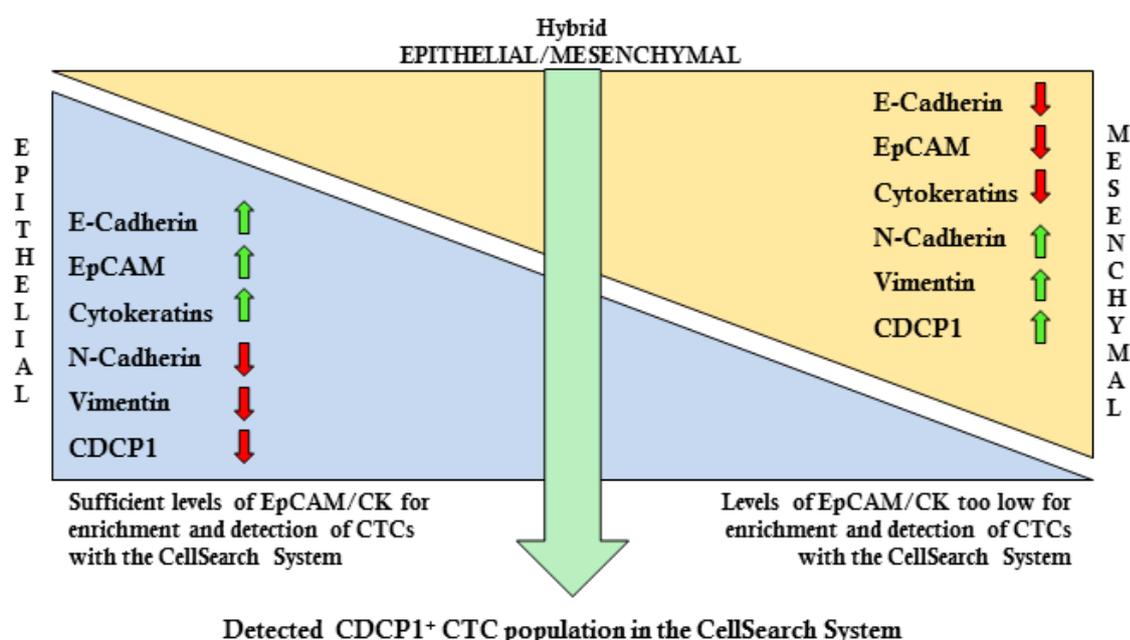


Figure 66: Suggested hypothetical model for the CDCP1⁺/CK⁺/CD45⁻ CTC subpopulation in the blood of metastatic breast cancer patients detected by EpCAM-dependent enrichment with the CellSearch System. The blue triangle represents tumor cells with an epithelial phenotype expressing the epithelial markers E-Cadherin, EpCAM and cytokeratins and lacking the mesenchymal markers N-Cadherin, Vimentin and the mesenchymal biomarker candidate CDCP1. The yellow triangle represents tumor cells with a mesenchymal phenotype expressing the mesenchymal markers N-Cadherin, Vimentin and CDCP1. For this subgroup, the expression levels of EpCAM and cytokeratins might be too low for enrichment and detection. The green arrow represents the suggested subgroup of CDCP1⁺/CK⁺/CD45⁻ CTCs.

The full length variant of CDCP1 can be cleaved proteolytically by different proteases (e.g. plasmin, trypsin) resulting in the generation of a 70 kDa membrane-bound fragment and a 65 kDa soluble variant [121]. The 70 kDa membrane-bound fragment is thought to be connected with increased invasiveness and aggressiveness by the formation of heterodimers with Integrin β 1. These biological properties can be very

important and necessary for the evaluation of the behavior of CTCs found in the blood of cancer patients and to gain information about therapy decision, treatment choice or disease progression. The existence of the cleaved membrane-bound 70 kDa-CDCP1-variant was observed in a set of breast cancer cell lines and the DTC cell line of a lung cancer patient by Western Blot analysis. Therefore it was assumed that the cell culture supernatant of cell lines exhibiting the cleaved, membrane-bound variant CDCP1 should contain the 65 kDa soluble fragment of CDCP1. For the establishment of a CDCP1-ELISA, the cell culture supernatant of MDA-231 was used as a positive control whereas the cell culture supernatant of MCF-7 was used as a negative control. The strong expression of the cleaved membrane-bound fragment of CDCP1 in MDA-231 was assumed to correlate with a high concentration of the soluble fragment of CDCP1 in the cell culture supernatant of this cell line. Therefore a Sandwich-ELISA was established which is capable of measuring the amount of the soluble CDCP1 fragment either in the supernatant of cancer cell or in the blood of breast cancer patients.

After establishment of an optimized protocol for the CDCP1-ELISA an internal control was established. A part of the extracellular domain of recombinant human-CDCP1 (Phe30-Ser341) was used as the internal control. The epitope of both applied ELISA antibodies were included in the sequence of the recombinant CDCP1. The capturing antibody as well as the detection antibody detected decreasing signal intensities with applied decreasing concentrations of recombinant CDCP1 at molecular weight at about 60 kDa. Both antibodies were not able to detect concentrations below 5 ng. The capturing antibody detected additional bands at a molecular weight of 130 kDa. The origin of this band could be caused by dimerization because the signal intensity also decreases with descending concentration and the molecular weight is twice as much as the expected molecular weight whereas the detection antibody recognizes a single band at the expected molecular weight of recombinant CDCP1. To guarantee the reliability of the Sandwich-ELISA blood plasma was analyzed with both applied antibodies by Western Blot analysis. This was a necessary test to exclude unspecific reactions with components of the blood plasma which might cause false-positive results. The Western Blot analysis with the capturing antibody revealed no detectable signals in the blood plasma of healthy female donors. The Western Blot analysis of blood plasma of healthy female donors with the detection antibody revealed weak signals at a molecular weight of 130 kDa and between 35 kDa and 55 kDa. These unspecific reactions with components of the blood plasma might be caused by the polyclonality of the detection antibody. The reliability of the Sandwich-ELISA should not be affected because the monoclonal capturing antibody does not bind unspecific products which the detection antibody

could potentially recognize. Another important parameter is the stability of CDCP1 in blood plasma of breast cancer patients regarding the sample processing. It was shown by Sandwich-ELISA that CDCP1 is stable for at least 48 h at 25 °C, for at least 2.5 h at 37 °C or for at least 96 h at 4 °C. These findings indicate that CDCP1 is at least stable for the time of sample processing and preparation for the Sandwich-ELISA. That fact that CDCP1 is stable for at least 2.5 h at 37 °C suggests a longer lifetime in the blood circulation of breast cancer patients which is an important parameter for the analysis and the informative value.

In MDA-468 the cleaved membrane-bound variant of CDCP1 was detected by Western Blot analysis whereas the corresponding soluble fragment could not be detected with an expected signal intensity by Sandwich-ELISA. Conceivably, MDA-468 produces a different cleavage product influencing the integrity of the epitope. This result could also be explained by an alternating glycosylation pattern and thereby masking the epitope. On the contrary, CDCP1 could be detected in the cell culture supernatant of the BC-M1 whereas in Western Blot analysis only the full length variant was detected. Potentially, the breast cancer secretes the full length variant of CDCP1 or the protein is detected on the cell surface of extracellular vesicles surrounded by the cell membrane containing the full length variant of CDCP1. CDCP1 was detected in the cell culture supernatant of CTC-ITB001 as well as in the supernatant of CTC-ITB001-MIND. This highlights the diagnostic and potentially prognostic value of CDCP1-expression on CTCs due to the fact that this protein is thought to be supporting metastasis formation [32].

The determination of CDCP1 in the cell culture supernatant of these two CTC cell lines derived from breast cancer patients [147] indicates that CDCP1 could be used as a liquid biopsy biomarker for breast cancer. Therefore the blood plasma of 771 breast cancer patients from the whole cohort and 426 healthy female donors was analyzed by CDCP1 ELISA. CDCP1 was detected in 22.2 % of the patients with a specificity of 100 % compared with healthy controls. If the specificity is reduced to 96.5 % the sensitivity increases up to 29.1 %. For the detection of early stages of breast cancer, all cT1-patients were analyzed for CDCP1 from the whole cohort and only minor changes in sensitivity were detectable compared to all cases from the whole cohort. To evaluate the ability of detecting precancerous lesions only patients with DCIS were analyzed. The CDCP1 ELISA is able to detect breast cancer patients with DCIS with a sensitivity of 19.7 % at a specificity of 100 %. By reducing the specificity to 97.4 % the sensitivity increases to 23.9 %. After correlation of the determined CDCP1 concentration with the clinicopathological parameters of the breast cancer patients, a significant difference was observed for the grading and the therapy. Tumors with grading of “G1” or “G2” exhibit higher concentrations of CDCP1 in the blood plasma whereas G3-tumors

have lower concentrations of CDCP1. Grade 1 tumors are well differentiated with low malignancy, grade 2 tumors are moderately differentiated intermediate malignancy and G3-tumors are poorly differentiated with high grade of malignancy [214, 215]. This indicates that poorly-differentiated tumor cells have lower levels than well-differentiated tumor cells which is contrary to the findings that the hybrid EMT cell line MDA-231 and the mesenchymal DTC cell lines exhibit the highest expression levels of CDCP1. This opposite finding could be explained by the invasive properties mediated by CDCP1. The well-differentiated primary tumors might benefit from the *in vivo* cleavage of CDCP1 resulting in suppression of apoptosis and thereby leading to CDCP1-dependent survival in the processes of dissemination and metastasis formation [216]. This mechanism might support the tumor cells to survive after leaving the organization of the primary tumor. It was shown *in vitro* that CDCP1 is involved in mediation of anoikis resistance of tumor cells [216] which might explain the high expression levels of CDCP1 in the cell lines with a mesenchymal phenotype. These mesenchymal cell lines were observed to grow with less cell-cell contacts than the well-differentiated cell lines with epithelial phenotype. Another significant correlation was observed between CDCP1 concentration and neoadjuvant therapy. These patients received chemotherapy prior to surgery and CDCP1 was determined with lower concentrations in these patients. The main reason of neoadjuvant chemotherapy is to decrease the tumor size and to facilitate the surgical practice [217]. The reduced tumor mass and the reduced number of tumor cells after successful neoadjuvant therapy could explain the decreased levels of CDCP1 whereas the higher levels of CDCP1 in the blood plasma of patients might correlate with the larger tumor mass caused by the enhanced growth period until surgical resection.

The “early detection” cohort was generated by exclusion of specific cases from the whole cohort. All cases with performed blood draw after surgery, neoadjuvant therapy, recurrence or missing data were excluded. The exclusion yielded a collective of breast cancer patients as found at the time of primary diagnosis. This procedure was necessary to avoid external influences on the CDCP1 levels in the blood of the cancer patients. Therefore the blood plasma of 440 breast cancer patients from the “early detection” cohort and 426 healthy female donors was analyzed by CDCP1 ELISA. CDCP1 was detected in 22.3 % of the patients with a specificity of 100 % compared with healthy controls. If the specificity is reduced to 97.4 % the sensitivity increases up to 28.2 %. For the detection of early stages of breast cancer, all cT1-patients were analyzed for CDCP1 from the “early detection” cohort. CDCP1 was detected in 24.9 % of the patients with a specificity of 100 % compared with healthy controls. If the specificity is reduced to 97.4 % the sensitivity increases up to 30.7 %. To evaluate the ability of detecting precancerous lesions in the “early

detection” cohort only patients with DCIS were analyzed. The CDPC1 ELISA is able to detect breast cancer patients with DCIS with a sensitivity of 20.0 % at a specificity of 100 %. By reducing the specificity to 97.4 % the sensitivity increases to 26.7 %. After comparison of the CDPC1 concentration with the clinicopathological parameters of the breast cancer patients from the “early detection” cohort, there was no significant correlation for the grading and therapy detectable. The explanation for the missing significant correlation of the neoadjuvant therapy might be due to the exclusion of all patients with pre-surgical treatment to generate an “early detection” cohort. There was no significant correlation detectable for the grading in the “early detection” cohort. This finding might be explained by the exclusion of cases resulting in the reduction of G1 cases by 26 %, G2 cases by 43 % and G3 cases by 44 % for the establishment of the early detection cohort.

CDPC1 might function as breast cancer liquid biopsy marker with clinical relevance. The serum tumor biomarkers CEA (carcinoembryonal antigen), CA19-9 (sialylated form of Lewis Antigen), CA-125 (Mucin-16) and CA15-3 (Mucin-1) belong to the commonly used tumor biomarkers [38-40]. For CEA, the American Society of Clinical Oncology has demonstrated that this marker is not reliable in breast cancer detection. The properties do not allow the use for screening and diagnosis in breast cancer because of low sensitivity for early stage detection. Fifty percent of patients with recurrence will not be detected and the false positivity rate is 12%. CEA is suitable for monitoring therapy response [39]. Furthermore, ASCO has shown that increased CEA levels were detected in 10% of stage I, 19% of stage II, 31% of stage III and 64% of stage IV breast cancer [39]. Wang *et al.* determined CEA as serum tumor marker in metastatic breast cancer. They revealed CEA with a sensitivity of 56.7% at a specificity of 92.0% (AUC: 0.806, 95% CI: 0.761-0.852) in the diagnosis of breast cancer [38]. For CA 15-3 (Mucin-1), ASCO has assessed this protein as a biomarker for breast cancer. A critical point is that low levels of CA 15-3 do not exclude metastasis which means that false positive results might occur [218]. Elevated levels of CA 15-3 were observed in 9% of stage I, 19% of stage II, 38% of stage III and 75% of stage IV of breast cancer patients [218]. Wang *et al.* also analyzed CA 15-3 regarding the performance as a serum tumor marker in breast cancer patients. They determined CA 15-3 with sensitivity of 44.5% at a specificity of 84.5% (AUC: 0.743, 95% CI: 0.691-0.795). They suggest a combination of different tumor markers to increase sensitivity and specificity [219]. Dolscheid-Pommerich *et al.* analyzed CA125 regarding the possible use as serum tumor marker for cancer patients. It was observed with a median of 9.8 U/mL in healthy individuals, 11.6 U/mL in stage I, 11.7 U/mL in stage II, 13.2 U/mL in stage III and 18.2 U/mL in stage 4 of breast cancer patients.

Compared with healthy individuals the sensitivity of CA 125 was 33% at a specificity of 95% (AUC: 0.63) [220]. Wang *et al.* even observed a lower performance for CA 125 with a sensitivity of 25.6% at a specificity of 97.0% (AUC: 0.64, 95% CI: 0.582-0.698) [219]. CA 19-9 was observed with a median of 4.1 U/mL in healthy individuals, 4.6 U/mL in stage I, 3.3 U/mL in stage II, 5.3 U/mL in stage III and 7.3 U/mL in stage IV of breast cancer patients. Compared with healthy individuals the sensitivity was 19% at a specificity of 95% (AUC: 0.61) [220]. Wang *et al.* have determined the detection rate of breast cancer patients with a sensitivity of 36% at a specificity of 82.5% (AUC: 0.599, 95% CI: 0.539-0.659) [219]. CA 19-9 and CA 125 are not able to detect breast cancer in an early stage (stage I) of disease progression, whereas CEA detects at least 10% of stage I cases and CA 15-3 9% [218] of breast cancer patients with a stage I diagnosis.

In summary, the conventional breast cancer biomarkers seem to be not sufficient in the detection of breast cancer stage I, which would be necessary to minimize metastasis formation and to prevent late disease relapse. Dolscheid-Pommerich *et al.* determined a sensitivity of 19% with a specificity of 95% for CA 19-9 for distinguishing breast cancer patients and healthy subjects [220]. For CA 125, Wang *et al.* examined a sensitivity of 25.6% with a specificity of 97% for distinguishing breast cancer patients and healthy subjects [219]. The markers CEA and CA 15-3 were analyzed by Wang *et al.* for the use as serum tumor markers. For CA 15-3, they have determined a sensitivity of 44.5% with a specificity of 84.5%. For CEA, Wang *et al.* detected a sensitivity rate of 56.7% with a specificity of 92% [219]. ASCO analyzed the biomarkers CEA and CA 15-3 for their suitability in breast cancer detection regarding the detection possibilities of different stages. For CA 15-3 the detection rate of stage I breast cancer was 9% and for CEA 10% [218].

These findings indicate that CDCP1 is a better early detection biomarker for stage I breast cancer cases than CEA and CA 15-3. Furthermore, CDCP1 detects breast cancer with a higher specificity and sensitivity than CA125 and CA19-9. Notably, the CDCP1 ELISA is able to detect breast cancer with a specificity of 100%, which is not the case for the above mentioned biomarkers. One remaining issue of the CDCP1 test is the moderate sensitivity. It might be possible to combine CDCP1 with CEA to increase the detection rate for early stages and to reduce the false positive rates of CEA. A combination of CDCP1 with CA 15-3 as well could be possible to increase the detection rate for early stages of breast cancer. Due to these results, patients with genetic predisposition and a higher risk for developing breast cancer could benefit from a blood-based analysis of CDCP1. Additionally, the CDCP1 test could prevent the

development of breast cancer in high-risk patients by detecting the precancerous disease DCIS with a specificity of 100%.

For the use of CDCP1 as an isolation or therapeutic target it is also crucial that the PBMC fraction is negative for CDCP1. In case of the use as a therapeutic target using cytotoxic treatment or personalized immune cell therapy (CAR-T cells) the expression of CDCP1 or NRP-1 on non-tumor cells would cause severe side effects and perhaps life-threatening consequences for the patient. The expression levels of CDCP1 and NRP-1 were investigated on erythrocytes of healthy female donors. There was no detectable expression of CDCP1 or NRP-1 on erythrocytes. The early detection of these tumor cells with metastasis-initiating (MIC) potential could result in the prevention of metastasis formation. Osteoclasts, the bone-degrading cells which are attracted by tumor cells settled in the bone, could be blocked by special therapeutics (bisphosphonates) if MICs are detected [221]. Additionally, MICs could be directly targeted by cytotoxic treatment or by genetically engineered immune cells (CAR-T cells) against CTC or DTC expressing CDCP1 or NRP-1 [117].

The expression of CDCP1 seems to play a role in other tumor entities and is not only relevant in breast cancer. It was demonstrated by Western Blot analysis that CDCP1 is also expressed in cell lines of pancreas-, prostate-, head and lung cancer. Furthermore the expression of CDCP1 on mesenchymal DTC cell lines isolated from patients of different entities like lung- (LC-M1) and prostate (PC-E1) cancer was determined by Western Blot analysis and immunofluorescence staining. This is a promising finding because these DTC cell lines do not express epithelial proteins (cytokeratins/EpCAM) which are common detection and enrichment markers.

The evaluation of the biomarker candidate expression levels on tumor cells exposed to different oxygen concentrations is necessary to assess the suitability of these markers for the detection of CTCs/DTCs that were exposed to hypoxic conditions. For the blood-based detection of CTCs/DTCs it is essential that the markers are stable expressed and not downregulated after subjection to hypoxia or reoxygenization. The analysis of tumor cell lines exposed to hypoxia followed by reoxygenization does not provide a connection to a single oxygen-dependent pathway. It was shown that the pathways mediated by HIFs, mTOR and UPR respond to changes in environmental oxygen levels [76]. To address specifically the HIF-pathway, tumor cells were exposed to CoCl_2 which is known to stabilize HIF- α [222]. HIF-1 α was stabilized after exposure to cobalt for four days in all analyzed cell lines. The weak expression levels of HIF-1 α in MDA-

231 and MDA-231 BO2 under cultivation without cobalt can be explained by the oxygen-independent activation by growth factors like insulin-like growth factor 2 (IGF2) and transforming growth factor- α (TGF- α). These factors induce the synthesis of HIF-1 α by the activation of the phosphatidylinositol 3-kinase (PI3K) or the mitogen-activated protein kinase (MAPK) pathway [83]. MDA-231 and MDA-231 BO2 exhibit an elevated expression of MAPK1 (Erk2) and MAPK3 (Erk1) and a strong expression of phosphorylated MAPK1 (Erk2) under normal culture conditions which was shown by Western Blot analysis (**figure 6**) whereas MCF-7 exhibits a weak expression of MAPK1 (Erk2), a strong expression of phosphorylated MAPK1/2 (Erk1/2) and no detectable expression of HIF-1 α . Additionally, HIF-1 α was not detectable in BC-M1 under normal culture conditions during the cobalt experiments due to the strong expression of HIF-1 α under exposure to cobalt. But a weak HIF-1 α expression in BC-M1 was determined for the cell line characterization (**figure 6**). Erk1/2 was detected in BC-M1 with expression levels similar to MDA-231 and MDA-231 BO2 and phosphorylated Erk2 was detected with a slightly weaker expression than MDA-231 and MDA-231 BO2. These findings suggest that the presence of MAPK1 (Erk2) and the absence of phosphorylated MAPK3 (Erk1) might be involved in the activation of the synthesis of HIF-1 α under normal oxygen conditions.

The evaluation of the expression of the marker candidates under hypoxia and exposure to cobalt will be confined to the cell lines MDA-231, MDA-231 BO2 and BC-M1 and the marker proteins CDCP1 and NRP-1.

The expression of the full length variant of CDCP1 in MDA-231 was significantly increased for all time periods of exposure to cobalt compared to normal culture conditions whereas the cleaved variant showed a significant decrease after eight and 14 days of exposure to cobalt. Increased expression levels were also detectable for the full length variant of CDCP1 of MDA-231 under hypoxic conditions but the changes were not significant. However, the decreased expression levels of the cleaved variant of CDCP1 in MDA-231 under hypoxia and reoxygenization were significant. These findings suggest a HIF-1 α mediated regulation of CDCP1 in MDA-231. The expression of the full length variant of CDCP1 in BC-M1 showed no significant changes under subjection to hypoxia as well as under exposure to cobalt for four and eight days. After 14 days of exposure to cobalt a significant reduction of full length CDCP1 is detectable in BC-M1 but the reduced expression level is still very well detectable. A significant increase of the full length variant of CDCP1 in MDA-231 BO2 was determined after eight and 14 days under exposure to cobalt whereas the cleaved variant showed a significant reduction after 14 days of exposure to cobalt. This result

suggests HIF-1 α mediated regulation of CDCP1 in MDA-231 BO2. In general, CDCP1 (especially the full length variant) retains detectable expression levels independent of the subject to hypoxia/reoxygenation or the exposure to cobalt. This indicates that CDCP1 is suitable as a detection marker for CTCs/DTCs developed in a hypoxic environment. Therapy and detection strategies might benefit from the possibility to detect and target hypoxic tumor cells because hypoxia is a potential activator of EMT [64, 65] and a partial EMT phenotype is thought to correlate with early dissemination and metastatic outgrowth [41, 62].

Additionally, there were no significant changes detectable for NRP-1 in MDA-231 under exposure to hypoxic conditions whereas the expression levels of NRP-1 decreased significantly after exposure to cobalt but in both cases NRP-1 remains well detectable in MDA-231 under all analyzed conditions. NRP-1 shows a significant decrease in MDA-231 BO2 after four and eight days of cobalt-exposure whereas no significant difference is detectable between 14 days of exposure to cobalt and standard culture conditions. The expression of NRP-1 is easily detectable at every point in MDA-231 BO2 under exposure to cobalt. The expression levels of NRP-1 in BC-M1 decrease significantly after exposure to hypoxia or cobalt but NRP-1 is still well detectable after exposure to cobalt in contrast to the hardly detectable signal of NRP-1 after exposure to hypoxia. These findings suggest an influence of HIF-1 α on the regulation of NRP-1 but the strong reduction under hypoxia suggests that there are additional oxygen-sensitive regulators of NRP-1. These facts indicate that NRP-1 might not function as a reliable detection marker for all types of hypoxic tumor cells in the context of tumor cell dissemination. In general, the biomarker candidates remain stable and detectable under the influence of hypoxia or cobalt with the exception of NRP-1 in BC-M1 under hypoxia. These results suggest that the biomarker candidates might be able to detect CTCs/DTCs from hypoxic or reoxygenated origin.

In conclusion, CDCP1 might be used as a novel marker protein for the specific detection of early stage breast cancer and NRP-1 and CDCP1 might function as suitable marker proteins for the detection of CTC/DTC with mesenchymal attributes. In future, the use of CDCP1 and NRP-1 could improve the early detection of breast cancer as well as metastasis and dissemination prevention.

6.1. Conclusion

The aim of this project was the identification of cell surface proteins whose properties allow the use as a diagnostic biomarker for mesenchymal or hybrid epithelial/mesenchymal (partial EMT) tumor cells (CTC/DTC) in breast cancer. The cell surface proteins CDCP1, NRP-1, NRP-2 and DCBLD2 were identified by SILAC-LC-tandem mass spectrometry in a proteome comparison analysis between the mesenchymal breast cancer DTC cell line BC-M1 and the hybrid epithelial/mesenchymal bone metastatic cell line MDA-231 BO2. The highest expression levels of the marker candidates were detected in the DTC cell lines. In general, marker candidate expression was observed in mesenchymal or hybrid epithelial/mesenchymal cell lines. These findings could be verified by immunofluorescence analysis for CDCP1 and NRP-1. For the use as a blood based biomarker the expression on PBMCs and erythrocytes had to be negative to prevent high rates of false-positive events during diagnosis. There was no detectable expression of CDCP1, NRP-1 and DCBLD2 in PBMCs in Western Blot analysis. NRP-2 was excluded as biomarker candidate at that time due to unspecific signals in PBMCs of healthy donors. Additionally, CD45⁺ PBMCs were negative for the expression of CDCP1 and NRP-1 in immunofluorescence staining. One of the erythrocyte samples of a healthy female donor showed a weak expression of the cleaved variant of CDCP1 after a long exposure time in Western Blot analysis. However, this would not affect the diagnostic procedure because erythrocytes are discarded prior to tumor cell detection. If CDCP1 would be used as a therapeutic target for cytotoxic treatment or personalized immune therapy, the determined expression of CDCP1 on erythrocytes could be critical. However, the potential side effects of such therapy approaches could be reduced by giving blood transfusions or dialysis. Additionally, a magnetic bead-based tumor cell isolation method (MACS) via the cell surface protein CDCP1 was established. The average recovery rate of cell lines spiked into blood was more than 50%, showing that more than half of the tumor cells expressing CDCP1 could be isolated and detected. This promising result might lead to the improvement of understanding the biological properties of metastasis initiating mesenchymal or hybrid mesenchymal/epithelial tumor cells. This improved understanding could increase the success of tumor therapy and metastasis prevention. Furthermore, the detection method for CTCs from breast cancer patients via CDCP1 with the CellSearch System was established. The expression of CDCP1 was observed on CTCs enriched via the epithelial marker EpCam on 4% of the analyzed cells. The expression of CDCP1 in tumor cell lines usually correlates with a mesenchymal or hybrid mesenchymal/epithelial phenotype and the expression on epithelial cells is weak or not detectable. CD45⁻/EpCam⁺/CK⁺/CDCP1⁺ CTCs could

originate from the rare subgroup of epithelial CK⁺/CDCP1⁺ tumor cells or from cells which have undergone a partial epithelial-to-mesenchymal transition and exhibit a hybrid mesenchymal/epithelial phenotype. Due to the biological properties of CDCP1 and the potential generation of soluble CDCP1 fragment under specific conditions additional blood based detection approaches are conceivable. Therefore a Sandwich-ELISA which is capable of detecting the soluble CDCP1 fragment either in the cell culture supernatant of breast cancer tumor cell lines or in the blood plasma of breast cancer patients was established. The analysis of CDCP1 in the whole breast cancer patient cohort (771 breast cancer patients) has yielded a sensitivity of 22.2 % with 100 % specificity or 29.1 % sensitivity with a specificity of 96.5 % compared with healthy controls. These results show that CDCP1 could be used as a liquid biopsy breast cancer biomarker with higher specificity and sensitivity than CA 125 and CA 19-9. The determination of CDCP1 in the “early detection” cohort concerning stage I breast cancer and DCIS cases revealed detection rates with a sensitivity of 24.9 % (specificity 100 %) for stage I breast cancer cases respectively 20.0 sensitivity (specificity 100 %) for DCIS. These results indicate that CDCP1 is a suitable marker for early stages of breast cancer as well as for precancerous diseases like DCIS. Additionally, CDCP1 is more specific and sensitive than the commonly used tumor markers CA 15-3 and CEA.

7. Outlook

To improve the detection and isolation techniques for CTCs and DTCs from breast cancer patients with a mesenchymal or a hybrid epithelial/mesenchymal phenotype it is necessary to establish enrichment techniques which are able to capture and detect such types of cells. It was shown that CDCP1 is a cell surface protein which is able to detect and isolate tumor cells with mesenchymal or hybrid epithelial/mesenchymal attributes. An improvement for the detection of mCTCs/emCTCs in the blood of breast cancer patients with the CellSearch system would be achieved by the replacement of the EpCAM antibody by the CDCP1 antibody. The EpCAM antibody is able to enrich epithelial CTCs and a portion of emCTCs but a large amount of emCTCs and mCTCs are not detected due to the lack of EpCAM expression. In this work, it was shown that CDCP1 is expressed on different breast cancer cell lines and DTC cell lines with mesenchymal properties. Therefore it would be beneficial if the enrichment of CTCs in the CellSearch system would be performed by CDCP1 followed staining for cytokeratins, DAPI and CD45.

Another possibility to isolate mesenchymal CTCs or CTCs that have undergone EMT is the CDCP1-dependent MACS isolation system. In this work it was shown that this technique is able to isolate 55 % of CDCP1⁺/hybrid EMT MDA-231 cells. It would be necessary to establish this system for the analysis of blood from breast cancer patients to isolate emCTCS or mCTCs via CDCP1. The isolation and downstream analysis of such cells could contribute to a better understanding of metastasis formation and tumor cell dissemination which is necessary to improve therapy strategies and clinically used detection methods. The average recovery rate of the CDCP1-dependent MACS isolation could be increased by the use of a direct magnetic bead labeled CDCP1 antibody. FACS experiments revealed that a directly labeled CDCP1 antibody is able to detect 97.5 % of MDA-231 cells compared to a detection efficiency of 37.8 % for an unlabeled CDCP1 antibody. The direct labeling with magnetic beads would increase the recovery rate by avoiding the secondary conjugation reaction which causes the loss of detection efficiency. The CDCP1 antibody could be directly labeled to superparamagnetic iron oxide nanoparticles (SPIOs) [223, 224] to achieve a higher recovery rate.

Furthermore it would be beneficial to investigate if the signal of NRP-2 in PBMCs of healthy female donors originates from a specific or unspecific reaction. For the use as a diagnostic biomarker it is essential that NRP-2 is not expressed on PBMCs of healthy donors. If the reaction was unspecific than it would be necessary to establish a new antibody for NRP-2 which does not recognize proteins expressed on PBCMs by a cross-reaction.

Furthermore the cell surface protein DCBLD2 is a promising biomarker candidate. Therefore it is necessary to establish an antibody which is able to produce reliable results for the immunofluorescence staining of tumor cells *in vitro* to use DCBLD2 as detection marker for CTCs/DTCs from cancer patients.

8. Methods and Materials

8.1. Methods

8.1.1. Patients

The human examinations were executed according to the Helsinki rules after acceptance was received by the ethics committee of the Medical Association Hamburg (reference number PV5392). From all patients, written informed consent was gained prior to any study-related investigation. Samples from women with breast cancer or healthy control persons treated at the University Medical Centre Hamburg-Eppendorf, Germany, were used. For analyses of blood plasma from breast cancer patients either fresh blood or stored aliquots from the department of Gynecology were used. Blood from healthy persons with an age of over 50 years was collected from the Institute for Transfusion Medicine, University Medical Center Hamburg-Eppendorf. Fresh clinical samples from breast cancer patients were drawn from breast cancer patients positive for distant metastases.

Blood plasma of the collected blood of breast cancer patients was used to determine the CDCP1 concentration by CDCP1-Sandwich ELISA. All samples from breast cancer patients that were present in the Institute of Tumour Biology, University Medical Centre Hamburg-Eppendorf were analyzed. All analyzed patients were female (six male breast cancer cases were later deleted from further analysis; furthermore three cases were deleted due to missing data). Thus, the remaining sample set consists of 777 breast cancer cases. The mean age of the breast cancer patients at the time of the sample acquisition was 60.0 ± 13.4 years. The oldest patient was 92 years and the youngest 27 years at the time of the blood collection.

The blood was drawn into EDTA containing S-Monovette 9 ml K3E (Sarstedt, Nümbrecht, Germany) and immediately processed. The plasma was produced by centrifugation of the blood at $2,000 \times g$ at 4°C for 20 min and transferred in aliquots and long-term stored at -80°C until the samples were analyzed for CDCP1.

The blood sample for the analysis of CDCP1 with the CellSearch System were obtained from University Hospital Heidelberg, Heidelberg, Germany and were kindly provided by PD Dr. Sabine Riethdorf (Institute of Tumour Biology, University Medical Centre Hamburg-Eppendorf) for the determination of CDCP1 on CTCs by the CellSearch System. Heid 716 to Heid 806 represent blood samples of female metastatic breast cancer patients. The results of the analysis of CDCP1 are listed in **table 2**. The positive

cases Heid 716, Heid 786 and Heid L11/172 were derived from patients with ER⁻/PR⁻/ErbB-2⁺ primary tumors whereas the positive cases Heid 796, Heid 781 and Heid 805 were derived from patients with ER⁺/PR⁺/ErbB-2⁻ primary tumors. The data of the patients was not further specified due to legal property.

8.1.2. Cultivation conditions of cell lines

All used cell lines were cultivated at 37 °C in a humidified environment. MDA-MB 231 (cell line service, Eppelheim, Germany), MDA-MB 468 (purchased September 2016 from Leibniz Institut-Deutsche Sammlung von Mikroorganismen und Zellkulturen, Braunschweig, Germany), Hs578t (kindly provided from Prof. Thomas Dittmar, University of Witten/Herdecke, Germany), MCF-7 (from ATCC) were cultured in DMEM (Pan-Biotech, Aidenbach, Germany) supplemented with 10 % fetal calf serum (Capricorn, Ebsdorfergrund, Germany), 2 mmol/L L-glutamine (Gibco/Thermo Fisher Scientific, Waltham, USA) and 100 U/mL Penicillin/Streptomycin (Gibco/Thermo Fisher Scientific, Waltham, USA). These cell lines were held under 10 % CO₂ and atmospheric oxygen concentration using the Heracell 150i incubator (Thermo Fisher Scientific, Waltham, USA). The cell lines BC-M1, LC-M1 and PC-E1 [193] were cultivated as described [194]. DTC cell lines were cultured in RPMI (Pan-Biotech, Aidenbach, Germany) supplemented with 10 % fetal calf serum, 2 mmol/L L-glutamine, 100 U/mL Penicillin/Streptomycin, 1 % IST (Insulin, Transferrin, Selen; Gibco/Thermo Fisher Scientific, Waltham, USA), 50 ng/mL human-EGF and 10 ng/mL human-bFGF (both from Miltenyi Biotec, Bergisch Gladbach, Germany) and were kept under 10 % oxygen and 5% CO₂ using the Heracell 15 incubator (Thermo Fisher Scientific, Waltham, USA). The cell lines MDA-MB 231 BO2 [195-197] (kindly provided from Prof. Philippe Clézardin, University of Sheffield, United Kingdom) and MDA- MB 231 SA [225-227] (kindly provided from Prof. Theresa Guise, Indiana University, USA) were cultured in RPMI supplemented with 10 % fetal calf serum, 2 mmol/L L-glutamine, 100 U/mL Penicillin/Streptomycin and were held under 10 % oxygen and 5 % CO₂ using the Heracell 15 incubator (Thermo Fisher Scientific, Waltham, USA) [69, 174]. The human pancreatic cell lines Panc-1, BxPC3 and MiaPaca2 [228] and the patient derived human pancreatic cancer cell line PaCa 5061 [229] were kindly provided by Dr. Cenap Güngör (University Medical Center Hamburg-Eppendorf, Department of General, Visceral- and Thoracic Surgery, Germany). Panc-1, BxPC3 and MiaPaca2 were cultivated in DMEM supplemented with 10 % fetal calf serum and 1 % Penicillin/Streptomycin and held under 10 % CO₂. The cell line PaCa 5061 was

cultivated in complete medium (TUM) [229] containing RPMI supplemented with 10 % fetal calf serum, 200 IU/mL penicillin/streptomycin, 0.1 mg/mL gentamycin (Sigma-Aldrich, Steinheim, Germany), 50 nmol/mL human transferrin (Sigma-Aldrich, Steinheim, Germany), 0.01 µg/mL bovine insulin (Sigma-Aldrich, Steinheim, Germany), 0.01 µg/mL recombinant human epidermal growth factor (Pepro Tech, London, UK) and 0.01 µg/mL human basic fibroblast growth factor (Pepro Tech, London, UK) and held under 5 % CO₂. The human prostate cancer cell lines LnCap, DU145 and PC3 were kindly provided by Dr. Stefan Werner (University Medical Center Hamburg-Eppendorf, Institute of Tumor Biology, Germany). These cell lines were cultured in RPMI supplemented with 10 % fetal calf serum, 2 mmol/L L-glutamine, 100 U/mL Penicillin/Streptomycin and were held under 5 % CO₂. The human lung cancer cell lines A549 and H1299 [230] were kindly provided by Prof. Dr. Harriet Wikman-Kocher (University Medical Center Hamburg-Eppendorf, Institute of Tumor Biology, Germany). These cell lines were cultured in RPMI supplemented with 10 % fetal calf serum, 2 mmol/L L-glutamine, 100 U/mL Penicillin/Streptomycin and were held under 5 % CO₂. The human head and neck cancer cell lines SAS, UM-SCC-5 and UT-SCC 42a were kindly provided by Dr. Tobias Gorges (University Medical Center Hamburg-Eppendorf, Institute of Tumor Biology, Germany). The cell line SAS cultivated with DMEM supplemented with 10 % fetal calf serum, 365 mg/L L-glutamine, 100 U/mL Penicillin, 100 µg/mL Streptomycin and were grown under 5 % CO₂ [231]. The cell line UM-SCC-5 was grown in complete DMEM (cDMEM) supplemented with 2 mM L-glutamine, 1 % non-essential amino acids, 1 % Penicillin/Streptomycin and 10 % fetal calf serum and was cultivated under 5 % CO₂ [232]. The cell line UT-SCC 42a was cultivated with DMEM supplemented with supplemented with 10 % fetal calf serum, 4 mM L-glutamine and 100 U/mL Penicillin/Streptomycin and were grown under 10 % CO₂ [233]. The CTC cell line CTC-ITB001 was generated by Dr. Andra Kuske (University Medical Center Hamburg-Eppendorf, Institute of Tumor Biology, Germany) [147]. The corresponding mouse xenograft cell line CTC-ITB001-MIND was kindly provided by Prof. Dr. Cathrin Brisken (École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland) and generated by Georgios Sflomos (École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland) according to the following procedure. A single cell suspension of cells of the cell line CTC-ITB001 were tagged with a luciferase gene (GFP-*luc2*) by lentiviral transfer (transduction efficiency: 30-40 %) and 300.000 cells were injected into milk duct of immunodeficient NOD scid gamma (NSG) female mice (n = 4). Mice were sacrificed after 8.5 months. The tumor size was examined at several time points followed by transplantation performed with Xenogen IVIS bioluminescence imaging system 200 (Caliper Life Science) according to manufacturer's protocol 12-

15 minutes after intraperitoneal injection of 150 mg/kg luciferin (Biosynth AG, catalogue number: L-8220). Analysis was performed with Living Imaging Software (Caliper Life Science). GFP-labeled tumor cells from the primary tumor caused by CTC-ITB001 were isolated and cultivation was established for this cell line named CTC-ITB001-MIND.

Both cell lines were cultivated in RPMI supplemented with 10 % fetal calf serum, 2 mmol/L L-glutamine, 200 U/mL Penicillin/Streptomycin, 1 % IST (Transferrin, Insulin, Selenium), 50 ng/mL human-EGF, 10 ng/mL human-bFGF, 100 ng/mL Hydrocortisone (Sigma-Aldrich, Steinheim, Germany), 100 ng/mL Cholera Toxin (Sigma-Aldrich, Steinheim, Germany) and were grown under 5 % CO₂. The Western Blot analysis (**figure 25**) for the proteins CK 18/19 and EpCAM was performed and kindly provided by Parinaz Mossahebi Mohammadi (University Medical Center Hamburg-Eppendorf, Institute of Tumor Biology, Germany).

8.1.3. Colloidal coomassie staining of polyacrylamide gels (Neuhoff staining)

For visualization of the diverse protein expression patterns of the different cell lines and for verification of the protein concentration determination SDS-polyacrylamide gels were analyzed by colloidal coomassie staining (Neuhoff staining) [234].

8.1.4. Cell harvest, SDS-Page, Western Blots and membrane stripping

Cells were grown in 75 cm²-flasks (Sarstedt, Nümbrecht, Germany) and harvested at required confluence in biological triplicates. Cells were washed three times with 37° C-prewarmed phosphate buffered saline (DPBS, life technologies, Carlsbad, USA) and harvested using 300 µL lysis buffer (9.8 M Urea, 30 mM Tris-Base and 15 mM EDTA). After ultrasound treatment (three times, amplitude 100 %, 10 seconds) with ultrasonic device UP50H (Hielscher, Teltow, Germany) cells were incubated at room temperature and shaking at 1400 rpm for two hours, followed by centrifugation (8.000 x g at room temperature for two minutes). The supernatant was taken and protein concentration was measured using Pierce BCA Protein Assay Kit (Pierce, Rockford, USA) according to the instruction manual. Bovine serum albumin (BSA; Sigma-Aldrich, Steinheim, Germany) was used to generate a standard curve used for conversion of the measured optical density into the protein concentration. All samples were stored at -80 °C.

Protein size separation was performed with Novex XCell Sure-Lock mini system (Invitrogen, Groningen, Netherlands) or Protean II xi cell (BioRad, Hercules, USA) using 8 % polyacrylamide separation gels and Laemmli-buffered system [235]. Prior to separation, protein samples were diluted in 5x sample buffer (2.73 g Tris-HCl, 50 g glycerin, 5 g sodium dodecyl sulfate, 50 mg bromphenol blue, 3.86 g dithiothreitol (DTT), H₂O ad 100 mL, pH6.8) and denatured by heat at 95° C for five minutes. As a molecular weight standard peqGOLD protein marker V (PeqLab, Erlangen, Germany) was used. 30 µg of total protein per sample was used for Western Blot analysis. After size separation, proteins were transferred to 0.2 µm pore size FluoroTrans[®] PVDF-membrane (VWR, Radnor, USA). The polyacrylamide gels were equilibrated with transfer buffer (5.82 g Tris-Base, 2.93 g glycine, 100 ml 99 % ethanol, H₂O ad 1 L) for 20 minutes. The PVDF-membrane was equilibrated in methanol for 10 minutes, followed by incubation in transfer buffer for 10 minutes. The protein transfer was performed using a tank blot system mini VE vertical electrophoresis combined with a tank blot transfer unit (GE Healthcare, Uppsala, Sweden). The transfer conditions were 30 V, 160 mA for 35 minutes per tank blot unit. Membranes were blocked with 5 % lowfat powdered milk (Roth, Karlsruhe, Germany) or 5 % BSA in 1x TBST (blocking buffer 10x: 48.4 g Tris-Base, 160 g sodium chloride, 20 mL Tween 20, pH 7.6 with HCl) for one hour at room temperature with gentle shaking. All used antibodies and applied concentrations are listed below in the section “antibodies used for Western Blot analysis”. Primary antibody dilutions were prepared in blocking buffer and membranes were incubated with antibody solution over night at 4° C with gentle shaking. Corresponding horseradish

peroxidase-conjugated secondary antibodies (DAKO, Glostrup, Denmark) were applied with dilutions from 1:1000 to 1:20.000 depending on used cell line and signal intensity. All secondary antibodies were diluted in 5 % lowfat powdered milk (Roth, Karlsruhe, Germany) in TBST (blocking buffer) and membrane was incubated for one hour at room temperature with smooth shaking. Protein bands were visualized with Signal Fire™ Plus ECL reagent (Cell Signaling Technology, Danvers, USA) and X-ray films CEA CR5 UV (Agfa Healthcare, Mortsel, Belgium) according to the supplied instruction manual. The development of the X-ray films were performed with Curix 60 device (Agfa). The digitization of the X-ray films was done with GS-700 imaging densitometer (BioRad) and Quantity one software (BioRad) was used for analysis. To remove membrane-bound antibodies the membranes were incubated with 25 mL stripping buffer (7.56 g Tris, 20 g SDS, 7.8 g 2-mercaptoethanol, adjustment to pH 9.5 using HCl, H₂O ad 1 L, prior to use 0.1 g DTT was added) for one hour at room temperature with smooth agitation. The membranes were washed with TBST for 3 hours at room temperature with gentle agitation and incubated with blocking buffer for one hour at room temperature with smooth agitation before a different primary antibody could be applied [174].

8.1.5. Culture conditions of cell lines under hypoxia (1% O₂)

Cell lines were cultivated under 1% O₂ (hypoxia) for 14 days. The following reoxygenation was performed for four hours under 10% O₂ using the Heracell 15 incubator (Thermo Fisher Scientific, Waltham, USA). The oxygen concentration was regulated by using N₂. Apart from that, the cell lines were cultivated as described in the chapter “cultivation conditions of cell lines”. The cell harvest was performed immediately and as described in the chapter “Cell harvest, SDS-Page, Western Blots and membrane stripping”. The cultivation under normal oxygen conditions is described as “0d”, cultivation under 1% O₂ for 14 days is described as “14d”, whereas “14 days+ 4h” means that the cell line was cultivated under 1% O₂ for 14 days and following cultivation under 10% O₂ concentration for four hours [69, 174].

8.1.6. Culture conditions of cell lines under cobalt-chloride (CoCl_2)

The regulation of alterations in target gene expression levels due to decreased oxygen concentrations is not only restricted to the HIF-pathway. The mTOR- and UPR-pathways are also known to be sensitive to oxygen [76]. To investigate the influence of HIF-1 mediated changes in gene expression levels it was shown that CoCl_2 mediates the stabilization of Hif-1 α under normoxia [222]. This approach makes it possible to determine the influence Hif-1 α on the expression levels of different proteins of interest (potential biomarkers). Cobalt mediates the stabilization of Hif-1 α in two ways. On the one hand the cobalt-ion replaces the iron-ion in the active center of the enzyme prolyl hydroxylase domain proteins (PHD) and thereby reducing the hydroxylase activity. As a result the PHD is not able to hydroxylate the proline residues in the oxygen-dependent degradation domain (ODD) of Hif-1 α . On the other hand, cobalt directly binds the ODD and thereby preventing the hydroxylation of the proline residues as well as the binding from the von Hippel-Lindau protein (VHL) to Hif-1 α (**figure 67**). Cobalt is able to bind to Hif-1 α independent of the hydroxylation situation of the proline residues in the ODD [222].

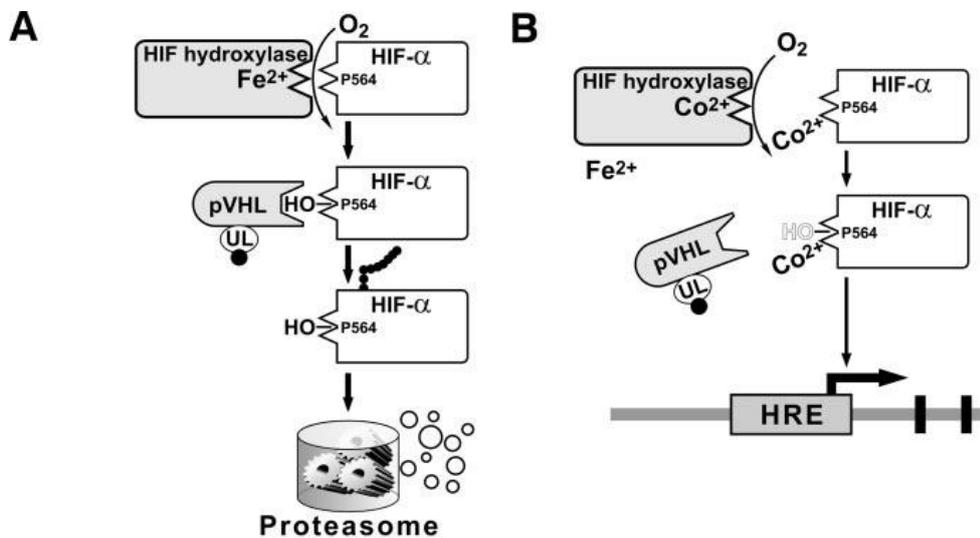


Figure 67: Detailed model for the degradation of Hif-1 α under normal oxygen conditions (normoxia) (A) and cobalt-mediated stabilization Hif-1 α (B). The image was edited according to the template figure 5 from “Cobalt Inhibits the Interaction between Hypoxia-inducible Factor- α and von Hippel-Lindau Protein by Direct Binding to hypoxia-inducible factor- α ” [222].

The cell lines were cultivated under normal culture conditions as described in the chapter “cultivation conditions of cell lines”. CoCl_2 (Sigma-Aldrich, Steinheim, Germany) was added to stabilize of HIF-1 α [222] and the cell lines were cultivated for four days, eight days, 14 days and without cobalt-chloride

(standard culture conditions). MCF-, MDA-231 and MDA-231 BO2 were cultivated with 150 μ M CoCl₂ added to normal cell culture medium. BC-M1 was cultivated with normal cell culture medium supplemented with 75 μ M CoCl₂. The medium was changed every 24 hours. CoCl₂ was diluted in ddH₂O to a final concentration of 150 mM. The stock solution was sterile filtrated and stored at -20°C. The microscope images were taken with Axiovert25 (Carl Zeiss, Oberkochen, Germany) before cell harvest and the cell harvest was performed as described in the chapter “Cell harvest, SDS-Page, Western Blots and membrane stripping”. Images generated with the software AxioVision (Carl Zeiss, version: 4.8.2). To ensure Hif-1 α -stabilization after cell harvest appropriate CoCl₂ concentration was added to the lysis buffer [69].

8.1.7. Antibodies used for Western blot analysis

Table 8: Antibodies for Western Blot analysis were acquired from following companies and were applied as described in **table 3** (mc = monoclonal, pc = polyclonal, M = mouse, R = rabbit, G = goat, S = sheep).

Primary antibody	Company	Clone	Clonality	Dilution	Species	Secondary antibody	Company	Dilution
anti-CK5	Abcam, Cambridge, United Kingdom	XM26	mc	1:5000	M	goat-anti-mouse-HRP	DAKO, Glostrup, Denmark	1:1000
anti-CD24	Abcam, Cambridge, United Kingdom	EPR3006(N)	mc	1:4000	R	goat-anti-rabbit-HRP	DAKO, Glostrup, Denmark	1:2000
anti-CXCR4	Abcam, Cambridge, United Kingdom	-	pc	1:4000	R	goat-anti-rabbit-HRP	DAKO, Glostrup, Denmark	1:1000
anti-NRP-1	Abcam, Cambridge, United Kingdom	EPR3113	mc	1:10000	R	goat-anti-rabbit-HRP	DAKO, Glostrup, Denmark	1:3000
anti-CK 19	Antibodies-online, Atlanta, USA	BA-17	mc	1:100000	M	goat-anti-mouse-HRP	DAKO, Glostrup, Denmark	1:10000
anti-Vimentin	BD Pharmingen, Erembodegem, Belgium	RV202	mc	1:200000	M	goat-anti-mouse-HRP	DAKO, Glostrup, Denmark	1:10000
anti-CDCP1	Cell Signaling Technology, Danvers, USA	D8M5K	mc	1:1000	R	goat-anti-rabbit-HRP	DAKO, Glostrup, Denmark	1:1000
anti-EGFR	Cell Signaling Technology, Danvers, USA	D38B1	mc	1:1000	R	goat-anti-rabbit-HRP	DAKO, Glostrup, Denmark	1:1000
anti-N-Cadherin	Cell Signaling Technology, Danvers, USA	D4R1H	mc	1:10000	R	goat-anti-rabbit-HRP	DAKO, Glostrup, Denmark	1:1000
anti-CD44	Cell Signaling Technology, Danvers, USA	-	pc	1:5000	R	goat-anti-rabbit-HRP	DAKO, Glostrup, Denmark	1:1000
anti-α-Tubulin	Cell Signaling Technology, Danvers, USA	11H10	mc	1:20000	R	goat-anti-rabbit-HRP	DAKO, Glostrup, Denmark	1:1000
anti-Grp78	Cell Signaling Technology, Danvers, USA	C50B12	mc	1:1000	R	goat-anti-rabbit-HRP	DAKO, Glostrup, Denmark	1:1000
anti-Integrin alpha 5	Cell Signaling Technology, Danvers, USA	-	pc	1:1000	R	goat-anti-rabbit-HRP	DAKO, Glostrup, Denmark	1:1000
anti-pSTAT3 (Y705)	Cell Signaling Technology, Danvers, USA	D3A7	mc	1:1000	R	goat-anti-rabbit-HRP	DAKO, Glostrup, Denmark	1:1000
anti-AKT	Cell Signaling Technology, Danvers, USA	-	pc	1:2000	R	goat-anti-rabbit-HRP	DAKO, Glostrup, Denmark	1:1000
anti-pAKT (S473)	Cell Signaling Technology, Danvers, USA	D9E	mc	1:500	R	goat-anti-rabbit-HRP	DAKO, Glostrup, Denmark	1:1000
anti-STAT3	Cell Signaling Technology, Danvers, USA	79D7	mc	1:2000	R	goat-anti-rabbit-HRP	DAKO, Glostrup, Denmark	1:1000

anti-HIF1α	Cell Signaling Technology, Danvers, USA	D2U3T	mc	1:1000	R	goat-anti-rabbit-HRP	DAKO, Glostrup, Denmark	1:1000
anti-Erk1/2	Cell Signaling Technology, Danvers, USA	-	pc	1:2000	R	goat-anti-rabbit-HRP	DAKO, Glostrup, Denmark	1:1000
anti-pErk1/2 (Thr202/Tyr304)	Cell Signaling Technology, Danvers, USA	D13.14.4E	mc	1:2000	R	goat-anti-rabbit-HRP	DAKO, Glostrup, Denmark	1:1000
anti-NRP-1^c	Cell Signaling Technology, Danvers, USA	D62C6	mc	1:1000	R	goat-anti-rabbit-HRP	DAKO, Glostrup, Denmark	1:1000
anti-DCBLD2^a	Cell Signaling Technology, Danvers, USA	-	pc	1:1000	R	goat-anti-rabbit-HRP	DAKO, Glostrup, Denmark	1:1000
anti-E-Cadherin	Epitomics, Burlingame, USA	EP700Y	mc	1:50000	R	goat-anti-rabbit-HRP	DAKO, Glostrup, Denmark	1:1000
anti-ErbB2	Leica Biosystems (Novocastra), Buffalo Grove, USA	CB11	mc	1:10000	M	goat-anti-mouse-HRP	DAKO, Glostrup, Denmark	1:10000
anti-EpCAM	Leica Biosystems (Novocastra), Buffalo Grove, USA	VU-1D9	mc	1:500	M	goat-anti-mouse-HRP	DAKO, Glostrup, Denmark	1:500
anti-pan cytokeratin	Leica Biosystems (Novocastra), Buffalo Grove, USA	AE1/AE3	mc	1:100000	M	goat-anti-mouse-HRP	DAKO, Glostrup, Denmark	1:10000
anti-pan cytokeratin	Micromet, Munich, Germany	A45/BB3	mc	1:100000	M	goat-anti-mouse-HRP	DAKO, Glostrup, Denmark	1:20000
anti-CK 18	Progen Biotechnik, Heidelberg, Germany	Ks18.04	mc	1:10000	M	goat-anti-mouse-HRP	DAKO, Glostrup, Denmark	1:1000
anti-CK 8	Progen Biotechnik, Heidelberg, Germany	Ks8.7	mc	1:5000	M	goat-anti-mouse-HRP	DAKO, Glostrup, Denmark	1:1000
anti-CDCP1	R&D systems, Minneapolis, USA	-	pc	1:1000	G	rabbit-anti-goat-HRP	DAKO, Glostrup, Denmark	1:1000
anti-NRP-2	R&D systems, Minneapolis, USA	-	pc	1:1000	G	rabbit-anti-goat-HRP	DAKO, Glostrup, Denmark	1:1000
anti-DCBLD2^b	R&D systems, Minneapolis, USA	-	pc	1:500	S	rabbit-anti-sheep-HRP	life technologies, Carlsbad, USA	1:500
anti-EGFR (1005)	Santa Cruz Biotechnology, Santa Cruz, USA	-	pc	1:500	R	goat-anti-rabbit-HRP	DAKO, Glostrup, Denmark	1:2000
anti-GAPDH	Santa Cruz Biotechnology, Santa Cruz, USA	-	pc	1:5000	R	goat-anti-rabbit-HRP	DAKO, Glostrup, Denmark	1:1000
anti-Connexin 43	Sigma-Aldrich, St. Louis, USA	-	pc	1:10000	R	goat-anti-rabbit-HRP	DAKO, Glostrup, Denmark	1:1000
anti-CDCP1	Thermo Fisher Scientific, Waltham, USA	309137	mc	1:1000	M	goat-anti-mouse-HRP	DAKO, Glostrup, Denmark	1:1000

^a This antibody was used in the experiments shown in the left panel of figure 24 and in figure 30. ^b This antibody was used in the experiments shown in the right panel of figure 24 and in figure 26. ^c This antibody was used in the experiments shown in figure 32 and in figure 33.

8.1.8. Isolation of mononuclear cells from peripheral blood

For the immunocytochemical detection, cells from human tumor cell lines were spiked into mononuclear cells of the peripheral blood (PBMC) of healthy female donors. The PBMC were isolated by density gradient (density gradient: 1.077 g/mL) centrifugation using Ficoll-Paque™ (GE Healthcare, Chicago, USA). Ficoll-Paque (20 mL) was covered with a solution consisting of whole blood (7.5 mL) from healthy female donor and DPBS (12.5 mL) and centrifuged with the Multifuge 3 S-R centrifuge (Heraeus, Hanau, Germany) at 1400 rpm ($399 \times g$) for 30 minutes with soft acceleration and braking. The supernatant (blood plasma) was discarded. The interphase which contains the PBMC fraction was transferred to a new tube, diluted with DPBS (ad 40 mL) and centrifuged at 1400 rpm for ten minutes. The supernatant was discarded and the pellet was diluted in DPBS (10 mL).

8.1.9. Generation of cytopins from cell lines

Adherent growing cells were detached with 0.25 % Trypsin-EDTA (Thermo Fisher Scientific, Waltham, USA) for five minutes at 37 °C, washed with appropriate cell culture medium and centrifuged with Rotofix32 centrifuge (Hettich, Kirchlengern, Germany) at 1200 rpm (zyto-rotor, six-fold) for three minutes and resuspended with DPBS. The isolated PBMCs (approximately 5×10^5 cells) were added to the cell suspension (approximately 5×10^4 cells) and applied to a funnel of large size which was placed on a microscope slide (Superfrost plus, Karl Hecht KG, Sondheim, Germany) and centrifuged with Rotofix32 centrifuge (Hettich, Kirchlengern, Germany) at 1200 rpm (swing-out rotor, six-fold) for seven minutes.

8.1.10. Antibodies used for immunocytochemical staining analysis

Antibodies were acquired from the following companies and were applied as follows: R&D systems, Minneapolis, USA: anti-human-CDCP1 (catalogue number: AF2666) antibody, goat polyclonal, dilution 1:100. Abcam, Cambridge, United Kingdom: anti-NRP-1 (clone: EPR3113) antibody, rabbit monoclonal, dilution: 1:300. Thermo Fisher Scientific (Invitrogen), Waltham, USA: anti-human-pan cytokeratin (AE1/AE3) Alexa Fluor 488-conjugated antibody, mouse monoclonal, dilution 1:300. Anti-goat IgG Alexa546-conjugated antibody (secondary antibody for R&D CDCP1 primary antibody), donkey polyclonal, dilution 1:200. Cell Signaling Technology, Danvers, USA: anti-human- pan cytokeratin (C11) Alexa Fluor 488-conjugated antibody, mouse monoclonal, dilution 1:200. Miltenyi Biotec, Bergisch

Gladbach, Germany: anti-human-CD45 (REA747) APC-conjugated antibody, dilution 1:200. Sigma-Aldrich, St. Louis, USA: DAPI (nucleic acid staining), dilution 1:1000.

8.1.11. Immunocytochemical staining of cell lines

For immunocytochemical staining and detection of cells from tumor cell lines cells were spiked into PBMC of healthy female donors. The antibodies goat polyclonal anti-human CDCP1 (AF2666, R&D systems, Minneapolis, USA), rabbit monoclonal anti-human-NRP-1 (EPR3113, Abcam, Cambridge, United Kingdom) and the established mouse monoclonal, Alexa488-conjugated anti-pan-Cytokeration antibodies AE1/AE3 (Invitrogen, Carlsbad, USA) and C11 (Cell Signaling Technology, Danvers, USA) were used. For the detection and exclusion of CD45-positive leukocytes the APC-conjugated anti-human-CD45 (REA737) antibody (Miltenyi Biotec, Bergisch Gladbach, Germany) was used. For the staining of the nucleus DAPI (Sigma-Aldrich, St. Louis, USA) was applied. The microscope slides were thawed 30 minutes prior to use. To prevent detaching of cells the cell-containing area on the microscope slide was surrounded with the DAKO-pen (Dako/Agilent, Santa Clara, USA). Cells were fixed using 250 μ L 2 % para-formaldehyde (pH 7.4) in DPBS for ten minutes at room temperature. Cells were washed one time with 150 μ L DPBS for three minutes and permeabilized for 10 minutes with 150 μ L 0.1 % Triton X 100™ (Sigma-Aldrich, St. Louis, USA) in DPBS. Cells were washed two times for three minutes with 150 μ L DPBS and were blocked to prevent unspecific signals for 20 minutes using 150 μ L 10 % AB-serum (Bio-Rad, Hercules, USA) in DPBS. 150 μ L of primary antibody against CDCP1 was added at a dilution of 1:100 in 10 % AB -serum in DPBS for 60 minutes at room temperature (For NRP-1: dilution 1:300 in 10 % AB -serum in DPBS for 60 minutes at room temperature). Microscope slide was washed three times for three minutes with 150 μ L DPBS and Alexa 546 donkey anti-goat secondary antibody (Invitrogen, Carlsbad, USA) was applied at a dilution of 1:200 in 10 % AB-serum in DPBS for 45 minutes at room temperature. After washing the cells three times for three minutes with 150 μ L an antibody-cocktail (Alexa488-conjugated anti-Cytokeration (AE1/AE3) antibody at a dilution of 1:300, Alexa488-conjugated anti-Cytokeration (C11) antibody at a dilution of 1:200, APC-conjugated anti-CD45 antibody at a dilution of 1:200 and DAPI at a dilution of 1:1.000 in 10% AB -serum in DPBS was added for 60 minutes at room temperature. Cells were washed three times for three minutes with 150 μ L DPBS and covered with one drop of mounting medium Prolong™ Gold Antifade (Thermo Fisher Scientific,

Waltham, USA). The cover slip was fixed with fixogum (Marabu, Ludwigsburg, Germany) and microscope slides were analyzed manually using the microscope Axioplan2 (Carl Zeiss, Oberkochen, Germany) with the software AxioVision (Carl Zeiss, Oberkochen, Germany, version: Rel.4.8.).

8.1.12. Cultivation and immunofluorescent analysis of cell lines in chamber slides

The cell line MDA-231 was seeded with a cell number of 50.000 cells/well into a two well-chamber slide (Lab-Tek chamber slides w/cover, Thermo Fisher Nunc, Waltham, USA). The cells were cultivated as described in the section “Cultivation conditions of cell lines” and were grown for 48 hours until a suitable confluence achieved. Medium (2 mL) was added to each well. After removal of the chamber immunofluorescent analysis for CDCP1, NRP-1 and cytokeratins (AE1/AE3) can be performed as described in the chapter “Immunocytochemical staining of cell lines”.

8.1.13. Stable isotope labeling by amino acids in cell culture (SILAC)

For the comparison of the protein expression profiles BC-M1 [194] and MDA-231 BO2 [195-197] were chosen as a model setup for disseminated tumor cells (BC-M1) and bone metastasis (MDA-231 BO2). These cell lines were cultivated as described in the section “culture conditions of cell lines”. The labeling procedure was performed using Pierce SILAC Protein Quantitation Kit (Pierce Biotechnology/Thermo Fisher Scientific, Waltham, USA) following the instruction manual (**figure 68**). BC-M1 and MDA-231 BO2 were cultivated in RPMI 1640 medium for SILAC containing 10 % dialyzed fetal bovine serum (Pierce Biotechnology/Thermo Fisher Scientific, Waltham, USA), 2 mM L-glutamine (Gibco/Thermo Fisher Scientific, Waltham, USA) and 100 U/mL Penicillin/Streptomycin (Gibco/Thermo Fisher Scientific, Waltham, USA). RPMI for BC-M1 was additionally supplemented with 1 % IST (Insulin, Transferrin, Selen; Gibco/Thermo Fisher Scientific, Waltham, USA), 50 ng/mL human-EGF and 10 ng/mL human-FGF (both from Miltenyi Biotec, Bergisch Gladbach, Germany). For labeling the proteins of the cell line MDA-231 BO2 $^{13}\text{C}_6$ -Arginine-HCl and $^{13}\text{C}_6$ -Lysine-2HCl (both Pierce Biotechnology) in a concentration of 100 mg/L RPMI medium were used (heavy). For labeling the proteins of the DTC cell line BC-M1 L-Arginine-HCl and L-Lysine-2HCl (both Pierce Biotechnology) in a

concentration of 100 mg/L RPMI was used (light). To minimize potential false labeling of the amino acid proline, cell culture medium was supplemented with 35 mg/L proline (Pierce Biotechnology) to a final proline concentration of 55 mg/L. All amino acids were sterile filtrated prior to addition to the culture medium. The microscopic analysis of the cells cultivated in the heavy- or light-labeled culture medium (SILAC-medium) did not cause any phenotypical changes compared to the cells growing under standard cell culture conditions.

Cells were harvested as described in the section “Cell harvest, SDS-Page, Western Blots and membrane stripping”. A small aliquot of the cell lysates was used for the determination of the correct incorporation of the heavy isotopes ($^{13}\text{C}_6$). In that case, 15 μg of the cell lysate was separated by SDS-PAGE as described and a Coomassie staining was performed [234], followed by excision of proteins bands. In-gel reduction, alkylation with iodoacetamide (IAA), tryptic digest, and extraction of the peptides were performed as described [236]. Mass spectrometry of these peptides was performed as described [69]. Signal intensities of 94.98 % (heavy), 2.95 % (light) and 2.07 % (heavy + proline) were determined for the samples. These values were considered as suitable for further analyses.

For preparative peptide analyses, the remaining cell lysate was applied. The samples were supplemented with 0.05 volume of 1 M DTT in 1 M Tris and incubated at 56 °C for 20 min with vigorous shaking. DTT also served as a scavenger for the potential formation of isocyanate ions from urea. The irreversible blocking of the SH-groups was carried out by addition of IAA to a final concentration of 200 mM. The pH values were adjusted to pH 9.5 using 2 M Tris. The samples were incubated at room temperature and protected from light for 30 min with vigorous shaking. Afterwards the proteins were purified by precipitation using 1.2 mL of precipitant (component of the 2-D Quant Kit, GE Healthcare) per 500 μL of the sample and co-precipitant (component of the 2-D Quant Kit) in the same amount as the precipitant [237]. The proteins were centrifuged (16.000 $\times g$ for 5 min) and washed with 1 ml of 80 mM Tris in 80 % acetone. An additional washing step with 1 mL of 80 mM acetone was used to remove residual contaminants. The purified proteins were dissolved in 100 μL of 9.8 M urea and solubilized for 1 h at room temperature. This was followed by determination of the protein concentration using the BCA test as described above. A small amount of the cell lysates was taken for a separate analysis of equal protein amounts. One to one mixtures of 10 μg each of BC-M1 and MDA-231 BO2 were separated by SDS-PAGE and stained by colloidal Coomassie as described. Peptides of excised gel slices were treated as described and analyzed by mass spectrometry as described below. The remaining samples were stored at -20 °C until the mass spectrometry results are available. Four couples of BC-M1- MDA-231 BO2 couples fulfilled the

requirements of equal protein loading and were further processed. The labeling (Supplementary methods from [69]) was performed by Dr. Kai Bartkowiak (University Medical Center Hamburg-Eppendorf, Institute of Tumor Biology, Germany).

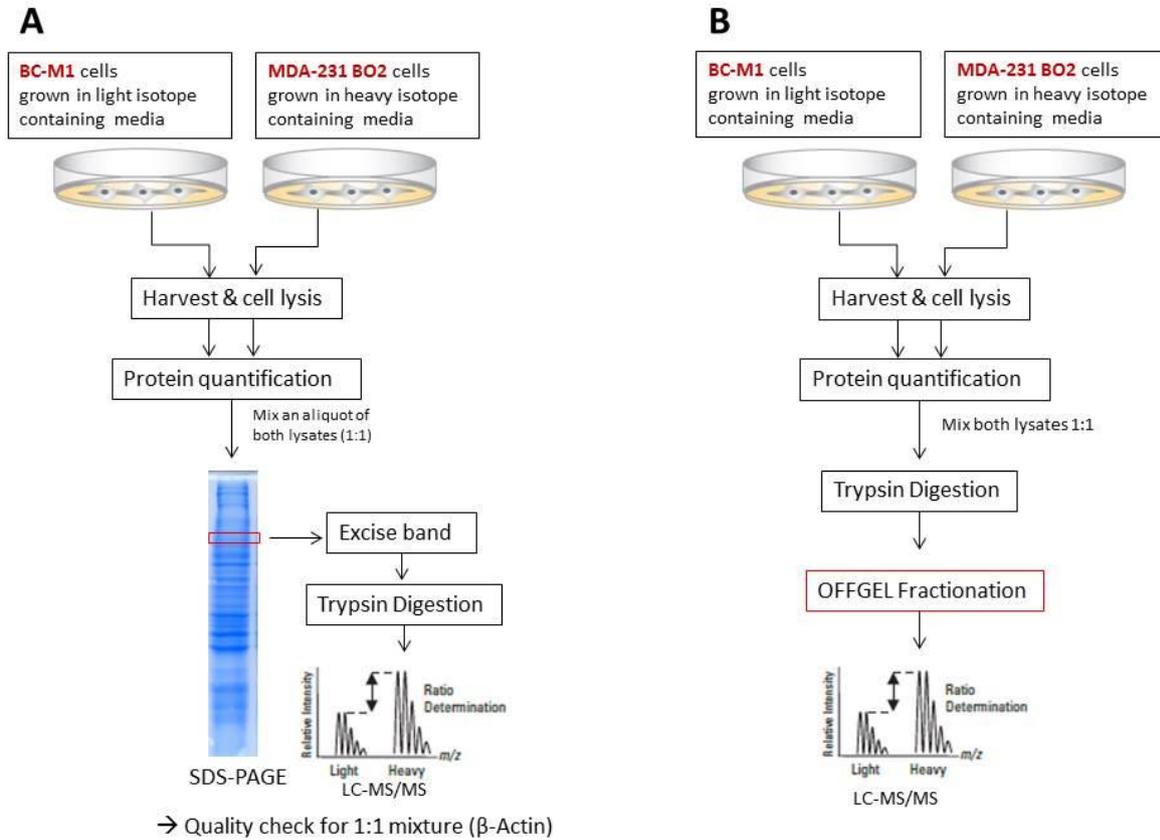


Figure 68: Overview of the workflow for SILAC-labeling of the cell lines MDA-231 BO2 and the DTC cell line BC-M1. A: Initial workflow for the analysis of the correct incorporation of the heavy isotopes and the mixture of both lysates with the same proportion. B: Final workflow for the preparative approach. The image was edited according to the template from Thermo Fisher Scientific (<https://www.thermofisher.com/de/de/home/life-science/protein-biology/protein-mass-spectrometry-analysis/protein-quantitation-mass-spectrometry/silac-metabolic-labeling-systems.html>).

8.1.14. Sample preparation for mass spectrometry

8.1.14.1. Tryptic digestion

Four hundred micrograms of each BC-M1 or MDA-231 BO2 samples were diluted with 9 volumes of 100 mM NH_4HCO_3 pH 8.3. The proteins were digested by Trypsin (Trypsin sequencing grade, Roche, Mannheim, Germany) by addition of 1 μg Trypsin per 20 μg of protein of the sample. The pH value of the samples was verified and for the subsequent analyses 380 μg of total protein of each cell line was combined in a 1:1 ratio. Thereafter the samples were incubated at 37 °C over night with vigorous shaking (Supplementary methods from [69]). This part was performed by Dr. Kai Bartkowiak (University Medical Center Hamburg-Eppendorf, Institute of Tumor Biology, Germany).

8.1.14.2. Peptide desalting

The desalting of the peptides after tryptic digestion was performed using the reversed phase Oasis HLB Plus Cartridge (225 mg, 60 μm ; Waters, Manchester, United Kingdom). The cartridge was conditioned with 3 mL 100 % methanol and equilibrated with 0.2 % formic acid in 3 mL 5 % methanol in HPLC- H_2O (Honeywell, Morris Plains, USA) using 5 mL syringes. Samples were diluted to a final volume of 1 mL using 1 % formic acid in 25 % methanol in HPLC- H_2O and the peptide-solution was transferred to the column using a 1 mL-syringe. The flow-through was kept and stored at -20° C. The column was washed with 3 mL 0.2 % formic acid in 5 % methanol in HPLC- H_2O . The peptides were eluted with 1.5 mL 50 % methanol in in HPLC- H_2O . The eluate was evaporated by using a vacuum speed vac for 3 hours.

8.1.14.3. OFFGEL fractionation

The fractionation of the desalted and evaporated peptides was performed using OFFGEL fractionator device (OFFGEL fractionator 3100, Agilent Technologies, Waldbronn, Germany) according to the isoelectric points of the peptides. A 24 well tray and IPG strips (pH 3-10 non-linear, 24 cm: GE Healthcare, Munich, Germany) were used for isoelectric focussing (IEF) of the peptides. The samples were diluted in 3.6 mL 20 % methanol in HPLC- H_2O containing 1 % IPG-buffer (pH 3-10: GE Healthcare, Munich, Germany). 40 μL of the focussing buffer was added to each well and the gel was rehydrated for 45 minutes.

200 μL focussing buffer was added to wells 1 and 24 and 100 μL focussing buffer was added to the wells 2 to 23. 150 μL of the samples were added to each well. Each electrode was covered with 10 μL H_2O double-distilled and 600 μL Dry Strip Cover Fluid (GE Healthcare, Munich, Germany) was added to each electrode. IEF was performed with a voltage range from 300 V to 2783 V and a maximum current of 50 μA until 50 kVh were reached after 33 hours and 50 minutes. After successful fractionation the samples were collected and evaporated using a vacuum speed vac.

8.1.15. LC-MS/MS analysis

The setup of the mass spectrometric analysis (**figure 69**) comprises a liquid chromatography system (reversed phase chromatography, RPC) used for separation and purification of the peptides followed by an electrospray ionization device (ESI) used for the generation of peptide ions. Thereafter, a quadrupole device is used for the specific precursor selection fragmentation or transmission of the whole mass range. The MS1-spectra are generated by a special ion trap mass analyzer (Orbitrap) with a coupled ion trap (C-trap). The fragmentation is performed by a HCD collision cell and the MS2-spectra are generated by an orbitrap.

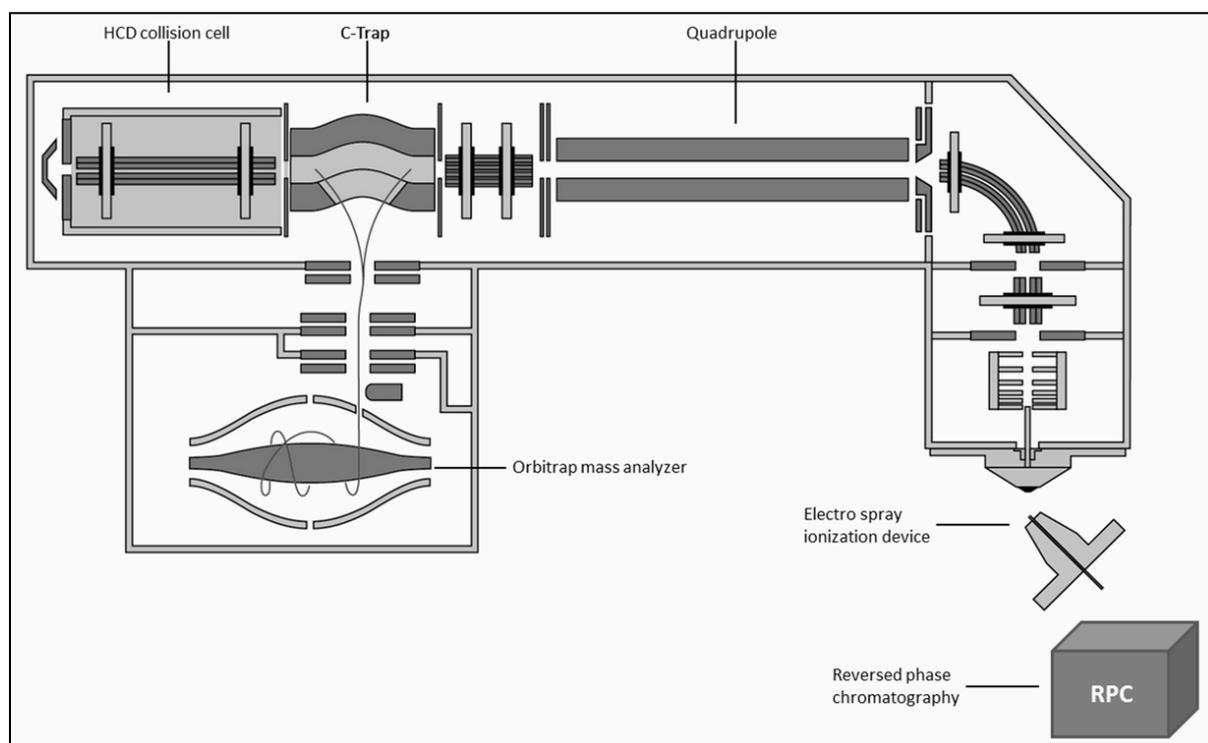


Figure 69: setup of the Quadrupole Orbitrap mass spectrometer (Q Exactive, Thermo Fisher Scientific) with coupled reversed phase chromatography. The image was edited according to the template figure 2 “Mass Spectrometry-based Proteomics Using Q Exactive, a High-performance Benchtop Quadrupole Orbitrap Mass Spectrometer” [238].

8.1.15.1. Reversed phase chromatography (RPC)

Adsorption chromatography is a technique depending on the hydrophobic interaction between molecules solubilized in the mobile phase and molecules immobilized on the column matrix (stationary phase). Reversed phase chromatography (RPC), a special adsorption chromatography variant, is an effective method for the analysis and purification of peptides. Peptides can be separated and purified by their degree of hydrophobicity with high recovery and resolution rates [239-242]. The stationary phase is composed of hydrophobic ligands like silica-alkyl [240] whereas the mobile phase consists of gradient composed of organic solvents like formic acid, acetonitrile and water [240, 241].

8.1.15.2. Electrospray ionization (ESI)

The peptide ions are directly generated from the solution by electrospray ionization (ESI). The generation of gas phase ions from the solution comprises three steps (**figure 70**): dispersion of a charge droplet spray (I), evaporation of the solvent (II) and release of the charged droplets. First, the analyte is guided through a capillary with applied high voltage. The ions move to the counter electrode. A fine mist which exits the capillary tip is generated by the formation of a Taylor cone due to the repulsion of the equally charged ions. A gas is commonly used to support the evaporation of the solvent and the nebulization of the analyte solution. Due to the step-wise evaporation of the solvent the droplet size decreases which in turn leads to an increased charge density. Below a certain drop size the droplets dissociate due to the repulsion of equal charges (Coulomb explosion) and smaller droplets are formed [243, 244].

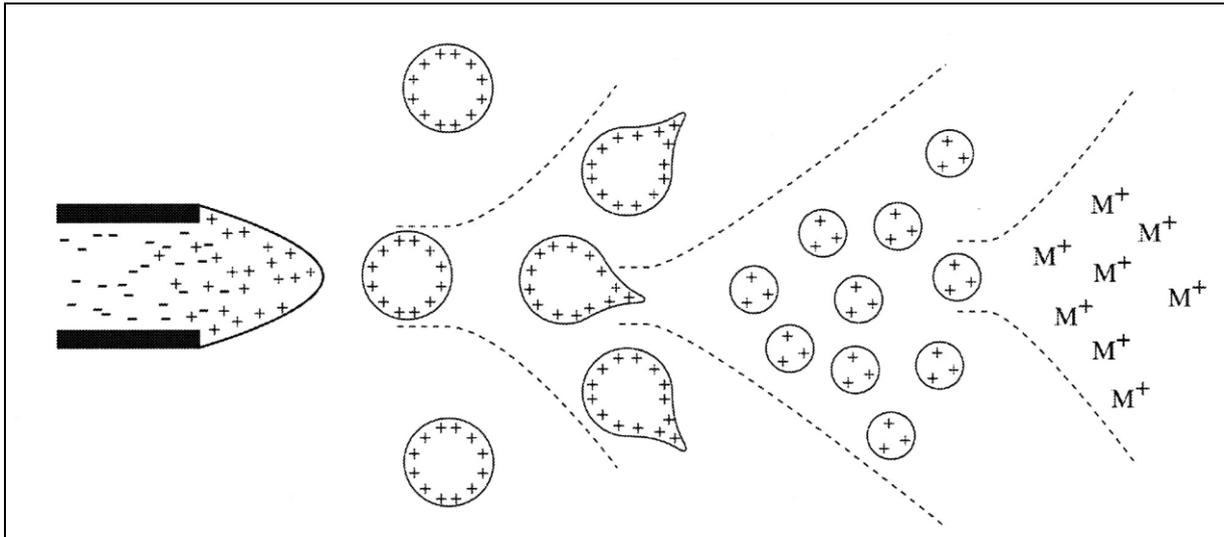


Figure 70: Formation of the charged gas phase ions by electrospray ionization. This image was edited according to the template figure 1 from “Electrospray Ionisation Mass Spectrometry: Principles and Clinical Applications” [244].

8.1.15.3. Mass analyzers

Proteomic analysis performed by mass spectrometry has become a powerful and sensitive tool to determine proteins or peptides contained in cells, tissues or body fluids [144, 238, 245-247]. The separation principle of mass spectrometry is based on the different velocity and deflection of the gas phase ions during their flight phase. The velocity and deflection is affected by the mass (m) and the charge (z) of the ions during the flight time in the electrical field [247] which means that the ions are separated by their mass-to-charge ratio (m/z) after initial acceleration. The kinetic energy (E_{kin}) of the ions can be described by the following equation (1) if the acceleration voltage (U) and the flight distance (L) are known:

$$E_{kin} = \frac{1}{2} \cdot m \cdot v^2 = z \cdot e \cdot U \quad (1)$$

m = ion mass, v = velocity after acceleration, z = charge number, e = elementary charge

$$v = \frac{L}{t} \quad (2)$$

The velocity (v) results from the division of the total flight time (t) by the flight distance (L). The equation 2 is inserted in equation which results in the following equation:

$$\frac{1}{2} \cdot m \cdot \frac{L^2}{t^2} = z \cdot e \cdot U \quad (3)$$

$$\frac{1}{2} \cdot m \cdot \frac{L^2}{t^2} = z \cdot e \cdot U \quad / \cdot 2 \cdot \frac{1}{L^2} \cdot t^2 \cdot \frac{1}{z}$$

Conversion of equation 3 generates equation 4:

$$\frac{m}{z} = \frac{2 \cdot e \cdot U}{L^2} \cdot t^2 \quad \rightarrow \quad \frac{m}{z} \propto t^2 \quad (4)$$

The ratio of mass and charge (m/z) is proportional to the square of flighttime (t^2).

There are different separation principles in mass spectrometry at the moment: quadrupole mass filters, time of flight (TOF) mass analyzers, linear ion traps, Orbitrap analyzers and combinations of these four mass analyzers [238]. A quadrupole mass analyzer (**figure 71**) is an arrangement of four parallel metal rods with equal distance. Two opposite rods are linked electrically whereas the diagonally arranged pair of rods exhibit an equal but opposite DC voltage combined with a radio frequency (RF) AC voltage. The ions move forward in z -direction with oscillation in the x - y -plane by the electrical field. The oscillation amplitude is related to the m/z ratio. The DC and RF voltages can be adjusted that only ions with specific

m/z ratios reach the detector [244]. The quadrupole TOF device in the Quadrupole Orbitrap mass spectrometer QExactive is used for the transmission of the whole range of generated ions for the full MS analysis (MS1 spectra). Additionally, the quadrupole transmits precursor ions with a defined m/z ratio for generation of the fragment ion spectrum (MS2) [238].

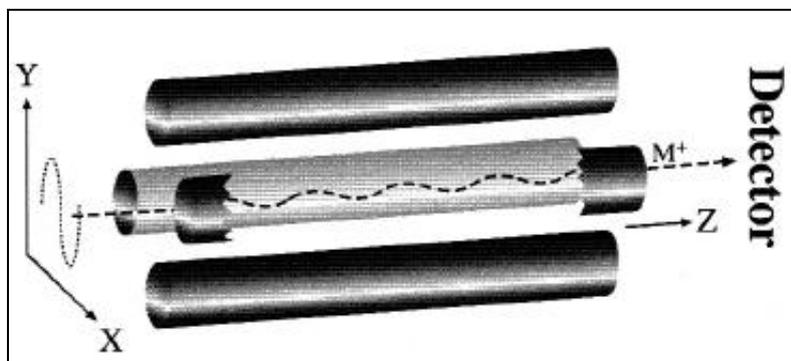


Figure 71: Typical arrangement of different parts of a Quadrupole mass analyzer. This image was edited according to the template figure 1 from “Electrospray Ionisation Mass Spectrometry: Principles and Clinical Applications” [244].

For the generation of fragment ions by collision-induced dissociation (CID) the peptides need to be fragmented using a higher energy collision-induced dissociation (HCD)[143, 248]. The HCD-cell is quadrupole filled with gas which is directly connected with the C-trap [238, 249]. The fragmentation of selected precursor ions is performed by the collision with neutral gas molecules [250]. HCD fragmentation is similar to the fragmentation procedure in triple quadrupole or quadrupole TOF devices. After successful fragmentation, the fragment ions are analyzed in the Orbitrap mass analyzer [143].

The Orbitrap mass analyzer is an electrostatic device with a quadro-logarithmic field where packets of ions are injected, trapped and analyzed. This field is generated by an axial central spindle-shaped electrode surrounded by a coaxial outer electrode (**figure 72**). The ions rotate around the inner electrode on stable trajectories with additional harmonic oscillations. The stability of the ion movement is guaranteed by the motion around the axial electrode. The frequency of the harmonic oscillations of the ions is proportional to $(m/z) - 1/2$. The frequencies of the axial harmonic oscillations of the ions are detected by image current detection on the two split halves of the outer electrode. The mass-to-charge ratio (m/z) is calculated by fast Fourier Transformation (FT) of the frequency signals of the harmonic oscillations. The C-trap is used as a storage device that guarantees that only small ion packets are injected into the Orbitrap mass analyzer [238, 249, 251, 252]

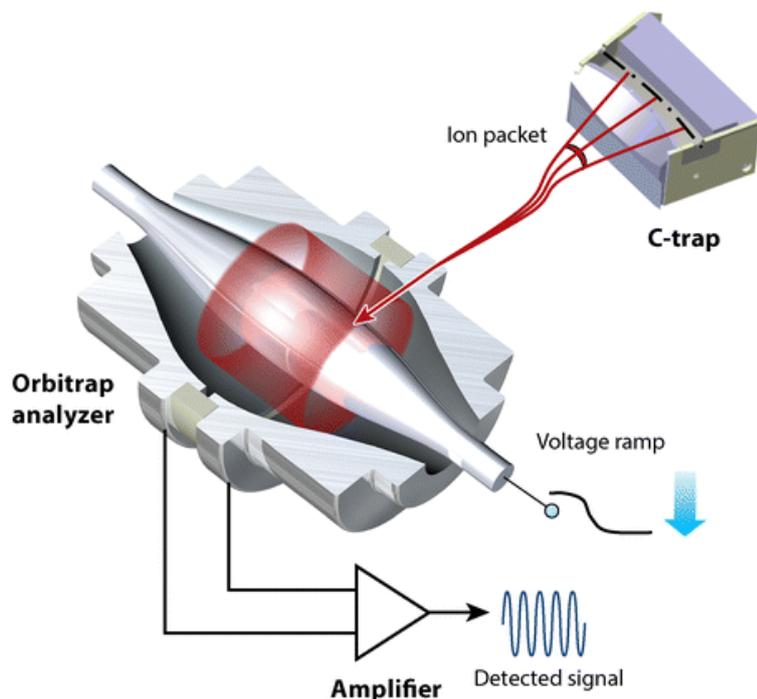


Figure 72: Cross section of the Orbitrap mass analyzer and the connected C-trap. This image was edited according to the template figure 1 from “Evolution of Orbitrap Mass Spectrometry Instrumentation” [249].

The parts “LC-MS/MS analysis” and “Data analysis” were performed by Dr. Marcel Kwiatkowski (University Medical Center Hamburg-Eppendorf, Institute for Clinical Chemistry and laboratory medicine, Germany) according to the protocol [253] (supplementary methods “Mass spectrometric proteome analysis”) with the following modifications. LC-MS/MS measurements were performed by injecting the samples on a nano liquid chromatography system (nanoACQUITYy, Waters, Manchester, UK) coupled via an electrospray ionization (ESI) interface to a quadrupole orbitrap mass spectrometer (Orbitrap QExactive, Thermo Scientific, Bremen, Germany). The samples were loaded (5 $\mu\text{L}/\text{min}$) on a trapping column (nanoACQUITY UPLC Symmetry C18 trap column, 180 $\mu\text{m} \times 20 \text{ mm}$, 5 μm , 100 \AA ; buffer A: 0.1% FA in HPLC-H₂O; buffer B: 0.1 % FA in ACN) with 2 % buffer B. After sample loading the trapping column was washed for 5 min with 2 % buffer B (5 $\mu\text{L}/\text{min}$) and the peptides were eluted (200 nL/min) onto the separation column (nanoAcquity UPLC column, BEH 130 C18, Waters; 75 $\mu\text{m} \times 250 \text{ mm}$, 1.7 μm , 100 \AA ; 200 nL/min, gradient: 2–30 % B in 30 min). The spray was generated from a fused-silica emitter (I.D. 10 μm , New Objective, Woburn, USA) at a capillary voltage of 1650 V. LC-MS/MS analysis were performed in positive ion mode. Precursor-ion scans were performed over an m/z range from 400–1500, with a resolution of 70000 FWHM at m/z 200 (transient length= 256 ms, injection time= 100 ms, AGC target= 3e6). Fragment-ion scans were carried out in data dependant acquisition (DDA) mode

(Top5), with a HCD collision energy of 30 %, a resolution of 17000 FWHM at m/z 200 (transient length= 64 ms, injection time= 100 ms, AGC target= $3e6$), an underfill ratio of 1 % and an isolation width of 2 m/z .

8.1.16. Data Analysis

Data analysis of the LC-MS/MS data were performed with MaxQuant (version 1.5.2.8) [254]. Peptide and protein identification was carried out with Andromeda against a human (Swiss-Prot, www.uniprot.org, downloaded November 10, 2014, 20,161 entries) and a contaminant database (cRAP-database, <http://www.thegpm.org/crap>, 298 entries). The search was performed with a tolerance of 10 ppm on MS- and 20 ppm on MS/MS-level. Carbamidomethylation was considered as a fixed modification on cysteine residues. Oxidation on methionine as well as $^{13}\text{C}_6$ -label on lysine and arginine were considered as variable modifications. The search was performed with a false discovery rate (FDR) of 0.01 on both peptide and protein level. SILAC quantification was carried out with MaxQuant [255] wherein only unique peptides were taken into account for SILAC quantification. Proteins for further analysis were only considered if unique peptides were detected in three out of four biological replicates and proteins were only included into the list of candidates if the protein was identified with at least two unique peptides. Proteins overexpressed in the cell line MDA-231 BO2 were considered to be differentially expressed if they have a fold change of two or higher for the intensity ratio heavy/light. Proteins overexpressed in the cell line BC-M1 were considered to be differentially expressed if they exhibit a fold change between 0 and 0.5 for the intensity ratio heavy/light. These expression ratio values for BC-M1 were converted into reciprocal values to facilitate the evaluation and the negative value was used to distinguish these values from those belonging to proteins overexpressed in the cell line MDA-231 BO2. The statistical analysis of the mass spectrometry data was performed with Student's t-test. The ratios of the quantified unique peptides were correlated with a reference value set to one to determine if the fold-change could be considered as significant. A p-value below 0.005 was considered as significant. The values for CDCP1 and for heat shock 70 kDa protein 14 (not differentially expressed) are shown in Table 1. The represented MS1- mass spectra were generated with the software Xcalibur, version 4.2.28.14 and the MS2- mass spectra were generated with Proteome Discoverer, version 2.0.0.802 (both Thermo Fisher, Waltham, USA).

8.1.17. Isolation and enumeration of tumor cells via CellSearch® System

The isolation, detection and enumeration of circulating tumor cells (CTCs) were performed by using the CellSearch® System (Menarini Silicon Biosystems, Pennsylvania, USA) [256, 257]. This system contains a semi-automated device for the blood sample preparation which is combined with the CellSearch Epithelial Kit [258]. The enrichment of the cells is obtained by ferrofluid magnetic particles [257] conjugated to anti-EpCAM antibodies and is performed using the Celltracks autprep system (Menarini Silicon Biosystems, Pennsylvania, USA). Detailed procedure for sample processing was done as described by Allard *et al.* [259]. Enriched cells are stained with a fluorescent nucleic dye (4,2-diamidino-2-phenylindole dihydrochloride, DAPI), an allophycocyan-labeled monoclonal antibody against CD45 used as exclusion marker for leukocytes and an phycoerythrin-labeled monoclonal antibody against keratins (pan keratin, K8/18/19 and others) is used to differentiate between peripheral blood and tumor cells. Scanning of the enriched cells is performed by the Celltracks analyzer II (Menarini Silicon Biosystems, Pennsylvania, USA), a semi-automated fluorescence-based microscope. CTCs were identified and evaluated as keratin-expressing round to oval nucleated cells without expression of CD45 by a trained senior scientist [258]. Further technical details of the CellSearch System and coupled devices including linearity, correctness, preciseness and reproducibility can be found elsewhere [259].

8.1.17.1. Antibody establishment for the CellSearch System with tumor cell lines

For establishing, testing the usability and determination of required concentrations and exposure times of the Alexa488-conjugated anti-human-CD45 antibody (NBP2-60224AF488) for the application to the CellSearch System 100 tumor cells from the cell lines MDA-231, MDA-468 (both serve as positive controls) and MCF-7 (negative control) were spiked into 7.5 mL EDTA-blood of a healthy female donor. Cells were detached as described in the section “Fluorescence-activated cell sorting (FACS)”. The EDTA-blood containing spiked tumor cells MDA-231/MDA-468/MCF-7 was transferred to CellSave preservative tube (Menarini Silicon Biosystems, Pennsylvania, USA). These samples were proceeded as mentioned above. The applied concentration was 1:100 with an exposure time of 1.6 seconds.

8.1.17.2. Isolation and enumeration of circulating tumor cells from breast cancer patients

For tumor cell analysis drawn blood from a breast cancer is transferred to a CellSave preservative tube. 7.5 mL of blood was added to 6.5 mL of dilution buffer, centrifuged at $800 \times g$ for ten minutes at room temperature and transferred to the CellTracks Autoprep device. After adding the previously described antibodies and reagents the following steps are performed automatically by the device. For the analysis of CDCP1 expression on tumor cells of patients Alexa488-conjugated anti-human-CDCP1 (NBP2-60224AF488, Novus Biologicals, Centennial, USA) antibody was added at a dilution of 1:100 in Dako REAL™ Antibody diluent (DAKO, Glostrup, Denmark) was added to the device. After magnetic separation magnetic-bead labeled cells are resuspended and labeled with the anti-CD45 antibody and the anti-cytokeratin antibody mentioned above. The cells will be separated again magnetically and then fixed by addition of a cell fixative. For evaluation of the analyzed samples the cartridge is transferred to the CellSpotter Analyzer and scanned four times with automatic change of fluorescence filters. The images of an epithelial CTC are shown in a gallery if all previously determined conditions are complied. The samples were analyzed by two trained persons [257-260].

8.1.18. Fluorescence-activated cell sorting (FACS)

Fluorescence-activated cell sorting (FACS) is a method which is used for the analysis of physical and chemical attributes of cell populations. This technique can also be used for cell separation procedures. The single cells pass a laser beam and the different types of scattered light give insight into different cell properties. The analysis of the plotting of the forward scattering (FSC) versus 90° side scattering (SSC) is used for the determination of a distinct cell population and for the exclusion of debris. FCS correlates with the size or volume of the cells whereas SSC correlates with internal structure or the granularity [148, 149, 261-265].

For additional validation of the properties of CDCP1 the Alexa488-conjugated antibody (NBP2-60224AF488, Novus Biologicals, Centennial, USA) was analyzed using the fluorescence-activated cell sorting (FACS) device FACSCanto II (BD Bioscience, Franklin Lakes, USA). The tumor cell lines MDA-231 served as a positive control and the cell line MCF-7 served as a negative control for FACS analysis with the anti-CDCP1-antibody. The results were analyzed with the software FACSDiva (BD Bioscience,

Franklin Lakes, USA, version 8.0.1). The adherent growing cell lines MDA-231 and MCF-7 were detached with 3 mL DPBS containing 10 mM EDTA for five minutes at 37° C. The cells were centrifuged with the Rotofix32 centrifuge (Hettich, Kirchlengern, Germany) at 1200 rpm (swing-out rotor, six-fold) for three minutes and resuspended in 5 mL DPBS. This step was repeated two times and the supernatant was discarded. 1×10^6 cells were resuspended in 300 μ L FACS-buffer (BD Bioscience, Franklin Lakes, USA). 2.5 μ g of the Alexa488-conjugated antibody (NBP2-60224AF488) was used for 1×10^6 cells. The samples were analyzed using the FACSCanto II and the analysis was kindly carried out by the FACS sorting Core Unit (University hospital Hamburg-Eppendorf, Hamburg, Germany).

8.1.19. Quantitative *Sandwich*-ELISA (Enzyme-linked Immunosorbent Assay)

The full length variant of CDCP1 (MW: 130 kDa) can be cleaved proteolytically by the proteases plasmin, trypsin and matriptase at Arginin-368 or Lysin-369 [216, 266] which leads to the generation of a membrane-bound CDCP1-fragment (MW: 70 kDa) and a soluble CDCP1-fragment (65 kDa) [121]. Cleaved CDCP1 is associated with survival of tumor cells at secondary sites in late phases of metastasis whereas the membrane-bound 70 kDa-CDCP1 fragment is thought to be relevant for cell escape from the primary tumor and intravasation [121]. A quantitative *Sandwich*-ELISA was developed and established to determine the concentration of the soluble 65 kDa CDCP1-fragment in supernatant of tumor cell lines and blood plasma of breast cancer patients. The cell culture supernatant of the breast cancer cell line MCF-7 was used as a negative control, the breast cancer cell line MDA-231 as a positive control for the method establishment and two antibodies recognizing different epitopes. The established optimal protocol is described: A 96 well microtiter plate (Sarstedt, Nümbrecht, Germany) was coated with monoclonal mouse anti-human CDCP1 antibody (Thermo Fisher, Waltham, USA, clone: 309137) as capturing antibody at a dilution of 1:150 in 0.02 % Tween-20 (Carl Roth, Karlsruhe, Germany) in 35 μ L PBS and incubated at room temperature for two hours with gentle agitation. Each well was washed one time with 100 μ L PBS and two times with 100 μ L PBS + 0.02 % Tween-20. To avoid unspecific binding each well was blocked with 5% Bovine serum albumin (BSA; Sigma-Aldrich, St. Louis, USA) in 100 μ L PBS+0.02 % Tween-20 and incubated at room temperature for two hours with gentle agitation. Each well was washed one time with 100 μ L PBS and two times with 100 μ L PBS + 0.02 % Tween-20. The samples were diluted in DMEM + 10 %FCS to a final volume of 100 μ L. For the breast cancer patients 40 μ L of blood plasma was

analyzed and the sample was incubated at 4° C over night with gentle agitation. Each well was washed one time with 100 µL PBS and two times with 100 µL PBS + 0.02 % Tween-20. Each well was incubated with polyclonal goat anti-human CDCP1 (R&D Systems, Minneapolis, USA) at a dilution of 1:150 in 35 µL PBS + 0.02 % Tween-20 and incubated at room temperature for two hours with gentle agitation. Each well was washed one time with 100 µL PBS and two times with 100 µL PBS + 0.02 % Tween-20. The assay was incubated with rabbit-anti-goat IgG antibody conjugated with horseradish peroxidase (DAKO, Glostrup, Denmark) as a detection antibody at a dilution of 1:1000 in 35 µL PBS + 0.02 % Tween-20 for 75 minutes at room temperature with gentle agitation. Each well was washed one time with 100 µL PBS and two times with 100 µL PBS + 0.02 % Tween-20. To start the reaction, 100 µL TMB-substrate solution (Cell Signaling Technology, Danvers, USA) was added to each well and incubated for 12 minutes in the dark. For termination of the reaction 100 µL STOP solution (Cell Signaling Technology, Danvers, USA) was added to each well and incubated for 30 minutes in the dark. The analysis was performed with the microplate reader Infinite NanoQuant M200 Pro (Tecan, Männedorf, Switzerland) and the microplate reader software i-Control™ (Tecan, Männedorf, Switzerland, software version 1.11). The absorption was measured at a wavelength of 450 nm and a reference wavelength of 620 nm. The recombinant human protein CDCP1 (Phe30-Ser341, Novoprotein, Summit, USA) was used as an internal control and as a protein standard to correlate protein concentration and optical density of an unknown sample of a cancer patient. The “early detection” cohort (**figure 53**) was generated by elimination of samples with unspecific or incomplete parameters in the patient data set. Therefore, 65 samples were eliminated because the blood draw was at the day of surgery or after surgery. 99 samples were deleted because the patients received a neoadjuvant therapy before blood draw. 58 samples were eliminated because a relapse was detected. 111 samples were deleted because there was no date for the biopsy or missing data for tumor staging. ROC (receiver operating characteristic) curves were calculated for the Sandwich-ELISA data from breast cancer patients and healthy individuals using the “plotROC” package version 2.2.1 in “R” (R Core Team, version 3.5.3) by Dr. Simon Joosse (University Medical Center Hamburg-Eppendorf, Department of Tumor Biology). The statistical analysis for the correlation between CDCP1 and the clinicopathological parameters of the breast cancer patients was performed with the software SPSS Statistics (version 24; IBM, Armonk, USA) by PD Dr. Leticia Oliveira-Ferrer (University Medical Centre Hamburg Eppendorf, Department of Gynecology).

8.1.20. Magnetic-activated cell sorting (MACS)

For the establishment of a MACS-based method for isolation or enrichment of CDCP1-positive CTCs/DTCs from the blood of cancer patients MCF-7 is used as a CDCP1-negative control. MDA-231 was chosen as a positive control because they are strongly positive for CDCP1 and have a characteristic cytokeratin-pattern (**figure 5 and 24**) which is important for microscopic analyses and detection of positively selected cells.

For the formation of the antibody-bead-complex 0.2 µg (1 µL) biotinylated polyclonal goat anti-human CDCP1 (catalogue number: BAF2666, Novus Biologicals, Centennial, USA) and 10 µL Streptavidin MicroBeads (diameter: 50 nm) were incubated with 100 µL autoMACS Running Buffer-MACS Separation Buffer (both Miltenyi Biotec, Bergisch Gladbach, Germany) and incubated for 30 minutes at 22 °C and shaking at 600 rpm. PBMC isolation from 7.5 mL blood of a healthy donor was performed with Ficoll-Paque as described in the section “Isolation of mononuclear cells from peripheral blood”. The PBMC pellet was diluted in 400 µL MACS Separation Buffer and 75 µL (corresponding approximately to 1.3×10^6 PBMCs) and 50 MDA-231 cells were added to preformed antibody-bead-complex and incubated for 60 minutes at 22 °C and shaking at 600 rpm. Cells were washed with 1.4 mL MACS Separation Buffer, centrifuged at 300 g for ten minutes. The supernatant was discarded and the cell pellet was diluted in 500 µL MACS Separation Buffer. Cells were loaded on a MS-column (Miltenyi Biotec, Bergisch Gladbach, Germany) which was placed in a magnetic rack before (Miltenyi Biotec, Bergisch Gladbach, Germany). The cells were washed three times with 500 µL MACS Separation Buffer and the column was removed from the magnet for elution of the cells. The elution was performed by adding 2 mL MACS Separation Buffer (1.5 mL by gravity flow and 0.5 mL by using a stamp) directly on microscope slide with small cytospin funnel. The microscope slide was dried over night and analyzed by immunofluorescent staining for cytokeratin, CD45 and DAPI as mentioned in the section “Immunocytochemical staining of cell lines”. The biotinylated capturing antibody CDCP1 was detected with Alexa546 donkey anti-goat secondary antibody at a dilution of 1:200 in 10 % AB-serum in DPBS for 45 minutes at room temperature.

8.2. Materials

8.2.1. Chemicals and solutions

Table 9: Overview of the used chemicals and solutions.

Description	Manufacturer
Acetone	Merck, Darmstadt, Germany
Acetonitrile, hypergrad for LC-MS	Merck, Darmstadt, Germany
Acrylamid	Carl Roth, Karlsruhe, Germany
Ammonium sulfate	Carl Roth, Karlsruhe, Germany
Ammonium bicarbonate	Sigma-Aldrich, Steinheim, Germany
Ammonium persulfate	Carl Roth, Karlsruhe, Germany
¹³ C ₆ -Arginine-HCl	Pierce Biotechnology, Rockford, USA
Bovine serum albumine (BSA)	Sigma-Aldrich, Steinheim, Germany
Bromphenol blue	Carl Roth, Karlsruhe, Germany
Cholera Toxin	Sigma-Aldrich, Steinheim, Germany
Cobalt (II) chloride	Sigma-Aldrich, Steinheim, Germany
Coomassie Brilliant Blue G250	Serva Electrophoresis GmbH, Heidelberg, Germany
Dithiothreitol (DTT)	Carl Roth, Karlsruhe, Germany
Dry Strip Cover Fluid PlusOne	GE Healthcare, Uppsala, Sweden
Ethanol	Theo Geyer GmbH & Co. KG, Renningen, Germany
Ethylenediaminetetraacetic acid (EDTA)	Carl Roth, Karlsruhe, Germany
Human epidermal growth factor	Miltenyi Biotec, Bergisch Gladbach, Germany
Human basic fibroblast growth factor	Miltenyi Biotec, Bergisch Gladbach, Germany
Formic acid	Merck, Darmstadt, Germany
Glycerol	Serva Electrophoresis GmbH, Heidelberg, Germany
Hydrochloric acid, 37 %	Carl Roth, Karlsruhe, Germany
Iodoacetamide	AppliChem, Darmstadt, Germany
Gentamycin	Sigma-Aldrich, Steinheim, Germany
L-glutamine	Gibco/Thermo Fisher Scientific, Waltham, USA
Glycine	Carl Roth, Karlsruhe, Germany
¹³ C ₆ -Lysine-2HCl	Pierce Biotechnology, Rockford, USA
Methanol	Theo Geyer GmbH & Co. KG, Renningen, Germany
Orthophosphoric acid (85%)	Carl Roth, Karlsruhe, Germany
Paraformaldehyde	Merck, Darmstadt, Germany
Penicillin/Streptomycin	Gibco/Thermo Fisher Scientific, Waltham, USA
Lowfat powdered milk	Carl Roth, Karlsruhe, Germany
L-proline	Pierce Biotechnology, Rockford, USA
2-propanol	Theo Geyer GmbH & Co. KG, Renningen, Germany
sodium chloride	Carl Roth, Karlsruhe, Germany
Sodium dodecyl sulfate	Carl Roth, Karlsruhe, Germany
N,N,N',N'-tetramethylethylenediamine	Sigma-Aldrich, Steinheim, Germany
2-Amino-2-(hydroxymethyl)propan-1,3-diol (TRIS)	Sigma-Aldrich, Steinheim, Germany
Triton X-100	Sigma-Aldrich, Steinheim, Germany

Trypsin (sequencing grade)	Roche, Mannheim, Germany
Tween 20	Carl Roth, Karlsruhe, Germany
Urea	Carl Roth, Karlsruhe, Germany
Urea for MS analysis	Sigma-Aldrich, Steinheim, Germany

8.2.3. Commercial kits and solutions

Table 10: Overview of the used commercially purchased solutions and kits.

Description	Manufacturer
AB-Serum, 10 %	Bio-Rad, Hercules, USA
CellSearch Circulating Tumor Cell Kit	Menarini Silicon Biosystems, Pennsylvania, USA
Dulbecco's Modified Eagle Medium (DMEM)	Pan-Biotech, Aidenbach, Germany
Fetal calf serum	Capricorn, Ebsdorfergrund, Germany
FACS-Buffer	BD Bioscience, Franklin Lakes, USA
Ficoll-Paque PLUS	GE Healthcare, Uppsala, Sweden
IST (Insulin, Transferrin, Selen)	Gibco/Thermo Fisher Scientific, Waltham, USA
MACS-Buffer (autoMACS running buffer)	Miltenyi Biotec, Bergisch Gladbach, Germany
PeqGOLD protein marker V	Pqlab, Erlangen, Germany
Pierce BCA Protein Assay Kit	Thermo Fisher Scientific, Waltham, USA
RPMI-1640	Pan-Biotech, Aidenbach, Germany
Signal Fire Plus ECL reagent	Cell Signaling Technology, Danvers, USA
STOP solution	Cell Signaling Technology, Danvers, USA
TMB substrate solution	Cell Signaling Technology, Danvers, USA
0.25 % Trypsin-EDTA	Gibco/Thermo Fisher Scientific, Waltham, USA

9. References

1. Hanahan, D. and R.A. Weinberg, *The hallmarks of cancer*. Cell, 2000. **100**(1): p. 57-70.
2. Renan, M.J., *How many mutations are required for tumorigenesis? Implications from human cancer data*. Mol Carcinog, 1993. **7**(3): p. 139-46.
3. Foulds, L., *The experimental study of tumor progression: a review*. Cancer Res, 1954. **14**(5): p. 327-39.
4. Nowell, P.C., *The clonal evolution of tumor cell populations*. Science, 1976. **194**(4260): p. 23-8.
5. Hanahan, D. and J. Folkman, *Patterns and emerging mechanisms of the angiogenic switch during tumorigenesis*. Cell, 1996. **86**(3): p. 353-64.
6. Sporn, M.B., *The war on cancer*. Lancet, 1996. **347**(9012): p. 1377-81.
7. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation*. Cell, 2011. **144**(5): p. 646-74.
8. Bray, F., et al., *Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries*. CA Cancer J Clin, 2018. **68**(6): p. 394-424.
9. Organization, W.H. *The top 10 causes of death*. [URL] 2018 April 30th, 2019; Available from: <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death>.
10. Pilleron, S., et al., *Global cancer incidence in older adults, 2012 and 2035: A population-based study*. Int J Cancer, 2019. **144**(1): p. 49-58.
11. Sorlie, T., et al., *Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications*. Proc Natl Acad Sci U S A, 2001. **98**(19): p. 10869-74.
12. Perou, C.M., et al., *Molecular portraits of human breast tumours*. Nature, 2000. **406**(6797): p. 747-52.
13. Cho, N., *Molecular subtypes and imaging phenotypes of breast cancer*. Ultrasonography, 2016. **35**(4): p. 281-8.
14. Ng, C.K., et al., *Breast cancer genomics from microarrays to massively parallel sequencing: paradigms and new insights*. J Natl Cancer Inst, 2015. **107**(5).
15. Dai, X., et al., *Breast cancer intrinsic subtype classification, clinical use and future trends*. Am J Cancer Res, 2015. **5**(10): p. 2929-43.
16. Cancer Genome Atlas, N., *Comprehensive molecular portraits of human breast tumours*. Nature, 2012. **490**(7418): p. 61-70.
17. Parker, J.S., et al., *Supervised risk predictor of breast cancer based on intrinsic subtypes*. J Clin Oncol, 2009. **27**(8): p. 1160-7.
18. Sotiriou, C., et al., *Breast cancer classification and prognosis based on gene expression profiles from a population-based study*. Proc Natl Acad Sci U S A, 2003. **100**(18): p. 10393-8.
19. Prat, A., et al., *Clinical implications of the intrinsic molecular subtypes of breast cancer*. Breast, 2015. **24 Suppl 2**: p. S26-35.
20. Wirapati, P., et al., *Meta-analysis of gene expression profiles in breast cancer: toward a unified understanding of breast cancer subtyping and prognosis signatures*. Breast Cancer Res, 2008. **10**(4): p. R65.
21. Sorlie, T., et al., *Repeated observation of breast tumor subtypes in independent gene expression data sets*. Proc Natl Acad Sci U S A, 2003. **100**(14): p. 8418-23.
22. Brenton, J.D., et al., *Molecular classification and molecular forecasting of breast cancer: ready for clinical application?* J Clin Oncol, 2005. **23**(29): p. 7350-60.
23. Coates, A.S., et al., *Tailoring therapies--improving the management of early breast cancer: St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2015*. Ann Oncol, 2015. **26**(8): p. 1533-46.
24. Arteaga, C.L., et al., *Treatment of HER2-positive breast cancer: current status and future perspectives*. Nat Rev Clin Oncol, 2011. **9**(1): p. 16-32.
25. Pohlmann, P.R., I.A. Mayer, and R. Mernaugh, *Resistance to Trastuzumab in Breast Cancer*. Clin Cancer Res, 2009. **15**(24): p. 7479-7491.

26. Prat, A., et al., *Molecular characterization of basal-like and non-basal-like triple-negative breast cancer*. *Oncologist*, 2013. **18**(2): p. 123-33.
27. Lehmann, B.D., et al., *Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies*. *J Clin Invest*, 2011. **121**(7): p. 2750-67.
28. Abd El-Rehim, D.M., et al., *High-throughput protein expression analysis using tissue microarray technology of a large well-characterised series identifies biologically distinct classes of breast cancer confirming recent cDNA expression analyses*. *Int J Cancer*, 2005. **116**(3): p. 340-50.
29. O'Brien, K.M., et al., *Intrinsic breast tumor subtypes, race, and long-term survival in the Carolina Breast Cancer Study*. *Clin Cancer Res*, 2010. **16**(24): p. 6100-10.
30. Fan, C., et al., *Concordance among gene-expression-based predictors for breast cancer*. *N Engl J Med*, 2006. **355**(6): p. 560-9.
31. Carey, L.A., et al., *Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study*. *JAMA*, 2006. **295**(21): p. 2492-502.
32. Wright, H.J., A.M. Police, and O.V. Razorenova, *Targeting CDCP1 dimerization in triple-negative breast cancer*. *Cell Cycle*, 2016. **15**(18): p. 2385-6.
33. Foulkes, W.D., I.E. Smith, and J.S. Reis-Filho, *Triple-negative breast cancer*. *N Engl J Med*, 2010. **363**(20): p. 1938-48.
34. Dias, K., et al., *Claudin-Low Breast Cancer; Clinical & Pathological Characteristics*. *PLoS One*, 2017. **12**(1): p. e0168669.
35. Joosse, S.A., T.M. Gorges, and K. Pantel, *Biology, detection, and clinical implications of circulating tumor cells*. *EMBO Mol Med*, 2015. **7**(1): p. 1-11.
36. Kang, Y. and K. Pantel, *Tumor cell dissemination: emerging biological insights from animal models and cancer patients*. *Cancer Cell*, 2013. **23**(5): p. 573-81.
37. Chambers, A.F., A.C. Groom, and I.C. MacDonald, *Dissemination and growth of cancer cells in metastatic sites*. *Nat Rev Cancer*, 2002. **2**(8): p. 563-72.
38. Uhr, J.W. and K. Pantel, *Controversies in clinical cancer dormancy*. *Proc Natl Acad Sci U S A*, 2011. **108**(30): p. 12396-400.
39. Ell, B., et al., *Tumor-induced osteoclast miRNA changes as regulators and biomarkers of osteolytic bone metastasis*. *Cancer Cell*, 2013. **24**(4): p. 542-56.
40. Lu, X., et al., *VCAM-1 promotes osteolytic expansion of indolent bone micrometastasis of breast cancer by engaging alpha4beta1-positive osteoclast progenitors*. *Cancer Cell*, 2011. **20**(6): p. 701-14.
41. Weidenfeld, K. and D. Barkan, *EMT and Stemness in Tumor Dormancy and Outgrowth: Are They Intertwined Processes?* *Front Oncol*, 2018. **8**: p. 381.
42. Celia-Terrassa, T., et al., *Epithelial-mesenchymal transition can suppress major attributes of human epithelial tumor-initiating cells*. *J Clin Invest*, 2012. **122**(5): p. 1849-68.
43. Thiery, J.P., et al., *Epithelial-mesenchymal transitions in development and disease*. *Cell*, 2009. **139**(5): p. 871-90.
44. Pantel, K., C. Alix-Panabieres, and S. Riethdorf, *Cancer micrometastases*. *Nat Rev Clin Oncol*, 2009. **6**(6): p. 339-51.
45. Becker, S., et al., *Primary systemic therapy does not eradicate disseminated tumor cells in breast cancer patients*. *Breast Cancer Res Treat*, 2007. **106**(2): p. 239-43.
46. Becker, S., et al., *Detection of cytokeratin-positive cells in the bone marrow of breast cancer patients undergoing adjuvant therapy*. *Breast Cancer Res Treat*, 2006. **97**(1): p. 91-6.
47. Muller, V., et al., *Circulating tumor cells in breast cancer: correlation to bone marrow micrometastases, heterogeneous response to systemic therapy and low proliferative activity*. *Clin Cancer Res*, 2005. **11**(10): p. 3678-85.
48. Reya, T., et al., *Stem cells, cancer, and cancer stem cells*. *Nature*, 2001. **414**(6859): p. 105-11.
49. Mani, S.A., et al., *The epithelial-mesenchymal transition generates cells with properties of stem cells*. *Cell*, 2008. **133**(4): p. 704-15.
50. Barriere, G., et al., *Circulating tumor cells and epithelial, mesenchymal and stemness markers: characterization of cell subpopulations*. *Ann Transl Med*, 2014. **2**(11): p. 109.
51. Thiery, J.P., *Epithelial-mesenchymal transitions in tumour progression*. *Nat Rev Cancer*, 2002. **2**(6): p. 442-54.

52. Ye, X. and R.A. Weinberg, *Epithelial-Mesenchymal Plasticity: A Central Regulator of Cancer Progression*. Trends Cell Biol, 2015. **25**(11): p. 675-686.
53. Kalluri, R. and R.A. Weinberg, *The basics of epithelial-mesenchymal transition*. J Clin Invest, 2009. **119**(6): p. 1420-8.
54. Nieto, M.A., *Epithelial plasticity: a common theme in embryonic and cancer cells*. Science, 2013. **342**(6159): p. 1234850.
55. Mitchell, M.J. and M.R. King, *Computational and experimental models of cancer cell response to fluid shear stress*. Front Oncol, 2013. **3**: p. 44.
56. Reymond, N., B.B. d'Agua, and A.J. Ridley, *Crossing the endothelial barrier during metastasis*. Nat Rev Cancer, 2013. **13**(12): p. 858-70.
57. Tam, W.L. and R.A. Weinberg, *The epigenetics of epithelial-mesenchymal plasticity in cancer*. Nat Med, 2013. **19**(11): p. 1438-49.
58. Bednarz-Knoll, N., C. Alix-Panabieres, and K. Pantel, *Plasticity of disseminating cancer cells in patients with epithelial malignancies*. Cancer Metastasis Rev, 2012. **31**(3-4): p. 673-87.
59. El Touny, L.H., et al., *Combined SFK/MEK inhibition prevents metastatic outgrowth of dormant tumor cells*. J Clin Invest, 2014. **124**(1): p. 156-68.
60. Bragado, P., et al., *TGF-beta2 dictates disseminated tumour cell fate in target organs through TGF-beta-RIII and p38alpha/beta signalling*. Nat Cell Biol, 2013. **15**(11): p. 1351-61.
61. Joosse, S.A., et al., *Changes in keratin expression during metastatic progression of breast cancer: impact on the detection of circulating tumor cells*. Clin Cancer Res, 2012. **18**(4): p. 993-1003.
62. Harper, K.L., et al., *Mechanism of early dissemination and metastasis in Her2(+) mammary cancer*. Nature, 2016. **540**(7634): p. 588-592.
63. Yu, M., et al., *Circulating breast tumor cells exhibit dynamic changes in epithelial and mesenchymal composition*. Science, 2013. **339**(6119): p. 580-4.
64. Joseph, J.P., et al., *Hypoxia induced EMT: A review on the mechanism of tumor progression and metastasis in OSCC*. Oral Oncol, 2018. **80**: p. 23-32.
65. Liu, Z.J., G.L. Semenza, and H.F. Zhang, *Hypoxia-inducible factor 1 and breast cancer metastasis*. J Zhejiang Univ Sci B, 2015. **16**(1): p. 32-43.
66. Maltepe, E. and O.D. Saugstad, *Oxygen in health and disease: regulation of oxygen homeostasis--clinical implications*. Pediatr Res, 2009. **65**(3): p. 261-8.
67. Liu, W., et al., *Targeted genes and interacting proteins of hypoxia inducible factor-1*. Int J Biochem Mol Biol, 2012. **3**(2): p. 165-78.
68. McKeown, S.R., *Defining normoxia, physoxia and hypoxia in tumours-implications for treatment response*. Br J Radiol, 2014. **87**(1035): p. 20130676.
69. Bartkowiak, K., et al., *Disseminated Tumor Cells Persist in the Bone Marrow of Breast Cancer Patients through Sustained Activation of the Unfolded Protein Response*. Cancer Res, 2015. **75**(24): p. 5367-77.
70. Ameri, K., et al., *Circulating tumour cells demonstrate an altered response to hypoxia and an aggressive phenotype*. Br J Cancer, 2010. **102**(3): p. 561-9.
71. Bristow, R.G. and R.P. Hill, *Hypoxia and metabolism. Hypoxia, DNA repair and genetic instability*. Nat Rev Cancer, 2008. **8**(3): p. 180-92.
72. Pantel, K. and C. Alix-Panabieres, *Bone marrow as a reservoir for disseminated tumor cells: a special source for liquid biopsy in cancer patients*. Bonekey Rep, 2014. **3**: p. 584.
73. Janni, W., et al., *Persistence of disseminated tumor cells in the bone marrow of breast cancer patients predicts increased risk for relapse--a European pooled analysis*. Clin Cancer Res, 2011. **17**(9): p. 2967-76.
74. Braun, S., et al., *A pooled analysis of bone marrow micrometastasis in breast cancer*. N Engl J Med, 2005. **353**(8): p. 793-802.
75. Braun, S., et al., *Lack of effect of adjuvant chemotherapy on the elimination of single dormant tumor cells in bone marrow of high-risk breast cancer patients*. J Clin Oncol, 2000. **18**(1): p. 80-6.
76. Wouters, B.G. and M. Koritzinsky, *Hypoxia signalling through mTOR and the unfolded protein response in cancer*. Nat Rev Cancer, 2008. **8**(11): p. 851-64.

77. Parmar, K., et al., *Distribution of hematopoietic stem cells in the bone marrow according to regional hypoxia*. Proc Natl Acad Sci U S A, 2007. **104**(13): p. 5431-6.
78. Wang, G.L., et al., *Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O₂ tension*. Proc Natl Acad Sci U S A, 1995. **92**(12): p. 5510-4.
79. Lando, D., et al., *Asparagine hydroxylation of the HIF transactivation domain a hypoxic switch*. Science, 2002. **295**(5556): p. 858-61.
80. Mahon, P.C., K. Hirota, and G.L. Semenza, *FIH-1: a novel protein that interacts with HIF-1alpha and VHL to mediate repression of HIF-1 transcriptional activity*. Genes Dev, 2001. **15**(20): p. 2675-86.
81. Semenza, G.L., *Hypoxia-inducible factors: mediators of cancer progression and targets for cancer therapy*. Trends Pharmacol Sci, 2012. **33**(4): p. 207-14.
82. Bache, M., et al., *Detection and specific targeting of hypoxic regions within solid tumors: current preclinical and clinical strategies*. Curr Med Chem, 2008. **15**(4): p. 322-38.
83. Semenza, G.L., *Targeting HIF-1 for cancer therapy*. Nat Rev Cancer, 2003. **3**(10): p. 721-32.
84. Hooper, J.D., et al., *Subtractive immunization using highly metastatic human tumor cells identifies SIMA135/CDCP1, a 135 kDa cell surface phosphorylated glycoprotein antigen*. Oncogene, 2003. **22**(12): p. 1783-94.
85. Brown, T.A., et al., *Adhesion or plasmin regulates tyrosine phosphorylation of a novel membrane glycoprotein p80/gp140/CUB domain-containing protein 1 in epithelia*. J Biol Chem, 2004. **279**(15): p. 14772-83.
86. Bhatt, A.S., et al., *Adhesion signaling by a novel mitotic substrate of src kinases*. Oncogene, 2005. **24**(34): p. 5333-43.
87. Scherl-Mostageer, M., et al., *Identification of a novel gene, CDCP1, overexpressed in human colorectal cancer*. Oncogene, 2001. **20**(32): p. 4402-8.
88. Wortmann, A., et al., *The cell surface glycoprotein CDCP1 in cancer--insights, opportunities, and challenges*. IUBMB Life, 2009. **61**(7): p. 723-30.
89. Xia, Y., S.G. Gil, and W.G. Carter, *Anchorage mediated by integrin alpha6beta4 to laminin 5 (epiligrin) regulates tyrosine phosphorylation of a membrane-associated 80-kD protein*. J Cell Biol, 1996. **132**(4): p. 727-40.
90. Buhring, H.J., et al., *CDCP1 identifies a broad spectrum of normal and malignant stem/progenitor cell subsets of hematopoietic and nonhematopoietic origin*. Stem Cells, 2004. **22**(3): p. 334-43.
91. Ikeda, J.I., et al., *Epigenetic regulation of the expression of the novel stem cell marker CDCP1 in cancer cells*. J Pathol, 2006. **210**(1): p. 75-84.
92. Uekita, T., et al., *CUB-domain-containing protein 1 regulates peritoneal dissemination of gastric scirrhous carcinoma*. Am J Pathol, 2008. **172**(6): p. 1729-39.
93. Chaudhary, B., et al., *Neuropilin 1: function and therapeutic potential in cancer*. Cancer Immunol Immunother, 2014. **63**(2): p. 81-99.
94. Cimato, T., et al., *Neuropilin-1 identifies endothelial precursors in human and murine embryonic stem cells before CD34 expression*. Circulation, 2009. **119**(16): p. 2170-8.
95. Pellet-Many, C., et al., *Neuropilins: structure, function and role in disease*. Biochem J, 2008. **411**(2): p. 211-26.
96. Nakamura, F. and Y. Goshima, *Structural and functional relation of neuropilins*. Adv Exp Med Biol, 2002. **515**: p. 55-69.
97. Giger, R.J., et al., *Neuropilin-2 is a receptor for semaphorin IV: insight into the structural basis of receptor function and specificity*. Neuron, 1998. **21**(5): p. 1079-92.
98. Schwarz, Q. and C. Ruhrberg, *Neuropilin, you gotta let me know: should I stay or should I go?* Cell Adh Migr, 2010. **4**(1): p. 61-6.
99. Gu, C., et al., *Characterization of neuropilin-1 structural features that confer binding to semaphorin 3A and vascular endothelial growth factor 165*. J Biol Chem, 2002. **277**(20): p. 18069-76.
100. Roth, L., et al., *Transmembrane domain interactions control biological functions of neuropilin-1*. Mol Biol Cell, 2008. **19**(2): p. 646-54.

101. Seerapu, H.R., et al., *The cytoplasmic domain of neuropilin-1 regulates focal adhesion turnover*. FEBS Lett, 2013. **587**(21): p. 3392-9.
102. Lanahan, A., et al., *The neuropilin 1 cytoplasmic domain is required for VEGF-A-dependent arteriogenesis*. Dev Cell, 2013. **25**(2): p. 156-68.
103. Janssen, B.J., et al., *Neuropilins lock secreted semaphorins onto plexins in a ternary signaling complex*. Nat Struct Mol Biol, 2012. **19**(12): p. 1293-9.
104. Djordjevic, S. and P.C. Driscoll, *Targeting VEGF signalling via the neuropilin co-receptor*. Drug Discov Today, 2013. **18**(9-10): p. 447-55.
105. Soker, S., et al., *VEGF165 mediates formation of complexes containing VEGFR-2 and neuropilin-1 that enhance VEGF165-receptor binding*. J Cell Biochem, 2002. **85**(2): p. 357-68.
106. Glinka, Y., et al., *Neuropilin-1 exerts co-receptor function for TGF-beta-1 on the membrane of cancer cells and enhances responses to both latent and active TGF-beta*. Carcinogenesis, 2011. **32**(4): p. 613-21.
107. Glinka, Y. and G.J. Prud'homme, *Neuropilin-1 is a receptor for transforming growth factor beta-1, activates its latent form, and promotes regulatory T cell activity*. J Leukoc Biol, 2008. **84**(1): p. 302-10.
108. Schuch, G., et al., *In vivo administration of vascular endothelial growth factor (VEGF) and its antagonist, soluble neuropilin-1, predicts a role of VEGF in the progression of acute myeloid leukemia in vivo*. Blood, 2002. **100**(13): p. 4622-8.
109. Yasuoka, H., et al., *Neuropilin-2 expression in breast cancer: correlation with lymph node metastasis, poor prognosis, and regulation of CXCR4 expression*. BMC Cancer, 2009. **9**: p. 220.
110. Nakamura, Y., et al., *Clinicopathological significance of vascular endothelial growth factor-C in breast carcinoma with long-term follow-up*. Mod Pathol, 2003. **16**(4): p. 309-14.
111. Stacker, S.A., et al., *Lymphangiogenesis and cancer metastasis*. Nat Rev Cancer, 2002. **2**(8): p. 573-83.
112. Chen, H., et al., *Neuropilin-2, a novel member of the neuropilin family, is a high affinity receptor for the semaphorins Sema E and Sema IV but not Sema III*. Neuron, 1997. **19**(3): p. 547-59.
113. Schmoker, A.M., A.M. Ebert, and B.A. Ballif, *The DCBLD receptor family: emerging signaling roles in development, homeostasis and disease*. Biochem J, 2019. **476**(6): p. 931-950.
114. Kobuke, K., et al., *ESDN, a novel neuropilin-like membrane protein cloned from vascular cells with the longest secretory signal sequence among eukaryotes, is up-regulated after vascular injury*. J Biol Chem, 2001. **276**(36): p. 34105-14.
115. Feng, H., et al., *EGFR phosphorylation of DCBLD2 recruits TRAF6 and stimulates AKT-promoted tumorigenesis*. J Clin Invest, 2014. **124**(9): p. 3741-56.
116. Kim, M., et al., *Epigenetic down-regulation and suppressive role of DCBLD2 in gastric cancer cell proliferation and invasion*. Mol Cancer Res, 2008. **6**(2): p. 222-30.
117. Brown, C.E. and C.L. Mackall, *CAR T cell therapy: inroads to response and resistance*. Nat Rev Immunol, 2019. **19**(2): p. 73-74.
118. Mussolino, C., et al., *Genome and Epigenome Editing to Treat Disorders of the Hematopoietic System*. Hum Gene Ther, 2017. **28**(11): p. 1105-1115.
119. Fesnak, A.D., C.H. June, and B.L. Levine, *Engineered T cells: the promise and challenges of cancer immunotherapy*. Nat Rev Cancer, 2016. **16**(9): p. 566-81.
120. Konigsberg, R., et al., *Detection of EpCAM positive and negative circulating tumor cells in metastatic breast cancer patients*. Acta Oncol, 2011. **50**(5): p. 700-10.
121. Casar, B., et al., *In vivo cleaved CDCP1 promotes early tumor dissemination via complexing with activated beta1 integrin and induction of FAK/PI3K/Akt motility signaling*. Oncogene, 2014. **33**(2): p. 255-68.
122. Grutzkau, A. and A. Radbruch, *Small but mighty: how the MACS-technology based on nanosized superparamagnetic particles has helped to analyze the immune system within the last 20 years*. Cytometry A, 2010. **77**(7): p. 643-7.
123. Haukanes, B.I. and C. Kvam, *Application of magnetic beads in bioassays*. Biotechnology (N Y), 1993. **11**(1): p. 60-3.
124. Miltenyi, S., et al., *High gradient magnetic cell separation with MACS*. Cytometry, 1990. **11**(2): p. 231-8.

125. Waks, A.G. and E.P. Winer, *Breast Cancer Treatment: A Review*. JAMA, 2019. **321**(3): p. 288-300.
126. Caughran, J., et al., *The Effect of the 2009 USPSTF breast cancer screening recommendations on breast cancer in Michigan: A longitudinal study*. Breast J, 2018. **24**(5): p. 730-737.
127. Early Breast Cancer Trialists' Collaborative, G., *Long-term outcomes for neoadjuvant versus adjuvant chemotherapy in early breast cancer: meta-analysis of individual patient data from ten randomised trials*. Lancet Oncol, 2018. **19**(1): p. 27-39.
128. Rastogi, P., et al., *Preoperative chemotherapy: updates of National Surgical Adjuvant Breast and Bowel Project Protocols B-18 and B-27*. J Clin Oncol, 2008. **26**(5): p. 778-85.
129. Mladenov, E., et al., *DNA double-strand break repair as determinant of cellular radiosensitivity to killing and target in radiation therapy*. Front Oncol, 2013. **3**: p. 113.
130. Haviland, J.S., et al., *The UK Standardisation of Breast Radiotherapy (START) trials of radiotherapy hypofractionation for treatment of early breast cancer: 10-year follow-up results of two randomised controlled trials*. Lancet Oncol, 2013. **14**(11): p. 1086-1094.
131. Whelan, T.J., et al., *Long-term results of hypofractionated radiation therapy for breast cancer*. N Engl J Med, 2010. **362**(6): p. 513-20.
132. Fisher, B., et al., *Twenty-year follow-up of a randomized trial comparing total mastectomy, lumpectomy, and lumpectomy plus irradiation for the treatment of invasive breast cancer*. N Engl J Med, 2002. **347**(16): p. 1233-41.
133. Hammond, M.E., et al., *American Society of Clinical Oncology/College Of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer*. J Clin Oncol, 2010. **28**(16): p. 2784-95.
134. Piccart-Gebhart, M.J., et al., *Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer*. N Engl J Med, 2005. **353**(16): p. 1659-72.
135. Early Breast Cancer Trialists' Collaborative, G., et al., *Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials*. Lancet, 2011. **378**(9793): p. 771-84.
136. Joosse, S.A. and K. Pantel, *Biologic challenges in the detection of circulating tumor cells*. Cancer Res, 2013. **73**(1): p. 8-11.
137. Armstrong, A.J., et al., *Circulating tumor cells from patients with advanced prostate and breast cancer display both epithelial and mesenchymal markers*. Mol Cancer Res, 2011. **9**(8): p. 997-1007.
138. Chavez, K.J., S.V. Garimella, and S. Lipkowitz, *Triple negative breast cancer cell lines: one tool in the search for better treatment of triple negative breast cancer*. Breast Dis, 2010. **32**(1-2): p. 35-48.
139. Brinkley, B.R., et al., *Variations in cell form and cytoskeleton in human breast carcinoma cells in vitro*. Cancer Res, 1980. **40**(9): p. 3118-29.
140. Cailleau, R., M. Olive, and Q.V. Cruciger, *Long-term human breast carcinoma cell lines of metastatic origin: preliminary characterization*. In Vitro, 1978. **14**(11): p. 911-5.
141. Cailleau, R., et al., *Breast tumor cell lines from pleural effusions*. J Natl Cancer Inst, 1974. **53**(3): p. 661-74.
142. Hackett, A.J., et al., *Two syngeneic cell lines from human breast tissue: the aneuploid mammary epithelial (Hs578T) and the diploid myoepithelial (Hs578Bst) cell lines*. J Natl Cancer Inst, 1977. **58**(6): p. 1795-806.
143. Olsen, J.V., et al., *Higher-energy C-trap dissociation for peptide modification analysis*. Nat Methods, 2007. **4**(9): p. 709-12.
144. Biemann, K., *Mass spectrometry of peptides and proteins*. Annu Rev Biochem, 1992. **61**: p. 977-1010.
145. Biemann, K., *Contributions of mass spectrometry to peptide and protein structure*. Biomed Environ Mass Spectrom, 1988. **16**(1-12): p. 99-111.
146. Roepstorff, P. and J. Fohlman, *Proposal for a common nomenclature for sequence ions in mass spectra of peptides*. Biomed Mass Spectrom, 1984. **11**(11): p. 601.

147. Kuske, A., *Detektion, molekulare und funktionale Charakterisierung zirkulierender Tumorzellen von Patienten mit Prostata- oder Mammakarzinom*, in *Institute of Tumor Biology*. 2016, University of Hamburg: Hamburg.
148. Loken, M.R. and L.A. Herzenber, *Analysis of cell populations with a fluorescence-activated cell sorter*. *Ann N Y Acad Sci*, 1975. **254**: p. 163-71.
149. Bernath, K., et al., *In vitro compartmentalization by double emulsions: sorting and gene enrichment by fluorescence activated cell sorting*. *Anal Biochem*, 2004. **325**(1): p. 151-7.
150. Youden, W.J., *Index for rating diagnostic tests*. *Cancer*, 1950. **3**(1): p. 32-5.
151. Quaranta, V., et al., *Invasion emerges from cancer cell adaptation to competitive microenvironments: quantitative predictions from multiscale mathematical models*. *Semin Cancer Biol*, 2008. **18**(5): p. 338-48.
152. Anderson, A.R., et al., *Tumor morphology and phenotypic evolution driven by selective pressure from the microenvironment*. *Cell*, 2006. **127**(5): p. 905-15.
153. Spencer, J.A., et al., *Direct measurement of local oxygen concentration in the bone marrow of live animals*. *Nature*, 2014. **508**(7495): p. 269-73.
154. Meng, S., et al., *Circulating tumor cells in patients with breast cancer dormancy*. *Clin Cancer Res*, 2004. **10**(24): p. 8152-62.
155. Pantel, K., R.H. Brakenhoff, and B. Brandt, *Detection, clinical relevance and specific biological properties of disseminating tumour cells*. *Nat Rev Cancer*, 2008. **8**(5): p. 329-40.
156. Eccles, S.A. and D.R. Welch, *Metastasis: recent discoveries and novel treatment strategies*. *Lancet*, 2007. **369**(9574): p. 1742-57.
157. Pantel, K. and R.H. Brakenhoff, *Dissecting the metastatic cascade*. *Nat Rev Cancer*, 2004. **4**(6): p. 448-56.
158. Gangnus, R., et al., *Genomic profiling of viable and proliferative micrometastatic cells from early-stage breast cancer patients*. *Clin Cancer Res*, 2004. **10**(10): p. 3457-64.
159. Pantel, K., et al., *Detection and clinical implications of early systemic tumor cell dissemination in breast cancer*. *Clin Cancer Res*, 2003. **9**(17): p. 6326-34.
160. Braun, S., et al., *Cytokeratin-positive cells in the bone marrow and survival of patients with stage I, II, or III breast cancer*. *N Engl J Med*, 2000. **342**(8): p. 525-33.
161. Zhang, S., et al., *Mesenchymal phenotype of circulating tumor cells is associated with distant metastasis in breast cancer patients*. *Cancer Manag Res*, 2017. **9**: p. 691-700.
162. Wilhelm, S.M., et al., *BAY 43-9006 exhibits broad spectrum oral antitumor activity and targets the RAF/MEK/ERK pathway and receptor tyrosine kinases involved in tumor progression and angiogenesis*. *Cancer Res*, 2004. **64**(19): p. 7099-109.
163. Sin, W.C. and C.L. Lim, *Breast cancer stem cells-from origins to targeted therapy*. *Stem Cell Investig*, 2017. **4**: p. 96.
164. Deng, X., et al., *CD24 Expression and differential resistance to chemotherapy in triple-negative breast cancer*. *Oncotarget*, 2017. **8**(24): p. 38294-38308.
165. Birzele, F., et al., *CD44 Isoform Status Predicts Response to Treatment with Anti-CD44 Antibody in Cancer Patients*. *Clin Cancer Res*, 2015. **21**(12): p. 2753-62.
166. Gotte, M. and G.W. Yip, *Heparanase, hyaluronan, and CD44 in cancers: a breast carcinoma perspective*. *Cancer Res*, 2006. **66**(21): p. 10233-7.
167. Shirure, V.S., et al., *CD44 variant isoforms expressed by breast cancer cells are functional E-selectin ligands under flow conditions*. *Am J Physiol Cell Physiol*, 2015. **308**(1): p. C68-78.
168. Ivaska, J., et al., *Novel functions of vimentin in cell adhesion, migration, and signaling*. *Exp Cell Res*, 2007. **313**(10): p. 2050-62.
169. Derycke, L.D. and M.E. Bracke, *N-cadherin in the spotlight of cell-cell adhesion, differentiation, embryogenesis, invasion and signalling*. *Int J Dev Biol*, 2004. **48**(5-6): p. 463-76.
170. Choi, Y., et al., *Epithelial-mesenchymal transition increases during the progression of in situ to invasive basal-like breast cancer*. *Hum Pathol*, 2013. **44**(11): p. 2581-9.
171. Chan, Y.M., et al., *Mutations in the non-helical linker segment L1-2 of keratin 5 in patients with Weber-Cockayne epidermolysis bullosa simplex*. *J Cell Sci*, 1994. **107 (Pt 4)**: p. 765-74.
172. Kanapathy Pillai, S.K., et al., *Triple-negative breast cancer is associated with EGFR, CK5/6 and c-KIT expression in Malaysian women*. *BMC Clin Pathol*, 2012. **12**: p. 18.

173. Abd El-Rehim, D.M., et al., *Expression of luminal and basal cytokeratins in human breast carcinoma*. J Pathol, 2004. **203**(2): p. 661-71.
174. Bartkowiak, K., et al., *Discovery of a novel unfolded protein response phenotype of cancer stem/progenitor cells from the bone marrow of breast cancer patients*. J Proteome Res, 2010. **9**(6): p. 3158-68.
175. Korsching, E., et al., *Basal carcinoma of the breast revisited: an old entity with new interpretations*. J Clin Pathol, 2008. **61**(5): p. 553-60.
176. Korsching, E., et al., *Cytogenetic alterations and cytokeratin expression patterns in breast cancer: integrating a new model of breast differentiation into cytogenetic pathways of breast carcinogenesis*. Lab Invest, 2002. **82**(11): p. 1525-33.
177. Filmus, J., et al., *MDA-468, a human breast cancer cell line with a high number of epidermal growth factor (EGF) receptors, has an amplified EGF receptor gene and is growth inhibited by EGF*. Biochem Biophys Res Commun, 1985. **128**(2): p. 898-905.
178. Balz, L.M., et al., *The interplay of HER2/HER3/PI3K and EGFR/HER2/PLC-gamma1 signalling in breast cancer cell migration and dissemination*. J Pathol, 2012. **227**(2): p. 234-44.
179. Paplomata, E. and R. O'Regan, *The PI3K/AKT/mTOR pathway in breast cancer: targets, trials and biomarkers*. Ther Adv Med Oncol, 2014. **6**(4): p. 154-66.
180. Yarden, Y. and M.X. Sliwkowski, *Untangling the ErbB signalling network*. Nat Rev Mol Cell Biol, 2001. **2**(2): p. 127-37.
181. Zhao, L. and P.K. Vogt, *Class I PI3K in oncogenic cellular transformation*. Oncogene, 2008. **27**(41): p. 5486-96.
182. Santen, R.J., et al., *The role of mitogen-activated protein (MAP) kinase in breast cancer*. J Steroid Biochem Mol Biol, 2002. **80**(2): p. 239-56.
183. Pearson, G., et al., *Mitogen-activated protein (MAP) kinase pathways: regulation and physiological functions*. Endocr Rev, 2001. **22**(2): p. 153-83.
184. Banerjee, K. and H. Resat, *Constitutive activation of STAT3 in breast cancer cells: A review*. Int J Cancer, 2016. **138**(11): p. 2570-8.
185. Segatto, I., G. Baldassarre, and B. Belletti, *STAT3 in Breast Cancer Onset and Progression: A Matter of Time and Context*. Int J Mol Sci, 2018. **19**(9).
186. Kuschel, A., P. Simon, and S. Tug, *Functional regulation of HIF-1alpha under normoxia--is there more than post-translational regulation?* J Cell Physiol, 2012. **227**(2): p. 514-24.
187. Lee, J.W., et al., *Hypoxia-inducible factor (HIF-1)alpha: its protein stability and biological functions*. Exp Mol Med, 2004. **36**(1): p. 1-12.
188. Lopez-Lazaro, M., *Role of oxygen in cancer: looking beyond hypoxia*. Anticancer Agents Med Chem, 2009. **9**(5): p. 517-25.
189. Ranasinghe, W.K., et al., *HIF1alpha expression under normoxia in prostate cancer--which pathways to target?* J Urol, 2015. **193**(3): p. 763-70.
190. Xu, C., et al., *CXCR4 in breast cancer: oncogenic role and therapeutic targeting*. Drug Des Devel Ther, 2015. **9**: p. 4953-64.
191. Dewan, M.Z., et al., *Stromal cell-derived factor-1 and CXCR4 receptor interaction in tumor growth and metastasis of breast cancer*. Biomed Pharmacother, 2006. **60**(6): p. 273-6.
192. Muller, A., et al., *Involvement of chemokine receptors in breast cancer metastasis*. Nature, 2001. **410**(6824): p. 50-6.
193. Putz, E., et al., *Phenotypic characteristics of cell lines derived from disseminated cancer cells in bone marrow of patients with solid epithelial tumors: establishment of working models for human micrometastases*. Cancer Res, 1999. **59**(1): p. 241-8.
194. Pantel, K., et al., *Establishment of micrometastatic carcinoma cell lines: a novel source of tumor cell vaccines*. J Natl Cancer Inst, 1995. **87**(15): p. 1162-8.
195. Kischel, P., et al., *Cell membrane proteomic analysis identifies proteins differentially expressed in osteotropic human breast cancer cells*. Neoplasia, 2008. **10**(9): p. 1014-20.
196. Bellahcene, A., et al., *Transcriptome analysis reveals an osteoblast-like phenotype for human osteotropic breast cancer cells*. Breast Cancer Res Treat, 2007. **101**(2): p. 135-48.

197. Peyruchaud, O., et al., *Early detection of bone metastases in a murine model using fluorescent human breast cancer cells: application to the use of the bisphosphonate zoledronic acid in the treatment of osteolytic lesions.* J Bone Miner Res, 2001. **16**(11): p. 2027-34.
198. Yang, L., et al., *Dysregulated expression of cell surface glycoprotein CDCP1 in prostate cancer.* Oncotarget, 2015. **6**(41): p. 43743-58.
199. Miura, S., et al., *CUB-domain containing protein 1 represses the epithelial phenotype of pancreatic cancer cells.* Exp Cell Res, 2014. **321**(2): p. 209-18.
200. Wright, H.J., et al., *CDCP1 cleavage is necessary for homodimerization-induced migration of triple-negative breast cancer.* Oncogene, 2016. **35**(36): p. 4762-72.
201. Mercurio, A.M., *VEGF/Neuropilin Signaling in Cancer Stem Cells.* Int J Mol Sci, 2019. **20**(3).
202. Zhang, L., et al., *VEGF-A/Neuropilin 1 Pathway Confers Cancer Stemness via Activating Wnt/beta-Catenin Axis in Breast Cancer Cells.* Cell Physiol Biochem, 2017. **44**(3): p. 1251-1262.
203. Chu, W., et al., *Neuropilin-1 promotes epithelial-to-mesenchymal transition by stimulating nuclear factor-kappa B and is associated with poor prognosis in human oral squamous cell carcinoma.* PLoS One, 2014. **9**(7): p. e101931.
204. Elaimy, A.L., et al., *VEGF-neuropilin-2 signaling promotes stem-like traits in breast cancer cells by TAZ-mediated repression of the Rac GAP beta2-chimaerin.* Sci Signal, 2018. **11**(528).
205. Nagai, H., et al., *CLCP1 interacts with semaphorin 4B and regulates motility of lung cancer cells.* Oncogene, 2007. **26**(27): p. 4025-31.
206. Roy, S., et al., *Macrophage-Derived Neuropilin-2 Exhibits Novel Tumor-Promoting Functions.* Cancer Res, 2018. **78**(19): p. 5600-5617.
207. Kallergi, G., et al., *Evaluation of Isolation Methods for Circulating Tumor Cells (CTCs).* Cell Physiol Biochem, 2016. **40**(3-4): p. 411-419.
208. Ring, A., et al., *EpCAM based capture detects and recovers circulating tumor cells from all subtypes of breast cancer except claudin-low.* Oncotarget, 2015. **6**(42): p. 44623-34.
209. Sieuwerts, A.M., et al., *Anti-epithelial cell adhesion molecule antibodies and the detection of circulating normal-like breast tumor cells.* J Natl Cancer Inst, 2009. **101**(1): p. 61-6.
210. Gorges, T.M., et al., *Circulating tumour cells escape from EpCAM-based detection due to epithelial-to-mesenchymal transition.* BMC Cancer, 2012. **12**: p. 178.
211. de Wit, S., et al., *The detection of EpCAM(+) and EpCAM(-) circulating tumor cells.* Sci Rep, 2015. **5**: p. 12270.
212. Bulfoni, M., et al., *In patients with metastatic breast cancer the identification of circulating tumor cells in epithelial-to-mesenchymal transition is associated with a poor prognosis.* Breast Cancer Res, 2016. **18**(1): p. 30.
213. Bulfoni, M., et al., *Dissecting the Heterogeneity of Circulating Tumor Cells in Metastatic Breast Cancer: Going Far Beyond the Needle in the Haystack.* Int J Mol Sci, 2016. **17**(10).
214. Elston, C.W. and I.O. Ellis, *Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up.* Histopathology, 1991. **19**(5): p. 403-10.
215. Bloom, H.J. and W.W. Richardson, *Histological grading and prognosis in breast cancer; a study of 1409 cases of which 359 have been followed for 15 years.* Br J Cancer, 1957. **11**(3): p. 359-77.
216. Casar, B., et al., *Blocking of CDCP1 cleavage in vivo prevents Akt-dependent survival and inhibits metastatic colonization through PARP1-mediated apoptosis of cancer cells.* Oncogene, 2012. **31**(35): p. 3924-38.
217. Zhou, X., et al., *Alterations of biomarker profiles after neoadjuvant chemotherapy in breast cancer: tumor heterogeneity should be taken into consideration.* Oncotarget, 2015. **6**(34): p. 36894-902.
218. *Clinical practice guidelines for the use of tumor markers in breast and colorectal cancer. Adopted on May 17, 1996 by the American Society of Clinical Oncology.* J Clin Oncol, 1996. **14**(10): p. 2843-77.
219. Wang, W., et al., *The diagnostic value of serum tumor markers CEA, CA19-9, CA125, CA15-3, and TPS in metastatic breast cancer.* Clin Chim Acta, 2017. **470**: p. 51-55.

220. Dolscheid-Pommerich, R.C., et al., *Clinical performance of LOCI-based tumor marker assays for tumor markers CA 15-3, CA 125, CEA, CA 19-9 and AFP in gynecological cancers*. *Tumour Biol*, 2017. **39**(10): p. 1010428317730246.
221. Oura, S., et al., *Clinical efficacy of bisphosphonate therapy for bone metastasis from breast cancer*. *Breast Cancer*, 2003. **10**(1): p. 28-32.
222. Yuan, Y., et al., *Cobalt inhibits the interaction between hypoxia-inducible factor-alpha and von Hippel-Lindau protein by direct binding to hypoxia-inducible factor-alpha*. *J Biol Chem*, 2003. **278**(18): p. 15911-6.
223. Schmidtke, C., et al., *Polymer-assisted self-assembly of superparamagnetic iron oxide nanoparticles into well-defined clusters: controlling the collective magnetic properties*. *Langmuir*, 2014. **30**(37): p. 11190-6.
224. Freund, B., et al., *A simple and widely applicable method to ⁵⁹Fe-radiolabel monodisperse superparamagnetic iron oxide nanoparticles for in vivo quantification studies*. *ACS Nano*, 2012. **6**(8): p. 7318-25.
225. Pollari, S., et al., *Enhanced serine production by bone metastatic breast cancer cells stimulates osteoclastogenesis*. *Breast Cancer Res Treat*, 2011. **125**(2): p. 421-30.
226. Kakonen, S.M., et al., *Transforming growth factor-beta stimulates parathyroid hormone-related protein and osteolytic metastases via Smad and mitogen-activated protein kinase signaling pathways*. *J Biol Chem*, 2002. **277**(27): p. 24571-8.
227. Yin, J.J., et al., *TGF-beta signaling blockade inhibits PTHrP secretion by breast cancer cells and bone metastases development*. *J Clin Invest*, 1999. **103**(2): p. 197-206.
228. Hofmann, B.T., et al., *COSMC knockdown mediated aberrant O-glycosylation promotes oncogenic properties in pancreatic cancer*. *Mol Cancer*, 2015. **14**: p. 109.
229. Kalinina, T., et al., *Establishment and characterization of a new human pancreatic adenocarcinoma cell line with high metastatic potential to the lung*. *BMC Cancer*, 2010. **10**: p. 295.
230. Korner, A., et al., *Enzastaurin inhibits invasion and metastasis in lung cancer by diverse molecules*. *Br J Cancer*, 2010. **103**(6): p. 802-11.
231. Okumura, K., et al., *Establishment of high- and low-invasion clones derived for a human tongue squamous-cell carcinoma cell line SAS*. *J Cancer Res Clin Oncol*, 1996. **122**(4): p. 243-8.
232. Brenner, J.C., et al., *Genotyping of 73 UM-SCC head and neck squamous cell carcinoma cell lines*. *Head Neck*, 2010. **32**(4): p. 417-26.
233. Mockelmann, N., et al., *Effect of sorafenib on cisplatin-based chemoradiation in head and neck cancer cells*. *Oncotarget*, 2016. **7**(17): p. 23542-51.
234. Neuhoff, V., et al., *Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250*. *Electrophoresis*, 1988. **9**(6): p. 255-62.
235. Laemmli, U.K., *Cleavage of structural proteins during the assembly of the head of bacteriophage T4*. *Nature*, 1970. **227**(5259): p. 680-5.
236. Shevchenko, A., et al., *Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels*. *Anal Chem*, 1996. **68**(5): p. 850-8.
237. Bartkowiak, K., et al., *Two-dimensional differential gel electrophoresis of a cell line derived from a breast cancer micrometastasis revealed a stem/ progenitor cell protein profile*. *J Proteome Res*, 2009. **8**(4): p. 2004-14.
238. Michalski, A., et al., *Mass spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole Orbitrap mass spectrometer*. *Mol Cell Proteomics*, 2011. **10**(9): p. M111 011015.
239. Biosciences, A. *Reversed Phase Chromatography - Principles and Methods*. 1999 September 2007; Available from: <http://wolfson.huji.ac.il/purification/PDF/ReversePhase/AmershamRPCManual.pdf>.
240. Heinrikson, R.L. and S.C. Meredith, *Amino acid analysis by reverse-phase high-performance liquid chromatography: precolumn derivatization with phenylisothiocyanate*. *Anal Biochem*, 1984. **136**(1): p. 65-74.

241. Wilson, K.J., et al., *The behaviour of peptides on reverse-phase supports during high-pressure liquid chromatography*. Biochem J, 1981. **199**(1): p. 31-41.
242. Mahoney, W.C. and M.A. Hermodson, *Separation of large denatured peptides by reverse phase high performance liquid chromatography. Trifluoroacetic acid as a peptide solvent*. J Biol Chem, 1980. **255**(23): p. 11199-203.
243. Roussel, C., et al., *Mechanistic aspects of on-line electrochemical tagging of free L-cysteine residues during electrospray ionisation for mass spectrometry in protein analysis*. Chemphyschem, 2003. **4**(2): p. 200-6.
244. Ho, C.S., et al., *Electrospray ionisation mass spectrometry: principles and clinical applications*. Clin Biochem Rev, 2003. **24**(1): p. 3-12.
245. Walther, T.C. and M. Mann, *Mass spectrometry-based proteomics in cell biology*. J Cell Biol, 2010. **190**(4): p. 491-500.
246. Yates, J.R., 3rd, et al., *Proteomics of organelles and large cellular structures*. Nat Rev Mol Cell Biol, 2005. **6**(9): p. 702-14.
247. Aebersold, R. and M. Mann, *Mass spectrometry-based proteomics*. Nature, 2003. **422**(6928): p. 198-207.
248. Olsen, J.V., et al., *A dual pressure linear ion trap Orbitrap instrument with very high sequencing speed*. Mol Cell Proteomics, 2009. **8**(12): p. 2759-69.
249. Eliuk, S. and A. Makarov, *Evolution of Orbitrap Mass Spectrometry Instrumentation*. Annu Rev Anal Chem (Palo Alto Calif), 2015. **8**: p. 61-80.
250. Sleno, L. and D.A. Volmer, *Ion activation methods for tandem mass spectrometry*. J Mass Spectrom, 2004. **39**(10): p. 1091-112.
251. Hardman, M. and A.A. Makarov, *Interfacing the orbitrap mass analyzer to an electrospray ion source*. Anal Chem, 2003. **75**(7): p. 1699-705.
252. Makarov, A., *Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis*. Anal Chem, 2000. **72**(6): p. 1156-62.
253. Myllynen, L., et al., *Quantitative proteomics unveiled: Regulation of DNA double strand break repair by EGFR involves PARP1*. Radiother Oncol, 2015. **116**(3): p. 423-30.
254. Cox, J. and M. Mann, *MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification*. Nat Biotechnol, 2008. **26**(12): p. 1367-72.
255. Cox, J., et al., *A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics*. Nat Protoc, 2009. **4**(5): p. 698-705.
256. Riethdorf, S., et al., *Clinical applications of the CellSearch platform in cancer patients*. Adv Drug Deliv Rev, 2018. **125**: p. 102-121.
257. Riethdorf, S., et al., *Detection of circulating tumor cells in peripheral blood of patients with metastatic breast cancer: a validation study of the CellSearch system*. Clin Cancer Res, 2007. **13**(3): p. 920-8.
258. Cristofanilli, M., et al., *Circulating tumor cells, disease progression, and survival in metastatic breast cancer*. N Engl J Med, 2004. **351**(8): p. 781-91.
259. Allard, W.J., et al., *Tumor cells circulate in the peripheral blood of all major carcinomas but not in healthy subjects or patients with nonmalignant diseases*. Clin Cancer Res, 2004. **10**(20): p. 6897-904.
260. de Bono, J.S., et al., *Circulating tumor cells predict survival benefit from treatment in metastatic castration-resistant prostate cancer*. Clin Cancer Res, 2008. **14**(19): p. 6302-9.
261. Picot, J., et al., *Flow cytometry: retrospective, fundamentals and recent instrumentation*. Cytotechnology, 2012. **64**(2): p. 109-30.
262. Givan, A.L., *Flow cytometry: an introduction*. Methods Mol Biol, 2011. **699**: p. 1-29.
263. Seidl, J., R. Knuechel, and L.A. Kunz-Schughart, *Evaluation of membrane physiology following fluorescence activated or magnetic cell separation*. Cytometry, 1999. **36**(2): p. 102-11.
264. Parks, D.R., et al., *Antigen-specific identification and cloning of hybridomas with a fluorescence-activated cell sorter*. Proc Natl Acad Sci U S A, 1979. **76**(4): p. 1962-6.
265. Herzenberg, L.A., R.G. Sweet, and L.A. Herzenberg, *Fluorescence-activated cell sorting*. Sci Am, 1976. **234**(3): p. 108-17.

266. He, Y., et al., *Proteolysis-induced N-terminal ectodomain shedding of the integral membrane glycoprotein CUB domain-containing protein 1 (CDCP1) is accompanied by tyrosine phosphorylation of its C-terminal domain and recruitment of Src and PKCdelta*. J Biol Chem, 2010. **285**(34): p. 26162-73.

9. Supplement

9.1. List of used hazardous substances (chemicals and solutions) with hazard statements according to GHS (Globally Harmonised System) classification

Description	H-phrases by GHS	P-phrases by GHS	Hazard pictogramm
Acetone	H225, H319, H336	P210, P240, P305+P351+P338, P403+P203	
Acetonitrile, hypergrad for LC-MS	H225, H302+H312+H332, H319	P210, P240, P302+P352, P305+P351+P338, P403+P233	
Acrylamid	H301, H312, H332, H315, H317, H319, H340, H350, H361f, H372	P201, P280, P302+P352, P304+P340, P305+P351+P338, P308+P310	
Ammonium sulfate	N/A	N/A	N/A
Ammonium bicarbonate	H302	P301+P312+P330	
Ammonium persulfate	H272, H302, H315, H317, H319, H334, H335	P220, P261, P280, P305+P351+P338, P342+P311	

¹³ C ₆ -Arginine-HCl	H319	P305+P351+P338	
Bromphenol blue	N/A	N/A	N/A
Cholera Toxin	H300+H310, H412	P273, P280, P301 P310+P330, P302+P352+P310	
Cobalt (II) chloride	H350i, H360F, H302, H317, H334, H341, H410	P201, P273, P280, P302+P352, P304+P340, P342+P311	
Coomassie Brilliant Blue G250	N/A	N/A	N/A
Dithiothreitol (DTT)	H302, H315, H319, H335	P261, P305+P351+P338	
Dry Strip Cover Fluid PlusOne	H304, H410	P273, P301+P310, P331, P405, P501	
Ethanol	H225, H319	P210, P240, P305+P351+P338, P403+P233	

Ethylenediaminetetraacetic acid (EDTA)	H319, H332, H373	P280, P304+P340, P312, P305+P351+P338, P337+P313	
Formic acid	H226, H302, H314, H331	P210, P280, P303+P361+P353, P304+P340+P310, P305+P351+P338, P403+P233	
Glycerol	N/A	N/A	N/A
Hydrochloric acid, 37 %	H290, H314, H335	P260, P280, P303+P361+P353, P304+P340+P310, P305+P351+P338	
Iodoacetamide	H301, H317, H334	P261, P280, P301+P310, P342+P311	
Gentamycin	H317	P280	
L-glutamine	N/A	N/A	N/A
Glycine	N/A	N/A	N/A

¹³ C ₆ -Lysine-2HCl	N/A	N/A	N/A
Methanol	H225, H331, H311, H301, H370	P210, P233, P280, P302+P352, P304+P340, P308+P310, P403+P235	  
Orthophosphoric acid (85%)	H290, H314	P280, P301+P330+P331, P305+P351+P338, P308+P310	
Paraformaldehyde	H228, H302+H332, H315, H317, H318, H335, H350	P201, P210, P280, P302+P352, P305+P351+P338	   
Penicillin/Streptomycin	H317, H334, H360, H371	P201, P261, P264, P272, P280, P302+P352, P304+P340, P333+P313, P342+P311	
Lowfat powdered milk	N/A	N/A	N/A
L-proline	N/A	N/A	N/A
2-propanol	H225, H319, H336	P210, P233, P240, P305+P351+P338, P403+P235	 

sodium chloride	N/A	N/A	N/A
Sodium dodecyl sulfate	H228, H302+H332, H315, H318, H335, H412	P210, P261, P280, P301+P312+P330, P305+P351+P338+P310, P370+P378	
N,N,N',N'-tetramethylethylenediamine	H225, H332, H302, H314	P210, P280, P305+P351+P338, P310	
2-Amino-2-(hydroxymethyl)propan-1,3-diol (TRIS)	H315, H319, H335	P280, P302+P352, P305+P351+P338, P312	
Triton X-100	H302, H315, H318, H410	P280, P301+P312+P330, P305+P351+P338+P310	
Tween 20	N/A	N/A	N/A
Urea	N/A	N/A	N/A

9.2. MS2-spectra (large scale)

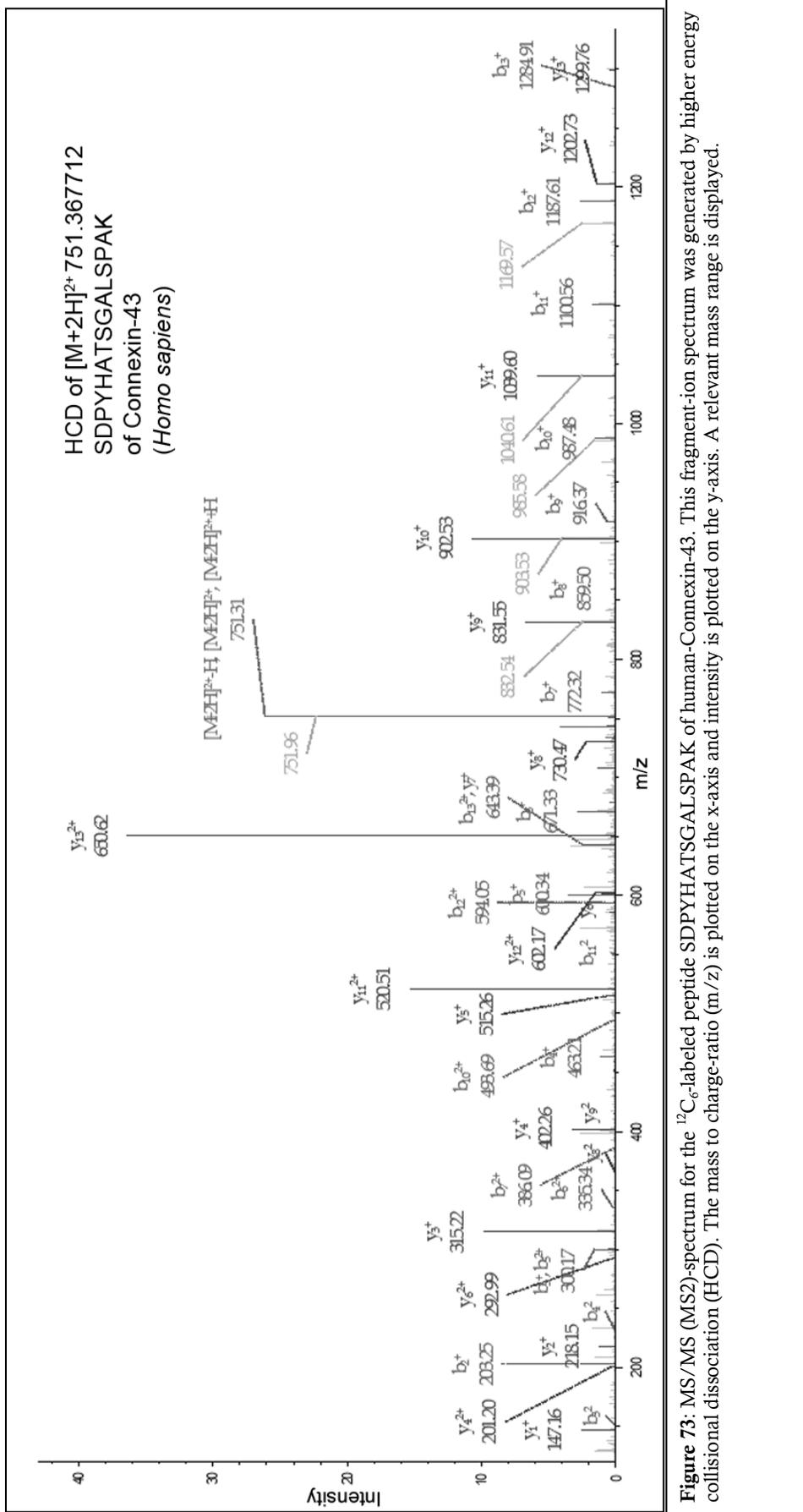


Figure 73: MS/MS (MS2)-spectrum for the ¹²C₆-labeled peptide SDPYHATSGALSPAK of human-Connexin-43. This fragment-ion spectrum was generated by higher energy collisional dissociation (HCD). The mass to charge-ratio (m/z) is plotted on the x-axis and intensity is plotted on the y-axis. A relevant mass range is displayed.

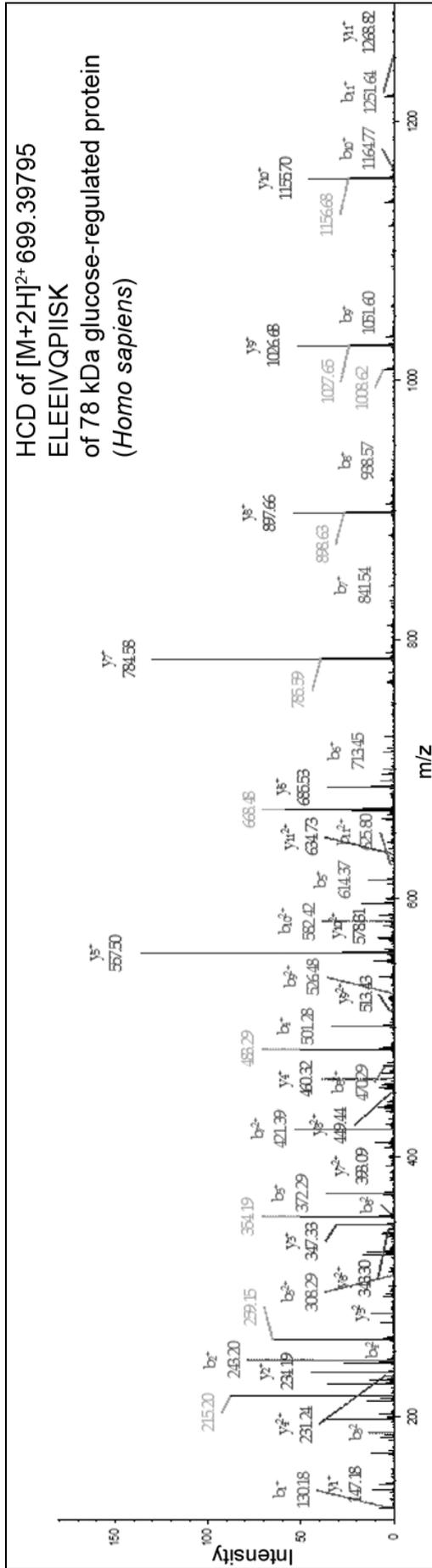


Figure 74: MS/MS (MS2)-spectrum for the $^{12}\text{C}_6$ -labeled peptide ELEEVQPIISK of human-Grp78. This fragment-ion spectrum was generated by higher energy collisional dissociation (HCD). The mass to charge-ratio (m/z) is plotted on the x-axis and intensity is plotted on the y-axis. A relevant mass range is displayed.

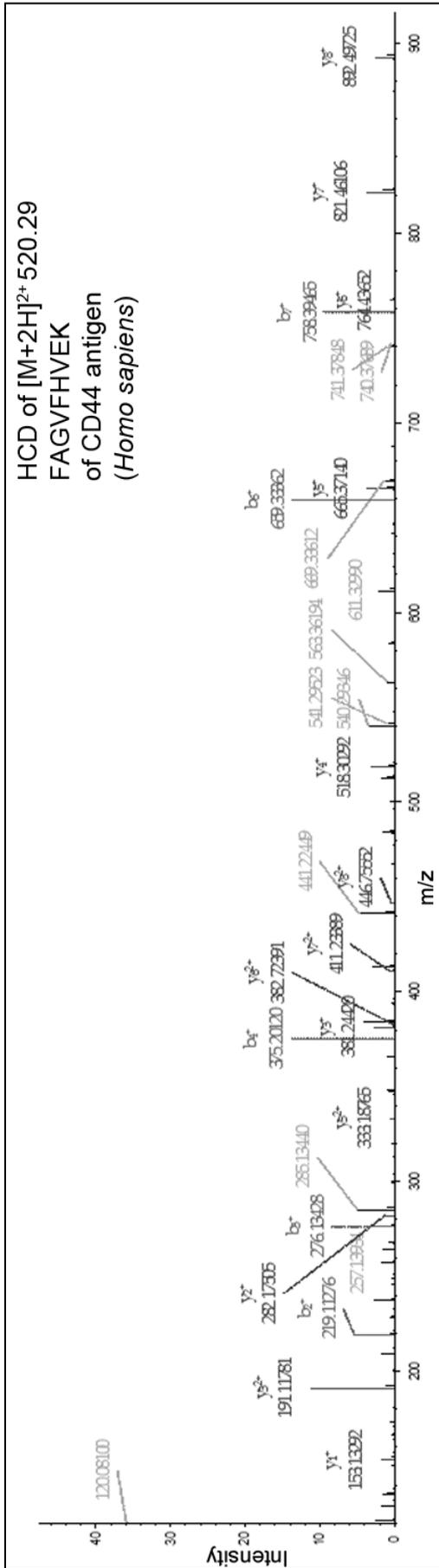


Figure 75: MS/MS (MS2)-spectrum for the $^{13}\text{C}_6$ -labeled peptide FAGVFHVEK of human-CD44. This fragment-ion spectrum was generated by higher energy collisional dissociation (HCD). The mass to charge-ratio (m/z) is plotted on the x-axis and intensity is plotted on the y-axis. A relevant mass range is displayed.

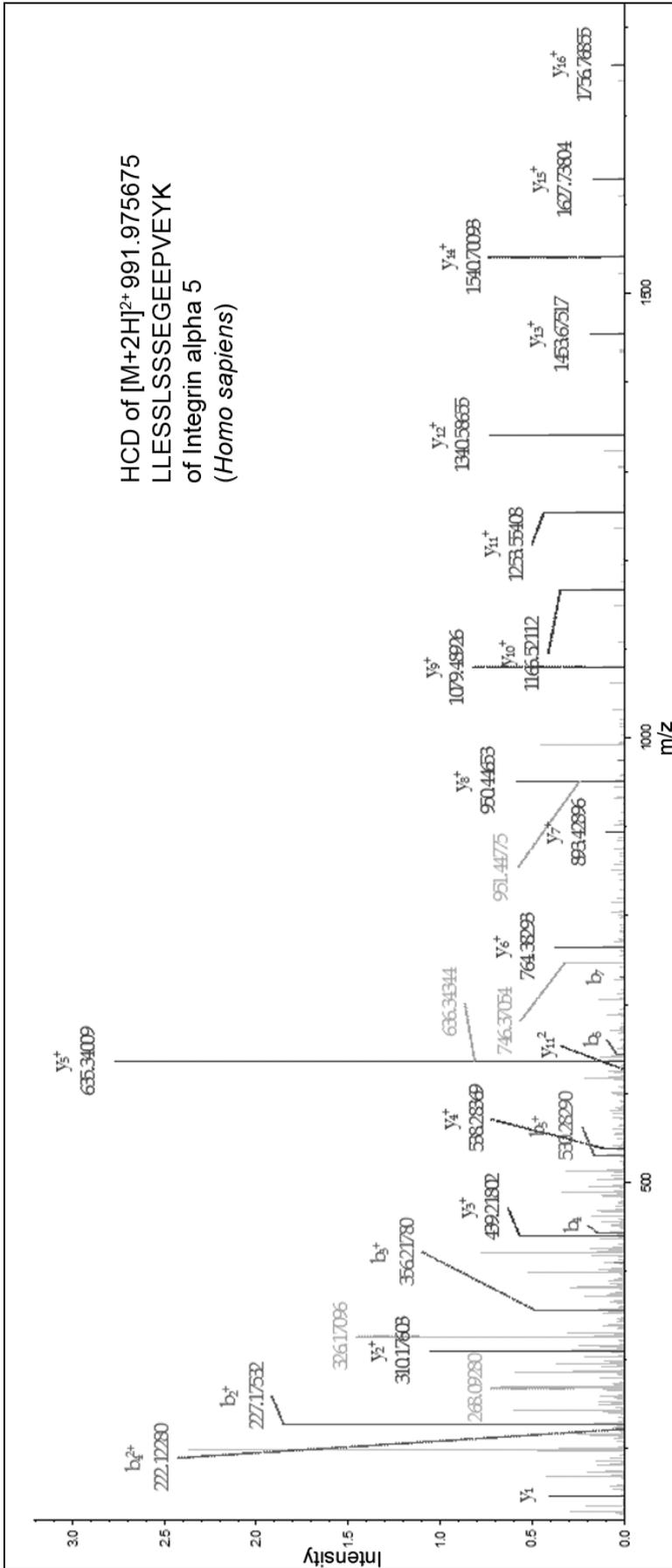


Figure 76: MS/MS (MS²)-spectrum for ¹²C₆-labeled the peptide LLESLSSEGEPEPVEYK of human-Integrin alpha 5. This fragment-ion spectrum was generated by higher energy collisional dissociation (HCD). The mass to charge-ratio (m/z) is plotted on the x-axis and intensity is plotted on the y-axis. A relevant mass range is displayed.

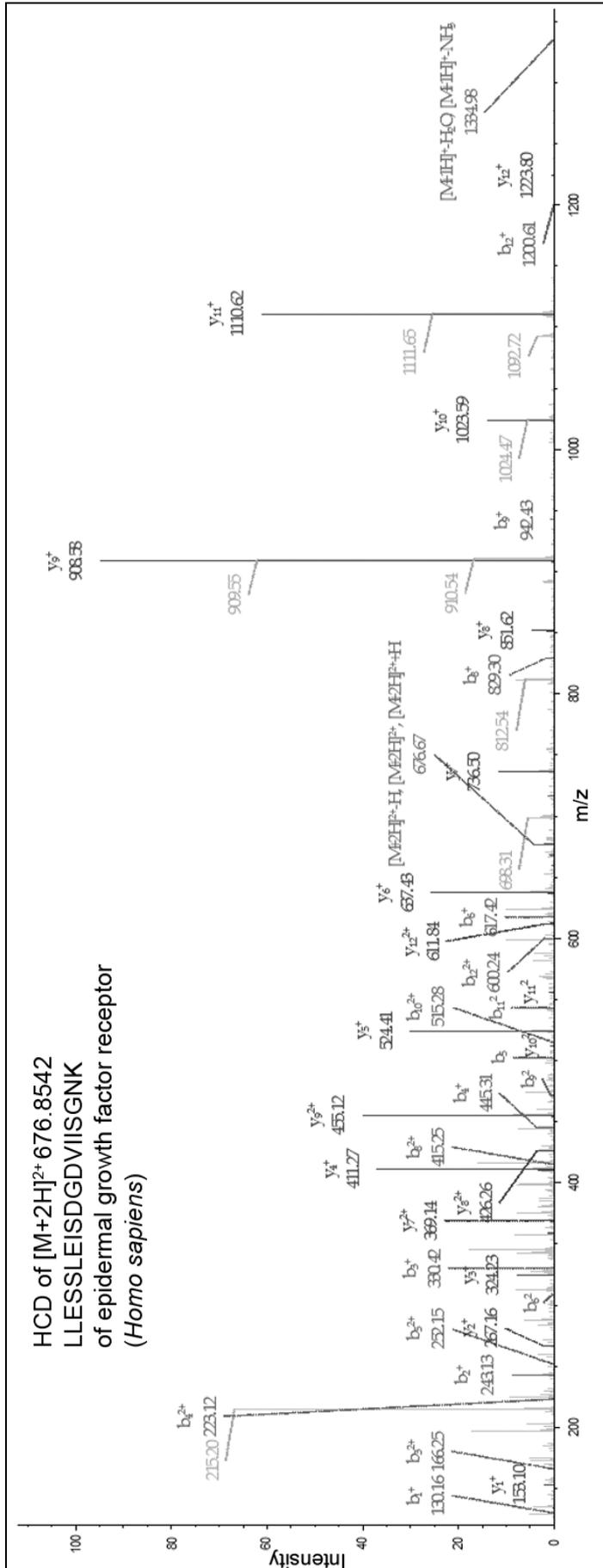


Figure 77: MS/MS (MS2)-spectrum for the ¹³C₆-labeled peptide EISDGDVIISGNK of human-EGFR. This fragment-ion spectrum was generated by higher energy collisional dissociation (HCD). The mass to charge-ratio (m/z) is plotted on the x-axis and intensity is plotted on the y-axis. A relevant mass range is displayed.

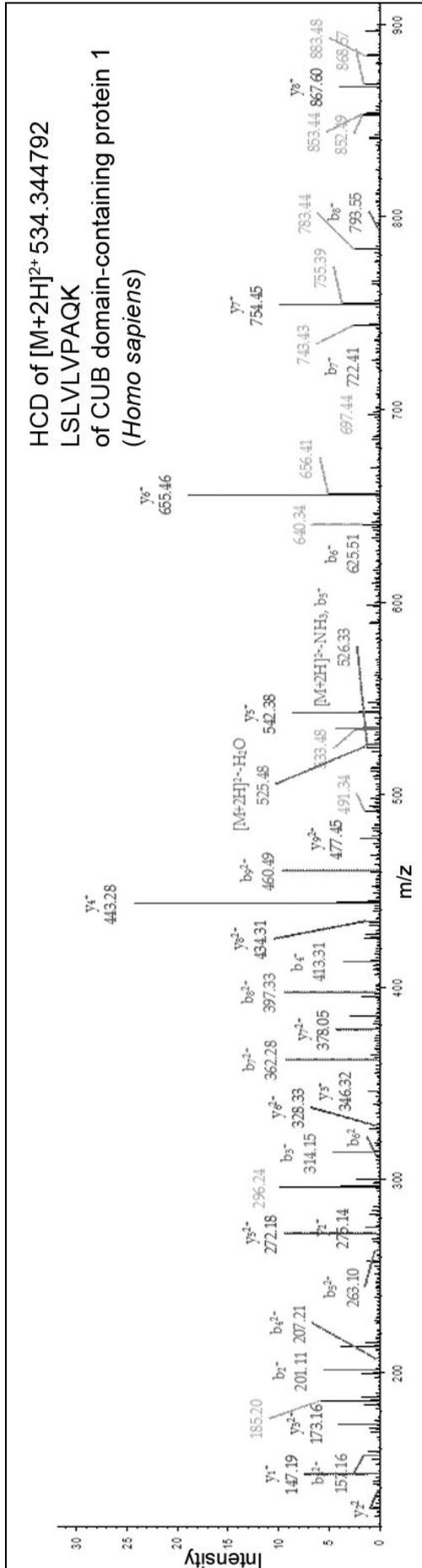


Figure 78: MS/MS (MS2)-spectrum for the $^{12}\text{C}_6$ -labeled peptide LSLVLVPAQK of human-CDCP1. This fragment-ion spectrum was generated by higher energy collisional dissociation (HCD). The mass to charge-ratio (m/z) is plotted on the x-axis and intensity on the y-axis. A relevant mass range is displayed.

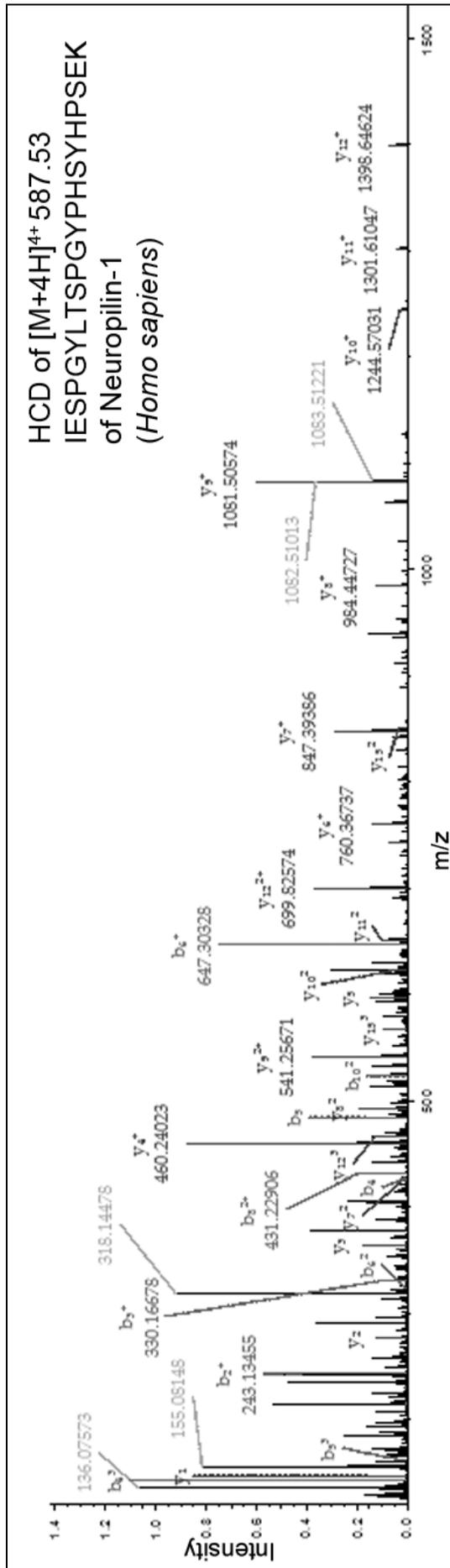


Figure 79: MS/MS (MS2)-spectrum for the $^{12}C_6$ -labeled peptide LSLVLVPAQK of human-CDCP1. This fragment-ion spectrum was generated by higher energy collisional dissociation (HCD). The mass to charge-ratio (m/z) is plotted on the x-axis and intensity is plotted on the y-axis. A relevant mass range is displayed.

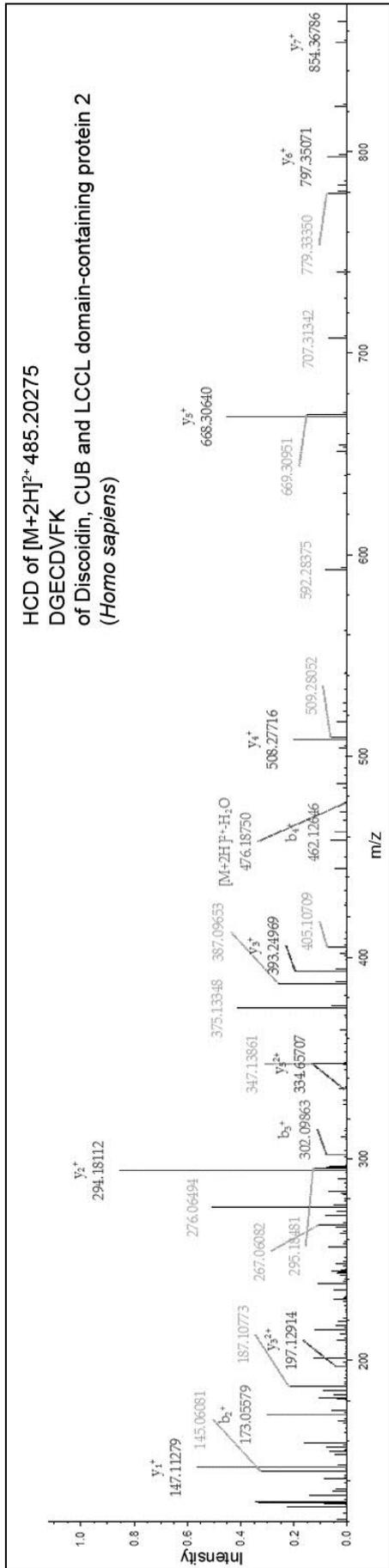


Figure 80: MS/MS (MS2)-spectrum for the $^{12}\text{C}_6$ -labeled peptide DGECDVFK of human-DCBLD2. This fragment-ion spectrum was generated by higher energy collisional dissociation (HCD). The mass to charge-ratio (m/z) is plotted on the x-axis and intensity is plotted on the y-axis. A relevant mass range is displayed.

9.3. Research reports for the analysis of CDCP1 with the CellSearch system.

celltracks
analyzer II[®]

FORSCHUNGSBERICHT

Einrichtung: Universitätskrankenhaus Hamburg Eppendorf
Institut fuer Tumorbiologie
Martinistrasse 52
Gerät 1
20246 Hamburg

Berichtsdatum: 21/05/2019 12:09

Proben-ID: Heid-716 **Patienten-ID:** **Kassetten-ID:** 01619125
Volumen: 7.5 ml **Scan Nr.:** 2

Geräte und Benutzer
CellTracks Analyzer II®
Seriennr.: CT0708035
Testprotokoll: Sebastian 1,6
Benutzer-ID des Scans: MS
Datum/Uhrzeit des Scans: 31/07/2018 14:14
ID des ersten Prüfers: sr
Datum/Uhrzeit der Überprüfung: 01/08/2018 08:59
ID des letzten Prüfers: sr
Datum/Uhrzeit der Überprüfung: 01/08/2018 10:23

CellTracks® AutoPrep® System
Seriennr.: AP0403005
Benutzer-ID: Service
Vorbereitungsdatum: 31/07/2018
Vorbereitungszeit: 12:16
Probenposition: 1
Datum der Entnahme:
Uhrzeit der Entnahme:

Chargendaten
Reagenzienset
Set-ID: CellSearch® CTC
Setcharge: S043
Verfallsdatum: 13/12/2018

Marker-Reagenz
Marker-ID: RUSS
Marker-Charge:
Verfallsdatum:

Ergebnisse
Nur für Forschungszwecke. Nicht zur Verwendung in der Diagnostik.

Ergebnis	Anz. Zellen	% von Zellen
CTC+	691	100.00
CTC+/R 1/100/1,6+	0	0.00
CTC+/R 1/100/1,6-	691	100.00
Nicht zugewiesen:	12715	

Anmerkungen

CellTracks Analyzer II® Kommentare - 5000 images evaluated
x2,7

in total about 1890 CTCs
about 12000 enucleated CK positive fragments

1460 relatively intact CTCs
about 430 apoptotic CTCs

Berichtsautorisierung: _____ Datum: _____

Figure 81: Research report for the breast cancer sample Heid-716. RUSS was used as pseudonym for CDCP1.



FORSCHUNGSBERICHT

Einrichtung: Universitätskrankenhaus Hamburg Eppendorf
 Institut für Tumorbiologie
 Martinistrasse 52
 Gerät 2
 20246 Hamburg

Berichtsdatum: 21/05/2019 13:37

Proben-ID: Heid/796
Volumen: 7.5 ml

Patienten-ID:

Kassetten-ID: 01775860
Scan Nr.: 2

Geräte und Benutzer

CellTracks Analyzer II®

Seriennr.: CT1010013
Testprotokoll: CTC Sebastian 1,6
Benutzer-ID des Scans: CC
Datum/Uhrzeit des Scans: 26/03/2019 14:24
ID des ersten Prüfers: sr
Datum/Uhrzeit der Überprüfung: 28/03/2019 08:11
ID des letzten Prüfers: sr
Datum/Uhrzeit der Überprüfung: 28/03/2019 08:19

CellTracks® AutoPrep® System

Seriennr.: AP1104005
Benutzer-ID: Service
Vorbereitungsdatum: 26/03/2019
Vorbereitungszeit: 11:05
Probenposition: 1
Datum der Entnahme:
Uhrzeit der Entnahme:

Chargendaten

Reagenzienset

Set-ID: CellSearch® CTC
Setcharge: S063
Verfallsdatum: 09/08/2019

Marker-Reagenz

Marker-ID: RUSS1
Marker-Charge:
Verfallsdatum:

Ergebnisse

Nur für Forschungszwecke. Nicht zur Verwendung in der Diagnostik.

Ergebnis	Anz. Zellen	% von Zellen
CTC+	168	100.00
CTC+/RUSS1+	0	0.00
CTC+/RUSS1-	168	100.00
Nicht zugewiesen:	270	

Anmerkungen

CellTracks Analyzer II® Kommentare - 25 relatively intact CTC
 143 apoptotic CTC
 a few with very weak RUSS1 staining

Berichtsautorisierung: _____ Datum: _____

Figure 82: Research report for the breast cancer sample Heid-796. RUSS1 was used as pseudonym for CDCP1.



FORSCHUNGSBERICHT

Einrichtung: Universitätskrankenhaus Hamburg Eppendorf
 Institut fuer Tumorbiologie
 Marlinistrasse 52
 Gerät 1
 20246 Hamburg

Berichtsdatum: 21/05/2019 11:27

Proben-ID: Heid 805
Volumen: 7.5 ml

Patienten-ID:

Kassetten-ID: 01745461
Scan Nr.: 2

Geräte und Benutzer CellTracks Analyzer II®

Seriennr.: CT0708035
Testprotokoll: Sebastian 1,6
Benutzer-ID des Scans: CC
Datum/Uhrzeit des Scans: 11/04/2019 15:43
ID des ersten Prüfers: CC
Datum/Uhrzeit der Überprüfung: 15/04/2019 17:46
ID des letzten Prüfers: CC
Datum/Uhrzeit der Überprüfung: 15/04/2019 17:50

CellTracks® AutoPrep® System

Seriennr.:
Benutzer-ID:
Vorbereitungsdatum:
Vorbereitungszeit:
Probe position: 0
Datum der Entnahme:
Uhrzeit der Entnahme:

Chargendaten Reagenzienset

Set-ID: CellSearch® CTC
Setcharge: S063
Verfallsdatum: 09/08/2019

Marker-Reagenz

Marker-ID: RUSS1
Marker-Charge:
Verfallsdatum:

Ergebnisse

Nur für Forschungszwecke. Nicht zur Verwendung in der Diagnostik.

Ergebnis	Anz. Zellen	% von Zellen
CTC+	26	100.00
CTC+/R1/100/1,6+:	0	0.00
CTC+/R1/100/1,6-:	26	100.00
Nicht zugewiesen:	240	

Anmerkungen

CellTracks Analyzer II® Kommentare - 15 relatively intact CTC
 11 apoptotic CTC
 Russ neg 25
 Russ very weakly 1

Berichtsautorisierung: _____ Datum: _____

Figure 83: Research report for the breast cancer sample Heid-805. RUSS1 was used as pseudonym for CDCP1.



FORSCHUNGSBERICHT

Einrichtung: Universitätskrankenhaus Hamburg Eppendorf
 Institut fuer Tumorbiologie
 Martinistrasse 52
 Gerät 1
 20246 Hamburg

Berichtsdatum: 21/05/2019 11:58

Proben-ID: Heid 786 1.BE
Volumen: 7.5 ml

Patienten-ID:

Kassetten-ID: 01749447
Scan Nr. 1

Geräte und Benutzer CellTracks Analyzer II®

Seriennr.: CT0708035
Testprotokoll: CTC Sebastian 0,8
Benutzer-ID des Scans: AA
Datum/Uhrzeit des Scans: 06/03/2019 13:59
ID des ersten Prüfers: AA
Datum/Uhrzeit der Überprüfung: 06/03/2019 18:35
ID des letzten Prüfers: AA
Datum/Uhrzeit der Überprüfung: 06/03/2019 19:01

CellTracks® AutoPrep® System

Seriennr.: AP0403005
Benutzer-ID: Service
Vorbereitungsdatum: 06/03/2019
Vorbereitungszeit: 11:45
Probenposition: 2
Datum der Entnahme:
Uhrzeit der Entnahme:

Chargendaten Reagenzienset

Set-ID: CellSearch® CTC
Setcharge: S063
Verfallsdatum: 09/08/2019

Marker-Reagenz

Marker-ID: RUSS1
Marker-Charge:
Verfallsdatum:

Ergebnisse

Nur für Forschungszwecke. Nicht zur Verwendung in der Diagnostik.

Ergebnis	Anz. Zellen	% von Zellen
CTC+:	587	100.00
CTC+/Russ1+:	0	0.00
CTC+/Russ1-:	587	100.00
Nicht zugewiesen:	280	

Anmerkungen

CellTracks Analyzer II® Kommentare - 450 relatively intact CTC
 about 140 apoptotic CTC
 ca 10% very weakly positive

Berichtsautorisierung: _____ Datum: _____

Figure 84: Research report for the breast cancer sample Heid-786. RUSS1 was used as pseudonym for CDCP1.



FORSCHUNGSBERICHT

Einrichtung: Universitätskrankenhaus Hamburg Eppendorf
 Institut fuer Tumorbiologie
 Martinistrasse 52
 Gerät 1
 20246 Hamburg

Berichtsdatum: 21/05/2019 12:05

Proben-ID: Heid 791 1.BE RU **Patienten-ID:**

Kassetten-ID: 01749086
Scan Nr. 2

Volumen: 7.5 ml

Geräte und Benutzer CellTracks Analyzer II®

Seriennr.: CT0708035
Testprotokoll: Sebastian 1,6
Benutzer-ID des Scans: AA
Datum/Uhrzeit des Scans: 21/02/2019 14:45
ID des ersten Prüfers: AA
Datum/Uhrzeit der Überprüfung: 21/02/2019 16:05
ID des letzten Prüfers: AA
Datum/Uhrzeit der Überprüfung: 21/02/2019 16:10

CellTracks® AutoPrep® System

Seriennr.: AP0403005
Benutzer-ID: Service
Vorbereitungsdatum: 21/02/2019
Vorbereitungszeit: 12:50
Probenposition: 3
Datum der Entnahme:
Uhrzeit der Entnahme:

Chargendaten Reagenzienset

Set-ID: CellSearch® CTC
Setcharge: S051
Verfallsdatum: 03/04/2019

Marker-Reagenz

Marker-ID: RUS 1
Marker-Charge: C392PA
Verfallsdatum: 13/07/2019

Ergebnisse

Nur für Forschungszwecke. Nicht zur Verwendung in der Diagnostik.

Ergebnis	Anz. Zellen	% von Zellen
CTC+:	0	0.00
CTC+/R 1/100/1,6+:	0	0.00
CTC+/R 1/100/1,6-:	0	0.00
Nicht zugewiesen:	1680	

Anmerkungen

CellTracks Analyzer II® Kommentare - ca 850 CTC
 davon 13 RUS1 sehr schwach positiv
 Probe ist Heid 781

Berichtsautorisierung: _____ Datum: _____

Figure 85: Research report for the breast cancer sample Heid-791. RUS 1 was used as pseudonym for CDCP1.



FORSCHUNGSBERICHT

Einrichtung: Universitätskrankenhaus Hamburg Eppendorf
 Institut fuer Tumorbiologie
 Marlinistrasse 52
 Gerät 1
 20246 Hamburg

Berichtsdatum: 21/05/2019 12:28

Proben-ID: Heid/nachL11/172 **Patienten-ID:**

Kassetten-ID: 01738791
Scan Nr.: 2

Volumen: 7.5 ml

Geräte und Benutzer

CellTracks Analyzer II®

Seriennr.: CT0708035
Testprotokoll: Sebastian 1,6
Benutzer-ID des Scans: CC
Datum/Uhrzeit des Scans: 08/03/2019 15:02
ID des ersten Prüfers: CC
Datum/Uhrzeit der Überprüfung: 08/03/2019 15:37
ID des letzten Prüfers: sr
Datum/Uhrzeit der Überprüfung: 11/03/2019 12:37

CellTracks® AutoPrep® System

Seriennr.: AP1104005
Benutzer-ID: Service
Vorbereitungsdatum: 08/03/2019
Vorbereitungszeit: 13:06
Probe position: 8
Datum der Entnahme:
Uhrzeit der Entnahme:

Chargendaten

Reagenzienset

Set-ID: CellSearch® CTC
Setcharge: S063
Verfallsdatum: 09/08/2019

Marker-Reagenz

Marker-ID: HER2
Marker-Charge:
Verfallsdatum:

Ergebnisse

Nur für Forschungszwecke. Nicht zur Verwendung in der Diagnostik.

Ergebnis	Anz. Zellen	% von Zellen
CTC+	368	100.00
CTC+/R1/100/1,6+:	15	4.08
CTC+/R1/100/1,6-:	353	95.92
Nicht zugewiesen:	800	

Anmerkungen

CellTracks Analyzer II® Kommentare - about 250 relatively intact CTC
 120 apoptotic CTC
 12 RUSS 1 very weakly

mit RUSS 1

Berichtsautorisierung: _____ Datum: _____

Figure 86: Research report for the breast cancer sample Heid-nachL11/172. RUS 1 was used as pseudonym for CDCP1.



FORSCHUNGSBERICHT

Einrichtung: Universitätskrankenhaus Hamburg Eppendorf
 Institut fuer Tumorbiologie
 Martinistrasse 52
 Gerät 1
 20246 Hamburg

Berichtsdatum: 27/05/2019 11:18

Proben-ID: Heid 790 1.BE
Volumen: 7.5 ml

Patienten-ID:

Kassetten-ID: 01739417
Scan Nr. 2

Geräte und Benutzer CellTracks Analyzer II®

Seriennr.: CT0708035
Testprotokoll: Sebastian 1,6
Benutzer-ID des Scans: AA
Datum/Uhrzeit des Scans: 13/03/2019 13:56
ID des ersten Prüfers: AA
Datum/Uhrzeit der Überprüfung: 13/03/2019 18:30
ID des letzten Prüfers: AA
Datum/Uhrzeit der Überprüfung: 13/03/2019 18:38

CellTracks® AutoPrep® System

Seriennr.: AP0403005
Benutzer-ID: Service
Vorbereitungsdatum: 13/03/2019
Vorbereitungszeit: 12:05
Probenposition: 1
Datum der Entnahme:
Uhrzeit der Entnahme:

Chargendaten Reagenzienset

Set-ID: CellSearch® CTC
Setcharge: S063
Verfallsdatum: 09/08/2019

Marker-Reagenz

Marker-ID: RUS1
Marker-Charge:
Verfallsdatum:

Ergebnisse

Nur für Forschungszwecke. Nicht zur Verwendung in der Diagnostik.

Ergebnis	Anz. Zellen	% von Zellen
CTC+	20	100.00
CTC+/R 1/100/1,6+	2	10.00
CTC+/R 1/100/1,6-	18	90.00
Nicht zugewiesen:	237	

Anmerkungen

CellTracks Analyzer II® Kommentare - 4 relatively intact CTC
 16 apoptotic CTC
 RUSS negative 18
 RUSS 1plus 2

Berichtsautorisierung: _____ Datum: _____

Figure 87: Research report for the breast cancer sample Heid-790. RUS 1/RUSS was used as pseudonym for CDCP1.

10. Danksagung

In diesem Abschnitt möchte ich mich bei allen bedanken, die diese Arbeit, auf welchem Wege auch immer, ermöglicht haben.

Als erstes möchte ich mich bei allen Patienten und Patientinnen bedanken, die mit der Bereitschaft ihr Blut zu spenden einen erheblichen Teil zu dieser Arbeit beigetragen haben. Worte wie Respekt oder Anerkennung reichen nicht aus, um die Einstellung der Patienten und Patientinnen zu beschreiben, die trotz einer sehr schweren Erkrankung, die Bereitschaft gezeigt haben die Forschung zu unterstützen und somit auch an zukünftige Generationen gedacht haben.

Ein besonders großer Dank gilt Herrn Prof. Dr. Klaus Pantel, der mir die Möglichkeit gegeben hat, meine Dissertation mit dem zugleich hochinteressanten Thema in seinem Institut anfertigen zu können. Ich kann mir nicht vorstellen, ob ich in einer anderen Arbeitsgruppe an so einer interessanten Fragestellung hätte arbeiten können, einen so tiefen Einblick in die Forschung bekommen hätte oder so einen großen Beitrag zum Fortschritt auf diesem Arbeitsgebiet hätte leisten können. Zudem hat Herr Prof. Pantel mit seinen Anregungen und Ratschlägen zum Erfolg dieser Arbeit beigetragen.

Der wohl größte Dank gilt meiner Frau, Linda Gärtner, die mich in dieser gesamten Zeit so selbstlos unterstützt, motiviert, aufgebaut, abgelenkt und ertragen hat. Ohne Dich wäre diese Doktorarbeit nicht möglich gewesen. Ich weiß, dass es mit mir in dieser Zeit nicht immer leicht war und bin umso glücklicher, dass Du immer noch bei mir bist. Du hast mich mit deiner herrlichen Art nach Arbeitstagen mit diversen experimentellen Rückschlägen beruhigt und aufgebaut und hast es immer geschafft Problemlösungen zu liefern, auf die ich nicht gekommen wäre. Schön, dass es dich gibt. Ich liebe Dich!

Ein ebenso großer Dank gilt Herrn Dr. Bartkowiak, der mir während der gesamten Zeit der Doktorarbeit mit Rat und Tat zur Seite stand und somit einen großen Anteil am Erfolg dieser Dissertation hat. Der permanente und zugleich intellektuelle Austausch mit Ihm hat die meisten Probleme schnell gelöst oder gar nicht erst aufkommen lassen. Zusätzlich bedanke ich mich für diverse spendierte Tassen Kaffee während der Feueralarme.

Mein nächster Dank geht an Antje Andreas, die mit Ihrer dauerhaften Unterstützung, Hilfsbereitschaft und Erfahrung ebenfalls einen großen Beitrag zu dieser Arbeit geleistet hat. Zusätzlich bedanke ich mich für die Versorgung mit Nervennahrung.

Zusätzlich möchte ich mich persönlich bei PD Dr. Sabine Riethdorf bedanken, die mich mit Ihrer großen Hilfsbereitschaft und Erfahrung unterstützt hat. Zusätzlich wurde mir von Ihr die Möglichkeit gegeben, mit speziellen Proben versorgt zu werden.

Ein riesiger Dank gilt meinen Eltern, die mich stets unterstützt haben und ohne die das Studium und die Doktorarbeit nicht möglich gewesen wären. Danke, dass Ihr mir diese Möglichkeit gegeben habt.

Bedanken möchte ich mich auch bei Herrn Prof. Dr. Horst Weller, der als Betreuer immer ansprechbar und immer sehr hilfsbereit war. Zusätzlich möchte ich mich für die Übernahme des Zweitgutachtens bedanken. Ebenso möchte ich mich bei Herrn Prof. Dr. Hartmut Schlüter für die Übernahme des Vorsitzes der Prüfungskommission meiner Disputation bedanken. In diesem Sinne möchte ich mich auch bei den weiteren Mitgliedern der Prüfungskommission bedanken, die zu diesem Zeitpunkt noch nicht feststanden.

Zusätzlich möchte ich mich bei den ITB Girls und allen ITB'lerinnen und ITB'lern für die schöne Zeit in der Mittagspause und in den Inkubationspausen bedanken!

Und natürlich danke ich allen meinen Freunden, die mich in dieser Zeit ertragen, aufgemuntert, unterstützt oder abgelenkt haben. Ohne Euch wäre die Arbeit wahrscheinlich nie fertig geworden.

11. Eidesstattliche Erklärung

Hiermit versichere ich an Eides statt, die vorliegende Dissertation selbst verfasst und keine anderen als die angegebenen Hilfsmittel benutzt zu haben. Die eingereichte schriftliche Fassung entspricht der auf dem elektronischen Speichermedium. Ich versichere, dass diese Dissertation nicht in einem früheren Promotionsverfahren eingereicht wurde.

Hamburg, den 28.05.2019

12. Declaration on oath

I hereby declare, on oath that I have written the available dissertation on my own and have not used other than the declared aid and resources. The submitted written version corresponds to the version on the electronic storage medium. I declare that this dissertation was not submitted in a former doctoral procedure.

Hamburg, 28th May, 2019