

Globale qualitative und quantitative Proteomanalyse mittels hochauflösender LC-MS

Zur Erlangung des akademischen Grades eines

Dr. rer. nat.

von der Fakultät Bio- und Chemieingenieurwesen

der Technischen Universität Dortmund

genehmigte Dissertation

vorgelegt von

Julia Maria Burkhart

aus

Schrobenhausen

Tag der mündlichen Prüfung: 15.05.2013

1. Gutachter: Prof. Dr. Albert Sickmann

2. Gutachter: Prof. Dr. Chris Meisinger

Dortmund 2013

1. Prüfer: Prof. Dr. Oliver Kayser

2. Prüfer: Prof. Dr.-Ing. Rolf Wichmann

Inhaltsverzeichnis

1. Einleitung	14
1.1 Globale Proteomstudien.....	14
1.2 Strategien der Proteomanalyse.....	14
1.2.1 Flüssigchromatographie	15
1.2.2 Präfraktionierungsmethoden.....	18
1.2.3 Massenspektrometrische Messverfahren	21
1.3 Quantitative Proteomforschung.....	25
1.3.1 Chemische Quantifizierungsstrategien	26
1.3.2 Metabolische Quantifizierungsstrategien.....	28
1.3.3 Label freie Quantifizierung.....	28
1.4 Qualitätskontrolle	29
1.4.1 Probenaufarbeitung	29
1.4.2 LC-MS.....	30
1.4.3 Datenanalyse	31
1.5 Modellproteome.....	32
1.5.1 Universal Proteomics Standard	32
1.5.2 Complex Proteomics Standard	32
1.5.3 Mitochondrien.....	32
1.5.4 Thrombozyten	34
1.6 HeLa	35
2. Zielsetzung	37
3. Material und Methoden	38
3.1 Geräte, Reagenzien, Software und Datenbanken	38
3.2 Aufreinigung humaner Thrombozyten aus Frischblut.....	41
3.3 Aufreinigung von HeLa Zellen.....	41
3.4 Aufreinigung von Mitochondrien aus <i>Saccharomyces cerevisiae</i>	41
3.4.1 BAX induzierte Freisetzung von IMS Proteinen	42
3.4.2 Proteinprozessierung in Mitochondrien	43
3.5 Probenaufarbeitung	44
3.6 Enzymatischer Proteinverdau und Optimierung.....	45
3.7 Verfahren zur Verdaukontrolle	45
3.7.1 1-DE-PAGE und Proteinvisualisierung	45
3.7.2 Monolithische Auftrennung	46
3.7.3 SRM Analysen zur Evaluierung des proteolytischen Verdaus.....	47
3.8 Entsalzung und Aufkonzentrierung von Peptidgemischen mittels C18 SPE	47
3.9 Stabil-Isotopen Markierung mittels iTRAQ.....	48
3.10 Verwendete Fraktionierungsmethoden	49
3.10.1 SCX.....	49
3.10.2 IEF	51
3.10.3 HILIC.....	51
3.10.4 COFRADIC	51
3.10.5 Phosphopeptid-Anreicherung mit TiO ₂	53
3.11 Massenspektrometrische Analyse mittels nano-LC-MS/MS	54
3.11.1 nano-LC.....	55

3.11.2	Verwendete Massenspektrometer	55
3.11.3	MS Akquisition.....	58
3.12	Qualitätskontrolle von nano-LC-MS Systemen mittels stabil-Isotopen markierter Peptide	59
3.12.1	Auswahl der stabil-Isotopen markierten Peptide und Komposition des Gemisches	59
3.12.2	Evaluierung grundlegender LC-MS Parameter.....	60
3.13	Datenbanksuchen	61
3.13.1	Datenbanksuchen zur Analyse von Degradationsprozessen	61
3.13.2	Optimierung von target/decoy Suchstrategien für verbesserte Identifikationsraten.....	63
3.13.3	Verwendung von UPS und <i>Pyrococcus fuoriosus</i> als Referenzproteome für target/decoy Strategien	64
3.14	Quantifizierung	65
3.14.1	SILAC – Proteome Discoverer.....	65
3.14.2	iTRAQ und NSAF – PeptideShaker	65
3.14.3	PFU als QC Instrument für die iTRAQ Quantifizierung.....	66
3.14.4	Optimierung und Qualitätskontrolle der iTRAQ - Arbeitsschritte	67
3.14.5	Labelfree - Nonlinear Progenesis	69
3.14.6	Stabil-Isotopen markierter Standard zur Qualitätskontrolle der LF Quantifizierung.....	70
3.15	Dateninterpretation.....	70
4.	Ergebnisse	72
4.1	Aufreinigung von Thrombozyten.....	72
4.2	Aufreinigung von Mitochondrien	73
4.3	Qualitätskontrolle des iTRAQ Workflows.....	73
4.3.1	Optimierung des proteolytischen Verdaus	74
4.3.2	Entsalzen und iTRAQ Markierungseffizienz	76
4.3.3	Dynamischer Bereich.....	78
4.3.4	Qualitätskontrollierte Datenauswertung von iTRAQ Analysen	78
4.3.5	Analyse der iTRAQ Interferenz	79
4.4	Vergleich verschiedener Trypsine bezüglich Effizienz und Reproduzier-barkeit	82
4.4.1	Monolith	82
4.4.2	Dreistufige Datenbanksuche	83
4.4.3	SCX.....	83
4.4.4	Direkte LC-MS.....	85
4.4.5	SRM	87
4.4.6	SDS-PAGE.....	88
4.5	Qualitätskontrolle von nano-LC-MS Systemen.....	88
4.6	Qualitätskontrollierte Datenauswertung von qualitativen proteomischen Daten / Optimierte statistische Datenauswertung	90
4.7	Charakterisierung des thrombozytären Proteoms.....	93
4.7.1	Abschätzung absoluter Proteinkopienzahlen in Thrombozyten.....	94
4.7.2	Quantitative Erfassung der inter-Donor Proteinvariabilität in Thrombozyten mittels iTRAQ.....	96
4.7.3	Erfassung der Variabilität über Kopiennummern	98
4.7.4	Intra-Donor Varianz	99

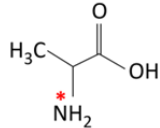
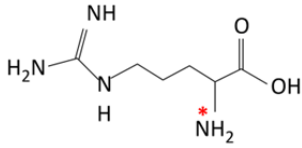
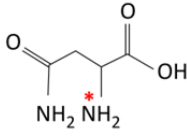
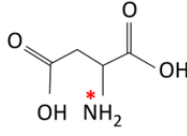
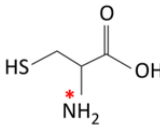
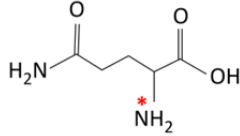
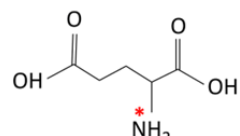
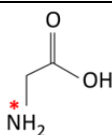
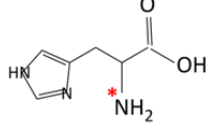
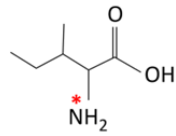
4.7.5	Globale Phosphorylierungsstudie zur Verbesserung der Proteomabdeckung..	99
4.7.6	Abgleich der Proteomdaten mit RNA Daten	100
4.7.7	Abgleich mit anderen Proteomdaten.....	100
4.8	Charakterisierung des mitochondrialen IMS Proteoms mittels SILAC	101
4.8.1	Validierung neuer IMS Proteine	102
4.9	LF Quantifizierung zur Charakterisierung von Proteindegradation in Mitochondrien	106
4.9.1	Qualitätskontrolle der LF Quantifizierung.....	106
4.9.2	Spectral counting.....	110
4.9.3	<i>Feature</i> -basierte LF Analyse	113
5.	Diskussion	118
5.1	Qualitätskontrollen in der qualitativen und quantitativen Proteomanalyse.....	118
5.2	Charakterisierung des humanen Thrombozytenproteoms.....	121
5.2.1	Das quantitative Thrombozytenproteom	122
5.2.2	Qualitativer wie quantitativer Abgleich des Thrombozytenproteoms mit globalen Proteomstudien humaner Zelllinien	123
5.2.3	Variabilität des Thrombozytenproteoms	125
5.3	Analyse von Organell-Subkompartimenten durch quantitative Proteomforschung am Beispiel des IMS.....	126
5.4	Mitochondriale Degradationsprozesse und die Definition von proteolytischen Schnittstellen.....	128
6.	Zusammenfassung.....	131
7.	Summary.....	132
8.	Literaturverzeichnis.....	132
9.	Danksagung.....	133
10.	Lebenslauf.....	144
11.	Erklärung.....	145
12.	Anhang.....	146
12.1	Verwendete Trypsine.....	146
12.2	SRM Übergänge zur Analyse der Verdaueffizienz	146
12.3	Modellproteom - UPS	148
12.4	Qualitätskontrolle von nano-LC-MS Systemen.....	149
12.5	Analyse des mitochondrialen IMS Proteoms (<i>S. cerevisiae</i>).....	149
	149
12.6	Charakterisierung der thrombozytären Proteoms	153
12.6.1	Abgleich von 68 CD Markern in RNA-, SAGE- und Proteomdaten	153

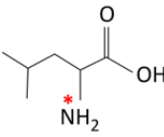
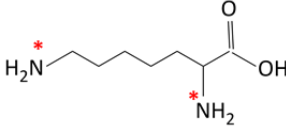
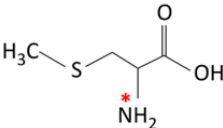
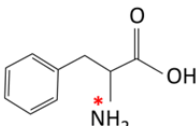
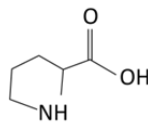
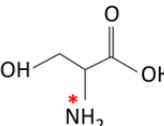
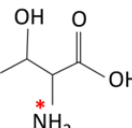
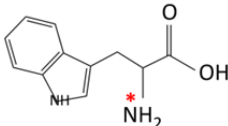
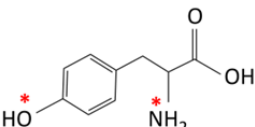
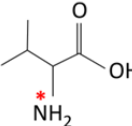
I Abkürzungsverzeichnis

1-DE	ein dimensionale Gelelektrophorese
2-DE	zwei dimensionale Gelelektrophorese
ANOVA	Varianzanalyse, engl. analysis of variance
AS	Aminosäuren
ACN	Acetonitril
BAX	Bcl-2-associated X protein
BCA	Bicinchoninsäure Assay
COFRADIC	Combined fractional diagonal chromatography
DDA	Daten abhängige Akquisition, engl. data dependant acquisition
DIA	Daten unabhängige Akquisition, engl. data independant acquisition
DTNB	Dithionitrobenzoesäure
DTT	Dithiotrethiol
EDTA	Ethylendiamintetraessigsäure
EGTA	Ethylenglycol-bis(aminoethylether)-tetraessigsäure
EtOH	Ethanol
FA	Ameisensäure, engl. formic acid
HEPES	2-(4-(2-Hydroxyethyl)-1-piperazinyl)-ethansulfonsäure
HILIC	Hydrophilic interaction liquid chromatography
HPLC	Hochleistungs-Flüssigchromatographie, engl. High performance liquid chromatography
IAA	Iodacetamid
ICAT	engl. Isotope-coded affinity tag
ICPL	engl. Isotope-coded protein label
ID	Peptid- bzw. Proteinidentifikation
IPG	immobilisierter pH-Gradient
IEF	Isoelektrische Fokussierung
IMS	Intermembranraum, engl. inter membrane space
iQuARI	engl. Integral quantification accuracy estimation for reporter ion-based quantitative proteomics
iTRAQ	engl. Isobaric tag for relative and absolute quantification
I.D.	Innendurchmesser

LF	Markierungsfrei, engl. label free
LOD	Detektionsgrenze, engl. limit of detection
LOI	Identifizierungsgrenze, engl. limit of identification
LOQ	Quantifizierungsgrenze, engl. limit of quantification
MAD	mittlere absolute Abweichung, engl. median absolute deviation
MOPS	3-(N-morpholino)propan sulfonsäure
MS	Massenspektrometrie
MW	Molekulargewicht, engl. molecular weight
m/z	Masse-zu-Ladungs-Verhältnis
PFU	Complex Proteomics Standard, entspricht dem Proteoms des Archaeobakteriums <i>Pyrococcus furiosus</i>
PMSF	Phenylmethylsulfonylfluorid
PSM	engl. Peptide to Spectrum Match
QC	Qualitätskontrolle, engl. QC
RPKM	engl. reads per kilobase per exon model per million mapped reads
SCX	Kationenaustauscher, engl. strong cation exchange
SIL	stabil-Isotopen markiert, engl. stable isotope labeled
SILAC	engl. Stable isotope labeling with amino acids in cell culture
SPE	Festphasenextraktion, engl. solid phase extraction
TCA	Trichloressigsäure, engl. trichloroacetic acid
TFE	Trifluorethanol
TMT	engl. Tandem mass tag
TFA	Trifluoressigsäure, engl. tri fluoro acetic acid
TiO ₂	Titandioxid
TPCK	L-1-Tosylamido-2-Phenylethyl-chloromethyl-keton
UPS	Universal Proteome Standard

II Aminosäuren

Name	Buchstabencodes	Strukturformel * potentielle iTRAQ Lokalisation	Monoisotopische Masse (-H ₂ O) / Da [2]
Alanin	Ala, A		71,03711
Arginin	Arg, R		156,10111
Asparagin	Asn, N		114,04293
Asparaginsäure	Asp, D		115,02694
Cystein	Cys, C		103,00919
Glutamin in N-terminaler Position Cyclisierung möglich	Gln, Q		128,05858
Glutaminsäure in N-terminaler Position Cyclisierung möglich	Gln, E		129,04259
Glycin	Gly, G		57,02146
Histidin	His, H		137,05891
Isoleucin	Ile, I		113,08406

Name	Buchstabencodes	Strukturformel	Monoisotopische Masse (-H ₂ O) / Da
Leucin	Leu, L		113,08406
Lysin	Lys, K		128,09496
Methionin	Met, M		131,04049
Phenylalanin	Phe, F		147,06841
Prolin	Pro, P		97,05276
Serin	Ser, S		87,03203
Threonin	Thr, T		101,04768
Tryptophan	Trp, W		186,07931
Tyrosin	Tyr, Y		163,06333
Valin	Val, V		99,06841

III Veröffentlichungen

Bisherige Veröffentlichungen:

- Vögtle, F.-N.* , Burkhardt, J.M.*, Rao, S., Gerbeth, C., Hinrichs, J., Martinou, J.-C., Chacinska, A., Sickmann, A. , Zahedi, R.P., Meisinger, C.; Intermembrane space proteome of yeast mitochondria. 2012, **Mol Cell Proteomics**, 11(10):1840-52.
- Vaudel, M. * , Burkhardt, J.M. *, Radau S., Zahedi R.P., Martens, L., Sickmann, A.; Integral Quantification Accuracy estimation for Reporter Ion based quantitative proteomics (iQuARI). 2012, **J Proteome Res**, 11(10):5072-80.
- Burkhardt, J.M.*, Vaudel, M.* , Gambaryan, S., Radau, S., Walter, U., Martens, L., Geiger, J., Sickmann, A., Zahedi, R.P.; The first comprehensive and quantitative analysis of human platelet protein composition allows the comparative analysis of structural and functional pathways in human platelets. 2012, **Blood**, 120(15):e73-82.
- Beck, F. * , Burkhardt, J.M. *, Geiger, J., Zahedi, R.P., Sickmann, A.; Robust workflow for iTRAQ based peptide and protein quantification. 2012, **Methods Mol Biol**, 893:101-13.
- Vaudel, M. * , Burkhardt, J.M., Zahedi, R.P., Martens, L., Sickmann, A.; iTRAQ data interpretation. 2012, **Methods Mol Biol**, 893:501-9.
- Vaudel, M. * , Burkhardt, J.M. *, Zahedi, R.P., Sickmann, A., Martens, L.; A complex standard for protein identification, designed by evolution. 2012, **J Proteome Res**, 11(10):5065-71.
- Burkhardt, J.M., Schumbrutzki, C., Wortelkamp, S., Sickmann, A., Zahedi, R.P.; Systematic and quantitative comparison of digest efficiency and specificity reveals the impact of trypsin quality on MS-based proteomics. 2012, **J Proteomics**, 75(4):1454-62.
- Vaudel, M., Burkhardt, J.M., Sickmann, A., Martens, L., Zahedi, R.P.; Peptide identification quality control. 2011, **Proteomics**, 11(10):2105-14.
- Burkhardt, J.M. *, Vaudel, M. * , Zahedi, R.P., Martens, L., Sickmann, A.; iTRAQ protein quantification: A quality controlled workflow. 2011, **Proteomics**, 11(6):1025-34.
- Burkhardt, J.M. *, Prensler T. * , Sickmann, A.; Quality control of nano-LC-MS systems using stable isotope-coded peptides. 2011, **Proteomics**, 11(6):1049-57.

* geteilte Erstautorenschaft

Übersichtsartikel und Buchkapitel:

Burkhardt, J.M., Sickmann, A.; Mass spectrometry-based proteomics: general overview and posttranslational modification analysis in the context of platelet research. 2011, **In: Platelet Proteomics: Principles, Analysis, and Applications**, Editors: Garcia, A., Senis, Y.;

IV Vorträge

'Quality control of nano-LC-MS systems using stable isotope-coded peptides', Proteomic Forum 2011, Berlin, April 2011.

'Quality control of nano-LC-MS systems using stable isotope-coded peptides', ISAS summer school, Dortmund 2011.

'Ein maßgeschneiderter Standard für die labelfreie Quantifizierung - und mehr!', 18. Arbeitstagung Mikromethoden der Proteinchemie, Martinsried, Juni 2011.

'Qualitätssicherung in der Proteomanalyse', Jahrestagung der Sektion Molekulare Diagnostik der DGKL 2011, Tutzing, Mai 2012.

'Quantitative proteome analysis – application to human platelets', 9. Jahrestagung der DGKL, Mannheim, September 2012.

V Posterpräsentationen

- 2008 7th HUPO World Congress Amsterdam, Niederlande
“Phosphorylation specific interaction and regulation of VASP in platelets”
J.M. Burkhart*, T. Premsler, U. Walter, T. Renné, A. Sickmann
- 2009 Proteomic Forum 2009, Berlin
“Phosphorylation specific interaction and regulation of VASP in platelets”
J.M. Burkhart*, T. Premsler, U. Walter, T. Renné, A. Sickmann
- 2009 ISAS, summer school
“Novel methods for high through put proteome analysis”
J.M. Burkhart*, R.P. Zahedi, A. Sickmann
- 2010 43. DGMS Jahrestagung, Halle
“Strategies to a complete organellar proteome”
A. Sickmann, J.M. Burkhart*, C. Meisinger, N. Pfanner, S. Wortelkamp, R. Zahedi
- 2010 ISAS, summer school
“Label free analysis of complexe peptide mixtures”
J.M. Burkhart*, R.P. Zahedi, A. Sickmann
- 2010 ISAS, Evaluierung der Leibniz Gemeinschaft
“Global and unbiased quantitative analysis of biological systems”
J.M. Burkhart*, M. Vaudel*, R.P. Zahedi, A. Sickmann
- “Targeted quantification of trace components in biological systems”
J.M. Burkhart*, S. Wortelkamp, R.P. Zahedi, A. Sickmann
- 2011 44. DGMS Jahrestagung, Dortmund
“Quality control of nano-LC-MS systems using stable isotope coded peptides”
J.M. Burkhart*, T. Premsler, A. Sickmann
- 2011 59th ASMS Denver, USA
“Quality control of nano-LC-MS systems using stable isotope coded peptides”
J.M. Burkhart*, T. Premsler, A. Sickmann
- “OpenMS / TOPP analysis pipeline for label free quantification”
H.Bauer, J. M. Burkhart, L. Nilse*, A. Sickmann, R.P. Zahedi, O. Kohlbacher
- 2011 1. Marktplatz bio.Dortmund
“Quality control of nano-LC-MS systems using stable isotope coded peptides”
J.M. Burkhart*, T. Premsler, A. Sickmann
- 2012 9. DGKL Jahrestagung, Mannheim
„Systematic characterization of human platelets by quantitative proteomics “
J.M. Burkhart*, M. Vaudel, S. Gambaryan, S. Radau, U. Walter, L. Martens, J. Geiger, A. Sickmann, R.P. Zahedi
- 2012 2. Marktplatz bio.Dortmund
„Systematic characterization of human platelets by quantitative proteomics “
F. Beck*, K. Becker*, J.M. Burkhart*, R.P. Zahedi, A. Sickmann

*Präsentierende/r Autoren

1. Einleitung

1.1 Globale Proteomstudien

Biologische Systeme weisen sich als komplexe Gefüge aus, deren heterogene Komponenten mit den analytischen Disziplinen der Genomik, Transkriptomik, Proteomik, Metabolomik und Fluxomik näherungsweise dargestellt werden können. Während das Genom weitgehend statischer Natur ist, spiegelt das Produkt der Genexpression, das Proteom eine Momentaufnahme wider. Das Proteom wird stark vom Zeitpunkt des Zellzyklus, von der Funktion und den zugrundeliegenden metabolischen Prozessen des jeweiligen Zellsystems sowie von vielen weiteren intrinsischen und extrinsischen Faktoren beeinflusst. Die Konzentration des Transkripts korreliert nur bedingt mit dem qualitativen und quantitativen Aufkommen von Proteinen in einem biologischen Netzwerk [3, 4]. Vielmehr sind es spezifische Stimuli, welche die Translation, Degradation, Translokation sowie Modifizierung von Proteinen bewirken und somit die Proteinkonzentration zeit- und orts aufgelöst in den zellulären Kompartimenten beeinflussen.

Eine humane Zelle mit einem bekannten Genom von mehr als 22.000 Genloci [5] weist zu einem Zeitpunkt unter definierten Bedingungen ca. 10.000 unterschiedliche Proteine auf [6, 7]. Der dynamische Bereich von Proteinen in einer Zelle wird dabei auf etwa sechs bis sieben Größenordnungen geschätzt [7] und wird deutlich im Vergleich von hochabundanten Strukturproteinen, die in mehreren Millionen Kopien pro Zelle vorkommen, z. B. zu niedrigabundanten Rezeptoren mit nur wenigen Kopien pro Zelle. Die Komplexität innerhalb einer Zelle wird zusätzlich durch dynamische Prozesse wie den posttranslationalen Modifikationen erhöht, die eine vielseitige Regulation und Modulation von Proteinaktivitäten und –funktionen bewirken.

1.2 Strategien der Proteomanalyse

Die Proteinsequenzierung einzelner Proteine nach Edman [8] sowie die Auftrennung komplexer Proteingemische mittels 2-DE [9] waren lange Zeit bewährte Ansätze in der Proteinchemie. In den letzten Jahren zeigte sich allerdings mehr und mehr, dass die Auflösung der 2-DE nicht ausreicht, um Proteingemische für differentielle Analysen zufriedenstellend aufzutrennen: Einerseits beinhalten einzelne Spots oft aufgrund von Ko-Migration bis zu 10 Proteine, andererseits verteilen sich Isoformen sowie Degradationsprodukte hochabundanter Proteine oft auf zahlreiche Spots [10].

Im Gegensatz dazu erlaubt die Verwendung von effizienten Präfraktionierungsmethoden in Kombination mit der hochauflösenden Massenspektrometrie heutzutage die Analyse von hoch komplexen Proben und ermöglicht über die deskriptive Darstellung hinaus die quantitative Erfassung von dynamischen Prozessen in zellulären Netzwerken. Selektivität und Sensitivität spielen zentrale Rollen in der Analyse von biologischen Proben, deren hochgestecktes Ziel die Abdeckung des kompletten Proteoms ist. Eine Schlüsselrolle nimmt hier die flüssigchromatographische (LC) Auftrennung von Peptidgemischen in Kombination mit der Massenspektrometrie ein.

1.2.1 Flüssigchromatographie

In der Flüssigchromatographie (LC) werden Substanzgemische in einer mobilen Phase gelöst und über Interaktion mit einer stationären Phase aufgetrennt. Die Komponenten des Gemisches interagieren entsprechend ihrer physiko-chemischen Eigenschaften unterschiedlich stark mit der stationären Phase und eluieren somit zu definierten Retentionszeiten. Auskunft über das Adsorptionsverhalten von Analyten an der stationären Phase geben spezifische Parameter wie der Retentionsfaktor k , die Selektivität α sowie die Anzahl der theoretischen Trennböden N und die Auflösung R [11].

Weiterhin ist die Trennbodenhöhe H als Maß für die Effizienz der Trennung durch die *van Deemter*-Gleichung beschrieben, welche ursprünglich für die Gaschromatographie entwickelt wurde [12]. Aus ihr geht hervor, dass die Trennbodenhöhe sowohl durch die Fließgeschwindigkeit u als auch durch ein Zusammenspiel von Diffusionsprozessen beeinflusst wird: je niedriger H , umso höher ist die Trennleistung einer chromatographischen Säule (Abbildung 1.1).

Bei optimaler Trennung und Wahl der Fließgeschwindigkeit u wird für ein chromatographisches System ein Minimum für H erreicht, welches der maximalen Anzahl an Trennböden entspricht.

Die Peakkapazität ist ein Maß für die Auflösung eines chromatographischen Systems und als die Anzahl der aufgelösten normalverteilten Peaks innerhalb eines Gradienten definiert [13, 14].

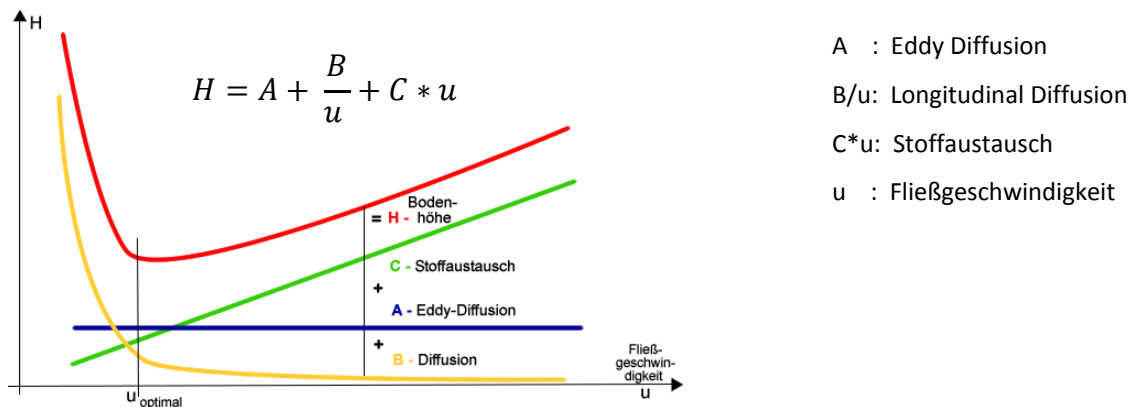


Abbildung 1.1: Van Deemter Gleichung. Die Abhängigkeit der einzelnen Terme der *van Deemter* Gleichung sowie die resultierende Trennbodenhöhe H sind in Abhängigkeit von der Flussgeschwindigkeit u angegeben. Quelle: http://www.chemgapedia.de/vsengine/media/vsc/de/ch/3/anc/croma/basics/saulen_chr/deemter/ht_u_bild3gr1002.gif

Durch die Miniaturisierung von LC Systemen konnte die Sensitivität von chromatographischen Trennungen verbessert werden (siehe Tabelle 1). Neben den Pumpen zur Förderung der mobilen Phase im Nanoliter-Maßstab mussten u.a. Mischkammern, Detektoren und entsprechende Flusszellen sowie die verwendeten Trennsäulen optimiert werden. Durch den Einsatz von Partikeln mit geringen Durchmessern von 1.7 – 5 µm und einer dichteren Packung der stationären Phase, in Kombination mit einem reduzierten Säuleninnendurchmesser von 25 – 100 µm, konnten verkürzte Retentionszeiten und intensivere Signale mit niedrigeren Peakhalbwertsbreiten bei stark erhöhtem Druck erzielt werden [15]. Die damit verbundenen geringeren Flussraten führen zu verringerten Bodenhöhen und limitieren das Probeninjektionsvolumen.

Fortan wurde der Begriff der Hochleistungsflüssigkeitschromatographie (HPLC) im nano Maßstab geprägt, deren Kombination mit der Massenspektrometrie den Nachweis von Proteinen / Peptiden im femtomol bis attomol Bereich ermöglicht [16].

Zur Auftrennung von Protein- und Peptidgemischen wird häufig die Umkehrphasenchromatographie (RP) herangezogen, bei welcher hydrophobe Wechselwirkungen an Kohlenwasserstoffketten auf der stationären Phase zu

Tabelle 1.1: Anstieg der Sensitivität von LC Systemen bezogen auf den Säulendurchmesser [1].

LC-Dimension	Säuleninnendurchmesser [mm]	Relative Sensitivität	Flussrate [µL/min]
Konventionelle LC	4,6	≅1	>1000
Analytische LC	2,0	5	200
Mikro LC	0,5-1,0	35	10-50
Kapillar LC	0,1-0,5	200	1-10
Nano LC	0,05-0,1	3500	0,05-0,2

Retardierung genutzt werden. Polare Gruppen der Analyten werden durch Zugabe eines Ionenpaarreagens wie Trifluoressigsäure (TFA) maskiert und somit die Hydrophobizität erhöht. Für die Auftrennung werden die Analyten im wässrigen Milieu auf die stationäre Phase geladen, um anschließend in einem binären Gradienten durch den steigenden Anteil eines organischen Lösungsmittels wie Acetonitril (ACN) von der Säule eluiert zu werden. Häufig wird die LC mit einer UV-Flusszelle kombiniert, die die Detektion der Peptidbindung bei 214 nm bzw. der aromatischen Seitenketten von Phe, Tyr und Trp bei 295 nm ermöglicht. Die bei der RP Chromatographie verwendeten Lösungsmittel ermöglichen eine direkte Kopplung mit der Elektrospray Ionisierung (ESI) für die LC-MS Analytik. Das von Fenn et al. etablierte ESI-Verfahren ermöglicht die schonende Überführung ionisierter Biomoleküle aus der wässrigen Phase in die Gasphase und erlaubt somit deren massenspektrometrische Detektion und Analyse [17].

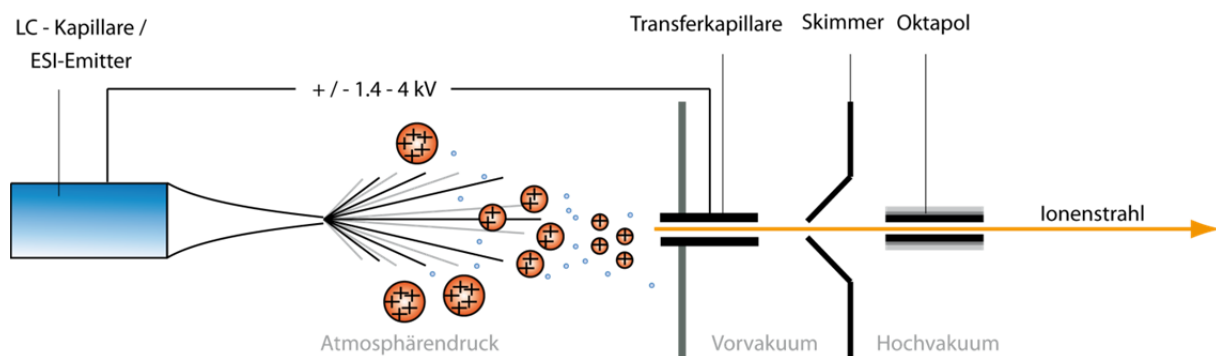


Abbildung 1.2: ESI Prozess. Zwischen ESI Emitter und Eingang des MS ist eine Hochspannung angelegt. Es bildet sich ein Flüssigkeitsfilament aus, von dem einzelne Tröpfchen abreißen. Aufgrund der Verdampfung des flüchtigen Lösungsmittels erhöht sich deren Ladungsdichte, sodass sie schließlich durch elektrostatische Abstoßung in sogenannten Coulomb-Explosionen zu kleinen Tröpfchen mit geringerer Ladungsdichte zerfallen. Dieser Prozess setzt sich fort, bis schließlich die Analytionen in die Gasphase und das MS überführt werden.

Für die Analyse von hochkomplexen Proben, wie z.B. Komplettverdauen von Zelllysaten, ist eine effiziente LC Auftrennung unerlässlich, um Koelutionen für die nachfolgende MS-Detektion zu minimieren [18] und somit den dynamischen Bereich zu vergrößern. Aus diesem Grund muss die Peakkapazität eines chromatographischen Systems erhöht werden, was durch Variation (a) der Lösungsmittelzusammensetzung, (b) des Säulenmaterials, (c) der Säulenlänge [19], (d) der Partikelgröße, (e) der Temperatur, (f) der Gradientendauer und (g) der Flussrate beeinflusst werden, wobei der Variation der einzelnen Faktoren durch

potentiellen Druckerhöhung, Peakverbreiterung oder verminderte Ionisierungseffizienz Grenzen gesetzt sind.

Ergänzend zur Optimierung eindimensionaler LC Systeme sind in der Literatur zweidimensionale RP-LC [20, 21] als auch orthogonale Systeme wie die *Multidimensional Protein Identification Technology* (MudPIT) beschrieben [22]. Diese können sowohl „online“ mit MS gekoppelt, als auch „offline“ zur Präfraktionierung und Reduzierung der Probenkomplexität eingesetzt werden (siehe Kapitel 1.2.2). Zudem ermöglicht die MS Kopplung, dank der in den letzten Dekaden eingeführten hochauflösenden Instrumente, mit erhöhter Scangeschwindigkeit die Analyse von mehr koeluierenden Analyten [23]. Hochauflösende MS Systeme können Masse-zu-Ladungs-Verhältnisse (m/z) sowie Ladungszustände von Analyten und daraus resultierend deren akkurate Masse auch aus höchstkomplexen Proben mit zahlreichen koeluierenden Analyten bestimmen. Dies vereinfacht wiederum die nachfolgenden Datenbanksuchen und erhöht die Validität der Ergebnisse enorm.

1.2.2 Präfraktionierungsmethoden

Die hochauflösende LC-MS Analytik hat in den letzten 10 Jahren aufgrund der enormen Erhöhung der Scangeschwindigkeiten (z.B. Orbitrap Analysatoren) stark zum Erfolg der Proteomanalyse beigetragen. So konnte die Abdeckung von Proteomen über einen großen dynamischen Bereich deutlich verbessert werden, sie ist jedoch nach wie vor eine Herausforderung für die LC-MS Analytik. Zur detailreichen Analyse von biologischen Proben mit großem dynamischen Bereich werden daher meist zweidimensionale Methoden angewendet, die auf orthogonalen Trennprinzipien basieren und so ergänzend zur eindimensionalen RP-Auftrennung den Trennraum vergrößern [24]. Die für die Präfraktionierung genutzten Charakteristika von Peptiden wie Ladung, isoelektrischer Punkt (pI) oder Hydrophobizität können weiterhin wertvolle Informationen für die Proteomanalyse liefern. Die in dieser Arbeit verwendeten Fraktionierungstechniken werden im Folgenden kurz vorgestellt.

1.2.2.1 Isoelektrische Fokussierung

Die isoelektrische Fokussierung (IEF) ist eine seit Mitte der 70er Jahre bekannte Fraktionierungsmethode, die ursprünglich mit der 2-DE eingeführt wurde [9, 25]. Zu Beginn

erfolgte die IEF von Proteinen und Peptiden gelbasiert mittels mobiler Ampholyte, die nach Anlegen von Spannung entsprechend ihres pI einen pH Gradienten im Gel aufbauen. Der ausgebildete pH Gradient führte aufgrund elektroosmotischer Kräfte zu zeitabhängigen Migrationen der Analyten ohne stabile Fokussierung. Mit der Einführung von immobilisierten pH Gradienten, die kovalent durch Kopolymerisation von Pufferkomponenten und Acrylamid inkorporiert sind, konnten stabile pH Gradienten erzielt werden, die, sobald ein Analyt seinen pI erreicht hat zu einer Fokussierung führen und ein Driften verhindern [26].

Mit der präoperativen IEF orientierte sich dieses Trennverfahren weg von der 2-DE hin zur Fraktionierungsmethode für die MS Analytik [27]. Die komplexe Probe wird in einem zwei phasigen System bestehend aus einem konventionellem IPG Gelstreifen und einer flüssigen ampholythaltigen Phase in 12 bzw. 24 Fraktionen aufgetrennt [28]. Die IEF ist orthogonal zur RP-Auftrennung und gibt mit der Fraktionierung Informationen zum pI eines Peptides preis, welcher charakteristisch ist und theoretisch zur Identifizierung desselben ergänzend genutzt werden kann.

1.2.2.2 Kationenaustausch-Chromatographie

In der Kationenaustausch-Chromatographie (engl. strong cation exchange, SCX) können positiv geladene Analyten durch Wechselwirkungen mit einer negativ geladenen stationären Phase anhand ihrer Ladungszustände aufgetrennt werden [29].

Die in der Proteomanalyse häufig verwendete Protease Trypsin generiert Peptide, die im sauren Milieu bei pH 2,7 charakteristischerweise protoniert vorliegen. Die große Mehrheit der tryptischen Peptide ist zweifach geladen, da sie eine positive Ladung am C-terminalen Lys/Arg-Rest und eine weitere am N-Terminus tragen. Hingegen können PTMs wie Phosphorylierung oder N-terminale Acetylierung ladungsreduzierend wirken, sodass entsprechende tryptische Peptide meist einfach geladen sind, genauso wie die meisten C-terminalen Peptide [30-32]. Falls weitere positive Ladungsträger, wie überlesene Schnittstellen oder His-Reste, in der Sequenz vorhanden sind, kann die Nettoladung drei und mehr betragen.

Mit Hilfe der SCX können derartige Unterschiede der Nettoladung von Analyten aufgelöst werden.

1.2.2.3 Hydrophile Interaktionschromatographie

Die Hydrophile Interaktionschromatographie (engl. hydrophilic liquid chromatography, HILIC) weist eine hohe Orthogonalität zur RP Chromatographie auf [24]. Ihr Retentionsmechanismus ist auf elektrostatische Wechselwirkungen zwischen polaren Wassermolekülen und geladenen Peptiden zurückzuführen [33]. Als stationäre Phase dienen polare Strukturen, die SCX, Anionenaustauscher (engl. strong anion exchange, SAX) oder ein zwitterionisches Gemisch [34] der beiden sein können. Die Retentionsmechanismen unterscheiden sich aufgrund der Heterogenität der stationären Phasen, wobei ihnen allen gemein die Ausbildung einer wässrigen Schicht ist, die die funktionalen Gruppen überzieht und durch Anordnung der Wassermoleküle eine einheitlich bzw. zwitterionische Ladung auf der Oberfläche aufweist. Die Retentionszeiten der Analyten verhalten sich dabei invers zu jenen der RP-Phase: Während hydrophobe Analyten zu Beginn des Gradienten bei hohem Anteil eines mit wassermischbaren organischen Lösungsmitteln eluieren, eluieren Analyten hoher Polarität spät mit dem ansteigenden Anteil eines wässrigen Lösungsmittels.

1.2.2.4 Combined Fractional Diagonal Chromatography

Die *Combined Fractional Diagonal Chromatography* (COFRADIC) verzichtet auf Orthogonalität der chromatographischen Trennung [35]. Es werden zunächst in der ersten RP-LC Auftrennung Fraktionen gesammelt, die anschließend selektiv an bestimmten Aminosäureresten, Amino-Termini oder posttranslationalen Modifikationen durch eine chemische oder enzymatische Reaktion derart derivatisiert werden, so dass sich ihr chromatographisches Verhalten von underivatisierten Peptidspezies der jeweiligen Fraktionen unterscheidet. In einer folgenden Rechromatographie unter identischen Bedingungen können derivatisierte und nicht-derivatisierte Peptide getrennt werden, da je nach Derivatisierungsart, der eine Teil im selben Zeitfenster wie zuvor eluiert, während der andere ein stark abweichendes Elutionsverhalten aufweist. COFRADIC wurde bisher u.a. für die gezielte Analyse von methionin- [35] und cysteinhaltigen Peptiden, Phospho- [36] und Glykopeptiden [37] sowie von N-terminalen Peptiden angewendet und kann die Probenkomplexität entsprechend des Vorkommens der Zielstrukturen massiv reduzieren. Der Ausschluss von Peptiden, die nicht im Fokus der Derivatisierung stehen ist für die globale Proteomanalyse kritisch zu betrachten, da hierbei wertvolle Information verloren geht.

1.2.3 Massenspektrometrische Messverfahren

Die in der Proteomanalytik verwendeten LC-MS Technologien verfolgen primär die Ziele einer höheren Sequenzabdeckung, einer größeren Sensitivität zur Detektion von niedrig abundanten in Anwesenheit von hochabundanten Peptidkomponenten und damit die Abdeckung eines großen dynamischen Bereichs. Generell basiert die Identifizierung von Peptiden auf der Generierung von Fragmentationenspektren (MS/MS) und deren Abgleich mit *in silico* verdauten und prozessierten Proteindatenbanken mittels etablierter Suchalgorithmen wie z.B. Mascot [38], Sequest [39], OMSSA [40] oder X!Tandem [41]. Nachfolgend beschriebene MS Operationen finden häufige Anwendung.

1.2.3.1 DDA und DIA

Die am häufigsten genutzte LC-MS Anwendung ist die datenabhängige Akquisition (engl. data-dependent acquisition, DDA). Hierbei werden, basierend auf einem MS-Übersichtsspektrum, die signalintensivsten Ionen für nachfolgende MS/MS Scans ausgewählt. Die Anzahl der MS/MS Spektren sowie die Dauer eines solchen Zyklus hängt stark von der Scangeschwindigkeit des Massenspektrometers sowie der angewandten Parameter ab, und liegt typischerweise bei ca. 1 bis 2 Sekunden.

Wenn die Zahl der zu analysierenden Signale in komplexen Proben die Kapazität des MS übersteigt, kommen dynamische Ausschlusslisten (engl. dynamic exclusion) zum Einsatz, welche bereits fragmentierte Vorläufer m/z Werte für eine definierte Dauer, z.B. 10 s (abhängig von den Peakbreiten des LC Systems), von der erneuten MS/MS Analyse ausschließen. Dies reduziert redundante Fragmentationenspektren und eröffnet mit der Analyse von Analyten geringerer Intensität einen größeren dynamischen Bereich der Messung. Aufgrund der stark unterschiedlichen Ionisierungs- und Detektionseffizienz verschiedener Analyten kann jedoch von Unterschieden in deren Signalintensitäten nicht direkt auf die vorliegenden Mengenverhältnisse geschlossen werden, sodass die Vergrößerung des biologischen und des intensitätsbasierten dynamischen Bereichs nicht zwangsläufig korrelieren.

Weiterführende Strategien setzen zeitabhängige Ausschluss- bzw. Einschlusslisten bekannter und unbekannter Vorläuferionen ein, die in aufeinanderfolgenden iterativen LC-MS Analysen ergänzt bzw. abgearbeitet werden und somit zur verbesserten Abdeckung des Proteoms führen sollen.

Neben ihrem Einsatz in der qualitativen Proteomanalyse wird die DDA Technologie ebenfalls in quantitativen Ansätzen verwendet, wobei hier Strategien unterschieden werden, die auf der Ebene des Übersichtsscans bzw. auf der Ebene des Fragmentationenspektrums quantifizieren. Diese Verfahren werden am Ende des Kapitels im Detail beschrieben.

Weiterhin werden in der Literatur qualitative wie auch quantitative Ansätze besprochen, die auf einer datenunabhängigen Akquisition (engl. data-independent acquisition, DIA) beruhen [42-44]. Ihnen allen gemein ist eine Unterteilung in m/z -Fenster variierender Breite von 1.5 – 10 m/z , in welchen alle vorhandenen Signale unabhängig ihrer Intensität und ohne selektive Isolation gemeinsam fragmentiert werden. DIA hat den Anspruch, sämtliche ko-eluierenden Analyten zu detektieren, diese reproduzierbar mittels der Information aus dem zeitlichen Abgleich von Ionenspuren der Vorläufer- und der dazugehörigen Fragmentationen (engl. precursor-/fragment-ion lineage) zu identifizieren und somit die Sequenzabdeckung zu erhöhen [45, 46]. Neben der hohen Selektivität aufgrund der kombinierten Zuordnung zeichnet sich diese Strategie mit einer hohen Sensitivität aus, da auf definierte Grenzwerte (engl. ion count threshold), die für eine Fragmentierung mittels DDA überschritten werden müssen, verzichtet wird. Um die Fülle an Signalen in einer akzeptablen Analysenzeit zu identifizieren, kommen MS Geräte mit sehr hohen Scangeschwindigkeiten zum Einsatz, die u.a. mit einer Ionen-Mobilitäts-Zelle ausgestattet sind (Synapt G2-S HDMS, Waters).

1.2.3.2 Labelfreie Quantifizierung

Bei der LF Quantifizierung können grundsätzlich zwei verschiedene Strategien unterschieden werden: (a) die „accurate mass and time tag“ (AMT) Technologie [47, 48], sowie (b) das „spectral counting“ [49].

AMT basiert auf dem relativen Vergleich identischer Signale in unabhängigen LC-MS Analysen, welcher bei Korrelation mit einem internen Standard auch für die absolute Quantifizierung herangezogen werden kann. Zur Quantifizierung werden MS^1 Spektren betrachtet, deren Information bestehend aus (a) Retentionszeit, (b) m/z - Verhältnis, (c) Isotopenmuster und (d) Signalintensität sogenannte ‚Features‘ beschreiben, die einen globalen quantitativen Überblick eines Proteoms zeichnen.

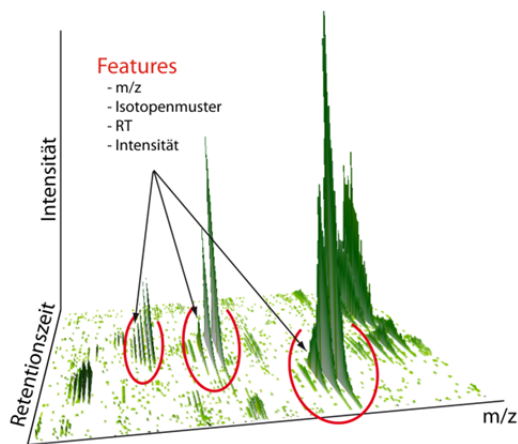


Abbildung 1.3: Schematische Darstellung von Features. Ein jedes Feature weist die eindeutige Charakteristika m/z , Isotopenmuster, RT und Intensität auf und ist in der gezeigten Abbildung mit einem roten Kreis gekennzeichnet.

Neben der Reproduzierbarkeit des LC-MS Systems und der Bestimmung der akkuraten Masse mittels hochauflösender MS spielt die Sensitivität in LF Analysen eine herausragende Rolle, da diese oft innerhalb einer globalen Messung - ohne Präfraktionierung zur Reduzierung der Komplexität - durchgeführt werden. Nachdem DDA Signale intensitätsabhängig für die Fragmentierung auswählt, können Strategien wie Einschluss- oder Ausschlusslisten angewendet werden, die in komplexen Proben die Analyse von niedrig abundanten Peptidspezies ermöglichen. Beim spectral counting wird die Anzahl identifizierter MS/MS Spektren eines Proteins als Maß für dessen Abundanz innerhalb einer Probe sowie zwischen differentiellen, reproduzierbar analysierten Proben verwendet. Grundsätzlich hängt die Güte der Quantifizierung bei allen Verfahren von der Qualität und Reproduzierbarkeit der Probengenerierung sowie –präparation ab. Während bei AMT die Reproduzierbarkeit und Stabilität des LC-MS Systems essentiell sind um eine gute Vergleichbarkeit der Proben zu gewährleisten, ist das spectral counting diesbezüglich robuster. Die Aussagekraft des spectral countings hängt vor allem davon ab, wie viele Spektren pro Protein identifiziert wurden (Statistik) und somit von Faktoren wie Probenkomplexität, Fraktionierung, Gradientenlänge, Scangeschwindigkeit sowie den Parametern für die dynamischen Ausschlusslisten.

1.2.3.3 SRM

Während die LF Quantifizierung einen globalen quantitativen und qualitativen Überblick von meist komplexen biologischen Proben liefert, fokussiert die Single Reaction Monitoring (SRM) Strategie allein die Quantifizierung eines Sets an bekannten Analyten. Dank der hohen Sensitivität ermöglicht SRM die Analyse von niedrig abundanten Peptiden mit nur wenigen Kopien in komplexen Gemischen wie Zelllysaten über einen dynamischen Bereich von 5 Größenordnungen [50]. SRM wird u.a. zur Validierung von quantitativen Ergebnissen aus

individuellen Transitions (engl. scheduled SRM) in Vorversuchen optimiert werden [51]. Kollisionsenergien müssen für ein jedes Peptid individuell durch Einsatz von synthetischen Peptiden oder alternativ und auch weniger arbeitsintensiv durch globale LC-MS-Analysen mit vergleichbarem Kollisionsverhalten evaluiert werden [52]. Für die Realisierung der scheduled SRM ist eine hochreproduzierbare chromatographische Auftrennung maßgebend.

1.3 Quantitative Proteomforschung

Die Quantifizierung in der Proteomforschung erfolgte in den Anfängen zunächst mittels Fluoreszenzfarbstoffen und radioaktiven Markierungen, die je nach Label und Detektor bereits eine gute Sensitivität, Linearität als auch einen dynamischen Bereich von fünf Größenordnungen abdeckten [53]. Diese Quantifizierungsstrategien sind jedoch auf die Bestimmung von Proteinkonzentrationen beschränkt, ohne die zugrundeliegenden Proteine zu identifizieren. Die häufig dazu angewandte Auftrennung von Proteingemischen mittels 2-DE [9, 25] erlaubt die Trennung und Visualisierung von bis zu 5000-10000 Proteinspots [53, 54]. Die Methode ist jedoch auf lösliche und abundante Proteine sowie – bei Kopplung mit MS – durch die hohe Anzahl zu analysierender Gelspots limitiert. Co-Migration von Proteinen führt zudem häufig zur Identifikation mehrerer Proteine in einem Spot und folglich zu einer Unschärfe der Quantifizierung.

Die Massenspektrometrie ermöglicht die Identifizierung von Proteinen, ist aber per se nicht quantitativ: Die Quantifizierung von biologischen Proben ist aufgrund der physikochemischen Vielseitigkeit der Analyten und der damit verbundenen variierenden Signalantwort nicht allein von der Konzentration des Analyten abhängig. Es werden daher chemische und metabolische Markierungsstrategien, die auf stabil-Isotopen Markierung (engl. stable isotope labeled, SIL) basieren, sowie markierungsfreie Ansätze eingesetzt, die eine relative Quantifizierung zweier differentieller Proben bzw. bei Einsatz eines internen Standards auch eine absolute Quantifizierung erlauben. Bei der Auswahl der Quantifizierungsstrategie müssen eine Vielzahl an Aspekten abgewägt werden: SIL basierte Verfahren erlauben das Vereinen unterschiedlich markierter Proben (engl. multiplexing) vor der LC-MS Analyse und weisen sich damit durch einen geringen Messaufwand aus. Durch den Einsatz einer Variation der Isotope ^{13}C , ^{15}N und / oder ^{18}O , die keinen Einfluss auf die chromatographische Trennung nehmen, eluieren identische Peptide unterschiedlicher Markierung zu einem Zeitpunkt. Dem stehen allerdings vielseitige Faktoren wie (a) die

Zugänglichkeit der biologischen Probe, (b) eine aufwendige Probenaufarbeitung, (c) hohe Kosten für stabil-Isotopen markierte Reagenzien, und u.U. (d) ein Quantifizierungslimit bzw. (e) eine erhöhte Probenkomplexität gegenüber. Die LF Quantifizierung fokussiert den relativen Vergleich von sogenannten Features in unabhängigen LC-MS Messungen, die durch die Charakteristika (a) Masse-zu-Ladungs-Verhältnis (m/z) und Isotopenverteilung, (b) Retentionszeit und (c) Intensität dreidimensional beschrieben werden. Da das Multiplexen von LF Proben nicht möglich ist, erfordert die LF Quantifizierung einen höheren zeitlichen Messaufwand. Die LF Quantifizierung erhebt für den quantitativen Vergleich weiterhin den Anspruch einer optimalen LC-MS Performance. Neben den genannten Aspekten unterscheiden sich die Quantifizierungsstrategien zusätzlich im Zeitpunkt der Probenkombination bzw. des quantitativen Vergleichs, so dass das Potential für Quantifizierungsfehler variiert.

Grundsätzlich muss bei Quantifizierungsansätzen auf eine hohe Reproduzierbarkeit, ausgehend von der Probengewinnung, über die Probenvorbereitung sowie Auftrennung, bis hin zur massenspektrometrischen Analyse geachtet werden, wobei sich der Aufwand für einen jeden Arbeitsschritt je nach Quantifizierungsstrategie (metabolisch, chemisch, LF) stark unterscheidet.

In den folgenden Kapiteln werden Verfahren beschrieben, die in dieser Arbeit für die MS-basierte quantitative Proteomforschung verwendet wurden.

1.3.1 Chemische Quantifizierungsstrategien

Chemische Quantifizierungsstrategien verwenden stabil-Isotopen markierte Reagenzien unterschiedlicher Isotopen-Komposition, die selektiv funktionelle Gruppen von Proteinen bzw. Peptiden derivatisieren und somit das Multiplexen verschiedener Proben zur relativen und absoluten Quantifizierung zulassen. Die in der Proteomik bekannten Markierungsreagenzien unterscheiden sich hinsichtlich der zu modifizierenden funktionellen Gruppe, der stabil-Isotopen Markierung, der Größe, und des chromatographischen Trennverhaltens. So fokussiert der isotope-coded affinity tag (ICAT) ausschließlich Cysteinreste und reduziert damit die Komplexität um ein zehnfaches, nimmt aber gleichzeitig den Verlust von 90 % der Information eines Proteoms in Kauf [55]. Andere Reagenzien derivatisieren primäre Amine, die im Proteom an freien N-termini und Lysinresten zu finden sind und somit die große Mehrheit der tryptischen Peptide einschließen [56-58]. Der

prominenteste Vertreter der chemischen stabil-Isotopen Markierungen ist der *isobaric Tag for relative and absolute quantitation* (iTRAQ) [57], dessen Verwendung den Vergleich von vier bzw. acht Proben ermöglicht. Dazu werden die Proben zunächst individuell prozessiert und proteolytisch verdaut, anschließend mit den unterschiedlichen iTRAQ Reagenzien markiert, bevor sie dann zu einer Probe vereint werden.

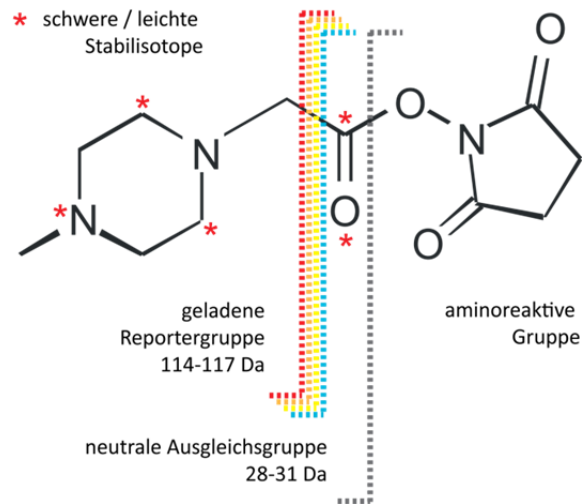


Abbildung 1.5: Aufbau des iTRAQ Labels.

Das iTRAQ Label besteht aus einer aminreaktiven-, einer Massenausgleichs-, sowie einer Reportergruppe; letztere ermöglicht die eigentliche Quantifizierung. Beim 4-plex iTRAQ gibt es unterschiedliche stabil-Isotopen Kompositionen der Reportergruppe mit den Massen 114, 115, 116 und 117 Da und dazu komplementär unterschiedliche Ausgleichsgruppen von 31, 30, 29 und 28 Da, so dass das gesamte Label isobar ist. Nachdem die markierten Peptide, wie unter 1.3. beschrieben, das gleiche Retentionsverhalten während der LC Auftrennung haben, werden die vier Varianten zusammen als ein Signal für das MS/MS isoliert und fragmentiert. Die eigentliche quantitative Information wird bei isobaren Tags über die Intensität der entsprechenden Reporterionen (hier 114, 115, 116 und 117 m/z) im MS/MS gewonnen. Folglich wird, im Gegensatz zu anderen Strategien wie ICPL oder SILAC, die Komplexität der LC-MS-Analyse weder durch die chemische Markierung noch durch die Kombination mehrerer biologischer Proben erhöht.

Nachteilig kann sich bei chemischen Quantifizierungsstrategien der relativ späte Zeitpunkt des Multiplexings, nämlich erst nach dem proteolytischen Verdau, auswirken. Zuvor erfolgt eine Reihe von Probenaufarbeitungsschritten, bei denen sich zufällige sowie systematische Fehler und folglich Varianzen manifestieren können. Außerdem liefert iTRAQ nur quantitative Informationen über fragmentierte Spezies, während bei Verfahren wie SILAC prinzipiell Unterschiede im MS¹ Scan auch ohne Fragmentierung und Identifizierung eines Peptids erkennbar sind. So können potentiell regulierte Peptide in einer nachfolgenden Analyse gezielt fragmentiert und identifiziert werden.

1.3.2 Metabolische Quantifizierungsstrategien

Das metabolische SILAC Labeling nutzt den Einbau von stabil-Isotopen markierten Aminosäuren in der Zellkultur oder in der Aufzucht von Versuchstieren. Es kommen überwiegend die SIL Aminosäuren Lysin, Arginin und/oder Leucin zum Einsatz, die eingebaut während der Proteinbiosynthese zu definierten Massenshifts führen. Nach vollständigem Einbau (nach durchschnittlich 5-8 Zellteilungszyklen) können die gelabelten Proben mit den natürlichen Vergleichsproben und damit früher als im Falle von chemischen Markierungsverfahren vereint und gemeinsam prozessiert werden.

Bei SILAC Ansätzen wird die Probenkomplexität aufgrund der variierenden Massen der gelabelten Aminosäuren und der Kombination mehrerer biologischer Proben erhöht: Durch den Einbau von definierten Massenshifts erscheint ein jedes Peptid entsprechend als Duplett- oder Triplettsignal, was aufgrund limitierter Scangeschwindigkeiten von MS Systemen zu einer verminderten Abdeckung des Proteoms führen kann, da die Anzahl der auf MS¹ Ebene zu quantifizierenden und anschließend zur Fragmentierung isolierten Signale aufgrund des Labelings stark ansteigt.

1.3.3 Label freie Quantifizierung

Die LF Quantifizierung verzichtet auf die beschriebenen Markierungsverfahren und zieht zur Quantifizierung die bereits beschriebenen Verfahren AMT oder spectral counting heran. Die Proben werden hierzu zunächst unabhängig voneinander aufgearbeitet und individuell bestenfalls in Replikaten mittels LC-MS analysiert, um anschließend in relativen Bezug gesetzt zu werden.

(1) *Spectral counting* Verfahren zur LF Quantifizierung wie der emPAI [59] oder NSAF-Index [60] evaluieren das Aufkommen von Peptiden anhand der für ein Protein aufgenommenen MS/MS Spektren und geben mit dem Bezug auf die Zahl der detektierbaren Peptide (emPAI) bzw. auf die jeweilige Proteinelänge (NSAF) einen näherungsweisen Überblick über die quantitative Verteilung eines Proteingemisches wider.

Eine Herausforderung ist hierbei die hohe Anfälligkeit für induzierte Varianzen, die sich potentiell im kompletten Workflow, von der Probegewinnung bis zur abschließenden LC-MS-Analyse in systematischen wie zufälligen Fehlern widerspiegeln können. Daher muss bei der LF Analyse höchstes Augenmerk auf die Reproduzierbarkeit während der Probenaufarbeitung sowie der LC-MS-Analyse gelegt werden. Essentiell ist dabei die

Stabilität des LC-MS-Systems, welches v.a. reproduzierbare Retentionszeiten, eine stabile Ionisierungseffizienz und eine hohe Massengenauigkeit liefern muss, um die oben beschriebenen Features eindeutig zuzuordnen, zu quantifizieren und darüber hinaus Interferenzen mit koeluiierenden Analyten zu vermeiden bzw. zu identifizieren.

Für eine präzise Quantifizierung und Validierung einzelner Komponenten bietet sich ergänzend zur LF Analytik die SRM Analytik an.

Die Kosten der verschiedenen Quantifizierungsmethoden sind beim Einsatz von stabil-Isotopen Reagenzien, auch in Abhängigkeit der zu markierenden Probenmenge, höher. Jedoch erfordert die Anwendung der LF Analytik ein hochreproduzierbares LC-MS System, welches häufig mit kostspieligen instrumentellen Aufwendungen verbunden ist. Durch die Notwendigkeit von Replikaten für eine jede einzelne zu vergleichende Probe ist der Messaufwand wesentlich höher.

1.4 Qualitätskontrolle

Qualitätskontrollen (QC) sind eine grundlegende Voraussetzung für die Validierung von analytischen Daten. Insbesondere in der Proteomanalyse, die sich auf globaler Ebene mit hochkomplexen Protein- bzw. Peptidgemischen beschäftigt, kann die QC zur Reproduzierbarkeit und Validität von analytischen Ergebnissen beitragen.

Für eine umfassende Aussage sollte die QC einen jeden Arbeitsschritt des Workflows ins Auge fassen und neben der Probenaufarbeitung, der LC-MS Performance und Messung auch die Datenauswertung mit robusten Methoden kontrollieren. Die QC trägt damit zur Qualitätssicherung bei. Bislang werden in der Proteomanalyse in vielen Laboren nur wenige optimierte QC Elemente routinemäßig eingesetzt.

1.4.1 Probenaufarbeitung

Die Probenaufarbeitung kann sich je nach analytischem Fokus vielseitig gestalten und bringt damit eine große Schar an Fehlermöglichkeiten mit, die sich sowohl auf qualitative aber vor allem auf quantitative Ergebnisse niederschlagen. Daher wird in der Proteomanalyse versucht Probenaufarbeitungsmethoden zu etablieren, die schnell und effizient sind, möglichst wenig Prozessierungsschritte umfassen, und weiterhin kompatibel zu nachfolgenden Analyseschritten sind. So eignet sich die Präzipitation, je nach verwendetem Fällungsreagenz, vor allem für Proteine mit hohem oder niedrigem Molekulargewicht [61] und

entfernt damit einen Teil der Proteine für die weitere Analyse. In der Literatur sind weiterhin durch die Acetonpräzipitation eingeführte Proteinmodifikationen bekannt [62]. Die Festphasenextraktion (engl. solid phase extraction, SPE) wird häufig eingesetzt zur Aufkonzentrierung und –trennung von biologischen Proben, aber auch zur Abreicherung der Matrix bzw. zur Umpufferung und kann, je nach Charakteristika / Kapazität des verwendeten Sorbents, zum Verlust von Proben führen.

Ein vor allem für die quantitative Proteomanalyse essentieller aber häufig vernachlässigter Schritt der Probenaufarbeitung ist der proteolytische Verdau von Proteinen bei dem die später zu analysierenden Peptidspezies generiert werden. In der quantitativen Proteomanalyse wird oft, ohne eigentliche Kontrolle des Prozesses, von einem vollständigen Verdau ausgegangen, bei welchem z.B. Trypsin nach einem jedem Lys oder Arg schneidet [63], es sei denn es folgt C-terminal die AS Prolin ([R/K].[[^]P]) [64]. Autoproteolyse, ungeeignete Enzymmengen oder Pufferbedingungen sowie besonders stabile Sekundärstrukturen können die Effizienz von Trypsin reduzieren und so zu überlesenen Schnittstellen führen. Weiterhin kann für Trypsin häufig eine chymotryptische Nebenaktivität beobachtet werden, welche durch Autoproteolyse aber auch Enzymverunreinigung hervorgerufen werden kann [65]. Da eine verminderte Verdauereffizienz negative Auswirkungen auf die qualitative und die quantitative Proteomanalyse hat, sollte idealerweise der Verdau für eine jede Probe individuell optimiert und die Effizienz sowie die Reproduzierbarkeit dediziert kontrolliert werden.

1.4.2 LC-MS

Insbesondere für die quantitative Proteomanalyse ist die tägliche Qualitätskontrolle der LC-MS Instrumente auf verschiedenen Ebenen - der LC, der MS sowie der Kombination der beiden - essentiell. Vor allem für die LF Quantifizierung, bei welcher u.a. Peakflächen identischer Features in unabhängigen LC-MS-Analysen miteinander verglichen werden, ist die Kontrolle hinsichtlich Stabilität und Reproduzierbarkeit grundlegend für die Qualität der quantitativen Daten, die häufig über viele Tage akquiriert werden.

In der Literatur sind Ansätze beschrieben, die Informationen aus Rohdaten zur Kontrolle (a) der Chromatographie, (b) der Ionisierungseffizienz, (c) des MS auf MS¹ oder MS² Ebene, (d) der Probenfragmentierung sowie (e) Peptididentifizierung nutzen [66-68]. Wesentliche Kriterien in LC-MS Analysen sind die Stabilität (1) der Retentionszeit, (2) der Signalintensität

bzw. der Peakflächen, (3) der Peakhalbwertsbreiten, (4) der Peakkapazität und (5) der Massengenauigkeit.

Meist werden zur Kontrolle von LC-MS Systemen proteolytische Verdauung von Standardproteinen wie Albumin oder Cytochrom C eingesetzt [69], ein Nachteil ist jedoch, dass proteolytische Verdauung äquimolare Peptidgemische erzeugen, welche auf die Charakteristika des jeweiligen Proteins beschränkt sind. Ein wesentlich vielseitigerer Ansatz ist die Verwendung von klar definierten Gemischen aus synthetischen Peptiden, da diese hinsichtlich der Aminosäurekomposition, dem Auftreten posttranslationaler Modifikationen, der Hydrophobizität und vor allem hinsichtlich des dynamischen Bereichs (nicht zwingend äquimolar wie beim proteolytischen Verdauung) eine Vielzahl von Gestaltungsmöglichkeiten bieten. Darüber hinaus können nicht reproduzierbare Verdauung von Proteinen zu unterschiedlichen Peptidmustern und damit zu Varianzen in der Langzeitkontrolle führen.

1.4.3 Datenanalyse

Aktuelle Massenspektrometer erlauben mit der Akquisition einer Vielzahl von MS/MS Spektren die detailreiche Darstellung von komplexen Proteomen mit hoher Sequenzabdeckung. Obwohl mit der Einführung von hochauflösenden Analysatoren wie der Orbitrap akkurate Massen mit wenigen ppm (engl. parts per million) Ungenauigkeit gewonnen werden können, bergen probabilistische Datenbanksuchen das Risiko einer falsch positiven Zuordnung von Spektren zu Peptidsequenzen (engl. Peptide to spectrum match, PSM). Die manuelle Evaluierung der großen und immer größer werdenden Datensätze auf ihre Richtigkeit ist nicht mehr durchführbar, sodass bioinformatische Werkzeuge für die Auswertung verwendet werden, welche eine statistische Einschätzung des Auftretens falsch positiver Peptididentifikationen liefern.

Hierzu wurde von Elias et al. die oft und auch in Variationen angewandte Falsch-Positiv-Rate (engl. false discovery rate, FDR) entwickelt, welche MS/MS Spektren sowohl gegen richtig-positive „target“ Sequenzen als auch gegen sogenannte falsch-positive „decoy“ Sequenzen in einer target/decoy Datenbank abgleicht, welche neben den originären Proteinsequenzen auch entsprechende rückwärtsgelesene bzw. randomisierte Proteinsequenzen aufweisen [70]. Hierbei wird für definierte Such- und Filterkriterien die Anzahl der decoy-Hits als Maß für das Auftreten zufälliger Hits in der target-Datenbank verwendet, sodass der Anteil an falsch-positiven Identifikationen abgeschätzt werden kann. Die FDR wird in der Proteomik

meist als Kriterium für die Datenauswertung herangezogen, allerdings nicht immer korrekt angewandt.

1.5 Modellproteome

Für die Analytik von globalen Proteomstudien wurden in der vorliegenden Arbeit Modellproteome unterschiedlicher Komplexität herangezogen, welche im folgenden Abschnitt kurz skizziert werden.

1.5.1 Universal Proteomics Standard

Beim Universal Proteomics Standard (UPS) [71] handelt es sich um eine kommerzielle, äquimolare Mischung von 48 humanen Proteinen (siehe Anhang 12.3). Aufgrund der bekannten Bestandteile und Konzentrationen eignet sich der niedrig komplexe UPS für die Entwicklung, Optimierung und Validierung von Methoden in der Proteomanalyse. Dem routinemäßigen Einsatz von UPS in diversen Arbeitsschritten der Proteomanalyse stehen allerdings die hohen Kosten der rekombinanten Herstellung sowie seine limitierte Komplexität gegenüber.

1.5.2 Complex Proteomics Standard

Der Complex Proteomics Standard ist ein Extrakt von lyophilisierten Proteinen des thermophilen Archaeobakteriums *Pyrococcus furiosus* (PFU). Mit etwa 2000 Proteinen weist PFU verglichen zu UPS eine höhere Komplexität auf. PFU wurde wie auch UPS innerhalb dieser Arbeit zur Entwicklung bioinformatischer und experimenteller QC-Verfahren sowie deren Validierung herangezogen.

1.5.3 Mitochondrien

Mitochondrien sind essentielle Organellen in Eukaryonten, deren Hauptaufgabe die Bereitstellung von Energie in Form von Adenosyl-Triphosphat (ATP) ist. 1856 von Kölliker zum ersten Mal beschrieben [72], wurden sie in den folgenden Jahren von Altmann hinsichtlich ihrer Struktur und Funktion näher charakterisiert [73], um schließlich im Jahr 1898 als Mitochondrium benannt zu werden [74].

Nach der Endosymbiontenhypothese liegt der Ursprung der Mitochondrien in der Endocytose von aeroben Prokaryonten, den Archaeobakterien, in Eukaryonten [75]. Die Hypothese stützt sich neben analogen Funktionen und Proteinstrukturen auch auf die Einteilung von Mitochondrien in vier Subkompartimente: (1) Die äußere Membran (OM, engl. outer membrane) und die (2) innere Membran (engl. inner membrane, IM) teilen das Organell in die weiteren Kompartimente (3) den Intermembranraum (engl. inter membrane space, IMS) und (4) die Matrix, die dem bakteriellen Periplasma und Cytosol entsprechen [76]. Aufgrund der Symbiose mit einem α - Proteobakterium sind Mitochondrien mit einer semi-autonomen, hochkonservierten DNA (engl. mitochondrial DNA, mtDNA) [77] und weiterhin einer Proteinbiosynthesemaschinerie mit prokaryotypischen 70S Ribosomen ausgestattet. Die mtDNA von *Saccharomyces cerevisiae* codiert dabei für nur 20 Proteine, welche der Atmungskette und der oxidativen Phosphorylierung zugeordnet werden können. 95% der mitochondrialen Proteine (*S. cerevisiae*) werden über fünf verschiedene Importwege in Mitochondrien entsprechend ihrer Funktion in den Kompartimenten translokalisiert [78, 79]. Zunächst müssen die im Nukleus kodierten Proteine anhand von Signalsequenzen von Oberflächenrezeptoren des Translokase-Komplexes (engl. translocase of the outer membrane, TOM-Komplex) der Außenmembran erkannt und in das Mitochondrium überführt werden.

(1) Proteine mit einer abspaltbaren N-terminalen Präsequenz werden daraufhin über den TIM23-Komplex (engl. translocase of the inner membrane) über die IM in die Matrix geleitet, wo TIM21 den Einbau der Proteine in die IM organisiert bzw. der PAM-Komplex (engl. presequence translocase associated motor) und nachgeschaltete Proteasen [80] die Lokalisation in der Matrix ermöglichen. (2) Hydrophobe Proteine mit internen, nicht spaltbaren Signalsequenzen passieren nach dem TOM-Transport, vermittelt durch Chaperon

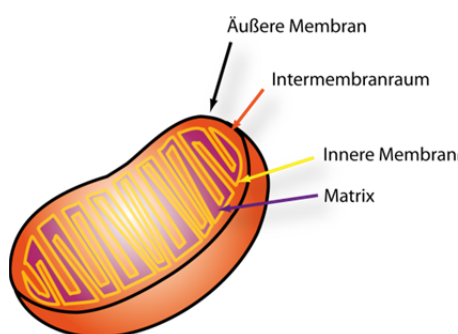


Abbildung 1.6: Schematische Darstellung eines Mitochondriums mit seinen vier Kompartimenten.

TIM Proteine, den polaren IMS und werden über den TIM22-Komplex in die IM inseriert. (3) Der MIA Importweg (engl. mitochondrial intermembrane space assembly) erkennt charakteristische Cysteine-reiche Motive in Proteinen und katalysiert deren Import und die Faltung in den IMS [81]. (4) Alternativ dazu werden Proteine mit einer β -Faltblatt Struktur durch den SAM-Komplex (engl. sorting and assembly

machinery) erkannt und ebenso wie in Bakterien ausschließlich in die OM integriert. (5) Weiterhin sind in der OM Proteine mit α -helicalen Strukturen zu finden, deren Biogenese noch nicht vollständig geklärt, jedoch auf die Aktivität des MIM-Komplexes (engl. mitochondrial import) zurückzuführen ist. Kleinen Proteinen mit einer Größe von max. 6 kDa ist der Zutritt über freie Diffusion über das porengebildende Transmembranprotein Porin in der OM ermöglicht [82].

Zur Hauptaufgabe der Mitochondrien gehört die aerobe Reoxidation von wasserstoffübertragenden Coenzymen unter der Bildung von Energie in Form von Adenosintriphosphat (ATP). Darüber hinaus spielen Mitochondrien in höheren Organismen eine zentrale Rolle im programmierten Zelltod, der Apoptose [83]. So werden durch die Freisetzung von Cytochrom C aus dem IMS in das Cytosol die Apoptose einleitende Caspasen aktiviert und die Integrität der OM aufgehoben.

Mutationen der mtDNA sind häufig letal oder werden mit neurodegenerativen und myopathischen Erkrankungen sowie Morbus Parkinson, Morbus Alzheimer und Krebs in Verbindung gebracht [84]. Fehlfunktionen im Proteinimport führen aufgrund der verminderten Energiegewinnung der Zelle u.a. zu Entwicklungsstörungen, Muskelerkrankungen, epileptischen Anfällen und zu weiteren klinischen Symptomen.

1.5.4 Thrombozyten

Thrombozyten sind cytoplasmatische Abschnürungen ihrer Vorläuferzellen, der Megakaryozyten, und haben einen Durchmesser von 2-4 μm . Im menschlichen Blut treten sie in einer durchschnittlichen Konzentration von $2-3 \times 10^{11} / \text{L}$ auf. Nach ihrer Abschnürung von den Megakaryozyten im Knochenmark zirkulieren sie für 7-10 Tage in der Blutlaufbahn, um anschließend in Leber und Milz abgebaut zu werden [85].

Die physiologische Aufgabe von Thrombozyten ist der Verschluss von vaskulären Läsionen, die sogenannte Hämostase. Zu diesem Zweck zirkulieren Thrombozyten in der Blutlaufbahn und interagieren im Falle einer Läsion u.a. über die Integrine GPIa-IIa und GPIV mit Kollagen sowie über weitere membranständige Proteine mit subendothelialen Matrixproteinen wie z.B. dem von-Willebrand-Faktor oder den Lamininen [86]. Dieser Prozess, Adhäsion genannt, leitet in die dreiphasige Thrombusbildung ein. Bei der Adhäsion kommt es durch nachfolgende intrazelluläre Signaltransduktion zur Aktivierung des Thrombozyten und damit zur Freisetzung von vesikulären Botenstoffen wie ADP oder Thromboxan A_2 . Dabei verändert

der Thrombozyt seine Struktur von der ruhenden diskoiden Form durch Ausprägung von Filopodien zur aktivierten Form. Die Botenstoffe führen zur Rekrutierung von weiteren Blutplättchen, die aggregieren, über das Plasmaprotein Fibrinogen quervernetzt werden und abschließend einen stabilen wundverschließenden Thrombus ausbilden.

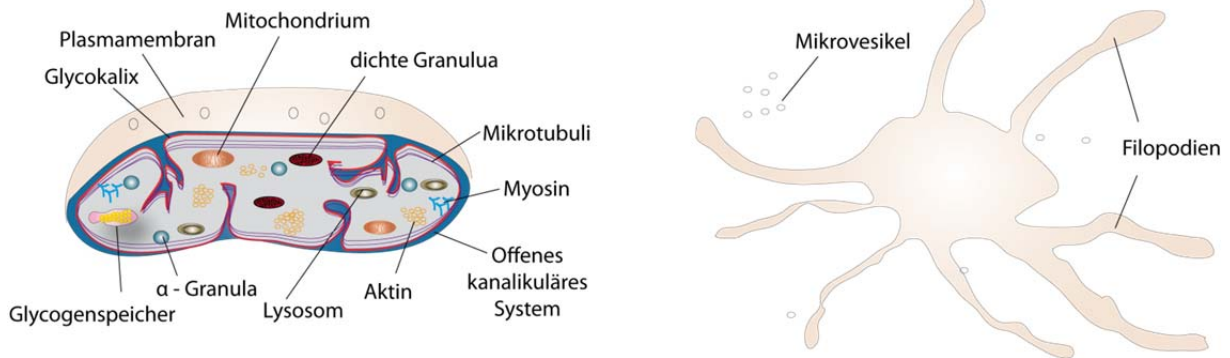


Abbildung 1.7: Schematische Darstellung eines Thrombozyten. Im ruhenden Zustand weist dieser eine diskoidale Form auf, während im aktivierten Zustand Filopodien ausgeprägt werden.

Thrombozyten spielen jedoch auch in akuten Stadien von pathophysiologischen Prozessen der chronischen Arteriosklerose eine Rolle, welche durch arteriosklerotische Plaquerupturen zu Gefäßverschlüssen und Infarkten führen können [87]. Aufgrund ihres Beitrags in der Entstehung und Progression von kardiovaskulären Erkrankungen, der Haupttodesursache weltweit (WHO 2011), sind Thrombozyten für die medizinische Forschung von größtem Interesse.

Ihr anukleärer Charakter befähigt sie zudem nur bedingt zu *de novo* Proteinsynthese. Einige globale Proteomstudien haben sich bereits mit dem thrombozytären Proteom beschäftigt: Noch im Laufe dieser Arbeit wurde die Anzahl der thrombozytären Proteine in der Literatur mit 1507 Proteinen beschrieben [88]. Basierend auf den Erkenntnissen einer Proteomstudie der humanen Thrombozytenplasmamembran, bei welcher 1282 Proteine identifiziert wurden [89], geht man jedoch von einem weitaus größeren thrombozytären Proteom aus – was durch die vorliegende Arbeit bestätigt werden konnte.

1.6 HeLa

HeLa Zellen sind die ersten menschlichen Zellen, aus denen eine permanente Zelllinie hergestellt wurde und werden häufig in der Proteomforschung als Modellsystem verwendet. Sie entstammen einem Gebärmutterhalskrebs und es sind in etwa 10000 Proteine quantitativ in HeLa Zellen nachgewiesen worden [7]. Im Vergleich zu humanen

Thrombozyten haben HeLa Zellen einen in etwa 10x größeren Durchmesser und enthalten einen Zellkern.

2. Zielsetzung

Biologische Netzwerke sind hochdynamisch und reagieren auf intra- bzw. interzelluläre Signale, die zu einer Veränderung zellulärer Prozesse führen können. Einfluss auf die Funktionalität nehmen post-translationale Modifikationen (PTM) oder die Regulation von Proteinmengen mittels *de novo* Synthese, Proteinimport, Internalisierung oder Degradation von Proteinen. Für ein verbessertes Verständnis von biochemischen Mechanismen, wie z.B. der Aktivierung von Thrombozyten, ist es daher enorm hilfreich, das vorhandene Proteinrepertoire/Proteom zu erfassen.

Im Rahmen dieser Arbeit sollten grundlegende Verfahren und Strategien erarbeitet und evaluiert werden, die die valide, reproduzierbare und leistungsfähige quantitative Proteomanalyse mittels hochauflösender LC-MS ermöglichen. Hierbei sollten verschiedene Ebenen des Analyseprozesses wie die Probenaufarbeitung, -auftrennung, -analyse und Datenauswertung systematisch optimiert, dedizierte Verfahren zur Qualitätskontrolle entwickelt und angewendet werden.

Die so erarbeiteten Verfahren sollten hinsichtlich ihrer universalen Einsetzbarkeit für die differentielle quantitative Analyse von Proteomen unterschiedlicher Komplexität – vom Subkompartiment einer Zellorganelle bis hin zum humanen Blutplättchen – evaluiert werden.

So sollten vor allem die folgenden Herausforderungen der modernen quantitativen Proteomanalytik adressiert werden: (1) eine möglichst hohe Abdeckung des Proteoms, (2) eine möglichst hohe Reproduzierbarkeit des gesamten Analyseprozesses, (3) die möglichst zeitsparende Analyse technischer sowie biologischer Replikate um die Signifikanz der Daten zu erhöhen, (4) die möglichst akkurate Quantifizierung bzw. die Abschätzung deren Verlässlichkeit, (5) die Quantifizierung von Spurenkomponenten in einer abundanten Matrix, sowie (6) die effiziente und verlässliche Auswertung der erlangten Daten.

3. Material und Methoden

3.1 Geräte, Reagenzien, Software und Datenbanken

Tabelle 3.1: Übersicht über die verwendeten Geräte.

Geräte	Hersteller	Kontakt
3100 Offgel fractionator	Agilent	Waldbronn, D
Electrophoresis Power Supply	GE Healthcare	USA
Immobiline DryStrips, pH 3-10	GE Healthcare	USA
Inkubator, INB	Memmert	Schwabach, D
LoBind Reaktionsgefäße 0.5 mL, 1.5 mL, 2.0 mL	Eppendorf	Hamburg, D
LTQ - Orbitrap XL	Thermo Fisher Scientific	Bremen, D
LTQ - Velos - Orbitrap	Thermo Fisher Scientific	Bremen, D
Mikrotiterplatten, 96-well PP 0.46 ml	Nunc Thermo Fisher	Roskilde, DK
NUPAGE Precast BisTris Gele	Invitrogen	Karlsruhe, D
Pipetten 0.1-2.5 µl, 0.5-10 µl, 10-100 µl, 20-200 µl, 50-1000 µl	Eppendorf	Hamburg, D
Pipetten, elektronisch, 0.5-10 µl, 10-100 µl, 20-300 µl, 50-1000 µl	Eppendorf	Hamburg, D
qExactive	Thermo Fisher Scientific	Bremen, D
Qstar Elite	Applied Biosystems	Darmstadt, D
Reaktionsgefäße 0.5 mL, 1.5 mL, 2.0 mL	Eppendorf	Hamburg, D
Sephadex NAP-5 Säulen	GE Healthcare	USA
Thermomixer, comfort	Eppendorf	Hamburg, D
TSQ Vantage	Thermo Fisher Scientific	Bremen, D
Ultimate HPLC System	Dionex	Germering, D
Ultimate 3000 HPLC System	Dionex	Germering, D
Ultimate 3000 Rapid Separation Liquid Chromatography (RSLC) systems	Dionex	Germering, D
Vakuumbstation	Agilent	Waldbronn, D
Zentrifuge 5424, 5417R und 5415R	Eppendorf	Hamburg, D
Zentrifuge Beckmann SW 41Ti	Beckmann	Krefeld, D

Tabelle 3.2: Übersicht über die verwendeten Chromatographiematerialien.

Produkt	Hersteller	Kontakt
Acclaim, PepMap (100 µm ID, 2 cm Länge, 100 Å Porengröße, 5 µm Partikeldurchmesser)	Dionex	Germering, D
Acclaim, PepMap, (75 µm ID, 15 cm / 25 cm / 50 cm Länge, 100 Å Porengröße, 2 µm Partikeldurchmesser)	Dionex	Germering, D
HydroRP C18, 2 µm und 4 µm Partikelgröße	Phenomenex	Aschaffenburg, D
OMIX Pipette Tips C18, 10 µl und 100 µl	Agilent	Waldbronn, D
Polysulfoethyl A (200 Å Porengröße, 5 µm Partikeldurchmesser)	PolyLC	Columbia, USA
Polysulfoethyl Aspartamid SCX Säule (1 mm ID, 15 cm Länge, 200 Å Porengröße, 5 µm Partikeldurchmesser)	Chromatographic Technologies	Basel, CH
Sephadex NAP-5 Säulen	GE Healthcare	USA
SPEC C18 AR, Kartuschen, 4mg und 15mg	Agilent	Waldbronn, D
Titandioxid - Beads, TiO ₂	GL Science	Tokio, JP
Titansphere® TiO ₂ Kit	GL Science	Tokio, JP
ZIC® -HILIC stationary phase	SeQuant	Umeå, S
Zorbax 300SB-C18 Säule (2,1mm ID, 15 cm Länge, 200 Å Porengröße, 5 µm Partikeldurchmesser)	Agilent	Waldbronn, D

Tabelle 3.3: Übersicht über die verwendeten Chemikalien.

Substanzen	Hersteller	Kontakt
Aceton	Merck	Darmstadt, D
Acetonitril	Biosolve	Valkenswaard, NL
Ameisensäure	Biosolve	Valkenswaard, NL
Ammoniumacetat	Fluka	Steinheim, D
Ammoniumhydrogencarbonat	Sigma-Aldrich	Steinheim, D
Ammoniumhydroxid	Fluka	Steinheim, D
Ampholyte pH 3-10	Sigma-Aldrich	Steinheim, D
Apyrase	Sigma-Aldrich	Steinheim, D
Bacto™ Peptone	BD BioScience	Heidelberg, D
Bax, Bcl2-associated protein, rekombinantes Protein	Universität de Genève, Martinou Laboratory	Genf, CH
BCA Assay	Pierce Thermo Fisher Scientific	Schwerte, D
BSA	Sigma-Aldrich	Steinheim, D
Calciumchlorid	Merck	Darmstadt, D
CompleteMini, Proteaseinhibitor Cocktail	Roche	Mannheim, D
Complex Proteomics Standard	Agilent	Waldbronn, D
Dinatriumhydrogenphosphat	Merck	Darmstadt, D
Dithionitrobenzoesäure	Sigma-Aldrich	Steinheim, D
Dithiotretiol	Roche	Mannheim, D
EDTA	Merck	Darmstadt, D
EGTA	Merck	Darmstadt, D
Ethanol	Merck	Darmstadt, D
Formaldehyd (37%)	Sigma-Aldrich	Steinheim, D
D-Glucose	Sigma-Aldrich	Steinheim, D
Glycerin	Merck	Darmstadt, D
Guanidinium Hydrochlorid	Merck	Darmstadt, D
Harnstoff	Sigma-Aldrich	Steinheim, D
Hydroxyethylpiperazinylethansulfonsäure (HEPES)	Sigma-Aldrich	Steinheim, D
Iodoacetamid	Sigma-Aldrich	Steinheim, D
iTRAQ - Kit, 4-plex	Applied Biosystems	Darmstadt, D
Kaliumchlorid, SupraPur	Merck	Darmstadt, D
Kaliumdihydrogenphosphat, SupraPur	Merck	Darmstadt, D
Mark 12 Protein Standard	Invitrogen	Karlsruhe, D
Methanol	Biosolve	Valkenswaard, NL
Natriumcarbonat	AppliChem	Darmstadt, D
Natriumcitrat	Merck	Darmstadt, D
Natriumchlorid, SupraPur	Merck	Darmstadt, D
Natriumdihydrogenphosphat	Merck	Darmstadt, D
Natriumhydroxid	Merck	Darmstadt, D
Natriumthiosulfat Pentahydrat	Merck	Darmstadt, D
NUPAGE 3-(N-morpholino)propan sulfonsäure	Invitrogen	Karlsruhe, D
NUPAGE Lithium dodecyl sulfat	Invitrogen	Karlsruhe, D
3-(N-morpholino)propan sulfonsäure (MOPS)	Sigma-Aldrich	Steinheim, D
Phenylmethylsulfonylfluorid (PMSF)	Sigma-Aldrich	Steinheim, D
Phosphorsäure	Merck	Darmstadt, D
Sequencing Grade modified Trypsin	Promega	Madison, USA
Silbernitrat	Sigma-Aldrich	Steinheim, D
Sorbitol	Roth	Karlsruhe, D
Sucrose	Sigma-Aldrich	Steinheim, D
Trichloressigsäure	Roth	Karlsruhe, D
Trifluoressigsäure	Biosolve	Valkenswaard, NL
Trifluoethanol	Sigma-Aldrich	Steinheim, D
Tris(hydroxymethyl)-aminomethan	Merck	Darmstadt, D
Trypsin *alte Charge - vormalis Fluka*	Sigma-Aldrich	Steinheim, D
Trypsin from bovine pancreas T1426 (≥10,000 BAEE units/mg protein)	Sigma-Aldrich	Steinheim, D
Trypsin from bovine pancreas T8658	Sigma-Aldrich	Steinheim, D
Trypsin from porcine pancreas 93614 (≥10000 units/mg)	Sigma-Aldrich	Steinheim, D
Trypsin from porcine pancreas T0303 (13,000-20,000 BAEE units/mg)	Sigma-Aldrich	Steinheim, D
Trypsin Gold	Promega	Madison, USA
Universal Proteomic Standard - UPS1	Sigma-Aldrich	Steinheim, D
Wasserstoffperoxid 30%	Merck	Darmstadt, D

Tabelle 3.4: Übersicht über die verwendeten Programme.

Software	Hersteller	Kontakt
Analyst 2.0 QS	Applied Biosystems	Darmstadt, D
Chromleon 6.8	Dionex	Germering, D
Dbtoolkit 4.1.5	L. Martens	Gent, B
Mascot 2.3.2	http://www.matrixscience.com/	
MS lms 7.5.1	L. Martens	Gent, B
MSOffice 2010	Microsoft	USA
OMSSA 2.1.9	http://pubchem.ncbi.nlm.nih.gov/omssa/browser.htm	
Ontologizer	http://sourceforge.net/projects/ontologizer/	
OpenMS 1.6	http://open-ms.sourceforge.net/	Tübingen
PeptideShaker 0.17.0	http://code.google.com/p/peptide-shaker/	
Progenesis 3.0	Nonlinear Dynamics	Newcastle
MS Convert 2.2.2954	http://sourceforge.net/	
Proteome Discoverer 1.3	Thermo Fisher Scientific	Bremen, D
Reactome Version 4.0	http://www.reactome.org/ReactomeGWT/entrypoint.html	
SearchGUI 1.9	http://code.google.com/p/searchgui/	
Sequest	http://fields.scripps.edu/sequest/	
XLTandem 2010.01.01	http://www.thegpm.org/tandem/	
Xcalibur 2.0	Thermo Fisher Scientific	Bremen, D

Tabelle 3.5: Übersicht über die verwendeten Puffer, inklusive der jeweiligen Pufferzusammensetzung und pH Werte.

Puffer	Pufferzusammensetzung	pH	Kapitel
Lyse	2 M Gu*HCl, 50 mM NaH ₂ PO ₄ bzw. 50 mM Tris	7.8	3.5
MOPS	50 mM MOPS, 50 mM Tris, 0.1 % SDS, 1 mM EDTA	7.7	3.7.1
LDS	Lithiumdodecylsulfat	8.4	3.7.1
SEM	250 mM Sucrose, 10 mM MOPS, 1 mM EDTA	7.2	3.4
Homogenisierungspuffer	0.6 M Sorbitol, 10 mM Tris, 0.2 % BSA (W/V), 1 mM PMSF, 1 mM EDTA	7.4	3.4
Verdaupuffer	50 mM NH ₄ HCO ₃	7.8	3.6
ACD Puffer	12 mM C ₆ H ₈ O ₇ , 15 mM C ₆ H ₇ NaO ₇ , 25 mM D-Glucose	6.5	3.2
HEPES Puffer	150 mM NaCl, 5 mM KCl, 1 mM MgCl ₂ , 10 mM D-Glucose, 10 mM HEPES	7.4	3.2
PBS Puffer	140 mM NaCl, 3 mM KCl, 10 mM Na ₂ HPO ₄ , 2 mM KH ₂ PO ₄	7.2	3.2
CGS Puffer	120 mM NaCl, 12.9 mM C ₆ H ₅ Na ₃ O ₇ , 30 mM D-Glucose	6.5	3.2
Assaypuffer	250 mM Sucrose, 10 mM MOPS, 1 mM KCl	7.2	3.4.1
Rehydratisierungspuffer	IPG Puffer pH 3-10 1:50, 10 % Glycerol	-	3.10.2
HILIC Puffer A	90 % ACN, 10 % 10 mM NH ₄ ⁺ CH ₃ COO ⁻	5	3.10.3
HILIC Puffer B	40 % ACN, 60 % 10 mM NH ₄ ⁺ CH ₃ COO ⁻	5	3.10.3
COFRADIC Puffer A	0.1 % TFA	2.1	3.10.4
COFRADIC Puffer B	0.08 % TFA, 84 % ACN	2.1	3.10.4
TiO ₂ Ladepuffer	80 % ACN, 6 % TFA, gesättigt mit Phthalsäure	0.1	3.10.5
TiO ₂ Waschpuffer	80 % ACN, 0.1% TFA	2.1	3.10.5
TiO ₂ Elutionspuffer 1	0.5 % NH ₄ OH in 250 mM NaHCO ₃	9.1	3.10.5
TiO ₂ Elutionspuffer 2	4.5 % NH ₄ OH, 0.3% H ₃ PO ₄ in 125 mM NaHCO ₃	10.5	3.10.5
TiO ₂ Elutionspuffer 3	1.7 % NH ₄ OH in H ₂ O	12	3.10.5
TiO ₂ Puffer A (GL Science)	Zusammensetzung nicht bekannt	-	3.10.5
TiO ₂ Puffer B (GL Science)	25 % Milchsäure, 75 % GL Puffer A	-	3.10.5
TiO ₂ Elutionspuffer 1 (GL Science)	5% NH ₄ OH	-	3.10.5
TiO ₂ Elutionspuffer 2 (GL Science)	5% Pyrrolidin	-	3.10.5
SCX Puffer A, selbst gepackt	10 mM KH ₂ PO ₄	2.7	3.10.1
SCX Puffer B, selbst gepackt	10 mM KH ₂ PO ₄ , 350 mM KCl, 15% ACN	2.7	3.10.1
SCX Puffer A, kommerzielle Säule	7 mM KH ₂ PO ₄	2.7	3.10.1
SCX Puffer B, kommerzielle Säule	7 mM KH ₂ PO ₄ , 350 mM KCl, 10% ACN	2.7	3.10.1
RP Puffer A	0.1 % FA	2.7	3.11.1
RP Puffer B	0.1 % FA, 84 % ACN	2.7	3.11.1

Tabelle 3.6: Übersicht über die verwendeten Datenbanken mit dem Datum des Downloads und der Herkunft.

Datenbank	Datum	Target / Decoy	Kontakt
N-terminal ragged SGD, 200 AS, tryptisch	05.01.2010	target	--
<i>Pyrococcus furiosus</i> (PFU)	03.08.2011	target	ftp://ftp.ncbi.nih.gov/genbank/genomes/Bacteria
Saccharomyces Genome Database (SGD)	05.01.2010	target	http://www.yeastgenome.org/
UniProt Archae	25.07.2011	target	http://www.uniprot.org/
UniProt Eukaryota	25.07.2011	target	http://www.uniprot.org/
UniProt Eukaryota - PFU - Kombination	--	target / decoy	--
UniProt Homo sapiens	04.11.2010	target	http://www.uniprot.org/
UniProt Homo sapiens	04.11.2010	target / decoy	--
UniProt Homo sapiens - PFU - Kombination	--	target / decoy	--
UniProt Mammalia	25.07.2011	target	http://www.uniprot.org/
UniProt Mammalia - PFU - Kombination	--	target / decoy	--
UniProt Pyrococcus	03.08.2011	target	http://www.uniprot.org/
UniProt Thermococci	03.08.2011	target	http://www.uniprot.org/
UniProt Vertebrata	25.07.2011	target	http://www.uniprot.org/
UniProt Vertebrata - PFU - Kombination	--	target / decoy	--
Universal proteomic standard (UPS)	16.03.2010	target	www.sigma-aldrich.com

3.2 Aufreinigung humaner Thrombozyten aus Frischblut

Für die differentielle Proteomstudie humaner Thrombozyten wurden die Proben von Dr. Stepan Gambaryan (IKBZ Würzburg) nach folgendem Protokoll aufgearbeitet und zur Verfügung gestellt.

200 mL Frischblut von vier Spendern wurde in ACD Puffer aufgenommen und mit 0,01 U/mL Apyrase und 3 mM EGTA ergänzt. Durch 5 minütiges Zentrifugieren bei 330 g wurde thrombozytenreiches Plasma erhalten, welches zur Entfernung von Leukozyten mit PBS Puffer im Verhältnis 1:1 verdünnt und wiederum für 10 min bei 240 g zentrifugiert wurde. Der Überstand wurde für weitere 10 min bei 430 g zentrifugiert und die nun pelletierten Thrombozyten in CGS Puffer gewaschen. Abschließend wurde zur Kontrolle der Leukozytenkontamination eine FACS Analyse durchgeführt. Hierfür wurden die Thrombozytenpellets in HEPES Puffer mit einer Konzentration von $3 \cdot 10^8$ Thrombozyten/mL aufgenommen und mit Hilfe des Leukocount Kits von BD Biosciences getestet.

3.3 Aufreinigung von HeLa Zellen

HeLa Zellen wurden von Hardy Schweigel (Arbeitsgruppe Dr. Peter Nollau, Eppendorf Universitätsklinikum Hamburg) kultiviert und nach Passage 10 geerntet. Dazu wurden die HeLa Zellen zunächst trypsinisiert und bei 13.300 rpm pelletiert.

3.4 Aufreinigung von Mitochondrien aus *Saccharomyces cerevisiae*

Hochaufgereinigte Mitochondrien aus Hefe-Zellkulturen wurden für diverse Fragestellungen von Dr. Frederike-Nora Vögtle (Arbeitsgruppe Prof. Dr. Chris Meisinger, Fakultät für Biochemie, Universität Freiburg) zur Verfügung gestellt.

Dazu wurden Hefen der Zelllinie YPH499 je nach experimenteller Fragestellung in fermentativem (YPD) bzw. respiratorischem (YPG) Medium (siehe Tabelle 3.7) über Nacht bei 30°C unter leichtem Schütteln (130 rpm) kultiviert und anschließend nach Erreichen einer Zelldichte von 1-3 OD bei $\lambda = 578$ nm durch differentielle Zentrifugation geerntet. Weitere Waschschriffe folgten bevor die in Homogenisierungspuffer resuspendierten Hefezellen, unter Anwendung von 15 Stößen, mit einem Glas-Teflon-Potter aufgespalten wurden.

Tabelle 3.7: Zusammensetzung der zur *S.cerevisiae* Kultur verwendeten Medien.

Medienbestandteile	YPG	YPD
Hefeextrakt	1% (w/v)	1% (w/v)
Bactopepton	2% (w/v)	2% (w/v)
Glycerin	3% (w/v)	--
Glucose	--	2% (w/v)
pH	5.0	5.0

In einer weiteren Zentrifugation bei 3.000 g wurden Zelltrümmer und Nuclei entfernt, um die vorgereinigten Mitochondrien anschließend bei 12.000 g zu pelletieren. Nach weiteren Waschschriffen in SEM Puffer erfolgte die Konzentrationsbestimmung mittels *Bradford* Assay.

Eine zweite Aufreinigung zur Entfernung von Kontaminationen wie Vakuolen oder endoplasmatischem Retikulum wurde unter Anwendung einer dreistufigen Sucroседichtegradientenzentrifugation (siehe Tabelle 3.8) erreicht.

Tabelle 3.8: Übersicht über den Sucroседichtegradienten.

Sucrose (w/v)	Volumen (mL)
60%	1.5
32%	4
23%	1.5
15%	1.5

Durch Zentrifugation von 50 mg vorgereinigter Mitochondrien für 1,5 h bei 134.000 g wurde eine hochreine Mitochondrienfraktion zwischen 32 % und 60 % Sucrose gewonnen, die anschließend mit SEM Puffer im Verhältnis 1:2 verdünnt und durch eine weitere Zentrifugation über 30 min bei 16.000 g pelletiert wurden.

3.4.1 BAX induzierte Freisetzung von IMS Proteinen

Für die SILAC Analyse wurden Hefen der Linie YPH499 auf Minimalmedium (6,7 % (w/v) Hefe-Stickstoff-Basis ohne AS, 2 % (w/v) Glucose, 0,77 % (w/v) Complete Supplement Mixture ohne Lys und Arg) kultiviert. Lys und Arg wurden entweder als leichte Komponenten ($^{12}\text{C}_6$) oder als schwere Komponenten mit den Isotopen L-Arginin HCl U- $^{13}\text{C}_6$ und L-Lysin U- $^{13}\text{C}_6$ bzw. L-Arginin HCl U- $^{13}\text{C}_6$ U- $^{15}\text{N}_4$ und L-Lysin 2HCl U- $^{13}\text{C}_6$ U- $^{15}\text{N}_2$ ergänzt.

Für die Analyse des mitochondrialen IMS Proteoms wurden 500 μg isolierte Mitochondrien für 1 h bei 37°C in 200 μL Assay-Puffer mit bzw. ohne 100 nM BAX inkubiert (Bcl-2 associated X protein, Abbildung 3.1). Die Mitochondrien wurden anschließend durch

Zentrifugation für 15 min bei 20.000 g und 4°C pelletiert und die Überstände (engl. supernatant, SN) SN_{BAX} und SN_{Con} zur detaillierten Analyse abgetrennt. Zum Vergleich mit kompletten Mitochondrienlysaten (Tot) wurden 500 µg Mitochondrien in Assay-Puffer aufgenommen und einer 20 s Ultraschallbehandlung mit fünf Wiederholungen im Abstand von 40 s unterzogen. Lösliche Proteine der Tot Fraktion wurden durch Zentrifugation für 45 min bei 100.000 g bei 4°C erhalten.

Die SNs des BAX-Release-Assays wurden mit den löslichen Mitochondrienlysaten in einem Verhältnis von 1:1 (v/v) vereint, so dass für jedes Replikat zwei Proben resultierten: SN_{BAX}/SN_{Con} (SN nach BAX Behandlung vs. SN ohne BAX Behandlung) und SN_{BAX}/Tot (SN nach BAX Behandlung vs. lösliches Mitochondrienlysat). Aus jedem der Labelvorgänge resultierten jeweils zwei Datensätze.

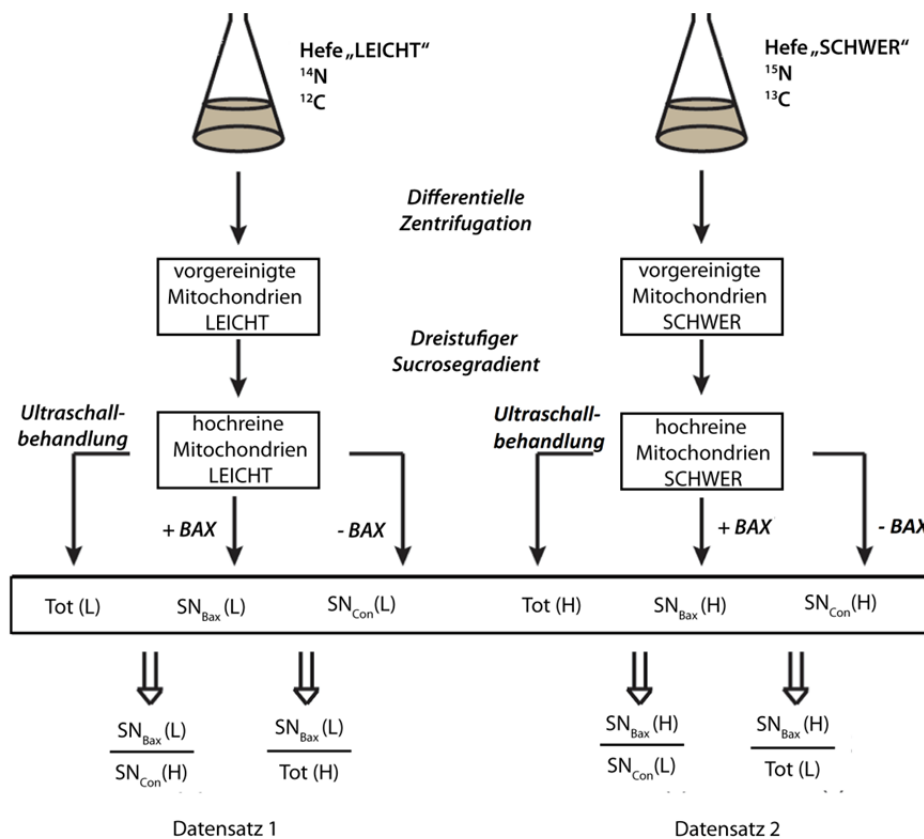


Abbildung 3.1: Übersicht über die Aufreinigung von Mitochondrien aus Hefekulturen, die entweder mit leichten oder schweren Aminosäurekomponenten (SILAC) kultiviert wurden, mit anschließender BAX Behandlung zur Analyse von IMS Proteinen (Abb. nach Vögtle, Burkhart et al., *Mol Cell Proteomics*, 2012).

3.4.2 Proteinprozessierung in Mitochondrien

Die mitochondriale Degradation wurde sowohl für die fermentative sowie für die respiratorische Kulturbedingung untersucht. Hierzu wurden jeweils 500 µg isolierte

Mitochondrien in 1 mL SEM Puffer bei 37°C inkubiert. Zu den Zeitpunkten 0, 6, 24, 72 und 168 h wurde entsprechend ein Aliquot der Mitochondriensuspension entnommen und mittels Zentrifugation pelletiert. Anschließend wurden die pelletierten Mitochondrien, wie unter 3.5 beschrieben, aufgearbeitet.

3.5 Probenaufarbeitung

Für die effiziente Lyse der jeweiligen Modellproteome wurden Probenpellets in Lysepuffer aufgenommen. Je nach Anforderung der nachfolgenden Probenaufarbeitung wurde ein 2M GuHCl Puffer in Kombination mit 50 mM NaH₂PO₄ oder 50 mM Tris, pH 7,8 verwendet. Die Lyse erfolgte durch mehrmaliges Auf- und Abpipettieren, sowie Ultraschallbehandlung, Kühlung und Vortexen.

Unter fortwährender mechanischer Behandlung wurde das Volumen des Lysepuffers bis zur vollständigen Transparenz der Lösung sukzessiv erhöht, wobei angesichts der folgenden Probenaufarbeitung ein minimales Volumen bevorzugt wurde. Die Ultraschallbehandlung wurde nur kurzweilig und mit maximal zehn Wiederholungen durchgeführt.

Daraufhin wurde die Proteinmenge mittels BCA Assay gegen einen Kalibrierstandard ermittelt (3-fach Bestimmung). Anschließend wurde die Probe zur reduktiven Spaltung von Disulfidbrücken für 30 min bei 56°C mit 10 mM DTT inkubiert. Freie Thiole wurden daraufhin durch Zugabe von 20 mM IAA für 30 min bei RT im Dunkeln alkyliert (Carbamidomethylierung).

Für die Probenaufarbeitung der Cystein COFRADIC Analysen (siehe Kapitel 3.10.4.2) wurden die freien Thiole reduziert und durch einstündige Inkubation bei 37°C mit 10 mM DTNB in 100 mM Tris, pH 8,7 maskiert. Zur Entfernung des überschüssigen DTNB Reagenzes wurden die Proben über Sephadex NAP-5-Säulen unter Verwendung von 30 mM Tris, pH 8,7 aufgereinigt.

UPS Proben wurden in 50 µL 0,2 % GuHCl, 5 mM NaH₂PO₄, 50 mM NH₄HCO₃ aufgenommen und, wie oben beschrieben, carbamidomethyliert, während der lyophilisierte Complex Proteomics Standard (PFU), in 50 µL Lysepuffer solubilisiert und anschließend einer TCA-Präzipitation unterzogen wurde. Hierfür wurde die Probe mit 10 % eiskalter TCA-Lösung fünffach verdünnt und für 30 min auf Eis inkubiert. Nach 50 minütiger Zentrifugation (25.000 g) bei 4°C wurde der Überstand abgenommen und das Pellet mit 100 µL Aceton gewaschen. Nach weiterer Zentrifugation für 10 min bei 4°C wurde der Überstand wiederum

abgenommen, das Pellet getrocknet und, wie im Protokoll des Herstellers beschrieben, in 40 μL 50 % TFE, 25 mM NH_4HCO_3 gelöst. Für die Carbamidomethylierung wurden Disulfidbrücken mit 4 mM DTT bei 60° für 60 min reduziert und anschließend mit 15 mM IAA bei RT für 60 min im Dunklen alkylirt. Für den Verdau wurde die finale TFE Konzentration durch Zugabe von 50 mM NH_4HCO_3 auf 5 % eingestellt.

3.6 Enzymatischer Proteinverdau und Optimierung

Für den effizienten Proteinverdau wurde die Konzentration von GuHCl mit Verdauopuffer auf maximal 0,2 mM reduziert. Zur Optimierung des Verdaus wurden variierende Anteile an CaCl_2 (1 mM, 10 mM) sowie organischem Lösungsmitteln (0 %, 5 %, 10 %, 20 % ACN) zur Verdaulösung gegeben und die Verdaueffizienz anhand verschiedener, unter 3.7 beschriebener Verfahren (SDS-PAGE, monolithische HPLC Säulen, SRM, LC-MS, SCX), evaluiert. Zusätzlich wurden ansteigende Inkubationsdauern (1 h, 4 h, 8 h, 12 h) bei 37°C getestet.

3.7 Verfahren zur Verdaukontrolle

Vor der weiteren Verarbeitung der prozessierten Proben wurde die Effizienz des durchgeführten Verdaus kontrolliert, um die Reproduzierbarkeit und Ausbeute von proteolytischen Verdauen für qualitative wie auch quantitative Proteomstudien zu garantieren. Der Verdau nimmt Einfluss auf die Komplexität und folglich auch auf die generelle Sequenzabdeckung, Sensitivität wie auch die Spezifität einer analytischen Messung. Weiterhin werden für quantitative Studien primär voll-tryptische Peptide in Betracht gezogen. Zur Verdaukontrolle wurden die 1-DE Polyacrylamidgelelektrophorese (PAGE), die monolithische HPLC Auftrennung, die Kationenaustauschchromatographie (siehe 3.9.1) sowie spezifische SRM Analysen herangezogen. Diese Verfahren werden in den folgenden Paragraphen beschrieben.

3.7.1 1-DE-PAGE und Proteinvisualisierung

Für die elektrophoretische Auftrennung von Proteinen bzw. Proteinverdaus anhand ihres MW wurden 12 % NuPAGE Precast Bis-Tris Gele verwendet. Hierfür wurden die Proben zunächst in LDS Puffer aufgenommen, für 10 min bei 85°C aufgeköcht und auf die Gele überführt. Ergänzend wurde Mark12 als MW Referenz mit 12 Proteinen im Bereich von 2.5 bis 200 kDa appliziert. Anschließend wurden die Proben in MOPS Puffer aufgetrennt.

Dazu wurde für 20 min eine Spannung von 50 V angelegt, um ein langsames Einlaufen der Proben zu ermöglichen, die daraufhin bei einer Spannung von 250 V über einen Zeitraum von ca. 60 min bei 4°C aufgetrennt wurden. Die Visualisierung der Proteine erfolgte durch Silberfärbung modifiziert nach Blum [90]. Zunächst wurden die Gele für 60 min in Fixierlösung (40 % EtOH, 10 % CH₃COOH) inkubiert. Nach zweimaligem Waschen für 20 min in 30 % EtOH wurden die Gele 20 min in reinem H₂O gewässert. Zur Sensibilisierung wurden die Gele für 1 min in einer 0,02 % Natriumthiosulfat-Lösung inkubiert und anschließend dreimal 20 s mit frischem Wasser gewaschen. Zur Färbung wurden die Gele für 20 min bei 4°C in Dunkelheit mit Färbelösung (0,1 % AgNO₃) abgesättigt. Überschüssige Ionen wurden in einem 1-minütigen Waschschriff mit 200 mL H₂O von der Geloberfläche entfernt. Zur Entwicklung wurden die Gele in basischem Formaldehyd (3 % Natriumcarbonat, 0,05 % Formaldehyd) inkubiert. Da die Silberfärbung keine Endpunktmethode ist, variierte die Entwicklungszeit und wurde auf die Signalintensität und das Signal/Rauschverhältnis abgestimmt. Abschließend wurde die Entwicklung mit 50 mM EDTA-Lösung gestoppt.

3.7.2 Monolithische Auftrennung

Monolithische Trennsäulen zeichnen sich durch ein kontinuierliches Säulenbett mit einer ausschließlich inneren Porosität aus. Es sind sowohl Monolithmaterialien auf Silica- als auch auf Polymerbasis bekannt, wobei letztere vor allem aufgrund ihrer Eigenschaften der schnellen und effizienten Auftrennung mit hoher Auflösung Anwendung bei Biomolekülen finden. Durch die Porenstruktur des Polystyrendivinylbenzenpolymers (Meso- und Macroporen) können sowohl Peptide als auch Proteine aufgetrennt werden, wobei Peptidstrukturen aufgrund niedriger Hydrophobizität und geringerer Größe das Säulenbett schneller durchwandern und zu einem früheren Zeitpunkt eluieren. Zur Verdaukontrolle wurden die Proben mit einer Konzentration von bis zu 0,3 µg/µL zunächst auf einer Vorsäule (PepSwift RP 200 µm x 5 mm, Dionex, Germering) mit 0,1 % TFA aufkonzentriert, um anschließend mit einem Fluss von 2,2 µL/min innerhalb eines binären Gradienten (LM A: 0,1 % TFA; LM B: 84 % ACN 0,1 % TFA) aufgetrennt zu werden (PepSwift RP, 200 µm x 5 cm, Dionex, Germering).

Tabelle 3.9: Gradient der monolithischen Auftrennung.

Monolithische Auftrennung	
Zeit [min]	Anteil Lösungsmittel B [%]
0.00	5
2.00	5
3.00	10
25.00	50
28.00	95
31.00	95
32.00	5
40.00	5

3.7.3 SRM Analysen zur Evaluierung des proteolytischen Verdaus

Zur Evaluierung des tryptischen Verdaus von Thrombozyten wurden außerdem gezielte SRM Analysen durchgeführt, um das Auftreten von vollständig verdauten tryptischen Peptiden sowie deren Analoga mit überlesener Schnittstelle semi-quantitativ zu erfassen.

Um geeignete SRM Übergänge zu erhalten, wurden 4 µg eines tryptischen Thrombozytenverdaus einer 3-stündigen LC-MS Analyse auf einem LTQ-Orbitrap Velos im HCD Modus unterzogen. Die HCD Fragmentierung ist eine Quadrupol-artige Fragmentierung, sodass identifizierte MS/MS Spektren dazu verwendet werden können, SRM Übergänge für eine zielgerichtete Triple Quadrupol Quantifizierung zu definieren. Es wurden Peptidsequenzen ausgewählt, die in diesem komplexen Gemisch sowohl mit als auch ohne überlesene Schnittstelle identifiziert wurden. Für diese prominenten Peptidpaare wurde dann mit Hilfe der fünf intensivsten HCD Fragmentationen pro Peptid eine Retentionszeit-abhängige SRM Methode generiert, welche dazu genutzt wurde, die Effizienz des tryptischen Verdaus semi-quantitativ zu evaluieren. Durch das SRM Messverfahren konnte gewährleistet werden, dass die ausgewählten Peptidpaare mit maximaler Sensitivität spezifisch detektiert und relativ quantifiziert wurden, wobei eine relative Quantifizierung der beiden korrespondierenden Paare untereinander zwar nicht möglich ist (unterschiedliche Ionisierungseffizienz, siehe Kapitel 1.3), aber sehr wohl die relative Quantifizierung eines Peptids zwischen den verschiedenen Verdauproben.

Die Verdaueffizienz wurde für sechs verschiedene Trypsinprodukte unterschiedlicher Hersteller mit jeweils drei Replikaten eines Thrombozytenverdaus untersucht. Hierfür wurden die Proben mittels SRM Analyse mit einer festgesetzten Zyklusdauer von 1,5 s und angepassten Kollisionsenergien von 20 – 35 % zum Scannen von 13 Vorläuferionenpaaren mit jeweils 5 Übergängen (siehe Anhang 12.4) und einer Auflösung von 0,2 im ersten und 0,7 im zweiten Quadrupol analysiert. Die erhaltenen Rohdaten wurden mit dem Programm PinPoint ausgewertet.

3.8 Entsalzung und Aufkonzentrierung von Peptidgemischen mittels C18 SPE

Zur Aufreinigung von salzhaltigen Proben mittels C18 SPE wurden je nach vorliegender Peptidmenge verschiedene Kartuschen mit entsprechend unterschiedlichen Kapazitäten verwendet. Exemplarisch soll der Gebrauch von C18 Kartuschen anhand von SPEC C18 AR 4

mg Tips dargestellt werden, wobei für andere Kartuschen lediglich die verwendeten Puffervolumina variierten.

SPEC C18 AR 4 mg Tips wurden für eine höhere Reproduzierbarkeit in Verbindung mit einer Vakuumstation verwendet. Die Tips wurden zunächst mit 1 ml ACN aktiviert und mit 1 ml 0,1 % TFA äquilibriert. Zur Durchführung der Ionenpaarchromatographie basierend auf hydrophoben Wechselwirkungen wurden die biologischen Proben mit 10 % TFA auf pH 2-3 angesäuert und für eine optimale Ausbeute dreimal auf die C18 Kartusche aufgeladen. Anschließend wurden ungebundene Komponenten durch Waschen mit 1 ml 0,1 % TFA entfernt und das Säulenbett getrocknet. Die Peptide wurden durch Einwirkung von Puffern mit ansteigenden Anteilen an ACN (jeweils, 200 µL, 20 %, 40 % bzw. 60 % ACN in 0,1 % TFA) eluiert. Für die weitere Behandlung der Proben wurden die vereinten Eluate im Vakuum eingeeengt.

Bei nachfolgender SCX Behandlung der Proben wurden der Waschschrift und die Elution mit 0,1 % FA durchgeführt.

3.9 Stabil-Isotopen Markierung mittels iTRAQ

Im Rahmen dieser Arbeit wurden zum einen der komplette Arbeitsablauf der iTRAQ Markierung hinsichtlich Probenaufarbeitung, LC-MS Analysen und Datenauswertung einer systematischen Qualitätskontrolle unterzogen (siehe 3.13.3 und 3.13.4). Weiterhin wurde die Variabilität des humanen Thrombozytenproteoms von vier gesunden Spendern unter Verwendung der chemischen Markierung iTRAQ analysiert.

Für die iTRAQ Markierung wurden die Proben nach dem Verdau entsalzt und zur Trockene eingeeengt, bevor sie in 30 µL 0,5 M TEAB aufgenommen wurden. Die iTRAQ Label wurden, jeweils mit 70 µL EtOH versetzt und vollständig zu den jeweiligen Peptidproben gegeben. Nach einstündiger Inkubation bei RT wurden die vier Proben mit entsprechenden Labels 114-117 in einem frischen LoBind Eppendorf Reaktionsgefäß vereint und anschließend zur Entfernung von EtOH unter Vakuumeinwirkung auf ein Maximum von 50 µL eingeeengt. Durch die darauffolgende Entsatzung auf C18 Basis wurde überschüssiges iTRAQ Reagenz entfernt.

Für die Analyse des Thrombozytenproteoms wurden die Proben abschließend in dem für die nachfolgende Fraktionierung geeigneten Puffer aufgenommen und mittels SCX, IEF, HILIC, Methionin- bzw. Cystein-COFRADIC aufgetrennt. Um einen markierungsbedingten Fehler

auszuschließen wurde für eine jede Fraktionierungsstrategie die Markierungsreihenfolge (Spender 1–4 / iTRAQ Label 114-117) variiert.

3.10 Verwendete Fraktionierungsmethoden

Zur Verringerung der Probenkomplexität wurden verschiedene Fraktionierungsmethoden angewendet, die im Folgenden erläutert werden. Die genutzten Fraktionierungsmethoden fokussierten dabei die Auftrennung der biologischen Proben (SCX, IEF, HILIC) bzw. die selektive Anreicherung von Peptiden definierter Charakteristika (COFRADIC, TiO₂).

3.10.1 SCX

Die SCX Fraktionierung wurde innerhalb dieser Arbeit für diverse Fragestellungen, sowohl zur Verringerung der Probenkomplexität von stabil-Isotopen markierten und markierungsfreien Proben, als auch zur Fraktionierung von Ladungszuständen zur Kontrolle von proteolytischen Verdauen, angewendet. Dabei wurden Fraktionen zeitlich orientiert in Minutenfraktionen bzw., wie in Abbildung 3.1 gezeigt, systematisch oder nach Ladungszuständen gesammelt. Exemplarisch werden die verwendeten Gradienten für die letzten beiden Aspekte vorgestellt.

Für die SCX-Auftrennung wurden die SILAC (leicht/schwer) markierten IMS Proben in 40 µL SCX Puffer A (10 mM KH₂PO₄) aufgenommen und auf einer selbstgepackten SCX Säule (150 mm x 550 µm PolySULFOETHYL A, 200 Å Porenweite, 5 µm Partikelgröße, PolyLC, USA) mittels eines Ultimate 3000 HPLC Systems (Dionex) aufgetrennt. Dazu wurden die Peptide mit einer Flussrate von 6 µL/min in einem binären Gradienten (SCX Puffer A: 10 mM KH₂PO₄, pH 2,7; SCX Puffer B: 10 mM KH₂PO₄, 350 mM KCl, 15% ACN, pH 2.7) separiert. Dieser ging innerhalb von 10 min von 1 bis 12 % SCX Puffer B, wo er für 5 min verblieb, um anschließend innerhalb von 10 min auf 50 % anzusteigen.

Für SN_{BAX}/Tot Proben wurde, wie in Abbildung 3.2 dargestellt, zunächst der Durchfluss und anschließend 5 systematisch fraktionierte Proben gesammelt: (1) 3-5 min, 13-15 min, 23-25 min; (2) 5-7 min, 15-17 min, 25-27 min; (3) 7-9 min, 17-19 min, 27-29 min; (4) 9-11 min, 19-21 min, 29-31 min; (5) 11-13 min, 21-23 min, 31-33 min. Aufgrund der niedrigeren Probenkonzentration und Komplexität wurde für die SN_{BAX}/SN_{Con} Proben der Replikate 1 und 2 zunächst auf eine Fraktionierung verzichtet. Durch die SCX Fraktionierung der SN_{BAX}/SN_{Con}

Proben der Replikate 3 und 4 in je 4 Fraktionen konnten jedoch signifikant mehr SILAC Verhältnisse quantifiziert werden.

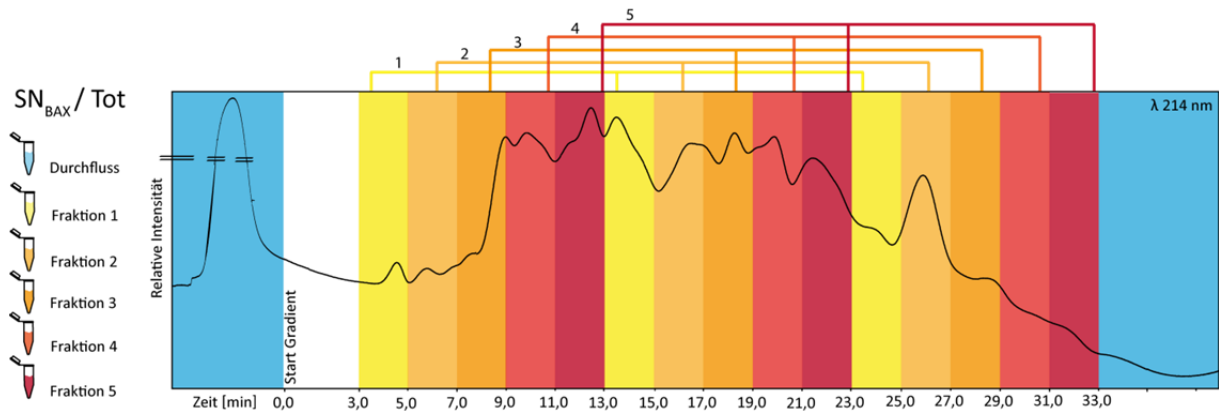


Abbildung 3.2: Übersicht über die systematische Auftrennung der Probe SN_{BAX}/Tot in sechs Fraktionen. Neben der Durchflussfraktion wurden 3 min nach Start des Gradienten Zeiteinheiten mit identischer Farbgebung zu Fraktionen vereint. Die Probe SN_{BAX}/SN_{CON} wurde aufgrund der niedrigeren Probenkomplexität analog in nur vier Fraktionen aufgetrennt.

Die SCX Auftrennungen für weitere Experimente wurden mit einer kommerziell erhältlichen 1 mm x 15 cm PolySULFOETHYL A Säule (200 Å Porenweite, 5 µm Partikelgröße, PolyLC, USA) in Kombination mit einem inerten Ultimate bzw. Ultimate 3000 HPLC System durchgeführt. Zur Kontrolle von tryptischen Verdauen wurden die Proben nach dem Entsalzen und Einengen in 7 mM KH_2PO_4 , pH 2.7 aufgenommen und innerhalb eines binären Gradienten mit steigendem Salzgehalt (SCX Puffer A: 7 mM KH_2PO_4 , pH 2,7; SCX Puffer B: 7 mM KH_2PO_4 , 350 mM KCl und 10 % ACN) bei einer Flussrate von 80 µL/ min über 45 min aufgetrennt. Dabei zeichneten sich deutlich Fraktionen mit Peptiden verschiedener Ladungszustände ab, z.B. +1 (ladungsreduzierte Peptidspezies wie semitryptische Peptide) sowie +2 bis +4 für tryptische Peptide mit und ohne überlesene Schnittstellen, von welchen, wie in Abbildung 4.8 gezeigt, die Ladungszustände +1 (SCX 1) und +2 (SCX 2 und 3) näher untersucht wurden. Die Auftrennung der iTRAQ Proben zur Bestimmung der Proteinvariabilität in Thrombozyten wurde analog durchgeführt, wobei die Fraktionierung in 66 ein-minütigen Fraktionen erfolgte. Währenddessen wurden die Proben zur Evaluierung der Proteinvariabilität in Thrombozyten (siehe 4.7.3), wie für die IMS Proben beschrieben, systematisch gesammelt. Das ACN in den SCX Fraktionen wurden anschließend unter Vakuum entfernt, die Proben mit Omix C18 10 µL Tips entsalzt und unter Vakuum eingengt. Die Proben wurden daraufhin in 0,1 % TFA resuspendiert und mittels RP-LC-MS analysiert.

3.10.2 IEF

Sowohl iTRAQ markierte als auch markierungsfreie Thrombozyten wurden mittels IEF in einem OFFGEL Fractionator in 24 Fraktionen separiert. Hierzu wurden immobilisierte pH-Gradienten (IPG) Streifen (Immobiline DryStrips) mit pH 3-7 und 24 cm Länge verwendet. Die IPG Streifen wurden zunächst für 15 min mit Rehydratisierungspuffer (IPG Puffer pH 3-10 1:50, 10 % Glycerol) rehydriert. Anschließend wurde in jede der 24 Vertiefungen des IEF Kamms 150 µL Probe (in IPG Puffer) gefüllt. Nach Behandlung der beiden Streifenenden mit Kontaktöl (Immobiline DryStrip Cover Fluid) wurde mit Stromstärken- und Leistungsmaxima von 50 µA und 200 mW fokussiert. Die Spannung wurde über 12h von 300 V auf 8000 V erhöht und anschließend bis zum Erreichen von 64 kVh gehalten. Nach erfolgter Fokussierung wurden die Fraktionen mit Omix C18 Tips entsalzt und für die RP-LC-MS Messung vorbereitet.

3.10.3 HILIC

Für die HILIC Auftrennung von iTRAQ markierten thrombozytären Peptiden wurde ein selbstgepacktes zwitterionisches ZIC®-HILIC Säulensystem (SeQuant, Umeå, Sweden) in Kombination mit einem inerten Ultimate HPLC System (Dionex, Germering, Germany) verwendet. Die Proben wurden zunächst in HILIC Puffer A aufgenommen und auf einer ZIC®-HILIC Vorsäule (240 µm x 4 cm, 5 µm Partikeldurchmesser, 200 Å Porengröße) aufkonzentriert. Die chromatographische Auftrennung erfolgte über eine 240 µm x 15 cm Trennsäule (3,5 µm Partikeldurchmesser, 200 Å Porengröße) mittels eines binären Gradienten (HILIC Puffer A: 90 % ACN, 10 % 10 mM NH₄⁺CH₃COO⁻, HILIC Puffer B: 40 % ACN, 60 % 10 mM NH₄⁺CH₃COO⁻) von 1 – 90 % B in 90 min bei einer Flussrate von 10 µL/min. Zwei-minütige Fraktionen wurden in jeweils mit 5 µL 0.1 % TFA beschickte Probengefäße gesammelt, zur Entfernung von ACN unter Vakuum eingengt und mittels LC-MS analysiert.

3.10.4 COFRADIC

Innerhalb dieser Arbeit wurden zwei COFRADIC Ansätze zur selektiven Anreicherung von Methionin (Met)- sowie von Cystein (Cys)-haltigen Peptidsequenzen für die differentielle Analyse der Proteinvariabilität in Thrombozyten eingesetzt. Während die Probenvorbereitung für Met-COFRADIC die beschriebene Carbamidomethylierung vorsah, wurde für Cys-COFRADIC eine alternative Maskierung der Cys, wie unter 3.9.4.2 beschrieben, verfolgt. Die Auftrennungen der ersten wie auch der zweiten Dimension erfolgte über eine

RP Trennsäule (Zorbax 300SB-C18, 2,1 mm x 15 cm, 5 µm Partikelgröße, 300 Å Porendurchmesser) innerhalb eines identischen binären Gradienten (Puffer A: 0,1 % TFA, Puffer B: 0,08 % TFA, 84 % ACN) von 5 % bis 80 % B in 80 min mit einer Flussrate von 80 µL/min. Es wurden 48 einminütige Fraktionen gesammelt und anschließend für die Rechromatographie, wie in Abbildung 3.2 für Met-COFRADIC gezeigt, zu 12 (Met-COFRADIC, [91]) bzw. 16 (Cys-COFRADIC, [92]) Proben kombiniert.

3.10.4.1 Methionin COFRADIC

Nach der ersten Dimension wurden die 12 kombinierten Proben durch Zugabe von 0.5 % H₂O₂ (w/v) und 30 minütiger Inkubation bei 30°C oxidiert (siehe Abbildung 3.3). In der Rechromatographie zeigten die nun zu Sulfoxiden modifizierten methioninhaltigen Peptidsequenzen einen hydrophilen Shift und eluieren, verglichen zur 1. Dimension im Mittel ca. acht Minuten früher. Für die automatisierte Fraktionierung dieser Peptidsequenzen wurde für eine jede kombinierte Rechromatographieprobe ein optimiertes HPLC Programm eingesetzt. Anschließend wurden die gesammelten Fraktionen wiederum zu neun Proben vereint, unter Vakuum eingengt und für die LC-MS Analyse vorbereitet.

3.10.4.2 Cystein – COFRADIC

Wie in 3.4 beschrieben sah die Probenvorbereitung für Cys-COFRADIC eine Variation in der Alkylierung von freien Thiolen durch Umsetzung mit DTNB und anschließender SPE Aufreinigung vor. Nach der RP Auftrennung der ersten Dimension wurde die Reduktion von DTNB zur

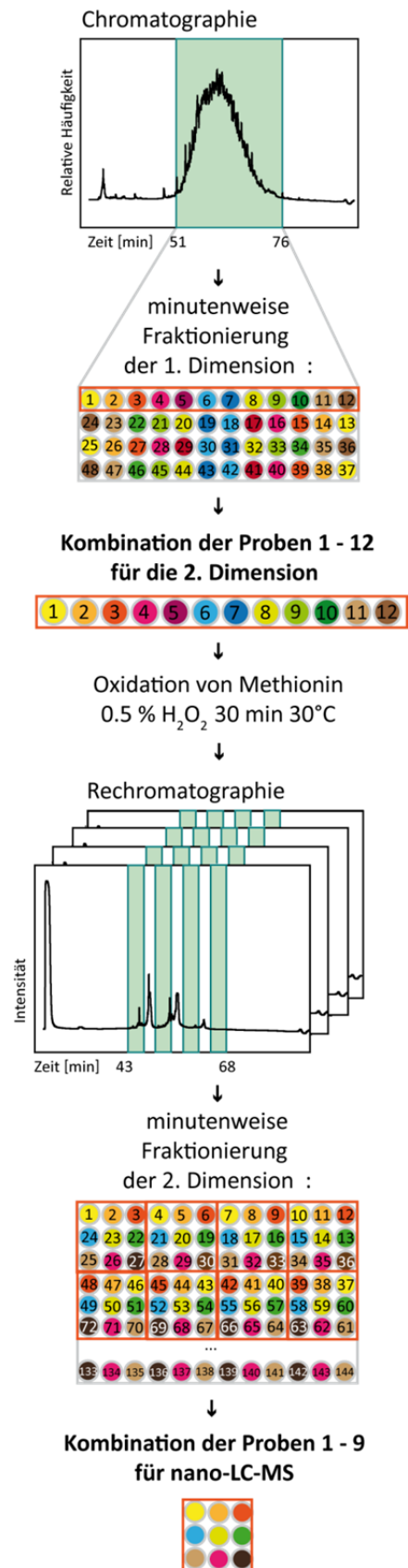


Abbildung 3.3: Übersicht über die selektive Anreicherung von methioninhaltigen Peptiden mittels COFRADIC.

Selektion von Cys-haltigen Peptiden genutzt. Hierfür wurden die 48 Fraktionen der ersten Dimension zu 16 Proben vereint, unter Vakuum eingengt und anschließend in 10 mM Tris, 50 mM DTT für 30 min bei 56°C zur kompetitiven Entfernung von DTNB reduziert. Nach Stoppen der Reaktion durch Zugabe von TFA (pH 2) wurden die 16 kombinierten Proben unter identischen Bedingungen rechromatographiert, wobei der Verlust von DTNB zu einem hydrophilen Shift im chromatographischen Verhalten der Cys-haltigen Peptide führte, die automatisiert fraktioniert wurden. Die gesammelten Fraktionen wurden abschließend zu 12 Proben vereint, eingengt und für die LC-MS Analyse vorbereitet.

3.10.5 Phosphopeptid-Anreicherung mit TiO₂

Um die Abdeckung des Thrombozytenproteoms zu verbessern, wurde eine Phosphopeptid-Anreicherung mittels TiO₂ durchgeführt. Ziel war es, den nicht phosphorylierten Hintergrund größtenteils in der globalen Analyse identifizierter Peptide abzureichern und schließlich über einen zweistufigen TiO₂ Ansatz (niedrige und hohe Spezifität) weitere (Phospho-)Peptide und somit Proteine zu identifizieren.

Jeweils 250 µg der vier Thrombozytenspender wurden zu einer Probe vereint und mittels C18 SPE entsalzt. Anschließend wurden die Proben einer zweistufigen TiO₂ Anreicherung unterzogen, wobei (1) zunächst eine Anreicherung geringer Spezifität und anschließend (2) eine zweite Anreicherung hoher Spezifität für Phosphopeptide durchgeführt wurde. Der Ablauf und die verwendeten Puffer sind in Tabelle 3.5 bzw. 3.10 dargestellt.

Für (1) wurde die Probe in Ladepuffer aufgenommen und nach 30 minütiger Inkubation unter Schütteln auf eine selbstgepackte TiO₂ Säule, bestehend aus einer Pipettenspitze mit Whatman-Papier Fritte und TiO₂ Beads (GL Science, Tokio, Japan) im Proben:Bead-Verhältnis von 1:4 (w:w) geladen. Nach mehrmaligem Waschen mit Ladepuffer und Waschpuffer (80 % ACN 0,1 % TFA) wurden die gebundenen Peptide in einer dreistufigen Elution gesammelt und sofort mit TFA angesäuert, um eine β-Eliminierung von Phosphatgruppen zu vermeiden. Ungebundene Peptide wurden wiederholt mit frischen TiO₂ Beads im Proben:Beadverhältnis von 1:4 inkubiert und dem beschriebenen Protokoll [93] unterzogen. Anschließend wurden die zwei resultierenden Proben mittels Omix C18 10 µL Tips entsalzt und jeweils die Hälfte davon mittels LC-MS analysiert.

50 % der in (1) angereicherten Proben wurden vereint und einer weiteren TiO₂ Anreicherung mittels (2) kommerziell erhältlicher TiO₂ Tips (Titansphere® Phos TiO₂ Kit) unterzogen. Dazu

wurde die vereinte Probe unter Vakuum eingengt und in 100 µL GL Puffer B aufgenommen. Nach zweimaligem Laden der Probe auf einem TiO₂ Tip wurde die Probe alternierend mit den GL Puffern B und A 4-mal gewaschen. Die Elution erfolgte in zwei Schritten. Die vereinten Eluate wurden umgehend mit 5 µL TFA angesäuert, anschließend unter Verwendung von Omix C18 10 µL Tips entsalzt und für die LC-MS Analyse in 15 µL 0,1 % TFA resuspendiert.

Tabelle 3.10: Übersicht über die zweistufige TiO₂-Anreicherung von Phosphopeptiden aus Thrombozytenlysaten mit den jeweiligen Protokollen.

TiO ₂ Anreicherung	(1) TiO ₂ , GL Science, BULK, Probe : Beads = 1 : 4, 2 x		(2) TiO ₂ Tips, GL Science, max. 20 µg Probe pro Tip	
	Anwendung	Puffer	Anwendung	Puffer
Laden	100 µl 30 min bei RT Schütteln	Ladepuffer 80 % ACN 6 % TFA Phtalsäure, gesättigt	100 µl 2 x 10 min bei RT 1.000 x g	GL Puffer B: 25 % Milchsäure 75 % GL Puffer A (0,5 % TFA in ACN)
Waschen	2 x 200 µl	Ladepuffer	20 µl 4 x 2 min bei RT 1.500 x g	GL Puffer B GL Puffer A GL Puffer B
	4 x 200 µl	Waschpuffer 80 % ACN 0,1 % TFA		GL Puffer A
Elution 1	50 µl Inkubation 2 min bei RT	Elutionspuffer 1 0,5 % NH ₄ OH in 250 mM NaHCO ₃ pH 9.1	50 µl 1 x 5 min bei RT 1.000 x g	Elutionspuffer 1 5% NH ₄ OH
Ansäuern des Eluats	15 µl	TFA	5 µl	TFA
Elution 2	50 µl Inkubation 2 min bei RT	Elutionspuffer 2 4,5 % NH ₄ OH 0,3 % H ₃ PO ₄ in 125 mM NaHCO ₃ pH 10.5	50 µl 1 x 5 min bei RT 1.000 x g	Elutionspuffer 2 5 % Pyrrolidin
Elution 3	50 µl Inkubation 2 min bei RT	Elutionspuffer 3 1,7 % NH ₄ OH in H ₂ O pH 12	--	--

3.11 Massenspektrometrische Analyse mittels nano-LC-MS/MS

Die massenspektrometrischen Analysen wurden stets in Kombination mit einem nano-HPLC System durchgeführt. Aufgrund vielseitiger, innovativer Entwicklungen auf Ebene der HPLC sowie auf Ebene der Massenspektrometrie in den letzten Jahren wurden innerhalb dieser Arbeit diverse instrumentelle Kombinationen eingesetzt. Im Folgenden soll exemplarisch die Anwendung grundlegender Einstellungen erläutert werden, während Details dieser in den Ergebnissen besprochen werden sollen.

3.11.1 nano-LC

Für die nano-LC Auftrennungen wurden drei Generationen von Dionex HPLC Systemen - das Ultimate classic, das Ultimate 3000 HPLC und das Ultimate 3000 RSLC HPLC System (Dionex, Germering, Deutschland) - eingesetzt. Generell wurden die Peptide dazu zunächst auf einer RP Vorsäule (Acclaim PepMap, 100 μm x 2 cm, 5 μm Partikelgröße, 100 Å Porendurchmesser, Dionex) in 0,1 % TFA aufkonzentriert [94] und anschließend über eine RP Trennsäule (Acclaim PepMap, 75 μm x 25 cm, 2 μm Partikelgröße, 100 Å Porendurchmesser, Dionex) innerhalb eines binären Gradienten (RP Puffer A: 0,1 % FA, RP Puffer B: 0,1 % FA, 84 % ACN) von 5 % bis 45 % RP Puffer B in variierenden Gradientenlängen von 50 min, 120 min, 200 min oder 240 min mit einer Flussrate von 300 nl/min aufgetrennt.

Ergänzend dazu wurden, im Rahmen der Optimierungen der LC-MS Systeme, zudem noch zahlreiche weitere Variationen der hier angegebenen Gradienten sowie spezifische Waschgradienten, die in 3.11 und dem Ergebnisteil dieser Arbeit näher erläutert werden, evaluiert. Die nano-LC Auftrennungen wurden in Kombination mit den genannten Massenspektrometern durchgeführt.

3.11.2 Verwendete Massenspektrometer

Im Rahmen dieser Arbeit wurden in der MS Analyse mit LTQ-Orbitrap-XL, LTQ Orbitrap Velos und qExactive (alle Thermo Fisher Scientific, Bremen, Deutschland) vorwiegend hochauflösende Massenspektrometer mit einer Orbitrap als MS Analysator/Detektor eingesetzt. Alternativ wurden auch Messungen auf einem hochauflösendem TOF Analysator, dem Qstar Elite, und einem Triple Quadrupole, dem TSQ Vantage, durchgeführt, die im Folgenden im Detail besprochen werden.

3.11.2.1 LTQ Orbitrap Velos

Das LTQ Orbitrap Velos [95] besteht aus zwei Massenanalysatoren: einer linearen Ionenfalle mit zwei Druckkompartimenten (engl. dual-pressure linear iontrap quadrupole, LTQ-Velos) und einem Orbitrap-Analysator. MS Spektren werden grundsätzlich hochauflösend (R bis zu 100,000) mit einer Massengenauigkeit von bis zu < 1 ppm in der Orbitrap akquiriert, während Fragmentationenspektren der intensivsten Ionen entweder mit hoher (Orbitrap) oder niedriger Auflösung (LTQ) gemessen werden. Bei hochauflösenden MS/MS erfolgt in der

Regel die Fragmentierung in dem über die C-trap mit der Orbitrap verbundenen Oktapol (HCD Zelle) mittels höherenergetischen Stößen mit Stickstoff (engl. higher collision induced dissociation, HCD). Die durch HCD generierten Fragmente werden hochauflösend in der Orbitrap ausgelesen. Niedrigaufgelöste MS/MS hingegen werden parallel zum Orbitrap-MS Spektrum in der Ionenfalle akquiriert. Hierzu erfolgt die Fragmentierung im Hochdruckkompartiment der LTQ mittels niedrigerenergetischer Stöße mit Helium (engl. collision induced dissociation, CID). Die durch CID generierten Fragmente werden typischerweise mit geringer Auflösung im Niederdruckkompartiment der LTQ detektiert. Die Kombination der optimierten Ionenfalle und der hochauflösenden Orbitrap ermöglicht eine Scangeschwindigkeit von bis zu 10 Hz. Der Oktapol, auch HCD-Zelle genannt, überwindet den niedrigen Massenausschluss (engl. low mass cut-off), wie ihn Ionenfallen aufweisen, und ermöglicht somit die Analyse von niedrig molekularen Vorläufer- und Fragmentionen. Essentiell ist dies vor allem für die Detektion von Reporterionen von chemischen Labels in der quantitativen Analyse. Während beim CID 1000e niedrig energetische Stöße zu einer Fragmentierung des Moleküls führen, sind es beim HCD einzelne höher energetische Stöße. Dies hat zur Folge, dass sich die Fragmentionenspektren stark unterscheiden können – was vor allem für modifizierte Peptide, wie z.B. Phosphopeptide, gilt.

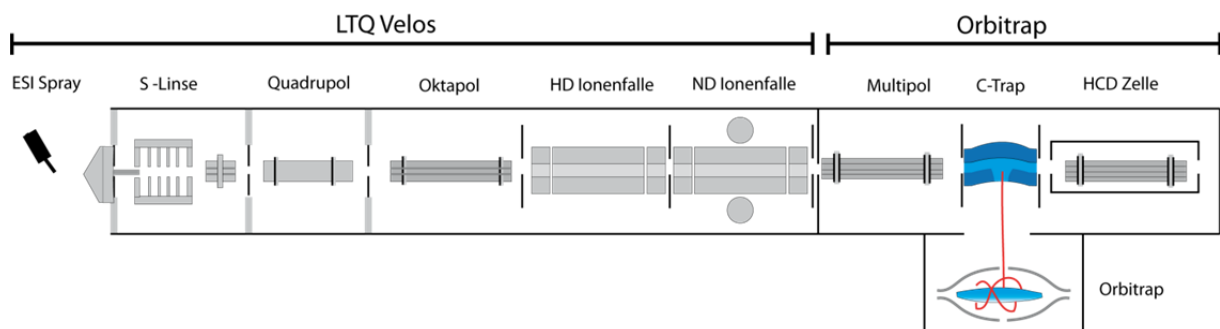


Abbildung 3.3: Schematische Darstellung eines LTQ-Orbitrap-Velos. Während die Fragmentierung durch CID im Hochdruckkompartiment (HD) der Ionenfalle erfolgt und die Fragmente im angrenzenden Niederdruckkompartiment (ND) ausgelesen werden, werden Fragmente der HCD Fragmentierung in der Orbitrap analysiert (Abb. in Anlehnung an Thermo Fisher Scientific).

3.11.2.2 TSQ Vantage

Bei der TSQ Vantage (Thermo Fisher Scientific) handelt es sich um ein sogenanntes *Triple Quadrupole Instrument*, welches aus drei aneinander gekoppelten Quadrupolen aufgebaut ist. Quadrupole finden in verschiedenen MS Systemen, z.B. zur Fokussierung von Ionen

Verwendung. Quadrupole bestehen aus vier parallel angeordneten Elektroden; die gegenüberliegenden Elektroden sind hinsichtlich ihrer Spannung in Reihe geschaltet und werden abwechselnd im Gleichstrom (engl. direct current, DC) bzw. Wechselstrom (engl. radio frequency oder alternate current, AC) betrieben [96]. Benachbarte Elektroden hingegen haben eine entgegengesetzte Polarität und sind um 180° phasenversetzt. Ionen erfahren aufgrund der alternierenden Spannungen sowohl Anziehung als auch Abstoßung und durchqueren die Quadrupole auf oszillierenden Bahnen. Die Veränderung der Spannung bei konstantem AC/DC-Verhältnis ermöglicht die Selektion von definierten m/z -Werten, welche den Quadrupol auf stabilen Bahnen passieren können – weshalb ein Quadrupol häufig als Massenfilter bezeichnet wird. *Triple Quadrupole* Instrumente können für spezielle Scanmodi wie Vorläuferionen-Scan (*Precursor Ion Scan*), Neutralverlust-Scan (*Neutral Loss Scan*) oder zur Quantifizierung mittels Single Reaction Monitoring (SRM, siehe 1.2.3.3) eingesetzt werden.

3.11.2.3 Qstar Elite

Die Qstar Elite zeigt den Aufbau eines Quadrupol-Flugzeitanalysators (engl. quadrupole-time-of-flight, qTOF): Hier dienen zwei Quadrupole dem Fokussieren der Ionen und einer, gefüllt mit Argon als Kollisionszelle. Peptide bzw. deren korrespondierende Fragmente werden schließlich zum

TOF Analysator geleitet und in einem elektrischen Feld mit einer kinetischen Energie von 2,2 eV beschleunigt, um anhand ihrer charakteristischen m/z -Werte in einer

feldfreien Reflektor-Driftstrecke aufgetrennt zu werden. Das m/z Verhältnis kann über die gemessene Flugzeit im Hochvakuum mit Hilfe der bekannten Parameter, der Beschleunigungsenergie und der Driftstrecke, errechnet werden. Eine Auflösung von bis zu

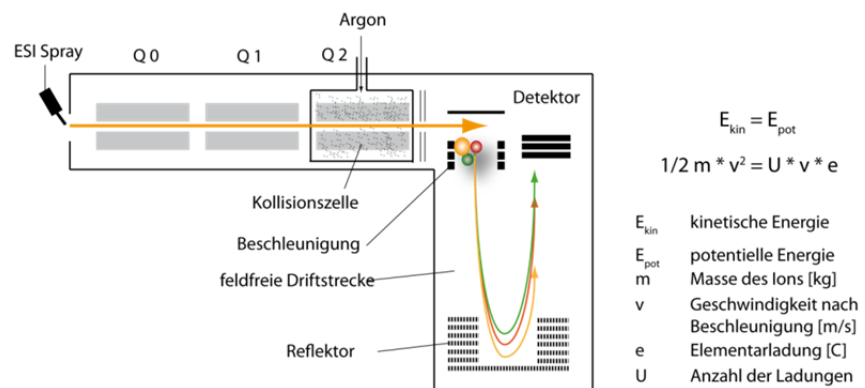


Abbildung 3.4: Schematische Darstellung der Qstar Elite (qTOF).

15.000 wird erzielt und ermöglicht damit die klare Darstellung von Isotopenmustern und Ladungszuständen.

3.11.2.4 QExactive

Das qExactive Massenspektrometer (Thermo Fisher Scientific) setzt sich aus einem Quadrupol und einem Orbitrap Massenanalysator zusammen (siehe Abbildung 3.5). Peptide werden ausschließlich in der HCD Zelle fragmentiert und die Fragmente anschließend mit einer hohen Präzision in der Orbitrap ausgelesen. Neben der hohen Massengenauigkeit weist sich das qExactive vor allem durch eine hohe Scangeschwindigkeit von 12 Hz aus, die weiterhin die Abdeckung eines großen dynamischen Bereichs erlaubt.

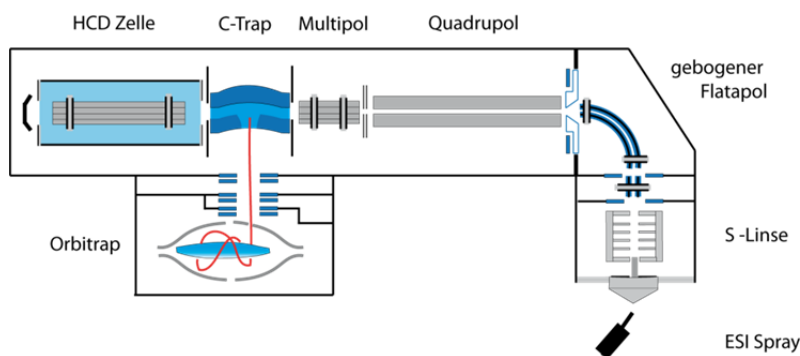


Abbildung 3.5: Schematische Darstellung des qExactive Massenspektrometers (in Anlehnung an Thermo Fisher Scientific).

3.11.3 MS Akquisition

Die massenspektrometrische Akquisition erfolgte in Abhängigkeit des verwendeten Massenanalysators und der zu analysierenden Probe bzw. Fragestellung mit verschiedenen Parametern. Generell unterteilt sich die MS Akquisition einer DDA Anwendung in einen Übersichtsscan auf MS Ebene und der Fragmentierung der intensivsten Signale in folgenden MS/MS Scans. Die Anzahl der in einem Zyklus fragmentierten Ionen wird in der Kurzform für z.B. zehn MS/MS mit "TOP10" angegeben. Von Interesse ist dabei die Art der Fragmentierung, sowie ob die MS/MS Spektren mit hoher (Orbitrap oder TOF) bzw. niedriger (Ionenfalle) Auflösung und Massengenauigkeit detektiert werden.

Tabelle 3.11: Übersicht über die verwendeten MS Methoden zur klassischen DDA.

Massen-spektrometer	Fragmentierung	Anzahl der MS/MS pro Zyklus	Auflösung MS	Auflösung MS/MS	Füllmengen MS [Ionen]	Füllmengen MS/MS [Ionen]	max. Füllzeiten MS [ms]	max. Füllzeiten MS/MS [ms]	normalisierte Kollisionsenergie
Orbitrap	CID	5	~60.000	~2.000	1×10^6	1×10^4	100	100	35
Velos	CID	10 bzw. 20	~60.000	~2.000	1×10^6	1×10^4	100	100	35
Velos	HCD	5	~30.000	~7.500	1×10^6	2×10^5	100	100	45
qExactive	HCD	15	~70.000	~12.500	2×10^5	5×10^4	100	250	27

Tabelle 3.12: Übersicht über die verwendeten MS Methoden zur Analyse von iTRAQ Proben.

Massen-spektrometer	Fragmentierung	Anzahl der MS/MS pro Zyklus	Auflösung MS	Auflösung MS/MS	Füllmengen MS [Ionen]	Füllmengen MS/MS [Ionen]	max. Füllzeiten MS [ms]	max. Füllzeiten MS/MS [ms]	Fragmentintensität* bzw. normalisierte Kollisionsenergie
Ostar Elite	CID	3	~10.000	~10.000	--	--	100	2 s	20*
Orbitrap	CID/HCD	3/3	~60.000	~2.000/~7.500	1 x 10 ⁶	2 x 10 ⁵	100	100	45/50
Velos	HCD	5	~30.000	~7.500	1 x 10 ⁶	2 x 10 ⁵	100	100	45

3.12 Qualitätskontrolle von nano-LC-MS Systemen mittels stabil-Isotopen markierter Peptide

3.12.1 Auswahl der stabil-Isotopen markierten Peptide und Komposition des Gemisches

Die für die Qualitätskontrolle verwendeten stabil-Isotopen markierten Peptide wurden von Thermo Fisher Scientific synthetisiert und lagen in einer definierten Konzentration von 5 pmol/μL vor. Es wurden tryptische Sequenzen (siehe Tabelle 3.13) ausgewählt, die über den kompletten LC-Gradienten verteilt eluieren und sich aufgrund ihrer Stabilität (z.B. keine Met-Reste, keine N-G Kombinationen) [97] für die Evaluierung von nano-LC-MS Parametern eignen. Aufgrund der stabil-Isotopenmarkierung können die Peptide zu jeder Realprobe hinzugegeben werden, ohne mit endogenen Peptiden zu interferieren und somit stets als interner QC-Standard verwendet werden.

Tabelle 3.13: Übersicht über die verwendeten stabil-Isotopen markierten Peptide (* stabil-Isotopen markierte Aminosäuren).

Retentionszeitfenster	relative Konzentration	m/z Verhältnis	M [Da]	Ladung	Peptidsequenz
1	1	416.7181	831.4217	2	LYGDSAAK*
	5	373.1791	744.3437	2	GPDGFSR*
	25	437.7826	873.5507	2	LGPGKPKLK*
	125	523.2731	1044.5317	2	AATAAADFTAK*
2	1	458.7411	915.4677	2	TPQEYLR*
	5	527.2471	1052.4797	2	QYFAETER*
	25	556.2966	1110.5787	2	LPRGPDGFSR*
	125	757.8426	1513.6707	2	ELNEDVSADVEER*
3	1	764.3613	1526.7057	2	ETCVVYTYGYNR*
	5	472.7391	943.4637	2	DFTPELGR*
	25	709.3749	1416.7498	2	VQDL*IEGHLTASQ
	125	416.2516	830.4887	2	TSVFLVR*
4	1	557.8206	1113.6267	2	VLISTDLTSR*
	5	606.3376	1210.6607	2	TNPVSLLQQR*
	25	610.3076	1218.6007	2	EFLVAGGEDFK*
	125	669.8279	1337.6427	2	CSDIISYTFK*P
5	1	536.2931	1070.5717	2	LFILDEADK*
	5	649.3471	1296.6797	2	LDVEEVDLSLR*
	25	710.3546	1418.6947	2	GPLEWQDYIYK*

Die Peptide wurden zur Evaluierung von LC-MS Parametern zu einer QC Mischung pipettiert, in welcher fünf Retentionszeitfenster definiert wurden, die jeweils einen dynamischen Bereich von mehr als zwei Größenordnungen umfassen (relative Verhältnisse 1:5:25:125). Auf diese Weise kann über die gesamte Analysezeit die Sensitivität sowie Performance des LC-MS Systems online kontrolliert werden.

3.12.2 Evaluierung grundlegender LC-MS Parameter

Die Kontrolle der Performance von LC-MS Systemen erfolgte auf den drei Ebenen (1) der LC, (2) der MS und (3) der LC-MS Kombination [98]. Mit Hilfe des internen QC Standards konnten in jeder Probe zahlreiche Parameter evaluiert werden: Peakformen und –halbwidths, RT-Stabilität auf Ebene der HPLC, sowie Signalintensität, Sensitivität, dynamischer Bereich und Massengenauigkeit im MS und MS/MS Modus auf Ebene des MS. Während stabile Retentionszeiten essentiell für die Definition von Features bei der LF Quantifizierung sind, nehmen Peakhalbwidths (engl. full width at half maximum, FWHM) u.a. Einfluss auf die Peakkapazität und Signalintensitäten und damit auf die Sensitivität eines Systems, welche wiederum bestimmend für die Detektionsgrenze (engl. limit of detection, LOD), die Identifikationsgrenze (engl. limit of identification, LOI) und die Quantifizierungsgrenze (engl. limit of quantification, LOQ) ist. Zusätzlich zu diesen Parametern kann mittels einer Datenbanksuche die Massengenauigkeit des Detektors evaluiert werden. Die Eignung des QC Standards zur Evaluierung dieser Parameter wurde intensiv durch die Analyse der Retentionszeit- und der Peakflächenstabilität in 1 µg einer thrombozytären Matrix untersucht. Zunächst wurde der QC Standard zur Kontrolle der LOD und der LOQ in ansteigenden Mengen mit einer minimalen Stoffmenge von $n_{\min} = 0,1 \text{ fmol} - 10 \text{ fmol}$ in einer 120 minütigen LC-MS Analyse auf einem LTQ-Orbitrap XL im CID TOP5 Modus gemessen.

In einer weiteren Messserie zur Kontrolle der LOQ wurde der QC Standard mit einem dynamischen Bereich von 1-125 fmol als stabile Komponente zu 1 µg thrombozytären Hintergrund gegeben, während drei Komponenten in den Mengen von 0,1 fmol, 0,5 fmol, 10 fmol, 100 fmol, 500 fmol bis 1000 fmol ergänzt wurden. Bei den ansteigenden Komponenten handelte es sich um (1) 9 stabil-Isotopen markierte Peptide des TIM22 - Komplexes (*S. cerevisiae*), (2) phosphorylierte und korrespondierende nicht phosphorylierte Peptide der Proteine Tom70, Mdm10 und Zeo1 (*S. cerevisiae*) sowie (3) bovines α-Casein (Uniprot

P02662); diese wurden äquimolar zugegeben (siehe Anhang 12.4). Die Messung der Proben erfolgte in Triplikaten.

Innerhalb von 1h LC-MS Analysen mit ansteigenden Mengen an QC Standard von 0,5 – 62,5 fmol bis zu 10 – 1250 fmol wurden weiterhin Memory Effekte geprüft. Der Verbleib des QC Standards auf den Säulen wurde durch Blankläufe ohne Injektion bzw. variierende Waschgradienten, wie in Abbildung 4.12 dargestellt, in Kombination mit einem Blanklauf untersucht.

3.13 Datenbanksuchen

Die Datenauswertung innerhalb dieser Arbeit wurde mit den Suchalgorithmen Mascot 2.3.2, Sequest, OMSSA Version 2.1.9 und X!Tandem Version 2010.01.01 durchgeführt. Die Suchalgorithmen wurden in diesem Zusammenhang (1) alleinstehend, (2) innerhalb der modulbasierten Software Proteome Discoverer (PD, Thermo Fisher Scientific, Bremen, Deutschland) bzw. (3) mit der Software SearchGUI [99] in Kombination mit PeptideShaker (<http://code.google.com/p/peptide-shaker/>) genutzt.

Tabelle 3.14: Übersicht über die durchgeführten Datenbanksuchen und die jeweils verwendeten Parameter.

Projekt	Standard	IMS		Thrombozytenproteom			Verdaueffizienz		
				Proteinkopien mit NSAF	Proteinvariabilität mit iTRAQ				
MS System	Velos bzw. Orbitrap	Velos	Velos	Velos	Orbitrap	Qstar Elite	Velos bzw. Orbitrap		
Software	PD und Mascot	PD	PeptideShaker	PeptideShaker			PD		
Datenbank	siehe Tabelle 3.4	SGD (target)	SGD (target/decoy)	Uniprot human database (target/decoy)			Uniprot human database (target)		
Suchalgorithmen	Mascot	Mascot	Mascot	Mascot			Mascot		
	OMSSA	Sequest	OMSSA	OMSSA			--		
	X!Tandem	--	X!Tandem	X!Tandem			--		
	Sequest	--		--			--		
Protease	Trypsin	Trypsin		Trypsin			Trypsin	Semistrypsin	Unspezifisch
überlesene Schnittstellen	2	0		2			0	0	0
MS Toleranz	10 ppm	10 ppm		60 ppm	60 ppm	0.5 Da	10 ppm		
MS/MS Toleranz	0.5 Da	0.5 Da (CID)		0.5 Da (CID)	0.5 Da (CID)	0.5 Da (CID)	0.5 Da (CID)		
fixe Modifikationen	CMC	CMC		CMC	CMC ¹	CMC ¹	CMC		
	--	--		--	iTRAQ (N-term, K)		--		
variable Modifikationen	Phosphorylierung (S,T,Y)	¹³ C ₆ Lys und ¹³ C ₆ Arg (beide +6.02 Da)		Phosphorylierung (S,T,Y)			Acetylierung ³ (N-term)		
	Oxidation (M)	bzw.		Oxidation (M)			--		
	Acetylierung	¹³ C ₆ ¹⁵ N ₂ Lys (+8.01 Da) und ¹³ C ₆ ¹⁵ N ₄ Arg (+10.01 Da)		Acetylierung (N-term)			--		
	--	--		Pyroglutamat (N-term, E und Q ²)			--		
Filterkriterien	FDR < 1 %	FDR < 1 %	FDR < 1 %	--			FDR < 1 %		
	--	Peptidrang max. 1	e-value < 10	e-value < 10			Peptidrang max. 1		
	Mascot Score ≥ 20	Mascot Score ≥ 20	--	--			Mascot Score ≥ 20		
	6 ≤ AS ≤ 22	Sequest XCorr +2 ≥ 2.0; +3 ≥ 2.0; +4 ≥ 2.5; +5 ≥ 2.75;	--	8 ≤ AS ≤ 20			6 ≤ AS ≤ 22		

¹ Ausgenommen für Cys-COFRADIC Proben

² Ausgenommen für X!Tandem

³ für die Analyse von SCX 1 und Trypsingelbanden

Für Datenbanksuchen mittels Mascot, OMSSA und X!Tandem mittels SearchGUI wurden die Rohdaten unter Verwendung von MS Convert Version 2.2.2954 (www.sourceforge.net) in

das Mascot generic file (mgf) -Datenformat umgewandelt. Für die Charakterisierung des thrombozytären Proteoms wurden die Rohdaten zunächst mit MS Convert in mzML Daten und anschließend mit OpenMS 1.6 [100] in das mgf-Format konvertiert. Die für SearchGUI genutzten target/decoy Datenbanken wurden mit Hilfe von DBtoolkit [101] erstellt. Die verwendeten Suchparameter und Filterkriterien sind zusammenfassend in Tabelle 3.14 dargestellt.

3.13.1 Datenbanksuchen zur Analyse von Degradationsprozessen

Für die Analyse von Degradationsprozessen mit Hilfe der Software Progenesis wurde die Standarddatenbanksuche zur Identifizierung von semitryptischen, unspezifisch verdauten sowie terminal-abgedauten Peptiden, zusammenfassend im Folgenden „atypische“ Peptide genannt, ergänzt und sukzessive in die Software eingelesen. Die Datenbanksuchen wurden nach dem in Tabelle 3.15 beschriebenen Vorgehen durchgeführt.

Tabelle 3.15: Übersicht über die mehrstufige Datenbank-suche zur Analyse von mitochondrialen Degradationsprozessen.

Projekt	mitochondriale Degradationsstudie				
MS System	Velos bzw. qExactive				
Software	NonLinear Progenesis				
Datenbank	SGD (target)	Promito / ragged 200 N- terminal AS	SGD (target)		
Suchalgorithmen	Mascot				
Suche #	1	2	3	4	5
Protease	Trypsin	Semitrypsin	No Enzyme	Trypsin	None
überlesene Schnittstellen	0	0	0	4	0
MS Toleranz	10 ppm				
MS/MS Toleranz	0.02 Da				
fixe Modifikationen	CMC				
variable Modifikationen	--				
	Acetylierung (N-term)				
	Oxidation (M)				
Filterkriterien	Pyroglutamat (D,E)				
	--				
	$\Delta m < 4$ ppm				
	--				
	Mascot Score ≥ 20				
	$6 \leq AS \leq 22$				

Während für mitochondriale Studien grunds grundsätzlich gegen die SGD Datenbank gesucht wurde, wurde zur Analyse von N-terminalen Proteinenden eine sogenannte „ragged“ Datenbank (siehe Abbildung 3.6) aus den N-terminal abgedauten Proteinenden der 851 bekannten mitochondrialen Proteine in Hefe [35] zusammengestellt. Die ersten 200 Aminosäuren je Protein wurden dazu tryptisch geschnitten und die entstandenen Peptide

des Weiteren N-terminal sukzessive abgedaut, so dass ein Gemisch aus Peptiden mit jeweils um eine AS verkürzten Sequenzlängen vorlag [80].

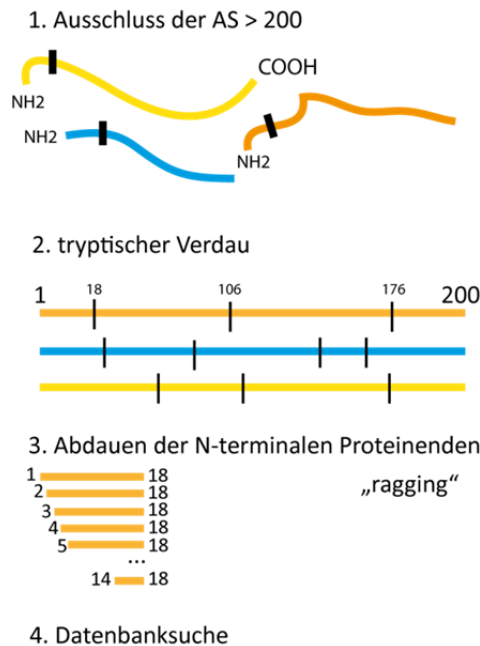


Abbildung 3.6: Darstellung des *in silico* ragging Prozesses von Datenbanken. (1) N-terminale Proteinenden bis zur AS 200 werden zunächst (2) tryptisch verdaut und die entstehenden Peptide anschließend (3) sukzessive N-terminal um eine AS verkürzt.

3.13.2 Optimierung von target/decoy Suchstrategien für verbesserte

Identifikationsraten

Aufgrund der hohen Datenmengen in globalen Proteomstudien sind für die automatisierte Datenanalyse mit Anwendung von diversen Suchalgorithmen Kontrollmechanismen erforderlich, die eine Beurteilung des Aufkommens falsch-positiver Spektrum/Peptid/Proteinidentifikationen (engl. false discovery rate, FDR) ermöglichen. Hierzu werden Datenbanken kreiert, die neben den richtig-positiven „target“ Proteineinträgen Variationen derselben, z.B. rückwärts gelesene oder zufällig variierte Proteinsequenzabfolgen enthalten, die als falsch-positive „decoy“ Einträge bezeichnet werden. Definiert durch den Quotienten der Anzahl der decoy IDs (n_{decoy}) und der Summe der Anzahl an target IDs (n_{target}) und an n_{decoy} , berechnet die FDR den prozentualen Anteil an falsch-positiven IDs in der Gesamtliste aller Identifikationen. Bei gleichbleibender FDR kann diese z.B. als Qualitätskriterium für die Optimierung von Datenbanksuchen fungieren.

Basierend auf jeweils fünf LC-MS Analysen von UPS und Mitochondrien (90 min, LTQ-Orbitrap XL CID TOP 5) wurden Suchalgorithmen (Mascot, X!Tandem, OMSSA) und Variationen der Suchparameter (Vorläuferion- und Fragmentionmassengenauigkeit)

untersucht. Weiterhin wurde die Präzision der Suchalgorithmen in der FDR Evaluierung sowie die für eine valide Aussage notwendige Datensatzgröße charakterisiert [102]. Als Qualitätskriterium wurde die Anzahl der validen Peptididentifikationen bei einer FDR von ≤ 1 % herangezogen.

3.13.3 Verwendung von UPS und *Pyrococcus furiosus* als Referenzproteom für target/decoy Strategien

Zur Kontrolle von target/decoy Strategien eignen sich Proben, deren Zusammensetzung bekannt ist und folglich Peptididentifikationen eindeutig als „echt-positiv“ bzw. „falsch-positiv“ klassifiziert werden können.

UPS ist ein Proteingemisch bestehend aus 48 humanen Proteinen (siehe Anhang 12.2), die durch rekombinante Expression sowie durch Aufreinigung humaner Gewebe gewonnen werden. Die lyophilisierten Proteine wurden für die Analyse in 100 μ L 50 mM NH_4HCO_3 resuspendiert und, wie unter 3.4 beschrieben, carbamidomethyliert und verdaut. Während UPS mit seinen 48 Proteinen nur bedingt als Referenzprobe geeignet ist, bietet der Organismus *Pyrococcus furiosus* (PFU) mit einem Proteom von ca. 2045 Proteinen (www.uniprot.org) eine ausreichende Komplexität zur Analyse der Protein FDR. Weiterhin zeigt PFU eine große evolutionäre Distanz zum Menschen: MS/MS Spektren von PFU können eindeutig dem Organismus zugeordnet werden und eignen sich daher, wie in dieser Arbeit gezeigt, hervorragend als Referenzspektren für die statistische Auswertung humaner Proteomstudien.

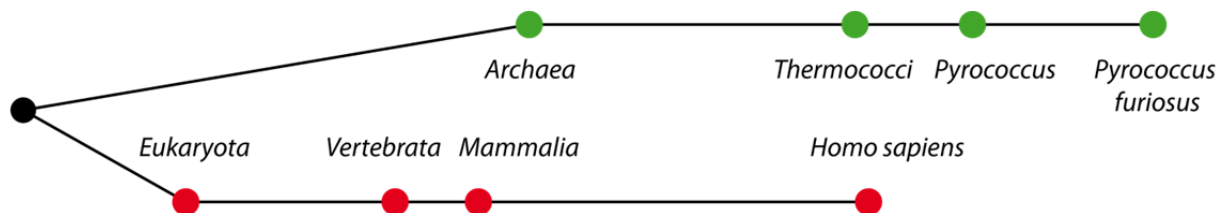


Abbildung 3.7: Aufgrund der stark unterschiedlichen Proteinsequenzen kann PFU für Eukaryoten, Vertebraten, Säuger und den Menschen als Referenzproteom dienen, wohingegen Archaea bereits Proteinsequenzen tragen, die auch in PFU zu finden sind.

3.14 Quantifizierung

3.14.1 SILAC – Proteome Discoverer

Die LC-MS Analysen der SILAC Experimente wurden über 4 h auf einem LTQ-Orbitrap-Velos System im CID TOP10 Modus durchgeführt. Für die Datenauswertung wurde die Software PD genutzt, welche die Identifizierung mit den Suchalgorithmen Mascot und Sequest mit der SILAC Quantifizierung kombiniert und die anschließende Datenprozessierung ermöglicht. Für die Quantifizierung wurden ausschließlich innerhalb der Datenbank einzigartige Peptide zugelassen, wobei bei Fehlen eines Doublettsignals ein maximales SILAC-Verhältnis von 100:1 angenommen wurde.

Zusätzlich zur Datenauswertung mittels PD wurde PeptideShaker zur Analyse von fehlenden SILAC-Verhältnissen unter Verwendung der Suchalgorithmen Mascot, OMSSA und X!Tandem genutzt. Für die hierdurch erhaltenen, ergänzenden Peptididentifikationen wurde die Quantifizierung manuell durchgeführt, da PeptideShaker keine SILAC Ratios berechnet. Die SILAC Verhältnisse wurden dazu, basierend auf monoisotopischen Signalintensitäten in den Rohdaten, manuell mit Xcalibur Qual Browser 2.1. erstellt und sind mit (*) markiert (siehe Anhang 12.6).

Weiterhin wurden die Rohdaten mit der frei erhältlichen Software MaxQuant [103] ausgewertet, welche den weiteren Suchalgorithmus Andromeda [104] nutzt. Diese Auswertung führte jedoch zu keinen zusätzlichen qualitativen wie quantitativen Informationen.

3.14.2 iTRAQ und NSAF – PeptideShaker

Die Datenauswertung der differentiellen Analyse von Thrombozyten erfolgte in enger Zusammenarbeit mit Marc Vaudel. In diesem Zusammenhang wurde das Programm PeptideShaker entwickelt, welches die kombinierte Datenauswertung der drei Suchalgorithmen, Mascot, OMSSA und X!Tandem mit einer optimierten statistischen Auswertung zur Bestimmung der FDR ermöglicht. PeptideShaker wurde sowohl für die Evaluierung der Proteinvariabilität mittels iTRAQ als auch für die auf NSAF basierte Quantifizierung des Thrombozytenproteoms genutzt. Weiterhin wurde die NSAF Quantifizierung zur Analyse von Degradationsprozessen in Mitochondrien angewendet.

Für die Quantifizierung der Proteinvariabilität in Thrombozyten wurden die iTRAQ Daten, nach Labelkorrektur basierend auf den Angaben des Herstellers, auf den Reporter 114

normalisiert. Soweit einem Protein mehr als sechs Peptidverhältnisse zugeordnet werden konnten, wurden die Proteinverhältnisse basierend auf Peptidratios mittels einer statistischen Funktion, dem „redescending M estimator“, bestimmt. Bei weniger als sechs Verhältnissen wurde der Median gebildet. Anschließend wurden die gebildeten Verhältnisse 117/114, 116/114 und 115/114 zur Darstellung von Regulationsereignissen auf dem Median aller evaluierten Ratios normalisiert.

Zur LF Quantifizierung des Thrombozytenproteoms wurden zunächst für alle identifizierten Proteine mittels PeptideShaker NSAF Werte errechnet, welche bezogen auf die Gesamtheit der identifizierten Proteine einer Analyse in etwa den prozentuellen Anteil eines jeden einzelnen Proteins widerspiegeln können. NSAF berücksichtigt für ein Protein die Anzahl an identifizierten Spektren in Relation zur Proteinelänge. Weiterhin wurden die NSAF Werte herangezogen, um für alle im Rahmen dieser Arbeit identifizierten Proteine die Kopienzahl pro Thrombozyt zu schätzen. Hierzu wurde die Literatur durchsucht und eine Referenzliste mit 24 thrombozytären Proteinen, die absolut quantifiziert wurden, erstellt und diese Werte gegen die entsprechenden NSAF Werte aufgetragen, um eine mathematische Beziehung aufzustellen.

3.14.3 PFU als QC Instrument für die iTRAQ Quantifizierung

Für die Quantifizierung von iTRAQ markierten Peptiden müssen deren Vorläuferionen, zur Freisetzung der quantitativen Information in Form von Reporterionen, zunächst fragmentiert werden. Da diese Reporterionen für alle iTRAQ gelabelten Peptide die Signale 114 m/z, 115 m/z, 116 m/z und 117 m/z sind, können „überlappende“ Vorläuferionen, die mit ähnlichen m/z eluieren und somit gemeinsam isoliert und fragmentiert werden, die Quantifizierung interferieren. Im Rahmen dieser Arbeit wurde eine Methode mit PFU als internem Standard entwickelt, die eine quantitative Bestimmung dieser Interferenz ermöglicht (Abbildung 3.8). PFU belegt innerhalb dieses Verfahrens einen Reporterkanal, wobei der interne Standard PFU aufgrund der großen evolutionären Distanz zu Mensch, Maus, Hefe usw. qualitativ eindeutig zugeordnet und weiter für die Evaluierung einer quantitativen FDR (qFDR) genutzt werden kann [105].

Zu diesem Zweck wurden Thrombozytenverdaue und PFU in zwei 4-plex iTRAQ Experimenten analysiert. Im *proof-of-principle* Experiment wurde die Interferenz zunächst rein qualitativ bestimmt. Hierzu wurden die zwei Reporterkanäle 114 und 115 mit

Thrombozytenverdauen und die Kanäle 116 und 117 mit PFU belegt. Nach erfolgreicher, eindeutiger Zuordnung der Signale beider Organismen wurde darauf aufbauend das Ausmaß der Interferenz in einem weiteren Experiment quantitativ erfasst. Dazu wurden 500 µg eines humanen Thrombozytenverdaus mittels IEF, wie unter 3.9.3 beschrieben, aufgetrennt.

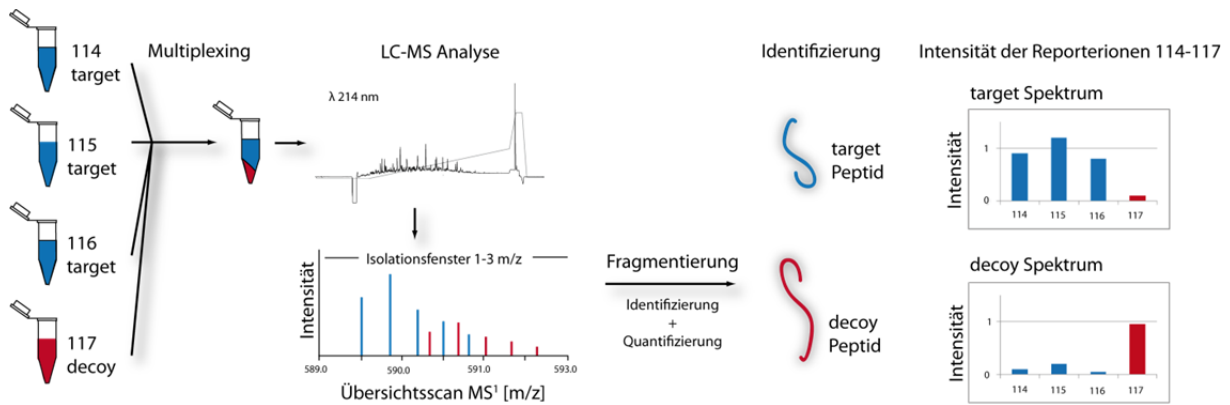


Abbildung 3.8: Schematische Darstellung zur QC der Reporterionen-basierten Quantifizierung mittels eines internen Standards. Ko-eluierende Peptide mit ähnlichen m/z im Übersichtsscan werden gemeinsam für die Fragmentierung isoliert. Nach erfolgter Identifizierung des signalintensiveren target Peptids werden alle Reporterion Signale - auch jenes des decoy Peptids - dem target Peptid (target Spektrum) zugeordnet. Durch Einsatz eines orthogonalen internen Standards wie PFU kann diese Interferenz identifiziert als auch quantifiziert werden.

Anschließend wurde ein Aliquot jeder Fraktion mittels LC-MS für 1 h mit einer LTQ-Orbitrap XL im CID TOP5 Modus analysiert und die Daten mittels PeptideShaker ausgewertet. Daraufhin wurden die Fraktionen mit den geringsten Überlappungen an Peptidentifikationen, nämlich Fraktion 3 (114), 9 (115) und 20 (116) gemeinsam mit PFU (117) als internem Standard, iTRAQ markiert, vereint und per LC-MS mit einer LTQ-Orbitrap Velos im HCD TOP5 Modus analysiert. Mit Hilfe dieser Probe wurden die präzise Isolierung und Fragmentierung der Vorläuferionen (1) bei unterschiedlichen Isolationsfenstern von 1 bzw. 2 m/z, (2) bei variierenden HPLC-Gradientenlängen von 90, 180 und 240 min sowie (3) bei Signalen niedriger Intensität untersucht.

3.14.4 Optimierung und Qualitätskontrolle der iTRAQ - Arbeitsschritte

Innerhalb einer Reihe von qualitätsorientierten Experimenten wurden grundlegende Probenaufarbeitungsschritte, die Quantifizierung mittels iTRAQ und die zugehörige Datenerhebung unter Verwendung der Quantifizierungsstrategien iTRAQ und LF untersucht [106].

Mit Hilfe von Hefe-Mitochondrien wurden verschiedene Parameter des iTRAQ Workflows systematisch quantitativ evaluiert, u.a. bezüglich der Probenaufarbeitung: (I) der Zusatz von divalenten Ca^{2+} Kationen und ACN während des Verdauens, (II) die Variation der Verdauzeit (1 h, 4 h, 8 h und 12 h), (III) die Auswirkung der C18 SPE auf die iTRAQ Markierungseffizienz bei Peptidgemischen.

Die Proben für (I) und (II) wurden, wie in Tabelle 4.1. angegeben, zusammengestellt. Während der Zusatz von Additiven in (I) mittels iTRAQ evaluiert wurde, wurde die Proben zur Variation der Verdauzeiten (II) mangels vollständiger iTRAQ Label 114-117 aufgrund wenig vergleichbarer Proben LF quantifiziert.

Weiterhin wurden (III) Verluste bei der C18 Entsalzung mittels LF Analyse kontrolliert, während die Markierungseffizienz, die stark von Störsubstanzen wie primären Aminen beeinflusst ist, innerhalb eines iTRAQ Experiments durchgeführt wurde. Hierfür wurden zwei Probentypen von jeweils 2 μg mitochondrialen Hintergrund und (III-1) 100 bzw. (III-2) 300 fmol eines äquimolaren Standardproteinverdau bestehend aus ADH (P00330, Hefe), BSA (P02769, bovin), β -casein (P02666, bovin), β -Lactoglobulin (P02754, bovin) und Phosphorylase B (P00489, Lepus) hergestellt. In der experimentellen Durchführung wurden 65 μg Mitochondrien mit (III-1) 6,5 pmol bzw. (III-2) 19,5 pmol des äquimolaren Standardproteinverdau ergänzt und nach bzw. ohne C18 SPE Aufreinigung wie in Tabelle 3.16 markiert.

Tabelle 3.16: Übersicht über die innerhalb der Qualitätskontrolle des iTRAQ Workflows untersuchten Aspekte und die dafür verwendeten Probenkompositionen. Die 4-plex iTRAQ Markierung der verschiedenen Proben ist jeweils farblich markiert.

QC Aspekte	(I) Verdaupuffer / Additiva	(II) Verdauzeit	(III) Markierungseffizienz / nach C18 SPE	(IV) Dynamischer Bereich
Proben- komposition	pro Ansatz 50 μg Mitochondrien Trypsin : Protein = 1: 30	pro Ansatz 50 μg Mitochondrien 1 mM CaCl_2 , 5 % ACN Trypsin : Protein = 1: 30	pro Injektion 2 μg Mitochondrien (III-1) + 100 fmol Standardproteinverdau bzw. (III-2) + 300 fmol Standardproteinverdau	pro Injektion 0,1 μg Mitochondrien + 5 / 25 / 250 / 500 fmol Standardproteinverdau
iTRAQ Markierung / LF Analyse	(A) 114 1 mM CaCl_2 (B) 115 5 % ACN (C) 116 1 mM CaCl_2 , 5 % ACN (D) 117 1 mM CaCl_2 , 10 % ACN	(A) LF 1 h (B) LF 4 h (C) LF 8 h (D) LF 12 h	(A) 114 + 100 fmol vor C18 SPE (B) 115 + 100 fmol nach C18 SPE (C) 116 + 300 fmol vor C18 SPE (D) 117 + 300 fmol nach C18SPE	(A) 114 + 5 fmol (B) 115 + 25 fmol (C) 116 + 250 fmol (D) 117 + 500 fmol

(IV) Zur Analyse des dynamischen Bereichs von iTRAQ Experimenten wurde derselbe äquimolare Proteinverdau in den Mengen (A) 2,5 pmol, (B) 12,5 pmol, (C) 125 pmol bzw. (D)

250 pmol zu 50 µg mitochondrialem Background geben, C18 SPE entsalzt, mit 4-plex iTRAQ markiert und vereint. In der LC-MS Analyse wurden entsprechend 5 fmol (114), 25 fmol (115), 250 fmol (116) bzw. 500 fmol (117) bei einem mitochondrialen Background von 0,1 µg analysiert.

Alle iTRAQ markierten Proben wurden mittels nano-LC über 4h von 3-45 % LM B aufgetrennt und mit einer LTQ-Orbitrap Velos im HCD TOP5 Modus analysiert. In der nachfolgenden Datenevaluierung wurde die Eignung der häufig eingesetzten statistischen Strategien, des (a) Mittelwerts und (b) des Medians sowie einer robusten statistischen Methode (c) dem "redescending M-estimator" systematisch getestet.

3.14.5 Labelfree - Nonlinear Progenesis

Für die LF Analyse von Degradationsprozessen wurden Mitochondrien von fermentativen und respiratorische Hefe Kulturen, YPD und YPG, isoliert und über einen Zeitraum von sieben Tagen in SEM Puffer bei 37°C inkubiert.

Die Isolierung der Mitochondrien unterbindet den Import von Proteinen und ermöglicht so die Analyse des *Status quo* bzw. der *in-Organello* Degradation mitochondrialer Proteine. Zu den Zeitpunkten 0 h, 6 h, 24 h, 72 h und 168 h wurden identische Probenvolumina abgenommen und, wie in 3.2 beschrieben, für die LC-MS Analyse vorbereitet. Die LC-MS Analysen erfolgten über LC Auftrennungen von 200 min auf einem Split-freien NCS HPLC System in Kombination mit einem qExactive Massenspektrometer im TOP15 Modus. Alle Proben wurden in einer Dreifachbestimmung erfasst. Die Messreihe wurde sowohl für die Abschätzung der prozentualen Proteinanteile in Mitochondrien mit Hilfe des NSAF Indexes als auch für die intensitätsbasierte Datenevaluierung auf Peptidebene unter Verwendung der Software Progenesis herangezogen.

Es wurden ausschließlich Features der Ladungszustände +2, +3 und +4, analog zur bereits in 3.12 beschriebenen Suchstrategie, sukzessive in einer mehrstufigen Mascotsuche berücksichtigt. Während für die Datenbanksuche die fixen und variablen Modifikationen identisch blieben, wurden die proteolytischen Parameter und z.T. die Datenbanken variiert und für eine klare Klassifizierung in der Software nacheinander eingepflegt (siehe Tabelle 3.15): (I) Trypsin, 0 überlesene Schnittstellen, SGD (II) Semitrypsin, 0 überlesene Schnittstellen, SGD (III) No cleavage, 0 überlesene Schnittstellen, N-terminal ragged Promito-Datenbank [107] (IV) Trypsin, 4 überlesene Schnittstellen, SGD (V) unspezifischer

Verdau/Peptidtermini, 0 überlesene Schnittstellen, SGD. In der weiteren Datenauswertung wurden (a) Met-haltige bzw. instabile Peptidsequenzen mit der AS Abfolge „NG“, welche leicht deamidiert wird [108], (b) Peptidsequenzen > 22 AS sowie (c) PSMs mit einer Massenabweichung ≥ 4 ppm ausgeschlossen.

Die LF Analyse mittels Progenesis ermöglichte die direkte Visualisierung der Peptidverläufe über die Zeit. Zusätzlich ließ die Klassifizierung der Peptididentifikationen der Datenbanksuchen (I-V) innerhalb der Progenesis Software die selektive Auswertung von Peptiden mit N- bzw. C-terminal semitryptischen Schnittstellen zu, die weiterhin den Kriterien der Eindeutigkeit innerhalb der Datenbank (unique) sowie des statistischen ANOVA Wertes von ≤ 0.05 genügten. Mit dem ANOVA wird die Wahrscheinlichkeit für eine falsch-positive Annahme eines Regulationsereignisses wiedergegeben, welche auf einem Vergleich der Varianz in technischen Replikaten mit der Varianz zwischen verschiedenen Proben beruht. Ein ANOVA Wert von 0,05 besagt, dass ein reguliertes Peptid mit einer Wahrscheinlichkeit von 5 % nicht reguliert ist.

Die weitere Analyse dieser isolierten Peptidsequenzen erfolgte mit Hilfe der Software iceLogo.

3.14.6 Stabil-Isotopen markierter Standard zur Qualitätskontrolle der LF

Quantifizierung

Zur Qualitätskontrolle von LC-MS Analysen für die LF Quantifizierung wurden einer jeden Probe gleiche Mengen des in 3.12.1 beschriebenen QC Standards mit einem dynamischen Bereich von 1 - 125 fmol hinzugefügt. Dies ermöglichte die fortwährende „online“ sowie spätere „offline“-Kontrolle der Retentionszeiten, Peakflächen und FWHM während der LC-MS Messungen. Dadurch konnte die Stabilität des LC-MS Systems während der gesamten Messserie überwacht und somit die Validität der erhaltenen Daten garantiert werden.

3.15 Dateninterpretation

Aufgrund der Vielfalt globaler proteomischer Daten werden für eine übersichtliche und detailreiche Dateninterpretation Softwares und Datenbanken verwendet, die durch Abgleich mit bekannten Signalwegen bzw. durch eine Einteilung nach Funktion oder biologischem Prozess sinnreiche Zusammenhänge darstellen können. In dieser Arbeit wurde zur Analyse von Signalwegen in Thrombozyten die frei zugängliche und manuell gepflegte Datenbank

Reactome (<http://www.reactome.org>) verwendet. Die Gesamtheit der identifizierten Proteine in Thrombozyten wurde zudem mit Hilfe der Software Ontologizer nach der Gene Ontology Annotation (<http://www.geneontology.org>, GO Annotation) klassifiziert.

Für die Analyse von Degradationsprozessen wurde die sequenzspezifische Auswertung von nicht tryptischen Peptiden, zur Charakterisierung einer potentiellen Konsensussequenz mit Hilfe der freien Software iceLogo (<http://code.google.com/p/icelogo/>) [109] durchgeführt.

4. Ergebnisse

4.1 Aufreinigung von Thrombozyten

Bei der Charakterisierung eines Proteoms ist die zugrunde liegende Reinheit der Probe ausschlaggebend für die Aussagekraft und Verlässlichkeit der qualitativen und quantitativen Ergebnisse. Thrombozyten, die im humanen Blut neben Erythrozyten, Leukozyten und Granulozyten mit in etwa 300.000 Exemplaren pro μL auftreten, müssen daher einem dedizierten Aufreinigungsprozess unterzogen werden. Um die Aktivierung der Thrombozyten zu vermeiden, welche leicht u.a. durch Scherstress oder Kontakt mit negativ geladenen Oberflächen, wie Glas, induziert wird, ist unbedingt auf eine schonende Aufarbeitung zu achten.

Thrombozyten aus Frischblut wurden nach einem optimierten Protokoll, u.a. mit aktivierungshemmenden Reagenzien wie dem Enzym Apyrase zur Hydrolyse von ATP, sowie EGTA zum Komplexbilden von Ca^{2+} Ionen aufgearbeitet. Die Reinheit der Thrombozyten wurde innerhalb von FACS Analysen mit Hilfe des Leukocount Kits (BD Bioscience) bestimmt. Während durch Apherese gewonnene Thrombozyten Leukozyten depletiert sind, ist das Auftreten letzterer ein Maß für die Reinheit aus Frischblut gewonnener Thrombozyten. Die Leukozytenkontamination wurde auf weniger als ein Leukozyt pro 10^6 Thrombozyten quantifiziert, zudem wurde eine Kontamination von ca. einem Erythrozyt pro 10^4 Thrombozyten festgestellt. Diese Werte entsprechen einem sehr hohen Reinheitsgrad der Thrombozytenisolation.

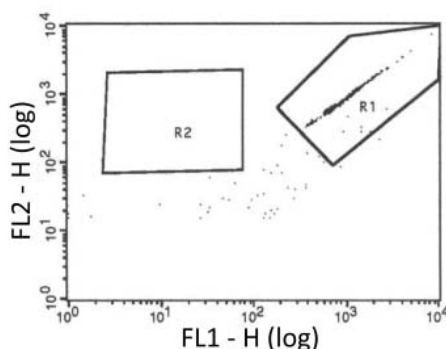


Abbildung 4.1: Leukocount FACS Analyse zur Kontrolle der Reinheit von Thrombozyten. Detektiert werden DNA tragende Zellen wie Leukozyten (Bereich R1) sowie BD Trucount™ Partikel (R2), die vom Hersteller als interne Referenz zur genauen Quantifizierung verbleibender Leukozyten ergänzt werden.

Zusätzlich wurde die Reinheit der Thrombozyten durch Abgleich einer Liste von 68 ausgewählten CD Markern (gelistet im Anhang, Tabelle 12.3), von welchen 28 in Thrombozyten auftreten, mit den proteomischen Ergebnissen evaluiert. Ausgenommen von CD23 konnten alle weiteren 27 in Thrombozyten bekannten CD Marker und zusätzlich zwei

weitere, wahrscheinlich aus Leukozyten stammende, CD81 und CD97, identifiziert werden. CD37, häufig Leukozyten zugeordnet, konnte, wie auch bereits von Dowal et al. in einer umfassenden Palmitoylierungsstudie in Thrombozyten [110], mit etwa 780 Kopien pro Thrombozyt bestätigt werden.

4.2 Aufreinigung von Mitochondrien

Mitochondrien wurden aus Hefekulturen (YPG499) mittels differentieller Zentrifugation und zweier aufeinanderfolgender Sucrosegradienten isoliert [111, 112]. Immunoblots mit kompletten Hefelysaten sowie vor- und hoch aufgereinigten Mitochondrien bestätigten einen steigenden Reinheitsgrad. Wie in Abbildung 4.2 dargestellt, können die mitochondrialen Markerproteine Por1 (OM) und Tim23 (IM) in allen drei Proben nachgewiesen, während Markerproteine anderer Organelle nur im kompletten Hefelysat bzw. in einer vorgereinigten Mitochondrienfraktion (P12) detektiert werden konnten.

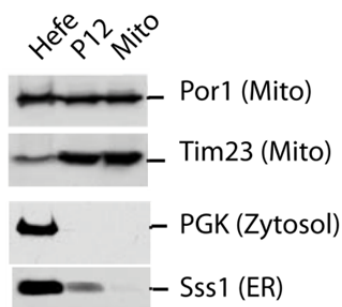


Abbildung 4.2: Westernblot zur Reinheit von hoch aufgereinigten Mitochondrien. Gezeigt sind Gesamthefelysat, eine vor- (P12) sowie eine hoch aufgereinigte Mitochondrienfraktion.

4.3 Qualitätskontrolle des iTRAQ Workflows

Die chemische Markierung von Peptiden mittels iTRAQ findet in der Proteomik häufige Anwendung. Die Vielseitigkeit, mit welcher diese Strategie auf alle Proben - auch jene, die für die metabolische Markierung nicht zugänglich sind, wie humanes Primärgewebe oder Körperflüssigkeiten - angewendet werden kann, stellt einen klaren Vorteil dar. Weiterhin steht die einfache und schnelle Durchführung eines iTRAQ Experiments in keinem Vergleich zur metabolischen Markierung. Essentiell für das Gelingen einer jeden quantitativen Studie ist jedoch die optimale Umsetzung der einzelnen Arbeitsschritte des Workflows – vom Proteinverdau bis hin zur Erstellung der Proteinverhältnisse aus den quantitativen Spektreninformationen. Die Ergebnisse der Kontrolle/Optimierung dieser Schritte sind im Folgenden zusammengefasst.

4.3.1 Optimierung des proteolytischen Verdaus

Die präzise Quantifizierung von biologischen Proben hängt stark von der Verdaueffizienz sowie von der Reproduzierbarkeit des Verdaus ab. Für eine jede Probe müssen daher die Verdaubedingungen und damit die Proteaseaktivität kontrolliert und optimiert werden, da die Zusammensetzung verschiedener Proben aus leicht- und schwerlöslichen Proteinen mit strukturellen Abwandlungen sowie weiteren Zellbestandteilen variieren kann. Die Verdaueffizienz wurde in diesem Zusammenhang für Mitochondrien, wie im Folgenden beschrieben, mittels 1-DE, LC-MS Analysen, iTRAQ und LF Quantifizierung evaluiert.

4.3.1.1 Additive

Die Verdaueffizienz wurde hinsichtlich verschiedener Zusammensetzungen des Verdauerpuffers untersucht. Bei gleichbleibendem Puffersystem, Verdauvolumen und Trypsin:Protein Verhältnis wurden in vier unterschiedlichen Proben das divalente Salz CaCl_2 sowie das organische Lösungsmittel ACN in variierenden Konzentrationen und Kombinationen A-D, wie in Tabelle 4.1 angegeben, ergänzt. Während die Anwesenheit von CaCl_2 zur Ausbildung eines aktiven Ca-Trypsin Komplexes führt und darüber hinaus die Autoprotolyse der Protease unterbindet [113], erhöht der Anteil von ACN die Löslichkeit von hydrophoben Komponenten bei gleichbleibender Aktivität von Trypsin [114].

Puffer	Volumen [μl]	Trypsin : Protein	Zeit [h]	CaCl_2 [mM]	ACN [%]
A	125	1:30	12	1	-
B	125	1:30	12	-	5
C	125	1:30	12	1	5
D	125	1:30	12	1	10

Tabelle 4.1: Übersicht über die Variationen des proteolytischen Verdaus.

Die Verdaukontrolle mittels 1-DE wies nur geringfügige Unterschiede zwischen den Proben A-D auf, hingegen zeigte die iTRAQ Analyse (A) für 1 mM CaCl_2 im Vergleich zu (B) 5 % ACN um 20 % höhere Signalintensitäten für Reporterionen identifizierter Peptide (siehe Abbildung 4.3). Die Kombination der beiden Additive 5 % ACN und 1 mM CaCl_2 in (C) erhöhte die Signale der Reporterionen im Vergleich zu (A) um 50 %. Die in der Literatur beschriebene bessere Verdaueffizienz bei höherem ACN Anteil (D) konnte innerhalb der iTRAQ Analyse nicht bestätigt werden [115].

Eine LF Analyse der Proben A-D bestätigte die in der iTRAQ Analyse evaluierten Trends, wobei die Variation zwischen den Proben geringer ausfiel und allgemein eine hohe Verdaueffizienz festgestellt wurde.

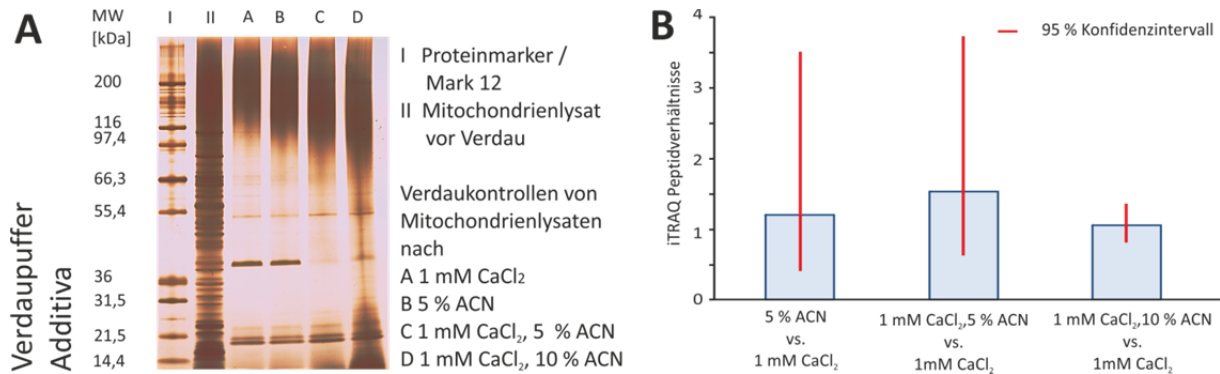


Abbildung 4.3: Kontrolle der tryptischen Verdaueffizienz bei variierenden Pufferadditiven anhand von Mitochondrien (*S. cerevisiae*). (A) Während innerhalb der 1-DE nur geringfügige Unterschiede bei variierenden Konzentrationen an CaCl_2 und ACN zu detektieren sind, zeigt (B) die iTRAQ Analyse beste Ergebnisse für den tryptischen Verdau mit 1 mM CaCl_2 in Kombination mit 5 % ACN. Wie in den 95 % Konfidenzintervallen dargestellt, konnte die Verbesserung des Verdau nicht für alle Peptidspezies beobachtet werden.

4.3.1.2 Zeit

Zur Kontrolle der Auswirkung unterschiedlicher Verdauezeiten auf die Effizienz wurde der Verdau von Mitochondrien mit der Pufferzusammensetzung (C) für Verdauezeiten von 1 h, 4 h, 8 h und 12 h, wie in Tabelle 4.2 (A-D) angegeben, durchgeführt.

Puffer	Volumen [μl]	Trypsin : Protein	Zeit [h]	CaCl_2 [mM]	ACN [%]
A	125	1:30	1	1	5
B	125	1:30	4	1	5
C	125	1:30	8	1	5
D	125	1:30	12	1	5

Tabelle 4.2: Übersicht über die Proben zur Kontrolle der Auswirkung unterschiedlicher Verdauezeiten.

Aufgrund der limitierten Sensitivität und Auflösung der 1-DE konnten in der gelelektrophoretischen Analyse der vier Verdauintervalle keine maßgeblichen Unterschiede festgestellt werden (Abbildung 4.4). In der iTRAQ Analyse konnten jedoch nur wenige identische Peptide in allen Proben A-D quantifiziert werden. Die iTRAQ Analyse konnte somit nur für wenige Analyten durchgeführt werden. Daher lieferte hier die iTRAQ Analyse nur unzureichende Ergebnisse und die Auswertung erfolgte ausschließlich mittels LF Analyse.

Die LF Analyse derselben Proben zeigte jedoch einen deutlichen Anstieg der voll-tryptischen Peptidentifikationen von 1030 nach 1 h auf ein Maximum von 1836 nach 12 h. Entgegen

der naheliegenden Annahme, dass die Zahl der Peptide mit überlesenen Schnittstellen über die Zeit abnimmt, stieg die Zahl der Peptide mit einer überlesenen Schnittstelle analog zu den voll-tryptischen Peptiden über die Zeit an. Bei näherer Untersuchung der überlesenen tryptischen Schnittstellen konnte ein Zusammenhang mit dem Auftreten von unmittelbar aufeinander folgenden tryptischen Schnittstellen [R/K] festgestellt werden, die wie in der Literatur beschrieben wesentlich schlechter verdaut werden [64]. Die Zahl der Peptide mit zwei überlesenen Schnittstellen stieg nach 8 h Verdau auf ein Maximum an, um danach wieder abzufallen.

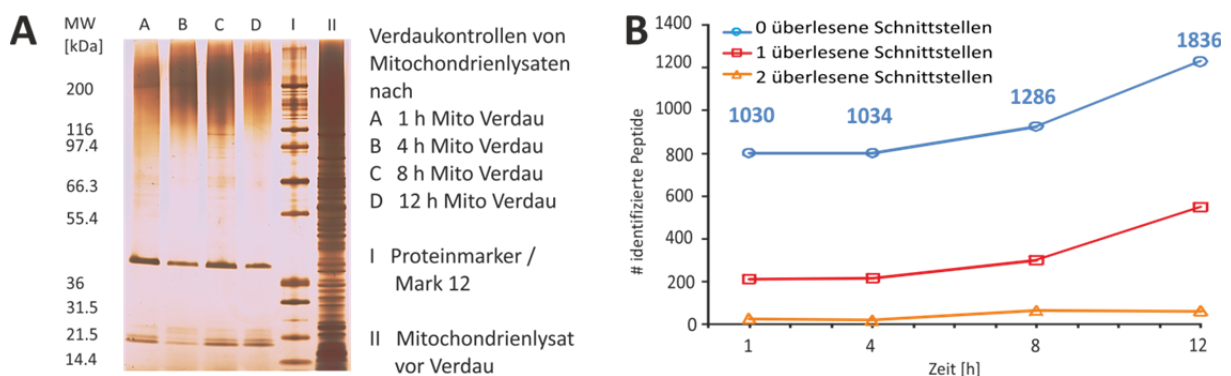


Abbildung 4.4: Kontrolle der tryptischen Verdaueffizienz über die Inkubationszeit. (A) Anhand der 1-DE konnten keine maßgeblichen Unterschiede zwischen den vier Verdauintervallen von 1 h, 4 h, 8 h und 12 h festgestellt werden. (B) Die LF Analyse zeigte jedoch einen deutlichen Anstieg an identifizierten voll-tryptischen Peptiden über die Zeit.

Dieses Ergebnis lässt das Vorhandensein von Peptidsequenzen mit einer Vielzahl von überlesenen Schnittstellen vermuten, die nicht identifiziert werden konnten. Erst nach 8 h Verdauzeit waren die korrespondierenden Sequenzbereiche auf Peptidlängen verdaut worden, die in der Datenauswertung mit zwei überlesenen Schnittstellen aufgeführt werden. Nachdem für den Verdau mit 1 mM CaCl_2 , 5 % ACN über 12 h die meisten Peptididentifikationen gefunden werden konnten, wurden alle weiteren Experimente unter diesen Bedingungen durchgeführt.

4.3.2 Entsalzen und iTRAQ Markierungseffizienz

Während laut Herstellerprotokoll für den iTRAQ Workflow eine Acetonpräzipitation vor der Markierung vorgesehen ist, wurde hier die alternative Verwendung von C18 SPE auf Peptidebene zur Entfernung von für die iTRAQ Markierung störenden Substanzen evaluiert. Wie in der Literatur beschrieben, weisen Präzipitationsprotokolle mit verschiedenen organischen Lösungsmitteln und Säuren variierende Spezifitäten auf und können so zu

undefinierbaren Proteinverlusten führen, mit welchen gleichzeitig eine quantitative Veränderung der Probenzusammensetzung einhergeht [61]. Weiterhin wurde bei der Verwendung von Aceton das Auftreten von Peptidmodifikationen festgestellt [62].

Die Proben III-1 und III-2 wurden auf Verluste bei der C18 Entsalzung mittels LF Analyse untersucht. Dabei konnten anhand der Signalintensitäten für 427 (III-1) bzw. 426 (III-2) der Peptide, die sowohl vor als auch nach der C18 Entsalzung identifiziert wurden, vernachlässigbare Peptidverluste von 4 % mit einer SD von 5 % festgestellt werden. Analog dazu waren die entsprechenden UV- sowie Base-Peak-Chromatogramme der LC-MS Auftrennung nahezu identisch (Abbildung 4.5 A und B).

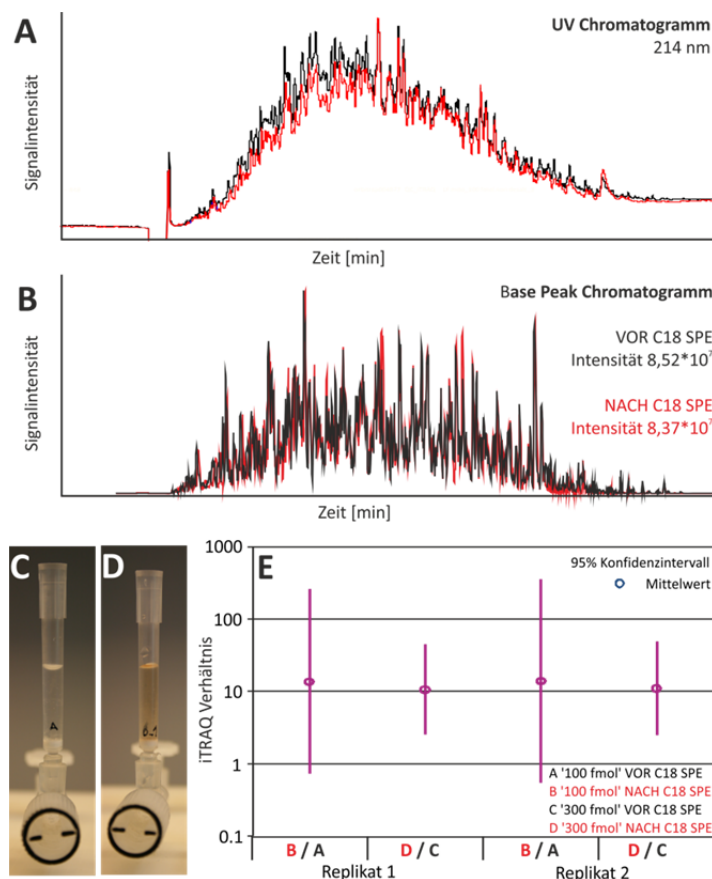


Abbildung 4.5: Kontrolle der C18 SPE auf Peptidverluste und die Auswirkung auf die iTRAQ Markierungseffizienz. In einer LC-MS Analyse zeigten entsalztes (rot) und nicht entsalztes (schwarz) Proben nur geringfügige Unterschiede sowohl im (A) UV- als auch im (B) Base Peak Chromatogramm. Die LF Analyse bestätigte vernachlässigbare Peptidverluste von 4 %. (E) Die Markierungseffizienz konnte nach C18 SPE Aufreinigung um einen Faktor von durchschnittlich 12.5 für variierende Probenkonzentrationen (100 und 300 fmol) und zwei Replikate erhöht werden. (C-D) Die verbesserte Markierungseffizienz konnte bereits während der Probenaufarbeitung in einer stärkeren Färbung der Probenlösung detektiert werden.

Für die Markierungseffizienz ist der Entsalzungsschritt von erheblicher Bedeutung: In einem Experiment bei welchem die Proben III-1 und III-2 nach bzw. ohne Entsalzung mit den iTRAQ Reagenz markiert wurden, zeigten sich für die entsalzten Proben in zwei Replikaten durchschnittlich 12.5-fach erhöhte Intensitäten. Dies konnte bereits während der Probenaufarbeitung anhand der unterschiedlich starken Färbungen der Proben gesehen werden (Abbildung 4.5 C-D).

4.3.3 Dynamischer Bereich

Der dynamische Bereich der iTRAQ Analyse ist von zwei Extremen begrenzt: Zum einen ist durch die Sättigung des Detektors eine maximale Signalintensität gegeben, zum anderen wird die minimale Signalintensität durch das Aufkommen von Rauschen im Detektor bestimmt. Zur Analyse der maximal detektierbaren Peptidverhältnisse in iTRAQ Experimenten wurde ein Probensatz mit definierten Mengen an zugegebenen Proteinverdau mit (1) 5 fmol, (2) 25 fmol (3) 250 fmol und (4) 500 fmol zu jeweils 0,1 µg mitochondrialem Background erstellt. Die dabei zu erwartenden Verhältnisse konnten in Abhängigkeit der durchschnittlichen Intensitäten mit variierenden SD detektiert werden, wobei niedrige Intensitäten unterhalb des Quantifizierungslimits der sogenannten „*Underestimation*“ zum Opfer fallen und daher die Verhältnisse nur unzureichend und mit hohen SDs widerspiegelt werden. Die Fehler für die Quantifizierung betragen 64 %, 56 % und 15 % für die entsprechenden Peptidverhältnisse der Proben 1:5 (5 fmol:25 fmol), 1:10 (25 fmol:250 fmol) und 1:2 (250 fmol:500 fmol).

Die Ergebnisse lassen folgern, dass die maximal detektierbare Peptidvarianz gleich dem Quotienten der beiden Quantifizierungslimits ist. Diese sind jedoch stark abhängig von den zu analysierenden Proben, dem verwendeten MS Analysator und der allgemeinen LC-MS Performance [116]. In dieser Analyse wurden die iTRAQ Reporterionen in der Sättigung mit durchschnittlich $5 \cdot 10^5$ Signalintensität und im Minimum mit $7 \cdot 10^3$ detektiert und damit das Quantifizierungslimit bei 20 festgestellt.

4.3.4 Qualitätskontrollierte Datenauswertung von iTRAQ Analysen

Allgemein ist die korrekte Zuordnung von Peptiden für die Proteinquantifizierung essentiell, wobei die Auswahl der Peptide, ob (a) *unique*, (b) vollständig tryptisch geschnitten bzw. (c) mit überlesenen Schnittstellen oder (d) Peptidmodifikationen, Einfluss auf die finale Quantifizierung hat. Ergänzend dazu gestalten sich Realproben als Kombination einer asymmetrischen Verteilung mit Ausreißern, so dass viele statistische Evaluierungsmethoden nur bedingt anwendbar sind. Daher wurden in dieser Arbeit in Zusammenarbeit mit Marc Vaudel verschiedene Verfahren für die Auswertung von iTRAQ Daten, nämlich der Mittelwert, der Median und der „redescending M-estimator“ evaluiert [117]. Der redescending M-estimator zeigte, anhand einer artifiziellen Verteilung verglichen zu den theoretischen Verhältnissen, die robustesten und akkuratesten Ergebnisse, wobei sich der

M-estimator vornehmlich auf die Form der Verteilung konzentriert. Ausreißer, die v.a. den statistischen Mittelwert erheblich beeinflussen, werden bei Anwendung des redescending M-estimators ab einer Überschreitung der 1.48-fachen mittleren absoluten Abweichung (engl. Median absolute deviation, MAD) vernachlässigt. Der redescending M-estimator ist auf die Zentren der Verteilung der Proteinverhältnisse fokussiert und gewährleistet eine hohe Qualität der Daten [117].

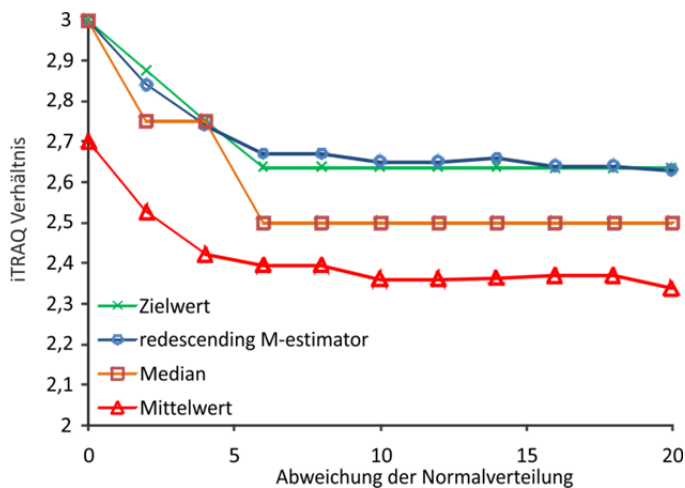


Abbildung 4.6: Untersuchung der Robustheit und der Genauigkeit der Proteinverhältnisbestimmung verschiedener statistischer Verfahren anhand einer simulierten Verteilung. Während der Mittelwert und der Median stark durch Ausreißer beeinflusst sind, schließt der redescending M-estimator Ausreißer ab einer $1.48 \times \text{MAD}$ aus. Stattdessen fokussiert der M-estimator die Verteilung und erzielt somit die beste Darstellung der Proteinverhältnisse.

4.3.5 Analyse der iTRAQ Interferenz

Die Anwendung von Reporterion-basierten Quantifizierungsstrategien wie iTRAQ oder TMT hat in der Proteomik aufgrund der einfachen Durchführung, des Multiplexings von bis zu acht unterschiedlichen Proben und der überschaubaren Kosten breite Anwendung gefunden. Die Freisetzung der quantitativen Information nach Fragmentierung wird jedoch von der Problematik der Interferenz begleitet, welche durch die Ko-Elution unabhängiger Vorläuferionen mit ähnlichen m/z Verhältnissen zur simultanen Isolierung in einem definierten Isolationsfenster von meist 1-3 m/z und anschließender gemeinsamer Fragmentierung auftritt. Die freigesetzten Reporterionen werden zur Quantifizierung herangezogen und einem einzigen identifizierten Peptid zugeordnet – mit teils folgenschweren Fehlern in der Quantifizierung (siehe Abbildung 3.6).

Die kürzlich in der Literatur beschriebenen Strategien zur Charakterisierung der Interferenz mittels MS^3 Fragmentierung [118] bzw. mittels Gasphasenextraktion durch Ladungsreduktion [119] ermöglichen deren Detektion und führen weiterhin zu weniger fehlerbehafteten quantitativen Ergebnissen. Allerdings sind der experimentelle Aufwand und auch die Datenprozessierung anspruchsvoll: Das Verfahren mittels MS^3 Fragmentierung

basiert darauf, dass zunächst ein CID MS/MS durchgeführt und anschließend das höchste Fragmentation einer weiteren HCD MS/MS/MS Fragmentierung unterzogen wird, bei der dann die Reporterionen mit geringerem Hintergrund detektiert werden und somit eine akkuratere Quantifizierung erreicht wird. Um Sicherzustellen, dass das entsprechende MS/MS Fragmentation iTRAQ markiert hat, muss für diesen Ansatz die Protease LysC (schneidet C-terminal von Lys-Resten) eingesetzt werden. Dadurch wird gewährleistet, dass „jedes“ Peptid sowohl am N- als auch am C-Terminus eine stabil-Isotopen Markierung trägt, da nur primäre Amine (N-Terminus sowie Lys-Reste) gelabelt werden. Y-Ionen voll-tryptischer Peptide, die mit Arg enden, sind nicht iTRAQ markiert und würden daher keine Reportersignale im MS³ generieren. Dies führt zu deutlich verringerten Identifikationsraten, da LysC Peptide im Durchschnitt länger als tryptische Peptide sind und im Schnitt auch höhere Ladungszustände aufweisen. Zudem werden für jedes Peptid zwei Data-dependent Spektren aufgezeichnet, was zu einer deutlichen Verlangsamung verglichen mit einem normalen MS/MS Experiment führt. Die Gasphasenextraktion hingegen erfordert die Dekonvolution der ko-isolierten Vorläuferionen mit anschließender Manipulation der zugrundeliegenden Ladungszustände bzw. Massen durch Protonentransfer. Ko-isolierte Vorläuferionen unterschiedlicher Ladungszuständen können somit in der Gasphasenextraktion getrennt und selektiv fragmentiert werden. Beide Verfahren entsprechen nicht der simplen Anwendung von Reporterion-basierten Quantifizierungsstrategien und sind zudem instrumentell eingeschränkt.

Die im Rahmen dieser Arbeit neuentwickelte target/decoy Strategie zur Bestimmung der qFDR ermöglicht, durch den Einsatz eines Referenzproteoms, die quantitative Abschätzung des durch die Interferenz verursachten Fehlers bei Reporterion-basierten Quantifizierungsstrategien (engl. Integral quantification accuracy estimation for reporter ion-based quantitative proteomics, iQuARI). Darüber hinaus folgt iQuARI der in dieser Arbeit beschriebenen Probenvorbereitung und verzichtet auf eine komplizierte bioinformatische Prozessierung.

Während in der qFDR Analyse ein Reporterionenkanal mit dem Referenzproteom belegt ist, können alle weiteren Kanäle für biologische Proben genutzt werden. Für das Referenzproteom ist eine ähnliche Komplexität und ein ähnlicher dynamischer Bereich bei größtmöglicher Orthogonalität zum analysierenden Proteom grundlegend für eine statistisch valide Durchführung. Hierfür wurde PFU als geeignetes Referenzproteom ausgewählt, da

dieser Organismus evolutionär weit vom humanen Genom entfernt ist. Ein Vergleich mit den entsprechenden Datenbanken Eukaryota, Vertebrata, Mammalia und *Homo sapiens* zeigte, dass in PFU nur 151 (Eukaryota), 16 (Vertebrata), neun (Mammalia) bzw. sieben (Human) identische, vollständig tryptische Peptide gefunden werden können. In 15.365 MS/MS PFU Spektren konnte kein einziger zuverlässiger Hit für ein humanes Peptid identifiziert werden, wobei der beste PSM einer decoy-Sequenz zugeordnet wurde [120] – dies entspricht einer maximal möglichen Orthogonalität, ohne dass hierfür ein artifizielles Proteom verwendet werden muss.

Zur Analyse der Interferenz wurden in einem *proof-of-principle* Experiment jeweils zwei Kanäle eines 4-plex iTRAQ Kits mit identischen Mengen humaner Thrombozyten (114, 115) bzw. PFU (116, 117) belegt. In der LC-MS Analyse, die folglich für Spektren humanen Ursprungs eine Reporterionenverteilung von 114:115:116:117 = 1:1:0:0 und für PFU Spektren entsprechend 114:115:116:117 = 0:0:1:1 aufzeigen sollte, wurden 68±8 % aller Spektren mit einem Fehler von weniger als 20 %, 26±6 % aller Spektren mit einem Fehler zwischen 20 % und 80 % und 6,5±2 % aller Spektren mit einem Fehler von über 80 % gemessen. Am stärksten betroffen waren niedrig abundante Peptide, bei welchen umgebende Rauschsignale mit fragmentiert wurden und daher zu einer stärkeren Ausprägung der Interferenz führten.

Zur quantitativen Evaluierung der Interferenz wurden in einem weiterführenden Experiment drei Kanäle eines 4-plex iTRAQ Kits mit den Fraktionen 3 (114), 9 (115) und 20 (116) einer mit humanen Thrombozyten durchgeführten IEF und nur ein Kanal mit PFU (117) belegt. Die IEF Fraktionen wiesen in einer separaten LC-MS Analyse, verglichen zum kompletten Proteom im *proof-of-principle* Experiment, aufgrund der niedrigeren Komplexität nur wenige überlappende Spektren auf. Im iTRAQ Experiment zeigten folglich 25 % aller Spektren keine Interferenz, 68,4 % aller Spektren einen Fehler von weniger als 20 %, 5,2 % aller Spektren einen Fehler zwischen 20 % und 80 % und 1,3 % aller Spektren einen Fehler von über 80 %. Die Interferenz konnte weiter mit der Verlängerung des LC Gradienten von 95 min auf 185 min, in Tabelle 4.3 gezeigt, geringfügig reduziert werden, während die Zahl der identifizierten Peptide wie zu erwarten stark anstieg.

Die Einschränkung des Isolationsfensters der Vorläuferionen von 2 m/z auf 1 m/z verringerte die Problematik der Interferenz und es konnten mit 6.998 Peptididentifikationen im 185 min Gradienten 15 % mehr Spektren ohne Interferenz detektiert werden. Allerdings führte die

Einschränkung des Isolationsfensters auch zu signifikant weniger Peptidentifizierungen bei 1 % FDR (7 % weniger PSM).

Anzahl der Peptid IDs bei 1% FDR	Gradientenlänge / Isolationsfenster					
	95 min / 2 m/z		185 min / 2 m/z		185 min / 1 m/z	
	absolut	relativ	absolut	relativ	absolut	relativ
total	11.766	-	21.801	-	20.478	-
Interferenz 0 %	2.945	25,1%	6.108	28,0%	6.998	34,2%
Interferenz 0 % - 20 %	8.047	68,4%	14.410	66,1%	12.286	60,0%
Interferenz 20 % - 80 %	617	5,2%	1.060	4,9%	936	4,6%
Interferenz > 80 %	157	1,3%	223	1,0%	258	1,2%

Tabelle 4.3: Übersicht über die Auswirkung der LC Gradientenlänge und des Isolationsfensters auf die Interferenz in der iTRAQ Analyse.

4.4 Vergleich verschiedener Trypsine bezüglich Effizienz und Reproduzierbarkeit

Die Quantifizierung von Proteinen basiert in der Proteomanalyse meist auf der Detektion von tryptischen Peptiden, wobei Aminosäuresequenzen mit überlesenen Schnittstellen bzw. inkorrekt geschnittene Peptide wie z.B. semitryptische Peptide oft von der Quantifizierung ausgeschlossen werden. Daher sind die Reproduzierbarkeit und die Vollständigkeit des proteolytischen Verdaus von maßgebender Bedeutung für eine verlässliche Quantifizierung. Zur Charakterisierung sechs kommerzieller Trypsine 1-6 (gelistet im Anhang, Tabelle 12.1) wurden jeweils 500 µg eines humanen Thrombozytenlysats bei identischen Bedingungen in Triplikaten aufgearbeitet und hinsichtlich der Verdaueffizienz und ihrer Auswirkung auf quantitative Ergebnisse analysiert.

4.4.1 Monolith

Die monolithische Auftrennung ermöglicht die parallele Separation von Proteinen und Peptiden innerhalb eines LC Laufs, wobei Proteine aufgrund ihres hydrophoben Charakters zu einem späteren Zeitpunkt von der Trennsäule eluieren als Peptide.

Wie in Abbildung 4.7 dargestellt, konnte die monolithische Auftrennung bereits Unterschiede in der Verdaueffizienz aufzeigen: Bei identisch eingesetzten Probenmengen variierte die Menge an eluierten Peptiden für Trypsin 6 stark von den Trypsinen 1-5 und zeigte im späten Gradienten weiterhin Signale, die nicht bzw. unvollständig geschnittenen Proteinen zugeordnet werden können. Die Reproduzierbarkeit konnte, wie in 4.7 B dargestellt, durch Triplikate bestätigt werden.

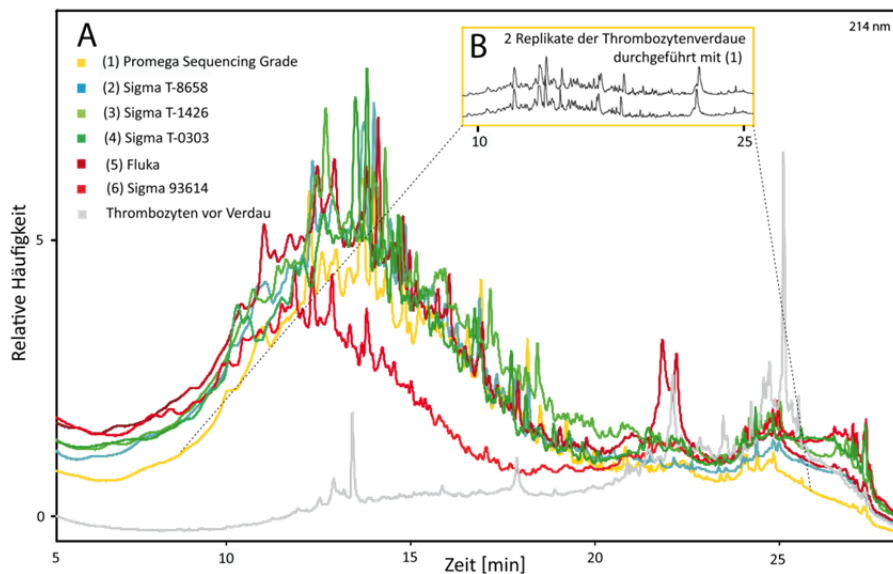


Abbildung 4.7: Verdaugkontrolle mittels monolithischer Auftrennung von unverdauten (grau) und mit 6 verschiedenen Trypsinen verdauten Thrombozytenproben. Während die verschiedenen Trypsine stark variieren, zeigt (B) der Verdauung mit Trypsin 1 reproduzierbare Signale.

4.4.2 Dreistufige Datenbanksuche

Die Datenevaluation der nachfolgenden SCX- und direkten nano-LC-MS/MS Analysen erfolgte in einer dreistufigen Datenbanksuche unter Verwendung der Software Proteome Discoverer in Kombination mit Mascot 2.2.3. Bei dem dreistufigen Prozess wurden die Rohdaten zunächst einer (1) tryptischen Datenbanksuche gegen die humane Uniprot Datenbank unterzogen, wobei mit nicht zugeordneten Fragmentationsspektren sukzessiv weitere Datenbanksuchen mit veränderten proteolytischen Parametern - (2) „Semitrypsin“ und abschließend (3) „None“, eine hinsichtlich der Protease unspezifische Suche - durchgeführt wurden. Nach dem Filtern der Suchergebnisse mit jenen in 3.12. angegebenen stringenten Kriterien, wurden anhand der Replikate der Verdauproben 1-6 Mittelwerte für die Anzahl an Spektren, Peptiden bzw. Proteinen gebildet und verglichen.

4.4.3 SCX

Die SCX Chromatographie bei pH 2,7 ermöglicht die Auftrennung von tryptischen Peptiden anhand ihrer positiven Ladungen, wobei höhere Ladungszustände zu einer verstärkten Retention auf der Trennsäule führen. Für die Charakterisierung der Verdaueffizienz und -spezifität verschiedener Trypsine wurden drei Fraktionen SCX 1, SCX 2 und SCX 3 zur Analyse der Ladungszustände +1 und +2 gesammelt und anschließend mittels nano-LC-MS/MS auf einer LTQ-Orbitrap XL untersucht (siehe Abbildung 4.8).

Tryptische Peptide haben bei pH 2,7 gewöhnlich, aufgrund ihres positiv geladenen N-Terminus sowie des C-terminalen Arg-/Lys-Restes, mindestens eine Ladung von +2. Während jede überlesene Schnittstelle und jeder zusätzliche His-Rest innerhalb der Sequenz diesen Ladungszustand um +1 erhöht, führen PTMs wie Phosphorylierung oder die Acetylierung des N-Terminus zu einer Verringerung der Nettoladung um 1.

Bei der Auftrennung der verschiedenen Trypsinverdauung und ihrer Replikate konnten starke Varianzen für ladungsreduzierte Peptidspezies (+1) detektiert werden. Nachdem ladungsreduzierte Peptide aufgrund von PTMs substöchiometrisch auftreten, korreliert der Anstieg im UV Signal im Bereich der +1 geladenen Peptide mit einer verminderten tryptischen Verdauspezifität und dem verstärkten Auftreten von semitryptischen und unspezifisch gespaltenen Peptiden, welche häufig einen positiv geladenen N-Terminus, aber kein C-terminales Arg bzw. Lys enthalten.

Während Trypsin 1 nur geringe Signalintensitäten im ladungsreduzierten Bereich aufwies, wurden für die weiteren Trypsine intensivere Signale detektiert. Die scheinbar hohe Spezifität von Trypsin 1 konnte auch in der anschließenden nano-LC-MS/MS Analyse bestätigt werden: Für die Fraktionen SCX 1 und SCX 2 konnten 84 % bzw. 83 % voll-tryptische Peptide identifiziert werden, unterdessen zeigten die Trypsine 2-6 verminderte tryptische Anteile, wie Trypsin 2, das für SCX 1 und SCX 2 entsprechend nur 66 % und 69 % aufwies. Trypsin 6 zeigte die schlechteste Spezifität mit nur 30 % und 37 % tryptischen Peptiden für SCX 1 und SCX 2.

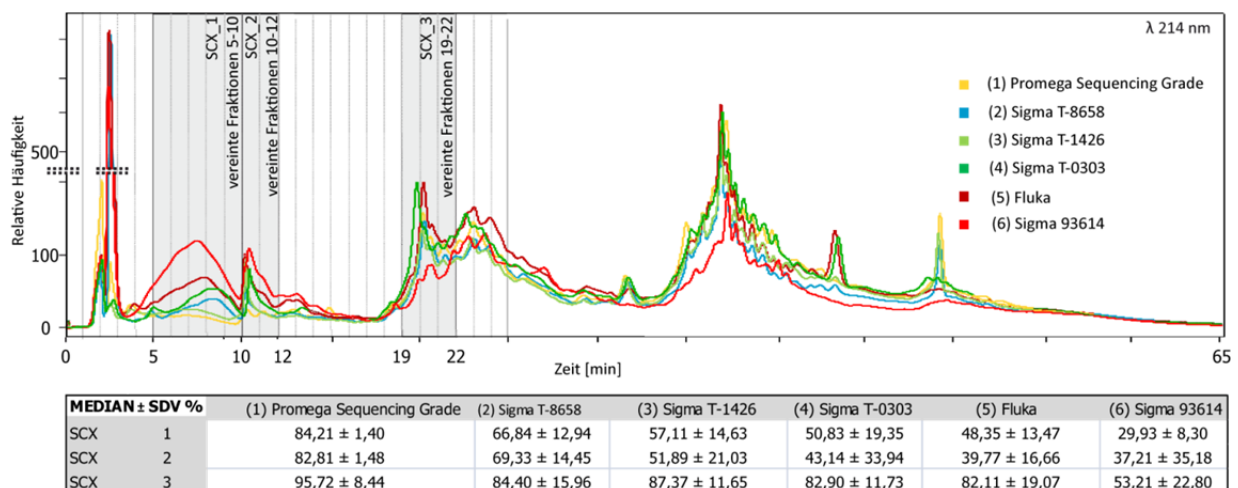


Abbildung 4.8: SCX Fraktionierung der Trypsinverdauung 1-6. Bereits im UV-Chromatogramm können starke Variationen für einfach geladene Peptidspezies in SCX 1 identifiziert werden. Die Tabelle zeigt für alle Trypsine und Fraktionen den in Triplikaten bestimmten Anteil an tryptischen Peptiden (Median±SD) bezogen auf die Gesamtheit aller identifizierten Peptide. Trypsin 1 weist selbst in den theoretisch ladungsreduzierten Fraktionen SCX 1 und SCX 2 hohe Anteile an vollständig tryptischen Peptiden auf, wohingegen für Trypsin 6 in allen Fraktionen S1-3 niedrige tryptische Anteile mit hohen SDs identifiziert wurden.

4.4.4 Direkte LC-MS

Für die Analyse der thrombozytären Verdauung mittels direkter nano-LC-MS/MS wurde auf eine Entsalzung der Proben verzichtet, um potentielle Verluste zu unterbinden. Die Datenauswertung erfolgte, wie unter 4.4.2. beschrieben, in einem dreistufigen Prozess (1-3).

(1) Die variierende Verdauungseffizienz kann bereits an der Anzahl der akquirierten Spektren und der identifizierten Peptide abgelesen werden: Ausgenommen von Trypsin 1 (18.592) und Trypsin 6 (16.224) beläuft sich die Anzahl der Spektren für alle weiteren Verdauungstests auf ca. 20.000, wobei für Trypsin 1 nach der tryptischen Datenbanksuche und stringentem Filtern mit durchschnittlich 7.630 Peptiden (41 % der MS/MS Spektren) verglichen zu den anderen Trypsinen die meisten Spektren identifiziert werden konnten. Wie auch bereits in der SCX Analyse abzusehen war, konnten für Trypsin 6 nur 1.399 tryptische Peptide (8 %) identifiziert werden. (2) Die Anzahl der nicht zugeordneten Spektren, welche einer semitryptischen Suche unterzogen wurden, zeigte ein inverses Bild zu den tryptischen Peptididentifikationen. Für Trypsin 1 wurden die wenigsten Spektren (5.003) semitryptisch gesucht und identifiziert (322), wohingegen Trypsin 6 mit 9.921 Spektren 829 semitryptische Peptide erzielte. Die größte Anzahl an semitryptischen Identifikationen konnte jedoch für die Trypsine 4 und 5 mit etwa 1.250 Peptiden und 200 Proteinen gefunden werden. (3) Für die unspezifische Datenbanksuche wurden für Trypsin 6 die meisten Spektren (6.789) gesucht wie auch die meisten unspezifischen Peptide (214) und Proteine (42) identifiziert, während für Trypsin 1 jeweils die geringste Anzahl an Treffern ermittelt wurde.

Ergänzend korrelierten die durchschnittlichen Protein- und Peptidscores, dargestellt als Mittelwert der drei Replikate, mit der Anzahl an Identifizierungen in Abbildung 4.9 D. Folglich wirkt sich die Qualität des verwendeten Enzyms auch auf qualitative Proteomstudien aus, wobei eine Chymotrypsin-ähnliche Nebenreaktivität für alle Trypsine detektiert werden konnte. Die Nebenreaktivität ist auf die autoproteolytische Prozessierung zu Pseudotrypsin zurückzuführen [121] und kann sowohl durch Methylierung der Lysinreste als auch durch die Zugabe von L-1-Tosylamido-2-Phenylethyl-chloromethyl-ke-ton (TPCK) als Chymotrypsin-Inhibitor reduziert, nicht aber vollständig unterbunden werden [122].

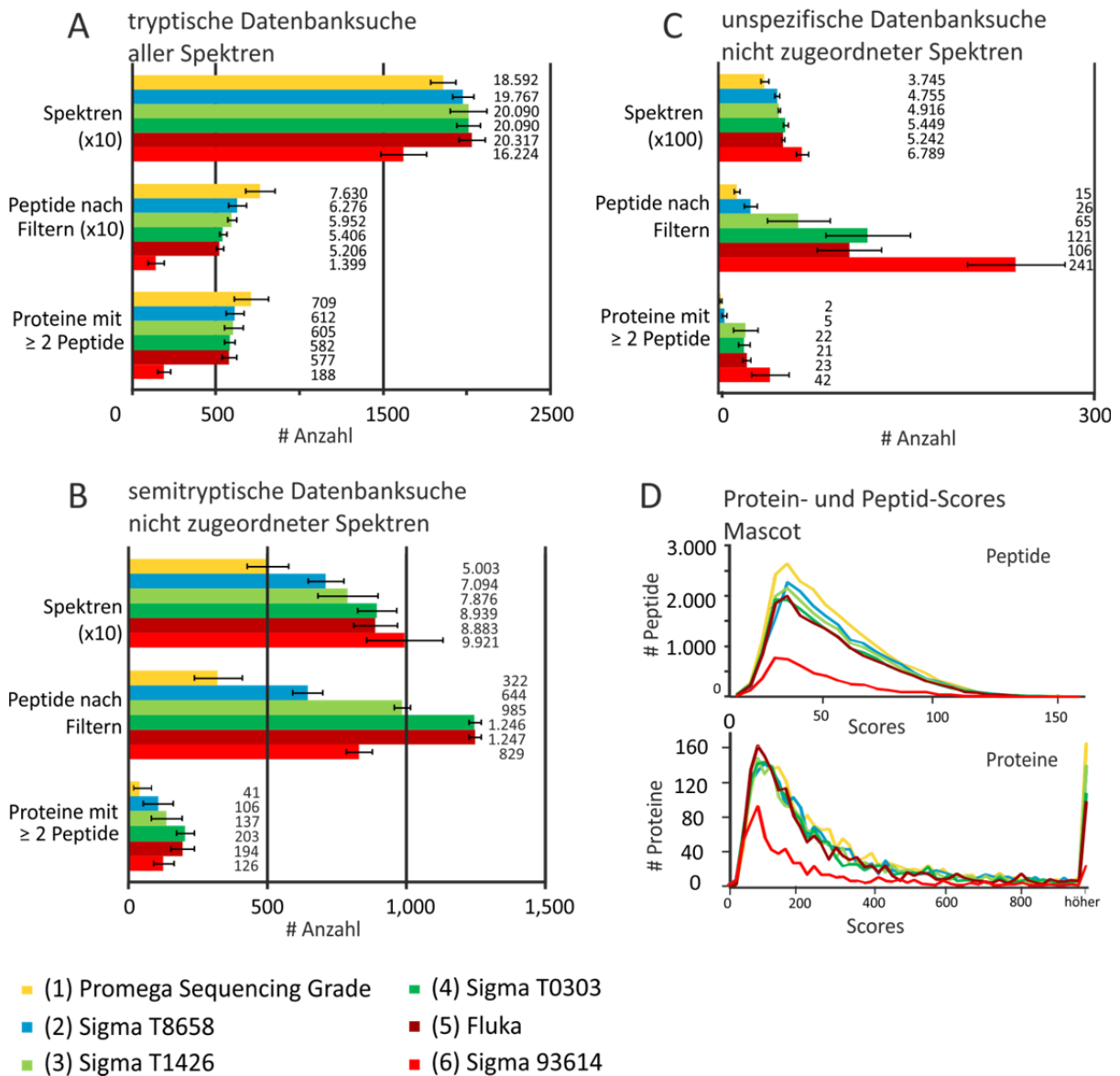


Abbildung 4.9: Direkte LC-MS Analysen von jeweils 1 µg Thrombozytenverdau aller Trypsine. In A-C ist die dreistufige Datenauswertung der drei Replikate mit Median±SD dargestellt: (A) Alle MS/MS Spektren wurden zunächst mit Trypsin als Protease gesucht, wobei die Zahl der Spektren nicht mit der Anzahl an identifizierten Peptiden korrelierte und für Trypsin 1 signifikant mehr Peptide identifiziert werden konnten. (B) Die in (A) nicht identifizierten Spektren wurden einer semitryptischen Datenbanksuche unterzogen; auffallend ist die niedrige Spektrenanzahl für Trypsin 1 und die zutreffende Korrelation zwischen den Spektrenzahlen und den semitryptischen Treffern. (C) Die in B nicht identifizierten Spektren wurden abschließend unspezifisch gesucht. Trypsin 1 zeigte hohe Spezifität und identifizierte nur 15 Treffer, während Trypsin 6 mit 241 Peptiden die meisten Treffer lieferte. (D) Die Verteilung der Peptid- und Protein-Scores (Mascot) zeigte für alle Trypsinverdaue denselben Trend. Die Verdaueffizienz führte offensichtlich zu global verminderten Peptidhäufigkeiten (Burkhart et al., *J Proteomics*, 2011).

4.4.5 SRM

Um die Auswirkung der Verdauereffizienz auf quantitative Studien zu evaluieren, bei welchen häufig Peptide mit überlesenen Schnittstellen von der Quantifizierung ausgeschlossen werden, wurde eine peptidorientierte SRM Studie durchgeführt. Für die Analyse wurden 13 Peptidpaare mit/ohne überlesener tryptischer Schnittstelle wie z.B. DFGSFDK/DFGSFDKFK ausgewählt, die in HCD MS/MS Läufen von Thrombozytenverdauen auf einer LTQ-Orbitrap Velos identifiziert wurden (siehe Anhang 12.2). Diese Peptidpaare – bzw. das Verhältnis der SRM-Intensitäten zwischen voll geschnittener und überlesener Variante eines Paares – wurden als Referenzen ausgewählt, um Effizienz, Spezifität und Reproduzierbarkeit der Thrombozytenverdauere (sowie Replikate) der Trypsine 1-6 semi-quantitativ zu evaluieren. Da die Partner eines Peptidpaares grundsätzlich unterschiedliche Ionisierungseffizienzen aufweisen, kann das tatsächliche Verhältnis von voll geschnittener zu überlesener Variante prinzipiell nur mit stabil-Isotopen markierten Versionen der jeweiligen Peptide bestimmt werden [123]. Stattdessen wurden die beiden Varianten zwischen den verschiedenen Verdauereproben relativ quantifiziert, um eine allgemeine Aussage über die Effizienz der einzelnen Trypsine zu erhalten. Wie in Abbildung 4.10 zu sehen, konnten für Trypsin 1 die intensivsten Signale mit der niedrigsten SD von 24 % detektiert werden, wobei die Gesamtintensität der evaluierten Signale proportional zur Kreisfläche ist. Darüber hinaus konnten für Trypsin 1 82 % der Signale vollständig tryptischen Peptiden und nur 18 % Peptiden mit überlesenen Schnittstellen zugeordnet werden.

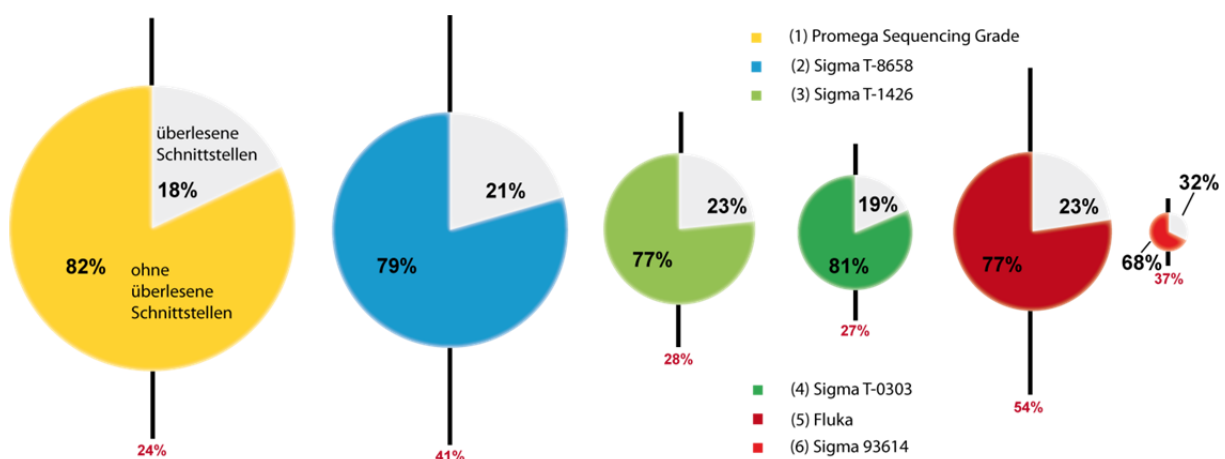


Abbildung 4.10: Zusammenfassung der SRM Ergebnisse. Peptidsequenzen mit und ohne überlesener proteolytischer Schnittstelle wie DFGSFDK / DFGSFDKFK wurden relativ für die sechs Thrombozytenverdauere quantifiziert. Die Kreisgröße korreliert mit dem Median der Peakflächen aller untersuchter SRM Übergänge für jeweils eine Verdauerebedingung und ihrer Replikate und setzt sich aus Peptiden mit und ohne überlesener Schnittstellen zusammen. Trypsin 1 wies die intensivsten SRM Signale mit der niedrigsten SD (24%) auf. Darüberhinaus konnten für Trypsin 1 82% aller Peptide vollständig tryptischen Spezies zugeschrieben werden. Diese Ergebnisse weisen Trypsin 1 als Protease mit einem hohen Maß an Spezifität, Reproduzierbarkeit und Verdauereffizienz aus.

Nur 0,2 % der überlesenen Schnittstellen aller Proben waren auf die verminderte Aktivität von Trypsin nach der Sequenzabfolge K-P bzw. R-P, der sogenannten „Keil-Rule“ zurückzuführen [64]. Auf eine detaillierte Quantifizierung der Peptidsequenzen mit überlesenen Schnittstellen wurde aufgrund des Auftretens von vielerlei Aminosäuresequenzvarianten u.a. mit PTMs verzichtet.

4.4.6 SDS-PAGE

Zur Analyse der in 4.4.3 detektierten chymotryptischen Aktivität der Trypsinpräparate wurde je 1 µg der Trypsine 1-6 innerhalb einer 1-D SDS-PAGE aufgetrennt, die mit Silberfärbung visualisierten Banden ausgeschnitten und einem in-Gel Verdauprotokoll unterzogen. In der nano-LC-MS/MS Analyse via LTQ-Orbitrap XL konnte ausschließlich für Trypsin 6 eine Kontamination mit Chymotrypsin festgestellt werden.

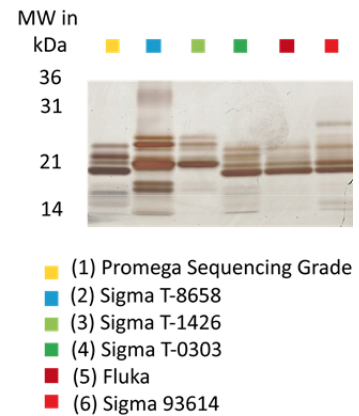


Abbildung 4.11: 1-DE-SDS-PAGE der Trypsine 1-6.

4.5 Qualitätskontrolle von nano-LC-MS Systemen

Die QC von nano-LC-MS Systemen wurde mit Hilfe der in 3.11 beschriebenen Stabil-Isotopen markierten QC Mixtur durchgeführt. Für eine detailreiche Studie des HPLC Systems inklusive LOD, LOI, LOQ und Memory Effekten wurden ansteigende Konzentrationen über drei Größenordnungen von $c_{\min} = 0,125$ fmol bis $c_{\max} = 1250$ fmol der QC Mixtur mit bzw. ohne biologischen Hintergrund innerhalb eines 90 minütigen LC-Gradienten aufgetrennt und in Kombination mit einem LTQ-Orbitrap XL Massenspektrometer analysiert [98].

Aufgrund von hochempfindlichen Massendetektoren konnten Memory Effekte selbst für niedrige Konzentrationen der QC Mixtur sowohl in den Variationen der LC Gradienten als auch in Kombination mit optimierten Waschgradienten detektiert werden (Abb. 4.12). Diese sind jedoch für quantitative Analysen vernachlässigbar, da eine quantitative Abweichung von ≤ 1 % aufgrund eines Memory Effekts in differentiellen Proteomstudien deutlich unter der normalen SD einer solchen Analyse liegt. Vielmehr tangiert der Memory Effekt vor allem qualitative Analysen.

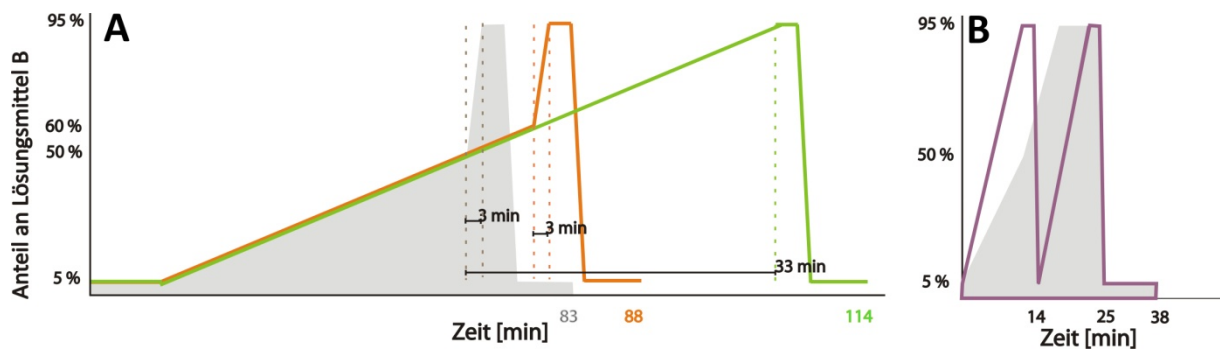


Abbildung 4.12: Zur Analyse von Memory Effekten wurden (A) unterschiedliche LC-Gradienten mit zeitlichen Variationen im Anstieg von 5-95 % LM B (84 % ACN 0,1 % TFA) sowie (B) Waschgradienten mit unterschiedlicher Anzahl an Anstiegen von 0-95 % LM B eingesetzt. Aufgrund der hohen Sensitivität von MS Instrumenten konnten selbst für niedrige Konzentrationen Memory Effekte in folgenden Blankläufen detektiert werden (Burkhart, Prensler et al., *Proteomics*, 2011).

Wie im Kapitel 4.9 in der Abbildung 4.22 zu sehen, wurde die QC Mixtur in dieser Arbeit v.a. in der LF Quantifizierung für die Kontrolle der Stabilität von RT, Signalintensitäten sowie von FWHMs eingesetzt.

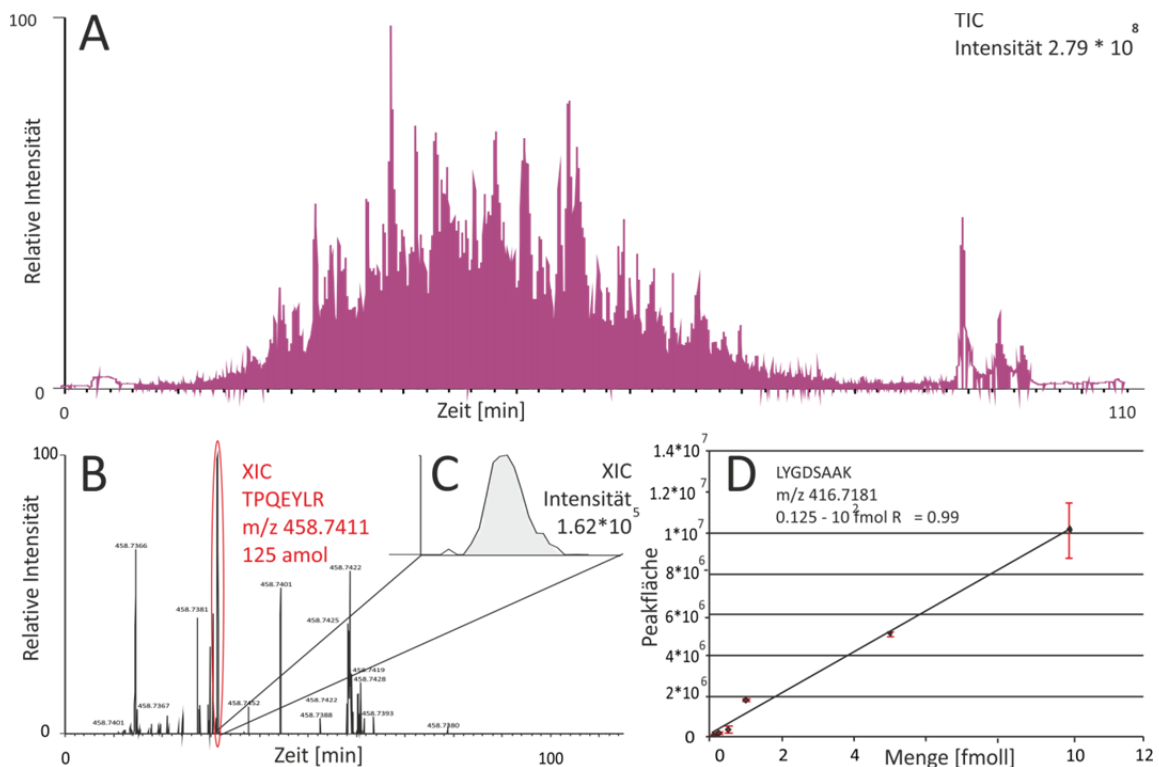


Abbildung 4.13: Bestimmung der LOD und LOQ. Ansteigende Mengen der stabil-Isotopen markierten QC Mixtur wurden zu jeweils 1 μg Thrombozytenverdau gegeben und innerhalb von 2 h in LC-MS/MS Läufen analysiert. (A) zeigt exemplarisch ein Total Ion Chromatogram (TIC) von 1 μg Thrombozytenverdau mit 0,125-62,5 fmol QC Mixtur. (B,C) Das QC Peptid TPQEYLR mit 125 amol kann im Extracted Ion Chromatogram (XIC) gut detektiert werden. (D) zeigt die hohe Linearität ($R^2 = 0,99$) der Quantifizierung des QC Peptids LYGDSAAG m/z 416,7181 in der komplexen Thrombozytenmatrix über einen dynamischen Bereich von 0,125 – 10 fmol (Burkhart, Prensler et al., *Proteomics*, 2011).

4.6 Qualitätskontrollierte Datenauswertung von qualitativen proteomischen Daten / Optimierte statistische Datenauswertung

Für die Analyse von komplexen Peptidgemischen, wie z.B. von kompletten Proteomen, durch hochauflösende und schnelle MS Instrumente ist die Automatisierung der Zuordnung von Peptidsequenzen zu Spektren (engl. peptide to spectrum match, PSM) unabdingbar [124]. Die hierfür genutzten Suchalgorithmen bewerkstelligen einen Abgleich von experimentell generierten Fragmentationsspektren mit theoretisch fragmentierten Spektren einer Datenbank.

Notwendig dafür ist zunächst die möglichst genaue Bestimmung der Peptidmasse anhand des m/z -Verhältnisses sowie der Ladung - ermittelt innerhalb des Übersichtscans. Die Peptidmasse wird herangezogen, um die Anzahl an theoretischen Peptiden, die *in silico* für den Datenbankabgleich generiert werden, auf ein Minimum einzuschränken und damit die Wahrscheinlichkeit von Zufallstreffern zu reduzieren. Die theoretischen Fragmentationsspektren dieser übereinstimmenden Peptide werden anschließend mit dem experimentellen MS/MS Spektrum abgeglichen. Hierbei können trotz hoher Massengenauigkeiten falsch-positive Identifikationen auftreten.

Die korrespondierende FDR kann mit der häufig angewendeten target/decoy Strategie geschätzt werden [70]. Ein Schwachpunkt bei der klassischen Anwendung dieser Strategie ist allerdings die fehlende Klassifizierung der Peptide in Subgruppen wie z.B. PTM tragende Peptide, die meist eine höhere FDR als unmodifizierte Peptide aufweisen und mit einer individuell berechneten FDR optimaler dargestellt werden.

Bei bekannter Probenzusammensetzung wie z.B. bei UPS, bei welchem die 48 bekannten Proteine durch experimentelle Kontaminationen wie Antikörper oder humane Keratine ergänzt wurden, kann die FDR Strategie daher als Robustheitsmaß zur Kontrolle der Datenevaluation herangezogen werden.

In dieser Arbeit wurden anhand von fünf identischen UPS LC-MS Analysen klassische Parameter der Datenauswertung wie (1) die Verwendung verschiedener Suchalgorithmen, (2) die Auflösung der FDR bei variierenden Datensatzgrößen, (3) variierende Suchparameter für Vorläufer- und Fragmentationen und auch verschiedene Modifikationen sowie (4) die Subgruppierung von Daten in Ladungszustände und Modifikationen, überprüft. Die erhaltenen Ergebnisse wurden anschließend mit der Analyse einer Probe unbekannter Zusammensetzung und höherer Komplexität - Mitochondrien der Bäckerhefe - verglichen.

Hierbei wurden identische Tendenzen festgestellt und damit UPS als aussagekräftiges Referenzproteom validiert.

Datenbanksuche bei 1% FDR		Mitochondrien	
		PSMs	Anstieg der PSMs [%]
--	Standarddatenbanksuche	873	--
(a)	Einteilung nach Ladungszuständen	928	6
(b)	Einteilung nach Ladungszuständen und Modifikationen	963	10
(c)	Einteilung nach Ladungszuständen und Modifikationen, mit optimierten	993	13,8

Tabelle 4.4: Übersicht über eine optimierte Datenbankauswertung von mitochondrialen Proben bei gleichbleibender FDR von 1 % als Qualitätskriterium.

(1) So zeigte eine Gegenüberstellung der Suchalgorithmen Mascot und OMSSA eine vergleichbare Anzahl an tryptischen Peptididentifikationen bei 1 % FDR, wobei das Ausmaß der Variabilität zwischen (a) den Suchalgorithmen und (b) den Replikaten eines Suchalgorithmus sich nicht unterschied. Dagegen konnte die Anzahl der Peptididentifikationen mit X!Tandem aufgrund eines zweiten inkorporierten multivariablen Datenabgleichs signifikant und damit die FDR zu einem niedrigeren Wert verändert werden. Beim Vereinen der Datensätze aller Suchalgorithmen zeigte sich jedoch eine erhöhte finale FDR von 3,3 %, da die Algorithmen bei richtig-positiven IDs eine starke Überlappung zeigen, wohingegen decoy-Peptide als Zufallstreffer nicht wiederholt identifiziert werden, sich statt dessen anhäufen.

(2) Weiterhin spielt die Größe des Datensatzes eine erhebliche Rolle für die Robustheit bzw. die Genauigkeit der FDR, da sie sowohl deren Auflösung als auch Fehleranfälligkeit bestimmt. Für die Validität der FDR muss der Datensatz zudem ausreichend komplex sein. So stellt eine falsch-positive (decoy) Peptididentifikation bei einer Gesamtdatensatzgröße von nur 10 Hits bereits eine FDR von 10 % dar, während bei einer Gesamtdatensatzgröße von 100 sich ein decoy Hit mit 1 % bzw. zwei decoy Hits mit 2 % FDR widerspiegeln.

(3) Von Interesse für den Routinegebrauch ist insbesondere die Qualitätskontrolle bzw. die Optimierung der Parameter für Suchalgorithmen, da hier wenig stringente Einstellungen den Einschluss von falsch-positiven target IDs sowie zu stringente Parameter den Verlust von richtig-positiven target IDs bewirken können. Die Balance zwischen beiden Extremen konnte für die genannten Datensätze, UPS und Mitochondrien, mit einer Vorläufermassentoleranz von 60 ppm und anschließendem Filtern auf 8 ppm und einer Fragmentionenmassentoleranz von 0,3 – 0,5 Da optimiert werden.

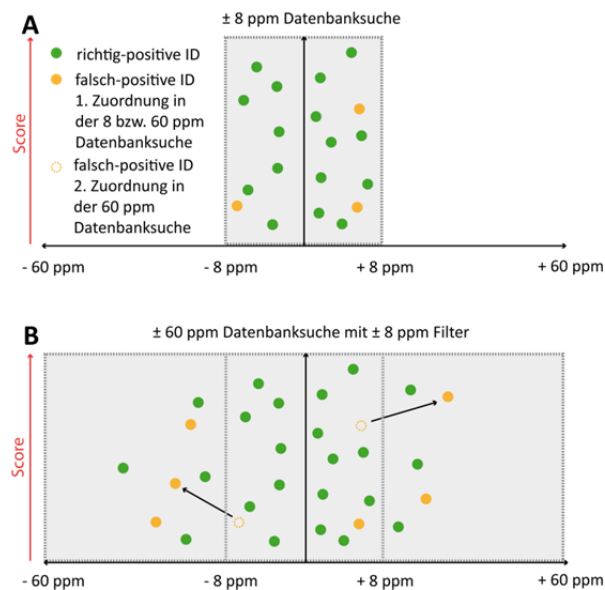


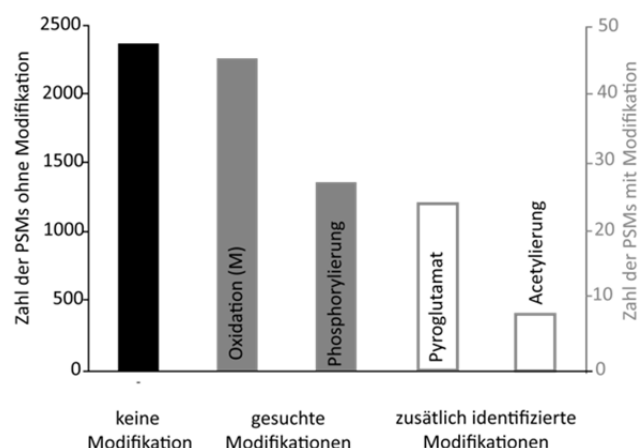
Abbildung 4.14: Hinsichtlich der Massentoleranz optimierte Datenbanksuche. (A) Während richtig-positive Identifikationen unabhängig von der Massentoleranz der Suche die beste Übereinstimmung liefern können Zufallstreffer „wandern“:

(A) Bei einer Datenbanksuche mit ± 8 ppm Massentoleranz treten richtig-positive (grün) sowie falsch-positive (orange) PSMs auf. (B) Bei einer erneuten Suche desselben Datensatzes mit ± 60 ppm Massentoleranz können für Spektren, welche einen falsch-positiven Zufallstreffer ergaben teils bessere Übereinstimmungen für andere falsch-positive PSMs mit einem höheren Score gefunden werden. Bei anschließendem Filtern auf 8 ppm können diese falsch-positiven IDs ausgeschlossen werden.

Für die Suche nach unbekanntem Modifikationen bietet der Suchalgorithmus X!Tandem einen sekundären Datenbankabgleich, in welchem für bereits identifizierte Protein IDs neben semitryptischen Peptiden auch nach zahlreichen variablen Modifikationen gesucht werden. Dieser Suchmodus muss für Mascot in einer sogenannten „Error Tolerant Search“ aktiv gewählt werden [125] und vermag auch hier zusätzliche Informationen zu liefern. So zeigte die Analyse mitochondrialer LC-MS Läufe unter Verwendung von X!Tandem neben der Identifizierung von nicht modifizierten, Met oxidierten und Ser/Thr/Tyr phosphorylierten Peptiden eine beträchtliche Anzahl weiterer Modifikationen (Pyroglutamat und Acetylierung am N-terminus [126]), die für weitere Datenbanksuchen berücksichtigt werden sollten (Abbildung 4.15).

(4) Da die Effizienz von Suchalgorithmen für Peptidspezies mit bzw. ohne PTM variieren kann, ist die Ausbildung eines peptidübergreifenden signifikanten Grenzwertes für die Identifikationen einer LC-MS Analyse nicht optimal und die Ausbildung von Subgruppierungen bei ausreichender Datensatzgröße zu bevorzugen.

Abbildung 4.15: Ein ergänzender multivariabler Datenbankabgleich ermöglicht die Identifizierung von unerwarteten Proteinmodifikationen. Im Beispiel gezeigt ist die ergänzende zweite Suche, durchgeführt mit X!Tandem, die neben den primär gesuchten Modifikationen, Oxidation (M) und Phosphorylierung (S/T/Y) im sekundären Datenbankabgleich weiterhin die Modifikationen Pyroglutamat und Acetylierung identifiziert (Vaudel, Burkhardt et al., *Proteomics*, 2011).



Bei Einteilung der UPS- und Mitochondrienspektren in Untergruppen bezüglich (a) Ladungszustand, (b) Ladungszustand und Modifikationen und (c) zusätzliche Kombination mit optimierten Vorläufertoleranzen, konnte, wie in Tabelle 4.4 dargestellt, ein Anstieg der PSMs um fast 14 % bei gleichbleibender Gesamt FDR erzielt werden.

4.7 Charakterisierung des thrombozytären Proteoms

Ziel der Darstellung des thrombozytären Proteoms war die reproduzierbare, valide und umfangreiche Analyse von Thrombozyten als Grundlage für fortführende krankheitsrelevante Studien. Aufgrund ihres anukleären Charakters und ihrer Entstehung als Abschnürungen ihrer Vorläuferzellen [85], den Megakaryozyten, ist die qualitative wie auch quantitative Zusammensetzung von Thrombozyten von großem Interesse. In der Literatur werden Proteinvariabilitäten mit einem Faktor von etwa zwei für Thrombozyten [127] und auch Resistenzen gegenüber einer Thrombozyten-orientierten medikamentösen Therapie beschrieben, für welche Unterschiede in der Proteinzusammensetzung als Ursache vermutet werden [128, 129]. Daher ist für eine quantitative proteomische Analyse vor allem die reproduzierbare Isolierung von Thrombozyten hoher Reinheit aus Frischblut eine Herausforderung und eine Grundvoraussetzung. Bedingt durch die Aufreinigung aus Frischblut ist davon auszugehen, dass mit steigender Abdeckung des thrombozytären Proteoms auch der Einfluss von Kontaminationen aus dem Blut zunimmt und den Blick auf die qualitative wie quantitative Zusammensetzung des Thrombozytenproteoms verzerrt.

Aus diesem Grund wurde ein optimiertes Aufreinigungsprotokoll verwendet, um Thrombozyten mit hoher Reinheit und geringer Kontamination durch Erythrozyten und Leukozyten zu isolieren (siehe Kapitel 4.1). Mit Hilfe dieses Protokolls wurden verschiedene Aufgabenstellungen adressiert: (a) es sollten möglichst viele thrombozytäre Proteine identifiziert werden, (b) es sollte für möglichst viele Proteine die Kopienzahl pro Thrombozyt abgeschätzt werden, (c) es sollte die Intra- sowie (d) inter-Spender-Varianz des thrombozytären Proteoms charakterisiert werden.

Hierzu wurde eine aufwendige Analysestrategie durchgeführt, welche in Abbildung 4.16 zusammengefasst ist.

4.7.1 Abschätzung absoluter Kopienzahlen für Thrombozytenproteine

Absolute Kopienzahlen von 3718 in Thrombozyten identifizierten Proteinen bzw. Proteingruppen wurden basierend auf dem NSAF Index ermittelt. Hierzu wurden je 50 µg tryptischer Verdau (nicht markiert) aller vier Spender zu einer kombinierten Probe vereint, da aufgrund der iTRAQ Daten von einer nur geringen Varianz des thrombozytären Proteoms ausgegangen werden konnte. Diese Probe wurde mittels IEF in 24 Fraktionen aufgetrennt und anschließend die erhaltenen Fraktionen per LC-MS auf einem LTQ Orbitrap Velos CID TOP10 analysiert.

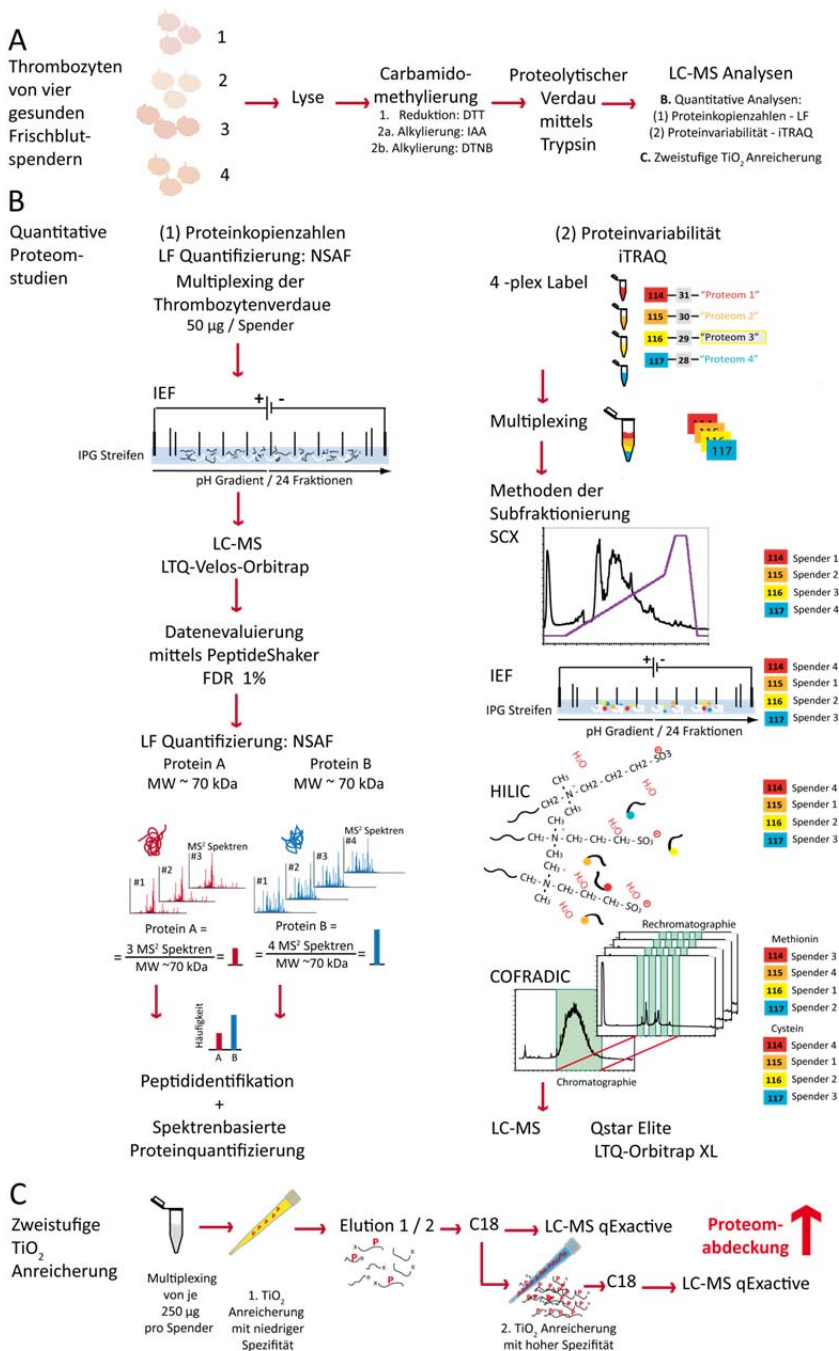


Abbildung 4.16: Übersicht über die analytische Strategie zur Charakterisierung des thrombozytären Proteoms. (A) Thrombozyten wurden aus Frischblut isoliert, lysiert, carbamidomethyliert und mit Trypsin für die folgenden quantitativen und phosphospezifischen Studien verdaut. Die Fraktionen wurden mittels LC-MS/MS auf einer LTQ Orbitrap Velos analysiert und auf Basis des NSAF Indexes und anschließender Korrelation mit Literaturwerten absolut quantifiziert. (2) Zur Analyse der Proteinvariabilität in Thrombozyten wurden jeweils 100 µg pro Spender mit den iTRAQ-Reagenzien 114, 115, 116 und 117 markiert und nach Subfraktionierung mit SCX, IEF, HILIC bzw. COFRADIC mit einem Orbitrap XL und Qstar Elite Massenspektrometer analysiert. (C) Für eine verbesserte Proteomabdeckung wurden 250 µg Thrombozyten aller Spender vereint und einer zweistufigen TiO₂ Anreicherung niedriger und hoher Spezifität mit anschließender Analyse auf einem qExactive Massenspektrometer unterzogen.

Die erhaltenen Daten wurden dann verwendet, um mit Hilfe der PeptideShaker Software die NSAF Werte für alle identifizierten Proteine (bei 0,8 % FDR) zu bestimmen. Ein Problem bei dieser Art der Quantifizierung stellen Peptide dar, die nicht eindeutig nur einem Protein zuordenbar sind – gerade im Falle von Proteinisoformen werden häufig keine eindeutigen Peptide identifiziert, sodass Proteingruppen wie z.B. α -Aktin bestehend aus den Isoformen α -Aktin 1 (Herzmuskel, P68032), α -Aktin 1 (Skelettmuskel, P68133), α -Aktin 2 (glatte Muskulatur der Aorta, P62736) und α -Aktin 3 (glatte Muskulatur, P63267) gebildet und als solche in die Proteinliste aufgenommen wurden. Nicht eindeutige Peptide, die unabhängigen Proteinen (d.h. keine Isoformen oder Homologe) zugeordnet werden konnten, wurden von der NSAF Auswertung ausgeschlossen.

Wie in Abbildung 4.17 dargestellt wurden 24 in der Literatur gefundene Werte für Proteinkopienzahlen in Thrombozyten mit den MS Daten korreliert, um die Abschätzung aller Proteinkopienzahlen per Regression zu ermöglichen. Für die 24 Referenzproteine zeigt sich eine gute Korrelation zwischen den geschätzten Kopienzahlen und den Literaturwerten, mit einem Bestimmtheitsmaß von $R^2 = 0,90$ (siehe Abbildung 4.17). Anhand von CD9 zeigt sich allerdings, dass einzelne Schätzwerte für Proteinkopienzahlen durchaus fehlerbehaftet sein können. Das ist vor allem auf die vier Transmembrandomänen von CD9, bei einer relativ kurzen Sequenzlänge von 228 AS, und der daraus resultierenden Generierung von nur zwei voll tryptischen Peptiden zwischen 7 und 20 AS (und drei weiteren mit $AS \geq 25$) zurückzuführen. Dies erklärt die Abweichung des experimentellen Wertes von 8.000 Kopien gegenüber dem Literaturwert von 49.000 Kopien pro Thrombozyt. [86, 130-144]

Differenzen zwischen Literatur und experimentellen Werten können weiterhin durch das gehäufte Auftreten von PTMs erklärt werden, welche sowohl den proteolytischen Verdau sowie die MS Analyse stören können und eventuell aufgrund der verwendeten Suchparameter erst gar nicht identifiziert werden. Dies kann zu einer Beeinflussung „spectral counting“ basierter Quantifizierungsmethoden wie NSAF führen. Ausgehend von den erhaltenen Kopienzahlen und den zugrundeliegenden MWs der jeweiligen Proteine wurde eine näherungsweise Proteingesamtmasse von 1,5 mg pro 10^9 Thrombozyten errechnet, die

gut mit dem typischen Laborwert von $1,8 \pm 0,2$ mg übereinstimmt.

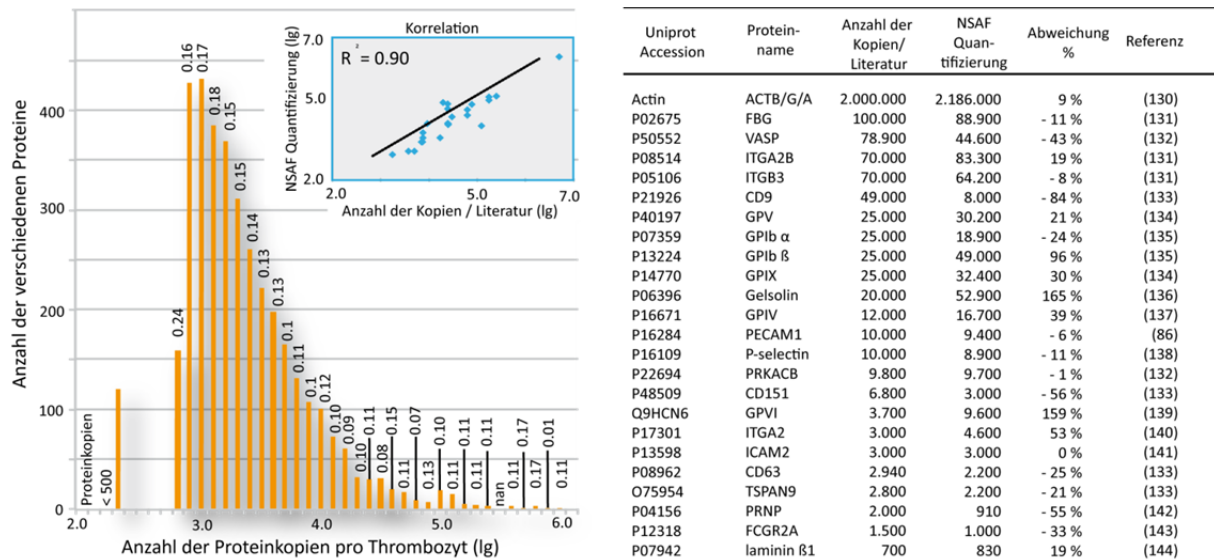


Abbildung 4.17: Der Vergleich von 24 Literaturwerten mit experimentell (NSAF) evaluierten Proteinkopienzahlen pro Thrombozyt (logarithmische Auftragung) zeigte eine gute Korrelation ($R^2 = 0,90$) über vier Größenordnungen. Angezeigt sind weiterhin die iTRAQ SDs für Proteinkopieneinheiten mit einer Weite von 0,1 (lg).

4.7.2 Quantitative Erfassung der Inter-Donor Proteinvariabilität in Thrombozyten mittels iTRAQ

Zur Analyse der Proteinvariabilität in Thrombozyten wurden für eine verbesserte Auflösung des Proteoms zur RP-LC-MS/MS orthogonale Präfraktionierungsstrategien angewendet und erhaltene Fraktionen mittels MS auf zwei verschiedenen Systemen analysiert. Die angewandte mehrschichtige Fraktionierungsstrategie wurde im Laufe dieser Arbeit, wie unter 4.7.3 dargestellt, durch den Einsatz von hochsensitiven und hochauflösenden MS Instrumenten mit Fragmentierungsfrequenzen von bis zu 12 Hertz abgelöst. Somit ist mit Hilfe einer *state-of-the-art* Ausstattung und optimaler Performance der einzelnen Arbeitsschritte - von der Probenvorbereitung, über die LC-MS Auftrennung / Akquisition bis hin zur Datenanalyse - ein wesentlich geringerer Aufwand zur Fraktionierung und Charakterisierung eines „kompletten“ Proteoms nötig.

Die Regulation der 1.900 quantifizierten Proteine wurde basierend auf den Standardabweichungen der iTRAQ-Proteinratios (engl. standard deviation, SD) evaluiert. Für zwischen den vier Spendern nicht regulierte Proteine würde man hier ein ideales Verhältnis von 1:1:1:1 erwarten. Tatsächlich ergab die Ermittlung der SDs dieser vier Werte pro Protein

und über alle 1.900 quantifizierten Proteine einen SD Median (SD_{median}) von 0,14. Da, wie in dieser Arbeit evaluiert [106], bei iTRAQ Studien von einer technischen Varianz von ca. 20 % ausgegangen werden kann, wurden all jene Proteine mit einer $SD > 2 \times SD_{\text{median}}$, also mit einer SD über 0,28 als potentiell reguliert eingestuft. Dieser Wert liegt deutlich unter der normalerweise erwarteten biologischen Varianz zwischen Menschen.

85 % der quantifizierten Proteine liegen unterhalb dieses Grenzwertes und werden daher als nicht bzw. nur sehr geringfügig differentiell angesehen. Da nur 25 % der potentiell regulierten Proteine niedrig abundant sind, d.h. weniger als 1.000 Kopien pro Thrombozyt besitzen, und ebenfalls 25 % der mehr als 3.600 quantifizierten Proteine unter 1.000 Kopien pro Thrombozyt liegen, kann die Hypothese der Regulation ausschließlich niedrig abundanter und damit schwer detektier- sowie quantifizierbarer Peptide vernachlässigt werden. Um den Einfluss von einzelnen Ausreißern auf die Analyse zu evaluieren, wurde die SD Analyse zudem unter Ausschluss der jeweils stärksten Abweichung mit nur drei der vier Proteinratios (SD_3 Patienten) durchgeführt und der SD_{median} von 0,14 als Grenzwert gesetzt. Hier zeigten 171 Proteine (9 %) eine potentielle Regulation.

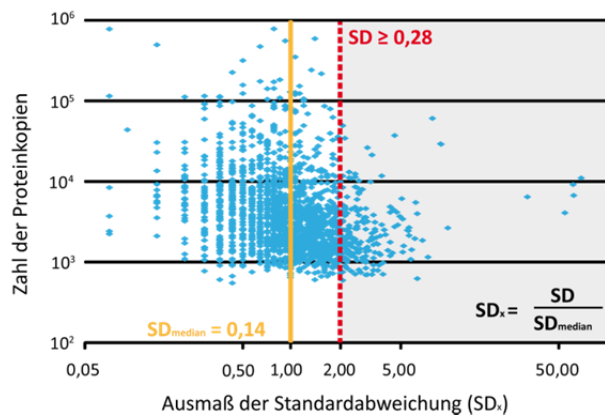


Abbildung 4.18: Zur Charakterisierung der Proteinvariabilität wurden in einem iTRAQ Experiment die individuellen SDs der iTRAQ Ratios durch den Median der SD des gesamten Datensatzes ($SD_{\text{median}} = 0,14$) dividiert. 85 % der 1900 mit iTRAQ quantifizierten Proteine wiesen eine SD von niedriger als 0,28 (gestrichelte rote Linie) auf und zeigten damit keine Variation zwischen den Spendern.

Die Gruppe der differentiellen Proteine umfasst u.a. potentielle bzw. bekannte Kontaminanten wie Hämoglobin, MHC und Apolipoproteine, die auf geringfügige Unterschiede in der Thrombozytenisolation zurückzuführen sind. Hingegen wiesen Proteine, die zentralen Funktionen in Thrombozyten zugeordnet werden können, unabhängig von ihrer Kopienzahl pro Thrombozyt, niedrige SDs auf, die höchstwahrscheinlich im technischen Fehler der iTRAQ basierten Proteinquantifizierung begründet sind: GPIb α 0,05; GPIb β 0,11; GPIV 0,16; GPVI 0,24; GPIX 0,10; PAR4 0,25; Platelet factor 4 0,13; STIM1 0,07; P2X1 0,17; VASP 0,11; LAP 0,10; PKA 0,20; PKG 0,10; Aktin 0,10.

4.7.3 Erfassung der Variabilität über Kopiezahl

Alternativ wurde die Proteinvariabilität zwischen den vier Spendern mittels NSAF evaluiert. Hierzu wurden die Proben sowohl individuell als auch als Mix der vier Spender, wie in 3.9.1. beschrieben, mittels SCX systematisch in sieben Fraktionen aufgetrennt. Die C18 aufgereinigten Fraktionen wurden anschließend innerhalb einer 4 h LC Auftrennung auf einer 50 cm x 75 µm C18 HPLC Säule in Kombination mit einem hochauflösenden und massengenauen qExactive Massenspektrometer analysiert. Analog zur iTRAQ Auswertung wurde auch für die LF Ermittlung der Proteinvariabilität die SD zur Charakterisierung der Regulation herangezogen. Hierzu wurden die ermittelten NSAF Werte der individuellen Proben zunächst auf den $NSAF_{Median}$ des Spendermixes normalisiert. 2457 Proteine konnten in allen vier Spendern mit einer $SD_{median} = 24,8 \%$ quantifiziert werden, von welchen 288 Proteine, entsprechend 11.7 %, das Kriterium für die Regulation, $SD > 2 \times SD_{median}$, erfüllten. 28 der in der NSAF basierten Untersuchung ermittelten, regulierten, Proteine wurden auch in der iTRAQ Auswertung als reguliert eingestuft (Tabelle 4.5).

Protein	Proteinname	Kopiezahl pro Thrombozyt	iTRAQ SD	LF SD
A6NDG6	PGP	2.100	0,32	0,63
O00461	GOLI4	700	0,46	0,52
O00571	DDX3X	3.100	0,19	0,50
O43264	ZW10	1.300	0,35	1,03
O60888	CUTA	1.600	0,32	0,54
O94874	UFL1	2.500	0,3	1,58
P00915	CAH1	2.500	0,77	1,09
P02774	VTDB	2.700	0,79	0,50
P11047	LAMC1	800	0,3	0,51
P16930	FAAA	7.100	0,38	0,56
P27701	CD82	1.400	0,42	0,83
P27986	P85A	2.000	0,31	0,70
P48163	MAOX	1.100	0,81	1,20
P49247	RPIA	2.600	0,48	0,92
P63241	IF5A1	<500	0,29	0,51
P69905	HBA	29.200	1,26	0,70
Q12841	FSTL1	2.900	0,67	0,76
Q14112	NID2	2.000	0,29	0,59
Q14152	EIF3A	1.200	0,6	0,56
Q15018	F175B	1.300	0,3	0,64
Q15149	PLEC	4.500	0,63	0,62
Q3ZQA7	VMA21	<500	0,3	0,57
Q53FA7	QORX	1.400	0,38	0,84
Q5BJH7	YIF1B	1.200	0,36	0,52
Q8N4L2	TM55A	4.000	0,3	0,53
Q96DZ9	CKLF5	2.600	0,36	0,65
Q96HY6	DDRGK	1.800	0,6	0,69
Q9UHQ4	BAP29	1.500	0,43	0,60

Tabelle 4.5: Thrombozytäre Proteine, die im quantitativen Vergleich sowohl in der iTRAQ Studie als auch in der LF Studie das Kriterium für die Regulation, $SD > 2 \times SD_{median}$, erfüllten.

4.7.4 Intra-Spender Varianz

Ergänzend wurden zur Analyse der Inter-Spender Proteinvariabilität drei unabhängige Blutproben eines Spenders untersucht. Für jede Blutspende wurden jeweils zwei technische Replikate unabhängig prozessiert und anschließend ohne jegliche Präfraktionierung mittels nano-LC-MS analysiert. Die fehlende Fraktionierung führte zu signifikant weniger MS/MS Spektren, die letztlich die NSAF Quantifizierung von 1.505 Proteinen in den drei Proben und den entsprechenden technischen Replikaten erlaubten ($SD_{\text{median}} = 12,2 \%$). 85 % der Proteine zeigten keine Variation in der NSAF Auswertung ($SD < 24,4 \%$). Nur ein Protein, DDX3X, konnte sowohl in der Intra-Spender als auch in der Inter-Spender Studie als reguliert eingestuft werden.

Die niedrigen SD der technischen Replikate für die drei unabhängigen Blutproben (12,7 %, 14,8 % und 13,8 %) zeigen die übergreifende Robustheit des kompletten LF Workflows auf.

4.7.5 Globale Phosphorylierungsstudie zur Verbesserung der Proteomabdeckung

Die Regulation von Thrombozyten erfolgt aufgrund ihres anukleären Charakters zu einem großen Anteil durch Phosphorylierung, welche an Proteinen über einen großen dynamischen Bereich zu finden ist. Während phosphorylierte Peptide nur substöchiometrisch neben nicht modifizierten Spezies vertreten sind, ermöglicht deren selektive Anreicherung aufgrund der enormen Verringerung der Probenkomplexität die Identifizierung von Proteinen eines großen dynamischen Bereichs.

Durch die zweistufige TiO_2 Anreicherung mit niedriger und anschließend hoher Spezifität konnten 2.900 Phosphorylierungsstellen auf 2.428 Peptiden identifiziert werden; dies ermöglichte bei einer FDR von 1 % die zusätzliche Charakterisierung von 263 Proteinen, die nicht in der globalen Proteomanalyse identifiziert wurden. 9,1 % der Phosphopeptide konnten ausschließlich im ersten Anreicherungsschritt identifiziert werden. Unterdessen konnte die zweite spezifischere Anreicherung 59,3 % zusätzliche Peptide identifizieren. Vor allem in der unspezifischen, ersten Anreicherung konnten neben phosphorylierten Peptiden auch zahlreiche unphosphorylierte Spezies niedrig abundanter Proteine wie P2Y12, das mit 600 Kopien pro Zelle angenommen wird [145], gefunden werden. P2Y12 wurde über die globale Analyse, höchstwahrscheinlich aufgrund seines 7-TMD Charakters in Kombination mit der niedrigen Abundanz, nicht identifiziert.

4.7.6 Abgleich der Proteomdaten mit RNA Daten

Die in der Proteomanalyse quantifizierten 3.718 Proteine wurden weiterhin mit Transkriptom- und SAGE RNA-Daten abgeglichen. Es wurde ein kombinierter Datensatz aus 3.598 Proteom, 6.202 RNAseq und 2.737 SAGE Einträgen erstellt. Es zeigte sich, dass nur 8% der Proteine in allen drei Datensätzen vorkamen, aber 12%, 38% und 9% ausschließlich in den Proteom-, RNAseq bzw. SAGE Daten zu finden sind. Sortiert nach Häufigkeit zeigte das 20. Perzentil der RNAseq Daten eine Übereinstimmung mit 71 % der Proteomdaten, während das 20. Perzentil der Proteomeinträge zu 85 % in den RNAseq Daten gefunden werden konnte. Jedoch konnten für die abundantesten Proteine / Transkripte keine Korrelation zwischen dem näherungsweise bestimmten Ausmaß der Genexpression (reads per kilobase per exon model per million mapped reads, RPKM) und der Proteinkopienzahl festgestellt werden ($R^2 < 0,1$).

4.7.7 Abgleich mit anderen Proteomdaten

Bei der Charakterisierung des Thrombozytenproteoms ist das Ausmaß an Kontaminationen durch andere Blutkomponenten ein wichtiger Faktor. Zum einen ist eine quantitative Entfernung von Leukozyten und Erythrozyten quasi unmöglich, zum anderen haben Thrombozyten aufgrund des offenen kanalikulären Systems eine schwammartige Struktur, welche massiv mit Plasma durchtränkt ist. Aus diesem Grund wurde das Thrombozytenproteom mit weiteren Proteomstudien von Plasma und Leukozyten verglichen.

Der Abgleich mit Kopienzahlen der bisher umfangreichsten, quantitativen Plasmaproteomstudie [146] zeigte, dass 149 der 150 abundantesten Plasmaproteine auch in den Thrombozytendaten gefunden werden. Serumalbumin ist das mit großem Abstand abundanteste Protein im Blutplasma und hat dort eine etablierte Konzentration von 50 mg/mL, was ca. 450.000 Kopien Serumalbumin pro fL Plasma entspricht. In dieser Arbeit konnten 52.000 Kopien Serumalbumin pro Thrombozyt nachgewiesen werden, was folglich einer Kontamination durch ca. 0,1 fL Plasma pro Thrombozyt entspricht. Da das mittlere Volumen eines Thrombozyten $9,7 \pm 0,5$ fL beträgt, ist davon auszugehen, dass ca. 1 Volumenprozent Plasmakontamination vorhanden ist. Analog lässt sich eine Kontamination von 1,5 Volumenprozent anhand des zweithäufigsten in Blutplasma auftretenden Proteins Transthyretin (0,77 mg/mL; 29.000 Kopien pro fL Plasma) bei einer Kopienzahl von 4.500 pro

Thrombozyt berechnen. Diese Zahlen unterstreichen den hohen Reinheitsgrad der analysierten Thrombozyten.

Hämoglobin, das mit Abstand abundanteste Protein in Erythrozyten (280.000.000 Kopien pro Erythrozyt), wurde mit ca. 61.000 Kopien pro Thrombozyt geschätzt. Dies deutet auf eine Kontamination von einem Erythrozyt pro 5.000 Thrombozyten hin, wobei die Verunreinigung mit Hämoglobin zumindest teilweise auf Hämolyse während der Thrombozytenisolation zurückzuführen ist. Die Ergebnisse werden durch die Abwesenheit des abundantesten Erythrozytenmembranproteins Band 3 (1.000.000 Kopien pro Erythrozyt) [147] gestützt.

4.8 Charakterisierung des mitochondrialen IMS Proteoms mittels SILAC

Während für die globale Charakterisierung von Zellorganellen in der Literatur Methoden beschrieben sind, welche Dichtegradientenzentrifugation mit anschließender quantitativer Proteomanalyse der erhaltenen Fraktionen kombinieren [148, 149], stellt die Analyse von Subkompartimenten eine neue Herausforderung dar. Bei Mitochondrien konnte bisher v.a. die äußere Membran von *S. cerevisiae* [150] und anderen Organismen [151] detailliert charakterisiert werden. Der Intermembranraum hingegen stellt das kleinste und am schwersten aufzureinigende Kompartiment in Mitochondrien dar. Er spielt eine essentielle Rolle im Transport und Modifizierung von Proteinen und Lipiden, in der Organisation der Atmungskettenkomplexe sowie in Redoxprozessen und der Apoptose.

Für die Darstellung des mitochondrialen IMS Proteoms (*S. cerevisiae*) wurde mit Hilfe von rekombinantem Säuger-BAX ein konservierter biochemischer Mechanismus ausgenutzt. BAX, ein der Bcl2-Familie zugeordnetes Protein (Bcl2-associated X protein), ist eine der Schlüsselkomponenten zur Initialisierung der Apoptose, indem es in die OM inseriert, was schließlich zur Ausbildung von Poren in der OM und damit zum Verlust der Impermeabilität der OM für Proteine führt [152]. Neben IMS Proteinen werden dabei ebenfalls geringe Mengen an Matrixproteinen in den Mitochondrienüberstand (engl. supernatant, SN) freigesetzt. Diese konnten jedoch mittels eines SILAC Ansatzes, in welchem zwei in Abbildung 3.1 dargestellten Ratios bestimmt wurden, differenziert werden: (1) Es wurde der Überstand nach BAX-Behandlung (SN_{BAX}) mit dem kompletten löslichen Mitochondrienlysate (Tot) verglichen. Potentielle IMS Proteine sollten hier ein SN_{BAX}/Tot von etwa 1 haben, während Matrixproteine eher kleinere Ratios haben sollten, da sie weniger effizient nach BAX-Behandlung in den Überstand freigesetzt werden. (2) Außerdem wurde SN_{BAX} mit

mitochondrialen SN ohne BAX Behandlung (SN_{Con}) verglichen. Hier sollten vor allem IMS Proteine hohe SN_{BAX}/SN_{Con} Ratios annehmen.

Für die Charakterisierung der im IMS lokalisierten Proteine wurden zunächst die SILAC Verhältnisse für 29 von 31 bekannten löslichen IMS Proteinen als auch von 59 bekannten Matrixproteinen bestimmt. IMS Proteine zeigten wie erwartet hohe Werte für SN_{BAX}/SN_{Con} (durchschnittlich 20) und für SN_{BAX}/Tot (durchschnittlich 0.74), während Matrixproteine signifikant niedrigere SILAC Verhältnisse (für SN_{BAX}/SN_{Con} 2.5 und für SN_{BAX}/Tot 0.08) aufwiesen (siehe Anhang 12.6 und 12.7).

Proteine, die in mindestens drei von vier Replikaten für SN_{BAX}/SN_{Con} ein Minimum von 7 und für SN_{BAX}/Tot ein Minimum von 0.5 erfüllten, wurden als neue potentielle IMS Proteine angesehen. Diese Kriterien konnten von 20 Proteinen erfüllt werden, wovon zehn bislang nicht in Mitochondrien lokalisiert waren, und weitere der Innenmembran zugeordnet waren (Tabelle 4.5). Drei Proteine sind durch die Open Reading Frames Ybl107c, Ybr056w und Ymr244c-a kodiert und konnten bislang keiner Funktion zugeordnet werden.

Potentiell neue IMS-Kandidaten wurden in biochemischen Verfahren, wie unter 4.8.1 beschrieben, von Dr. Frederike-Nora Vögtle (Arbeitsgruppe Prof. Dr. Chris Meisinger, Institut für Biochemie und Molekulare Biologie, Freiburg) validiert.

Um eine Verfälschung der Ergebnisse durch die verwendeten Label auszuschließen, wurden schwere und „leichte“ SILAC Label zwischen den vier biologischen Replikaten variiert und auch die Art der stabil Isotopen Markierung geändert: $Lys^{13}C_6$ und $Arg^{13}C_6$ (beide +6.02 Da) oder $Lys^{13}C_6^{15}N_2$ (+8.01 Da) und $Arg^{13}C_6^{15}N_4$ (+10.01 Da).

4.8.1 Validierung neuer IMS Proteine

Zur Verifizierung der potentiell neuen IMS Proteine wurde eine Reihe biochemischer Experimente von Frau Dr. Nora Vögtle durchgeführt. Dazu gehörten BAX-Release Studien mit spezifischen Antikörpern, der *in organello* Import von radioaktiv markierten Präproteinen in Kombination mit Manipulation des mitochondrialen Transmembranpotentials, sowie die Zugänglichkeit für den Proteinase K Verdau nach osmotischem Schwellen. Die biochemischen Studien erlaubten die Validierung von 15 IMS Proteinen, während fünf weitere potentielle Kandidaten aufgrund des Fehlens von Antikörpern und der wenig effizienten Herstellung von radioaktiv markierten Präproteinen nicht biochemisch verifiziert werden konnten.

Tabelle 4.5: Übersicht über die in der Studie neu lokalisierten mitochondrialen IMS Proteine. Dargestellt sind die Mittelwerte der evaluierten SILAC Verhältnisse der 4 biologischen Replikate. Die Lokalisation wurde entsprechend der SGD entnommen.

ORF	Genname	Genannotation	Lokalisation	Molekulargewicht (kDa)	Sequenzabdeckung (%)	Twin C(X) _n C Motiv	SN _{BAX} /SN _{Con}	SN _{BAX} /Tot	IMS Lokalisation biochemisch validiert
YBL107C	Mic23	Putative protein of unknown function	Zytosol	23,0	55,1	CX ₁₄ C, CX ₁₃ C	33,9	0,8	x
YDR155C	Cpr1	Cytoplasmic peptidyl-prolyl cis-trans isomerase	Zytosol, Mitochondrien	17,4	41,4		8,2	0,9	x
YDR353W	Trr1	Cytoplasmic thioredoxin reductase	Zytosol	34,2	42,0		19,7	0,6	x
YER057C	Hmf1	Member of the p14.5 protein family	Zytosol, Nukleus	13,9	57,4		20,1	0,5	x
YFR033C	Qcr6	Subunit 6 of the ubiquinol cytochrome-c reductase complex	Integral / IM	17,3	32,0		29,8	0,5	x
YIR037W	Hyr1	Thiol peroxidase	intrazellulär	18,6	41,1		7,3	0,6	x
YJL066C	Mpm1	Mitochondrial membrane protein of unknown function	mitochondriale Membran	28,5	65,9		24,6	0,5	x
YJR034W	Pet191	Protein required for assembly of cytochrome c oxidase	Integral / IM	12,4	57,4	twin CX ₉ C	9,3	0,7	x
YKL152C	Gpm1	Tetrameric phosphoglycerate mutase	Zytosol, Mitochondrien	27,6	65,9		45,4	0,7	x
YLR038C	Cox12	subunit VIB of cytochrome c oxidase	Integral / IM	9,8	79,5	CX ₉ C, CX ₁₀ C	14,9	0,7	x
YLR043C	Trx1	Cytoplasmic thioredoxin	Zytosol	11,2	54,4		29,0	0,9	x
YMR244C-A	Coa6	Putative protein of unknown function	Zytosol, Nukleus	12,4	32,7	CX ₉ C, CX ₁₀ C	7,8	0,7	x
YNL036W	Nce103	Carbonic anhydrase	Zytosol, Nukleus	24,9	19,9	twin CX ₈ C	14,4	0,6	x
YOL143C	Rib4	Lumazine synthase	Zytosol, Nukleus	18,6	26,0		52,3	0,8	x
YOR090C	Ptc5	Mitochondrial type 2C protein phosphatase	Mitochondrien	63,7	35,1		14,6	0,6	x
YBR035C	Pdx3	Pyridoxine phosphate oxidase	unbekannt	26,9	25,0		20,0	0,6	
YBR056W	YBR056W	Putative protein of unknown function	Zytosol	57,8	6,6		7,0	0,6	
YIL160C	Pot1	3-ketoacyl-CoA thiolase	Peroxisom	44,7	38,4		7,0	0,6	
YPL132W	Cox11	Protein required for delivery of copper to Cox1 subunit	Integral / IM	34,0	16,3		24,4	0,8	
YMR228W	Mtf1	Mitochondrial RNA polymerase specificity factor	Mitochondrien	39,7	10,6		11,8	1,0	

Wie in Abbildung 4.19 A dargestellt konnte für zehn Proteine die Freisetzung nach BAX-Release mit spezifischen Antikörpern nachgewiesen werden. Interessanterweise waren darunter auch die Proteine Qcr6 und Ptc5, die bislang der IM zugeordnet wurden. Für weiterführende Experimente zur Bestimmung der genauen Lokalisation der Proteine wurden durch osmotisches Schwellen Mitoplasten erzeugt, wodurch die Proteine des IMS freigesetzt wurden. Proteine der OM, des IMS und der IM sind somit einem Proteinase K Verdau

zugänglich, während Matrixproteine von der intakten IM umschlossen sind und daher nicht verdaut werden.

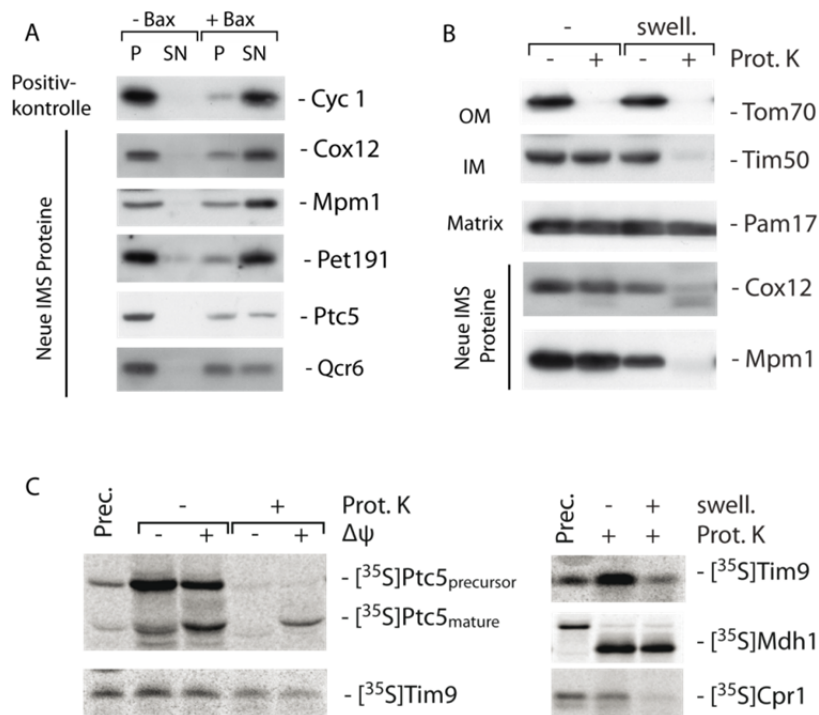


Abbildung 4.19: Übersicht über die Kombination biochemischer Experimente zur Validierung neuer IMS Proteine. (A) Mittels spezifischer Antikörper konnten IMS Proteine analog zur Positivkontrolle Cyc1 nach BAX Freisetzung visualisiert werden. (B) Das osmotische Schwellen erlaubt durch Kombination mit Proteinase K Verdau die Lokalisation der Proteine in Mitochondrien: Während Proteine der OM, wie Tom70, ohne Schwellen für den Proteinase K Verdau zugänglich sind, können Proteine der IM, wie Tim50, erst nach Schwellen verdaut werden. Matrixproteine, die selbst nach Schwellen in den Mitoplasten verbleiben, sind für den Proteinase K Verdau nicht zugänglich. IMS Proteine, hier Cox12 und Mpm1, werden durch das Schwellen freigesetzt und verdaut. (C) Der *in organello* Import radioaktiv markierter Präproteine erlaubt sowohl eine Aussage zur Lokalisation als auch zum Import. (Vögtle, Burkhart et al., MCP 2012).

Für elf neue IMS Proteine konnten radioaktiv markierte Präproteine durch *in vitro* Transkription/ Translation hergestellt und *in organello* importiert werden. Proteinase K Verdaustudien, sowohl an Mitochondrien als auch an korrespondierenden Mitoplasten, verifizierten deren Lokalisation im IMS. Weiterhin lieferte die Kontrolle des Membranpotentials $\Delta\psi$ nähere Informationen zum Importweg der Proteine.

Für zehn Proteine, die zusätzlich sensitiv für den Proteinase K Verdau bei Mitoplasten waren, konnte eine von Präsequenzen und Membranpotential unabhängige Lokalisation im IMS definiert werden. Ptc5 hingegen unterliegt einem Membranpotential $\Delta\psi$ abhängigen Import, was darauf schließen ließ, dass Ptc5 dem Präsequenzimportpfad folgt und von der IM

lokalisierten Inner Membrane Protease (IMP) prozessiert wird. Dies konnte in einem IMP Knock-out Experiment verifiziert werden.

Alternativ wurden vier neue Substrate des MIA Importpfads (Pet191, Cox12, Ybl107c, Ymr244c-a) charakterisiert, von welchen ausschließlich Pet191 die für MIA Substrate typischen, wiederkehrenden Cystein-reiche Motive, CX₃C und CX₉C, trägt. Hingegen konnten für die Proteine Cox12, Ybl107c und Ymr244c-a untypische CX_nC Motive festgestellt werden, deren Zugehörigkeit zum MIA Importweg - analog zu Pet191 - mittels thermosensitiven Mutanten *mia-40e* verifiziert wurde. Darüber hinaus konnte Ymr224c-a als Assemblierungsfaktor in der Organisation des Cytochrom C Oxidase Komplexes charakterisiert werden.

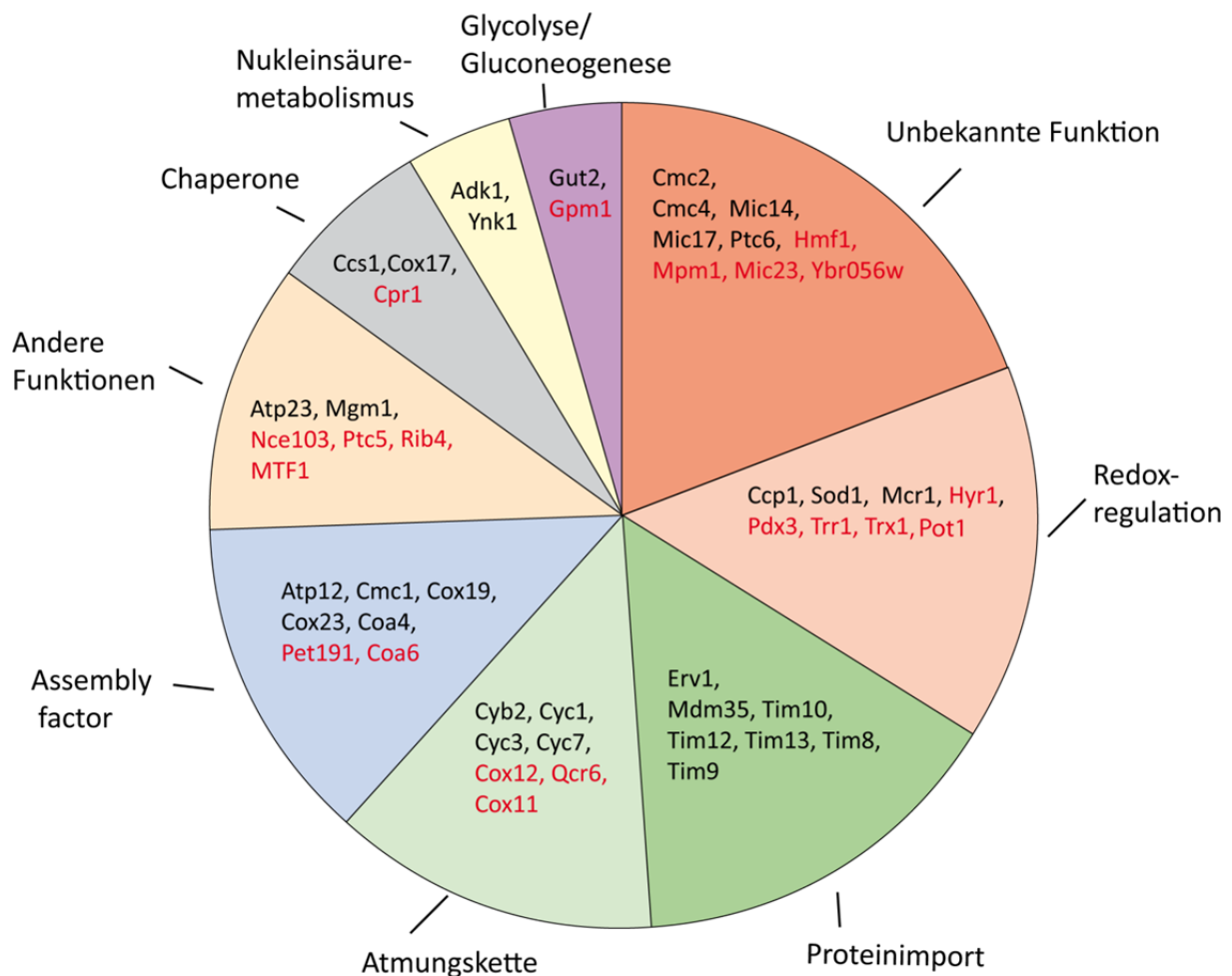


Abbildung 4.20: Übersicht über die funktionale Einordnung der 51 IMS Proteine in *S. cerevisiae*. Neu charakterisierte IMS Proteine sind rot dargestellt.

Mit der Analyse des IMS Proteoms konnten 20 Proteine neu lokalisiert und 29 von 31 bekannten IMS Proteinen reproduzierbar validiert werden. Die IMS Proteine erfüllen, wie in Abbildung 4.20 dargestellt, vornehmlich die Funktion des Proteinimports, der Redoxregulation, der Atmungskette sowie von Assemblyfaktoren. Die funktionale Einordnung von 20 % der 51 IMS Proteine ist unbekannt.

4.9 LF Quantifizierung zur Charakterisierung von Proteindegradation in Mitochondrien

Während der Proteinimport in Mitochondrien bereits eingehend untersucht wurde und neue Wege charakterisiert werden konnten, ist über die Stabilität und den Verbleib der ca. 1.000 mitochondrialen Proteine in *S. cerevisiae* bislang wenig bekannt. Die grundlegenden Degradationsprozesse, die zum Abbau von Proteinen in Mitochondrien führen, konnten noch nicht ausreichend anhand der Aktivität von AAA-ATPasen beschrieben werden [153]. Aktivität und Spezifität der etwa 20 bekannten Proteasen in Mitochondrien sind zudem noch nicht vollständig geklärt und erfordern eine weitere Charakterisierung. Im Rahmen dieser Arbeit sollte die Eignung der LF Quantifizierung zur globalen und zeitaufgelösten Analyse von Degradationsprozessen in Proben mittlerer Komplexität evaluiert werden. Die LF Quantifizierung zeichnet sich im Vergleich zu alternativen Quantifizierungsstrategien in der MS Analytik v.a. aufgrund der sensitiven Detektion von Features im niedrigen fmol Bereich aus.

4.9.1 Qualitätskontrolle der LF Quantifizierung

Eine grundlegende Voraussetzung für die LF Quantifizierung ist die Stabilität und Reproduzierbarkeit der Probenvorbereitung sowie der LC-MS Systeme. Die LF Quantifizierung von individuellen Proben nutzt Features mit identischen m/z und RT Charakteristika in unabhängigen LC-MS Messungen und vergleicht diese hinsichtlich ihrer dreidimensionalen Peakflächen. Dadurch wird der relative Vergleich der Mengenverhältnisse in individuellen LC-MS Messungen ermöglicht.

Für die LF Quantifizierung sind stabile RT und Signalintensitäten von essentieller Bedeutung. In Abbildung 4.21 sind die RT bzw. Signalintensitäten unabhängiger LC-MS Messungen identischer Proben eines humanen Thrombozytenverdaus dargestellt. Der Verlauf der mehr

als 3.000 identifizierten Peptide eines humanen Thrombozytenverdaus unabhängiger LC-MS Läufe zeigt mit einer $SD_{RT} = 0,08 \%$ für Retentionszeiten und einer $SD_{Peakfläche} < 12 \%$ für Peakflächen die hohe Reproduzierbarkeit der LC-MS Analysen.

Diese Ergebnisse zeigen, dass die hohen Anforderungen an die Reproduzierbarkeit von LC-MS Analysen bei der LF Quantifizierung komplexer Peptidgemische prinzipiell erfüllbar sind.

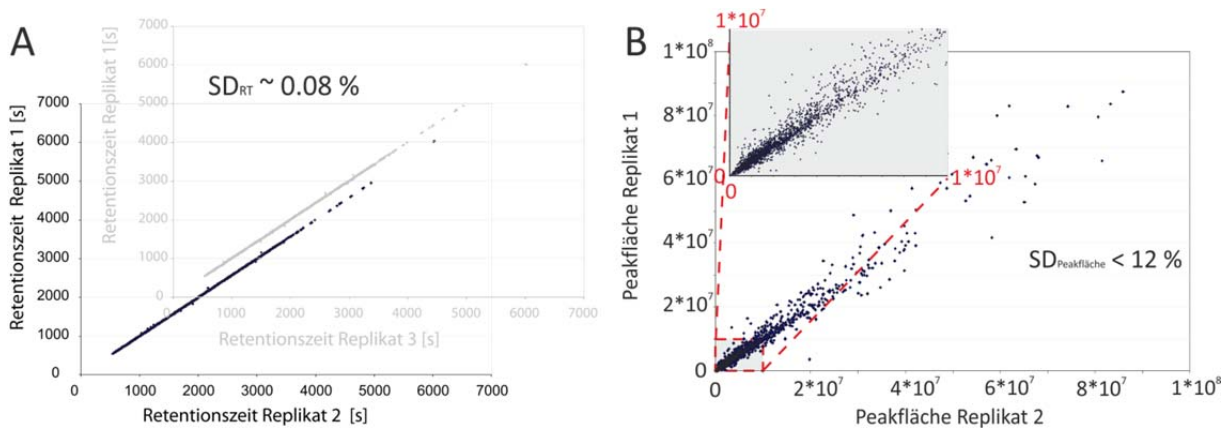


Abbildung 4.21: Die Reproduzierbarkeit von RT und Signalintensitäten wurde anhand unabhängiger LC-MS Messungen eines humanen Thrombozytenverdaus evaluiert. Die C18 Auftrennung erfolgte über 2 h auf einem Split-freien NCS System, die Akquisition an einem LTQ Orbitrap XL Massenspektrometer im CID TOP5 Modus. Die Messungen wiesen bei mehr als 3.000 identifizierten Peptiden mit einer $SD_{RT} = 0,08\%$ und $SD_{Peakfläche} < 12 \%$ eine hohe Reproduzierbarkeit auf, sodass grundsätzlich unter den verwendeten Bedingungen die LF Quantifizierung komplexer Proben möglich ist.

Um die Möglichkeiten und Limitierungen der LF Quantifizierung systematisch zu erfassen, wurde eine Probenreihe erstellt, welche aus folgenden Komponenten bestand: (1) 1 μg Thrombozytenverdaus als Matrix, (2) einer durchweg stabilen Komponente, der QC Mixtur, mit einem dynamischen Bereich von 1-125 fmol (siehe Kapitel 3.12) und (3) einer differentiellen Komponente, bestehend aus einer äquimolaren Mixtur definierter Peptide in den ansteigenden Mengen 0,1 fmol, 10 fmol, 100 fmol, 500 fmol bzw. 1.000 fmol (siehe 3.11.2 und Anhang 12.4). Die Proben wurden über 120 min innerhalb eines binären Gradienten von 5-45 % LM B aufgetrennt und im CID TOP5 Modus auf einer LTQ-Orbitrap-XL in Triplikaten analysiert. Während der Messserie von etwa vier Tagen konnten für alle QC Peptide reproduzierbare Retentionszeiten mit einer $SD_{RT/Median} = 0,30 \%$ und Peakflächen mit einer $SD_{Peakfläche/Median} = 10,44 \%$ ermittelt werden. QC Peptide der Konzentration 125 fmol wiesen dabei, verglichen zu QC Peptiden der Konzentration 1-25 fmol, wie in Abbildung 4.22 C gezeigt, die niedrigste $SD_{125 \text{ fmol/Peakfläche/Median}} = 8,74 \%$ auf.

Die linear ansteigenden Komponenten (3) konnten nur in wenigen Fällen über den dynamischen Bereich von vier Größenordnungen bis zur minimalen Peptidmenge von 0,1 fmol nachgewiesen werden. Gründe für die unterschiedlichen Signalantworten der Peptide können variierende Ionisierungseffizienzen, aber auch variierende Ausgangsmengen sein: Während die Konzentration der synthetisierten Peptide mittels Aminosäureanalyse bestimmt wurde, wurde α -Casein, ausgehend von einer konzentrierten Stocklösung, mehreren Aufarbeitungsschritten unterzogen, die potentiellen Pipettierfehlern unterliegen können, bevor es der komplexeren Probenkomposition hinzugefügt wurde.

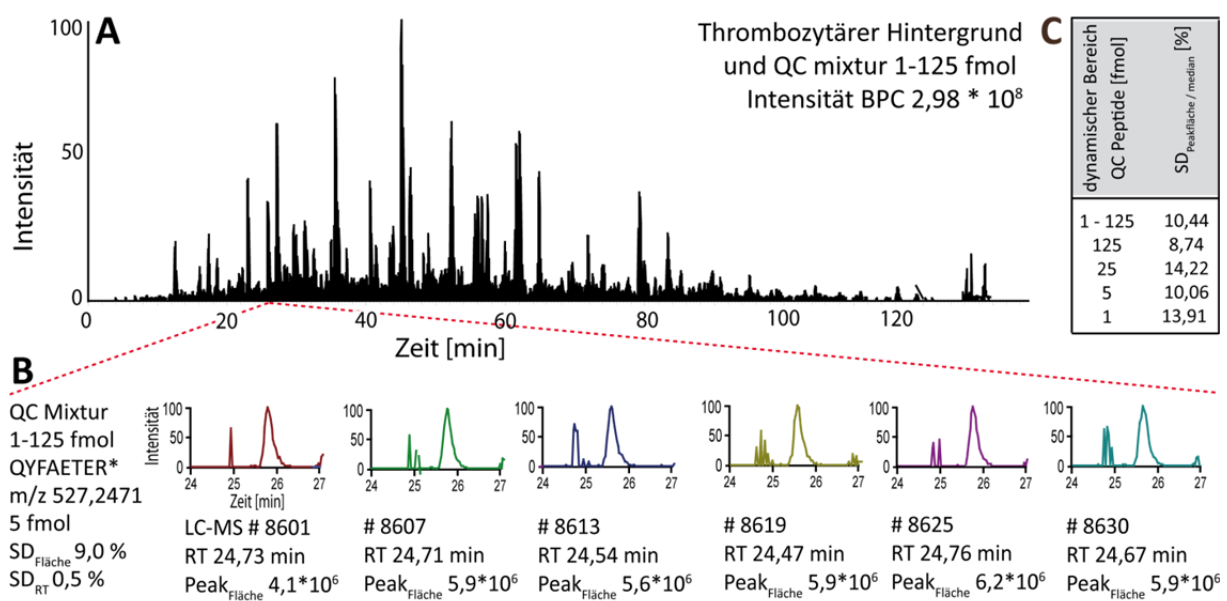


Abbildung 4.22: Qualitätskontrolle von 30 LC-MS Analysen bestehend aus thrombozytärem Hintergrund, regulierten Komponenten und der QC Mixtur im dynamischen Bereich von 1-125 fmol. Gezeigt sind (A) das Base Peak Chromatogramm und (B) die Detailanalyse des QC Peptids QYFAETER* bei einer Konzentration von 5 fmol/Injektion. Exemplarisch sind Signale von sechs unabhängigen LC-MS Analysen gezeigt. Über 30 Proben konnte für QYFAETER* eine SD_{Fläche} von 9.0 % und eine SD_{RT} von 0.5 % erreicht werden. (C) SDs der Peakflächen von QC Peptiden übergreifend (1-125 fmol) und aufgeteilt nach jeweils injizierten Mengen (1, 5, 25, 125 fmol).

Alle drei hinzugegebenen Phosphopeptide konnten bis zu einer Stoffmenge von 0,5 fmol mit hoher Linearität nachgewiesen werden (siehe Abbildung 4.23).

Für alle linear ansteigenden Peptide konnten, basierend auf den Mittelwerten der in der Dreifachbestimmung ermittelten Peakflächen, hohe Bestimmtheitsmaße (siehe Tabelle 4.6) ermittelt werden.

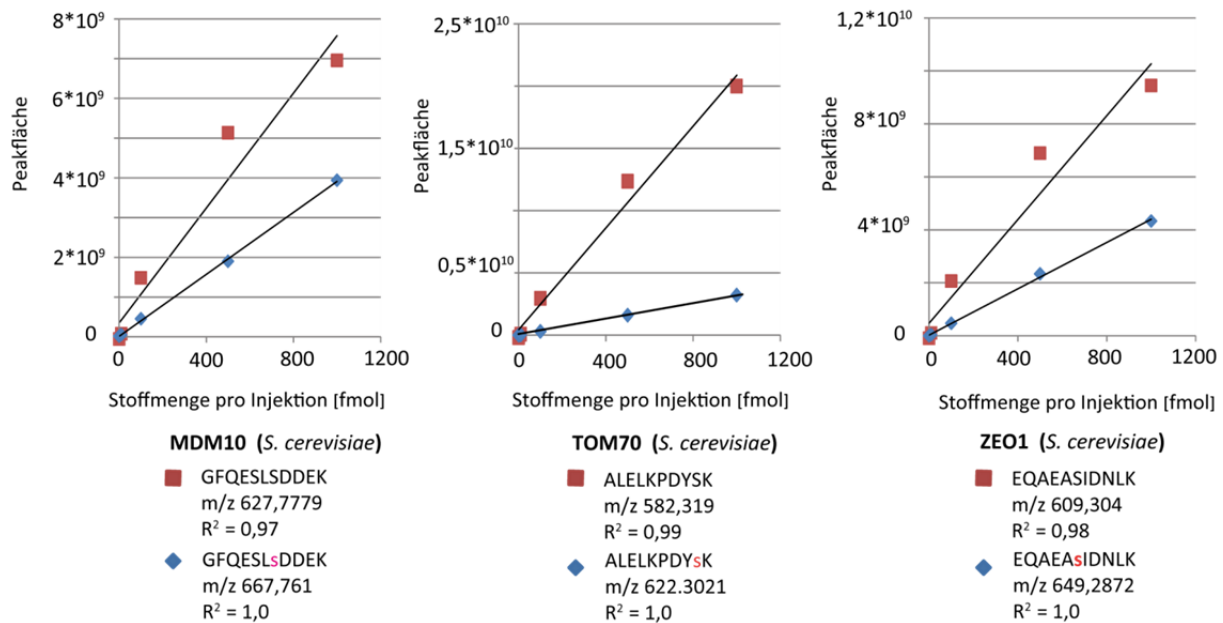


Abbildung 4.23: Ansteigende Signalintensitäten ergänzter mitochondrialer Peptide in phosphorylierter und nicht-phosphorylierter Form.

Tabelle 4.6: Übersicht über die Bestimmtheitsmaße linear ansteigender Peakflächen von Peptiden, die einem humanen Thrombozytenverdaus in Stoffmengen von 0,1 – 1.000 fmol hinzugegeben wurden. Basierend auf den Mittelwerten der Signalintensitäten wurde das Bestimmtheitsmaß über einem dynamischen Bereich von vier Größenordnungen ermittelt.

Proteine	Peptide	m/z	Peakflächen						Bestimmtheitsmaß R ²
			Stoffmenge pro Injektion [fmol]						
			0.1	0.5	10	100	500	1000	
MDM10	ALELKPDYSK	627,773	9,76 * 10 ⁶	1,07 * 10 ⁷	4,53 * 10 ⁷	4,49 * 10 ⁸	1,89 * 10 ⁹	3,93 * 10 ⁹	1.00
MDM10	ALELKPDYSK	667,761	1,77 * 10 ⁶	7,6 * 10 ⁶	1,38 * 10 ⁸	1,48 * 10 ⁹	5,14 * 10 ⁹	6,95 * 10 ⁹	0.97
TOM70	GFQESLSDDEK	582,318	-	8,62 * 10 ⁵	2,27 * 10 ⁷	2,86 * 10 ⁸	1,60 * 10 ⁹	3,22 * 10 ⁹	1.00
TOM70	GFQESLSDDEK	622,302	2,85 * 10 ⁶	1,48 * 10 ⁷	2,72 * 10 ⁸	2,95 * 10 ⁹	1,24 * 10 ¹⁰	2,00 * 10 ¹⁰	0.99
ZEO1	EQAEASIDNLK	609,304	-	1,72 * 10 ⁶	4,01 * 10 ⁷	4,74 * 10 ⁸	2,34 * 10 ⁹	4,33 * 10 ⁹	1.00
ZEO1	EQAEASIDNLK	649,287	5,47 * 10 ⁵	9,83 * 10 ⁶	2,00 * 10 ⁸	2,06 * 10 ⁹	6,89 * 10 ⁹	9,45 * 10 ⁹	0.97
TIM54	ISSLLPFNK*	513,808	-	9,23 * 10 ⁶	2,33 * 10 ⁷	2,59 * 10 ⁸	1,42 * 10 ⁹	2,56 * 10 ⁹	1.00
TIM54	NSEWTQELVCDPR*	822,370	-	1,93 * 10 ⁷	-	5,47 * 10 ⁸	2,59 * 10 ⁹	4,67 * 10 ⁹	1.00
TIM54	EKNSEWTQELVCDPR*	950,941	9,87 * 10 ⁴	9,39 * 10 ⁵	2,70 * 10 ⁷	2,89 * 10 ⁸	1,79 * 10 ⁹	3,85 * 10 ⁹	1.00
TIM 22	VYTGFGLEQISPAQK*	823,439	-	9,76 * 10 ⁵	2,38 * 10 ⁷	2,92 * 10 ⁸	1,61 * 10 ⁹	3,10 * 10 ⁹	1.00
TIM 18	NNGFVDLVK*	507,279	-	-	3,30 * 10 ⁷	3,38 * 10 ⁸	1,76 * 10 ⁹	3,37 * 10 ⁹	1.00
TIM 18	LWNNDDHLVIF GR*	901,433	-	-	3,74 * 10 ⁶	6,11 * 10 ⁸	4,36 * 10 ⁸	9,28 * 10 ⁸	1.00
TIM 9	FQE QNAALGQGLGR*	749,883	-	2,12 * 10 ⁶	4,86 * 10 ⁷	4,84 * 10 ⁸	2,47 * 10 ⁹	4,45 * 10 ⁹	1.00
TIM 9	LYSNLVER*	502,277	1,59 * 10 ⁷	1,54 * 10 ⁷	7,08 * 10 ⁷	s	3,29 * 10 ⁹	5,42 * 10 ⁹	0.99
TIM 12	LGGFVQTR*	500,794	4,70 * 10 ⁵	1,42 * 10 ⁶	3,86 * 10 ⁷	4,01 * 10 ⁸	2,33 * 10 ⁹	4,48 * 10 ⁹	1.00
	HQGLPQEVLNENLLR	880,476	-	-	8,03 * 10 ⁶	1,4 * 10 ⁸	4,77 * 10 ⁸	1,11 * 10 ⁹	1.00
α-Casein	YLGYLEQLLR	634,356	-	3,77 * 10 ⁵	1,58 * 10 ⁷	1,78 * 10 ⁸	4,72 * 10 ⁸	1,33 * 10 ⁹	0.99
	YKVPQLEIVPNSAEER	976,480	-	-	5,56 * 10 ⁶	6,59 * 10 ⁷	3,32 * 10 ⁸	6,48 * 10 ⁸	1.00

Um die Reproduzierbarkeit der Probenaufarbeitung zu bestimmen wurden drei Proben eines Mitochondrienlysats (YPG) zeitgleich aber unabhängig aufgearbeitet. Kontrolliert wurden verschiedene Proteinkonzentrationen während des Verdaus (0,1 µg/µL, 0,5 µg/µL und 1 µg/µL) und Verluste bzw. die Reproduzierbarkeit der C18 Entsalzung. Die Proben wurden in Triplikaten jeweils für 2 h mit einer LTQ-Orbitrap-XL im CID TOP5 Modus gemessen und die

anschließende LF Datenauswertung mit der Software Progenesis von Nonlinear Dynamics durchgeführt.

Es zeigten sich durchweg niedrige Abweichungen auf Proteinebene ($\Delta_{\text{Median}} = 8,4 \%$), wobei für eine Konzentration von $0,5 \mu\text{g}/\mu\text{L}$ und anschließender C18 Entsalzung reproduzierbar die meisten Proteine mit mehr als zwei unigen Peptiden quantifiziert werden konnten.

Tabelle 4.7: Zur Kontrolle der reproduzierbaren Probenaufarbeitung wurden identische Mitochondrienlysate unabhängig aufgearbeitet, in Triplikaten gemessen und einer LF Quantifizierung unterzogen. Der Vergleich der Mediane der Signalintensitäten zwischen den unabhängig bearbeiteten Proben (Δ_{Median}) zeigt eine hohe Reproduzierbarkeit bei geringen SDs auf.

0.5 $\mu\text{g}/\mu\text{L}$ C18	Anzahl	Δ_{Median}	SD [%]
alle identifizierten Proteine	792	-	-
alle quantifizierten Proteine, mit unigen Peptiden	618	1,06	5,5
quantifizierte Proteine, mit ≥ 2 unigen Peptiden	419	1,03	2,55
quantifizierte Proteine, mit nur 1 unigen Peptid	199	1,13	3,26

4.9.2 Spectral counting

Nachdem die Qualität sowohl der Probenaufarbeitung als auch der Analyse den hohen Anforderungen der LF Quantifizierung entsprach, wurde die LF Quantifizierung von Degradationsprozessen in Mitochondrien durchgeführt.

Hierzu wurden aufgereinigte Mitochondrien über einen Zeitraum von sieben Tagen in SEM Puffer inkubiert, wobei zu den Zeitpunkten 0 h, 6 h, 24 h, 72 h und 168 h ein definiertes Volumen der Mitochondriensuspension abgenommen wurde. Die Proben wurden dann entsprechend der im Rahmen dieser Arbeit optimierten Verfahren aufgearbeitet und analysiert. Zunächst wurde eine globale Quantifizierung mittels des NSAF Index [154] durchgeführt. Untersucht wurden Degradationsprozesse in Mitochondrien von Hefen, die unter fermentativen bzw. respiratorischen Bedingungen kultiviert wurden. Nach Aufarbeitung der über die Zeit abgenommenen Proben wurden jeweils LC-MS Analysen in Triplikaten durchgeführt. Zwecks einer verbesserten chromatographischen Auflösung wurde für die Auftrennung über 4 h eine 50 cm C18 Trennsäule mit I.D. $75 \mu\text{m}$ auf einem Split-freien NCS System verwendet und dieses mit einem qExactive Massenspektrometer gekoppelt, welches sich durch eine sehr hohe Scangeschwindigkeit bei hoher Massengenauigkeit ausweist. Es galt hier möglichst viele Spektren und Peptide pro Protein über einen hohen dynamischen Bereich zu identifizieren.

Die mittels PeptideShaker berechneten NSAF Indizes wurden auf dem Median aller für einen Zeitpunkt ermittelten NSAF Werte normalisiert. Ein Vergleich der normalisierten NSAF Werte

zwischen den Zeitpunkten 0 h bis 168 h zeigt deutlich abfallende aber auch stabile Proteinmengen auf, korrespondierend zu degradierten und stabilen Proteinen.

Bei Auftragung der NSAF Werte gegen die Zeit lässt sich die Steigung (Δm) der Regressionsgeraden und damit der generelle Trend der Proteinmengen anschaulich darstellen. Die lineare Regression wurde gewählt, da die NSAF Trends der ca. 1.000 quantifizierten Proteine sehr unterschiedlich waren, sodass diese Art der Auswertung zwar die tatsächlichen Verläufe im Einzelfall nur ungenügend widerspiegelt, aber eine schnelle und verlässliche Unterscheidung von abgebauten und stabilen Proteinen ermöglicht.

Die Berechnung der SD_{NSAF} eines Proteins für die fünf Zeitpunkte gibt das Ausmaß der quantitativen Änderung wieder. Proteine, die degradiert werden, weisen abfallende NSAF Indizes und folglich eine negative Δm sowie eine hohe SD_{NSAF} auf (Abbildung 4.24 A).

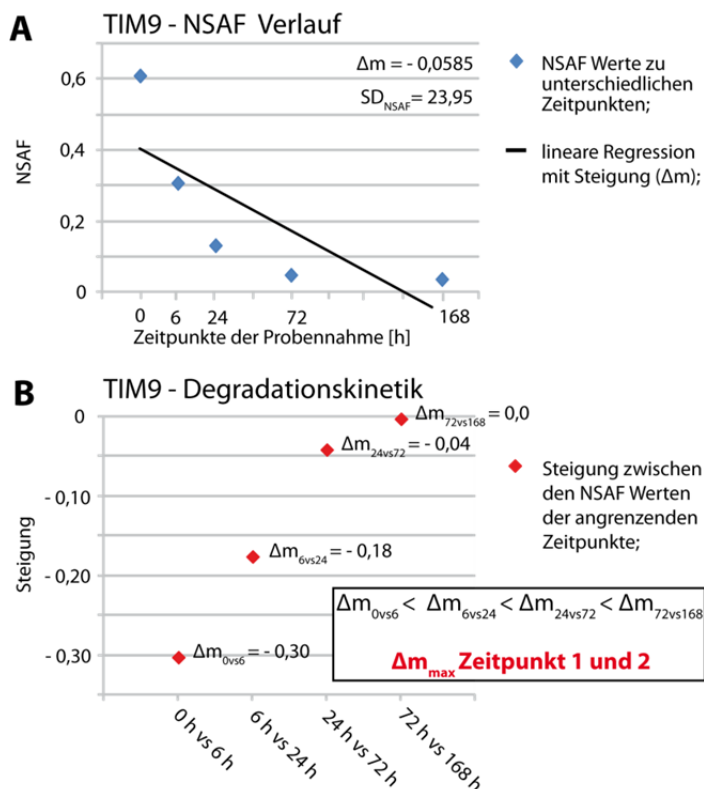


Abbildung 4.24: (A) Aufgetragen gegen die Zeit zeigen die NSAF Indizes den Abfall der Proteinkonzentration von TIM9 auf. Charakteristisch für einen Konzentrationsabfall ist die negative Steigung (Δm). Die hohe SD_{NSAF} weist auf eine starke Änderung der Signalintensität hin. (B) Anhand der näheren Betrachtung der Steigung Δm zwischen den aufeinanderfolgenden Zeitpunkten lässt sich eine Aussage zur Degradationskinetik treffen.

Eine Detailanalyse der Steigungen aufeinanderfolgender Zeitpunkte ermöglicht eine Aussage zur zugrundeliegenden Abbaukinetik der Proteine und erlaubt, wie in Abbildung 4.24 (B) dargestellt, die Bestimmung des Zeitpunkts des maximalen Abfalls der Proteinmenge.

Während für die YPG Zeitreihe nur 269 mitochondriale Proteine (von 652 identifizierten) den oben beschriebenen Kriterien für Degradation genügten, wurden für die YPD Zeitreihe mit 421 (von 693 identifizierten) wesentlich mehr abgebaute Proteine gefunden. Soweit in der Literatur bekannt bzw. in der Saccharomyces Genome Database gelistet, konnte für 56 %

bzw. 60 % der für YPG und YPD identifizierten Proteine eine Lokalisation innerhalb der vier mitochondrialen Subkompartimente zugeordnet werden.

Für beide Probenreihen konnten für eine hohe Zahl an IMS Proteinen z.T. stark abfallende Tendenzen bestimmt werden - für YPG 40 (82 %) und für YPD 33 (67 %) von 51 bekannten bzw. im Rahmen dieser Arbeit identifizierten IMS Proteinen (siehe Kapitel 4.8).

Quantifizierung	NSAF - Proteinebene			
	abfallend		stabil	
Verlauf Probenstet	YPG	YPD	YPG	YPD
mitochondrial IDs	269	421	384	273
Lokalisation	56%	60%	57%	62%
OM	44 (24,9%)	50 (20,6%)	27 (12,3 %)	31 (16 %)
IMS	40 (22,6 %)	33 (13,6 %)	2 (0,9 %)	33 (17 %)
IM	65 (36,7 %)	93 (38,3 %)	107 (48,9 %)	87 (44,8 %)
Matrix	28 (15,8 %)	67 (27,6 %)	83 (37,9 %)	43 (22,2 %)

Tabelle 4.8: Zusammenfassung mitochondrialer Degradationsprozesse unter fermentativen (YPD) und respiratorischen (YPG) Bedingungen. Angegeben ist die Anzahl der abfallenden bzw. stabilen mitochondrialen Proteine, die zu mehr als der Hälfte lokalisiert und den entsprechenden Kompartimenten (%) zugeordnet werden konnten.

Die globale Betrachtung der fünf Zeitpunkte zeigte jedoch auch eine systematische Drift der Auswertung auf: Nachdem die Gesamtproteinmenge bei fortwährender Inkubation aufgrund von Degradation und quasi nicht vorhandener Proteinbiosynthese kontinuierlich abnimmt, sollten keine Proteine ansteigende NSAF Indizes aufweisen. Da der prozentuale Anteil von stabilen Proteinen an der Gesamtproteinmasse über die Zeit jedoch ansteigt und die korrespondierenden Peptide folglich eine höhere Wahrscheinlichkeit haben MS/MS Spektren zu generieren, nehmen ihre NSAF Indizes entsprechend zu. Dies führt zu einer Verzerrung der relativen Quantifizierung des Proteoms, die, aufgrund der komplett unterschiedlichen Degradationsverläufe der mitochondrialen Proteine, nicht durch einen Korrekturfaktor rechnerisch kompensiert werden kann. So zeigen Proteine mit einer stabilen Tendenz, wie die Kontaminanten HTB1 oder HHF1 und PMA1, tatsächlich über die Zeit ansteigende NSAF Indizes (Abbildung 4.25).

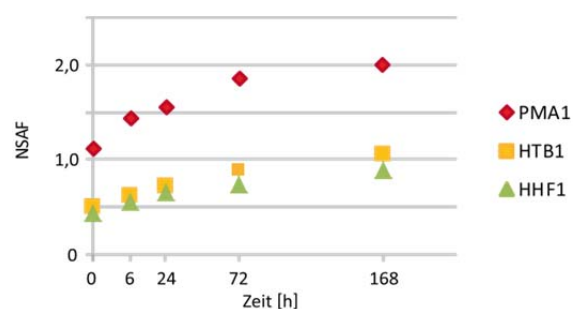
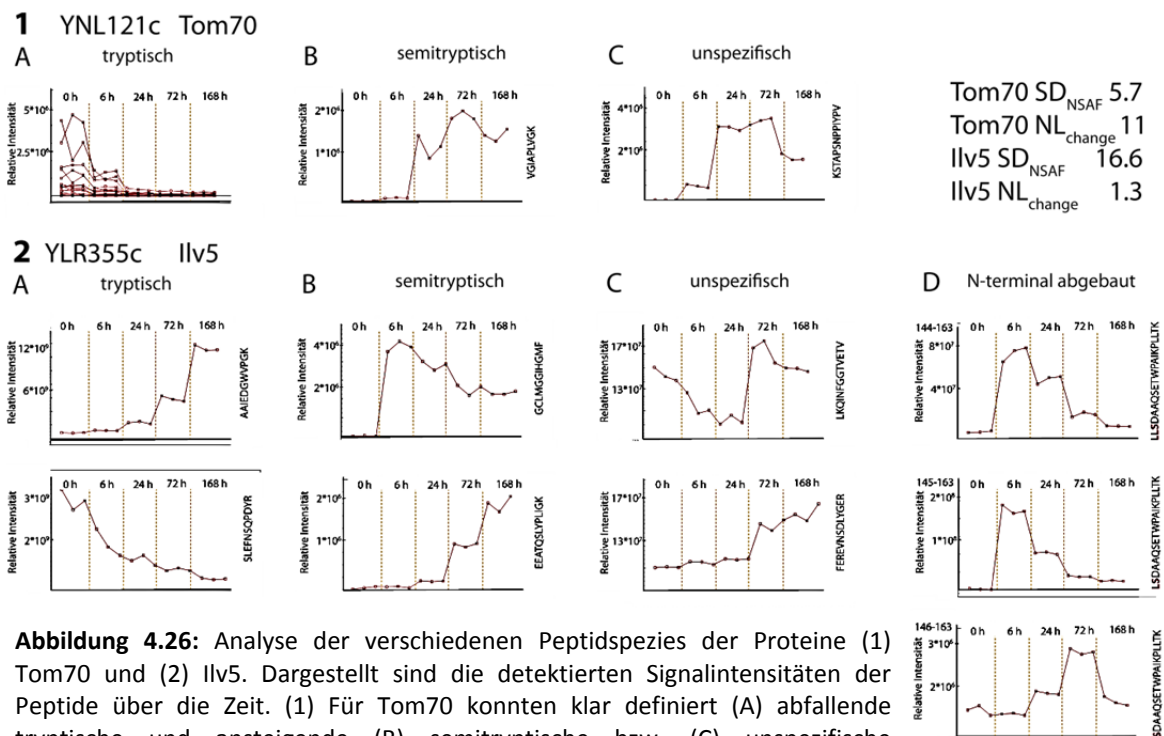


Abbildung 4.25: Durch die Normalisierung der NSAF Werte zeigen stabile Proteine wie die Kontaminanten PMA1, HTB1 und HHF1 über die Zeit ansteigende NSAF Indizes.

4.9.3 Feature-basierte LF Analyse

Die Feature-basierte LF Analyse wurde mit der Software Progenesis von Nonlinear Dynamics durchgeführt. Die proteolytische *in vivo* Degradation von Proteinen führt beim tryptischen Verdau im Rahmen der Probenvorbereitung zur Entstehung neuer, nicht tryptischer Peptide, die C- und/oder N-terminal keine echte tryptische Schnittstelle aufweisen. Dies bewirkt, dass die Menge des korrespondierenden tryptischen Peptids, im Vergleich zur „nicht degradierten“ Kontrollprobe (Zeitpunkt 0 h), abnimmt, während die des neu entstehenden nicht-tryptischen Peptids zunimmt. Somit kann bei dieser Art der LF Quantifizierung nicht nur, wie beim NSAF Ansatz, die grundlegende Abnahme der Proteinmenge quantifiziert, sondern auch die eigentliche Schnittstelle eines proteolytischen Prozesses ermittelt werden. Während für die Mehrheit der Proteomanalysen ausschließlich tryptische Peptide herangezogen werden, vermag die Analyse semitryptischer, unspezifischer oder terminal abgebauter Peptide [80], die im Folgenden übergreifend als „atypische Peptide“ bezeichnet werden, Licht in die bislang wenig erforschten Degradationsprozesse in Mitochondrien zu bringen. Neben den für die LF Analyse grundlegenden Kriterien bezüglich Reproduzierbarkeit von Probenaufarbeitung und LC-MS Analyse ist für die Analyse von Degradationsprozessen weiterhin die Effizienz des proteolytischen Verdaus von großer Bedeutung.



Eine, wie unter 3.12.1 beschriebene, mehrstufige Datenbanksuche ermöglichte eine detailreiche Analyse der verschiedenen genannten Peptidspezies für individuelle Proteine, wie in Abbildung 4.26 (1) am Beispiel von Tom70 dargestellt. Während die Signalintensitäten tryptischer Tom70 Peptide über den Zeitverlauf abnehmen, steigen Signale, die semitryptischen oder unspezifisch geschnittenen Sequenzen zugeordnet werden können, an. Dieser Verlauf ließ sich jedoch nur für eine Minderheit der Proteine derart deutlich nachweisen. Die Mehrheit der Proteine mit allgemein abfallender Tendenz zeigte eine Aneinanderreihung von proteolytischen Ereignissen, die zu einem komplexen Peptidgemisch führen, in welchem tryptische Peptide abnehmen, korrespondierende semitryptische oder unspezifische Peptide ansteigen, um schließlich inmitten der Zeitreihe aufgrund eines erneuten proteolytischen Ereignisses wieder abzufallen und somit die Entschlüsselung von Degradationsprozessen zu erschweren (Abbildung 4.26 (2)).

Aus diesem Grund wurden für beide Zustände atypische Peptidsequenzen mit Hilfe der mehrstufigen Datenbanksuchen extrahiert, um diese auf sequenzspezifische proteolytische Schnittstellen zu untersuchen. Da die Kinetik von Degradationsprozessen nicht zwingend linear verläuft und außerdem LC-MS Messungen Schwankungen unterliegen können, müssen die Mittelwerte der detektierten Peakflächen ($\text{Fläche}_{\text{Mittelwert}}$) aufeinanderfolgender Zeitpunkte, v.a. bei geringfügigen Veränderungen, nicht zwangsläufig entsprechend der allgemeinen Tendenz zu- bzw. abnehmen (Abbildung 4.27).

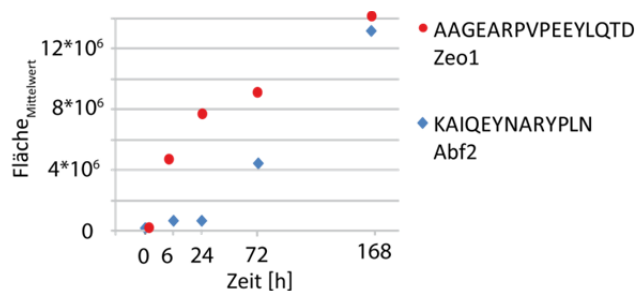


Abbildung 4.27: Die Mittelwerte der Peakflächen müssen nicht zwangsläufig zwischen den Zeitpunkten an- bzw. absteigen. Wie hier für KAIQEYNARYPLN (Abf2; blau) gezeigt, ist zwischen den Signalen zu den Zeitpunkten 6 und 24 h kein Anstieg zu sehen; die allgemeine Tendenz des Peptids entspricht jedoch AAGEARPVPEEYLQTD (Zeo1; rot).

Zur Detektion von Degradationsereignissen wurde daher, neben den bereits in 3.14.5 beschriebenen Anforderungen, ein weiteres Kriterium gewählt: Mindestens drei von vier $\text{Fläche}_{\text{Mittelwert}}$ - Werten aufeinanderfolgender Zeitpunkte müssen der allgemeinen Tendenz folgend zu- bzw. abnehmen.

Atypische Peptide wurden gemäß der genannten Kriterien in Gruppen ansteigender und abfallender Tendenz und weiterhin hinsichtlich der Orientierung der atypischen Schnittstelle,

nämlich N- bzw. C-terminal, sortiert. Diese Gruppen wurden verwendet, um potentielle Schnittstellen bzw. Konsensussequenzen für Proteasen zu identifizieren.

Für die Auswertung wurden hierzu dekamere Peptidsequenzen, bestehend aus der Schnittstelle und den 4 bzw. 5 benachbarten C- und N-terminalen Aminosäuren, berücksichtigt. Die proteolytische Schnittstelle befindet sich somit immer im Zentrum der zehn AS (Abbildung 4.28).

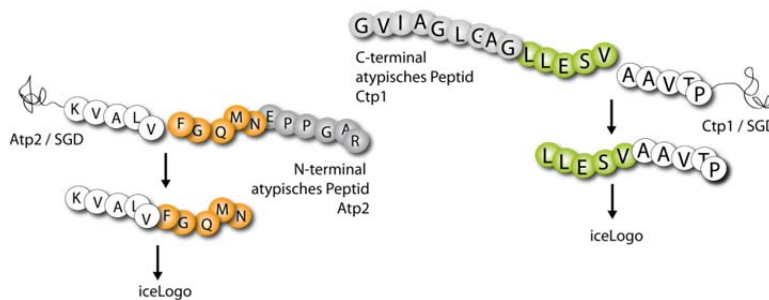


Abbildung 4.28: Jeweils fünf AS der C- (grüne Kreise) bzw. N-terminalen (orange) atypischen Peptide wurden mit fünf AS der zugrundeliegenden Proteinsequenz (weiß) ergänzt. Anschließend wurden die Peptide mit einer Länge von zehn AS einer statistischen Sequenzanalyse mit Hilfe von iceLogo unterzogen.

Durchgeführt wurde die Analyse auf charakteristische Konsensussequenzen mit Hilfe der freien Software iceLogo [109], welche die positionsspezifische Häufigkeit von Aminosäuren unter Berücksichtigung ihrer Prävalenz im respektiven Proteom statistisch evaluiert.

Durch die statistische Analyse zeichnen sich deutlich Häufigkeitsprofile der AS für die Positionen -5 bis +5 ab, wobei die untersuchten Zustände (YPG, YPD) ähnliche Profile aufweisen, die sich hinsichtlich abfallender und ansteigender Peptide marginal unterscheiden (Abbildung 4.29). Als Referenzproteom wurde die Promito-Datenbank 851 mitochondrialer Proteine herangezogen, welche die bisher umfangreichste Studie mitochondrialer Proteine in Hefe darstellt [107].

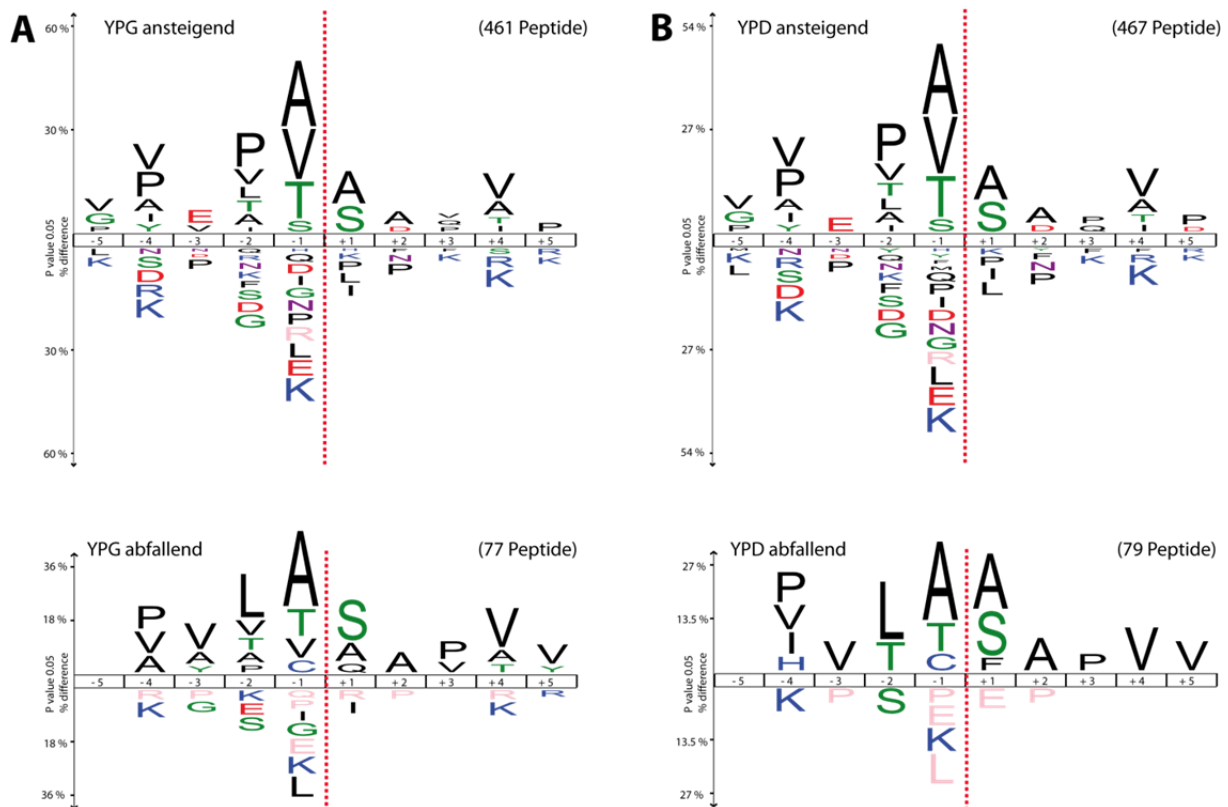


Abbildung 4.29: Sequenzanalyse von atypischen mitochondrialen Peptiden mit ansteigenden bzw. abfallenden Trends, mittels iceLogo dargestellt. Überrepräsentierte und unterrepräsentierte AS werden entsprechend ober- bzw. unterhalb der x-Achse angegeben, wobei die Größe der Buchstaben ein Maß für die Stärke der Über-/Unterrepräsentierung ist. (A) YPG und (B) YPD Mitochondrien weisen für an- bzw. absteigende Peptide eine große Ähnlichkeit auf - insbesondere an der proteolytischen Schnittstelle in Position -1 (rote gestrichelte Linie).

Eine eindeutige Aussage bezüglich einer definierten Protease-Erkennungssequenz gestaltet sich allerdings schwierig, da möglicherweise Sequenzprofile unterschiedlicher Proteasen überlappen und somit eine diffuse Ansicht dargestellt wird. Eine Klassifizierung der zugehörigen Proteinidentifikationen in die vier mitochondrialen Subkompartimente bietet sich an, da sich die Proteome dieser unterscheiden und weiterhin der Wirkungsraum von proteolytisch aktiven Proteasen aufgrund der IM-Barriere begrenzt sein sollte. Wie in Abbildung 4.30 am Beispiel von YPG / ansteigend dargestellt, lassen sich dadurch Sequenzprofile verdeutlichen, die jedoch teilweise auf nur wenigen Peptiden beruhen.

Weiterführende sequenzspezifische Analysen proteomischer und biochemischer Art sind daher für die detailreiche Charakterisierung von mitochondrialen Degradationsprozessen notwendig. Die im Rahmen dieser Arbeit erstellten potentiellen Konsensussequenzen bzw.

5. Diskussion

Im Rahmen dieser Arbeit wurden Proteome unterschiedlicher Komplexität hinsichtlich ihrer qualitativen als auch quantitativen Zusammensetzung unter Zuhilfenahme dreier verschiedener quantitativer Strategien, der chemischen bzw. der metabolischen stabil-Isotopen Markierung und der LF Quantifizierung, untersucht. Grundlage einer umfassenden und validen Proteomanalyse sind sowohl Reproduzierbarkeit als auch Sensitivität, mit welcher Proben aufgearbeitet, Daten erhoben und evaluiert werden. Die quantitativen Proteomstudien wurden daher durch neue Strategien zur Qualitätskontrolle, welche die Probenaufarbeitung, die LC-MS Performance sowie die Datenauswertung hinsichtlich Identifizierung und Quantifizierung betreffen, ergänzt.

5.1 Qualitätskontrollen in der qualitativen und quantitativen Proteomanalyse

Die QC ist in analytischen Disziplinen wie der Proteomanalyse ein essentieller Bestandteil, der die Reproduzierbarkeit und Validität von Analysen und von analytischen Ergebnissen garantiert. Eine geeignete QC umfasst daher alle Teilbereiche der Proteomanalyse und erfordert für einen zeitlich akzeptablen Aufwand etablierte und robuste Strategien. In dieser Arbeit wurden vielseitige Vorgehensweisen und Strategien, die einfach anzuwenden sind und die, entsprechend neuer analytischer Fragestellungen, einfach abgewandelt bzw. ergänzt werden können, entwickelt und auch erfolgreich angewendet.

Die Schwierigkeit komplexe Proteome vollständig mit einer hohen Sequenzabdeckung quantitativ darzustellen kann u.a., wie in der Verdaustudie (Kapitel 4.4) dargestellt, an ineffizienten und nicht reproduzierbaren proteolytischen Verdauen liegen. Trypsin ist die in der Proteomanalytik am häufigsten eingesetzte Protease und Datenarchive wie PeptideAtlas basieren auf der Listung von tryptischen Peptiden [155]. Jedoch können Aminosäuresequenzen wie etwa von Membranproteinen hydrophobe Bereiche (z.B. Transmembrandomänen) aufweisen, die jegliche tryptische Schnittstelle vermissen lassen und somit ungeeignet für den Einsatz von Trypsin sind. Analog zur in dieser Arbeit gezeigten Verdaustudie haben daher Peng et al. [156] die Eignung alternativer Proteasen wie Chymotrypsin, Lys-N und Lys-C für die Proteinquantifizierung mittels *Spectral counting* und die Auswirkung auf die Sequenzabdeckung eines Proteoms untersucht.

Die Etablierung einer stabil-Isotopen gelabelten Peptid-Qualitätsmischung mit dynamischem Bereich und Verteilung über den gesamten Elutionszeitraum, die für sich oder als interner Standard innerhalb eines realen Experiments die (online) Kontrolle vielerlei LC-MS Parameter erlaubt, fand innerhalb der Proteomik-Gemeinde Anklang. Alternative/ähnliche Mischungen wurden nach Veröffentlichung des hier erarbeiteten QC Standards sowohl seitens der akademischen Forschung [157] als auch der Industrie zum kommerziellen Erwerb (Thermo Scientific, Sigma Aldrich) präsentiert. Die Synthese der entsprechenden Peptide inklusive einer tryptischen Schnittstelle, und gegebenenfalls einem Reporter-Peptid, würde die Zugabe dieser QC Mischung vor dem proteolytischen Verdau ermöglichen, sodass dieser quantitativ auf Effizienz und Reproduzierbarkeit evaluiert werden könnte. Alternativ wäre eine Zugabe von SIL Proteinen denkbar. In jedem Fall wäre eine Kombination eines solchen Verdau-QC-Standards mit dem in dieser Arbeit entwickelten QC Standard sinnvoll, da dann tatsächlich alle Ebenen des Analyseprozesses komplett intern validiert werden könnten. Obwohl derzeit die meisten quantitativen Studien derartige Kontrollen vermissen lassen, führt mit Sicherheit in der Zukunft kein Weg daran vorbei, um bei der großen und wachsenden Anzahl proteomischer Studien Qualität von Quantität zu unterscheiden.

Während die Anwendung einer QC Mischung in ihrer Strategie eine simple aber effektive Qualitätskontrolle darstellt, ist das Auslesen der vollständigen QC Mischung für eine LF Analyse, die mit Replikaten eine Vielzahl an LC-MS Messungen umfasst, ein zeitaufwendiges Unterfangen. Eine automatisierte Auswertung der QC Peptide wäre daher für die Routine von großangelegten LF Studien erstrebenswert. Erste Versuche, eine solche QC Software im Rahmen einer Kooperation mit der Arbeitsgruppe von Prof. Dr. Oliver Kohlbacher (Universität Tübingen) in die frei-erhältliche Software OpenMS zu integrieren, wurden bereits unternommen. Weiterführende MS datenverarbeitende Programme wie das kürzlich publizierte SimpatiQCo [158] (<http://ms.imp.ac.at/?goto=simpatiqco>) erlauben die Kontrolle von MS spezifischen Parametern wie z.B. die Analyse der Füllzeiten oder der Massengenauigkeit des Detektors, die auf die Sensitivität von LC-MS Analysen erhebliche Auswirkungen haben können.

Neben der QC der eigentlichen Laborarbeit/Messung wurden in dieser Arbeit auch Konzepte entwickelt und umgesetzt, die eine QC der Datenauswertung ermöglichen bzw. diese bedeutend verbessern und erleichtern. Die etablierte Strategie der Abschätzung von falsch positiven Identifikationen mittels der Bestimmung der FDR wurde hinsichtlich verschiedener

Parameter der Datenbanksuche, wie z.B. der Vorläuferionentoleranz, optimiert und die Anwendbarkeit auf verschiedene Datensatzgrößen getestet. So konnte z.B. die Peptidentifizierungsrate von PSMs durch die Gruppierung von Ladungszuständen und PTM tragenden Peptiden bei gleichbleibender Qualität signifikant um 13 % gesteigert werden [102]. Aufgrund der Gruppierung in Subdatensätze geringerer Größe stellte sich allerdings die Frage der Anwendbarkeit und Allgemeingültigkeit der FDR Strategie. So kann die Auflösung der FDR bei Datensätzen mit nur wenigen Einträgen unzureichend sein, wenn die Anzahl an target Identifikationen zwischen zwei aufeinanderfolgenden decoy Hits zu gering ist; ein weiterer Punkt, der oftmals nicht beachtet wird. In diesem Fall ist der Einsatz der FDR Strategie nicht geeignet und andere stringente Qualitätskriterien sollten Verwendung finden. Im Zuge der detailreichen Kontrolle der Probenaufarbeitung wurde die Quantifizierungsstrategie iTRAQ sowohl eingesetzt als auch der zugrundeliegende Arbeitsablauf der iTRAQ Quantifizierung getestet und optimiert. Neben Optimierungen in der Probenaufarbeitung wurde die iTRAQ Quantifizierung v.a. hinsichtlich ihrer Präzision untersucht. Es zeigte sich, dass die statistische Methodik des redescending M-estimators, welcher v.a. die globale Verteilung von Datensätzen detektiert und Ausreißer von der weiteren Auswertung ausschließt, iTRAQ Verhältnisse am besten darstellen kann. Zur Charakterisierung des dynamischen Bereichs der iTRAQ Quantifizierung wurden die bekannten Peptidverhältnisse 1:5 (5 fmol:25 fmol), 1:10 (25 fmol:250 fmol) und 1:2 (250 fmol:500 fmol) von Standardproteinen in einem mitochondrialen Hintergrund untersucht. V.a. für niedrige Signalintensitäten, die vom Rauschen kaum zu unterscheiden waren, konnten große Abweichungen in der Darstellung der bekannten Peptidverhältnisse detektiert werden. Für höhere Signalintensitäten wie im Verhältnis 1:2 (250 fmol:500 fmol) konnte der Fehler auf 15 % reduziert werden. Die 15 % wurden zunächst einem systematischen Fehler in der Probenaufarbeitung zugeordnet, konnten jedoch in dieser Arbeit innerhalb der Analyse des quantitativen Fehlers von Reporterionen-basierten Quantifizierungsstrategien nach MS-Fragmentierung als Problematik der Interferenz identifiziert werden [105]. Nachdem die Erfassung des quantitativen Fehlers target/decoy Suchen ähnlich ist, wurde die Analyse einer quantitativen FDR (qFDR) gleichgesetzt. Einer der vier iTRAQ Kanäle ist hierzu mit einem zum analysierenden Organismus unterscheidbaren Referenzproteom belegt, welches eindeutig hinsichtlich Identität und Quantität selbst bei Ko-Fragmentierung zugeordnet werden kann. Als Referenzproteom eignet sich hierzu der

thermophile Organismus *Pyrococcus furiosus* (PFU), welcher bei einem *in silico* Abgleich mit der humanen Datenbank nur sieben identische tryptische Peptide aufwies und dessen Spektren einer LC-MS Messung nur ein humaner decoy Hit zugeordnet werden konnte [120]. Dies konnte ebenso in Experimenten ansteigender Komplexität bestätigt werden: Ein Vergleich der durchgeführten Experimente (a) PFU vs. IEF Fraktionen von Thrombozyten, (b) PFU vs. Gesamtthrombozyten und (c) PFU vs. HeLa wies iQuARI als robuste Methode zur Bestimmung des Fehlers von Reporterionen-basierten Quantifizierungsstrategien aus, wobei die Interferenz proportional zur Komplexität der zu analysierenden Probe abnimmt (Abbildung 5.1). Wenn auch iQuARI nur drei von vier iTRAQ Kanälen für die eigentliche quantitative Analyse bereitstellt, so überzeugt die Methodik mit ihrer einfachen Anwendung und unkomplizierten Auswertung, die neben der Identifizierung der Interferenz auch die Quantifizierung derselben erlaubt. Dies sind entscheidende Vorteile gegenüber den von Ting et al. [118] und Wenger et al. [119] entwickelten Strategien zur Analyse der Interferenz, da diese instrumentell bzw. bioinformatisch sehr anspruchsvoll sind. Weiterhin ist iQuARI auch problemlos auf die 8-plex iTRAQ Variante anwendbar.

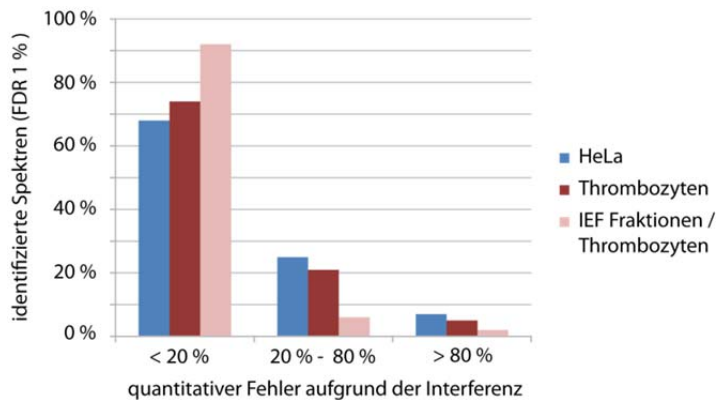


Abbildung 5.1: iQuARI erwies sich als robuste Methodik zur Analyse des quantitativen Fehlers in iTRAQ Experimenten. Die Interferenz nimmt proportional zur Komplexität der zu analysierenden Probe ab: Hochkomplexe HeLa Proben mit ca. 10.000 Proteinen (blau) weisen weniger Spektren mit einem quantitativen Fehler < 20 % auf als Thrombozyten mit ca. 5.000 Proteinen (rot) und geringkomplexe IEF Fraktionen von Thrombozyten (rosa) (Vaudel, Burkhardt et al., *J Proteome Res*, 2012).

5.2 Charakterisierung des humanen Thrombozytenproteoms

Thrombozyten spielen eine zentrale Rolle in der Hämostase als auch in der Entstehung und Progression kardiovaskulärer Erkrankungen, welche weltweit die Haupttodesursache sind (WHO 2012). Um die Prävalenz von kardiovaskulären Ereignissen zu reduzieren, ist die Thrombozyten fokussierte Medikation häufig Mittel der Wahl. In Studien konnte jedoch ein signifikanter Anteil der Patienten charakterisiert werden, der nicht auf die Medikation anspricht und somit einem erhöhtem Risiko für Herz-Kreislauf- und direkt korrelierte

sekundäre Erkrankungen ausgesetzt ist. Warum diese Patienten eine unzureichende Signalantwort liefern konnte bislang nicht geklärt werden, ist aber möglicherweise eine direkte Konsequenz unterschiedlicher Proteinexpressionsmuster und/oder auf Variationen im PTM-Muster individueller Patienten zurückzuführen.

5.2.1 Das quantitative Thrombozytenproteom

Mit der systematischen, quantitativen und umfangreichen Proteomanalyse von hochreinen humanen Thrombozyten konnte somit in dieser Arbeit die Grundlage für weitere biochemische Untersuchungen zur Analyse von Thrombozytenfunktionen und Mechanismen gelegt werden. Von ca. 4.000 identifizierten Proteinen konnten für ca. 3.700 die absoluten Proteinkopienzahlen geschätzt und damit das mit Abstand umfassendste Bild der Komposition humaner Thrombozyten präsentiert werden. Basierend auf Leukocount Analysen und den quantitativen proteomischen Daten konnte die Kontamination mit weiteren Blutkomponenten auf weniger als einen Leukozyten pro 10^6 Thrombozyten, ca. einen Erythrozyten pro 5.000 Thrombozyten und 1-1,5 % (V/V) Plasma pro Thrombozyt (entsprechend 0,1 fL pro ~ 10 fL Thrombozytenvolumen) eingeschränkt werden.

Die evaluierten Proteinkopienzahlen korrelieren stark mit den Werten aus der Literatur ($R^2=0.90$), obwohl die auf *spectral counting* basierende absolute Quantifizierung von Proteinen in Abhängigkeit von Strukturmerkmalen, wie Membranproteinen oder häufigem Auftreten von PTMs durchaus einen großen Fehler aufweisen kann. Weiterhin zeigt die Analyse mittels Reactome einen hohen Grad der Abdeckung für eine Vielzahl von Signalwegen in Thrombozyten, wie z.B. Ubiquitin-Proteasom (Abbildung 5.2), Aminosäuremetabolismus-, Calmodulin-, Zytoskelett-, Integrin-, MAPKinase- und PI3Kinase Netzwerke. Auch für essentielle Signalwege wie der Thrombozytenaktivierung/ Signaling/Aggregation (78,6 %) oder Thrombozytendegranulation (83,3 %) konnten hohe Abdeckungen erzielt werden.

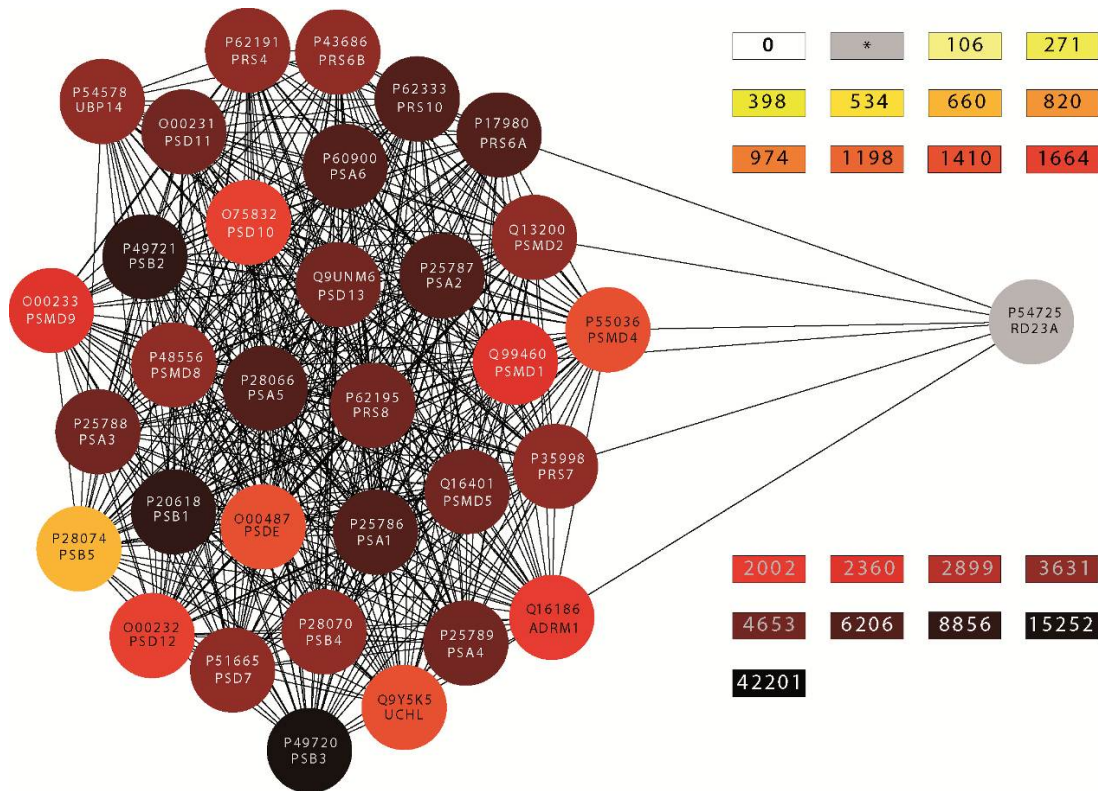


Abbildung 5.2: Abdeckung des Ubiquitin-Proteasom Netzwerks. Die Farben korrelieren mit Proteinkopienzahlen, wobei das grau unterlegte Protein HR23a als einziges nur qualitativ erfasst wurde.

5.2.2 Qualitativer wie quantitativer Abgleich des Thrombozytenproteoms mit globalen Proteomstudien humaner Zelllinien

Um mögliche Abweichungen der Proteinexpression und der Proteinkopienzahlen zwischen dem präsentierten Thrombozytenproteom und anderen quantitativen Proteomen humaner Zelllinien zu analysieren, wurden die Thrombozytendaten mit entsprechenden Daten von HeLa [7] und U2OS [6] Zellen abgeglichen. Die Kombination aller Datensätze umfasst bei Ausschluss von Splicevarianten eine Gesamtheit von 9.696 Proteinen. Nur 4 % der Proteine konnten ausschließlich in Thrombozyten identifiziert werden, während 69 % in mindestens zwei und 27 % in allen drei Datensätzen, Thrombozyten, HeLa- und U2OS-Zellen, gefunden wurden (Abb. 5.3 A). 32 % der thrombozytären Proteine sind in mindestens einer der humanen Zelllinien zu finden. Hingegen liegt die Überschneidung der abundantesten Proteine der drei Datensätze (95. Perzentil, 671 Proteine) entsprechend höher bei 58 % in allen drei Datensätzen und 88 % in mindestens zwei Datensätzen. Eine Minderheit von 28 Proteinen (4 %) konnte ausschließlich in Thrombozyten identifiziert werden: Zu diesen zählen γ -Aktin, Thrombozyten-Glykoproteine, Thrombozyten- und Koagulationsfaktoren.

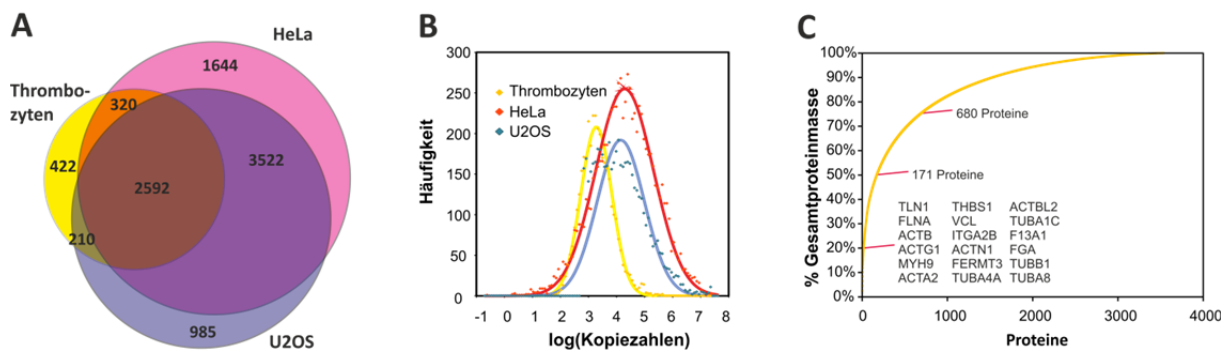


Abbildung 5.3: Vergleich der Proteome von Thrombozyten, HeLa und U2OS Zellen, bei einer Gesamtheit von 9.696 Proteinen. (A) Alle drei Proteome weisen eine starke Kongruenz auf, die offensichtlich unabhängig von den individuellen Kopiezahl ist. (B) Die Verteilung der Kopiezahl zeigt in allen drei Proteomen eine große Ähnlichkeit, wobei sie in Thrombozyten, aufgrund des geringeren Zellvolumens, zu niedrigeren Kopiezahl verschoben ist. (C) Prozentualer Beitrag zur Gesamtproteinmasse des thrombozytären Proteoms: 20 % der Gesamtmasse setzen sich aus nur 18 Proteinen, 50 % aus 171 Proteinen und 75 % aus 680 Proteinen zusammen (Burkhart et al., *Blood*, 2012).

Für alle drei untersuchten Zellen, HeLa, U2OS und Thrombozyten, konnten große, von Proteinkopiezahl unabhängige Ähnlichkeiten festgestellt werden, wobei das Maximum der Häufigkeitsverteilung der logarithmierten Kopiezahl μ und die Halbwertsbreite σ in Thrombozyten zu niedrigeren Werten verschoben sind (HeLa: $\mu = 4,4$, $\sigma = 1,03$; U2OS: $\mu = 4,2$, $\sigma = 0,85$; Thrombozyten $\mu = 3,16$, $\sigma = 0,55$, Abbildung 5.3 B). Dies entspricht der in der Literatur beschriebenen Korrelation zwischen Zellvolumina und Proteinexpressionslevels [159], da Thrombozyten nur einen Durchmesser von $\sim 3 \mu\text{m}$ im Vergleich zu HeLa mit $\sim 30 \mu\text{m}$ haben. Hoch abundante Proteine, die ausschließlich in Thrombozyten identifiziert wurden, konnten thrombozytenspezifischen Signalwegen zugeordnet werden. Die Verteilung der MW der Proteine ist für alle drei Datensätze ähnlich zum kompletten humanen Proteom (Swissprot) und weist die nicht-gel-basierte Proteomanalyse als geeignete Methodik ohne systematischen Fehler aus.

Bei Berechnung des prozentualen Beitrags zur Gesamtproteinmasse des humanen Thrombozytenproteoms zeigt sich, dass nur 18 Proteine 20 % der Gesamtproteinmasse ausmachen, bzw. 171 Proteine 50 % und 680 Proteine 75 % (Abbildung 5.3 C). Der Großteil des thrombozytären Proteoms ist damit im niedrig abundanten Bereich angesiedelt.

5.2.3 Variabilität des Thrombozytenproteoms

Die umfassende Thrombozytenstudie wurde durch einen quantitativen Vergleich von vier individuellen Blutspendern ergänzt. Hierzu wurden auf Proteinebene SDs zwischen den iTRAQ Ratios bzw. den NSAF Werten berechnet. Um Abweichungen während (a) der Probenaufarbeitung und (b) der LC-MS Analyse Rechnung zu tragen, wurde sowohl für die iTRAQ- als auch für die NSAF-Auswertung eine $SD > 2 \times SD_{\text{Median}}$ als Kriterium für eine potentielle Regulation definiert. Gleichwohl die biologische Varianz zwischen Menschen höher angenommen wird ($>>2x$), liegt die Mehrheit der Proteine in beiden quantitativen Analysen unterhalb des gewählten Grenzwertes und zeigt fast keine Variation zwischen unabhängigen Blutspendern. Innerhalb der iTRAQ basierten Analyse konnten ca. 1.900 Proteine mit Proteinkopienzahlen über vier Größenordnungen in allen vier Spendern quantifiziert werden, von welchen nur 15 % (234 Proteine) das Kriterium der $SD > 2 \times SD_{\text{median}}$ erfüllten und damit als potentiell reguliert definiert wurden. Innerhalb der auf *spectral counting* basierenden LF Analyse konnten in individuellen LC-MS Messungen von sieben SCX Fraktionen pro Spender hingegen 2.400 Proteine für alle vier Spender quantifiziert werden, wovon 12 % (288 Proteine) reguliert waren. Die regulierten Proteine beider Analysen können u.a. den Signalwegen Apoptose, adaptives Immunsystem und dem Proteinmetabolismus zugeordnet werden. Übergreifend konnten nur 28 Proteine in beiden Analysen als reguliert identifiziert werden (siehe Tabelle 4.5). Ähnliche Ergebnisse konnten auch in einer 2D-DIGE (fluoreszenzbasierte differentielle 2-DE) Analyse des Proteinexpressionsmuster von Patienten verschiedenen Alters aus dem Jahr 2008 gefunden werden [127]. Gel-basierte Proteomstudien sind jedoch durch die Ko-Migration von Proteinen und Proteinisoformen, die in einer, wie in der Studie angewandten, kombinierten MALDI-TOF/TOF Analyse nicht aufgelöst werden können, qualitativ und quantitativ fehlerbehaftet. Die Information der quantitativen Analyse von 898 Proteinen bei einer FDR von 5 % ist daher nur von begrenztem Wert.

Neben der Proteinvariabilität zwischen vier individuellen Spendern wurde weiterhin die Intravariabilität eines Spenders in drei unabhängigen Blutspenden untersucht. Bei dieser Analyse wurde auf eine Fraktionierung der Proben verzichtet. Die insgesamt 1.505 Proteine, die in allen drei Proben quantifiziert wurden, wiesen analog zur inter-Spender Analyse mit 15 % regulierten Proteinen eine hohe Stabilität auf.

Die Bildung der Thrombozyten durch Abschnürung aus ihren Vorläuferzellen, den Megakaryozyten, stellt somit einen hoch reproduzierbaren und regulierten Prozess dar.

Während für die iTRAQ basierte Studie eine Variation an Präfraktionierungstechniken (SCX, IEF, HILIC, COFRADIC) mit einer Vielzahl an Fraktionen angewendet wurde, stützte sich die LF Analyse auf sieben Fraktionen einer SCX Auftrennung von 50 µg eines jeden der vier Spender. Die fortschrittlichen Entwicklungen in der HPLC- und MS- Technologie erlaubten durch eine höhere Auflösung, eine höhere Frequenz in der Fragmentierung sowie ein akkurates Auslesen der Massen trotz geringeren Messaufwandes die Quantifizierung von ca. 25 % mehr Proteinen innerhalb der LF Analysen.

Zusammenfassend zeigen die Daten, dass bei einer sorgfältigen und kontrollierten Aufarbeitung, Prozessierung, Analyse und Auswertung von Thrombozyten aus Frischblut ein globaler Vergleich des Proteoms zwischen

- normalen und klar-definierten dysfunktionellen Thrombozyten
- Thrombozyten mit normaler sowie pathologischer Antwort auf Thrombozytenaggregationshemmer
- Thrombozyten von Patienten unterschiedlicher Stadien chronischer kardiovaskulärer und inflammatorischer Erkrankungen

möglich ist und bilden somit eine völlig neue Grundlage der krankheitsorientierten Thrombozytenforschung. Insbesondere erlaubt die optimierte Probenaufarbeitung in Kombination mit leistungsstarker LC-MS Instrumentation die Miniaturisierung der Analyse auf ein Minimum an Blut, Thrombozyten, Proteinmenge und Messzeit. Damit wird die Untersuchung von Erkrankungen, die nur eine geringe Blutabgabe zulassen, wie z.B. der Hämophilie (Bluterkrankheit) in der klinischen Anwendung möglich (bench to bedside).

5.3 Analyse von Organell-Subkompartimenten durch quantitative

Proteomforschung am Beispiel des IMS

Bislang war die Analyse von mitochondrialen Subproteomen im Prinzip auf die OM beschränkt [150]. Ein Grund hierfür war zumeist die unzureichende Reinheit der submitochondrialen Proteome für die LC-MS basierte Proteomanalyse [160]. Im Falle des IMS können kontaminierende Matrixproteine aufgrund von Transportprozessen nicht ausgeschlossen werden und müssen daher anhand dedizierter analytischer Strategien als solche definiert und vom IMS Proteom differenziert werden.

Dazu wurde eine Strategie entwickelt, in welcher ein konservierter biochemischer Mechanismus aus Säugerzellen auf Hefen transferiert wurde und in Kombination mit einer metabolischen stabil-Isotopen Markierung die Differenzierung zwischen Matrix- und IMS Proteinen erlaubte. Das Protein BAX - ein Schlüsselprotein in der Regulation der Apoptose bei Säugerzellen - bildet dabei Poren in der OM von Mitochondrien aus und führt somit zur Freisetzung von Proteinen aus dem IMS. So konnten von den derzeit 31 bekannten löslichen bzw. lose an eine Membran assoziierten IMS Proteinen 29 reproduzierbar identifiziert werden: Mit einem SN_{BAX}/Tot Verhältnis von theoretisch 1 unterschieden sie sich deutlich von Matrixproteinen, welche entsprechende Verhältnisse $\lll 1$ aufwiesen. Im ergänzenden SN_{BAX}/SN_{CON} Verhältnis konnten Matrixproteine nicht derart deutlich hinsichtlich ihrer Lokalisation unterschieden werden, wobei sie, verglichen mit den bekannten IMS Proteinen, auch hier tendenziell niedrigere Verhältnisse zeigten. Ein theoretisches SN_{BAX}/Tot Verhältnis von 1 stützt sich auf die Annahme, dass die Freisetzung der IMS Proteine vollständig stattfindet was in der Realität jedoch nicht immer der Fall ist. Dies zeigte sich v.a. für das Replikat 4, für welches aufgrund einer unzureichenden Freisetzung niedrige SN_{BAX}/Tot und SN_{BAX}/SN_{CON} Verhältnisse bestimmt werden konnten (siehe Anhang 12.6-8). Dank der Information aus den Replikaten 1 bis 3 und dank des Bezugs auf das Referenzproteom der bekannten IMS Proteine konnten dennoch 20 neue IMS Kandidaten identifiziert werden, von welchen 15 in der anschließenden biochemischen Validierung eindeutig als IMS Proteine bestätigt wurden.

Zehn der insgesamt 20 neu lokalisierten IMS Proteine wurden bislang in keiner proteomischen oder biochemischen Studie Mitochondrien zugeordnet, während zehn weitere anderen Kompartimenten als dem IMS zugeordnet waren. Darunter waren Qcr6 und Cox12, die als Komponenten der Atmungskettenkomplexe III und IV der IM zugeschrieben wurden. Gründe für das bivalente Auftreten von Qcr6 und Cox12 könnten eine lose Anheftung an der IM oder auch lösliche Formen sein.

Mit dem präsentierten analytischen Ansatz konnten weiterhin Informationen zum Importverhalten der IMS Proteine gewonnen werden. Für IMS Proteine sind sowohl der Präsequenzsignalweg als auch die MIA Importmaschinerie in der Literatur beschrieben [81, 161]. Dem Präsequenzimportweg werden mit Gut2, Cyb2, Mcr1 und dem in dieser Studie charakterisierten Ptc5 nur wenige IMS Proteine zugeschrieben, während die MIA Maschinerie (charakteristische Cystein-reiche Motive) mit insgesamt 16 IMS Proteinen -

darunter die vier neuen Kandidaten Pet191, Cox12, Coa6 und Mic23 - den vorwiegenden Importweg darstellt. Da die Mehrheit der IMS Proteine weder dem Präsequenzimportweg noch der MIA Maschinerie zugeordnet werden können, ist das Vorhandensein weiterer ungeklärter Importwege denkbar.

Innerhalb dieser Studie konnte für Coa6 (bisher als ORF Ymr244c-a geführt), weiterhin eine Funktion in der Ausbildung von Atmungskettenkomplexen nachgewiesen werden. Die Hypothese, dass es sich dabei um einen Assemblierungsfaktor handelt, liegt nahe, da Coa6 keine stöchiometrische Komponente eines Superkomplexes ist, *coa6Δ* Mutanten aber eine verminderte Ausprägung der Untereinheit des Komplexes IV aufweisen. Dies steht in Analogie zu bereits bekannten Cytochrom C Oxidase Assemblierungsdefiziten [162].

Es zeigt sich, dass die moderne Analytik nicht nur die Analyse von (1) Zellen, (2) Kompartimenten wie der Plasmamembran, Mitochondrien oder dem ER, (3) Proteinkomplexen sowie (4) PTMs ermöglicht. Stattdessen ist durch die geeignete Kombination von dedizierten, interdisziplinären Strategien, akkurater biochemischer Umsetzung sowie moderner quantitativer Analytik auch die Charakterisierung kleinster Subproteome mit hoher Abdeckung und Validität möglich. Die hier verwendete Strategie ist ein Prototyp für völlig neue Ansätze der räumlichen Auflösung bei der vollständigen Charakterisierung „des Proteoms“ in Raum und Zeit und bietet somit nie dagewesene Einblicke in den grundlegenden Aufbau der Zelle.

5.4 Mitochondriale Degradationsprozesse und die Definition von proteolytischen Schnittstellen

Auf Zellebene werden Proteine kontinuierlich in die zugrundeliegenden AS degradiert, um durch neu synthetisierte Proteine ersetzt zu werden [163]. Generell ist die Degradation in zellulären Systemen schwer zu messen. Mitochondrien, die mehr als 99 % ihrer Proteine aus dem Zytosol importieren weisen sich im isolierten Zustand daher als ideales System zur Analyse der globalen Degradation aus. In dieser Arbeit wurde alternativ zu Methoden wie COFRADIC die LF Quantifizierung gewählt, um die Stabilität des gesamten mitochondrialen Proteoms auf globaler Ebene quantitativ über die Zeit zu erfassen. Die LF Quantifizierung erlaubt dabei die quantitative Untersuchung von kompletten Proteinen, während Methoden wie COFRADIC nur das Auftreten einzelner Peptide aber nicht die Expression des Proteins berücksichtigen. Zudem erfordert COFRADIC die Verwendung erheblich größerer

Probenmengen (Faktor 5-10) bei einem deutlich höheren Zeitaufwand (ca. eine Woche pro Zeitpunkt) und ist weniger reproduzierbar.

Untersucht wurden Mitochondrien, die aus respiratorischen (YPG) bzw. fermentativen (YPD) Hefekulturen isoliert und über sieben Tage inkubiert wurden. Mittels der auf *Spectral counting* basierenden NSAF Quantifizierung konnte auf Proteinebene ein globaler Überblick über Abbauprozesse gewonnen werden. Weiterhin konnte durch die Berechnung der Steigung zwischen zwei zeitlich aufeinanderfolgenden NSAF Werten Informationen über die zugrundeliegende Abbaukinetik erlangt werden. Auch wenn hier klar gezeigt werden konnte, dass die gewählte Strategie für die globale Quantifizierung von Degradationsprozessen geeignet ist, so sind für eine detailreiche Validierung derselben jedoch weitere biochemische Untersuchungen notwendig, die (a) die Charakterisierung der beteiligten Proteasen sowie (b) weitere Aussagen zur Kinetik ermöglichen. Für die Mitochondrien der respiratorischen und fermentativen Zellkulturen zeigten sich deutliche Unterschiede. Für die Mitochondrien der YPG Kultur konnten mit 652 insgesamt geringfügig weniger Proteine identifiziert bzw. NSAF Werte ermittelt werden als für Mitochondrien der YPD Kultur (693 Proteine). Während für YPG 269 Proteine über die Zeit abfallende Tendenzen und 393 Protein einen stabilen Charakter aufwiesen, waren es für Mitochondrien der YPD Kultur mit 421 Proteinen deutlich mehr mit abfallendem Charakter. Dies deutet auf eine Stabilisierung des mitochondrialen Proteoms im Falle respiratorischer Wachstumsbedingungen hin, d.h. dass die Zellatmung in ihrer essentiellen Funktion zu einer Stabilisierung des mitochondrialen Proteoms führt, während unter fermentativen Bedingungen, d.h. unter Anwesenheit von Glucose als Energielieferant, im Vergleich eine Destabilisierung des Proteoms vorliegt. Nachdem das mitochondriale Proteom, wie bereits gezeigt, durch Phosphorylierungsprozesse metabolisch reguliert wird [164, 165], könnte hier eine ergänzende globale Phosphorylierungsstudie von großem Interesse sein. Zeitaufgelöste und quantitative Phosphorylierungsmuster könnten weitere Informationen über die Aktivierungszustände von Proteasen liefern und somit zu deren Identifizierung und Charakterisierung beitragen. Ein Proteasom-artiger Abbau über Ubiquitylierung konnte durch eine entsprechende Mascotsuche mit der variablen Modifikation GlyGly an Lys nahezu ausgeschlossen werden. Die Ubiquitylierung führt zur kovalenten Verknüpfung der Sequenz Ubiquitin-ArgGlyGly-Lys an die ϵ -Aminogruppe von Lysinen, welche somit für den proteolytischen Verdau nicht zugänglich sind [166]. Dies führt dazu, dass stattdessen beim proteolytischen Verdau die Sequenz GlyGly an der Seitenkette

des entsprechenden Lys-Rests über eine Isoeptidbindung verbleibt, die entsprechend massenspektrometrisch nachgewiesen werden kann.

Von den charakterisierten Proteinen konnten, mit Hilfe der *Saccharomyces* Genome Database (www.yeastgenome.org), für beide Wachstumsbedingungen jeweils etwas mehr als die Hälfte einem der vier Kompartimente zugeordnet werden. Auffallend dabei war die starke Präsenz der IMS Proteine unter den abfallenden Proteinen: Mit 80 % (YPG) bzw. 67 % (YPD) der in dieser Arbeit 50 identifizierten IMS Proteine scheint der IMS besonders stark proteolytischen Prozessen zu unterliegen. Der einheitliche Charakter dieser Proteine ist wiederum ein Beleg für eine gemeinsame Organisation in einem Kompartiment.

Im Fokus der *Feature*-basierten Quantifizierung stand insbesondere die Charakterisierung der Proteaseaktivität. Während die Einteilung der atypischen Peptide in Gruppen mit ansteigender bzw. abfallender Tendenz bereits charakteristische Sequenzmotive aufzeigt, konnte die anschließende Einteilung in die vier Kompartimente entsprechend deutlichere Trends darstellen. Beispielsweise scheinen im IMS befindliche Proteasen auch mit IM assoziierten Proteinen zu interagieren, deren terminale Sequenzen in den IMS reichen. Eine nähere Definition der proteolytischen Schnittstellen ist daher notwendig. Neben der Untersuchung von Proben mit kürzeren Zeitintervallen als 24 h könnten mit Hilfe von synthetischen Peptidanaloga der atypischen Peptide weitere Einblicke über (a) proteolytische Schnittstellen bzw. Konsensussequenzen, (b) Lokalisation und (c) Effizienz der Proteasen gewonnen werden.

Zusammenfassend konnte gezeigt werden, dass ein alternativer quantitativer MS Ansatz genutzt werden kann, um Degradationsprozesse in Mitochondrien zeitaufgelöst und global zu erfassen. Diese Grundlage muss nun in weiteren umfassenden Arbeiten ausgenutzt werden, um neben dem mitochondrialen Proteinimport auch neue Erkenntnisse über die Proteindynamik / -stabilität zu gewinnen.

6. Zusammenfassung

In der vorliegenden Arbeit konnten, basierend auf niedrig komplexen Modellproteomen, Methoden für die qualitätsbasierte und quantitative Proteomanalyse mittels hochauflösender LC-MS entwickelt und angewendet werden. (1) Der proteolytische Verdau wurde bezüglich Effizienz, Spezifität und Reproduzierbarkeit evaluiert und optimiert. (2) Ein stabil-Isotopen markierter Peptidstandard, welcher die online-Kontrolle der LC-MS Performance bei LF Experimenten erlaubt, wurde erstellt. (3) Der Arbeitsablauf von iTRAQ Experimenten sowie die dazugehörige Datenauswertung wurden evaluiert und optimiert. (4) Eine Strategie zur Bestimmung des Ausmaßes der Interferenz bei Reporterionen-basierten Markierungsstrategien, iQuARI, wurde entwickelt. (5) Weiterhin wurde die Auswertung von proteomischen Datensätzen mit Hilfe der FDR untersucht und optimiert.

Diese entwickelten Strategien wurden erfolgreich zur Untersuchung komplexer Proteome angewendet. So konnte mit ca. 4.000 Proteinen die bisher umfassendste und einzige quantitative Studie des Thrombozytenproteoms realisiert werden. Die Analyse der Intra- sowie Inter-Proteinvariabilität zeigte mit verschiedenen quantitativen Strategien (iTRAQ, LF) einen hohen Grad an Stabilität des Thrombozytenproteoms.

Die Anwendung von SILAC ermöglichte die Charakterisierung des Proteoms des mitochondrialen IMS von *S. cerevisiae*; dieser war erst durch die Anwendung einer dedizierten Strategie (Freisetzung von IMS Proteinen mittels BAX) zugänglich. So konnten insgesamt 50 Proteine eindeutig dem IMS zugeordnet und für 15 von 20 neuen IMS Proteinen diese Lokalisation biochemisch verifiziert werden.

Für die Analyse von Degradationsprozessen in Mitochondrien konnten zwei verschiedene Ansätze zur Kontrolle der Proteinmengen mittels LC-MS etabliert werden. Vorläufige Ergebnisse weisen auf funktionale räumliche Einheiten für den Wirkungsraum von Proteasen hin, die eine weitere, detailreiche Charakterisierung erfordern.

Folglich konnte in dieser Arbeit gezeigt werden, dass die qualitätskontrollierte, quantitative LC-MS Analyse vollkommen neuartige Einblicke ermöglicht: von der Zusammensetzung klinisch hochrelevanter Patientenproben und bisher nicht zugänglicher Organell-Subkompartimente bis hin zu komplexen Mechanismen der Proteindegradation.

7. Summary

In the present work, methods for quality based and quantitative proteome analysis using high resolution LC-MS were developed based on low complex model proteomes, and applied to mitochondria (*S.cerevisiae*) and human platelets. (1) The proteolytic digest was quality controlled (QC) and optimized regarding efficiency, specificity, and reproducibility. (2) A stable isotope coded peptide standard was developed, allowing for online QC of LC-MS performance in LF experiments. (3) The iTRAQ workflow, including statistical data interpretation, was QC and optimized. (4) A method to assess interference in reporter ion-based quantification strategies, named iQuARI, was developed. (5) Moreover, data evaluation of complex proteomic datasets using the FDR approach was investigated and optimized.

The developed strategies were successfully applied to investigate complex proteomes. Thus, the human platelet proteome could be characterized with ca. 4,000 proteins on a qualitative and quantitative level, representing the most comprehensive and first quantitative study on this primary tissue to date. Additionally, a focused analysis of the intra- and inter-protein variability with different quantitative strategies (iTRAQ, LF), revealed that the human proteome is remarkably stable.

The characterization of the mitochondrial intermembrane space proteome of *S. cerevisiae* was accomplished using SILAC in combination with a dedicated strategy, including the pro-apoptotic protein BAX. By doing so, 50 proteins could be localized unambiguously to the IMS, and for 15 out of 20 new IMS proteins this localization could be validated biochemically. For analyzing protein degradation in mitochondria, two different strategies using LC-MS were developed to account for varying protein amounts over time. Preliminary results indicate functional and spatial entities with differences in regional protease activity; for a more detailed characterization further studies have to be conducted.

To sum up, this work demonstrates that utilizing quality control-based quantitative LC-MS analysis can provide novel insights, not only into the composition of patient samples of high clinical relevance and almost inaccessible sub-compartments, but also into highly complex mechanisms of protein degradation.

8. Literaturverzeichnis

1. Sickmann A, Mreyen M, Meyer HE: Mass spectrometry--a key technology in proteome research. *Adv Biochem Eng Biotechnol* 2003, 83:141-176.
2. <http://prospector.ucsf.edu>
3. Vogel C, Marcotte EM: Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat Rev Genet* 2012, 13(4):227-232.
4. Schwanhausser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M: Global quantification of mammalian gene expression control. *Nature* 2011, 473(7347):337-342.
5. Finishing the euchromatic sequence of the human genome. *Nature* 2004, 431(7011):931-945.
6. Beck M, Schmidt A, Malmstroem J, Claassen M, Ori A, Szymborska A, Herzog F, Rinner O, Ellenberg J, Aebersold R: The quantitative proteome of a human cell line. *Mol Syst Biol* 2011, 7:549.
7. Nagaraj N, Wisniewski JR, Geiger T, Cox J, Kircher M, Kelso J, Paabo S, Mann M: Deep proteome and transcriptome mapping of a human cancer cell line. *Mol Syst Biol* 2011, 7:548.
8. Edman P: A method for the determination of the amino acid sequence in peptides. *Arch Biochem*, 1949, 22:475-476.
9. O'Farrell PH: High resolution two-dimensional electrophoresis of proteins. *J Biol Chem* 1975, 250(10):4007-4021.
10. Rabilloud T, Lelong C: Two-dimensional gel electrophoresis in proteomics: a tutorial. *J Proteomics* 2011, 74(10):1829-1841.
11. Lottspeich F, Engels JW (Eds.): *Bioanalytik. Volume 2*: Spektrum Akademischer Verlag; 2006.
12. van Deemter JJ, Zuiderweg FJ, Klinkenberg A: Longitudinal diffusion and resistance to mass transfer as causes of nonideality in chromatography. *Chem Eng Sci* 1956, 5:271-289.
13. Horvath CG, Preiss BA, Lipsky SR: Fast liquid chromatography: an investigation of operating parameters and the separation of nucleotides on pellicular ion exchangers. *Anal Chem* 1967, 39(12):1422-1428.
14. Giddings JC: Maximum number of components resolvable by gel filtration and other elution chromatographic methods. *Analytical chemistry* 1967, 39:1027-1028.
15. MacNair JE, Patel KD, Jorgenson JW: Ultrahigh-pressure reversed-phase capillary liquid chromatography: isocratic and gradient elution using columns packed with 1.0-micron particles. *Anal Chem* 1999, 71(3):700-708.
16. Wilm M, Shevchenko A, Houthaeve T, Breit S, Schweigerer L, Fotsis T, Mann M: Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. *Nature* 1996, 379(6564):466-469.
17. Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM: Electrospray ionization for mass spectrometry of large biomolecules. *Science* 1989, 246(4926):64-71.
18. Michalski A, Cox J, Mann M: More than 100,000 detectable peptide species elute in single shotgun proteomics runs but the majority is inaccessible to data-dependent LC-MS/MS. *J Proteome Res* 2011, 10(4):1785-1793.

19. Kocher T, Swart R, Mechtler K: Ultra-high-pressure RPLC hyphenated to an LTQ-Orbitrap Velos reveals a linear relation between peak capacity and number of identified peptides. *Analytical chemistry* 2011, 83(7):2699-2704.
20. Delmotte N, Lasaosa M, Tholey A, Heinzle E, Huber CG: Two-dimensional reversed-phase x ion-pair reversed-phase HPLC: an alternative approach to high-resolution peptide separation for shotgun proteome analysis. *J Proteome Res* 2007, 6(11):4363-4373.
21. Giddings JC: Sample dimensionality: a predictor of order-disorder in component peak distribution in multidimensional separation. *J Chromatogr A* 1995, 703(1-2):3-15.
22. Washburn MP, Wolters D, Yates JR, 3rd: Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol* 2001, 19(3):242-247.
23. Makarov A: Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. *Analytical chemistry* 2000, 72(6):1156-1162.
24. Gilar M, Olivova P, Daly AE, Gebler JC: Orthogonality of separation in two-dimensional liquid chromatography. *Analytical chemistry* 2005, 77(19):6426-6434.
25. Klose J: Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals. *Humangenetik* 1975, 26(3):231-243.
26. Bjellqvist B, Ek K, Righetti PG, Gianazza E, Gorg A, Westermeier R, Postel W: Isoelectric focusing in immobilized pH gradients: principle, methodology and some applications. *J Biochem Biophys Methods* 1982, 6(4):317-339.
27. Michel PE, Reymond F, Arnaud IL, Josserand J, Girault HH, Rossier JS: Protein fractionation in a multicompartiment device using Off-Gel isoelectric focusing. *Electrophoresis* 2003, 24(1-2):3-11.
28. Horth P, Miller CA, Preckel T, Wenz C: Efficient fractionation and improved protein identification by peptide OFFGEL electrophoresis. *Mol Cell Proteomics* 2006, 5(10):1968-1974.
29. Alpert AJ, Andrews PC: Cation-exchange chromatography of peptides on poly(2-sulfoethyl aspartamide)-silica. *J Chromatogr* 1988, 443:85-96.
30. Beausoleil SA, Jedrychowski M, Schwartz D, Elias JE, Villen J, Li J, Cohn MA, Cantley LC, Gygi SP: Large-scale characterization of HeLa cell nuclear phosphoproteins. *Proc Natl Acad Sci U S A* 2004, 101(33):12130-12135.
31. Lewandrowski U, Zahedi RP, Moebius J, Walter U, Sickmann A: Enhanced N-glycosylation site analysis of sialoglycopeptides by strong cation exchange prefractionation applied to platelet plasma membranes. *Mol Cell Proteomics* 2007, 6(11):1933-1941.
32. Mohammed S, Heck A, Jr.: Strong cation exchange (SCX) based analytical methods for the targeted analysis of protein post-translational modifications. *Curr Opin Biotechnol* 2011, 22(1):9-16.
33. Alpert AJ: Hydrophilic-interaction chromatography for the separation of peptides, nucleic acids and other polar compounds. *J Chromatogr* 1990(499):177-196.
34. Di Palma S, Boersema PJ, Heck AJ, Mohammed S: Zwitterionic hydrophilic interaction liquid chromatography (ZIC-HILIC and ZIC-cHILIC) provide high resolution separation and increase sensitivity in proteome analysis. *Anal Chem* 2011, 83(9):3440-3447.
35. Gevaert K, Goethals M, Martens L, Van Damme J, Staes A, Thomas GR, Vandekerckhove J: Exploring proteomes and analyzing protein processing by mass

- spectrometric identification of sorted N-terminal peptides. *Nat Biotechnol* 2003, 21(5):566-569.
36. Gevaert K, Vandekerckhove J: Reverse-phase diagonal chromatography for phosphoproteome research. *Methods Mol Biol* 2009, 527:219-227, ix.
37. Ghesquiere B, Buyl L, Demol H, Van Damme J, Staes A, Timmerman E, Vandekerckhove J, Gevaert K: A new approach for mapping sialylated N-glycosites in serum proteomes. *J Proteome Res* 2007, 6(11):4304-4312.
38. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS: Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 1999, 20(18):3551-3567.
39. Yates JR, 3rd, Eng JK, McCormack AL, Schieltz D: Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database. *Analytical chemistry* 1995, 67(8):1426-1436.
40. Geer LY, Markey SP, Kowalak JA, Wagner L, Xu M, Maynard DM, Yang X, Shi W, Bryant SH: Open mass spectrometry search algorithm. *J Proteome Res* 2004, 3(5):958-964.
41. Fenyo D, Beavis RC: A method for assessing the statistical significance of mass spectrometry-based protein identifications using general scoring schemes. *Analytical chemistry* 2003, 75(4):768-774.
42. Venable JD, Dong MQ, Wohlschlegel J, Dillin A, Yates JR: Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra. *Nat Methods* 2004, 1(1):39-45.
43. Silva JC, Denny R, Dorschel CA, Gorenstein M, Kass IJ, Li GZ, McKenna T, Nold MJ, Richardson K, Young P *et al*: Quantitative proteomic analysis by accurate mass retention time pairs. *Analytical chemistry* 2005, 77(7):2187-2200.
44. Panchaud A, Scherl A, Shaffer SA, von Haller PD, Kulasekara HD, Miller SI, Goodlett DR: Precursor acquisition independent from ion count: how to dive deeper into the proteomics ocean. *Analytical chemistry* 2009, 81(15):6481-6488.
45. Geromanos SJ, Vissers JP, Silva JC, Dorschel CA, Li GZ, Gorenstein MV, Bateman RH, Langridge JI: The detection, correlation, and comparison of peptide precursor and product ions from data independent LC-MS with data dependant LC-MS/MS. *Proteomics* 2009, 9(6):1683-1695.
46. Blackburn K, Mbeunkui F, Mitra SK, Mentzel T, Goshe MB: Improving protein and proteome coverage through data-independent multiplexed peptide fragmentation. *J Proteome Res* 2010, 9(7):3621-3637.
47. Conrads TP, Anderson GA, Veenstra TD, Pasa-Tolic L, Smith RD: Utility of accurate mass tags for proteome-wide protein identification. *Anal Chem* 2000, 72(14):3349-3354.
48. Strittmatter EF, Ferguson PL, Tang K, Smith RD: Proteome analyses using accurate mass and elution time peptide tags with capillary LC time-of-flight mass spectrometry. *J Am Soc Mass Spectrom* 2003, 14(9):980-991.
49. Liu H, Sadygov RG, Yates JR, 3rd: A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal Chem* 2004, 76(14):4193-4201.
50. Picotti P, Bodenmiller B, Mueller LN, Domon B, Aebersold R: Full dynamic range proteome analysis of *S. cerevisiae* by targeted proteomics. *Cell* 2009, 138(4):795-806.
51. Lange V, Picotti P, Domon B, Aebersold R: Selected reaction monitoring for quantitative proteomics: a tutorial. *Mol Syst Biol* 2008, 4:222.

52. Maclean B, Tomazela DM, Abbatiello SE, Zhang S, Whiteaker JR, Paulovich AG, Carr SA, Maccoss MJ: Effect of collision energy optimization on the measurement of peptides by selected reaction monitoring (SRM) mass spectrometry. *Analytical chemistry* 2010, 82(24):10116-10124.
53. Gorg A, Weiss W, Dunn MJ: Current two-dimensional electrophoresis technology for proteomics. *Proteomics* 2004, 4(12):3665-3685.
54. Klose J, Kobalz U: Two-dimensional electrophoresis of proteins: an updated protocol and implications for a functional analysis of the genome. *Electrophoresis* 1995, 16(6):1034-1059.
55. Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R: Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol* 1999, 17(10):994-999.
56. Dayon L, Hainard A, Licker V, Turck N, Kuhn K, Hochstrasser DF, Burkhard PR, Sanchez JC: Relative quantification of proteins in human cerebrospinal fluids by MS/MS using 6-plex isobaric tags. *Analytical chemistry* 2008, 80(8):2921-2931.
57. Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S *et al*: Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics* 2004, 3(12):1154-1169.
58. Schmidt A, Kellermann J, Lottspeich F: A novel strategy for quantitative proteomics using isotope-coded protein labels. *Proteomics* 2005, 5(1):4-15.
59. Ishihama Y, Oda Y, Tabata T, Sato T, Nagasu T, Rappsilber J, Mann M: Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol Cell Proteomics* 2005, 4(9):1265-1272.
60. Paoletti AC, Parmely TJ, Tomomori-Sato C, Sato S, Zhu D, Conaway RC, Conaway JW, Florens L, Washburn MP: Quantitative proteomic analysis of distinct mammalian Mediator complexes using normalized spectral abundance factors. *Proc Natl Acad Sci U S A* 2006, 103(50):18928-18933.
61. Kay R, Barton C, Ratcliffe L, Matharoo-Ball B, Brown P, Roberts J, Teale P, Creaser C: Enrichment of low molecular weight serum proteins using acetonitrile precipitation for mass spectrometry based proteomic analysis. *Rapid Commun Mass Spectrom* 2008, 22(20):3255-3260.
62. Simpson DM, Beynon RJ: Acetone precipitation of proteins and the modification of peptides. *J Proteome Res* 2010, 9(1):444-450.
63. Olsen JV, Ong SE, Mann M: Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Mol Cell Proteomics* 2004, 3(6):608-614.
64. Rodriguez J, Gupta N, Smith RD, Pevzner PA: Does trypsin cut before proline? *J Proteome Res* 2008, 7(1):300-305.
65. Page MJ, Di Cera E: Serine peptidases: classification, structure and function. *Cell Mol Life Sci* 2008, 65(7-8):1220-1236.
66. Rudnick PA, Clauser KR, Kilpatrick LE, Tchekhovskoi DV, Neta P, Blonder N, Billheimer DD, Blackman RK, Bunk DM, Cardasis HL *et al*: Performance metrics for liquid chromatography-tandem mass spectrometry systems in proteomics analyses. *Mol Cell Proteomics* 2010, 9(2):225-241.
67. Tabb DL, Vega-Montoto L, Rudnick PA, Variyath AM, Ham AJ, Bunk DM, Kilpatrick LE, Billheimer DD, Blackman RK, Cardasis HL *et al*: Repeatability and reproducibility in

- proteomic identifications by liquid chromatography-tandem mass spectrometry. *J Proteome Res* 2010, 9(2):761-776.
68. Piening BD, Wang P, Bangur CS, Whiteaker J, Zhang H, Feng LC, Keane JF, Eng JK, Tang H, Prakash A *et al*: Quality control metrics for LC-MS feature detection tools demonstrated on *Saccharomyces cerevisiae* proteomic profiles. *J Proteome Res* 2006, 5(7):1527-1534.
69. Xu H, Freitas MA: Automated diagnosis of LC-MS/MS performance. *Bioinformatics (Oxford, England)* 2009, 25(10):1341-1343.
70. Elias JE, Haas W, Faherty BK, Gygi SP: Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations. *Nat Methods* 2005, 2(9):667-675.
71. Arnott DP, al. e: A Proteomics Standard. *ABRF 2006 Conference, Long beach, CA* 2006.
72. Kölliker A: *Zeitschrift für wissenschaftl Zoologie* 1856, VIII:311-318.
73. Altmann R: Die Elementarorganismen und ihre Beziehungen zu den Zellen. *Veit & Comp, Leipzig* 1890.
74. Benda C: *Arch Anat Physiol* 1898, 73:393-398.
75. Sagan L: On the origin of mitosing cells. *J Theor Biol* 1967, 14(3):255-274.
76. Kutik S, Stroud DA, Wiedemann N, Pfanner N: Evolution of mitochondrial protein biogenesis. *Biochim Biophys Acta* 2009, 1790(6):409-415.
77. Nass MM, Nass S: Intramitochondrial Fibers with DNA Characteristics. *J Cell Biol* 1963, 19:593-629.
78. Milenkovic D, Muller J, Stojanovski D, Pfanner N, Chacinska A: Diverse mechanisms and machineries for import of mitochondrial proteins. *Biol Chem* 2007, 388(9):891-897.
79. Becker T, Wenz LS, Kruger V, Lehmann W, Muller JM, Goroncy L, Zufall N, Lithgow T, Guiard B, Chacinska A *et al*: The mitochondrial import protein Mim1 promotes biogenesis of multispanning outer membrane proteins. *J Cell Biol* 2011, 194(3):387-395.
80. Vogtle FN, Wortelkamp S, Zahedi RP, Becker D, Leidhold C, Gevaert K, Kellermann J, Voos W, Sickmann A, Pfanner N *et al*: Global analysis of the mitochondrial N-proteome identifies a processing peptidase critical for protein stability. *Cell* 2009, 139(2):428-439.
81. Stojanovski D, Muller JM, Milenkovic D, Guiard B, Pfanner N, Chacinska A: The MIA system for protein import into the mitochondrial intermembrane space. *Biochim Biophys Acta* 2008, 1783(4):610-617.
82. Zalman LS, Nikaido H, Kagawa Y: Mitochondrial outer membrane contains a protein producing nonspecific diffusion channels. *J Biol Chem* 1980, 255(5):1771-1774.
83. Martinou JC, Youle RJ: Mitochondria in apoptosis: Bcl-2 family members and mitochondrial dynamics. *Dev Cell* 2011, 21(1):92-101.
84. de Moura MB, dos Santos LS, Van Houten B: Mitochondrial dysfunction in neurodegenerative diseases and cancer. *Environ Mol Mutagen* 2010, 51(5):391-405.
85. Hartwig J, Italiano J, Jr.: The birth of the platelet. *J Thromb Haemost* 2003, 1(7):1580-1586.
86. Varga-Szabo D, Pleines I, Nieswandt B: Cell adhesion mechanisms in platelets. *Arterioscler Thromb Vasc Biol* 2008, 28(3):403-412.
87. Ruggeri ZM: Platelets in atherothrombosis. *Nat Med* 2002, 8(11):1227-1234.
88. Qureshi AH, Chaoji V, Maignel D, Faridi MH, Barth CJ, Salem SM, Singhal M, Stoub D, Krastins B, Ogihara M *et al*: Proteomic and phospho-proteomic profile of human

- platelets in basal, resting state: insights into integrin signaling. *PLoS One* 2009, 4(10):e7627.
89. Lewandrowski U, Wortelkamp S, Lohrig K, Zahedi RP, Wolters DA, Walter U, Sickmann A: Platelet membrane proteomics: a novel repository for functional research. *Blood* 2009, 114(1):e10-19.
 90. Winkler C, Denker K, Wortelkamp S, Sickmann A: Silver- and Coomassie-staining protocols: detection limits and compatibility with ESI MS. *Electrophoresis* 2007, 28(12):2095-2099.
 91. Gevaert K, Van Damme J, Goethals M, Thomas GR, Hoorelbeke B, Demol H, Martens L, Puype M, Staes A, Vandekerckhove J: Chromatographic isolation of methionine-containing peptides for gel-free proteome analysis: identification of more than 800 *Escherichia coli* proteins. *Mol Cell Proteomics* 2002, 1(11):896-903.
 92. Gevaert K, Ghesquiere B, Staes A, Martens L, Van Damme J, Thomas GR, Vandekerckhove J: Reversible labeling of cysteine-containing peptides allows their specific chromatographic isolation for non-gel proteome studies. *Proteomics* 2004, 4(4):897-908.
 93. Beck F, Lewandrowski U, Wiltfang M, Feldmann I, Geiger J, Sickmann A, Zahedi RP: The good, the bad, the ugly: validating the mass spectrometric analysis of modified peptides. *Proteomics* 2011, 11(6):1099-1109.
 94. Mitulovic G, Smoluch M, Chervet JP, Steinmacher I, Kungl A, Mechtler K: An improved method for tracking and reducing the void volume in nano HPLC-MS with micro trapping columns. *Anal Bioanal Chem* 2003, 376(7):946-951.
 95. Olsen JV, Schwartz JC, Griep-Raming J, Nielsen ML, Damoc E, Denisov E, Lange O, Remes P, Taylor D, Splendore M *et al*: A dual pressure linear ion trap Orbitrap instrument with very high sequencing speed. *Mol Cell Proteomics* 2009, 8(12):2759-2769.
 96. Lottspeich F, Engels JW (Eds.): *Bioanalytik. Volume 2*. Heidelberg: Spektrum Akademischer Verlag; 2006.
 97. Peptide Stability and Potential Degradation
 98. Burkhardt JM, Premisler T, Sickmann A: Quality control of nano-LC-MS systems using stable isotope-coded peptides. *Proteomics* 2011, 11(6):1049-1057.
 99. Vaudel M, Barsnes H, Berven FS, Sickmann A, Martens L: SearchGUI: An open-source graphical user interface for simultaneous OMSSA and X!Tandem searches. *Proteomics* 2011, 11(5):996-999.
 100. Sturm M, Bertsch A, Gropl C, Hildebrandt A, Hussong R, Lange E, Pfeifer N, Schulz-Trieglaff O, Zerck A, Reinert K *et al*: OpenMS - an open-source software framework for mass spectrometry. *BMC Bioinformatics* 2008, 9:163.
 101. Martens L, Vandekerckhove J, Gevaert K: DBToolkit: processing protein databases for peptide-centric proteomics. *Bioinformatics* 2005, 21(17):3584-3585.
 102. Vaudel M, Burkhardt JM, Sickmann A, Martens L, Zahedi RP: Peptide identification quality control. *Proteomics* 2011, 11(10):2105-2114.
 103. Cox J, Mann M: MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* 2008, 26(12):1367-1372.
 104. Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M: Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res* 2011, 10(4):1794-1805.

105. Vaudel M, Burkhardt JM, Radau S, Zahedi RP, Martens L, Sickmann A: Integral Quantification Accuracy estimation for Reporter Ion based quantitative proteomics (iQuARI). *Journal of Proteome Research* 2012.
106. Burkhardt JM, Vaudel M, Zahedi RP, Martens L, Sickmann A: iTRAQ protein quantification: A quality-controlled workflow. *Proteomics* 2011, 11(6):1125-1134.
107. Reinders J, Zahedi RP, Pfanner N, Meisinger C, Sickmann A: Toward the complete yeast mitochondrial proteome: multidimensional separation techniques for mitochondrial proteomics. *J Proteome Res* 2006, 5(7):1543-1554.
108. Sigma-Aldrich: Storage & Handling Synthetic Peptides Guidelines. Castle Hill; 2009.
109. Colaert N, Helsens K, Martens L, Vandekerckhove J, Gevaert K: Improved visualization of protein consensus sequences by iceLogo. *Nat Methods* 2009, 6(11):786-787.
110. Dowal L, Yang W, Freeman MR, Steen H, Flaumenhaft R: Proteomic analysis of palmitoylated platelet proteins. *Blood* 2011, 118(13):e62-73.
111. Meisinger C, Ryan MT, Hill K, Model K, Lim JH, Sickmann A, Muller H, Meyer HE, Wagner R, Pfanner N: Protein import channel of the outer mitochondrial membrane: a highly stable Tom40-Tom22 core structure differentially interacts with preproteins, small tom proteins, and import receptors. *Mol Cell Biol* 2001, 21(7):2337-2348.
112. Sickmann A, Reinders J, Wagner Y, Joppich C, Zahedi R, Meyer HE, Schonfisch B, Perschil I, Chacinska A, Guiard B *et al*: The proteome of *Saccharomyces cerevisiae* mitochondria. *Proc Natl Acad Sci U S A* 2003, 100(23):13207-13212.
113. Sipos T, Merkel JR: An effect of calcium ions on the activity, heat stability, and structure of trypsin. *Biochemistry* 1970, 9(14):2766-2775.
114. Proc JL, Kuzyk MA, Hardie DB, Yang J, Smith DS, Jackson AM, Parker CE, Borchers CH: A quantitative study of the effects of chaotropic agents, surfactants, and solvents on the digestion efficiency of human plasma proteins by trypsin. *J Proteome Res* 2010, 9(10):5422-5437.
115. Russell WK, Park ZY, Russell DH: Proteolysis in mixed organic-aqueous solvent systems: applications for peptide mass mapping using mass spectrometry. *Analytical chemistry* 2001, 73(11):2682-2685.
116. Ow SY, Salim M, Noirel J, Evans C, Rehman I, Wright PC: iTRAQ underestimation in simple and complex mixtures: "the good, the bad and the ugly". *J Proteome Res* 2009, 8(11):5347-5355.
117. Shevlyakov G, Morgenthaler S, Shurygin A: Redescending M-estimators. *Journal of Statistical Planning and Inference* 2008, 138:2906-2917.
118. Ting L, Rad R, Gygi SP, Haas W: MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics. *Nat Methods* 2011, 8(11):937-940.
119. Wenger CD, Lee MV, Hebert AS, McAlister GC, Phanstiel DH, Westphall MS, Coon JJ: Gas-phase purification enables accurate, multiplexed proteome quantification with isobaric tagging. *Nat Methods* 2011, 8(11):933-935.
120. Vaudel M, Burkhardt JM, Breiter D, Zahedi RP, Sickmann A, Martens L: A complex standard for protein identification, designed by evolution. *J Proteome Res* 2012.
121. Keil-Dlouha VV, Zylber N, Imhoff J, Tong N, Keil B: Proteolytic activity of pseudotrypsin. *FEBS Lett* 1971, 16(4):291-295.
122. Kostka V, Carpenter FH: Inhibition of Chymotrypsin Activity in Crystalline Trypsin Preparations. *J Biol Chem* 1964, 239:1799-1803.
123. Gerber SA, Rush J, Stemman O, Kirschner MW, Gygi SP: Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc Natl Acad Sci U S A* 2003, 100(12):6940-6945.

124. Sadygov RG, Cociorva D, Yates JR, 3rd: Large-scale database searching using tandem mass spectra: looking up the answer in the back of the book. *Nat Methods* 2004, 1(3):195-202.
125. Creasy DM, Cottrell JS: Error tolerant searching of uninterpreted tandem mass spectrometry data. *Proteomics* 2002, 2(10):1426-1434.
126. Arnesen T, Van Damme P, Polevoda B, Helsens K, Evjenth R, Colaert N, Varhaug JE, Vandekerckhove J, Lillehaug JR, Sherman F *et al*: Proteomics analyses reveal the evolutionary conservation and divergence of N-terminal acetyltransferases from yeast and humans. *Proc Natl Acad Sci U S A* 2009, 106(20):8157-8162.
127. Winkler W, Zellner M, Diestinger M, Babeluk R, Marchetti M, Goll A, Zehetmayer S, Bauer P, Rappold E, Miller I *et al*: Biological variation of the platelet proteome in the elderly population and its implication for biomarker research. *Mol Cell Proteomics* 2008, 7(1):193-203.
128. Gorog DA, Sweeny JM, Fuster V: Antiplatelet drug 'resistance'. Part 2: laboratory resistance to antiplatelet drugs-fact or artifact? *Nat Rev Cardiol* 2009, 6(5):365-373.
129. Sweeny JM, Gorog DA, Fuster V: Antiplatelet drug 'resistance'. Part 1: mechanisms and clinical measurements. *Nat Rev Cardiol* 2009, 6(4):273-282.
130. Watson SP, Harrison P: The vascular function of platelets. In *Postgraduate Haematology*. Edited by Hoffbrand AV, Catovsky D, Tuddenham EGD, Green AR: Wiley-Blackwell; 2011.
131. Fischer TH, Merricks EP, Bode AP, Bellinger DA, Russell K, Reddick R, Sanders WE, Nichols TC, Read MS: Thrombus formation with rehydrated, lyophilized platelets. *Hematology* 2002, 7(6):359-369.
132. Eigenthaler M, Nolte C, Halbrugge M, Walter U: Concentration and regulation of cyclic nucleotides, cyclic-nucleotide-dependent protein kinases and one of their major substrates in human platelets. Estimating the rate of cAMP-regulated and cGMP-regulated protein phosphorylation in intact cells. *Eur J Biochem* 1992, 205(2):471-481.
133. Prottly MB, Watkins NA, Colombo D, Thomas SG, Heath VL, Herbert JM, Bicknell R, Senis YA, Ashman LK, Berditchevski F *et al*: Identification of Tspan9 as a novel platelet tetraspanin and the collagen receptor GPVI as a component of tetraspanin microdomains. *Biochem J* 2009, 417(1):391-400.
134. Lopéz JA, Berndt MC: The GPIb-IX-V complex. In *Platelets*. Edited by Michelson AD: Elsevier; 2002.
135. Berndt MC, Gregory C, Kabral A, Zola H, Fournier D, Castaldi PA: Purification and preliminary characterization of the glycoprotein Ib complex in the human platelet membrane. *Eur J Biochem* 1985, 151(3):637-649.
136. Hartwig JH: The platelet: form and function. *Semin Hematol* 2006, 43(1 Suppl 1):S94-100.
137. van Schravendijk MR, Handunnetti SM, Barnwell JW, Howard RJ: Normal human erythrocytes express CD36, an adhesion molecule of monocytes, platelets, and endothelial cells. *Blood* 1992, 80(8):2105-2114.
138. Shattil SJ, Ginsberg MH, Brugge JS: Adhesive signaling in platelets. *Curr Opin Cell Biol* 1994, 6(5):695-704.
139. Best D, Senis YA, Jarvis GE, Eagleton HJ, Roberts DJ, Saito T, Jung SM, Moroi M, Harrison P, Green FR *et al*: GPVI levels in platelets: relationship to platelet function at high shear. *Blood* 2003, 102(8):2811-2818.

140. Cosemans JM, Iserbyt BF, Deckmyn H, Heemskerk JW: Multiple ways to switch platelet integrins on and off. *J Thromb Haemost* 2008, 6(8):1253-1261.
141. Diacovo TG, deFougerolles AR, Bainton DF, Springer TA: A functional integrin ligand on the surface of platelets: intercellular adhesion molecule-2. *J Clin Invest* 1994, 94(3):1243-1251.
142. Holada K, Mondoro TH, Muller J, Vostal JG: Increased expression of phosphatidylinositol-specific phospholipase C resistant prion proteins on the surface of activated platelets. *Br J Haematol* 1998, 103(1):276-282.
143. Kelton JG: The pathophysiology of heparin-induced thrombocytopenia: biological basis for treatment. *Chest* 2005, 127(2 Suppl):9S-20S.
144. Geberhiwot T, Ingerpuu S, Pedraza C, Neira M, Lehto U, Virtanen I, Kortessmaa J, Tryggvason K, Engvall E, Patarroyo M: Blood platelets contain and secrete laminin-8 (alpha4beta1gamma1) and adhere to laminin-8 via alpha6beta1 integrin. *Exp Cell Res* 1999, 253(2):723-732.
145. Woulfe D, Yang J, Provost N, O'Brien P, Brass L: Signal Transduction during Initiation, Extension, and Perpetuation of Platelet Plug Formation. In *Platelets*. Edited by A.D. M. San Diego; 2002:197-213.
146. Farrah T, Deutsch EW, Omenn GS, Campbell DS, Sun Z, Bletz JA, Mallick P, Katz JE, Malmstrom J, Ossola R *et al*: A high-confidence human plasma proteome reference set with estimated concentrations in PeptideAtlas. *Mol Cell Proteomics* 2011, 10(9):M110 006353.
147. Hogh B, Petersen E, Crandall I, Gottschau A, Sherman IW: Immune responses to band 3 neoantigens on Plasmodium falciparum-infected erythrocytes in subjects living in an area of intense malaria transmission are associated with low parasite density and high hematocrit value. *Infect Immun* 1994, 62(10):4362-4366.
148. Andersen JS, Wilkinson CJ, Mayor T, Mortensen P, Nigg EA, Mann M: Proteomic characterization of the human centrosome by protein correlation profiling. *Nature* 2003, 426(6966):570-574.
149. Dunkley TP, Watson R, Griffin JL, Dupree P, Lilley KS: Localization of organelle proteins by isotope tagging (LOPIT). *Mol Cell Proteomics* 2004, 3(11):1128-1134.
150. Zahedi RP, Meisinger C, Sickmann A: Two-dimensional benzyldimethyl-n-hexadecylammonium chloride/SDS-PAGE for membrane proteomics. *Proteomics* 2005, 5(14):3581-3588.
151. Schmitt S, Prokisch H, Schlunck T, Camp DG, 2nd, Ahting U, Waizenegger T, Scharfe C, Meitinger T, Imhof A, Neupert W *et al*: Proteome analysis of mitochondrial outer membrane from *Neurospora crassa*. *Proteomics* 2006, 6(1):72-80.
152. Sanjuan Szklarz LK, Kozjak-Pavlovic V, Vogtle FN, Chacinska A, Milenkovic D, Vogel S, Durr M, Westermann B, Guiard B, Martinou JC *et al*: Preprotein transport machineries of yeast mitochondrial outer membrane are not required for Bax-induced release of intermembrane space proteins. *J Mol Biol* 2007, 368(1):44-54.
153. Meyer H, Bug M, Bremer S: Emerging functions of the VCP/p97 AAA-ATPase in the ubiquitin system. *Nat Cell Biol* 2012, 14(2):117-123.
154. Sardiù ME, Cai Y, Jin J, Swanson SK, Conaway RC, Conaway JW, Florens L, Washburn MP: Probabilistic assembly of human protein interaction networks from label-free quantitative proteomics. *Proc Natl Acad Sci U S A* 2008, 105(5):1454-1459.
155. Desiere F, Deutsch EW, King NL, Nesvizhskii AI, Mallick P, Eng J, Chen S, Eddes J, Loevenich SN, Aebersold R: The PeptideAtlas project. *Nucleic Acids Res* 2006, 34(Database issue):D655-658.

156. Peng M, Taouatas N, Cappadona S, van Breukelen B, Mohammed S, Scholten A, Heck AJ: Protease bias in absolute protein quantitation. *Nat Methods* 2012, 9(6):524-525.
157. Ivanov AR, Colangelo C, Dufresne C, Friedman DB, Lilley KS, Mechtler K, Phinney B, Rose KL, Rudnick P, Searle BC *et al*: Development and Characterization of a Comprehensive Standard for Analysis of Post-translational Modifications. In *ASMS*. Edited by ABRF. Denver, Colorado, USA; 2011.
158. Pichler P, Mazanek M, Dusberger F, Weilnbock L, Huber CG, Stingl C, Luider TM, Straube WL, Kocher T, Mechtler K: SIMPATIQCO: A Server-Based Software Suite Which Facilitates Monitoring the Time Course of LC-MS Performance Metrics on Orbitrap Instruments. *J Proteome Res* 2012, 11(11):5540-5547.
159. Lundberg E, Gry M, Oksvold P, Kononen J, Andersson-Svahn H, Ponten F, Uhlen M, Asplund A: The correlation between cellular size and protein expression levels--normalization for global protein profiling. *J Proteomics* 2008, 71(4):448-460.
160. Martin H, Eckerskorn C, Gartner F, Rassow J, Lottspeich F, Pfanner N: The yeast mitochondrial intermembrane space: purification and analysis of two distinct fractions. *Anal Biochem* 1998, 265(1):123-128.
161. Mossmann D, Meisinger C, Vogtle FN: Processing of mitochondrial presequences. *Biochim Biophys Acta* 2011.
162. Mick DU, Fox TD, Rehling P: Inventory control: cytochrome c oxidase assembly regulates mitochondrial translation. *Nat Rev Mol Cell Biol* 2011, 12(1):14-20.
163. Varshavsky A: The N-end rule pathway and regulation by proteolysis. *Protein Sci* 2011.
164. Reinders J, Wagner K, Zahedi RP, Stojanovski D, Eyrich B, van der Laan M, Rehling P, Sickmann A, Pfanner N, Meisinger C: Profiling phosphoproteins of yeast mitochondria reveals a role of phosphorylation in assembly of the ATP synthase. *Mol Cell Proteomics* 2007, 6(11):1896-1906.
165. Schmidt O, Harbauer AB, Rao S, Eyrich B, Zahedi RP, Stojanovski D, Schonfisch B, Guiard B, Sickmann A, Pfanner N *et al*: Regulation of mitochondrial protein import by cytosolic kinases. *Cell* 2011, 144(2):227-239.
166. Peng J, Schwartz D, Elias JE, Thoreen CC, Cheng D, Marsischky G, Roelofs J, Finley D, Gygi SP: A proteomics approach to understanding protein ubiquitination. *Nat Biotechnol* 2003, 21(8):921-926.

9. Danksagung

Meine Promotion wurde von vielen Personen begleitet, die allesamt ein Stück zum Gelingen derselben beigetragen haben. An dieser Stelle möchte ich mich herzlich bedanken. Im Besonderen gilt mein Dank...

... Prof. Dr. Albert Sickmann, dem Institutsdirektor des ISAS, Doktorvater und Supervisor der Arbeit, für die interessanten, vielseitigen Fragestellungen, die ich während meiner Doktorarbeit selbständig und mit großem Vertrauen bearbeiten konnte.

... Prof Dr. Chris Meisinger, für die erfolgreiche Zusammenarbeit in mehreren interessanten Projekten, die motivierenden Telefonkonferenzen und auch für die Bereitschaft als Gutachter meiner Promotion zu fungieren.

... Prof. Dr. Kayser und Prof. Dr. Wichmann für ihre Bereitschaft als Prüfer der Promotion mitzuwirken.

... Dr. Rene Zahedi, der mit großem Interesse viele Projekte dieser Arbeit begleitete und Dank der ausgiebigen Diskussionen, kritischen Korrekturen und heiteren Gesellschaft stark zum Gelingen der Arbeit beigetragen hat.

... Marc Vaudel für die produktive Zusammenarbeit in unzähligen Projekten und für diverse Pétanque-, Käse- und Weinlehrgänge.

... Dr. Nora Vögtle für die Kooperation und den Austausch in vielen Mitochondrienprojekten.

... allen Mitarbeitern des Bereichs Systemanalyse in Dortmund wie auch in Würzburg für das lebendige Miteinander im Labor als auch in der Freizeit, insbesondere Sonja Radau und Clarissa Dickhut für ihre wertvollen Korrekturen.

... Dr. Stepan Gambaryan, Dr. Jörg Geiger und Prof. Dr. Ulrich Walter für die Isolierung von Thrombozyten und viele Diskussionen.

... Carola, Clarissa und Jan für die sportliche Ablenkung vom Laboralltag.

... Conny und Vroni, die mich schon seit dem Studium begleiten.

... meiner Familie, die mir das Studium der Pharmazie ermöglicht und mich die ganze Zeit über unterstützt hat.

10. Lebenslauf

Name: Julia Maria Burkhart

Adresse: Schillingstraße 42
44139 Dortmund

Geburtsdatum: 07.10.1981

Geburtsort: Schrobenhausen, Oberbayern

Eltern: Manfred Burkhart
Therese Burkhart, geb. Ottilinger

Schulbildung: 1988 – 1992 Grundschule, Langenmosen
1992 – 2001 Gymnasium Schrobenhausen

Studium: ab 01.04.2002 Studium der Pharmazie an der
Julius – Maximilians - Universität Würzburg

März 2004	Erste Pharmazeutische Prüfung Staatsexamen
Oktober 2006	Zweite Pharmazeutische Prüfung Staatsexamen
Mai 2008	Dritte Pharmazeutische Prüfung Staatsexamen / Approbation

**Auslands-
aufenthalt:** März – August 2005 Université de Limoges, Frankreich
Centre hospitalier universitaire (CHU)

11. Erklärung

Hiermit erkläre ich an Eides statt, dass ich die Dissertation „Globale qualitative und quantitative Proteomanalyse mittels hochauflösender LC-MS“ selbständig angefertigt und keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt habe.

Ich erkläre außerdem, dass diese Dissertation weder in gleicher oder anderer Form bereits in einem anderen Prüfungsverfahren vorgelegen hat.

Ich habe früher außer den mit dem Zulassungsgesuch urkundlich vorgelegten Graden keine weiteren akademischen Grade erworben oder zu erwerben versucht.

Dortmund, den

12. Anhang

12.1 Verwendete Trypsine

Tabelle 12.1: Übersicht über die in der Verdaustudie verwendeten Trypsine.

Kennung	Hersteller	Bestellnummer	Herstellerangaben zu Nebenreaktivitäten	Bemerkung
1	Promega	V5111	keine	modifiziertes Trypsin geeignet für die Proteinsequenzierung, TPCK behandelt
2	Sigma	T-8658	keine	geeignet für die Proteinsequenzierung
3	Sigma	T-1426	Chymotrypsin ≤ 0.1 Einheiten /mg Protein	TPCK behandelt
4	Sigma	T-0303	Chymotrypsin ≤ 1.0 Einheiten/mg Protein	
5	Fluka	nicht mehr kommerziell erhältlich		
6	Sigma	93614	Chymotrypsin $\leq 0.2\%$	zum Ablösen von adherierten Zellen (Trypsinierung von Zellfragmenten)

12.2 SRM Übergänge zur Analyse der Verdaueffizienz

Tabelle 12.2: SRM Übergänge der evaluierten Peptidpaare mit und ohne überlesener Schnittstelle. Analysiert wurde die Auswirkung der Verdaueffizienz sechs verschiedener Trypsine auf quantitative Studien.

UniProt Accession	Vorläuferion m/z	Fragmention m/z	Peptidsequenz
P02675	768.365	575.278	AHYGGFTVQNEANK
P02675	768.365	703.336	AHYGGFTVQNEANK
P02675	768.365	732.847	AHYGGFTVQNEANK
P02675	768.365	1164.564	AHYGGFTVQNEANK
P02675	768.365	1327.627	AHYGGFTVQNEANK
P02675	790.061	447.256	AHYGGFTVQNEANKYQISVNK
P02675	790.061	766.382	AHYGGFTVQNEANKYQISVNK
P02675	790.061	970.497	AHYGGFTVQNEANKYQISVNK
P02675	790.061	999.008	AHYGGFTVQNEANKYQISVNK
P02675	790.061	1080.539	AHYGGFTVQNEANKYQISVNK
P02679	717.692	631.32	ASTPNGYDNGIIWATWKTR
P02679	717.692	665.002	ASTPNGYDNGIIWATWKTR
P02679	717.692	840.928	ASTPNGYDNGIIWATWKTR
P02679	717.692	897.95	ASTPNGYDNGIIWATWKTR
P02679	717.692	946.476	ASTPNGYDNGIIWATWKTR
P02679	947.46	434.239	ASTPNGYDNGIIWATWK
P02679	947.46	505.276	ASTPNGYDNGIIWATWK
P02679	947.46	691.356	ASTPNGYDNGIIWATWK
P02679	947.46	804.44	ASTPNGYDNGIIWATWK
P02679	947.46	817.901	ASTPNGYDNGIIWATWK
P04179	408.182	248.624	DFGSFDK
P04179	408.182	409.208	DFGSFDK
P04179	408.182	496.24	DFGSFDK
P04179	408.182	553.261	DFGSFDK
P04179	408.182	700.33	DFGSFDK
P04179	545.764	422.276	DFGSFDKFK
P04179	545.764	488.25	DFGSFDKFK
P04179	545.764	684.371	DFGSFDKFK
P04179	545.764	771.403	DFGSFDKFK
P04179	545.764	828.425	DFGSFDKFK
P04406	681.041	715.885	LVINGNPITIFQERDPSK
P04406	681.041	801.418	LVINGNPITIFQERDPSK
P04406	681.041	858.439	LVINGNPITIFQERDPSK
P04406	681.041	859.426	LVINGNPITIFQERDPSK
P04406	681.041	914.981	LVINGNPITIFQERDPSK
P04406	807.454	579.288	LVINGNPITIFQER
P04406	807.454	793.42	LVINGNPITIFQER
P04406	807.454	1003.557	LVINGNPITIFQER
P04406	807.454	1174.621	LVINGNPITIFQER
P04406	807.454	1288.664	LVINGNPITIFQER
P21333	543.793	477.23	VAQPTITDNK
P21333	543.793	590.314	VAQPTITDNK
P21333	543.793	691.362	VAQPTITDNK
P21333	543.793	788.414	VAQPTITDNK
P21333	543.793	987.51	VAQPTITDNK
P21333	907.984	632.372	VAQPTITDNKDGTVTVR
P21333	907.984	758.902	VAQPTITDNKDGTVTVR
P21333	907.984	989.537	VAQPTITDNKDGTVTVR
P21333	907.984	1104.564	VAQPTITDNKDGTVTVR
P21333	907.984	1205.611	VAQPTITDNKDGTVTVR

UniProt Accession	Vorläuferion m/z	Fragmention m/z	Peptidsequenz
P21333	540.283	433.727	VNVGAGSHPNK
P21333	540.283	639.32	VNVGAGSHPNK
P21333	540.283	710.358	VNVGAGSHPNK
P21333	540.283	767.379	VNVGAGSHPNK
P21333	540.283	866.447	VNVGAGSHPNK
P21333	653.865	547.309	VNVGAGSHPNKVK
P21333	653.865	585.371	VNVGAGSHPNKVK
P21333	653.865	866.484	VNVGAGSHPNKVK
P21333	653.865	994.542	VNVGAGSHPNKVK
P21333	653.865	1093.611	VNVGAGSHPNKVK
P23528	760.346	941.403	HELQAN[C]Carboxyamidomethyl]YEEVK
P23528	760.346	1012.44	HELQAN[C]Carboxyamidomethyl]YEEVK
P23528	760.346	1140.498	HELQAN[C]Carboxyamidomethyl]YEEVK
P23528	760.346	1253.583	HELQAN[C]Carboxyamidomethyl]YEEVK
P23528	760.346	1382.625	HELQAN[C]Carboxyamidomethyl]YEEVK
P23528	895.91	418.24	HELQAN[C]Carboxyamidomethyl]YEEVKDR
P23528	895.91	1098.488	HELQAN[C]Carboxyamidomethyl]YEEVKDR
P23528	895.91	1212.531	HELQAN[C]Carboxyamidomethyl]YEEVKDR
P23528	895.91	1283.568	HELQAN[C]Carboxyamidomethyl]YEEVKDR
P23528	895.91	1411.626	HELQAN[C]Carboxyamidomethyl]YEEVKDR
P35579	508.94	420.722	TDLLLEPYNKYR
P35579	508.94	541.785	TDLLLEPYNKYR
P35579	508.94	598.327	TDLLLEPYNKYR
P35579	508.94	654.869	TDLLLEPYNKYR
P35579	508.94	840.436	TDLLLEPYNKYR
P35579	603.324	521.271	TDLLLEPYNK
P35579	603.324	650.314	TDLLLEPYNK
P35579	603.324	763.398	TDLLLEPYNK
P35579	603.324	876.482	TDLLLEPYNK
P35579	603.324	989.566	TDLLLEPYNK
P62937	528.274	480.245	VSFELFADK
P62937	528.274	593.329	VSFELFADK
P62937	528.274	722.371	VSFELFADK
P62937	528.274	869.44	VSFELFADK
P62937	528.274	956.472	VSFELFADK
P62937	690.382	597.332	VSFELFADKVPK
P62937	690.382	640.848	VSFELFADKVPK
P62937	690.382	657.393	VSFELFADKVPK
P62937	690.382	804.461	VSFELFADKVPK
P62937	690.382	917.545	VSFELFADKVPK
P62937	560.271	446.224	ITIADC[Carboxyamidomethyl]GQLE
P62937	560.271	453.205	ITIADC[Carboxyamidomethyl]GQLE
P62937	560.271	606.255	ITIADC[Carboxyamidomethyl]GQLE
P62937	560.271	721.282	ITIADC[Carboxyamidomethyl]GQLE
P62937	560.271	792.319	ITIADC[Carboxyamidomethyl]GQLE
P62937	624.319	446.224	KITIADC[Carboxyamidomethyl]GQLE
P62937	624.319	503.729	KITIADC[Carboxyamidomethyl]GQLE
P62937	624.319	560.271	KITIADC[Carboxyamidomethyl]GQLE
P62937	624.319	606.255	KITIADC[Carboxyamidomethyl]GQLE
P62937	624.319	792.319	KITIADC[Carboxyamidomethyl]GQLE
Q9Y490	523.796	446.272	KKEITGTLR
Q9Y490	523.796	547.319	KKEITGTLR
Q9Y490	523.796	660.403	KKEITGTLR
Q9Y490	523.796	789.446	KKEITGTLR
Q9Y490	523.796	918.489	KKEITGTLR
Q9Y490	587.843	523.795	KKEITGTLRK
Q9Y490	587.843	675.414	KKEITGTLRK
Q9Y490	587.843	788.498	KKEITGTLRK
Q9Y490	587.843	917.541	KKEITGTLRK
Q9Y490	587.843	1046.583	KKEITGTLRK
Q9Y490	805.069	767.374	IPEAPAGPPSDFGLFLSDDDPKK
Q9Y490	805.069	804.373	IPEAPAGPPSDFGLFLSDDDPKK
Q9Y490	805.069	917.457	IPEAPAGPPSDFGLFLSDDDPKK
Q9Y490	805.069	1001.991	IPEAPAGPPSDFGLFLSDDDPKK
Q9Y490	805.069	1064.525	IPEAPAGPPSDFGLFLSDDDPKK
Q9Y490	1143.052	676.278	IPEAPAGPPSDFGLFLSDDDPK
Q9Y490	1143.052	825.388	IPEAPAGPPSDFGLFLSDDDPK
Q9Y490	1143.052	937.944	IPEAPAGPPSDFGLFLSDDDPK
Q9Y490	1143.052	973.462	IPEAPAGPPSDFGLFLSDDDPK
Q9Y490	1143.052	1253.604	IPEAPAGPPSDFGLFLSDDDPK
Q9HBI1	673.346	611.324	DAFDLFDHAPDKLSVVK
Q9HBI1	673.346	678.369	DAFDLFDHAPDKLSVVK
Q9HBI1	673.346	785.435	DAFDLFDHAPDKLSVVK
Q9HBI1	673.346	842.949	DAFDLFDHAPDKLSVVK
Q9HBI1	673.346	916.483	DAFDLFDHAPDKLSVVK
Q9HBI1	746.341	567.288	DAFDLFDHAPDK
Q9HBI1	746.341	829.383	DAFDLFDHAPDK
Q9HBI1	746.341	942.467	DAFDLFDHAPDK
Q9HBI1	746.341	1043.515	DAFDLFDHAPDK
Q9HBI1	746.341	1158.542	DAFDLFDHAPDK

12.3 Modellproteom - UPS

Tabelle 12.3: Der kommerziell erhältliche Universal Proteome Standard (UPS) setzt sich aus 48 humanen Proteinen zusammen, die entweder natürlichen Ursprungs sind bzw. rekombinant hergestellt werden. Die Proteine liegen in äquimolarer Konzentration vor. UPS eignet sich aufgrund der bekannten Zusammensetzung für die Evaluierung von Methoden in der Proteomanalyse.

UniProt Accession	Proteinname	MW (Da)	Sequenz	Ursprung	Wirt	Tag zur Aufreinigung	mögliche PTM
P00709	Alpha-lactalbumin	14,078	20-142	Milch			Glycosylierung
P08758	Annexin A5	35,806	2-320	Placenta			Acetylierung Phosphorylierung
P01008	Antithrombin-III (ATIII)	49,039	22-464	Plasma			Phosphorylierung Glycosylierung
P61769	Beta-2-Microglobulin	11,731	21-119	Urin			
P55957	BH3-interacting domain death agonist	21,995	1-195		<i>E. coli</i>		
P00915	Carbonanhydrase 1	28,739	2-261	Erythrozyten			Acetylierung
P00918	Carbonanhydrase 2	29,115	2-260	Erythrozyten			Acetylierung
P04040	Katalase	59,625	2-527	Erythrozyten			Phosphorylierung
P01031	Complement C5/C5a anaphylatoxin	8,563	678-751		<i>E. coli</i>	Glutathione auf Cys705	
P02741	C-reactives Protein	23,047	19-224	Plasma			
P06732	Kreatinkinase M-Typ	43,101	1-381	Herz			
P00167	Cytochrom B5	16,022	2-134		<i>E. coli</i>	N-terminal 6-His	
P99999	Cytochrom C	11,618	2-105		<i>E. coli</i>		
P05413	Fatty acid-binding Protein	14,727	2-133	Herz			Acetylierung Phosphorylierung
O76070	Gamma-Synuclein		1-127		<i>E. coli</i>	N-terminal 6-His	
P06396	Gelsolin			Plasma			Phosphorylierung
P08263	Glutathion S-Transferase A1 (GTH1, HA subunit 1)	82,959	28-782		<i>E. coli</i>		
P09211	Glutathion S-Transferase P	25,500	2-222	Placenta			
P01112	GTPase HRas	21,298	1-189		<i>E. coli</i>		
P69905	Hämoglobin Subunit alpha	15,126	2-142	Erythrozyten			Glycosylierung Phosphorylierung
P68871	Hämoglobin Subunit beta	15,867	2-147	Erythrozyten			Acetylierung Glycosylierung Nitrosylierung Phosphorylierung
P12081	Histidyl-tRNA Synthetase, cytoplasmatisch	58,233	1-509		<i>E. coli</i>	C-terminal 6-His	
P01344	Insulin-like growth factor II (IGF-II)	7,475	25-91		<i>E. coli</i>		
P01579	Interferon gamma (IFN-gamma)	16,879	23-166		<i>E. coli</i>		
P10145	Interleukin-8 (IL-8)	8,386	28-99		<i>E. coli</i>		
P02788	Lactotransferrin	76,165	20-710	Milch			Glycosylierung Phosphorylierung
P41159	Leptin	16,158	22-167	Rekombinant	<i>E. coli</i>		
P61626	Lysozyme C	14,701	19-148	Milch			
P10636-8	Microtubule-associated Protein tau	45,719	2-441	Rekombinant	<i>E. coli</i>		
P02144	Myoglobin	17,053	2-154	Heart			
P15559	NAD(P)H dehydrogenase [quinone] 1	30,736	2-274	Rekombinant	<i>E. coli</i>		
Q15843	NEDD8	9,072	1-81	Rekombinant	<i>E. coli</i>		
P62937	Peptidyl-prolyl cis-trans Isomerase A (PPIase A, Rotamase A)	20,176	1-165	Rekombinant	<i>E. coli</i>	N-terminal 6-His	
Q06830	Peroxiredoxin 1	21,979	2-199	Rekombinant	<i>E. coli</i>		
P01127	Platelet-derived growth factor Subunit B (PDGF Subunit B)	12,294	82-190	Rekombinant	<i>E. coli</i>		
P01133	Pro-epidermal growth factor (EGF)	6,353	971-1023	Rekombinant	<i>E. coli</i>		
P02753	Retinol-binding Protein 4	21071	19-201	Urine			
P16083	Ribosyl-dihydropyridinyl Dehydrogenase [quinone]	25,821	2-231	Rekombinant	<i>E. coli</i>		
P02787	Serotransferrin (Transferrin)	75,181	20-698	Plasma			Glycosylierung Phosphorylierung
P02768	Serumalbumin	66,357	26-609	Rekombinant	<i>Pichia pastoris</i>		
P63165	Small ubiquitin-related modifier 1 (SUMO-1)	38,815	1-158	Rekombinant	<i>E. coli</i>	N-terminal GST	
P63279	SUMO-conjugating Enzym UBC9	18,007	1-158	Rekombinant	<i>E. coli</i>		
P00441	Superoxid-Dismutase [Cu-Zn]	15,805	2-154	Erythrozyten			Acetylierung Phosphorylierung
P10599	Thioredoxin (Trx)	12,429	2-105	Rekombinant	<i>E. coli</i>	N-terminal 6-His	
P01375	Tumornekrosefaktor, löslich	17,353	77-233	Rekombinant	<i>E. coli</i>		
P62988	Ubiquitin	10,597		Rekombinant	<i>E. coli</i>	N-terminal 6-His	
O00762	Ubiquitin-conjugating Enzym E2 C	20,475	1-179	Rekombinant	<i>E. coli</i>	N-terminal 6-His	
P51965	Ubiquitin-conjugating Enzym E2 E1	22,227	1-193	Rekombinant	<i>E. coli</i>	N-terminal 6-His	

12.4 Qualitätskontrolle von nano-LC-MS Systemen

Tabelle 12.4: Ergänzte Komponenten zur Analyse der LF Quantifizierung und der LOQ über einen dynamischen Bereich von 1-1000 fmol.

Synthetische Peptide (<i>S. cerevisiae</i>)				
Peptidsequenz	Modifikation	Protein	Ladung	m/z
ALELKPDYSK	Phosphorylierung	TOM 70	2	622.3021
ALELKPDYSK	-	TOM 70	2	582.319
GFQESLSDDEK	Phosphorylierung	MDM 10	2	667.761
GFQESLSDDEK	-	MDM 10	2	627.7779
EQAEASIDNLK	Phosphorylierung	ZEO 1	2	649.2872
EQAEASIDNLK	-	ZEO 1	2	609.304
ISS LLP FN*K	13C (6) 15N (2)	TIM 54	2	513.8076
NSE WTQ ELV CDP*R	13C (6) 15N (4)	TIM 54	2	822.3701
EKN SEW TQE LVC DP* R	13C (6) 15N (4)	TIM 54	2	950.9411
VYT GFG LEQ ISP AQ*K	13C (6) 15N (2)	TIM 22	2	823.4386
NNG FVD LV*K	13C (6) 15N (2)	TIM 18	2	507.2786
LWN END DHL YIF G*R	13C (6) 15N (4)	TIM 18	2	901.4326
FQE QNA ALG QGL G*R	13C (6) 15N (4)	TIM 9	2	749.8831
LYSNLV E*R	13C (6) 15N (4)	TIM 9	2	502.2766
LIG GFV QT*R	13C (6) 15N (4)	TIM 12	2	500.7936
P02662, α - Casein, bovine, MW 24.5 kDa; tryptischer Verdau				
MKLLILTLCLVAALARPKHPKHKHQLPQEVLNENLLRFVAPFPEVFGKEKVNELSKDIG SESTEDQAMEDIKQMEAESISSSEEIVPNSVEQKHQKEDVPSERYLGYLEQLLRKYYK VPQLEIVPNSAEERLHSMKEGIHAQQKEPMIGVNQELAYFYPELFRQFYQLDAYPSGAWY YVPLGTQYTDAPSFSDIPNPIGSENSEKTTMPLW				
Peptidsequenz	Modifikation	Protein	Ladung	m/z
HQGLPQEVLNENLLR	-	α - Casein	2	880.4759
YLGYLEQLLR	-	α - Casein	2	634.3557
YKVPQLEIVPNSAEER	-	α - Casein	2	976.48
TVDMESTEVFTK	-	α - Casein	2	733.8099
ALNEINQFYQK	-	α - Casein	2	684.351

12.5 Analyse des mitochondrialen IMS Proteoms (*S. cerevisiae*)

Tabelle 12.5: IMS Probenkombinationen der vier Replikate mit variierenden SILAC Markierungen. Gezeigt sind weiterhin die in der Auswertung dargestellten SILAC Verhältnisse.

Replikat	SILAC Kombination der Proben				Dargestellte SILAC Verhältnisse
	Probe	SILAC Label	Probe	SILAC Label	
1	SN _{BAX}	Lys ¹³ C ₆ und Arg ¹³ C ₆	vs. Total	kein SIL Label	Heavy (H) / Light (L)
1	SN _{BAX}	Lys ¹³ C ₆ und Arg ¹³ C ₆	vs. SN _{CON}	kein SIL Label	H / L
2	SN _{BAX}	kein SIL Label	vs. Total	Lys ¹³ C ₆ und Arg ¹³ C ₆	L / H
2	SN _{BAX}	kein SIL Label	vs. SN _{CON}	Lys ¹³ C ₆ und Arg ¹³ C ₆	L / H
3	SN _{BAX}	kein SIL Label	vs. Total	Lys ¹³ C ₆ ¹⁵ N ₂ und Arg ¹³ C ₆ ¹⁵ N ₄	L / H
3	SN _{BAX}	kein SIL Label	vs. SN _{CON}	Lys ¹³ C ₆ ¹⁵ N ₂ und Arg ¹³ C ₆ ¹⁵ N ₄	L / H
4	SN _{BAX}	Lys ¹³ C ₆ ¹⁵ N ₂ und Arg ¹³ C ₆ ¹⁵ N ₄	vs. Total	kein SIL Label	H / L
4	SN _{BAX}	Lys ¹³ C ₆ ¹⁵ N ₂ und Arg ¹³ C ₆ ¹⁵ N ₄	vs. SN _{CON}	kein SIL Label	H / L

Tabelle 12.7: SN_{BAX}/SN_{Con} und SN_{BAX}/Tot Verhältnisse von 29 bekannten Intermembranraumproteinen, die reproduzierbar im Überstand von Mitochondrien nach BAX Behandlung identifiziert werden konnten. Die Zahlen 1-4 beziehen sich auf die verschiedenen Replikate.

ORF	Genname	Sequenz- abdeckung [%]	1. SN_{BAX}/Tot		2. SN_{BAX}/Tot		2. SN_{BAX}/SN_{Con}		3. SN_{BAX}/Tot		3. SN_{BAX}/SN_{Con}		4. SN_{BAX}/Tot		4. SN_{BAX}/SN_{Con}											
			Ratio	H/L counts	Ratio	H/L counts	Ratio	H/L counts	Ratio	H/L counts	Ratio	H/L counts	Ratio	H/L counts	Ratio	H/L counts	Ratio	H/L counts								
YBR212C	POB1	43.44	0.1	27	1072	12	15	41.9	0.1	43	56.8	4.7	11	27.6	0.2	23	91.0	2.7	18	12.5	0.0	46	30.4	1.1	21.0	12.0
YGD17C	NF51	49.7	0.2	3	3.2	1.9	6	11.9	0.1	16	34.9	3.5	10	18.3	0.2	10	23.0	2.2	13	21.8	0.0	38	41.4	2.2	10.0	15.6
YCR026CA	RIM1	69.63	0.1	7	2.81	1.7	6	6.7	0.1	18	96.2	2.1	2	19.4	0.2	13	367.6	1.9	11	15.0	0.0	38	36.9	1.1	10.0	9.0
YDL178W	DLD2	60.38	0.1	34	90.5	3.1	11	14.2	0.0	42	97.9	5.1	16	39.7	0.1	24	83.2	3.8	16	25.6	0.0	65	49.8	2.2	15.0	4.8
YDR148C	KGD2	39.96	0.1	19	83.7	1.6	15	31.8	0.1	42	62.6	3.5	11	65.7	0.3	17	426.6	2.2	16	40.8	0.1	45	37.1	1.3	23.0	20.3
YDR175C	RSM24	38.24	0.0	7	41.59	0.0	14	127.0	2.21	1	0.0	0.1	7	183.0	1.1	2	38.4	0.0	18	51.4	1.8	3.0	14.9
YDR430C	CYH1	48.33	0.0	30	42.0	1.8	3	36.2	0.0	72	67.1	5.0	5	35.9	0.0	23	98.8	3.7	6	22.6	0.0	84	53.2	2.0	12.0	16.6
YER048W-A	ISD11	71.28	0.1	2	10.45	2.2	1	0.0	0.0	15	76.4	2.8	2	6.8	0.1	6	15.6	2.1	6	13.5	0.0	19	35.1	1.3	6.0	19.5
YER099W	ARG5.6	51.1	0.1	32	64.1	1.6	26	94.9	0.1	70	86.6	4.2	30	94.9	0.2	23	174.6	2.7	51	22.1	0.1	102	50.1	1.5	49.0	166.4
YER078C	ICP55	31.12	0.1	13	38.2	1.5	2	94.6	0.1	22	74.0	4.4	2	62.4	0.2	2	52.8	4.7	1	0.0	0.0	15	48.7	1.8	5.0	22.7
YFL018C	LPF1	58.32	0.1	28	52.5	1.4	26	33.2	0.1	47	60.5	2.3	16	21.4	0.1	24	120.4	2.0	40	14.0	0.1	65	43.2	0.9	31.0	11.5
YFR049W	YFR31	74.8	0.0	6	27.6	1.0	2	14.6	0.1	15	20.3	2.8	4	2.0	0.1	6	33.2	1.9	4	8.1	0.0	26	47.9	0.8	7.0	5.8
YGL018C	JAC1	34.78	0.1	5	14.0	2.1	2	35.7	0.1	10	29.2	3.1	3	39.8	0.1	4	23.6	2.1	6	11.3	0.0	11	41.7	2.1	4.0	15.2
YGL193C	PDX1	59.51	0.1	16	55.7	1.0	5	39.8	0.1	27	39.8	6.2	5	57.2	0.1	7	23.9	2.4	10	26.5	0.0	45	44.1	1.3	9.0	29.6
YGR234W	YH81	74.44	0.1	34	51.6	2.3	23	75.1	0.1	75	60.6	8.7	26	59.6	0.1	12	16.5	6.1	26	24.3	0.1	97	38.5	2.2	15.0	5.0
YHR008C	SOJ2	56.65	0.1	13	21.60	2.6	6	4.5	0.0	16	71.5	3.7	4	32.1	0.1	11	17.3	3.9	7	11.2	0.0	32	37.1	1.9	6.0	6.8
YHR024C	MAS2	45.44	0.0	11	37.8	6.4	1	0.0	0.0	36	67.7	0.0	13	495.3	4.2	1	0.0	0.0	29	63.9	4.3	3.0	33.8
YHR037W	FUT2	42.09	0.0	13	74.6	1.9	1	0.0	0.0	25	63.3	4.3	5	70.2	0.1	2	12.9	4.0	2	10.5	0.0	53	74.2	1.2	11.0	15.4
YHR208W	BAT1	68.7	0.1	52	74.8	1.7	37	57.9	0.1	80	70.9	11.0	23	77.6	0.1	50	127.0	4.2	29	19.3	0.1	94	34.4	3.1	32.0	20.2
YIL051C	MFM1	58.62	0.1	8	11.95	2.4	4	33.1	0.1	14	43.0	4.6	5	56.9	0.1	9	544.4	2.9	5	2.7	0.0	26	31.8	2.0	8.0	10.7
YIL070C	MAM33	34.59	0.1	11	23.7	2.1	2	5.7	0.0	19	84.4	2.9	7	15.7	0.1	7	71.3	2.9	4	7.6	0.0	36	43.6	1.2	4.0	7.1
YIL125W	KGD1	57.69	0.1	70	135.8	1.6	58	48.3	0.1	167	80.5	4.5	49	47.1	0.1	72	81.5	2.4	39	20.5	0.0	207	30.3	1.9	34.0	13.8
YIR049C	SSC1	60.55	0.0	68	26.2	2.1	30	28.7	0.0	140	241.4	6.8	17	49.9	0.1	57	808.9	2.6	41	14.7	0.0	131	51.3	2.4	46.0	28.3
YIR144W	MGM101	53.53	0.1	15	46.7	1.1	10	51.9	0.1	25	45.6	2.3	8	43.5	0.1	11	128.5	1.7	13	7.2	0.0	28	26.3	1.0	7.0	14.0
YKD40C	NFU1	75	0.1	16	27.1	4.0	13	24.0	0.0	38	57.7	12.1	8	35.3	0.2	21	708.2	5.9	11	111.1	0.0	56	35.0	3.5	10.0	6.6
YKL085W	MDH1	78.44	0.1	86	65.7	1.6	31	60.1	0.1	126	74.9	5.1	27	22.5	0.1	99	201.8	3.4	70	25.3	0.0	142	31.3	1.5	67.0	19.2
YKL134C	OCT1	24.68	0.3	1	0.0	0.1	14	53.6	7.8	1	0.0	0.7	3	964.2	3.1	1	0.0	0.0	16	57.1	1.4	2.0	50.7
YUR069C	MEF1	47.44	0.1	21	58.6	5.1	11	35.8	0.1	44	68.1	15.0	14	67.8	0.1	12	57.0	8.2	17	26.9	0.0	76	30.7	6.1	17.0	28.3
YUR163C	MAS1	41.56	0.0	6	36.19	0.0	23	99.0	3.3	1	0.0	0.1	11	1082.3	2.7	2	54.0	0.0	30	48.8	1.6	2.0	64.6
YUR529C	HSF60	60.84	0.0	83	93.9	1.5	36	55.2	0.0	132	54.8	3.9	18	35.2	0.1	93	91.3	3.0	67	19.6	0.0	135	39.9	1.2	55.0	15.0
YUR394C	ACO1	69.67	0.1	92	71.9	2.4	62	37.8	0.1	239	69.5	4.5	43	47.1	0.2	119	373.8	3.3	60	16.3	0.0	234	33.3	1.6	83.0	17.7
YUS313W-A	MRP15	35.18	0.2	1	0.0	1.9	1	0.0	0.1	8	34.1	5.2	3	36.2	0.2	3	41.7	3.2	2	2.6	0.1	3	8.5	2.2	3.0	22.6
YUR355C	ILU5	72.15	0.0	85	88.3	2.2	41	51.4	0.0	169	77.8	3.8	28	24.5	0.2	108	749.3	2.7	58	16.1	0.0	170	37.4	2.1	52.0	15.6
YML078W	CFR3	40.33	0.1	12	23.5	3.8	11	17.8	0.2	25	30.7	11.6	4	12.8	0.1	15	73.5	6.2	11	11.6	0.1	45	40.7	3.4	12.0	14.5
YMR024W	MRFL3	45.38	0.2	5	37.9	1.3	1	0.0	0.1	14	55.5	7.64	2	5.6	0.3	5	57.2	3.0	2	23.0	0.0	18	27.0	2.3	5.0	4.5
YMR072W	ABP2	49.18	0.1	8	35.09	1.0	2	59.5	0.0	14	28.6	1.7	2	0.3	0.2	3	34.4	1.5	5	19.0	0.0	31	34.8	0.9	4.0	5.1
YMR083W	ADP3	50.93	0.1	26	77.0	1.6	21	7.9	0.1	74	71.0	3.0	14	65.1	0.1	29	30.6	2.2	33	12.5	0.1	78	24.1	1.3	34.0	11.4
YMR158W	MRPS8	39.35	0.0	2	74.7	0.1	6	185.1	4.9	1	0.0	0.1	4	53.4	2.1	3	19.7	0.0	7	24.5	1.1	1.0	0.0
YMR189W	GCV2	47.1	0.1	15	69.9	0.9	1	0.0	0.1	27	69.5	2.9	9	77.7	0.1	6	23.4	3.9	5	4.4	0.0	76	43.2	1.2	16.0	21.1
YND071C	IDH1	55.56	0.1	27	111.8	1.0	15	32.1	0.0	62	66.5	3.5	12	3.3	0.1	8	60.8	2.4	10	18.3	0.0	71	40.4	1.2	16.0	7.2
YNL071W	LAT1	34.65	0.1	8	57.1	1.0	8	33.3	0.1	23	54.2	3.3	6	58.9	0.2	18	36.9	2.2	19	12.9	0.1	34	35.7	1.1	21.0	16.8
YNR0031C	CIT1	59.29	0.1	58	63.4	1.1	45	50.8	0.1	185	89.9	3.1	52	72.2	0.1	84	118.4	2.0	53	13.9	0.1	180	27.8	1.2	63.0	12.8
YNR022C	MRPL50	38.95	0.2	2	26.1	0.1	9	18.7	2.8	2	6.8	0.0	0.0	6	76.4
YNR037C	RSM19	64.84	0.0	1	0.0	0.1	7	80.0	3.5	1	0.0	0.3	6	93.8	1.9	4	14.4	0.0	11	34.9	0.6	4.0	13.6
YOR020C	HSF10	67.92	0.1	19	22.5	1.2	12	20.1	0.1	36	62.6	3.9	5	12.1	0.1	35	94.8	1.1	20	12.1	0.0	39	32.0	0.8	21.0	12.0
YOR040W	GLO4	40.35	0.1	4	26.9	1.6	3	13.5	0.1	11	121.4	4.3	2	15.3	0.1	6	55.3	3.2	5	15.5	0.1	14	44.3	1.5	4.0	40.4
YOR138W	IDN2	73.17	0.1	34	79.0	1.4	13	38.5	0.0	65	78.1	3.1	11	50.2	0.1	7	25.7	2.4	16	13.5	0.0	76	35.4	1.2	19.0	15.5
YOR142W	LSC1	84.8	0.1	54	137.3	3.6	32	39.2	0.1	108	69.8	13.5	20	37.9	0.1	42	195.6	6.8	41	36.5	0.0	121	37.8	3.7	35.0	21.6
YOR188W	PET123	38.99	0.1	3	16.43	0.0	7	376.0	1.4	1	0.0	0.1	6	107.5	1.2	2	17.3	0.0	14	158.7	0.7	2.0	14.8
YOR187W	TUF1	68.65	0.1	29	60.3	3.4	22	66.9	0.1	59	55.3	16.6	30	84.5	0.1	35	51.6	7.9	34	23.7	0.0	99	31.3	5.4	30.0	29.4
YOR321W	MGE1	27.63	0.0	7	22.1	2.9	4	2.6	0.0	17	50.1	0.1	10	330.4	1.9	5	15.4	0.0	22	24.4	1.2	6.0	5.7
YPL059W	GRU5	64	0.0	11	42.3	2.4	4	13.2	0.1	18	88.1	7.0	4	7.4	0.1	13	135.7	3.1	9	10.7	0.0	30	39.5	1.3	7.0	3.9
YPL118W	MRP51	37.5	0.1	2	25.3	0.0	10	275.0	3.6	1	0.0	0.0	6	509.9	1.4	1	0.0						

Tabelle 12.8: SN_{BAX}/SN_{Con} und SN_{BAX}/Tot Verhältnisse der 59 Matrixproteine, die reproduzierbar im Überstand von BAX behandelten Mitochondrien innerhalb einer SILAC Studie identifiziert wurden. Die Zahlen 1-4 beziehen sich auf die verschiedenen Replikate.

ORF	GeneName	Genenotation	Sequenz-abdeckung [%]	1. SN _{BAX} /Tot Ratio [H/L counts] (Variabilität)	1. SN _{BAX} /SN _{Con} Ratio [H/L counts] (Variabilität)	2. SN _{BAX} /Tot Ratio [H/L counts] (Variabilität)	2. SN _{BAX} /SN _{Con} Ratio [H/L counts] (Variabilität)	3. SN _{BAX} /Tot Ratio [H/L counts] (Variabilität)	3. SN _{BAX} /SN _{Con} Ratio [H/L counts] (Variabilität)	4. SN _{BAX} /Tot Ratio [H/L counts] (Variabilität)	4. SN _{BAX} /SN _{Con} Ratio [H/L counts] (Variabilität)
Y0S22W	ADK1	Adenylylase kinase	4189	0.6 3 20.1 6.3 7	67.5 2	1.2 6 61.2 15.6 2	76.7 2	1.2 6 35.2 15.6 8	39.6 8	0.3 17 28.3 4.1 10	25.4
Y1L9C	ATP12	Assembly factor of ATP synthase	48	0.8 4 68.3 7.5 2	285.4 6	0.3 8 10.3 13.0 6	38.3 0.4	0.4 5 16.0 5.8 14	26.4 0.1	0.1 21 23.8 6.0 11	17.9
YK064C	COF1	Mitochondrial cytochrome-c oxidase	5734	1.1 19 42.8 32.8 27	54.7 13	1.6 53 41.0 106.2 13	43.6 1.8	30 25.7 51.0 39	73.9 0.4	62 32.4 23.3 2.8	24.2
YK038C	CCS1	Copper chaperone for superoxide dismutase SOD1	2691	1.7 1 0.0 15.9 3	31.9 5	1.7 4 40.2 110.1 5	86.5 1.8	3 7.1 37.4 4	70.4 0.5	5 15.0 7.4 2	58.0
YK137W	CIM1	Mitochondrial protein necessary for full assembly of Cytochrome c oxidase	3423	0.8 1 0.0 6.4 4	49.6 4	0.7 6 19.5 9.1 3	18.0 0.9	5 9.5 9.3 8	28.5 0.3	12 17.3 3.6 10	18.8
YK039C-A	CIM2	Cy9C Motif-Containing protein	4495	0.8 2 28.3 6.3 6	11.8 10	1.0 7 60.0 14.1 4	110.2 0.9	2 4.3 5.6 11	39.9 0.4	13 15.0 3.3 5	9.1
YK141C-B	CIM3	Cy9C Motif-Containing protein	5753	0.7 3 14.2 24.4 5	24.3 8	5 13.7 93.3 4	66.2 1.0	11 19.3 38.0 11	28.7 0.3	10 7.4 11.8 7	54.9
YK121C	COM1/CIM3	Cytochrome Oxidase Assembly factor	2867	1.0 1 0.0 9.1 3	24.0 8	0.9 8 25.6 107.6 3	10.4 -	- 48.7 3 2.6	0.4 6	16.6 10.1 2	5.3
YK099C	COI17	Copper metallochaperone	4058	1.0 2 1.3 26.3 1	0.0 0	0 2.0 0.0 0	0.0 0.9	2 5.5 30.5 2	0.2 0.4	5 14.9 16.8 3	9.1
YK018C-A	COX19	metal ionophore required for cytochrome c oxidase assembly	6122	0.7 1 0.0 39.3 4	67.5 4	0.8 5 6.9 25.7 1	0.0 1.0	1 0.0 35.3 7	26.0 0.4	7 10.4 16.4 6	18.2
YK116W	COX23	Protein required for cytochrome oxidase assembly	35.1	- - 19.6 2	65.7 2	1.0 5 2.8 18.3 2	147.8 1.2	3 1.0 21.3 6	101.5 0.3	5 9.0 18.2 4	3.5
YK054C	CV82	L-lactate cytochrome-c oxidoreductase	6142	0.9 37 41.8 39.7 51	51.7 49	1.2 142 33.0 84.5 49	73.9 1.0	58 12.8 41.6 72	63.4 0.3	112 19.4 16.8 5.6	60.7
YK048W	CV1	Cytochrome c isoform 1	6789	1.1 6 17.2 8.0 7	22.3 3	1.0 17 45.3 14.5 3	13.4 1.5	5 15.0 10.7 8	24.2 0.6	15 5.1 8.2 10	20.4
YK039C	CV3	Cytochrome c heme lyase	3271	0.4 1 0.0 1.5 1	0.0 4	0.5 4 33.7 13.2 2	697.1 0.8	3 13.7 1.4 2	3.5 0.1	8 17.8 0.9 3	4.8
YK039C	CV7	Cytochrome c isoform 2	6106	0.9 4 35.8 4.8 3	105.3 3	1.3 6 2.1 69.3 1	0.0 1.9	2 13.0 13.7 4	23.7 0.7	9 7.5 8.7 5	9.0
YK023W	CV1	Sulfhydryl oxidase involved in hMA protein import pathway	5979	0.5 4 33.5 7.1 5	33.4 5	0.7 7 8.9 12.9 5	24.8 0.6	3 6.3 9.4 11	26.7 0.2	13 11.4 3.2 6	8.6
Y1L35C	GU72	Mitochondrial (E)-over-3-phosphate dehydrogenase	5871	0.6 35 19.1 3.0 51	17.7 40	0.4 89 35.3 4.0 53	22.0 0.8	22 20.3 3.4 38	22.4 0.2	94 12.2 2.4 51	21.2
YK115W	ME1	Mitochondrial NADH-Cytochrome b5 Reductase**	7384	0.5 31 59.8 2.5 27	45.3 21	1.0 57 53.5 5.8 21	54.3 1.0	22 15.1 5.3 45	24.9 0.4	82 20.7 2.0 41	17.0
YK135C-A	MDM35	Import factor for Ucp proteins	7442	0.7 5 29.8 29.3 4	17.4 4	0.7 10 10.7 37.8 3	13.5 1.0	3 3.9 15.5 8	91.6 0.4	9 18.2 8.1 10	133.0
YK021C	MEM1	Mitochondrial GTPase, required for mt morphology and genome maintenance**	1669	17.0 2 35.2 9.7 6	28.7 9	13 177.7 70.2 9	151.8 8.0	4 44.1 8.4 9	39.3 1.6	6 18.1 8.4 6	24.8
YK031W	MC14	Intermembrane space Cysteine motif protein	3802	0.8 2 7.7 36.7 6	29.8 6	0.8 7 32.9 87.7 2	15.8 1.0	7 13.0 25.0 11	55.8 0.3	10 13.1 19.1 8	55.0
YK002W	MC17	Intermembrane space Cysteine motif protein	3269	0.9 4 4.3 13.7 6	17.5 6	0.6 8 4.8 120.1 2	1391.7 0.9	4 4.6 10.4 6	3.8 0.4	8 5.6 7.3 4	10.0
YK104C	SOD1	copper-zinc superoxide dismutase	6623	0.6 7 1.6 18.5 8	27.8 8	1.4 8 27.3 66.6 4	48.0 1.1	5 8.6 28.3 14	196.4 0.3	20 15.3 11.2 15	20.4
YK005C-A	TM10	Protein required for import of hydrophobic proteins	8387	0.8 8 10.0 23.2 13	37.1 9	0.5 19 10.0 39.1 9	69.4 0.7	11 33.8 24.9 11	24.7 0.2	19 19.0 7.9 11	48.9
YK091C	TM12	Protein required for import of hydrophobic proteins	1651	0.2 1 0.0 2.4 1	0.0 1	0.3 1 0.0 -	-	1 0.0 3.2 4	23.8 0.1	6 19.5 2.2 4	22.6
YK181W	TM13	Protein required for import of hydrophobic proteins	819	0.9 11 20.4 14.7 8	24.8 8	0.5 17 27.9 17.5 1	0.0 0.4	7 29.6 18.0 21	37.5 0.2	26 32.1 6.7 15	17.2
YK135W-A	TM8	Protein required for import of hydrophobic proteins	7586	0.7 5 10.1 14.7 5	6.5 6	0.6 14 13.7 27.4 5	35.5 0.7	10 38.0 15.6 11	18.2 0.2	23 26.0 8.7 10	12.2
YK020W-A	TM9	Protein required for import of hydrophobic proteins	4828	0.7 6 4.1 17.1 5	10.3 5	0.5 17 10.5 23.5 4	6.6 0.6	18 29.2 16.1 14	25.2 0.2	22 10.0 8.1 16	15.3
YK027W	YK1	Nucleoside diphosphate kinase	3072	- - - -	- -	1.7 1 0.0 8.9 1	0.0 1.1	1 0.0 4.5 4	21.7 0.3	2 0.6 4.4 3	69.4

12.6 Charakterisierung des thrombozytären Proteoms

12.6.1 Abgleich von 68 CD Markern in RNA-, SAGE- und Proteomdaten

CD Marker (Cluster of differentiation) sind Oberflächenmerkmale von Zellen, die z.T. zellspezifisch exprimiert werden und in vielen Fällen strukturell den membrangebundenen Glykoproteinen zugeordnet werden können. Ihre Aufgaben erstrecken sich über Rezeptor- oder Signalfunktionen, unterdessen wurde auch enzymatische Aktivität dokumentiert. Eingeführt während des Human Leucocyte Differentiation Antigen (HLDA) Workshops im Jahr 1982 wurden CD Marker ursprünglich zur Charakterisierung von Leukozyten herangezogen mit dem Ziel einer standardisierten Analyse unter Verwendung von monoklonalen Antikörpern.

Tabelle 12.8: Die Tabelle führt die 68 in dieser Studie mittels RNA-, SAGE- und Proteomdaten abgeglichenen CD Marker auf, wovon 28 (grün) in Thrombozyten bekannt sind. In den jeweiligen Datensätzen identifizierten CD Marker sind mit (+) gekennzeichnet, während CD Marker mit (°) markiert sind.

CD Marker	Zelle	Thrombozyt	RNA	SAGE	Kopienzahlen aus der LC-MS Analyse
CD19	B cell	-	+	-	-
CD20	B cell	-	+	+	-
CD146	endothelial cell	-	-	-	-
CD35	erythrocyte	-	-	-	-
CD44	erythrocyte	-	+	-	-
CD66b	granulocyte	-	-	-	-
CD49a	monocyte	-	-	-	-
CD117	stem cell	-	-	-	-
CD243	stem cell	-	-	-	-
CD33	stem cell	-	-	-	-
CD34	stem cell	-	-	-	-
CD90	stem cell	-	-	-	-
CD133	stem cell	-	-	-	-
CD135	stem cell	-	-	-	-
CD4	T cell	-	-	-	-
CD8a	T cell	-	+	-	-
CD8b	T cell	-	-	-	-
CD81	T cell	-	+	+	+
CD97	granulocyte	-	+	+	1000
CD37	B cell	+	+	-	580
CD93	neutrophils	+	+	-	750
CD92	platelet	+	+	+	3.300
CD41	platelet	+	+	+	83.300
CD61	platelet	+	+	+	64.200
CD62P	platelet	+	+	-	8.900
CD9	platelet	+	+	+	8.000
CD23	platelet	+	-	-	-
CD31	platelet	+	+	-	9.400
CD36	platelet	+	+	+	16.700
CD42a	platelet	+	+	+	32.400
CD42b	platelet	+	-	-	18.900
CD42c	platelet	+	+	+	49.000
CD42d	platelet	+	+	-	30.200
CD49b	platelet	+	+	-	4.600
CD49f	platelet	+	+	-	11.500
CD51	platelet	+	-	-	1.400
CD84	platelet	+	+	-	2.900
CD109	platelet	+	-	-	1.700
CD110	platelet	+	+	+	1.600
CD147	platelet	+	+	-	2.900
CD151	platelet	+	+	+	3.000
CD226	platelet	+	+	-	2.800
CD107a	platelet (activated)	+	+	-	3.000
CD107b	platelet (activated)	+	+	+	2.000
CD63	platelet (activated)	+	-	+	2.200
CD111	stem cell	+	-	-	570
CD112	stem cell	+	+	-	750
CD11c	dendritic cell	0	+	-	-
CD123	dendritic cell	0	-	-	-
CD62E	endothelial cell	0	-	+	-
CD236	epithelial cell	0	+	-	-
CD235a	erythrocyte	0	-	-	-
CD238	erythrocyte	0	-	-	-
CD247	lymphocyte	0	+	-	-
CD62L	lymphocyte	0	+	-	-
CD131	megakaryocyte	0	+	-	-
CD203c	megakaryocyte	0	-	-	-
CD14	monocyte	0	+	-	-
CD56	NK cell	0	-	-	-
CD228	stem cell	0	-	-	-
CD280	stem cell	0	-	+	-
CD338	stem cell	0	-	-	-
CD277	stem cell	0	+	-	670
CD318	stem cell	0	-	-	-
CD227	stem cell	0	-	-	-
CD3d	T cell	0	+	-	-
CD3e	T cell	0	+	-	-
CD3g	T cell	0	-	-	-

