

**Development and application of high-
performance liquid chromatography- and mass
spectrometry- based two-dimensional methods
for proteome analysis**

Dissertation

zur Erlangung des Grades
des Doktors der Naturwissenschaftlichen Fakultät III
Chemie, Pharmazie und Werkstoffwissenschaften
der Universität des Saarlandes

von

Katja MELCHIOR

Saarbrücken

2008

Tag der Einreichung:	09.09.2008
Tag des Kolloquiums:	21.11.2008
Dekan:	Prof. Dr. Uli Müller
Mitglieder des Prüfungsausschusses:	Prof. Dr. Christian G. Huber PD. Dr. Andreas Tholey Dr. Markus Martin Prof. Dr. Uli Kazmaier

Acknowledgements

I have been supported for the work of this dissertation by many people. My warmest thank first to **Prof. Dr. Christian G. Huber** who gave me the opportunity to graduate in his lab and guided me through the last three years.

PD. Dr. Andreas Tholey I want to thank for the interesting discussions and introduction into MALDI mass spectrometry as well as **Prof. Dr. Elmar Heinzle** for providing the MALDI-TOF/TOF instrument for my research.

I am further grateful to **Klaus Hollemeyer**, **Maria Lasaosa** and **Nathalie Selevsek** for corporation in terms of the MALDI mass spectrometer.

I also want to thank **Prof. Dr. Eckart Meese** and **Sabrina Heisel** for the supply of the human brain tumor samples.

My thanks go further to **Christa Göllen** and **Gabriele Krug** for their support in administrative matters.

Dr. Markus Martin I want to thank for patiently answering my questions on different theoretical chromatographic topics.

Reiner Wintringer always was very helpful in technical support and producing of monolithic columns.

Furthermore I want to thank **Andreas Keller** for bioinformatic support.

I am very grateful to the whole group of Prof. Dr. Christian G. Huber for numerous discussions and assistance. These are **Nathanaël Delmotte**, **Patrick Eschhold**, **Andreas Leinenbach**, **Jens Mohr** and **Christian Schley** who shared an office with me. I do not want to forget **Verena Fraaß**, **David Gode**, **Iris Gostomski**, **Thomas Hoffmann**, **Manuela Hügel**, **Anis Mahsuna**, **Volker Neu**, **Sascha Quinten**, **Silke Ruzek** and **Bilgin Vatansever** who worked with me every day and shared so much time with me in the lab.

Moreover, I want to thank the apprentices **Eva Luxemburger**, **Devid Hero** and **Robin Adolph**.

For technical support I am grateful to **Norbert Ochs**, **Hans-Peter Skohoutil** and **Jens Wiegert**.

Last but not least my deepest thank goes to my **family** and **friends** who laughed with me in good times and supported me in bad times.

Table of contents

Acknowledgements	III
Table of contents	V
Index of abbreviations, acronyms and symbols	XI
Index of figures	XIII
Index of tables	XVI
Preliminaries	XIX
English version	XXI
Deutsche Version	XXII
Chapter I: Principles of chromatography- and mass spectrometry- based human brain proteomics	1
1 Introduction	3
2 Proteome research	6
2.1 Structure of proteins and peptides	6
2.2 Functions of proteins in the human organism	8
2.3 The Human Proteome Organization (HUPO)	10
2.4 Strategies for proteome research	11
2.5 Biomarker research	13
2.5.1 SEREX as strategy for biomarker identification	13
2.5.2 Biomarker qualification process	15
3 Chromatographic separation of proteins and peptides	17
3.1 Chromatographic principles	17
3.1.1 Optimization of resolution	17
3.1.2 Peak capacity	19
3.1.3 Gradient volume concept	20
3.2 Chromatographic modes	22

3.2.1 Reversed-phase and ion-pair reversed-phase HPLC	24
3.2.2 Ion-exchange high-performance liquid chromatography.....	26
3.3 Stationary phase materials for separation of proteins and peptides.....	28
3.3.1 Monolithic columns for peptide and protein separation.....	29
3.3.2 Strong cation exchange columns for peptide separation	31
4 Mass spectrometry	33
4.1 MALDI as ionization technique for biopolymers	33
4.2 Mass analyzers	35
4.2.1 Overview	35
4.2.2 Time of flight mass analyzer	36
4.2.3 Precursor selection and peptide fragmentation in a MALDI-TOF/TOF mass analyzer	40
4.3 Detectors in mass spectrometry.....	41
4.3.1 Overview of secondary electron multipliers	42
4.3.2 Microchannel plate detector	42
5 Two-dimensional HPLC-MS for proteome analyses	44
5.1 Overview of chromatographic mode combinations	44
5.2 Nano-flow HPLC-MALDI-MS as platform for proteome research.....	45
5.2.1 Interfaces for nano-flow HPLC-MALDI mass spectrometry	46
5.2.2 Proteomes analyzed by 2D-HPLC-MALDI tandem mass spectrometry....	47
5.3 Software tools for result generation, validation and interpretation.....	50
5.3.1 Peptide fragmentation	51
5.3.2 Mascot search engine	52
5.3.3 Scaffold Proteome software	53
5.3.4 Gene ontology-based interpretation software.....	54
6 Brain tumors	55
6.1 Overview of brain tumors	55
6.1.1 Primary brain tumors	56
6.1.2 Secondary brain tumors	57
6.2 Diagnosis, therapy and prognosis.....	57
6.3 Glioblastoma multiforme	59

6.4 Human brain proteomics-where are we standing?	60
Chapter II: Aim of the thesis	65
Chapter III: Development and validation of HPLC-MS approaches for 2D proteome analysis.....	69
1 Introduction.....	71
2 Experimental	73
2.1 Chemicals and Materials.....	73
2.2 Tryptic digest of the proteins	73
2.3 SCX-HPLC for peptide separation	75
2.4 IP-RP-HPLC for separation of intact proteins.....	76
2.5 Nano-flow IP-RP-HPLC separation of peptides	76
2.6 MALDI-TOF/TOF data acquisition.....	77
2.7 Data processing	78
3 Results and Discussion	80
3.1 IP-RP-HPLC for intact proteins as first dimension in 2D analysis	80
3.1.1 Influence of column length on separation performance	80
3.1.2 Mobile phase additives	83
3.1.3 Repeatability of IP-RP-HPLC intact protein separations.....	85
3.2 Validation of a nano-flow HPLC MALDI-TOF/TOF system as second dimension in 2D analysis	87
3.2.1 Configuration of the separation and identification system	87
3.2.2 Variation of spotting time	88
3.2.3 Preconcentration of peptides.....	93
3.2.4 Fragmentation of tryptic peptides	97
3.2.5 Repeatability of nano-flow IP-RP-HPLC peptide separations.....	102
3.3 Comparison of two buffer systems for SCX-HPLC of peptides in the first dimension	105

3.4 Mascot search parameters for proteome analysis	110
4 Summary	113
5 Conclusions	114
Chapter IV: Application of the established approaches to proteome analysis of <i>Glioblastoma multiforme</i>	115
1 Introduction	117
2 Experimental section	119
2.1 Chemicals and materials.....	119
2.2 Preparation of human brain tumor tissue protein extracts.....	119
2.3 IP-RP-HPLC prefractionation of intact proteins as first dimension of the semi top-down approach.....	120
2.4 Tryptic digest of the extracted proteins for the semi top-down approach	120
2.5 Tryptic digest of the extracted proteins for the bottom-up approach	122
2.6 SCX-HPLC prefractionation of peptides for the bottom-up approach.....	122
2.7 Nano-flow IP-RP-HPLC MALDI-MS of peptides.....	123
2.8 Data processing	124
2.9 Data interpretation.....	125
3 Results and discussion	126
3.1 The bottom-up approach as standard method in chromatography- and mass spectrometry-based proteomics	126
3.1.1 Two-dimensional fractionation of peptides	126
3.1.2 Protein identification	127
3.2 A semi top-down approach as alternative method for proteome analysis	129
3.2.1 Two-dimensional fractionation of proteins and peptides.....	129
3.2.2 Protein identification	131
3.3 Comparison of the semi top-down and the bottom-up approach.....	135
3.3.1 Overlap of peptide and protein identifications.....	135

3.3.2 Coverage of mass and pI range	136
3.3.3 Characterization of retention behavior	139
3.3.3.1 <i>First dimension separation of peptides or proteins</i>	139
3.3.3.2 <i>Second dimension separation of peptides</i>	141
3.3.4 Sequence coverage.....	143
3.3.5 Dimension orthogonality and peak capacity	145
3.4 Biological relevance of the identified proteins in the human brain tumor tissue.....	148
3.4.1 Validation of the obtained protein identifications	148
3.4.2 Molecular function of the identified proteins	151
3.4.3 Protein retention behavior in the first separation dimension	152
3.4.4 Confirmation of previously identified potential biomarkers.....	154
4 Summary	158
5 Conclusions	160
 Chapter V: Reference list	 161
 Appendix	 XXIII

Index of abbreviations, acronyms and symbols

BPC	base peak chromatogram
BU	bottom-up
cDNA	complementary DNA
CID	collision induced dissociation
1D	one-dimensional
2D	two-dimensional
2D-GE	two-dimensional gel electrophoresis
DIGE	differential gel electrophoresis
<i>E. coli</i>	<i>Escherichia coli</i>
e.g.	<i>exempli gratia</i>
ESI	electrospray ionization
ExPASy	expert protein analysis system
FA	formic acid
FDA	US Food and Drug Administration
glu ¹ -fib	(Glu ¹)-fibrinopeptide B (human)
GO	gene ontology
GRAVY	grand average of hydropathy
HFBA	heptafluorobutyric acid
HPLC	high-performance liquid chromatography
HUPO	human proteome organization
ICAT	cleavable isotope-coded affinity tag
IEF	isoelectric focusing
IPG	immobilized pH gradient
IP-RP	ion pair-reversed phase
iTRAQ	isobaric tag for relative and absolute quantitation
k	retention factor
laser	light amplification by simulated emission of radiation
LHRH	luteinizing hormone releasing hormone
MALDI	matrix assisted laser desorption/ionization
MCP	microchannel plate
MOWSE	molecular weight search
MudPIT	multidimensional protein identification technology

MWCO	molecular weight cut off
m/z	mass-to-charge ratio
N	plate number
Nd/YAG laser	neodymium endowed yttrium aluminum garnet laser
P	peak capacity
Panther	protein analysis through evolutionary relationships
pH	<i>pondus hydrogenii</i>
pI	isoelectric point
PEEK	polyetheretherketone
PSD	post source decay
PS-DVB	poly-(sterene-divenylbenzene)
RP	reversed phase
R_s	peak resolution
SEREX	serological identification of antigens by recombinant expression screening
STD	semi top-down
TAA	tumor-associated antigen
TFA	trifluoroacetic acid
TFE	trifluoroethanol
t_G	gradient time
t_R	total retention time
US	United States (of America)
vs	<i>versus</i>
w_b	peak width at the basis
w_h	peak width at half hight
w/o	without

Index of figures

Figure 1. Strategies in proteome analysis.....	12
Figure 2. SEREX technique ^[52]	14
Figure 3 ^[59] . Biomarker qualification process for drug-development defined by the US Food and Drug Administration (FDA).	15
Figure 4. Desorption behavior of a small molecule (nitrobenzene, dashed lines) and a protein (lysozyme, solid lines) in RP-HPLC ^[83]	24
Figure 5. IP-RP- HPLC ^[89,90]	25
Figure 6. Ion-exchange HPLC of proteins and peptides ^[89]	27
Figure 7. Morphology of four different stationary phase configurations ^[90]	28
Figure 8. PS-DVB-based monolithic columns and performance ^[106;107]	30
Figure 9. Structure of poly(2-sulfoethyl aspartamide) (=PolySulfoethyl A) as stationary phase material in SCX-HPLC.....	32
Figure 10. Scheme of ion generation during MALDI process.....	34
Figure 11. Principle of the reflector in TOF-MS.....	38
Figure 12. Time-lag focusing technique.....	39
Figure 13. Schematic sketch of precursor passing controlled by the timed ion selector (TIS).....	41
Figure 14. Microchannel plate detector ^[140]	43
Figure 15. Nomenclature for peptide fragmentation sites introduced by Roepstorff and Fohlman ^[169]	52
Figure 16. Side view of the human brain ^[180]	55
Figure 17. Image of <i>Glioblastoma multiforme</i> in a patient's brain.	60
Figure 18. Separation of ten intact proteins on monolithic column(s).	81
Figure 19. Separation of the ten-protein standard with different mobile phase additives.	84
Figure 20. Separation repeatability of the ten-protein mixture.	86
Figure 21. Nano-flow IP-RP-HPLC-MALDI-TOF/TOF setup.	87
Figure 22. Separation of the ten-protein digest peptides.	88
Figure 23. Mascot MOWSE Scores of the identified ten proteins, spotted with 5 s and 10 s per spot, respectively (average of three replicates).	90
Figure 24. Characterization of the identified ten proteins for 5-s and 10-s spotting, respectively (average of three replicates).....	92

Figure 25. UV chromatograms of three-fold separation of the ten-protein digest.	94
Figure 26. Mascot MOWSE Scores of the identified ten proteins, without and with utilizing a trap column, respectively (average of three replicates).	95
Figure 27. Characterization for the ten proteins separated either with or without trap column (average of three replicates).	96
Figure 28. MS/MS spectra for the peptide GAGAFGYFEVTHDITR obtained from catalase, acquired (a) with PSD and (b) with CID.	98
Figure 29. Mascot MOWSE Scores using PSD and CID for peptide fragmentation, respectively (average of three replicates).	100
Figure 30. Characterization of the ten proteins with peptide fragmentation performed either with PSD or CID (average of three replicates).	101
Figure 31. Repeatability of five chromatographic runs of the ten-protein standard mixture digest.	102
Figure 32. Repeatability of peptide identification for three replicates.	104
Figure 33. Repeatability of peptide identification for the single proteins.	104
Figure 34. SCX separation of the ten-protein digest.	106
Figure 35. Mascot MOWSE Scores for the identified ten proteins with SCX separation in the first dimension and two different buffer systems.	107
Figure 36. Characterization of the ten-protein digest separated with two different buffer systems for SCX in the first dimension.	108
Figure 37. Identified unique peptides for the ten proteins with both buffer systems and 1D analysis in comparison.	109
Figure 38. Settings of Mascot parameters for proteome analysis.	111
Figure 39. Two-dimensional fractionation of peptides in the bottom-up approach.	126
Figure 40. Non-redundant protein and peptide hits identified with the bottom-up approach.	128
Figure 41. First dimension fractionation (280 nm) of intact proteins followed by peptide fractionation.	1300
Figure 42. Proteins and peptides identified with the semi top-down approach.	132
Figure 43. Peptides per protein of the 1,642 with the semi top-down approach in triplicate fraction analysis identified proteins.	134
Figure 44. Venn diagram showing the the overlapping of both approaches.	136

Figure 45. LgM _r /pI plots of the identified proteins with at least two peptides of (a) bottom-up, (b) semi top-down approach and (c) proteins identified with both methods.	1377
Figure 46. Comparison of the semi top-down and the bottom-up approach according to (a) the theoretical calculated pI for the proteins and (b) the computed molecular weight of the intact proteins.	138
Figure 47. Retention behavior of first-dimension analytes.	140
Figure 48. Elution of six peptides of the human protein E41L3 (band 4.1-like protein 3) in fraction 9 of the semi top-down approach.	141
Figure 49. GRAVY Indices for unique peptides in fraction 09 of the semi top-down approach.	143
Figure 50. Sequence coverage for the identified proteins of (a) the bottom-up approach and (b) the semi top-down approach.	144
Figure 51. Distribution of peptide length for the semi top-down and the bottom-up approach.	145
Figure 52. Orthogonality of separation dimension in both two-dimensional separation schemes tested.	146
Figure 53. Protein probability plotted as function of peptide probability for the semi top-down approach.	149
Figure 54. Histogram of discriminant scores of the identified peptides for the pool of both approaches.	150
Figure 55. Venn diagram of the semi top-down and bottom-up identifications validated with the Scaffold proteome software.	150
Figure 56. Molecular function of the 1, 429 proteins identified in the pool of the semi top-down and the bottom-up approach (unknown and miscellaneous function excluded).	151
Figure 57. Exemplary retention behavior of six selected membrane proteins.	153

Index of tables

Table 1. Structure, three- and one-letter code for the 20 proteinogenic amino acids.	7
Table 2. Classification systems of proteins.	9
Table 3. Classification of enzymes ^[42]	9
Table 4. Adjustment of gradient volume to application.	22
Table 5. High-performance liquid chromatography techniques for protein and peptide separation.	23
Table 6. Common base materials and functional groups for SCX-HPLC ^[92;120]	31
Table 7. Most common matrices suitable for MALDI.	35
Table 8. Common mass analyzers ^[128]	36
Table 9. Overview of SEMs as detectors in mass spectrometry.	42
Table 10. Common combinations of chromatographic modes in 2D HPLC and selected applications.	45
Table 11. Proteome analysis employing chromatography and MALDI mass spectrometry.	48
Table 12. On-line available search engines for protein identification.	50
Table 13. Overview of primary brain tumors.	56
Table 14. Symptomatic of brain tumors referring to their local occurrence ^[182]	57
Table 15. Average five-year survival rates of primary brain tumors.	59
Table 16. Overview of the recent brain proteome studies.	61
Table 17. Ten-protein standard mixture for digestion.	74
Table 18. Six-peptide standard for calibration of the MALDI-mass spectrometer.	78
Table 19. Peak width at half height (w_H) of the ten proteins.	82
Table 20. Peak capacities (P) for two column lengths and different gradient times.	82
Table 21. Peak width at half height for the ten-protein mix separated with different mobile phase additives.	85
Table 22. Repeatability of retention time for the ten proteins of the standard mixture.	86
Table 23. Ion type distribution of an example spectrum for PSD and CID, respectively.	99

Table 24. Characterization of repeatability of retention times for the established RP-HPLC system using a trap column, 10-s spotting and PSD as peptide fragmentation method.....	103
Table 25. Distribution of trypsin solution over 28 fractions obtained from SCX-HPLC separation of peptides in the first dimension.	121
Table 26. Results from the bottom-up approach (Two replicates in the second separation dimension).	129
Table 27. Results from the semi top-down approach (Three replicates in the second separation dimension).....	133
Table 28. Theoretical peak capacity for both two-dimensional chromatographic separation systems.	147
Table 29. Detection for five of 13 potential antigens for <i>Glioblastoma multiforme</i> directly identified with 2D-HPLC-MALDI-MS/MS.	155
Table 30. Distribution of peptides observed for SEREX proteins. White numbers belong to bottom-up and black numbers to semi top-down results.....	156

Preliminaries

Abstract

Key-words

German translation

English version

Development and application of high-performance liquid chromatography- and mass spectrometry- based two-dimensional methods for proteome analysis

The availability of robust, sensitive and reliable analytic methods is a highly relevant requirement for proteome analysis. In this work two gel-free two-dimensional high-performance liquid chromatography-matrix assisted laser desorption/ionization mass spectrometry (2D HPLC-MALDI-MS) platforms suitable for proteome analysis were developed and compared.

To provide an alternative approach to the classical bottom-up method combining strong cation-exchange- and ion-pair reversed phase HPLC (SCX x IP-RP-HPLC), in which peptides are separated in both dimensions, a semi top-down method was elaborated. This new approach employed separation of proteins in the first and of peptides in the second dimension, respectively. Both separation dimensions of the new method were operated in the ion-pair reversed phase (IP-RP) chromatographic mode using poly-(styrene-divinylbenzene) (PS-DVB)-based monolithic columns and hydro-organic eluents at a pH value of 2.1. Both methods were validated using the same 10-protein mixture.

The elaborated semi top-down- and the classical bottom-up approach were applied to the proteome analysis of a *Glioblastoma multiforme* tissue protein extract including the comparison of the two approaches. Fewer proteins were identified with the semi top-down approach. However, the sequence coverage of the proteins identified was slightly higher. The major advantage of this new method is the elution of proteins in less fractions of the first separation dimension than for the bottom-up approach. Thus, only the fractions containing the target proteins have to be digested in further experiments. Five of 13 potential biomarkers previously identified applying the SEREX (serological identification of antigens by recombinant expression screening) approach, could be confirmed on molecular level by combining the results of both chromatographic approaches.

Key-words: proteome analysis, *Glioblastoma multiforme*, nano-flow HPLC, MALDI mass spectrometry, multidimensional high-performance liquid chromatography, biomarker, monolithic separation columns, orthogonality, peak capacity, retention behavior

Deutsche Version

Entwicklung und Anwendung von zweidimensionalen Methoden für die Proteomanalyse basierend auf Hochleistungsflüssigchromatographie und Massenspektrometrie

Robuste, sensitive und zuverlässige Analysenmethoden sind eine enorm wichtige Voraussetzung für die Proteomanalyse. In dieser Arbeit wurden zwei zweidimensionale gelfreie Plattformen basierend auf Hochleistungsflüssigchromatographie und matrixunterstützter Laser Desorptions-/Ionisations Massenspektrometrie (2D HPLC-MALDI-MS), die *off-line* gekoppelt wurden, entwickelt und miteinander verglichen.

Eine Alternativmethode zur klassischen *bottom-up* Methode, welche starken Kationenaustausch- mit Ionenpaar-Umkehrphasenchromatographie verbindet (SCX x IP-RP-HPLC) und bei der in beiden Dimensionen Peptide getrennt werden, wurde erarbeitet. Diese alternative *semi top-down* Methode beinhaltet die Trennung von Proteinen in der ersten und von Peptiden in der zweiten Dimension. In beiden Dimensionen der neuen Methode wurde der chromatographische Ionen-Umkehrphasen Modus (IP-RP-HPLC) in Verbindung mit Polystyrol-Divenylbenzol (PS-DVB) basierten Monolithen und hydro-organischen Eluenten bei einem pH Wert von 2.1 verwendet. Beide Methoden wurden mit einem 10-Proteinstandard validiert.

Die erarbeitete *semi top-down* und die klassische *bottom-up* Methode wurden auf die Proteomanalyse eines *Glioblastoma multiforme* Gewebe Proteinextraktes angewandt und verglichen. Mit der *semi top-down* Methode wurden zwar weniger Proteine identifiziert, aber deren Sequenzabdeckung war höher. Der Vorteil dieser Methode ist die Elution der Proteine in weniger Fraktionen der ersten Dimension der Trennung. Nur die Fraktionen, die die Targetproteine enthalten, müssen in weiteren Experimenten verdaut werden. Fünf von 13 Biomarkern, die vorher mit der SEREX (*serological identification of antigens by recombinant expression screening*) - Methode identifiziert wurden, konnten auf der molekularen Ebene unter Verbindung der Ergebnisse beider chromatographischer Methoden, nachgewiesen werden.

Schlagwörter: Proteomanalyse, *Glioblastoma multiforme*, nano-Fluss HPLC, MALDI-Massenspektrometrie, Biomarker, monolithische Trennsäulen, Orthogonalität, Peakkapazität, Retentionsverhalten

Chapter I

Principles of chromatography- and
mass spectrometry-based
human brain proteomics

1 Introduction

The term “proteome” was introduced by Wilkins in 1995 who described it as “the total protein complement of a genome” ^[1]. Lottspeich expanded this definition explaining the proteome as the total protein content of an organism, a cell or organelle at a special time under exactly defined surrounding conditions ^[2]. These information are used to compare different states of tissues (e.g. healthy and diseased) as well as cell lines treated with drugs (treated and untreated), for instance. It is hoped to find any hints for target proteins which can be employed for diagnosis or even therapy of a disease.

Three milestones made large-scale proteome analysis possible: the development of two-dimensional gel electrophoresis (2D-GE), mass spectrometry and multi-dimensional chromatography.

In 1975, Klose and O’Farrel independently published a new method for the separation of protein samples with high complexity ^[3;4]. They utilized isoelectric focusing in combination with gel electrophoresis. Today, 2D-GE (in combination with mass spectrometry) is still the most employed methodology in terms of proteome research ^[5].

By operating 2D-GE, only a protein pattern of one condition can be compared to another one but no efficient identification is possible. Therefore, mass spectrometry has to be employed to fragment peptides generated by digestion of the in-gel separated proteins. The fragmented peptides then are identified using the knowledge of the fragment masses. Hence, it is important to know the underlying fragmentation rules depending on the applied mass spectrometer ^[6;7]. Although the first mass spectrometer was developed in 1899 by Thompson ^[8], it took almost one century to operate mass spectrometry as a routine method for protein analysis. A huge problem was the lack of “soft” ionization techniques able to transfer zwitterionic and polar macromolecules in the gaseous phase ^[9]. The immense technical achievements in the eighties and the nineties of the last century offered new opportunities to cope with this problem. In 1968, the concept of electrospray ionization (ESI) was already known ^[10;11] but it took until the late eighties for Fenn et al. to demonstrate the ionization of biopolymers by intact protein analysis with ESI-mass spectrometry (ESI-MS) ^[12]. Matrix assisted laser desorption/ionization (MALDI) as ionization technique for large

and labile biological components followed only few years later ^[13]. Now the road was clear for large-scale analysis of protein samples.

Due to the immensely complex data sets generated computer-based data interpretation is compulsive for proteome analysis. The storage of protein sequences in databases permits rapid protein identification ^[14;15]. Database search engines use the fragment ion pattern of tandem mass spectra to identify the amino acid sequence of the fragmented peptides. The first developed database search engine (SEQUEST) ^[16-18] is still successful in proteome analysis. Other search engines like Mascot, OMSSA or Protein Prophet followed ^[19-21].

Mass spectrometry is an important factor in large-scale protein analysis according to its sensitivity and selectivity. However, efficient multi-dimensional separation methods prior to mass spectrometry are necessary to reduce the enormous complexity of the sample. The most prominent method in this context is the multi-dimensional protein identification technology (MudPIT) ^[22] employing ESI-MS. Here the capillary electrospray needle is directly filled with two stationary phase materials for cation-exchange chromatography and reversed-phase chromatography. This setup permits on-line two-dimensional separation and on-line hyphenation to an ESI-MS analysis. Another successful method for proteome analysis is protein mapping based on pI value and molecular mass of the proteins ^[23-24]. The first dimension employs chromatofocusing (CF) which is based on ion-exchange chromatography of intact proteins. To perform separation of the proteins a pH-gradient is employed ^[25;26]. Subsequently the obtained fractions from the first dimension are separated utilizing reversed phase chromatography with on-line ESI-MS detection to obtain the exact intact protein masses. Recently, a micro-CF approach was introduced employing CF of intact proteins in the first dimension and separation of peptides in the second dimension with the advantage of low sample consumption ^[27].

The introduction of monolithic capillary columns contributed relevantly to highly efficient chromatographic separation prior to mass spectrometry. These separation columns are suitable for miniaturization which is necessary when small sample volumes and low concentrations of compounds are analyzed ^[28]. Moreover, these monolithic columns offer faster mass transfer with lower backpressure than observed with porous particle packed capillary columns since there is no diffusion of the analytes in and out of the pores. Especially for high resolution separation of peptides and proteins, which possess a low diffusion coefficient, stationary phases offering

fast mass transfer are required ^[29;30]. Today, chromatographic approaches in combination with mass spectrometry are well established and applied for proteome analysis ^[31]. Especially the so-called bottom-up or shotgun approach is used successfully for proteome analysis beside 2D-GE ^[32]. This procedure includes peptide separation in both chromatographic dimensions. Monolithic capillary columns are very convenient for the second dimension according to the mentioned advantages.

This chapter provides information about strategies in proteome research including the underlying principles of chromatography and mass spectrometry. The focus is set on the human proteome, especially brain tumor proteomics with the search of efficient biomarkers improving diagnosis and therapy of diseases.

2 Proteome research

2.1 Structure of proteins and peptides

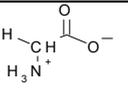
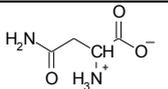
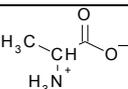
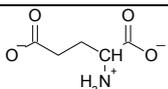
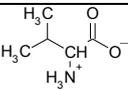
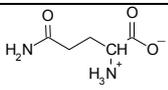
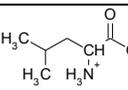
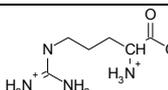
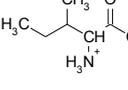
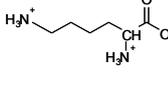
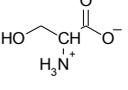
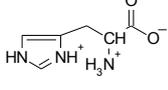
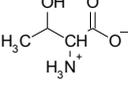
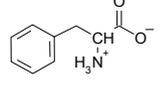
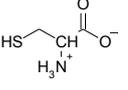
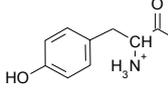
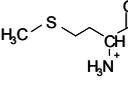
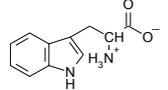
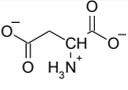
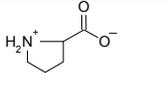
Proteins are macromolecules possessing a complex structure that may be subdivided into four layers: primary, secondary, tertiary and quaternary structure. The primary structure is determined by the order of α -amino acids in the protein sequence. These α -amino acids possess a central carbon atom (α) binding four different substituents: an amino group, a carboxyl group, a hydrogen atom and a so-called side chain. This side chain distinguishes the different amino acids and determines their chemical properties. As four different substituents are linked to the α -carbon atom, amino acids are chiral (except glycine) resulting in D- and L-isomers. In proteins only L-amino acids are incorporated. A very important property of amino acids is their zwitterionic character. At a neutral pH value, the amino acids are mostly existent as dipolar ions with a protonated amino group ($-\text{NH}_3^+$) and a dissociated carboxyl group ($-\text{COO}^-$). Peptides and proteins consist of amino acids that contain both basic and acidic functional groups, therefore they are ampholytes. Their charge depends not only on the amino acid composition, but also on the pH value of the surrounding medium. The pH value where a protein holds no net electrical charge is called isoelectric point (pI).

As an abbreviation system for amino acids a three letter code was introduced. Because of troubles for presenting long protein sequences, a one letter code followed ^[33]. In Table 1 abbreviations, full names as well as structures of the 20 proteinogenic amino acids are listed.

The spatial arrangement of amino acids which are situated near each other in the linear amino acid chain is called secondary structure. The most important periodic structures are α -helices and β -sheets suggested by Pauling and Corey in 1951 ^[34-36]. They are formed by a regularly appearing pattern of hydrogen bonds between $-\text{N}-\text{H}-$ and $-\text{C}=\text{O}-$ connections of the peptide bonds of amino acids laying near each other in the linear amino acid sequence. Hydrogen bonds between carboxyl and amino groups in the inner of the α -helices with a distance of four amino acids stabilize this structure, whereas the side chains are directed to the outside. One turn of an α -helix corresponds to 3.6 amino acids which causes direct vicinity of every fourth amino acid. Those with a distance of two amino acids between them are located on

opposite sides of the α -helix turn and can not interact with each other. All α -helices in natural proteins occur dextrorotatory because this direction of rotation is energetically more favorable due to less steric collisions between side chains and peptide backbone. The β -sheet construction was also discovered by Pauling and Corey as second structure; hence, it was called β -sheet. In contrast to α -helices this structure is craned and not coiled. The distances between adjacent amino acids are higher than for α -helices (0.35 nm and 0.15 nm, respectively) and their side chains are oppositely directed. While all the atoms of the peptide bond are situated in the same plane, the α -carbon atoms linger simultaneously in two planes.

Table 1. Structure, three- and one-letter code for the 20 proteinogenic amino acids.

amino acid	3-letter code (1-letter code)	sequence	amino acid	3-letter code (1-letter code)	sequence
glycine	Gly (G)		asparagine	Asn (N)	
alanine	Ala (A)		glutamic acid	Glu (E)	
valine	Val (V)		glutamine	Gln (Q)	
leucine	Leu (L)		arginine	Arg (R)	
isoleucine	Ile (I)		lysine	Lys (K)	
serine	Ser (S)		histidine	His (H)	
threonine	Thr (T)		phenylalanine	Phe (F)	
cysteine	Cys (C)		tyrosine	Tyr (Y)	
methionine	Met (M)		tryptophan	Trp (W)	
aspartic acid	Asp (D)		proline	Pro (P)	

The tertiary structure is specified by the spatial arrangement of the polypeptide chains. X-ray and nuclear magnetic resonance analyses for the first time permitted the investigation of the spatial arrangement with atomic resolution. The first protein structure which was determined by X-ray crystallography was that of myoglobin ^[37]. Every protein possesses a specific three-dimensional structure which is influenced by the surrounding medium. It depends, for instance, on polarity, viscosity, concentration and ionic strength of the surrounding solvent. Moreover, the presence of prosthetic groups has an impact on protein confirmation. These are molecules mostly covalently linked with the protein, but not consisting of amino acids. Biotin in carboxylases, heme in hemoglobin (not covalently linked) and cytochrome C as well as flavins in flavoproteins are ranked among prosthetic groups. Also post translational modifications, such as phosphorylation, acetylation or glycosylation are important parameters for protein confirmations. Besides hydrogen bonds, ionic, hydrophobic and van-der-Waals forces as well as covalent disulfide bonds between sulfhydryl groups of two cysteine moieties contribute to the stabilization of the three-dimensional protein structure.

Finally, the quaternary structure is formed by building of protein complexes. Proteins possessing more than one polypeptide chain are organized in so-called subunits. The quaternary structure of a protein describes the spatial arrangement of its subunits and the interactions among them which are mostly non-covalent.

For protein quantification the Bradford assay is often employed. Here, the Coomassie Brilliant Blue G-250 dye binds mainly to basic and aromatic amino acid residues of the protein, especially to arginine ^[38]. By binding to these amino acids the color of the dye is shifting from red (465 nm) to blue (595 nm) ^[39;40]. The extent of color change correlates with the protein concentration ^[41]. With this method a high sensitivity is achieved because of the high extinction coefficient of the protein-dye complex.

2.2 Functions of proteins in the human organism

Proteins are involved in almost all processes in the human body. They are classified in several ways regarding structure, source and function (Table 2). The differentiation according to function allows a concrete view to the special task of every protein. Thus, examples are given in Table 2 and are mentioned in particular.

Table 2. Classification systems of proteins.

structure	function	examples	source
fibrous	structure	elastin, fibroin, collagen, keratin	plant
globular	enzymes	extra classification	animal
	transport	hemoglobin, lipoproteins, albumin	human
	agitation	actin, myosin, tubulin	
	reservoir	casein, ovalbumin, ferritin	
	defense	immunoglobulins, cytokines, interferons,	
	regulation	fibrinogen, thrombin human growth hormone, repressors/activators	

Structure proteins determine the assembly of cells and inter-cellular compartments.

As enzymes themselves are a very complex group, several classification systems exist for them. One with respect to their type of reaction is listed in Table 3. Hydrolases are of special interest in protein identification analysis, especially peptidases which catalyze cleavage of the peptide bonds. The most prominent endopeptidase for proteomic applications is trypsin which cleaves C-terminally after lysine and arginine when no proline follows in the amino acid sequence. Other peptidases often employed are pepsin, chymotrypsin, Glu-C-protease and Asp-N-protease.

Table 3. Classification of enzymes ^[42].

class	subclasses
oxidoreductases	dehydrogenases, oxidases, peroxidases, reductases, monooxygenases, dioxygenases
transferases	C ₁ -transferases, glycosyl-transferases, aminotransferases, phosphotransferases
hydrolases	esterases, glycosidases, peptidases (proteases), amidases
lyases (synthases)	C-C-lyases, C-O-lyases, C-N-lyases, C-S-lyases
isomerases	epimerases, cis-trans-isomerases, intra molecular transferases
ligases (synthetases)	C-C-ligases, C-O-ligases, C-N-ligases, C-S-ligases

Transport proteins carry substances through cell membranes, and are then able to release these substances. Hemoglobin, for instance, carries oxygen and lipoproteins lipids. Albumins transfer hydrophobic substances which can not be transported by the

aqueous blood itself, such as fatty acids, bilirubin, vitamins, hormones, and micronutrients. Proteins of motion permit movement of tissues. Actin and myosin appear in filaments of the skeletal muscle cells where they realize contraction under energy consumption. Tubulin is a structure unit of microtubules as well as responsible for motility of the cell.

Proteins storing components are called reservoir proteins. Casein, for instance, contains high amounts of calcium, and ferritin stores iron.

In the human organism two different kinds of defense proteins are distinguished: Proteins of the immune system and proteins of blood clotting. To the latter group fibrinogen and thrombin are allotted and to the first immune globulins, cytokines and interferons. Their task is to maintain the healthy condition of the human body.

The last group is the sub class of regulating proteins. To this class protein hormones like human growth hormone and parathormone as well as repressors and activators controlling the biosyntheses of proteins are classified.

2.3 The Human Proteome Organization (HUPO)

The human genome consists of 20,000-25,000 genes ^[43] which are stored in 23 chromosome pairs. These genes encode approximately 500,000 different proteins, but the exact number still is unknown. The concentration of expressed proteins in the cell may vary in a dynamic range of $1-10^{10}$ due to their different functions ^[44].

A large-scale project dedicated to human proteome analyses was started in 2001 by the Human Proteome Organization (HUPO). This organization aims to identify all human proteins and clarify their interactions, fostering the development of new proteomic approaches. Seven sub-projects are already initiated; inter alia brain, liver, and plasma proteome projects.

Worldwide, 18 laboratories are working on the Brain Proteome Project (BBP) which is focused on brain development and associated diseases. In 2006, the pilot phase of the BBP was finished. Two-dimensional gel electrophoresis as well as approaches based on chromatography and mass spectrometry were applied both using bioinformatic data interpretation ^[45]. Initial information about the normal human brain as well as the mouse brain should be gained to yield a comprehensive comparison with the diseased brain to be investigated in the future. As a result 1,804 proteins could be identified and implicated in biological context ^[46]. Three main points

associated with sample and data handling were elaborated during this starting phase ^[47]:

1. Sample handling, preparation, separation and identification have to be standardized to trace back differences and discrepancy in analytical results.
2. Two or more different approaches and search engines have to be acclaimed as complementary. On the one hand identified proteins appearing in both methods can be confirmed and on the other hand additional proteins can be added to the data set.
3. Only highly confident protein lists generated by rigorous database search parameters should be included in the search for biomarkers.

2.4 Strategies for proteome research

The aim of proteome research is to analyze preferably the entire protein content of cells, tissues or body fluids. Therefore, two main strategies have been developed successfully: The top-down and the bottom-up (= shotgun) approach. Whereas in the top-down approach the separation is performed at protein level, it is realized at peptide level in the bottom-up approach as outlined in Figure 1.

Top-down proteome analysis is mostly realized by applying 2D electrophoresis (2D-GE). This means in a first dimension, proteins are separated according to their isoelectric point, employing isoelectric focusing (IEF). In a second dimension, they are separated by size-dependent sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE). Bottom-up proteomics in contrast utilizes mostly 2D HPLC-MS approaches.

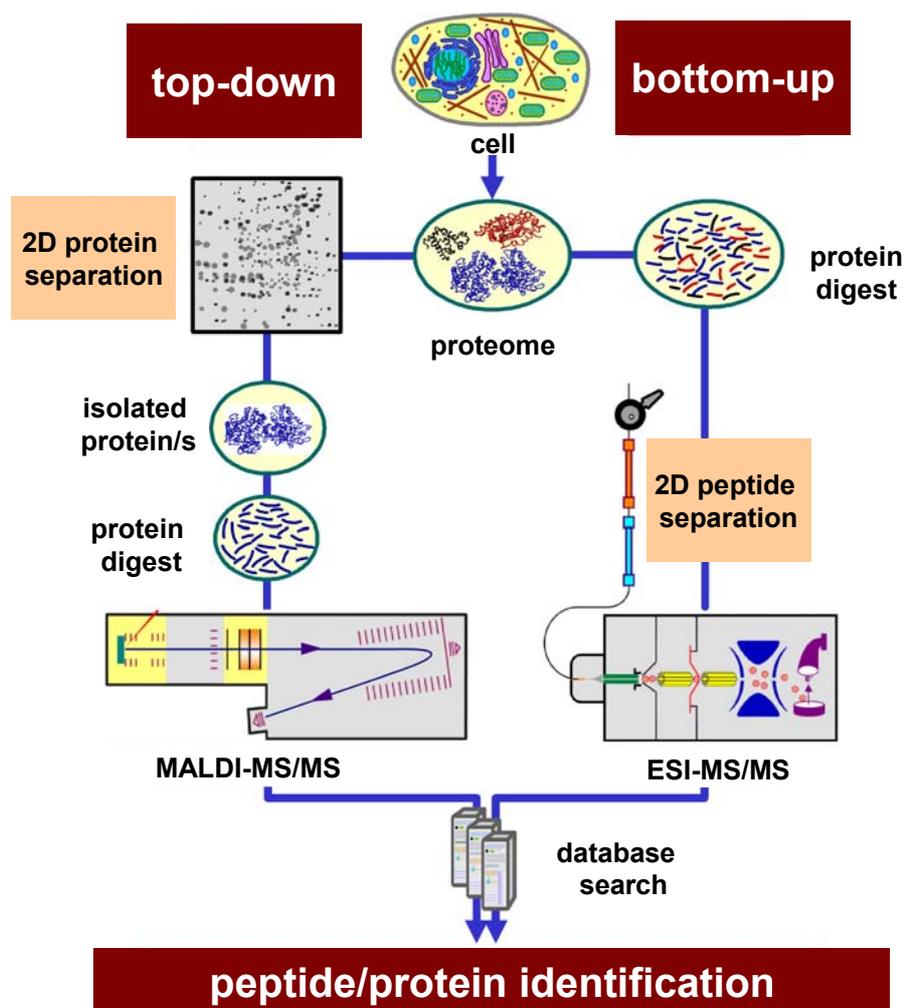


Figure 1. Strategies in proteome analysis.

After digestion and 2D separation, peptides are identified using mass spectrometric analysis with subsequent computer-based data interpretation. Therefore, the entire known proteins of a given organism are digested *in silico* and the calculated peptide masses are compared to the masses obtained by mass spectrometry. This procedure is called peptide mass fingerprint (PMF). The peptide mass information however is not sufficient to identify a peptide reliably. Peptides containing the same amino acids but in another order will result in the same peptide mass. For more specific identification peptide ions are fragmented in a tandem MS experiment in the mass spectrometer. The differences between the fragment masses are assigned to amino acid residues. With this so-called peptide fragment fingerprint (PFF) the amino acid sequence of the identified peptide as additional information is obtained. In recent proteome studies it is compulsive to perform tandem MS experiments to minimize false positive identified proteins.

2.5 Biomarker research

Various definitions exist for the term biomarker. The National Institutes of Health (NIH) working group elaborated a comprehensive definition in 2001. According to them a biomarker is an objectively measured feature to indicate a normal physiological or pathogenic process or pharmacological feedback to a therapeutic intervention ^[48]. The easiest available biomarkers in practice are physiological measures like blood pressure for monitoring of cardiovascular diseases or blood glucose level for diabetes diagnosis and monitoring. However, for drug development more complex features have to be accessible for analysis like patterns of genotypes, disturbance of gene expression profiles or changes of protein and metabolite levels. In summary four entities today are described as biomarkers in drug development ^[49]:

1. DNA - reflecting the genetic variability
2. mRNA - displaying changes of gene expression
3. proteins - representing cellular and enzymatic changes
4. metabolites - highlighting physiological endpoints

The aims after identification of biomarkers are usually validation, characterization, qualification and routinely measuring of these markers before they can be applied to diagnosis or drug development for therapy.

2.5.1 SEREX as strategy for biomarker identification

Serological identification of antigens by recombinant expression screening was developed by Pfreundschuh et al. 1995 ^[50]. The advantage of this method is overcoming the obstacle of tumor cell cultivation. Except of some special tumor types it is very difficult to establish cultured cell lines. By combination of autologous typing of cancer cells and a molecular cloning procedure cell cultures are no longer necessary. In Figure 2 the workflow of SEREX is depicted. First a cDNA library has to be constructed by purification of the tissue sample mRNA, amplification and reversed transcription to cDNA. The entire cDNA library is transferred into λ -phages, and the encoded proteins are expressed in *E. coli*. After a lytic infection of the bacteria the

proteins are blotted onto nitro cellulose membranes. Those membranes were incubated with autologous patient serum. Clones which are reactive with patient's IgG antibodies are identified by a staining reaction using an enzyme- conjugated secondary antibody specific for human IgG as illustrated in Figure 2.

Positive clones are then subcloned to isolate single plaques containing a single insert. The nucleotide sequence of the inserted DNA is accordingly determined ^[51]. For validation of the identified antigens spot assays are applied. Here, only the phage clones with the proteins of interest are expressed in *E. coli* instead of the whole cDNA library.

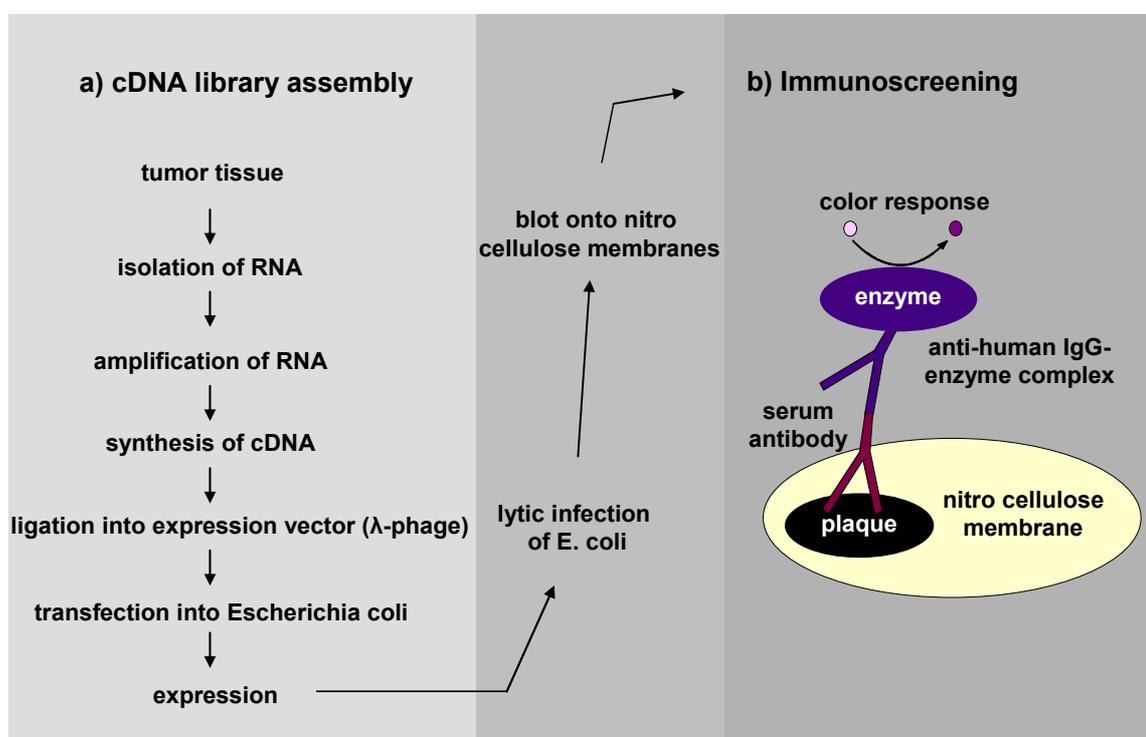


Figure 2. SEREX technique ^[52].

More than 2,000 tumor-associated antigens (TAA) of 15 different tumor types were identified by SEREX ^[53]. Subsequently to the first SEREX phase (identification) a validation step follows by screening sera from healthy individuals and cancer patients by spot assays to show cancer-restricted immune response. The identification of 13, 32 and 40 antigens reacting exclusively with cancer patient's sera of renal ^[54], colon ^[55] and breast cancer ^[56], respectively, was achieved employing SEREX, for instance. Recently, the detection of 13 tumor-associated antigens for *Glioblastoma multiforme* was published ^[57].

2.5.2 Biomarker qualification process

The US Department of Health and Human Services, Food and Drug Administration (FDA) developed a “Guidance for Industry Pharmacogenomic Data Submissions” where they described the regulatory aspects of biomarkers [58]. These are summarized in Figure 3.

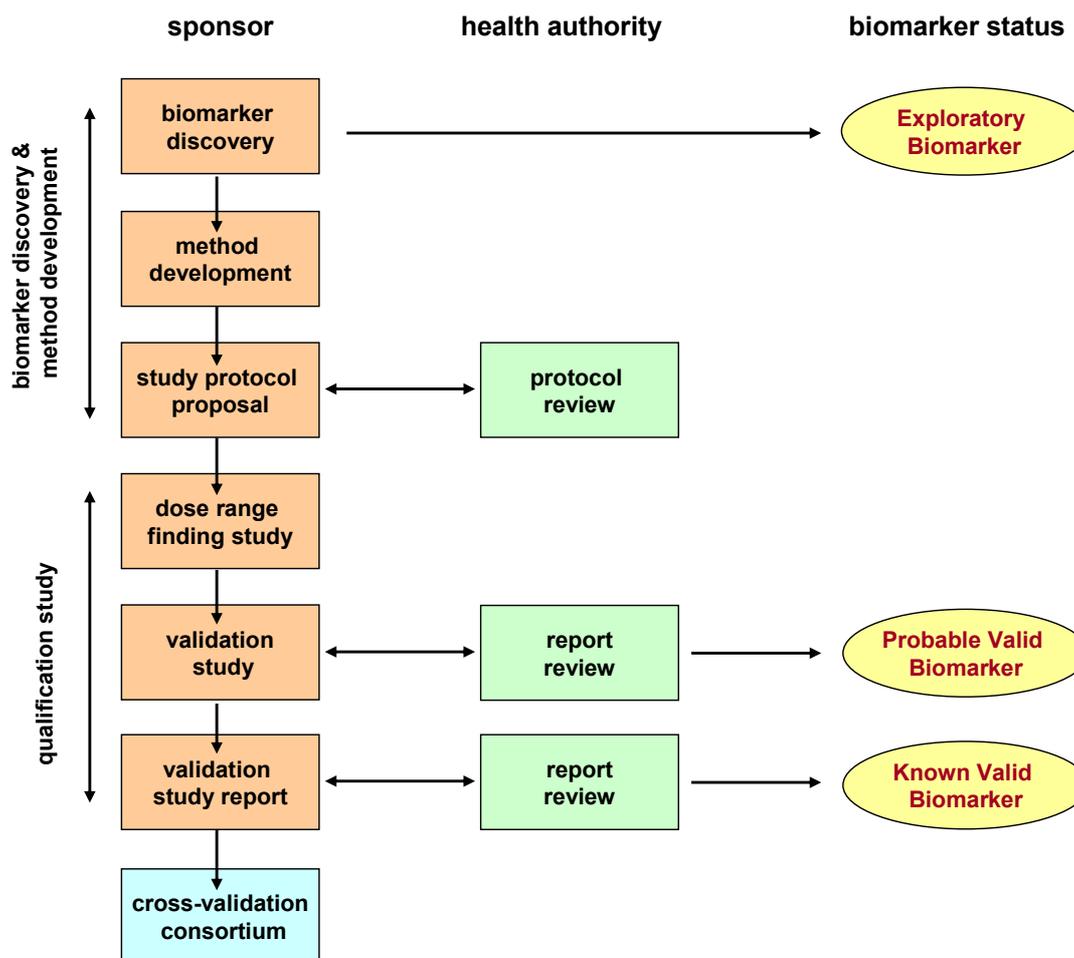


Figure 3 [59]. Biomarker qualification process for drug-development defined by the US Food and Drug Administration (FDA).

Three types of biomarkers are distinguished: “Exploratory Biomarkers”, “Probable Valid Biomarkers” and “Known Valid Biomarkers”. A certain proof of concept exists for “Exploratory Biomarkers” from in-house experiments or literature. Two requirements have to be fulfilled for such a biomarker to get the status of a “Probable Valid Biomarker”:

1. It has to be analyzed with a system of well-established performance characteristics.
2. There has to be an established scientific framework that explains the physiological, pharmacological, toxicological or clinical significance of the test results.

When the success of the requirements above is repeatable and confirmed in cross-validation experiments, the biomarker finally attains the status of “Known Valid Biomarker”.

3 Chromatographic separations of proteins and peptides

3.1 Chromatographic principles

Several parameters are defined for the characterization of HPLC separations. First, it has to be distinguished between isocratic and gradient elution. In isocratic separations the plate number per meter is the most important factor for characterization of separation efficiency while for gradient separations the peak capacity is the crucial parameter. The higher these values are, the higher is the efficiency of the separation column. The complete set of HPLC separation characterization parameters is summarized in the IUPAC nomenclature compiled by Engelhardt and Rohrschneider ^[60].

3.1.1 Optimization of resolution

The parameter for characterizing chromatographic separations aimed to be optimized in terms of efficiency is the peak resolution, describing the relative position of two analyte peaks to each other. To emphasize the influence of selectivity and plate number on the peak resolution, the equations for isocratic HPLC separation are given in equation (1) and (2).

$$R_s = \frac{1}{4} \sqrt{N} \cdot \frac{k_2}{1+k_2} \cdot \frac{\alpha-1}{\alpha}$$

efficiency
retention
selectivity

Equation (1)

R_speak resolution

Nplate number

k retention factor

αselectivity

The retention term is not affected by changing the column length as the flow time for an inert substance and the net retention time are changing in the same order of magnitude, too. Hence, the ratio of both (= k) stays the same. The selectivity term of the equation is not affected as well because the stationary phase, the mobile phase and the temperature are not changed. The only factor which alters by varying the column length is the efficiency term caused by the raising plate number with column length according to equation (2):

$$N = 16 \cdot \left(\frac{t_R}{w_b} \right)^2$$

Equation (2)

N.....plate number

t_Rtotal retention time [min]

w_bpeak width at basis [min]

For instance, the plate number is increasing almost to square value as t_R is doubling by doubling the column length (when w_b remains nearly the same).

Beside the variation of the efficiency term by changing the column length, the selectivity α (= k_2/k_1) can be changed by variation of the mobile phase, the stationary phase or the separation temperature ^[59].

However, these considerations are valid strictly only for isocratic HPLC separations. Nevertheless, the resolution for gradient separation is affected, too. A mathematical expression of N, k and α for gradient elution with the equations listed above is not strictly valid because no thermodynamic equilibration can adjust due to the permanently changing composition of the mobile phase. A formula for calculating the peak resolution of isocratic and gradient separations for two compounds (analyte 1 and analyte 2) is revealed in equation (3):

$$R_s = \frac{2(t_{R2} - t_{R1})}{w_{b1} + w_{b2}}$$

Equation (3)

R_speak resolution

t_{R1}total retention time of analyte 1 [min]

t_{R2}total retention time of analyte 2 [min]

w_{b1}peak basis width of analyte 1 [min]

w_{b2}peak basis width of analyte 2 [min]

3.1.2 Peak capacity

The theoretical peak capacity (P) is an important parameter to characterize a two-dimensional separation system. It describes how many components can be separated by an HPLC method ($R_s = 1$) in a particular gradient time. It is computed according to the equation ^[61]:

$$P = 1 + \frac{t_G}{w_b}$$

Equation (4)

Ppeak capacity

t_Ggradient time

w_bpeak width at the basis

P is calculated by the ratio of the gradient time (t_G) and the average peptide peak width (w_b) at 4σ (= 13.4 % of peak height) ^[62]. The total peak capacity of a multi-dimensional HPLC system is estimated by multiplying the peak capacities from the single dimensions ^[63]:

$$P_{\text{total}} = P_m \times P_n \dots \times P_z$$

Equation (5)

P_{total}peak capacity for the whole multi-dimensional system

P_mpeak capacity of the first dimension

P_npeak capacity of the second dimension

P_zpeak capacity of the last dimension

3.1.3 Gradient volume concept

For the transfer of an optimized gradient to another column length it is necessary to adapt the gradient volume which is calculated according to equation (6):

$$V_G = t_G \cdot F$$

Equation (6)

V_Ggradient volume

t_Gprogrammed gradient time [min]

Fsolvent flow-rate [mL min^{-1}]

When a liner gradient is applied, the inner diameter and the stationary phase of the separation column remain the same, and the gradient volume is constant, then each analyte is eluted at a certain solvent concentration, and the separation is only changing in terms of analysis time ^[64]. Furthermore the start and end conditions of the gradient separation have to remain the same.

Increasing the peak capacity can be achieved by increasing the column length. To employ the gradient volume concept first the hold up time of both columns has to be determined. The gradient time of the optimized gradient is known. The particular volumes are calculated by multiplication with the solvent flow-rate. To adapt the gradient volume for a column twice as long as the column where the gradient was optimized, the following computation is necessary:

$$\frac{V_{G1}}{V_{M1}} = \frac{V_{G2}}{V_{M2}}$$

Equation (7)

V_{G1}gradient volume for optimized gradient [mL]

V_{M1}hold up volume for column where the gradient was optimized [mL]

V_{G2}adapted gradient volume for the longer column [mL]

V_{M2}hold up volume for the longer column [mL]

$$V_{G2} = \frac{V_{G1}}{V_{M1}} \cdot V_{M2}$$

Equation (8)

$$t_{G2} = \frac{V_{G2}}{F}$$

Equation (9)

Employing a two-fold longer column would result in a doubling hold up time:

$$V_{M2} = 2 \cdot V_{M1}$$

Equation (10)

In Equation (8):

$$V_{G2} = V_{G1} \cdot 2$$

Equation (11)

This means that when the column length is doubled, the gradient volume has to be doubled, too. This results in a doubled gradient time when the solvent flow rate remains the same.

Furthermore, the gradient volume can be adapted to the application. In Table 4 the recommendations for chosen gradient volumes in terms of the employed application are summarized ^[64].

Table 4. Adjustment of gradient volume to application.

V_G	application	characterization
$5 V_M$	trace analysis of low complex samples	small gradient volume leads to small peak volume, low dilution, high sensitivity
$10-15 V_M$	optimized resolution of medium complex samples	optimal resolution for moderate analysis time
$> 15 V_M$	multi component analysis	optimization of peak capacity

3.2 Chromatographic modes

Since the term and the technique of chromatography have been introduced in 1903 by M. Tswett^[65;66], over one century of development in this field has passed. A great variety of chromatographic modes are established today as a result of the enormous progress in technology. Beside gas chromatography (GC), high-performance liquid chromatography (HPLC) has become the most important technique in bioanalytics. As GC-MS dominates the field of metabolite analysis, HPLC-MS is the prevalent choice for proteomic approaches. Peptides and proteins are accessible to a variety of HPLC methods because of their various properties like charge, polarity, biological and metal affinity^[67], as listed in Table 5^[68;69].

Size exclusion chromatography was introduced in 1956^[70]. Here, the analytes are not interacting with the material of the stationary phase, but are separated according to the accessible pore volume of the stationary phase.

In 1990, the term hydrophilic interaction chromatography (HILIC) was introduced describing the combination of hydrophilic stationary and hydrophobic mobile phase more accurately than the historical term normal-phase chromatography^[71]. This mode is especially used for separating sugars and oligosaccharides^[72-77]. In recent studies HILIC has been also applied to proteome analysis in multidimensional peptide separation due to the orthogonality towards RPC^[78;79].

Hydrophobic interaction chromatography (HIC) represents a special case of RPC and uses salt gradients of decreasing concentration to elute the adsorbed proteins. This method has been known since 1948^[80]. In contrast to ion exchange chromatography, a hydrophobic stationary phase material is applied, and the salt gradient starts at high ionic strength. The moderate eluting conditions are not affecting the tertiary and

secondary structure of proteins. As a consequence this chromatographic mode is highly suitable for protein purification without loss of protein activity.

Table 5. High-performance liquid chromatography techniques for protein and peptide separation.

method	mobile phase	principle of separation	stationary phase	abbrev-iation	alternative names /special cases
<i>non interactive methods</i>					
size exclusion chromatography	gradient of organic solvents or water	varying molecule size	silica or polymer gels with specific pore width	SEC	gel filtration (GF) gel permeation chromatography (GPC)
<i>interactive methods</i>					
normal phase chromatography	gradient of organic solvents (mixed)	polar interactions	silica gel aluminum oxide dextran gel	NPC	hydrophilic interaction chromatography (HILIC)
reversed phase chromatography	gradient of organic solvents with water or buffer	solvophobic interactions	silanized silica gel, silica or PS-DVB	RPC	hydrophobic interaction chromatography (HIC)
ion-pair reversed phase chromatography	gradient of organic solvents with water or buffer and ion pair modifier	solvophobic and electrostatic interactions	silanized silica gel, silica or PS-DVB	IP-RPC	
ion-exchange chromatography	gradient of buffer solution	electrostatic interactions	silica or polymers with anionic or cationic functional groups	IEC	strong- or weak anion-exchange chromatography (SAX/WAX) strong- or weak cation-exchange chromatography (SCX/WCX)
affinity chromatography	gradient of buffer solution	biospecific interactions, complex formation	affinity carrier with specific ligands	AC	immobilized metal ion affinity chromatography (IMAC)

Compared to small molecules, proteins and peptides differ mainly in their more complex structure and larger molecule size. This means that the analyte molecules can interact with the stationary phase of the separation column with more than one moiety. That effect is called multisite adsorption^[81]. Therefore, a protein elutes only when the solvent strength has reached a special value and hence is high enough to desorb all groups of the protein interacting with the stationary phase. This phenomenon has been introduced as “on-off model”^[82]. As a consequence, a steep

negative slope can be observed in a semilogarithmic plot of the retention factor versus the solvent strength, as depicted for lysozyme in Figure 4 in comparison to nitrobenzene. Furthermore, multisite adsorption leads to focusing of the analytes at the column head which allows the injecting of large volumes of biopolymer samples.

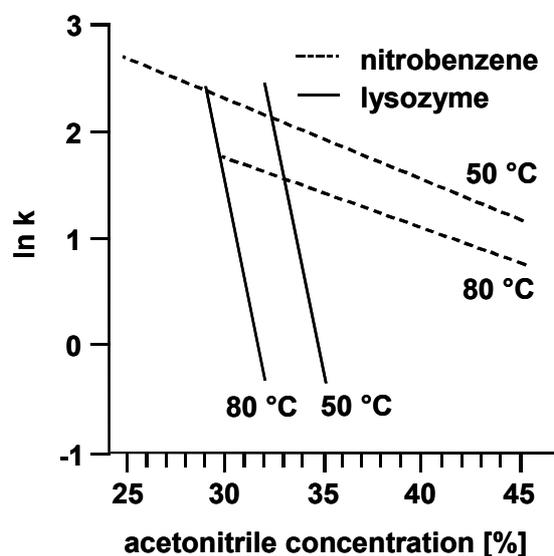


Figure 4. Desorption behavior of a small molecule (nitrobenzene, dashed lines) and a protein (lysozyme, solid lines) in RP-HPLC ^[83].

3.2.1 Reversed-phase and ion-pair reversed-phase HPLC

In reversed-phase high-performance liquid chromatography (RP-HPLC), proteins and peptides are separated with respect to their hydrophobic properties. The hydrophobic moieties undergo solvophobic interactions to a hydrophobic stationary phase in a polar surrounding medium (mostly water). Elution of the analytes is realized by raising the concentration of organic solvents (e.g. methanol or acetonitrile). Small polar acids like formic, hydrochloric or phosphoric acid are often added to denature the proteins and circumvent secondary unspecific interactions with remaining silanol groups of silica-based stationary phases ^[84]. As a consequence, peaks with higher symmetry can be observed. Under the described conditions, elution processes take place mainly caused by solvophobic interactions ^[85]. Protein charge only affects separation by adjusting the hydrophobicity.

In ion-pair reversed-phase chromatography additional electrostatic interactions are induced by adding amphiphilic compounds like trifluoroacetic- or heptafluorobutyric acid or bases like triethylammonium acetate to the mobile phase (=ion-pair modifiers). As a result, the hydrophobic moieties of the ion pair modifier adsorb on the hydrophobic stationary phase and generate an electrical surface potential, which attracts charged functional groups of the analytes ^[86]. Since the concentration of the amphiphile is usually not high enough to cover the complete surface of the stationary phase solvophobic interactions are still contributing to retention of the analytes ^[87;88] as revealed in Figure 5a.

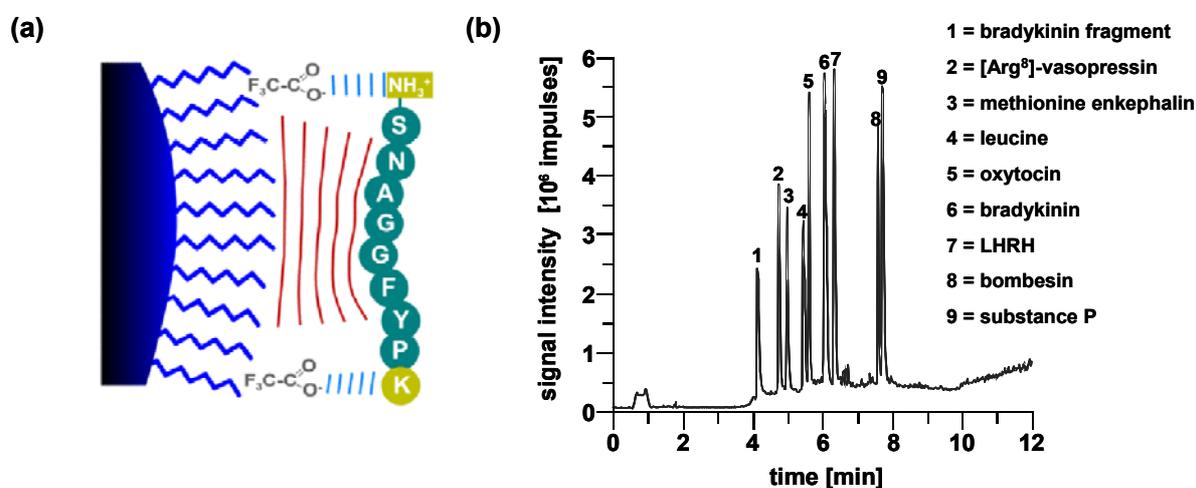


Figure 5. IP-RP- HPLC ^[89;90].

(a) Principle of IP-RP-HPLC, solvophobic as well as electrostatic interactions shown on the example of a peptide adsorbing to a hydrophobic octyl silica stationary phase applying TFA as ion-pair modifier. (b) IP-RP-HPLC separation of nine peptides. Column, PS-DVB monolith, 60 x 0.2 mm i.d.; mobile phase, (A) 0.05 % TFA in water, (B) 0.05 % TFA in acetonitrile; gradient, 0-50 % B in 15 min; flow rate, 2.5 $\mu\text{L}/\text{min}$, temperature, 50 $^{\circ}\text{C}$, detection, ESI-MS; peptide concentration, 1ng μL^{-1} ; injection volume, 1 μL .

RP-HPLC and IP-RP-HPLC are the most efficient chromatographic separation methods because of protein denaturation. They provide the highest peak capacities caused by suppression of conformational isomers and maximization of the molecule surface available for interaction. Under the denaturing conditions proteins are unfolded and losing their three-dimensional structure but also their biological activity. To optimize the separations, selectivity can be influenced by temperature and choice

of stationary phase as well as mobile phase adjustment using other organic solvents or other ion-pair modifiers. An exemplary separation of nine peptides in a monolithic capillary column is demonstrated in Figure 5b.

3.2.2 Ion-exchange high-performance liquid chromatography

The principle of ion-exchange high-performance liquid chromatography (IEX-HPLC) is based on charge-charge interactions ^[91;92]. The positively or negatively charged analytes can electrostatically interact with charged functional groups of the stationary phase. For proteins and peptides such interactions are promoted by the amino- and carboxy-terminus, the charged side chains of acidic (aspartic and glutamic acid) and basic (lysine and arginine, histidine - depending on the pH value) amino acid residues as well as by post translational modifications like phosphate or sulfate groups. Depending on the pH value of the surrounding medium proteins and peptides can either be separated with anion exchange (AEX-HPLC) or with cation exchange high-performance liquid chromatography (CEX-HPLC) due to their ampholyte character ^[93].

In Figure 6a and c exemplary protein separations applying AEX- and CEX-HPLC are shown and Figure 6b reveals schematically the principle of AEX-HPLC.

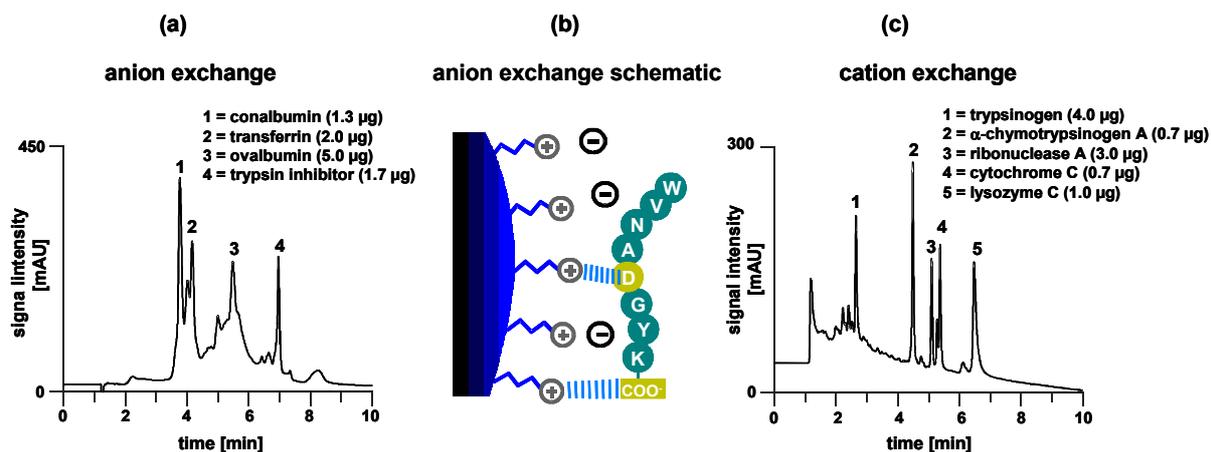


Figure 6. Ion-exchange HPLC of proteins and peptides ^[89].

(a) Column, 250 x 0.4 mm i.d. ProPac SAX-10, 10 μm ; gradient, 0-0.5 M NaCl in 20 mM Tris-HCl within 10 min, pH 8.0; flow rate, 1.0 mL min^{-1} ; temperature, ambient; UV-detection, 220 nm; injection volume, 10 μL . (b) Scheme of anion exchange for an octapeptide. (c) Column, 250 x 0.4 mm i.d. ProPac SCX-10, 10 μm ; gradient, 0-0.5 M NaCl in 50 mM Na_2HPO_4 within 10 min, pH 6.0; flow rate, 1.0 mL min^{-1} ; temperature, ambient; UV-detection, 220 nm; injection volume, 10 μL .

Although the net electrical charge is zero at the isoelectric point (pI) of a protein, retention can still be observed in IEX-HPLC. This phenomenon occurs because not exclusively the net electrical charge contributes to separation, but also the charge distribution. For separation of peptides and proteins with AEX-HPLC the pH value of the surrounding medium has to be higher and for CEX separations lower than the pI of the analyte. In general, the pH value should differ one pH unit to the pI of the analyte to get repeatable retention of peptides and proteins.

The elution of analytes in IEX-HPLC is controlled by increasing the ionic strength of the solvent. As long as no extreme pH conditions are created the three-dimensional structure and the biological function of proteins are not impaired. However, the maintained complex structure leads to broader or even multiple peaks especially for proteins keeping their isomeric forms. As a result, in IEX-HPLC lower peak capacities are obtained than in RP- and IP-RP-HPLC. Furthermore, the high amounts of salt hamper online hyphenation to mass spectrometry. For protein purification ^[94] ion-exchange chromatography is a wide spread methodology.

3.3 Stationary phase materials for separation of proteins and peptides

One main difficulty of HPLC separations of proteins and peptides is the slow mass transfer caused by the low diffusion constant of large molecules. The slow movement inside and out of the pores of a porous stationary phase material delays the mass transfer resulting in band broadening. To improve separation efficiency different phase morphologies were suggested as displayed in Figure 7.

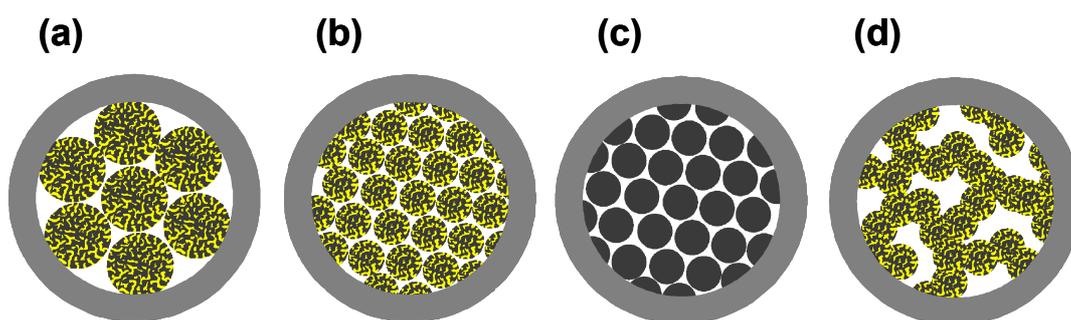


Figure 7. Morphology of four different stationary phase configurations ^[90].

(a) Spherical, porous and large particles, (b) spherical, porous and small particles, (c) spherical, non-porous and small particles and (d) monolithic stationary phase.

In Figure 7a a porous stationary phase is shown. It consists of large spherical and porous particles with high permeability and large interparticle voids. However, this design results in low separation efficiency caused by the long diffusion paths in the pores. To shorten them, smaller spherical, porous particles were introduced (Figure 7b). They actually achieved higher efficiency, caused by small interparticle voids and shorter pores, admittedly at the cost of lower permeability. As a result, the required pressure for pumping mobile phase through the column increases. Another stationary phase design, so-called pellicular stationary phases ^[95], is totally non-porous as schematically demonstrated in Figure 7c. For this phase design the efficiency is also very high because of the small interparticle voids, but low permeability is caused by small particle sizes necessary to obtain sufficient surface area. Furthermore, only the surface material of the particles contributes to retention of analytes at the cost of separation area and so on sample loadability. The last of these four stationary phases is the monolithic design which offers a high permeability and efficiency discussed in detail in the next paragraph.

3.3.1 Monolithic columns for peptide and protein separation

Monolithic separation columns characteristically consist of a single porous medium^[96,97]. One advantage of this structure is the absence of interparticle voids. Thus, the mobile phase is pumped through the macro-pores of the separation medium^[98]. The mass transfer hereby is supported by the convective flow through the macro pores (> 50 nm)^[99]. Diffusion dependent mass transfer is restrained only to the meso- (2-50 nm) and micro- (< 2 nm) pores. A further advantage is the high permeability of monolithic columns allowing fast separation of biomolecules at high flow rates^[100-103]. Caused by their flat van-Deemter curves at high linear velocities, monolithic columns combine excellent separation performance and short analysis times^[104].

There are two main strategies for monolith preparation: Polymerization of organic monomers and polycondensation of alkylsilane monomers. The first monolithic columns were organic, synthesized in 1974 based on polyurethane^[96]. Since then a variety of organic monolithic materials based on the polymerization of acrylates, methacrylates, norbornene, styrene and other monomers have been shown^[105]. PS-DVB monoliths were recognized as monolithic columns with especially high efficiency for biopolymers (Figure 8). Compared to columns packed with 2 µm PS-DVB particles, they offer a three times higher efficiency demonstrated on nucleotides as biopolymers^[104]. In these micropellicular monoliths no accessible micro- or mesopores are present for large biomolecules as nucleotides, proteins or peptides^[101].

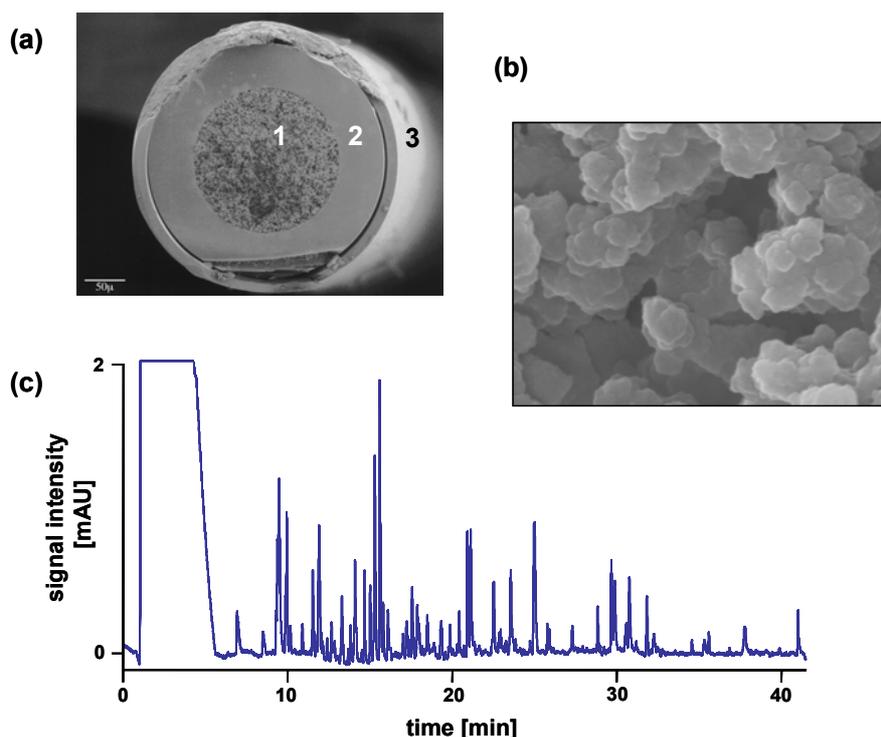


Figure 8. PS-DVB-based monolithic columns and performance ^[106;107].

(a) 200 µm i.d. monolithic column, outwards: (1) monolithic stationary phase, (2) fused silica, (3) polyimide layer (scanning electron microscope image). (b) Globules of the monolithic phase (scanning electron microscope image). (c) Separation of 1.0 pmol human transferrin peptides. Column, monolithic PS-DVB 60 x 0.2 mm i.d.; mobile phase, (A) 2.0 % acetonitrile and 0.05 % TFA in water, (B) 80 % acetonitrile and 0.05 % TFA in water, gradient, 0-40 % B in 30 min, 40-100 % B in 10 min; flow rate, 1.70 µL min⁻¹; temperature, 50 °C; UV-detection, 214 nm ^[107].

In the middle of the 90s, silica-based monoliths were introduced for HPLC separations ^[108] which are synthesized by a combination of sol-gel reaction and agglomeration ^[109]. The most commonly adopted silica monomer is tetramethoxysilane (TMOS) which was first employed by Tanaka et al. ^[110-112]. Later, further monomers were tested separately or in combination: tetraethoxysilane (TEOS) ^[113], methyltrimethoxysilane (MTMS) ^[114], 2-cyanoethyltriethoxysilane (CEOS) ^[115], (3-aminopropyl) triethoxysilane (APTES) ^[116] and diglycerylsilane (DGS) ^[113;116]. First, silica monoliths were particularly valuable in the field of small molecules ^[117] and became then more and more applied in the field of peptidomics and proteomics ^[118;119].

3.3.2 Strong cation exchange columns for peptide separation

For 2D proteome analyses based on chromatography and mass spectrometry SCX is often applied in the first dimension for peptide separation and pre-fractionation. The SCX chromatography mode is chosen because the stationary phase materials are protonated over a broad pH range and hence keep their ability of cation exchange. A suitable pH for peptide separation would be around three where all carboxyl groups are protonated and not disturbing the interactions of positively charged peptide residues and the stationary phase. Both, silica as well as polymeric base materials are suitable for SCX as indicated in Table 6.

Table 6. Common base materials and functional groups for SCX-HPLC ^[92;120].

materials		
property	polymer-based	silica-based
ion exchange capacity	high	high
loadability	high	medium
pH range	2-14	2-8
efficiency	low	high
functional groups for ion exchange		
name	formula	
sulfoethyl groups	$-\text{O}-\text{CH}_2\text{CH}_2\text{SO}_3^- \quad \text{H}^+$	
sulfopropyl groups	$-\text{O}-\text{CH}_2\text{CH}_2\text{CH}_2\text{SO}_3^- \quad \text{H}^+$	

Poly(2-sulfoethyl aspartamide) was shown to be advantageous as silica-based column packing material for cation exchange ^[121]. Its structure is shown in Figure 9. Compared to sulfopropyl groups, undesirable secondary hydrophobic interactions are decreased. This is a critical point especially for hydrophobic peptides as they form interactions with the sulfopropyl moieties of the stationary phase.

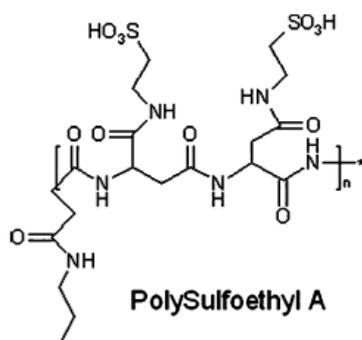


Figure 9. Structure of poly(2-sulfoethyl aspartamide) (= PolySulfoethyl A) as stationary phase material in SCX-HPLC.

4 Mass spectrometry

Mass spectrometry has become an important analytical technique for proteome research and bioanalysis at all. Due to its high selectivity, sensitivity and resolution power it is applied to 2D HPLC proteome analysis. Since a mass spectrometer consists of three modules (ion source, mass analyzer and detector) this chapter is arranged in the same way. First matrix assisted laser desorption/ionization (MALDI) as a “soft” ionization method for biopolymers is discussed. This is, beside electrospray ionization (ESI), the most common technique to transfer biopolymers like peptides, proteins and nucleic acids into the gaseous phase without fragmentation.

The analytes enter the mass spectrometer charged by the pH value of the dissolution buffer. In the case of chromatography coupled to mass spectrometry the charge of proteins and peptides depends on their pI and the pH value of the mobile phase. After passing the ion source they travel the high vacuum region of the mass analyzer. Depending on the ionization technique and the employed setup the ion source either operates under vacuum or at atmospheric pressure. In a complex sample more than one analyte enter the mass analyzer simultaneously. Here they are separated according to their mass to charge ratio finally reaching the detector.

4.1 MALDI as ionization technique for biopolymers

Since the first publications of MALDI employed for protein ionization in 1988 by Karas, Hillenkamp ^[122] and Tanaka ^[123], it has become a widely spread technique for transferring large and non volatile molecules into the gaseous phase without fragmentation.

The matrix employed in the MALDI process is usually a small organic molecule with the ability to absorb UV light. The matrix fulfills three tasks: it separates the sample molecules from each other, absorbs the laser energy partly (because otherwise the molecules would fragment immediately), and the third task of the matrix is to serve as a proton-acceptor or -donor to the analyte molecules for ionization.

Analyte and matrix are separately dissolved in a suitable solvent like water, methanol, acetone, acetonitrile or a mixture of these for sample preparation. Then

both solutions are mixed to contain excess of matrix in the final solution (usually 1,000:1 or higher ^[124]). Subsequently, both solutions are transferred to the surface of the sample carrier consisting mostly of stainless steel, but also gold or other metal surfaces are possible ^[125]. In Figure 10 the principle of MALDI is depicted schematically.

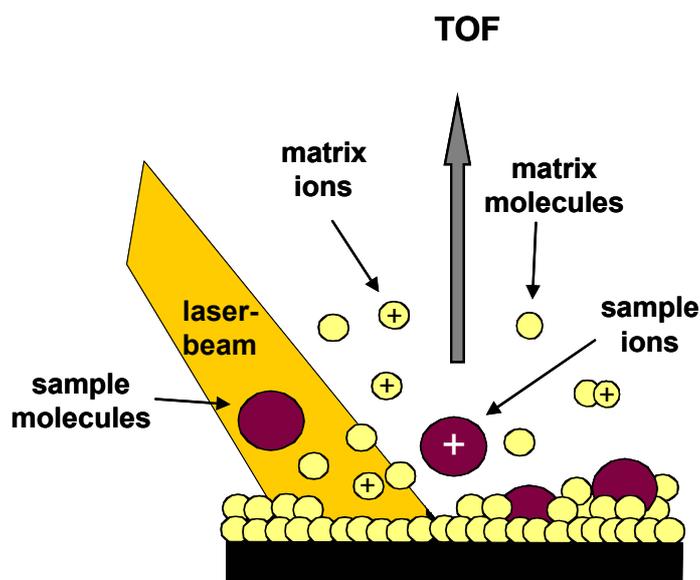
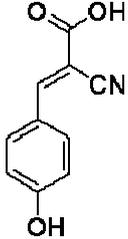
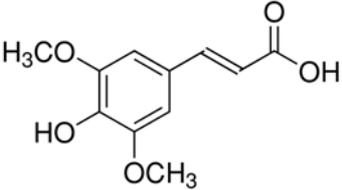
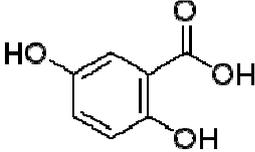
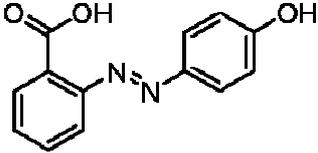


Figure 10. Scheme of ion generation during MALDI process.

After co-crystallization of matrix and analyte, a pulsed short wave laser shoots on the co-crystal whereas one shot lasts only few nano seconds. Its energy is sufficient to create analyte ions and transfer them into the gaseous phase. The whole ion-generation process is very complex and yet not totally understood. It is suggested ^[92] that in the matrix molecules stored laser energy relaxes the solid state grid resulting in expansion of it. Before reaching a thermal equilibrium a phase transition takes place which resolves explosively a part of the solid state surface. In a specific range of laser fluence (10^6 - 10^7 W cm⁻²) ^[126] even large labile molecules, such as peptides and proteins, survive this process unfragmented. Most common applied UV lasers are nitrogen (337 nm) and Nd-YAG (266 nm or 355 nm) lasers. The fluence as well as the laser wavelength influence the analysis. The matrix absorbance should not overlap with the wave length of the laser. The mostly relevant matrices for proteomics and other fields in bioanalysis are listed in Table 7.

Table 7. Most common matrices suitable for MALDI.

matrix	abbreviation	common applications	structure
α -cyano-4-hydroxycinnamic acid	HCCA	peptides proteins	
sinapinic acid	SA	peptides proteins	
2,5-dihydroxybenzoic acid	2,5-DHB	peptides phosphopeptides proteins carbohydrates glycolipids	
2-(4-hydroxyphenyl azo)-benzoic acid	HABA	proteins lipopolysaccharides	

4.2 Mass analyzers

4.2.1 Overview

After generating an ion beam, the sample ions have to be separated which is realized in mass analyzers. A variety of them are available for bioanalytical research. A lot of different physical principles have been applied to construct mass analyzing systems. Some of them became very successful from the beginning of their existence, for others it took decades until their potential was recovered again. Although the perfect mass analyzer is exactly outlined ^[127], in spite of all recent improvements still none is ideal. However, very useful tools for structure determination, identification and

quantification of biological important substances, such as proteins, peptides, nucleotides and metabolites, are disposable today. The most common mass analyzers are listed in Table 8.

Table 8. Common mass analyzers ^[128].

type	acronym	ion properties	separation principle
time of flight	TOF	time dispersion of a pulsed ion beam	time of flight
magnetic sector	B	deflection of continuous ion beam	momentum in magnetic field due to Lorentz force
linear quadrupole	Q	continuous ion beam in linear radio frequency quadrupole field	stability of trajectories
linear quadrupole ion trap	LIT	continuous ion beam and trapped ions	storage and eventually separation in linear radio frequency quadrupole field due to stability of trajectories
quadrupole ion trap	QIT	trapped ions	in three-dimensional radio frequency quadrupole field due to stability of trajectories
ion cyclotron resonance	ICR	trapped ions	cyclotron frequency (Lorentz force) in magnetic field

In the next paragraph, the time of flight mass analyzer will be described in detail because it is the applied system in the research of this thesis and very widespread in combination with ESI and MALDI ionization in bioanalysis.

4.2.2 Time of flight mass analyzer

Times of flight (TOF) instruments were first commercially available in the middle of the 1950s ^[129]. Due to poor performance they were soon pushed out of the market by linear quadrupole analyzers. Their revival took until the late 1980s, when MALDI appeared as ionization technique for large biomolecules ^[130;131].

The basic principle of TOF-MS is the acceleration of analyte ions and subsequent drifting in a field free flight tube where they are separated according to their velocities. In the ion source all analyte ions are accelerated with the same kinetic energy. Based on their different masses however, the velocity of different ion species

will be diverse because the applied kinetic energy is equal for all ions and directly proportional to the mass and the velocity of an ion as revealed in Equation (12):

$$E_{kin} = \frac{1}{2} mv^2$$

Equation (12)

E_{kin} kinetic energy [$m^2 \text{ kg s}^{-2}$]

m mass of the analyte [kg]

v velocity of the analyte ion [$m \text{ s}^{-1}$]

To yield the same kinetic energy, an accelerating voltage is applied to the analyte ions. Due to their different velocities this principle is used for determination of the mass-to-charge-ratio and in the end the mass of the analyte in a specific time domain and fixed flight path distance as shown in Equation (13):

$$m/z = \frac{2Ut^2}{D^2}$$

Equation (13)

m/z mass-to-charge-ratio

U acceleration voltage [kV]

t time of flight [s]

D flight path distance [m]

ESI-TOF analyzers are very suitable for on-line hyphenation of nano-flow HPLC to mass spectrometry. Hyphenation of nano-flow HPLC to MALDI-TOF mass spectrometry however is normally realized in off-line mode caused by the need of co-crystallization of the sample with the solid matrix.

To reduce peak broadening and loss of resolution caused by initial time-, kinetic-energy- and space distribution effects, two strategies were developed. First, a so-called reflector is installed at the end of the drifting path of the ions established by Mamyryn in 1994 ^[132]. The reflector consists of a series of metal rings or grids at rising potential, is located opposite the ion source and a voltage around 1.1 times higher than the ion source voltage is applied. Analyte ions penetrate the electric field of the reflector until their kinetic energy is reduced to zero and are then expelled in opposite direction. Ions of the same species subjected to time distribution effects in the ion

source can thus be focused. Summarized, three main factors support adoption of a reflector:

1. substantially improved resolving power due to a longer flight path of the ions
2. corrected time dispersion due to initial kinetic energy distribution of ions of the same mass
3. neutral molecules are filtered out

The reflector can be assembled in single-, two- and multiple-stage design. An additional detector is often installed behind the reflector, and by switching it on or off both a linear and a reflectron TOF can be realized in one instrument. The reflector process is schematically shown in Figure 11.

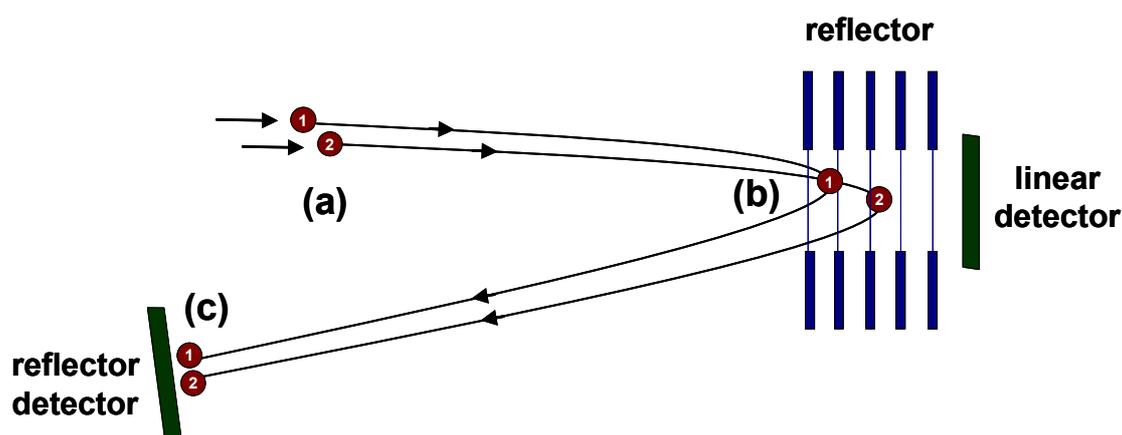


Figure 11. Principle of the reflector in TOF-MS.

(a) Ion 1 and 2 possess the same mass and charge, but different velocities; (b) faster ions (ion 2) penetrate deeper in reflector field than slower ones (ion 1); (c) simultaneous arrival at the detector.

The second strategy for sharpening ion distributions in the ion source was already introduced in 1955 - the time-lag focusing ^[133]. Caused by established patents every manufacturer has its own synonym for it: Time lag focusing (TLF, Micromass), Delayed extraction (DE, Applied Biosystems) and Pulsed ion extraction (PIE, Bruker Daltonik). However, the principle remains the same: A variable voltage plate is installed between the ion source and the entrance to the flight tube, as demonstrated in Figure 12 for a linear TOF (without reflector). When the sample is irradiated by the

short laser shot the created ions are not accelerated immediately but with a delay time of 100 to 500 ns ^[134]. During this time no voltage is applied to both sample plate and variable voltage plate. The ions spread out in the space between both plates according to their initial energy distribution as drafted in Figure 12a. Then voltage is switched on with a fast pulse at both plates but on the variable voltage plate only a percentage of the sample plate; in the example of Figure 12b 90 %. As a consequence, a voltage gradient is formed which accelerates slow ions more because they remain in the stronger voltage field. As a result, the initial energy distribution is adjusted and all ions of one species offering the same mass-to-charge ratio reach the detector to the same time (Figure 12c). With the next laser pulse the high voltage is switched off again and the cycle starts anew. This pulsed ion generation permits very fast data acquisition.

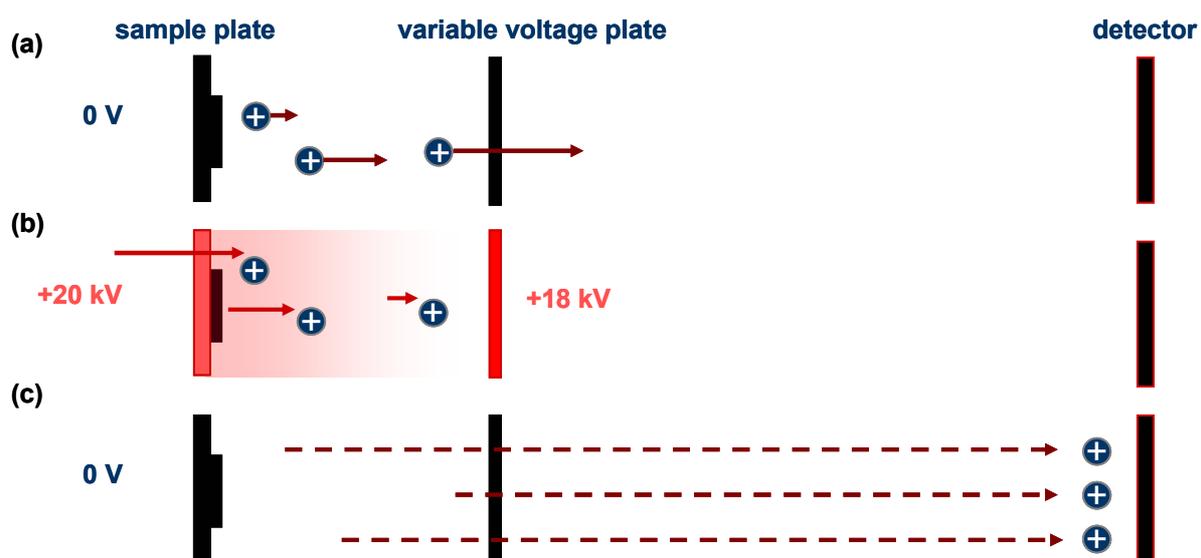


Figure 12. Time-lag focusing technique.

(a) No electric field is applied, ions of same mass and charge with different velocities; (b) an electric field gradient is applied to focus ions of the same mass; (c) ions of the same species arrive the detector simultaneously after traveling the field free drifting zone.

4.2.3 Precursor selection and peptide fragmentation in a MALDI-TOF/TOF mass analyzer

The pulsed laser shoots on the sample/matrix spots of the target which is situated in the ion source. Thereby, generated ions are focused with delayed extraction (paragraph 4.2.2) and accelerated for mass analysis. After traveling the field free drifting passage in the flight tube the ions reach the reflector (paragraph 4.2.2) where a further focusing follows decreasing the time- and energy haziness. Then the analyte ions are accelerated again and impinge onto the detector. There, the signal is amplified and converted into an electric current which can be observed with a PC-based operating system. Now, the masses for fragmentation are chosen automatically by the software. A set number of precursors (e.g. five) are selected for one spot whereas the strongest signal is fragmented first (or last when selected). This does not mean that the five strongest signals of one spot are fragmented, but refers to the whole first analysis of intact peptide masses. When the highest peptide signal in one spot is not fragmented it is even stronger in another spot and there selected for fragmentation. In a second mass spectrometric run the precursors can be fragmented without using the CID chamber as post source dissociation (PSD) with subsequent acceleration ^[135] or by additional collision with gas molecules (CID). The metastable decomposition of the peptides takes place in the field free drifting passage they pass after the first acceleration. By then the precursor and fragment ions possess the same velocity because they were accelerated as one ion, and during the metastable decay process no decelerating effect occurred according to the vacuum applied in the flight tube. Therefore, they pass the timed ion selector (TIS) at the same time as illustrated in Figure 13.

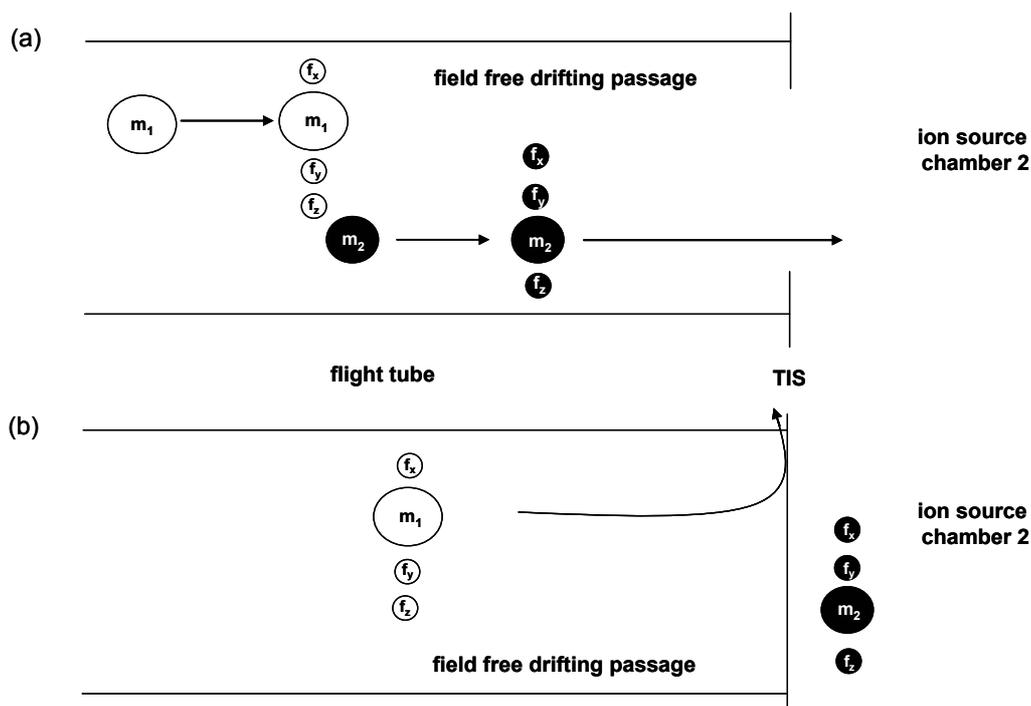


Figure 13. Schematic sketch of precursor passing controlled by the timed ion selector (TIS). (a) A selected precursor passes the TIS and (b) a non selected ion is deflected.

The TIS lets only the selected precursor mass-to-charge-ratio (m/z) pass by including all fragment masses; here demonstrated with the black ion of a mass m_2 with its symbolic fragments f_{x-z} (Figure 13a). As the white ion with a mass m_1 has a higher mass, it is slower than the black ones after acceleration with the same energy and deflected as shown in Figure 13b. Hence, the TIS has to be triggered with a start signal when the ions enter the field free drifting passage to let the selected precursor pass at exactly its arrival time at the TIS. This time is calculated by the software for each precursor ion.

4.3 Detectors in mass spectrometry

The last module of a mass spectrometer is devoted to ion detection and digitalization of the signals for computer processing. All currently employed detectors are secondary electron multipliers (SEM).

4.3.1 Overview of secondary electron multipliers

An overview to the state of the art detectors for mass spectrometry is given in Table 9. All SEMs are working with the same principle: An ion impinges onto the metal or semiconductor surface of a dynode leading to the emission of secondary electrons. An electrode is installed opposite of this dynode, holding a more positive potential. The emitted electrons are accelerated then and hit the opposite electrode, emitting in turn several electrons each. After passing 12-18 discrete dynodes an electric current is created large enough to be detected by a sensitive preamplifier. The SEMs have to be operated in vacuum due to air sensitivity of the emission layer and prevention of arcing caused by the high voltage applied.

Table 9. Overview of SEMs as detectors in mass spectrometry.

detector	synonym	acronym	principle of operation
discrete dynode electron multipliers	-	classic SEM	arriving ions emit electrons from dynode surface which are amplified in a cascade
channel electron multipliers	channeltron	CEM	SEM in continuous tube
microchannel plates	multichannel plate, channel electron multiplier array	MCP	CEM in reduced size to μm -diameters; bundled to millions of those tubes
post-acceleration and conversion dynode	-	-	acceleration prior to detector (10-30 kV) to circumvent discrimination of slower ions
focal plane detectors	array detectors	FPD	ions arriving at focal plane on surface of MCP are converted to electrons; on the backside of MCP electrons converted into photons by phosphor screen; light image guided onto photo array detector

The microchannel plate detector is the most commonly employed detector in mass spectrometry and also integrated in the mass spectrometer used for this thesis. Therefore, it is described in the next paragraph.

4.3.2 Microchannel plate detector

In channeltron electron multiplier (CEM) detectors materials withstanding the high voltages around 2 kV are required. This is realized by an emissive layer of silicon dioxide. This coats a conductive layer of lead oxide on the supporting heavily lead-

doped glass tube ^[136;137]. By reducing the size of a linear CEM to μm -diameters the cross section would be much too small to generate sufficient electric currents. Therefore, millions of these are bundled. Employing this array design an effective detector is created occupying as little space as possible – the microchannel plate detector (MCP) which is displayed in Figure 14. The channel diameters ranging from 8 to 25 μm with a length 40-80 times the diameter and the center-to-center distance in the range from 6 to 32 μm ^[138]. A voltage of 1-2 kV is applied. The amplification of the electrons can reach up to 10^5 ^[139].

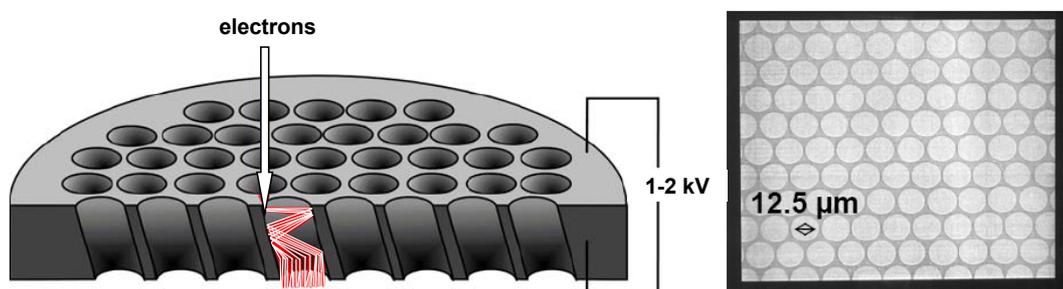


Figure 14. Microchannel plate detector ^[140].

In MCP detectors slow ions are discriminated. They are not arriving at the detector because a threshold energy density is required for electron emission. To gain higher sensitivity a conversion dynode is installed prior to the detector. These are rugged electrodes set on high voltages (5-20 kV) of opposite polarity of the ions leaving the mass analyzer. Furthermore, this dynode removes photons and neutral particles from the ion current when not installed in-line.

5 Two-dimensional HPLC-MS for proteome analyses

5.1 Overview of chromatographic mode combinations

Due to the enormous complexity of biological samples especially in proteome research, 1D chromatographic approaches are not sufficient to achieve the demanded resolving power and removing of matrix effects. To optimize selectivity and separation efficiency comprehensive 2D-chromatography was advocated. Here, two columns supporting different chromatographic modes are hyphenated on-line. Another opportunity is the so-called heart-cutting mode where only selected fractions obtained from the first dimension are transferred on-line into the second dimension ^[141]. Off-line coupling is also possible, providing the advantage of storing the fractions obtained from the first dimension. Furthermore, multiplex analyses of them by only one sample injection into the first dimension are possible. There are several criteria which ideally should be fulfilled in a truly comprehensive 2D HPLC system ^[142;143]. The first criterion, also called the orthogonal criterion, implies the separation of the substances with two different separation mechanisms. This is realized either by two different stationary phases or by changing the mobile phase. As second point, all analytes have to pass both dimensions at equal percentages (100 % or less) and finally reach the detector. This is only relevant for on-line and heart-cutting mode where a split ratio has to be defined. The last criterion is the maintenance of separation resolution, or at least the resolution reduction should not exceed 10 %. However, in most current applications, the sampling rate is not as high as recommended because the more fractions collected in the first dimension the higher becomes the time-consumption of the total 2D analysis.

The possibilities of 2D HPLC setups are multifarious. A selection of common combinations is listed in Table 10.

Table 10. Common combinations of chromatographic modes in 2D HPLC and selected applications.

first dimension chromatography mode	second dimension chromatography mode	analytes	Example(s) of application
change of stationary phase			
IEX-HPLC	(IP)-RP-HPLC	peptides	proteome of <i>S. cerevisiae</i> ^[144] ; proteome of <i>M. xanthus</i> ^[145]
AC	(IP)-RP-HPLC	peptides	human plasma/serum ^[146]
IMAC	(IP)-RP-HPLC	peptides	phosphopeptides ^[147;148]
SEC	(IP)-RP-HPLC	proteins, peptides	proteome of <i>E. Coli</i> ^[149] ; human serum ^[150]
HILIC	(IP)-RP-HPLC	peptides	murine erythroleukemia (MEL) cells ^[151]
change of mobile phase			
RP-HPLC	(IP)-RP-HPLC	peptides	proteome of <i>C. glutamicum</i> ^[152]

As detector for identification of peptides a mass spectrometer is usually employed providing excellent selectivity, resolution and sensitivity, but no efficient separation. This is why a 2D HPLC system has to be installed prior to mass spectrometric sample analysis.

5.2 Nano-flow HPLC-MALDI-MS as platform for proteome research

The development of efficient TOF mass spectrometers with the capability of precursor fragmentation at the beginning of this century^[153-155] made MALDI-MS appealing for proteome analysis. As alternative to nano-flow HPLC-ESI-MS, methods were introduced employing nano-flow HPLC-MALDI-MS. Although the mixing of a matrix and subsequent sample/matrix co-crystallization hampers an on-line hyphenation to mass spectrometry both, on-line and off-line approaches were developed.

5.2.1 Interfaces for nano-flow HPLC-MALDI mass spectrometry

The simplest and most straightforward method for off-line HPLC-MALDI mass spectrometry is the collection of a series of individual spots directly onto the MALDI target ^[156]. A heated droplet interface was introduced ^[157] where the MALDI target as well as the HPLC effluent transfer tubes are heated. The droplets are hanging at the outlet of the transfer tube and as much of the carrier solvent as possible is evaporated without spraying. The temperature of the MALDI plate is adjusted above the boiling point of the least volatile solvent. When the droplet lands on the plate, the remaining carrier solvent is evaporated immediately, depositing the analyte/matrix spot in a small area of the MALDI target. Miliotis et al. introduced a method for spotting the HPLC effluent employing a piezoelectric flow-through micro-dispenser ^[158]. By applying a short voltage pulse at a frequency of 15 Hz, pressure pulses are generated that eject 60-nL droplets onto the target. This is entirely precoated by air-brushing with a matrix (α -cyano-4-hydroxycinnamic acid)/nitrocellulose layer. On the contrary to discrete spot generation continuous sample deposition is described ^[159;160]. Here, the target is precoated with matrix and the analytes eluting off the HPLC system are deposited onto a track on the MALDI target. The continuous deposited tracks preserve chromatographic integrity. Resolution and sensitivity are determined by the drive speed.

Another approach applies electrospray for deposition of the sample/matrix mixture onto the target ^[161;162]. The advantage of this method is the fast solvent evaporation and more homogeneous sample/matrix co-crystal due to small droplets generated by electrospray. However, the electrospray is distributed over a relatively broad circular area (approximately 2.5 cm). An approach also based on electrical fields was developed by Ericson et al. ^[163]. The end tip of the capillary separation column is fixed in a capillary clamp which is installed 2-5 mm above the MALDI target. Here, an electrically grounded droplet is generated by the sample/matrix solution. Then a short (20 ms) negative voltage impulse (-2 kV) is applied to the target stage. This polarizes the droplet which is then forced towards the target surface by the electrical field. Electrospray deposition in this approach is prevented by applying relatively low electrical field strength at the droplet surface and the pulsed nature of the applied voltage. Therefore, the sample/matrix mixture is concentrated in discrete spots.

In the previously presented interfaces the analyte/matrix deposition is operated off-line to the mass spectrometric analysis. Also on-line approaches were reported.

The rotating ball inlet (ROBIN) was used for on-line MALDI MS ^[164]. Here, a stainless steel ball (10 mm diameter) is installed in the MS vacuum chamber and is rotated by a drive shaft connected to a gear motor which is positioned outside the vacuum chamber. The sample/matrix solution is delivered from the end of a capillary to the ball surface. The ball rotates past a polymer gasket and the volatile solvents are evaporated into the vacuum region. The laser ablates the sample directly from the ball surface. However, this approach is limited in terms of resolution due to the relatively high pressure (5×10^{-5} torr) and nonparallel fields in the MALDI ion source. Furthermore, signal intensities are influenced by irregularities in the ball surface.

A MALDI-MS interface especially used for on-line coupling with capillary electrophoresis (CE) was developed by Karger et al. ^[165]. In principle, the sample/matrix mixture is deposited onto a rotating quartz wheel and then transferred to the repeller, where the MALDI process takes place.

Today, the collection of discrete spots and subsequent off-line transfer to the mass spectrometer is the most employed interface in terms of nano-flow HPLC-MALDI-MS ^[166]. Automatic spotting units are commercially available while no on-line spotting approach is commercialized today.

5.2.2 Proteomes analyzed by 2D-HPLC-MALDI tandem mass spectrometry

The most approaches applying MALDI mass spectrometry are gel-based ^[167] according to low volumes necessary for sample preparation (usually 0.5 μ L of sample solution) and higher salt tolerance than required for ESI mass spectrometry. However, in the last decade gel-free nano-flow HPLC-MALDI-MS methods also became relevant with emerging of mass spectrometers capable of tandem MS experiments. In the following table, proteome analysis based on chromatography and MALDI mass spectrometry are summarized in the order they were published divided into human, non-human and quantitative analysis.

All of these studies were realized employing bottom-up proteomics. This means, prior to separation all extracted proteins were enzymatically digested and separation was performed exclusively at peptide level.

Table 11. Proteome analysis employing chromatography and MALDI mass spectrometry (PSD = post source decay, CID = collision induced fragmentation, ICAT = cleavable isotope-coded affinity tag, iTRAQ = isobaric tag for relative and absolute quantitation).

species	material	aim of the work	HPLC-MALDI setup	number of ident. proteins	reference
non-human					
neat	mito-chondrial ribo-somes	protein identification from the large, 39s subunit of bovine mitochondrial ribosomes	discrete spot collection, C ₁₈ precolumn: 1.0 x 0.3 mm i.d. 3 µm particle size C ₁₈ separation column: 150 x 0.18 mm i.d. 3 µm particle size Quadrupole/ TOF (CID)	43	Bodnar et al. J. Am. Soc. Mass Spectrom. 2003 14, 971-9
<i>Saccharomyces cerevisiae</i> (yeast)	strain YP H499	extract maximum amount of information from HPLC-MALDI_MS experiments	discrete spot collection, C ₁₈ precolumn: 1.0 x 0.3 mm i.d. 3 µm particle size C ₁₈ separation column: 150 x 0.075 mm i.d. 3 µm particle size TOF/TOF (CID)	271 (at least two peptides per protein)	Karger et al. Anal. Chem. 2004 76, 6017-28
<i>Escherichia coli</i>	strain HM22	optimization of precursor exclusion list strategy	discrete spot collection, C ₁₈ separation column: 150 x 0.1 mm 5 µm particle size TOF/TOF (CID)	362	Karger et al. Anal. Chem. 2005 77, 7816-25
<i>Corynebacterium glutamicum</i>	whole cell lysate	large-scale proteome analysis	discrete spot collection, monolithic precolumn (PS-DVB-based): 10 x 0.2 mm i.d. Monolithic separation column: 60 x 0.1 mm i.d. TOF/TOF (CID)	1,208 (at least two peptides per protein)	Lasaosa et al. Anal. Bioanal. Chem. 2008 submitted

Table 11 - continued

species	material	aim of the work	HPLC-MALDI setup	number of ident. proteins	reference
human					
human	tears	combination of several proteome analysis techniques	heated droplet interface, C ₁₈ separation column: 150 x 1.0 mm i.d. Quadrupole/ (CID) TOF	44 (HPLC-MALDI-MS)	Li et al. J. Proteome Res. 2005 4, 6, 2052-61
human	multi-potent adult progenitor cells from bone marrow	four stage HPLC selection of methionyl peptides for peptide centric proteome-analysis	discrete spot collection, C ₁₈ precolumn: 5.0 x 0.3 mm i.d. C ₁₈ separation column: 3 µm particle size 10.0 x 0.15 mm i.d. TOF/TOF (PSD)	2,151	Gevaert et al. J. Proteome Res. 2006 5, 1415-28
human	liver tissue	capillary array (18 channels) RP-HPLC MALDI-MS for high throughput proteome analysis	discrete spot collection, precolumn: silica-based monolith 5.0 x 0.32 mm i.d. C ₁₈ separation column: 5 µm particles, 250 and 320 µm i.d. TOF/TOF (CID)	462	Zhang et al. J. Proteome Res. 2006 5, 3186-96
quantitative					
<i>Saccharomyces cerevisiae</i> (yeast)	two strains: HFY1200 and HFY871	quantification strategy employing ICAT	discrete spot collection, C ₁₈ precolumn: 3 µm particle size, 5.0 x 0.3 mm i.d. C ₁₈ separation column: 3 µm particle size 150 x 0.1 mm i.d. TOF/TOF (CID)	quantification of 700 proteins	Parker et al. Mol. Cell. Proteomics 2004 3, 625-59
human	serum and plasma samples	relative quantitation in large-scale proteomics without isotope tagging reagents	discrete spot collection, C ₁₈ precolumn: 3 µm particle size, 5.0 x 0.3 mm i.d. C ₁₈ separation column: 3 µm particle size 150 x 0.1 mm i.d. TOF/TOF (CID)	126	Hattan and Parker Anal. Chem. 2006 78, 7986-96
mouse	hippocampal synapses of brain	optimization of an iTRAQ-based labeling technique for quantification of membrane proteins	discrete spot collection, C ₁₈ precolumn: 1.0 x 0.3 mm i.d. 3 µm particle size C ₁₈ separation column: 150 x 0.1 mm i.d. 3 µm particle size TOF/TOF (CID)	1,122 (at least three peptides per protein)	Li et al. J. Proteome Res. 2007 6, 3127-33

Table 11 – continued

species	material	aim of the work	HPLC-MALDI setup	number of ident. proteins	reference
human	cerebro-spinal liquid	quantitative proteomics with the use of isotopic-labeled synthetic peptides as reference	discrete spot collection, precolumn, C ₁₈ separation column: 3 µm particle size 150 x 0.1 mm i.d. TOF/TOF (CID)	absolute quantitation of 5 peptides from 5 proteins	Zhang et al. J. Proteome Res. 2008 7, 720-30

No comprehensive proteome study for *Glioblastoma multiforme* based on chromatography and mass spectrometry was accomplished so far. The highest number of identified proteins (2,151) was reported by Gevaert and colleagues. However, it is mostly difficult to find out the validation procedure for the identified proteins, especially in older publications.

5.3 Software tools for result generation, validation and interpretation

To cope with the enormous complexity of datasets, computer-based methods for handling them are essentially.

For *in silico* digest and comparison with experimentally gained mass spectrometric data, several search engines are available on-line. The most prominent are listed in Table 12.

Table 12. On-line available search engines for protein identification.

search engine	location
Mascot	http://www.matrixscience.com
Protein Prospector	http://prospector.ucsf.edu/
Profound	http://prowl.rockefeller.edu/prowl-cgi/profound.exe
OMSSA	http://pubchem.ncbi.nlm.nih.gov/omssa/omssacgi.cgi?searchsettings=etd.xml

For obtaining a correct search result the treatment of the protein sample has to be known to set parameters in the identification software such as the employed enzyme for digestion as well as fixed- and variable modifications.

The obtained list of protein hits then has to be validated with suitable software. Subsequently, the confirmed proteins have to be put in biological context employing appropriate software tools again.

5.3.1 Peptide fragmentation

As basis for search engines the peptide fragmentation rules have to be known. Most mass analyzers today are equipped with the opportunity of tandem mass spectrometric (MS/MS) experiments where a selected analyte ion is fragmented after detection of the intact (precursor) ion mass-to-charge ratio. To identify peptides not only with the help of their intact masses, but also determining the primary structure of them, fragmentation is necessary. The intact peptide mass is not a unique property for a peptide. In order to identify peptides confidently on the basis of their amino acid sequence, peptides are subjected to sequence-specific fragmentation reactions. The most commonly applied technique is the so-called collision induced dissociation (CID) or collisionally activated dissociation (CAD), respectively. Hereby, the peptide ions are transferred into a collision chamber where gas molecules such as air, helium, or argon are present. These molecules collide with the peptide ions and induce translational energy conversion to oscillation energy of the peptide ions ^[168]. This process leads to structure specific fragmentation. Especially the peptide bonds (= peptide backbone) along the sequence are susceptible to such fragmentations. Three typical cleavage sites are observed for peptides:

1. cleavage prior to peptide bond: a- and x-ions are created
2. cleavage past peptide bond: c- and z-ions are created
3. cleavage of peptide bond: b- and y-ions are created

By the cleavage of one peptide ion two fragments emerge. If the peptide charge remains at the N-terminus a-, b- and c-ion series are created and for remaining at the C-terminus x-, y- and z-ion series, respectively. This scheme is depicted in Figure 15.

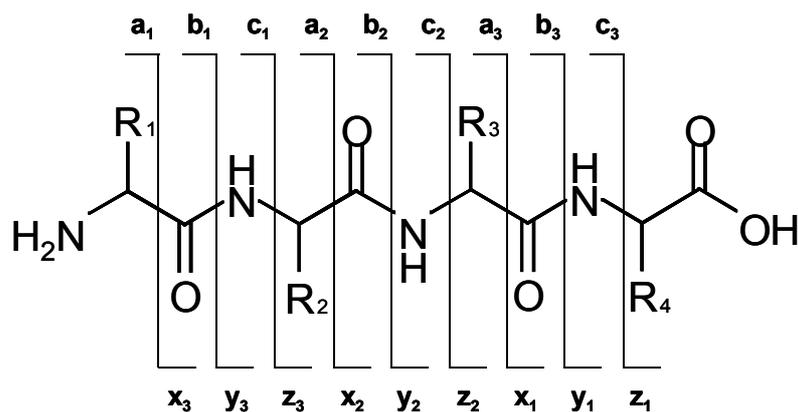


Figure 15. Nomenclature for peptide fragmentation sites introduced by Roepstorff and Fohlman ^[169].

The knowledge of defined fragment patterns in peptide dissociation permits the assignment of the amino acid sequence to the peptide and further the protein. This is exploited in bioinformatic algorithms for peptide- and protein identification in mass spectrometry-based proteomics.

5.3.2 Mascot search engine

The most prominent search engine beside SEQUEST is the Mascot software. Proteins are identified by comparing experimental obtained peptide fragment mass spectra with *in silico* generated ones from protein sequences stored in a database. The so-called MOWSE Score (molecular weight search) is generated by a probability-based implementation of the MOWSE algorithm ^[170] in the Mascot software. Here, the probability for an assigned hit to be a random event is computed. The lower this probability, the better is the identification quality. Because this is quite confusing the scores are reported as $-10 \times \log_{10}(P)$. The higher these scores the lower is the probability of a false assigned hit and as a consequence the higher is the probability of correct peptide and protein identification.

For a small number of peptide queries the protein MOWSE Score consists of the sum of the ion scores of unique peptides achieved for an individual protein. To reduce the weight of low-scoring random hits, a slight correction is applied. This is a function of

the total number of hits for a protein and the width of the set peptide tolerance window^[171]. Usually this correction is very small.

5.3.3 Scaffold Proteome software

For result validation, Scaffold proteome software was developed using statistical algorithms evolved in the Institute for Systems Biology (Seattle, WA, USA). These algorithms are implemented in software known as Peptide- and Protein Prophet. In Scaffold these algorithms are implemented independently.

Mascot offers probability estimation with respect to the database size and not on sample characteristics. As Scaffold involves the sample specific distribution by creating discriminant scores for peptides, it provides a more real estimation of correct hit probabilities^[172].

This software uses Mascot “.dat” files for further processing. The X!tandem search engine can be browsed additionally to Mascot, and Peptide- and Protein Prophet are used for validation. Peptide identifications are accepted if they could be established at higher than 95 % probability as specified by the Peptide Prophet algorithm^[173]. This algorithm first creates a discriminant score for all peptides which is independent of the used search engine. Then Bayesian statistics is applied to calculate that the identification is correct or incorrect. Protein identifications are accepted if they could be established at higher than 99 % probability and contained at least two identified peptides. Protein probabilities are assigned by the Protein Prophet algorithm^[174]. The final protein identifications are less than the Mascot identification hits, because only the minimum list of proteins adequate to explain the peptide imputations using the expectation-maximization algorithm is gained with Scaffold.

The number of peptide mass spectra observed is referred to as spectrum count^[40]. The spectrum count can be used for relative protein quantification of two samples in comparison or for an estimation of protein abundance in one sample.

5.3.4 Gene ontology-based interpretation software

The gene ontology project (GO) ^[175] makes controlled vocabulary for the description of gene and gene product attributes of any organism available. Three ontologies are distinguished, each providing a key concept of molecular biology: molecular function, biological process the genes or gen products are involved in, and their localization to cellular components. To categorize the identified proteins to gene ontologies **GeneTrail** software (<http://genetrail.bioinf.uni-sb.de>) can be employed. The gene products can be active in more than one cellular component as well as in more biological processes and molecular functions. Cellular components describe the locations where the gene products are situated. These can be subcellular structures and macromolecular complexes. Biological processes represent sequences of events or molecular functions of the gene products like metabolic processes, but no pathways. By matching a molecular function the activity of a gene product is meant at molecular level, such as enzyme activities.

PANTHER (protein analysis through evolutionary relationships) ^[176] is a classification system based on gene functions as well as GeneTrail (www.pantherdb.org). This software uses the terms biological process and molecular function similar to the GO, but much simplified to allow high throughput analysis.

6 Brain tumors

6.1 Overview of brain tumors

Tumors arise when oncogenes become activated and tumor suppressor genes inactivated within neoplastic cells. The factors and mechanisms leading to such regulation are rarely understood and purpose of intense research today. Brain tumors are intracranial tumors evolving either in the brain itself (primary) or spread from other tissues (secondary). Affected areas of the brain can be the cranial nerves, the brain envelope, skull, pituitary and pineal gland and brain cells itself, such as neurons, glial cells (astrocytes, oligodendrocytes and epidymal cells), lymphatic tissue and blood vessels. In Germany, every year 436,000 new patients suffer from cancer, and 211,500 are going to die each year ^[177]. In 2006 a percentage of 3.2 % of the male and 3.4 % of the female cancer patients died of brain tumors ^[178]. In the United States, specific cancer statistics are published regularly. In 2008 ^[179], the number of new cancer cases in the United States is estimated to 1.44 million whereas 21,810 (1.5 %) of them are going to be brain tumors; 13,070 patients of them are evaluated to die. The five-year survival rate increased from 24 % to 35 % since 1977 for brain tumors, in general.

Subsequently, the most frequently appearing brain tumors are briefly summarized. To get a glimpse of their location the human brain assembly is depicted in Figure 16.

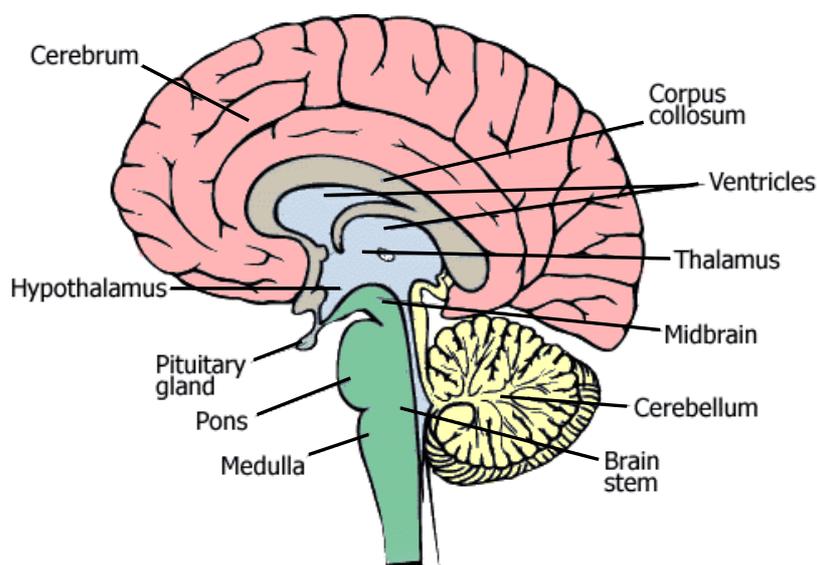


Figure 16. Side view of the human brain ^[180].

6.1.1 Primary brain tumors

The primary brain tumors in principle can appear in the entire brain. Most commonly, they are distinguished due to their origin cells, from which they developed as listed in Table 13. The widespread classification system of grade I-IV was defined by the World Health Organization (WHO) ^[181].

Table 13. Overview of primary brain tumors.

tumor	WHO-grades appearing
<i>astrocytic</i>	I-IV
<i>oligodendroglial</i>	II+III
<i>oligoastrocytic</i>	II+III
<i>epidymal</i>	I-III
<i>choroid plexus</i>	I-III
<i>other neuroepithelial</i>	I+II
<i>neuronal and mixed neuronal-glia</i>	I-III
<i>pineal</i>	I-IV
<i>embryonal</i>	IV
<i>cranial and paraspinal nerves</i>	I-IV
<i>meningeal</i>	I-III
<i>sellar region</i>	I
tumor grade	description
I	least malignant, slowly growing, microscopically almost normal appearance
II	slightly faster growing and abnormal microscopic appearance, may recur as grade III or IV tumor
III	malignant, actively reproducing abnormal cells, invade surrounding normal tissue, recur frequently, often as grade IV
IV	most malignant, invade wide areas of surrounding tissue, reproduce rapidly, necrotic in center

Other distinction criterions often applied are the location in the brain where the tumor appears or the differentiation of benign and malign tumors.

6.1.2 Secondary brain tumors

Malignant tumor cells are able to spread across the entire body with the bloodstream or lymphatic system and form secondary tumors (metastasis). Their incident in brain tumors is higher than for primary ones. The most frequently tumor afflicted origin organs are kidney, lung, skin (melanoma, other skin tumors scarcely metastasize), breast and colon. The cells in the secondary tumor are very similar to those in the origin tumor. Therefore, the cell type can be examined microscopically for identification of the origin tissue. However, in some cases the primary tumor remains unidentified.

6.2 Diagnosis, therapy and prognosis

A lot of unspecific symptoms like headache, seizures, gastrointestinal symptoms and even strokes appear with the occurrence of brain tumors. Depending on the location in the brain, different symptomatic manifestations can be observed as indicated in Table 14.

Table 14. Symptomatic of brain tumors referring to their local occurrence ^[182].

tumor location	symptoms
cerebrum	
<i>frontal lobe</i>	dementia, personality changes, gait disturbances, generalized or focal seizures, expressive aphasia (loss of speech)
<i>parietal lobe</i>	receptive aphasia, sensory loss, hemianopsia (bisection of visual field), spatial disorientation
<i>temporal lobe</i>	complex partial or generalized seizures, quadrantanopsia (loss of one quarter of the visual field), behavioral alterations
<i>occipital lobe</i>	contralateral hemianopsia
others	
<i>thalamus</i>	contralateral sensory loss, behavioral changes, language disorder
<i>cerebellum</i>	ataxia (motion disorder), dysmetria (disorder of haphazard trajectories), nystagmus (irrepressible rhythmic motion of muscles, mostly ocular muscle)
<i>brain stem</i>	cranial nerve dysfunction, ataxia, papillary abnormalities, nystagmus, hemiparesis (partial paralysis of one side of the body), autonomic dysfunction

First parts of the diagnosis are common checks of vision, hearing, balance, coordination and reflexes. Remains the suspicion of brain cancer, one or a combination of the following methods has to be applied:

- **Computerized tomography (CT) scan:** Here, detailed two-dimensional images of the brain are taken using X-rays. A special dye can be injected to visualize the brain tumor. This method usually lasts less than ten minutes.
- **Magnetic resonance imaging (MRI) scan:** Magnetic fields and radio waves are applied for brain image acquisition. A special dye to distinguish tumor and healthy cells can be injected into the bloodstream. This procedure can last between fifteen minutes and an hour.
- **Angiogram:** Here, a special dye is injected into the bloodstream which can be observed with X-ray radiography. Blood vessels in and around the tumor are visualized.
- **X-ray radiography of head and skull:** With this test, bone alterations which can be associated with brain tumors are identified. However, this test is not as sensitive as a brain imaging.
- **Other brain scans:** Tests like single-photon emission computerized tomography (SPECT), magnetic resonance spectroscopy (MRS) and positron emission tomography (PET) can be applied to observe brain activity, such as metabolism, chemistry and blood flow within the brain. Often they are used in combination with MRI for further analysis of the tumor but not for initial diagnosis.

The therapy of brain cancers depends on the age and overall health situation of the patient as well as location, size and grade of the tumor. When the tumor is accessible, surgery is one common treatment. Hereby, the strategy includes the attempt to remove carefully as much as possible of the tumor tissue and as little as necessary of healthy tissue. Some tumors can be removed totally, others only partly. Secondary therapy strategy is high energy radiation to destroy the tumor cells. Last of the three main treatments is chemotherapy where drugs are introduced orally or intravenously into the body, helping to kill cancer cells. Mostly, all of these strategies are adopted in combination. Once vanquished, a brain tumor can recur any time when just a few cells have survived the initial treatment.

Prognosis of survival depends of numerous patient-specific details, such as type of the tumor, extent of surgical tumor removal and age of the patient, for instance ^[183]. An average survival time dependent on patient's age is published by the central brain tumor registry of the United States (CBTRUS) ^[184]. This study is based on data from 1973 to 2004 shown in Table 15. As expected, the chances of five-year survival drop with rising age.

Table 15. Average five-year survival rates of primary brain tumors.

age at diagnosis	percentage of five-year survival [%]
0-19	66.0
20-44	49.2
45-54	24.0
55-64	11.1
65-74	6.7
75 or older	4.7

However, not only the age is important for prognosis. The kind and grade of brain tumor are also pivotal. For type II tumors a surviving over five years is common whereas for grade III tumors usually two to three years are predicted ^[185]. Prognosis of grade IV patients depends on the specific tumor properties. For instance, *cerebellar medulloblastomas* as grade IV tumors are expeditiously fatal if not immediately treated, but with an appropriate radiation and chemotherapy the five-year survival chance can exceed 60 %.

6.3 *Glioblastoma multiforme*

The most aggressive and most common primary brain tumor is *Glioblastoma multiforme* (GBM). This tumor occurs in 52 % of all primary brain tumors and 20 % of total intracranial tumors ^[185]. After its appearance, it rapidly infiltrates other parts of the brain resulting in difficult treatment. In men it is more often diagnosed than in women ^[186] with an average age of 55. This grade IV *astrocytoma* often appears abruptly and sporadic without any genetic disposition. No associations could be verified for *Glioblastoma multiforme* and smoking, diet, cellular phones, electromagnetic fields or viral infection ^[187-192]. The average survival time of these

thuggish tumors is less than 16 months with an estimated five-year survival rate of 5 % ^[193]. An image of a *Glioblastoma multiforme* in a patient's brain is shown in Figure 17.

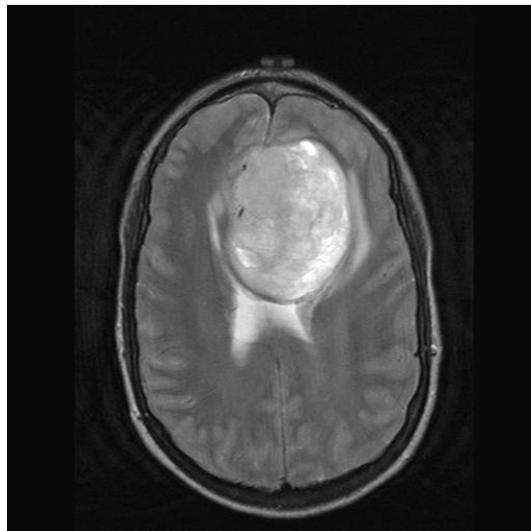


Figure 17. Image of *Glioblastoma multiforme* in a patient's brain.

6.4 Human brain proteomics-where are we standing?

In the last two decades numerous reports of proteomic research with efforts on brain and brain diseases were published. Especially the comparison between the healthy and the diseased brain is challenging because healthy brain tissue samples are very rarely available. Non tumor tissue is often yielded at epilepsy brain surgeries. Furthermore, the investigation of the fetal brain is very promising to get information about the development of the very complex composition and function of the human brain. In 1991, the HUPO Brain Proteome Project (BPP) was launched to achieve a deeper understanding of the healthy and diseased brain. The following selection of brain proteomic studies listed in Table 16 should give a short overview of studies focused on the human brain proteome in healthy and tumor affected tissue. Identification of proteins in all studies was performed with mass spectrometry. All projects involved in the HUPO Brain Proteome Project are marked with BPP. The content of Table 16 is classified into the tissue origin of the tumor or normal tissue. All studies implied exclusively brain or brain tumor tissue, no cell line studies are

considered. The rare application of 2D HPLC-MS is noticeable; almost all studies were accomplished adopting 2D-GE. A method for relative quantification of proteins in the 2D-GE is named DIGE (differential gel electrophoresis). Two polyacrylamide gels with two different sample states (for instance tumor and normal tissue) are generated, and both are treated with a different dye. The proteins expressed only in one stage will appear in one of both dyes clearly whereas proteins occurring in both states emerge as mixture of both dyes. This way the differential expression of proteins can be observed.

Table 16. Overview of the recent brain proteome studies.

analysis method	material	biopsy (b) or autopsy (a)	aims	results	reference
normal fetal brain					
2D-GE	cerebellum of human fetal brain of aborted fetuses in the early second trimester (18.2 ± 2.2 weeks)	a	comprehensive information of early neuronal processes	437 proteins identified, assembly of fetal brain database	Fountoulakis et al. <i>Proteomics</i> 2002, 2, 1547-76
2D-GE	whole brain of human fetal brain of aborted fetuses in the early second trimester (18.2 ± 2.2 weeks)	a	updating the map of fetal brain proteins	identification of 543 proteins, 314 of them never found before in fetal brain	Shin et al. <i>Electrophoresis</i> 2005, 26, 2759-78
children brain tumor tissue					
DIGE	pediatric primitive neuroectodermal tumors (PNET) and ependymoma tissue	b	identification of biomarkers for PNET and ependymomas	biomarker for PNET: stathmin, for ependymomas: annexin A1 and calcyphosine	De Bont et al. <i>J. Neuropathol. Exp. Neurol.</i> 2007, 66, 505-16
normal adult brain					
MUDPIT	enriched nuclear extracts of eight human key tissues: brain, heart, liver, lung, muscle, pancreas, spleen, testis	a+b	access level of proteome diversity in different tissues	most proteins are detected in single tissue	Cagney et al. <i>J. proteome Res.</i> 2005, 4, 1757-67
2D HPLC	a: prostate cancer, no signs of neurodegeneration, b: epilepsy surgery Both: superior temporal gyrus (temporal lobe of cerebrum)	a+b	protein identification using shotgun approach	identification of 209 proteins	Dumont et al. <i>Proteomics</i> 2006, 6, 4967-77 (BPP)

Table 16 - continued

analysis method	material	biopsy (b) or autopsy (a)	aims	results	reference
2D-GE	cerebrum temporal lobe tissue from non tumor patients, b: epilepsy patient a: no neurodegenerative disease	a+b	understand function of central nerve system in normal and pathologic state	267 proteins identified, 6 down and 23 up regulated in autopsy sample	He et al. Proteomics, 2006, 6, 4987-96 (BPP)
2D-GE	white matter of human brain (frontal parietal, temporal, occipital lobe of cerebrum), no neurodegeneration	a	protein expression profile analysis	identification of 64 proteins for initial reference map and database for proteins of white matter	Mucelli et al. J. Chromatogr. B 2006, 833, 80-90
2D-GE MS/ 2D HPLC	brain tissue	a+b	putting results of initial BPP in biological context	identification of 1,804 human brain proteins and their biological impact	Müller et al. or rather Reidegeld et al. Proteomics 2006,6, 5059-75 and 4997-5014, respectively (BPP)
1D GE with prior functional fractionation	temporal lobe of cerebrum fractionation in cyto- and nucleoplasmic, membrane and other structural and DNA-binding proteins	b	identify as much as possible proteins from human brain	1,553 proteins identified	Park et al. Proteomics 2006, 6 4978-86 (BPP)
1D GE micro-HPLC	neuromelanin (NM) granules, no disease referring to brain	a	Isolate intact highly pure NM granules from human substantia nigra for subcellular proteomics	72 proteins identified for insight in NM granule building processes	Tribl et al. Mol. & Cell. Proteomics 2005, 4, 945-57
2D-GE	hippocampus tissue (part of cerebrum), no disease referring to brain	a	protein map for hippocampus proteins	165 proteins identified	Yang et al. Electrophoresis 2004, 25, 1169-74
tumor adult brain					
2D-GE	fibrillary astrocytoma (grade II) and Glioblastoma multiforme (grade IV)	b	comparison of protein pattern between low- and high-grade astrocytomas	few differences reproducibly observable, initial reference map for astrocytomas	Odreman et al. J. Proteome Res. 2005, 4, 698-708

Table 16 - continued

analysis method	material	biopsy (b) or autopsy (a)	aims	results	reference
capillary isoelectr. focusing (CIEF) nano-flow-HPLC	<i>Glioblastoma multiforme</i>	b	comparison CIEF/SCX as first dimension	1,820 proteins identified, CIEF less fraction-to-fraction peptide carry over than SCX (10-15 % and 40-80 %, respectively)	Wang et al. Anal. Chem. 2005, 77, 6549-56
comparison healthy and tumor adult brain					
DIGE	non tumor (epilepsy) and grade II-IV astrocytoma tissue	b	provide information about differentially expressed proteins of healthy and tumor tissue	91 differentially expressed proteins, 20 of them identified for the first time	Khalil Cancer Sci., 98, 201-13
MALDI MS protein profiling	grade II-IV astrocytomas	b	use of protein patterns for classification of astrocytomas and survival diagnosis	distinction of normal and astrocytoma as well as grade-to-grade distinction, short time and long time survival prediction	Schwartz et al. Cancer Res. 2005, 65, 7674-81

In the following work the focus is set on identification of proteins expressed in *Glioblastoma multiforme* tissue with methods based on chromatography and mass spectrometry. An additional aspect is the confirmation of previous indirectly identified potential biomarkers for this type of tumor.

Chapter II

Aim of the thesis

Since the late 80s, the mapping and sequencing of the human genome stood in the focus of the global research interest ^[194]. In 1920 the term genome was established by H. Winkler who generated a portmanteau of the words gene and chromosome. The first encoded DNA-genome was the one of phage phi x 174 by Sanger et al. in 1977 ^[195;196]. In 2004, the human genome was decoded totally resulting in 20,000-25,000 protein encoding genes ^[43]. From this point proteomic research became more and more important. Now the proteins expressed from these genes playing a major role in maintenance of physiological and pathological activity should be investigated. Their structure, function and location in the cell were of major importance as they tell the story of human life. Today, the total amount of human proteins still can only be estimated.

In this thesis the focus is set on proteome analysis based on chromatography and mass spectrometry as revealed in the previous chapter. Beside 2G-GE gel-free methods have become enormous relevant in the last few years. Especially for shotgun proteomics applying MALDI mass spectrometry after the second separation dimension, a rapid development was observable during the last five years. The most straightforward of the numerous nano-flow HPLC-MALDI mass spectrometry interfaces is the position of discrete sample/matrix spots directly onto the MALDI target and off-line transfer of the target into the mass spectrometer ^[156]. Different proteomes were analyzed using this design in the second separation dimension: *Saccharomyces cerevisiae*, *Escherichia coli*, *Corynebacterium glutamicum*, hippocampal synapses of mouse brain, mitochondrial ribosomes (neat) and human liver tissue as well as human multi-potent adult progenitor cells from bone marrow, serum and plasma samples and cerebrospinal liquid ^[138;176;197-203]. No report of a proteome analysis of *Glioblastoma multiforme* employing this setup is known so far.

The aim of this thesis was to develop a semi top-down approach as alternative to the classical bottom-up approach for proteome analysis. Therefore an efficient HPLC method for intact protein separation in the first dimension should be established. To accomplish peptide separation in the second dimension an efficient nano-flow HPLC-MALDI-MS/MS method should be developed. Therefore, polymer-based monolithic capillary columns should be used as separation- and preconcentration columns for high speed, resolution and sensitivity. The new semi top-down- and the standard bottom-up approach should be established with moderate complex standard

samples. Then the optimized methods should be applied to a proteome study of a human brain tumor tissue (*Glioblastoma multiforme*) for comparison with each other. Previously, a set of biomarker proteins was identified indirectly by SEREX for the investigated kind of tumor. Biomarkers are proteins specific for one kind of disease as for *Glioblastoma multiforme* brain tumors, for instance. These should be employed for early disease diagnosis which is of urgent importance especially in the case of cancer. Not only an individual protein can function as such biomarker but rather a combination of proteins up to ≥ 50 proteins for one organism ^[204]. Such a set consisting of 13 potential biomarker proteins was identified for *Glioblastoma multiforme* recently ^[205]. To confirm at least some of these proteins on molecular level with the developed methods, was a further goal of this thesis.

Chapter III

Development and validation of
HPLC-MS approaches for
2D proteome analysis

1 Introduction

Reversed-phase high-performance liquid chromatography (RP-HPLC) is the mostly used chromatographic mode in terms of hyphenation to mass spectrometry. The advantage in comparison to other chromatographic modes is the compatibility of the used eluents with electrospray ionization (ESI) and matrix assisted laser desorption/ionization (MALDI) techniques, permitting on-line coupling to mass spectrometry. For increase of resolution, speed and sensitivity of the separation prior to mass spectrometric analysis two strategies were introduced. In ultra-pressure liquid chromatography (UPLC) small C₁₈ particles (sub 2 μm) are applied for column packing, and the columns are operated at very high pressures (about 10,000 psi). Hence, separation efficiency is maintained while analysis time is reduced [206;207]. Another approach employs capillary separation columns with monolithic stationary phases on silica or polymer (mostly PS-DVB) basis instead of microparticles [208-210]. Both methods have been shown to be very suitable for separation of highly complex peptide mixtures due to their high speed and resolution.

After separation, peptides are fragmented by mass spectrometry and computer-based identified. On-line nano-flow HPLC-ESI-MS is a very efficient and fully automated combination for separating and identifying peptides sensitively from a highly complex mixture as required for proteome analysis [211;212]. However, in on-line nano-flow HPLC-ESI-MS only in the short peptide retention window of typically 10-30 s the precursor ion is available for fragmentation [213]. The introduction of MALDI mass spectrometers with MS/MS capability [214-216] stimulated the interest in nano-flow HPLC-MALDI-MS interfaces for proteome analysis. In nano-flow HPLC-MALDI-MS the sample is deposited and stored on a target prior to MS and MS/MS analysis. Thus, in contrast to real-time nano-flow HPLC-ESI-MS off-line nano-flow HPLC-MALDI-MS permits a more detailed precursor selection [217] with no time restriction by completing the MS data set before starting MS/MS data generation. Furthermore, the whole sample analysis can be repeated for several times until the sample is consumed. Especially when only small sample amounts are available this approach becomes very advantageous.

In this context the development of an alternative approach to the classical bottom-up method in proteome analysis should be developed. No defined term exists for such a

method including intact protein separation in the first dimension, subsequent digestion of the proteins in the individual fractions and separation of peptides in the second dimension. We called this method “semi top-down” approach because it stands between top-down and bottom-up proteome analysis. Chromatographic semi top-down approaches were introduced in proteome analysis to circumvent the distribution of peptides over several fractions collected in the first chromatographic dimension. The development of the first dimension separation method of intact proteins should be presented in this chapter.

Furthermore an efficient nano-flow IP-RP-HPLC-MALDI-MS method should be presented for second dimension separations of proteome analysis, employing the newest generation of MALDI-TOF/TOF instruments. Thus, a high sensitivity, efficient precursor selection and fragmentation should be achieved. The application of polymer-based monolithic separation columns should assure high resolution, speed, sensitivity and repeatability of the separation prior to mass spectrometry.

2 Experimental

2.1 Chemicals and Materials

The water used for all experiments was prepared with a Purelab Ultra system (Elga, Siershahn, Switzerland). The following materials were purchased from Sigma-Aldrich Steinheim, Germany): acetonitrile (E Chromasolv), dithiothreitol (min. 99 %), 2-mercaptoethanol (> 98 %), benzoylated dialysis tubing (MWCO 2,000 Da), cytochrome C (type V-A, bovine heart), carbonic anhydrase (bovine), serum albumin (bovine, \geq 97 %), β -lactoglobulin A (bovine milk), catalase (bovine liver), lysozyme (chicken egg, 95 %), myoglobin (horse, 95-100 %), ribonuclease A (bovine pancreas, 85 %), transferrin (bovine, 98 %), α -lactalbumin (type I, bovine milk, 85 %), α -cyano-4-hydroxycinnamic acid (4-HCCA), \geq 98 %) and human (Glu¹)-fibrinopeptide B (glu¹-fib). Urea (\geq 99.5 %), ammonium hydrogen carbonate (\geq 99.5 %), iodoacetic acid (\geq 99.5 %), formic acid (FA, p.a.) trifluoroacetic acid (TFA, \geq 99.5 %), heptafluorobutyric acid (HFBA, \geq 99.5 %) and ammonium formate ((NH₄)OAc, \geq 97 %) were obtained from Fluka (Buchs, Switzerland) and sodium dihydrogen phosphate (p.a.) from Merck (Darmstadt, Germany). Modified trypsin for protein digestion was ordered from Promega (Madison, WI, USA). The capillary tubing utilized for device connections was purchased from Polymicro Technologies (Phoenix, AZ; USA) and the PEEK capillary tubing, tubing sleeves, microtight unions and microtight fittings from Upchurch Scientific (Oak Harbor, WA, USA). The thermo mixer was purchased from Eppendorf AG (Comfort, Hamburg, Germany) as well as the vacuum concentrator (Concentrator 5301). The utilized centrifuge was procured from Heraeus (Biofuge13, Heidelberg, Germany).

2.2 Tryptic digest of the proteins

To establish an nano-flow IP-RP-HPLC-MALDI-TOF/TOF platform for proteome analysis a ten-protein mixture digest prepared by C. Schley and H.J. Toll (Saarland University, Saarbruecken, Germany) was employed. Protein concentrations in the mixture for digest, 1D and 2D analysis are listed in Table 17.

Table 17. Ten-protein standard mixture for digestion.

protein	abbreviation	no dilution	dilution 1:9	dilution 1:4
		concentration for digest [pmol μL^{-1}]	concentration for 1D analysis [pmol μL^{-1}]	concentration for 2D analysis [pmol 50 μL^{-1}]
cytochrome C	CYC_HORSE	16.18	1.62	161.80
carbonic anhydrase 2	CAH2_BOVIN	6.89	0.69	68.90
β -lactoglobulin A*	LACB_BOVIN	10.89	1.09	108.90
catalase	CATA_BOVIN	3.47	0.35	34.70
lysozyme C	LYSC_CHICK	13.98	1.40	139.80
myoglobin	MYG_HORSE	11.80	1.18	118.00
ribonuclease A	RNAS1_BOVIN	14.62	1.46	146.20
transferrin	TRFE_BOVIN	2.51	0.25	25.10
α -lactalbumin	LALBA_BOVIN	14.10	1.41	141.00
bovine serum albumin	BSA_BOVIN	7.51	0.75	75.10

* In the Swiss Prot database only β -lactoglobulin B is listed. The variants A and B differ only by two amino acids at position 64: glycine is substituted by aspartic acid in variant A and in position 118: instead of alanine valine is incorporated in the protein sequence ^[218;219]. None of both peptides including these positions was identified in this study.

Proteins were dissolved in a mixture of 8 M urea and 0.5 M ammonium hydrogen carbonate and denatured for 30 min with gentle shaking at a temperature of 37 °C. Disulfide bonds were broken by adding 50 μL of 0.3 M aqueous dithiothreitol solution and incubating at 37 °C for 4 h in the thermo mixer. Subsequently the cysteine sulfhydryl groups were carboxymethylated with 20 μL of 2 M iodoacetic acid for 5 min at room temperature. The excess of alkylation reagent was removed by adding 40 μL of 0.1 M 2-mercaptoethanol and incubated for 15 min. Dialysis of the treated protein mixture followed as final step prior to digestion (12 h, 1 L distilled water). In the thermo mixer 30 μg trypsin were activated for 30 min at 37 °C with 30 μL acetic acid (50 mM) and were then added to the modified protein mixture. Sequencing grade modified trypsin (modified by reductive methylation) was used to minimize auto proteolysis ^[220]. After incubation over night the reaction was quenched by addition of 1 vol. % TFA. Prior to injection of 1 μL the peptide mixture was diluted 1:9 (10 μL digest plus 90 μL eluent A) in 0.05 % aqueous TFA solution for 1D analysis. For injection of 50 μL into the first dimension of 2D analysis the peptide mixture was

diluted 1:4 (20 μL digest plus 80 μL eluent A) in sodium dihydrogen phosphate (5 mM, pH 3) or ammonium formate (10 mM at pH 3), respectively.

2.3 SCX-HPLC for peptide separation

Two different buffer systems were compared for SCX-HPLC of peptides. Sodium dihydrogen phosphate (5 mM, pH 3) as eluent A and sodium dihydrogen phosphate (5 mM, pH 3) plus sodium chloride (500 mM, pH 3) as eluent B were utilized as buffer system (1), both eluents contained 20 % acetonitrile. Buffer system (2) consisted of ammonium formate (10 mM, pH 3 for eluent A and 500 mM, pH 6.8 for eluent B, 25 % acetonitrile for A and B). To both eluent solutions, acetonitrile was added to suppress unspecific secondary solvophobic interactions with the separation column. An analytical HPLC system (Model 1100, HP, Waldbronn, Germany) was used to perform peptide separation at room temperature. It was equipped with an external six-port injection valve (Model 7125, Rheodyne, Rohnert Park, CA, USA) endowed with a 100- μL sample loop (Rheodyne). The sample was diluted 1:4 with eluent A (20 μL digest and 80 μL eluent A) prior to injection of 50 μL of the protein digest (= 1.02 nmol corresponding to 22.99 μg). The gradient started with 0-100 % B in 10 min and was then held for 10 min at 100 % B. The polymeric Polysulfoethyl A 200 x 2.1 mm i.d. column (PolyLC, 5 μm , 200 \AA , Columbia, MD, USA) was operated at a flow rate of 200 $\mu\text{L min}^{-1}$. UV chromatograms were acquired at 214 nm. Two-min fractions were collected except for fraction 1 (1.5 min), resulting in 10 fractions obtained in 19.5 min. The fractions were frozen at a temperature of -30 $^{\circ}\text{C}$. Prior to injection into the second separation dimension the eluents remaining from the SCX separation were evaporated until approximately 10 μL remained in the vial. For analysis the fractions were completed with 90 μL of 0.1 % aqueous HFBA (= solvent of loading pump). To fraction 01 only 65 μL were added to a final volume of 75 μL corresponding to 25 % less original fraction volume than fractions 02-10 (only 1.5 min collected).

2.4 IP-RP-HPLC for separation of intact proteins

An analytical HPLC system (Model 1050, HP), equipped with an external six-port switching valve (Model 7125, Rheodyne) with a 5- μ l sample loop (Rheodyne) was employed for protein separation. Furthermore an in-house made column oven was used to maintain a temperature of 50 °C. UV chromatograms were recorded at 280 nm with an UV/VIS detector (Model 433, Kontron, Eching, Germany). A linear gradient of 15-50 % acetonitrile containing FA, TFA or HFBA as mobile phase additives in the eluent A and B was used on one or two 50 x 4.6 mm i.d. ProSwift-RP-1S PS-DVB based monolithic columns (Dionex, Sunnyvale, CA, USA). When using two columns, they were connected in series. The gradient time for separation of ten standard proteins was set to 10 min and 20 min for one column and 20 min for two columns, respectively.

2.5 Nano-flow IP-RP-HPLC separation of peptides

Ten microliters of the fractions collected and concentrated after SCX-separation were injected into the ion-pair reversed-phase nano-flow HPLC system (= second dimension) in partial-loop injection mode. In 1D peptide separation without prefractionation, 1 μ L (= 10.20 pmol corresponding to 229.92 ng) was injected in full-loop injection mode. The setup consisted of a capillary HPLC system (Ultimate, LC Packings, Amsterdam, The Netherlands), an automatic injection unit (Famos, LC-Packings) and a loading pump (Model K-1001, Knauer, Berlin, Germany) with a 10-port switching valve (Model C2-1000D (stator) and 06A-8029C (rotor), VICI, Schenkon, Switzerland). The detector was equipped with a 3 nL Z-shaped flow cell (Model Ultimate), and the UV chromatogram was recorded at 214 nm. A short precolumn (10 x 0.2 mm i.d.) was used to desalt and concentrate the samples obtained from the first dimension. The separation column (60 x 0.1 mm i.d.), as well as the precolumn, consisted of a monolithic PS-DVB based stationary phase material ^[221] (available from LC Packings, Dionex Corporation, Sunnyvale, USA). When a precolumn was installed, peptides were isocratically concentrated and desalted for 3 min. The flow rate of the loading pump delivering 0.1 % aqueous HFBA was set to 10 μ L min⁻¹. After valve switching, peptides were eluted in back-flush mode onto the

separation column. Then, a 40-min gradient of 0-40 % acetonitrile in 0.05 % aqueous TFA was applied at a flow rate of $0.7 \mu\text{L min}^{-1}$ and a temperature of $55 \text{ }^{\circ}\text{C}$. The eluting analytes were spotted onto an Opti-TOF 123 x 81 mm stainless steel target (ABI, Framingham, MA, USA) with an automatic fractionation unit (Probot, Dionex Corporation, Sunnyvale, CA, USA). Alpha-cyano-4-hydroxycinnamic acid in 70 % acetonitrile and 0.1 % aqueous TFA was utilized as matrix with human glu¹-fibrino-peptide B (= glu¹-fib, sequence given in Table 18) in a concentration of $15 \text{ fmol } \mu\text{L}^{-1}$. The matrix flow was adjusted to $2.9 \mu\text{L min}^{-1}$ for a spotting time of 5 s per spot (242 nL matrix per spot) or to $1.2 \mu\text{L min}^{-1}$ for a spotting time of 10 s per spot (200 nL matrix per spot), respectively. Altogether, 45 min of the HPLC run were spotted (min 5 to 50). Spectra for peptide and protein identification were acquired using a MALDI-TOF/TOF mass spectrometer (4800 TOF/TOF Analyzer, ABI, Framingham, MA, USA).

2.6 MALDI-TOF/TOF data acquisition

The spectra were generated in positive reflector mode within a mass range of 800-4,000 m/z . In MS mode 25 laser shots were summarized to one sub-spectrum, and 20 sub-spectra were accumulated to the final spectrum (500 shots) with a frequency of 200 Hz and a laser wavelength of 355 nm (Nd/YAG laser). Prior to analysis, the mass spectra were calibrated using a six-peptide calibration mix (ABI) as revealed in Table 18, on eight external calibration spots for each MALDI plate. The manufacturer recommended procedure (Plate Model Calibration) was applied resulting in a mass accuracy of 50 ppm (= default calibration). For optimized mass accuracy (5 ppm) an internal calibration on the mass-to-charge ratio of glu¹-fib (m/z 1570.677) was performed during data acquisition. When this additional calibration failed due to signal suppression especially in spectra with high sample peptide signal intensity, default calibration was automatically used (50 ppm). The mass-to-charge ratio of glu¹-fib was set on the precursor selection exclusion list to circumvent fragmentation of it. Calibration of the instrument for MS/MS experiments was performed using four fragments of glu¹-fib. In one sub-spectrum 40 laser shots were accumulated, and 25 sub-spectra were summarized to the final spectrum (1,000 shots in total).

Table 18. Six-peptide standard for calibration of the MALDI-mass spectrometer.

peptide name	concentration on target [pmol]	sequence	monoisotopic mass (m/z , $z = 1$)
des-Arg ¹ -bradykinin	1.02	PPGFSPFR	904.4681
angiotensin I	2.01	DRVYIHPFHL	1296.6853
glu ¹ -fibrinopeptide B	1.30	EGVNDNEEGFFSAR	1570.6774
ACTH* (1-17 clip)	2.01	SYSMEHFRWGKPVGKKR	2093.0867
ACTH* (18-39 clip)	1.51	RPVKVYPNGAEDESAEAFPLEF	2465.1989
ACTH* (7-38 clip)	3.01	FRWGKPVGKKRRPVKYPNGAEDESAEAFPLE	3657.9294

* = adrenocorticotrophic hormone, sequence given in Estivariz et al. ^[222]. Remaining peptide sequences supplied by ABI.

For peptide fragmentation, precursors were selected in the mass range of m/z 900-3,000. Fragmentation of five precursors per spot at maximum was realized without using the collision induced dissociation (CID) chamber as a metastable decay process ^[223]. Here peptide fragmentation is realized by post source decay (PSD) and subsequent acceleration without additional collision of the analytes with gas molecules. The CID fragmentation of peptides was tested as well. Air as collision gas for peptide fragmentation was used, applying a pressure of 1.7×10^{-9} bar in the source chamber 2 (= collision chamber). In MS/MS mode 25 sub-spectra containing 40 laser shots each, were summarized to finally 1,000 shots for peptide fragment detection (same conditions for CID and PSD). The precursor selection window was set to 280, which corresponds to a 3.6-Da-window for a precursor with the m/z of 1,000, for instance.

2.7 Data processing

The recorded MS/MS spectra were smoothed with Savitsky-Golay algorithm using three points across the peak and a polynomial order of four. For exporting data to create Mascot generic files (".mgf") the following settings were used in MS/MS peak filter: mass range: 60 m/z to precursor mass -35 Da; peak density: maximum 20 peaks per 200 Da; minimum S/N: 10; minimum area: 200 and maximum 65 peaks/precursor. To identify proteins, the ".mgf"s were sent to Mascot software (version 2.2.03, Matrix Science, London, UK) which uses the MOWSE (molecular weight search) algorithm to

create scores based on identification quality. The following settings were applied: database: Swiss Prot (Version 54.7, January 15th, 2008); taxonomy: chordata (vertebrates and relatives; 70,931 sequences); enzyme: trypsin; variable modification: methionine oxidation + cysteine carboxymethylation; peptide tolerance: 100 ppm; MS/MS tolerance: 0.5 Da; maximal missed cleavages: 1 and ion-score cut off for peptides: 0.05 (95 %). This means that only peptide hits identified with a probability to be a correct hit ≥ 95 % remain in the identification list. The probability of an identification as a random event for proteins was set to 0.05, too. Thus, only proteins with ≥ 95 % probability for correct identification were considered. When the parameter settings differ from the listed above it is mentioned in the text.

3 Results and Discussion

3.1 IP-RP-HPLC for intact proteins as first dimension in 2D analysis

3.1.1 Influence of column length on separation performance

To establish an HPLC method for intact protein separation as first dimension of semi top-down proteome analysis a ten-protein standard was used. The proteins were obtained from different organisms: chicken (lysozyme C), neat (bovine serum albumin, transferrin, catalase, β -lactoglobulin, cytochrome C, carbonic anhydrase 2, α -lactalbumin and ribonuclease A) and horse (myoglobin). For protein separation a PS-DVB-based monolithic column was utilized (ProSwift-RP-1S monolith, 50 x 4.6 mm i.d., Dionex, Sunnyvale, CA, USA). The gradient was chosen to start at 15 % acetonitrile in 0.05 % aqueous TFA and not as generally handled for proteins at 20 %. This was decided because later for the proteome analysis it was expected that not only intact, but also fragments of proteins are present in the tissue protein sample.

To achieve higher resolution, the column length was doubled by coupling two columns in series. Thus, the gradient time was doubled too, according to the gradient volume concept (explained in chapter I, paragraph 3.1.3). The resulting UV chromatograms are depicted in Figure 18. A clear increase in resolution is observable for coupling two columns and applying a gradient time of 20 min. For one column using a gradient time of 10 min (Figure 18a), three unresolved peaks are observed, whereas the third contains even three proteins. By increasing the gradient time by a factor of two (Figure 18b), only two unresolved peaks with each containing two proteins remain. By increasing the column length by a factor of two (Figure 18c), only one peak containing two compounds of the ten-protein standard (catalase and carbonic anhydrase 2) is left due to better resolution.

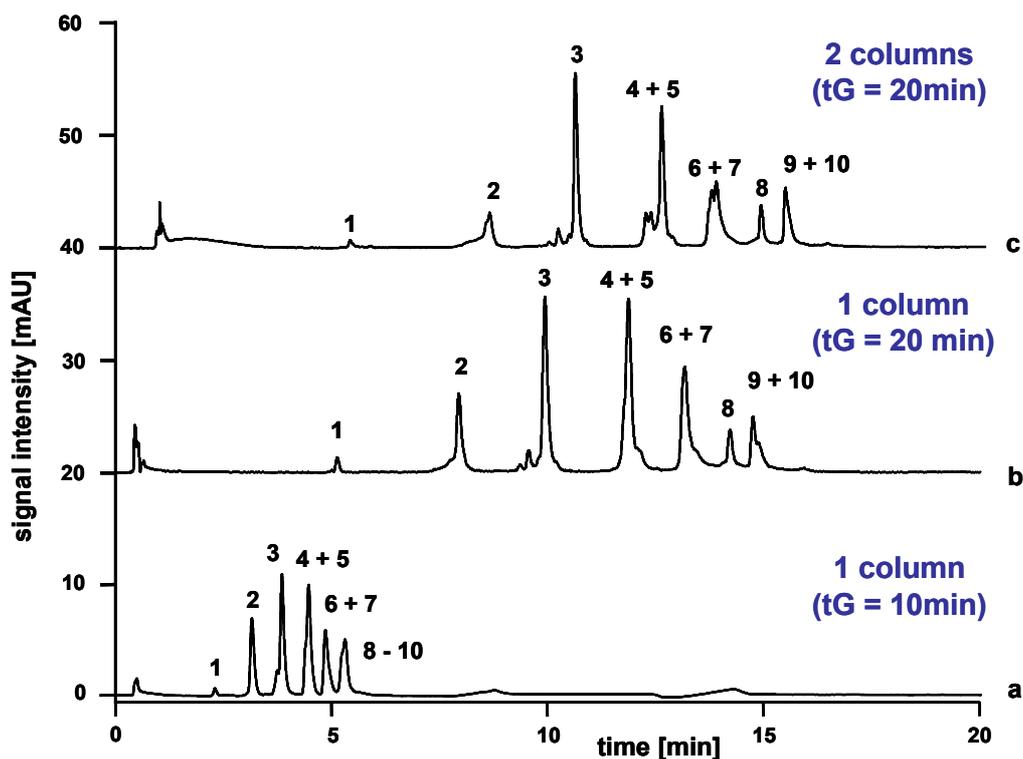


Figure 18. Separation of ten intact proteins on monolithic column(s). Gradient, 15-50 % acetonitrile + 0.05 % aqueous TFA; flow rate, 0.8 mL min^{-1} ; temperature, $55 \text{ }^{\circ}\text{C}$; injection volume, $5 \text{ }\mu\text{L}$; c_{protein} , $0.2 \text{ }\mu\text{g }\mu\text{L}^{-1}$; c_{BSA} , $0.5 \text{ }\mu\text{g }\mu\text{L}^{-1}$; wavelength, 280 nm . (a) One column; $t_{\text{G}} = 10 \text{ min}$. (b) One column; $t_{\text{G}} = 20 \text{ min}$. (c) Two columns, $t_{\text{G}} = 20 \text{ min}$.

The corresponding peak width at half height (w_{h}) for every protein peak in Figure 18 is given in Table 19. Except for catalase and carbonic anhydrase 2, all protein peaks were at least partly separated using two columns and a gradient time of 20 min. Furthermore, w_{h} for the proteins remains almost constant, being 0.13 min for one column with a gradient time of 10 min, 0.11 min for one column applying a gradient time of 20 min and 0.12 min for two columns with a gradient time of 20 min. To compute average w_{h} only ribonuclease, cytochrome C and lysozyme C were included as they were base-line separated in all three cases.

Table 19. Peak width at half height (w_h) of the ten proteins.

protein	peak number	one column		two columns
		(a) w_h [min] ($t_G = 10$ min)	(b) w_h [min] ($t_G = 20$ min)	(c) w_h [min] ($t_G = 20$ min)
RNAS1	1	0.11	0.09	0.08
CYC	2	0.15	0.12	0.18
LYSC	3	0.14	0.13	0.09
TRFE	4	0.17	0.14	0.23
LALBA	5			0.10
BSA + MYG	6 + 7	0.17	0.17	0.28
LACB	8	0.25	0.11	0.09
CATA + CAH2	9 + 10		0.12	0.13

Although the gradient time was doubled the average w_h remained almost the same. For a 60-min gradient of 15-60 % acetonitrile in 0.05 % aqueous TFA (as applied for proteome analysis in the next chapter) an average w_h for the three proteins of 0.25 min was obtained. The changing peak capacities (P) (Chapter I, paragraph 3.1.2) for different gradient times and the two column lengths are given in Table 20. Gradient times for one and two columns were chosen considering the gradient volume concept.

Table 20. Peak capacities (P) for two column lengths and different gradient times.

one column (50 mm)			two columns (100 mm)		
gradient time [min]	average w_b [min]	peak capacity	gradient time [min]	average w_b [min]	peak capacity
7.5	0.19	40	15	0.23	66
15	0.24	62	30	0.31	80
30	0.37	82	60	0.49	124

It can be observed from Table 20 that by increasing the gradient time of a factor F the peak capacity rises about the square root of F. By increasing the gradient time from 15 to 60 min, for instance, the peak capacity increases about two-fold.

With doubling the column length the peak capacity was increased about one third. Peak capacity of 249 was obtained applying a 50-min gradient for peptide separation on a 150 mm C_{18} RP column and 113 for peptides separated with a 50 mm SCX column and a gradient time of 80 min, reported by Gilar et al. ^[224]. However, proteins generate broader peaks due to their more complex structure. As a result w_b at 4σ

(13.4 % of peak height) is increased and thus, the peak capacity is decreased for proteins in comparison to peptides.

In two-dimensional proteome analysis based on chromatography peak capacity of the whole chromatographic system is determined by the number of fractions collected in the first dimension. It is not possible to collect as much fractions in the first dimension as peak capacity is obtained due to time and sample concentration restrictions. Nevertheless an efficient separation in the first dimension is required to circumvent distribution of the collected proteins or peptides over plenty of fractions.

3.1.2 Mobile phase additives

The selectivity of a separation system can be influenced by temperature and choice of stationary and mobile phases. After optimizing the efficiency of separation, the selectivity should be considered. This was realized by testing different mobile phase additives. FA was used only to adjust a low pH for protein denaturation and HFBA as well as TFA additionally as ion-pair modifiers. These are amphiphilic compounds which are added to the mobile phase and adsorb on the surface of the stationary phase with their hydrophobic moiety. The hydrophilic residue (in this case negatively charged) initiates electrostatic interactions to the positively charged analytes ^[225]. Figure 19 reveals the UV chromatograms of separations applying two different ion-pair modifiers (TFA and HFBA) in comparison to addition of a small polar acid (FA) in similar concentration to the mobile phase: HFBA (410 mg), TFA (375 mg) and FA (305 mg) which were added as 0.05 vol. % in 1L mobile phase.

The best selectivity was achieved using TFA as ion-pair modifier. Obviously only with TFA even the critical peak pair bovine serum albumin and myoglobin (peaks 6+7) was at least partly separated.

Originally, it was assumed that the proteins separated with FA as mobile phase additive would elute with the lowest retention time because FA has no hydrophobic moiety and is added in the lowest concentration. By proving the elution window it could be observed that separations performed with FA actually generated the shortest retention window; 6.3 min in comparison to 10.0 min for TFA. The proteins which were separated with HFBA as ion-pair modifier offered the longest retention times as expected.

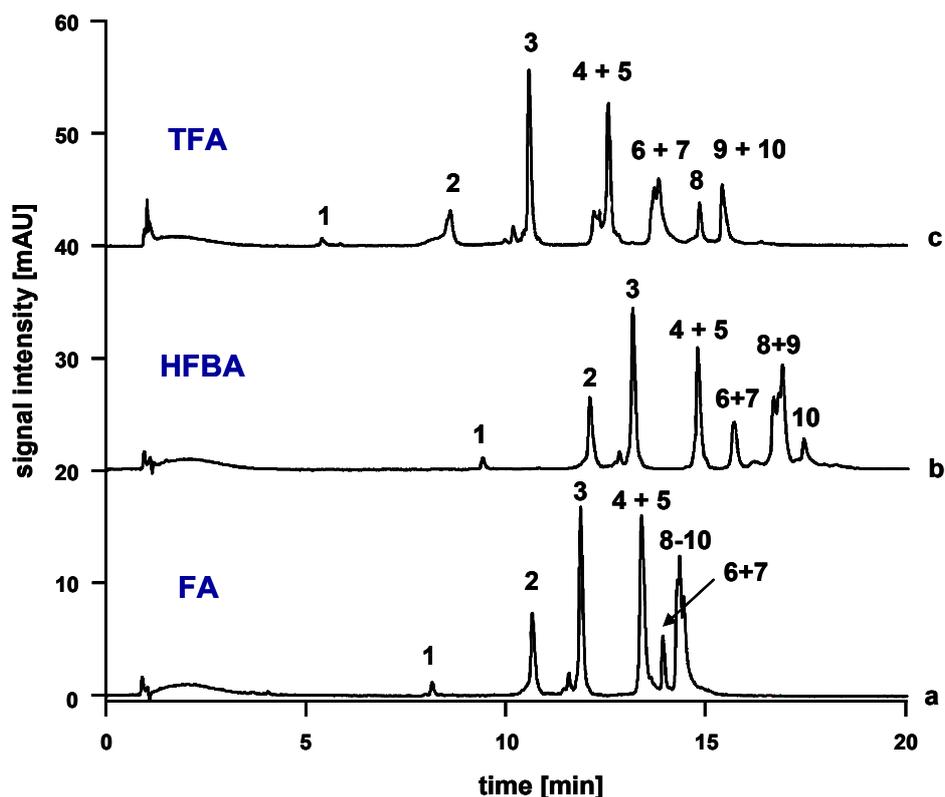


Figure 19. Separation of the ten-protein standard with different mobile phase additives. Gradient, 15-50 % acetonitrile + 0.05 % ion-pair modifier or FA; flow rate, 0.8 mL min^{-1} ; temperature, $55 \text{ }^\circ\text{C}$; injection volume, $5 \text{ }\mu\text{L}$; c_{protein} , $0.2 \text{ }\mu\text{g }\mu\text{L}^{-1}$; c_{BSA} , $0.5 \text{ }\mu\text{g }\mu\text{L}^{-1}$; wavelength, 280 nm ; (a) 0.05 % FA, (b) 0.05 % HFBA and (c) 0.05 % TFA.

Peak widths at half height for the protein peaks are listed in Table 21. The obtained average peak width at half height for proteins separated is the lowest with FA added to the mobile phase (0.12 min). For separation of intact proteins with HFBA and TFA as ion-pair modifiers average peak widths at half maximum were almost the same (0.14 min and 0.15 min , respectively). As the attained selectivity for TFA is the highest, this ion-pair modifier was used for the following separations.

Table 21. Peak width at half height for the ten-protein mix separated with different mobile phase additives.

protein	peak number	(a) w_h [min] 0.05 % FA	(b) w_h [min] 0.05 % HFBA	(c) w_h [min] 0.05 % TFA
RNAS1	1	0.08	0.10	0.08
CYC	2	0.10	0.11	0.18
LYSC	3	0.10	0.10	0.09
TRFE	4	0.13	0.10	0.23
LALBA	5	0.08	0.14	0.10
BSA + MYG	6 + 7	0.23	0.32	0.28
LACB	8			0.09
CATA + CAH2	9 + 10		0.11	0.13

These three mobile phase additives were chosen because they are compatible with MALDI as ionization technique for mass spectrometry due to their high volatility. As TFA provided the best selectivity it was applied as ion-pair modifier to the proteome analysis of *Glioblastoma multiforme* in the next chapter.

3.1.3 Repeatability of IP-RP-HPLC intact protein separations

Another very important parameter to discuss is the repeatability of the established IP-RP-HPLC method for protein separation. In Figure 20 the UV chromatograms of three consecutive separations are displayed. The previously optimized column length (100 mm) and ion-pair modifier (0.05 % TFA) were applied as well as 20 min gradient time. The high repeatability of retention times demonstrated in Figure 20 is proven by computing standard deviations for retention times of the protein peaks in these three runs shown in Table 22.

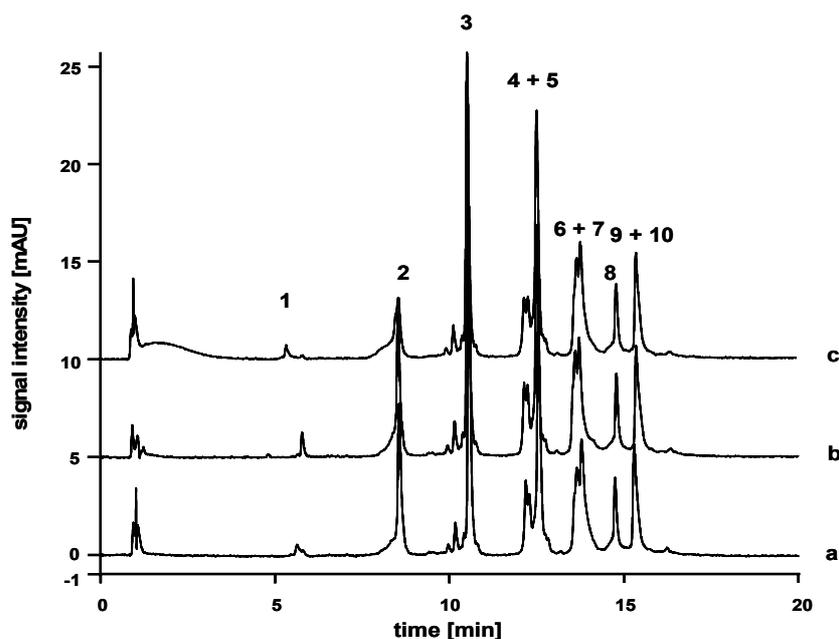


Figure 20. Separation repeatability of the ten-protein mixture.

Gradient, 15-50 % acetonitrile + 0.05 % aqueous TFA ; flow rate, 0.8 mL min⁻¹; temperature, 55 °C; injection volume, 5 µL; c_{protein} , 0.2 µg µL⁻¹; c_{BSA} , 0.5 µg µL⁻¹; wavelength 280 nm.

Table 22. Repeatability of retention time for the ten proteins of the standard mixture.

protein	peak number	average retention time [min]	standard deviation [%]
RNAS1	1	5.6	3.7
CYC	2	8.5	0.4
LYSC	3	10.5	0.1
TRFE	4	12.2	0.7
LALBA	5	12.5	0.2
BSA	6	13.6	0.3
MYG	7	13.8	0.3
LACB	8	14.8	0.3
CATA + CAH2	9 + 10	15.3	0.4

The standard deviation was below 0.8 % for all peaks except ribonuclease A. For this protein a higher variation is observable. It is quite difficult to explain why only ribonuclease A elutes so inconstantly. It might be this way because it elutes as first of the proteins (at 24.8 % acetonitrile + 0.05 % aqueous TFA) and there is a similar variation of the hold-up time of the column, too (3.5 %), caused by the manual injection.

3.2 Validation of a nano-flow HPLC MALDI-TOF/TOF system as second dimension in 2D analysis

3.2.1 Configuration of the separation and identification system

The setup for the off-line HPLC MALDI-TOF/TOF system is depicted in Figure 21. After injection (1) the peptides were flushed onto the trap column (5) when installed. Subsequently to 3-min accumulation time the eluents mixed by the gradient pump transferred the peptides onto the analytical column (6). Here, the peptides were separated using a 40-min gradient (0-40 %) of acetonitrile in 0.05 % aqueous TFA. After UV detection the eluting peptides were mixed with matrix (α -cyano-4-hydroxycinnamic acid) in a T-piece and spotted onto a stainless steel target. Subsequently to co-crystallization of matrix and peptides, the target was manually transferred into the MALDI-TOF/TOF mass spectrometer (Figure 21b). Here the peptides were ionized in the ion source (8) by MALDI, passed the reflectron (12) and reached the microchannel plate detector (13) in a first analysis. Then the precursors were selected and in a second analysis fragmented either with or without using the CID chamber (11). Protein identification was performed by database search against the Swiss Prot database using the Mascot search engine.

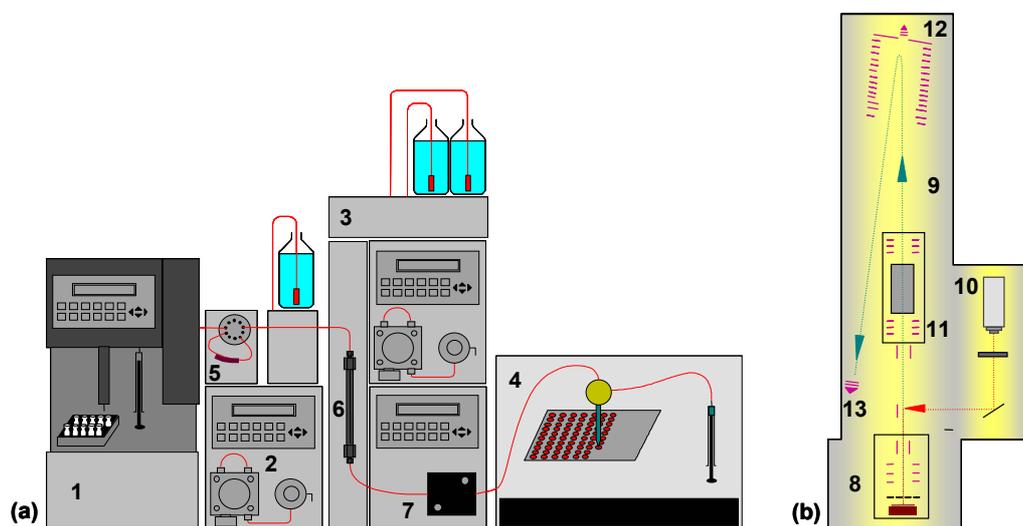


Figure 21. Nano-flow IP-RP-HPLC-MALDI-TOF/TOF setup. (a): 1 = automatic injection unit, 2 = loading pump, 3 = gradient pump, 4 = spotter, 5 = trap column, 6 = analytical column, 7 = UV detector, (b): MALDI-TOF/TOF: 8 = sample loading chamber/ion source 1, 9 = flight tube, 10 = Nd/YAG laser, 10 = CID chamber/ion source 2, 11 = reflectron, 12 = microchannel plate detector

3.2.2 Variation of spotting time

For all validation steps, the same ten-protein standard mixture as for establishing of the HPLC method for intact proteins was used, this time digested. The protein concentration before digest was adjusted to 0.3-1.6 pmol μL^{-1} (inj. volume: 1 μL). Trypsin digestion of this mixture resulted in 395 different peptides in the set mass range of m/z 900-3,000, considering one missed cleavage. Peptide masses including the selected modifications carboxymethylation of sulfhydryl groups in cysteine moieties and methionine oxidation were not counted as individual peptide masses.

In the literature the spotting time varies from 5 s ^[226] to 30 s ^[227] per spot. It was decided to test 5 and 10 s because the ten-protein standard digest is a moderately complex sample in comparison to sample complexity in proteome analysis. For this test, a loading pump and a trap column were not installed. This was possible because no desalting step was necessary prior to analysis, since salts and excessive reagents were removed by dialysis prior to digestion. The peptides of the ten-protein standard mixture digest were separated applying a 40-min gradient of 0-40 % acetonitrile in 0.05 % aqueous TFA. As shown in Figure 22 for both spotting times, 5 s and 10 s respectively, 45 min (min 5 to 50) were spotted of which 35 min (min 5 to 40) were analyzed with the MALDI-TOF/TOF mass spectrometer.

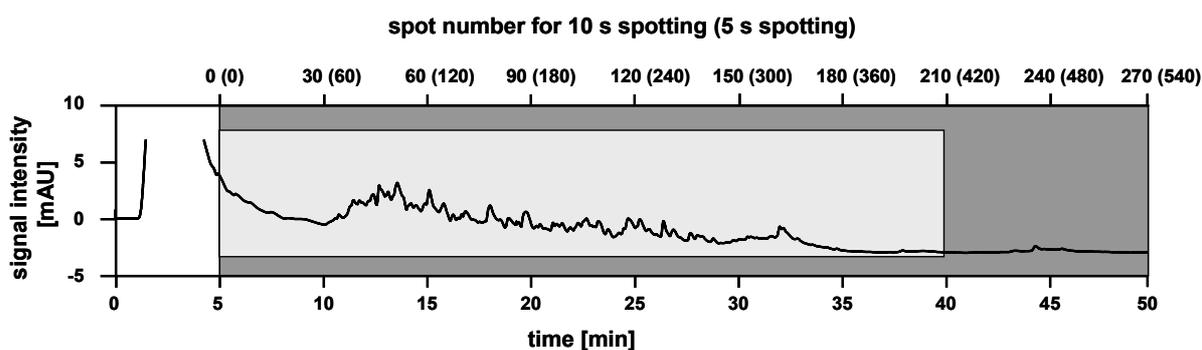


Figure 22. Separation of the ten-protein digest peptides; grey area (dark): spotted with Probot (min 5-50); grey area (light): analyzed with MALDI-TOF/TOF (min 5-40).

For each spotting time three replicates were performed whereas for every analysis the minimum MOWSE Score for a positive identification was 33 for proteins.

The following four parameters characterizing best the quality of analysis will be discussed for each validation step: The obtained MOWSE Score for the proteins and

peptides, the number of identified unique peptides and the sequence coverage for the proteins. The protein MOWSE Score indicates the quality of protein identification, but since it is a sum of the peptide scores of one protein, identifying five peptides with a low score (for example 20) would result in a similar score for the same protein identified with less peptides assigned with a high score (for example two peptides with the score of 50). Therefore the average peptide MOWSE Score is considered, too. Furthermore, the number of identified peptides shows the quality of the digest and also of the separation and identification system. However, all proteins differ in size and number of tryptic cleavage sites. To account for this fact the sequence coverage is always given in combination with the number of identified peptides.

The more peptides identified for one protein and the better the identification of these peptides the higher the MOWSE Score will be. Identification quality for the peptides is determined by the quality of the underlying MS/MS spectra. Clear signals without intense noise peaks and completeness of complementary b- and y-ion series are important factors for the quality of peptide identification ^[228]. With a score of 44 for 5-s spotting and a score of 46 for 10-s spotting ribonuclease A was the protein with the lowest computed MOWSE Score (Figure 23a). This protein was identified for both spotting times in only two of the three replicates. For all proteins and peptides the MOWSE Score was higher for the 10-s spotting, observing an average increase of 15.0 %. The average peptide score was computed by averaging all scores for one protein in one replicate. Then the three values for one protein obtained from the three replicates were averaged again. As for protein scores before, peptide scores are also increased with the 10-s spotting (Figure 23b). The highest average peptide score was obtained for myoglobin (73 for 5-s spotting and 90 for 10-s spotting, respectively). The clearest difference in the average peptide score can be observed for carbonic anhydrase 2. Here, for 10-s spotting a 35 % higher peptide MOWSE Score was obtained. Altogether, a 17.0 % higher average peptide MOWSE Score was obtained for 10-s spotting.

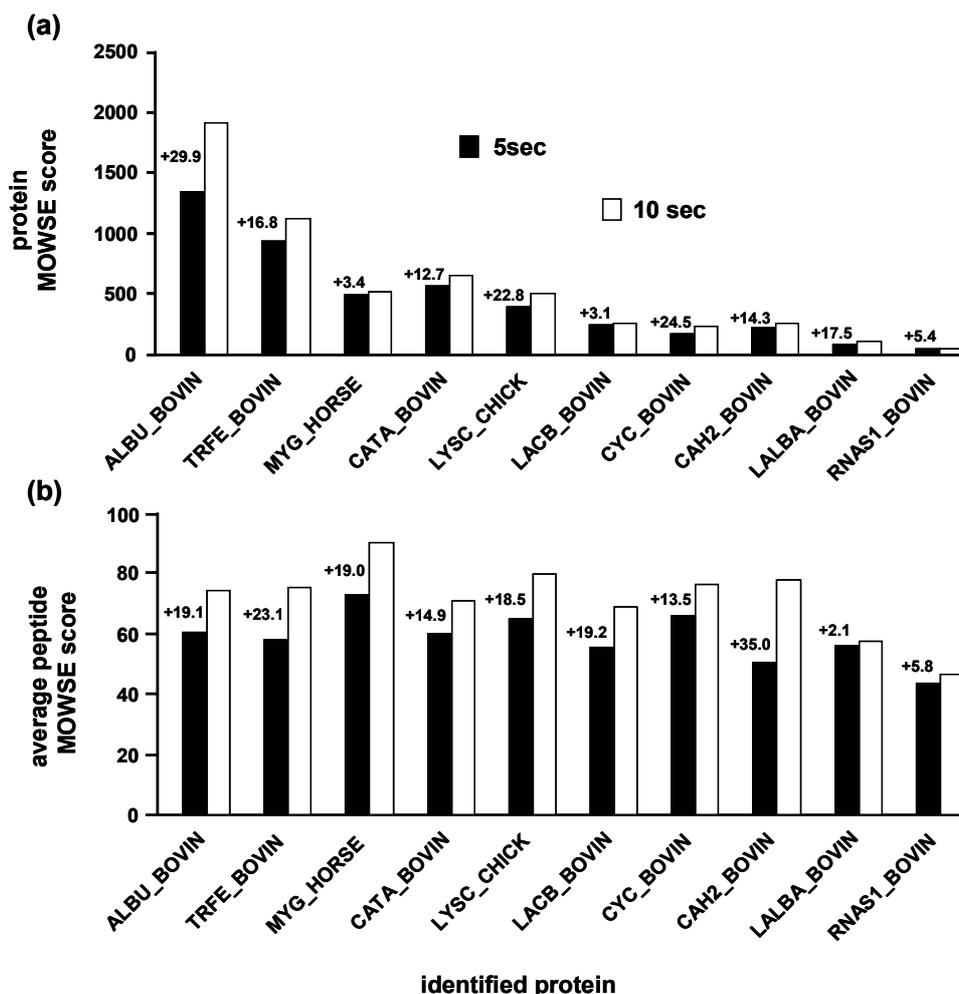
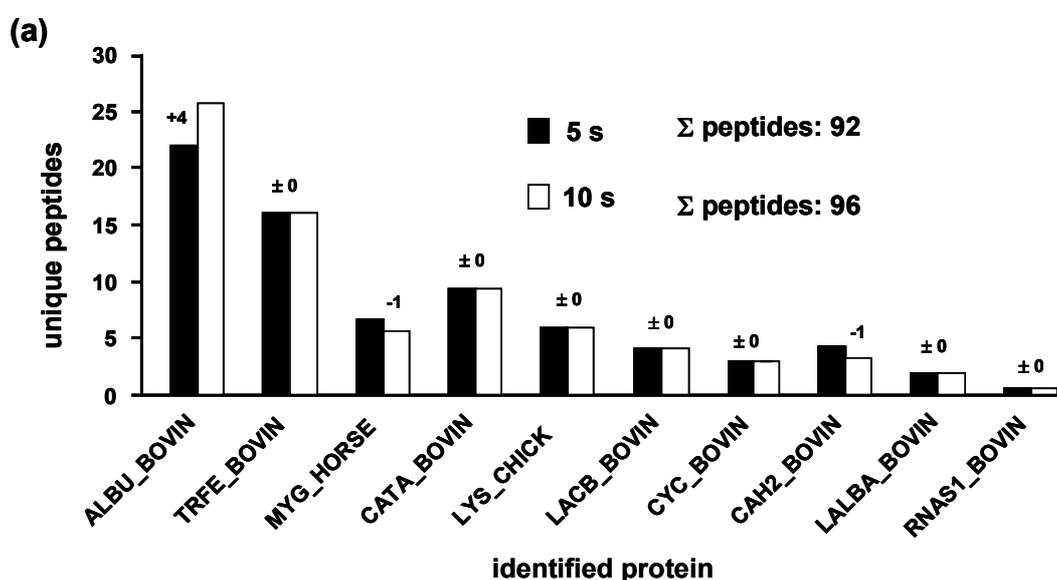


Figure 23. Mascot MOWSE Scores of the identified ten proteins, spotted with 5 s and 10 s per spot, respectively (average of three replicates). Numbers show percentage of additional scores for 10-s spotting. (a) Protein scores, (b) peptide scores.

Looking at Figure 24a, the low MOWSE Score of ribonuclease A becomes reasonable. This small protein was identified with only one peptide. As the MOWSE Score is a cumulative one, proteins with a high number of identified peptides are more likely to yield a high MOWSE Score. In the case of ribonuclease A, a small protein consisting of only 150 amino acids resulting in a molecular mass of 16.5 kDa is digested. This protein produces only four peptides with zero missed cleavages in the considered mass range and ten peptides with one missed cleavage site. In the relatively low complex ten-protein digest mixture consisting of 395 peptides (referring to the sum of peptides without and with one missed cleavage) ribonuclease A could be identified, but for a more complex sample it might be difficult. Moreover in a real proteome analysis this protein identification would be discarded because it would be

considered as “one-hit wonder”. These are proteins identified by only one peptide. In Figure 24b it can be observed that this one identified peptide for ribonuclease A is still sufficient to achieve sequence coverage of 15.3 %. α -Lactoglobulin identified with two peptides attains for the 5-s spotting only a sequence coverage of 10.8 %.

A remarkable fact can be observed in Figure 24. Although the protein scores seen in Figure 23 are all higher for the 10-s spotting, this is not the case for the identified peptides per protein and the sequence coverage. For almost all proteins an equal amount of peptides is identified with the 5-s spotting. For myoglobin and carbonic anhydrase 2 even more peptides can be assigned with the 5-s spotting, and only for BSA more peptides were observed with a spotting time of 10 s. Altogether, only four peptides more were identified with 10-s spotting. The sum given for the peptides identified is not equal to the sum of the bar values (because they are average values), but shows the sum of unique peptides identified for the ten-proteins in the standard mixture digest.



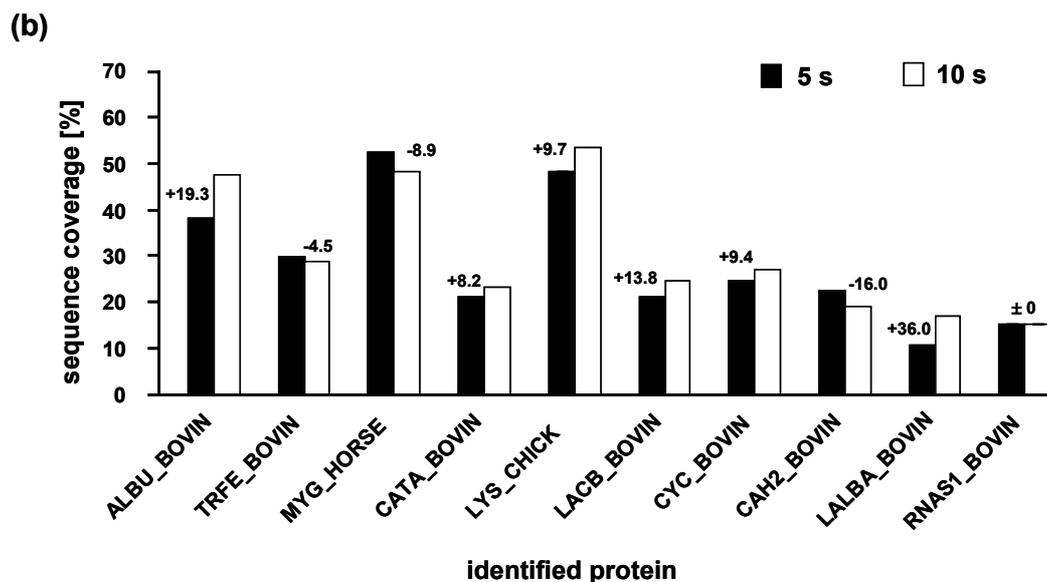


Figure 24. Characterization of the identified ten proteins for 5-s and 10-s spotting, respectively (average of three replicates). (a) Identified peptides for the single proteins, numbers show additional peptide identification for 10-s spotting. (b) Sequence coverage of the single proteins, numbers show percentage of rise in sequence coverage for 10-s spotting.

Although the number of identified peptides for both spotting times of the most proteins is equal, as displayed in Figure 24a, the sequence coverage for these proteins in Figure 24b is not. This effect occurs because the sequence coverage is averaged for the three replicates, too. For instance α -lactalbumin is identified twice with two peptides and once with one peptide with 5-s spotting whereas for 10-s spotting this protein is observed once with one, once with two and once with three peptides. The average peptides of both are two, but the sequence coverage for three identified peptides is 24.6 % and for one peptide 7.0 % in the case of α -lactalbumin. As a result the average sequence coverage varies for both spotting times. The protein with the highest sequence coverage is lysozyme C with 53.5 % obtained for 10-s spotting.

All ten-proteins could be identified with this setup by spotting 5 s per spot as well as 10 s per spot. Although only small differences were observed between 5 s and 10-s spotting in terms of sequence coverage (averaged 6.7 % higher with 10-s spotting) and peptides per protein (four peptides more with 10-s spotting) the MOWSE Scores for the proteins as well as for the peptides were higher applying 10-s spotting (15.0 % and 17.0 % higher, respectively). This effect is caused by the higher peptide

concentration per spot due to the double spotting time and hence better spectra quality. Therefore, the 10-s spotting was chosen for further system validation consistent with the moderate complexity of the sample compared to a real proteome analysis. To enrich enough material for MS analysis 10 s per spot were more suitable. On the other hand, 5 s per spot were used for second dimension of proteome analysis to reduce the tremendous sample complexity.

3.2.3 Preconcentration of peptides

Strong cation-exchange high-performance liquid chromatography (SCX-HPLC) is a proper retention mode for proteome analysis, applying reversed-phase high-performance liquid chromatography (RP-HPLC) in the second dimension. A high orthogonality is achieved with this combination. However, utilizing SCX in the first dimension leads to a high salt load in the collected fractions which is not compatible with mass spectrometry used for identification after the second separation dimension. Therefore a short desalting column is coupled on-line prior to the separation column which in our case consists of the same stationary phase as the separation column (PS-DVB monolith). In recent approaches two-dimensional systems utilizing RP-HPLC in both dimensions were introduced ^[229]. The retention modes vary mainly by applying first a basic pH (10.0) and second an acidic one (2.1) in the mobile phase of peptide separation. Even for this method, the desalting column is advantageous because the sample is concentrated (“trapped”) prior to separation. This is why the pre-column is also called trap column. Moreover, the trap column setup is very robust and therefore suitable for high-throughput analysis as necessary in proteome research ^[230]. In the developed nano-flow IP-RP-HPLC setup installing of the trap column was realized by positioning a further pump (Knauer, Berlin, Germany) connected to a 10-port switching valve (VICI, Schenkon, Switzerland) between the automatic injection unit and the gradient pumping system (Figure 21). This pump was run isocratically with 10 $\mu\text{L min}^{-1}$ 0.1 % aqueous HFBA for loading the trap column. Figure 25 illustrates the difference of acquired signals in the UV chromatogram without trap column prior to separation and with an installed trap column.

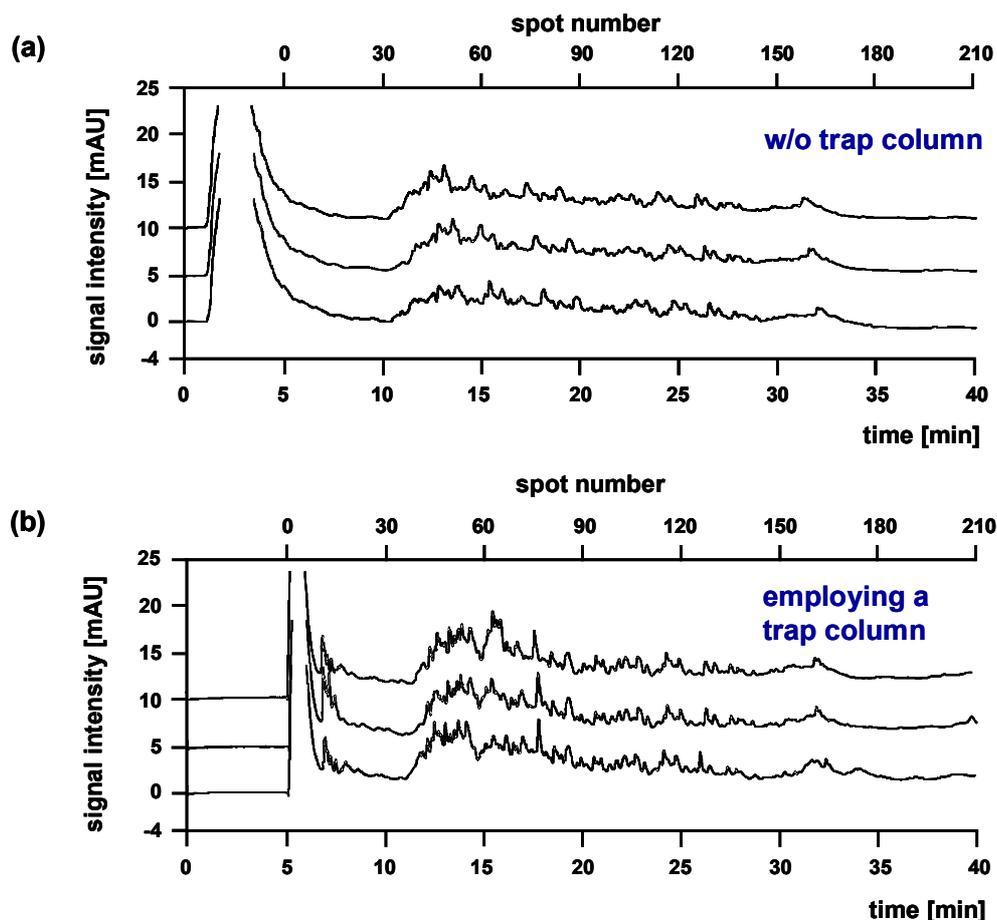


Figure 25. UV chromatograms of three-fold separation of the ten-protein digest (a) without trap column and (b) employing a trap column.

Obviously, using a trap column sharpens the peptide peaks and rises the signal intensity due to preconcentration on this small monolith. Whereas without the trap column peptide peak intensity never rises above 6.0 mAU it increases to 9.5 mAU for the chromatograms recorded for the separations using the trap column for each of the three replicates.

In Figure 26 the average protein and peptide MOWSE Scores are shown. A remarkable increase of protein scores for the three replicates using the trap column is observed in Figure 26a except for cytochrome C, up to 85 % for ribonuclease A. This protein is now identified in all three analyses. The average increase of the Mascot MOWSE Scores of the ten-proteins is 35.0 %.

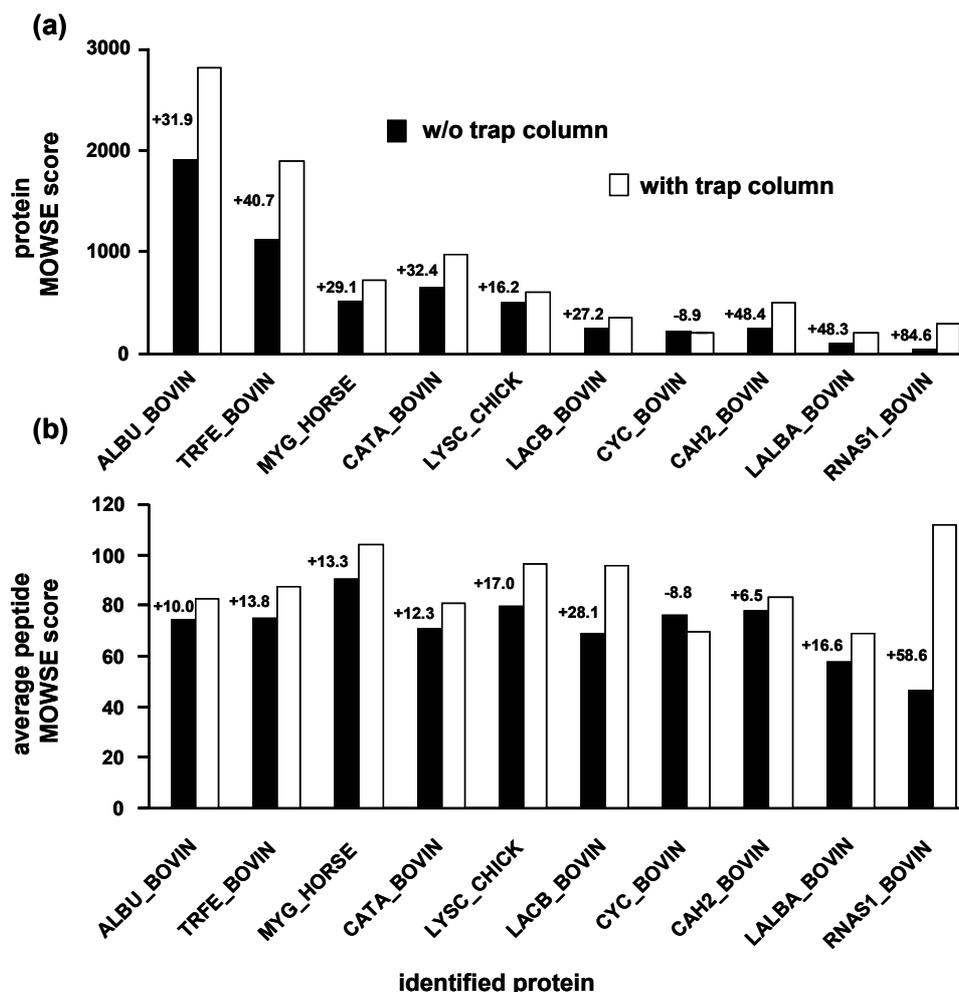


Figure 26. Mascot MOWSE Scores of the identified ten proteins, without and with utilizing a trap column, respectively (average of three replicates). Numbers show percentage of additional scores applying the trap column. (a) Protein scores, (b) average peptide scores.

An increase is also observable for the average peptide scores, but more moderate than for the protein MOWSE Scores (percentual rise of 17.0 %). Only ribonuclease A shows a remarkable increase of 58.6 % for the average peptide MOWSE Score when a trap column was applied.

Figure 27 depicts the identified peptides for the particular proteins and the sequence coverage of the ten proteins. A striking increase of peptide identification for the replicates involving the trap column is seen (Figure 27a). Altogether, 24 additional peptides could be identified. Especially for ribonuclease A which is the most critical protein two peptides instead of one were observed. As a result this protein would no longer be considered as one-hit wonder in a real proteome analysis.

Now BSA is the protein with the highest sequence coverage (62.5 %, Figure 27b) and not longer lysozyme C (59.6 %), resulted from 34 identified peptides for BSA and six peptides for lysozyme C.

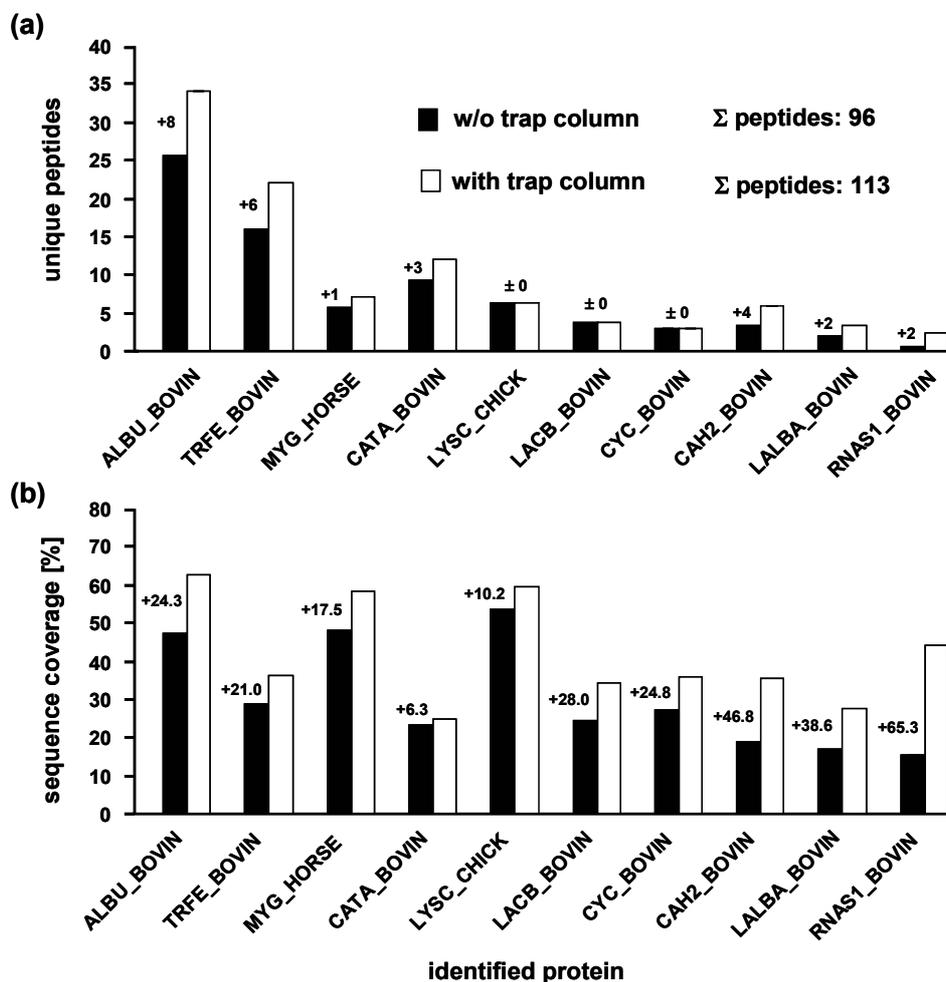


Figure 27. Characterization for the ten proteins separated either with or without trap column (average of three replicates). (a) Identified peptides for the individual proteins, numbers show different peptide identification for the runs analyzed using a trap column. (b) Sequence coverage, numbers show percentage of difference in sequence coverage for runs analyzed using the trap column.

For further validation of the separation and identification platform, the trap column was installed applying 10-s spotting.

3.2.4 Fragmentation of tryptic peptides

An alternative to the collision induced dissociation (CID) is using post source decay (PSD) with re-acceleration for peptide fragmentation. The peptide ions are generated and accelerated in the ion source 1 and also activated. During the field free drifting passage the metastable ions decay into b and y ion fragment series. Then the timed ion selector isolates the chosen precursor together with the fragment ions because they have the same velocities at this moment (however different kinetic energies). The CID chamber remains evacuated during PSD experiments and is only used for a second acceleration of the passing ions.

The utilized MALDI-TOF/TOF analyzer is equipped with a CID chamber, offering the opportunity of additional CID tandem mass spectrometric experiments. The CID chamber is vented with air molecules; the peptides collide with them and break mainly on their peptide backbone using medium collision energy.

After deceleration for precursor selection the ion package is accelerated to the reflector. There the precursor and the fragment ions are separated with respect to their mass. This means that the ions possessing a higher mass infiltrate deeper in the reflectron field than the lighter ones. Finally, the precursor and the fragment ions reach the microchannel plate detector at different times; first the fragment ions, later the precursor ions.

There is no clear evidence which fragmentation method is more suitable for proteome analysis for peptide and protein identification. Here, both methods were tested to find possible differences. Therefore, three replicates of the previous described ten-protein standard mixture digestion were performed for each approach.

Figure 28 shows two representative spectra for CID and PSD generated fragmentation patterns for the peptide GAGAFGYFEVTHDITR obtained from catalase. In Figure 28a the PSD spectrum and in Figure 28b the CID spectrum is depicted with only the b- and y-ion series marked which were recognized by Mascot search engine to keep clearness of the spectra. For the PSD spectrum an ion score of 143 was obtained and for the CID spectrum a score of 135. In the CID spectrum in the low mass range more signals appear, caused by the higher fragmentation energy applied for CID. In this region immonium ions and amino-acid signals are located. Although more fragments could be assigned to the peptide using CID the score for the PSD spectrum is higher.

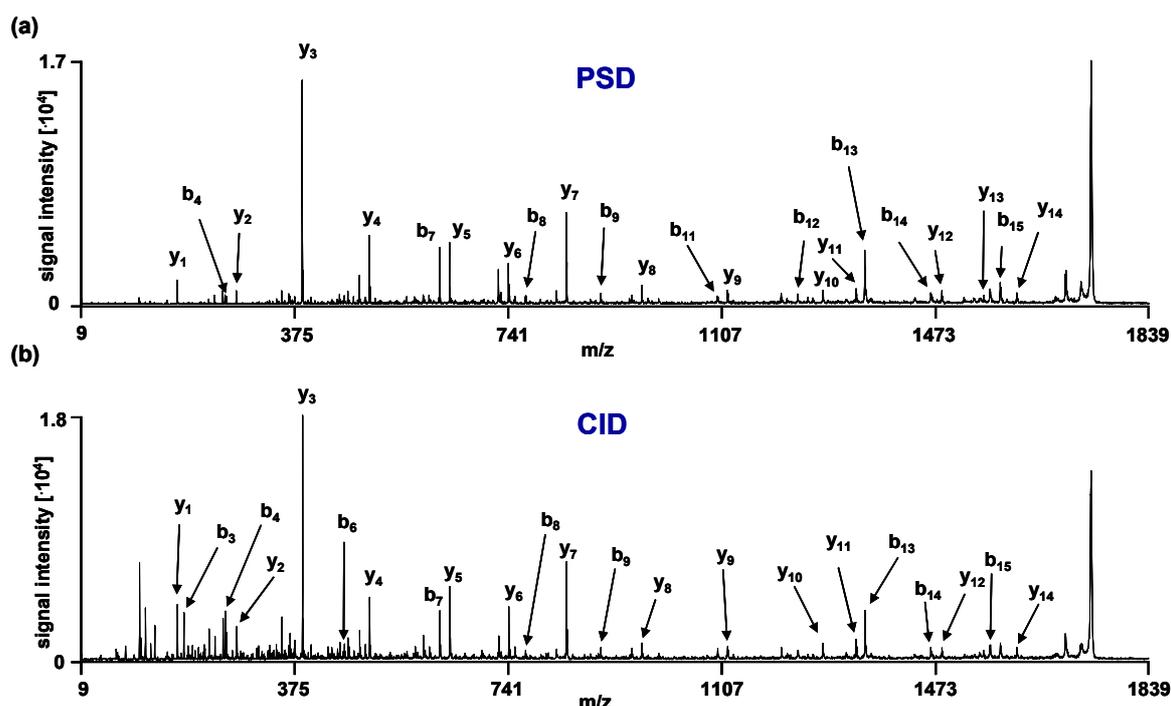


Figure 28. MS/MS spectra for the peptide GAGAFGYFEVTHDITR obtained from catalase, acquired (a) with PSD and (b) with CID.

Using only the absolute spectra intensity for estimation of quality is critical in MALDI mass spectrometry because so-called hot spots appear during analysis. This means that the co-crystallization of matrix and sample is not homogeneous at all. In the hot-spot areas the sample is higher concentrated than in others. The laser shoots arbitrarily on the spot and so even in one replicate spectrum the intensity can be higher than in another independently of any device parameters or sample concentrations. Therefore the intensity of the fragment base peak (y_3 , m/z 389.35) was taken to normalize both spectra. For the CID spectrum the absolute intensity of y_3 was 7.2 % higher than for the PSD spectra. Now comparing the normalized intensity of the most sequence relevant ion types (b and y) which appear in both spectra the normalized intensities for the CID spectrum signals are 3.8 % lower than for the PSD spectrum. This might explain the slightly higher score for the PSD spectrum. In Table 23, the ion types are listed corresponding to the example spectra for both fragmentation methods.

Table 23. Ion type distribution of an example spectrum for PSD and CID, respectively. 0 = fragment minus H₂O, * = fragment minus NH₃

ion type	PSD	CID
immonium	1	4
a	2	4
a ⁰	1	1
b	9	9
y	14	13
y*	6	4
y ⁰	1	0
internal fragments < 700 Da	8	18

Altogether 42 fragments were assigned to this exemplary peptide using PSD and 53 using CID. However, the fragmentation methods differ mainly in the number of internal fragments < 700 Da which do not provide as much sequence information as b- and y-ion series [231]. Surprisingly losses of ammonia and water mainly appeared in the PSD spectrum. It was expected because of the higher energy applied for fragmentation using CID that higher neutral losses would emerge employing this fragmentation method. Due to theory with CID fragmentation more immonium ions, which crop up in the low mass range, could be observed compared to PSD fragmentation. Finally, 30 b- and y-ions could be observed with PSD and 26 with CID. This might be an additional reason for the higher PSD spectrum score. However, it is difficult to say which fragmentation method is more suitable for proteome analysis, in general. Following, the known ten-protein standard mixture digest is used to characterize identification of peptides applying both approaches.

Figure 29 depicts the obtained average protein and peptide Mascot MOWSE Scores applying CID and PSD for peptide fragmentation, respectively. It is observed that no protein score for peptide fragmentation generated with CID is higher than for PSD spectra. Only one average peptide MOWSE Score (carbonic anhydrase 2) is higher applying CID. However, the average variation differs only slightly with 15.6 % higher scores for the proteins and 5.7 % for the peptides.

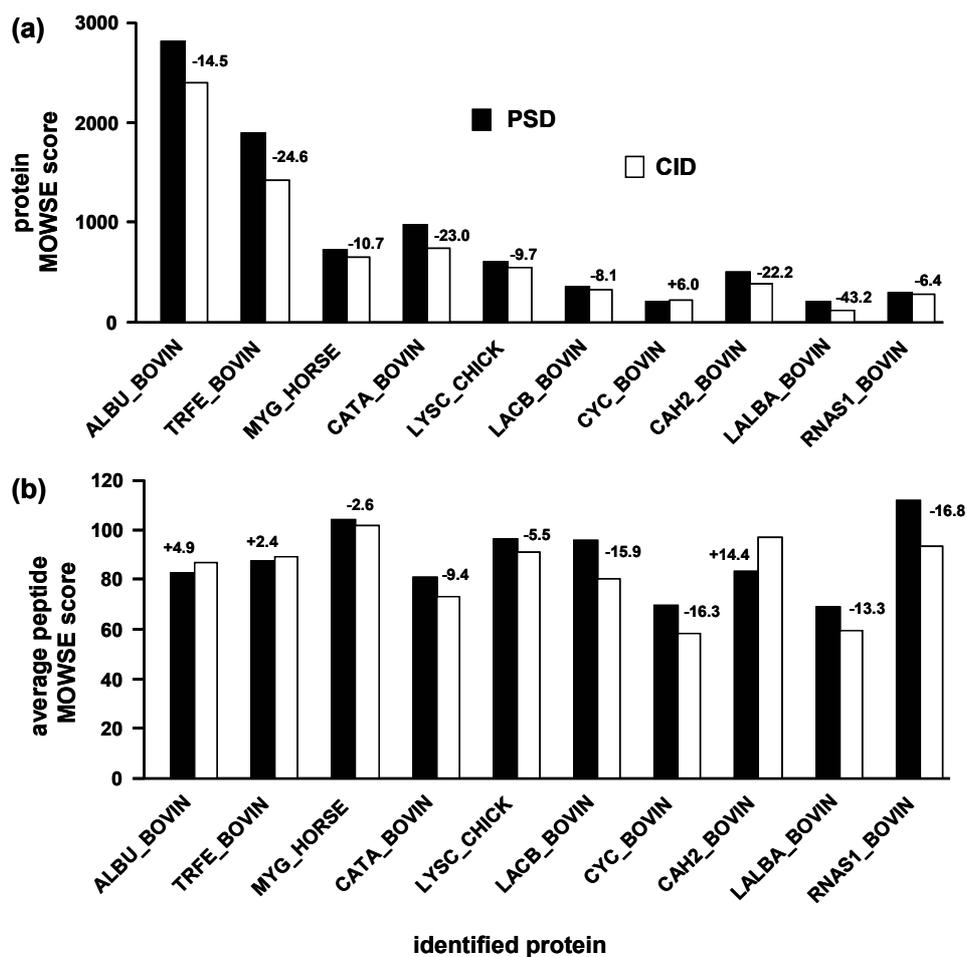


Figure 29. Mascot MOWSE Scores using PSD and CID for peptide fragmentation, respectively (average of three replicates). Numbers show percentage of varying scores applying CID. (a) Protein scores and (b) peptide scores.

In Figure 30, the identified peptides per protein as well as the sequence coverage are revealed. Altogether, 113 peptides could be identified with PSD and 101 with CID. Only for ribonuclease A one peptide more was obtained with CID, resulting in a higher sequence coverage.

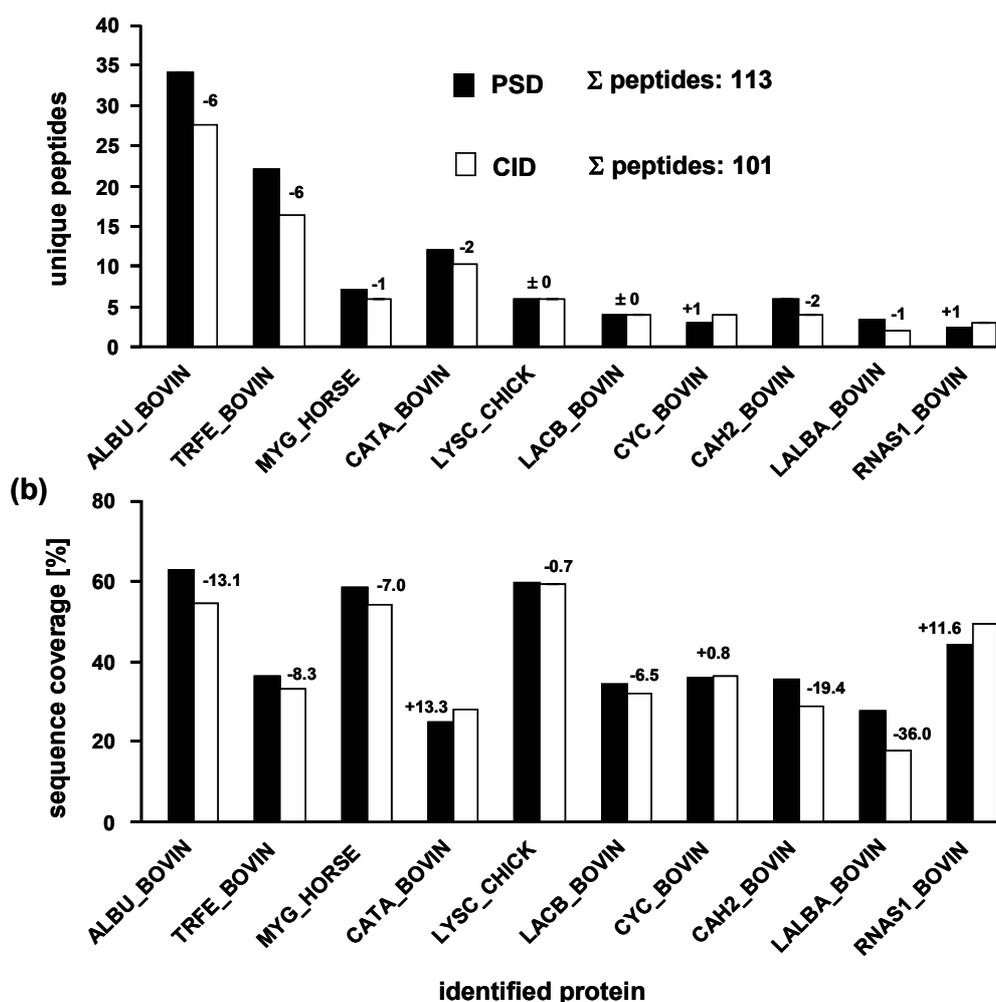


Figure 30. Characterization of the ten-proteins with peptide fragmentation performed either with PSD or CID (average of three replicates). (a) Unique peptides for the single proteins, numbers show different peptide identifications for CID. (b) Sequence coverage, numbers show percentage of difference in sequence coverage for peptides fragmented with PSD.

For proteome analysis aiming the identification of proteins, it is difficult to say which method is more suitable for peptide fragmentation. For special purposes like structure analysis of carbohydrates CID is preferred, permitting the determination of linkages in polysaccharides [232;233]. In proteomic content, determination of glycosylation moieties of post-translationally modified proteins is a prominent application [234;235]. Moreover, quantitation using the iTRAQ reagent strictly requires CID because the reporter ions, used for quantitation are low mass ions [236]. PSD with subsequent acceleration was applied for remaining validation of the nano-flow IP-RP-

HPLC-MALDI-MS platform because the focus was not set on identification of post-translational modifications but on peptide identification.

3.2.5 Repeatability of nano-flow IP-RP-HPLC peptide separations

For further analysis it was decided to use 10-s spotting for low complex samples as the ten-protein standard mixture digestion and 5-s spotting for higher complex samples. Furthermore, a trap column was installed and PSD was utilized for peptide fragmentation. Now the repeatability of the validated nano-flow IP-RP-HPLC off-line MALDI-TOF/TOF system was verified using the ten-protein standard mixture digest again. As seen in Figure 31, an adequate repeatability was achieved using the established separation platform. However, retention shifts are observable.

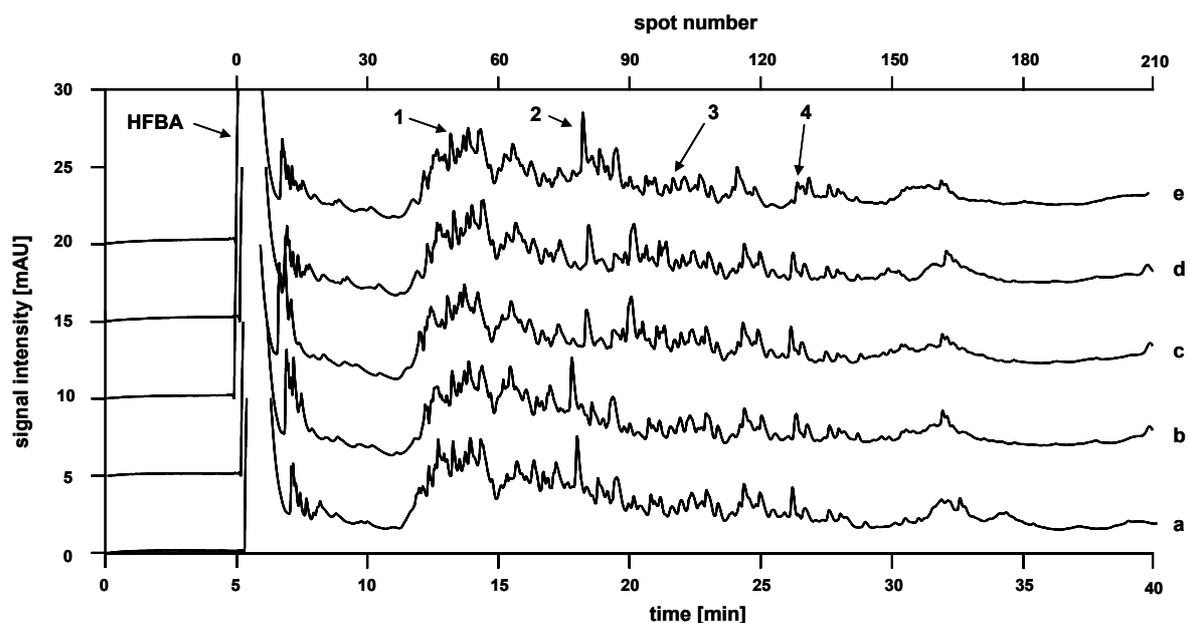


Figure 31. Repeatability of five chromatographic runs of the ten-protein standard mixture digest.

For four peptide peaks retention time was determined to evaluate the repeatability of retention time. Although it is difficult to integrate the peaks correctly when no baseline is available it was done by using the start and end point of the peak. So this is no absolute retention time assignment, but an approximation to show repeatability. To be sure the chosen peaks are peptide signals they were confirmed with mass

spectrometry analysis. The HFBA peak elutes with a relative standard deviation of 0.2 % as revealed in Table 24. Because of this precision the absolute peptide retention times were taken for evaluation repeatability. Usually, the run-to-run repeatability as an indicator for the system quality of the used HPLC platform should be better than 1.0 % [237]. The relative standard deviation is below this threshold for three peaks but for peak number 2 it is higher. Shifting of retention time for this peak can already be observed in the UV chromatograms in Figure 31 without integration. In contrast to the other considered three signals, peak 2 has a high intensity, so there might exist column overloading effects, and therefore a shift in retention time occurs. The four peptides were assigned from three of the five shown chromatograms (Figure 31c-e) with help of MALDI-TOF/TOF mass spectrometry to prove the peaks at the given retention times actually are peptide signals. The retention time revealed in Table 24 is the average of the three replicates. As the spotting time is 10 s per spot the retention time of MS analysis can only be assigned to the decimal place precision of 0.16 min. Peak 4 varies more than that 10 s offset what might result from the difficulty of integration. Moreover peptide KPVTDAENCHLAR from transferrin was identified in only two of the three MS runs.

Table 24. Characterization of repeatability of retention times for the established RP-HPLC system using a trap column, 10-s spotting and PSD as peptide fragmentation method.

peak	retention time [min]	relative standard deviation [%]	sequence	protein origin	retention time MS [min]
HFBA	5.77	0.20			
1	13.27	0.62	KPVTDAENCHLAR	transferrin	13.25
2	18.23	1.88	LVNELTEFAK	albumin	18.36
3	21.97	0.53	LVQFHFHWGSSDDQGSEHTVDR	carbonic anhydrase 2	21.81
4	26.33	0.70	NLCNIPCSALLSSDITASVNCAK	lysozyme C	25.92

It was further determined how many of the identified peptides elute in all three replicates taken for the chosen method using 10-s spotting, the trap column and PSD fragmentation for the peptides. Of the identified 113 peptides, 81 (= 71.7 %) were

identified in all three replicates as indicated in Figure 32. This high repeatability is another indicator for the quality of the established HPLC-MS platform.

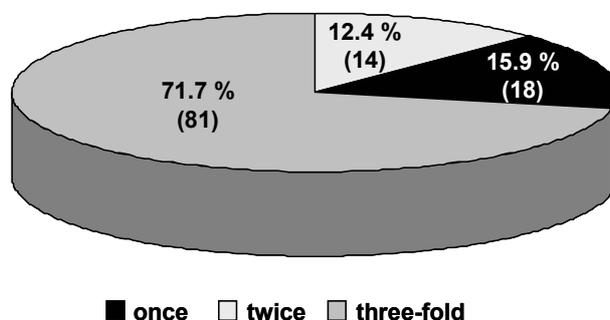


Figure 32. Repeatability of peptide identification for three replicates.

For further validation the repeatability for the optimized nano-flow IP-RP-HPLC-MALDI-TOF/TOF approach applying 10-s spotting, a trap column and PSD to fragment the peptides with respect to peptide identification repeatability for the individual proteins was tested.

As given in Figure 33 for three proteins (bovine serum albumin, cytochrome C and carbonic anhydrase 2), all three replicates delivered the same number of peptides.

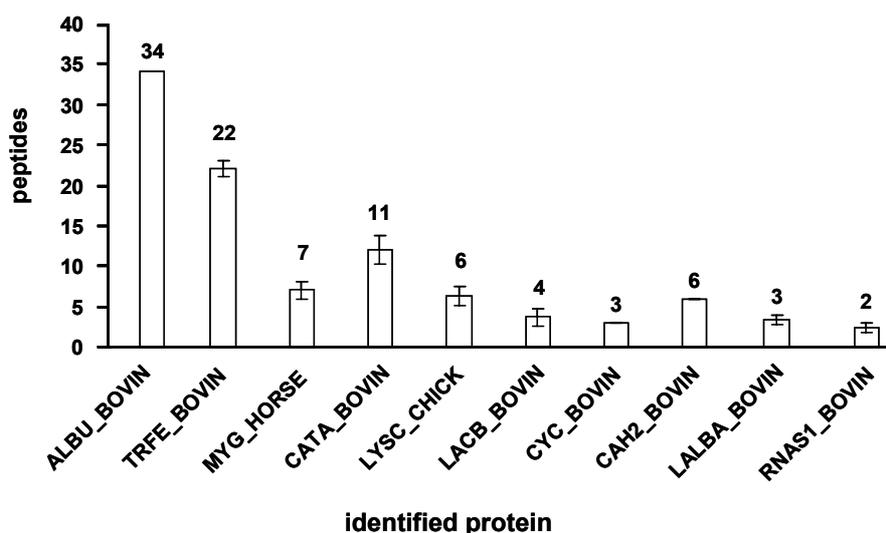


Figure 33. Repeatability of peptide identification for the single proteins.

The remaining seven proteins were assigned with only a slight variation in the number of peptides. These differences can be mainly explained by the

heterogeneous composition of the matrix/sample co-crystal and resulting hot-spot formation. This leads to poor shot-to-shot and sample-to-sample repeatability [238]. By shooting arbitrarily on the sample/matrix spot and summarizing shots to sub spectra which are again summarized to the final spectrum the repeatability is improved.

Diverse sample preparation methods were elaborated to circumvent sample concentration heterogeneity: applying liquid matrices [239;240], forming a binary matrix by adding co-matrices like fucose [241], using highly volatile solvents for fast evaporation approaches [242], using 4-HCCA-thin layer preparation with aerospray sample deposition [243], systems on sol-gel basis [244] or sample preparations without solvents [245]. However, all these methods are developed for special purpose. Today the prevalent approach for peptide identification still is using crystalline α -cyano-4-hydroxycinnamic acid as matrix.

3.3 Comparison of two buffer systems for SCX-HPLC of peptides in the first dimension

In the previous sections of this chapter the suitability of a nano-flow HPLC system using a trap column, 10-s spotting and PSD for peptide fragmentation for the separation and identification of a ten-protein standard mixture digest was demonstrated. In proteome analysis however, two-dimensional methods are necessary to cope with the tremendous complexity of the sample. Now the established nano-flow HPLC-MALDI-TOF/TOF approach should be evaluated in combination with SCX-HPLC of peptides in the first separation dimension employing two different buffer systems. Buffer system (1) consisted of sodium dihydrogen phosphate whereas the ionic strength through the gradient was increased with additional sodium chloride. In system (2) ammonium formate was utilized, increasing the ionic strength by elevating the ammonium ion concentration and simultaneously changing the pH. As ammonium possesses a pK_A of 9.24, there are still ammonium ions left at the adjusted pH of 6.8 whereas acidic analytes which have a pK_A lower than 6.8 are less charged and therefore elute off the column. Hence, in buffer system (1) the separation is performed by changing the ionic strength only while for system (2) an extra effect by changing the pH is prevalent. Figure 34 depicts the UV chromatogram of the ten-protein standard mixture digest separation utilizing buffer system (1). The chromatogram for

separation with ammonium formate buffer is not shown because at 214 nm it is UV absorbent itself and so no discrete peaks are observable from the UV chromatogram.

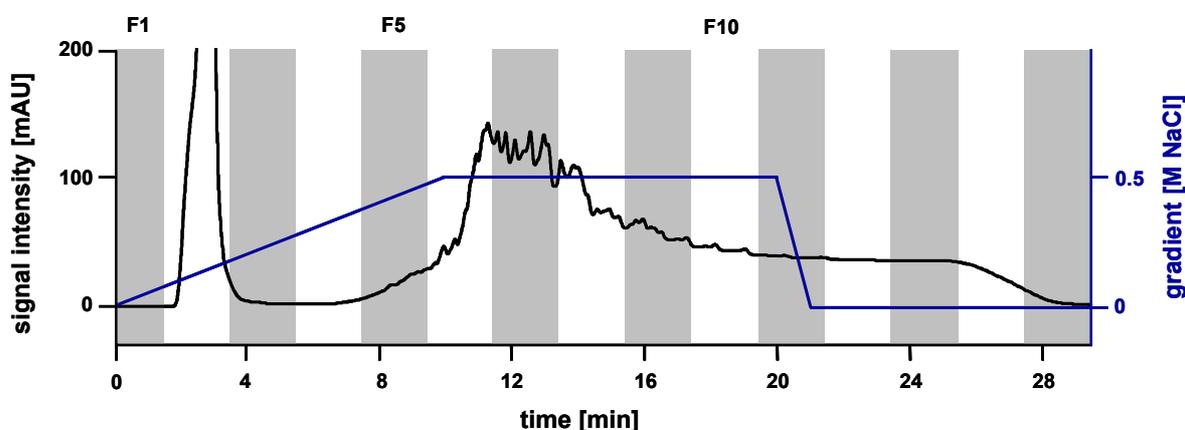


Figure 34. SCX separation of the ten-protein digest. Separation was performed using a sodium dihydrogen phosphate buffer system. The blue line shows the gradient profile.

Fractions were collected every two minutes for the fractionation of the ten-protein digest in the first dimension except for fraction 01 (1.5 min, 300 μ L). As observed in Figure 34 UV signals were obtained only in fractions 06-10 and only in these fractions peptides were analyzed by nano-flow IP-RP MALDI-MS/MS. The same parameters as for the validation steps before will be considered. The recorded MS raw data for fractions 06-10 were merged into one single file for each buffer system and submitted to Mascot for peptide and protein identification. Here 50 ppm instead of 100 ppm was applied for precursor mass tolerance, in order to be closer to requirements for proteome analysis. All other definitions remained the same as for 1D analysis. First, the MOWSE Scores for proteins and peptides are shown in Figure 35. In Figure 35a the protein scores are obviously higher than for 1D analysis where the highest achieved score was 2,805. This occurs because for multidimensional analysis another protein scoring is adopted by the Mascot software. The peptides count no longer only for their score, but additionally the average of their homology threshold is added (or identity when no homology threshold is available).

On average, the MOWSE Scores for all proteins are 24.1 % higher applying the ammonium formate buffer system. Only myoglobin is identified with a lower score. However, for peptide MOWSE Scores no significant advantage was observed with the

ammonium formate method. Employing the ammonium formate buffer system 1.6 % higher scores for the peptides were obtained on average.

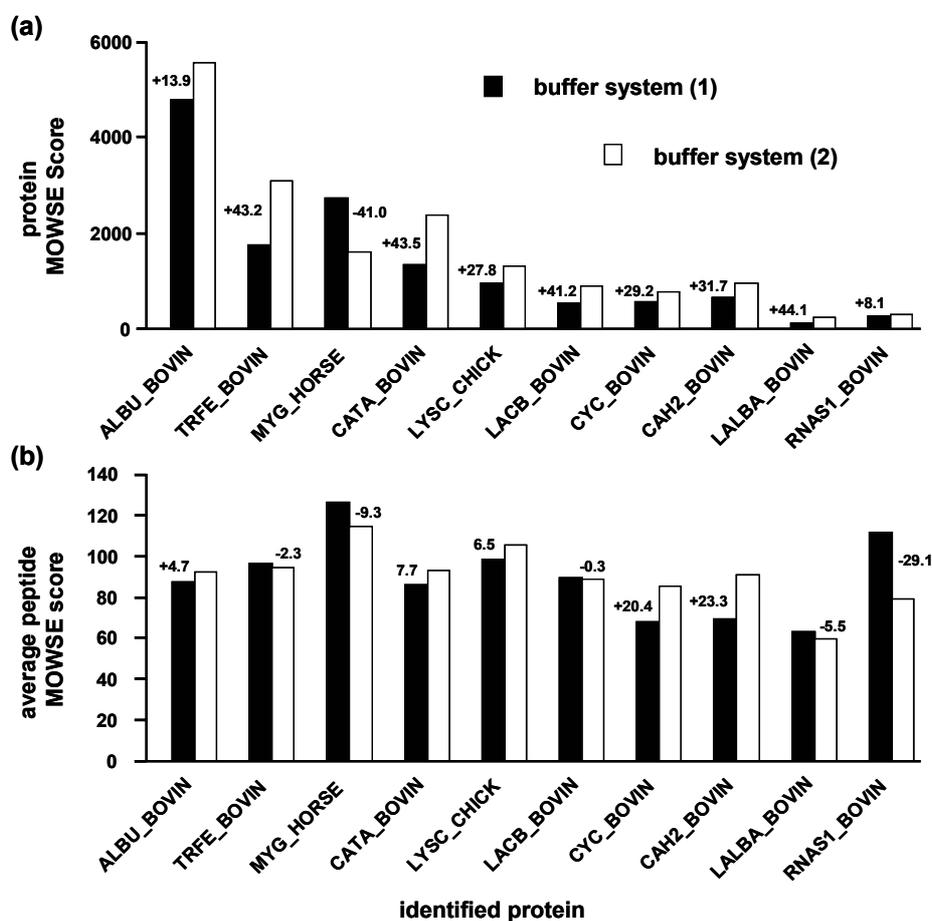


Figure 35. Mascot MOWSE Scores for the identified ten proteins with SCX separation in the first dimension and two different buffer systems. Numbers show additional scores for buffer system (2). (a) Protein scores and (b) average peptide scores. Buffer system (1) = sodium dihydrogen phosphate, buffer system (2) = ammonium formate.

Furthermore, the number of peptides per protein and the sequence coverage were evaluated as shown in Figure 36.

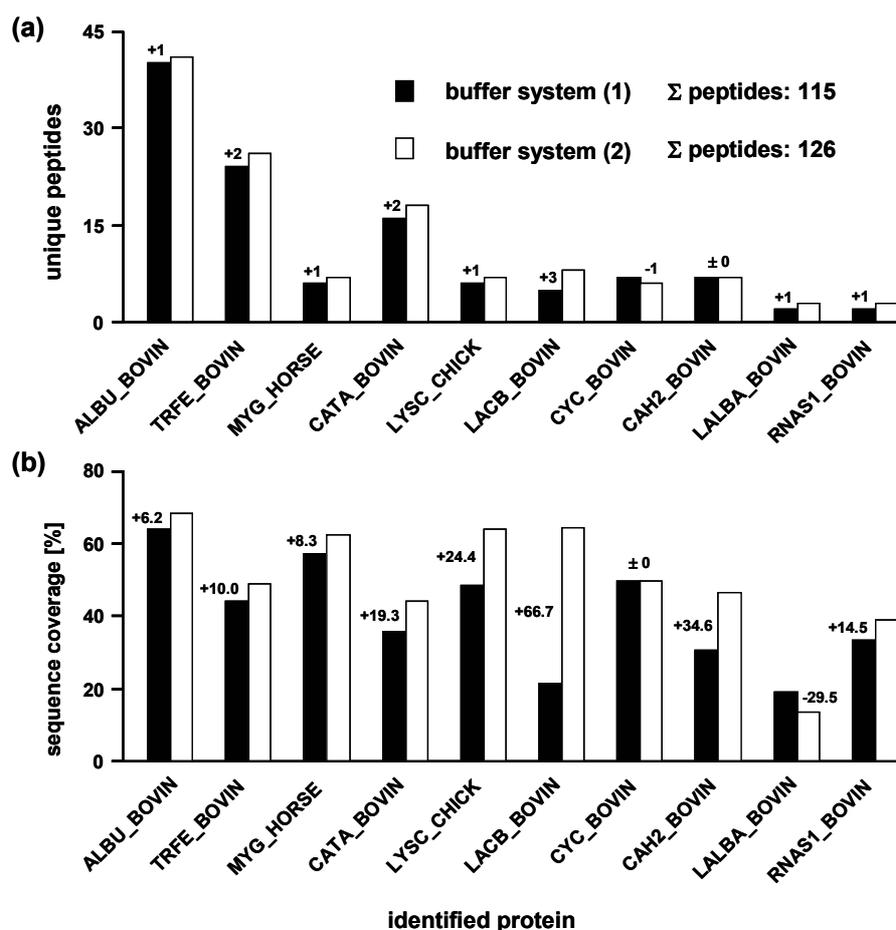


Figure 36. Characterization of the ten-protein digest separated with two different buffer systems for SCX in the first dimension. (a) Unique peptides for the individual proteins, numbers show difference in peptide identifications for buffer system (2) and (b) sequence coverage, numbers show percentage of difference in sequence coverage for buffer system (2).

For all proteins except cytochrome C and carbonic anhydrase 2, more peptides were identified using the ammonium formate buffer. Altogether, 11 peptides more were assigned utilizing buffer system (2) which corresponds to 8.7 % increase. A higher sequence coverage was obtained for α -lactalbumin with buffer system (1) in spite of a slightly higher number of peptide identifications. This effect occurs because using buffer system (1) two peptides were identified: one with 10 amino acids and one with 17. When employing buffer system (2) three peptides were assigned to α -lactalbumin: one was a hit with 10, the second with 8 and the last with 9 amino acids, but the last was the same as the second peptide, containing one missed cleavage. So only the additional sequence information of one amino acid was obtained. Altogether, there

were 27 amino acid residues identified with buffer system (1) and only 19 with buffer system (2). This explains the higher sequence coverage for α -lactalbumin using sodium dihydrogen phosphate buffer in spite of more identified peptides using ammonium formate. The most notable difference was found for β -lactoglobulin, where three more identified peptides resulted in an increase of sequence coverage of 66.7 % with respect to buffer system (2).

Altogether 115 peptides could be identified with the sodium dihydrogen phosphate buffer system, which is equal to 1.8 % more identified peptides in comparison to 1D analysis. With the ammonium formate buffer system 126 peptides were assigned to proteins corresponding to 11.5 % additional peptide hits compared to 1D separation, although here the MS mass window in Mascot software was narrower than in 1D analysis (50 ppm and 100 ppm, respectively). The precursor mass tolerance was set as for 1D (100 ppm) for both methods, to achieve a more accurate comparison. No change of peptide identification was observable for buffer system (2) where still 126 peptides were assigned to the proteins. However, with buffer system (1) 120 instead of 115 peptides could be assigned to the ten-protein standard which increases the percental additional identification in comparison to 1D analysis from 1.8 % to 6.2 %. The comparison of 1D analysis and the two different buffer systems is shown in Figure 37.

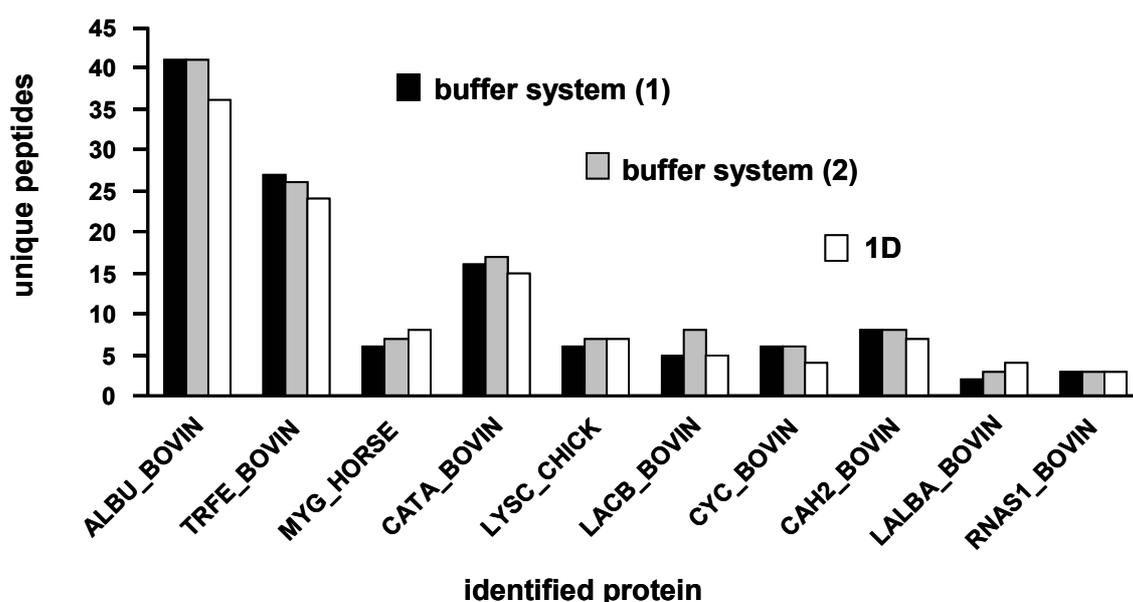


Figure 37. Identified unique peptides for the ten proteins with both buffer systems and 1D analysis in comparison.

Two proteins were identified with less peptides than in 1D analysis: myoglobin and α -lactalbumin. This might be caused by the triplicate analysis of 1D whereas for 2D only two replicates were measured. The remaining proteins were identified with an equal or higher number of peptides for buffer system (2). However, by employing buffer system (1) additionally less peptide hits for lysozyme C were obtained.

These results indicate that the extra separation effect caused by the simultaneously applied pH and salt gradient leads to higher peptide identification. This resulted in 11 peptides assigned additionally in comparison to buffer system (1), applying a precursor mass tolerance of 50 ppm. Five peptides more were identified, setting it to 100 ppm, respectively. Altogether, with the formate buffer system 32 % of the *in silico* calculated peptides for the considered mass range could be covered. Finally, the ammonium formate buffer system was utilized for the proteome analysis of a *Glioblastoma multiforme* tissue protein extract following in the next chapter.

3.4 Mascot search parameters for proteome analysis

There are a variety of possibilities to set Mascot parameters for proteome analysis. Five different parameter sets were chosen for the ammonium formate buffer system as shown in Figure 38. The MOWSE Score of proteins is not considered in this case. The precursor tolerance was modified (50 and 100 ppm) as well as the confidence interval of proteins and peptides. For peptides the ion score cut-off can be set as confidence interval or as fixed value.

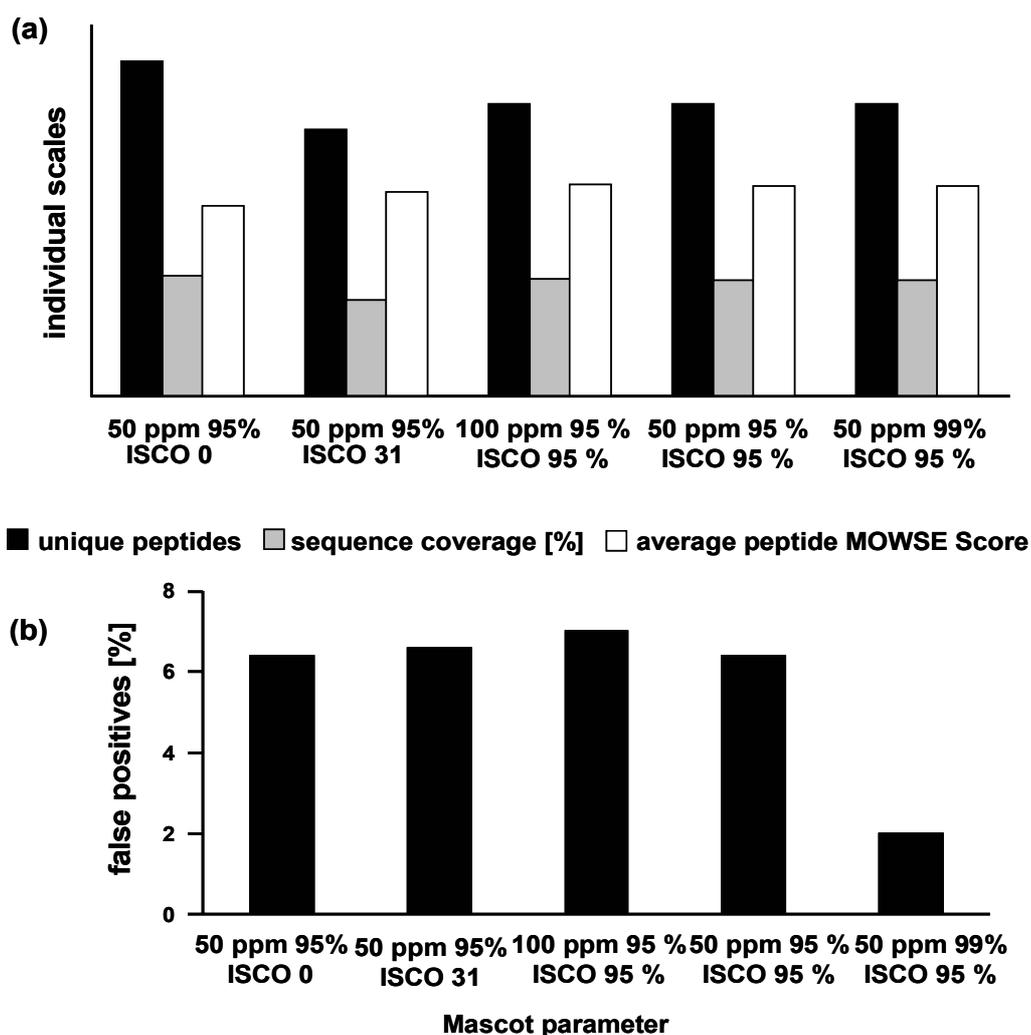


Figure 38. Settings of Mascot parameters for proteome analysis. (a) Unique peptides, sequence coverage and average peptide MOWSE Score for different Mascot settings and (b) false positives rate for different Mascot settings. ISCO = ion score cut-off.

The highest unique peptide identification and therefore the highest average sequence coverage were achieved with 50 ppm precursor tolerance and an ion score cut-off of zero. In contrast, only slight differences between 50 ppm and 100 ppm precursor tolerance are seen with an ion score cut-off of 95 %. This indicates that the impact of the ion score cut-off is more pronounced than that of the precursor tolerance. Choosing a fixed value of 31 for the ion score cut-off (suggested by Mascot) leads to a decrease in the identified unique peptides. This is reasonable because Mascot computes an appropriate ion score for every individual peptide hit above which the set confidence interval is reached (in this case 95 %). By setting a fixed ion score cut-off

all peptides below this value are dismissed whether or not reaching their specific ion score for the set confidence level.

The false positives rate should be below 5 % and 1 %, respectively. Actually the value is always above 6 % for protein confidence of 95 %, and 2 % for 99 %. These high false positives rates are probably caused by the small dataset. For 100 ppm the highest level with 7 % is obtained because of the lower mass accuracy. Hence, for the false positives rate the precursor tolerance seems to be more relevant than the ion score cut-off.

For proteome analysis a precursor tolerance of 50 ppm was chosen, a confidence interval for proteins of 99 % and for peptides of 95 % resulting in a false positives rate of 1.2 % (shown in the next chapter).

The mass tolerance for the peptide fragment ions was set to 0.5 Da for this 10 protein standard because the proteins to be identified were already known. In proteome analysis however, a stricter mass tolerance has to be applied because the results of protein identification have to be reliable and it is mostly unknown which proteins are present in the sample. Therefore, a mass tolerance for fragment ions of 0.2 Da was applied in proteome analysis.

4 Summary

In this chapter two approaches for 2D proteome analysis based on chromatography and mass spectrometry were established. First, an IP-RP-HPLC approach for intact protein separation as first dimension of semi top-down proteomics was optimized. Here, the focus was especially on separation efficiency and selectivity as well as the repeatability of the intact protein separation.

Furthermore, a nano-flow HPLC-MALDI-TOF/TOF platform for peptide separation in the second dimension of proteome analysis was elaborated. The impact of spotting time, preconcentration and fragmentation method for tryptic peptides to the protein and peptide identification were investigated. The most distinct effect on system performance was obtained by installing a trap column and consequent preconcentration of the peptides.

The established nano-flow IP-RP-HPLC MALDI-TOF/TOF approach was applied as second dimension of a bottom-up approach including SCX-HPLC of peptides of a moderately complex (395 peptides) standard protein digest sample in the first dimension. Two buffer systems were compared: a sodium dihydrogen phosphate buffer system with increasing ionic strength and an ammonium formate buffer system offering additional changing of pH. The ammonium formate buffer system was chosen for application in proteome analysis due to 8.7 % more identified peptides in comparison to the sodium dihydrogen phosphate buffer system. Compared to 1D analysis 11.5 % additional peptide identification could be achieved applying the 2D bottom-up strategy with ammonium formate as buffer system.

5 Conclusions

The established bottom-up and semi top-down approaches should be applied to a proteome analysis in the next chapter. According to the elaborated optimization in this chapter the following conditions were used for intact protein separation as first dimension for the semi top-down proteome analysis:

1. Two 50 x 4.6 mm monolithic columns were coupled in series to achieve higher separation efficiency.
2. As ion-pair modifier for IP-RP-HPLC 0.05 % TFA was added to the mobile phase.
3. The gradient was expanded due to the immense complexity of the sample to 15-60 % acetonitrile in 0.05 % aqueous TFA in 60 min.

The following parameters were elaborated for the second dimension nano-flow IP-RP-HPLC peptide separation of both methods:

1. The spotting time was set to 5 s for one spot due to the tremendous complexity of the sample.
2. The peptides were concentrated and desalted prior to separation realized by installing a monolithic trap column.
3. The tryptic peptides were fragmented by PSD with subsequent acceleration.

Finally, the evaluation of Mascot search parameters resulted in the following settings for proteome analysis:

1. The precursor mass tolerance was set to 50 ppm.
2. The fragment mass tolerance was set to 0.2 Da.
3. The confidence interval for peptide identification was set to 95 %.
4. The confidence interval for protein identification was set to 99 %.

Chapter IV

Application of the established
approaches to proteome analysis of
Glioblastoma multiforme

1 Introduction

The analysis of complex proteome samples is a challenging task. Until today, interactions in the human proteome including hundred thousands of proteins in an extremely broad dynamic range are rarely understood ^[246]. Therefore, it is essential to develop efficient methods for proteome analysis to get a deeper understanding of the functionality and tremendous complexity of the human proteome.

Three fundamental developments have made large scale proteome analyses possible: Two-dimensional gel electrophoresis, high-performance liquid chromatography and mass spectrometry.

Gel-based approaches are very powerful but suffer from some drawbacks: the observable mass and pI range are restricted; especially basic, very large and small as well as very hydrophobic proteins are underrepresented in the analysis results ^[3]. Chromatographic approaches are advantageous since they offer higher coverage of mass- and pI range.

The classical chromatographic-based strategy in proteomics is the bottom-up approach. Here, in both dimensions peptides are separated using different chromatographic modes. The peptides of the proteins detected with the bottom-up approach are distributed over the set of fractions obtained from the first dimension. Thus, even when only few proteins are of interest, as in diagnostics, the whole proteome analysis has to be repeated. By applying semi top-down approaches the distribution of peptides over the first separation dimension is circumvented. Here peptides are concentrated in distinct fraction(s) where the corresponding protein elutes in the first dimension. In further analysis only the fractions of interest containing the demanded proteins have to be digested. In the case of diagnostic application the proteome analysis in total has to be operated once and then only distinct fractions from the first dimension have to be investigated. This approach is advantageous in terms of micropreparation for proteins and sequence determination e.g. for biomarkers.

Recently, a micro semi top-down approach was introduced which employs chromatofocusing (CF) for intact protein separation in the first dimension applying weak anion exchange (WAX)-HPLC applying a pH gradient ^[247]. Fractions collected in the first dimension are subsequently digested, and the peptides are analyzed in

the second dimension. This approach provides the important information of the pI value for the identified proteins.

The aim of the presented study should be to apply the elaborated semi top-down approach to proteome analysis of a human *Glioblastoma multiforme* tissue sample. In this new semi top-down approach the retention time as characteristic information of a protein is obtained alternatively to the pI in CF. This alternative method should be compared to the classical shotgun proteomics approach.

2 Experimental section

2.1 Chemicals and materials

Sodiumdihydrogen phosphate (NaH_2PO_4 , min. 99 %), was purchased from Merck (Darmstadt, Germany). Trifluoroethanol (TFE, ≥ 99.5 %), tributylphosphine (TBP, 97 %), acetonitrile (E Chromasolv), α -cyano-4-hydroxycinnamic acid (4-HCCA), ≥ 98 % and human (Glu¹)-fibrinopeptide B (glu¹-fib) were obtained from Sigma-Aldrich (Steinheim, Germany). Trifluoroacetic acid (TFA, ≥ 99.5 %), heptafluorobutyric acid (HFBA, ≥ 99 %), ammonium formate ($(\text{NH}_4)\text{OAc}$, ≥ 97 %) and ammonium hydrogen carbonate (NH_4HCO_3 , ≥ 99.5 %) were purchased from Fluka (Buchs, Switzerland). Trypsin (sequencing grade modified) was purchased from Promega (Madison, WI, USA) and fused silica tubing from Polymicro Technologies (Phoenix, AZ, USA). Polyetheretherketone (PEEK) capillary tubing, tubing sleeves, microtight unions and microtight fittings were obtained from Upchurch Scientific (Oak Harbor, WA, USA). The utilized water was purified by a Purelab Ultra system (Elga, Siershahn, Switzerland). The thermo mixer was purchased from Eppendorf AG (Model Comfort, Hamburg, Germany) as well as the vacuum concentrator (Model Concentrator 5301). The centrifuge was procured from Heraeus (Model Biofuge13, Heidelberg, Germany). The *Glioblastoma multiforme* tissue was obtained from the group of Prof. Dr. Meese (Saarland University, Institute of Human Genetics, University Hospital, Homburg, Germany). The biopsy samples were immediately deep frozen after surgery at -70 °C. For protein quantification prior to digestion or injection of intact proteins into the first dimension the Bradford Assay was applied (Bio-Rad, Hercules, CA, USA).

2.2 Preparation of human brain tumor tissue protein extracts

To pieces of 10-15 mg of *Glioblastoma multiforme* tissue, 100 μL phosphate buffer (5 mM NaH_2PO_4) at pH 7.0 were added. After shaking (vortex genie 2, Bender&Hobein AG, Zurich, Switzerland) for 3 min the samples were incubated for 1 h at room temperature with gentle shaking and then sonicated for 5 min in ice-water. Addition of 100 μL trifluoroethanol (TFE), incubation for 2 h at 60 °C and

sonication for 2 min in ice-water followed^[248]. Subsequently the protein concentration of the samples was determined using the Bradford Assay. The proteins were digested on the same day for bottom-up approach (paragraph 2.5). The samples for the semi top-down approach were stored two days at -30 °C prior to digestion (paragraph 2.4). To reduce the disulfide bonds 50 mM tributylphosphine (TBP) was added to a final concentration in the solution of 4.5 mM (20 µL TBP to 200 µL protein extract) and incubated for 30 min at 60 °C directly before injection into the first dimension of the semi top-down approach.

2.3 IP-RP-HPLC prefractionation of intact proteins as first dimension of the semi top-down approach

Approximately 860 µg of proteins (quantified with Bradford assay) in 500 µL were injected into an analytical HPLC system (Model 1050, HP, Waldbronn, Germany) to perform separation with IP-RP-HPLC. The proteins were separated employing two 50 x 4.6 mm i.d. monolithic reversed-phase columns (ProSwift RP-1S, Dionex Corporation, Sunnyvale, CA, USA) connected in-series using a flow rate of 0.8 mL min⁻¹ and operating at 50 °C. The external six port injection valve (Model 7125, Rheodyne, Rohnert Park, CA, USA) was equipped with a 1 mL sample loop. Proteins were eluted with a 60-min gradient of 15-60 % acetonitrile in 0.05 % aqueous TFA. UV chromatograms were recorded at 280 nm, fractions were collected every second minute and stored at -30 °C.

2.4 Tryptic digest of the extracted proteins for the semi top-down approach

For the semi top-down approach every fraction obtained from the first dimension was evaporated to dryness in the vacuum concentrator and 100 µL of 50 mM ammonium hydrogencarbonate were added. The data from Bradford assay were used to adjust the enzyme/protein ratio to approximately 1/20 (40 µg trypsin to 860 µg proteins). Therefore two vials of trypsin each containing 20 µg were dissolved in 40 µL acetic acid (50 mM) and combined to 80 µL trypsin solution. The mixture was activated at 37 °C for 30 min in the thermo mixer (550 rpm) and then added to the proteins in the

fractions of the first dimension. The added trypsin mass was estimated based on the UV signals observed from the chromatogram of the fractionation as revealed in Table 25 (chromatogram depicted in Figure 41a).

Table 25. Distribution of trypsin solution over 28 fractions obtained from SCX-HPLC separation of peptides in the first dimension.

fraction number	added trypsin solution [μL]	added trypsin [μg]
01	4	2.0
02	4	2.0
03	1	0.5
04	1	0.5
05	2	1.0
06	2	1.0
07	2	1.0
08	2	1.0
09	2	1.0
10	4	2.0
11	4	2.0
12	4	2.0
13	4	2.0
14	4	2.0
15	4	2.0
16	4	2.0
17	4	2.0
18	4	2.0
19	4	2.0
20	4	2.0
21	4	2.0
22	4	2.0
23	2	1.0
24	2	1.0
25	1	0.5
26	1	0.5
27	1	0.5
28	1	0.5

The solutions were incubated for 15 h at 37 °C and 550 rpm in the thermo mixer. Subsequently, the reaction was quenched by adding 10 μL aqueous trifluoroacetic acid (1 % v/v) to each fraction. The fractions then were frozen at -30 °C. Prior to injection into the nano-flow HPLC-MALDI-TOF/TOF system 100 μL of aqueous 0.1 % HFBA solution were added to every fraction.

2.5 Tryptic digest of the extracted proteins for the bottom-up approach

After protein quantification utilizing Bradford Assay 20 μL of 50 mM TBP (4.5 mM final concentration in the solution) were added to 200 μL (100 μL phosphate buffer plus 100 μL TFE) of the protein extract. This protein solution was diluted 1:4 (220 μL protein solution plus 880 μL NH_4HCO_3) with ammonium hydrogencarbonate (50 mM) at pH 7.9 to reduce the TFE concentration. The enzyme/protein ratio was adjusted to approximate 1/50 using the protein concentration determined with Bradford assay. Twenty micrograms trypsin were dissolved in 20 μL of 50 mM acetic acid and activated in the thermo mixer for 30 min at 37 °C and 550 rpm. Of this solution 8.5 μL (= 8.5 μg) were added to the protein extract and incubated for 15 h in the thermo mixer at 37 °C and 550 rpm. The digest was quenched by addition of 10 μL aqueous TFA (~1 % v/v). The peptide solution was evaporated to dryness in the vacuum concentrator and was then stored at -30 °C. Prior to injection into the first dimension the peptides were dissolved in 200 μL ammonium formate (10 mM, pH 3 = eluent A in SCX-HPLC).

2.6 SCX-HPLC prefractionation of peptides for the bottom-up approach

About 426 $\mu\text{g}/200 \mu\text{L}$ of peptides (quantified with Bradford assay) were injected into an analytical HPLC system (Model 1050, HP) to perform peptide separation by SCX-HPLC. The peptides were separated on a SCX-column (Polysulfoethyl A, 200 x 2.1 mm i.d., 5 μm , 200 Å, PolyLC, Columbia, MD, USA) using a 10 x 2.1 mm i.d. guard column and a flow rate of 0.2 mL min⁻¹ operating at 25 °C. The external six-port injection valve (Model 7125, Rheodyne) was equipped with a 200- μL sample loop (Rheodyne). Eluent A consisted of 10 mM ammonium formate at pH 3 and eluent B of 500 mM ammonium formate at pH 6.8. Both mobile phase solvents contained 25 % acetonitrile. After injection of the peptide sample the system was held for 10 min at 10 mM ammonium formate (pH 3). Then the gradient for peptide separation started raising within 40 min to 250 mM ammonium formate (pH 4.9) followed by ramping to 500 mM ammonium formate (pH 6.8) within 10 min. For the last 10 min the system was held at 500 mM ammonium formate (pH 6.8). The UV chromatogram was recorded at 214 nm. Fractions were collected every two minutes and then frozen at

-30 °C. Prior to injection into the second dimension 100 µL of 0.1 % aqueous HFBA solution were added to the fractions.

2.7 Nano-flow IP-RP-HPLC MALDI-MS of peptides

Ten micoliters of sample (prepared as described in paragraphs 2.2-2.6) were injected into the second-dimension ion-pair reversed-phase nano-flow HPLC system. The experiment was performed in triplicate for the semi top-down approach and in duplicate for the bottom-up approach, respectively. The setup consisted of a nano-flow HPLC unit (Model Ultimate, LC Packings, Amsterdam, The Netherlands), an automatic injection unit (Model Famos, LC-Packings) and a loading pump (Model K-1001, Knauer, Berlin, Germany) with a 10-port switching valve (Model C2-1000D (stator) and 06A-8029C (rotor), VICI, Schenkon, Switzerland). The detector was equipped with a 3 nL Z-shaped flow cell (Model Ultimate). The UV chromatogram was recorded at 214 nm. A short precolumn (10 x 0.2 mm i.d.) was used to desalt and concentrate the peptide samples obtained from the first dimension. The separation column (60 x 0.1 mm i.d.) as well as the precolumn contained a monolithic PS-DVB-based stationary phase material synthesized according to the published protocol ^[221] (available from LC Packings, Dionex Corporation, Sunnyvale, USA). After injection into the second dimension the peptides were isocratically concentrated and desalted on the precolumn for 3 minutes. The flow rate of 0.1 % aqueous HFBA solvent, delivered by the loading pump was set to 10 µL min⁻¹. After switching the valve the peptides were eluted in back-flush mode onto the separation column. A 50-min gradient of 0-30 % acetonitrile in 0.05 % aqueous TFA was applied followed by ramping to 100 % acetonitrile in 10 min at a flow rate of 0.7 µL min⁻¹. Separation was performed at a temperature of 55 °C. The eluate from the HPLC system was spotted onto a stainless steel target (Opti-TOF 123 x 81 mm, ABI, Framingham, MA, USA) with a spotting unit (Model Probot, LC Packings, Dionex Corporation). As matrix 3 mg mL⁻¹ α-cyano-4-hydroxycinnamic acid in 70 % acetonitrile and 0.1 % aqueous TFA was utilized. The matrix solution contained human (Glu¹)-fibrinopeptide B (glu¹-fib) in a concentration of 15 fmol µL⁻¹. The matrix flow was set to 3.1 µL min⁻¹ with a spotting time of 5 s per spot (258 nL per spot). Sixty-seven minutes of the HPLC run (min 8 to 75) were spotted. For protein identification a MALDI-TOF/TOF mass spectrometer

(ABI 4800 TOF/TOF Analyzer, ABI) was used. The spectra were generated in positive reflector mode in a mass range of m/z 800–4,000. In MS mode 25 laser shots were summarized to one sub-spectrum and 40 sub-spectra were accumulated to the final spectrum (1,000 shots) with a frequency of 200 Hz and a laser wavelength of 355 nm. For optimized mass accuracy an internal calibration on the m/z of glu¹-fib (m/z 1570.677) was performed during sample acquisition. The m/z of glu¹-fib was excluded from precursor selection for the peptide ions. Fragmentation of maximum six precursors was realized in a metastable decay process with re-acceleration [223]. No additional collision gas was used for fragmentation. Stop conditions were set to achieve optimal quality of the MS/MS spectra: Between 750 and 3,000 laser shots were recorded for one spectrum. 50 laser shots were acquired for one sub-spectrum, minimum 15 sub-spectra and maximum 60 sub-spectra were recorded. When 3,000 laser shots or a signal to noise ratio (S/N) of 35 for at least 10 peaks were achieved the fragment ion data acquisition stopped.

Calibration of the mass spectrometer was performed as described in Chapter III (paragraph 2.6). MS calibration was initiated prior to each analysis, MS/MS calibration prior to every second analysis.

2.8 Data processing

The MS/MS spectra were smoothed with Savitsky-Golay algorithm using three points across the peak and a polynomial order of four. For exporting data to create Mascot generic files (mgf) the following settings were used in the MS/MS peak filter:

Mass range: m/z 60 to precursor mass -35 Da; peak density: maximum 20 peaks per 200 Da; minimum S/N: 10; minimum area: 200 and maximum 65 peaks/precursor. To identify proteins the “.mgf”s were sent to Mascot software (version 2.2.03, Matrix Science, London, UK) which uses the MOWSE (Molecular weight search) algorithm. These settings were applied: database: Swiss Prot (Version 54.7, January 15th, 2008); taxonomy: *homo sapiens* (18,117 sequences); enzyme: trypsin; variable modification: methionine oxidation; peptide tolerance: 50 ppm; MS/MS tolerance: 0.2 Da; maximal missed cleavages: 1 and ion-score cut off for peptides: 0.05 (95 %). The probability of an identification as a random event for proteins was set to 0.01 (99 %).

2.9 Data interpretation

To categorize the identified proteins of the *Glioblastoma multiforme* tissue to gene ontologies GeneTrail software was used. Here, the subset of membrane proteins was selected for characterization of the retention behavior for intact proteins in the first dimension of the semi top-down approach. The following parameters were set: The chosen organism was human; identifier type: Swiss Prot; significance threshold: 99 % and the minimal genes of a category shown: 100. The GeneTrail software is based on gene functions of the Gene Ontology (GO) terms. PANTHER software was utilized to divide the identified proteins into main molecular functions based on their gene function.

3 Results and discussion

3.1 The bottom-up approach as standard method in chromatography- and mass spectrometry-based proteomics

3.1.1 Two-dimensional fractionation of peptides

In two-dimensional shotgun proteomics peptides are separated in both dimensions. The applied method in this study includes strong cation-exchange chromatography in the first dimension. Figure 39a provides the UV chromatogram of this separation.

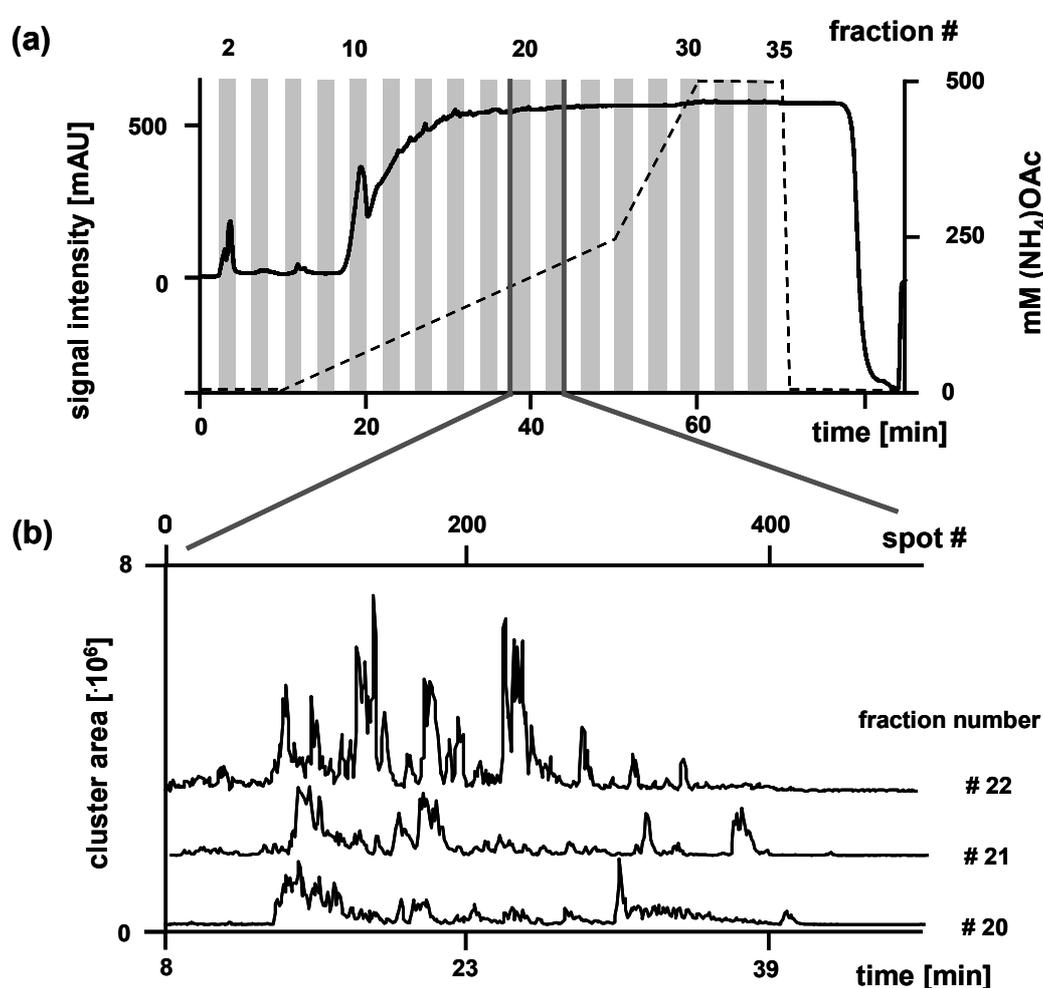


Figure 39. Two-dimensional fractionation of peptides in the bottom-up approach. UV chromatogram (214 nm) of the first dimension (a) and BPCs of fraction 20-22 in the second dimension (b). Chromatographic conditions are given in the experimental section.

A shallow gradient with increasing ionic strength and pH was selected for the first dimension. No defined peaks are observable because the signal was recorded at 214 nm and ammonium formate also shows UV activity at this wavelength.

Fractions were collected in a time range of 80 min every second minute and the 40 obtained fractions were evaporated to dryness before tryptic digestion. Subsequently peptides of every fraction were injected twice into the second dimension, concentrated and desalted on a monolithic trap column, and a monolithic column was also used for peptide separation. A shallow 50-min gradient was applied to perform peptide separation. The eluted peptides were analyzed employing the previously optimized nano-flow HPLC-MALDI-TOF/TOF platform. Figure 39b illustrates the base peak chromatograms (BPC) of three adjacent fractions from obtained from the second dimension, acquired by MALDI-TOF/TOF. A time range of 50 minutes (minute 8-58) was analyzed including 600 spots. In the exemplary shown fractions (Figure 39b) peptide signals until spot 450 were observable which corresponds to an acetonitrile concentration of 27.3 % (including 0.05 % TFA).

3.1.2 Protein identification

The recorded mass spectrometric data were sent to the Mascot search engine which compares tandem mass spectra simulated from *in silico* digestion with the recorded mass spectra. Altogether, 2,056 proteins (Table 26) were identified in 34 fractions obtained from the first dimension peptide separation. As shown in Figure 40 the typical pattern for SCX separation in the first dimension is visible. In ion-exchange chromatography the analytes elute as clusters depending on their overall charge ^[249;250]. Only the proteins and peptides which appear for the first time in an individual fraction are included in the bar chart. Furthermore, the average peptides per protein are displayed in Figure 40. In fraction 19 the highest number of peptides (603) was observed, the highest number of proteins were identified in fraction 2 (190) followed by fraction 19 (184).

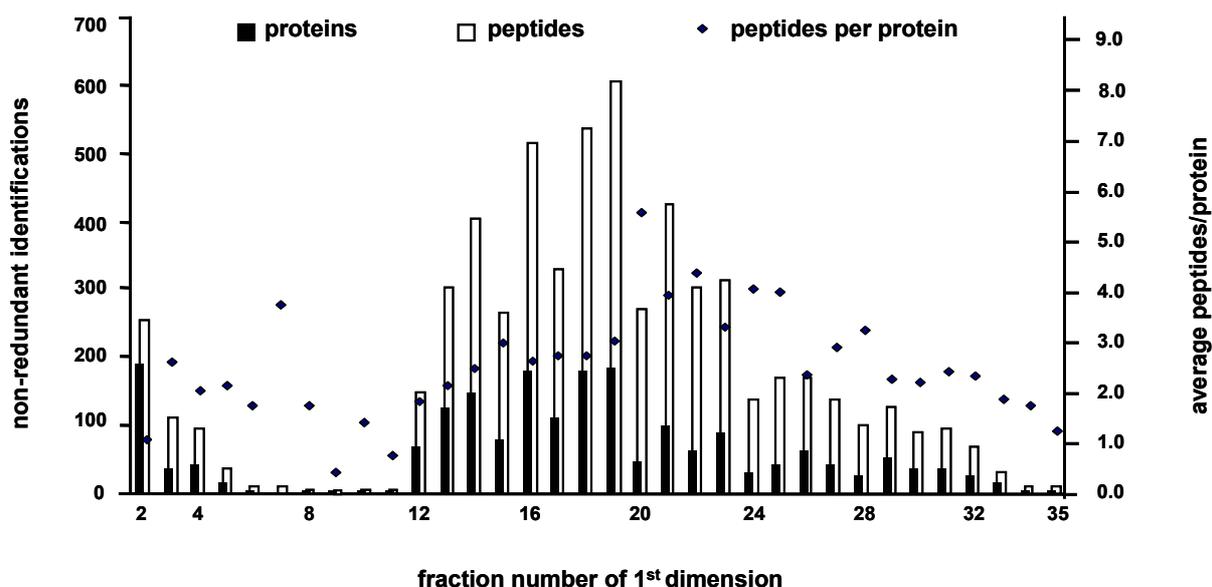


Figure 40. Non-redundant protein and peptide hits identified with the bottom-up approach. Distribution of identifications over the 34 protein- and peptide containing fractions (left scale) as well as average peptides per protein for these fractions (right scale).

In Table 26 characteristic data for the bottom-up approach are given. Eight hundred-eleven proteins, so-called one-hit wonders, were identified by only one peptide which corresponds to 39.4 % of the 2,056 identified proteins. There are various reasons for the appearance of one-hit wonders: the other peptides of one protein are too large or too small to be detected in the set mass range. This can be caused either by the protein sequence itself (high or low numbers of tryptic cleavage sites) or by incomplete digestion. Another reason for observing one-hit wonders are mass spectra of low quality due to sample contamination. Further a bad ionization efficiency caused by the amino acid composition of the peptide or its low abundance result in the appearance of one-hit wonders. Another well known phenomenon is that some peptides do not undergo sufficient fragmentation, hence not allowing sequence determination and identification (weak signal intensity). One-hit wonders were excluded from further evaluation. Although a significance threshold of 99 % for the proteins is given for this analysis one-hit wonders hold the highest probability to be false positive protein identifications. By excluding them, of course, a high number of correct peptide hits are excluded, too. It is even assumed that for one false positive identification 19 correct ones are eliminated^[251]. A similar rate of one-hit wonders as

in this study was observed by Yates et al ^[252] in the proteomic study of a whole cell lysate of yeast. They reported a one-hit wonder rate of 34 %.

As expected for human samples a high homology between the proteins is observed. As a result 594 peptides can be assigned to more than one protein. The percentage of annotated spectra (28.3 %) lies in the typical range for tandem mass spectra ^[253].

Table 26. Results from the bottom-up approach (Two replicates in the second separation dimension).

parameter	value
identified proteins	6,883
non-redundant proteins	2,056
“one-hit wonders”	811
identified proteins without “one-hit wonders”	1,245
peptides in non-redundant proteins	7,122
non-redundant peptides	6,528
annotated spectra [%]	28.3
false positives [%]	1.3
average mass accuracy (MS mode) [ppm]	18
average peptide score	66.5

The remaining results revealed in Table 26 are discussed in detail for the semi top-down approach in the paragraph 3.2.2 of this chapter.

3.2 A semi top-down approach as alternative method for proteome analysis

3.2.1 Two-dimensional fractionation of proteins and peptides

Prior to first dimension separation for semi top-down approach disulfide bonds of the intact proteins were reduced with tributylphosphine to unfold protein structure and promote better interaction with the stationary phase of the monolithic separation column. In Figure 41 the UV chromatogram of the fractionation of the intact proteins is shown as well as the BPCs of three adjacent fractions of the second-dimension separation.

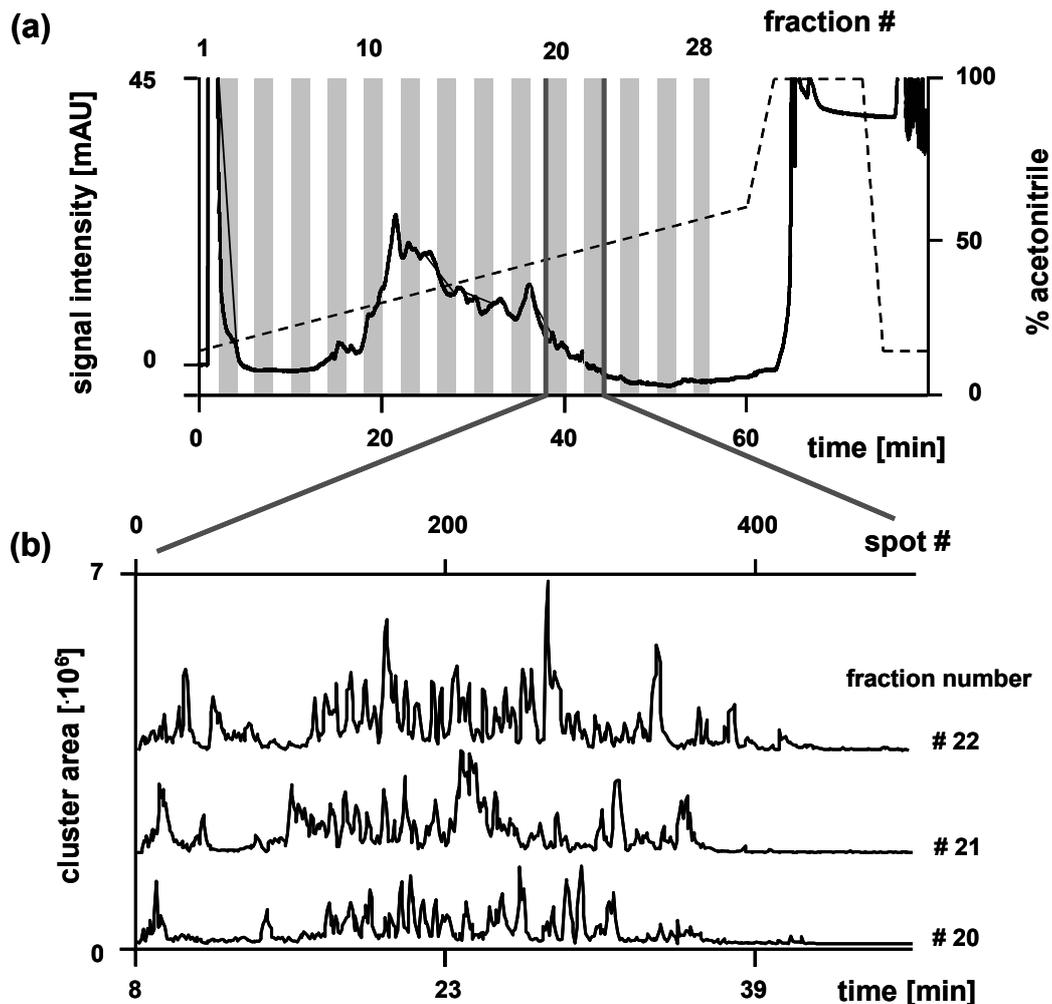


Figure 41. First dimension fractionation (280 nm) of intact proteins followed by peptide fractionation. UV chromatogram of first dimension (a) and BPCs of fraction 20-22 in second dimension (b). Chromatographic conditions are given in the experimental section.

Fractionation for further separation has always to cope with some aspects: conservation of the accomplished separation, dilution of the peptides or proteins gained in the individual fractions, time consumption for separation in the second dimension and carryover of peptides or proteins in contiguous fractions.

Especially MALDI-TOF/TOF analyses after second-dimension separation are very time consuming. One analysis of a 50-min HPLC run can last up to 15 or even 20 hours for complex samples. Therefore, two minute fractions were taken over a time range of 80 min and these 40 fractions were analyzed in the second dimension. However, the dilution of the proteins is still moderate and the high sensitivity of the

MALDI-TOF/TOF instrument still provides detection of low abundant proteins. Furthermore, the precursor selection procedure of the mass spectrometer supplies a pertinent decision for precursor selection: the spot with the highest intensity for the precursor ion of interest of the whole spotted HPLC run is chosen for fragmentation. The fraction-to-fraction carryover for proteins may be higher than for peptides due to the larger chromatographic peak width for proteins. This will be discussed later in detail (paragraph 3.4.3). For the fractionation of the intact proteins (Figure 41a) a shallow 60-minutes gradient was used to achieve high peak capacity resulting in observable discrete peaks containing numerous of different proteins.

Into the second dimension the samples were injected in triplicate. Peptides were concentrated on a PS-DVB-based monolithic trap column. Subsequently the analytes were transferred for separation onto an equal monolithic capillary column and separated with the same gradient as used for the second separation dimension of the bottom-up approach. Typical base peak chromatograms (BPC) of three contiguous fractions are depicted in Figure 41b. In each case the first of the three replicates for one fraction is shown. In the fractions 20, 21 and 22, obtained from the first separation dimension 176, 174 and 178 proteins were identified, respectively.

3.2.2 Protein identification

The acquired mass spectrometric data were launched to the Mascot search engine using the same settings as for the previously discussed bottom-up approach. Proteins could be identified in the first 28 fractions as displayed in Figure 42.

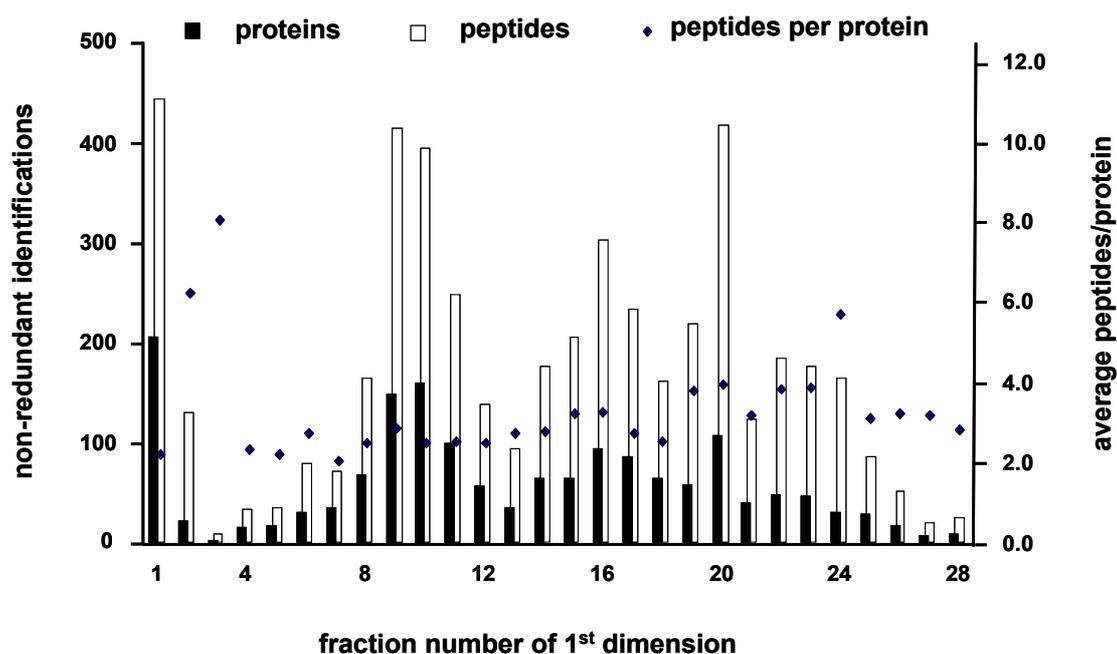


Figure 42. Proteins and peptides identified with the semi top-down approach.

Proteins and peptides which appear for the first time in a fraction are plotted (left scale) as well as average peptides per protein of the individual fractions (right scale).

The highest number of proteins was identified in the first fraction which indicates that fragments of proteins and small hydrophilic proteins elute in this fraction because the gradient was started at a low concentration of acetonitrile (15 %) in 0.05 % aqueous TFA. The rest of the proteins was distributed over the remaining 27 fractions. Fraction 9 contained the highest number of non-redundant proteins (159) beside fraction 1. In fraction 20 the highest number of peptide hits was obtained (417). The average number of identified peptides per protein for the individual fractions is also shown in Figure 42. In fraction 3, a very high protein/peptide ratio is observed because there was only one non-redundant protein identified, including eight peptides which is similar for fraction 2. An average peptide/protein ratio of 3.3 could be observed over all fractions as shown in Table 27. Altogether, 5,270 protein hits were obtained for the 28 fractions analyzed using the Mascot search engine and the Swiss Prot database (release 54.7, same as for the bottom-up approach). After merging the raw data files of the triplicate analyses for all fractions, 1,642 non-redundant proteins (Table 27) were identified. With this semi top-down method altogether 5,439 peptides were assigned to the 1,642 proteins. After removal of redundant identifications, 4,945

unique peptides were identified. This means that 494 peptides occur in more than one protein.

Table 27. Results from the semi top-down approach (Three replicates in the second separation dimension).

parameter	value
identified proteins	5,270
non-redundant proteins	1,642
“one-hit wonders”	638
identified proteins without “one-hit wonders”	1,004
peptides in non-redundant proteins	5,439
non-redundant peptides	4,945
annotated spectra [%]	30.5
false positives [%]	1.2
average mass accuracy (MS mode) [ppm]	17
average peptide score	70.3
average protein/peptide ratio	3.3

As expected, false-positives rate, average mass accuracy and average peptide score are similar to those of the bottom-up approach because the same instrumental setup and the same settings for the Mascot search engine were used.

The percentage of 30.5 % annotated spectra lies in the typically reported range of 25 % for tandem mass spectra. The relatively low percentage is caused either by poor quality of the spectra according to contamination of the sample or experimental noise. Another explanation for poor mass spectra assignment is searching with the incorrect sequence for the protein in the database. Protein sequence errors in the databases are caused generally by post translational modifications, polymorphisms or splice variants ^[254]. Then the peptide fragments can not be interpreted correctly by the search engine.

The utilized Mascot version (2.2.2.) offers the automatic search in a decoy database. After browsing the database containing the protein sequences, a decoy database is applied for estimation of the false-positive identification rate. This decoy database is randomized which means that for the tested peptide or protein sequence a random sequence is generated with the same length of the sequences and probability of

average amino acid composition [255]. In that way the decoy database is built up and the containing sequences are browsed after generating the results for the “real” database. As given in Table 27 a false positives rate of 1.2 % is observed which is equivalent to the setting of a 99 % probability for correct protein identification. The peptide mass tolerance which was used to perform Mascot search was set to 50 ppm. The average obtained mass accuracy resulted in 17 ppm for MS mode (Table 27). By decreasing the value for mass tolerance in the search settings the average mass accuracy value would drop, too. The mass spectrometer was specified for 50 ppm employing default calibration and 5 ppm for internal calibration according to the manufacturer. Internal calibration was performed using glu¹-fib in a concentration of 15 fmol μL^{-1} resulting in a concentration of 3.9 fmol spot⁻¹. However, this calibration peptide could not be recorded in every spectrum. Especially in spectra with high peptide signal intensity glu¹-fib was subjected to peak suppression and hence was not detectable. For those spectra the default calibration was automatically applied resulting in a lower mass accuracy (50 ppm instead of 5 ppm). The Mowse Score developed with Mascot search algorithm is a scale for quality estimation of the identified peptides and proteins which means higher scores characterize more confident identification. The average Mowse Score for successful protein identification for this analysis was computed at 32. Table 27 shows 70.3 as average peptide score whereas 25.2 % of the peptides were identified with a score above 90 and 22.2 % with a score below 40.

On the pie chart of Figure 43 the percental parts of peptides per protein for the 1,642 identified proteins are depicted.

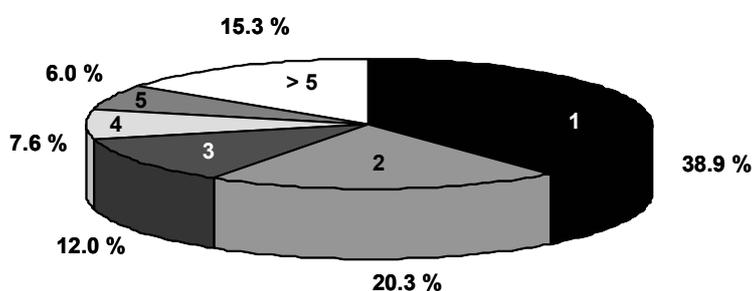


Figure 43. Peptides per protein of the 1,642 with the semi top-down approach in triplicate fraction analysis identified proteins.

3.3 Comparison of the semi top-down and the bottom-up approach

3.3.1 Overlap of peptide and protein identifications

Both methods of proteome analysis described above are based on two-dimensional chromatography and mass spectrometry. Whereas the bottom-up or shotgun approach is the most common method for proteome analysis utilizing chromatography and mass spectrometry, the semi top-down approach as shown here was newly established. In the following, a comparison of both methods is described to elaborate the advantages and disadvantages for both methods. To set both approaches in appropriate relation, for the semi top-down approach only two of the three technical replicates of the second separation dimension were considered because only two technical replicates were available for the classical bottom-up approach. This results in 920 instead of 1,004 protein hits with more than one peptide compared to 1,245 identified proteins obtained from the bottom-up approach. As depicted in Figure 44, 533 proteins were identified in both approaches.

There are several possible explanations for the higher number of protein identification employing the bottom-up approach: In the semi top-down approach intact proteins are separated in the first dimension which can cause problems of solubility as they are much more complex in their structure than peptides. Furthermore large proteins can adsorb irreversibly at the stationary phase of the separation column which causes further protein losses. Another aspect to consider is the protein loss occurring during transfer of the analytes from the first- to the second dimension. Proteins can adsorb at the wall of the Eppendorf vials more probable than peptides during shaking and are then not available for the tryptic digest.

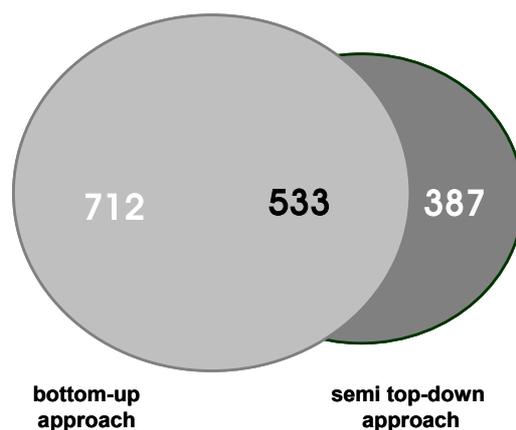


Figure 44. Venn diagram showing the the overlapping of both approaches. Identified proteins of the human *Glioblastoma multiforme* tissue with more than one peptide are considered.

Altogether 1,632 proteins of the human *Glioblastoma multiforme* tumor tissue were identified with more than one peptide. In the overlapping area of the Venn diagram 43 % of the proteins identified with bottom-up and 58 % of the proteins identified with semi top-down appeared. Hence, only about half of the identified proteins for one method were also identified with the other one. This may be caused inter alia by the extreme heterogeneity of the studied *Glioblastoma multiforme* tissue ^[256] since two different slices of the same tumor were used for the analysis.

3.3.2 Coverage of mass and pI range

An advantage of chromatographic methods in proteome analysis in comparison to gel-based approaches is the higher mass and pI range coverage of the identified proteins ^[3;257;258].

To investigate the covered mass and pI range for both chromatographic approaches, the logarithmic theoretical computed molecular protein mass was plotted as a function of the theoretical calculated pI value for each protein identified with more than one peptide (Figure 45). A wide pH and mass range could be covered with both methods as shown in Figure 45a+b. For the bottom-up approach 30.7 % proteins with a theoretical pI>7.5 were observed and for the semi top-down approach 26.5 %, respectively. As visible in Figure 45c, a similar distribution of pI and mass range as

for the single methods is obtained for the overlapping proteins. Here 26.5 % basic proteins with a pI value above 7.5 were observed. A large number of high molecular proteins were identified, for the bottom-up approach 18.3 % with a molecular weight >100 kDa, for the semi top-down approach 11.4 % and for the proteins identified in both approaches 12.4 %, respectively. Small proteins were also identified (semi top-down approach 13.6 % < 20 kDa, bottom-up approach 9.0 % and intersection 10.7 %, respectively). The inset at pH 7.5 is observable because this is the physiological pH of body fluids whereas the proteins have to be soluble.

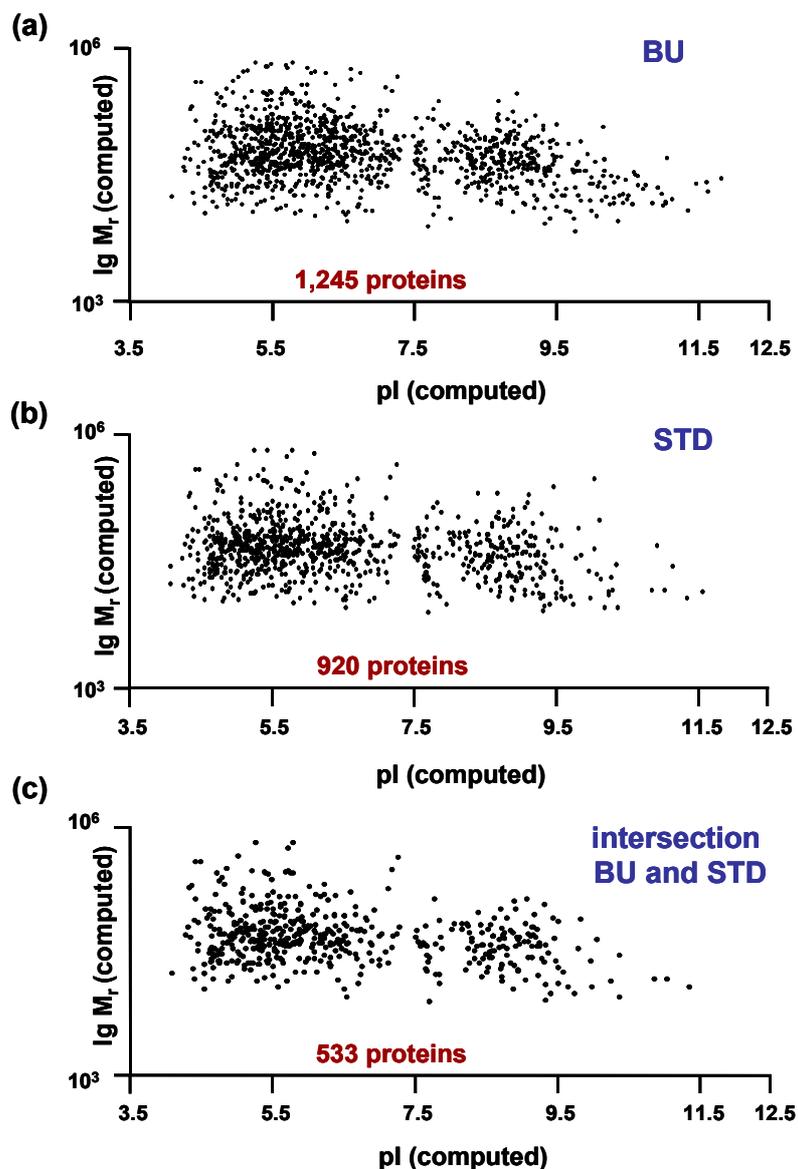


Figure 45. LgM_r/pI plots of the identified proteins with at least two peptides of (a) bottom-up, (b) semi top-down approach and (c) proteins identified with both methods. BU = bottom-up; STD = semi top-down.

In Figure 46a and Figure 46b the percentage of the theoretical computed pI and mass of the proteins is depicted. As far as the semi top-down approach is concerned slightly more acidic proteins were identified, consequently the bottom-up approach provided more basic ones (Figure 46a). A trend was also observed for the small proteins (below 70 kDa) in Figure 46b. Smaller proteins are more presented by the semi top-down approach whereas with the bottom-up method more high-molecular proteins are observable. All in all the small acidic proteins are higher represented in the semi top-down approach and the high molecular basic proteins by utilizing the bottom-up method.

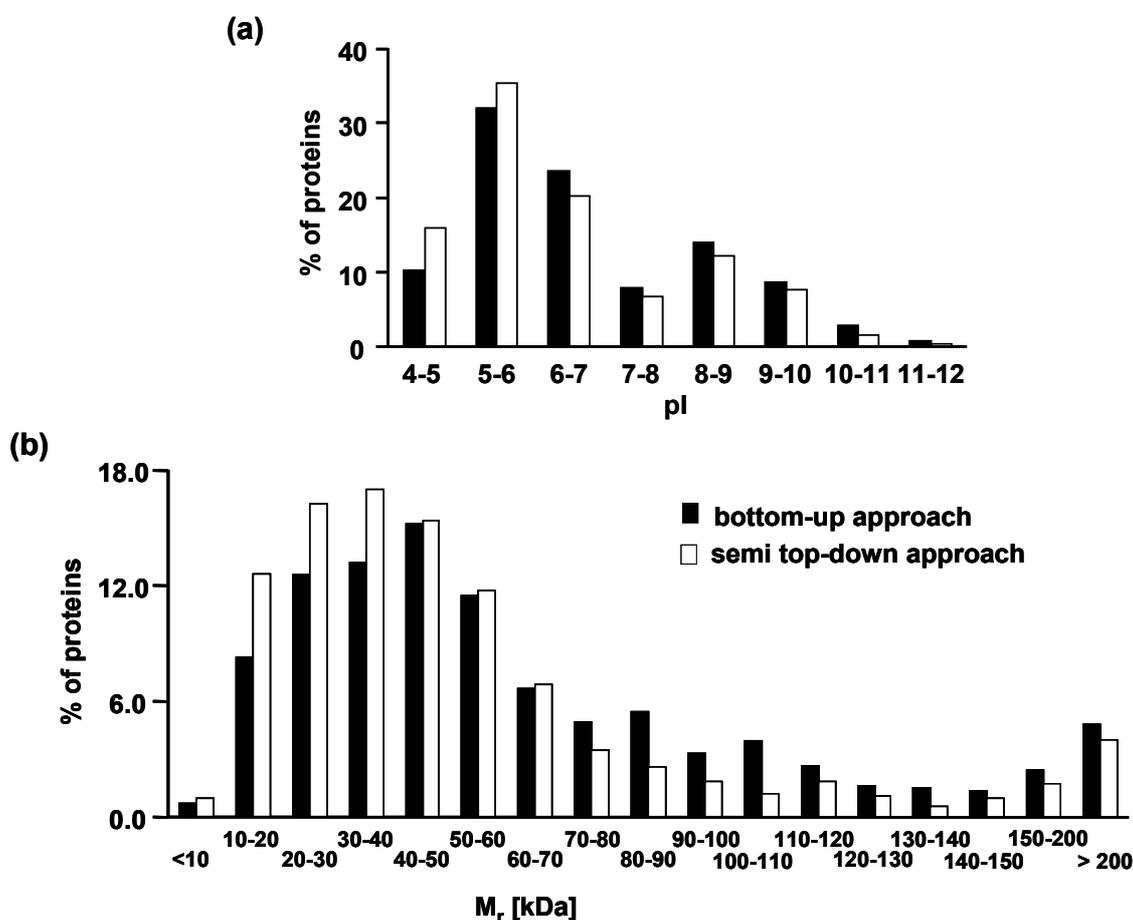


Figure 46. Comparison of the semi top-down and the bottom-up approach according to (a) the theoretical calculated pI for the proteins and (b) the computed molecular weight of the intact proteins.

3.3.3 Characterization of retention behavior

3.3.3.1 First dimension separation of peptides or proteins

In strong cation-exchange chromatography peptides should elute as function of their net charge, and their separation is influenced by the charge distribution. In order to retain a peptide on a cation-exchange column, the pH of the eluent needs to be lower than the isoelectric point of the peptide. A pH gradient was applied in SCX-HPLC additional to a gradient of ionic strength to perform peptide separation in the bottom-up approach. Furthermore, 25 % acetonitrile were added to both eluents to suppress unspecific secondary hydrophobic interactions of the peptides with the stationary phase. The range of the calculated theoretical pI values for the peptides (computed with MW/pI tool on www.expasy.org) eluting in the different fractions of the first separation dimension is shown in Figure 47a. The values were arithmetically averaged for every individual fraction. Only unique peptides were included for one fraction. There is no linear correlation observable because of the isocratic step at the beginning of elution. However, the computed pI of the eluting peptides increases with the ionic strength during gradient elution (fraction 10 to 33). A similar effect was recognized before in a proteome study of *C. glutamicum* [259].

The retention of proteins in the first dimension for the semi top-down approach is achieved by solvophobic interactions between the hydrophobic stationary phase and the hydrophobic moieties of the proteins as well as electrostatic effects. Thus, the hydrophobicity of the eluted proteins should rise with increasing fraction number. To investigate the retention behavior the Grand Average of Hydropathy (GRAVY) Index was calculated [260;261] for every protein identified in the 28 fractions (Protein GRAVY tool on www.bioinformatics.org) and was then arithmetically averaged. Figure 47b shows the distribution of the computed average GRAVY Indices as a function of the fraction number in the first dimension.

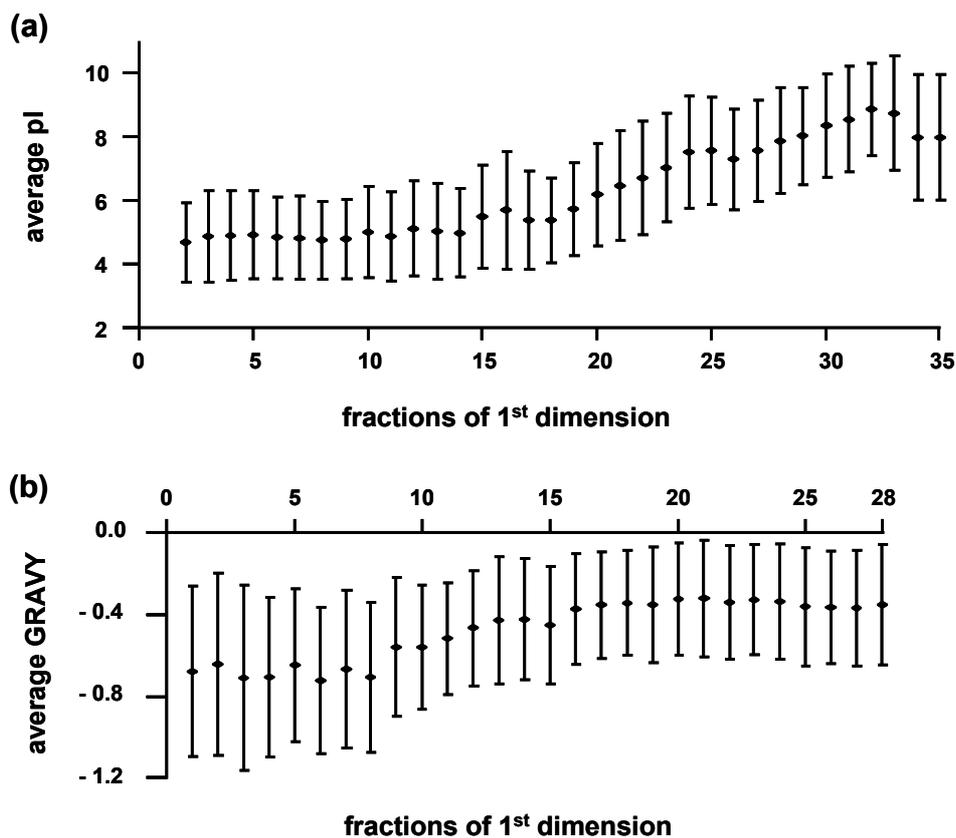


Figure 47. Retention behavior of first-dimension analytes. (a) pI as a function of peptide retention in the first dimension. (b) GRAVY Index as a function of protein retention in first dimension.

In the first eight fractions the GRAVY Indices range between -0.7 and -0.6. From fraction 9 to 20 a slight increase of GRAVY values was observed. This means, as expected, that more hydrophobic proteins elute in later fractions. However, the dependency was not highly significant. Potentially, this indicates that the hydrophobicity of a protein is not sufficiently described by calculating the GRAVY Index from the primary structure. Moreover, electrostatic interactions also have to be considered due to the effect of the ion-pair modifier (0.05 % TFA) which was added to both eluents. In the case of working under acidic conditions with 0.05 % TFA at pH 2.1 the alkaline amino-acid residues histidine, arginine, lysine and the N-termini are protonated. This means that not only hydrophobic amino-acid residues can interact with the stationary phase, but also charged moieties leading to an increase of selectivity for the protein separation ^[262]. These secondary interactions are favored (in contrast to unspecific solvophobic interactions in ion-exchange HPLC) because they abet multi-point adsorption of the proteins.

3.3.3.2 Second dimension separation of peptides

For both chromatographic approaches the same nano-flow HPLC system with the same column dimensions and equal gradient (0-30 % acetonitrile in 0.05 % aqueous TFA within 50 min) was used. In the second dimension the retention behavior for both methods is characterized by computing the GRAVY Indices for the individual peptides. The principle of separation is the same as for proteins eluting in the first dimension of the semi top-down approach. Solvophobic interactions and electrostatic effects contribute to the elution of the peptides. To investigate the correlation of hydrophobicity and retention time one protein from fraction 9 of the semi top-down approach, which is represented by six peptides, was chosen (Figure 48). The most hydrophilic peptide HQTNSELK elutes first, the most hydrophobic peptide LMDGSEIFSLLESAR last and the moderate hydrophobic peptide VVQETVLVEER between them, as expected. However, the peptides of moderate hydrophobicity do not follow this concept strictly. For peptide QEEQSAAIHISSETLEQKPHFESSTVK a GRAVY Index of -0.96 was computed which is more hydrophilic than the peptide KPTEFIGGVTSTSQSWVQK (GRAVY Index -0.55), but the former peptide elutes later.

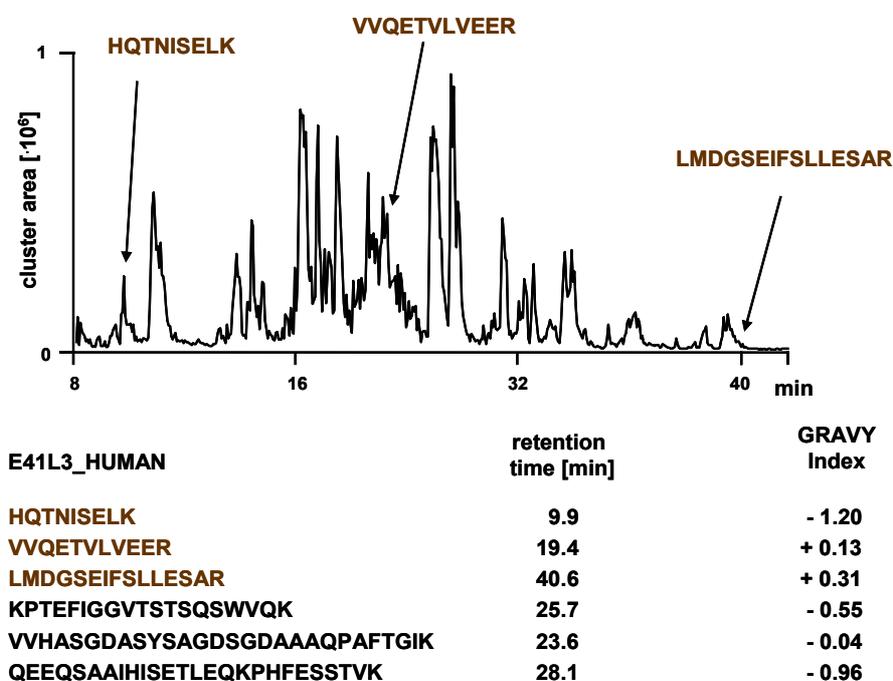
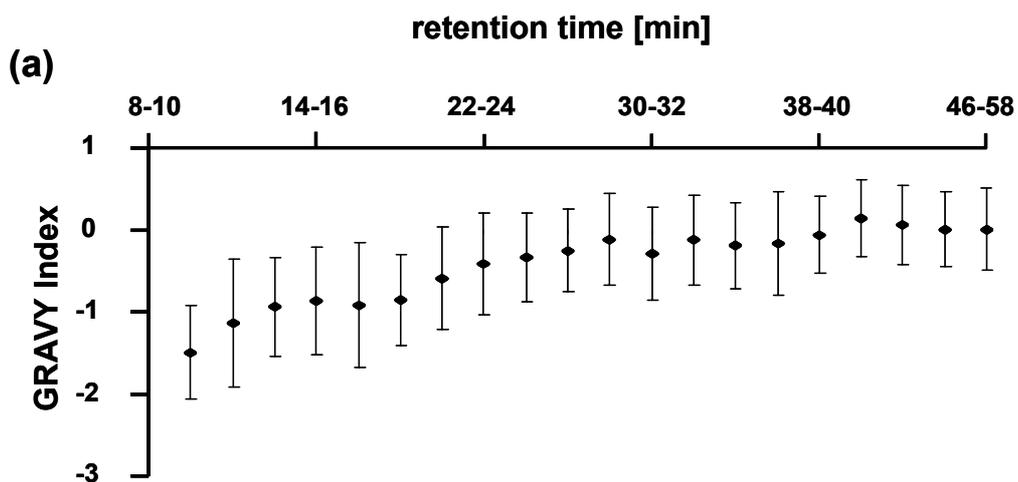


Figure 48. Elution of six peptides of the human protein E41L3 (band 4.1-like protein 3) in fraction 9 of the semi top-down approach.

The ionic interactions which also contribute to peptide retention prevent a linear increase of the GRAVY Index as a function of retention time. This is illustrated further by plotting the average GRAVY Index of all unique peptides of fraction 9 as a function of their retention times in the second dimension (Figure 49a). The indices are summarized to two-minute sets of peptides to emphasize the slightly increase of GRAVY Indices with the retention time. However, no linear correlation is observable. The separated peptides in the second dimension are more likely to be hydrophobic than the intact proteins separated in the first dimension of the semi top-down approach. In the aqueous medium surrounding the proteins in the cell (and later in the vial), the majority of the hydrophobic moieties are situated in the inner part of the protein whereas at the surface mainly hydrophilic residues are positioned. When proteins are digested, hydrophobic as well as hydrophilic peptides are present in the digest. The cumulative appearance of hydrophilic or hydrophobic amino acids in a short sequence (= peptide) is more probable while in proteins the GRAVY Index is averaged for plenty of amino acids. Actually, higher GRAVY Indices were observed for the peptides than for the proteins. In Figure 49b GRAVY Indices up to a value of 1.88 are obtained for the peptides whereas the highest computed GRAVY Index for a protein in fraction 9 is 0.209.



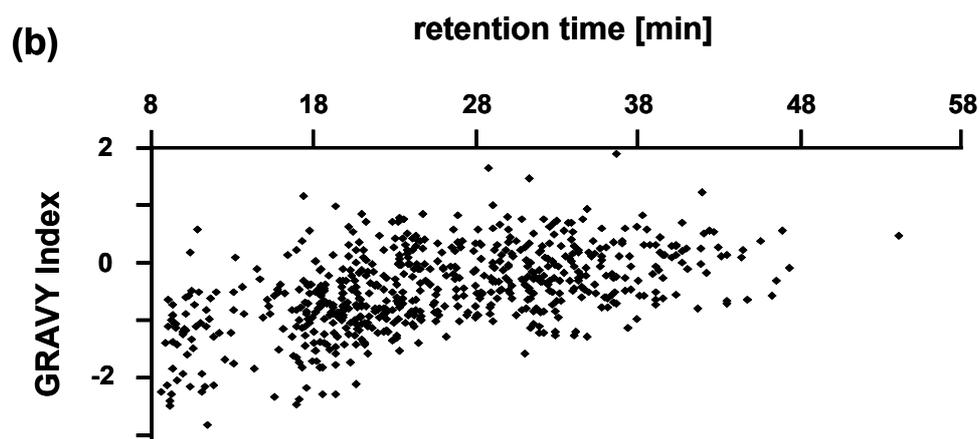


Figure 49. GRAVY Indices for unique peptides in fraction 09 of the semi top-down approach. (a) Peptides were summarized every two minutes and (b) all peptides of fraction 09 were plotted.

In Figure 49a the summarized GRAVY Indices emphasize the trend of increasing hydrophobicity of the peptides with the concentration of organic solvent, and the GRAVY Indices of the unique peptides are revealed in Figure 49b to show their individual distribution.

3.3.4 Sequence coverage

For the semi top-down and the bottom-up approach the sequence coverage of the individual proteins was examined. Only proteins identified with more than one peptide were considered for this study. The sequence coverage of a protein is computed by multiplying the ratio of identified and possible amino-acid residues with 100 %. Referring to Figure 50a for the bottom-up approach 10.6 % more proteins with a sequence coverage of less than 10 % were found. This induces a 3.2 % higher protein coverage for proteins identified with more than 40 % sequence coverage for the semi top-down approach in comparison to the bottom-up method (Figure 50).

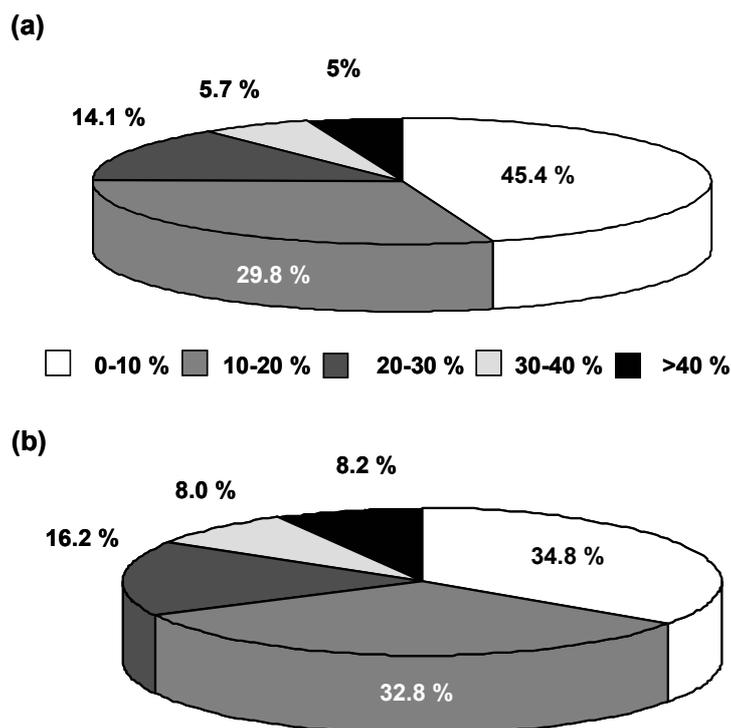


Figure 50. Sequence coverage for the identified proteins of (a) the bottom-up approach and (b) the semi top-down approach.

To get a more accurate comparison, not only the sequence coverage for individual proteins was ascertained but also the average sequence coverage of the identified proteins to prove the higher sequence coverage for the established semi top-down approach. For this, the sum of all amino acids listed in the database protein sequences of the identified proteins was set to 100 %. Then the number of amino acids of the identified peptides was summed and the percentage of them was computed to the total number of amino acids in the identified peptides. Again, only proteins identified with more than one peptide were included in this computation. For the semi top-down approach an average sequence coverage of 11.4 % and for the bottom-up approach of 9.9 % were obtained. This means that 11.4 % (and 9.9 %, respectively) of the possibly identifiable amino acids of the observed proteins (without one-hit wonders) were found. Still, the semi top-down approach offers a slightly increased average sequence coverage of 1.5 %.

Furthermore, the distribution of the peptide length was considered (Figure 51). For both methods the maximum of the obtained distribution is 11 amino acids per peptide. For the semi top-down approach more peptides with higher amino acid

numbers were identified resulting in the slightly higher sequence coverage for the identified proteins.

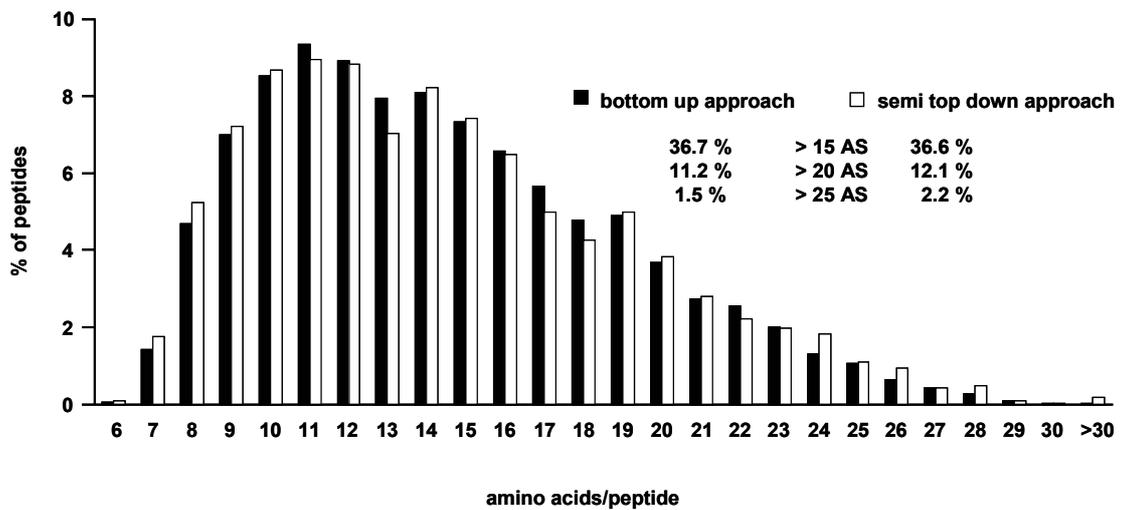


Figure 51. Distribution of peptide length for the semi top-down and the bottom-up approach.

3.3.5 Dimension orthogonality and peak capacity

It is important for a two-dimensional chromatographic separation scheme that both dimensions follow orthogonal separation mechanisms. This means that almost the entire separation space should be covered. The orthogonality of a two-dimensional separation system for proteome analysis can be proven by plotting each hit as a function of the retention times for both dimensions. The abscissa represents the retention time in the first dimension whereas the ordinate shows the retention time in the second dimension. Only non-redundant peptides within one fraction were included. Proteins were identified in 34 fractions (fraction 2 to 35) of the bottom-up approach and in 28 fractions of the first dimension of the semi top-down approach. Therefore, in Figure 52a 70 min are plotted and in Figure 52b 56 min, respectively.

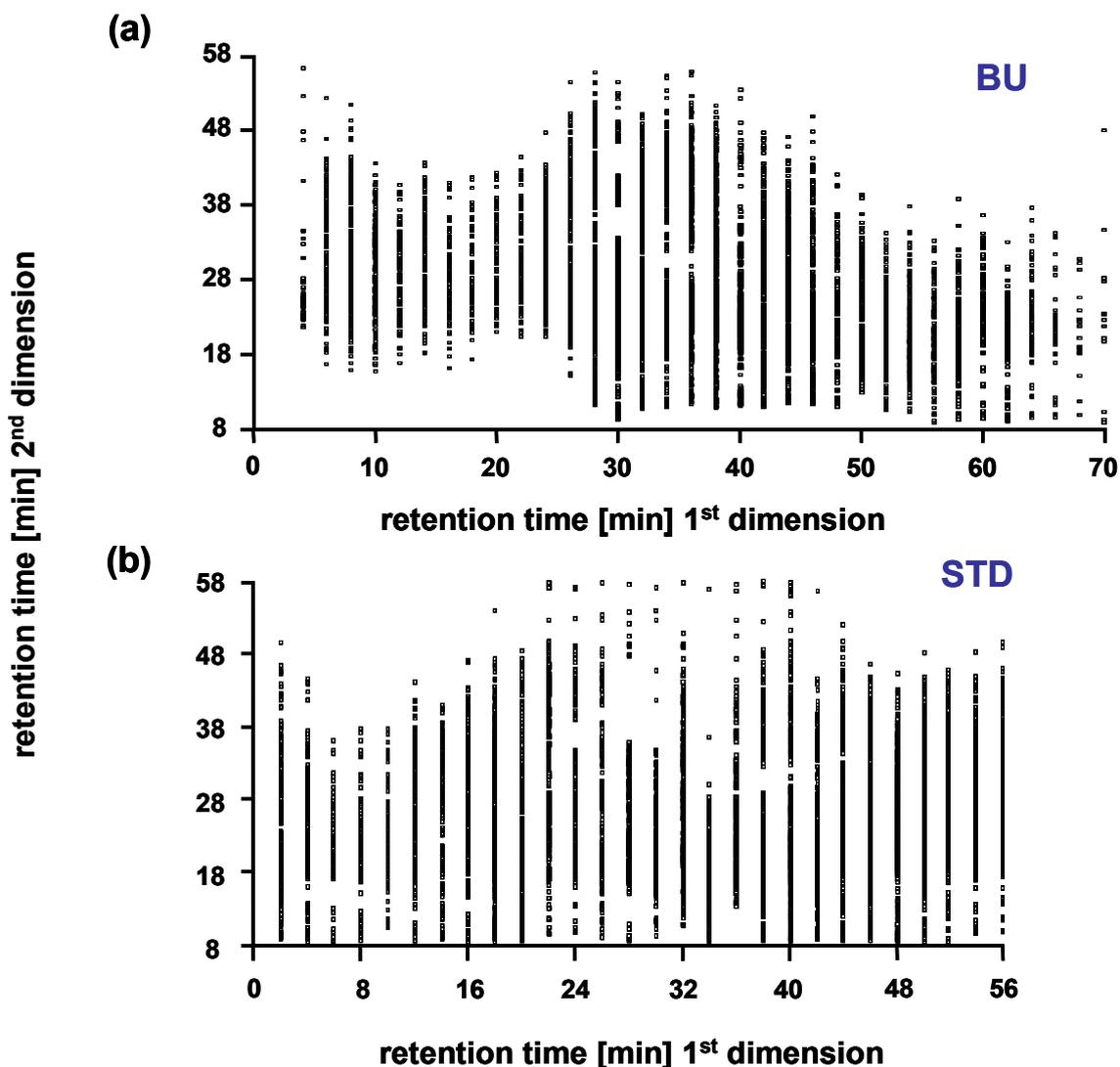


Figure 52. Orthogonality of separation dimensions in both two-dimensional separation schemes tested. The retention time of the second dimension is plotted as a function of the retention time of the first dimension for (a) the bottom-up approach (= BU) and (b) the semi top-down approach (= STD).

As seen in Figure 52, nearly the whole separation space for both methods is covered. It seems that for the semi top-down approach the coverage is even higher. For this method intact proteins were separated in the first dimension and peptides in the second dimension. This implies that one data point in the first dimension (= one protein) results in several data points in the second dimension when the protein is identified by more than one peptide. This is why for the semi top-down approach the term “pseudo”-orthogonality is more suitable. In contrast, for the bottom-up approach peptides are separated in both dimensions. As a result one data point in the first

dimension corresponds to one data point in the second dimension. In Figure 52b the charge-dependent distribution of the eluted peptides is visible. Peptides are quasi not individually separated but clusters of peptides with similar charges (+1, +2, +3, +4) are separated in SCX-HPLC.

A further important parameter to characterize a two-dimensional separation system is the theoretical peak capacity (P) (Chapter I, paragraph 3.1.2). The total peak capacity of a multi-dimensional chromatographic system is estimated by multiplying the peak capacities from the single dimensions ^[263;264]. The peak capacities of both tested approaches have been computed as shown in Table 28. Here, the peak capacities for the first dimension were not computed but determined by the number of fractions in which proteins were identified.

Table 28. Theoretical peak capacity for both two-dimensional chromatographic separation systems. BU = bottom-up, STD = semi top-down, t_G = gradient time.

approach	t_G	average w_b	P
1st dimension SCX of peptides = BU (a)	50		34
1st dimension IP-RP of intact proteins = STD (b)	60		28
2nd dimension IP-RP peptides	50	0.26	190
2D-HPLC (a)			6,460
2D-HPLC (b)			5,320

The peak capacities for both methods of 6,460 and 5,320, respectively, lay within the range of previously described peak capacities for 2D-HPLC systems ^[224].

The second dimension peak capacity was calculated using the average peak width at 4σ (=13.4 % of peak height) for four standard peptides at the same 50-min gradient as for the second dimension of the proteome analysis (0-30 % acetonitrile in 0.05 % aqueous TFA). A 17.5 % higher peak capacity was accomplished using the classical bottom-up approach, resulting in an increased protein identification of 26 % (1,245 vs. 920 proteins). By comprising three replicates for the semi top-down method, 19.4 % more proteins were identified with the bottom-up method (1,240 vs. 1,004).

A higher peak capacity of the 2D-HPLC system was achieved for the bottom-up approach because in more fractions of the first dimension proteins were identified (34 vs. 28). The number of protein- and peptide containing fractions is determined by the gradient time and the time interval for fractionation as well as the final fraction volume (after evaporation to dryness and dissolving) due to dilution of the analytes. To increase the peak capacity for the semi top-down approach the following experimental changes or a combination of them could be applied:

1. increased gradient time
2. shorter collection interval for one fraction
3. dissolving the analytes in less solvent volume

Aspect 1. and 2. can only be applied at the cost of time consumed for the proteome analysis. The last point however, could be employed with the risk of precipitation of proteins. Furthermore, the replicate injection is limited due to reduced sample volume.

3.4 Biological relevance of the identified proteins in the human brain tumor tissue

3.4.1 Validation of the obtained protein identifications

For further validation of Mascot search results the software Scaffold 2.0 was applied to the obtained protein identifications. Peptide Prophet and Protein Prophet are the underlying validation algorithms. Additionally X!tandem search engine was run to obtain extra information.

By validating the Mascot result files with Scaffold for the semi top-down approach 802 proteins could be identified instead of 1,004 and for the bottom-up approach 1,065 instead of 1,245. The reduction in identifications is caused by the sorting principle of Protein Prophet. Proteins that contained the identical set of identified peptides could not be differentiated based on MS/MS analysis alone and were grouped utilizing the expectation-maximization algorithm. Then one group with an identical subset of identified peptides is counted as one hit.

Figure 53 demonstrates exemplary for the semi top-down approach the necessity of

eliminating proteins identified by only one peptide. Even with a peptide probability of 95 % the protein probability does not rise above 78 %, whereas for two identified peptides per protein already with 95 % peptide probability a protein probability of almost 100 % can be achieved. Proteins identified with five peptides need only 65 % peptide probability to achieve close to 100 % protein probability.

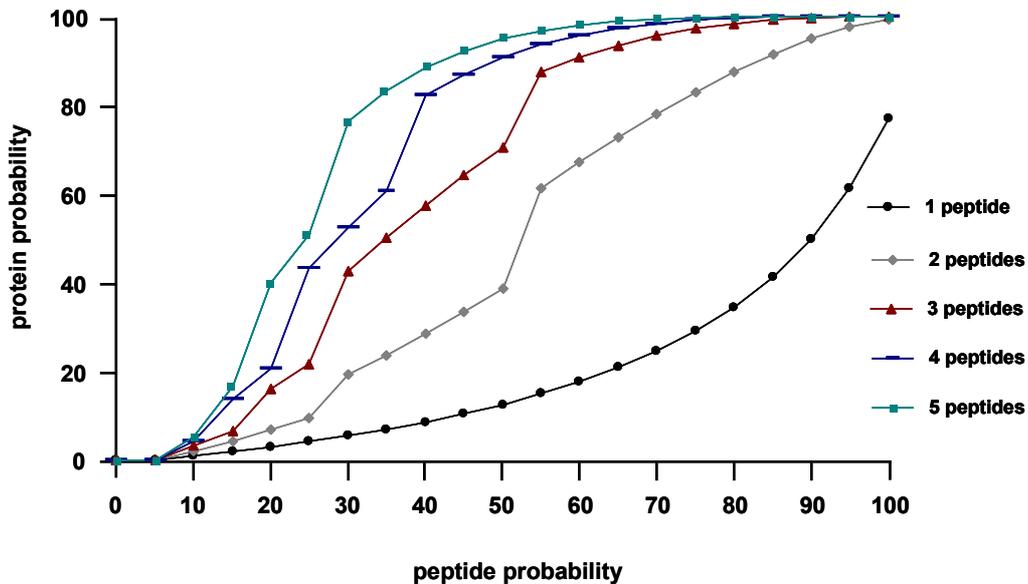


Figure 53. Protein probability plotted as function of peptide probability for the semi top-down approach.

The distribution of the discriminant scores computed for the peptides of the semi top-down approach are shown in Figure 54. The black distribution shows the incorrect assignments and the white one the correctly identified peptides with a probability equal to or higher than 95 %. In the overlapping area the probability of a false assignment declines with the increase of the discriminant score. Peptides which are identified with a discriminant score above 2.8 in this case are counted as positive identification.

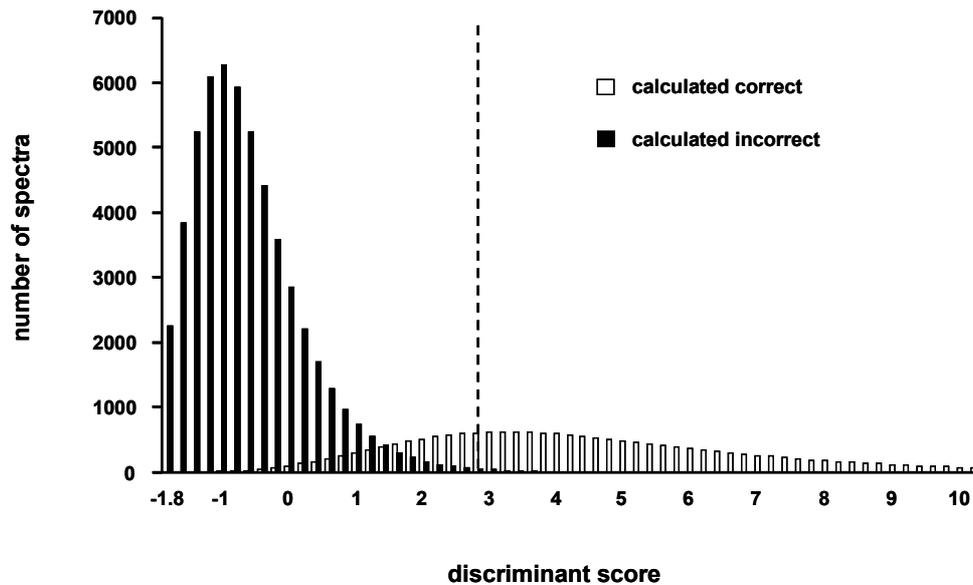


Figure 54. Histogram of discriminant scores of the identified peptides for the pool of both approaches.

By pooling the results of both chromatographic methods 1,401 proteins were identified with more than one peptide instead of 1,670 with Mascot only. In this case all three replicates for the semi top-down approach were included because now the total proteins of the investigated tumor tissue are of interest and not only those suited for a comparison between the methods. Thus, 466 proteins were identified with both methods which is equal to 44 % of the bottom-up identifications and 58 % of the semi top-down identifications, respectively (Figure 55).

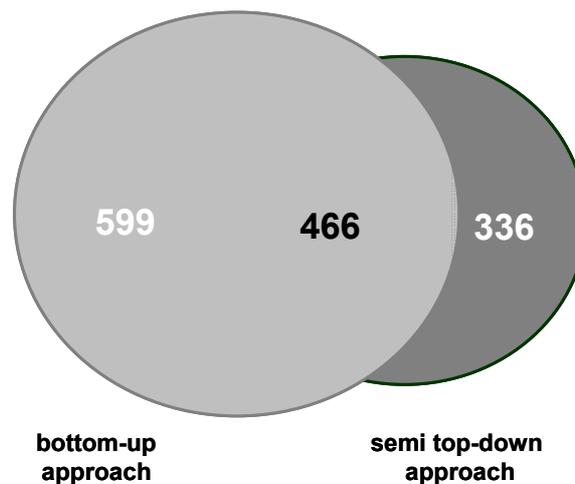


Figure 55. Venn diagram of the semi top-down and bottom-up identifications validated with the Scaffold proteome software.

3.4.2 Molecular function of the identified proteins

The molecular function of the 1,401 protein hits confirmed with the Scaffold proteome software was disposed using the PANTHER classification system. Dissolving the Scaffold grouping of proteins including the same subset of identified peptides resulted in 1,429 protein hits which were uploaded to PANTHER. To these proteins 1,761 molecular functions were allotted. For 186 hits the molecular function was unknown (10.6 %). These were eliminated as well as the 85 hits with miscellaneous function (4.8 %). The distribution of the remaining 1,490 molecular function hits is depicted in Figure 56.

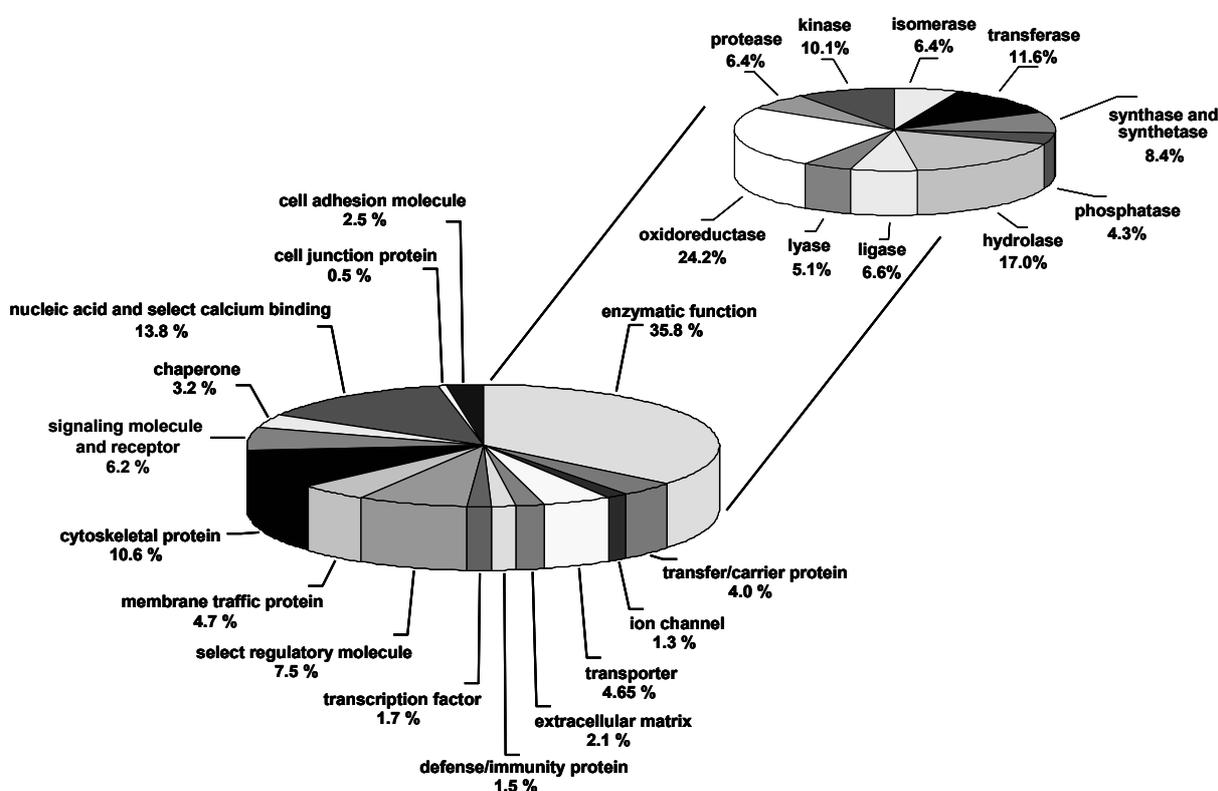


Figure 56. Molecular function of the 1, 429 proteins identified in the pool of the semi top-down and the bottom-up approach (unknown and miscellaneous function excluded).

More than one third (35.8 %) of the molecular function hits were assigned to enzymes. These are shown in Figure 56 as an extra chart. Altogether 10 different classes of enzymes could be identified in which the oxidoreductases with 24.2 % are most highly represented.

3.4.3 Protein retention behavior in the first separation dimension

To characterize the retention pattern of the intact proteins in the first dimension of the semi top-down approach, the subset of identified membrane proteins was chosen. Identifying high amounts of membrane proteins normally demands detergent-based methods including a strong denaturant applied for dissolving the membrane proteins. However, the main drawback of those methods is the necessity to remove the detergents prior to mass spectrometric analysis which comes along with sample losses. To prevent this effect TFE was added in the sample preparation method instead of detergents. This organic co-solvent enhances protein solubility and contributes to protein denaturation. It easily evaporates during the sample concentration process; hence no cleanup step is required prior to HPLC-MS/MS analysis. Furthermore a high concentration of TFE (40-50 % v/v) probably decreases the dielectric constant of the solubilization medium resulting in superior protein extraction performance [265]. In this study the focus was not on the identification of membrane proteins strictly. However, due to their poor solubility, membrane proteins are usually underrepresented, especially in top-down approaches. To investigate these proteins exclusively, sedimentation or partition techniques are required to separate membrane components from the cytosolic bulk [266].

The retention behavior of the proteins in the first dimension is discussed with the membrane protein subset of the identified proteins. Altogether, 290 membrane proteins were identified for the pool of both chromatographic methods, of which 236 were confirmed with Scaffold. Here, the retention pattern of the intact proteins in the first dimension is discussed. Thus, only the semi top-down results are taken in account. For this method 142 intrinsic membrane proteins were identified utilizing the Mascot search engine with the parameters given in the experimental section. All of them were integral, no anchored membrane proteins were observed. Of the 142 membrane proteins identified with Mascot, 104 proteins were confirmed with the Scaffold proteome software. Forty-two percent of the 104 proteins eluted in one or two fractions and 58 % were distributed over more than two fractions. Figure 57 shows the retention behavior of six selected membrane proteins on which the retention pattern will be discussed. The retention behavior of all membrane proteins matched with the semi top-down approach is listed in the appendix.

fraction and ends in the next one. Though protein 4 with 586 amino acids (62.3 kDa) is also a moderate protein in size as well as protein 3, it is distributed over two isolated fractions. It is very likely that the heavy and the light chain of this heterodimer were separated or the protein was degraded by proteases of the cell prior to fractionation.

Two exemplary proteins were also chosen for proteins which are dispersed over more than two fractions. Protein 1, a protein of medium size (592 amino acids and 67.6 kDa), is distributed over seven consecutive fractions. Protein 2 however, a large protein composed of 1,347 amino acids resulting in a molecular mass of 150 kDa, is distributed over nine fractions, not all contiguous. There are some reasonable explanations for this behavior. At first the proteins 1 and 2 are highly abundant as can be estimated from their spectrum count. For protein 1 a spectrum count of 118 was computed and 89 for protein 2, respectively. As for only 9.9 % of the 802 proteins identified with the semi top-down approach a spectrum count >80 was computed these proteins are highly abundant ones. For such highly abundant proteins the separation column may have been overloaded and therefore the proteins spread over different fractions. This could explain why the proteins elute over more than two fractions. The distribution over isolated fractions could be explained as before for protein 4 with rest proteases activity. Furthermore, the very high sequence homology of many human proteins is an important factor. Especially polymorphisms, alternative splicing and post-translational modifications are increasing the complexity of the sample. It is estimated that 40 to 60 % of the human genes are alternatively spliced ^[267]. In fact, for protein 2 thirteen isoforms are known (www.expasy.org). It is possible that by identifying one protein in several isolated or even adjacent fractions rather different isoforms with the same subset of peptides are identified. The presence of highly abundant, degraded and highly homologous proteins likely explains the distribution over more than two fractions of the first chromatographic separation dimension.

3.4.4 Confirmation of previously identified potential biomarkers

Serological identification of antigens by recombinant expression screening (SEREX) was developed to identify human tumor antigens. In this approach molecular cloning

techniques are implemented into serological analysis ^[268]. By applying SEREX to pooled *Glioblastoma multiforme* patient's sera 13 potential antigens were obtained ^[205]. Now it should be proven, if these proteins indirectly identified in the patient serum based on antibodies directed against them, are also present in molecular form in the *Glioblastoma multiforme* tissue. Table 29 shows the 13 antigens indirectly identified with SEREX and those which were confirmed by chromatography- and mass spectrometry-based proteome analysis.

Table 29. Detection for five of 13 potential antigens for *Glioblastoma multiforme* directly identified with 2D-HPLC-MALDI-MS/MS. BU = bottom-up; STD = semi top-down. For GO annotations see appendix.

antigen	description	detected with	
		STD	BU
ACTN4	actinin, alpha 4		x
ING4	inhibitor of growth family, member 4		
RTN4	reticulon 4	x	x
CLIP2	cytoplasmic linker 2	x	
HCLS1	hematopoietic lineage cell-specific protein	x	
U2AFL	U2(RNU2) small nuclear RNA auxiliary factor 1-like 1		
ZN232	zinc finger protein 232		
HS105	heat-shock protein 105 kDa	x	x
PHF3	PHD finger protein 3		
PHF20	PHD finger protein 20		
TPR	translocated promoter region (to activated MET oncogene)		
CYTSA	cytospin-A (NY-REN-22 antigen)		
GOGA1	golgi autoantigen, golgin subfamily a		

According to their spectrum count ACTN4, RTN4 and HCLS are relatively high abundant proteins whereas CLIP2 and HS105 are represented by a low spectrum count. ACTN4 also was identified by Mascot in the semi top-down approach, but categorized to ACTN1 in Scaffold, following the expectation-maximization algorithm.

In Table 30, the distribution of the identified SEREX peptides is displayed. Only the proteins confirmed with the Scaffold software are shown in the table. The annotation of peptide distribution was performed with Mascot. The protein HS105 in the table is represented by just one peptide in Mascot. Caused by this fact it has to be excluded from the list following strictly proteomics identification rules or at least discussed critical [269].

Table 30. Distribution of peptides observed for SEREX proteins. White numbers belong to bottom-up and black numbers to semi top-down results.

fraction	1	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	25	27	28	31	unique peptides	sequence coverage	
protein	identified peptides																								
RTN4		1	1	1				1	1	1	1	1			1	1	1	1		1			5	5.7	
	4					2	2																6	9.7	
HS105															1								1	1.7	
																2	1	1					3	4.5	
ACTN4													4	4	3	2				1		2	1	14	23.1
CLIP2							1	1															2	3.3	
HCLS1	1		1	1	4		1																6	22.2	

As Table 30 demonstrates clearly, the advantage of the semi top-down method rests within the possibility to focus on potential tumor markers found in selected fractions of intact protein separation without the need to analyze the entire set of fractions again. Moreover, identified tumor antigens may be isolated micropreparatively for further investigations.

The narrower retention window for one protein applying the semi top-down approach becomes especially clear by looking at the protein reticulon 4 (RTN4). This potential biomarker is identified by both methods. Its observed peptides are distributed over 13 fractions employing the bottom-up approach while for the semi top-down method the peptides elute in just three fractions. It is very likely that isoforms are present. Actually four isoforms are known for reticulon 4 (www.expasy.org). This could explain why the peptides elute in isolated fractions identified with the semi top-down approach. Another reason could be a degradation of the protein prior to fractionation as discussed for the membrane proteins before.

The hematopoietic lineage cell-specific protein (HS105) was identified by three peptides for the semi top-down approach. These peptides were distributed over three

fractions. However, this protein was identified by only one peptide in the bottom-up approach and has to be excluded from the protein identification list as one-hit wonder for this method.

4 Summary

It was shown by comparison to the classical bottom-up approach that the developed semi top-down platform is suitable for proteome research. First, the classical bottom-up approach was validated using a human *Glioblastoma multiforme* tissue sample. The extracted proteome was tryptic digested and then separated using strong cation-exchange chromatography. Then the fractions were injected into the second dimension: the previously established off-line nano-flow IP-RP-HPLC MALDI-TOF/TOF platform. Here, peptides were flushed after injection onto a monolithic trap column for concentration and desalting and then in back-flush mode onto the monolithic separation column. The eluting analytes were mixed with α -cyano-4-hydroxycinnamic acid employed as matrix and spotted onto a stainless steel target automatically. The target subsequently was transferred off-line into the MALDI-TOF/TOF system and there the intact peptide mass as well as the mass of the peptide fragment ions were analyzed. With this method 1,245 proteins were identified using the Mascot search engine. By validation with the Scaffold proteome software, using Peptide- and Protein Prophet to filter peptide and protein hits, 1,065 proteins still passed acceptance.

This classical bottom-up method was compared to the new semi top-down approach. In this approach intact proteins were first separated using IP-RP-HPLC and then each fraction was digested and separated in the second dimension in the same chromatographic system as the classical bottom-up approach. We called this method semi top-down approach. Here, applying the Mascot search engine 1,004 protein hits for three and 920 for two replicates, respectively, were obtained. As a comparison with the bottom-up approach should be performed, only two replicates of the second separation dimension were considered first. Validation of this method with Scaffold resulted in a protein acceptance of 763 instead of 920 for two replicates and 802 instead of 1,004 hits for three replicates, respectively.

With both approaches 533 proteins of 1,632 total hits were identified (Mascot) which is not surprising due to the high heterogeneity of the used tumor tissue sample. With each method a broad pH (4-12) and mass range (6.4-628.7 kDa) could be covered. Interestingly, with the bottom-up approach more basic proteins were identified whereas a higher level of acidic proteins was detected with the semi top-down method. Moreover, the retention modes for the first and the second dimension of both

methods were discussed. A slight increase of the GRAVY Index with the retention time in IP-RP-HPLC was observed for the proteins identified with the semi top-down approach. The same was valid for peptides separated with IP-RP-HPLC with both methods in the second dimension. For the pI of the peptides separated with SCX-HPLC in the first dimension of the bottom-up approach also a slight increase was observed with the retention time.

Furthermore, to mirror the entire protein hits for this brain tumor, the results of both methods were pooled comprising three second-dimension replicates for the semi top-down approach. In this pool 1,401 proteins were confirmed with Scaffold proteome software. These proteins were classified using PANTHER software according to their functions. This software categorized the most highly part (35.8 %) of the proteins to be enzymes.

Finally, the set of identified tumor proteins was compared to a set of biomarkers previously discovered with SEREX. Of 13 with this indirect method detected biomarkers, five were identified as molecular proteins. A clear advantage was shown for the semi top-down approach in terms of focused protein elution in few or ideally one fraction. Thus, this new method is suitable for biomarker isolation circumventing the analysis of the whole set of fractions obtained from the first dimension in further experiments.

5 Conclusions

By applying the established proteomic platforms to a highly complex protein mixture (proteins of a human *Glioblastoma multiforme* tissue) a comprehensive protein overview could be obtained. Moreover, the presence of biomarkers which were identified indirectly utilizing SEREX was now confirmed directly.

Obviously, more proteins were identified with the classical bottom-up approach; nevertheless advantages of the alternative approach could be demonstrated:

1. The sequence coverage for proteins with > 40 % sequence coverage was 3.2 % higher for the semi top-down approach. As different proteins were identified for both methods the average sequence coverage was computed for a more adequate comparison. So all amino acids in the sequence of the identified proteins were summed for each method and set to 100 %. For the semi top-down approach a 1.5 % higher average sequence coverage was obtained.
2. The semi top-down approach offered a higher orthogonality than the bottom-up method, although it is more a “pseudo”-orthogonality because every data point of the first dimension results in several data points for the second dimension when the protein is identified with more than one peptide.
3. For further analyses there is no need to analyze the entire set of fractions from first dimension, but only those fractions which contain the demanded proteins.

The gradient time could be extended or the fraction collection time interval could be shortened for further optimization in terms of peak capacity for the semi top-down approach. Moreover, the final volume of the fractions after evaporating and refilling could be reduced. To prevent problems with the solubility of the proteins a detergent could be added which would be removed by the chromatographic separation.

Chapter V

Reference list

1. Wasinger, V. C.; Cordwell, S. J.; Cerpa-Poljak, A.; Yan, J. X.; Gooley, A. A.; Wilkins, M. R.; Duncan, M. W.; Harris, R.; Williams, K. L.; Humphery-Smith, I. *Electrophoresis* **1995**, *16*, 1090-4.
2. Lottspeich, F. *Angew. Chem. Int. Ed.* **1999**, *38*, 2476-92.
3. O'Farrell, P. H. *J. Biol. Chem.* **1975**, *250*, 4007-21.
4. Klose, J. *Humangenetik* **1975**, *26*, 231-43.
5. Gorg, A.; Weiss, W.; Dunn, M. J. *Proteomics* **2004**, *4*, 3665-85.
6. Hunt, D. F.; Yates III, J. R.; Shabanowitz, J.; Winston, S.; Hauer, C. *Proc. Natl. Acad. Sci. USA* **1986**, *83*, 6233-37.
7. Biemann, K.; Papayannopoulos, I. A. *Acc. Chem. Res.* **1994**, *27*, 370-78.
8. Borman S. ; Russell H. ; Siuzdak G. *Today's Chemist at Work* **2003**.
9. Biemann, K. *Annu. Rev. Biochem.* **1992**, *61*, 977-1010.
10. Dole, M.; Mack, L. L.; Hines, R. L.; Mobley, R. C.; Ferguson, L. D.; Alice, M. B. *J. Chem. Phys.* **1968**, *49*, 2240-49.
11. Mack, L. L.; Kralic P.; Rheude A.; Dole M. *J. Chem. Phys.* **1970**, *52*, 4977-86.
12. Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989**, *246*, 64-71.
13. Karas, M.; Bachmann, D.; Bahr, U.; Hillenkamp, F. *Int. J. Mass Spectrom. Ion Processes* **1987**, *78*, 53-68.
14. Pappin, D. J.; Hojrup, P.; Bleasby, A. J. *Curr. Biol.* **1993**, *3*, 327-32.
15. Mann, M.; Hojrup, P.; Roepstorff, P. *Biol. Mass Spectrom.* **1993**, *22*, 338-45.
16. Eng, J. K.; McCormack, A. L.; Yates, J. R. III *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 976-89.
17. Yates III, J. R.; Eng, J. K.; McCormack, A. L.; Schieltz, D. *Anal. Chem.* **1995**,

- 67, 1426-36.
18. Yates, J. R. 3rd; Eng, J. K.; McCormack, A. L. *Anal. Chem.* **1995**, *67*, 3202-10.
 19. Perkins, D. N.; Pappin, D. J.; Creasy, D. M.; Cottrell, J. S. *Electrophoresis* **1999**, *20*, 3551-67.
 20. Geer, L. Y.; Markey, S. P.; Kowalak, J. A.; Wagner, L.; Xu, M.; Maynard, D. M.; Yang, X.; Shi, W.; Bryant, S. H. *J. Proteome Res.* **2004**, *3*, 958-64.
 21. Clauser, K. R.; Baker, P.; Burlingame, A. L. *Anal. Chem.* **1999**, *71*, 2871-82.
 22. Wolters, D. A.; Washburn, M. P.; Yates, J. R. 3rd *Anal. Chem.* **2001**, *73*, 5683-90.
 23. Zhao, J.; Zhu, K.; Lubman, D. M.; Miller, F. R.; Shekhar, M. P.; Gerard, B.; Barder, T. J. *Proteomics* **2006**, *6*, 3847-61.
 24. Yan, F.; Subramanian, B.; Nakeff, A.; Barder, T. J.; Parus, S. J.; Lubman, D. M. *Anal. Chem.* **2003**, *75*, 2299-308.
 25. Sluyterman, L. A. AE.; Wijdenes, J. *J. Chromatogr.* **1978**, *150*, 31-44.
 26. Sluyterman, L. A. AE.; Elgersma, O. *J. Chromatogr.* **1978**, *150*, 17-30.
 27. Kim, H.; Lubman, D. M. *J. Chromatogr. A* **2008**, *1194*, 3-10.
 28. Andren, P. E.; Emmett, M. R.; Caprioli, R. M. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 867-69.
 29. Chen, H.; Horváth, C. *J. Chromatogr. A* **1995**, *705*, 3-20.
 30. Unger, K. K. *Packings and Stationary Phases in Chromatographic Techniques*, Unger, K. K., Ed.; Marcel Dekker: New York, 1990.
 31. Frohlich, T.; Arnold, G. J. *J. Neural Transm.* **2006**, *113*, 973-94.
 32. Aebersold, R.; Mann, M. *Nature* **2003**, *422*, 198-207.
 33. International Union of Pure and Applied Chemistry and International Union of Biochemistry *Pure and Appl. Chem.* **1983**, *56*, 595-625.

34. PAULING, L.; COREY, R. B. *Proc. Natl. Acad. Sci. U S A* **1951**, 37, 251-6.
35. PAULING, L.; COREY, R. B.; BRANSON, H. R. *Proc. Natl. Acad. Sci. U S A* **1951**, 37, 205-11.
36. Pauling, L.; Corey, R. B. *Proc. Natl. Acad. Sci. U S A* **1951**, 37, 729-40.
37. Borovikov, V.-A.; Vainshtein, B. K Gel'fand I. M.; Kalinin, D. I. *Kristallografiya* **1979**, 24, 227-38.
38. Compton, S. J.; Jones, C. G. *Anal. Biochem.* **1985**, 151, 369-74.
39. Reisner, A.H.; Nemes, P.; Bucholtz, C. *Anal. Biochem.* **1975**, 64, 509-16.
40. Sedmak, J. J.; Grossberg, S. E. *Anal. Biochem.* **1977**, 79, 544-52.
41. Bradford, M. M. *Anal. Biochem.* **1976**, 72, 248-54.
42. Koolman, J.; Röhm, K.-H. *Taschenatlas der Biochemie*, Georg Thieme Verlag, Stuttgart: 1998.
43. International Human Genome Sequencing Consortium *Nature* **2004**, 931-45.
44. Anderson, N. L.; Anderson, N. G. *Mol Cell Proteomics* **2002**, 1, 845-67.
45. Reidegeld, K. A.; Muller, M.; Stephan, C.; Bluggel, M.; Hamacher, M.; Martens, L.; Korting, G.; Chamrad, D. C.; Parkinson, D.; Apweiler, R.; Meyer, H. E.; Marcus, K. *Proteomics* **2006**, 6, 4997-5014.
46. Mueller, M.; Martens, L.; Reidegeld, K. A.; Hamacher, M.; Stephan, C.; Bluggel, M.; Korting, G.; Chamrad, D.; Scheer, C.; Marcus, K.; Meyer, H. E.; Apweiler, R. *Proteomics* **2006**, 6, 5059-75.
47. Hamacher, M.; Apweiler, R.; Arnold, G.; Becker, A.; Bluggel, M.; Carrette, O.; Colvis, C.; Dunn, M. J.; Frohlich, T.; Fountoulakis, M.; van Hall, A.; Herberg, F.; Ji, J.; Kretzschmar, H.; Lewczuk, P.; Lubec, G.; Marcus, K.; Martens, L.; Palacios Bustamante, N.; Park, Y. M.; Pennington, S. R.; Robben, J.; Stuhler, K.; Reidegeld, K. A.; Riederer, P.; Rossier, J.; Sanchez, J. C.; Schrader, M.; Stephan, C.; Tagle, D.; Thiele, H.; Wang, J.; Wiltfang, J.; Yoo, J. S.; Zhang, C.; Klose, J.; Meyer, H. E. *Proteomics* **2006**, 6, 4890-8.

-
48. Biomarkers Definitions Working Group *Clin. Pharmacol. Therapeut* **2001**, 69, 89-95.
 49. Dieterle, F.; Marrer, E. *Anal. Bioanal. Chem.* **2008**, 390, 141-54.
 50. Sahin, U.; Tureci, O.; Schmitt, H.; Cochlovius, B.; Johannes, T.; Schmits, R.; Stenner, F.; Luo, G.; Schobert, I.; Pfreunds Schuh, M. *Proc. Natl. Acad. Sci. U S A* **1995**, 92, 11810-3.
 51. Obata, Y.; Takahashi, T.; Sakamoto, J.; Tamaki, H.; Tominaga, S.; Hamajima, N.; Chen, Y. T.; Old, L. J. *Cancer. Chemother. Pharmacol.* **2000**, 46 Suppl, S37-42.
 52. Tureci, O.; Sahin, U.; Pfreunds Schuh, M. *Mol. Med. Today* **1997**, 3, 342-9.
 53. Lee, S. Y.; Jeoung, D. *J. Microbiol. Biotechnol.* **2007**, 17, 879-90.
 54. Scanlan, M. J.; Gordan, J. D.; Williamson, B.; Stockert, E.; Bander, N. H.; Jongeneel, V.; Gure, A. O.; Jager, D.; Jager, E.; Knuth, A.; Chen, Y. T.; Old, L. J. *Int. J. Cancer* **1999**, 83, 456-64.
 55. Scanlan, M. J.; Welt, S.; Gordon, C. M.; Chen, Y. T.; Gure, A. O.; Stockert, E.; Jungbluth, A. A.; Ritter, G.; Jager, D.; Jager, E.; Knuth, A.; Old, L. J. *Cancer Res.* **2002**, 62, 4041-7.
 56. Scanlan, M. J.; Gout, I.; Gordon, C. M.; Williamson, B.; Stockert, E.; Gure, A. O.; Jager, D.; Chen, Y. T.; Mackay, A.; O'Hare, M. J.; Old, L. J. *Cancer Immun.* **2001**, 1, 4.
 57. Ludwig, N.; Keller, A.; Comtesse, N.; Rheinheimer, S.; Pallasch, C.; Fischer, U.; Fassbender, K.; Steudel, W. I.; Lenhof, H. P.; Meese, E. *Clin. Cancer Res.* **2008**, 14, 4767-74.
 58. US Department of Health and Human Services, Food and Drug Administration 2004.
 59. Snyder, L. R. *J. Chromatogr. B Biomed. Sci. Appl.* **1997**, 689, 105-15.
 60. Engelhardt, H., Rohrschneider, and M.

61. Neue, U. *J. of Chromatogr. A* **2005**, 1079, 153-61.
62. Gilar, M.; Daly, A. E.; Kele, M.; Neue, U. D.; Gebler, J. C. *J. Chromatogr. A* **2004**, 1061, 183-92.
63. Giddings, J. C. *J. Chromatogr. A* **1995**, 703, 3-15.
64. Engelhardt, H.; Elgass, H. *J. of Chromatogr.* **1978**, 158, 249-59.
65. Hesse, G.; Weil, H. M. Woelm: Eschwege, 1954.
66. Berezkin, V. G. Ellis Horwood: New York, 1990.
67. Creighton, T. E. *Proteins: Structures and molecular properties*, 2nd ed.; W. H. Freeman and Company: New York, 1993.
68. Kastner, M. *Protein liquid chromatography*, Elsevier: Amsterdam, 2000.
69. Katz, E. D. *High-Performance Liquid Chromatography: Principles and Methods in Biotechnology*, John Wiley and Sons: Chichester, 1996.
70. LATHE, G. H.; RUTHVEN, C. R. *Biochem. J.* **1956**, 62, 665-74.
71. Alpert, A. J. *J. Chromatogr.* **1990**, 499, 177-96.
72. Linden, J. C.; Lawhead, C. L. *J. Chrom. A* **1975**, 105, 125-33.
73. Palmer, J. K. *Anal. Letters* **1975**, 8, 215-24.
74. Rabel, F. M.; Caputo, A. G.; Butts, E. T. *J. Chromatogr.* **1976**, 126, 731-40.
75. Brons, C.; Olieman, C. *J. Chromatogr.* **1983**, 259, 79-86.
76. Bendiak, B.; Orr, J.; Brockhausen, I.; Vella, G.; Phoebe, C. *Anal. Biochem.* **1988**, 175, 96-105.
77. Koizumi, K.; Utamura, T.; Okada, Y. *J. Chrom.A* **1985**, 321, 145-57.
78. Boersema, P. J. ; Mohammed, S.; Heck, A. J. *Anal. Bioanal.Chem.* **2008**, 391, 151-9.
79. Omaetxebarria, M. J.; Hagglund, P.; Elortza, F.; Hooper, N. M.; Arizmendi, J.

- M.; Jensen, O. N. *Anal. Chem.* **2006**, 78, 3335-41.
80. Tiselius, A. *Mineralogi Geologi* **1948**, 26B, 1-5.
81. Geng, X.; Regnier, F. E. *J. Chromatogr.* **1984**, 296, 15-30.
82. Heftman, E. *Chromatography, 5th edition*, Elsevier: Amsterdam, 1992.
83. Chen, H.; Horváth, C. *Anal. Meth. Instr.* **1993**, 1, 213-22.
84. Thevenon, G.; Regnier, F. E. *J. Chromatogr.* **1989**, 476, 499-511.
85. Horvath, C.; Melander, W.; Molnar, I. *J. Chromatogr.* **1976**, 125, 129-56.
86. Patthy, M. *J. Chromatogr. A* **1994**, 660, 17-23.
87. Melander, W. R.; El Rassi, Z.; Horváth, C. *J. Chromatogr.* **1989**, 469, 3.
88. Chen J.-G.; Weber, S. G.; Glavina, L. L.; Cantwell, F. F. *J. Chromatogr. A* **1993**, 656, 549-76.
89. Huber, C. G.; Schley, C.; Delmotte, N. *Wilson & Wilson's Comprehensive Analytical Chemistry Volume 46, Proteomics and Peptidomics: Technology Driving in Biology*, 1 ed.; Elsevier B.V.: Amsterdam, The Netherlands, 2005.
90. Schley, C. Saarland University Saarbruecken Germany 2006 .
91. Yamamoto, Sh.; Ishihara, T. *J. Chromatogr. A* **1999**, 852, 31-36.
92. Lottspeich, F.; Zorbas, H. *Bioanalytik*, ed.; Spektrum: Heidelberg, 2006.
93. Kopaciewicz, W.; Rounds, M. A.; Fausnaugh, J.; Regnier, F. E. *J. Chromatogr.* **1983**, 266, 3-21.
94. Levison, P. R. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **2003**, 790, 17-33.
95. Kalghatgi, K.; Horváth, C. *Analytical Biotechnology-Capillary Electrophoresis and Chromatography*, Horváth, C.; Nikelly, J. G., Eds.; American Chemical Society: Washington, DC, 1990.

-
96. Hansen, L. C.; Sievers, R. E. *J. Chromatogr.* **1974**, *99*, 123-33.
 97. Svec, F.; Fréchet, J. M. J. *Anal. Chem.* **1992**, *64*, 820-22.
 98. Viklund, C.; Svec, F.; Fréchet, J. M. J. *Chem. Mater.* **1996**, *8*, 744-50.
 99. Petro, M.; Svec, F.; Fréchet, J. M. J. *J. Chromatogr. A* **1996**, *752*, 59-66.
 100. Huber, C. G.; Kleindienst, G.; Bonn, G. K. *Chromatographia* **1997**, *44*, 438-48.
 101. Oberacher, H.; Huber, C. G. *TRAC* **2002**, *21*, 166-73.
 102. Walcher, W.; Oberacher, H.; Troiani, S.; Hölzl, G.; Oefner, P.; Zolla, L.; Huber, C. G. *J. Chromatogr. B* **2002**, *782*, 111-25.
 103. Huber, C. G.; Oefner, P. J.; Bonn, G. K. *Chromatographia* **1993**, *37*, 653-58.
 104. Oberacher, H.; Premstaller, A.; Huber, C. G. *J. Chromatogr. A* **2004**, *1030*, 201-8.
 105. Svec, F.; Tennikova, T.; Deyl, Z. *Monolithic Materials: Preparation, Properties and Applications*, Elsevier: Amsterdam, 2003.
 106. Mohr, J. Saarland University Saarbruecken Germany 2007 .
 107. Premstaller, A.; Oberacher, H.; Walcher, W.; Timperio, A.-M.; Zolla, L.; Chervet, J.-P.; Cavusoglu, N.; Van Dorsselaer, A.; Huber, C. G. *Anal. Chem.* **2001**, *73*, 2390-96.
 108. Minakuchi, H.; Nakanishi, K.; Soga, N.; Ishizuka, N.; Tanaka, N. *Anal. Chem.* **1996**, *68*, 3498-501.
 109. Kaji, H.; Nakanishi, K.; Soga, N. *J. Non-Cryst. Solids* **1995**, *185*, 18-30.
 110. Ishizuka, N.; Minakuchi, H.; Nakanishi, K.; Soga, N.; Tanaka, N. *J. Chromatogr. A* **1998**, *797*, 133-37.
 111. Minakuchi, H.; Nakanishi, K.; Soga, N.; Ishizuka, N.; Tanaka, N. *J. Chromatogr. A* **1998**, *797*, 121-31.
 112. Ishizuka, N.; Minakuchi, H.; Nakanishi, K.; Soga, N.; Nagayama, H.; Hosoya,

- K.; Tanaka, N. *Anal. Chem.* **2000**, *72*, 1275-80.
113. Brook, M. A.; Chen, Y.; Guo, K.; Zhang, Z.; Jin, W.; Deisingh, A.; Cruz-Aguado, J.; Brennan, J. D. *J. Sol-Gel Sci. Technol.* **2004**, *31*, 343-8.
114. Ishizuka, N.; Kobayashi, H.; Minakuchi, H.; Nakanishi, K.; Hirao, K.; Hosoya, K.; Ikegami, T.; Tanaka, N. *J. Chromatogr. A* **2002**, *960*, 85-96.
115. Goradia, D.; Cooney, J.; Hodnett, B. K.; Magner, E. *J. Mol. Catal. B: Enzymatic* **2005**, *32*, 231-9.
116. Hodgson, R. J. ; Chen, Y.; Zhang, Z.; Tleugabulova, D.; Long, H.; Zhao, X.; Organ, M.; Brook, M. A.; Brennan, J. D. *Anal. Chem.* **2004**, *76*, 2780-90.
117. Saito, Y.; Jinno, K.; Greibrokk, T. *J. Sep. Sci.* **2004**, *27*, 1379-90.
118. Barroso, B.; Lubda, D.; Bischoff, R. *J. Proteome Res.* **2003**, *2*, 633-42.
119. Machtejevas, E.; Andrecht, S.; Lubda, D.; Unger, K. K. *J. Chromatogr. A* **2007**, *1144*, 97-101.
120. Lindsay, S. *High Performance Liquid Chromatography*, ed.; Friedr. Vieweg & Sohn Verlagsgesellschaft mbH: Braunschweig/ Wiesbaden, 1996.
121. Crimmins, D. L.; Gorka, J.; Thoma, R. S.; Schwartz, B. D. *J. Chromatogr.* **1988**, *443*, 63-71.
122. Karas, M.; Hillenkamp, F. *Anal. Chem.* **1988**, *60*, 2299-301.
123. Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y. *Rapid Commun. Mass Spectrometry* **1988**, *2*, 151-3.
124. Stump, M. J.; Fleming, R. C.; Gong, W.-H.; Jaber, A. J.; Jones, J. J.; Surber, C. W.; Wilkins, C. L. *Appl. Spec. Rev.* **2002**, *37*, 275-303.
125. Yao, J.; Scott, J. R.; Young, M. K.; Wilkins, C. L. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 805-13.
126. Gong, W.; Elitzin, V. I.; Janardhanam, S.; Wilkins, C. L.; Fritsch, I. *J. Am. Chem. Soc.* **2001**, *123*, 769-70.

-
127. Brunnee *Int. J. Mass Spectrom. Ion Proc.* **1987**, 76, 125.
 128. Gross, J. H. *Mass Spectrometry*, ed.; Springer: Germany, 2004.
 129. Wiley, W. C. *Science* **1956**, 124, 817-20.
 130. Guilhaus, M. *Adv. Mass Spectrom.* **1995**, 13, 213-26.
 131. Guilhaus, M.; Mlynski, V.; Selby, D. *Rapid Commun. Mass Spectrom.* **1997**, 11, 951-62.
 132. Mamyryn, B. A. *Int. J. mass Spectrom. Ion Proc.* **1994**, 131, 1027A-39A.
 133. Wiley, W. C.; McLaren, I. H. *Rev. Sci. Instrum.* **1955**, 26, 1150-7.
 134. Jensen, O. N.; Podtelejnikov, A.; Mann, M. *Rapid. Commun. Mass Spectrom.* **1996**, 10, 1371-78.
 135. Samyn, B.; Debyser, G.; Sergeant, K.; Devreese, B.; Van Beeumen, J. *J. Am. Soc. Mass. Spectrom.* **2004**, 15, 1838-52.
 136. Busch, K. L. *Spectroscopy* **2000**, 15, 28-33.
 137. Kurz, E. A. *Am. Laboratory* **1979**, 11, 67-82.
 138. Boerboom, A. J. H. *Org. Mass Spectrom.* **1991**, 26, 929-35.
 139. De Hoffmann, E., Charette, J., and Stroobant, V. John Wiley & Sons: Paris, 96.
 140. Huber, C. G. Saarland University 2006 .
 141. Pol, J.; Hyotylainen, T. *Anal. Bioanal. Chem.* **2008**, 391, 21-31.
 142. Giddings, J. C. *Multidimensional chromatography techniques and applications*, Cortes, H. J., Ed.; Marcel Dekker Inc.: New York, 1990.
 143. Shoenmakers, P.; Marriott, P.; Beens, J. *Lc GC Eur.* **2003**, 16.
 144. Vollmer, M.; Horth, P.; Nagele, E. *Anal. Chem.* **2004**, 76, 5180-5.
 145. Schley, C.; Altmeyer, M.; Müller, R.; Swart, R.; Huber, C. G. *J. Proteome Res.*

- 2006**, 5, 2760-68.
146. Zolotarjova, N.; Mrozinski, P.; Chen, H.; Martosella, J. *J. Chromatogr. A.* **2008**, 1189, 332-8.
147. McLachlin, D. T.; Chait, B. T. *Curr. Opin. Chem. Biol.* **2001**, 5, 591-602.
148. Reinders, J.; Sickmann, A. *Proteomics* **2005**, 5, 4052-61.
149. Ihling, C.; Sinz, A. *Proteomics* **2005**, 5, 2029-42.
150. Gundry, R. L.; Fu, Q.; Jelinek, C. A.; Van Eyk, J. E.; Cotter, R. J. *Proteomics Clin. Appl.* **2007**, 1, 73-88.
151. Boersema, P. J.; Divecha, N.; Heck, A. J.; Mohammed, S. *J. Proteome Res.* **2007**, 6, 937-46.
152. Delmotte, N.; Lasaosa, M.; Tholey, A.; Heinzle, E.; Huber, C. G. *J. Proteome Res.* **2007**, 6, 4363-73.
153. Medzihradszky, K. F.; Campbell, J. M.; Baldwin, M. A.; Falick, A. M.; Juhasz, P.; Vestal, M. L.; Burlingame, A. L. *Anal. Chem.* **2000**, 72, 552-8.
154. Loboda, A. V.; Krutchinsky, A. N.; Bromirski, M.; Ens, W.; Standing, K. G. *Rapid Commun. Mass Spectrom.* **2000**, 14, 1047-57.
155. Bienvenut, W. V.; Deon, C.; Pasquarello, C.; Campbell, J. M.; Sanchez, J. C.; Vestal, M. L.; Hochstrasser, D. F. *Proteomics* **2002**, 2, 868-76.
156. Hsieh, S.; Dreisewerd, K.; van der Schors, R. C.; Jimenez, C. R.; Stahl-Zeng, J.; Hillenkamp, F.; Jorgenson, J. W.; Geraerts, W. P.; Li, K. W. *Anal. Chem.* **1998**, 70, 1847-52.
157. Zhang, B.; McDonald, C.; Li, L. *Anal. Chem.* **2004**, 76, 992-1001.
158. Miliotis, T.; Kjellstrom, S.; Nilsson, J.; Laurell, T.; Edholm, L. E.; Marko-Varga, G. *J. Mass. Spectrom.* **2000**, 35, 369-77.
159. Wall, D. B.; Berger, S. J.; Finch, J. W.; Cohen, S. A.; Richardson, K.; Chapman, R.; Drabble, D.; Brown, J.; Gostick, D. *Electrophoresis* **2002**, 23,

- 3193-204.
160. Chen, H. S.; Rejtar, T.; Andreev, V.; Moskovets, E.; Karger, B. L. *Anal. Chem.* **2005**, *77*, 2323-31.
161. Hensel, R. R.; King, R. C.; Owens, K. G. *Rapid Commun. Mass Spectrom.* **1997**, *11*, 1785-93.
162. Lou, X.; van Dongen, J. L. *J. Mass Spectrom.* **2000**, *35*, 1308-12.
163. Ericson, C.; Phung, Q. T.; Horn, D. M.; Peters, E. C.; Fitchett, J. R.; Ficarro, S. B.; Salomon, A. R.; Brill, L. M.; Brock, A. *Anal. Chem.* **2003**, *75*, 2309-15.
164. Orsnes, H.; Graf, T.; Degn, H.; Murray, K. K. *Anal. Chem.* **2000**, *72*, 251-4.
165. Preisler, J.; Foret, F.; Karger, B. L. *Anal. Chem.* **1998**, *70*, 5278-87.
166. Wehr, T. *LCGC North America* **2003**, *21*, 974-82.
167. Gluckmann, M.; Fella, K.; Waidelich, D.; Merkel, D.; Krufft, V.; Kramer, P. J.; Walter, Y.; Hellmann, J.; Karas, M.; Kroger, M. *Proteomics* **2007**, *7*, 1564-74.
168. Wang, P.; Kish, M. M.; Wesdemiotis, C. 2004.
169. Roepstorff, P.; Fohlman, J. *Biomed. Mass Spectrom.* **1984**, *11*, 601.
170. Pappin, D. J.; Hojrup, P.; Bleasby, A. J. *Curr. Biol.* **1993**, *3*, 327-32.
171. www.matrixscience.com 2008.
172. Scaffold Proteome Software User's Guide 2007.
173. Keller, A.; Nesvizhskii, A. I.; Kolker, E.; Aebersold, R. *Anal. Chem.* **2002**, *74*, 5383-92.
174. Nesvizhskii, A. I.; Keller, A.; Kolker, E.; Aebersold, R. *Anal. Chem.* **2003**, *75*, 4646-58.
175. Ashburner, M.; Ball, C. A.; Blake, J. A.; Botstein, D.; Butler, H.; Cherry, J. M.; Davis, A. P.; Dolinski, K.; Dwight, S. S.; Eppig, J. T.; Harris, M. A.; Hill, D. P.; Issel-Tarver, L.; Kasarskis, A.; Lewis, S.; Matese, J. C.; Richardson, J. E.;

- Ringwald, M.; Rubin, G. M.; Sherlock, G. *Nat.Genet.* **2000**, *25*, 25-9.
176. Mi, H.; Lazareva-Ulitsky, B.; Loo, R.; Kejariwal, A.; Vandergriff, J.; Rabkin, S.; Guo, N.; Muruganujan, A.; Doremieux, O.; Campbell, M. J.; Kitano, H.; Thomas, P. D. *Nucleic Acids Res.* **2005**, *33*, D284-8.
177. Kalbheim, E. www.krebshilfe.de .
178. Becker, N. and Holzmeier, S. www.dkfz.de .
179. Jemal, A.; Siegel, R.; Ward, E.; Hao, Y.; Xu, J.; Murray, T.; Thun, M. J. *CA. Cancer J. Clin.* **2008**, *58*, 71-96.
180. Atlas of the Body. American Medical Association: www.medem.com .
181. Kleihues, P.; Louis, D. N.; Scheithauer, B. W.; Rorke, L. B.; Reifenberger, G.; Burger, P. C.; Cavenee, W. K. *J. Neuropathol. Exp. Neurol.* **2002**, *61*, 215-25; discussion 226-9.
182. Newton, H. B. *Am.Fam.Physician* **1994**, *49*, 787-97.
183. Nicolato, A.; Gerosa, M. A.; Fina, P.; Iuzzolino, P.; Giorgiutti, F.; Bricolo, A. *Surg. Neurol.* **1995**, *44*, 208-21; discussion 221-3.
184. Central Brain Tumor Registry of the United States: www.cbtrus.org .
185. Louis, D. N.; Ohgaki, H.; Wiestler, O. D.; Cavenee, W. K.; Burger, P. C.; Jouvett, A.; Scheithauer, B. W.; Kleihues, P. *Acta Neuropathol.* **2007**, *114*, 97-109.
186. Ohgaki, H.; Kleihues, P. *J. Neuropathol. Exp. Neurol.* **2005**, *64*, 479-89.
187. Zheng, T.; Cantor, K. P.; Zhang, Y.; Chiu, B. C.; Lynch, C. F. *Cancer Epidemiol. Biomarkers Prev.* **2001**, *10*, 413-4.
188. Huncharek, M.; Kupelnick, B.; Wheeler, L. *J. Environ. Pathol. Toxicol. Oncol.* **2003**, *22* , 129-37.
189. Inskip, P. D.; Tarone, R. E.; Hatch, E. E.; Wilcosky, T. C.; Shapiro, W. R.; Selker, R. G.; Fine, H. A.; Black, P. M.; Loeffler, J. S.; Linet, M. S. *N. Engl. J.*

- Med.* **2001**, 344, 79-86.
190. Savitz, D. A.; Checkoway, H.; Loomis, D. P. *Epidemiology* **1998**, 9, 398-404.
191. Vilchez, R. A. ; Kozinetz, C. A.; Arrington, A. S.; Madden, C. R.; Butel, J. S. *Am. J. Med.* **2003**, 114, 675-84.
192. Cavenee, W. K. *J. Neurosurg.* **2000**, 92, 1080-1.
193. Prados, M. D.; Gutin, P. H.; Phillips, T. L.; Wara, W. M.; Larson, D. A.; Sneed, P. K.; Davis, R. L.; Ahn, D. K.; Lamborn, K.; Wilson, C. B. *Int. J. Radiat. Oncol. Biol. Phys.* **1992**, 23, 3-8.
194. Baldi, P.; Hatfield, G. W. Cambridge University Press: Cambridge, United Kingdom, 2002.
195. Sanger, F.; Air, G. M.; Barrell, B. G.; Brown, N. L.; Coulson, A. R.; Fiddes, C. A.; Hutchison, C. A.; Slocombe, P. M.; Smith, M. *Nature* **1977**, 265, 687-95.
196. Smith, M.; Brown, N. L.; Air, G. M.; Barrell, B. G.; Coulson, A. R.; Hutchison, C. A. 3rd; Sanger, F. *Nature* **1977**, 265, 702-5.
197. Parker, K. C.; Patterson, D.; Williamson, B.; Marchese, J.; Graber, A.; He, F.; Jacobson, A.; Juhasz, P.; Martin, S. *Mol. Cell. Proteomics* **2004**, 3, 625-59.
198. Li, K. W.; Miller, S.; Klychnikov, O.; Loos, M.; Stahl-Zeng, J.; Spijker, S.; Mayford, M.; Smit, A. B. *J. Proteome Res.* **2007**, 6, 3127-33.
199. Bodnar, W. M.; Blackburn, R. K.; Krise, J. M.; Moseley, M. A. *J. Am. Soc. Mass Spectrom.* **2003**, 14, 971-79.
200. Gu, X.; Deng, C.; Yan, G.; Zhang, X. *J. Proteome Res.* **2006**, 5, 3186-96.
201. Gevaert, K.; Pinxteren, J.; Demol, H.; Hugelier, K.; Staes, A.; Van Damme, J.; Martens, L.; Vandekerckhove, J. *J. Proteome Res.* **2006**, 5, 1415-28.
202. Hattan, S. J.; Parker, K. C. *Anal. Chem.* **2006**, 78, 7986-96.
203. Pan, S.; Rush, J.; Peskind, E. R.; Galasko, D.; Chung, K.; Quinn, J.; Jankovic, J.; Leverenz, J. B.; Zabetian, C.; Pan, C.; Wang, Y.; Oh, J. H.; Gao, J.; Zhang,

- J.; Montine, T.; Zhang, J. *J. Proteome Res.* **2008**, *7*, 720-30.
204. Caprioli, R. M. *Cancer Res.* **2005**, *65*, 10642-5.
205. Ludwig, N.; Keller, A.; Comtesse, N.; Rheinheimer, S.; Pallasch, C.; Fischer, U.; Fassbender, K.; Steudel, W. I.; Lenhof, H. P.; Meese, E. *Clin. Cancer Res.* **2008**, *14*, 4767-74.
206. Shen, Y.; Zhao, R.; Belov, M. E.; Conrads, T. P.; Anderson, G. A.; Tang, K.; Pasa-Tolic, L.; Veenstra, T. D.; Lipton, M. S.; Udseth, H. R.; Smith, R. D. *Anal. Chem.* **2001**, *73*, 1766-75.
207. Tolley, L.; Jorgenson, J. W.; Moseley, M. A. *Anal. Chem.* **2001**, *73*, 2985-91.
208. Premstaller, A.; Oberacher, H.; Walcher, W.; Timperio, A.-M.; Zolla, L.; Chervet, J.-P.; Cavusoglu, N.; Van Dorsselaer, A.; Huber, C. G. *Anal. Chem.* **2001**, *73*, 2390-96.
209. Ishizuka, N.; Minakuchi, H.; Nakanishi, K.; Soga, N.; Nagayama, H.; Hosoya, K.; Tanaka, N. *Anal. Chem.* **2000**, *72*, 1275-80.
210. Tanaka, N.; Kimura, H.; Tokuda, D.; Hosoya, K.; Ikegami, T.; Ishizuka, N.; Minakuchi, H.; Nakanishi, K.; Shintani, Y.; Furuno, M.; Cabrera, K. *Anal. Chem.* **2004**, *76*, 1273-81.
211. Wolters, D. A. ; Washburn, M. P.; Yates, J. R. 3rd *Anal. Chem.* **2001**, *73*, 5683-90.
212. Link, A. J.; Eng, J.; Schieltz, D. M.; Carmack, E.; Mize, G. J.; Morris, D. R.; Garvik, B. M.; Yates, J. R. III *Nature Biotechnology* **1999**, *17*, 676-82.
213. Wall, D. B.; Berger, S. J.; Finch, J. W.; Cohen, S. A.; Richardson, K.; Chapman, R.; Drabble, D.; Brown, J.; Gostick, D. *Electrophoresis* **2002**, *23*, 3193-204.
214. Krutchinsky, A. N.; Zhang, W.; Chait, B. T. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 493-504.
215. Shevchenko, A.; Loboda, A.; Shevchenko, A.; Ens, W.; Standing, K. G. *Anal.*

- Chem.* **2000**, 72, 2132-41.
216. Medzihradszky, K. F.; Campbell, J. M.; Baldwin, M. A.; Falick, A. M.; Juhasz, P.; Vestal, M. L.; Burlingame, A. L. *Anal. Chem.* **2000**, 72, 552-8.
217. Andreev, V. P. ; Rejtar, T.; Chen, H. S.; Moskovets, E. V.; Ivanov, A. R.; Karger, B. L. *Anal. Chem.* **2003**, 75, 6314-26.
218. Léonil, J.; Molle, D.; Gaucheron, F.; Arpino, P.; Guénot, P.; Maubois, J. L. *Lait* **1995**, 75, 193-210.
219. Turula, V. E.; Bishop, R. T.; Ricker, R. D.; de Haseth, J. A. *J. Chromatogr. A* **1997**, 763, 91-103.
220. Rice, R. H.; Means, G. E.; Brown, W. D. *Biochim. Biophys. Acta* **1977**, 492, 316-21.
221. Premstaller, A.; Oberacher, H.; Huber, C. G. *Anal. Chem.* **2000**, 72, 4386-93.
222. Estivariz, F. E.; Friedman, T. C.; Chikuma, T.; Loh, Y. P. *J. Biol. Chem.* **1992**, 267, 7456-63.
223. Campbell, J. M.; Vestal M.L. *Proceedings of the 50th ASMS Conference on Mass Spectrometry and Allied Topics* **2002**.
224. Gilar, M.; Daly, A. E.; Kele, M.; Neue, U. D.; Gebler, J. C. *J. Chromatogr. A* **2004**, 1061, 183-92.
225. Bartha, A.; Stahlberg, J. *J. Chromatogr. A* **1994**, 668, 255-84.
226. Pashkova, A.; Chen, H. S.; Rejtar, T.; Zang, X.; Giese, R.; Andreev, V.; Moskovets, E.; Karger, B. L. *Anal. Chem.* **2005**, 77, 2085-96.
227. Chen, H. S.; Rejtar, T.; Andreev, V.; Moskovets, E.; Karger, B. L. *Anal. Chem.* **2005**, 77, 2323-31.
228. Marcotte, E. M. *Nat. Biotechnol.* **2007**, 25, 755-7.
229. Delmotte, N.; Lasaosa, M.; Tholey, A.; Heinzle, E.; Huber, C. G. *J. Proteome Res.* **2007**, 6, 4363-73.

-
230. Schley, C.; Swart, R.; Huber, C. G. *J. Chromatogr. A* **2006**, 1136, 210-20.
231. Grossmann, J.; Roos, F. F.; Cieliebak, M.; Liptak, Z.; Mathis, L. K.; Muller, M.; Gruissem, W.; Baginsky, S. *J. Proteome Res.* **2005**, 4, 1768-74.
232. Mechref, Y.; Novotny, M. V.; Krishnan, C. *Anal. Chem.* **2003**, 75, 4895-903.
233. Spina, E.; Sturiale, L.; Romeo, D.; Impallomeni, G.; Garozzo, D.; Waidelich, D.; Glueckmann, M. *Rapid Commun. Mass Spectrom.* **2004**, 18, 392-8.
234. Stephens, E.; Maslen, S. L.; Green, L. G.; Williams, D. H. *Anal. Chem.* **2004**, 76, 2343-54.
235. Morelle, W.; Slomianny, M. C.; Diemer, H.; Schaeffer, C.; van Dorsselaer, A.; Michalski, J. C. *Rapid Commun. Mass Spectrom.* **2004**, 18, 2637-49.
236. Enoksson, M.; Li, J.; Ivancic, M. M.; Timmer, J. C.; Wildfang, E.; Eroshkin, A.; Salvesen, G. S.; Tao, W. A. *J. Proteome Res.* **2007**, 6, 2850-8.
237. Oberacher, H.; Premstaller, A.; Huber, C. G. *J. Chromatogr. A* **2004**, 1030, 201-8.
238. Garden, R. W.; Sweedler, J. V. *Anal. Chem.* **2000**, 72, 30-6.
239. Turney, K.; Harrison, W. W. *Rapid Commun. Mass Spectrom.* **2004**, 18, 629-35.
240. Tholey, A.; Heinzle, E. *Anal. Bioanal. Chem.* **2006**, 386, 24-37.
241. Gusev, A. I.; Wilkinson, W. R.; Proctor, A.; Hercules, D. M. *Anal. Chem.* **1995**, 67, 1034-41.
242. Nicola, A. J.; Gusev, A. I.; Proctor, A.; Jackson, E. K.; Hercules, D. M. *Rapid Commun. Mass Spectrom.* **1995**, 9, 1164-71.
243. Dally, J. E.; Gorniak, J.; Bowie, R.; Bentzley, C. M. *Anal. Chem.* **2003**, 75, 5046-53.
244. Lin, Y. S.; Chen, Y. C. *Anal. Chem.* **2002**, 74, 5793-8.
245. Trimpin, S.; Rouhanipour, A.; Az, R.; Rader, H. J.; Mullen, K. *Rapid Commun.*
-

- Mass Spectrom.* **2001**, *15*, 1364-73.
246. Pearson, H. *Nature* **2008**, *452*, 920-1.
247. Kim, H.; Lubman, D. M. *J. Chromatogr. A* **2008**, *1194*, 3-10.
248. Wang, H.; Qian, W. J.; Chin, M. H.; Petyuk, V. A.; Barry, R. C.; Liu, T.; Gritsenko, M. A.; Mottaz, H. M.; Moore, R. J.; Camp li, D. G.; Khan, A. H.; Smith, D. J.; Smith, R. D. *J. Proteome Res.* **2006**, *5*, 361-9.
249. Gilar, M.; Olivova, P.; Daly, A. E.; Gebler, J. C. *Anal. Chem.* **2005**, *77*, 6426-34.
250. Gilar, M.; Olivova, P.; Daly, A. E.; Gebler, J. C. *J. Sep. Sci.* **2005**, *28*, 1694-703.
251. Veenstra, T. D.; Conrads, T. P.; Issaq, H. J. *Electrophoresis* **2004**, *25*, 1278-9.
252. Lu, B.; Motoyama, A.; Ruse, C.; Venable, J.; Yates, J. R. 3rd *Anal.Chem.* **2008**, *80*, 2018-25.
253. Shadforth, I.; Xu, W.; Crowther, D.; Bessant, C. *J.Proteome Res.* **2006**, *5*, 2849-52.
254. Wilkins, M. R.; Gasteiger, E.; Gooley, A. A.; Herbert, B. R.; Molloy, M. P.; Binz, P.-A.; Ou, K.; Sanchez, J.-C.; Bairoch, A.; Williams, K. L.; Hochstrasser, D. F. *Journal of Molecular Biology* **1999**, *289*, 645-57.
255. 7th annual Mascot workshop and user meeting 2007 55 th ASMS, Indianapolis, Ind, USA.
256. Jung, V.; Romeike, B. F.; Henn, W.; Feiden, W.; Moringlane, J. R.; Zang, K. D.; Urbschat, S. *J. Neuropathol. Exp. Neurol.* **1999**, *58*, 993-9.
257. Liu, H.; Lin, D.; Yates, J. R. 3rd *Biotechniques* **2002**, *32*, 898, 900, 902 passim.
258. Plant-Microbe Genomics Faciity (PMGF) at Ohio State University 2007.
259. Delmotte, N., Saarland University, Saarbruecken, 2007.

260. Kyte, J.; Doolittle, R. F. *J. Mol. Biol.* **1982**, *157*, 105-32.
261. [http: and www.bioinformatics.org/sms2/protein_gravy.html](http://andwww.bioinformatics.org/sms2/protein_gravy.html) .
262. Hancock, W. S. ; Bishop, C. A.; Prestidge, R. L.; Harding, D. R.; Hearn, M. T. *Science* **1978**, *200*, 1168-70.
263. Giddings, J. C. *J. Chromatogr. A* **1995**, *703*, 3-15.
264. Giddings, J. C. *J. High Res. Chromatogr.* **1987**, *10*, 319--323.
265. Wang, H.; Qian, W. J.; Mottaz, H. M.; Clauss, T. R.; Anderson, D. J.; Moore, R. J.; Camp, D. G. 2nd; Khan, A. H.; Sforza, D. M.; Pallavicini, M.; Smith, D. J.; Smith, R. D. *J. Proteome Res.* **2005**, *4*, 2397-403.
266. Santoni, V.; Molloy, M.; Rabilloud, T. *Electrophoresis* **2000**, *21*, 1054-70.
267. Brett, D.; Pospisil, H.; Valcarcel, J.; Reich, J.; Bork, P. *Nat. Genet.* **2002**, *30*, 29-30.
268. Tureci, O.; Sahin, U.; Pfreundschuh, M. *Mol. Med. Today* **1997**, *3*, 342-9.
269. Wilkins, M. R. ; Appel, R. D.; Van Eyk, J. E.; Chung, M. C.; Gorg, A.; Hecker, M.; Huber, L. A.; Langen, H.; Link, A. J.; Paik, Y. K.; Patterson, S. D.; Pennington, S. R.; Rabilloud, T.; Simpson, R. J.; Weiss, W.; Dunn, M. J. *Proteomics* **2006**, *6*, 4-8.

Appendix

- list of 1,401 non-redundant proteins identified in a human *Glioblastoma multiforme* tissue with more than one peptide
 - membrane proteins
 - SEREX proteins
-

identified proteins of <i>Glioblastoma multiforme</i> # 1-52			
nr.	protein description	accession numbers	mass [Da]
1	Hemoglobin subunit beta	HBB_HUMAN	15,980
2	Glial fibrillary acidic protein	GFAP_HUMAN	49,863
3	Serum albumin precursor	ALBU_HUMAN	69,349
4	Tubulin beta-2C chain	TBB2C_HUMAN	49,813
5	Actin, cytoplasmic 2	ACTB_HUMAN, ACTG_HUMAN	41,776
6	Keratin, type II cytoskeletal 1	K2C1_HUMAN	66,001
7	Tubulin alpha-1B chain	TBA1B_HUMAN	50,134
8	ATP synthase subunit beta, mitochondrial precursor	ATPB_HUMAN	56,543
9	Hemoglobin subunit alpha	HBA_HUMAN	15,240
10	Creatine kinase B-type	KCRB_HUMAN	42,627
11	Myelin basic protein	MBP_HUMAN	33,100
12	Vimentin	VIME_HUMAN	53,635
13	Keratin, type I cytoskeletal 9	K1C9_HUMAN	62,113
14	Spectrin alpha chain, brain	SPTA2_HUMAN	284,525
15	Heat shock cognate 71 kDa protein	HSP7C_HUMAN	70,882
16	Dihydropyrimidinase-related protein 2	DPYL2_HUMAN	62,276
17	Synapsin-1	SYN1_HUMAN	74,093
18	14-3-3 protein epsilon	1433E_HUMAN	29,157
19	Keratin, type I cytoskeletal 10	K1C10_HUMAN	59,494
20	Alpha-enolase	ENOA_HUMAN	47,152
21	14-3-3 protein zeta/delta	1433Z_HUMAN	27,728
22	Fructose-bisphosphate aldolase A	ALDOA_HUMAN	39,403
23	Spectrin beta chain, brain 1	SPTB2_HUMAN	274,595
24	Brain acid soluble protein 1	BASP_HUMAN	22,675
25	Annexin A5	ANXA5_HUMAN	35,921
26	Synaptosomal-associated protein 25	SNP25_HUMAN	23,297
27	Stress-70 protein, mitochondrial precursor	GRP75_HUMAN	73,663
28	Ferritin light chain	FRIL_HUMAN	20,003
29	78 kDa glucose-regulated protein precursor	GRP78_HUMAN	72,317
30	Microtubule-associated protein 2	MAP2_HUMAN	199,523
31	Annexin A2	ANXA2_HUMAN	38,588
32	Alpha-1-antichymotrypsin precursor	AACT_HUMAN	47,635
33	Serotransferrin precursor	TRFE_HUMAN	77,032
34	Calnexin precursor	CALX_HUMAN	67,552
35	Keratin, type II cytoskeletal 2 epidermal	K22E_HUMAN	65,848
36	Glyceraldehyde-3-phosphate dehydrogenase	G3P_HUMAN	36,035
37	Tenascin-R precursor	TENR_HUMAN	149,528
38	Tubulin beta-2B chain	TBB2B_HUMAN	49,935
39	Gamma-enolase	ENOG_HUMAN	47,252
40	Microtubule-associated protein tau	TAU_HUMAN	78,860
41	Malate dehydrogenase, mitochondrial precursor	MDHM_HUMAN	35,514
42	Ferritin heavy chain	FRIH_HUMAN	21,208
43	2',3'-cyclic-nucleotide 3'-phosphodiesterase	CN37_HUMAN	47,563
44	Carbonic anhydrase 1	CAH1_HUMAN	28,852
45	Guanine nucleotide-binding protein G(o) subunit alpha 2	GNAO1_HUMAN, GNAO2_HUMAN	40,070
46	Phosphoglycerate mutase 1	PGAM1_HUMAN	28,787
47	Phosphoglycerate kinase 1	PGK1_HUMAN	44,597
48	Neurofilament medium polypeptide	NFM_HUMAN	102,429
49	L-lactate dehydrogenase B chain	LDHB_HUMAN	36,621
50	Alpha-1-antitrypsin precursor	A1AT_HUMAN	46,720
51	Clathrin coat assembly protein AP180	AP180_HUMAN	92,486
52	Fibrinogen beta chain precursor [Contains: Fibrinopeptide B]	FIBB_HUMAN	55,911

identified proteins of <i>Glioblastoma multiforme</i> # 53-97			
nr.	protein description	accession numbers	mass [Da]
53	Microtubule-associated protein 1B	MAP1B_HUMAN	270,602
54	Clathrin heavy chain 1	CLH1_HUMAN	191,601
55	Fibrinogen alpha chain precursor [Contains: Fibrinopeptide A]	FIBA_HUMAN	94,955
56	Triosephosphate isomerase	TPIS_HUMAN	26,651
57	Catalase	CATA_HUMAN	59,739
58	Endophilin-A1	SH3G2_HUMAN	39,945
59	Annexin A1	ANXA1_HUMAN	38,698
60	Ankyrin-2	ANK2_HUMAN	430,319
61	Sodium/potassium-transporting ATPase subunit beta-1	AT1B1_HUMAN	35,045
62	60 kDa heat shock protein, mitochondrial precursor	CH60_HUMAN	61,038
63	Dynamin-1	DYN1_HUMAN	97,392
64	Fructose-bisphosphate aldolase C	ALDOC_HUMAN	39,438
65	Myelin-associated glycoprotein precursor	MAG_HUMAN	69,050
66	Fibronectin precursor	FINC_HUMAN	262,581
67	14-3-3 protein gamma	1433G_HUMAN	28,285
68	Phosphatidylethanolamine-binding protein 1	PEBP1_HUMAN	21,039
69	Superoxide dismutase [Mn], mitochondrial precursor	SODM_HUMAN	24,705
70	Tubulin beta-4 chain	TBB4_HUMAN	49,567
71	Plectin-1	PLEC1_HUMAN	531,708
72	Apolipoprotein A-I precursor	APOA1_HUMAN	30,761
73	Neuromodulin	NEUM_HUMAN	24,784
74	Protein disulfide-isomerase A3 precursor	PDIA3_HUMAN	56,767
75	Visinin-like protein 1	VISL1_HUMAN	22,125
76	Peroxiredoxin-1	PRDX1_HUMAN	22,093
77	Pyruvate kinase isozymes M1/M2	KPYM_HUMAN	57,920
78	Clusterin precursor	CLUS_HUMAN	52,477
79	14-3-3 protein beta/alpha	1433B_HUMAN	28,065
80	ATP synthase subunit alpha, mitochondrial precursor	ATPA_HUMAN	59,734
81	Actin, alpha cardiac muscle 1	ACTC_HUMAN	42,002
82	Microtubule-associated protein 1A	MAP1A_HUMAN	306,456
83	Heat shock protein HSP 90-alpha	HS90A_HUMAN	84,645
84	Haptoglobin precursor [Contains: Haptoglobin alpha chain; Haptoglobin beta chain]	HPT_HUMAN	45,187
85	Neurofascin precursor	NFASC_HUMAN	150,010
86	Filamin-A	FLNA_HUMAN	280,711
87	Vacuolar ATP synthase subunit B, brain isoform	VATB2_HUMAN	56,484
88	Versican core protein precursor	CSPG2_HUMAN	372,795
89	Hornerin	HORN_HUMAN	282,355
90	Septin-11	SEP11_HUMAN	49,381
91	Heterogeneous nuclear ribonucleoproteins A2/B1	ROA2_HUMAN	37,412
92	Protein bassoon	BSN_HUMAN	416,480
93	Neuroblast differentiation-associated protein AHNK	AHNK_HUMAN	629,086
94	Protein S100-A9	S10A9_HUMAN	13,224
95	Hemoglobin subunit delta	HBD_HUMAN	16,037
96	Myosin-9	MYH9_HUMAN	226,520
97	Ubiquitin carboxyl-terminal hydrolase isozyme L1	UCHL1_HUMAN	24,806

identified proteins of <i>Glioblastoma multiforme</i> # 98-142			
nr.	protein description	accession numbers	mass [Da]
98	Glutaminase kidney isoform, mitochondrial precursor	GLSK_HUMAN	73,444
99	Talin-1	TLN1_HUMAN	269,747
100	Contactin-1 precursor	CNTN1_HUMAN	113,305
101	Fibrinogen gamma chain precursor	FIBG_HUMAN	51,495
102	Dynein heavy chain, cytosolic	DYHC_HUMAN	532,388
103	Tyrosine-protein phosphatase non-receptor type substrate 1 precursor	SHPS1_HUMAN	54,793
104	Isocitrate dehydrogenase [NAD] subunit alpha, mitochondrial precursor	IDH3A_HUMAN	39,575
105	Cofilin-1	COF1_HUMAN	18,485
106	Sodium/potassium-transporting ATPase subunit alpha-1 precursor	AT1A1_HUMAN	112,882
107	Beta-soluble NSF attachment protein	SNAB_HUMAN	33,540
108	Thy-1 membrane glycoprotein precursor	THY1_HUMAN	17,917
109	N(G),N(G)-dimethylarginine dimethylaminohydrolase 1	DDAH1_HUMAN	31,104
110	Neuron-specific calcium-binding protein hippocalcin	HPCA_HUMAN	22,411
111	Syntaxin-1B	STX1B_HUMAN	33,227
112	Aspartate aminotransferase, cytoplasmic	AATC_HUMAN	46,230
113	Band 4.1-like protein 3	E41L3_HUMAN	120,662
114	14-3-3 protein theta	1433T_HUMAN	27,747
115	Collagen alpha-1(I) chain precursor	CO1A1_HUMAN	138,893
116	Alpha-crystallin B chain	CRYAB_HUMAN	20,141
117	Glutamine synthetase	GLNA_HUMAN	42,047
118	Alpha-actinin-1	ACTN1_HUMAN	103,043
119	Spectrin alpha chain, erythrocyte	SPTA1_HUMAN	279,903
120	Calmodulin	CALM_HUMAN	16,820
121	Band 3 anion transport protein	B3AT_HUMAN	101,778
122	Rab GDP dissociation inhibitor alpha	GDIA_HUMAN	50,566
123	Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A alpha isoform	2AAA_HUMAN	65,292
124	Dihydropyrimidinase-related protein 3	DPYL3_HUMAN	61,946
125	Neural cell adhesion molecule L1 precursor	L1CAM_HUMAN	139,985
126	Glutathione S-transferase P	GSTP1_HUMAN	23,339
127	Peptidyl-prolyl cis-trans isomerase A	PPIA_HUMAN	17,995
128	A-kinase anchor protein 12	AKA12_HUMAN	191,457
129	Acyl-CoA-binding protein	ACBP_HUMAN	10,027
130	Syntaxin-binding protein 1	STXB1_HUMAN	67,554
131	Ubiquitin	UBIQ_HUMAN	8,547
132	NAD-dependent deacetylase sirtuin-2	SIRT2_HUMAN	43,166
133	Protein FAM49B	FA49B_HUMAN	36,731
134	Nucleoside diphosphate kinase B	NDKB_HUMAN	17,280
135	Annexin A6	ANXA6_HUMAN	75,860
136	Glutamate dehydrogenase 1, mitochondrial precursor	DHE3_HUMAN	61,382
137	Heat shock protein beta-1	HSPB1_HUMAN	22,765
138	Vesicle-fusing ATPase	NSF_HUMAN	82,545
139	Tubulin beta-3 chain	TBB3_HUMAN	50,415
140	Aconitate hydratase, mitochondrial precursor	ACON_HUMAN	85,410
141	Alpha-internexin	AINX_HUMAN	55,374
142	Amphiphysin	AMPH_HUMAN	76,239

identified proteins of <i>Glioblastoma multiforme</i> # 143-185			
nr.	protein description	accession numbers	mass [Da]
143	Isocitrate dehydrogenase [NADP], mitochondrial precursor	IDHP_HUMAN	50,892
144	Heat shock 70 kDa protein 1	HSP71_HUMAN	70,036
145	Peroxiredoxin-6	PRDX6_HUMAN	25,018
146	Protein S100-A8	S10A8_HUMAN	10,817
147	Cadherin-2 precursor	CADH2_HUMAN	99,794
148	14-3-3 protein eta	1433F_HUMAN	28,202
149	CD44 antigen precursor	CD44_HUMAN	81,535
150	Heterogeneous nuclear ribonucleoprotein K	HNRPK_HUMAN	50,961
151	Tubulin polymerization-promoting protein	TPPP_HUMAN	23,676
152	Histone H4	H4_HUMAN	11,350
153	Endoplasmic precursor	ENPL_HUMAN	92,454
154	4F2 cell-surface antigen heavy chain	4F2_HUMAN	57,929
155	ATP synthase subunit b, mitochondrial precursor	AT5F1_HUMAN	28,891
156	Vesicle-associated membrane protein-associated protein A	VAPA_HUMAN	27,876
157	Spectrin beta chain, erythrocyte	SPTB1_HUMAN	246,307
158	Keratin, type I cytoskeletal 16	K1C16_HUMAN	51,251
159	Protein disulfide-isomerase precursor	PDIA1_HUMAN	57,100
160	Collagen alpha-2(I) chain precursor	CO1A2_HUMAN	129,271
161	Aldehyde dehydrogenase, mitochondrial precursor	ALDH2_HUMAN	56,363
162	L-lactate dehydrogenase A chain	LDHA_HUMAN	36,671
163	Thioredoxin-dependent peroxide reductase, mitochondrial precursor	PRDX3_HUMAN	27,675
164	Carbonic anhydrase 2	CAH2_HUMAN	29,229
165	Acid ceramidase precursor	ASAH1_HUMAN	44,633
166	Synapsin-2	SYN2_HUMAN	62,951
167	Ig gamma-1 chain C region	IGHG1_HUMAN	36,087
168	Large proline-rich protein BAT3	BAT3_HUMAN	119,389
169	Serine/threonine-protein phosphatase 2B catalytic subunit alpha isoform	PP2BA_HUMAN	58,672
170	Carbonyl reductase [NADPH] 1	CBR1_HUMAN	30,357
171	Vacuolar ATP synthase catalytic subunit A	VATA_HUMAN	68,287
172	Sideroflexin-1	SFXN1_HUMAN	35,602
173	Mitochondrial inner membrane protein	IMMT_HUMAN	83,661
174	Histone H2B type 1-B	H2B1B_HUMAN, H2B1J_HUMAN, H2B1O_HUMAN, H2B2E_HUMAN, H2B3B_HUMAN	13,933
175	Myoglobin	MYG_HUMAN	17,166
176	Septin-7	SEPT7_HUMAN	50,662
177	Ig alpha-1 chain C region	IGHA1_HUMAN	37,636
178	Ankyrin-1	ANK1_HUMAN	206,246
179	Citrate synthase, mitochondrial precursor	CISY_HUMAN	51,696
180	10 kDa heat shock protein, mitochondrial	CH10_HUMAN	10,914
181	Tubulin alpha-4A chain	TBA4A_HUMAN	49,907
182	Calcineurin subunit B isoform 1	CANB1_HUMAN	19,282
183	Vacuolar proton pump subunit d 1	VA0D1_HUMAN	40,313
184	Lysosome membrane protein 2	SCRB2_HUMAN	54,274
185	Serine/threonine-protein phosphatase 2A regulatory subunit B'	PTPA_HUMAN	40,650

identified proteins of <i>Glioblastoma multiforme</i> # 186-227			
nr.	protein description	accession numbers	mass [Da]
186	Myristoylated alanine-rich C-kinase substrate	MARCS_HUMAN	31,536
187	Ermin	ERMIN_HUMAN	32,765
188	Astrocytic phosphoprotein PEA-15	PEA15_HUMAN	15,023
189	Limbic system-associated membrane protein precursor	LSAMP_HUMAN	37,375
190	Adenylate kinase isoenzyme 1	KAD1_HUMAN	21,617
191	Complement C3 precursor [Contains: Complement C3 beta chain; Complement C3 alpha chain; C3a anaphylatoxin; Complement C3b alpha' chain; Complement C3c alpha' chain fragment 1; Complement C3dg fragment; Complemen...	CO3_HUMAN	187,131
192	T-complex protein 1 subunit beta	TCPB_HUMAN	57,472
193	Histone H2AV	H2AV_HUMAN, H2AZ_HUMAN	13,491
194	Ubiquitin-like modifier-activating enzyme 1	UBA1_HUMAN	117,832
195	Malate dehydrogenase, cytoplasmic	MDHC_HUMAN	36,409
196	Brevican core protein precursor	PGCB_HUMAN	99,100
197	Myosin-10	MYH10_HUMAN	228,927
198	Ubiquitin thioesterase OTUB1	OTUB1_HUMAN	31,267
199	Profilin-1	PROF1_HUMAN	15,036
200	Myc box-dependent-interacting protein 1	BIN1_HUMAN	64,681
201	Receptor-type tyrosine-protein phosphatase zeta precursor	PTPRZ_HUMAN	254,512
202	Protein NDRG2	NDRG2_HUMAN	40,781
203	Protein S100-B	S100B_HUMAN	10,695
204	Neurogranin	NEUG_HUMAN	7,601
205	Tubulin beta chain	TBB5_HUMAN	49,653
206	Neutral alpha-glucosidase AB precursor	GANAB_HUMAN	106,858
207	Keratin, type I cytoskeletal 14	K1C14_HUMAN	51,605
208	Contactin-associated protein 1 precursor	CNTP1_HUMAN	156,250
209	Fumarate hydratase, mitochondrial precursor	FUMH_HUMAN	54,620
210	Cathepsin D precursor	CATD_HUMAN	44,535
211	Ig kappa chain C region	KAC_HUMAN	11,591
212	Sodium/potassium-transporting ATPase subunit beta-3	AT1B3_HUMAN	31,496
213	Voltage-dependent calcium channel subunit alpha-2/delta-1 precursor	CA2D1_HUMAN	123,169
214	Protein DJ-1	PARK7_HUMAN	19,873
215	Transthyretin precursor	TTHY_HUMAN	15,869
216	Ectonucleotide pyrophosphatase/phosphodiesterase family member 6 precursor	ENPP6_HUMAN	50,225
217	Superoxide dismutase [Cu-Zn]	SODC_HUMAN	15,917
218	Cytosol aminopeptidase	AMPL_HUMAN	56,150
219	Neuronal cell adhesion molecule precursor	NRCAM_HUMAN	143,877
220	Nicotinamide phosphoribosyltransferase	NAMPT_HUMAN	55,505
221	Myosin light polypeptide 6	MYL6_HUMAN	16,912
222	ADP/ATP translocase 2	ADT2_HUMAN	32,879
223	Aspartate aminotransferase, mitochondrial precursor	AATM_HUMAN	47,459
224	Dematin	DEMA_HUMAN	45,498
225	Vitronectin precursor	VTNC_HUMAN	54,288
226	Stress-induced-phosphoprotein 1	STIP1_HUMAN	62,624
227	Protein kinase C and casein kinase substrate in neurons protein 1	PACN1_HUMAN	50,948

identified proteins of <i>Glioblastoma multiforme</i> # 228-270			
nr.	protein description	accession numbers	mass [Da]
228	Creatine kinase, ubiquitous mitochondrial precursor	KCRU_HUMAN	47,020
229	Hepatocyte cell adhesion molecule precursor	HECAM_HUMAN	46,010
230	Delta-1-pyrroline-5-carboxylate dehydrogenase, mitochondrial precursor	AL4A1_HUMAN	61,702
231	Tripeptidyl-peptidase 1 precursor	TPP1_HUMAN	61,230
232	ATP synthase-coupling factor 6, mitochondrial precursor	ATP5J_HUMAN	12,570
233	Reticulon-1	RTN1_HUMAN	83,602
234	NADH dehydrogenase [ubiquinone] iron-sulfur protein 3, mitochondrial precursor	NDUS3_HUMAN	30,224
235	Syntaxin-1A	STX1A_HUMAN	33,006
236	Reticulon-4	RTN4_HUMAN	129,917
237	Membrane-associated progesterone receptor component 1	PGRC1_HUMAN	21,654
238	Galectin-3-binding protein precursor	LG3BP_HUMAN	65,314
239	Septin-8	SEPT8_HUMAN	55,738
240	Synaptotagmin-1	SYT1_HUMAN	47,556
241	Tropomyosin alpha-1 chain	TPM1_HUMAN	32,692
242	Lysosomal alpha-glucosidase precursor	LYAG_HUMAN	105,321
243	HLA class I histocompatibility antigen, Cw-12 alpha chain precursor	1C12_HUMAN	40,867
244	Lysosome-associated membrane glycoprotein 2 precursor	LAMP2_HUMAN	44,943
245	Filamin-C	FLNC_HUMAN	290,934
246	AP-2 complex subunit beta-1	AP2B1_HUMAN	104,537
247	Ezrin	EZRI_HUMAN	69,397
248	Peroxiredoxin-2	PRDX2_HUMAN	21,874
249	Alpha-1-acid glycoprotein 1 precursor	A1AG1_HUMAN	23,494
250	Acetyl-CoA acetyltransferase, mitochondrial precursor	THIL_HUMAN	45,182
251	Elongation factor Tu, mitochondrial precursor	EFTU_HUMAN	49,524
252	Myelin proteolipid protein	MYPR_HUMAN	30,060
253	Nestin	NEST_HUMAN	177,419
254	Glucose-6-phosphate isomerase	G6PI_HUMAN	63,130
255	Rho GDP-dissociation inhibitor 1	GDIR_HUMAN	23,190
256	Alpha-synuclein	SYUA_HUMAN	14,441
257	Integrin alpha-V precursor	ITAV_HUMAN	116,023
258	Protein-L-isoaspartate(D-aspartate) O-methyltransferase	PIMT_HUMAN	24,633
259	Sorcin	SORCN_HUMAN	21,659
260	Ras-related protein Rab-11B	RB11B_HUMAN	24,471
261	Microtubule-associated protein 4	MAP4_HUMAN	121,003
262	Basement membrane-specific heparan sulfate proteoglycan core protein precursor	PGBM_HUMAN	468,788
263	Keratin, type II cytoskeletal 6A	K2C6A_HUMAN	60,028
264	Cytochrome c oxidase subunit 4 isoform 1, mitochondrial precursor	COX41_HUMAN	19,559
265	Calreticulin precursor	CALR_HUMAN	48,125
266	Sodium/potassium-transporting ATPase subunit beta-2	AT1B2_HUMAN	33,349
267	Excitatory amino acid transporter 1	EAA1_HUMAN	59,556
268	Protein disulfide-isomerase A4 precursor	PDIA4_HUMAN	72,916
269	Trifunctional enzyme subunit alpha, mitochondrial precursor	ECHA_HUMAN	82,984
270	Collagen alpha-3(VI) chain precursor	CO6A3_HUMAN	343,532

identified proteins of <i>Glioblastoma multiforme</i> # 271-318			
nr.	protein description	accession numbers	mass [Da]
271	Breast carcinoma-amplified sequence 1	BCAS1_HUMAN	61,691
272	Enoyl-CoA hydratase, mitochondrial precursor	ECHM_HUMAN	31,370
273	Heterogeneous nuclear ribonucleoprotein H	HNRH1_HUMAN	49,212
274	Beta-adducin	ADDB_HUMAN	80,836
275	Basigin precursor	BASI_HUMAN	42,182
276	Myeloperoxidase precursor	PERM_HUMAN	83,853
277	Neural cell adhesion molecule 1, 140 kDa isoform precursor	NCA11_HUMAN	93,343
278	Heat shock protein HSP 90-beta	HS90B_HUMAN	83,249
279	Cytochrome b-c1 complex subunit 7	QCR7_HUMAN	13,513
280	Cysteine and glycine-rich protein 1	CSRP1_HUMAN	20,549
281	Calumenin precursor	CALU_HUMAN	37,090
282	Heterogeneous nuclear ribonucleoprotein A3	ROA3_HUMAN	39,577
283	Cell cycle control protein 50A	CC50A_HUMAN	40,668
284	HLA class I histocompatibility antigen, Cw-7 alpha chain precursor	1C07_HUMAN	40,630
285	Mitochondrial 28S ribosomal protein S36	RT36_HUMAN	11,448
286	3-hydroxyacyl-CoA dehydrogenase type-2	HCD2_HUMAN	26,905
287	Septin-9	SEPT9_HUMAN	65,384
288	Band 4.1-like protein 1	E41L1_HUMAN	98,487
289	Ras-related protein Rab-7a	RAB7A_HUMAN	23,472
290	Toll-interacting protein	TOLIP_HUMAN	30,263
291	2-oxoglutarate dehydrogenase E1 component, mitochondrial precursor	ODO1_HUMAN	115,919
292	Rho GTPase-activating protein 1	RHG01_HUMAN	50,420
293	Cytochrome b-c1 complex subunit 2, mitochondrial precursor	QCR2_HUMAN	48,425
294	T-complex protein 1 subunit theta	TCPQ_HUMAN	59,603
295	S-phase kinase-associated protein 1A	SKP1_HUMAN	18,640
296	Beta-synuclein	SYUB_HUMAN	14,269
297	Stathmin	STMN1_HUMAN	17,285
298	Neuronal-specific septin-3	SEPT3_HUMAN	40,687
299	Secretogranin-2 precursor	SCG2_HUMAN	70,925
300	Peroxiredoxin-5, mitochondrial precursor	PRDX5_HUMAN	22,008
301	Alpha-adducin	ADDA_HUMAN	80,938
302	Peptidyl-prolyl cis-trans isomerase B precursor	PPIB_HUMAN	22,725
303	Sulfatase-modifying factor 2 precursor	SUMF2_HUMAN	33,839
304	D-3-phosphoglycerate dehydrogenase	SERA_HUMAN	56,633
305	Lactotransferrin precursor	TRFL_HUMAN	78,164
306	SH3 domain-binding glutamic acid-rich-like protein	SH3L1_HUMAN	12,757
307	Plasma protease C1 inhibitor precursor	IC1_HUMAN	55,138
308	Transitional endoplasmic reticulum ATPase	TERA_HUMAN	89,307
309	Neurofilament light polypeptide	NFL_HUMAN	61,500
310	Splicing factor, proline- and glutamine-rich	SFPQ_HUMAN	76,132
311	Alpha-aminoacidic semialdehyde dehydrogenase	AL7A1_HUMAN	55,349
312	Cytosolic non-specific dipeptidase	CNDP2_HUMAN	52,862
313	Serum amyloid P-component precursor	SAMP_HUMAN	25,370
314	ATP synthase subunit e, mitochondrial	ATP5I_HUMAN	7,916
315	Histidine triad nucleotide-binding protein 1	HINT1_HUMAN	13,784
316	ATP synthase subunit g, mitochondrial	ATP5L_HUMAN	11,411
317	Elongation factor 1-alpha 1	EF1A1_HUMAN	50,123
318	Vesicle-associated membrane protein 3	VAMP3_HUMAN	11,291

identified proteins of <i>Glioblastoma multiforme</i> # 319-364			
nr.	protein description	accession numbers	mass [Da]
319	Brain-specific angiogenesis inhibitor 1-associated protein 2	BAIP2_HUMAN	60,851
320	Apolipoprotein E precursor	APOE_HUMAN	36,136
321	HLA class I histocompatibility antigen, A-69 alpha chain	1A69_HUMAN	40,958
322	Reticulocalbin-1 precursor	RCN1_HUMAN	38,873
323	Vacuolar proton pump subunit E 1	VATE1_HUMAN	26,128
324	Protein NDRG1	NDRG1_HUMAN	42,817
325	Dihydropyrimidinase-related protein 1	DPYL1_HUMAN	62,167
326	Neuroplastin precursor	NPTN_HUMAN	31,274
327	Macrophage migration inhibitory factor	MIF_HUMAN	12,459
328	cAMP-regulated phosphoprotein 19	ARP19_HUMAN	12,305
329	Mitogen-activated protein kinase 1	MK01_HUMAN	41,374
330	Transgelin-2	TAGL2_HUMAN	22,374
331	Calcium/calmodulin-dependent protein kinase type II alpha chain	KCC2A_HUMAN	54,013
332	Ig gamma-2 chain C region	IGHG2_HUMAN	35,865
333	Guanine nucleotide-binding protein G(s) subunit alpha isoforms XLas	GNAS1_HUMAN	111,008
334	Gelsolin precursor	GELS_HUMAN	85,680
335	Chromogranin-A precursor	CMGA_HUMAN	50,712
336	Nucleophosmin	NPM_HUMAN	32,557
337	Catenin delta-2	CTND2_HUMAN	132,641
338	Protein S100-A13	S10AD_HUMAN	11,454
339	Ig lambda chain C regions	LAC_HUMAN	11,218
340	NADH dehydrogenase [ubiquinone] iron-sulfur protein 6, mitochondrial precursor	NDUS6_HUMAN	13,693
341	Vacuolar proton translocating ATPase 116 kDa subunit a isoform 1	VPP1_HUMAN	96,399
342	Amine oxidase [flavin-containing] B	AOFB_HUMAN	58,746
343	Neurosecretory protein VGF precursor	VGF_HUMAN	67,269
344	Antithrombin-III precursor	ANT3_HUMAN	52,586
345	Cytochrome b-c1 complex subunit Rieske, mitochondrial precursor	UCRI_HUMAN	29,650
346	Sideroflexin-3	SFXN3_HUMAN	35,486
347	Rabphilin-3A	RP3A_HUMAN	76,856
348	Programmed cell death protein 6	PDCD6_HUMAN	21,851
349	Myosin regulatory light chain 2, nonsarcomeric	MLRM_HUMAN	19,777
350	ATP synthase subunit d, mitochondrial	ATP5H_HUMAN	18,474
351	UMP-CMP kinase	KCY_HUMAN	22,205
352	Hypoxia up-regulated protein 1 precursor	HYOU1_HUMAN	111,319
353	Protein phosphatase 1 regulatory subunit 1B	PPR1B_HUMAN	22,945
354	Glycogen phosphorylase, brain form	PYGB_HUMAN	96,680
355	NADH-cytochrome b5 reductase 3	NB5R3_HUMAN	34,218
356	Annexin A11	ANX11_HUMAN	54,374
357	Programmed cell death 6-interacting protein	PDC6I_HUMAN	96,007
358	Protein disulfide-isomerase A6 precursor	PDIA6_HUMAN	48,104
359	Ubiquitin carboxyl-terminal hydrolase 5	UBP5_HUMAN	95,770
360	Cell adhesion molecule 2 precursor	CADM2_HUMAN	47,536
361	Src substrate cortactin	SRC8_HUMAN	61,617
362	Monoglyceride lipase	MGLL_HUMAN	33,244
363	Elongation factor 1-alpha 2	EF1A2_HUMAN	50,453
364	Septin-2	SEPT2_HUMAN	41,470

identified proteins of <i>Glioblastoma multiforme</i> # 365-407			
nr.	protein description	accession numbers	mass [Da]
365	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 67 kDa subunit precursor	RIB1_HUMAN	68,553
366	Sodium/potassium-transporting ATPase subunit alpha-2 precursor	AT1A2_HUMAN	112,251
367	Serine/threonine-protein kinase PAK 1	PAK1_HUMAN	60,630
368	LIM and SH3 domain protein 1	LASP1_HUMAN	29,699
369	Sodium channel subunit beta-3 precursor	SCN3B_HUMAN	24,684
370	Heterogeneous nuclear ribonucleoprotein A1	ROA1_HUMAN	38,828
371	TOM1-like protein 2	TM1L2_HUMAN	55,539
372	Inorganic pyrophosphatase	IPYR_HUMAN	32,643
373	Synaptojanin-1	SYNJ1_HUMAN	173,332
374	Matrin-3	MATR3_HUMAN	94,609
375	Acetyl-CoA acetyltransferase, cytosolic	THIC_HUMAN	41,332
376	Aquaporin-4	AQP4_HUMAN	34,812
377	Synaptic vesicle glycoprotein 2A	SV2A_HUMAN	82,679
378	Heterogeneous nuclear ribonucleoprotein U	HNRPU_HUMAN	90,496
379	Inter-alpha-trypsin inhibitor heavy chain H2 precursor	ITIH2_HUMAN	106,421
380	Angiotensinogen precursor	ANGT_HUMAN	53,137
381	Prohibitin	PHB_HUMAN	29,787
382	Alpha-soluble NSF attachment protein	SNAA_HUMAN	33,216
383	Neurocan core protein precursor	CSPG3_HUMAN	143,072
384	Flavin reductase	BLVRB_HUMAN	22,101
385	Immunoglobulin superfamily member 8 precursor	IGSF8_HUMAN	65,015
386	Protein NDRG4	NDRG4_HUMAN	38,440
387	Opioid-binding protein/cell adhesion molecule precursor	OPCM_HUMAN	37,990
388	Tropomyosin alpha-4 chain	TPM4_HUMAN	28,504
389	Tubulin alpha-1A chain	TBA1A_HUMAN	50,118
390	Osteopontin precursor	OSTP_HUMAN	35,405
391	Clathrin light chain B	CLCB_HUMAN	25,173
392	Ezrin-radixin-moesin-binding phosphoprotein 50	NHERF_HUMAN	38,850
393	Heat shock-related 70 kDa protein 2	HSP72_HUMAN	70,005
394	PH and SEC7 domain-containing protein 3	PSD3_HUMAN	116,018
395	Galectin-3	LEG3_HUMAN	26,171
396	Glutathione transferase omega-1	GSTO1_HUMAN	27,549
397	Erlin-2 precursor	ERLN2_HUMAN	37,822
398	Synaptic vesicle membrane protein VAT-1 homolog	VAT1_HUMAN	41,902
399	Plasma membrane calcium-transporting ATPase 4	AT2B4_HUMAN	137,906
400	Microtubule-associated protein RP/EB family member 2	MARE2_HUMAN	37,014
401	Transgelin-3	TAGL3_HUMAN	22,455
402	Plasminogen activator inhibitor 1 RNA-binding protein	PAIRB_HUMAN	44,948
403	ADP-ribosylation factor 1	ARF1_HUMAN, ARF3_HUMAN	20,680
404	AP-2 complex subunit alpha-2	AP2A2_HUMAN	103,945
405	Annexin A7	ANXA7_HUMAN	52,723
406	Synaptopodin	SYNPO_HUMAN	99,446
407	Cytochrome c1, heme protein, mitochondrial precursor	CY1_HUMAN	35,373

identified proteins of <i>Glioblastoma multiforme</i> # 408-450			
nr.	protein description	accession numbers	mass [Da]
408	Hepatoma-derived growth factor	HDGF_HUMAN	26,771
409	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 10	NDUBA_HUMAN	20,759
410	Cystatin-B	CYTB_HUMAN	11,121
411	Protein TFG	TFG_HUMAN	43,416
412	Pyruvate carboxylase, mitochondrial precursor	PYC_HUMAN	129,617
413	Phosphate carrier protein, mitochondrial precursor	MPCP_HUMAN	40,078
414	Cytochrome b-c1 complex subunit 1, mitochondrial precursor	QCR1_HUMAN	52,628
415	Dihydrolipoyllysine-residue succinyltransferase component of 2-oxoglutarate dehydrogenase complex, mitochondrial precursor	ODO2_HUMAN	48,622
416	S-formylglutathione hydrolase	ESTD_HUMAN	31,446
417	Transforming protein RhoA precursor	RHOA_HUMAN	21,750
418	Calpain small subunit 1	CPNS1_HUMAN	28,299
419	ATP-dependent RNA helicase A	DHX9_HUMAN	140,944
420	Importin subunit beta-1	IMB1_HUMAN	97,153
421	Calpastatin	ICAL_HUMAN	76,557
422	Ras GTPase-activating-like protein IQGAP1	IQGA1_HUMAN	189,241
423	Protein NipSnap2	NIPS2_HUMAN	33,725
424	Nebulette	NEBL_HUMAN	116,438
425	Methylmalonate-semialdehyde dehydrogenase [acylating], mitochondrial precursor	MMSA_HUMAN	57,822
426	Drebrin	DREB_HUMAN	71,407
427	Chitinase-3-like protein 1 precursor	CH3L1_HUMAN	42,609
428	Cell adhesion molecule 3 precursor	CADM3_HUMAN	43,283
429	FK506-binding protein 3	FKBP3_HUMAN	25,159
430	Integrin beta-1 precursor	ITB1_HUMAN	88,447
431	4-aminobutyrate aminotransferase, mitochondrial precursor	GABT_HUMAN	56,423
432	Serpin H1 precursor	SERPH_HUMAN	46,424
433	Apolipoprotein C-III precursor	APOC3_HUMAN	10,834
434	Plastin-2	PLSL_HUMAN	70,274
435	Alpha-actinin-4	ACTN4_HUMAN	104,839
436	MAP6 domain-containing protein 1	MA6D1_HUMAN	20,987
437	WAS/WASL-interacting protein family member 2	WIPF2_HUMAN	46,272
438	6-phosphogluconolactonase	6PGL_HUMAN	27,530
439	Prolargin precursor	PRELP_HUMAN	43,794
440	Tropomodulin-2	TMOD2_HUMAN	39,578
441	Adenylate kinase isoenzyme 4, mitochondrial	KAD4_HUMAN	25,251
442	GTP:AMP phosphotransferase mitochondrial	KAD3_HUMAN	25,548
443	Transaldolase	TALDO_HUMAN	37,524
444	Vacuolar proton pump subunit G 2	VATG2_HUMAN	13,586
445	Lactoylglutathione lyase	LGUL_HUMAN	20,761
446	NSFL1 cofactor p47	NSF1C_HUMAN	40,555
447	GTP-binding nuclear protein Ran	RAN_HUMAN	24,405
448	Gamma-synuclein	SYUG_HUMAN	13,312
449	Voltage-dependent anion-selective channel protein 1	VDAC1_HUMAN	30,756
450	Endoplasmic reticulum protein ERp29 precursor	ERP29_HUMAN	28,977

identified proteins of <i>Glioblastoma multiforme</i> # 451-493			
nr.	protein description	accession numbers	mass [Da]
451	Ras-related protein Rap-1b precursor	RAP1B_HUMAN	20,807
452	Lambda-crystallin homolog	CRYL1_HUMAN	35,401
453	Mannose-6-phosphate isomerase	MANA_HUMAN	46,639
454	Cytochrome b5 type B precursor	CYB5B_HUMAN	16,314
455	Trifunctional enzyme subunit beta, mitochondrial precursor	ECHB_HUMAN	51,278
456	Neuronal growth regulator 1 precursor	NEGR1_HUMAN	38,701
457	Drebrin-like protein	DBNL_HUMAN	48,188
458	Fatty acid synthase	FAS_HUMAN	273,382
459	Pyridoxal kinase	PDXK_HUMAN	35,084
460	F-actin-capping protein subunit beta	CAPZB_HUMAN	31,334
461	Dipeptidyl-peptidase 2 precursor	DPP2_HUMAN	54,311
462	Proactivator polypeptide precursor [Contains: Saposin-A	SAP_HUMAN	58,094
463	Peptidyl-prolyl cis-trans isomerase NIMA-interacting 1	PIN1_HUMAN	18,226
464	Glutathione S-transferase Mu 3	GSTM3_HUMAN	26,543
465	Spectrin beta chain, brain 2	SPTN2_HUMAN	271,278
466	Succinate dehydrogenase [ubiquinone] iron-sulfur subunit, mitochondrial precursor	DHSB_HUMAN	31,613
467	Far upstream element-binding protein 2	FUBP2_HUMAN	73,129
468	EH domain-containing protein 3	EHD3_HUMAN	61,881
469	Nuclease sensitive element-binding protein 1	YBOX1_HUMAN	35,906
470	Excitatory amino acid transporter 2	EAA2_HUMAN	62,088
471	NADH dehydrogenase [ubiquinone] flavoprotein 1, mitochondrial precursor	NDUV1_HUMAN	50,800
472	BAG family molecular chaperone regulator 3	BAG3_HUMAN	61,575
473	Sodium/potassium-transporting ATPase subunit alpha-3	AT1A3_HUMAN	111,734
474	Adenylyl cyclase-associated protein 1	CAP1_HUMAN	51,838
475	Vesicular integral-membrane protein VIP36 precursor	LMAN2_HUMAN	40,212
476	Cytochrome c oxidase subunit 5B, mitochondrial precursor	COX5B_HUMAN	13,678
477	Secernin-1	SCRN1_HUMAN	46,364
478	Actin-related protein 2/3 complex subunit 4	ARPC4_HUMAN	19,649
479	Crk-like protein	CRKL_HUMAN	33,759
480	Dynamin-1-like protein	DNM1L_HUMAN	81,861
481	Transketolase	TKT_HUMAN	67,861
482	Glucosidase 2 subunit beta precursor	GLU2B_HUMAN	59,408
483	Actin-related protein 3	ARP3_HUMAN	47,354
484	Protein enabled homolog	ENAH_HUMAN	66,493
485	FK506-binding protein 2 precursor	FKBP2_HUMAN	15,632
486	Transmembrane protein 65	TMM65_HUMAN	25,526
487	Protein S100-A10	S10AA_HUMAN	11,186
488	Paralemmin	PALM_HUMAN	42,057
489	SH3-containing GRB2-like protein 3-interacting protein 1	SGIP1_HUMAN	89,093
490	SRA stem-loop-interacting RNA-binding protein, mitochondrial precursor	SLIRP_HUMAN	12,331
491	26S proteasome non-ATPase regulatory subunit 2	PSMD2_HUMAN	100,184
492	Hexokinase-1	HXK1_HUMAN	102,470
493	Puromycin-sensitive aminopeptidase	PSA_HUMAN	103,261

identified proteins of <i>Glioblastoma multiforme</i> # 494-537			
nr.	protein description	accession numbers	mass [Da]
494	Calcium-binding mitochondrial carrier protein Aralar1	CMC1_HUMAN	74,740
495	Hematopoietic lineage cell-specific protein	HCLS1_HUMAN	53,979
496	Vacuolar protein sorting-associated protein 35	VPS35_HUMAN	91,692
497	Hippocalcin-like protein 4	HPCL4_HUMAN	22,186
498	Cytochrome c oxidase subunit 2	COX2_HUMAN	25,548
499	Nidogen-2 precursor	NID2_HUMAN	151,377
500	Elongation factor 1-gamma	EF1G_HUMAN	50,101
501	Mitochondrial precursor proteins import receptor	TOM70_HUMAN	67,439
502	Heterogeneous nuclear ribonucleoprotein D0	HNRPD_HUMAN	38,417
503	Mannose-6-phosphate receptor-binding protein 1	M6PBP_HUMAN	47,028
504	Sorbin and SH3 domain-containing protein 1	SRBS1_HUMAN	142,441
505	Succinyl-CoA ligase [ADP-forming] beta-chain, mitochondrial precursor	SUCB1_HUMAN	50,300
506	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 8, mitochondrial precursor	NDUB8_HUMAN	21,748
507	Microtubule-associated protein RP/EB family member 1	MARE1_HUMAN	29,982
508	ES1 protein homolog, mitochondrial precursor	ES1_HUMAN	28,152
509	Talin-2	TLN2_HUMAN	271,535
510	Eukaryotic translation initiation factor 5A-1	IF5A1_HUMAN	16,815
511	Glucose-6-phosphate 1-dehydrogenase	G6PD_HUMAN	59,240
512	Cholecystokinins precursor	CCKN_HUMAN	12,652
513	Heat shock 70 kDa protein 12A	HS12A_HUMAN	74,962
514	Neurocalcin-delta	NCALD_HUMAN	22,228
515	6-phosphofructokinase, muscle type	K6PF_HUMAN	85,166
516	Epoxide hydrolase 1	HYEP_HUMAN	52,933
517	Mimitin, mitochondrial precursor	MIMIT_HUMAN	19,839
518	Endonuclease domain-containing 1 protein precursor	ENDD1_HUMAN	55,000
519	Ribonuclease inhibitor	RINI_HUMAN	49,956
520	Aquaporin-1	AQP1_HUMAN	28,509
521	Catenin beta-1	CTNB1_HUMAN	85,479
522	High mobility group protein B1	HMGB1_HUMAN	24,877
523	Endophilin-B2	SHLB2_HUMAN	43,957
524	Heat shock protein 105 kDa	HS105_HUMAN	96,848
525	Carbonic anhydrase 4 precursor	CAH4_HUMAN	35,015
526	26S protease regulatory subunit 6B	PRS6B_HUMAN	47,350
527	Septin-5	SEPT5_HUMAN	42,760
528	Alpha-1B-glycoprotein precursor	A1BG_HUMAN	54,254
529	Dynactin subunit 2	DCTN2_HUMAN	44,214
530	Glycolipid transfer protein	GLTP_HUMAN	23,833
531	Neuronal membrane glycoprotein M6-a	GPM6A_HUMAN	31,192
532	Calcium/calmodulin-dependent protein kinase type II delta chain	KCC2D_HUMAN	56,353
533	14 kDa phosphohistidine phosphatase	PHP14_HUMAN	13,815
534	Plasminogen precursor	PLMN_HUMAN	90,549
535	Pro-low-density lipoprotein receptor-related protein 1 precursor	LRP1_HUMAN	504,543
536	Succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial precursor	DHSA_HUMAN	72,674
537	Tubulin polymerization-promoting protein family member 3	TPPP3_HUMAN	18,968

identified proteins of <i>Glioblastoma multiforme</i> # 538-578			
nr.	protein description	accession numbers	mass [Da]
538	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 9, mitochondrial precursor	NDUA9_HUMAN	42,492
539	Lysosomal acid phosphatase precursor	PPAL_HUMAN	48,327
540	Guanine nucleotide-binding protein G(i), alpha-2 subunit	GNAI2_HUMAN	40,434
541	Flotillin-1	FLOT1_HUMAN	47,337
542	Collagen alpha-2(VI) chain precursor	CO6A2_HUMAN	108,563
543	Lamina-associated polypeptide 2, isoforms beta/gamma	LAP2A_HUMAN, LAP2B_HUMAN	50,653
544	Alcohol dehydrogenase [NADP+]	AK1A1_HUMAN	36,556
545	Protein FAM123A	F123A_HUMAN	69,532
546	ATPase inhibitor, mitochondrial precursor	ATIF1_HUMAN	12,231
547	Vacuolar proton pump subunit F	VATF_HUMAN	13,353
548	Membrane-associated progesterone receptor component 2	PGRC2_HUMAN	23,801
549	Pyruvate dehydrogenase E1 component subunit beta, mitochondrial precursor	ODPB_HUMAN	39,215
550	Platelet-activating factor acetylhydrolase IB subunit beta	PA1B2_HUMAN	25,552
551	Nucleosome assembly protein 1-like 1	NP1L1_HUMAN	45,357
552	26S proteasome non-ATPase regulatory subunit 13	PSD13_HUMAN	42,901
553	Ubiquitin-conjugating enzyme E2 N	UBE2N_HUMAN	17,121
554	Palmitoyl-protein thioesterase 1 precursor	PPT1_HUMAN	34,176
555	LanC-like protein 1	LANC1_HUMAN	45,267
556	Cytochrome c oxidase polypeptide VIIa-liver/heart, mitochondrial precursor	CX7A2_HUMAN	9,379
557	Lumican precursor	LUM_HUMAN	38,414
558	Galectin-1	LEG1_HUMAN	14,698
559	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 12	NDUAC_HUMAN	17,096
560	PC4 and SFRS1-interacting protein	PSIP1_HUMAN	60,086
561	NADH dehydrogenase [ubiquinone] iron-sulfur protein 8, mitochondrial precursor	NDUS8_HUMAN	23,688
562	Myelin-oligodendrocyte glycoprotein precursor	MOG_HUMAN	28,162
563	Uncharacterized protein C10orf35	CJ035_HUMAN	13,220
564	10-formyltetrahydrofolate dehydrogenase	FTHFD_HUMAN	98,812
565	Ubiquilin-2	UBQL2_HUMAN	65,679
566	Cytoskeleton-associated protein 4	CKAP4_HUMAN	66,004
567	6-phosphofructokinase, liver type	K6PL_HUMAN	85,001
568	Vacuolar ATP synthase subunit S1 precursor	VAS1_HUMAN	52,009
569	Phospholipid transfer protein precursor	PLTP_HUMAN	54,723
570	Polypyrimidine tract-binding protein 1	PTBP1_HUMAN	57,205
571	UTP--glucose-1-phosphate uridylyltransferase	UGPA_HUMAN	56,924
572	Tenascin precursor	TENA_HUMAN	240,845
573	Sodium/calcium exchanger 2 precursor	NAC2_HUMAN	100,351
574	NADH-ubiquinone oxidoreductase 75 kDa subunit, mitochondrial precursor	NDUS1_HUMAN	79,451
575	Alpha-centractin	ACTZ_HUMAN	42,597
576	Ras-related protein Rab-31	RAB31_HUMAN	21,551
577	ATP synthase subunit O, mitochondrial precursor	ATPO_HUMAN	23,259
578	Heterogeneous nuclear ribonucleoprotein G	HNRPG_HUMAN	42,316

identified proteins of <i>Glioblastoma multiforme</i> # 579-620			
nr.	protein description	accession numbers	mass [Da]
579	Melanoma inhibitory activity protein 3 precursor	MIA3_HUMAN	213,687
580	Protein RUFY3	RUFY3_HUMAN	52,949
581	Protein ERGIC-53 precursor	LMAN1_HUMAN	57,531
582	Dipeptidyl aminopeptidase-like protein 6	DPP6_HUMAN	97,573
583	Beta-hexosaminidase alpha chain precursor	HEXA_HUMAN	60,672
584	Mitochondrial import receptor subunit TOM22 homolog	TOM22_HUMAN	15,504
585	Hepatoma-derived growth factor-related protein 3	HDGR3_HUMAN	22,603
586	Cation-dependent mannose-6-phosphate receptor precursor	MPRD_HUMAN	30,975
587	Casein kinase II subunit alpha	CSK21_HUMAN	45,127
588	T-complex protein 1 subunit zeta	TCPZ_HUMAN	58,007
589	Thrombospondin-1 precursor	TSP1_HUMAN	129,364
590	Ganglioside-induced differentiation-associated protein 1	GDAP1_HUMAN	41,235
591	Serine/threonine-protein kinase DCLK2	DCLK2_HUMAN	83,595
592	Cathepsin B precursor	CATB_HUMAN	37,803
593	Synapsin-3	SYN3_HUMAN	63,285
594	Mitochondrial carrier homolog 2	MTCH2_HUMAN	33,314
595	Adipocyte plasma membrane-associated protein	APMAP_HUMAN	46,464
596	Medium-chain specific acyl-CoA dehydrogenase, mitochondrial precursor	ACADM_HUMAN	46,572
597	Lysosomal Pro-X carboxypeptidase precursor	PCP_HUMAN	55,783
598	Sorting nexin-3	SNX3_HUMAN	18,745
599	Gamma-soluble NSF attachment protein	SNAG_HUMAN	34,729
600	D-dopachrome decarboxylase	DOPD_HUMAN	12,694
601	Galectin-related protein	LEGL_HUMAN	18,968
602	La-related protein 1	LARP1_HUMAN	123,495
603	Selenium-binding protein 1	SBP1_HUMAN	52,374
604	Neuronal calcium sensor 1	NCS1_HUMAN	21,862
605	Kininogen-1 precursor	KNG1_HUMAN	71,940
606	Ras-related protein Rab-10	RAB10_HUMAN	22,524
607	Thioredoxin domain-containing protein 4 precursor	TXND4_HUMAN	46,955
608	PDZ and LIM domain protein 5	PDLI5_HUMAN	63,953
609	SH3 and multiple ankyrin repeat domains protein 3	SHAN3_HUMAN	186,282
610	Emerin	EMD_HUMAN	28,977
611	Dynamin-like 120 kDa protein, mitochondrial precursor	OPA1_HUMAN	111,643
612	Elongation factor 2	EF2_HUMAN	95,322
613	WD repeat-containing protein 1	WDR1_HUMAN	66,175
614	NAD(P) transhydrogenase, mitochondrial precursor	NNTM_HUMAN	113,881
615	Neural cell adhesion molecule 2 precursor	NCAM2_HUMAN	92,916
616	Very long-chain specific acyl-CoA dehydrogenase, mitochondrial precursor	ACADV_HUMAN	70,374
617	60S ribosomal protein L15	RL15_HUMAN	24,129
618	Complement C4-B precursor	CO4A_HUMAN, CO4B_HUMAN	192,777
619	60S ribosomal protein L10	RL10_HUMAN	24,587
620	Estradiol 17-beta-dehydrogenase 12	DHB12_HUMAN	34,307

identified proteins of <i>Glioblastoma multiforme</i> # 621-662			
nr.	protein description	accession numbers	mass [Da]
621	Major vault protein	MVP_HUMAN	99,308
622	Dystroglycan precursor	DAG1_HUMAN	97,563
623	Rab GDP dissociation inhibitor beta	GDIB_HUMAN	50,648
624	Kinectin	KTN1_HUMAN	156,258
625	Serine/threonine-protein kinase DCLK1	DCLK1_HUMAN	82,208
626	Growth factor receptor-bound protein 2	GRB2_HUMAN	25,189
627	Cytochrome c oxidase polypeptide VIc precursor	COX6C_HUMAN	8,764
628	2,4-dienoyl-CoA reductase, mitochondrial precursor	DECR_HUMAN	36,051
629	Proto-oncogene C-crk	CRK_HUMAN	33,813
630	Heat shock protein beta-6	HSPB6_HUMAN	17,117
631	Splicing factor, arginine/serine-rich 1	SFRS1_HUMAN	27,727
632	Metaxin-2	MTX2_HUMAN	29,745
633	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 2	NDUA2_HUMAN	10,904
634	Neurabin-2	NEB2_HUMAN	89,174
635	Dual specificity mitogen-activated protein kinase kinase 1	MP2K1_HUMAN	43,422
636	Serine/threonine-protein phosphatase 2B catalytic subunit beta isoform	PP2BB_HUMAN	59,007
637	Plasma membrane calcium-transporting ATPase 1	AT2B1_HUMAN	138,741
638	Poliovirus receptor-related protein 1 precursor	PVRL1_HUMAN	57,140
639	Copine-3	CPNE3_HUMAN	60,114
640	p130Cas-associated protein	SNIP_HUMAN	112,450
641	Nucleoside diphosphate kinase A	NDKA_HUMAN	17,131
642	Adenine phosphoribosyltransferase	APT_HUMAN	19,591
643	RNA-binding protein FUS	FUS_HUMAN	53,408
644	Nicastrin precursor	NICA_HUMAN	78,394
645	Syntaxin-7	STX7_HUMAN	29,798
646	ADP-ribosylation factor GTPase-activating protein 1	ARFG1_HUMAN	44,649
647	NADH dehydrogenase [ubiquinone] iron-sulfur protein 5	NDUS5_HUMAN	12,500
648	SH3 domain-binding glutamic acid-rich-like protein 3	SH3L3_HUMAN	10,420
649	Uncharacterized protein C2orf32	CB032_HUMAN	18,630
650	Dynein light chain 1, cytoplasmic	DYL1_HUMAN, DYL2_HUMAN	10,348
651	Guanine nucleotide-binding protein G(q) subunit alpha	GNAQ_HUMAN	41,451
652	T-complex protein 1 subunit alpha	TCPA_HUMAN	60,327
653	Target of Myb protein 1	TOM1_HUMAN	53,801
654	Succinyl-CoA:3-ketoacid-coenzyme A transferase 1, mitochondrial precursor	SCOT_HUMAN	56,141
655	Mannosyl-oligosaccharide glucosidase	GCS1_HUMAN	91,901
656	Phospholipase D3	PLD3_HUMAN	54,688
657	Leucine-rich PPR motif-containing protein, mitochondrial precursor	LPPRC_HUMAN	157,894
658	Integrin alpha-IIb precursor	ITA2B_HUMAN	113,373
659	Protein kinase C alpha type	KPCA_HUMAN	76,747
660	Ras-related protein Rab-27B	RB27B_HUMAN	24,591
661	Voltage-dependent anion-selective channel protein 2	VDAC2_HUMAN	31,549
662	Cold-inducible RNA-binding protein	CIRBP_HUMAN	18,630

identified proteins of <i>Glioblastoma multiforme</i> # 663-702			
nr.	protein description	accession numbers	mass [Da]
663	Oligodendrocyte-myelin glycoprotein precursor	OMGP_HUMAN	49,592
664	Chloride intracellular channel protein 4	CLIC4_HUMAN	28,756
665	Solute carrier family 12 member 5	S12A5_HUMAN	126,168
666	Cytoplasmic dynein 1 light intermediate chain 1	DC1L1_HUMAN	56,562
667	Platelet-activating factor acetylhydrolase IB subunit gamma	PA1B3_HUMAN	25,716
668	Single-stranded DNA-binding protein, mitochondrial precursor	SSB_HUMAN	17,242
669	Leukocyte elastase inhibitor	ILEU_HUMAN	42,726
670	Heterogeneous nuclear ribonucleoproteins C1/C2	HNRPC_HUMAN	33,653
671	Synaptophysin	SYPH_HUMAN	33,827
672	SPARC-like protein 1 precursor	SPRL1_HUMAN	75,198
673	5'(3')-deoxyribonucleotidase, cytosolic type	NT5C_HUMAN	23,365
674	Heterogeneous nuclear ribonucleoprotein A0	ROA0_HUMAN	30,823
675	Band 4.1-like protein 2	E41L2_HUMAN	112,570
676	HLA class II histocompatibility antigen, DRB1-1 beta chain precursor	2B11_HUMAN, 2B14_HUMAN, 2B1A_HUMAN, HB2C_HUMAN	29,896
677	Protein NipSnap1	NIPS1_HUMAN	33,293
678	Gamma-glutamyltransferase 5 precursor	GGT5_HUMAN	62,243
679	FXD domain-containing ion transport regulator 6 precursor	FXD6_HUMAN	10,525
680	Erythrocyte band 7 integral membrane protein	STOM_HUMAN	31,714
681	Sorting and assembly machinery component 50 homolog	SAM50_HUMAN	51,960
682	DmX-like protein 2	DMXL2_HUMAN	339,740
683	Transmembrane protein 43	TMM43_HUMAN	44,859
684	Immunity-related GTPase family Q protein	IRGQ_HUMAN	62,701
685	40S ribosomal protein S4, X isoform	RS4X_HUMAN	29,581
686	NADH-cytochrome b5 reductase 1	NB5R1_HUMAN	34,078
687	UV excision repair protein RAD23 homolog B	RD23B_HUMAN	43,153
688	Cell cycle exit and neuronal differentiation protein 1	CEND_HUMAN	14,936
689	Dihydrolipoyl dehydrogenase, mitochondrial precursor	DLDH_HUMAN	54,132
690	ADP/ATP translocase 1	ADT1_HUMAN	33,047
691	Cdc42 effector protein 4	BORG4_HUMAN	37,963
692	Renin receptor precursor	RENH_HUMAN	38,991
693	Heme oxygenase 1	HMOX1_HUMAN	32,801
694	Gap junction alpha-1 protein	CXA1_HUMAN	42,991
695	Nucleobindin-1 precursor	NUCB1_HUMAN	53,862
696	Tryptophanyl-tRNA synthetase, cytoplasmic	SYWC_HUMAN	53,150
697	COP9 signalosome complex subunit 6	CSN6_HUMAN	36,145
698	Integrin beta-2 precursor	ITB2_HUMAN	84,764
699	Polyadenylate-binding protein 1	PABP1_HUMAN	70,653
700	Arsenical pump-driving ATPase	ARSA1_HUMAN	38,776
701	Glycerol-3-phosphate dehydrogenase [NAD+], cytoplasmic	GPDA_HUMAN	37,550
702	Proline synthetase co-transcribed bacterial homolog protein	PROSC_HUMAN	30,326

identified proteins of <i>Glioblastoma multiforme</i> # 703-748			
nr.	protein description	accession numbers	mass [Da]
703	3-hydroxyisobutyrate dehydrogenase, mitochondrial precursor	3HIDH_HUMAN	35,312
704	Chloride intracellular channel protein 1	CLIC1_HUMAN	26,905
705	Gamma-adducin	ADDG_HUMAN	79,138
706	Alpha-2-antiplasmin precursor	A2AP_HUMAN	54,549
707	Serine/threonine-protein phosphatase PP1-gamma catalytic subunit	PP1A_HUMAN, PP1G_HUMAN	36,968
708	Aflatoxin B1 aldehyde reductase member 2	ARK72_HUMAN	39,571
709	Ubiquitin carboxyl-terminal hydrolase 14	UBP14_HUMAN	56,052
710	Protein piccolo	PCLO_HUMAN	566,639
711	26S protease regulatory subunit 6A	PRS6A_HUMAN	49,187
712	Alpha-2-macroglobulin precursor	A2MG_HUMAN	163,259
713	ADP-ribosylation factor-like protein 3	ARL3_HUMAN	20,438
714	Collagen alpha-1(VI) chain precursor	CO6A1_HUMAN	108,513
715	NADH dehydrogenase [ubiquinone] iron-sulfur protein 2, mitochondrial precursor	NDUS2_HUMAN	52,529
716	3-ketoacyl-CoA thiolase, mitochondrial	THIM_HUMAN	41,906
717	AP-2 complex subunit alpha-1	AP2A1_HUMAN	107,540
718	Neurofilament heavy polypeptide	NFH_HUMAN	112,463
719	Translationally-controlled tumor protein	TCTP_HUMAN	19,578
720	Centrosomal protein of 170 kDa	CE170_HUMAN	175,278
721	Activated RNA polymerase II transcriptional coactivator p15	TCP4_HUMAN	14,378
722	Hemopexin precursor	HEMO_HUMAN	51,659
723	Beta-2-microglobulin precursor [Contains: Beta-2-microglobulin form pI 5.3]	B2MG_HUMAN	13,697
724	Cysteine-rich protein 2	CRIP2_HUMAN	22,475
725	Nitrilase homolog 1	NIT1_HUMAN	35,879
726	Heterogeneous nuclear ribonucleoprotein H3	HNRH3_HUMAN	36,910
727	Syntaxin-12	STX12_HUMAN	31,625
728	Apolipoprotein B-100 precursor	APOB_HUMAN	515,554
729	Lamin-A/C	LMNA_HUMAN	74,123
730	Disks large homolog 4	DLG4_HUMAN	80,479
731	Protein KIAA0284	K0284_HUMAN	171,671
732	Junctional adhesion molecule C precursor	JAM3_HUMAN	35,002
733	DNA-dependent protein kinase catalytic subunit	PRKDC_HUMAN	469,078
734	Guanine nucleotide-binding protein alpha-13 subunit	GNA13_HUMAN	44,033
735	Beta-Ala-His dipeptidase precursor	CNDP1_HUMAN	56,675
736	6-phosphofructokinase type C	K6PP_HUMAN	85,579
737	Fascin	FSCN1_HUMAN	54,512
738	COP9 signalosome complex subunit 7a	CSN7A_HUMAN	30,259
739	Acylphosphatase-2	ACYP2_HUMAN	11,121
740	Ubiquitin-associated protein 2-like	UBP2L_HUMAN	114,516
741	AP-2 complex subunit mu-1	AP2M1_HUMAN	49,638
742	LanC-like protein 2	LANC2_HUMAN	50,838
743	Myosin light chain kinase, smooth muscle	MYLK_HUMAN	210,754
744	40S ribosomal protein S8	RS8_HUMAN	24,188
745	NADH dehydrogenase [ubiquinone] flavoprotein 3, mitochondrial precursor	NDUV3_HUMAN	11,923
746	Ras GTPase-activating protein-binding protein 2	G3BP2_HUMAN	54,102
747	Cytoplasmic dynein 1 light intermediate chain 2	DC1L2_HUMAN	54,082
748	ProSAAS precursor	PCSK1_HUMAN	27,355

identified proteins of <i>Glioblastoma multiforme</i> # 749-788			
nr.	protein description	accession numbers	mass [Da]
749	Interferon-inducible double stranded RNA-dependent protein kinase activator A	PRKRA_HUMAN	34,387
750	Secretory carrier-associated membrane protein 1	SCAM1_HUMAN	37,903
751	Heterogeneous nuclear ribonucleoprotein L	HNRPL_HUMAN	60,169
752	Tubulin folding cofactor B	TBCB_HUMAN	27,308
753	Transcription elongation factor A protein-like 5	TCAL5_HUMAN	23,289
754	Transcription intermediary factor 1-beta	TIF1B_HUMAN	88,531
755	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1	GBB1_HUMAN	37,360
756	T-complex protein 1 subunit epsilon	TCPE_HUMAN	59,654
757	Protein FAM49A	FA49A_HUMAN	37,296
758	Receptor-type tyrosine-protein phosphatase-like N precursor	PTPRN_HUMAN	105,830
759	Heterogeneous nuclear ribonucleoprotein M	HNRPM_HUMAN	77,499
760	Poly(rC)-binding protein 1	PCBP1_HUMAN	37,480
761	Tropomyosin alpha-3 chain	TPM3_HUMAN	32,802
762	Cytochrome b5	CYB5_HUMAN	15,312
763	Dihydropyrimidinase-related protein 4	DPYL4_HUMAN	61,859
764	Tyrosine-protein phosphatase non-receptor type 11	PTN11_HUMAN	68,418
765	40S ribosomal protein S19	RS19_HUMAN	16,043
766	E2-induced gene 5 protein	E2IG5_HUMAN	17,325
767	Splicing factor, arginine/serine-rich 3	SFRS3_HUMAN	19,312
768	T-complex protein 1 subunit delta	TCPD_HUMAN	57,908
769	26S proteasome non-ATPase regulatory subunit 8	PSMD8_HUMAN	29,989
770	Ras GTPase-activating protein-binding protein 1	G3BP1_HUMAN	52,145
771	Hydroxyacylglutathione hydrolase	GLO2_HUMAN	28,842
772	Electron transfer flavoprotein subunit alpha, mitochondrial precursor	ETFA_HUMAN	35,062
773	TATA-binding protein-associated factor 2N	RBP56_HUMAN	61,813
774	Glutathione S-transferase theta-2	GSTT2_HUMAN	27,490
775	Zinc-alpha-2-glycoprotein precursor	ZA2G_HUMAN	33,854
776	Isoamyl acetate-hydrolyzing esterase 1 homolog	IAH1_HUMAN	27,582
777	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 5	NDUA5_HUMAN	13,441
778	Prostaglandin E synthase 3	TEBP_HUMAN	18,680
779	Caldesmon	CALD1_HUMAN	93,233
780	Collagen alpha-2(IV) chain precursor [Contains: Canstatin]	CO4A2_HUMAN	167,540
781	Zyxin	ZYX_HUMAN	61,258
782	General transcription factor II-I	GTF2I_HUMAN	112,400
783	Pyridoxine-5'-phosphate oxidase	PNPO_HUMAN	29,970
784	60S ribosomal protein L6	RL6_HUMAN	32,711
785	Thymidine phosphorylase precursor	TYPH_HUMAN	49,938
786	Leucine zipper-EF-hand-containing transmembrane protein 1, mitochondrial precursor	LETM1_HUMAN	83,338
787	Laminin subunit gamma-1 precursor	LAMC1_HUMAN	177,587
788	Glycerol-3-phosphate dehydrogenase, mitochondrial precursor	GPDM_HUMAN	80,818

identified proteins of <i>Glioblastoma multiforme</i> # 789-832			
nr.	protein description	accession numbers	mass [Da]
789	Hydroxyacyl-coenzyme A dehydrogenase, mitochondrial precursor	HCDH_HUMAN	34,260
790	Acyl-coenzyme A thioesterase 1	ACOT1_HUMAN	46,260
791	Alanyl-tRNA synthetase, cytoplasmic	SYAC_HUMAN	106,795
792	Armadillo repeat-containing protein 10	ARM10_HUMAN	37,523
793	T-cell immunomodulatory protein precursor	TIP_HUMAN	68,092
794	60S ribosomal protein L19	RL19_HUMAN	23,449
795	Solute carrier family 2, facilitated glucose transporter member 3	GTR3_HUMAN	53,908
796	Protachykinin 1 precursor	TKN1_HUMAN	14,986
797	Vitamin D-binding protein precursor	VTDB_HUMAN	52,947
798	Filamin-B	FLNB_HUMAN	278,172
799	Protein tyrosine kinase 2 beta	FAK2_HUMAN	115,860
800	Adenylate kinase isoenzyme 2, mitochondrial	KAD2_HUMAN	26,461
801	NADH dehydrogenase [ubiquinone] flavoprotein 2, mitochondrial precursor	NDUV2_HUMAN	27,374
802	Serine/threonine-protein phosphatase 2A catalytic subunit alpha isoform	PP2AA_HUMAN, PP2AB_HUMAN	35,577
803	Solute carrier family 12 member 2	S12A2_HUMAN	131,434
804	COP9 signalosome complex subunit 3	CSN3_HUMAN	47,857
805	Proteasome activator complex subunit 2	PSME2_HUMAN	27,344
806	Intercellular adhesion molecule 1 precursor	ICAM1_HUMAN	57,807
807	Inter-alpha-trypsin inhibitor heavy chain H1 precursor	ITIH1_HUMAN	101,372
808	Ceruloplasmin precursor	CERU_HUMAN	122,190
809	Cytochrome c oxidase subunit 5A, mitochondrial precursor	COX5A_HUMAN	16,757
810	Proline-rich transmembrane protein 2	PRRT2_HUMAN	34,926
811	Heterogeneous nuclear ribonucleoprotein H2	HNRH2_HUMAN	49,246
812	Cytochrome b-c1 complex subunit 8	QCR8_HUMAN	9,889
813	3-mercaptopyruvate sulfurtransferase	THTM_HUMAN	33,161
814	Heme-binding protein 1	HEBP1_HUMAN	21,079
815	Neuronal pentraxin-1 precursor	NPTX1_HUMAN	47,105
816	Serum deprivation-response protein	SDPR_HUMAN	47,155
817	Sodium/hydrogen exchanger 6	SL9A6_HUMAN	74,146
818	Neuropeptide Y precursor [Contains: Neuropeptide Y	NPY_HUMAN	10,834
819	Diablo homolog, mitochondrial precursor	DBLOH_HUMAN	27,113
820	Hypoxanthine-guanine phosphoribosyltransferase	HPRT_HUMAN	24,562
821	Radixin	RADI_HUMAN	68,548
822	Splicing factor 1	SF01_HUMAN	68,313
823	Isochorismatase domain-containing protein 1	ISOC1_HUMAN	32,219
824	UDP-glucose:glycoprotein glucosyltransferase 1 precursor	UGGG1_HUMAN	174,965
825	26S protease regulatory subunit 4	PRS4_HUMAN	49,168
826	Lethal(2) giant larvae protein homolog 1	L2GL1_HUMAN	115,024
827	Isocitrate dehydrogenase [NAD] subunit beta, mitochondrial precursor	IDH3B_HUMAN	42,166
828	Ubiquilin-1	UBQL1_HUMAN	62,502
829	Interleukin enhancer-binding factor 2	ILF2_HUMAN	43,045
830	DnaJ homolog subfamily C member 5	DNJC5_HUMAN	22,131
831	Acylphosphatase-1	ACYP1_HUMAN	11,243
832	Obg-like ATPase 1	OLA1_HUMAN	44,727

identified proteins of <i>Glioblastoma multiforme</i> # 833-873			
nr.	protein description	accession numbers	mass [Da]
833	Actin-related protein 2/3 complex subunit 2	ARPC2_HUMAN	34,316
834	ATP-dependent DNA helicase 2 subunit 2	KU86_HUMAN	82,689
835	TBC1 domain family member 13	TBC13_HUMAN	32,143
836	ATP-binding cassette sub-family D member 3	ABCD3_HUMAN	75,461
837	Transforming growth factor-beta-induced protein ig-h3 precursor	BGH3_HUMAN	74,665
838	ATP-citrate synthase	ACLY_HUMAN	120,825
839	28 kDa heat- and acid-stable phosphoprotein	HAP28_HUMAN	20,613
840	B-cell receptor-associated protein 31	BAP31_HUMAN	27,975
841	Complexin-2	CPLX2_HUMAN	15,377
842	Neutrophil defensin 1 precursor	DEF1_HUMAN, DEF3_HUMAN	10,183
843	Hematological and neurological expressed 1 protein	HN1_HUMAN	15,997
844	Plasma membrane calcium-transporting ATPase 2	AT2B2_HUMAN	136,862
845	Reticulon-3	RTN3_HUMAN	112,595
846	Tight junction protein ZO-1	ZO1_HUMAN	195,442
847	DnaJ homolog subfamily B member 2	DNJB2_HUMAN	35,563
848	Contactin-associated protein-like 2 precursor	CNTP2_HUMAN	148,151
849	Serum paraoxonase/arylesterase 2	PON2_HUMAN	39,381
850	Na(+)/H(+) exchange regulatory cofactor NHE-RF2	NHRF2_HUMAN	37,395
851	Ig kappa chain V-III region SIE	KV302_HUMAN, KV305_HUMAN	11,757
852	Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex, mitochondrial precursor	ODP2_HUMAN	65,764
853	Actin-binding LIM protein 2	ABLM2_HUMAN	67,795
854	Protein 4.1	41_HUMAN	97,000
855	Ubiquitin-conjugating enzyme E2 variant 2	UB2V2_HUMAN	16,345
856	IQ motif and Sec7 domain-containing protein 1	IQEC1_HUMAN	108,300
857	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 4	NDUA4_HUMAN	9,352
858	Delta-aminolevulinic acid dehydratase	HEM2_HUMAN	36,277
859	Tripartite motif-containing protein 2	TRIM2_HUMAN	81,513
860	Beta-hexosaminidase beta chain precursor	HEXB_HUMAN	63,095
861	G-protein coupled receptor 56 precursor	GPR56_HUMAN	77,721
862	Laminin subunit beta-1 precursor	LAMB1_HUMAN	198,045
863	Dynactin subunit 1	DCTN1_HUMAN	141,680
864	Ubiquitin-conjugating enzyme E2 L3	UB2L3_HUMAN	17,844
865	Pyruvate dehydrogenase E1 component subunit alpha, somatic form, mitochondrial precursor	ODPA_HUMAN	43,279
866	Arfaptin-2	ARFP2_HUMAN	37,839
867	ERO1-like protein alpha precursor	ERO1A_HUMAN	54,377
868	Amyloid beta A4 protein precursor	A4_HUMAN	86,923
869	Apolipoprotein C-I precursor	APOC1_HUMAN	9,314
870	Integrin-linked protein kinase	ILK_HUMAN	51,402
871	Delta(3,5)-Delta(2,4)-dienoyl-CoA isomerase, mitochondrial precursor	ECH1_HUMAN	35,798
872	Microtubule-actin cross-linking factor 1, isoforms 1/2/3/5	MACF1_HUMAN	620,397
873	LIM domain and actin-binding protein 1	LIMA1_HUMAN	85,208

identified proteins of <i>Glioblastoma multiforme</i> # 874-916			
nr.	protein description	accession numbers	mass [Da]
874	Proteasome activator complex subunit 1	PSME1_HUMAN	28,706
875	C-terminal-binding protein 1	CTBP1_HUMAN	47,517
876	AP-3 complex subunit sigma-1	AP3S1_HUMAN	21,715
877	Neurotrimin precursor	NTRI_HUMAN	37,953
878	COP9 signalosome complex subunit 4	CSN4_HUMAN	46,252
879	E3 ubiquitin-protein ligase HUWE1	HUWE1_HUMAN	481,874
880	Afadin	AFAD_HUMAN	205,592
881	Vacuolar proton pump subunit G 1	VATG1_HUMAN	13,740
882	GDH/6PGL endoplasmic bifunctional protein precursor [Includes: Glucose 1-dehydrogenase	G6PE_HUMAN	88,876
883	Exportin-1	XPO1_HUMAN	123,371
884	Eukaryotic initiation factor 4A-I	IF4A1_HUMAN	46,137
885	Atlastin-3	ATLA3_HUMAN	60,525
886	Peroxisomal multifunctional enzyme type 2	DHB4_HUMAN	79,670
887	Calretinin	CALB2_HUMAN	31,524
888	Purine nucleoside phosphorylase	PNPH_HUMAN	32,100
889	Extended-synaptotagmin-1	ESYT1_HUMAN	122,841
890	cAMP-dependent protein kinase, alpha-catalytic subunit	KAPCA_HUMAN	40,573
891	Protein tyrosine phosphatase-like protein PTPLAD1	PTAD1_HUMAN	43,143
892	ATP-dependent DNA helicase 2 subunit 1	KU70_HUMAN	69,828
893	Ubiquitin carboxyl-terminal hydrolase 7	UBP7_HUMAN	128,257
894	Presequence protease, mitochondrial precursor	PREP_HUMAN	117,439
895	Keratin, type II cytoskeletal 5	K2C5_HUMAN	62,362
896	Tricarboxylate transport protein, mitochondrial precursor	TXTP_HUMAN	33,995
897	Solute carrier family 2, facilitated glucose transporter member 1	GTR1_HUMAN	54,067
898	Calcium-regulated heat stable protein 1	CHSP1_HUMAN	15,874
899	Nicotinate phosphoribosyltransferase	PNCB_HUMAN	57,561
900	F-actin-capping protein subunit alpha-2	CAZA2_HUMAN	32,931
901	Importin subunit alpha-3	IMA3_HUMAN	57,793
902	FK506-binding protein 1A	FKB1A_HUMAN	11,933
903	Histone H1.4	H12_HUMAN, H13_HUMAN, H14_HUMAN	21,849
904	Putative RNA-binding protein 3	RBM3_HUMAN	17,153
905	Bisphosphoglycerate mutase	PMGE_HUMAN	29,988
906	40S ribosomal protein SA	RSSA_HUMAN	32,836
907	3-ketoacyl-CoA thiolase, peroxisomal precursor	THIK_HUMAN	44,274
908	Metaxin-1	MTX1_HUMAN	35,760
909	3,2-trans-enoyl-CoA isomerase, mitochondrial precursor	D3D2_HUMAN	32,799
910	Desmuslin	DMN_HUMAN	172,747
911	Charged multivesicular body protein 4b	CHM4B_HUMAN	24,933
912	Protein phosphatase 1F	PPM1F_HUMAN	49,812
913	Thioredoxin domain-containing protein 5 precursor	TXND5_HUMAN	47,611
914	Ig kappa chain V-IV region Len	KV402_HUMAN	12,622
915	Probable serine carboxypeptidase CPVL precursor	CPVL_HUMAN	54,148
916	Catenin delta-1	CTND1_HUMAN	108,154

identified proteins of <i>Glioblastoma multiforme</i> # 917-961			
nr.	protein description	accession numbers	mass [Da]
917	Fatty acid-binding protein, brain	FABPB_HUMAN	14,871
918	Apolipoprotein A-IV precursor	APOA4_HUMAN	45,381
919	Calponin-3	CNN3_HUMAN	36,397
920	Vasodilator-stimulated phosphoprotein	VASP_HUMAN	39,811
921	Caskin-1	CSKI1_HUMAN	149,797
922	Acyl-coenzyme A oxidase 1, peroxisomal	ACOX1_HUMAN	74,407
923	Putative tyrosine-protein phosphatase auxilin	AUXI_HUMAN	99,980
924	Actin-related protein 2/3 complex subunit 1A	ARC1A_HUMAN	41,551
925	Double-stranded RNA-specific adenosine deaminase	DSRAD_HUMAN	135,981
926	Interleukin enhancer-binding factor 3	ILF3_HUMAN	95,321
927	Phosphoserine phosphatase	SERB_HUMAN	24,990
928	DnaJ homolog subfamily A member 1	DNJA1_HUMAN	44,851
929	F-actin-capping protein subunit alpha-1	CAZA1_HUMAN	32,905
930	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 6	NDUA6_HUMAN	15,118
931	Muscleblind-like protein 1	MBNL1_HUMAN	41,799
932	Mitochondrial glutamate carrier 2	GHC2_HUMAN	33,832
933	Histidine triad nucleotide-binding protein 2	HINT2_HUMAN	17,144
934	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 8	NDUA8_HUMAN	20,087
935	DCN1-like protein 1	DCNL1_HUMAN	30,108
936	PDZ and LIM domain protein 4	PDLI4_HUMAN	35,380
937	N(G),N(G)-dimethylarginine dimethylaminohydrolase 2	DDAH2_HUMAN	29,626
938	Small glutamine-rich tetratricopeptide repeat-containing protein A	SGTA_HUMAN	34,046
939	Translocon-associated protein subunit alpha precursor	SSRA_HUMAN	32,218
940	Farnesyl pyrophosphate synthetase	FPPS_HUMAN	40,516
941	Mu-crystallin homolog	CRYM_HUMAN	33,757
942	Ras-related protein Rab-1B	RAB1B_HUMAN	22,154
943	Histidine-rich glycoprotein precursor	HRG_HUMAN	59,559
944	26S protease regulatory subunit S10B	PRS10_HUMAN	44,157
945	Cytosolic acyl coenzyme A thioester hydrolase	BACH_HUMAN	41,777
946	Heat shock protein 75 kDa, mitochondrial precursor	TRAP1_HUMAN	80,095
947	Actin-related protein 2	ARP2_HUMAN	44,744
948	Nuclear mitotic apparatus protein 1	NUMA1_HUMAN	238,242
949	Poly [ADP-ribose] polymerase 1	PARP1_HUMAN	113,070
950	Protein phosphatase methylesterase 1	PPME1_HUMAN	42,298
951	ATP synthase subunit f, mitochondrial	ATPK_HUMAN	10,900
952	Vigilin	VIGLN_HUMAN	141,424
953	Adipophilin	ADFP_HUMAN	48,057
954	Tubulin-specific chaperone A	TBCA_HUMAN	12,837
955	Plastin-3	PLST_HUMAN	70,421
956	Amine oxidase [flavin-containing] A	AOFA_HUMAN	59,665
957	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 63 kDa subunit precursor	RIB2_HUMAN	69,267
958	Vacuolar proton pump subunit H	VATH_HUMAN	55,865
959	Arylamide deacetylase-like 1	ADCL1_HUMAN	45,791
960	Prohibitin-2	PHB2_HUMAN	33,280
961	Biglycan precursor	PGS1_HUMAN	41,638

identified proteins of <i>Glioblastoma multiforme</i> # 962-1,003			
nr.	protein description	accession numbers	mass [Da]
962	Aspartyl/asparaginyl beta-hydroxylase	ASPH_HUMAN	85,845
963	Aspartyl-tRNA synthetase, cytoplasmic	SYDC_HUMAN	57,119
964	Mitochondrial 2-oxoglutarate/malate carrier protein	M2OM_HUMAN	34,045
965	Myosin-Va	MYO5A_HUMAN	215,411
966	Vesicle-trafficking protein SEC22b	SC22B_HUMAN	24,723
967	Glutathione synthetase	GSHB_HUMAN	52,368
968	Sarcoplasmic/endoplasmic reticulum calcium ATPase 2	AT2A2_HUMAN	114,741
969	MAP7 domain-containing protein 1	MA7D1_HUMAN	92,802
970	Protein BAT5	BAT5_HUMAN	63,226
971	Monocyte differentiation antigen CD14 precursor	CD14_HUMAN	40,059
972	Uncharacterized protein KIAA1671	K1671_HUMAN	196,693
973	60S ribosomal protein L7a	RL7A_HUMAN	29,978
974	PRA1 family protein 3	PRAF3_HUMAN	21,598
975	Misshapen-like kinase 1	MINK1_HUMAN	149,793
976	Synaptotagmin-7	SYT7_HUMAN	45,547
977	CDGSH iron sulfur domain-containing protein 1	CISD1_HUMAN	12,181
978	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 3	NDUA3_HUMAN	9,261
979	Protein dpy-30 homolog	DPY30_HUMAN	11,232
980	Phytanoyl-CoA hydroxylase-interacting protein	PHYIP_HUMAN	37,555
981	CD63 antigen	CD63_HUMAN	25,618
982	Abl interactor 1	ABI1_HUMAN	55,064
983	Glucosamine-6-phosphate isomerase	GNPI_HUMAN	32,651
984	UPF0424 protein C1orf128	CA128_HUMAN	24,160
985	Stomatin-like protein 2	STML2_HUMAN	38,517
986	Complexin-1	CPLX1_HUMAN	15,013
987	Dystrobrevin alpha	DTNA_HUMAN	83,917
988	Trans-2-enoyl-CoA reductase, mitochondrial precursor	MECR_HUMAN	40,410
989	60S acidic ribosomal protein P2	RLA2_HUMAN	11,647
990	Probable oxidoreductase KIAA1576	K1576_HUMAN	45,882
991	Haloacid dehalogenase-like hydrolase domain-containing protein 3	HDHD3_HUMAN	27,982
992	Interferon-induced guanylate-binding protein 2	GBP2_HUMAN	67,167
993	Protein canopy homolog 2 precursor	CNPY2_HUMAN	20,635
994	Actin-related protein 2/3 complex subunit 3	ARPC3_HUMAN	20,530
995	Importin subunit beta-3	IMB3_HUMAN	123,614
996	Transmembrane emp24 domain-containing protein 9 precursor	TMED9_HUMAN	25,087
997	3-hydroxyisobutyryl-CoA hydrolase, mitochondrial precursor	HIBCH_HUMAN	43,466
998	Glutathione S-transferase kappa 1	GSTK1_HUMAN	25,480
999	Transmembrane protein 126A	T126A_HUMAN	21,511
1,000	Fatty acid-binding protein, heart	FABPH_HUMAN	14,841
1,001	NADH dehydrogenase [ubiquinone] iron-sulfur protein 4, mitochondrial precursor	NDUS4_HUMAN	20,090
1,002	HLA class I histocompatibility antigen, alpha chain H precursor	HLAH_HUMAN	40,831
1,003	Bifunctional aminoacyl-tRNA synthetase	SYEP_HUMAN	170,632

identified proteins of <i>Glioblastoma multiforme</i> # 1,004-1,043			
nr.	protein description	accession numbers	mass [Da]
1,004	Epidermal growth factor receptor precursor	EGFR_HUMAN	134,261
1,005	FK506-binding protein 4	FKBP4_HUMAN	51,788
1,006	Phosphatidylinositol 4-kinase alpha	PI4KA_HUMAN	231,282
1,007	Synaptosomal-associated protein 29	SNP29_HUMAN	28,953
1,008	26S proteasome non-ATPase regulatory subunit 7	PSD7_HUMAN	37,008
1,009	Suppressor of G2 allele of SKP1 homolog	SUGT1_HUMAN	41,007
1,010	Glycogen phosphorylase, muscle form	PYGM_HUMAN	97,077
1,011	ATP-dependent RNA helicase DDX1	DDX1_HUMAN	82,415
1,012	Carnitine O-acetyltransferase	CACP_HUMAN	70,909
1,013	Isoleucyl-tRNA synthetase, mitochondrial precursor	SYIM_HUMAN	113,776
1,014	Staphylococcal nuclease domain-containing protein 1	SND1_HUMAN	101,981
1,015	Guanine nucleotide-binding protein G(z) subunit alpha	GNAZ_HUMAN	40,908
1,016	Serine/threonine-protein kinase PAK 2	PAK2_HUMAN	58,026
1,017	Cytoplasmic FMR1-interacting protein 2	CYFP2_HUMAN	148,385
1,018	ATPase family AAA domain-containing protein 3A	ATD3A_HUMAN	71,352
1,019	Isoleucyl-tRNA synthetase, cytoplasmic	SYIC_HUMAN	144,944
1,020	40S ribosomal protein S9	RS9_HUMAN	22,575
1,021	C-1-tetrahydrofolate synthase, cytoplasmic	C1TC_HUMAN	101,544
1,022	Vacuolar proton pump subunit C 1	VATC1_HUMAN	43,925
1,023	Pre-mRNA-processing-splicing factor 8	PRP8_HUMAN	273,591
1,024	Core histone macro-H2A.1	H2AY_HUMAN	39,601
1,025	Lymphocyte-specific protein 1	LSP1_HUMAN	37,173
1,026	Platelet-activating factor acetylhydrolase IB subunit alpha	LIS1_HUMAN	46,619
1,027	Erythrocyte membrane protein band 4.2	EPB42_HUMAN	76,956
1,028	Retinaldehyde-binding protein 1-like protein 2	RLBL2_HUMAN	37,984
1,029	Protein FAM114A2	F1142_HUMAN	56,255
1,030	Agrin precursor	AGRIN_HUMAN	214,820
1,031	T-complex protein 1 subunit eta	TCPH_HUMAN	59,350
1,032	ATP-dependent RNA helicase DDX3X	DDX3X_HUMAN	73,228
1,033	Methylcrotonoyl-CoA carboxylase beta chain, mitochondrial precursor	MCCC2_HUMAN	61,317
1,034	Glycogen [starch] synthase, muscle	GYS1_HUMAN	83,769
1,035	Bifunctional purine biosynthesis protein PURH [Includes: Phosphoribosylaminoimidazolecarboxamide formyltransferase	PUR9_HUMAN	64,599
1,036	Procollagen-lysine,2-oxoglutarate 5-dioxygenase 1 precursor	PLOD1_HUMAN	83,535
1,037	Exportin-2	XPO2_HUMAN	110,404
1,038	G protein-regulated inducer of neurite outgrowth 1	GRIN1_HUMAN	102,353
1,039	Protein CREG1 precursor	CREG1_HUMAN	24,057
1,040	40S ribosomal protein S13	RS13_HUMAN	17,205
1,041	Synaptogyrin-3	SNG3_HUMAN	24,537
1,042	1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase beta-1	PLCB1_HUMAN	138,553
1,043	Cathepsin L1 precursor	CATL1_HUMAN	37,546

identified proteins of <i>Glioblastoma multiforme</i> # 1,044-1,084			
nr.	protein description	accession numbers	mass [Da]
1,044	Complement component 1 Q subcomponent-binding protein, mitochondrial precursor	C1QBP_HUMAN	31,345
1,045	Glyoxalase domain-containing protein 4	GLOD4_HUMAN	34,776
1,046	Rap1 GTPase-activating protein 1	RPGP1_HUMAN	73,374
1,047	cAMP-dependent protein kinase type II-alpha regulatory subunit	KAP2_HUMAN	45,501
1,048	Bullous pemphigoid antigen 1, isoforms 6/9/10	BPAAE_HUMAN	590,974
1,049	Fibromodulin precursor	FMOD_HUMAN	43,230
1,050	Catenin alpha-2	CTNA2_HUMAN	105,296
1,051	Tyrosyl-tRNA synthetase, cytoplasmic	SYYC_HUMAN	59,127
1,052	Twinfilin-2	TWF2_HUMAN	39,531
1,053	Calcium/calmodulin-dependent protein kinase type II beta chain	KCC2B_HUMAN	72,710
1,054	Putative adenosylhomocysteinase 3	SAHH2_HUMAN, SAHH3_HUMAN	66,703
1,055	STIP1 homology and U box-containing protein 1	STUB1_HUMAN	34,839
1,056	Reticulocalbin-2 precursor	RCN2_HUMAN	36,860
1,057	AP-2 complex subunit sigma-1	AP2S1_HUMAN	17,000
1,058	Nuclear protein localization protein 4 homolog	NPL4_HUMAN	68,103
1,059	Coatomer subunit epsilon	COPE_HUMAN	34,465
1,060	Sodium/hydrogen exchanger 1	SL9A1_HUMAN	90,748
1,061	Acyl carrier protein, mitochondrial precursor	ACPM_HUMAN	17,400
1,062	60S ribosomal protein L17	RL17_HUMAN	21,379
1,063	Protein FAM3C precursor	FAM3C_HUMAN	24,663
1,064	Casein kinase II subunit beta	CSK2B_HUMAN	24,925
1,065	Kinesin light chain 1	KLC1_HUMAN	64,769
1,066	Paraspeckle component 1	PSPC1_HUMAN	58,727
1,067	WW domain-binding protein 2	WBP2_HUMAN	28,069
1,068	Serine racemase	SRR_HUMAN	36,548
1,069	Thioredoxin domain-containing protein 12 precursor	TXD12_HUMAN	19,188
1,070	Calpain-1 catalytic subunit	CAN1_HUMAN	81,875
1,071	Synembryn-A	RIC8A_HUMAN	59,595
1,072	NADH dehydrogenase [ubiquinone] 1 subunit C1, mitochondrial precursor	NDUC1_HUMAN	8,717
1,073	Actin-binding LIM protein 1	ABLM1_HUMAN	87,628
1,074	Probable ubiquitin carboxyl-terminal hydrolase FAF-X	USP9X_HUMAN	289,527
1,075	Vitamin K-dependent protein S precursor	PROS_HUMAN	75,105
1,076	Nuclear pore complex protein Nup205	NU205_HUMAN	227,909
1,077	Neuroblastoma-amplified gene protein	NAG_HUMAN	268,573
1,078	Carboxypeptidase E precursor	CBPE_HUMAN	53,134
1,079	Import inner membrane translocase subunit TIM50, mitochondrial precursor	TIM50_HUMAN	39,630
1,080	Abhydrolase domain-containing protein 14B	ABHEB_HUMAN	22,328
1,081	T-complex protein 1 subunit gamma	TCPG_HUMAN	60,517
1,082	Choline dehydrogenase, mitochondrial precursor	CHDH_HUMAN	65,383
1,083	Eukaryotic translation initiation factor 3 subunit A	EIF3A_HUMAN	166,557
1,084	Signal transducer and activator of transcription 1-alpha/beta	STAT1_HUMAN	87,319

identified proteins of <i>Glioblastoma multiforme</i> # 1,085-1,125			
nr.	protein description	accession numbers	mass [Da]
1,085	Coatomer subunit beta	COPB_HUMAN	107,128
1,086	Heterogeneous nuclear ribonucleoprotein Q	HNRPQ_HUMAN	69,586
1,087	Coatomer subunit alpha	COPA_HUMAN	138,317
1,088	Heme oxygenase 2	HMOX2_HUMAN	36,016
1,089	Sorting nexin-18	SNX18_HUMAN	68,863
1,090	Thioredoxin domain-containing protein 1 precursor	TXND1_HUMAN	31,774
1,091	Protein phosphatase inhibitor 2	IPP2_HUMAN	22,998
1,092	Pentraxin-related protein PTX3 precursor	PTX3_HUMAN	42,002
1,093	Fructosamine-3-kinase	FN3K_HUMAN	35,154
1,094	Nucleosome assembly protein 1-like 4	NP1L4_HUMAN	42,806
1,095	Ras-related protein Rab-2A	RAB2A_HUMAN	23,528
1,096	Factor VIII intron 22 protein	F8I2_HUMAN	39,086
1,097	Golgi-associated plant pathogenesis-related protein 1	GAPR1_HUMAN	17,200
1,098	Moesin	MOES_HUMAN	67,804
1,099	N-terminal EF-hand calcium-binding protein 1	NECA1_HUMAN	40,554
1,100	40S ribosomal protein S16	RS16_HUMAN	16,428
1,101	Inter-alpha-trypsin inhibitor heavy chain H3 precursor	ITIH3_HUMAN	99,833
1,102	Electrogenic sodium bicarbonate cotransporter 1	S4A4_HUMAN	121,449
1,103	Collagen alpha-1(III) chain precursor	CO3A1_HUMAN	138,547
1,104	Tubulin gamma-2 chain	TBG1_HUMAN, TBG2_HUMAN	51,075
1,105	Thiosulfate sulfurtransferase	THTR_HUMAN	33,411
1,106	Glycogen debranching enzyme	GDE_HUMAN	174,750
1,107	Transmembrane protein 16F	TM16F_HUMAN	106,151
1,108	Pre-B-cell leukemia transcription factor-interacting protein 1	PBIP1_HUMAN	80,625
1,109	Eukaryotic translation initiation factor 3 subunit F	EIF3F_HUMAN	37,546
1,110	Coronin-1A	COR1A_HUMAN	51,008
1,111	Eukaryotic initiation factor 4A-III	IF4A3_HUMAN	46,854
1,112	Dynamin-3	DYN3_HUMAN	96,667
1,113	Monocarboxylate transporter 1	MOT1_HUMAN	53,942
1,114	E3 ubiquitin-protein ligase UBR4	UBR4_HUMAN	573,825
1,115	Secretogranin-1 precursor	SCG1_HUMAN	78,228
1,116	Adaptin ear-binding coat-associated protein 1	NECP1_HUMAN	29,720
1,117	Ras-related protein Rab-3C	RAB3C_HUMAN	25,935
1,118	Apoptosis-inducing factor 1, mitochondrial precursor	AIFM1_HUMAN	66,884
1,119	Protein TBRG4	TBRG4_HUMAN	70,722
1,120	Calcyclin-binding protein	CYBP_HUMAN	26,192
1,121	Ectonucleoside triphosphate diphosphohydrolase 2	ENTP2_HUMAN	53,648
1,122	Histone H3.1t	H31T_HUMAN, H31_HUMAN, H32_HUMAN, H33_HUMAN	15,491
1,123	Signal transducer and activator of transcription 3	STAT3_HUMAN	88,052
1,124	Potassium voltage-gated channel subfamily A member 3	KCNA2_HUMAN, KCNA3_HUMAN	58,287
1,125	C4b-binding protein alpha chain precursor	C4BP_HUMAN	67,015

identified proteins of <i>Glioblastoma multiforme</i> # 1,126-1,167			
nr.	protein description	accession numbers	mass [Da]
1,126	Atlastin-1	ATLA1_HUMAN	63,528
1,127	Succinate-semialdehyde dehydrogenase, mitochondrial precursor	SSDH_HUMAN	57,197
1,128	Uncharacterized protein C1orf198	CA198_HUMAN	36,329
1,129	Brain protein 44	BR44_HUMAN	14,262
1,130	40S ribosomal protein S26	RS26_HUMAN	12,998
1,131	Guanine nucleotide-binding protein subunit alpha-11	GNA11_HUMAN	42,107
1,132	Retinol dehydrogenase 11	RDH11_HUMAN	35,369
1,133	Proteasome subunit alpha type-5	PSA5_HUMAN	26,393
1,134	Low molecular weight phosphotyrosine protein phosphatase	PPAC_HUMAN	18,025
1,135	Nicotinate-nucleotide pyrophosphorylase [carboxylating]	NADC_HUMAN	30,797
1,136	Peroxisomal biogenesis factor 19	PEX19_HUMAN	32,789
1,137	Transmembrane protein 109 precursor	TM109_HUMAN	26,193
1,138	40S ribosomal protein S23	RS23_HUMAN	15,790
1,139	Prenylcysteine oxidase 1 precursor	PCYOX_HUMAN	56,624
1,140	BTB/POZ domain-containing protein KCTD16	KCD16_HUMAN	49,121
1,141	Isocitrate dehydrogenase [NADP] cytoplasmic	IDHC_HUMAN	46,643
1,142	FK506-binding protein 10 precursor	FKB10_HUMAN	64,228
1,143	Glucosylceramidase precursor	GLCM_HUMAN	59,700
1,144	Vacuolar protein sorting-associating protein 4B	VPS4B_HUMAN	49,286
1,145	F-box only protein 2	FBX2_HUMAN	33,309
1,146	Calcium/calmodulin-dependent protein kinase type 1	KCC1A_HUMAN	41,320
1,147	Tapasin precursor	TPSN_HUMAN	47,609
1,148	Protein FAM98A	FA98A_HUMAN	55,383
1,149	Macrophage-capping protein	CAPG_HUMAN	38,500
1,150	Serine/threonine-protein kinase PAK 3	PAK3_HUMAN	62,293
1,151	Ubiquinone biosynthesis protein COQ9, mitochondrial precursor	COQ9_HUMAN	35,491
1,152	Alpha-1-acid glycoprotein 2 precursor	A1AG2_HUMAN	23,585
1,153	Lin-7 homolog C	LIN7C_HUMAN	21,816
1,154	Eukaryotic translation initiation factor 4B	IF4B_HUMAN	69,209
1,155	Coiled-coil domain-containing protein 56	CCD56_HUMAN	11,714
1,156	Serine/threonine-protein kinase TAO1	TAOK1_HUMAN	116,054
1,157	Dihydropteridine reductase	DHPR_HUMAN	25,772
1,158	Ubiquitin-like modifier-activating enzyme 6	UBA6_HUMAN	117,955
1,159	AP2-associated protein kinase 1	AAK1_HUMAN	93,560
1,160	Plexin-B2 precursor	PLXB2_HUMAN	205,109
1,161	NAD-dependent malic enzyme, mitochondrial precursor	MAOM_HUMAN	65,428
1,162	NADH dehydrogenase [ubiquinone] 1 subunit C2	NDUC2_HUMAN	14,171
1,163	Uncharacterized protein C2orf55	CB055_HUMAN	102,145
1,164	Vesicle-associated membrane protein-associated protein B/C	VAPB_HUMAN	27,211
1,165	Valyl-tRNA synthetase	SYV_HUMAN	140,460
1,166	ATP-dependent RNA helicase DDX39	DDX39_HUMAN, UAP56_HUMAN	49,112
1,167	60S ribosomal protein L3	RL3_HUMAN	46,092

identified proteins of <i>Glioblastoma multiforme</i> # 1,168-1,209			
nr.	protein description	accession numbers	mass [Da]
1,168	Lon protease homolog, mitochondrial precursor	LONM_HUMAN	106,473
1,169	AP-3 complex subunit delta-1	AP3D1_HUMAN	130,144
1,170	Ephrin-B3 precursor	EFNB3_HUMAN	35,818
1,171	Ras-related protein Rab-5C	RAB5C_HUMAN	23,465
1,172	Hematological and neurological expressed 1-like protein	HN1L_HUMAN	20,046
1,173	ADP-ribosylation factor-like protein 8B	ARL8B_HUMAN	21,522
1,174	Fatty aldehyde dehydrogenase	AL3A2_HUMAN	54,832
1,175	Cytochrome c oxidase subunit VIb isoform 1	CX6B1_HUMAN	10,174
1,176	EH domain-containing protein 1	EHD1_HUMAN	60,611
1,177	ATP synthase subunit gamma, mitochondrial precursor	ATPG_HUMAN	32,980
1,178	60S ribosomal protein L4	RL4_HUMAN	47,681
1,179	Phosphatidate cytidyltransferase 2	CDS2_HUMAN	51,402
1,180	26S proteasome non-ATPase regulatory subunit 1	PSMD1_HUMAN	105,821
1,181	Annexin A3	ANXA3_HUMAN	36,359
1,182	KH domain-containing, RNA-binding, signal transduction-associated protein 1	KHDR1_HUMAN	48,210
1,183	ADP-ribose pyrophosphatase, mitochondrial precursor	NUDT9_HUMAN	39,108
1,184	Acylglycerol kinase, mitochondrial precursor	AGK_HUMAN	47,120
1,185	UBX domain-containing protein 8	UBXD8_HUMAN	52,606
1,186	Extended synaptotagmin-2	ESYT2_HUMAN	102,343
1,187	Tubulin beta-1 chain	TBB1_HUMAN	50,309
1,188	Mitochondrial carrier homolog 1	MTCH1_HUMAN	41,527
1,189	S-adenosylmethionine synthetase isoform type-2	METK2_HUMAN	43,643
1,190	Phenylalanyl-tRNA synthetase beta chain	SYFB_HUMAN	66,115
1,191	26S proteasome non-ATPase regulatory subunit 6	PSMD6_HUMAN	45,515
1,192	40S ribosomal protein S3a	RS3A_HUMAN	29,927
1,193	Bcl-2-like 13 protein	B2L13_HUMAN	52,705
1,194	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit STT3A	STT3A_HUMAN	80,457
1,195	Long-chain-fatty-acid--CoA ligase 3	ACSL3_HUMAN	80,405
1,196	von Willebrand factor A domain-containing protein 1 precursor	VWA1_HUMAN	46,786
1,197	Tetratricopeptide repeat protein 1	TTC1_HUMAN	33,509
1,198	Brain-specific polypeptide PEP-19	PEP19_HUMAN	6,773
1,199	Transgelin	TAGL_HUMAN	22,593
1,200	Laminin subunit alpha-5 precursor	LAMA5_HUMAN	399,725
1,201	TSC22 domain family protein 4	T22D4_HUMAN	41,009
1,202	Parvalbumin alpha	PRVA_HUMAN	12,042
1,203	Microsomal glutathione S-transferase 3	MGST3_HUMAN	16,499
1,204	Histone H1x	H1X_HUMAN	22,470
1,205	BRO1 domain-containing protein BROX	BROX_HUMAN	46,460
1,206	Guanylate kinase	KGUA_HUMAN	21,708
1,207	Ubiquitin fusion degradation protein 1 homolog	UFD1_HUMAN	34,483
1,208	Intracellular hyaluronan-binding protein 4	HABP4_HUMAN	45,767
1,209	Coiled-coil-helix-coiled-coil-helix domain-containing protein 6	CHCH6_HUMAN	26,439

identified proteins of <i>Glioblastoma multiforme</i> # 1,210-1,253			
nr.	protein description	accession numbers	mass [Da]
1,210	Protein unc-84 homolog B	UN84B_HUMAN	80,294
1,211	Retinal dehydrogenase 1	AL1A1_HUMAN	54,845
1,212	NudC domain-containing protein 3	NUDC3_HUMAN	40,804
1,213	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A member 5	SMCA1_HUMAN, SMCA5_HUMAN	121,893
1,214	Protein phosphatase 1 regulatory subunit 7	PP1R7_HUMAN	41,548
1,215	60S ribosomal protein L18a	RL18A_HUMAN	20,745
1,216	Proteasome subunit alpha type-7	PSA7_HUMAN	27,869
1,217	Vacuolar protein sorting-associated protein 18 homolog	VPS18_HUMAN	110,170
1,218	Pirin	PIR_HUMAN	32,096
1,219	Pigment epithelium-derived factor precursor	PEDF_HUMAN	46,326
1,220	CAP-Gly domain-containing linker protein 2	CLIP2_HUMAN	115,821
1,221	Growth arrest-specific protein 7	GAS7_HUMAN	47,249
1,222	Intercellular adhesion molecule 5 precursor	ICAM5_HUMAN	97,311
1,223	COP9 signalosome complex subunit 8	CSN8_HUMAN	23,208
1,224	Sorting nexin-5	SNX5_HUMAN	46,800
1,225	Protein S100-A11	S10AB_HUMAN	11,723
1,226	Leucine-rich repeat-containing protein 57	LRC57_HUMAN	26,738
1,227	Isocitrate dehydrogenase [NAD] subunit gamma, mitochondrial precursor	IDH3G_HUMAN	42,776
1,228	60S ribosomal protein L26	RL26_HUMAN	17,240
1,229	Homer protein homolog 1	HOME1_HUMAN	40,260
1,230	Acyl-coenzyme A thioesterase 9	ACOT9_HUMAN	46,337
1,231	ARF GTPase-activating protein GIT1	GIT1_HUMAN	84,324
1,232	Galactokinase	GALK1_HUMAN	42,254
1,233	Clathrin interactor 1	EPN4_HUMAN	68,243
1,234	40S ribosomal protein S7	RS7_HUMAN	22,110
1,235	Propionyl-CoA carboxylase alpha chain, mitochondrial precursor	PCCA_HUMAN	77,336
1,236	Unc-112-related protein 2	URP2_HUMAN	75,937
1,237	Calcium-binding protein 39	CAB39_HUMAN	39,853
1,238	Spectrin beta chain, brain 3	SPTN4_HUMAN	288,966
1,239	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 48 kDa subunit precursor	OST48_HUMAN	48,793
1,240	Serine hydroxymethyltransferase, mitochondrial precursor	GLYM_HUMAN	55,977
1,241	Tubulin-specific chaperone D	TBCD_HUMAN	132,513
1,242	Apolipoprotein A-I-binding protein precursor	AIBP_HUMAN	31,657
1,243	Flotillin-2	FLOT2_HUMAN	41,667
1,244	N-terminal EF-hand calcium-binding protein 2	NECA2_HUMAN	43,176
1,245	Ras-related protein Rab-32	RAB32_HUMAN	24,979
1,246	Prolyl 4-hydroxylase subunit alpha-1 precursor	P4HA1_HUMAN	61,034
1,247	Orphan sodium- and chloride-dependent neurotransmitter transporter NTT4	S6A17_HUMAN	80,986
1,248	26S proteasome non-ATPase regulatory subunit 11	PSD11_HUMAN	47,448
1,249	Myeloid leukemia factor 2	MLF2_HUMAN	28,130
1,250	Pre-mRNA-processing factor 19	PRP19_HUMAN	55,163
1,251	Glycyl-tRNA synthetase	SYG_HUMAN	83,124
1,252	THO complex subunit 4	THOC4_HUMAN	26,871
1,253	Exocyst complex component 7	EXOC7_HUMAN	83,367

identified proteins of <i>Glioblastoma multiforme</i> # 1,254-1,299			
nr.	protein description	accession numbers	mass [Da]
1,254	Abhydrolase domain-containing protein 12	ABD12_HUMAN	45,080
1,255	Phosphofurin acidic cluster sorting protein 1	PACS1_HUMAN	104,881
1,256	Tubulin--tyrosine ligase-like protein 12	TTL12_HUMAN	74,386
1,257	Copine-1	CPNE1_HUMAN	59,041
1,258	Endoplasmic reticulum aminopeptidase 1	ERAP1_HUMAN	107,220
1,259	40S ribosomal protein S15a	RS15A_HUMAN	14,822
1,260	NADPH:adenodoxin oxidoreductase, mitochondrial precursor	ADRO_HUMAN	53,819
1,261	Regulator of nonsense transcripts 1	RENT1_HUMAN	124,329
1,262	Probable phospholipid-transporting ATPase IA	AT8A1_HUMAN	131,356
1,263	Matrix metalloproteinase-9 precursor	MMP9_HUMAN	78,411
1,264	Phosphatidylinositol-5-phosphate 4-kinase type-2 beta	PI42B_HUMAN	47,361
1,265	ADP-dependent glucokinase	ADPGK_HUMAN	54,071
1,266	Triple functional domain protein	TRIO_HUMAN	341,600
1,267	Collagen alpha-1(XIV) chain precursor	COEA1_HUMAN	193,498
1,268	Opioid growth factor receptor-like protein 1	OGRL1_HUMAN	51,235
1,269	40S ribosomal protein S14	RS14_HUMAN	16,255
1,270	Protein-glutamine gamma-glutamyltransferase 2	TGM2_HUMAN	77,311
1,271	Prolactin regulatory element-binding protein	PREB_HUMAN	45,450
1,272	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 6	NDUB6_HUMAN	15,472
1,273	RalBP1-associated Eps domain-containing protein 1	REPS1_HUMAN	80,751
1,274	COMM domain-containing protein 9	COMD9_HUMAN	21,802
1,275	Cleft lip and palate transmembrane protein 1	CLPT1_HUMAN	76,080
1,276	40S ribosomal protein S15	RS15_HUMAN	17,023
1,277	Heterogeneous nuclear ribonucleoprotein D-like	HNRDL_HUMAN	46,421
1,278	Copine-5	CPNE5_HUMAN	65,717
1,279	Sulfide:quinone oxidoreductase, mitochondrial precursor	SQRD_HUMAN	49,944
1,280	Programmed cell death protein 5	PDCD5_HUMAN	14,267
1,281	Mitochondrial 39S ribosomal protein L4	RM04_HUMAN	34,902
1,282	60S ribosomal protein L5	RL5_HUMAN	34,346
1,283	Procollagen-lysine,2-oxoglutarate 5-dioxygenase 3 precursor	PLOD3_HUMAN	84,769
1,284	RuvB-like 1	RUVB1_HUMAN	50,211
1,285	Thioredoxin-like protein 1	TXNL1_HUMAN	32,233
1,286	Copine-6	CPNE6_HUMAN	61,973
1,287	Ras-related protein Rab-21	RAB21_HUMAN	24,330
1,288	Methyltransferase-like protein 7B precursor	MET7B_HUMAN	27,757
1,289	SEC14-like protein 2	S14L2_HUMAN	46,128
1,290	DnaJ homolog subfamily B member 6	DNJB6_HUMAN	36,070
1,291	Microtubule-associated protein 1S	MAP1S_HUMAN	112,195
1,292	Quinone oxidoreductase	QOR_HUMAN	35,189
1,293	Surfeit locus protein 4	SURF4_HUMAN	30,377
1,294	Methyl-CpG-binding protein 2	MECP2_HUMAN	52,423
1,295	Cathepsin H precursor	CATH_HUMAN	37,360
1,296	Ribose-phosphate pyrophosphokinase 2	PRPS2_HUMAN	34,752
1,297	60 kDa SS-A/Ro ribonucleoprotein	RO60_HUMAN	60,654
1,298	MAP7 domain-containing protein 2	MA7D2_HUMAN	81,945
1,299	Ribonuclease T2 precursor	RNT2_HUMAN	29,463

identified proteins of <i>Glioblastoma multiforme</i> # 1,300-1,341			
nr.	protein description	accession numbers	mass [Da]
1,300	40S ribosomal protein S27	RS27_HUMAN	9,443
1,301	Succinyl-CoA ligase [GDP-forming] subunit alpha, mitochondrial precursor	SUCA_HUMAN	35,030
1,302	Peripheral plasma membrane protein CASK	CSKP_HUMAN	105,106
1,303	Metabotropic glutamate receptor 5 precursor	GRM5_HUMAN	132,454
1,304	Transcriptional activator protein Pur-alpha	PURA_HUMAN	34,893
1,305	Calcium-dependent secretion activator 1	CAPS1_HUMAN	152,771
1,306	Ephrin-B2 precursor	EFNB2_HUMAN	36,906
1,307	PDZ domain-containing protein 11	PDZ11_HUMAN	16,114
1,308	Rap1 GTPase-GDP dissociation stimulator 1	GDS1_HUMAN	66,386
1,309	Autophagy-related protein 3	ATG3_HUMAN	35,846
1,310	CaM kinase-like vesicle-associated protein	CAMKV_HUMAN	54,336
1,311	Exonuclease 3'-5' domain-like-containing protein 2	EXDL2_HUMAN	56,329
1,312	Tetratricopeptide repeat protein 9C	TTC9C_HUMAN	19,995
1,313	NADH-cytochrome b5 reductase 2	NB5R2_HUMAN	31,442
1,314	Proliferation-associated protein 2G4	PA2G4_HUMAN	43,769
1,315	Phosphoglucomutase-2-like 1	PGM2L_HUMAN	70,439
1,316	Ribosomal protein S6 kinase alpha-3	KS6A3_HUMAN	83,721
1,317	Lupus La protein	LA_HUMAN	46,821
1,318	Oxysterol-binding protein-related protein 8	OSBL8_HUMAN	101,181
1,319	Heparin-binding growth factor 1 precursor	FGF1_HUMAN	17,442
1,320	Zinc-binding alcohol dehydrogenase domain-containing protein 1	ZADH1_HUMAN	38,481
1,321	Ankyrin repeat and FYVE domain-containing protein 1	ANFY1_HUMAN	128,384
1,322	Casein kinase II subunit alpha'	CSK22_HUMAN	41,197
1,323	rRNA 2'-O-methyltransferase fibrillarin	FBRL_HUMAN	33,766
1,324	Alpha-aminoacidic semialdehyde synthase, mitochondrial precursor	AASS_HUMAN	102,117
1,325	UPF0368 protein Cxorf26	CX026_HUMAN	26,040
1,326	Catechol O-methyltransferase	COMT_HUMAN	30,020
1,327	Coatmer subunit beta'	COPB2_HUMAN	102,471
1,328	Active breakpoint cluster region-related protein	ABR_HUMAN	97,682
1,329	Transferrin receptor protein 1	TFR1_HUMAN	84,856
1,330	Succinyl-CoA ligase [GDP-forming] beta-chain, mitochondrial precursor	SUCB2_HUMAN	46,494
1,331	SH3 domain-binding glutamic acid-rich-like protein 2	SH3L2_HUMAN	12,308
1,332	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 11, mitochondrial precursor	NDUBB_HUMAN	17,299
1,333	Translin	TSN_HUMAN	26,165
1,334	Coiled-coil domain-containing protein 47 precursor	CCD47_HUMAN	55,857
1,335	Glycine dehydrogenase [decarboxylating], mitochondrial precursor	GCSP_HUMAN	112,697
1,336	Zinc finger CCCH type antiviral protein 1	ZCC2_HUMAN	101,428
1,337	ADP-ribosyl cyclase 1	CD38_HUMAN	34,310
1,338	Coronin-1C	COR1C_HUMAN	53,232
1,339	Torsin-1A-interacting protein 2	TOIP2_HUMAN	51,245
1,340	Signal recognition particle 68 kDa protein	SRP68_HUMAN	70,714
1,341	Amyloid-like protein 1 precursor	APLP1_HUMAN	72,158

identified proteins of <i>Glioblastoma multiforme</i> # 1,342-1,381			
nr.	protein description	accession numbers	mass [Da]
1,342	Twinfilin-1	TWF1_HUMAN	42,192
1,343	Kinesin-like protein KIF3A	KIF3A_HUMAN	80,370
1,344	Enhancer of rudimentary homolog	ERH_HUMAN	12,241
1,345	Transportin-1	TNPO1_HUMAN	101,296
1,346	Mitochondrial import inner membrane translocase subunit Tim13	TIM13_HUMAN	10,482
1,347	BTB/POZ domain-containing protein KCTD12	KCD12_HUMAN	35,684
1,348	Mesoderm development candidate 2	MESD2_HUMAN	26,060
1,349	Importin-7	IPO7_HUMAN	119,502
1,350	Oligopeptide transporter, kidney isoform	S15A2_HUMAN	81,925
1,351	Retinol-binding protein 1, cellular	RET1_HUMAN	15,832
1,352	Serine/threonine-protein kinase WNK1	WNK1_HUMAN	250,732
1,353	SAPK substrate protein 1	SAKS1_HUMAN	33,307
1,354	NADP-dependent leukotriene B4 12-hydroxydehydrogenase	LTB4D_HUMAN	35,853
1,355	GRB2-associated-binding protein 1	GAB1_HUMAN	76,599
1,356	Choline transporter-like protein 2	CTL2_HUMAN	80,138
1,357	Leucine-rich repeat-containing protein 8A	LRC8A_HUMAN	94,186
1,358	Tripartite motif-containing protein 47	TRI47_HUMAN	69,513
1,359	Eukaryotic translation initiation factor 3 subunit E-interacting protein	IF3EI_HUMAN	66,711
1,360	ETHE1 protein, mitochondrial precursor	ETHE1_HUMAN	27,855
1,361	Rho-related GTP-binding protein RhoG precursor	RHOG_HUMAN	21,290
1,362	1-acyl-sn-glycerol-3-phosphate acyltransferase epsilon	PLCE_HUMAN	42,056
1,363	OTU domain-containing protein 7B	OTU7B_HUMAN	92,509
1,364	Aspartoacylase	ACY2_HUMAN	35,718
1,365	1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase delta-1	PLCD1_HUMAN	85,747
1,366	Gamma-butyrobetaine dioxygenase	BODG_HUMAN	44,698
1,367	Derlin-1	DERL1_HUMAN	28,785
1,368	Thyroid hormone receptor-associated protein 3	TR150_HUMAN	108,651
1,369	Nuclear pore complex protein Nup107	NU107_HUMAN	106,359
1,370	Multiple coagulation factor deficiency protein 2 precursor	MCFD2_HUMAN	16,373
1,371	39S ribosomal protein L12, mitochondrial precursor	RM12_HUMAN	21,330
1,372	Phenylalanyl-tRNA synthetase alpha chain	SYFA_HUMAN	57,547
1,373	Integrin beta-5 precursor	ITB5_HUMAN	88,037
1,374	Bifunctional coenzyme A synthase	COASY_HUMAN	62,312
1,375	Peptidyl-prolyl cis-trans isomerase G	PPIG_HUMAN	88,602
1,376	Multifunctional protein ADE2 [Includes: Phosphoribosylaminoimidazole-succinocarboxamide synthase	PUR6_HUMAN	47,062
1,377	HLA class II histocompatibility antigen, DQ(5) alpha chain precursor	HA25_HUMAN	27,982
1,378	AsparaginyI-tRNA synthetase, cytoplasmic	SYNC_HUMAN	62,926
1,379	40S ribosomal protein S18	RS18_HUMAN	17,701
1,380	Putative pre-mRNA-splicing factor ATP-dependent RNA helicase DHX15	DHX15_HUMAN	90,917
1,381	Vacuolar protein sorting-associated protein 11 homolog	VPS11_HUMAN	107,821

identified proteins of <i>Glioblastoma multiforme</i> # 1,382-1,401			
nr.	protein description	accession numbers	mass [Da]
1,382	Protein phosphatase 1A	PPM1A_HUMAN	42,429
1,383	Tescalcin	TESC_HUMAN	24,733
1,384	Adenosylhomocysteinase	SAHH_HUMAN	47,699
1,385	Vacuolar protein sorting-associated protein 41 homolog	VPS41_HUMAN	98,521
1,386	Calcium-independent phospholipase A2-gamma	PLPL8_HUMAN	88,462
1,387	D-beta-hydroxybutyrate dehydrogenase, mitochondrial precursor	BDH_HUMAN	38,140
1,388	WD repeat-containing protein 7	WDR7_HUMAN	163,795
1,389	ATP-dependent RNA helicase DDX19A	DD19A_HUMAN, DD19B_HUMAN	53,958
1,390	Dihydropyrimidinase-related protein 5	DPYL5_HUMAN	61,403
1,391	Glutathione reductase, mitochondrial precursor	GSHR_HUMAN	56,239
1,392	Heat shock 70 kDa protein 4	HSP74_HUMAN	94,283
1,393	Alpha-actinin-2	ACTN2_HUMAN	103,840
1,394	Gamma-aminobutyric acid receptor subunit gamma-2 precursor	GBRG2_HUMAN	54,146
1,395	Myosin-11	MYH11_HUMAN	227,327
1,396	[Pyruvate dehydrogenase [lipoamide]] kinase isozyme 3, mitochondrial precursor	PDK3_HUMAN	46,923
1,397	Double-strand break repair protein MRE11A	MRE11_HUMAN	80,577
1,398	Arylsulfatase A precursor	ARSA_HUMAN	53,571
1,399	Myosin-Ic	MYO1C_HUMAN	118,024
1,400	Beta-glucuronidase precursor	BGLR_HUMAN	74,715
1,401	Spartin	SPG20_HUMAN	72,815

Retention behavior of all 142 membrane proteins identified with the semi top-down approach. Proteins in bold and italic are not confirmed with Scaffold proteome software. KCNA2 and KCNA 3 were grouped by Scaffold.

protein	fraction																												sequence coverage [%]	different peptides
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28		
1 ATPB																													53.3	20
2 TBB3	-																												34.9	17
3 CALX																													24.3	11
4 THY1																													16.1	3
5 VAMP2	-																												28.4	2
6 MAG	-																												19.5	9
7 AT1B1																													27.4	9
8 NFASC																													9.1	15
9 SHPS1																													10.7	4
10 CD44	-																												6.3	4
11 VAPA																													33.7	7
12 SFXN1																													18	5
13 CADH2																													18.7	12
14 PTPRZ																													3.8	7
15 VAMP3	-																												33	2
16 ENPP6																													22	7
17 L1CAM																													12.5	13
18 4F2																													29.9	14
19 AT1B3																													31.2	8
20 SCRB2																													1.6	6
21 HECAM																													15.4	4
22 TOLIP	-																												23.7	4
23 PALM	-																												25.1	7
24 IMMT																													11.7	5

protein	fraction																												sequence coverage [%]	different peptides
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28		
25 CNTP1																													6	6
26 STX1B	—																												20.1	4
27 CA2D1																													20.1	4
28 ITAV																													14	13
29 VAT1																													30	6
30 SYT1																													20.1	6
31 MPRD																													11.6	2
32 AT1B2																													25.5	7
33 FXD6																													38.9	3
34 SFXN3																													27.4	7
35 PGRC1																													37.4	6
36 CJ035																													32.2	3
37 VATA																													9.2	5
38 CADM3																													23.1	5
39 LAMP2																													15.7	4
40 TOM22																													7.1	3
41 AQP4																													62	5
42 RTN1																													8	2
43 CC50A																													13	7
44 NRCAM																													14.7	7
45 UCRI																													6.1	5
46 MYH9																													29.9	5
47 PGRC2																													2	2
48 RAN																													22.9	3

protein	fraction																												sequence coverage [%]	different peptides
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28		
49 SCN3B																													24.2	4
50 CADM2																													5.3	2
51 ITB1																													9.9	6
52 TMM65																													20	4
53 IMA4																													7.1	3
54 BASI																													15.3	4
55 STX1A																													36.1	7
56 CYB5B																													27.4	3
57 RENR																													6.9	2
58 SYPH																													8.3	2
59 STX7																													16.1	3
60 RTN4																													9.7	6
61 PPAL																													11.3	5
62 LAP2A																													9.5	4
63 VAS1																													10	5
64 TFG																													17.3	3
65 NPTN																													7.8	2
66 LMAN2																													15.7	5
67 VDAC1																													13.8	2
68 OMGP																													7.7	2
69 SL9A6																													4.6	2
70 GGT5																													7.3	4
71 PTPRN																													5.5	3
72 ZYX																													9.6	4

protein	fraction																												sequence coverage [%]	different peptides
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28		
73							—	—	—																				4.0	3
74	—						—																		—				14.6	3
75									—																				13.2	2
76																—													13.2	3
77										—	—												—	—					2.0	2
78																	—						—						14.1	6
79																							—	—					3.8	4
80																—													7.2	4
81																				—									21.2	5
82																				—									14.2	3
83																				—	—								8.0	2
84																				—	—								5.2	3
85																							—						23.3	5
86																										—	—		11.6	2
87																—		—											4.4	2
88																									—				7.6	5
89																					—								15.3	4
90																						—				—			4.6	2
91	—								—	—																			7.2	3
92																—													4.8	3
93																								—	—				5.5	2
94																									—	—			2.1	3
95																—													17.6	3
96																	—				—		—	—					8.7	3

protein	fraction																												sequence coverage [%]	different peptides	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28			
97																							—		—				12.0	4	
98																							—							5.4	3
99														—															4.0	3	
100						—																							3.6	4	
101										—																			4.9	2	
102																				—						—			11.0	3	
103									—	—																			4.2	3	
104								—	—																				10.3	2	
105																										—			17.8	4	
106						—																				—			4.8	3	
107										—																			5.4	2	
108																								—					11.7	2	
109																								—					4.4	2	
110																								—					4.2	2	
111																													27.5	2	
112							—										—												8.6	2	
113								—																					2.3	2	
114	—							—																					10.9	3	
115									—																				21.3	2	
116													—																41.7	2	
117																								—	—				9.0	2	
118	—																												12.0	2	
119																													17.9	2	
120								—																					3.6	2	

protein	fraction																												sequence coverage [%]	different peptides	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28			
121 T126A																				—									14.9	2	
122 NSDHL																								—						8.6	2
123 K0152										—																—				14.4	3
124 S12A4																				—										2.1	2
125 CY1																	—									—				11.7	3
126 AT2B2																											—			2.7	2
127 PCDH1						—																								3.4	2
128 RANB3										—																				5.5	2
129 CADM1																					—	—	—	—	—	—	—	—		5.0	2
130 RTN3	—																													4.3	2
131 EFNB3											—																			7.1	2
132 S12A2									—																					2.6	2
133 HA25																							—							9.0	2
134 CT103																					—									9.6	2
135 EFNB2																						—	—	—	—	—	—	—		9.3	2
136 ICAM5												—		—																2.7	2
137 CNTP2																					—									1.7	2
138 CAD23															—										—		—			0.7	2
139 NPTXR																					—									6.4	2
140 HNRPM													—																	4.5	2
141 SYNE1	—																													0.3	3
142 LPHN1	—																													1.3	2

Distribution of peptides observed for SEREX proteins. Not for all 13 proteins GO annotations are available (www.expasy.org).

anti- gen	name	Gene ontology term		
		cellular component	molecular function	biological process
ACTN4	actinin, alpha 4	nucleus, perinuclear region of cytoplasm, protein complex, pseudopodium	actin filament binding, integrin binding, nucleoside binding, protein homodimerization activity	positive regulation of cell motility, positive regulation of sodium:hydrogen antiporter activity, regulation of apoptosis
ING4	inhibitor of growth family, member 4	nucleus	protein binding	apoptosis, cell cycle arrest, negative regulation of cell proliferation, negative regulation of growth, negative regulation of transcription, DNA-dependent protein amino acid acetylation
RTN4	reticulon 4	integral to endoplasmic reticulum membrane, nuclear envelope	protein binding	negative regulation of anti-apoptosis, negative regulation of axon extension, regulation of cell migration
CLIP2	cytoplasmic linker 2	microtubule associated complex		
HCLS1	hematopoietic cell-specific Lyn substrate 1	DNA-directed RNA polymerase II, core complex; cytoplasm	protein binding, transcription factor activity	erythrocyte differentiation, positive regulation of cell proliferation, positive regulation of tyrosine phosphorylation of STAT protein, regulation of transcription, DNA-dependent, response to hormone stimulus
U2AFL	U2(RNU2) small nuclear RNA auxillary factor 1-like 1	nucleus	RNA binding	
ZN232	zinc finger protein 232	Nucleus, intracellular	zinc ion binding, DNA binding	regulation of transcription, DNA dependent
HS105	Heat-shock protein 105 kDa	cytoplasm	nucleotide binding	response to unfolded protein
PHF3	PHD finger protein 3	-	zinc ion binding, DNA binding	multicellular organismal development
PHF20	PHD finger protein 20	-	-	-
TPR	translocated promoter region (to activated MET oncogene)	Cytoplasm, nuclear pore	nucleotide binding, ATP binding, serine-tRNA-ligase activity	protein import into nucleus
CYTSA	Cytospin-A (NY-REN-22 antigen)	-	-	-

anti- gen	name	Gene ontology term		
		cellular component	molecular function	biological process
GOGA 1	golgi autoantigen, golgin subfamily a	golgi apparatus	-	-