Highly sensitive global phosphorylation analysis of complex biological samples using ERLIC in conjunction with LC-MS

Dissertation

zur Erlangung des akademischen Grades Dr. rer. nat.

eingereicht am Fachbereich Chemie

der Fakultät für Mathematik, Informatik und Naturwissenschaften

der Universität Hamburg

vorgelegt von

Stefan Loroch

aus Bielefeld

Dortmund, 2019

Gutachter

Prof. Dr. Albert Sickmann

- Leibniz-Institut für Analytische Wissenschaften – ISAS – e.V. , Dortmund
- Medizinische Fakultät der Ruhr-Universität Bochum
- Department of Chemistry, University of Aberdeen

Prof. Dr. Hartmut Schlüter

- Institut für Klinische Chemie und Laboratoriumsmedizin des Universitätsklinikums Hamburg-Eppendorf
- Institut für Biochemie der Universität Hamburg

Disputation und Druckfreigabe am 14.08.2020

Angefertigt im Zeitraum von April 2015 bis November 2019
am Leibniz-Institut für Analytische Wissenschaften – ISAS – e.V.
in der Otto-Hahn-Straße 6b, 44227 Dortmund

(I) Peer-reviewed articles

First author research articles

Loroch, S.; Zahedi, R. P.; Sickmann, A., *Highly sensitive phosphoproteomics by tailoring solid-phase extraction to electrostatic repulsion-hydrophilic interaction chromatography*. **Analytical chemistry 2015, 87 (3), 1596-604.**

Loroch, S.; Schommartz, T.; Brune, W.; Zahedi, R. P.; Sickmann, A., *Multidimensional electrostatic repulsion-hydrophilic interaction chromatography (ERLIC) for quantitative analysis of the proteome and phosphoproteome in clinical and biomedical research*. **Biochimica et biophysica acta 2015, 1854 (5), 460-468.**

Loroch, S.; Trabold, K.; Gambaryan, S.; Reiss, C.; Schwierczek, K.; Fleming, I.; Sickmann, A.; Behnisch, W.; Zieger, B.; Zahedi, R. P.; Walter, U.; Jurk, K., *Alterations of the platelet proteome in type I Glanzmann thrombasthenia caused by different homozygous delG frameshift mutations in ITGA2B*. **Thrombosis and haemostasis 2017, 117 (3), 556-569.**

Shared first author research articles

Gonczarowska-Jorge, H.; **Loroch, S.**; Dell'Aica, M.; Sickmann, A.; Roos, A.; Zahedi, R. P., *Quantifying missing (phospho-) proteome regions with the broad-specificity protease subtilisin.* **Analytical chemistry 2017, 89 (24), 13137-13145.**

Shared first author review articles

L**oroch, S.**; Dickhut, C.; Zahedi, R. P.; Sickmann, A., *Phosphoproteomics--more than meets the eye.* **Electrophoresis 2013, 34 (11), 1483-92.**

Pagel, O**.; Loroch, S.**; Sickmann, A.; Zahedi, R. P., *Current strategies and findings in clinically relevant post-translational modification-specific proteomics.* **Expert review of proteomics 2015, 12 (3), 235-53.**

Co-authored articles

Schommartz, T.; **Loroch, S.**; Alawi, M.; Grundhoff, A.; Sickmann, A.; Brune, W., *Functional dissection of an alternatively spliced herpesvirus gene by splice site mutagenesis.* **Journal of virology 2016, 90 (9), 4626-36.**

Di Conza, G.; Trusso Cafarello, S.; **Loroch, S.**; Mennerich, D.; Deschoemaeker, S.; Di Matteo, M.; Ehling, M.; Gevaert, K.; Prenen, H.; Zahedi, R. P.; Sickmann, A.; Kietzmann, T.; Moretti, F.; Mazzone, M., *The mTOR and PP2A pathways regulate PHD2 phosphorylation to fine-tune HIF1alpha levels and colorectal cancer cell survival under hypoxia.* **Cell reports 2017, 18 (7), 1699-1712.**

Beck, F.; Geiger, J.; Gambaryan, S.; Solari, F. A.; Dell'Aica, M.; **Loroch, S.**; Mattheij, N. J.; Mindukshev, I.; Potz, O.; Jurk, K.; Burkhart, J. M.; Fufezan, C.; Heemskerk, J. W.; Walter, U.; Zahedi, R. P.; Sickmann, A., *Temporal quantitative phosphoproteomics of ADP stimulation reveals novel central nodes in platelet activation and inhibition.* **Blood 2017, 129 (2), e1-e12.**

Bruning, U.; Morales-Rodriguez, F.; Kalucka, J.; Goveia, J.; Taverna, F.; Queiroz, K. C. S.; Dubois, C.; Cantelmo, A. R.; Chen, R.; **Loroch, S.**; Timmerman, E.; Caixeta, V.; Bloch, K.; Conradi, L. C.; Treps, L.; Staes, A.; Gevaert, K.; Tee, A.; Dewerchin, M.; Semenkovich, C. F.; Impens, F.; Schilling, B.; Verdin, E.; Swinnen, J. V.; Meier, J. L.; Kulkarni, R. A.; Sickmann, A.; Ghesquiere, B.; Schoonjans, L.; Li, X.; Mazzone, M.; Carmeliet, P., *Impairment of angiogenesis by fatty acid synthase inhibition involves mTOR malonylation*. **Cell metabolism 2018, 28 (6), 866-880.e15.**

Ostermann, E.; **Loroch, S**.; Qian, Z.; Sickmann, A.; Wiebusch, L.; Brune, W., *Activation of E2F-dependent transcription by the mouse cytomegalovirus M117 protein affects the viral host range.* **PLoS pathogens 2018, 14 (12), e1007481.**

Shema, G.; Nguyen, M. T. N.; Solari, F. A.; **Loroch, S.**; Venne, A. S.; Kollipara, L.; Sickmann, A.; Verhelst, S. H. L.; Zahedi, R. P., *Simple, scalable, and ultrasensitive tip-based identification of protease substrates.* **Molecular & cellular proteomics 2018, 17 (4), 826-834.**

Rodziewicz, P.; **Loroch, S.**; Marczak, L.; Sickmann, A.; Kayser, O., *Cannabinoid synthases and osmoprotective metabolites accumulate in the exudates of Cannabis sativa L. glandular trichomes.* **Plant science 2019, 284, 108-116.**

Nguyen, C. D. L.; Malchow, S.; Reich, S.; Steltgens, S.; Shuvaev, K. V.; **Loroch, S.**; Lorenz, C.; Sickmann, A.; Knobbe-Thomsen, C. B.; Tews, B.; Medenbach, J.; Ahrends, R., *A sensitive and simple targeted proteomics approach to quantify transcription factor and membrane proteins of the unfolded protein response pathway in glioblastoma cells.* **Scientific reports 2019, 9 (1), 8836.**

Research articles accepted

Muscolino, E.; Schmitz, R.; **Loroch, S.;** Caragliano, E.; Schneider, C.; Rizzato, M.; Kim, Y.; Krause E.; Lisnić, V. J.; Sickmann, A..; Reimer, R.; Ostermann, E.; Brune, W., *Induced protein aggregation and selective autophagy as a viral immune evasion mechanism.* **accepted**

Research articles submitted

Brandt, S.; Reginskaya, I.; Schilling, M.; Janasek, D., Ahrends, R.; Sickmann, A.; Franzke, J.; **Loroch, S.,** *Dielectric Barrier-nanoESI overcomes the ion flux problem of modern LC-MS/MS.* **submitted**

Research articles in preparation

Loroch, S.; Schommartz, T.; Poetz, O.; Brune, W.; Zahedi, R. P.; Sickmann, A., *Time-Resolved Proteomics and Phosphoproteomics analysis of MCMV-infected mouse fibroblast cells.* **in preparation**

Loroch, S.; Schommartz, T.; Brune, W.; Zahedi, R. P.; Sickmann, A., *Excellent sensitivity through excellent recovery – ERLIC outperforms TiO_2 for phosphoproteomics with low abundant samples.* **in preparation**

(II) Awards

Innovation Award for Young Scientists of the German Society for Biochemistry and Molecular Biology (Gesellschaft für Biochemie und Molekularbiologie) at the **Annual Meeting of the Signal Transduction Society 2016** in Weimar for the oral presentation: *Specific but antibody-free absolute quantification of signaling events in cancer patient tissues using ERLIC and targeted LC-MS*

PhD Student Award at the **ISAS - PhD Summerschool 2016** in Dortmund for the oral presentation: *Excellent sensitivity through excellent recovery - ERLIC outperforms TiO_2-MOAC in quantitative phosphoproteomics with minute amounts of sample*

Poster Award at the **ANAKON 2013** in Essen (Analytikkonferenz der Gesellschaft Deutscher Chemiker) for: *ERLIC - a powerful tool in multidimensional LC-MS-based approaches for global proteome and phosphoproteome analysis of human cells*

(III) Oral conference presentations (selection)

03/2019: **Proteomic Forum 2019 and XIII. Annual Congress of the European Proteomics Association** in Potsdam: *DB-nESI overcomes the ion flux problem of modern nanoLC-MS/MS*

03/2019: **52nd Annual Conference of the German Society for Mass Spectrometry** (Deutsche Gesellschaft für Massenspektrometrie) in Rostock: *Dielectric Barrier – nanoESI overcomes the ion flux problem of modern LC-MS/MS*

03/2017: **50th Annual Conference of the German Society for Mass Spectrometry** (Deutsche Gesellschaft für Massenspektrometrie) in Kiel: *Excellent sensitivity through excellent recovery – ERLIC for absolute quantification of low abundant protein phosphorylation events in cancer patient tissue*

10/2016: **20th Annual Conference of the Signal Transduction Society** in Weimar: *Specific but antibody-free absolute quantification of signaling events in cancer patient tissues using ERLIC and targeted LC-MS*

04/2016: **49th Annual Conference of the German Society for Mass Spectrometry** (Deutsche Gesellschaft für Massenspektrometrie) in Hamburg *Quantitative phosphoproteomics on a µg-scale – A straightforward and highly sensitive ERLIC-SCX/RP-LC-MS strategy*

06/2015: **22. Arbeitstagung - Mikromethoden in der Proteinchemie** in Dortmund – invited speaker: *Electrostatic repulsion-hydrophilic interaction chromatography*

03/2015: **Proteomic Forum 2015** in Berlin: *Quantitative proteomics meets highly sensitive phosphoproteomics - a simple workflow that combines both for low sample amount experiments*

10/2014: **13th Human Proteome World Congress** (HUPO) in Madrid, Spain: *Quantitative phosphoproteomics on a µg-scale – A straightforward and highly sensitive ERLIC-SCX/RP-LC-MS strategy*

(IV) Table of content

(V) List of abbreviations and protein names

ACN	Acetonitrile
AIP1	Programmed cell death 6-interacting protein
ATF6	Cyclic AMP-dependent transcription factor ATF6
BAD	Bcl-2 antagonist of cell death
BAK	Bcl-2 homologous antagonist/killer
BAX	Apoptosis regulator BAX
BCA	Bicinchoninic acid assay
BET1	BET1-like protein
BID	BH3-interacting domain death agonist
CALR	Calreticulin
CAMK2D	Calcium/calmodulin-dependent protein kinase type II subunit delta
CASP8	Caspase-8
CDC23	Cell division cycle protein 23 homolog
CDK	Cyclin-dependent kinase
CDKN1B	Cyclin-dependent kinase inhibitor 1B
CMV	Cytomegalovirus
CREB1	cAMP-responsive element-binding protein 1
CYC	Cyclin
DDA	Data-dependant acquisition
DNA	Deoxyribonucleic acid
DTT	Dithiotreithol
DUSP	Dual specificity protein phosphatase
E	Early (expression kinetic of viral proteins)
E2F4	Transcription factor E2F4
EDTA	Ethylenediaminetetraacetic acid
EEF1D	Elongation factor 1-delta
EGLN-1	Egl nine homolog 1
EIF1AX	Eukaryotic translation initiation factor 1A, X-chromosomal
EIF2A	Eukaryotic translation initiation factor 2A
EIF2AK4	Translation initiation factor 2-alpha kinase 4
EIF3B	Translation initiation factor 3 subunit B
EIF3J2	Eukaryotic translation initiation factor 3 subunit J-B
EIF4G2	Translation initiation factor 4 gamma 2
ER	Endoplasmic reticulum
ERK1	Mitogen-activated protein kinase 3
ERK2	Mitogen-activated protein kinase 1
ERK5	Mitogen-activated protein kinase 7
ERLIC	Electrostatic repulsion-hydrophilic interaction chromatography
ESI	Electrospray ionization
EtOH	Ethanol
FA	Formic acid
FAM20C	Extracellular serine/threonine protein kinase FAM20C
FDR	False discovery rate
FWHM	Full width at half maximum
GRP78	Endoplasmic reticulum chaperone BiP
GRP94	Endoplasmin, HSP90B1
H2AX	Histone H2AX

List of abbreviations and protein names

H2AY	Core histone macro-H2A.1
HCMV	Human cytomeglovirus
HIF1?	Hypoxia-inducible factor 1-alpha
HILIC	Hydrophilic interaction liquid chromatography
hpi	Hours post-infection
HPLC	High-performance liquid chromatography
HSP40	DnaJ homolog subfamily B member 1
HSP90B1	Endoplasmin, GRP-94
HSPA5	78 kDa glucose-regulated protein
HYOU1	Hypoxia up-regulated protein 1
I.D.	Inner diameter
IAA	Iodoacetamide
ID	Identification (peptides or proteins)
IE	Immediate early (expression kinetic of viral proteins)
IE1	Immediate-early protein 1 , vira
IE3	Immediate-early protein 3 , viral
IEX	Ion exchange chromatography
IFI-204	Interferon-activable protein 204
IMAC	Immobilized metal ion-affinity chromatography
iTRAQ	Isobaric tag for relative and absolute quantification
JAK	Tyrosine-protein kinase JAK1
JIP4	C-Jun-amino-terminal kinase-interacting protein 4
KIP1	Cyclin-dependent kinase inhibitor 1B
KMT2B	Histone-lysine N-methyltransferase 2B
LC-MS	Mass spectrometry online-coupled to liquid chromatography
LIG1	DNA ligase 1
LL	Leaky late (expression kinetic of viral proteins)
LLOQ	Lower limit of quantification
LOD	Limit of detection
m/z	Mass-to-charge ratio
MAP2K4	Dual specificity mitogen-activated protein kinase kinase 4
MAP4K4	Mitogen-activated protein kinase kinase kinase kinase 4
MAP4K5	Mitogen-activated protein kinase kinase kinase kinase 5
MAPK	Mitogen-activated protein kinase
MAPK7	Mitogen-activated protein kinase 7
MAPKAPK2	MAP kinase-activated protein kinase 2
MCM	Minichromosom maintenance, syn: DNA replication licensing factor MCM
MCMV	Murine cytomegalovirus, mouse cytomegalovirus
MDC1	Mediator of DNA damage checkpoint protein 1
MOAC	Metal oxide-affinity chromatography
MS	Mass spectrometry, mass spectrometer or MS survey scan
MS/MS	Tandem MS, MS2
MS2	Tandem MS, MS/MS
MSK2	Ribosomal protein S6 kinase alpha-4
MW	Molecular weight
NCBP1	Nuclear cap-binding protein subunit 1
nESI	Nano-electrospray ionization
NF-κB	Nuclear factor NF-kappa-B
OST	Oligosaccharyl transferase complex (ER)
OST	Oligosaccharyl transferase complex

P53	Cellular tumor antigen p53, Tp53
PCD	Programmed cell death
PERK	Eukaryotic translation initiation factor 2-alpha kinase 3
pI	Isoelectric point
PIK3C	Phosphatidylinositol 4-phosphate 3-kinase C2 domain-containing
PIK3R	Phosphatidylinositol 3-kinase regulatory subunit gamma
PLC	Phospholipase C
POLR2A	DNA-directed RNA polymerase II subunit RPB1
ppm	Parts per million
PPM1G	Protein phosphatase 1G
PPP1R12A	Protein phosphatase 1 regulatory subunit 12A
PPP4R2	Serine/threonine-protein phosphatase 4 regulatory subunit 2
PRKRA	Interferon-inducible double stranded RNA-dependent protein kinase activator A
PRM	Parallel reaction monitoring
pS, pSer	Phosphoserine
PSM	Peptide spectrum match
pT, pThr	Phosphothreonine
PTM	Post-translational modification
pY, pTyr	Phosphotyrosine
QC	Quality control
R^2, R2	Pearson product-moment correlation coefficient
RAD23	UV excision repair protein RAD23 homolog
RBMX	RNA-binding motif protein
RIP3	Receptor-interacting serine/threonine-protein kinase 3
RNPS1	RNA-binding protein with serine-rich domain 1
RP	Reversed phase
RPB1	DNA-directed RNA polymerase II subunit RPB11
RPL23A	60S ribosomal protein L23a
RPL27	60S ribosomal protein L27a
RPL32	Putative 60S ribosomal protein L32'
RPL35	60S ribosomal protein L35a
RPL39	60S ribosomal protein L39
RPN1	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 1
RPN2	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 2
RPS14	40S ribosomal protein S14
RPS15	40S ribosomal protein S15a
RPS18	40S ribosomal protein S18
RPS24	40S ribosomal protein S24
RPS27A	Ubiquitin-40S ribosomal protein S27a
RPS6KA3	Ribosomal protein S6 kinase alpha-3
RPS6KS4	Ubiquitin-40S ribosomal protein S27a
RSK2	Ribosomal protein S6 kinase alpha-3
RT	Retention time
SAX	Strong anion exchange chromatography
SCX	Strong cation exchange chromatography
SEC20	Vesicle transport protein SEC20
SEC61	Protein transport protein Sec61 subunit alpha isoform 1
SEC61B	Protein transport protein Sec61 subunit beta
Ser	Serine
SF3B1	Splicing factor 3B subunit 1

SIL	Stable isotope labelled
SILAC	Stable isotope labelling by amino acids in cell culture
SNAP23	Synaptosomal-associated protein 23
SNRNP70	U1 small nuclear ribonucleoprotein 70 kDa
SPE	Solid-phase extraction
SPP	Superficial porous particles
SRRM1	Serine/arginine repetitive matrix protein 1
SSR3	Translocon-associated protein subunit gamma
STAT	Signal transducer and activator of transcription
STMN1	Stathmin
STX18	Syntaxin-18
STX4	Syntaxin-4
STX6	Syntaxin-6
STX7	Syntaxin-7
sub-2µ	HPLC column particles < 2 µm
TAO1	Serine/threonine-protein kinase TAO1
TFA	Trifluoroacetic acid
TiO2	Titandioxid
TL	True late (expression kinetic of viral proteins)
TMT	Tandem mass tag
TNF	Tumor necrosis factor
TNFR	Tumor necrosis factor receptor
TOP1	DNA topoisomerase 1
TOP2	DNA topoisomerase 2
TP53	Cellular tumor antigen p53
TRAP	Translocon-associated protein
UPR	Unfolded protein response
USE1	Vesicle transport protein USE1
VAMP2	Vesicle-associated membrane protein 2
VAMP3	Vesicle-associated membrane protein 3
VAMP4	Vesicle-associated membrane protein 4
YWHAE	14-3-3 protein epsilon

1. Abstract in English

With the advent of mass spectrometry online-coupled to liquid chromatography (LC-MS) as high-content-screening method for proteomes and phosphoproteomes of complex biological samples, quantification of thousands to ten thousands of proteins and protein phosphorylation events became possible. Whereas deep proteome analyses can nowadays be conducted with comparably low sample amounts < 50 µg (100,000 HeLa cells), deep phosphoproteomics still requires ~100-times more starting material to reach an appreciable depth. This is not only attributed to the sub-stoichiometric nature of protein phosphorylation events, but also to poor analyte recovery during sample processing – a critical problem for application in clinical and biomedical research, which is often restricted in sample material.

The present work demonstrates the advantages of electrostatic repulsion-hydrophilic interaction chromatography (ERLIC) for phosphoproteomics with low sample amounts in the µg-range. ERLIC proved to allow for highly efficient phosphopeptide enrichment from proteolytic digests with trypsin or subtilisin (which proved superior for covering proline-rich regions) and to exhibit 1.7-fold higher analyte recovery compared to the widely used TiO_2-affinity enrichment. Unlike TiO_2, which exhibits poor recovery for a substantial fraction of analytes, ERLIC does not alter the phosphopeptide stoichiometry. It further outperformed TiO_2 in the detection of significantly regulated phosphorylation events (by means of p-values) from aliquots of a biological sample, namely cytomegalovirus (CMV) infected fibroblasts, but at the expense of increased LC-MS analysis time. In addition, ERLIC in conjunction with SIL-peptides and targeted LC-MS enabled the detection of a significant 3.5-fold downregulation of Ser125 phosphorylation on EGLN-1 (HIF1α-mediated hypoxia response) in cancerous tissue of ten colon cancer patients.

In the second part, ERLIC was exploited for time-resolved deep proteomics and phosphoproteomics with minute amounts of murid cytomegalovirus- (MCMV-) infected mouse fibroblasts (120 µg).

Abstract in English

MCMV serves as mouse model for human CMV (HCMV) infection, which is a major death cause of immunocompromised individuals and connected to sever neurological impairments of infants. The resulting dataset provided deep insight into the expression dynamics of 95 viral proteins and 142 viral protein phosphorylation events. 63 viral proteins were quantified with complete time-courses and classified according to expression profiles, which substantially differed from the traditional classification. Hierarchical clustering, functional enrichment analysis and pathway mapping of 6,618 mouse proteins and 2,311 phosphopeptides (complete time-courses) revealed tremendous manipulation of the host cell. Reorganization of the host cell comprised (I) systematic shutdown of MAPK and TNFα signalling, DNA replication and programmed cell-death in parallel to (II) extensive upregulation of cellular resources for splicing, protein synthesis and vesicular transport to enable efficient viral replication. Subsequent comparison with published datasets on HCMV revealed a high degree of similarity on the level of cellular functions. However, the underlying molecular mechanisms significantly differed such as the fine-tuning of MAPK signalling-shutdown, suppression of protease-mediated apoptosis and manipulation of cell cycle proteins including DNA replication. Hence, the results underscore the value of MCMV as mouse model for HCMV, but also unveiled pathways, which do not allow a transfer of findings.

In summary, this work successfully demonstrates the strength of ERLIC as tool for sensitive phosphoproteomics and its successful application to study viral infection in fibroblast cells with low sample amounts in the µg-range.

2. Abstract in German (Zusammenfassung)

Mit dem Einzug der *Flüssigchromatographie mit Massenspektrometrie-Kopplung* (LC-MS) als High-Content-Screening-Methode für die Proteom- und Phosphoproteomanalyse komplexer biologischer Proben ist eine Quantifizierung von Tausend oder Zehntausend Proteinen und Proteinphosphorylierungen möglich geworden. Während umfassende Proteomanalysen mit vergleichsweise geringen Probenmengen < 50 µg (100,000 HeLa-Zellen) durchgeführt werden können, wird für eine umfassende Phosphoproteomanalyse immer noch ~100-mal mehr benötigt – nicht nur wegen der substöchiometrischen Natur der Proteinphosphorylierungen, sondern auch wegen der schlechten Wiederfindungsrate der Analyten in der Probenvorbereitung. Dies ist ein wesentliches Problem für die Anwendung in der klinischen und biomedizinischen Forschung, welche oft mit niedrigen Probenmengen durchgeführt wird.

Der erste Teil dieser Arbeit zeigt die Vorteile der *Elektrostatische Repulsions-Hydrophile Interaktions-Chromatographie* (ERLIC) für die Phosphoproteomanalysen mit geringen Probenmengen. Durch eine Reihe methodischer Experimente wurde gezeigt, dass ERLIC eine effiziente Anreicherung von Phosphopeptiden aus proteolytischen Verdauen mit Trypsin oder auch Subtilisin (welches nachweislich prolinreiche Regionen besser abdeckt) ermöglicht und eine 1,7-fach verbesserte Wiederfindungsrate aufweist als die vielverwendete TiO_2-Affinitätsanreicherung. Im Gegensatz zu TiO_2, welches für einen wesentlichen Teil an Analyten schlechte Wiederfindungsraten zeigt, bleibt die Stöchiometrie der Analyten bei ERLIC unverändert. Weiterhin ermöglicht ERLIC (im Vergleich zu TiO_2) eine verbesserte Detektion differentiell regulierter Proteinphosphorylierungen, was mit identischen Aliquots einer biologischen Probe gezeigt wurde, nämlich Cytomegalovirus- (CMV-) infizierter Fibroblastenzellen aber auf Kosten höherer LC-MS-Analysezeiten. Ebenfalls erlaubte ERLIC die gezielte Anreicherung eines spezifischen Phosphopeptids aus gesundem und erkranktem Gewebe von zehn Kolonkarzinompatienten. Dabei wurde unter Verwendung von SIL-Peptiden und targeted LC-MS eine

signifikante, 3,5-fache Verringerung der Ser125 Phosphorylierung an EGLN-1 im Krebsgewebe nachgewiesen.

Im zweiten Teil wurde mittels ERLIC eine umfassende und zeitaufgelöste Analyse des Proteoms und Phosphoproteoms mit geringen Mengen (120 µg) mauscytomegalovirus- (MCMV-) infizierter Fibroblasten durchgeführt – dabei handelt es sich um das Mausmodell für das humane CMV (HCMV), welches eine Haupttodesursache für immunsupprimierte Menschen ist und schwere Geburtsschäden an Neugeborenen hervorruft. Der Datensatz gibt tiefe Einblicke in die Dynamik der Expression 95 viraler Proteine und 142 viral Proteinphosphorylierung. Für 63 virale Proteine wurde ein vollständiges Expressionsprofil erstellt, was eine Klassifizierung ermöglichte, die sich wesentlich von der traditionellen Klassifizierung unterscheidet. Hierarchische Clustering-, Functional-Enrichment- und Pathway-Analysen von 6,618 mouse proteinen und 2,311 phosphopeptiden (vollständiges Zeitprofil) deckten eine erhebliche Manipulation der Wirtszelle auf. Diese umfasst (I) die Unterdrückung des MAPK- und TNFα- Signalwegs, der DNA-Replikation und den programmierten Zelltod, parallel zur (II) Hochregulation zellulärer Ressourcen für das Spleißen, die Proteinsynthese und den intrazellulären Transport um eine effiziente Virusreplikation zu ermöglichen. Ein Vergleich mit publizierten Daten zu HCMV-infektion zeigte, dass die Manipulationen auf Ebene der Zellfunktionen sehr ähnlich sind, aber dass es auf molekular Ebene wesentliche Unterschiede gibt, wie bei der Feinabstimmung der MAPK-Signalunterdrückung, der Unterdrückung Protease-abhängiger Apoptose und der Manipulation von Proteinen des Zellzyklus und der DNA-Replikation. Die Ergebnisse unterstreichen somit den Wert von MCMV als Mausmodell für HCMV, deuten aber auch auf Signal- und Stoffwechselwege, die keinen direkten Erkenntnistransfer erlauben.

Zusammengefasst: Die Arbeit demonstriert die Vorteile von ERLIC für die sensitive Phosphoproteomanalyse und ihren Einsatz zur Analyse viraler Infektion in Fibroblastenzellen mit geringen Probenmenge im µg-Bereich.

3. Introduction

3.1 Omics technologies

In the past three decades, technological advancements have led to a tremendous number of high-content screening-methods, which allow researchers system-wide monitoring of biological samples. Hence, experiments do not need to focus on known pathways, but can identify novel mechanisms and their interconnection without prior knowledge. Application of such holistic approaches leads to a fundamental change from hypothesis-driven to a non-hypothesis-driven research, which allows conducting experiments without any expectations about the outcome. As system-wide effects can occur on multiple levels, terms such as *genome*, *transcriptome* and *proteome* have been introduced to differentiate between the entities of biomolecule species within biological systems. In 1920, Hans Winkler coined the term *genome,*[1] which is nowadays used for the entity of DNA in a biological system, and the related term *genomics* refers to its analysis. Since genomes of individual organisms are comparably stable, and do not undergo major changes upon most external stimuli, the attention of researchers has increasingly focused on the abundance and composition of proteins, the so-called *proteome* – a term introduced in 1994 by Marc Wilkins.[2]

Already in 1975, 2D gel electrophoresis gave first impressions about the dynamics and complexity of proteomes in organisms leading to the widely accepted hypothesis that proteomes are suitable to reflect the current status of biological systems.[3, 4] With the advent of soft ionization methods, which allowed online-coupling of liquid chromatography with tandem mass spectrometry (LC-MS), qualitative and quantitative analysis of thousands of proteins in a sample became feasible[5] and proteomics became a principal discipline for monitoring changes on the proteome level upon perturbation or alteration of biological systems. Owing to numerous technological advancements, analysis time per sample has been tremendously reduced allowing for time-efficient screening within a few hours of analysis time. Hence,

quantitative proteomics can nowadays be conducted at a reasonable expense of time and costs rendering it a promising future in routine research and, even though not yet within reach, in clinical diagnostics.

3.2 LC-MS-based global proteomics and phosphoproteomics

[...]. While rapid and deep whole proteome-level characterization is currently within reach, comprehensive PTM-level characterization, such as a complete phosphoproteome, appears a decade or more away. [...]

Nicholas M. Riley and Joshua J. Coon, Anal. Chem., 2016, 88 (1), pp 74–94

With the advent of LC-MS, a detection of thousands of proteins in a single sample became possible, which allowed in-depth characterizations of (nearly) entire proteomes. In a given cell type or tissue, ~10,000-12,000 proteins are estimated to be expressed at a given point in time, rendering state-of-the art proteomics studies to cover between 10 % to 80 % of the proteome (1,000-9,000 proteins).[6, 7] Notably, these estimations rely on the idea of "one gene – one protein – one function" and do not consider the full complexity of the proteome. It is important to note that the concept of *protein speciation* – a term introduced by Jungblut and Schlüter to describe the large variety of different protein species originating from one gene[8, 9] – leads to much higher estimates in the range of 1 billion protein species in a cell.[10] Unfortunately, protein species cannot be readily distinguished neither by large-scale proteomics nor by most antibody-based assays, rendering protein numbers in most studies greatly simplified. Hence, it is important to keep in mind that the frequently found term of "proteome coverage" reduces the entity of protein species from one single gene to one single protein.

The depth of proteomics studies usually depends on several parameters. Most importantly, (I) the total sample amount, (III) the dynamic range of the sample and (III) the total LC-MS analysis time available. A deep proteome analysis can nowadays be conducted with standard cell culture amounts or tissue

samples (e.g. human or mouse) as the required starting material is usually < 50 µg of protein (~100,000 HeLa cells). Less comprehensive analyses can be already conducted with considerably lower amounts in the range of 1 µg.[11] Importantly, the dynamic range of biological systems has a critical impact on the required sample material. It usually spans several orders of magnitude – for instance 4 orders in platelets,[12] 7 orders in HeLa cells,[6] and 10 orders in blood plasma[13] – in other words, the most abundant proteins are up to 10 billion times more intense than the lowest ones. Hence, an in-depth proteomics analysis of a high dynamic range-sample requires larger amounts of sample material, otherwise low abundant analytes are below the lower limit of detection (LLOD, Pagel & Loroch et al).[14] Besides, LC-MS analysis time often represents a bottleneck of many research facilities. However, the depth of analyses correlates with the total analysis time rendering it a major limiting factor for analysis depth. However, this problem might be (partially) solved with the advent of faster instruments,[15] which allow for realization of different acquisition concepts (such as data-independent acquisition or BoxCar).[16-19]

In 2014, the groups of Akhilesh Pandey and Bernhard Küster gave impressions about the theoretical depth that can be achieved in proteomics by analyzing dozens of human cell types and tissue samples and further including datasets from public repositories.[20, 21] Both groups postulated the existence of nearly 18,000 different proteins as gene products, hence, presented that 90 % of the ~20,000 annotated genes in human encode for proteins.[22] However, it is important to note that Oliver Serang and Lukas Käll convincingly demonstrated that classical approaches for controlling the false discovery rate (FDR) cannot be readily applied to such extraordinary large dataset. Hence, these datasets might comprise a substantial fraction of false positive identifications.[23] However, the search for yet unknown or only predicted proteins is ongoing and the Nextprot repository (nextprot.org, July-2019) currently lists about 2,500 human proteins without validation of existence – ~1000 predicted and ~1500 just with evidence on transcript level.

Even though proteomics has reached a tremendous depth, standard proteomics experiments do not exhibit information on protein activity, conformation and subcellular localization, which are widely regulated by patterns of post-translational protein modifications (PTMs). However, as a versatile tool in life science, LC-MS has proven to enable deciphering of PTM profiles using specialized workflows. Especially, dynamic protein phosphorylation is one of the most extensively studied PTMs and has been found to facilitate a wide-range of signaling processes and subsequent activation or fine-tuning of biochemical processes. Notably, phosphorylation has been shown to play a central regulatory role in the activity of the majority of eukaryotic proteins and can be found on ~70 % of all proteins in a human cell.[24] Nevertheless, the lack of knowledge about the regulation, interplay and involvement in disease state is rather disconcerting,[14] rendering the development of simple and robust methods for comprehensive analysis of *phosphoproteomes* an ongoing task.

In comparison to a conventional proteomics experiment, LC-MS-based phosphoproteomics needs to overcome certain challenges. First, the labile and dynamic nature of the phosphorylation can lead to artificial dephosphorylation during sample preparation and renders MS-based identification more challenging. Second, the sub-stoichiometric nature of most phosphorylation events renders additional sample preparation steps for purification of phosphorylated analytes crucial.[25] In other words, the purification aims at depletion of the high-abundant background of non-modified analytes, as usually only a minor fraction of a give phosphorylation site (< 20 % non-stimulated cells) is actually phosphorylated (also referred to as site occupancy). This is especially critical when aiming at analyzing very low-abundant phosphorylation events, such as the majority of tyrosine phosphorylation, which usually requires specific antibody-based enrichment.[24, 26] Consequently, the sample material required to achieve an appreciable depth is 10-100-fold higher than for a proteomics experiment and careful optimization of purification strategies towards highest analyte recovery (sensitivity) is mandatory, if experiments have to be conducted with limited sample material. Studies conducted with minute amounts

of sample do usually not exceed a depth of ten-thousand phosphorylation sites and, since signaling proteins (such as kinases and phosphates) are often of low abundance and the respective phosphorylation events within signaling pathways are often below the LLOD.[14] Notably, according to studies from the Mann group and the Olsen group, at least 40,000 phosphorylation sites can be expected in human cell lines when employing vast amounts of cells to yield milligrams of protein as starting material.[24, 27] However, to identify such high numbers, experiments usually require thorough stimulation of the cells' intracellular signaling pathways (e.g. using epidermal growth factor), blocking of phosphatases (e.g. using pervanadate)[24] or cell phase arrest[28] leading to extensive intracellular protein phosphorylation events.

In summary, limitations in proteomics rather result from restricted LC-MS analysis time – the bottleneck of most research facilities. In contrast, limitations in phosphoproteomics rather originate from insufficient amounts of sample material. Hence, for application of phosphoproteomics in clinical and biomedical research, workflows require careful optimization towards highest analyte recovery to reach appreciable depths.

3.2.1 Bottom-up proteomics – a key-strategy for protein detection

The most comprehensive proteome analyses in terms of detected proteins can be conducted using so-called LC-MS-based *bottom-up proteomics* approaches, which combine proteolytic digestion of protein fractions followed by peptide sequencing via determination of the exact mass and acquisition of tandem MS (MS/MS, MS^2)[29] for subsequent database search (Figure 1).[30, 31] Digest conditions are usually adjusted to yield peptides in a range of 6-50 amino acids (aa) with an ideal length of 8-14 aa. If a detected peptide is *unique* to a given protein, hence its sequence can only derive from one specific protein, unambiguous protein identification is directly possible. For instance, an *in silico* digest of the human database (Uniprot, taxonomy: human) with the most-widely used enzyme trypsin,[32] reveals 66 % of all

peptides as unique - accounting for ~600,000 peptides. Hence, these peptides can serve as surrogates for the proteins and, in addition, the signal intensity in MS reflects the protein's abundance. In contrast, non-unique peptides (shared between two or more proteins) only provide ambiguous information about protein composition of a sample, hence, such peptides are usually only of limited value and are often not considered in data analysis. To determine whether peptides are unique, a comprehensive and well-annotated sequence databases is required which, in an ideal case, exclusively contains protein from the organism to be analyzed. Otherwise, the additional protein entries increase the likelihood of peptides being falsely assigned as non-unique.

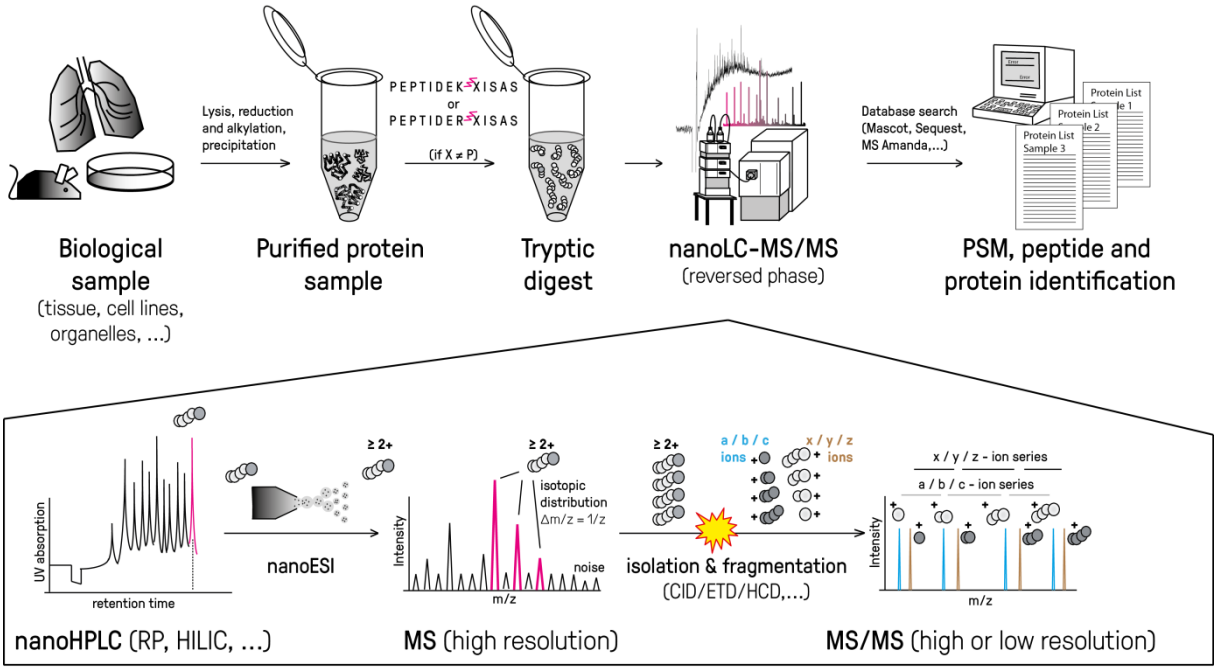


Figure 1: Overview of a typical bottom-up proteomics workflow. Peptides from proteolytic digests are separated by nanoHPLC and the effluent is introduced to the MS by nESI, facilitating ionization and subsequent transfer to the gas phase (see also section 3.2.1.2 and 3.2.1.3). The exact peptide masses are determined by a high resolution MS (survey scan), whereas isolation, fragmentation of peptides and subsequent acquisition of MS/MS (in parallel or serial) allows to derive the amino acid sequence.

3.2.1.1 Peptide sequencing using MS

MS-based peptide identification is usually conducted via determination of (I) the peptide mass (MS) and (II) fragmentation of the peptide in the gas-phase to yield an MS/MS spectrum which reflects the amino acid sequence (Figure 1). Depending on the type of gas-phase fragmentation, a peptide usually generates defined types of fragment ions. Depending on the type of fragmentation, peptide backbone fragmentation can occur at different positions leading to the formation of amino-terminal a-, b-, c- or carboxy-terminal x-, y-, z-ions according to the nomenclature by Roepstorff and Fohlmann in 1984 (Figure 2A).[33] The most-widely used collision-based fragmentation methods result in peptide backbone cleavage at the amide group – usually the bond of lowest energy. This leads to generation of b-ions if a charge (e.g. a proton) is retained at the amino-terminal fragment and to y-ions if a charge is retained at the carboxy-terminal fragment (Figure 2B, C). The resulting ion-series reflect the peptide's amino acid sequence, hence, the resulting MS/MS enables sequencing of the peptide as 19 out of 20 proteinogenic amino acid differ in their nominal mass. Notably, only isoleucine and leucine are isobaric and cannot be differentiated solely by mass spectrometry, but peptide sequences might still be determined with the help of a sequence databases.

For peptide identification, search engines such as Sequest,[31] Mascot,[30] MS Amanda[34] or X!Tandem[35, 36] systematically compare the acquired MS/MS against theoretical spectra, generated by *in-silico* digest of reference databases. Implemented scoring algorithms assign a score to each hit, which reflects the likelihood of a correct match. Notably, identification solely based on scoring does not allow for discrimination between false positive and false negative hits,[37] hence, database shuffling and other statistical tools are usually employed to determine adequate score threshold to control the *false discovery rate* (FDR).[38] In most projects the FDR is adjusted to 1%, in other words, one out of hundred hits can be expected to be false positive and the likelihood of false positive identification increases with lower scores. Hence, proteins detected by only one unique peptide are not reliably identified, especially if the

peptide is close to the score threshold. In contrast, proteins identified with two or more unique peptides are widely-accepted as truly identified.[5]

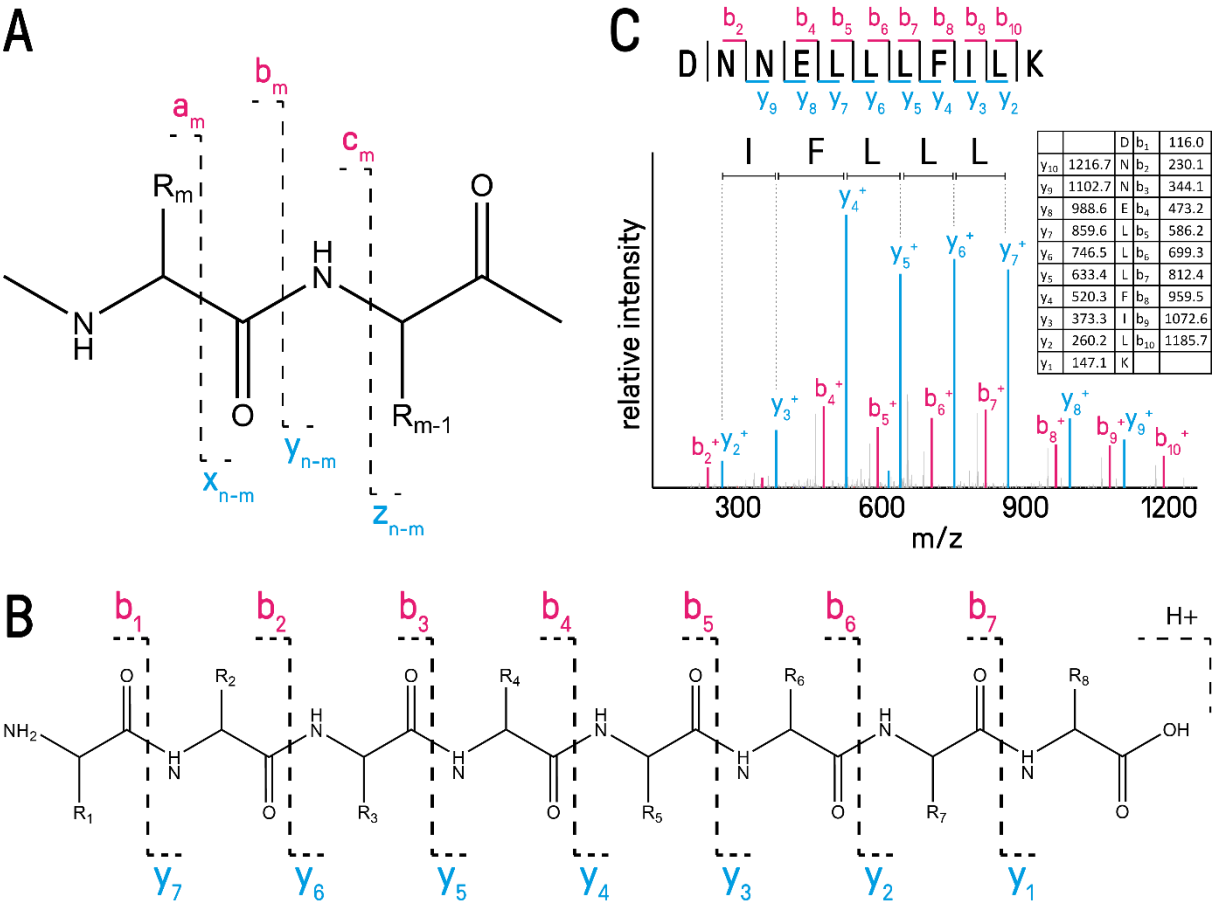


Figure 2: Fragmentation of the peptide backbone by imparting energy to a peptide yields different types of ions. (A) Designation of fragment ions according to Roepstorff-Fohlmann-Biemann. (B) Collision-induced dissociation methods usually results in a dominant y- and a less dominant b-ion series. (C) The example MS/MS of the peptide DNNELLLFILK shows that ion series reflect the peptides amino acids sequence (only singly charged ions are annotated). Notably, the two isobaric amino acids leucine and isoleucine cannot be differentiated by mass spectrometry, hence, unambiguous sequence determination is not possible without additional information (such as a sequence databases).

3.2.1.2 Identification of peptides by mass spectrometry online-coupled to liquid chromatography (LC-MS)

For peptide identification, modern hybrid mass spectrometers determine (I) the mass usually within 2 parts per million (ppm) in a *survey scan* (MS or MS1) followed by gas-phase isolation and fragmentation of the respective peptide to yield MS/MS. State-of-the-art instruments allow for acquisition of one MS and several MS/MS spectra on a second to millisecond timescale.[39, 40] The automated duty cycles start with an initial survey scan (MS) followed by consecutive MS/MS of the most intense ions detected (topN with N usually 1-20). A *dynamic exclusion* is often employed which prevents re-fragmentation of the same precursor ion multiple times to avoid redundant information. Protein identification with such instruments does not require fractionation on the protein level (for instance by 2D-PAGE or size exclusion chromatography) as the peptides generated upon digest can be separated by *high-performance liquid chromatography* (HPLC) online-coupled to the MS for peptide identification (section 3.2.1.3). This *bottom-up* strategy exhibits several advantages compared to the analysis of full length proteins (*top-down*): (I) MS/MS spectra are of lower complexity allowing for straightforward and automated sequencing, (II) peptides are more homogenous in terms of physicochemical properties (especially if trypsin is used) and can be well separated by HPLC, (III) the improved ionization efficiency of most peptides allows for higher sensitivity and (IV) the lower molecular mass allows for higher mass accuracy and baseline separation of isotope signals.[5]

Eluting peptides are introduced to the MS by electrospray ionization (ESI)[41] or, for lower flowrates and higher sensitivity, by nano-ESI (nESI).[42] Both play a pivotal role as techniques for ionization and evaporation of large biomolecules (such as peptides, proteins and lipids) from solution, as the soft ionization does usually not result in analyte fragmentation. If used as an LC-MS interface, the electric field of a few kilovolts is established between the HPLC effluent and the orifice of the mass spectrometer generating small solvent droplets comprising the charged analytes. With continuous solvent evaporation

(in the electric field) the charge density increases to a critical level (Rayleigh-limit) leading to coulomb-explosions, which generate smaller droplets releasing single ions into the gas phase. The exact mechanisms are still not fully understood,[43] however, soft ionization of analytes via ESI was one of the major breakthroughs for biomolecule detection (nobel prize in chemistry 2002 for John Fenn) and is a fundamental prerequisite for employing MS as a versatile detector for HPLC.

3.2.1.3 State-of-the-art LC-MS instrumentation

LC-MS-based proteomics is widely conducted using reversed phase (RP) nanoHPLC with columns of inner diameters (I.D.) of 75-100 µm packed with 2-5 µm silica-based particles. This allows for flow rates of ~ 250-500 nL / min resulting in considerably higher sensitivity than conventional analytical HPLC systems (4.6 mm ID columns and 1 mL / min flow rate).[44, 45] Since sensitivity increases linearly with decreasing flow rates and is approximately inversely proportional to the square of the column diameter, columns with diameters ≤ 50 µm have been shown to increase sensitivity even further, but these columns are not yet in routine use.[46, 47] As separation efficiency (mass transfer) depends on the surface area of the stationary phase, column particles < 2 µm (sub-2µ) became employed more frequently. As a drawback, smaller particles form a denser column of high backpressure, rendering sub-2µ columns only applicable in Ultra-HPLC systems (≥ 800 bar).[48] As an alternative, superficial porous particles (SPP), made of a non-porous silica core and a porous shell layer, can be operated at modest pressure with conventional HPLC systems (Figure 3A-C).[49] In comparison to porous particles, SPP exhibits low intraparticle mass transfer, hence, show reduced peak broadening as analyte diffusion through the particle cannot occur. Most recently even sub-2µ SPP particles have been introduced to the market (Waters), which might exhibit even higher separation efficiency. However, it remains questionable whether the benefit of SPP can be further enhanced by particle size reduction as the accessible surface area might be quite similar. As an important alternative to particles, monolithic columns consist of a fiber mesh, which can be generated by co-polymerization of styrene and

divinylbenzene inside a glass capillary.[50] These columns demonstrate high separation efficiency at very low backpressure and exhibit excellent resistance to sample impurities, harsh conditions or undigested protein, which usually clogs or dissolves other column types. Notably, even though monolithic columns have a comparably low loading capacity (Figure 3D),[50] they serve as excellent tool for quality control of proteolytic digests.[51]

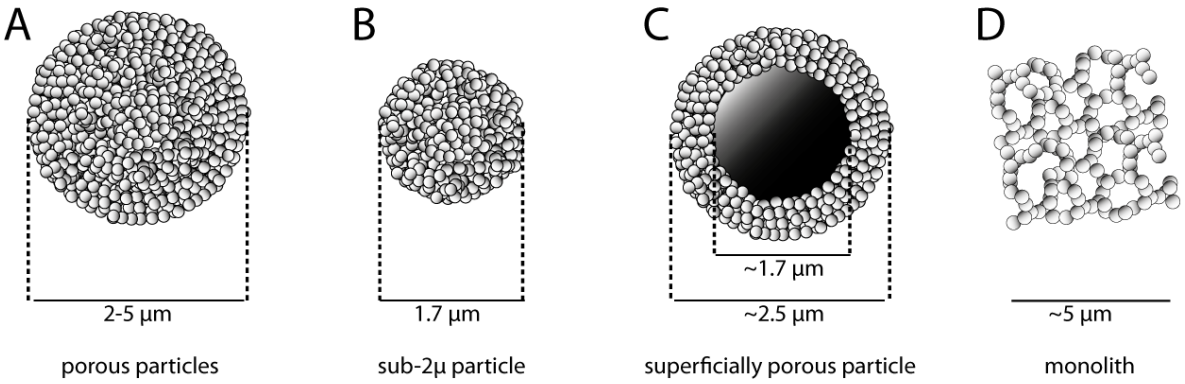


Figure 3: Columns with a wide variety of modified silica particles are frequently employed for peptide chromatography. Compared to classical porous particles (A) sub-2µ particles (B) have a higher separation efficiency due to higher surface area, but also considerable higher backpressure. In contrast, superficially porous particles (C) exhibit similar efficiency (due to the reduced intraparticle mass transfer) and can be operated at modest pressure, but have a ~30 % lower sample capacity. In contrast, the stationary phase of monolithic columns (D) consists of a copolymerized fibre mesh, which allows operation at very low backpressure, but exhibits limitations in loading capacity. Notably, monolithic columns tolerate a wide variety of sample impurities and harsh conditions, which would clog or dissolve other types of columns.

NanoHPLC is usually online-coupled via a nanoESI interface to hybrid MS instruments such as quadrupole-Orbitrap (e.g. Q-Exactive HF-X),[40, 52] linear ion trap-Orbitrap (e.g. LTQ Orbitrap Elite, Figure 4a),[53, 54] or even quadrupole-Orbitrap-linear trap (e.g. Fusion Lumos Tribrid) instruments.[55] Such instruments allow for acquisition rates in the range of 5-40 Hz, hence, generate ~20,000 - 150,000 spectra per hour of which 10-50% usually have sufficient quality for peptide identification (using automated database search algorithms, section 3.2.1.1). This usually allows for the identification of thousands to ten thousands non-redundant unique peptides from thousands of proteins per hour (e.g. in a human, mouse or yeast sample).[56, 57] However, high scan rates require short ion injection times (time of ion accumulation for spectra acquisition), which can reduce overall identification when working with

low sample amounts, highlighting the demand for ionization methods of higher flux.[40] Importantly, a dielectric barrier-nESI can overcome this problem by applying a well-triggered rectangular voltage pulse via a dielectric-barrier to the analyte solution (Stark et al.[58] and Brandt et al., Loroch, submitted).

Most instruments are operated in a topN data-dependent acquisition (DDA) mode with one high resolution MS (R=60,000-120,000 with resolution $R = m/\Delta m$ and a transient of 128-256 ms) to determine the exact mass of the peptides (usually ≤ 2 ppm) followed by N MS/MS scans (N = 5-20) of the most abundant peptide ions. In case of linear ion trap-Orbitrap instruments, the linear ion trap can be used for isolation, fragmentation via collision induced dissociation (CID) and acquisition of low resolution MS/MS in parallel to the acquisition of the high resolution MS in the Orbitrap (or high field Orbitrap, Figure 4). In case of quadrupole-Orbitrap instruments, a high resolution MS scan is acquired in the Orbitrap followed by 10-20 consecutive high resolution MS/MS. With these instruments, precursor ions are isolated in the quadrupole and fragmented by higher energy collision induced dissociation (HCD) in the collision cell before ions are injected into the Orbitrap. However, since CID usually leads to a dominant dissociation at the weakest chemical bond of a precursor ion, spectra are often dominated by a few intense fragment ions. In contrast, HCD spectra often exhibit a more complete ion-series increasing the probability of identification.[59, 60] In case of phosphopeptides, HCD spectra usually exhibit a reduced neutral loss of phosphoric acid (the weakest bond) rendering identification of the peptide including correct localization of phosphosites more likely.[59] However, with the advent of dual pressure linear ion traps (Figure 4), which are divided into a higher-pressure collision compartment and a lower-pressure scanning compartment, CID spectra annotation has improved considerably.[61]

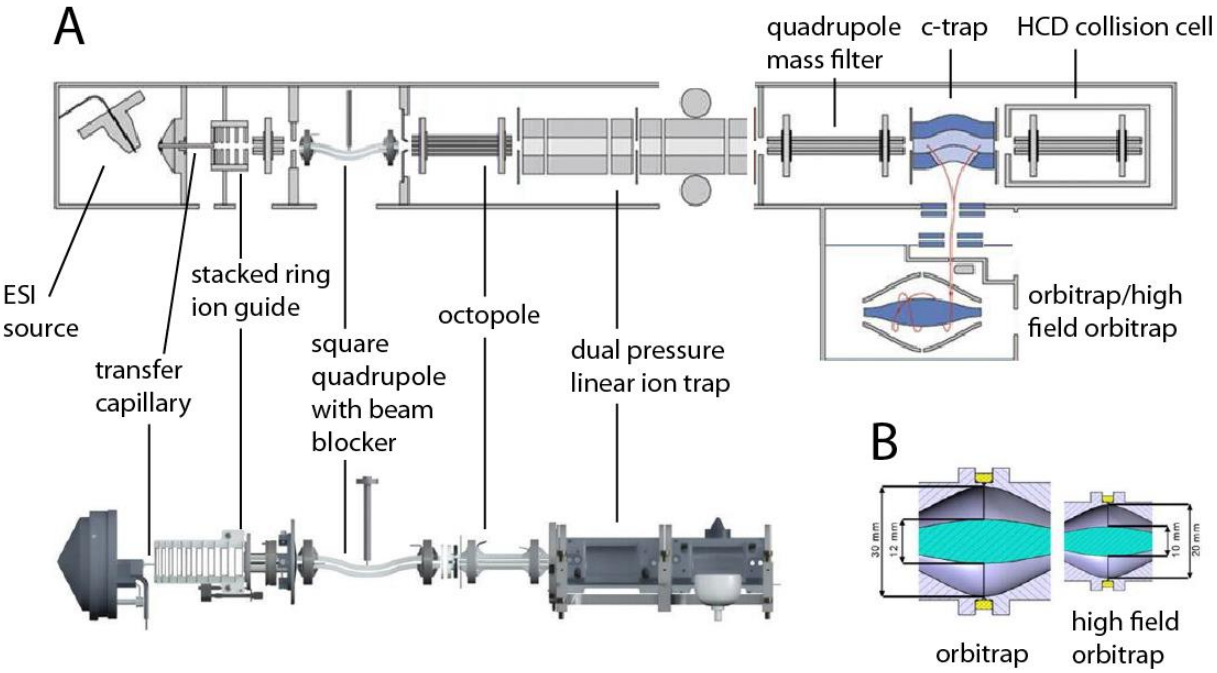


Figure 4: Schematic representation of a hybrid linear ion trap-Orbitrap mass spectrometer (LTQ Orbitrap Velos Pro, LTQ Orbitrap Elite). In DDA, the ion trap can be used to acquire multiple MS/MS in the dual pressure linear ion trap in parallel to the acquisition of a high resolution MS in the Orbitrap (A). Newer instruments (LTQ Orbitrap Elite, Q Exactive HF) are equipped with a high field Orbitrap, which exhibit higher scan rates at equal resolution (B). Figure from Michalski et al. with slight modifications.[54]

3.2.1.4 Quantification strategies in proteomics

Although the numbers of detected features in LC-MS are quite impressive, qualitative analysis of biological samples usually provides only little information about samples. On one hand, complete absence of proteins occurs rather seldom, whereas up- or down- regulation can be frequently observed (however, a downregulation below the LLOD is often misinterpreted as absence). On the other hand, peptides are identified in a probabilistic manner (the most intense are triggered for MS/MS), hence, a missing identification is never an evidence for absence. Hence, most proteomics experiments focus on comparison of protein abundance over different biological samples based on comparison of signal intensities – which is often described with the ambiguous or even misleading term *relative quantification.*

Notably, the IUPAC Gold Book (www.goldbook.iupac.org, August-2019) defines *quantitative analysis* as “Analyses in which the amount or concentration of an analyte may be determined (estimated) and expressed as a numerical value in appropriate units”. Consequently, a quantitative analysis requires previous determination of detector response or calibration curves as MS signal intensities do not necessary scale proportionally with abundance. As comparison of intensities across multiple samples does neither allow determination of absolute nor relative amounts, the term *semi-quantitative* appears more appropriate. In other words, results only give a rough estimation of up- or downregulation and the frequently used term *fold-change* (of a protein) simply refers to signal intensities but not to protein abundance.

Label-free analysis of samples: Several *semi-quantitative* strategies based on comparison of signal intensities have been developed over the past decades. Label-free approaches are straightforward strategies comprising parallel processing of all samples and subsequent analysis by LC-MS in several consecutive runs. Importantly, a very stable instrument set up is prerequisite and samples should be randomized to evade any systemic bias originating from variation in LC-MS performance (HPLC separation, sensitivity, mass accuracy). Additionally, multiple analyses of individual samples are beneficial to allow discrimination of technical variance from biological variance. Subsequent data analysis relies on MS level feature comparison, by extracting the peptide ion trace at the specific RT. If the precursor ion shows an appreciable signal intensity and has at least one identified MS/MS, the respective peptide is considered as identified and quantified (Figure 5). This so-called *label-free* quantification is regarded as comparably accurate in terms of relative changes, but prone to technical variance from sample preparation and LC-MS instrumentation. Another drawback is that multidimensional approaches (section 3.2.1.5) induce additional variance rendering label-free approaches more applicable with 1D-LC-MS. Notably, 1D-LC-MS is limited in terms of study depth as

peptide identification rates (ID / min) decrease with increasing gradient lengths rendering gradients above 4 h not remunerative .[62]

Peptide labelling for quantification on the MS level: Further increase of the depth is possible, when using peptide labelling approaches in conjunction with multidimensional LC-MS in form of pre-fractionation of samples to reduce the complexity (section 3.2.1.5). Peptide labelling allows to combine (*multiplex*) multiple samples in a single reaction tube, but relative signal intensities can still be determined for each peptide from each sample. The most basic approach is the introduction of different isotopes into the peptides of each sample. For instance di-methylation of primary amines with cyanoborohydrate and formaldehyde of different $^{13}C/^{12}C$ and $^{1}H/^{2}H$ isotope compositions results in mass shifts between +28 ($^{12}C_2$, $^{1}H_6$) and + 36 Da ($^{13}C_2$, $^{2}H_6$) allowing for differentiation and comparison of precursor intensities on the MS level.[63] Notably, the mass difference of differential labels should be ≥ 4 Da to reduce interference between the natural isotope pattern of peptide signals (attributed to the natural ^{13}C occurrence of 1.1 %). Analysis of multiplexed samples (usually up to 3 samples with di-methylation) considerably reduces technical variances of sample preparation and LC-MS instrumentation. However, as the labelling is conducted comparably late in the workflow, variance originating from cell/tissue lysis, protein extraction and digestion cannot be omitted.

Stable isotope labelling by amino acids in cell culture (SILAC): As a further refinement, *stable isotope labelling by amino acids in cell culture* (SILAC) allows for incorporation of isotopes into living cells by supplementing the unlabelled amino acids in the media by isotope-labelled analogues (heavy amino acids).[64] A wide variety of cells, including yeast, animal and human cells, can incorporate heavy labelled amino acids to a sufficient extend and ≥ 97 % is usually regarded as threshold for complete labelling. SILAC reduces technical variance to a minimum as multiplexing is possible at the very beginning of sample preparation, either after harvesting (by combining equal numbers of cells) or after cell lysis and protein concentration determination (by combining equal amounts of protein fractions).

Notably, SILAC is usually conducted with heavy lysine and arginine ensuring that all peptides in a tryptic digest exhibit at least one labelled amino acid (except for C-terminal peptides, which usually do not comprise arginine or lysine). For most cells, sufficient incorporation can be achieved within five passages and unspecific labelling can be widely omitted when adding proline in higher concentrations to suppress biochemical arginine-to-proline conversion. Hence, labelled cells can be usually generated within a few weeks of cultivation and allow multiplexing of up to three conditions when using arginine with +6 Da / +10 Da ($^{13}C_6$ / $^{13}C_6$, $^{15}N_4$) and lysine with +4 Da / +8 Da ($^{2}H_4$ and $^{13}C_6$,$^{15}N_2$). As with dimethyl labelling, quantification is based on comparison of MS precursor signal intensities (Figure 5). As a major drawback, cells supplemented with heavy amino acids might exhibit a slightly altered biochemistry, perhaps originating from the effect of isotope fractionation by organisms,[65, 66] rendering label-switch experiments absolutely mandatory.

Isobaric labels for MS/MS-based quantification: Experiments with multiplexing of up to eleven samples can be conducted when employing isobaric labeling approaches such as iTRAQ (isobaric tag for relative and absolute quantification) or TMT (tandem mass tag).[67, 68] Isobaric labels are usually introduced on the peptide levels by derivatization of primary amines after proteolytic digestion. Labels of all conditions exhibit the same nominal mass shift, but the positions of the isotopes are specifically altered allowing for the release of a reporter ion with a specific mass upon MS/MS. Relative intensities of reporter ions represent the abundance of each peptide in each condition allowing quantification over all conditions with a single fragment spectrum. In contrast, only one precursor signal can be observed on the MS level representing the sum of the signals from each sample (Figure 5). Isobaric labelling approaches have become widely used for all kind of experiments aiming at in-depth quantification of different conditions and allow coverage of nearly-full proteomes if samples are extensively pre-fractionated. Nevertheless, isobaric labeling approaches suffer from reduced accuracy in relative quantification especially for analytes with low signal intensity.[69, 70] Besides, quantification with ion trap

mass spectrometers is difficult since reporter ion signals are usually below the scan range of the MS/MS. These issues have not yet been resolved to a satisfying extend, though numerous attempts have been made in the past (MS^3 and multinotch-$MS^{69, 70}$ or PQD).[71, 72]

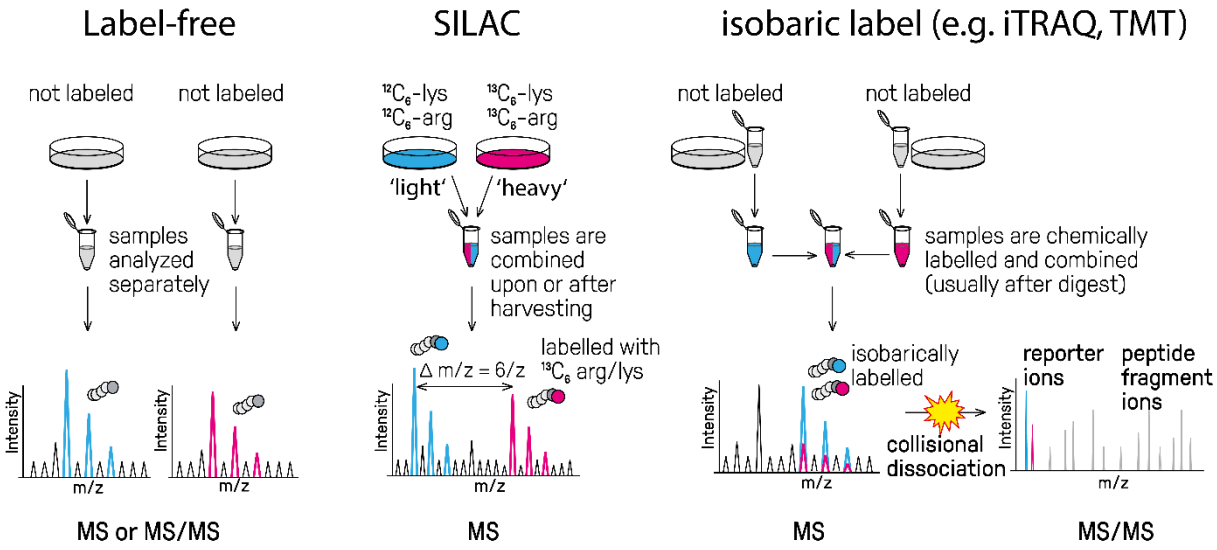


Figure 5: Quantification strategies in bottom-up proteomics. Label-free analysis allows for straightforward quantification of numerous samples, usually via 1D-LC-MS, but is comparably prone to variations from sample preparation and LC-MS instrumentation. These variations can be drastically reduced by incorporation of heavy labelled amino acids into cells (SILAC), which allows to combine samples directly upon harvesting. Isobaric labels, usually introduced after digestion, can be employed for multiplexing of up to 11 samples. However, the MS/MS-level quantification with such labels often suffers from low accuracy – especially for low abundant analytes.

Quantification using targeted LC-MS and stable isotope-labelled (SIL) peptides: In addition to large-scale approaches, which enable relative comparison of signal intensities of thousands to ten thousands of peptides, a comparably precise quantification is possible when using targeted methods.[73] Targeted LC-MS allows for monitoring of up to hundreds of analytes at their specific retention time and mass-to-charge (m/z) ratio by a series of consecutive MS/MS scans over the chromatographic peak.[74] Ion traces of individual fragments, which are termed *transitions*, are subsequently extracted from the MS/MS spectra and usually the best three to five interference-free signals are used for quantification. For targeted LC-MS, quantification accuracy improves tremendously when employing synthetic reference peptides of known concentration, which can be used for external calibration. However, with

the advent of isotopic labelled reference peptides,[75, 76] internal calibration became possible, which allows omitting of matrix effects in the sample and technical variation from LC-MS. Such approaches are considered as quite robust quantification methods with high accuracy and precision.[77]

3.2.1.5 Large-scale analysis using multidimensional workflows

The depth of proteomics studies can be considerably increased by extensive fractionation of a sample (*large-scale proteomics*), either online[78-81] or offline,[82-84] to reduce complexity in the final LC-MS run (Figure 6). Appropriate pre-fractionation techniques (first dimension) separate analytes by different mechanisms than the HPLC coupled to MS (second dimension) to preserve good separation for final LC-MS (*orthogonality*).[85] Usually, 6-24 fractions are sufficient with a well-optimized two-dimensional separation to reach an appreciable depth of analysis. To further increase the depth or enrich certain sub-proteomes, additional dimensions can be employed but at the expense of a tremendous increase in total LC-MS time.[84, 86]

Since tryptic peptides exhibit homogeneous physicochemical properties compared to protein fractions from biological samples, multidimensional peptide separation by HPLC has become a key-method for in-depth analysis of biological samples – also owing to the fact that peptide chromatography benefits from high peak capacity. Ion exchange chromatography (IEX),[87-89] hydrophilic interaction liquid chromatography (HILIC)[83] and electrostatic repulsion-hydrophilic interaction liquid chromatography (ERLIC)[84, 90] exhibit high orthogonality to low pH-ion pairing-RP, which usually serves as the last dimension, but suffer from comparably low peak capacity. Therefore, many studies are conducted with high pH-RP as pre-fraction step, which exhibits high peak capacity, even though the orthogonality is only moderate.[82, 91, 92] To overcome the limited orthogonality problem, the eluate is collected in short intervals and non-adjacent intervals are subsequently combined to yield final *concatenated fractions* for LC-MS analysis. For instance, when collecting eluate in 40 1-min intervals to yield 10 concatenated

fractions, the eluate from the RT intervals 0, 10, 20, 30 are combined in a first fraction, 1, 11, 21, 31 are combined in a second fraction and so on (see also section 5.4.3). Notably, the short RT intervals (usually 0.5-2 min) tremendously increase the probability that peptides distribute over multiple fractions, which, at least in theory, drastically reduces the chance of detection. Nevertheless, excellent separation efficiency and high (pseudo-) orthogonality render concatenated fractionation via high-pH-RP as excellent pre-fraction method, which has been employed in numerous in-depth proteomics[40, 93, 94] and phosphoproteomics studies.[95, 96]

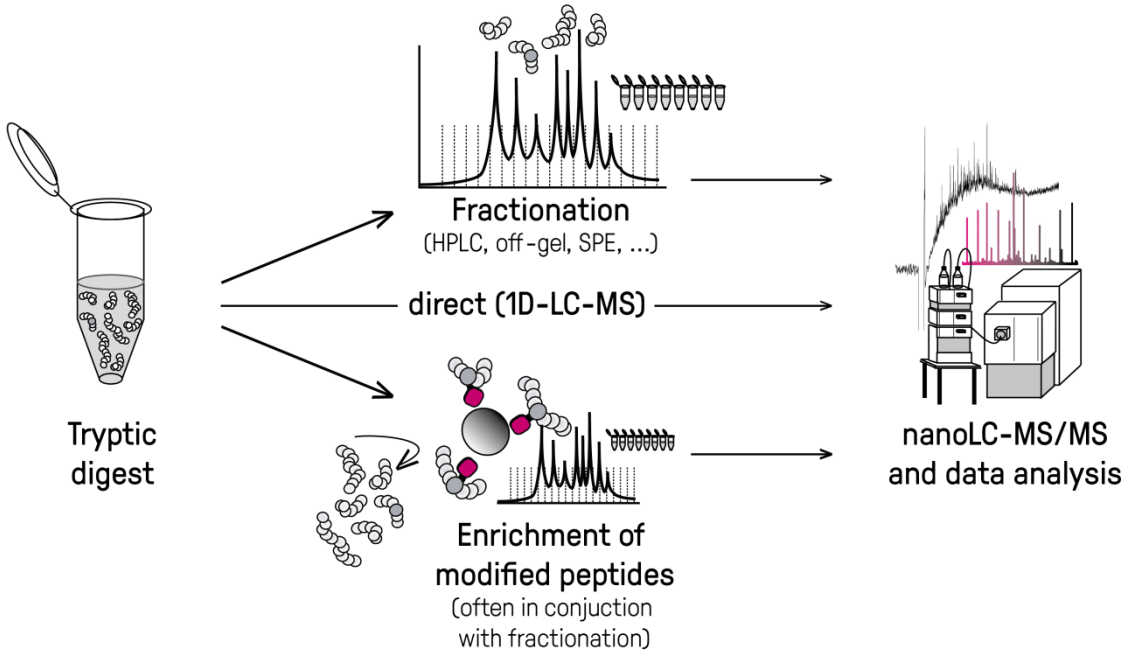


Figure 6: Analysis strategies in LC-MS-based bottom-up proteomics. After proteolytic digestion with trypsin, peptides can be analyzed either by direct LC-MS (1D) or subjected to fractionation to increase the depth of analysis (2D). For analysis of PTMs (e.g. phosphopeptides), enrichment methods need to be employed (often in conjunction with additional fractionation) to deplete the vast of non-modified peptides

Certain HPLC-based-fractionation methods also allow for isolation/enrichment of peptide subsets if physicochemical properties are appreciably different from the vast majority of peptides in the digest. Tryptic phosphorylated, N-acetylated and sialylated peptides, which exhibit a reduced net charge at lower pH (e.g. pH ≤ 2.7), can be enriched in the first fractions of strong cation exchange chromatography (SCX)[97-99] or in the later fractions of strong anion exchange chromatography (SAX).[86, 89, 100] However,

both methods exhibit drawbacks: (I) already minor salt contaminations in SCX can result in poor separation and (II) the specificity of SAX is usually not sufficient for an appreciable enrichment.[89] As phosphopeptides and glycopeptides exhibit higher hydrophilicity than non-modified ones, HILIC has been extensively tested for enrichment and fractionation,[101, 102] but appeared to lack specificity if not combined with additional enrichment steps.[103, 104]

3.3 Strategies for global phosphoproteomics

Phosphoproteomics workflows are usually conducted with comparably high amounts of starting material usually subjected to affinity enrichment (to deplete non-modified peptides) in conjunction with extensive fractionation. The next sections give an introduction into the key-techniques and strategies for phosphoproteomics.

3.3.1 Phosphopeptide enrichment techniques

Numerous phosphopeptide enrichment strategies have been introduced in the past decade. Most of them predominantly rely on affinity purification, which usually exploits the specific chemical features of the phosphate group or the phosphorylated amino acids. Especially the low isoelectric point (pI) and the high hydrophilicity are properties allowing for immobilization on a functionalized surface or by a tailored phosphate-specific antibody. Although acidic amino acids often exhibit similar affinities and acidic peptides are frequently found co-enriched, a variety of displacers can help to reduce these unspecific binders. However, the problem of co-enrichment renders optimization of binding and washing conditions especially critical to a successful purification. Consequently, establishing such protocols often takes years.[14, 25]

Nowadays, affinity purification is usually conducted via immunoprecipitation (using anti-pTyr or anti-kinase motive antibodies),[105, 106] immobilized metal-ion affinity chromatography (IMAC)[106, 107] or metal-

oxide affinity chromatography (MOAC).[108, 109] These techniques have been all well-refined over decades to allow for a detection of thousands to ten thousands analytes in complex samples (Figure 7).[25, 110 14]

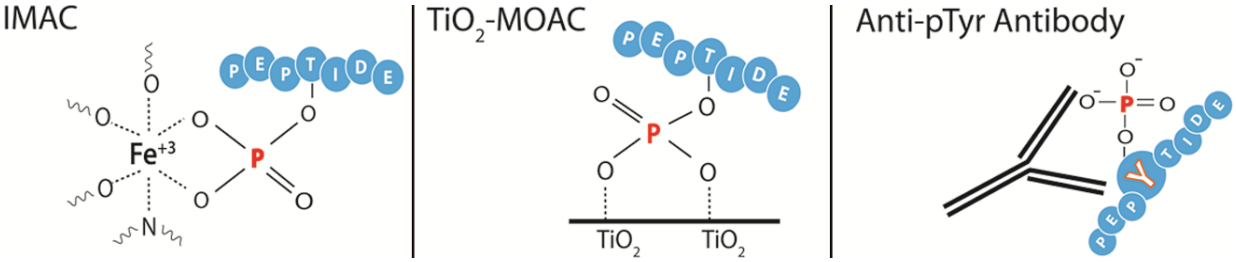


Figure 7: The high affinity of phosphate groups towards immobilized metal ions or metal oxides can be exploited for specific enrichment of phosphopeptides via IMAC or MOAC. In addition, specific antibodies can be employed for the enrichment of pTyr containing peptides. Figure from Grimsrud with slight modifications.[111]

IMAC is conducted with metal cations such as Ti^{4+}, Fe^{3+} or Ga^{3+} usually immobilized on a carrier matrix.[107, 112] Combination of these materials enables even more comprehensive phosphoproteomics as they exhibit complementarity in phosphopeptide enrichment. For instance, Fe^{3+}-IMAC shows high affinity towards basic phosphopeptides, whereas acidic ones can be enriched using Ga^{3+} rendering sequential enrichment strategies highly efficient.[107, 113] In the past years, Ti^{4+}-IMAC has been successfully used in conjunction with HPLC-based fractionation for in-depth phosphoproteomics[28] and the high specificity renders it promising for 1D-label-free quantification towards higher sample throughput.[114, 115] However, lifetime of the carrier matrix is usually limited (according to vendors) rendering robustness of the Ti^{4+}-IMAC method as problematic.

One of the gold-standards of phosphopeptide enrichment represents MOAC (conducted with metal oxides), which allows phosphate groups to arrange on the surface in form of a Helmholtz layer. TiO_2-MOAC using porous TiO_2-particles has been demonstrated as highly versatile for straightforward enrichment with high specificity[108, 116] and represents the central enrichment step in numerous large-scale studies.[95, 104, 117] As already demonstrated some years ago, a well-optimized TiO_2 purification can yield specificities of ≥ 98 % as further described in the next section (section 3.3.2).

3.3.2 15 years of phosphopeptide enrichment via TiO_2-affinity purification

TiO_2-affinity purification is nowadays a widely technique as it exhibits extraordinary sensitivity and specificity and is applicable for low sample amounts in the µg-range.[104] In previous studies by Pinkse et al., nanoHPLC trap columns were used to allow for an online-enrichment of phosphopeptides prior to RP separation[108] and combinations with SCX pre-fractionation allowed to further increase the depth of such analyses.[116] Approximately at the same time, Larsen et al. demonstrated that specificity of micro column-based offline enrichment can be drastically enhanced using 2,5-dihydroxybenzoic acid (DHB) or similar additives during loading, which compete with non-modified peptides for TiO_2 binding positions .[109] Later, the Larsen group demonstrated that glycolic acid is a well-suitable DHB substitute with higher compatibility to nanoLC-MS.[104] Notably, another considerable increase of specificity was achieved by adjusting the TiO_2 bead-to-sample ratio to ~6:1 (w/w), which limits the available surface to the phosphopeptides, whereas non-specific binders remain predominantly in the supernatant. However, as the resulting supernatant still comprises phosphopeptides, it can be subjected to additional rounds of enrichment with ~3:1 and afterwards ~1.5:1 bead-to-sample ratios. Mixing of the three TiO_2-bead fractions followed by stringent washing and basic elution at high pH (~ pH 11) yields a highly enriched phosphopeptide sample with usually ≥ 98 % phosphopeptide IDs upon 1D-LC-MS. The total number of IDs can be further increased by combination with IMAC (for separation of mono- and multiphosphorylated peptides) or fractionation via offline HILIC-,[104, 118] SCX-[119] or high-pH-RP pre-fractionation[95] (as already described in the previous sections 3.2.1.5 and 3.3.1).

In the past years, combinations of TiO_2-based enrichment with extensive fractionation and state-of-the-art LC-MS instrumentation enabled extremely comprehensive phosphoproteome studies resulting in the detection of ten thousands of different phosphorylation events in human cells (inferred from more than 50,000 different phosphopeptides). For instance, Sharma et al. employed SCX-TiO_2 and complementary

pTyr immunoprecipitation for detecting ~40,000 sites, inferred from ~50,000 tryptic phosphopeptides of more than 10,000 proteins[24] emphasizing that nearly every cellular protein might become phosphorylated under certain conditions – at least if the cells are stimulated. In a recent study by the Olsen group, more than 40,000 sites were detected in NIH-3T3 mouse fibroblasts (~7,000 phosphoproteins) using automated high-pH-RP fractionation prior to affinity enrichment.[39] Both studies demonstrate the extraordinary versatility of TiO_2-based phosphopeptide purification, even though such depths can only be achieved when employing milligrams of starting material. However, applicability in low sample amount-experiments has been successfully shown in numerous other studies such as from the Larsen group (Engholm-Keller et al.) with ~6,600 phosphopeptides from 300 µg of digest[104] or the Sickmann group (Beck et al.) with ~2,700 phosphopeptides from 4-plex iTRAQ samples with 200 µg of digest per channel.[120]

3.3.3 Phosphopeptide enrichment via electrostatic repulsion-hydrophilic interaction liquid chromatography (ERLIC)

In 2008, Andrew Alpert introduced the term electrostatic repulsion-hydrophilic interaction liquid chromatography (ERLIC) as mixed-mode chromatography, which employs ion exchange columns eluted with buffers of high organic solvent concentration as usually used for HILIC. The ion exchange column for ERLIC has the same charge as the target analyte, which would lead to early elution under aqueous conditions (electrostatic repulsion), but this mechanism is counteracted by the superimposed HILIC mechanism. Hence, the high percentage of organic solvent in the mobile phase (>> 40 % acetonitrile, ACN) results in retention of the analyte, even though it is repulsed from stationary phase (Figure 8). In fact, it has been demonstrated that negatively charged adenosine triphosphate can be highly retained on a negatively charged SCX column, if the ACN concentration of the eluent is ~75 %.[121]

For the selective enrichment of phosphopeptides, which exhibit a reduced net charge and a higher hydrophilicity at low pH (~2-3) than their non-modified counterparts, an ion exchange column can be eluted with a mobile phase of 70 % ACN (Figure 8).[121] When injecting tryptic digests, non-modified peptides predominantly elute within the first minutes. In contrast, phosphopeptides elute somewhat later over a time frame of 10-20 minutes allowing for parallel enrichment and fractionation with fairly separated singly and multiphosphorylated species. Even though some research groups employed mobile phases with volatile salts to overcome time consuming SPE of ERLIC fractions,[122, 123] Alpert et al. convincingly demonstrated that selectivity improves when using Na-methylphosphonate as additive. It reduces retention of acidic non-modified peptides, whereas phosphopeptide retention remains unaffected.[89] As a drawback, addition of Na-methylphosphonate renders a well-optimized SPE step mandatory, but this can be exploited for additional enrichment of phosphopeptides as shown by Loroch et al..[124] Notably, especially SCX solid-phase extraction has been demonstrated well-suitable for early fractions, as phosphopeptides predominantly elute in the flow-through (Loroch et al.),[84, 124] but also SCX-HPLC can be employed (as shown by Zarei et al.).[125]

In first studies about ERLIC-based phosphopeptide enrichment from complex biological samples, only a few hundred unique phosphopeptides were detected,[122, 126, 127] partially attributed to the use of not state-of-the-art LC-MS instrumentation. In 2013, Zarei et al. demonstrated that combinations of ERLIC-SPE, SCX-SPE and TiO_2 affinity purification can be used to identify nearly ten thousands phosphopeptides from a few milligrams of starting material,[128] but the workflow still relied on TiO_2 to enhance the specificity. In contrast, Loroch et al. demonstrated in 2015 that a well-optimized ERLIC-HPLC in conjunction with simple RP- or SCX-SPE is sufficient to identify ~7,500 phosphopeptides from as little as 100 µg non-stimulated HeLa[124] and that the workflow can be further modified to allow for a simultaneous analysis of the proteome [84] without requiring additional sample material. The depth of these ERLIC-based phosphoproteomics studies can be explained by the well-optimized HPLC system

with low injection and dead volume and the tailored combination with SCX-SPE to further enrich for phosphopeptides in the early fractions, where non-phosphorylated peptide co-elute. Hence, ERLIC ranges with TiO_2- and IMAC-based workflows among the most sensitive methods for global phosphoproteomics.

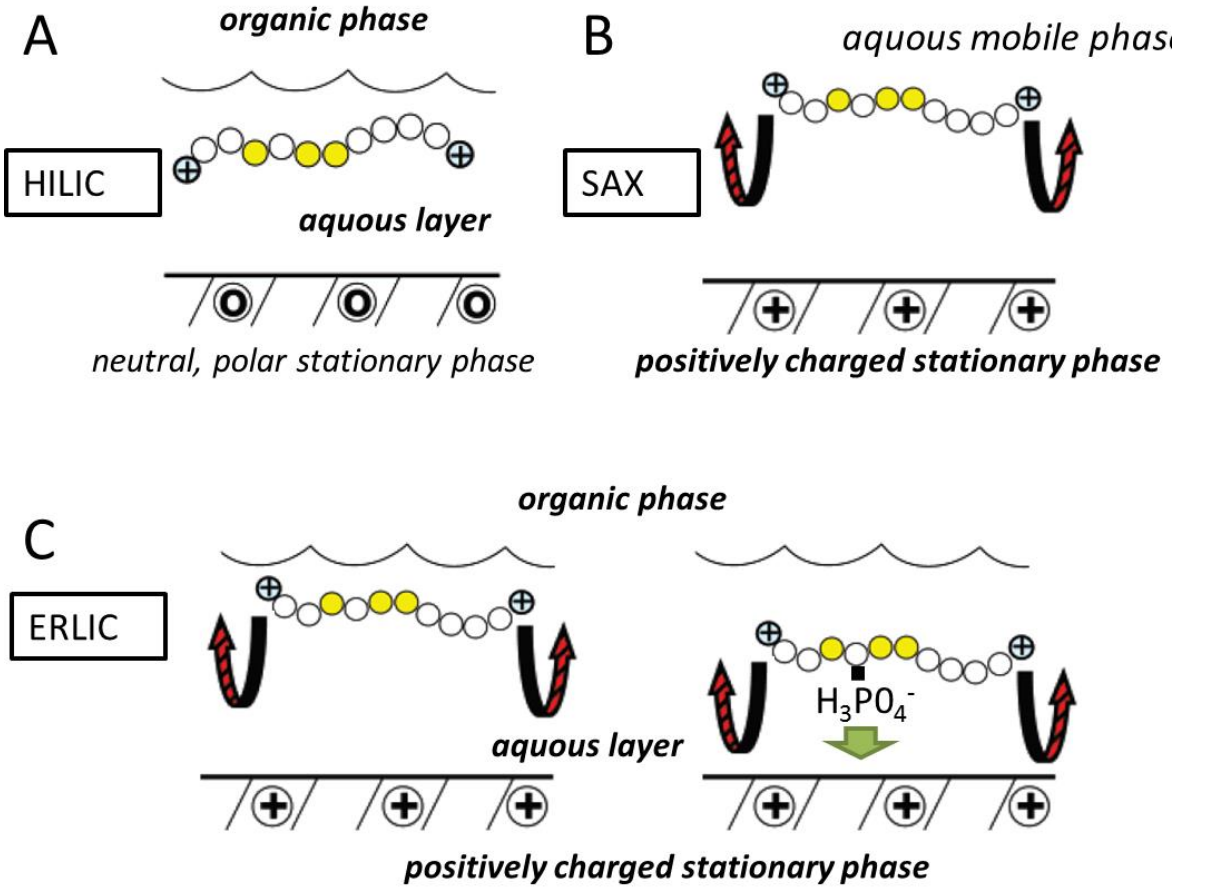


Figure 8: HILIC (A) and SAX (B) are the principal mechanisms of ERLIC (C). In HILIC, columns are eluted with high percentage of organic solvents leading to the formation of an aqueous layer along the stationary phase, which retains analytes by hydrophilicity (A). SAX at pH 2 would not retain a tryptic peptide due to electrostatic repulsion (B). However, if the SAX column is eluted with ~70 % ACN (superimposed HILIC mechanism), non-modified peptides are only slightly retained, whereas phosphopeptides are highly retained (higher hydrophilicity and lower net-charge). Hence, ERLIC can be exploited for enrichment of phosphopeptides. Figure from Alpert 2008 with modifications[121]

In summary, ERLIC allows for isolation of phosphopeptides by exploiting increased retention on an anion exchange column eluted with a predominantly organic mobile phase. ERLIC-based workflows enable parallel enrichment and fractionation of phosphopeptides and exhibit high sensitivity with specificities of up to 80% (% phosphopeptide IDs). In contrast to the very sensitive and specific TiO_2-MOAC, where unspecific binding is prevented by displacement with DHB or glycolic acid, which might attenuate phosphopeptide recovery, the principal mechanism of isolation is chromatographic retention. However, the high sensitivity of both methods, TiO_2 and ERLIC, renders their future promising for application in biomedical and clinical studies.

3.4 Herpesviridae – a family of opportunistic pathogens

Herpesviridae – the herpesvirus family of mammals, birds and reptiles – are highly disseminated in nature.[129] Several members infect humans as their primary host such as *human cytomegalovirus* (HCMV), *herpes simplex viruses* (HSV), *varicella-zoster virus* (VZV), *Epstein-Barr virus* (EBV), *Roseolovirus* (HHV-6A, -6B and -7) and *Kaposi's sarcoma-associated herpesvirus* (HHV-8). These herpesviruses developed an array of immune evasion mechanisms to escape killing by the host immune system and have the capacity to hide from it by establishing *latency.*[130] As opportunistic pathogens, herpesviruses persist lifelong in their hosts and take advantage of a weak immune system to switch from latency to a *lytic replication* cycle, which serves the production and release of infectious progeny *virions* to infect other cells and hosts. The accompanied destruction of the cell and spreading of virions is often connected to diseases, such as glandular fever or glandular fever-like symptoms (EBV, HCMV), herpes (HSV), chicken pocks and shingles (VZV) or roseola (HHV-6A and -6B). Whereas healthy individuals are usually less affected, herpesvirus infection can be life-threatening for immunocompromised individuals.[131]

3.4.1 The human cytomegalovirus (HCMV)

HCMV is worldwide distributed among the human population. Depending on the country, the HCMV seroprevalence ranges from 40 to nearly 100 % (especially high in South America, Africa and Asia). Depending on the health state, symptoms range from slight febrile illness or mononucleosis-like symptoms to severe inflammation of various tissues including retina, heart and colon. Most importantly, HCMV is a leading death cause for transplantation patients and HIV-infected individuals.[132] Furthermore, congenital HCMV infection is associated with severe neurological impairments ranging from hearing loss to infant death. In fact, it is the most common congenital virus infection known so far, affecting 20,000-30,000 infants each year only in the United States.[133]

Primary infection occurs mostly at mucosal sites if the virus overcomes the innate immune barriers and sustains replication. Subsequent dissemination to salivary glands and kidneys usually result in systemic infection.[129] The primary infection is mostly asymptomatic in healthy individuals, even though the virus is readily detectable in body fluids. Once infected, HCMV persists lifelong within the host in a latent form and only little is known about effects of the latent state on pathologies such as cancer[134, 135] or vascular diseases.[136] Re-activation can cause disease pathogenesis, especially in immunocompromised individuals, which is predominantly attributed to virus replication, the accompanied cell death upon virion release and the spreading.

3.4.2 Structure and biological cycle of HCMV

The structure and biological cycle of HCMV is characteristic for most herpesviruses. Virion particles are about 200-230 nm and harbour a comparably large double-stranded DNA (dsDNA) genome of ~230 kb inside the 130 nm nucleocapsid.[129] The biological cycle on the cellular level can be divided in three phases: infection, latency and lytic replication. Infection starts with virion entry, mediated by a variety of tissue-specific surface receptors (glycoprotein complexes), via direct fusion with the plasma

membrane or by endocytosis. Subsequently, the nucleocapsid hitch rides the microtubules for active transport to the nuclear pores and the DNA is translocated into the nucleus.[137] The biological decision between establishing latency or lytic replication is predominantly based on the chromatinization and the availability of transcription regulators. In case of latency, the viral DNA forms a closed circular molecule from which only a small subset of genes are active whereas genes for lytic replication are suppressed. The genome retains the ability to switch to the lytic replication cycle, which involves activation of viral genes, genome replication, virion assembly and virion release. This requires a ruthless exploitation of cellular resources and a parallel suppression of host defence mechanisms. Initially, *immediate-early (IE)* genes are activated and the respective gene products mediate modulation of the cell cycle, chromatin organization, RNA transport, splicing and adaptive immune response. IE proteins activate expression of *early* (*E*) proteins, which are predominantly involved in virus genome replication to several thousand copies from a single starting template. These genes modulate the nucleotide metabolism and formation of the replication machinery, which involves a set of six well-conserved viral proteins: a DNA polymerase including its processivity subunit, a single-stranded DNA binding protein and three helicase-primase components.[138] *Leaky late (LL)* and *true late (TL)* proteins are predominantly structural virion components. LL expression is augmented during DNA synthesis, whereas TL expression completely depends on viral DNA synthesis. Virion assembly starts in the nucleus where nascent capsids are filled with newly replicated DNA.[139] After egress of these *nucleocapsids* including some associated tegument proteins (inner tegument), the final complement of tegument proteins and surrounding envelope proteins are added in the cytoplasm. This process involves extensive exploitation of cellular systems for synthesis, processing and trafficking of proteins and vesicular transport systems to deliver the required virion subcomponents. Underlying processes are not completely understood, but tegument proteins appear to interact with each other to form subassemblies that associate with the virion during transport to the exit organelles. Depending on the herpesvirus, different cellular structures are

exploited as exit organelles for transportation to the plasma membrane. For instance, HCMV exploits the recycling endosomes as exit organelles to release the newly synthesized infectious virions.[129]

3.4.3 Murine cytomegalovirus – a valuable mouse model

Cytomegaloviruses (CMVs) are highly species-specific and replicate exclusively in cells of their own or closely related host species. For instance, HCMV replicates only in cells from humans or chimpanzees, hence, small animals cannot be employed for studying infection mechanisms.[140] However, the murine CMV (MCMV) serves as valuable mouse model to study pathogenesis *in vivo* or *in vitro* – using mice or mouse cell lines – as numerous proteins have been found to be structural or functional homologues. Even though HCMV and MCMV differ in their dependence on the cell cycle stage (HCMV's initial gene expression requires stage G1, whereas MCMV genes appear independent of the cell cycle),[141] the exploitation of host resources and suppression of host defence mechanisms including immune evasion appears widely conserved.[142, 143]

Traditionally, HCMV and MCMV proteins are not only classified according to their expression profiles, but also to their onset of expression relative to DNA synthesis (section 3.4.2). The major IE (MIE) genes encode for transcriptional regulatory proteins, such as the well-described splice variants IE1/IE3 of MCMV and their homologues IE1/IE2 of HCMV, which are required for activation of the early genes. MCMV IE3 and HCMV IE2 are the major switches for subsequent gene expression whereas IE1 – differing by a single exon, is of diverse functions – such as binding of chromatin, enhancing activation by IE2, disruption of *promyelocytic leukemia protein nuclear bodies* to undermine activation of interferon and orchestration of transcription signalling. In addition, re-activation from latency is under control of the MIE genes[144], which represents a central research topic in the field of virology.[145] Notably, a recent study provides evidence for additional IE3-similar splice variants of yet unknown function.[146]

The early gene promotor controls expression of numerous proteins involved in initiation of viral DNA replication, including M112/M113 in MCMV and the homologues UL112/UL113 in HCMV.[147] Their gene products (E1), comprising at least four different splice variants, are expressed shortly after the IE proteins suggesting that a few copies of IE proteins are already sufficient for activating the respective E promotor.[148] However, MIE and E1 proteins are supposed to serves the orchestration of the viral DNA replication machinery comprising the DNA polymerase UL54[149] and associated proteins UL44, UL57, UL70, UL102 and UL105 (M54, M44, M57, M70 and M102 in MCMV, respectively).[150, 151] As E1 proteins bind DNA and accumulate in viral replication compartments, a central role in DNA replication can be assumed (Loroch, Schommartz et al.).[147] With ongoing viral DNA replication, the so-called *late proteins* reach their expression maximum. The group of late proteins comprises structural proteins or proteins involved in assembly, maturation and virion egress. The respective genes are under control of M92 and M79 in MCMV, which are homologue to UL92 and UL79 in HCMV, and localize to viral replication compartments during infection.[152] A recent study identified these two proteins to further interact with M49, M87, M95 and the host's RNA polymerase II to form a multi-component transcription complex.[153] Notably, exact regulation and involvement of viral transactivating factors is still a topic of ongoing research.[154]

The traditional classification of viral proteins relies on biological assays, which usually employ detection on the transcript level relative to cellular events. For instance, the *cycloheximide-actinomycin D chase* allows differentiation of IE and E genes as cycloheximide blocks viral protein synthesis and does not allow expression of E proteins.[155] Similarly, inhibition of viral DNA replication allows identification of TL genes,[156] as their expression requires viral DNA replication. In contrast to this classification, a recent large-scale proteomics study by Weekes et al. reported that K-Means clustering of protein expression profiles during the course of HCMV-infection in fibroblasts suggests rather 5 to 7 classes.[94] This appears likely when considering that protein expression depends on both transcriptional and translational

control. For instance, *m142* and *m143* transcripts follow IE kinetics but the proteins follow an E kinetic.[157] The presence of more classes appears even more likely when considering that virus replication involves numerous processes of host-cell manipulation such as suppression of cell death, hijacking of transcription and replication machinery and exploitation of the endoplasmic reticulum (ER) for protein processing and vesicular transport. The next sections (3.4.4 and 3.4.5) give two examples of host cell exploitations, which further highlight the complexity of host-pathogen interaction.

3.4.4 Suppression of host defence mechanisms

Host-defence mechanisms are widely suppressed by CMV as they aim at attenuating or abolishing viral replication. For instance, the *programmed cell death* (PCD) is a controlled cellular mechanism, which can be initiated upon cellular stress and usually serves the protection of a multicellular organism. This is meaningful as cell suicide is more favorable for a multicellular organism than viral replication and subsequent spreading to other cells. However, CMV acquired mechanisms to suppress this host defense.[158] A principal form of PCD is *apoptosis*, which is usually not accompanied by an inflammatory response since the resulting apoptotic bodies, comprising cellular constituents, are degraded by phagocytosis. Apoptosis can be triggered via an extrinsic pathway, which is induced by stress signal recognition via surface receptors, leading to activation of caspase-8, a trigger for mitochondria-mediated apoptosis through cleavage of the protein BH3-interacting domain death agonist (Bid).[159] The subsequent *mitochondrial outer membrane permeabilization* (MOMP) leads to a release of pro-apoptotic factors, which finally trigger apoptosis.[160] Alternatively, an intrinsic pathway can be triggered via intracellular stress signals (such as oxidative stress, DNA damage, unfolded protein response (UPR)) or a sudden increase in cytosolic Ca^{2+}. This leads to an accumulation of the proteins Bcl-2 homologous antagonist (Bak) and apoptosis regulator BAX (Bax) resulting in MOMP and induction of caspase signaling.[161, 162] To suppress apoptosis, the MCMV protein M36 and the HCMV protein UL36 can inhibit Caspase-8 mediated signaling.[163, 164] Besides, the MCMV proteins m38.5 and m41.1 bind

cytosolic Bax and Bak leading to an inhibition of MOMP. Similarly, the HCMV protein UL37x1 inhibits mitochondria-mediated apoptosis and appears to be a functional analogue to m38.5 emphasizing high functional homology of apoptosis suppression in HCMV and MCMV.[158, 165]

Necroptosis, is a caspase-independent PCD mechanism triggered via stress signals such as alkylation-induced DNA damage[166] or ligation of surface receptors including the tumor necrosis factor receptor 1 (TNFR1). Upon TNFR1 ligation numerous proteins such as TNFR1-associated DEATH domain protein (TRADD), FAS-associated death domain protein (FADD) and receptor-interacting serine/threonine-protein kinase 1 (RIP1) are recruited into a protein complex, which subsequently induce caspase 8-mediated apoptosis. Alternatively, if this pathway is blocked, necroptosis can be executed via RIP3-mediated activation of the mixed-lineage kinase domain-like protein.[167] There is strong evidence that complex formation of RIP1 and RIP3 is attenuated by the MCMV protein M45.[168] Other studies have unveiled additional effects of M45 on RIP3 signaling such as inhibition of DAI-induced nuclear factor NF-kappa-B (Nf-κB) activation, a sensor for DNA/RNA.[169, 170] Notably, no M45 homologue has been identified in HCMV, but the UL39-encoded proteins of herpes simplex virus (HSV) 1 and 2 exhibit similar inhibitory function towards RIP3[171] indicating countermeasures for host-defense well-conserved across herpes virus-species.

3.4.5 Exploitation of the host cell machinery towards efficient virus replication

Successful CMV replication relies on systematic exploitation of the host DNA-replication, translation and transcription machinery for the synthesis of viral proteins. Major instances for controlling such cellular functions are cyclins in conjunction with their related cyclin-dependent kinases (CDK) which are strongly connected to the induction of CMV gene expression. It has been shown already two decades ago that HCMV-infection induces upregulation of the cyclins A, B, E and related kinases Cdk1 and

Cdk2 resulting in the typical cell cycle arrest upon infection.[172, 173] Accordingly, selective inhibition of CDKs results in shutdown of early viral gene expression or, with CDK inhibition at later time points of infection, in substantial virus titer reduction.[174] Most interestingly, CDK inhibition during latency has an opposite effect since it relieves the block of IE expression towards reactivation.[175] The complex role of cyclins and CDKs for orchestration of viral reactivation and replication render these proteins potential candidates for functional studies and as potential drug targets.

The extensive viral protein synthesis upon replication is connected to considerable stress in the endoplasmic reticulum (ER). Under normal conditions, ER-stress triggers the UPR[176] resulting in: (I) degradation of polypeptides in the 26S-proteasome via ubiquitination, (II) proline-rich receptor-like protein kinase- (PERK-) mediated phosphorylation of the transcription factor eIF2 to shutdown protein synthesis and (III) chaperone synthesis via activation of the transcription factors ATF-4 (mediated by PERK), ATF6 and the kinase IRE1 to enhance protein folding. However, a number of studies, conducted with HCMV and partially with MCMV, indicate extensive manipulation within the three UPR branches. Ubiquitin ligases have been described to undergo substantial de-regulation during infection[94] and, surprisingly, the proteasome activity appears to increase during the course of HCMV infection. In fact, treatment of HCMV-infected cells with proteasome inhibitors results in attenuated viral protein synthesis.[177] Notably, this phenomenon might not be conserved among CMV species since replication of HSV-1 and HSV-2, is not affected by proteasome inhibition.[178] HCMV-infection has further been reported to activate PERK, but the expected extensive eIF2 phosphorylation cannot be observed.[179] Consequently, translation can continue unaffected to further exploit the host cell's resources for efficient viral protein synthesis. Most interestingly, in contrast to normal conditions, PERK activation does not result in the expected shutdown of synthesis, but ATF4-mediated transcription remains widely active. Since ATF-4-dependent genes are highly involved in metabolism and redox regulation, activation might help to reduce the ER-stress and maintain extensive viral protein synthesis.[180] Notably, ATF6 is

suppressed under HCMV infection whereas ATF6 dependent genes are active indicating an alternative mechanism for chaperone regulation.[179] UL38 of HCMV has been identified as a central modulator in this pathway, but it is unknown if M38 of MCMV exhibits similar function.

3.4.6 Proteomics and phosphoprotomics to study CMV infection

The worldwide distribution, the absence of suitable vaccine[181-183] and the growing resistance to antiviral drugs[184] render extensive research on HCMV infection a central task of virology and immunology. As infection mechanisms have been well-conserved on multiple levels, MCMV represents an adequate mouse model for HCMV (sections 3.4.3, 3.4.4 and 3.4.5). However, on the level of molecular mechanisms, several differences can be observed emphasizing that translation of findings requires comprehensive knowledge about the exact similarities and differences. Such insights can be provided by large-scale quantitative proteomics and phosphoproteomics, which allow monitoring of a substantial fraction of both viral and host proteins (as shown for instance by Loroch et al.[84] or Weekes et al.[94]) and systematic comparison of data from different CMVs. As previous studies in the field of virology rather focussed on limited numbers of host proteins, connection with state-of-the-art omics technologies has the potential to unravel host-pathogen interaction on a system-wide level.[84, 94, 185] The resulting comprehensive pictures on reprogramming of the host cell might help to identify antiviral drug targets[186] and antigens suitable for vaccination.[187-189] Hence, these high-content screening technologies represent promising starting points for subsequent functional and immunological studies.

4. Aims & Objectives

Sensitive phosphoproteomics techniques are a prerequisite for studies with limited sample material. If experiments are conducted with animal or patient tissue, highly sensitivity workflows for deep coverage of the proteome and phosphoproteome are usually favourable to enhance the likelihood of detecting signalling mechanisms even if mediated by low abundant proteins. Experiments conducted in this work aimed at a characterization and an improvement of the quantitative phosphoproteomics toolbox, with the focus on ERLIC as central enrichment and fractionation techniques. The first part aims at a detailed characterization of ERLIC for phosphopeptide enrichment and whether employing the alternative protease subtilisin improves the outcome of a study in terms of phosphoproteome coverage (experiment 1-1). In a next step, experiments for systematic comparison against TiO_2-MOAC, the gold-standard for enrichment, were designed to allow for addressing efficiency in terms of phosphopeptide recovery - an important benchmark for evaluating applicability in quantitative phosphoproteomics with limited sample material (experiments 1-2 and 1-3). These preliminary experiments were conducted with aliquots of ordinary HeLa and A431 digests, which can be generated in vast amounts in standard cell culture labs and are frequently used as model sample for benchmarking in proteomics method development. A final methodological experiment, aimed at the comparison of both enrichment techniques for reliable detection of differentially regulated phosphopeptides using minute amount of a biomedical sample, namely 25 µg of cytomegalovirus-infected mouse fibroblast cells against a non-infected control (experiment 1-4). This experiment was conducted in technical triplicates with aliquots from a single biological sample to ensure that the comparison is not compromised by biological variation.

The subsequent part aimed at evaluation of ERLIC's applicability in clinical research with limited sample material, namely 32 µg of colon tissue digest (experiment 1-5). In this experiment, ERLIC was employed for purification of a specific phosphopeptide to infer the phosphorylation site occupancy (% phosphorylation) of Ser125 on EGLN-1, a key-player in the hypoxia-response pathway, in healthy

and cancerous colon tissue. The experiment was conducted with 20 samples from 10 patients, which allowed addressing robustness of the ERLIC system over multiple injections of tissue samples. A subsequent targeted LC-MS assay was designed for quantification of the phosphopeptide and its non-modified counterpart in the presence of synthetic SIL reference peptides as internal standards. This experimental design allowed to demonstrate the use of ERLIC in conjunction with targeted LC-MS for inferring phosphorylation stoichiometries of cancer-relevant pathways in clinical samples.

The second part of the work aimed at application of ERLIC for a combined time-resolved global characterization of the proteome and phosphoproteome of MCMV-infected mouse fibroblasts (experiment 2). From the technical point, the experiment was supposed to demonstrate that ERLIC does not only allow for phosphopeptide enrichment, but that non-modified peptides can be recovered for a global 2D-proteome analysis in parallel without the need for additional sample material. The experiment was conducted with minute amount of infected fibroblasts cells, namely 120 µg of protein per condition, to highlight the future potential for experiments with limited sample material such as animal or patient tissue. The MCMV sample was chosen as it serves as important animal model for HCMV, which has never been extensively characterized on the proteome and phosphoproteome level. The study was supposed to unveil time courses of protein abundances and protein phosphorylation dynamics of both viral and host proteins. Furthermore, the proteomics dataset was generated to allow for systematic comparison with other datasets such as Weekes et al.[94], comprising comprehensive quantitative proteomics data over the time course of HCMV-infection in fibroblast cells (NIH-3T3). In addition, the phosphoproteomics dataset was used for a comparison with the dataset of Oberstein et al., which reports UL97-dependent phosphorylation events in HCMV-infected lung fibroblast cells (MRC5).[190] This extensive data analysis aimed at deciphering similarities and differences of protein regulation patterns to evaluate MCMV's applicability as mouse model for HCMV.

5. Materials & Methods

5.1. Instrumentation, Chemicals, Consumables

The following tables give a complete overview of instrumentation, software, chromatographic systems, chemicals (including buffers) and consumables required for conducting the experiments described in this work. GHS hazard and precautionary statements are in the appendix (section 10.5).

Table 1: Instrumentation

Product	Vendor	Contact
Cell culture dishes 10 cm and flaks (75, 175 cm^2)	VWR	Hannover, Germany
Centrifuge 5424, 5417R, 5415R	Eppendorf	Hamburg, Germany
Cool trap – RVT4104 (for vacuum centrifuge)	Thermo Scientific	Bremen, Germany
Falcon tubes, 15 mL and 50 mL	VWR	Hannover, Germany
Incubator	INB Memmert	Schwabach, Germany
LoBind reaction tubes 0.5 mL, 1.5 mL, 2.0 mL and LoBind mikrotiterplates, 96-well	Eppendorf	Hamburg, Germany
LTQ-OrbitrapVelos Pro, LTQ-Orbitrap Elite, Q Exactive HF	Thermo Scientific	Bremen, Germany
Online degasser	Knauer	Berlin, Germany
Pipettes 0.1-2.5 µl, 0.5-10 µl, 10-100 µl, 20-200 µl, 50-1000 µl	Eppendorf	Hamburg, Germany
Pipette tips	Eppendorf	Hamburg, Germany
Reaction tubes 0.5 mL, 1.5 mL, 2.0 mL	Eppendorf	Hamburg, Germany
Sterile filters 0.22 µm	Thermo Scientific	Hamburg, Germany
Thermomixer, comfort	Eppendorf	Hamburg, Germany
Ultimate HPLC system	Thermo Scientific (former LC-Packings)	Bremen, Germany
Ultimate 3000 HPLC systems	Thermo Scientific	Bremen, Germany
Ultimate 3000 Rapid Separation Liquid Chromatography (RSLC) systems	Thermo Scientific	Bremen, Germany
Ultrasonic bath – RK52	Bandelin Sonorex	Berlin, Germany
Ultrasonic tip – Vibra-Cell – 75022	Sonics via VWR	Hannover, Germany
Vacuum centrifuge – Savant SPD121P	Thermo Scientific	Bremen, Germany
Vacuum manifold (for SPE)	Agilent	Waldbronn, Germany
Vacuum pump RC6 (for vacuum centrifuge)	Vacuubrand	Wertheim, Germany

Table 2: Stationary phases for HPLC and solid-phase extraction products

Product	Vendor	Contact
Acclaim, PepMap (100 µm x 2 cm, particle size 5 µm, pore size 100 Å)	Thermo Scientific	Dreieich, Germany
Acclaim, PepMap, (75 µm ID x 15 cm / 25 cm / 50 cm, 2 µm particle size, pore size 100 Å)	Thermo Scientific	Bremen, Germany
Bond Elut C18 SPE cartridges, 1 g	Agilent	Waldbronn, Germany
PolyWAX (4.6 x 100 mm, particle size 5 µm, pore size 300 Å)	PolyLC	Columbia, USA
SPEC C18 AR SPE cartridges, 4 mg / 15mg or 30mg	Agilent	Waldbronn, Germany
TSK gelamide 80, 150 µm x 15 cm, 5 µm, self-packed	Tosoh Bioscience	King of Prussia, PA, USA
Biobasic C18, 0.5 mm x 15 cm	Thermo Scientific	Bremen, Germany
ACE C18 column 75 µm x 25 cm, 3 µm, 100 Å, self-packed	Hichrom	Leicestershire, UK
PepSwift monolithic trap column, 200 µm × 5 mm	Thermo Scientific	Bremen, Germany
PepSwift monolithic capillary column 200 µm × 5 cm	Thermo Scientific	Bremen, Germany
Hypersep RP (C18), 10-200 µL spin tips	Thermo Scientific	Bremen, Germany
Hypersep POROS SCX, 10-200 µL spin tips	Thermo Scientific	Bremen, Germany

Table 3: Chemicals; GHS hazard and precautionary statements are listed in the appendix (section 10.5)

Product	Vendor	Contact
Acetone	Merck	Darmstadt, Germany
Acetonitrile, ACN, ULC/MS	Biosolve	Valkenswaard, Netherlands
Ammonium hydroxide, NH_4OH	Sigma Aldrich	Steinheim, Germany
$^{13}C_6$-arginine	Sigma Aldrich	Steinheim, Germany
$^{13}C_6$-$^{15}N_4$-arginine	Sigma Aldrich	Steinheim, Germany
Arginine	Sigma Aldrich	Steinheim, Germany
Formic acid, FA	Biosolve	Valkenswaard, Netherlands
Ammonium acetate	Fluka	Steinheim, Germany
Ammonium bicarbonate, ABC	Sigma Aldrich	Steinheim, Germany
Benzonase nuclease (25 units/µL)	Merck	Darmstadt, Germany
BCA protein assay kit (Pierce)	Thermo Scientific	Bremen, Germany
Calcium chloride, $CaCl_2$	Merck	Darmstadt, Germany
Complete mini (protease inhibitors)	Roche	Mannheim, Germany
D-MEM high glucose (SILAC medium)	Thermo Scientific	Bremen, Germany
D-MEM low glucose	Thermo Scientific	Bremen, Germany
Dithiothreitol, DTT	Roche	Mannheim, Germany
Ethanol, EtOH	Merck	Darmstadt, Germany
Fetal calf serum (FCS) for HeLa S3	Thermo Scientific	Bremen, Germany
Fetal calf serum (FCS) for NIH-3T3	GE Healthcare	Freiburg, Germany
Fetal calf serum (FCS) – dialyzed (for SILAC medium)	Thermo Scientific	Bremen, Germany
Glutamine	Thermo Scientific	Bremen, Germany
Glycolic acid	Sigma Aldrich	Steinheim, Germany
Guanidinium hydrochlorid, GuHCl	Merck	Darmstadt, Germany
Iodoacetamide, IAA	Sigma Aldrich	Steinheim, Germany
$^{2}H_4$-lysine (SILAC)	Sigma Aldrich	Steinheim, Germany
$^{13}C_6$-$^{15}N_2$-lysine (SILAC)	Sigma Aldrich	Steinheim, Germany
Lysine	Sigma Aldrich	Steinheim, Germany
Methylphosphonic acid	Sigma Aldrich	Steinheim, Germany
Phosphate-buffered saline, PBS	VWR	Hannover, Germany
Phosphoric acid, H_3PO_4 (HPLC grade)	Sigma Aldrich	Steinheim, Germany
PhosSTOP (phosphatase inhibitors)	Roche	Mannheim, Germany
Potassium chloride, KCl (SupraPur)	Merck	Darmstadt, Germany
Potassium dihydrogen phosphate, KH_2PO_4 (SupraPur)	Merck	Darmstadt, Germany
Proline	Thermo Scientific	Bremen, Germany
Sodium dodecyl sulfate, SDS	Roth	Karlsruhe, Germany
Subtilisin P5380	Sigma Aldrich	Steinheim, Germany
Sodium hydroxide, NaOH (HPLC grade)	Sigma Aldrich	Steinheim, Germany
TiO_2, Titansphere, 5 µm particles	GL-Science	Eindhoven, Netherlands
Triethylamine, TEA	Sigma Aldrich	Steinheim, Germany
Trifluoroacetic acid, TFA	Biosolve	Valkenswaard, Netherlands
Trypsin T1426 (from bovine pancreas)	Sigma Aldrich	Steinheim, Germany
Trypsin (sequencing grade, modified)	Promega	Madison, USA
Trypsin 0.025% with 0.02 % EDTA (cell culture)	Sigma Aldrich	Steinheim, Germany

Table 4: Software and programs

Product	Vendor/Developer	Contact
Chromeleon v6.8	Thermo Scientific	Bremen, Germany
Mascot v2.4	Matrix science	London, UK
MSOffice 2010	Microsoft	USA
phosphoRS v3.1	Mechtler-Lab, IMP	Vienna, Austria
Proteome discoverer (PD) v1.3, v1.4 and v2.2	Thermo Scientific	Bremen, Germany
R v3.2.0 “Full of Ingredients” und R v3.5.2 “Great Truth”	R Foundation for Statistical Computing	Vienna, Austria
Sequest implementation in Proteome discoverer v1.3 and v1.4	Yates-Lab, TSRI	California, USA
Xcalibur v2.0	Thermo Scientific	Bremen, Germany

Table 5: Buffers for cell culture

Buffer name	Composition
Medium	D-MEM low glucose, 10 % FBS, 4 mM glutamine, 100 U/mL penicillin and 100 µg/mL streptomycin
SILAC-Medium	D-MEM high glucose, 10 % dialyzed FBS, 100 U/mL penicillin and 100 µg/mL streptomycin, 584 mg/L glutamine, 200 mg/L proline, 143 mg/L lysine (differentially labelled), 84 mg/L arginine (differentially labelled), medium was pushed through a 0.22 µm sterile filter
PBS	137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4
SDS lysis buffer	1 % SDS, 50 mM Tris-Cl pH 7.6, 150 mM NaCl, 1 tablet Complete mini and 1 tablet PhosSTOP per 10 mL
Trypsin-EDTA solution	0.05% Trypsin, 0.02% ethylenediaminetetraacetic acid (EDTA)

Table 6: Buffers for solid-phase extraction (unless otherwise stated). A: Loading buffer, B: Elution buffer

Buffer name	Composition
Bond Elute RP A	0.5 % FA
Bond Elute RP B	60 % ACN
SPEC C18 RP A	0.5 % FA
SPEC C18 RP B	60 % ACN
Hypersep-RP A	0.1 % TFA
Hypersep-RP B	60 % ACN
Hypersep-SCX A	25 mM KH_2PO_4, 20 % ACN, pH 2.7 (using H_3PO_4)
Hypersep-SCX B	25 mM KH_2PO_4, 600 mM KCl, 20 % ACN, pH 2.7 (using H_3PO_4)
Oligo-R3-RP A	0.1 % TFA
Oligo-R3-RP B	70 % ACN

Table 7: Buffers for HPLC

Buffer name	Composition
Monolithic column HPLC A	0.1 % TFA
Monolithic column HPLC B	0.08 % TFA, 84 % ACN
ERLIC A	20 mM sodium methylphosphonate, pH 2 (adjusted with NaOH), 70 % ACN,
ERLIC B	200 mM triethylammonium phosphate (adjusted with TEA), pH 2, 60 % ACN
NanoHPLC loading buffer	0.1 % TFA
NanoHPLC A	0.1% FA
NanoHPLC B	0.1% FA, 84 % ACN
High-pH-RP A	10 mM ammonium formate, pH8
High-pH-RP B	10 mM ammonium formate, pH8 , 84 % ACN

Table 8: Buffers for TiO_2-affinity purification

Buffer name	Composition
TiO_2 loading buffer 1	90 % ACN, 5 % TFA, 1 M glycolic acid
TiO_2 loading buffer 2	70 % ACN, 2 % TFA
TiO_2 washing buffer 1	80 % ACN, 1 % TFA
TiO_2 washing buffer 2	10 % ACN, 0.1 % TFA
TiO_2 washing buffer 3	50 % ACN 0.1 % TFA
TiO_2 elution buffer	1 % NH_4OH (pH11.3)

5.2. Solid-phase extraction (SPE) protocols

The following Table 9 summarizes the elution scheme of all SPE procedures conducted in this work.

Table 9: Protocols for solid-phase extraction of peptides (unless otherwise stated in the text). If no volume is given, one cartridge volume was used in each step (Bond Elute: 6 mL, SPEC C18 15mg/30mg: 3 mL, SPEC C18 4 mg: 400 µL, Hypersep and Oligo-R3: 200 µL). Buffers are listed in Table 6.

Cartridge	Equilibration	Loading	Washing	Elution
Bond Elute RP 1 g	1 X ACN, 3x A	2x in A (flow-through reloaded)	3x A	3x B
SPEC C18 4, 15 or 30 mg	1 X ACN, 3x A	2x in A (flow-through reloaded)	3x A	3x B
Hypersep-RP	2x B, 2x A	2x in A (flow-through reloaded)	3x A	4x 150µl B
Hypersep-SCX	3x B, 3x A	2x in A (flow-through reloaded)	---	For recovery of phosphopeptides, the flow-through of the loading step was collected. Non-modified peptides were eluted with 3x B
Oligo-R3-RP	3x B, 3x A	2x in A (flow-through reloaded)	3x A	3x B

5.3. Cell culture and sample preparation

5.3.1. Cell culture

NIH-3T3 and HeLa S3: Cells were grown in Dulbecco's Modified Eagle Medium (D-MEM, low glucose) supplemented with 10% fetal calf serum (FCS), 4 mM glutamine, 100 U/mL penicillin and 100 µg/mL streptomycin. Incubation conditions were kept constant at 37 °C and 5 % CO_2. Cells were grown in 75 cm^2 cell culture flasks and passaged 1:10 or 1:20 every 3-4 days at 80-90 % confluence as determined by light microscopy. For passaging, cells were washed three times in 10 mL PBS and detached by incubation with 3 mL trypsin solution at 37 °C, 5 % CO_2 for 7 min. Trypsin was inhibited by adding twice the volume of medium and cells were spun down at 200 x g for 7 min. The supernatant was discarded and cells were re-suspended in fresh medium and adjusted to ~500,000 cells / mL. 1 mL of cell suspension was added to a new 75 cm^2 cell culture flask containing 19 mL of fresh medium. Cells were grown up to a maximum of 20 passages.

For stable isotope incorporation, cells were grown in D-MEM (high-glucose) SILAC medium supplemented with 10 % dialyzed FCS, 4 mM glutamine, 100 U/mL penicillin and 100 µg/mL streptomycin in the presence of either $^{13}C_6$-$^{15}N_2$-lysine / $^{13}C_6$-$^{15}N_4$-arginine (SILAC heavy), $^{2}H_4$-lysine / $^{13}C_6$-arginine (SILAC medium) or unlabelled lysine / arginine (SILAC light). All SILAC media were filtered before usage and stored at 4°C for up to 6 month. To ensure complete incorporation of SILAC amino acids (≥ 97 %), cells were grown for at least five passages (1:10 or 1:20) in SILAC medium.

5.3.2. MCMV infection *in vitro* (NIH-3T3), cell lysis (NIH-3T3, HeLa S3, A431) and colon tissue lysis

NIH-3T3: Cells were infected with MCMV strain Smith in 6-well cell culture dishes at a multiplicity of infection of 3 ($TCID_{50}$/cell) using centrifugal enhancement (1000 g, 30 min).[143] Uninfected controls (mock) were treated identically, but in the absence of virus particles, and were harvested after 4 h to avoid further cell proliferation, but to allow cells to recover from centrifugal enhancement. For harvesting, cells were washed with ice-cold PBS and detached by adding 1 % SDS, 50 mM Tris-Cl pH 7.6, 150 mM NaCl (+ 1 tablet Complete mini, and 1 tablet PhosSTOP per 10 mL) using a cell scraper. The suspension was transferred to a reaction tube and homogenized using an ultrasonic probe (10 s cycles, 40 watts and 3-4 repetitions) with incubation on ice in between.

HeLa S3: Before harvesting, cells were grown in 75 cm^2 or 175 cm^2 cell culture flask for 3 days in D-MEM (low glucose) supplemented with 10% fetal calf serum (FCS), 4 mM glutamine, 100 U/mL penicillin and 100 µg/mL streptomycin at 37°C and 5 CO_2 to 75 % confluence. Before harvesting, cells were washed three times with ice-cold PBS. HeLa S3 cells were detached with a cell scraper in SDS lysis buffer (200 µL / 10^6 cells) and transferred to a 15 mL-falcon tube. Cells were lysed using an ultrasonic probe (10 s cycles, 40 watts, 3-4 repetitions) with incubation on ice in between. In case of A431, cells were detached with 0.05 % trypsin/0.02 % ethylenediaminetetraacetic acid in PBS and the

suspension was transferred to a 15 mL-falcon tube. Cells were spun down by centrifugation (300 x g, 5 min, RT) and 200 µl SDS lysis buffer was added per 10^6 cells. For degradation of nucleic acids, A431 lysate was incubated with benzonase (150 U / mL lysate) for 30 min at 37°C.

Colon tissue samples: Patient tissue was snap- frozen (by Prof. Massimiliano Mazzone and co-workers, Vesalius Research Center, Leuven, Belgium), shipped on dry ice and stored at -80°C until further processing. The tissue material was lysed in 1 % SDS, 50 mM Tris-Cl pH 7.6, 150 mM NaCl (+ 1 tablet Complete mini, and 1 tablet PhosSTOP per 10 mL) and homogenized using an ultrasonic probe (10 s cycles, 40 watts, 3-4 repetitions) with incubation on ice in between.

5.3.3. Purification of the soluble protein fraction from lysates

Lysates were centrifuged at 20,000 x g for 15 minutes to remove debris and the clear supernatant, containing the soluble protein fraction, was transferred to a new tube. The protein concentration was determined by a bicinchoninic acid assay (Pierce kit) against a 5 point calibration curve of bovine serum albumin (5, 25, 50, 125, 250 µg BSA/mL, included in the kit). Samples were diluted 1:10, 1:20 and 1:40 resulting usually in two out of three dilutions being in the linear range of the calibration curve. Samples and calibration curve were always measured in triplicates.

Cysteines were reduced using 10 mM DTT at 56 °C for 30 min and alkylated using 20 mM IAA for 30 min at room temperature in the dark. Protein were precipitated by diluting the sample 10-fold in ice-cold ethanol (-40 °C) followed by vigorous mixing and incubation at -40 °C for 1 h. Precipitated proteins were spun down at 20,000 x g at 4 °C for 40 min and the pellet was washed with ice-cold acetone (-40 °C) using 10 % of the EtOH volume used for precipitation. After 15 minutes of incubation at -40 °C, proteins were spun down for 20 min at 20,000 x g and 4 °C. The sediment was re-solubilized in 6 M guanidinium hydrochloride (GuHCl) by vigorous mixing, ultrasonication and repeated freezing and thawing.

5.3.4. Digestion of proteins and quality control

For digestion, samples were diluted 1:20 using 50 mM ammonium bicarbonate (ABC) to a final GuHCl concentration of 300 mM and $CaCl_2$ was added to a final concentration of 1mM. Trypsin was added to a 1:20 or 1:33 enzyme-to-sample ratio and samples were incubated at 37 °C for 14-16 h under slight agitation (thermomixer, 300 rpm). For digestion of A431 with subtilisin, samples were diluted 1:30 with 50 mM ABC and subtilisin was added in a 1:20 enzyme-to-sample ratio followed by incubation at 56°C for 20 min. Digestion reactions were stopped by acidification with formic acid (FA) to a final concentration of 1%. Samples were desalted by SPE using a vacuum manifold and SPEC C18 AR cartridges (4 mg cartridges for up to 200 µg of digest, 15 mg cartridges for up to 750 µg of digest and 30 mg cartridges for up to 1.5 mg digest) or Bond elute C18 1 g cartridges (for up to 50 mg of digest). Eluates were dried in a vacuum centrifuge and peptides were stored at -80°C for subsequent experiments.

For quality control, 400 ng of digested sample were injected to a monolithic column-HPLC to verify a complete digestion and to identify potential contaminations (Figure 9). When using a monolithic column-HPLC, peptide and protein RT roughly correlate with molecular weights, hence tryptic peptide elute earlier (Figure 9, UV traces 1A-C and 2 A-C) than undigested or partially digested proteins (Figure 9, UV trace 3A). Beside, the monolithic column-HPLC allowed identification of potential contaminations (Figure 9, UV traces 1A-C).[51]

5.4. Phosphopeptide enrichment and pre-fractionation methods

Phosphopeptide enrichment was conducted using either ERLIC or TiO_2, two methods of extraordinary sensitivity and good reproducibility as shown in previous studies.[104, 124] ERLIC was performed using an HPLC allowing for simultaneous enrichment and fractionation of retained phosphopeptides, whereas non-modified peptides eluted early (within the first 5 minutes). Strongly retained phosphopeptides were eluted using a linear gradient of slightly decreasing ACN concentration (70 % to 60 %) and a parallel

change of the salt from 20 mM Na-methylphosphonate to 200 mM triethylammonium phosphate at pH 2.[121]

For TiO_2, digests were incubated with TiO_2 beads in 2 mL-reaction tubes followed by consecutive washing steps to remove unbound peptides. After release of TiO_2-bound phosphopeptides using a basic elution buffer (pH 11.3), TiO_2-particles were removed using self-made C8-SPE-tips. In most experiments, the phosphopeptide-enriched fraction was subjected to a second round of enrichment to further enhance the specificity. The following sections give the experimental details of both methods ERLIC and TiO_2.

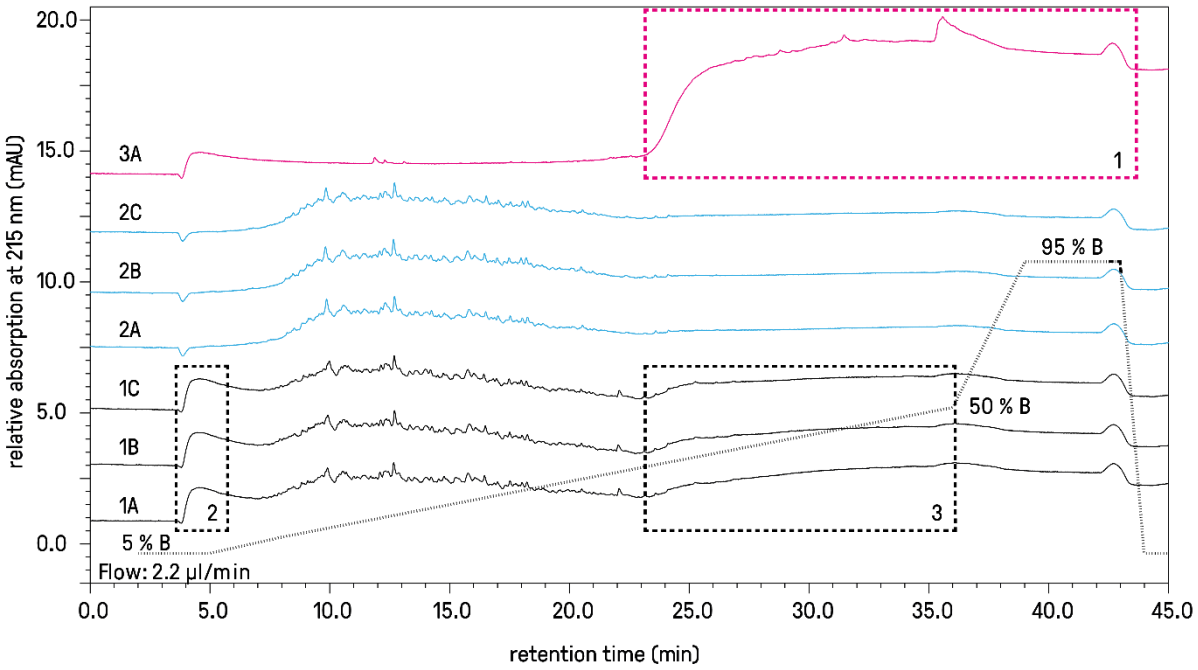


Figure 9: Quality control after tryptic digestion and subsequent RP-SPE using a monolithic column-HPLC. Triplicate digests of the NIH-3T3 soluble protein fraction were separated before (1A-C) and after SPE (2A-C). In comparison, non-digested proteins (3A) appear at later RT (≥ 23 min, dashed box 1). Potential contaminations (4 min and 25-35 min, 1A-C, dashed box 2 and 3) were successfully removed by RP-SPE. The dashed line indicates the buffer gradient. Column: PepSwift monolithic capillary column (200 µm × 5 cm) and PepSwift monolithic trap column, (200 µm × 5 mm). Buffer A/trap column loading buffer: 0.1 % TFA, Buffer B: 0.08 % TFA, 84 %.

5.4.1. Electrostatic repulsion-hydrophilic interaction liquid chromatography (ERLIC)

ERLIC-based phosphopeptide enrichment was conducted with an Ultimate HPLC, equipped with a low-pressure-gradient system with an inert micropump, a Famos autosampler, a variable wavelength detector (up to 4 wavelength, 190 – 740 nm) and an online degassing system for buffers (Figure 10A). Buffers were used for up to 10 days and prepared according to Alpert 2008.[121] ERLIC buffer A: 20 mM Na-methylphosphonate, 70 % ACN, pH 2 and ERLIC B: 200 mM triethylammonium phosphate, 60 % ACN, pH 2. Except for experiment 1-5 (section 5.7.5), all ERLIC separations were conducted using a PolyWAX 4.6 x 100 mm, 5 µm, 300 Å column, which exhibits improved selectivity in terms of separating phosphorylated peptides and non-phosphorylated counterparts, chromatographic resolution and peak width (Figure 10 B, Figure 11A, comparison with PolyWAX 2.1 x 200 mm, 5 µm, 300 Å). Experiment 1-5 was conducted with a 2.1-mm column to reduce the volume of the collected 4-min fraction (800 µl instead of 4 mL), as presence of large quantities of ERLIC buffer residuals increases the chance of clogging SPE cartridge and nanoHPLC columns. Samples were injected in 20 µL A and separated using a binary gradient program starting with 100 % ERLIC A (usually for 5 to 10 min), followed by a linear gradient to 100 % ERLIC B (for 5 to 15 minutes) followed by an isocratic elution step with 100 % ERLIC B for 5 to 10 minutes. Afterwards, the column was re-equilibrated for at least 20 minutes with 100 % ERLIC A. The flow rates were 1 mL/min (4.6 mm column) and 200 µl/min (2.1 mm column). UV traces were acquired at a wavelength of either 214 or 230 nm resulting in ERLIC typical chromatograms as shown in Figure 10, with three artefact signals (a1-a3) at RT minute 12, 18 and 19, presumably derived from methylphosphonic acid impurities (≥ 98 % according to vendor).

For quality control, up to 5 nmol of the synthetic phosphopeptide DyVPML (phosphorylated amino acid indicated by small letter) were frequently injected to monitor the stability of RT (4.4 ± 0.5 min) and the peak shape (≤ 0.25 min full width at half maximum, FWHM). The peptide eluted in the isocratic part of

the gradient (100% A) and allowed monitoring of RT independently of the employed gradient-program (Figure 11). Besides, baseline separation of the artefact signals a2 and a3 served as indicator for a good chromatographic resolution (Figure 10).

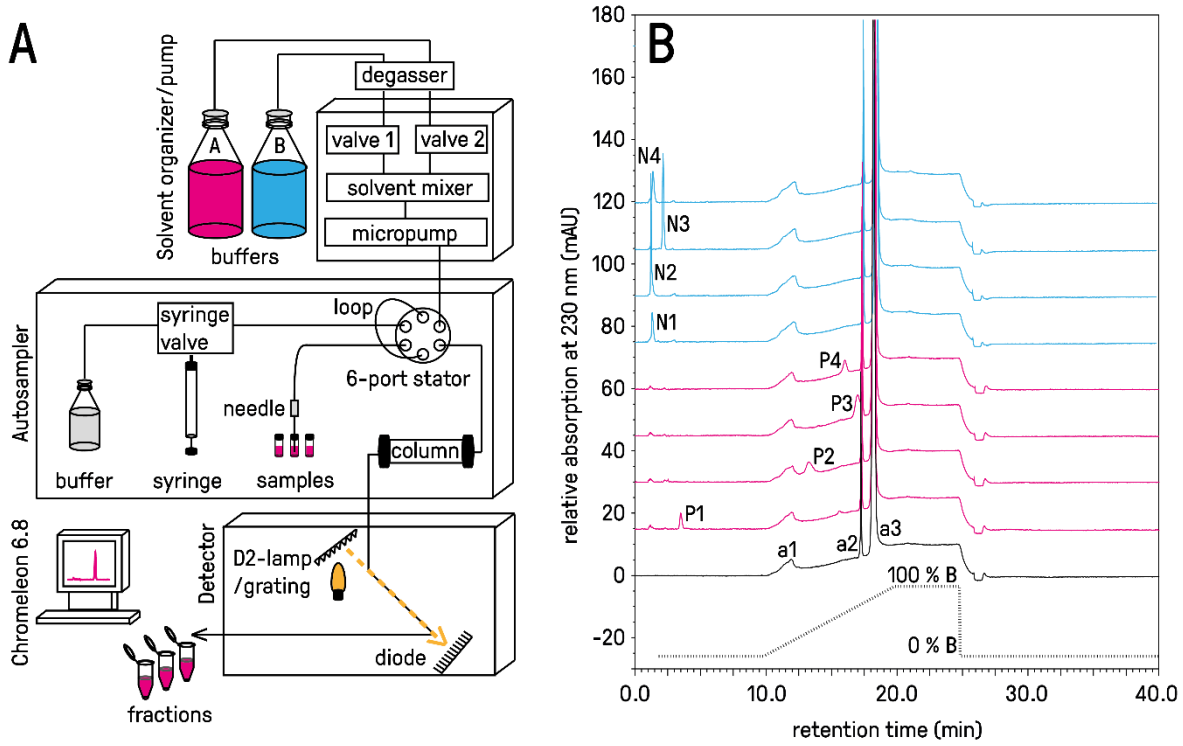


Figure 10: A: Schematic representation of the HPLC setup used for ERLIC. Gradients were generated by a low pressure mixing system connected to an HPLC micropump. An autosampler equipped with a 6-port injection valve (stator) was employed to inject samples into the system and UV-traces were aquired using a variable wavelength detector. B: ERLIC chromatograms show the separation of phosphopeptides (P1-P4) and their non-phosphorylated counterparts (N1-N4) using the indicated gradient (dashed line). The black trace shows a blank run (no sample injected) with ERLIC-typcial artifact signals (a1-a3) at 12, 18 and 19 min. Peptide sequences: (P/N) 1. SLAsIEAK, 2. sLTRsPPAIR, 3. EYFDDsTEER, 4. SFLsEPSsPGR (phosphorylated amino acid as small letters); column: PolyWAX 4.6 x 100 mm, 5 µm, 300 Å; buffer A: 20 mM methylphosphonic acid, 70 % ACN, pH2; buffer B: 200 mM triethylammonium phosphate, 60 % ACN, pH 2; flow rate: 1 mL/min

Peptide standards were separated by injection of up to 5 nmol in 20 µL buffer A. For separation of complex peptide samples (digested proteins), dried samples were re-solubilized in 20 µL ERLIC A and incubated for 10 seconds in an ultrasonic bath. In case of fractionation, ERLIC fractions were collected manually according to RT. At least one blank gradient was in between injections of complex samples to reduce carry over. Fractions were dried in a vacuum centrifuge, followed by RP-SPE to remove ERLIC

buffer residuals. Therefor fractions were re-solubilized in 200 µL 0.1 % TFA and loaded (two times) onto a Hypersep RP 10 -200 µL cartridge. After 3 washing steps with 200 µL 0.1 % TFA, peptides were eluted using 150 µL 60 % ACN (3 times). Eluates were dried completely *in vacuo.*

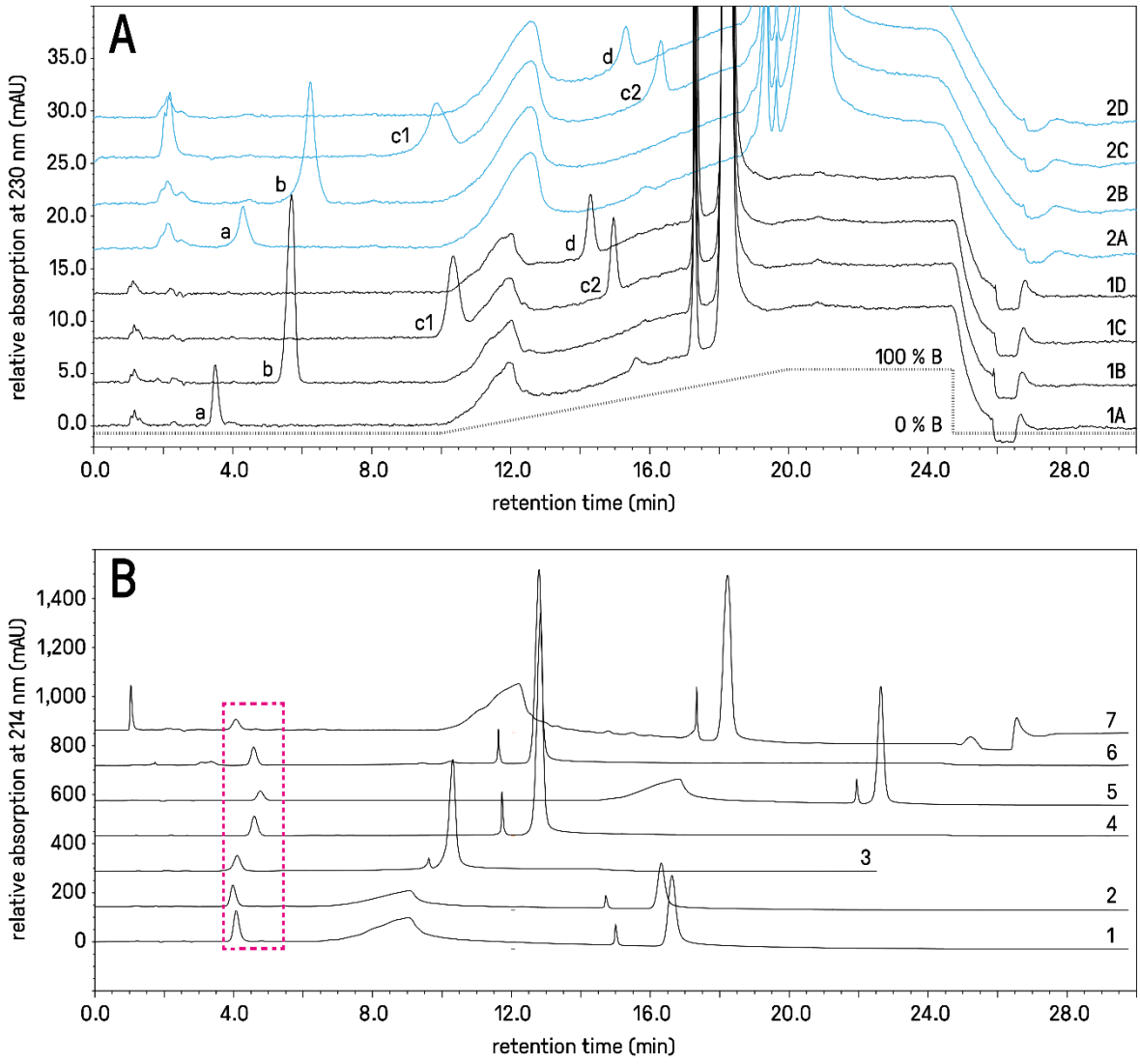


Figure 11A: A PolyWAX 4.6 x 100 mm column (black, 1 mL/min) was clearly superior to a 2.1 x 200 mm column (blue, 250 µL/min) for the separation of five selected phosphopeptides (A-D) using the indicated gradient (dashed line). B: Quality control of 4.6 x 100 mm columns was done regularly (over 2 years) using peptide E (red dashed box). The peptide eluted reproducibly after 4 min in the isocratic part (100 % A) with a stable RT even when the gradient-program was changed (UV-traces 1-7). Peptide A: SLAsIEAK, B: sTFSTNYR, C: FNYSGsGGR, D: sFSIsPVR, E: DyVPML (phosphorylated amino acid as small letters); buffer A: 20 mM methylphosphonic acid, 70 % ACN, pH2; buffer B: 200 mM triethylammonium phosphate, 60 % ACN, pH2.

For the experiment 2 aiming at a parallel analysis of the proteome and phosphoproteome of MCMV-infected fibroblasts, the ERLIC fractions 1-8 were subjected to SCX-SPE yielding (I) the phosphopeptide-enriched flow-through and (II) an eluate of predominantly non-modified peptides. For SCX-SPE, fractions were re-constituted in 200 µL Hypersep SCX A and samples were loaded two times on a Hypersep POROS SCX 10-200 µL cartridge. The flow-through was collected, dried *in vacuo* and up to 50 % were subjected to LC-MS analysis. The peptide fraction which remained on the SCX cartridge (predominantly non-phosphorylated)[124] was eluted using 2 x 200 µL Hypersep SCX B buffer. Eluates of all 8 ERLIC fractions and the remaining material from the flow-through (not used for LC-MS) were combined and subjected to RP-SPE as described in the previous paragraph.

5.4.2. Metal-oxide affinity chromatography using TiO_2-microparticles (TiO_2)

To enrich phosphopeptides by TiO_2, peptide samples were re-dissolved in 1 mL TiO_2 loading buffer 1 and transferred to a 2 mL LoBind-tube. TiO_2 microparticles were added in a 1:6 ratio (sample:TiO_2, w/w) followed by incubation at room temperature for 10 minutes under intense agitation. Afterwards microparticles were spun down (1 min, 5000 x g) and the supernatant was transferred to a new 2 mL reaction tube containing half the amount of TiO_2 (1:3, sample:TiO_2) and incubated under the same condition. After spinning down, the supernatant was transferred once again to a new 2 mL tube containing again half the amount of TiO_2 (1:1.5 sample:TiO_2) followed by another round of incubation. The supernatant was discarded and the TiO_2-microparticles from all three incubations were combined in a new 2 mL-tube using 100 µL of TiO_2-loading buffer 1. For washing, TiO_2-microparticles were re-suspended in 100 µL TiO_2 washing buffer 1, vigorously stirred for 30 seconds, spun down (1 min, 5000 x g) and the supernatant was discarded. After a second round of washing using TiO_2-washing buffer 2, beads were completely dried *in vacuo* for approximately 10 min followed by addition of 100 µL elution buffer and incubation for 10 minutes under intense agitation. Beads were spun down and the supernatant

was filtered through a C8-disk-plugged tip into a reaction tube by applying positive air pressure. Beads were incubated a second time with 100 µL elution buffer before the entire suspension was filtered using the same tip. The entire tip was flushed with 30 µL 50% ACN to elute peptides bound to the C8-disk. The eluate was dried using a vacuum centrifuge.

Prior LC-MS analysis, TiO_2-eluates were desalted by RP-SPE using self-packed Oligo-R3 tips (containing 5 µL of a 50 % v/v suspension of Oligo-R3 in ACN) plugged with a C18-disk. All liquid was pushed through the tip by applying positive air pressure. Eluates were reconstituted in 50 µL 0.1 % TFA and loaded onto a tip, which was equilibrated 2 times with 50 µL Oligo-R3-RP elution buffer and 2 times 50 µL Oligo-R3-RP loading buffer. All liquid was pushed through the tip, and the flow-through was loaded again 2 times, before washing 3 times with 50 µL loading buffer. Peptides were eluted using 3 times 50 µL Oligo-R3-RP elution buffer and eluates were dried in a vacuum centrifuge for subsequent LC-MS analysis.

5.4.3. HPLC-based pre-fractionation using high-pH-RP

To further reduce the complexity of samples, peptides were pre-fractionated by HPLC-based high-pH-RP (pH 8) using an U3000 HPLC equipped with a micropump, an autosampler capable of fraction collection, a column oven and a UV detector. Peptides were separated at a flow rate of 12.5 µL/min on a Biobasic-C18 column (0.5 mm x 15 cm, 5 µm particles with 300 Å pore size) at 30°C using high-pH-RP buffer A: 10 mM ammonium formate, pH8 and high-pH-RP buffer B: 10 mM ammonium formate, pH8, 84 % ACN. The following gradient was used: 15 minute isocratic at 3 % B, followed by a linear gradient to 50 % B in 65 minutes. The column was washed with 95 % B for 5 minutes followed by re-equilibration using 3 % B for 15 minutes. Samples were injected in 15 µL A and fractions were collected by concatenation of 2 minute windows (Figure 12) at the RT where the vast majority of

peptides were expected (usually ~15-75 min) using the automated fraction collection option. All samples were dried *in vacuo* before subsequent LC-MS analysis.

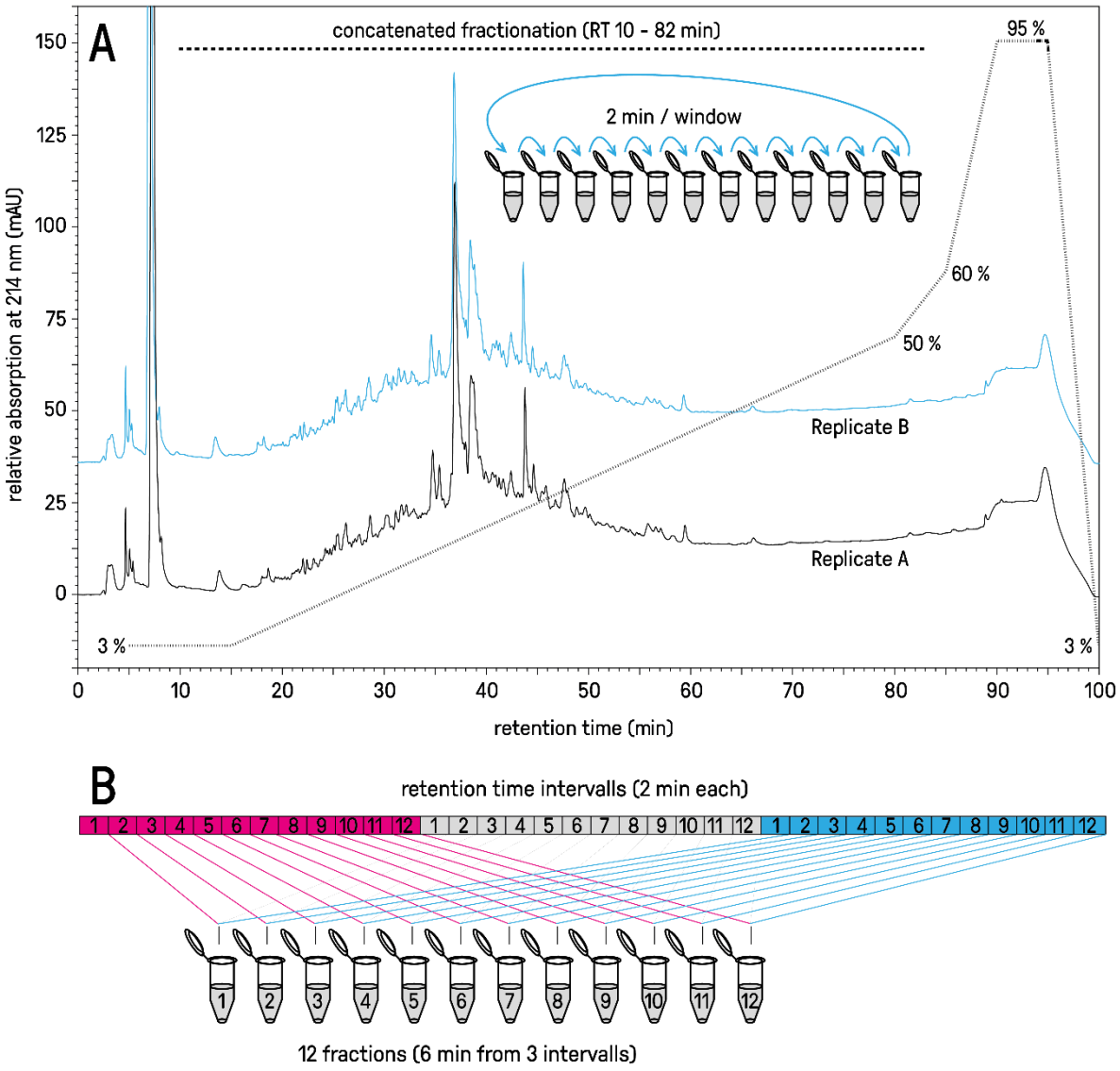


Figure 12: A phosphopeptide-enriched sample (two technical replicates) was separated by high-pH-RP in duplicates (A) with concatenated fraction collection (B). The vast majority of peptides elute in the retention time window from 20 – 60 min, the fractionation window was from 10-82 minutes. Buffer A: 10 mM ammonium formate, pH 8; buffer b: 10 mM ammonium formate, pH 8, 84 % ACN; column: Biobasic-C18, 0.5 mm x 15 cm, 5 µm.

5.5. LC- MS

For LC-MS, hybrid MS instruments (LTQ Orbitrap Velos Pro, LTQ-Orbitrap Elite or Q Exactive HF) were online-coupled to a U3000 HPLC or a U3000 RSLC nanoHPLC. Peptides were separated by ion-pairing low-pH-RP allowing for high chromatographic resolution. Peptides were ionized at ambient pressure using a commercial nanoESI source as interface before MS. MS instruments were operated in data-dependant acquisition-mode (DDA) in case of large-scale experiments and in parallel reaction monitoring-mode (PRM) for targeted detection of peptides.

5.5.1. nanoHPLC

nanoHPLC was conducted using either a U3000 equipped with two micropumps and a split system to deliver nanoflow and low pressure gradient system or a U3000 RSLCnano systems equipped with a nano-flow selector and a high pressure gradient system and an additional micropump. Systems were further equipped with an autosampler, a column oven and a UV detector for acquisition of chromatograms. 15 µL of peptide sample were loaded in 0.1 % TFA on a C18 trap column (100 µm x 2 cm) at a flow rate of 10-20 µL / min using a micropump. After 10 minutes the trap column was switched in-line with the C18 main column (75 µm x 30 cm or 75 µm x 50 cm) operated at a flow rate of 250-270 nL / min and 60 °C (using either a nanopump or a micropump connected to a 1:1000 flow splitter cartridge). Peptides were separated using a linear gradient-program from 3 % up to 50 % nanoHPLC buffer B, followed by a 3-5 min washing step at 95 %, before re-equilibration with nanoHPLC loading buffer and nanoHPLC buffer A. Gradient slopes and lengths were adjusted according to the properties of the sample. Especially peptides from late ERLIC fractions, comprising predominantly early eluting hydrophilic peptides, were separated by gradients of reduced slope and lower percentage of ACN (e.g. 3-25 % nanoHPLC buffer B). For MS-independent quality control, UV-traces were acquired at 214 nm.

5.5.2. Mass spectrometry

All instruments were directly coupled to the UV-detector of an HPLC via a nanoESI source equipped with a metal-coated nanoESI emitter (10 µm in diameter) operated at 1.5-1.8 kV in positive ion-mode with the ion transmission optimized for maximum intensity at 640 m/z (tuning). Instruments were calibrated once a week to maintain a mass accuracy of ≤ 5 ppm and polysiloxane signals of 371.10124 m/z or 445.12003 m/z were used as lock masses for internal mass calibration.[191] For quality control of peptide retention, resolution and signal intensity, a 7-peptide standard with a concentration range of 1-30 fmol was measured daily[192] and a complex peptide sample of 50 ng HeLa digest was measured at least once a week.

DDA was performed using topN CID when using LTQ-Orbitrap instruments or topN HCD for quadrupole-Orbitrap instruments with N = 10 or 15. A high resolution survey scan (MS) was acquired in the Orbitrap at a resolution of 60,000 (@ m/z 400 for LTQ-Orbitrap instruments or m/z 200 for quadrupole-Orbitrap instruments) and the N most intense ions were isolated and fragmented to acquire low resolution MS/MS spectra (R≈3,000) in the linear ion trap (LTQ-Orbitrap instruments) or high resolution MS/MS spectra (R=15,000) in the orbitrap (quadrupole-Orbitrap instruments). Fragmented ions were excluded from re-fragmentation for at least 10 sec, typically for 25 or 30 sec. The maximum ion injection time was set to 100-250 ms and the automatic gain control (AGC) target values were usually 10^6 for MS and 10^4 for MS/MS in the linear trap (LTQ-Orbitrap instrument) or $5x10^4$ for MS/MS in the Orbitrap (Quadrupole-Orbitrap instrument). Ions with a charge of +1 and ions with undetermined charge states were excluded from fragmentation.

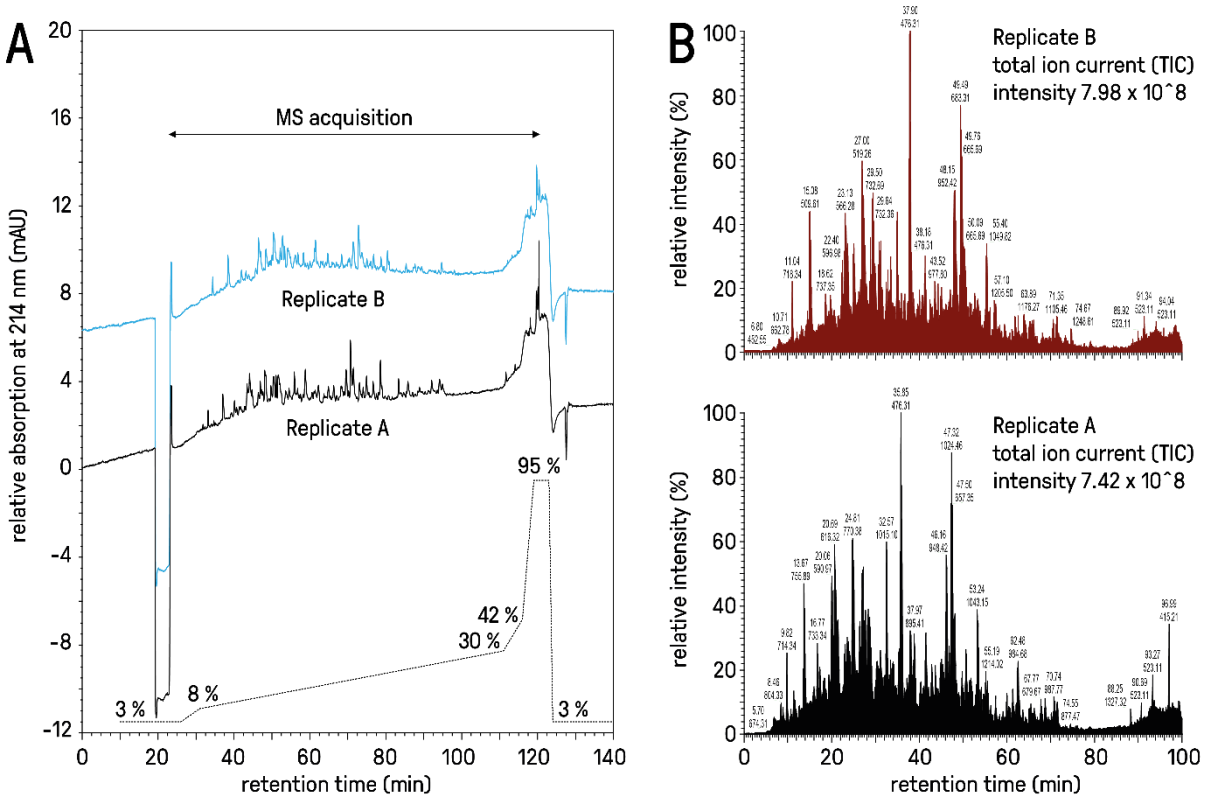


Figure 13: LC-MS analysis of a high-pH-RP fraction of a phosphopeptide-enriched sample (2 replicates). The UV chromatogram (A) allows for MS-independent quality control, but exhibits lower sensitivity than the chromatogram resulting from the total ion current (TIC) of the LTQ Orbitrap Velos Pro mass spectrometer (B). The optimized gradient (dashed line in A) enabled improved separation in the range from 8-30 % buffer B, where the vast majority of peptides eluted. nanoHPLC loading buffer: 0.1 % TFA; nanoHPLC A: 0.1 % FA; nanoHPLC B: 0.1 % FA, 84 % ACN; precolumn: ACE C18 column 100 µm x 20 mm, 3 µm, 100 Å, self-packed; main column: 75 µm x 25 cm, 3 µm, 100 Å, self-packed

5.6. Data analysis

5.6.1. Database search and phosphorylation site validation

Raw-files were processed using Mascot 2.4 implemented in Proteome Discoverer (PD) v1.3, v1.4 or v2.2 (Thermo Scientific). For identification of peptides in HeLa and A431 samples, database (DB) searches were performed against Uniprot (www.uniprot.org), taxonomy *human* (10-2014 for experiment 1-1; 07-2015 for experiments 1-2 and 1-3). For identification of peptides in MCMV-infected

mouse fibroblasts, a custom database was generated comprising the Uniprot entries of the taxonomy *Mus musculus*, all non-redundant entries for mouse cytomegalovirus from Uniprot/TrEMBL (both 01-2013) and the 6-frame *in silico*-translated murid herpesvirus 1 (MCMV) genome[193] with the reference sequence NC_004065.1 from the National Center for Biotechnology (NCBI, www.ncbi.nlm.nih.gov/nuccore/21716071). DB search parameters were: trypsin with two missed cleavage sites, carbamidomethylation of cysteines as fixed modification and oxidation of methionines and N-terminal acetylation as variable modification. In case of phosphopeptide-enriched samples, phosphorylation of serine, threonine and tyrosine was included as variable modification. In case of SILAC samples, stable isotope-labelled lysine and arginine were included as variable modifications ($^{13}C_6$-lysine and $^{13}C_6$-arginine: +6.020 Da, $^{2}H_4$-lysine: +4.024 Da, $^{13}C_6$-$^{15}N_2$-lysine: +8.014 Da and $^{13}C_6$-$^{15}N_4$-arginine: +10.008 Da) and the precursor ion quantifier node was used to determine the area of the respective peptide peaks (MS level). Percolator or target-decoy PSM validator were used for spectra validation and phosphoRS 3.1[194] was employed for phosphorylation site validation. A false discovery rate of ≤ 1 % was used to filter for confident PSMs in all methodological experiments. In the time-resolved MCMV experiment, peptides from PSMs with ≤ 5% FDR were considered, if identified in multiple replicates. Resulting PSM lists were exported as excel-tables and converted to comma separated value-files using Microsoft excel.

In case of digestion with the broad specificity protease subtilisin, enzyme specificity was set to *none* with zero missed cleavage sites. For the systematic comparison of trypsin and subtilisin, N-terminal acetylation was not selected as variable modification. Other parameters were as described.

5.6.2. Mapping phosphopeptides to proline dense regions

Proteins in the Uniprot human database were screened for proline-rich regions by determining the proline-density at each position within a protein as follows: Using the position vector of all prolines

within a protein, the kernel-density function was calculated and multiplied by the number of prolines (in the protein) resulting in an absolute measure of proline-density (bandwidth = 10, n = 16 384). Regions with a density above 0.2 were regarded as proline-rich (roughly accounting for at least 25% of proline residues within 25 amino acids). Proteins were plotted according to the relative positions of the proline dense regions and proline density was indicated as colour gradient. Peptides within proline-dense regions were indicated as black bars.

5.6.3. Quantification of proteins and phosphopeptides

PSM lists were further processed using R v3.2.0 and entries without lysine or arginine (indistinguishable channels), with inconsistent labelling (e.g. lysines/arginine from different SILAC channels), redundant sequences or phosphorylation sites < 90 % phosphoRS probability were removed. Areas of PSMs, if identified in different fractions, were summed to calculate peptide areas and Mascot scores of each top-scored PSM were assigned as peptide scores. For quantification of proteins and protein groups the summed area of all corresponding non-phosphorylated peptides was used. The sum of all areas in every SILAC channel was used for normalization of proteins, protein groups and phosphopeptides (except for experiments to determine quantitative loss during enrichment). The binary logarithm (log2), calculated for each possible channel pair, was used for further analyses.

5.6.4. Statistical analysis for MCMV-infected fibroblasts using limma and rank product

To identify differentially regulated features (DFRs) in MCMV-infected fibroblasts, 2 out of 4 conditions (0, 4, 8, 24 hours post-infection, hpi) were quantified against a pooled reference using a 3-plex SILAC approach allowing for a relative comparison of signal intensities over all four time points. For each replicate, the area of each feature was divided by the area of the respective feature from the pooled-

reference giving relative abundances with a theoretical range from 0 to 400 %. This approach further offers the advantage that the sum of a feature over all time points is theoretically 400 % and irregularities in quantification (problems with feature assignment) can be directly inferred by inspection of plotted time-courses or quantitative values. For further statistical testing, values were used to calculate ratios for each time point against the uninfected control (0 h) resulting in the log ratios: 4h/0h, 8h/0h, 24h/0h. For analysis of changes in the host cell upon infection, features originating from MCMV proteins were removed to reduce the number of features with infinite ratios (originating from missing values in the mock sample). All MCMV features (proteins and phosphopeptides) were processed separately.

For identification of differentially regulated proteins and phosphopeptides (differentially regulated features, DRFs), the R-script provided in the supplemental material of Schwämmle et al.[195] was used for assigning p-values according to the Student's t-test. Differentially regulated features were identified based on rank product and limma (moderated t-test based on empirical bayes)[196] and p-values were further corrected to meet a 5 % false discovery rate using functions from the q-value package (Benjamini Hochberg). Phosphopeptides were only considered, if the respective protein were quantified to allow for phosphopeptide ratio correction using the protein ratio.

Most String Network analyses (www.string-db.org) were performed with proteins regulated at least 2-fold and a q-value ≤ 0.05 (limma or rank product) as the complexity of the dataset would have exceeded the limits of the platform. Protein and phosphopeptide expression patterns were analyzed by hierarchically clustering (method: "ward.D" - minimum variance method) using the heatmap.2 function of the R-package "gplots" and functional enrichment analysis of clusters was performed using the DAVID online tool (https://david.ncifcrf.gov/).[197] The data were screened for systemic regulation within the KEGG-pathways (http://www.genome.jp/kegg/pathway.htmL)[198] using the R-package "pathview" (https://bioconductor.org/packages/pathview).[199]

5.6.5. Analysis of viral protein expression in MCMV-infected fibroblasts

Proteins and phosphopeptides were quantified as described for the host (section 5.6.3), but non-unique peptides not shared by host proteins (but not to host proteins) were used to generate viral protein groups. For mock the relative expression (0 h) was set to 0 % indicating absence of any viral protein expression. Expression profile plots were drawn using R giving the expression determined for each replicate and the mean expression determined as the average from all three replicates. The mean expression profiles of all viral proteins without any missing value were hierarchically clustered (method: "ward.D" - minimum variance method) using the heatmap.2 function of the R-package "gplots".

5.7. Experimental Design

5.7.1. Experiment 1-1: ERLIC of trypsin and subtilisin digests

For the characterization of ERLIC, 300 µg trypsin or subtilisin digest (A431, unlabelled) were fractionated in 12 fractions using the following gradient: 100 % A for 10 min, a linear gradient to 100 % B in 10 min, 100 % B constant for 10 min. Fractions 1-7 were processed by SCX-SPE to further enrich phosphopeptides (Loroch et al. 2015)[124] – fractions 8-12 were processed by RP-SPE. For analysis, a Q-Exactive HF mass spectrometer online coupled to U3000 RSLCnano was employed and the experiment was performed in triplicates. This experiment aimed at a comparison of phosphopeptide IDs, physicochemical properties and complementarity of both proteases.

5.7.2. Experiment 1-2: Relative quantitative recovery of phosphopeptides in ERLIC and 2-step TiO_2

To estimate the relative quantitative recovery of phosphopeptides in ERLIC and TiO_2, 400 µg SILAC light and heavy digests of HeLa where subjected to phosphopeptide enrichment with either method in

order to mix the differentially labelled phosphopeptide enriched samples of ERLIC and TiO_2 afterwards (Figure 14). A large ERLIC fraction was collected from 3 – 20 min, dried *in vacuo* and reconstituted in 0.1 % TFA. 50 % of the sample was combined with 50 % of a differentially labelled eluate from 2-step TiO_2 (after drying and reconstitution in 0.1 % TFA). After removal of salts using RP-SPE, the sample was pre-fractionated by high-pH-RP in 12 fractions (Figure 14) by concatenation using 2-minute intervals between RT 10 and 82 min. In addition, the flow-through from 0-10 min and a "late"-fraction from 82-95 min were collected. After drying in vacuum centrifuge, fraction were reconstituted in 0.1 % TFA for subsequent LC-MS analysis (LTQ-Orbitrap-Velos Pro online coupled to U3000 HPLC). As technical replicate, a label-switch experiment was performed in parallel.[120]

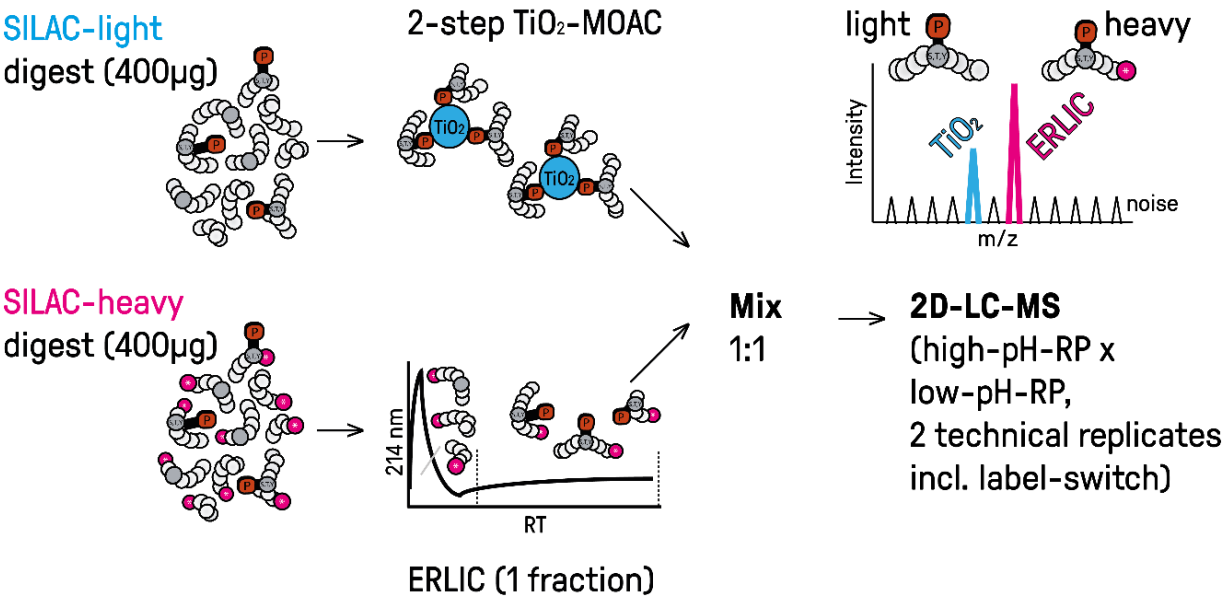


Figure 14: Experiment 1-2: Systematic comparison of the quantitative recovery of phosphopeptides during enrichment with ERLIC or TiO_2. Phosphopeptide enriched fractions from SILAC "light" or "heavy" digests were mixed 1:1 and subjected to 2D-LC-MS. Heavy-to-light ratios revealed relative quantitative recovery of each detected phosphopeptide.

5.7.3. Experiment 1-3: Quantitative loss during enrichment – a comparison of ERLIC and TiO_2

To determine the quantitative loss in one round of ERLIC-based phosphopeptide enrichment, ERLIC was used to generate a large fraction from minute 2 – 12 from a 100 µg-aliquot of SILAC digest using the following gradient: 0 % ERLIC B for 4 min, 0 % to 100 % B in 5 min and 100 % B constant for 5 min. The dried fraction was reconstituted in 0.1 % FA and subjected to RP-SPE (Hypersep spin tips with 0.1 % FA as loading buffer). The dried eluate, was reconstituted in ERLIC A and 50 % (corresponding to 50 µg of starting material) were mixed with 50 µg of a differentially labelled digest. The mixture was subjected to ERLIC and six 2-minute fractions were collected from RT minute 2-14 and subjected to RP-SPE (Hypersep spin tips) for subsequent LC-MS analysis (Figure 15).

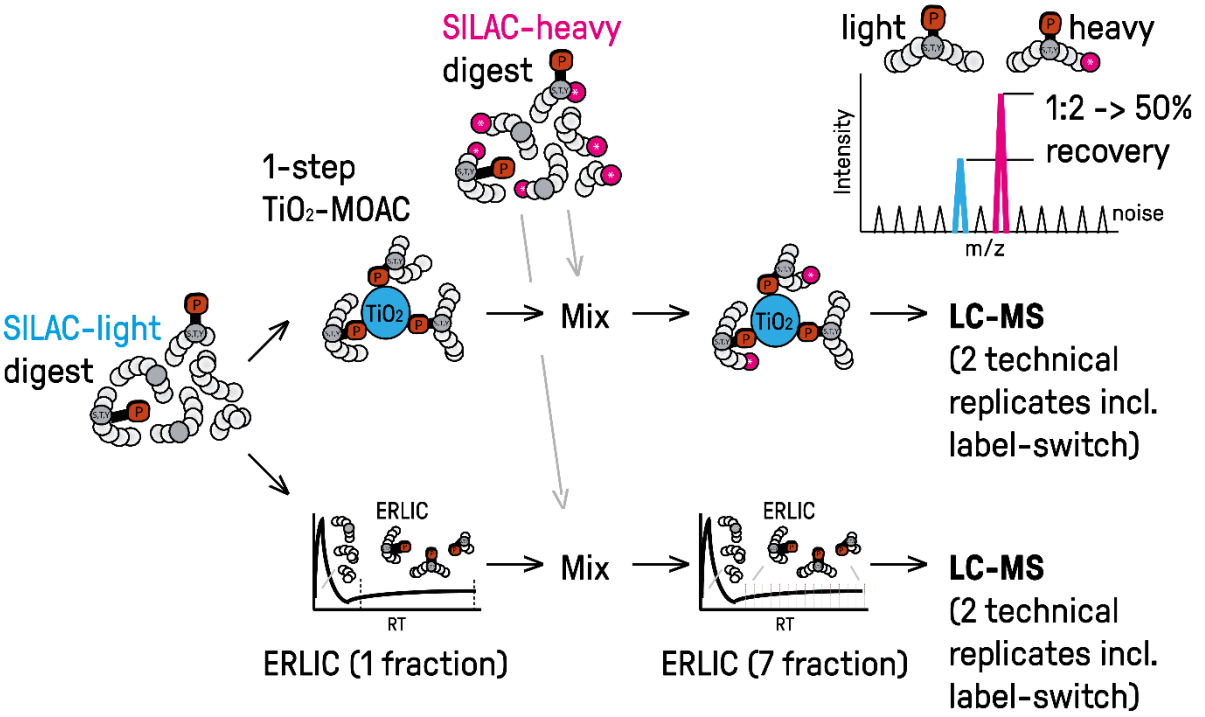


Figure 15: Experiment 1-3: Quantitative loss upon phosphopeptide enrichment using either TiO_2 or ERLIC. A phosphopeptide-enriched fraction, corresponding to 50 µg of a SILAC-light digest, was mixed with 50 µg of a SILAC-heavy digest. Samples were subjected to another round of enrichment for depletion of non-phosphorylated peptides. Heavy-to-light ratios (as determined by LC-MS) revealed the quantitative loss of each method during one round of enrichment.

To determine the quantitative loss in a single round of TiO_2, a 100 µg-aliquot of SILAC digest was subjected to 1-step TiO_2-based phosphopeptide enrichment. The dried eluate was reconstituted in TiO_2-loading buffer and 50 % (corresponding to 50 µg starting material) were mixed with 50 µg of a differentially labelled digest. The sample was subjected to another round of TiO_2 and subjected to RP-SPE (Oligo-R3) for subsequent LC-MS analysis (Figure 15). LC-MS analysis was conducted with 50 % of the sample from each approach (LTQ Orbitrap Velos Pro online coupled to U3000 RSLCnano). As technical replicate, a label-switch experiment was performed in parallel.

5.7.4. Experiment 1-4: Benchmarking ERLIC and TiO_2 for quantitative phosphoproteomics using minute amounts of sample

For a subsequent systematic comparison for large-scale quantitative phosphoproteomics, ERLIC and TiO_2 were employed for phosphopeptide enrichment from minute amounts of a complex biological sample. Identical aliquots of a 2-plex SILAC sample of MCMV-infected fibroblasts (5 hours post-infection, hpi) against non-infected control with 25 µg of digest per condition were used as starting material (Figure 16). **ERLIC:** 1-minute fractions were collected from minute 3-13, except for a 2-minute fraction from minute 7-9 (retention window of low phosphopeptide IDs), using a flow rate of 1 mL/min and the following gradient: 0 % B for 7 minutes, linear gradient to 100 % B in 7 minutes and 100 % B for 5 minutes. All fractions were dried and re-constituted in loading buffer for subsequent RP-SPE (Hypersep). **TiO_2:** Samples were processed using 2-step TiO_2 affinity purification and dried eluates were reconstituted in loading buffer for RP-SPE (oligo-R3-RP).

50 % of each ERLIC fraction and TiO_2 sample were subjected to LC-MS (LTQ-Orbitrap-Velos Pro online-coupled to a U3000 RSLCnano). The MS acquisition time was 30 min for each ERLIC fraction and 200 min for each TiO_2 samples resulting in a sampling rate of ~75 % (acquired MS/MS relative to maximum scan rate). Experiments were conducted in triplicates using aliquots of the same fibroblast

digests. For identification of differentially expressed phosphopeptides, a two-sample Student's t-test was used and p-values were adjusted to a 5 % FDR according to the Benjamini-Hochberg procedure.

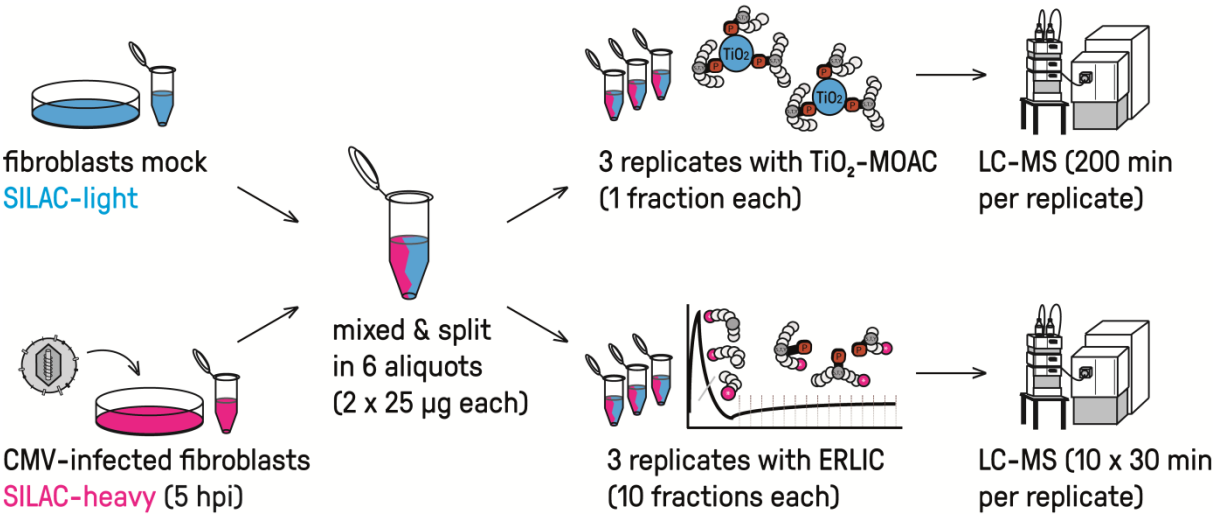


Figure 16: Experiment 1-4: Evaluating ERLIC and TiO_2 for quantitative phosphoproteomics with low sample amounts. Aliquots of 25 µg of a 2-plex SILAC sample of MCMV-infected fibroblasts (5 hpi) against non-infected control cells were subjected to TiO_2 or ERLIC. The experiment aimed at addressing the number of phosphopeptide IDs and robustness of quantification.

5.7.5. Experiment 1-5: ERLIC in conjunction with targeted LC-MS to monitor aberrant phosphorylation in cancer patient tissue

This experiment aimed at an absolute quantification of S125 phosphorylation levels of Egl nine homologue 1 (EGLN-1) in healthy tissue and colon cancer tissue samples of ten patients provided by Prof. Dr. Massimiliano Mazzone (Vesalius Research Center, Leuven, Belgium). All samples were processed and analysed in randomized order to avoid systematic errors. Stable isotope labelled (SIL) versions of the phosphorylated and the non-phosphorylated EGLN-1 peptide AKPPADPAAAAsPCR (small letter gives phosphorylated amino acid) comprising S125 (small letter) were synthesized in-house and subjected to quantification by amino acid analysis. The phosphopeptide was used to determine the RT in ERLIC. Subsequently 32 µg of colon cancer digest were injected including a spike-in of the

synthetic phosphopeptide as internal standard and a 4-minute fraction was collected at the specific retention time (Figure 17). The ERLIC fraction was subjected to RP-SPE for subsequent LC-MS analysis. In addition, 1.2 µg of digest including a spike-in of the synthetic non-phosphorylated peptide as internal standard were subjected to LC-MS.

ERLIC was conducted using a PolyWAX 2.1 x 200 mm to reduce the fraction volume to 800 µL since an excess of ERLIC-buffer residuals might hamper robust solid-phase extraction and reduce nanoLC-column lifespan (see section 5.4.1).[124] All samples were analysed in PRM mode (Q Exactive HF online coupled to a U3000 RSLCnano) triggering the masses of the +2 and +3 ion of the endogenous peptide and the stable isotope labelled standards. The concentration of the endogenous peptides was estimated via the peak area ratio of the endogenous peptide compared to the spike-in, which further allowed estimation of the degree of EGLN-1 S125 phosphorylation (site occupancy). Since healthy and cancerous tissues pairs were from the same individual a paired Student's t-test was adequate to test for significance.

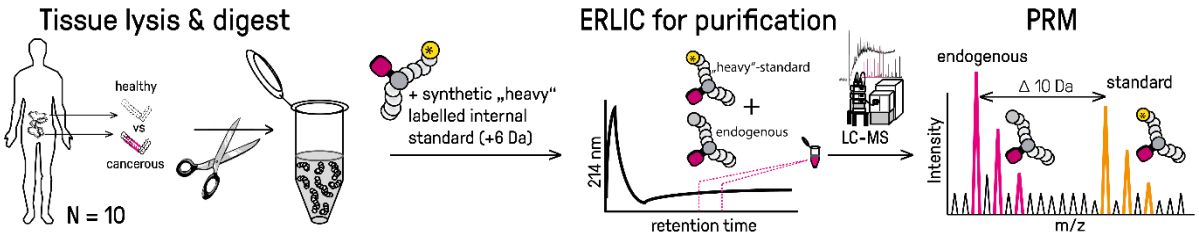


Figure 17: Experiment 1-5: Quantification of a phosphopeptide in colon tissue digests. Healthy and cancerous tissue from 10 patients (20 samples) were digested using trypsin. After spike-in of an isotope-labelled synthetic standard of known concentration, 32 µg of digest were subjected to ERLIC and a 4-minute fraction was collected at the expected retention time. The fraction was analysed by PRM and the absolute amount of the endogenous phosphopeptide was determined via peak area ratios (vs spike-in).

5.7.6. Experiment 2: Time-resolved characterization of the proteome and phosphoproteome of MCMV-infected mouse fibroblasts

A time-resolved characterization of the proteome, phosphoproteome in MCMV-infected mouse fibroblasts was conducted by analysing cells 4, 8 and 24 hpi against an uninfected control (mock or 0 h). To allow a quantification between the four conditions using a SILAC 3-plex approach, a reference sample was generated by pooling equal amounts of digest from all conditions. One aliquot of the reference was mixed with two time points and a second aliquot with the other two time points. Hence, the pooled reference with constant protein levels enabled a comparison of all four conditions. For all three replicates, 120 µg of protein (according to BCA) were used for each SILAC channel.

Digests were processed using a multi-step fractionation strategy, employing ERLIC in conjunction with RP- and SCX-SPE for analysis of the phosphoproteome, and high-pH-RP for analysis of the proteome. The RP- and SCX-SPE steps after ERLIC were designed for recovery of non-phosphorylated peptides. Recovered peptides were subjected to high-pH-RP fractionation (section 5.4.3) to allow a full proteome analysis in parallel without the need for additional sample material (Loroch et al., [84] Figure 18).

LC-MS analysis of the 18 ERLIC fractions was conducted with an LTQ-Orbitrap Velos Pro online coupled to a U3000 HPLC (self-packed columns) using a 190 minutes-acquisition method for the fractions SCX 1-4 and a 125 minutes-acquisition method for the fractions SCX 5-8 and RP 9-18. All 24 high-pH-RP fractions were analysed using an LTQ-Orbitrap Elite online coupled to a U3000 RSLCnano using a 125 minutes-acquisition method. Raw-files were processed as described in section 5.6 using the percolator node. Phosphorylation was included as variable modification for database search of ERLIC fractions. Since the percolator node reported “unhandled exception errors”, for database searches with all raw-files of a replicate, the ERLIC fractions SCX 1-3, SCX 4-8 and RP 9-18 as well as the high-pH-

RP fractions 1-12 and 13-24 were processed as separate packages. Results (PSM lists) were merged in PD for further processing (sections 5.6.3-5.6.5).

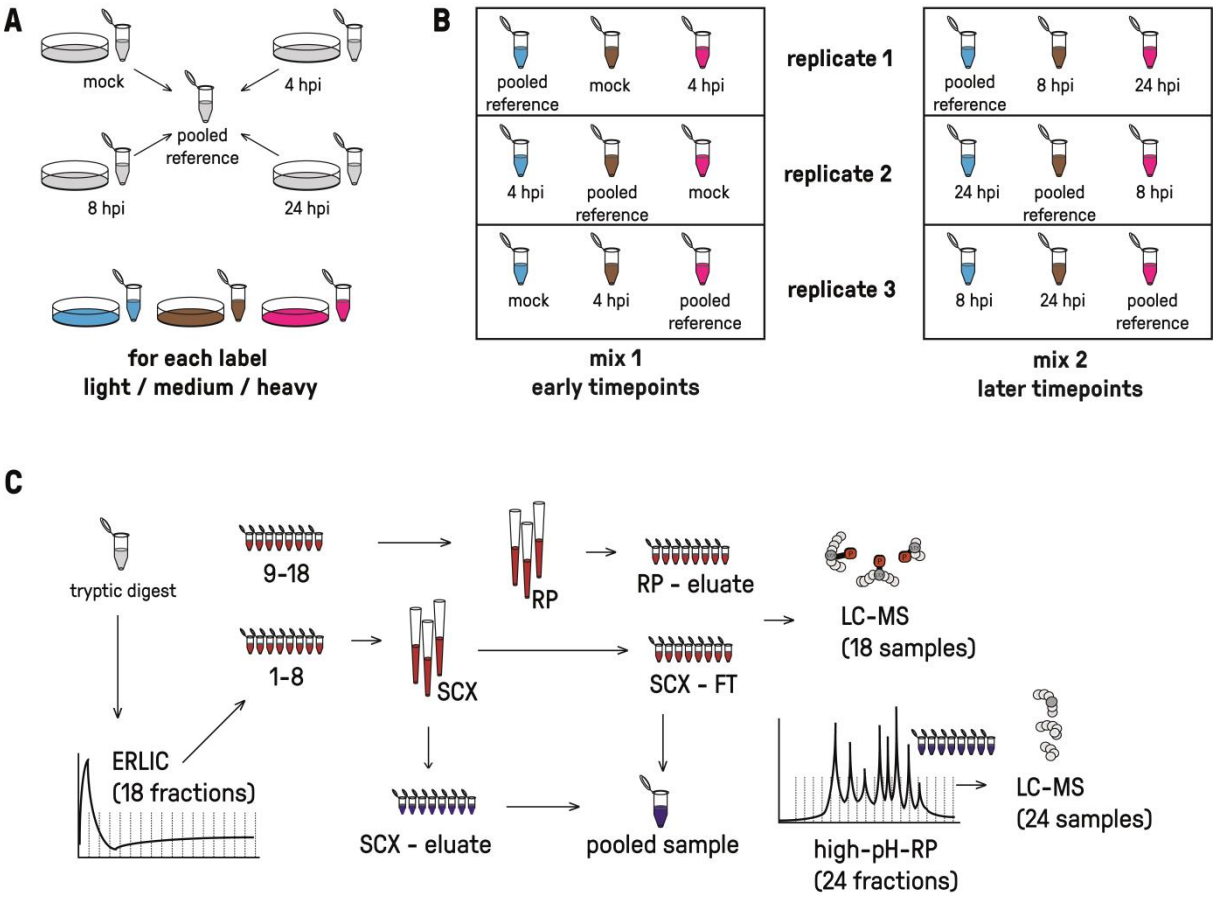


Figure 18: Experiment 2: Workflow for time-resolved characterization of the proteome and phosphoproteome in MCMV-infected mouse fibroblasts. A: Fibroblast cells (NIH-3T3) were cultured in presence of SILAC light, medium or heavy amino acids in 4 dishes in triplicates. 3 dishes were used for MCMV infection (timepoints 4 h, 8 h, 16 h), the fourth dish served as uninfected control (mock). After lysis, a pooled reference sample was generated for each replicate by combining equal protein amounts of every condition. B: The first sample (mix 1) was the pooled reference with mock and 4 hpi. The second sample (mix 2) was the pooled reference with 8 hpi and 24 hpi. C: Digests were subjected to ERLIC for enrichment and fractionation of phosphopeptides. For analysis of the phosphoproteome (red tubes), SCX-SPE flow-through of fractions 1-8 and RP-SPE eluates of fractions 9-18 were subjected to LC-MS. For analysis of the proteome (purple tubes), SCX-SPE salt eluates of the fractions 1-9 and the remaining material of the SCX-SPE flow-through were combined, subjected to high-pH-RP fractionation followed by LC-MS.

6. Results

6.1. Experiment 1-1: ERLIC of trypsin and subtilisin digests

300 µg of a soluble protein fractions were digested with either trypsin or subtilisin for subsequent phosphopeptide enrichment and fractionation using ERLIC. This comparison aimed at (I) evaluation of subtilisin for large-scale phosphoproteomics, (II) evaluation of ERLIC with non-tryptic digests and (III) analysis of complementarity of both proteases with respect to phosphoproteome coverage. The experiment was conducted in technical duplicates with replicate digests of the soluble protein fraction from A431 cells. 12 fractions were collected, starting at RT 2 min, since phosphopeptides are barely detected earlier (as demonstrated by Loroch et al.).[84, 124] 1 min-fractions were collected from RT 2–6 min, 2 min-fractions were collected from RT 6–20 min and a 6 min-fraction was collected from RT 20–26 min. After drying, fractions 1-7 were subjected to SCX-SPE, which allows desalting and further enrichment of phosphopeptides and fractions 8-12 were subjected to RP-SPE for desalting prior LC-MS.

In two replicates, 5,718 ± 49 and 6,259 ± 49 unique non-redundant phosphopeptides were identified with trypsin and subtilisin, respectively (Figure 19 A, all sites ≥ 90 % probability using phosphoRS). Trypsin digests yielded slightly more multiphosphorylated peptides (13.9 % doubly and 0.6 % triply or higher) compared to subtilisin (11.1 % doubly and 0.3 % triply or higher). With trypsin, 11 % more phosphorylation sites were detected (5,476 vs 4,913) and both proteases displayed substantial complementarity with 37 % of all detected sites not detected with trypsin, corresponding to an overlap of just 19 %. With respect to reproducibility between replicates, both proteases performed well with 63 % (trypsin) and 59 % (subtilisin, Figure 19 B).

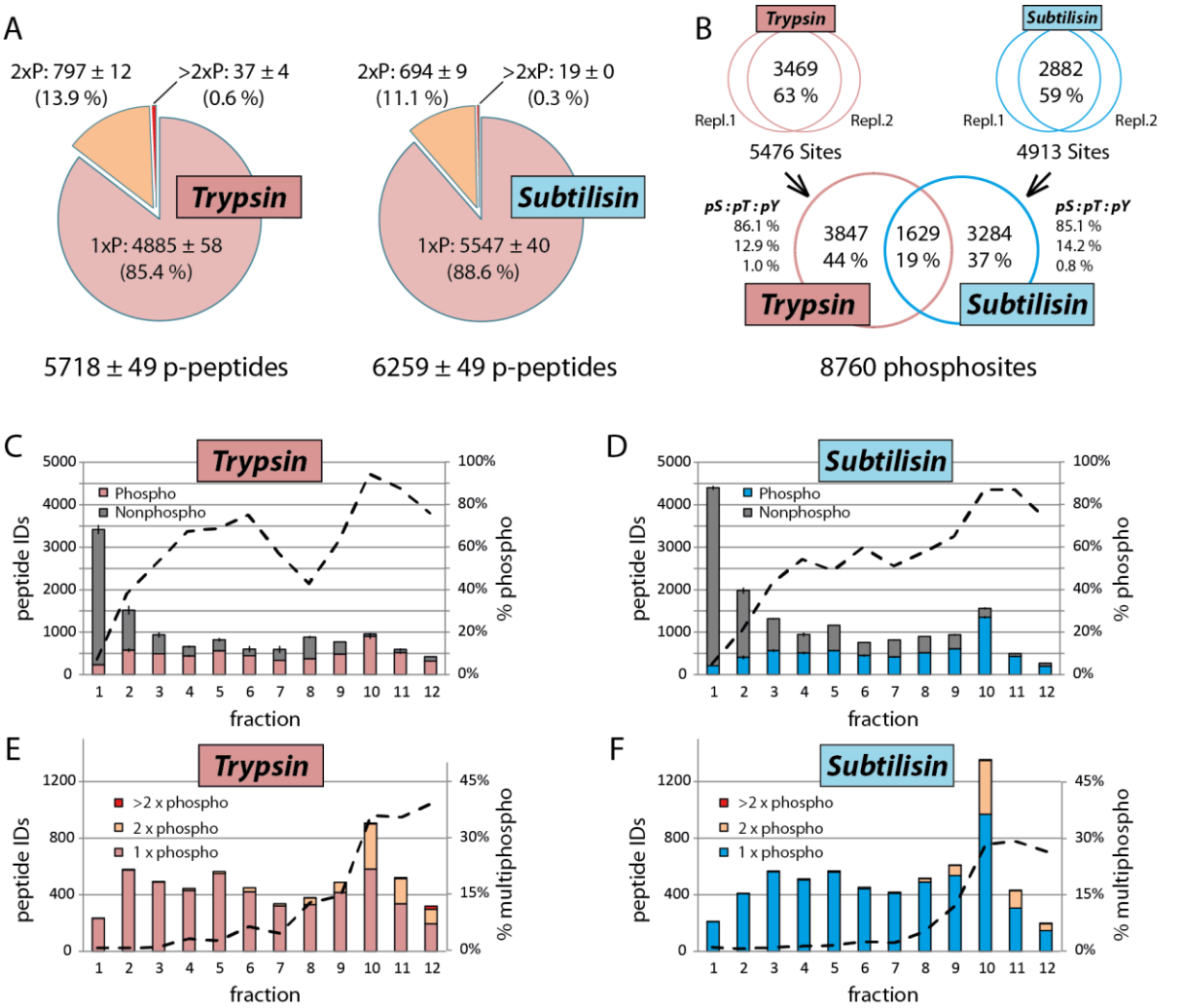


Figure 19: Systematic comparison of ERLIC-based phosphopeptide enrichment and fractionation using trypsin or subtilisin digests. Both proteases performed comparable in terms of phosphopeptide identification with slightly more IDs for subtilisin (A). In contrast, trypsin enabled the detection of 10 % more phosphorylation sites, but 37 % of the total detected sites were not covered by trypsin (B). Phosphopeptide distribution revealed a very similar pattern with enrichment specificities of up to 80 % in the later fractions (C, D). Multiphosphorylated peptides were predominantly identified in the fraction 9-12 (E, F). In case of redundant phosphopeptide identification (over multiple fractions), the phosphopeptide was only counted in the fraction with the top-scored PSM.

Analysis of the distribution of phosphopeptides among ERLIC fractions revealed a very similar enrichment pattern over the chromatographic run with an overall specificity of ~ 50 % (portion of phosphopeptide IDs relative to total IDs) and up to 85 % in the later fractions. In the first two fractions, only minor portions of peptides were phosphorylated (≤ 30 %) and the absolute number of phosphopeptide IDs in fraction 1 was comparably low. In contrast, fractions 2-11 yielded approximately

400-600 phosphopeptide IDs rendering the fractionation scheme as well-chosen. Notably, with the start of the linear gradient of the trypsin run (fraction 8) a larger share of non-modified peptides was detected (~50 %), which was not observed for subtilisin. Besides, a fair enrichment of multiphosphorylated peptides was observed in the fractions 9-12 with a share of 15 % - 40 % (Figure 19C-F).

To explain the complementarity of both proteases in terms of phosphorylation site detection, physicochemical properties of all identified phosphopeptides were analyzed (Figure 20). A systematic comparison revealed that subtilisin phosphopeptides are (I) slightly more hydrophilic, (II) more acidic and (III) shorter with a rather narrow length distribution (Figure 20 A and B). Higher acidity can be partially explained by the theoretical net-charge distribution of phosphopeptides at pH 2 (Figure 20C) as the theoretical net charge considers each basic amino acid (and the N-terminus) with +1. As subtilisin phosphopeptides displayed lower net-charges (e.g. 0 and +1), less basic amino acids are present. In contrast, tryptic phosphopeptides had predominantly higher net charges of +1, +2 and +3, hence, had higher shares of basic amino acids.

Accordingly, analysis of the amino acid frequencies (Figure 21) revealed that subtilisin yielded more phosphopeptides without the basic amino acids K and R. Besides, a larger portion with ≥ 5 D or ≥ 5 E, the two acidic amino acids, was detected rendering subtilisin not only more acidic, but also less basic. The slightly higher hydrophilicity of subtilisin digests can be explained by the higher share of phosphopeptides without any of the hydrophobic amino acids A, F, I, L and V (Figure 20 A). Controversially, less phosphopeptides with high share of the hydrophilic hydroxy-amino acids S and T were detected with subtilisin, rendering hydrophilic motives from hydroxylated amino acids less efficiently cleaved.[200]

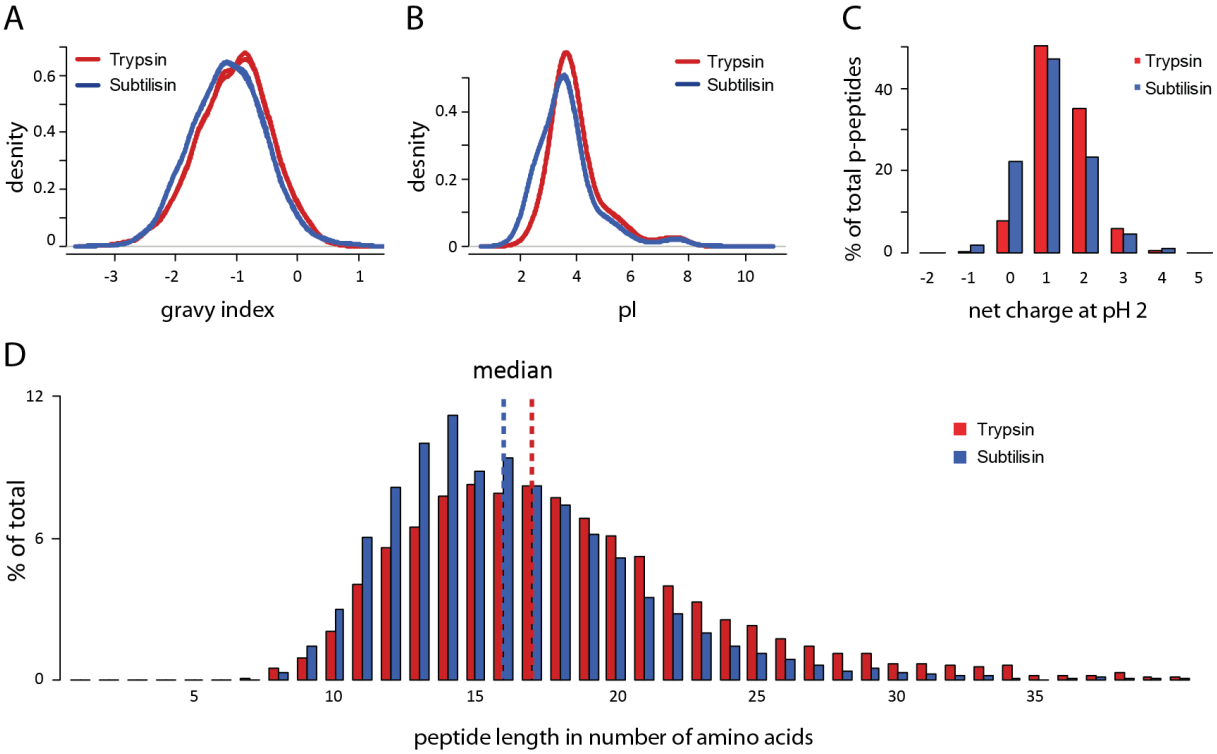


Figure 20: Analysis of physicochemical properties revealed that subtilisin generates phosphopeptides of slightly lower hydrophobicity (lower gravy index, A), lower acidity (pI and net charge, B and C) in comparison to trypsin. In addition, phosphopeptides from subtilisin digests display a narrower length distribution with more phosphopeptides of 9-16 amino acids but less with ≥ 18 (D).

As a major difference, subtilisin seems to generate considerably more phosphopeptides with ≥ 3 proline in the sequence, whereas the tryptic digest had a higher share of phosphopeptides with only 1-2 proline. This considerable difference in proline occurrence might originate from improved proteolytic cleavage within regions of high proline density, which are described to form proline-type helices[201] and might be less accessible for trypsin. To validate this, entries of the Uniprot human database were screened for regions with proline densities (≥ 0.2, roughly accounting for ≥ 25 % proline within 25 amino acids) and phosphopeptides were mapped to these regions. Among 20,194 protein sequences in the human protein database, 6,733 proteins (33 %) exhibited at least one proline-rich region. All phosphopeptides of this dataset and of a corresponding experiment from Humberto Gonczarowska-Jorge[200] (TiO_2-HILIC-based enrichment of subtilisin vs trypsin digests) were mapped to these regions. As assumed, the coverage was nearly 2-fold higher with subtilisin compared to trypsin (36,857 amino acids covered by

subtilisin-phosphopeptides vs 20,484 amino acids by trypsin-phosphopeptides). This result was visualised by plotting all 6,733 proline-rich proteins according to relative position of their proline-rich region. Proline density was indicated as colour gradient and phosphopeptides were indicated as black bars (Figure 22).

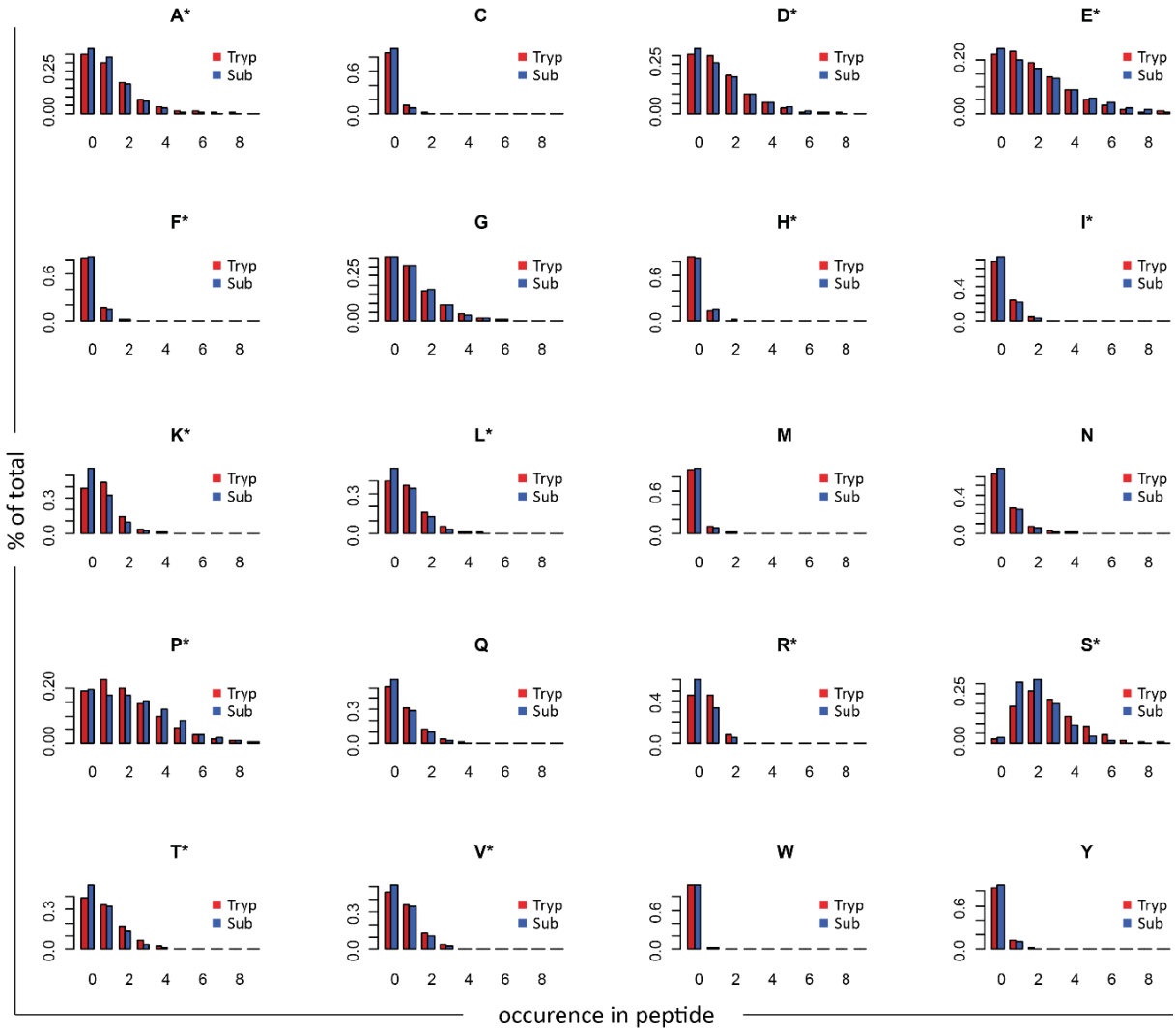


Figure 21: Analysis of amino acid occurence revealed that subtilisin generates more phosphopeptide with ≥ 5 D or E, without any K or R (basic) or A, F, I, L and V (hydrophobic) explaining the lower gravy and the lower pI. In addition, more phosphopeptides with ≥ 3 prolines were detected.*indicates appreciable differences which are further discussed in the text.

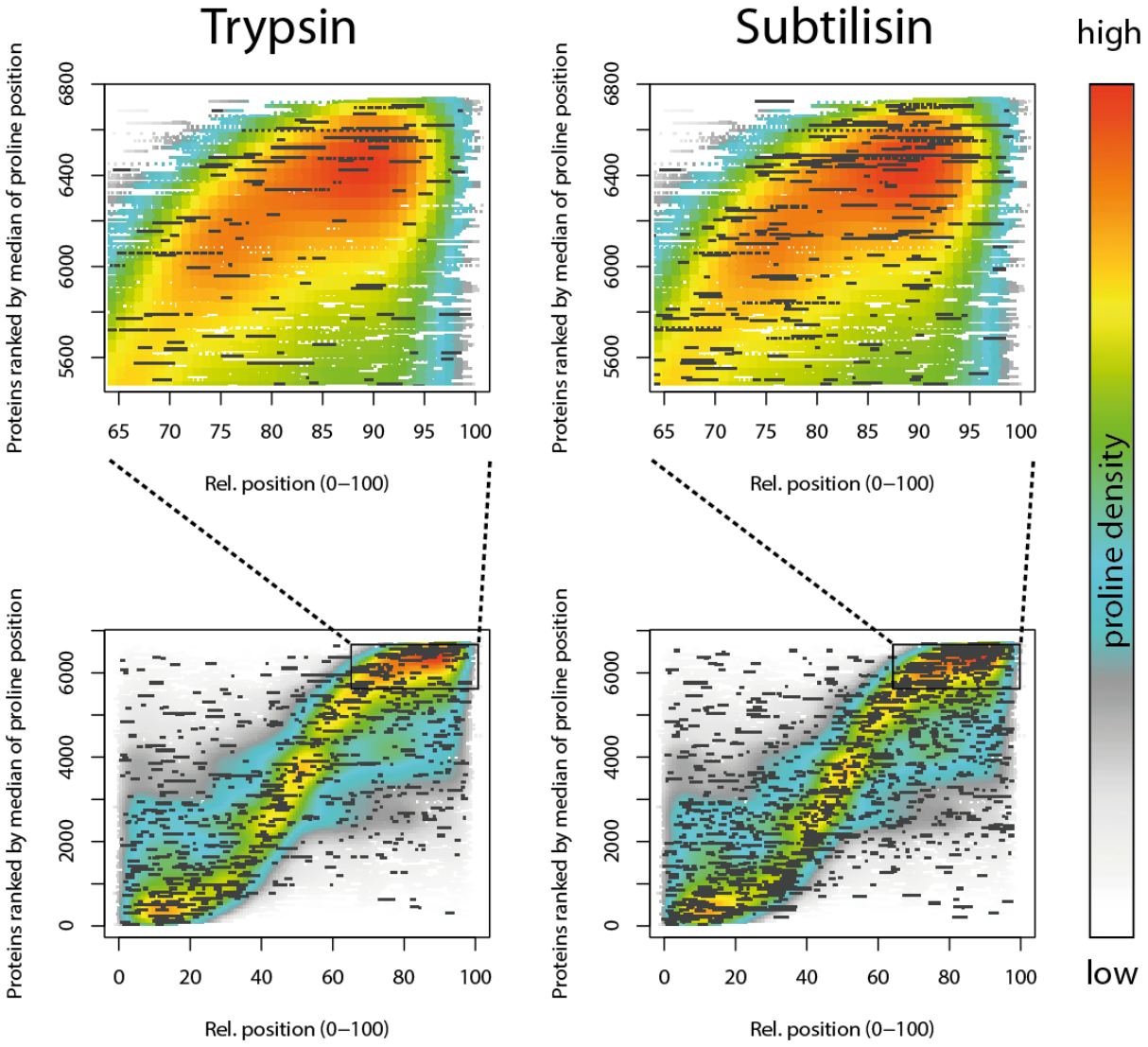


Figure 22: The human proteome comprises 6,733 proteins, which exhibit regions of high proline density (color gradient). Phosphopeptides of a subtilisin digest display a considerably higher coverage within these regions (36,857 amino acids with subtilisin vs 20,484 with trypsin, black bars) explaining the high complementarity. Figure from Gonczarowska-Jorge & Loroch et al..[200]

6.2. Experiment 1-2: Relative quantitative recovery of phosphopeptides in ERLIC and 2-step TiO_2

In this experiment, phosphopeptide-enriched fractions of ERLIC and TiO_2 (differentially labelled) were mixed and subjected to 2D-LC-MS analysis. Heavy-to-light ratios reflect the relative recovery of each phosphopeptide with either methods (ERLIC vs TiO_2). In a duplicate experiment, ratios were assigned to 2,908 ± 87 phosphopeptides of which 94 % were singly phosphorylated. Calculating ratio distributions revealed 80.4 % of all phosphopeptides (2,339 ± 33) to exhibit higher recovery with ERLIC whereas only 19.6 % (569 ± 54) exhibited higher recovery with TiO_2 (Figure 23A). Hence, ERLIC exhibited an improved recovery for the vast majority of all quantified phosphopeptides. The median log2-ratio (TiO_2/ERLIC) was found to be -0.8 corresponding to a 1.74-fold higher recovery ($2^{0.8} = 1.74$) with ERLIC. Interestingly, only singly phosphorylated peptides (94 % of total IDs, Figure 23 C) exhibited an appreciably improved recovery, whereas doubly phosphorylated ones (6 % of total) exhibited only a 1.1-fold improvement (Figure 23 D). Notably, ratios were also assigned to 16,970 ± 327 unphosphorylated peptides which derived almost entirely (~ 99.5 %) from ERLIC (Figure 23). This is attributed to the experimental design, which required the collection of a single ERLIC fraction (RT 3-20 min) to yield a maximum of phosphopeptide IDs at the expense of collecting co-eluting unphosphorylated ones.

Notably, log2-ratio distributions were plotted for both replicates separately (Figure 23) demonstrating excellent correlation of both replicates. As a label-switch replicate was conducted, findings cannot derive from a systematic error between the SILAC channels, but ratios in deed reflect true differences between methods. In addition, the low variance of total IDs (<< 10 %) demonstrates that the entire workflow including high-pH-RP-fractionation was robust and is suitable for future experiments.

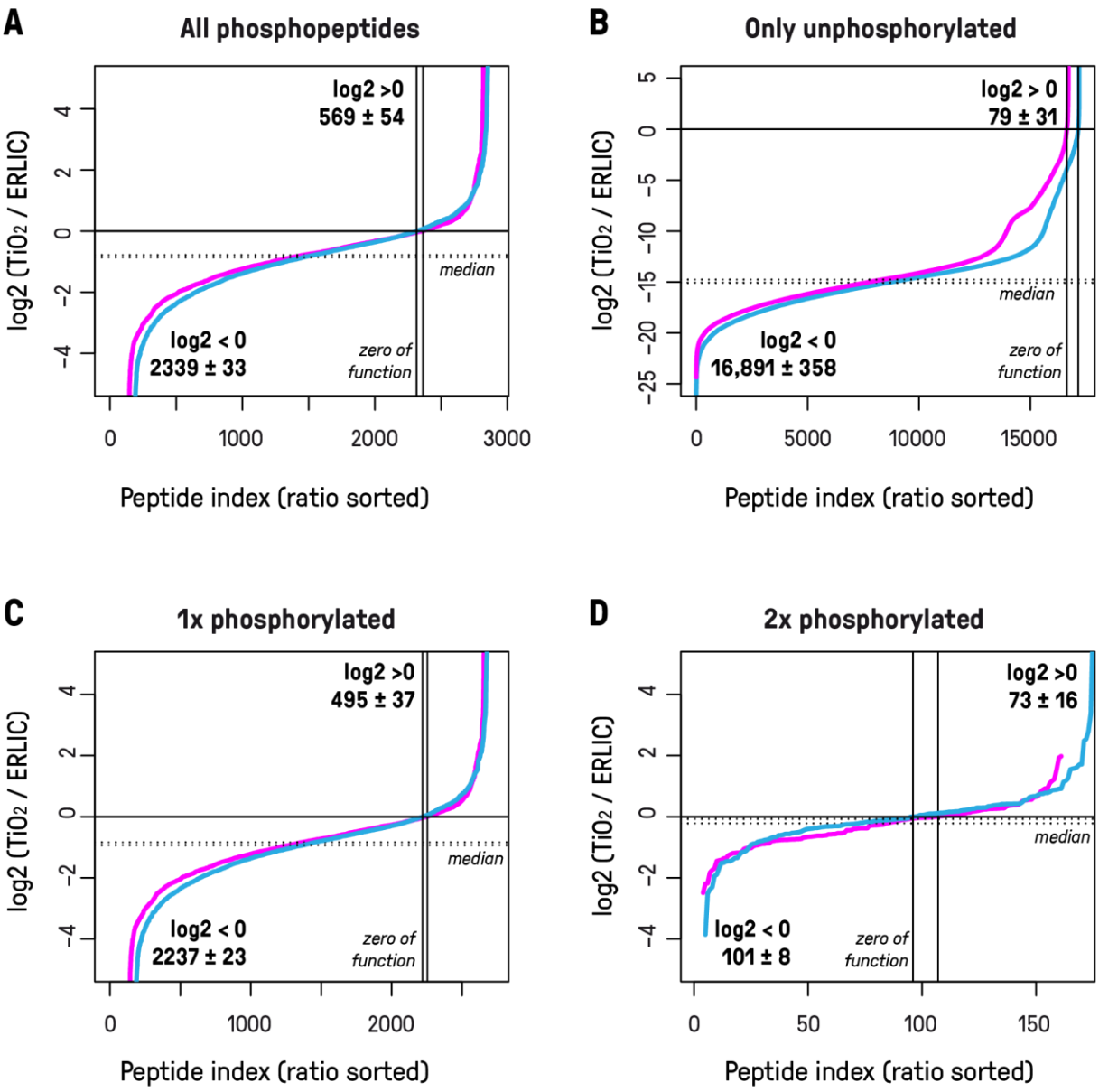


Figure 23: Relative comparison of quantitative recovery in ERLIC and TiO_2 using a 2-plex SILAC strategy. Log2-ratios of y-axes reflect whether a phosphopeptide was better recovered by ERLIC (< 0) or by TiO_2 (> 0). In a duplicate experiment, ratios were assigned to 2,908 ± 21 phosphopeptides of which 2,339 ± 33 (80.4 %) had a higher recovery with ERLIC. The median phosphopeptide recovery (dashed line) was 1.74-fold higher using ERLIC (A). According to the experimental design (see text), 99.5 % of the 16,970 ± 327 unphosphorylated peptides had ratios < 0, hence, derived almost exclusively from co-enrichment in ERLIC (B). Notably, improved recovery of phosphopeptides with ERLIC is well pronounced for singly (C), but not for doubly phosphorylated peptides (D).

6.3. Experiment 1-3: Quantitative loss during enrichment – a comparison of ERLIC and TiO_2

In a follow-up experiment, the quantitative loss during one round of enrichment was investigated. A phosphopeptide-enriched sample (e.g. light) was mixed with a corresponding non-enriched digest with a different label (e.g. heavy). Heavy-to-light ratios directly reflected the quantitative loss of each phosphopeptide in one round of enrichment. Before LC-MS, a second round of enrichment was conducted to remove the complex background of unphosphorylated peptides to avoid interference in LC-MS. In case of ERLIC, 1,372 ± 94 quantified phosphopeptides revealed a median recovery of 61 %, whereas TiO_2 exhibited a median recovery of only 39 % as determined from 2,136 ± 40 phosphopeptide ratios (Figure 24). This corresponds to an improved recovery with ERLIC of 1.6-fold (61% / 39 % = 1.6) and is in good accordance with the previous experiment 1-2 (section 6.2), which determined the recovery to be 1.7-fold higher.

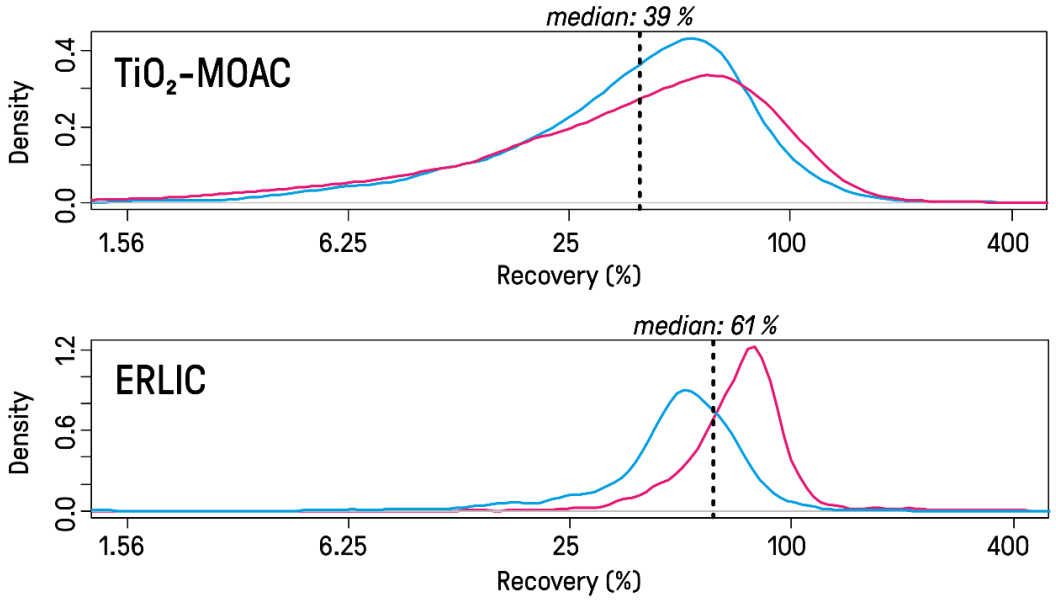


Figure 24: Kernel density estimation of phosphopeptide recovery during enrichment (duplicate experiment). In TiO_2, the median recovery was found to be 39 % with a broad, left-tailing distribution indicating recovery to be a function of phosphopeptide properties. In contrast, phosphopeptide recovery in ERLIC shows a narrow, Gaussian-like distribution around the median of 62 % indicating recovery to be independent of physicochemical properties.

In case of ERLIC, the density plot of phosphopeptide ratios displays a narrow gaussian-like distribution around the median. Only ~5 % of all phosphopeptides exhibited a recovery below 25 %. This indicates stable recovery of any phosphopeptide and the normal distribution most likely reflects the typical variance around the specific recovery rate. In contrast, TiO_2 exhibited a broad, left-tailed distribution and ~30 % of all phosphopeptides exhibited a poor recovery below 25 %. Most interestingly, the pronounced left-tailing of the distribution indicates that recovery might depend on physicochemical phosphopeptide properties.

A closer look on physicochemical properties in TiO_2 revealed that recovery of singly phosphorylated peptides is much lower (median of 36 %) compared to the recovery of doubly phosphorylated ones (76 %), which is in god accordance with the findings of experiment 1-2 (section 6.2). Further examination revealed a positive correlation of recovery with acidity and length (Figure 25), in other words, shorter and less acidic peptides were poorly recovered by TiO_2 (<< 25 %). In ERLIC, no notable difference in recovery was observed for singly and doubly phosphorylated peptides (63 % and 61 %, respectively) as well as no correlation with length or acidity (Figure 25).

In a next step, occurrence of all 20 proteinogenic amino acids in each phosphopeptide sequences were calculated (Figure 26, Figure 27) and expressed as percentage of length in amino acids. In case of TiO_2, scatter-plots display an asymmetric distribution for various amino acids (such as D, E, Q, K) similar as observed for the overall ratio distribution (Figure 24). Pearson product-moment correlation coefficients (R^2) revealed a positive correlation of recovery with the occurrence of Q, E, N, D, C, S, G (R^2 of 0.07 – 0.22) and negative correlation with L, F, R, V, P, K, I, Y, T (R^2 of -0.21 – -0.08). Hence, phosphopeptide recovery rates appear to critically depend on amino acid composition. In contrast, ERLIC, phosphopeptide recovery appears to be independent of amino acid occurrence. R^2 were close to zero (within -0.1 and +0.1) and data points (scatter plots) appear symmetrically distributed around the median (Figure 27).

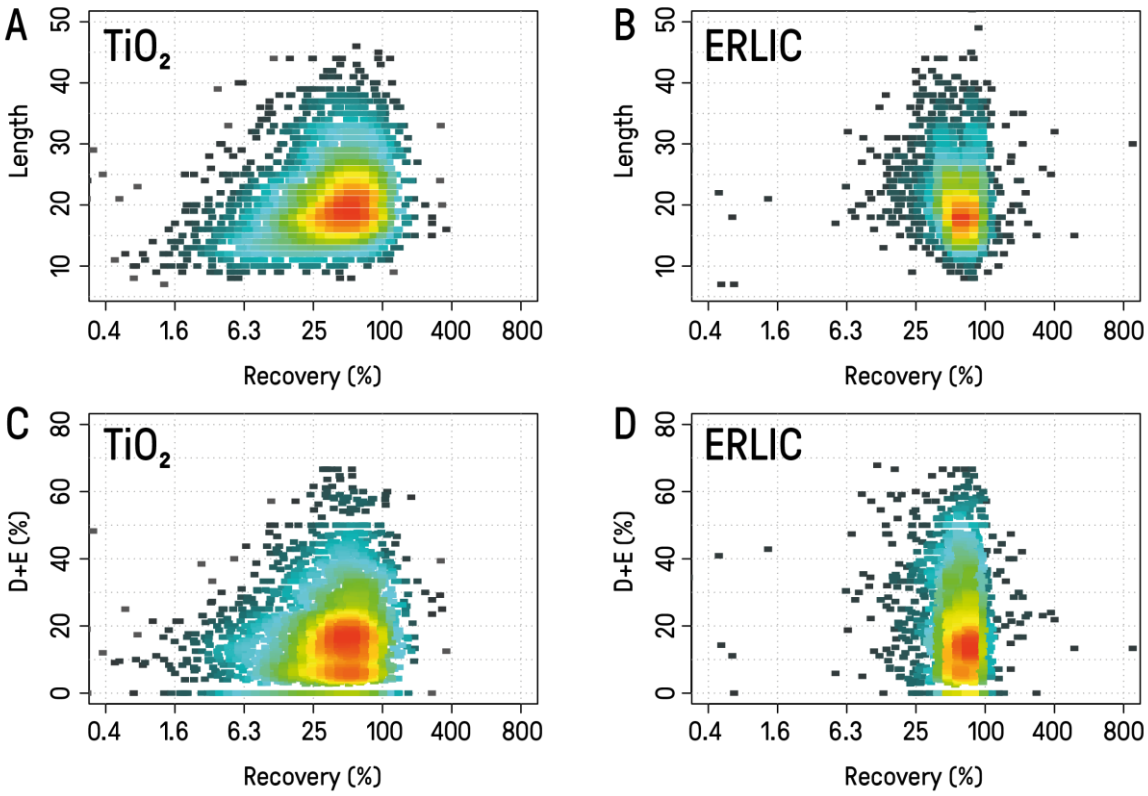


Figure 25: Recovery of phosphopeptides depends on phosphopeptide properties such as length (A, B) and share of acidic amino acids (C, D). With TiO_2, short and less acidic phosphopeptides exhibit poor quantitative recovery (median 39 %, A, C). In ERLIC (B, D), >95 % of phosphopeptides show at least 25 % recovery and data points are gaussian-like distributed around the median (62 %). Colour gradients indicate data point density.

As recovery in TiO_2 appeared to depend on peptide length and occurrence of 13 amino acids (I, K, P, V, F, R, L, Q, E, N, D, C, S), 3D-models were built to estimate recovery as a function of length and amino acid composition. For each phosphopeptides, the sum of all negatively correlated amino acids (I, K, P, V, F, R, L) was subtracted from the sum of all positively correlated amino acids (Q, E, N, D, C, S) resulting in *amino acid composition values (AACVs)*, which were mostly within -10 to +10. The median quantitative recovery was calculated for each possible combination of AACVs within -10 and +15 and length within 9 and 30 amino acids (low values were set to -4; missing values were replaced by the minimum value in the matrix). The resulting matrix was smoothed (theta = 2) using (image.smooth function of package "fields") and visualized as a 3D-plot (Figure 28). As expected, TiO_2 displayed poor recovery for short peptides (< 13 amino acids) with low AACV (< -5), but recovery improved slightly

with increasing length. Optimal recovery was found for all phosphopeptides of AACVs > 5. As phosphopeptide length did not influence recovery of phosphopeptides with AACVs > 5, AACVs appeared to be of higher impact. Most interestingly, recovery in ERLIC did not seem to correlate with AACV or length as recovery rates for all AACV-length combinations were ~61 % (Figure 28).

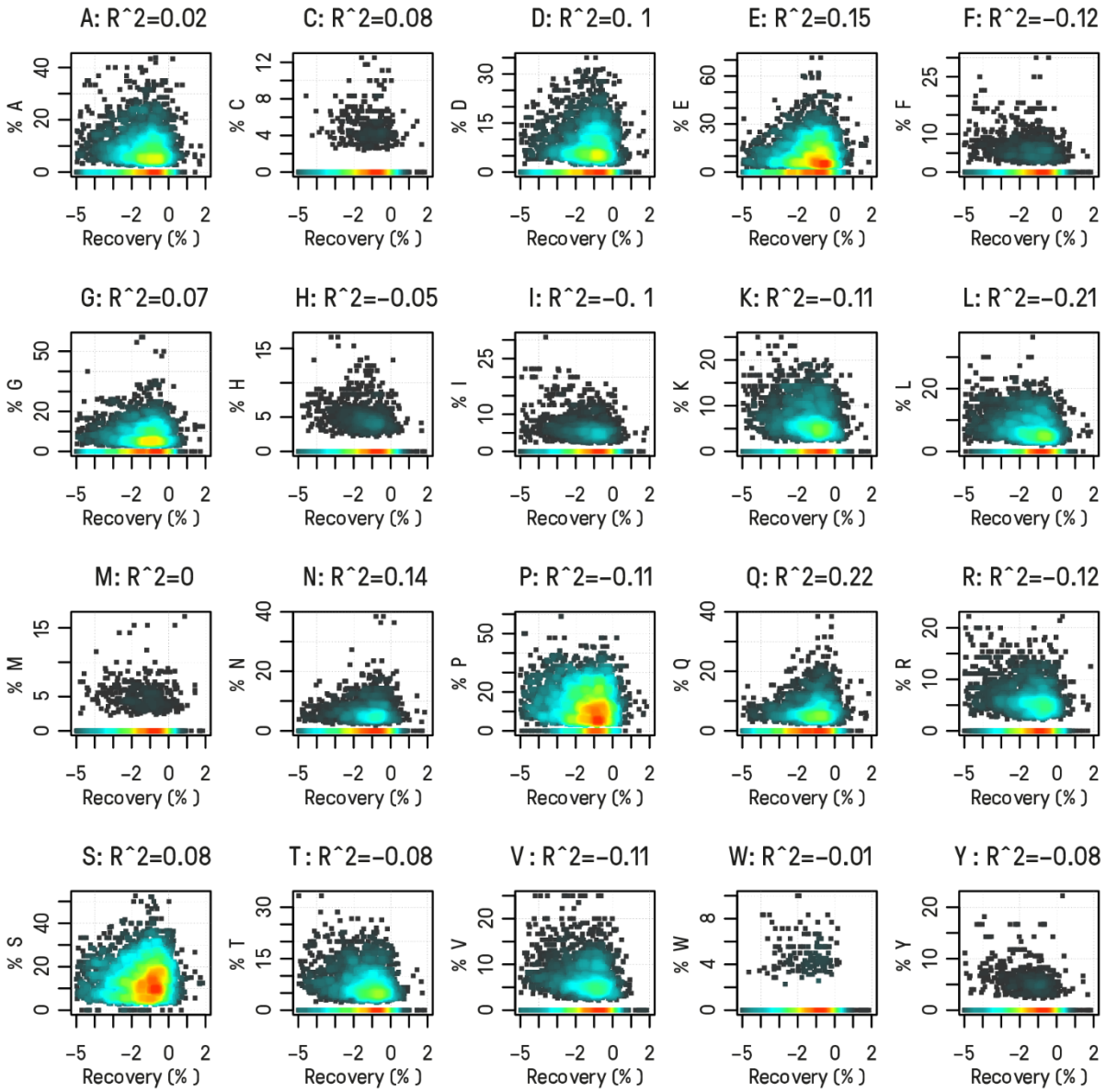


Figure 26: Phosphopeptide recovery rates of TiO_2 plotted against the share of each amino acid including Pearson product moment correlation (R^2 in caption). A positive correlation was found for the amino acids Q, E, N, D, C, S, G (R^2 within +0.07/+0.22), whereas the amino acids L, F, R, V, P, K, I, Y, T exhibit a negative one (coefficient within -0.08/-0.21); colour gradients (gray-to-red) give density of data points.

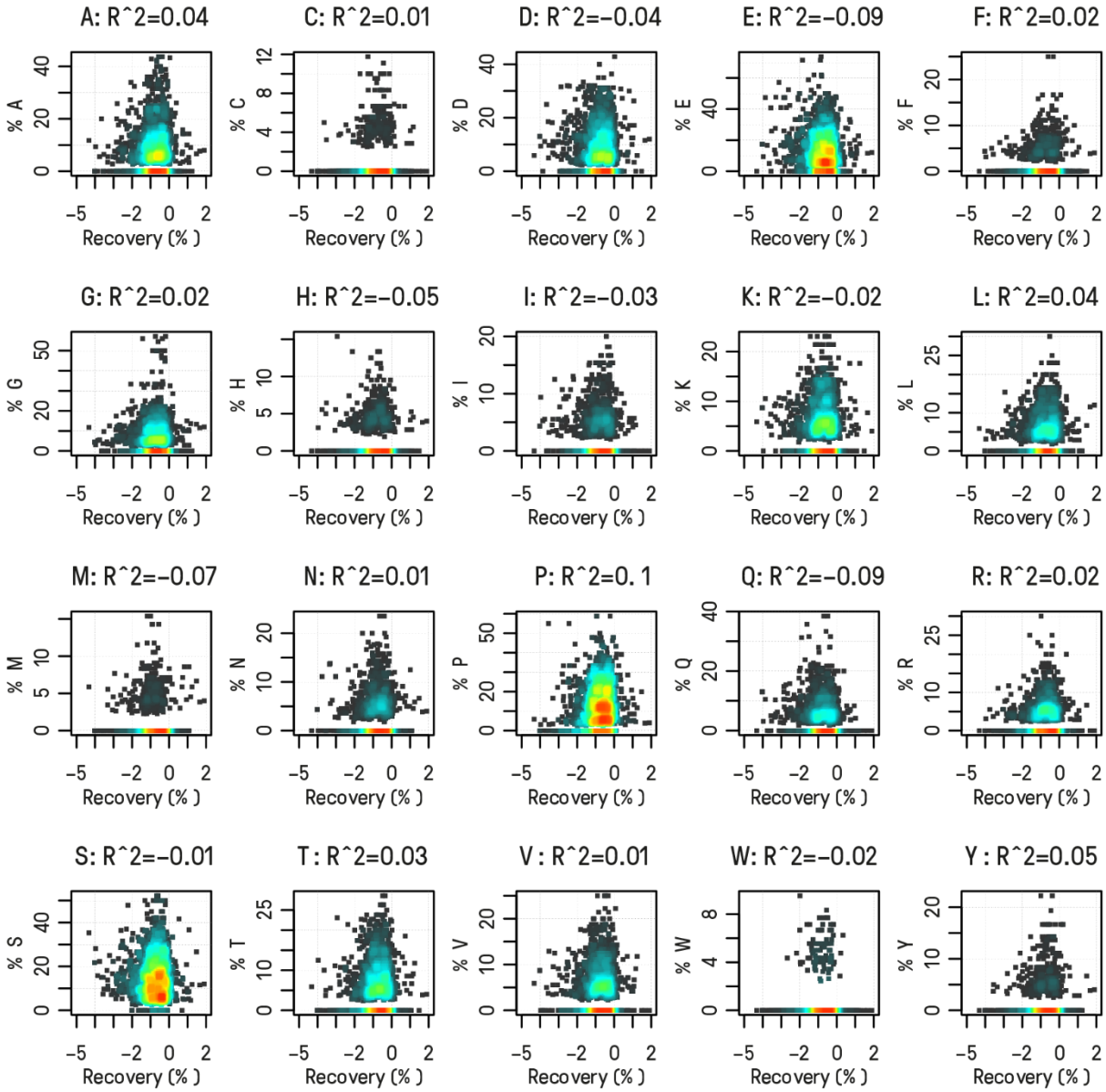


Figure 27: Phosphopeptide recovery rates of ERLIC plotted against the share of each amino acid including Pearson product moment correlation (R^2 in caption). Correlation coefficients (R^2 in caption) are distributed around 0 within the range of -0.1/+0.1 indicating no correlation; colour gradients (gray-to-red) give density of data points.

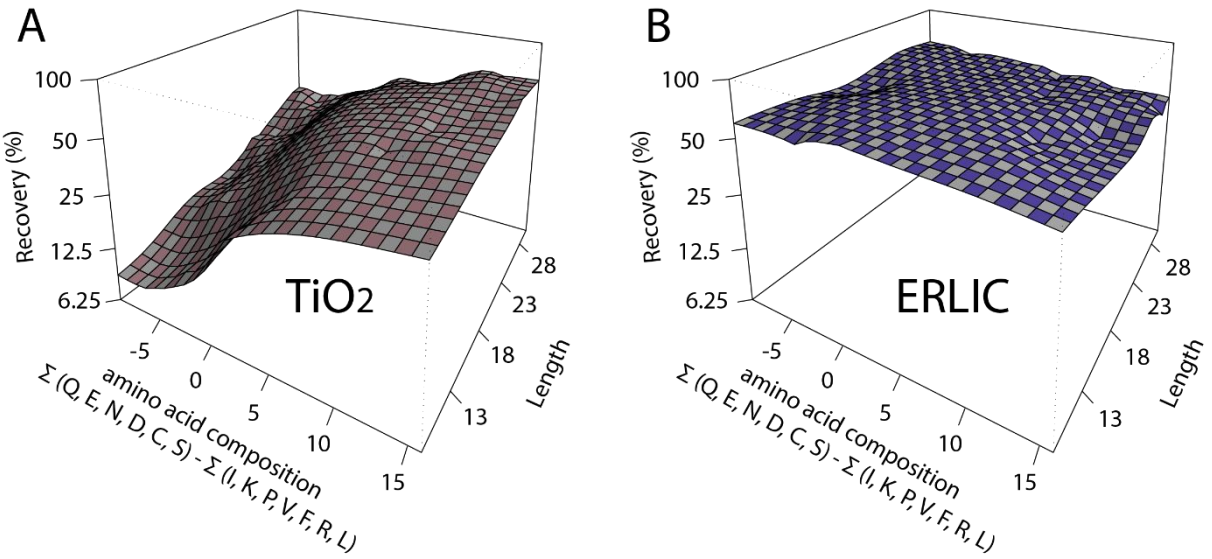


Figure 28: 3D-model visualizing the median recovery of phosphopeptide enrichment methods in dependence of physicochemical properties. Whereas median recovery rates of TiO_2 are below 25% for peptides containing mostly I, K, P, V, F, R, L and few Q, E, N, D, C, S, recovery rates in ERLIC appear stable (median 61 %) irrespectively of length or amino acid composition.

In summary, ERLIC exhibited 1.6-fold higher phosphopeptide recovery than TiO_2, which is in good accordance with the previous experiment 1-2 (1.7-fold, section 6.2). In addition, ERLIC does not show any correlation between amino acid composition, phosphorylation state, length or acidity, hence, a phosphopeptide-enriched sample appears to maintain its initial phosphopeptide stoichiometry. In contrast, TiO_2 considerably alters the phosphopeptide stoichiometry since short phosphopeptides and peptides of a low AACV are only poorly recovered.

6.4. Experiment 1-4: Benchmarking ERLIC and TiO_2 for quantitative phosphoproteomics using minute amounts of sample

In the next experiment, ERLIC and TiO_2 were evaluated for applicability in low sample amount phosphoproteomics and the experimental design aimed at a fair comparison with respect to strengths and weaknesses of either method. The experiment was conducted with a SILAC 2-plex sample of 25 µg MCMV-infected fibroblasts (SILAC heavy) against uninfected fibroblasts (SILAC light). The sample was split in six identical aliquots and triplicates were performed with either method. These identical

aliquots were chosen in order to reduce sample variability to a minimum. Therefore, reproducibility of both methods could be addressed allowing to evaluate the power for detecting differentially regulated phosphopeptides. For a fair comparison, both methods were optimized to yield maximum phosphopeptide IDs with the least quantitative loss possible: (I) LC-MS gradients and gradient lengths were adjusted in a way that avoids undersampling (complexity of analytes excels the scan rate of the MS) and (II) TiO_2 samples were not further pre-fractionated, which would have led to additional loss. Hence, TiO_2-samples were directly analyzed using a 200 min LC-gradient, whereas the 10 ERLIC fractions were analyzed with 30 min gradients each – resulting in 300 min total MS-acquisition time. Hence, ERLIC fraction acquisition time was 50 % higher (counting just MS-acquisition time), but 250 % higher when considering the total required LC-MS instrument time (including wash blanks and sample loading).

Using TiO_2, 2,645 ± 59 phosphopeptides were identified and ratios were assigned to 2,463 ± 55 phosphopeptides with 2,276 ± 72 above a reasonable signal intensity threshold for quantification (500,000 counts for heavy + light). Using ERLIC, 3,277 ± 205 phosphopeptides were identified and ratios assigned for 2,991 ± 160 with 2623 ± 160 above the threshold. With ERLIC, 1,749 phosphopeptides were reproducibly found in all replicates in comparison to 1,494 using TiO_2. Hence, ERLIC enabled the detection of 15-23 % more phosphopeptides than TiO_2 (Figure 29 A, B). Notably, the overlap between methods was only 35 %, though the overlap of replicates was ~60 % for either method (57 % for ERLIC and 62 % TiO_2) indicating considerable complementarity.

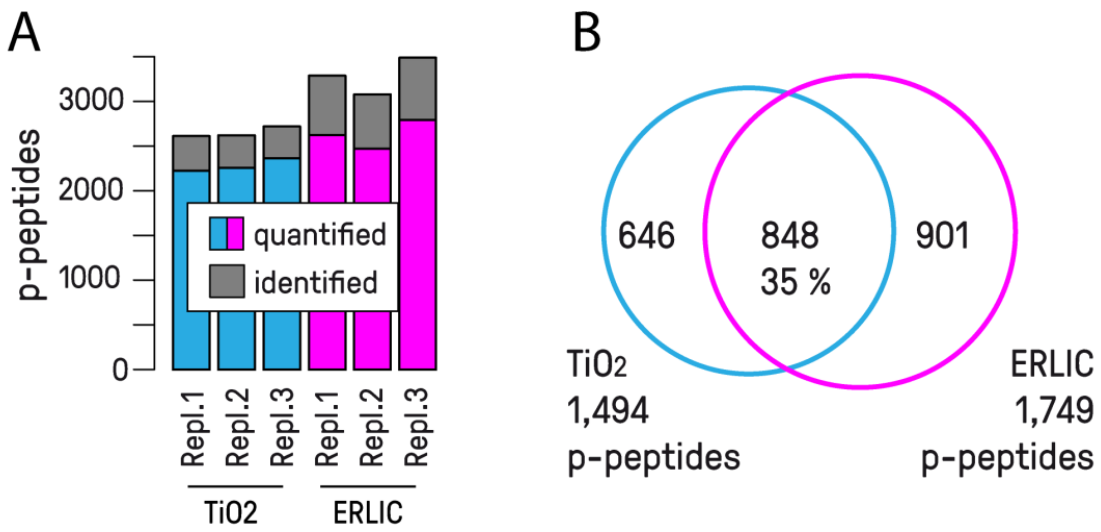


Figure 29: Phosphopeptides detected in a SILAC 2-plex experiment using 25 µg MCMV-infected fibroblasts against a non-infected control. Using ERLIC, ~20 % more phosphopeptides were identified and quantified in each replicate (A) and 17 % more phosphopeptides were reproducibly found in all three replicates (B). Notably, only 35 % off all reproducibly found phosphopeptides were detected with either method (B).

Phosphopeptide heavy-to-light ratios were calculated for each replicate separately to allow for a comparison between replicates. Pearson product-moment correlation coefficients were R^2 = 0.929 ± 0.006 (for TiO_2) and R^2 = 0.936 ± 0.002 (ERLIC) indicating excellent reproducibility of quantification with either method (Figure 30). However, reproducibility usually improves when using higher signal intensity thresholds (here 500,000 counts), hence, correlation coefficients usually depend on the number of included features (here: phosphopeptides).[84] To get deeper insights into this dependency, correlation coefficients were calculated as a function of the number of included phosphopeptides ranked by intensity. Figure 31 shows that R^2 = 0.98 is possible with either method when only considering the top 500 (most intense) phosphopeptides. However, to meet an R^2 of ≥ 0.90, TiO_2 allowed to include the top ~1,600 phosphopeptides in the quantification, but ERLIC allowed to include 20 % more, namely ~1,900. Hence, ERLIC enabled the robust quantification of ~300 additional phosphopeptides at a comparable "robustness"-level.

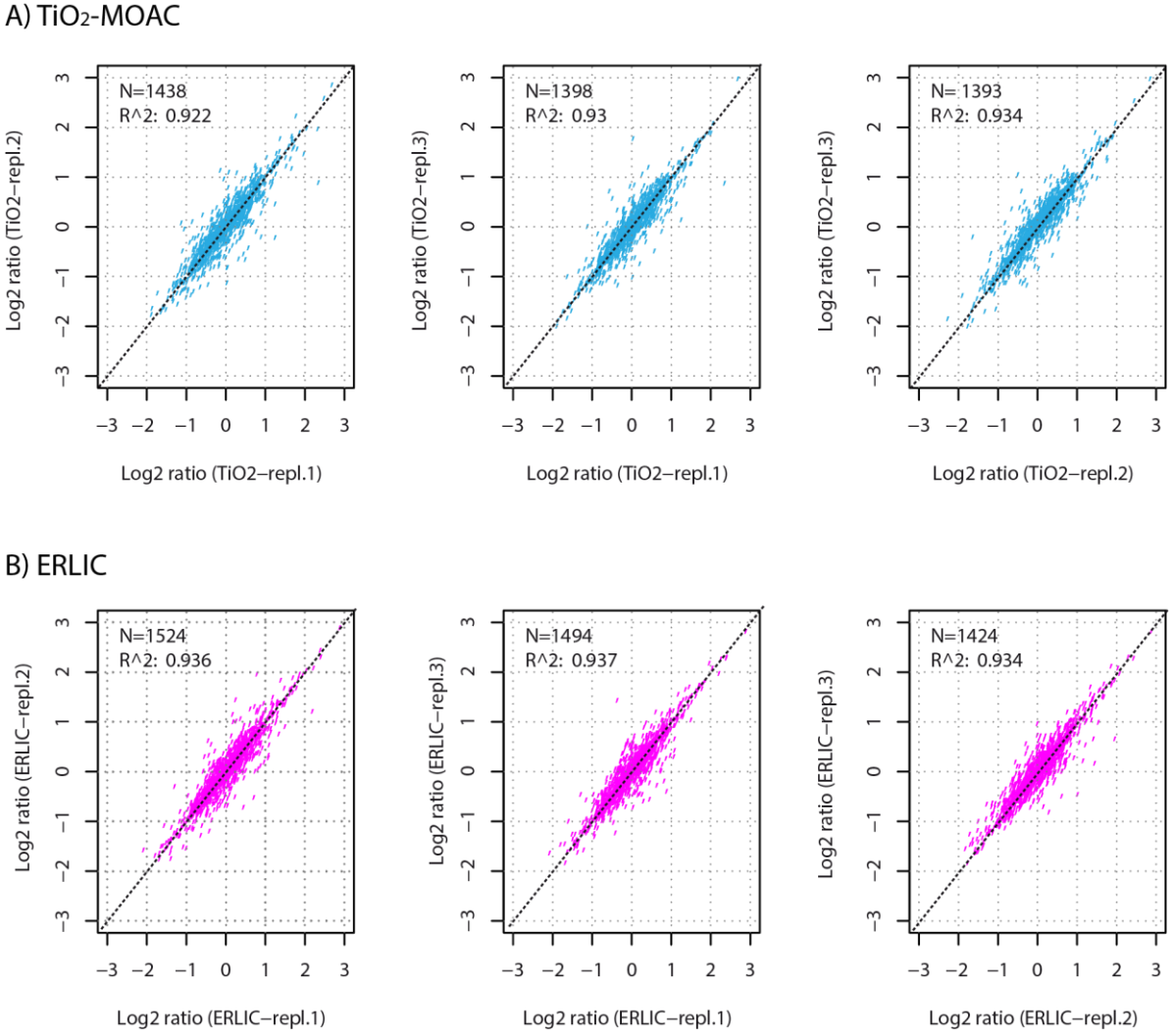


Figure 30: Reproducibility of quantification using TiO_2 and ERLIC for enrichment of phosphopeptides from a 2-plex SILAC sample of 25 µg MCMV-infected fibroblasts against a non-infected control. Both methods demonstrated excellent reproducibility with Pearson product-moment correlation coefficients (R^2) around 0.93 for all replicate comparisons.

To identify differentially regulated phosphopeptides, heavy-to-light ratios were log2-transformed for a two-sample Student's t-test. Resulting p-values were corrected for multiple testing using the procedure of Benjamini-Hochberg to adjust the false-discovery rate to 5 %. Interestingly, TiO_2 enabled identification of 10 % more differentially regulated phosphopeptides compared to ERLIC using the uncorrected Student's t-test ($p \leq 0.05$, Figure 32A). In contrast, ERLIC enabled the identification of 80 % more differentially regulated phosphopeptides after Benjamini-Hochberg correction (36 vs 20) and even 82 % (31 vs 17) when applying a fold-change cut-off of 1.5. As Benjamini-Hochberg

correction aims at a reduction of false-positives, ERLIC appears to identify more highly significant phosphopeptides (by means of p-values, Figure 32 B) whereas TiO_2 identifies more regulated phosphopeptides of lower significance ("maybe-candidates").

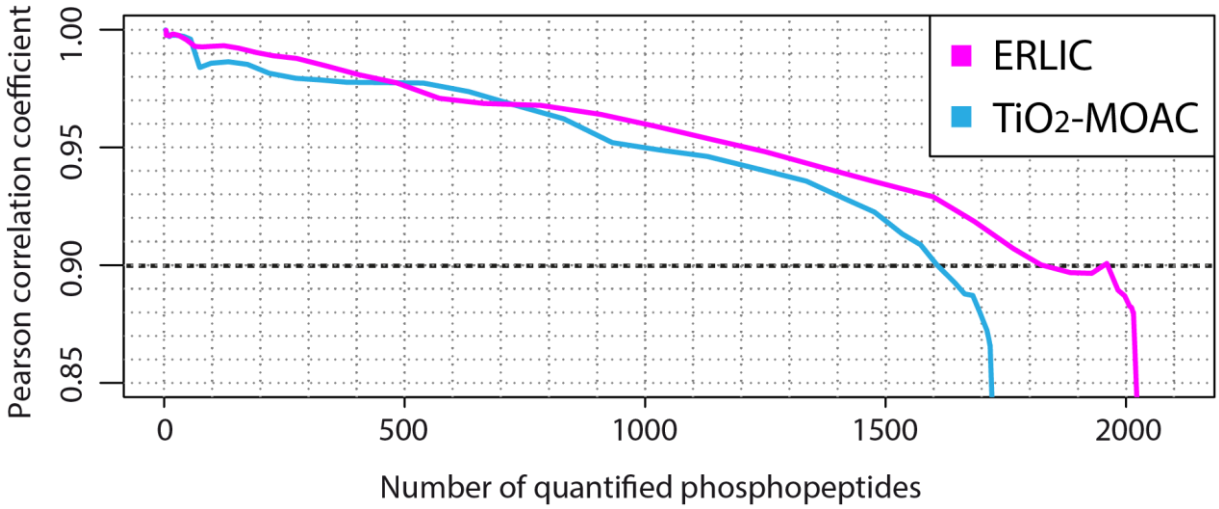


Figure 31: Pearson correlation coefficient as a function of the number of phosphopeptides included for quantification (intensity ranked). At a correlation of $R^2 = 0.90$ (dashed line), ERLIC enables quantification of additional ~300 phosphopeptides compared to TiO_2. Correlation drastically drops (<< 0.9) for both methods when including too many phosphopeptides of low intensity.

The relation of the number of regulated phosphopeptides and p-value threshold (significance level α), was further determined by calculating the numbers of differentially regulated phosphopeptides at different p-value threshold (Figure 33). Using a simple two-sample Student's t-test, ERLIC enabled the identification of more differentially regulated phosphopeptides at very low thresholds (p-value ≤ 0.005, Figure 33A). Since these are the phosphopeptides, which pass the Benjamini-Hochberg correction, ERLIC appeared more beneficial when aiming at avoiding false-positives (Figure 33B).

Subsequent comparison of differentially regulated phosphopeptides revealed a huge complementarity of both methods with only 1 overlapping out of 55 phosphopeptides. Most interestingly, the 36 regulated phosphopeptides of ERLIC had an AACV of 0.25 ± 3.4 with an average length of 15.8 ± 5.3 amino acids whereas the 20 phosphopeptides of TiO_2 had an AACV of 2.15 ± 4.2 and an average length of 17.8 ± 5.3 amino acids. This result is good accordance with the observation from the previous

experiment 1-3 (section 6.3), which demonstrated that TiO_2 exhibits improved recovery for longer phosphopeptides and phosphopeptides of high AACV. In contrast, the unbiased recovery in ERLIC allows for a reliable identification of shorter phosphopeptides with lower AACV, which leads to extremely complementary results.

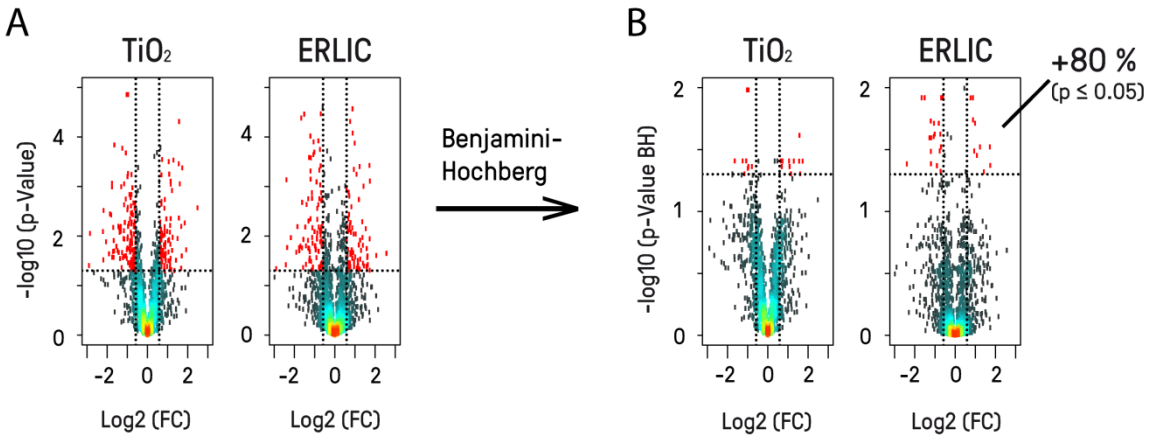


Figure 32: Analysis of differentially regulated phosphopeptides using a SILAC 2-plex sample of 25 µg MCMV-infected fibroblasts against a non-infected control in triplicates. Using an uncorrected two-sample t-test with a 5 % p-value threshold, both methods led to comparable numbers of DRFs (A). After controlling the FDR using the Benjamini-Hochberg procedure, ERLIC led to 80 % additional DRFs (36 vs 20). Notably, only one differentially regulated phosphopeptide was shared between methods (B).

In summary, ERLIC enabled (I) identification and quantification of more phosphopeptides and (II) detection of more differentially regulated phosphopeptides when controlling the FDR (Benjamini-Hochberg-corrected p-values of Student's t-test). For instance, at an FDR of 5 %, ERLIC detected ~80 % more regulated phosphopeptides. However, improved quantification is only possible at the expense of analysis time, namely 50 % more MS-acquisition time and 250% total instrument time. Most interestingly, both methods exhibited tremendous complementarity with respect to quantified phopshopeptides (35 %) and differentially regulated phosphopeptides (1 out of 55) as length and AACVs of phosphopeptides differ with methods.

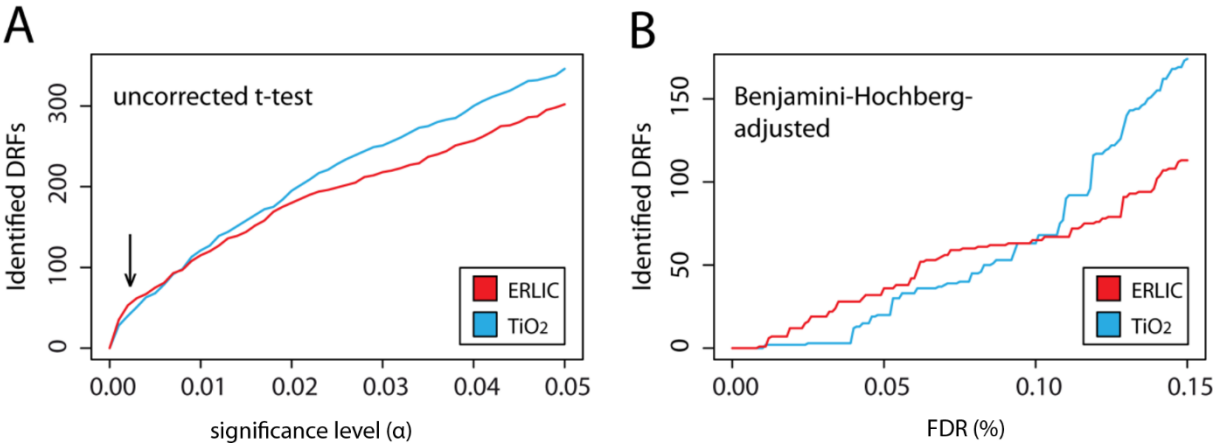


Figure 33: Number of differentially regulated phosphopeptides (irrespectively of fold-change) as a function of significance level (p-value threshold). With ERLIC, more phosphopeptides of low p-value are found (arrow) using an uncorrected two-sample t-test (A) and hence, more phosphopeptides are found at an FDR ≤ 0.1 using the method of Benjamini and Hochberg (B).

6.5. Experiment 1-5: ERLIC in conjunction with targeted LC-MS to monitor aberrant phosphorylation in cancer patient tissue

In this experiment, an ERLIC-based workflow was set up for quantification of the S125 phosphorylation level of the hydroxylase EGLN1 (syn. PHD2) in colon tissue samples from 10 colon cancer patients (10 healthy vs 10 cancerous). The SIL reference phosphopeptide of known concentration was used (I) to determine retention time of the phosphopeptide in ERLIC and (II) as internal standards for quantification of the endogenous phosphopeptide in the tissue digests. Digests were subjected to ERLIC to collect a 4-min fraction at the specific retention time of the phosphopeptide and fractions were analyzed by PRM. In addition, the unphosphorylated counterpart of the SIL reference peptide was used for quantification of the non-phosphorylated version directly in the digest.

To determine the retention time of the phosphopeptide in ERLIC, 130 pmol of the heavy synthetic phosphopeptide AKPPADPAAAAsPC*R (small letter indicates phosphorylated amino acid, * indicates carbamidomethylated cysteine) were injected, which resulted in a clear UV-signal at RT 14 min (Figure

34 A). To purify the phosphopeptide from the tissue samples, 32 µg of digest including 780 amol of the SIL peptide (as internal standard) were subjected to ERLIC and a 4-minute fraction was collected at the specific retention time (12-16 min, Figure 34 B, dashed vertical lines). Inspection of the UV traces revealed that the vast majority of peptides eluted within the first 7 minutes (Figure 34B) emphasizing a substantial purification of the phosphopeptide. Indeed, PRM with 50 % of the ERLIC fraction (corresponding to 16 µg and 390 amol of the internal standard) resulted in appreciable detector response of ~10^6 counts (SIL) and ~$3*10^5$ counts (endogenous) for the most intense transition ($[y13\text{-}H_3PO_4]^{2+}$, (Figure 35).

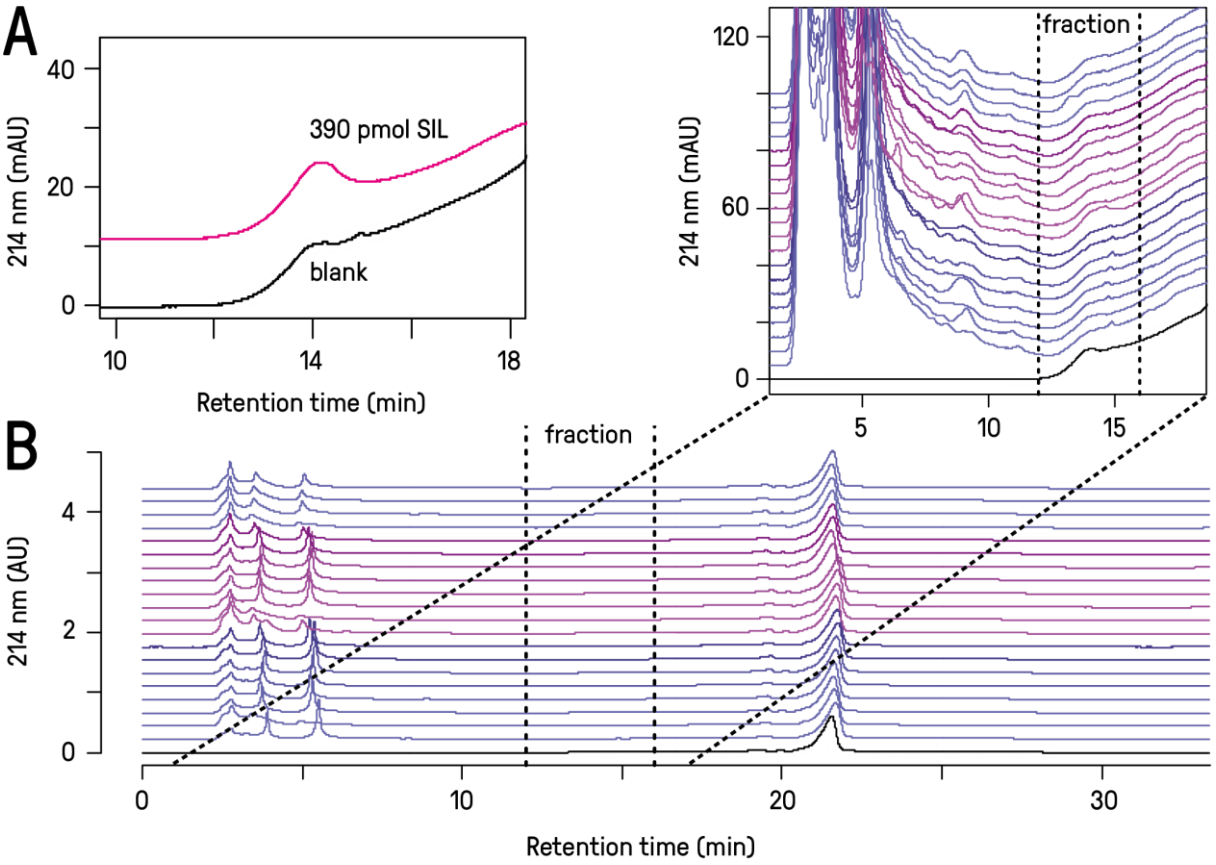


Figure 34: A: 390 pmol of the SIL peptide AKPPADPAAAAsPC*R (small letter indicates phosphorylated amino acid, * indicates carbamidomethylated cysteine) were injected to determine the retention time in ERLIC. B: Digests of 20 colon cancer tissue samples including 780 amol of the SIL peptide were injected and a 4-min fraction was collected at the specific RT (12-16 min – indicated as dashed vertical line). The UV-chromatogram shows, that the vast of peptides eluted within the first 7 minutes, emphasizing appreciable enrichment of the specific phosphopeptide. The black UV-trace gives a blank injection.

Overall, the concentration of the endogenous phosphopeptide was determined to be 6.7 ± 4.5 amol / µg digest in healthy and 3.5 ± 2.4 amol / µg digest in cancerous tissue. The coefficient of variation over all individuals was quite high with ~70 %, and phosphopeptide levels were only found to be lower in cancerous tissue from 8 out of 10 patients (on average 3.5-fold, exemplified in Figure 35A). Quantification of the non-phosphorylated peptide using 1.2 µg of digest including 240 amol of the internal standard revealed no appreciable regulation of the non-phosphorylated counterpart with 253 ± 109 amol / µg digest in healthy and 354 ± 152 amol / µg digest in cancerous tissue (as exemplified in Figure 35B). The coefficient of variation of the unphosphorylated peptides was high as well (~45 %) rendering EGLN1 levels rather unstable among samples and maybe even among individuals.

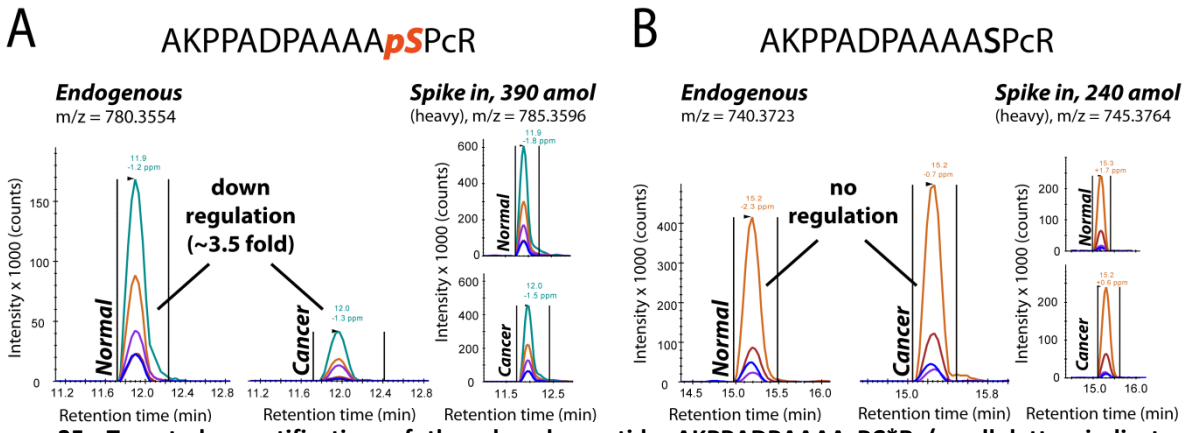


Figure 35: Targeted quantification of the phosphopeptide AKPPADPAAAAsPC*R (small letter indicates phosphorylated amino acid, * indicates carbamidomethylated cysteine) from a phosphopeptide-enriched ERLIC fraction using targetd LC-MS in presence of a synthetic heavy reference peptide. The injected sample amount corresponds to 16 µg digest + 390 amol of SIL (A). In addition, the non-phosphorylated counterpart was quantified using 1.2 µg digest + 240 amol of the SIL (B). Figure from Di Conza, Trusso Cafarello, Loroch et al. 2017.[202]

However, quantification of both, the phosphopeptide and its non-modified counterpart, allowed to calculate the site occupancies (% of phosphorylation). Site occupancy in healthy tissue was 2.5 ± 1.2 %, but in cancerous tissue only 0.9 ± 0.4 % – a difference of high statistical significance with a p-value = 0.007 (paired Student's t-test). The downregulation was detected in all 10 patients and the average downregulation was 3.5-fold (Figure 36A, healthy vs cancerous tissue).

Results of immunodetection experiments (Western-blotting, conducted by Prof. Massimiliano Mazzone and co-workers, Vesalius research center, Leuven, Belgium) were in good accordance. The downregulation of the S125 phosphorylation on EGLN-1 was accompanied by elevated HIF1α-levels. As EGLN-1 stabilizes Hif1α by hydroxylation, the results strongly emphasize a regulatory role of the phosphorylation with respect to activity (Figure 36B).[202] In summary, the ERLIC-based workflow for purifying a phosphopeptide from patients tissue digest allowed a reliable quantification of S125 phosphorylation on EGLN-1 using only 32 µg of digest. Most likely, even 5 times less starting material would have been sufficient for the analysis when considering the appreciable detector response of the enriched phosphopeptide.

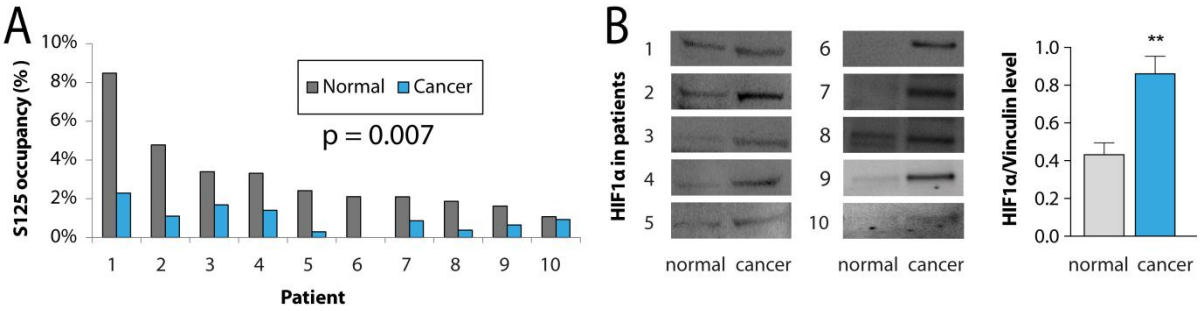


Figure 36: Targted LC-MS revealed reduced S125 phosphorylation site occupancy on EGLN1 in cancerous tissue of all ten patients (A). The concomitant stabilization of HIF1α indicates a regulatory role of EGLN-1 phosphorylation in the cancer-typical hypoxia response (B). Immunoblots from Di Conza, Trusso Cafarello, Loroch et al. 2017.[202] ** indicates a p-value ≤ 0.01.

6.6. Experiment 2: Time-resolved characterization of the proteome and phosphoproteome of MCMV-infected mouse fibroblasts

To study the manipulation of fibroblast cells by MCMV, a time-resolved characterization of the proteome and phosphoproteome was conducted using cells 4, 8 and 24 h after MCMV infection against a non-infected control. All 4 conditions were quantified against a pooled reference using a 3-plex SILAC approach, hence, each multiplexed sample comprised 2 conditions plus the reference. As the experiment

was performed in biological triplicates, 6 multiplexed samples were processed using the following workflow: ERLIC fractionation (18 fractions per sample), recovery of non-modified peptides from early fractions via SCX-SPE for high-pH-RP fractionation (24 fractions per sample) and LC-MS (252 runs in total, Figure 18).

6.6.1. Depth of analysis

The entire dataset displays expression of 99 MCMV proteins when considering all identifications over all replicates and conditions. Viral proteins were regarded 'well-quantified', if they exhibited quantitative values from all replicates at each time point (except for 0 hpi – the mock sample), which was the case for 63 viral proteins. Further, when considering all time points of all replicates, 8,752 host proteins were identified. Since quantitative values from two or more replicates at each time point were considered as minimum requirement, 6,618 host proteins remained well-quantified. Besides, 9,491 host and 259 viral phosphopeptides were identified and viral phosphopeptides were assigned to 192 phosphorylation sites on 41 viral phosphoproteins. For quantification, phosphopeptides were only considered if (I) the respective protein was well-quantified and (II) if the phosphopeptide was quantified in two or more replicates at each time point. This resulted in 2,311 well-quantified host phosphopeptides (of 1,011 phosphoproteins) and 251 viral phosphopeptides (of 41 viral proteins). These criteria allowed a ratio correction of each phosphopeptide by the respective protein ratio. This was mandatory for data analysis as the proteome underwent substantial alteration. Notably, without ratio correction phosphopeptide regulations are quite ambiguous as they might originate from differential phosphorylation as well as from differential protein abundance.

6.6.2. Dynamics of viral protein expression upon infection

Time courses were generated for 95 viral proteins (Figure 37) by calculating the abundance of a protein against the pooled reference. Relative expression levels ranged from 0 % (no expression) to 400 %

(maximum expression possible) which originates from calculating expression levels relative to the average expression. The 63 well-quantified viral proteins were subjected to hierarchical clustering analysis, which allowed to assign viral proteins to four different groups. These groups represent temporal profiles with respect to the onset of protein expression and the time point of augmented expression. Proteins with an early expression after infection (4 hpi) were designated immediate early (IE, 8 proteins), followed by early proteins (E, 4 – 8 hpi, 16 proteins), leaky late proteins (LL, 8 -24 hpi, 21 proteins) and true late proteins (TL, 24 hpi, 18 proteins).

Immediate early: The IE cluster comprised proteins with a strong onset directly after infection, well-expressed already 4 hpi. For instance, IE1 was found to be expressed at maximum 4 hpi followed by a continuous decrease at all later time points. In contrast, IE2 showed as well high levels at 4 hpi, but expression reached maximum at 24 hpi. Interestingly, IE3, the major transcriptional activator of the E1 promotor, showed a zig-zag course with two maxima, 4 and 24 hpi. Notably, a recent study demonstrated the expression of different IE3 isoforms during the course of infection, hence, this unusual time course might be attributed to interference with other IE3 isoforms with different kinetics if those isoforms share the IE3-specific exon 5.[146] Other proteins found in the IE cluster were M31, M84, immunoevasin M152 and M164, which all showed maximum expression 4 or 8 hpi. Notably, protein kinase M97[151, 203] somewhat differed from others in the cluster as it had stable levels at all time points (no clear maximum).

Early: The second cluster, termed early (E), comprised 16 proteins with dominant expression 8 hpi, slightly later than the proteins in the IE cluster. E1, the enhancer of viral protein expression, showed the most pronounced onset of expression at 4 - 8 hpi (Figure 37-1), but only a slight increase afterward. Various other viral proteins, mostly of unknown function, showed a similar expression pattern with a maximum onset between 4 and 8 hpi, such as m03, immunoevasin m06, m25.1, m42, the seven transmembrane receptor (7 TMR) M78, the primase associated factor M102, M133, m138, m139, m140, m141, m143, m145, m153 and m166. Although some of the proteins in the E cluster had maximum

expression levels 24 hpi, all proteins exhibited a very pronounced onset of expression (slope) at an earlier time point than LL or TL proteins. The expression of these proteins might depend on the E1 promotor.

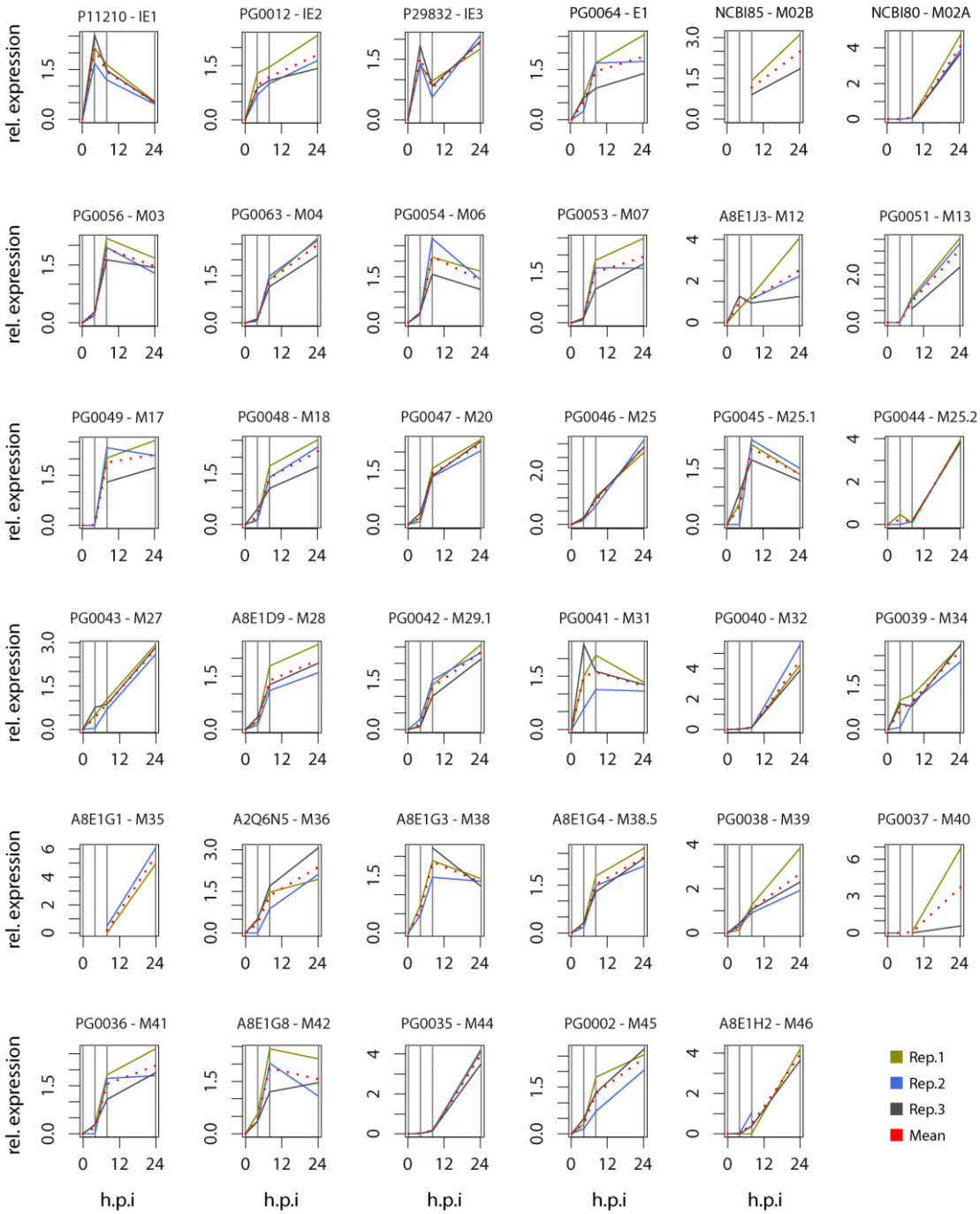


Figure 37-1: Time courses of 95 virus proteins quantified in this study. Y-axis gives the relative expression within 0: no expression and 4: maximum expression possible. Figure continues on next page.

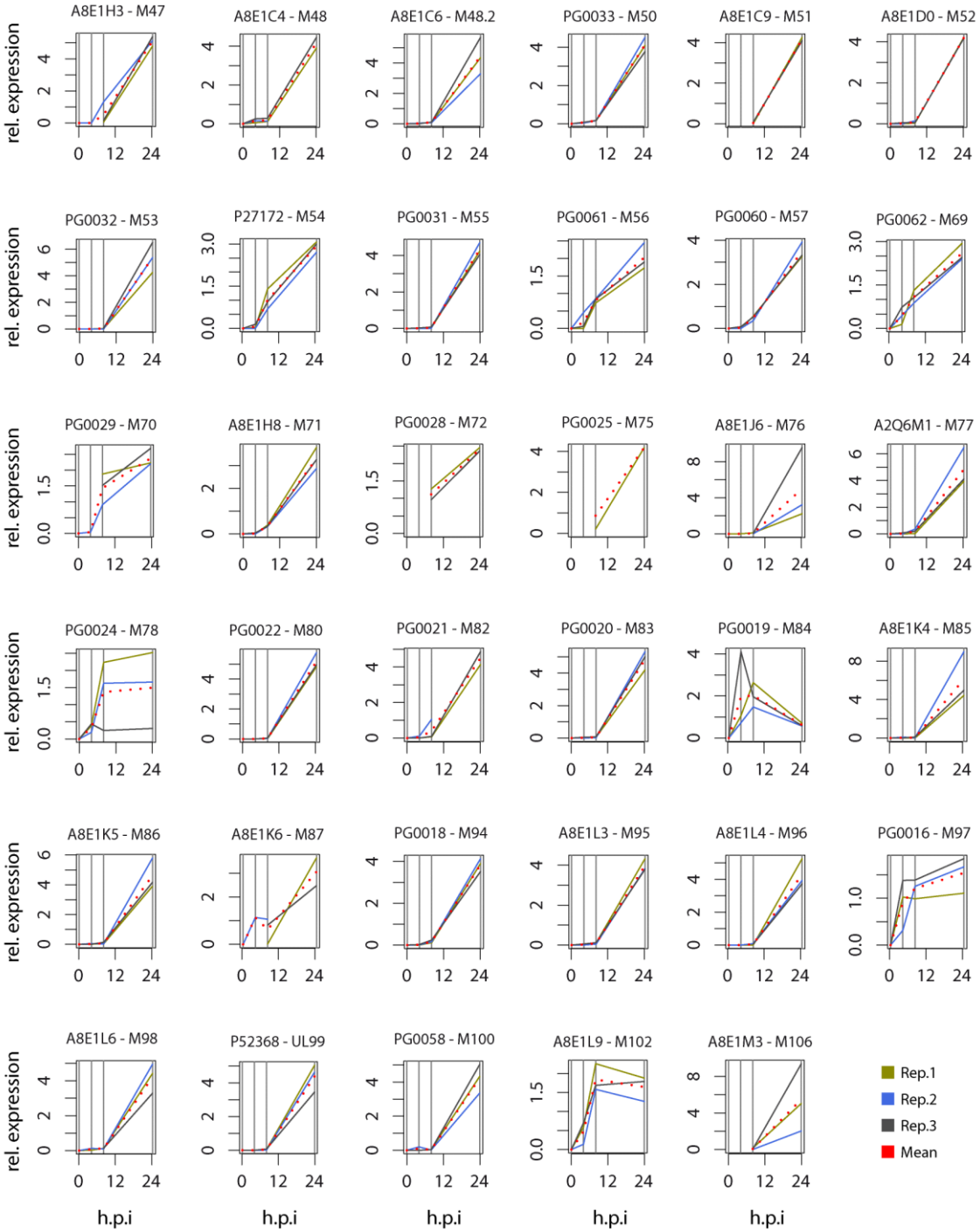


Figure 37-2: Time courses of 95 virus proteins quantified in this study. Y-axis gives the relative expression within 0: no expression and 4: maximum expression possible. Figure continues on next page.

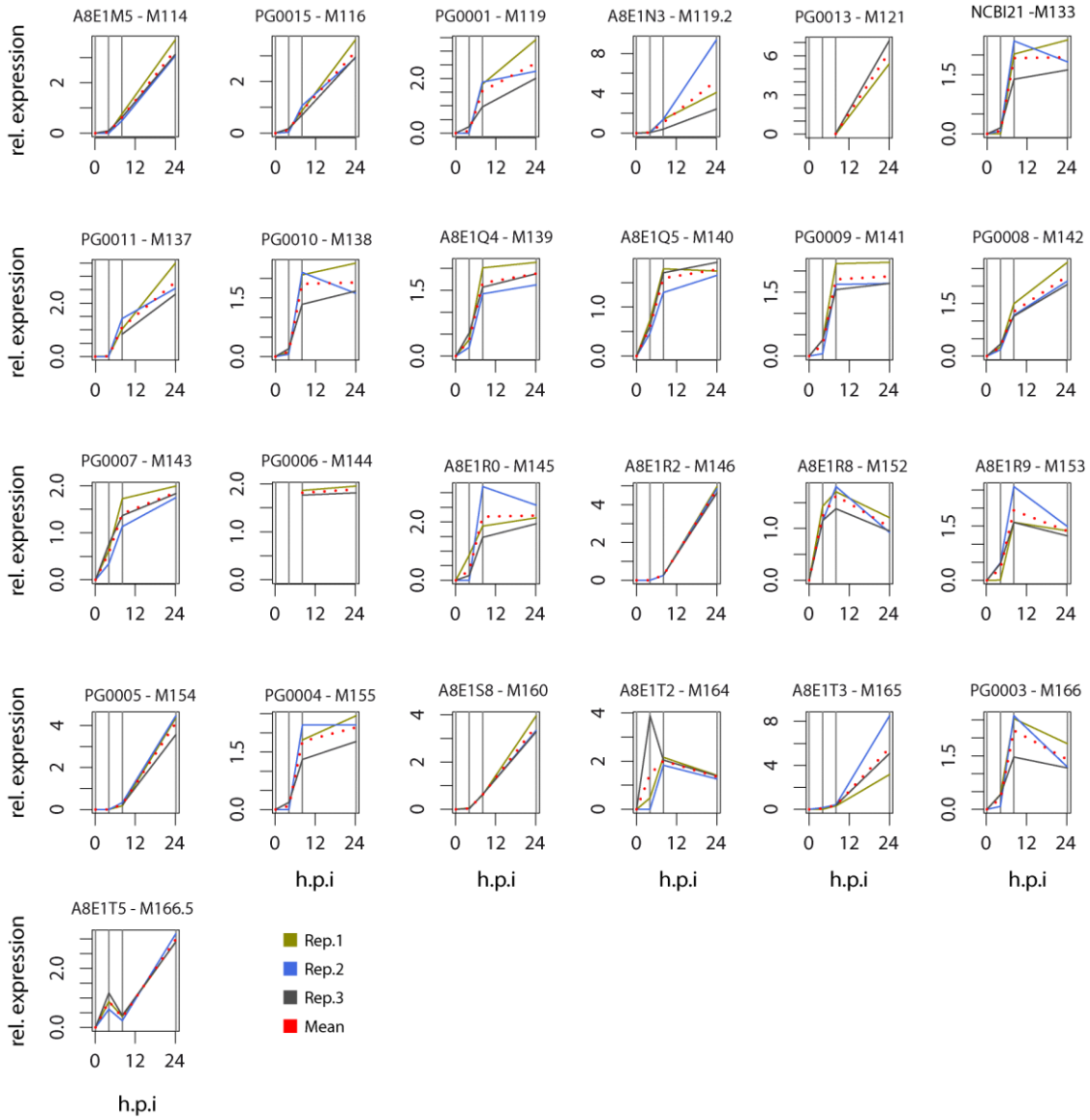


Figure 37-3: Time courses of 95 virus proteins quantified in this study. Y-axis gives the relative expression within 0: no expression and 4: maximum expression possible

Leaky late: The majority of proteins were detected in the leaky late (LL) cluster, with slight expression 8 hpi but a clear maximum 24 hpi. The LL cluster comprised the 21 proteins m4, m7, m18, m20, M25, M27, M28, m29.1, M34, M36, m38.5, M39, M41, M45, M54, M56, M69, m116, m119, m142, m166.5, which are partially involved in cell death inhibition (M36, m38.5, m41 and M45),[158] but also in viral DNA replication (DNA-polymerase M54 and the DNA-packing protein M56).[151] Interestingly, m142 and m143 were in different clusters (LL and E, Figure 37-3 and Figure 38), even though both are

described to form a complex to shutdown PKR-induced protein synthesis.[204] Notably, although both exhibit similar profiles, m143 had an earlier onset.

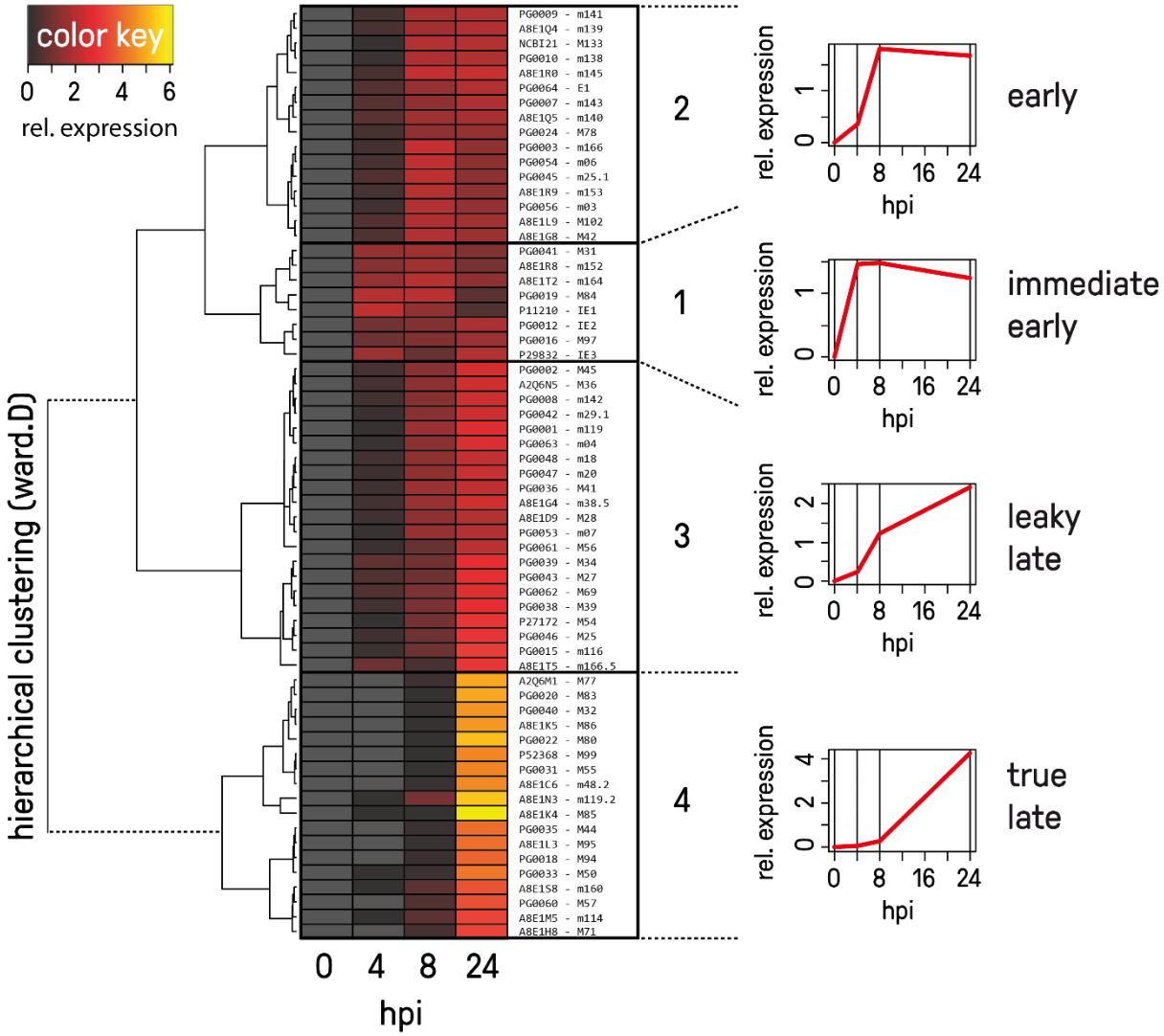


Figure 38: Hierarchical clustering of viral protein expression profiles using all 63 proteins with complete time course. With respect to onset of expression and expression maxima, proteins were grouped in immediate early (4 hpi), early (8 hpi), leaky late (8-24 hpi) and true late (24 hpi).

True late: The true late (TL) cluster comprised 17 proteins with maximum expression at 24 hpi and almost no expression at earlier time points (M32, M44, m48.2, M50, M55, M57, M71, M77, M80, M83, M85, M86, M94, M95, M99, m114, m119.2 and m160). Proteins of this cluster most likely facilitate the

final steps of virus replication. For instance, M77, M94 and M99 are associated with virus particle assembly and release and M48.2, M55, M83, M85, M86 and M100 have been described to be virion-associated.[151] It is important to note, that some proteins do not match to the traditional classification.

6.6.3. Phosphorylation sites and dynamics on viral proteins

Analysis of phosphoproteome data revealed 192 different phosphorylation sites on 41 different viral proteins (Figure 39, Figure 40). With site stoichiometry of 73.4 % for phosphoserine (pS), 21.9 % for phosphothreonine (pT) and 4.7 % for phosphotyrosine (pY, Figure 39B), pT and pY are double as frequent as expected for murine fibroblast proteins (11 % pT and 1.6 % pY according to literature).[96] Notably, 41 % of all viral phosphoproteins were only detected with one or two phosphorylation sites, whereas M18, M25, M31, M69 and M97 were detected as highly phosphorylated (12 to 20 sites, Figure 39 A). Mapping of phosphorylation sites to protein sequences revealed that phosphorylation was predominantly found in distinct clusters – probably highly phosphorylated protein domains. Such domains might be responsible for phosphorylation-mediated modulation of viral protein activity, function or localization. For instance, M18 was detected with 8 phosphorylation sites within S710 to T724, M25 with 5 sites within S907 to S915 and M31 with 5 sites within S157 to S169. In addition, the major transcriptional activator IE3 had two out of four detected sites in close proximity, namely T295 and S298 (Figure 40).

Ratios of all 251 viral phosphopeptides were corrected by protein ratios (section 6.6.1). As strong changes of protein expression render phosphopeptide ratio correction prone to errors, regulated candidates were only considered if the correction factor was ≤ 5, hence, a protein fold-change was supposed to be less than 5-fold. Phosphopeptides were regarded significant if the corrected fold-change was above (I) 1.5-fold with ≥ 1.33-fold in all three replicates or (II) ≥ 2-fold in two out of three replicates in case of a missing value. These comparably conservative criteria aimed at reducing false positives as

no further validation was conducted. 34 phosphopeptides exhibited a pronounced up- and 10 phosphopeptides a pronounced downregulation. These regulated phosphopeptides derived from 14 viral proteins – 3 from IE, 3 from E, 7 from LL proteins (Figure 40).

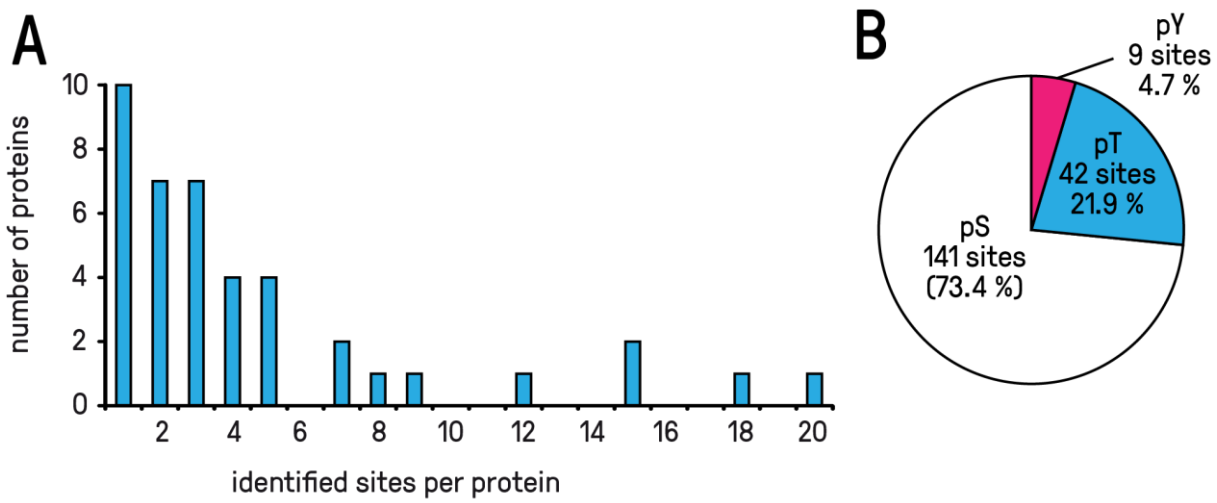


Figure 39: The majority of phosphoproteins were detected with 1-5 sites, but 5 proteins were highly phosphorylated with up to 20 sites (A). Overall site stoichiometry was 73.4 % pS, 21.9 % pT and 4.7 % pY (B).

Proteins from the IE cluster (IE3, M31 and M97): The major activator of the E1 promotor was phosphorylated on T295 and S298 as detected by a doubly phosphorylated peptide which was nearly 5-fold upregulated 24 hpi (in comparison to 8 hpi). Whether both sites become dephosphorylated remains unknown but appears likely, since no corresponding singly phosphorylated peptide was found upregulated. Protein kinase M97 exhibited numerous phosphorylations at S2, T6, S13, S86, S107 and T111. Notably, T6 was upregulated 8 hpi, but downregulated 24 hpi. Two other sites in close proximity, S2 and S13 (covered by a doubly phosphorylated peptide) were 2-fold upregulated 24 hpi indicating regulatory function of N-Terminal sites. M31, a protein of unknown function,[151] was detected with numerous phosphopeptides upregulated 8 hpi, poiting towards differential phosphorylation on S10, S87, S157, S160, T205, S209. Notably, S10 and S87 appeared downregulated 24 hpi, whereas all others remained upregulated. The exact contribution of each site in particular remains unknown, but the complex phosphorylation pattern on the N-terminus of M31 emphasizes an important regulatory role.

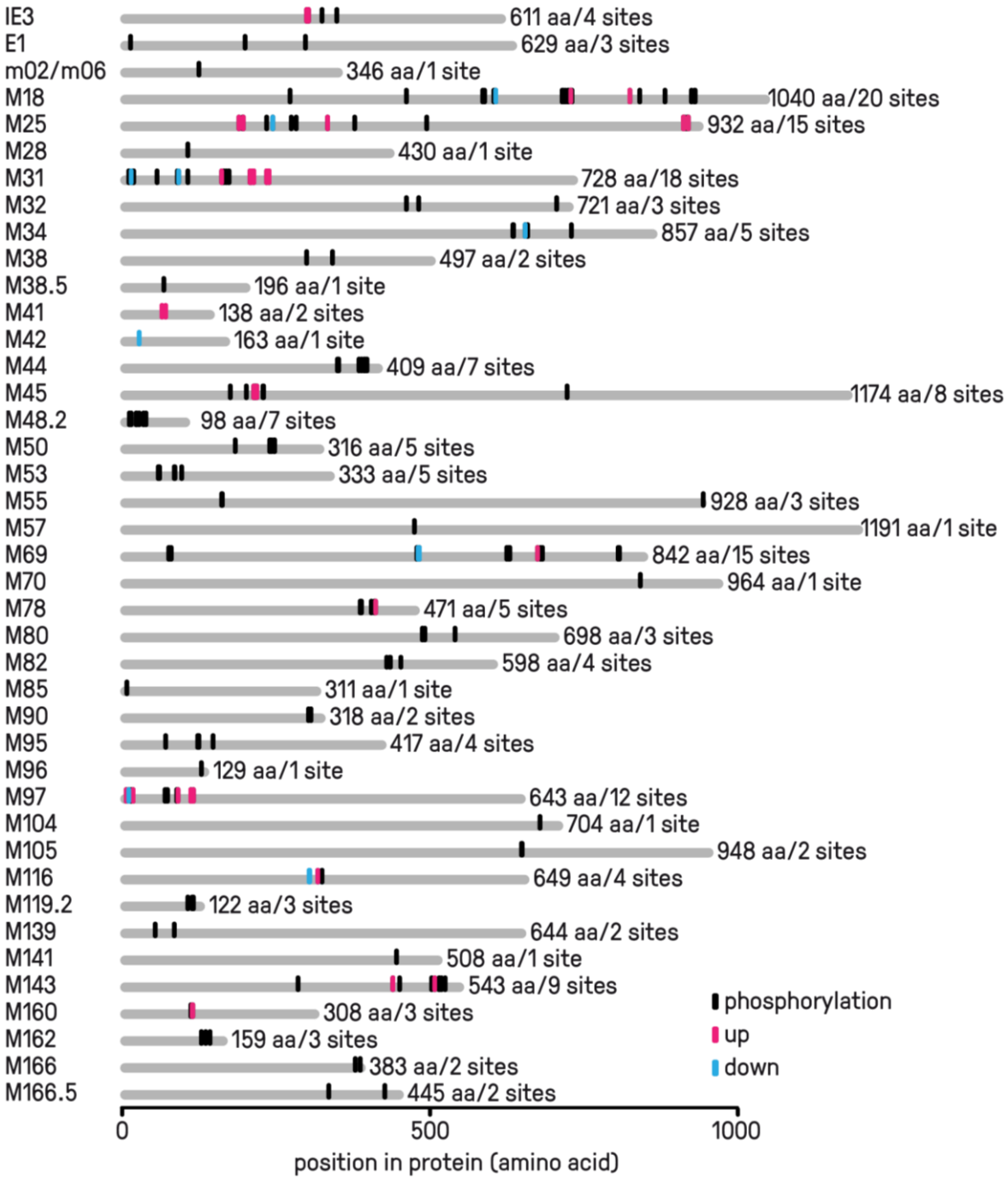


Figure 40: The data set comprises 192 confident phosphorylation sites on 41 viral phosphoproteins (phosphoRS ≥ 90 %) predominantly found in clusters. Sites of regulation were inferred from phosphopeptide ratios (4 vs 8 hpi and 4 vs 24 hpi. Regulated phosphopeptides, exact sites and protein sequences are given in the appendix (sections 10.1, 10.2 and 10.3).

Proteins from the E cluster (M42, M78, M143): M42, a protein of unknown function, showed a 2-fold downregulation of Y23. For the 7TMR M78,[205] a 5-fold upregulation of T407 from 8 to 24 hpi was

observed, whereas protein levels remained stable. M143, which forms a complex with m142 to suppress PKD-mediated synthesis shutoff,[204] showed 1.5- and 2.5-fold upregulation on S435 and S503 from 4 hpi to 24 hpi. In addition, the C-terminus was found highly phosphorylated (6 sites in close proximity) indicating that the C-terminal domain might be of regulatory function.

Proteins from the LL cluster (M18, M25, M34, M41, M45, M69, M116): M18, as well a protein of unknown function, was detected with three regulated phosphorylation sites: S602 was 2-fold down- and S724 and S820 were 2-fold and 3-fold upregulated 24 hpi. For M34, a singly phosphorylated peptide comprising S650 was 2-fold downregulated 24 hpi. Similarly, a singly phosphorylated peptide of M69 covering S670 was 2-fold up and a doubly phosphorylated peptide covering T476 and S478 was 1.5-fold downregulated. For M116, two singly phosphorylated peptides covering S299 and S300 were 3-fold downregulated, whereas another singly phosphorylated peptide indicated T313 to be extremely upregulated 24 hpi.

M25 was detected as extensively phosphorylated with the majority of sites being located in the regions S185-Y192 and S908-S915. Several sites displayed differential phosphorylation between 8 and 24 hpi. Singly-, doubly and triply phosphorylated peptides indicated dephosphorylation on S908, T909, S913, S915, but the close proximity did not allow to determine the exact pattern of regulation. Similarly, a singly and a doubly phosphorylated peptide, covering S185 and Y192, were 2-fold downregulated indicating either dephosphorylation of both sites or dephosphorylation of S185 with concomitant phosphorylation of Y192. In addition, the site T191 was detected as phosphorylated but without sufficient quantitative information. However, the complex phosphorylation pattern might be connected to the described re-localization of M25 during the course of infection.[206]

For the two proteins involved in host-defense manipulation, M41 and M45,[158, 168, 169, 171, 207] multiple regulated phosphopeptides were detected. A doubly phosphorylated peptide covering S60 and S66 of

M41 was 2-fold upregulated 24 hpi. Similarly, two phosphopeptides of M45, covering S209 and T214, were upregulated. Surprisingly, the identified site in between, S211, did not appear regulated. However, these phosphorylation sites might be directly involved in regulation of protein activity or localization.

No phosphopeptides of TL proteins passed the criteria for regulation: As the strong onset of TL protein expression at 24 hpi (usually >> 5-fold) renders quantification of phosphopeptides prone to errors, no phosphopeptide was considered regulated. Irrespectively of this criterion, the two sites S109 and S110 on M160 appeared potentially regulated as the corresponding doubly phosphorylated peptide increased by 13-fold, but the protein only by 6-fold.

6.6.4. Proteomics and Phosphoproteomics characterization of fibroblasts upon MCMV infection

In the first 4 hours after infection, during expression of IE and partially E proteins, the proteome was exceptionally stable and only two proteins were found regulated. At 8 hpi when E and the first LL proteins were detected, fundamental changes of the proteome were observed pointing towards a comprehensive re-organization of the entire proteome. At 24 hpi, the host cell displayed a fundamentally altered proteome with numerous downregulations (Figure 41). In comparison, regulations on the phosphoproteome level were already observed at 4 hpi. Similarly to the proteome, substantial changes were observed at 8 hpi. Interestingly, only minor differences were detected at 24 hpi (Figure 42) rendering phosphorylation events to respond earlier to MCMV infection than changes on the proteome level.

Considering at least 1.5-fold regulation as threshold, 9.2 % of all quantified proteins were regulated during the course of infection (252 up and 358 down). Notably, during subsequent data analysis, systemic alteration of numerous pathways was observed, but proteins were predominantly below the threshold. This becomes apparent when including proteins of at least 1.33-fold regulation, revealing

already 43 % of all quantified proteins as significantly regulated (2,606). Differential phosphorylation was detected on 284 phosphoproteins (5 % of all quantified proteins) inferred from 174 upregulated and 241 downregulated phosphopeptides (at least 1.5-fold).

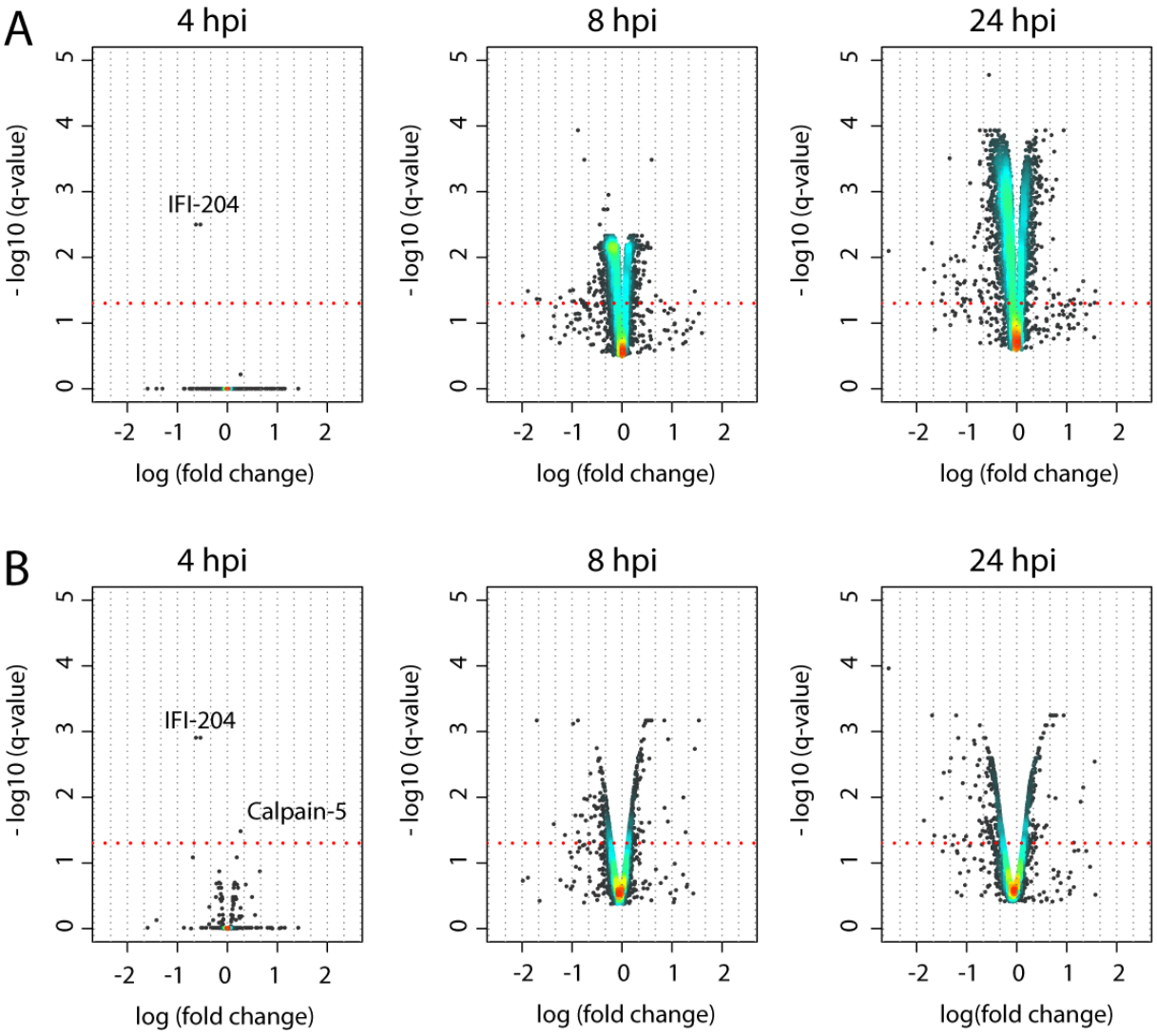


Figure 41: Altered protein expression in fibroblasts cells during the course of MCMV infection against non-infected control. Rank product (A) and Limma (B), known to yield complementary results, were used to determine significantly regulated features.[195] At 4 hpi, the proteome remains quite stable though IFI-204 and Calpain-5 are regulated, perhaps initial events upon infection. At 8 hpi, drastic regulations can be observed towards a considerably reorganization of the proteome at 24 hpi.

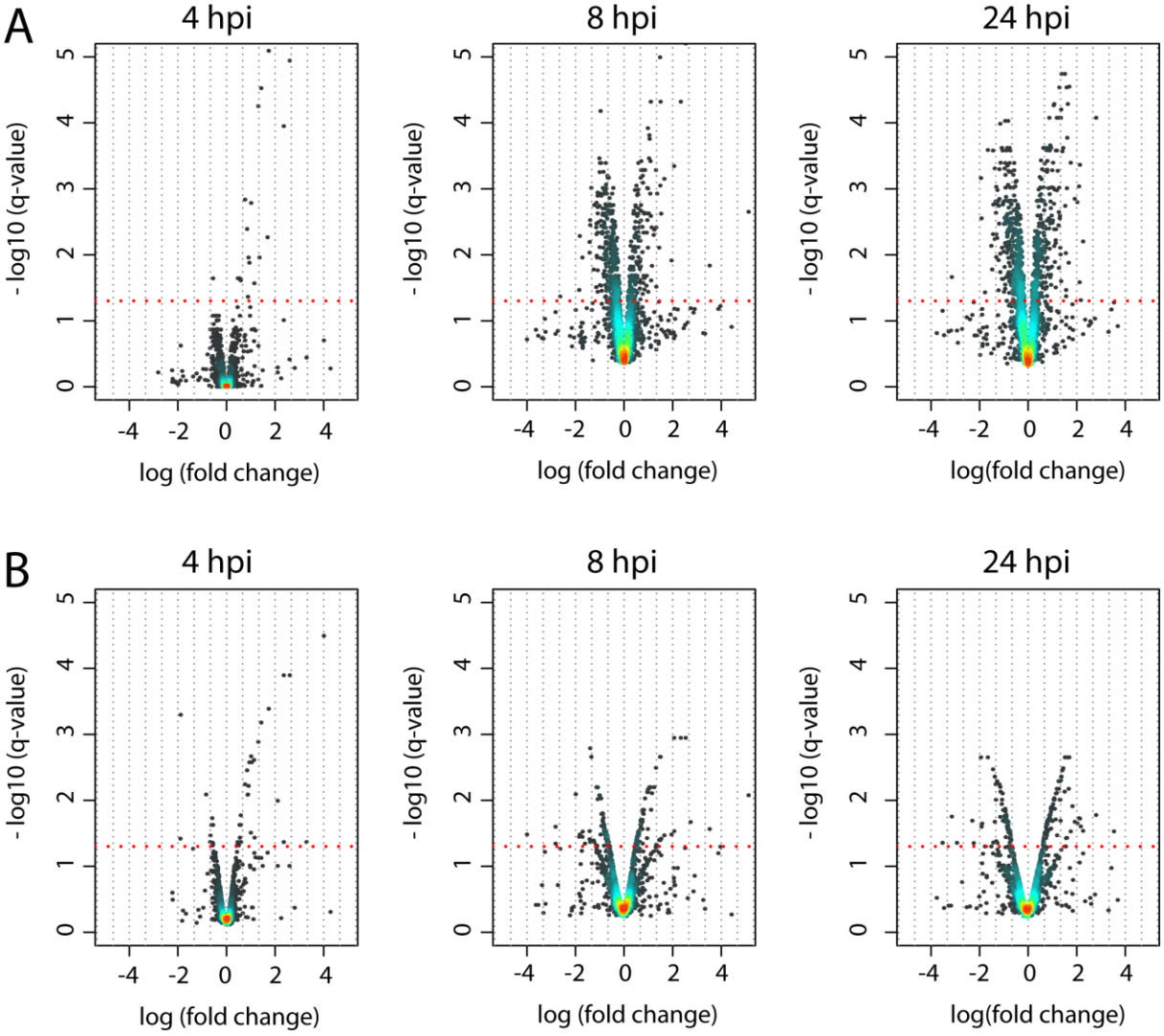


Figure 42: Phosphopeptide levels in fibroblasts cells upon CMV infection against non-infected control cells. As in the previous figure (Figure 41) q-values of both rank product (A) and limma (B) were plotted.

6.6.4.1. Hierarchical clustering to determine global protein regulation

Proteome: Quantified proteins were structured according to their regulation by hierarchical clustering and 6 clusters were assigned: up, slightly up, background, slightly down, moderately down, down (Figure 43). Clusters were subjected to a functional enrichment analysis using the DAVID online tool

(david.ncifcrf.gov,[197] accessions of all quantified proteins were used as background). In cluster 1, comprising distinctly upregulated proteins, functional terms were related to "RNA-binding / splicing", "oxidative phosphorylation" and "protein processing in endoplasmic reticulum" (Table 10) emphasizing upregulation of protein synthesis capabilities. In cluster 2, comprising slightly upregulated background proteins, similar terms were enriched complementing the picture of a global upregulation of protein synthesis capabilities. However, the two background clusters 3 and 4 (stable and slightly downregulated) did not yield any enriched terms noteworthy. In contrast, cluster 5 and cluster 6 (moderately down and pronounced downregulation) displayed terms related to cellular signalling such as "JAK-STAT signalling", "Calcium signalling", "ERBB signalling" and WNT signalling". Notably, also the term "cell cycle" was enriched in cluster 5.

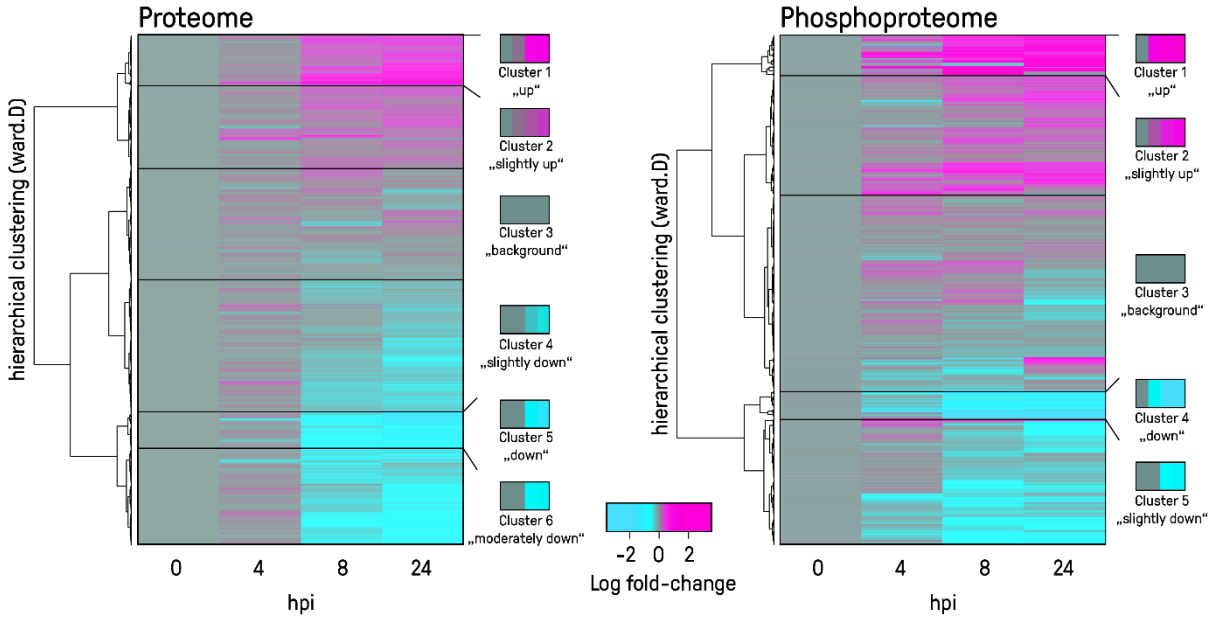


Figure 43: Heatmap of hierarchical cluster analyses of quantified proteins (A) and phosphopeptides (B). Proteins were grouped into 6 clusters, 3 of which showed a pronounced regulation (1, 5, 6), the others were regarded as background (2, 3, 4). Among 5 phosphopeptide clusters, two clusters exhibited a distinct regulation (1, 4), the remaining three were regarded as background (2, 3, 5). Clusters were subjected to a functional enrichment analysis to decipher patterns related to host cell manipulation (Table 10).

Phosphoproteome: Clustering of phosphopeptides resulted in 5 clusters (Figure 43) and protein accessions were subjected to functional enrichment analysis as well (Table 10). In cluster 1, comprising

108 phosphoproteins with upregulated phosphopeptides, the terms “regulation of cell migration” and “intermediate filament” were found enriched, likely to be related to vesicle transport. In addition, “histone acetyltransferase activity” was enriched, probably related to re-organization of histone proteins. In cluster 2 (slightly and moderately upregulated), comprising 358 protein accessions, the terms “RNA-binding” / “processing” / “transport”, “alternative splicing” and “spliceosome” were found enriched. In contrast, cluster 4 and 5 (downregulated and slightly downregulated) were annotated with terms such as “cell cycle”, “DNA-replication” and “DNA-repair” pointing towards a systemic shut-down of these processes.

In summary, the clustering analysis according to regulation profiles revealed an increase in protein and protein phosphorylation levels in pathways related to splicing and oxidative phosphorylation. In contrast, diminished protein levels were observed for numerous signalling processes accompanied by dephosphorylation events within the DNA-replication machinery. These alterations point towards (I) an abrogation of normal cell homeostasis by shutdown of signalling, (II) the inhibition of cell proliferation by dephosphorylation within the replication machinery, (III) the transformation of the cell into a protein synthesis machinery by upregulation of splicing-related proteins and (IV) an overall increase in the energy metabolism (oxidative phosphorylation).

Table 10: Enrichment of functional terms among protein and phosphoprotein clusters (Figure 43) using the online-tool DAVID.[197] All redundant and protein structure-related terms were removed. Terms in bold indicate pathways discussed in the text.

Proteome (6 clusters)	Phosphoproteins (5 clusters)
Pronounced upregulation (cluster 1) poly(A) RNA binding (25.9%), **Oxidative phosphorylation** (5.5%), Mitochondrion inner membrane (7.5%), Spliceosome (5.5%), **mRNA splicing** (7.9%), Parkinson's disease (4.3%), Alzheimer's disease (4.7%), Huntington's disease (5.1%), **Ribosomal protein** (6.3%), **Protein processing in endoplasmic reticulum** (4.1%), ATP synthesis (1.2%), Electron transport (3.1%)	**Pronounced upregulation (cluster 1)** regulation of cell migration (3.7%), negative regulation of neuron projection development (3.7%), PcG protein complex (2.8%), Intermediate filament (2.8%), nucleoplasm (34.9%), ubiquitin ligase complex (3.7%), **histone acetyltransferase activity** (2.8%), nuclear matrix (5.5%)
Background - slightly upregulated (cluster 2) extracellular exosome (31.7%), mitochondrial inner membrane (7.4%), **Ribonucleoprotein** (7.5%), **mRNA splicing** (6.2%), Spliceosome (4.3%), Huntington's disease (4.4%), Proteasome (2.3%), **Oxidative phosphorylation** (3.3%), **Chaperone** (4.5%), chaperone binding (2.1%), Electron transport (2.6%), mitochondrion (20.7%), **respiratory chain** (1.8%)	**Background - slightly upregulated (cluster 2)** poly(A) RNA binding (31.7%), Translation, ribosomal structure and biogenesis (2.2%), **Spliceosome** (4.9%), Methylation (21.9%), RNA binding (16.9%), peroxisome (2.2%), Ribonucleoprotein (4.6%), acyl-CoA metabolic process (1.4%), Peroxisome (1.9%), RNA transport (5.2%), **mRNA splicing** (8.2%), cellular protein localization (2.2%)

Proteome (6 clusters)	Phosphoproteins (5 clusters)
Background (cluster 3) extracellular exosome (30%), ribosome (4.6%), Ribosomal protein (4.4%), Prenylation (2.3%), cell-cell adhesion (4.4%), cell surface (4.1%)	**Background (cluster 3)** Ubl conjugation (27.9%), actin binding (6.7%), regulation of transcription from RNA polymerase II promoter (10.4%), RNA-binding (14%), regulation of cell proliferation (3.6%), intracellular ribonucleoprotein complex (6.1%), poly(A) RNA binding (27.2%), Methylation (20%), apoptotic process (6.5%), DNA-binding (15.3%)
Background - slightly down (cluster 4) NADP (2.9%), oxidoreductase activity (6.9%), Nucleotide-binding (17.5%), Arginine and proline metabolism (1.1%), Actin-related protein (0.7%), Glutathione metabolism (1.1%), rRNA processing (2.2%), Cytoplasm (40.8%)	**Background - slightly downregulated (cluster 5)** Methylation (23.3%), Ubl conjugation (28.7%), Ankyrin repeat (3.3%), Pleckstrin homology domain (5.7%), cell differentiation (7.6%), **DNA repair** (5.7%), nervous system development (4.9%), **DNA damage** (6.2%), cell junction (9.5%), DNA-binding (16.3%), small nucleolar ribonucleoprotein complex (1.4%), chromosome (6.5%), regulation of transcription, DNA-templated (22%), cytoskeleton (16.5%)
Moderately down (cluster 5) ATP binding (17.2%), protein kinase activity (6.2%), Chromosome (5.8%), **Cell cycle** (9.1%), Mitosis (4.9%), Ubl conjugation (17.4%), cellular response to DNA damage stimulus (5.8%), HTLV-I infection (3.1%), **phosphatidylinositol 3-kinase binding** (0.9%), **Jak-STAT signaling pathway** (1.2%), DNA binding (12.4%), **signal transduction** (5.5%), microtubule (4.7%), DNA methylation (0.9%)	-No 6th cluster assigned-
Pronounced downregulation (cluster 6) Pleckstrin homology-like domain (11.1%), **signal transduction** (9.7%), **Kinase** (10.8%), **Calcium signaling pathway** (2.7%), activation of GTPase activity (2.4%), cell junction (8.1%), SH2 domain (2.4%), PLC-like phosphodiesterase / Rho protein signal transduction (2.2%), regulation of vesicle fusion (1.6%), ErbB signaling pathway (2.4%), Wnt signaling pathway (2.4%)	**Pronounced downregulation (cluster 4)** Translocation (5.6%), Nuclear pore complex (4.5%), Ribosomal protein L2 domain 2 (3.4%), NLS-bearing protein import into nucleus (3.4%), **DNA replication initiation** (3.4%), kinetochore (6.7%), histone binding (5.6%), Lipid-binding (6.7%), DNA unwinding involved in DNA replication (3.4%), **DNA replication** (3.4%), Protein transport (10.1%)

6.6.4.2. Temporal changes of the proteome and phosphoproteome

This section highlights the changes observed over time with a special focus on signalling pathways and cellular functions found significantly altered.

Early changes at 4 hpi: The cell growth inhibitor Interferon-activable protein (Ifi-204), a key-player of interferon-mediated signalling, was downregulated 4 hpi. The endopeptidase Calpain-5, a protease of undetermined function, but known to co-localize with promyelocytic leukemia-nuclear bodies, important in cellular processes such as apoptosis and viral defence,[208] was upregulated. Analysis of the time course of these proteins revealed that the regulation persists over all later time points and proteins do not return to the initial levels (Figure 44).

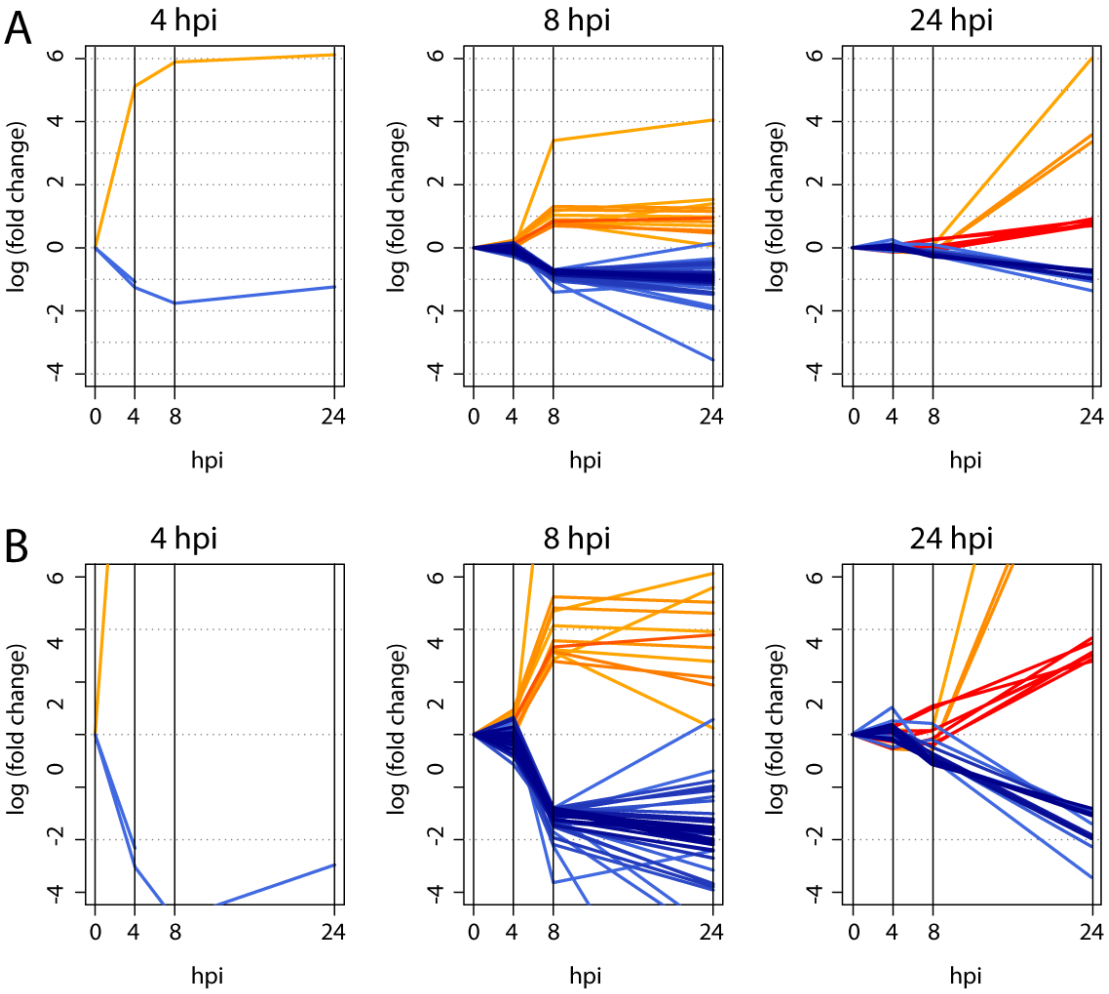


Figure 44: A: Time courses of regulated proteins with a distinct onset either 4, 8 or 24 hpi (at least 2-fold, $\log_e \approx 0.69$) in reference to non-infected control cells. Plots indicate that regulations predominantly persist or might be even more distinct 24 hpi. B: Y-axis zoom

Tim Schommartz from the Heinrich Pette Institute (HPI, Leibniz Institute for Experimental Virology, Hamburg, Germany) successfully validated the pronounced early regulation of Ifi-204 by Western blotting (Figure 45A). In contrast, antibody-based detection of Calpain-5 repeatedly failed, but preliminary results of semi-quantitative RT-PCR-experiments indicate an increase of Calpain-5 transcript levels (Figure 45B) supporting the finding of an early Calpain-5 upregulation. Both validation experiments render Ifi-204 and Calpain-5 as true early responders to MCMV infection.

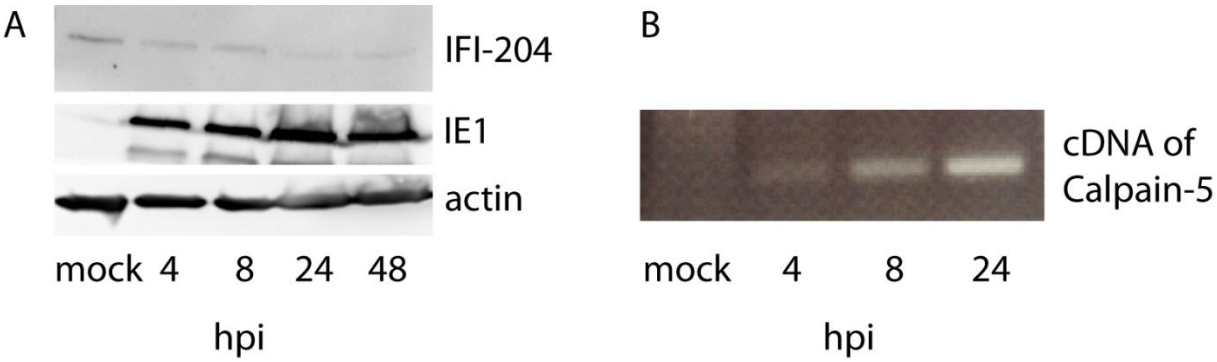


Figure 45: Validation of early protein expression changes in fibroblast cells upon MCMV for IFI-204 using Western blot (A). Antibody-based assays for Calpain-5 repeatedly failed to work, but increased mRNA levels detected by semi-quantitative RT-PCR strongly support the LC-MS results (B). The experiments were conducted by Tim Schommartz from the HPI.

Signalling (e.g. MAPK, TNFα): Among differentially phosphorylated proteins at 4 hpi, the GO-term "nucleus" was found significantly enriched (21 proteins, category: *cellular component*, FDR: $3.5*10^{-3}$). Interestingly, numerous regulations were connected to kinase and phosphatase signalling. The interferon-inducible double-stranded RNA-dependent protein kinase activator A (Prkra, S18, 1.5-fold) and the protein phosphatase 1G (Ppm1g, S524, 2.5-fold up) were detected with elevated phosphorylation levels. Further, the ribosomal protein S6 kinase alpha-4 (Rps6ka4 *syn.* Msk2) exhibited alteration on protein phosphorylation levels (S343 and S347 covered by a doubly-phosphorylated peptide, 10-fold upregulated). The central role of MSK2 in both tumor necrosis factor alpha (TNFα) and MAPK signalling[209, 210] renders this phosphorylation potentially important as initial event of global abrogation of these pathways (section 6.6.5.1).

At 8 hpi, numerous proteins connected to signalling were considerably downregulated – such as proteins connected to phospholipase C- (PLC) and JAK/STAT- signalling (Figure 46 A, B). Aberrant phosphorylation was detected on a wide range of signalling proteins, especially mitogen-activated protein kinases (MAPKs) such as Mapk7 (*syn.* Erk5, T723, 2-fold up), Mapk4k5 (T400 2-fold up) and the serine/threonine-protein kinase Tao1 (Taok1, S421, 2.5-fold down), an initial activator of Mapk14

(*syn.* p38). In addition, a specific interactor and scaffold protein of MAPKs, C-Jun-amino-terminal kinase-interacting protein 4 (Jip4, S815 2-fold down), was detected with aberrant phosphorylation levels, which might further contribute to MAPK pathway alterations. Interestingly, aberrant phosphorylation of signalling proteins was predominantly observed 4 and 8 hpi, but barely 24 hpi. Only the calcium/calmodulin-dependent protein kinase type II subunit delta (Camk2d, S330, 1.5-fold up) and the ribosomal protein S6 kinase alpha-3 (Rps6ka3 *syn* Rsk2, S715, 1.5-fold down) were found as late responder to the infection. Noteworthy, Rps6ka3 acts downstream of Mapk1/Erk2 and Mapk3/Erk1 signalling pathways and is involved in TNFα mediated stress-induced activation of transcription factors like Creb1,[211, 212] but also involved in histone dephosphorylation.[210]

Proteins involved in cell cycle regulation: Among proteins detected with diminished phosphorylation at 4 hpi, the regulatory subunit 12 A of protein phosphatase 1 (Ppp1r12a) was detected with elevated S507 phosphorylation level (but returned to the initial phosphorylation state at 24 hpi). As Ppp1r12a is involved in cell cycle regulation,[213] this phosphorylation might be a promising target for future studies. Accordingly, GO-related to cell cycle ("cell cycle", "cell division" and "mitotic cell cycle") were not only enriched among downregulated proteins at 8 hpi (Figure 46 D, 63 proteins, category "biological process", FDR: $1.8*10^{-16}$, $1.4*10^{-13}$, $4.1*10^{-12}$, respectively), but also enriched among differentially phosphorylated proteins ("cell cycle", 23 proteins, category: *biological process*, FDR: $5.7*10^{-7}$). Three topoisomerase isoforms were detected with aberrant phosphorylation (TOP1: S10, 2.5-fold down; Top2a: S1211 and T1245, 3-fold and 2-fold down; Top2b: S1568, 1.5-fold down) as well as proteins involved in DNA-replication and -repair, such as DNA damage checkpoint 1 (Mdc1, S929 and T1480, 2-fold down), ligase I (Lig1, S51, 1.5-fold down), mini-chromosome maintenance deficient 2 (Mcm2, doubly phosphorylated peptides comprising S26/S27; 2-fold) and 3 (Mcm3, two doubly phosphorylated peptides comprising S672/S681 and T729/S732; 2-fold down). In addition the serine/threonine-protein phosphatase 4 regulatory subunit 2 (Ppp4r2, S226 1.5-fold) was differentially phosphorylated which is

known to be involved in maturation of spliceosomal snRNPs and dephosphorylation of the histone H2ax.[214, 215] Although H2ax was not detected, several other histones were differentially phosphorylated such as H2ay (T178 phosphorylation, 1.5-fold up), H1.4 (T18, 2-fold) and H1.5 (S2, S18, 1.5-fold). In addition, other histone modifiers such as the histone-lysine N-methyltransferase 2B (Kmt2b, S866, 2-fold down), described to methylate histone H3[216], and cullin-4B (doubly phosphorylated peptide comprising T106 and S110, 2.5-fold down), involved in polyubiquitination of H2a, H3 and H4, were detected with aberrant phosphorylation. As both proteins are involved in cell cycle regulation, such aberrant histone phosphorylation might be an important link to host DNA replication shut down or viral DNA replication.

Proteins involved in mRNA processing and mRNA splicing: Among proteins upregulated at 8 hpi, the GO-terms "RNA metabolic process" were clearly enriched (22 proteins, category "biological process", $6.9*10^{-4}$, respectively) and, similarly, "RNA-binding" (19 proteins, category "molecular function" (FDR: $1.2*10^{-6}$ and). A small cluster of 5 proteins connected to "mRNA-processing" was found. This cluster included Y-Box protein 1 (Ybx1), the small nuclear ribonucleoprotein 70 (Snrnp70), the ribonucleic acid binding protein S1 (Rnps1) and the RNA binding-motif protein (Rbmx), all known to be involved in splicing. Additionally, the general transcription factor IIF was upregulated (Gtf2f1), which is involved in the recruitment of RNA polymerase II into the initiation complex (Figure 46). This tremendous regulation of RNA-related processes is further reflected by a variety of aberrant phosphorylation events: DNA-directed RNA polymerase II subunit RPB1 (Polr2a, doubly phosphorylated peptide comprising S1913 and S1920 or 1927 and 1934, 1.5-fold up), splicing factor 3B subunit (Sf3b1, S194 1.5-fold up), serine/arginine repetitive matrix protein 1 (Srrm1, T600, S602, S624, S626 via doubly phosphorylated peptides, 1.5-fold up; S1151, S1216, S2224, 1.5-fold down) and nuclear cap-binding protein subunit 1 (Ncbp1, 2-fold down). Overall, these regulations point towards a global upregulation of transcription and RNA-processing capabilities towards a systemic upregulation

of protein synthesis capabilities of the cell. A subsequent pathway-centric analysis of these regulations is described in section 6.6.5.3.

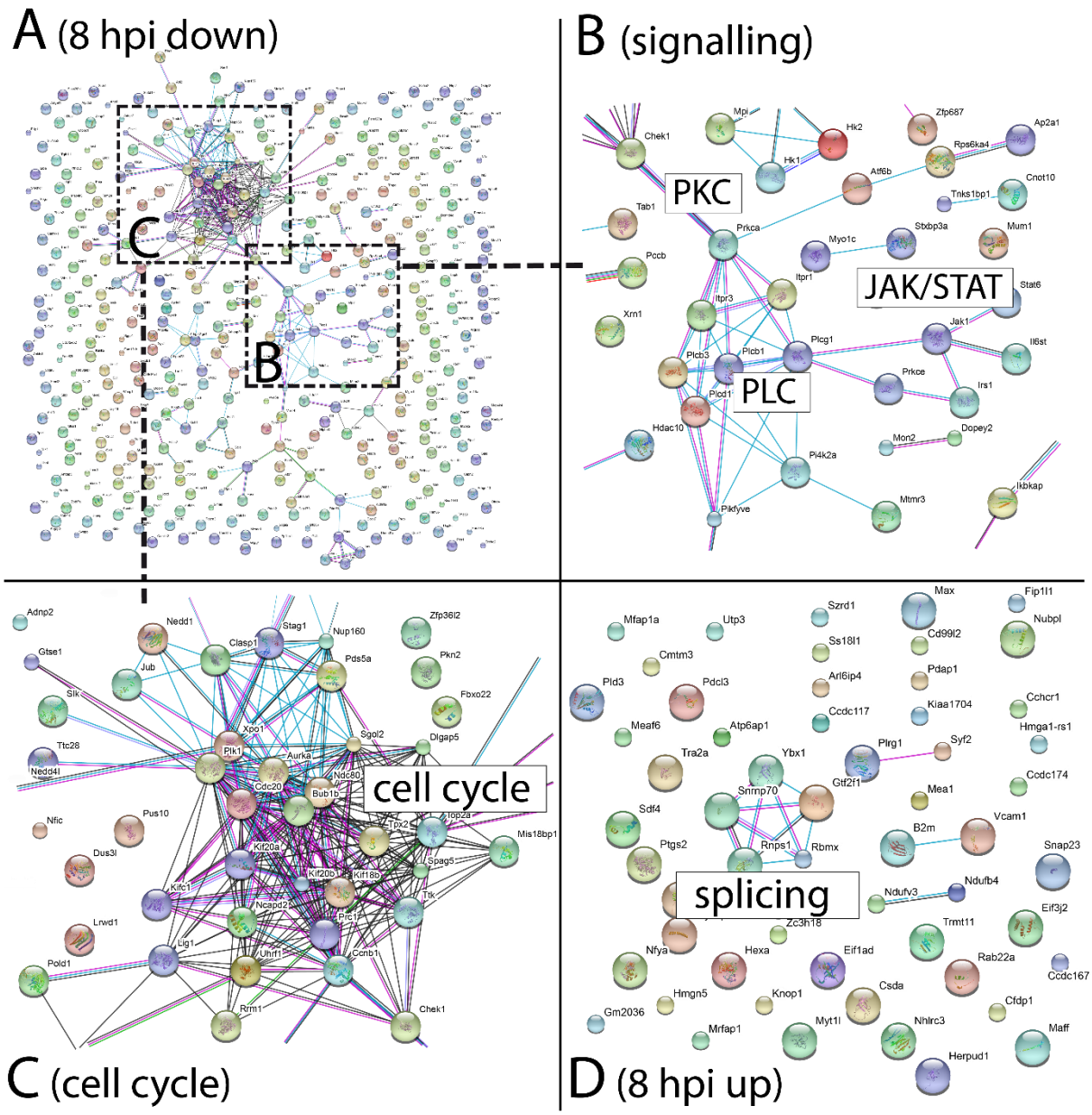


Figure 46: String network analysis of 449 host-proteins 2-fold downregulated at 8 hpi (A). A larger cluster was detected involved in cellular signaling for instance via JAK/STAT (Jak1, Stat6) and PLC/PKC (B). This cluster was connected to a seond cluster of 63 proteins involved in cell cycle regulation and mitosis (C). Among 58 proteins 2-fold upregulated, splicing proteins were found enriched (D). Settings were "confidence high" and "no textmining".

Protein processing in the ER: At 24 hpi, 27 proteins annotated with "endomembrane system" (GO category "Cellular component", FDR: $8.2*10^{-8}$) were upregulated such as transport protein Sec61 subunit beta (Sec61b), signal sequence receptor gamma (Ssr3 *syn* Trap) and ribophorin II (Rpn2 *syn* OST). Both are annotated with "protein processing in endoplasmic reticulum" and form a cluster with the ribosomal protein S27A (Rps27a), which links to a second cluster of chaperons such as calreticulin (Calr *syn* Crt) and hypoxia up-regulated protein 1 (Hyou1 *syn* Nef) and the co-chaperon Dnaj homolog subfamily B member 11 (Dnajb11 *syn* Hsp40). Although only slightly upregulated (~1.5-fold), these proteins might reflect increased capabilities of cellular protein synthesis, which is a logical consequence of the systemic upregulation of the splicing machinery

6.6.5. Pathway centric analysis of host signalling and cellular functions using KEGG pathway maps

As GO-term and functional enrichment analyses revealed numerous signalling pathways and cellular functions to be significantly altered, the data were mapped in a next step to different KEGG-pathways (www.genome.jp/kegg) to visualize changes in a pathway-centric way. Mapping was performed irrespectively of fold change cut-offs as also minor systemic regulations (e.g. ± 10 % of an entire pathway) might result in tremendous system perturbations. However, the pathway maps just display the average regulation observed and do not reflect time-courses. Notably, the analysis was conducted with the R-package "pathview",[199] which required conversion of KEGG identifiers to Uniprot accessions. This conversion was quite error-prone as in ~20 % of all cases KEGG identifiers linked to wrong or non-reviewed Uniprot accessions. Consequently, all annotations were manually revised and error-corrected. In addition, differential protein phosphorylation was added manually into each map in an isoform-specific manner. The following sections give the results of the systematic analysis of six selected KEGG-pathway maps: MAPK signalling, TNFα signalling, cell cycle signalling, apoptosis

signalling, protein processing in ER and vesicular transport. These maps were chosen based on the results from section 6.6.4 to display the substantial manipulation upon MCMV infection.

6.6.5.1. Global abrogation of MAPK- and TNFα-signaling

A closer look on proteins involved in MAPK signaling revealed a global downregulation of proteins involved in (I) the classical MAPK pathway via MEK1/2-ERK, but also via (II) ERK5 and (III) JNK/p38 (Figure 47). The downregulation was observed on all levels including a substantial fraction of MAP4K-members (Hgk *syn.* Map4k4), MAP3K-members (Mekk3 *syn.* Map3k3; Mekk4 *syn.* Map3k4), MAP2K-members (Mek1 *syn.* Map2k1; Mek2 *syn.* MAP2K2; Mek5 *syn.* Map2k5; Mkk3 *syn.* Map2k3; Mkk4 *syn.* Map2k4; Mkk6 *syn.* Map2k6) and MAPK-members (Erk5 *syn.* Mapk7; p38-kinase isoform Mapk14; JNK isoform Mapk8, and ERK isoforms Mapk1 and Mapk3). In addition, the PTP isoform dual specificity protein phosphatase 3 (Dusp3), known to inactivate Erk1/2,[217] was found upregulated. This indicated a substantial abrogation of MAPK-signaling, not only via downregulation of positive regulators, but also via upregulation of negative regulators.

In addition, aberrant phosphorylation was observed in all branches of MAPK signaling. Dephosphorylation was detected on classical MAPK pathway effectors Rsk2 (isoform Rps6ka3, S715), Stathmin (Stmn1, S25) and Nf-κb (isoform Nf-κb2, T414 and T425) and increased phosphorylation levels were detected within the JNK/p38 pathways on Flna (isoforms Flna, S2327 and S1084), Msk2 (S343/347), Taok1 (S421) and Taok2 (S324). Notably, in the ERK5 branch, Erk5 (syn. Mapk7) was detected with elevated phosphorylation levels (T723). Notably, the MAPK downstream effectors p53 (S309) and Atf2 (T53) were also detected with elevated phosphorylation levels, but did not exhibit any downregulation (including Stmn1) indicating that the protein levels are not affected by the substantial MAPK pathway modulation.

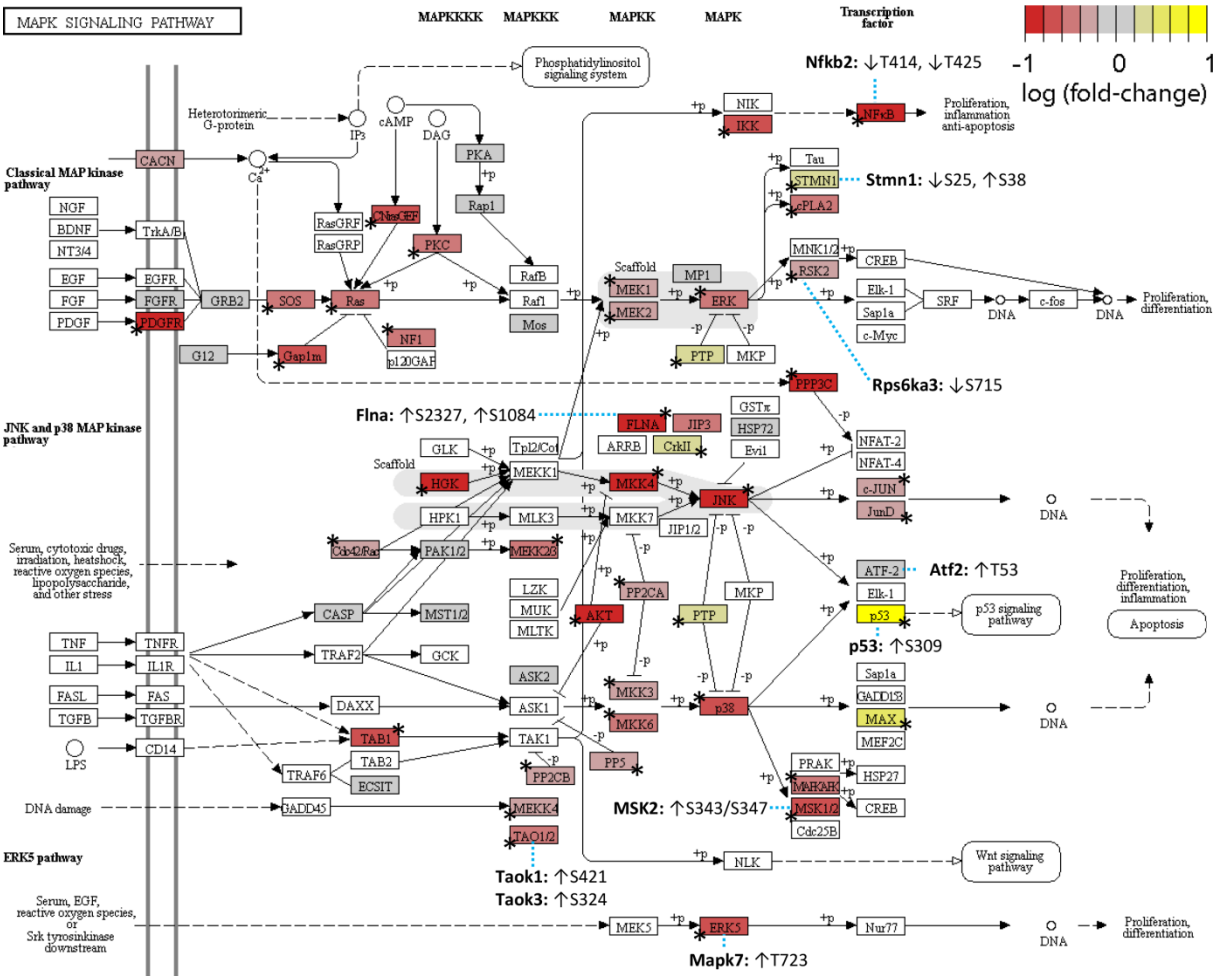


Figure 47: Systematic downregulation of MAPK signalling in fibroblasts upon MCMV infection. All three branches were found downregulated (1) via Mek1+2/Erk, (2) via p38 (Mapk14) and (3) via Erk5 (Mapk8). In addition, Dusp3, known to inactivate Erk, Jnk and p38, was slightly upregulated. Box colours represent the average log fold-change of proteins and asterisks indicate statistical significance. Altered protein phosphorylation levels were indicated isoform-specifically if regulation of the phosphopeptide was statistically significant (q ≤ 0.05). An enlarged map is in the appendix (section 10.4).

As MAPK signalling represents a downstream pathway of TNFα signalling, the respective KEGG map was supposed to get a more comprehensive picture of the global signalling shutdown. In fact, signalling via TNF receptor 1 (Tnfr1), widely mediated via MAPK and NF-κB signalling, was completely downregulated including proteins from the associated necroptosis pathway (Tradd, Rip1, Mlkl and Drp1 *syn.* Dnm1l, Figure 48). The extent of global downregulation even affected Tnfr2, the second major receptor of TNFα signalling. Several phosphoinositide 3-kinase (PI3K) isoforms were downregulated including Pik3r1, Pik3r2 and Pik3r4. However, the isoforms Pik3c2a, Pik3c3, Pik3r6 did not show

regulation and Pik3c2g was even upregulated indicating a complex pattern among PI3K proteins. Further, also Aip1 (*syn.* Dab2ip) and the Ikk isoform NF-κB essential modulator (*syn.* Nemo, Ikbkg) were considerably downregulated (Figure 48) as already described in literature.[207]

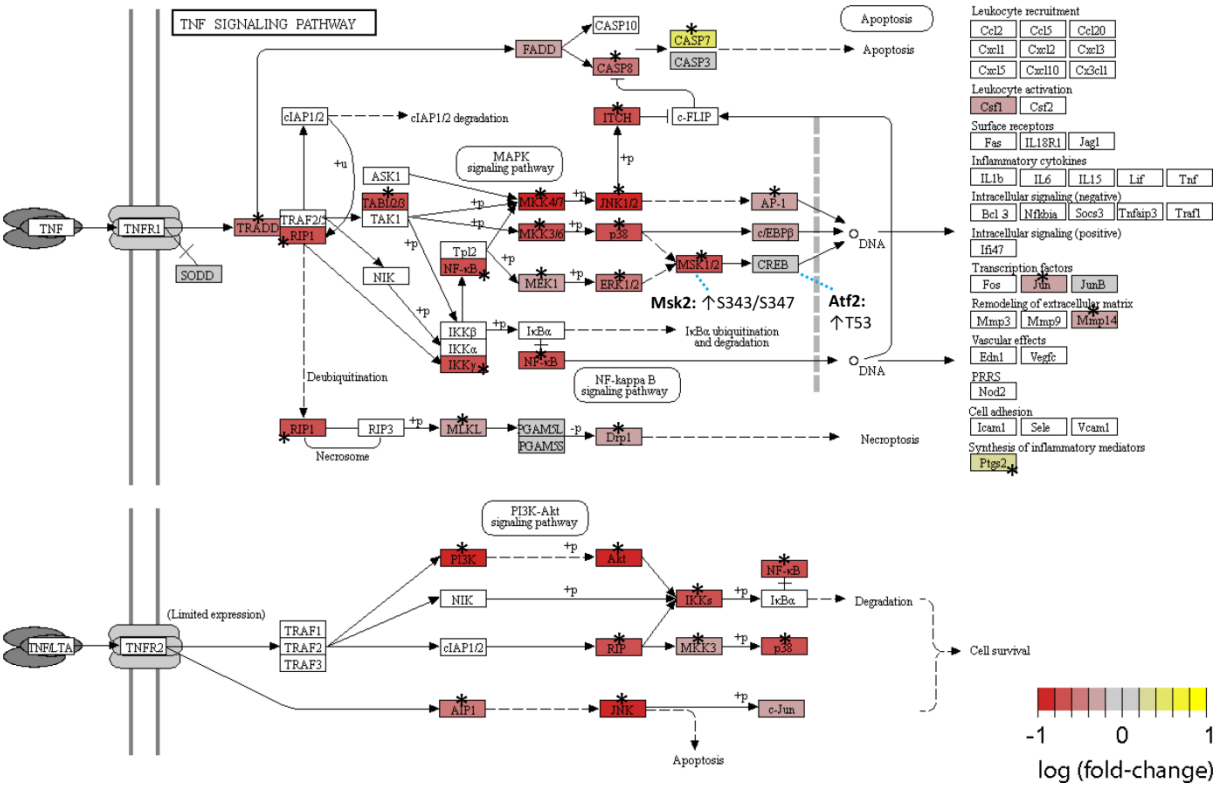


Figure 48: Modulation of TNFα signalling upon MCMV infection. Box colours represent the average log fold-change of proteins and asterisks indicate statistical significance. Altered protein phosphorylation levels were indicated isoform-specifically if regulation of the phosphopeptide was statistically significant (q ≤ 0.05). An enlarged map is in the appendix (section 10.4).

Notably, apoptosis signalling was an associated pathway not fully affected by the global downregulation (Figure 48, Figure 49). Caspase 8 (Casp8) was downregulated, Casp3 was found stable and Casp7 was even slightly but significantly upregulated indicating a more complex regulation pattern in this pathway independent of the TNFα and MAPK signalling shutdown. To investigate this manipulation in more detail, a KEGG pathway map for caspase-mediated apoptosis was generated as described in the next section 6.6.5.2.

In summary, the observed systematic downregulation of positive regulators (numerous MAPKs on all levels) and stabilization of the negative regulator Dusp3 emphasizes a comprehensive shutdown of MAPK signalling on the protein level. Notably, protein phosphorylations appeared upregulated. In addition, superordinated TNFα signalling displayed a concomitant downregulation of associated pathways, such as NF-κB and necroptosis signalling. In contrast, p53 as well as the apoptosis mediators Casp3 and Casp7 were not affected by the downregulation.

6.6.5.2. Manipulation of apoptosis and cell cycle signalling

In a next step, the data set was examined for regulations in apoptosis signalling. Unlike for MAPK and TNFα signalling, a substantial fraction of proteins was found upregulated including Cathepsin (isoforms D and L1, but not B and Z), Bcl2-associated agonist of cell death (Bad), Diablo, Casp7. The remaining proteins did not exhibit any regulation or were found downregulated such as eukaryotic translation initiation factor 2A (Eif2a), BH3-interacting domain death agonist (Bid), apoptosis regulator BAX (Bax), Bcl-2 homologous antagonist/killer (Bak), cytochrom C, apoptotic protease-activating factor 1 (Apaf-1), apoptosis-inducing factor (Aif), E3 ubiquitin-protein ligase XIAP (XIAP) and Casp3. Diminished phosphorylation was detected on the inositol 1,4,5-trisphosphate receptor (Itpr3, S1832) and on the Pdpk1 isoform of 3-phosphoinositide-dependent protein kinase-1. Bad was detected with elevated S155 phosphorylation levels (Figure 49). This rather complex regulation pattern renders the apoptosis pathway as target for extensive functional studies. A first explanation might be an equilibrium of the host's attempt to induce apoptosis and MCMVs countermeasures to attenuate induction of cell death.

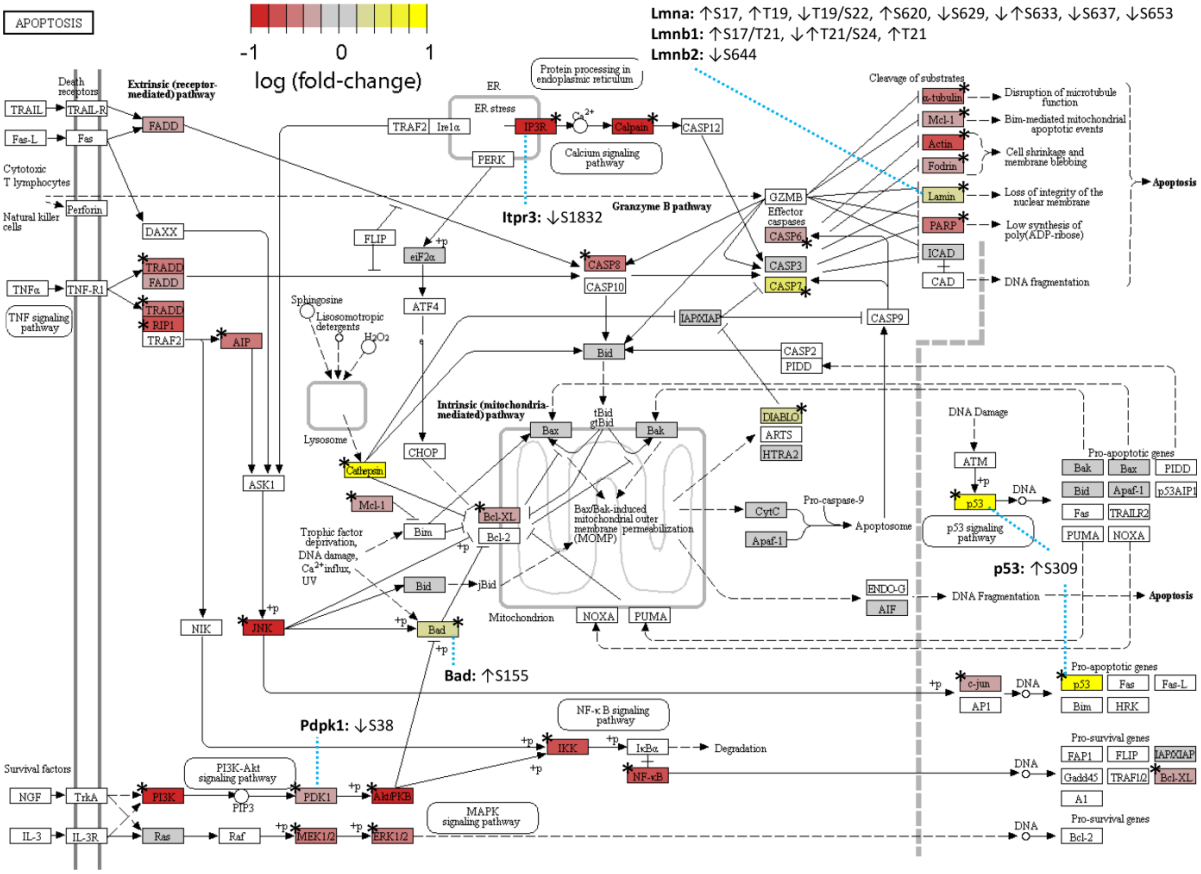


Figure 49: Modulation of apoptosis signalling upon MCMV infection. Box colours represent the average log fold-change of proteins and asterisks indicate statistical significance. Altered protein phosphorylation levels were indicated isoform-specifically if regulation of the phosphopeptide was statistically significant (q ≤ 0.05). An enlarged map is in the appendix (section 10.4).

As already expected from the GO-term enrichment analysis (section 6.6.4.1), systematic screening for proteins involved in cell cycle regulation revealed tremendous regulation. Downregulation was observed for various key-proteins such as cycline D (Ccnd1), the cyclin-dependent kinases (Cdk4, Cdk6), Mcm3, Mcm4, Mcm5, Mcm6. A small fraction of proteins was significantly upregulated including Cdk1, Cdk2, cycline A (Ccna2), Kip1/2 (Cdkn1b) and the 14-3-3 family member Ywhae. In addition, differential phosphorylation levels were found for the anaphase-promoting complex (APC/C) isoform Cdc23 (S588/T596), for Ywhae (S210), Kip1 (S178), transcription factor E2f4 (E2f4, T14), Smc3 (S1067) and numerous sites of Mcm2 (S26/S27) and Mcm3 (S672/S681 and T729/S732) (Figure 50). Notably, downregulation of cell cycle regulators during the course of MCMV infection has been shown on the transcript level,[218] however, a dependence of viral gene transcription and cell cycle is only known for

HCMV.[219] To understand these species-specific differences in manipulating the cell cycle, extensive future studies are required.

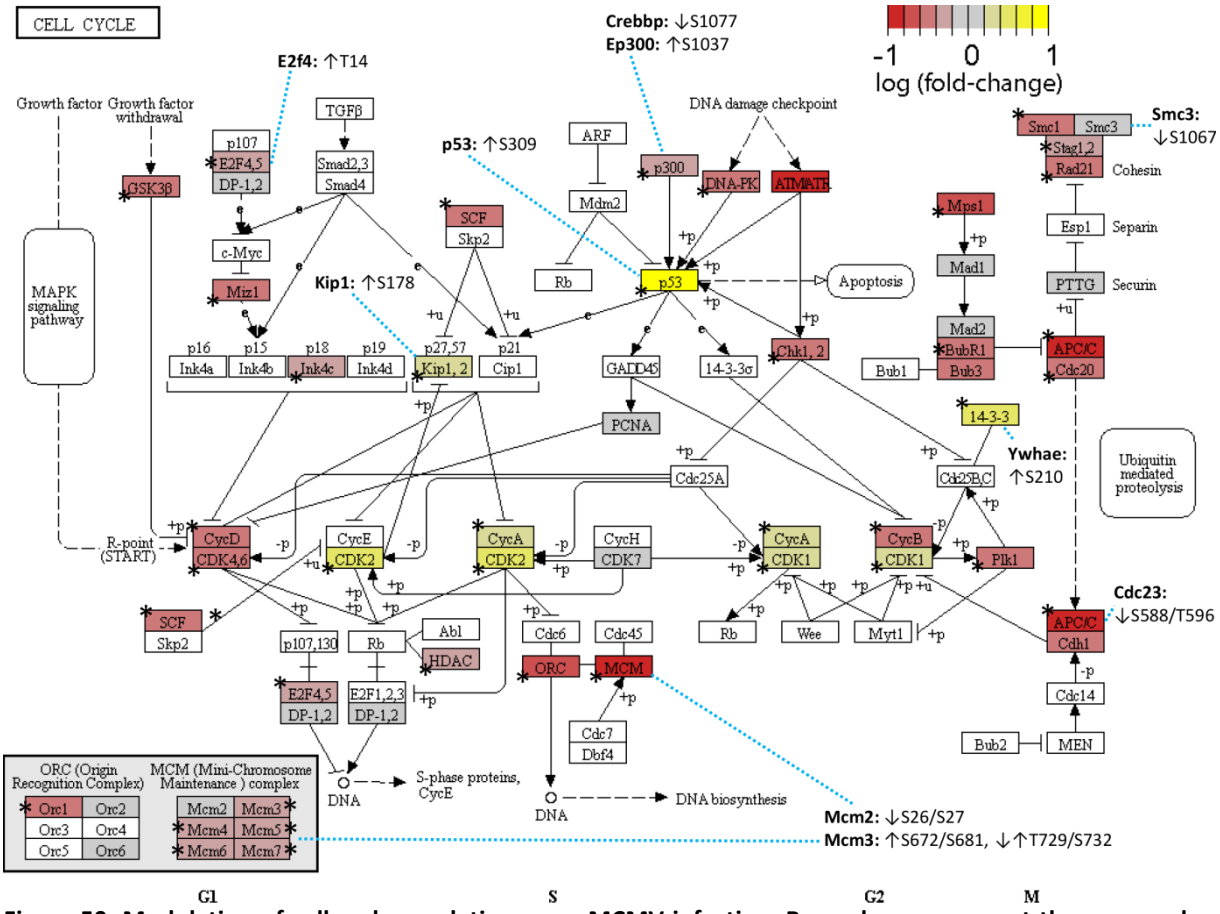


Figure 50: Modulation of cell cycle regulation upon MCMV infection. Box colours represent the average log fold-change of proteins and asterisks indicate statistical significance. Altered protein phosphorylation levels were indicated isoform-specifically if regulation of the phosphopeptide was statistically significant (q ≤ 0.05). An enlarged map is in the appendix (section 10.4).

6.6.5.3. Systemic onset of protein synthesis, protein processing and vesicular transport

In addition to the results of the GO-term enrichment analysis (section 6.6.4.1), String Network analysis of substantially upregulated proteins (≥ 2-fold) revealed systemic increase of proteins involved in

splicing- and ribosomal activity (Figure 51). As this upregulation points towards enhanced protein synthesis capabilities of the cell, the KEGG-pathways “protein processing in endoplasmic reticulum” and “SNARE proteins in vesicular transport” were used for visualizing regulations in these pathways. Notably, both can be regarded as downstream processes of synthesis, as they involve protein folding, assembly and glycosylation in ER as well as trafficking (vesicular transport between cellular compartments).

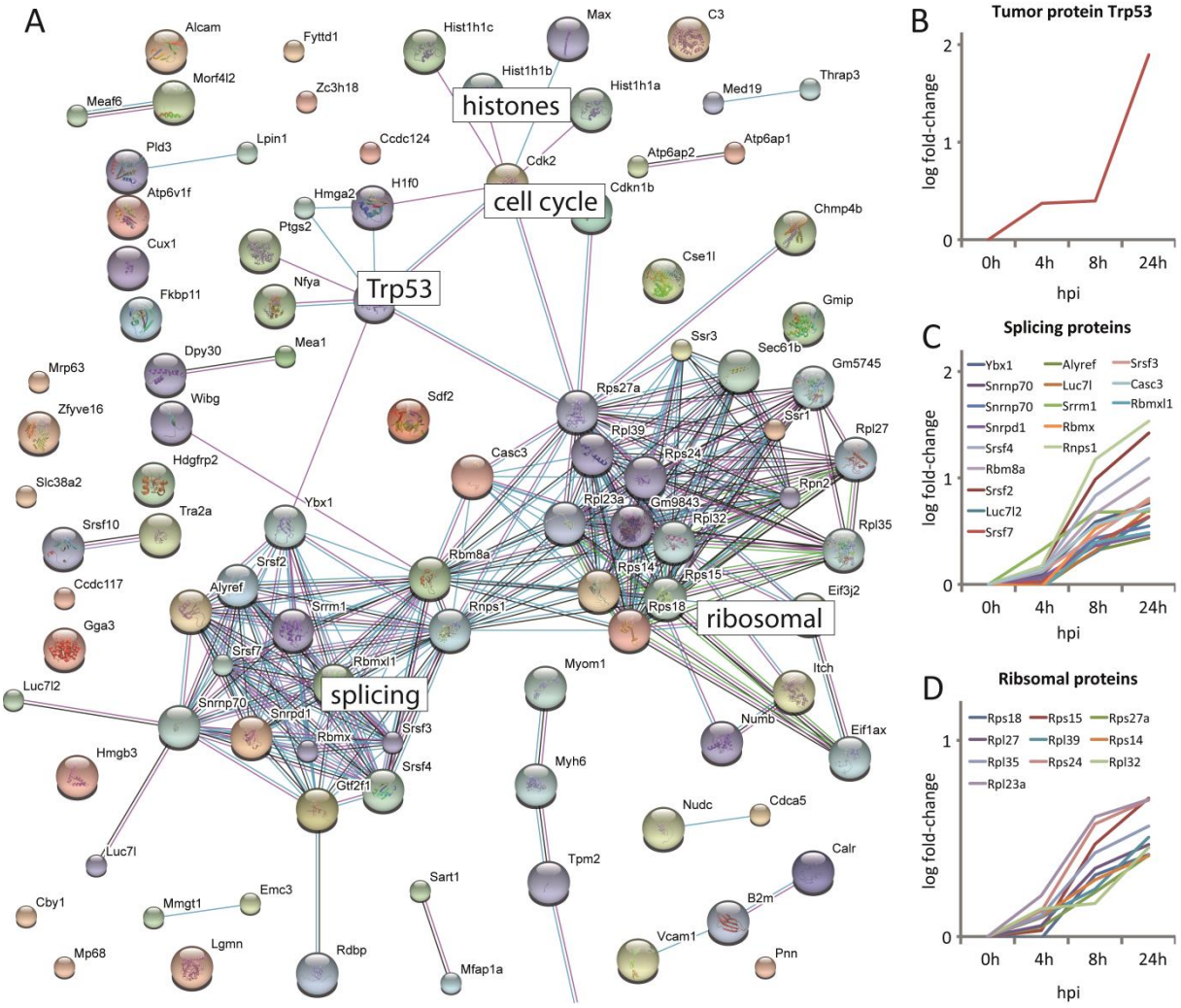


Figure 51: String network analysis of 281 significantly upregulated proteins in fibroblasts upon MCMV infection (A). Trp53, a key-player of cellular signalling, was strongly upregulated (B) and might mediate upregulation of RNA-processing proteins involved in mRNA splicing (C) and ribosomal proteins (D). Settings were “confidence high” and “no textmining”.

Among proteins linked to "protein processing in ER" numerous key-players connected to glycosylation and folding, were found upregulated, such as Sec61 (isoform Sec61b, but not Sec61a1), members of the oligosaccharyltransferase (OST, isoforms Rpn1, Rpn2, Dad1, Ddost, but not Stt3a and Stt3b), hypoxia up-regulated protein 1 (NEF, isoform Hyou1), heat shock protein 40 (Hsp40, isoforms Dnajb11, Dnajc1, Dnajc3, but not Dnajc10), calreticulin (Crt, Calr) and endoplasmic reticulum mannosyl-oligosaccharide 1,2-alpha-mannosidase (isoform Man1a2). However, other proteins were predominantly found to be stable except for mannosyl-oligosaccharide glucosidase (GlcI, isoform Mogs), which was found slightly downregulated at 8 hpi (Figure 52, top-left). In addition, numerous dephosphorylation events were observed in these functional groups such as S14 and S17 of Sec61b, S650 of Hspa5, S306 of Hsp90b1 (isoform of GRP94), S582 of calnexin (Canx) and S158 of wolframin (Wfs1), pointing towards global dephosphorylation.

In stark contrast, members of the unfolded protein response (UPR, section 3.4.4) were widely downregulated including PERK (isoforms Eif2ak2, Eif2ak4), ATF6 (isoform Atf6b) and Wfs1. Concomitantly, proteins involved in lysosomal degradation, which reduces ER stress by decreasing synthesis product levels, were found downregulated such as the NEF isoforms Hsph1, Hspa4l (but not Hspbp1, Bag1, Bag2), p97 (isoform Vcp), Npl4 (isoform Nploc4) and DOA1 (isoform Plaa). In contrast, DSK2 (isoform Ubqln1, Ubqln2, but not Ubqln4) and RAD23 (Rad23b, but not Rad23a), both involved in proteasomal degradation,[220] were upregulated. Surprisingly, no aberrant phosphorylation was detected in this branch of the KEGG map, except for a downregulation of S158 on Wfs1.

Concurrently, proteins involved in anterograde vesicular transport from the ER to Golgi network and plasma membrane were detected with a slight but significant upregulation including synaptosomal-associated protein 23 (Snap23), vesicle-associated membrane proteins (Vamp2, Vamp3), syntaxins (Stx6, Stx7, Stx18), protein transport protein Bos1 (Bos1) and BET1 homolog (Bet1). In addition, aberrant phosphorylation levels were detected on Stx7 (S126/S129 and S129), Vamp4 (S90) Snap23

(S110) and Stx4 (S117). Only Stx8, which is predominantly associated with endosomes,[221] and Use1 and Sec20, which are involved in retrograde transport,[222, 223] were found downregulated indicating diminished transportation from ER back to the Golgi and the plasma membrane.

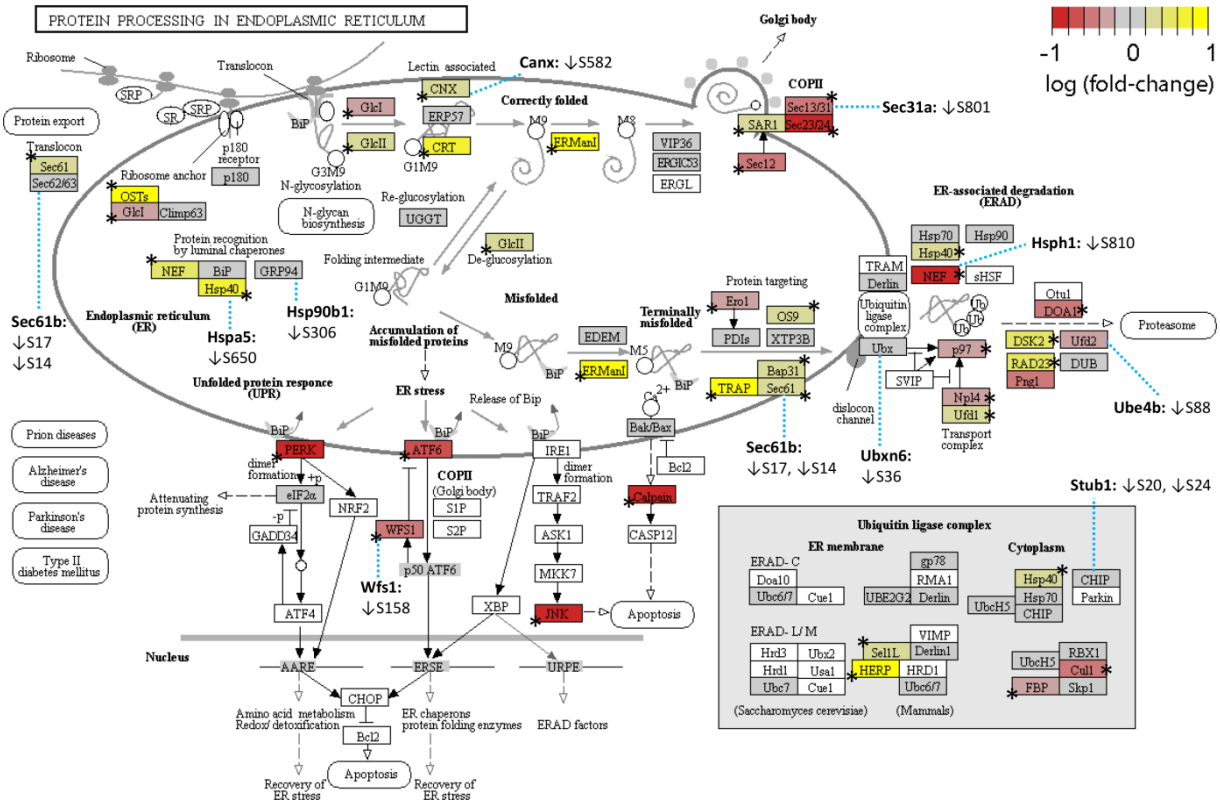


Figure 52: Modulation of the protein processing machinery upon MCMV infection. Box colours represent the average log fold-change of proteins and asterisks indicate statistical significance. Altered protein phosphorylation levels were indicated isoform-specifically if regulation of the phosphopeptide was statistically significant (q ≤ 0.05). NEF in glycosylation (top-left) and NEF in lysosomal degradation (right) differ by isoforms. An enlarged map is in the appendix (section 10.4).

To summarize this analysis, MAPK and TNFα signalling were systematically downregulated (on protein level) whereas MAPK phosphorylation levels appeared increased. Concomitantly, a substantial upregulation of the splicing machinery, the translation machinery, the protein processing capabilities and the anterograde trafficking system was observed. The unfolded protein response and the retrograde trafficking systems appeared shutdown. Interestingly, regulation patterns in cell cycle and apoptosis signalling were extremely complex an did not allow to draw a conclusion about system wide effects of these manipulations.

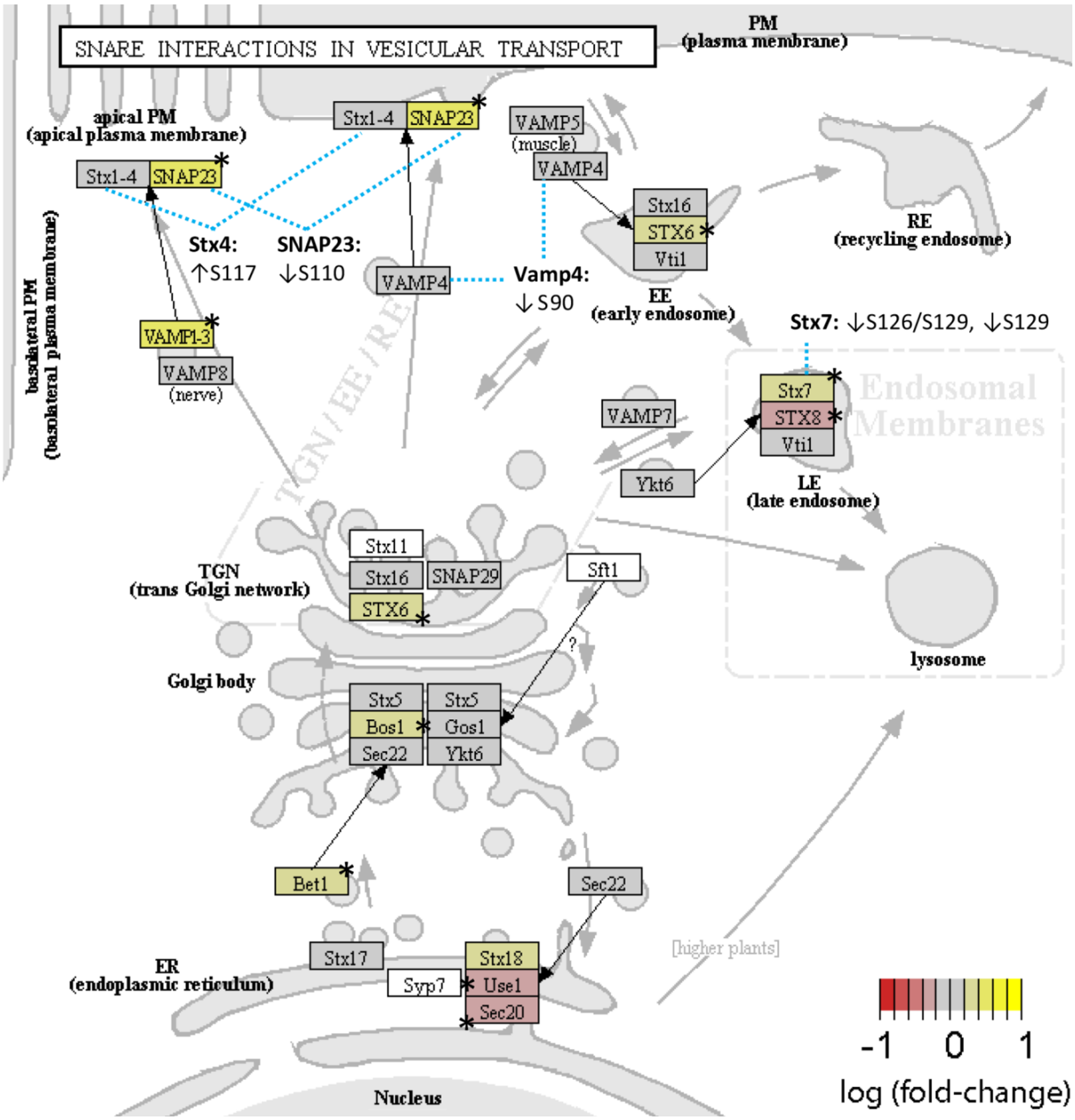


Figure 53: Modulation of the vesicular transport machinery upon MCMV infection. Box colours represent the average log fold-change of proteins and asterisks indicate statistical significance (q ≤ 0.05, limma or rank product). Altered protein phosphorylation levels are indicated isoform-specifically if regulation of the respective phosphopeptides was statistically significant (q ≤ 0.05, limma or rank product). An enlarged map is in the appendix (section 10.4).

7. Discussion

7.1 ERLIC as a powerful tool for phosphoproteomics

In this work, ERLIC was systematically evaluated for LC-MS-based phosphoproteomics with a special focus on sensitive detection of protein phosphorylation events in complex biological samples. Using both discovery and targeted approaches, ERLIC has proven to be an easy-to-establish method requiring a simple HPLC setup including autosampler, micropump and UV-detector. In addition, automatic fraction collection might be employed to reduce laboratory time when processing larger sample cohorts (section 6.5). As demonstrated, ERLIC-HPLC allows for robust sample processing owing to its reproducibility in terms of RT stability, but also of the subsequent SPE-step either via SCX or RP (Loroch et al.)[124]. Quality control can be performed by injecting single phosphopeptides for monitoring retention time and peak areas via UV-detection (214 – 280 nm). As a drawback, the strong UV-absorption of the ERLIC buffers (20 mM sodium methylphosphonate / 200 mM triethylammonium phosphate) requires injection of comparably high phosphopeptide amounts when employing columns of 4.6 mm I.D. and flow rates of up to 1 mL/min. Hence, UV-based determination of RT via injection of SIL peptides in the upper pmol or lower nmol range (experiment 1-5, section 6.5) requires subsequent cleaning of the HPLC-system as even a slight carry-over in the ppm-range might interfere with quantification usually conducted in the lower fmol range or below (section 6.5). Notably, to overcome this problem, two non-isobaric standards (different isotope compositions) might be employed– one for retention time evaluation (quality control), the other as spike-in for quantification.

In a systematic comparison with a well-established TiO_2-affinity purification protocol, the employed ERLIC set-up has demonstrated unbiased quantitative recovery rates of 62 % (median) for any phosphopeptide. In contrast, TiO_2-based enrichment exhibited an average quantitative recovery of only 39 % with poor recovery for a substantial fraction of phosphopeptides (short and low AACV). This bias

of recovery for certain phosphpeptide species emphasizes that binding to the TiO_2 surface cannot only be explained by complexation, but emphasizes that other mechanisms take place: (I) a HILIC mechanism due to the high percentage of ACN in loading and washing buffers,[224] (II) a displacement mechanism by adding glycolic acid to supress unspecific binders, which also affects acidic, glycosylated or multi-phosphorylated peptides[81, 225] and (III) a cooperativity of molecule adsorption – a process barely understood yet but well-investigated by Eriksson et al..[226] This multiplicity of interfering effects, renders unbiased recovery extraordinary challenging or even impossible and might explain the long time required for establishing such protocols. In addition, adjusting bead-to-sample ratio is quite challenging for low sample amounts and represents an additional source for errors. Consequently, controlling the conditions and the recovery upon TiO_2-based enrichment appears less promising when working with minute amounts of sample.

Table 11 summarizes some key-results and key-conclusions about enrichment using ERLIC and TiO_2 and provides a guide for the decision which method to choose for certain experimental setups. The following sections provide an even-handed discussion of both methods.

Table 11: Summary of key-findings and key-conclusions for the use of ERLIC and TiO_2 in LC-MS-based phosphoproteomics

	TiO_2	**ERLIC**
Workflow / Setup	TiO_2-particles in reaction tubes, elution in self-made microcolumns, RP-SPE for desalting	Standard HPLC set-up with anion-exchange column, SPE for desalting (RP) or further enrichment (SCX)
ID rate (discovery)	793 IDs / h	655 IDs / h
DRF rate (discovery)	0.085 DRFs / h	0.103 DRFs / h
Quantitative recovery	~40 % (without RP-SPE), recovery drastically varies with physicochemical properties of phosphopeptides	~60 % (including RP-SPE), recovery independent of physicochemical properties of phosphopeptides
Enrichment specificity	Up to 99 %, depending on the use of competitors (phtalic acid, DHB, glycolic acid) and optimization of bead-to-sample ratio	Up to 85 %, depending on the fractionation window and fraction size
Recommended starting material	Higher sample amounts (≥ 100 µg)	Low sample amounts (≤ 100 µg) and ultra-low sample amounts (≤ 10 µg)
Potential for targeted phosphoproteomics	The higher complexity of the phosphopeptide-enriched fraction might interfere with quantification of target phosphopeptides.	The complexity of fractions depends on the RT and the fraction size. The RT can be determined using synthetic reference standards.

	TiO_2	ERLIC
Potential for automation	Repeated spin-down cycles and removal of supernatant render automation in 96-well plates rather challenging.	High potential in conjunction with automated fraction collection and 96-well SPE. ERLIC-SPE (Zarei et al. [128]) in 96-well format represents a powerful alternative for well-retained phosphopeptides.
Expenses (chemicals / consumables)	Costs depend on the amount of sample to be processed: ~11:1 TiO_2-bead-sample-ratio (w/w). 1x RP-SPE tip per sample	Costs are independent of sample amount. 39 mL ACN + 83 mg of methylphosphonic acid / h (4.6 mm column). 1 x RP-SPE tip per fraction.

7.1.1 ERLIC and TiO_2 for large-scale discovery approaches with minute amounts of sample

Despite these aforementioned disadvantages (section 7.1), a fine-tuned TiO_2-protocol allows for excellent enrichment specificity of up to 99 % (phosphopeptide IDs / total IDs), when employing competitors such as dihydroxybenzoic acid, phtalic acid[227] or glycolic acid[104] in the loading buffer, which strongly supresses unspecific binding. Hence, TiO_2 is an excellent choice for large-scale discovery approaches when aiming at high identification rates. Experiment 1-4 (6.4) demonstrated that TiO_2 exhibits higher phosphopeptide identification rates with respect to acquisition time (793 IDs / h and 739 quantified / h), approximately 20-25 % higher than with ERLIC (655 IDs / h and 598 quantified / h). If considering total instrument time (including wash blanks), TiO_2 exhibits even higher ID rates – approximately 3-fold higher than ERLIC. It is important to note that LC-MS gradients of experiment 1-4 were adjusted to an appreciable sampling rate (~ 75 % of the MS duty cycle, section 6.4) to have an unbiased comparison which does not favour any method. However, despite the higher ID rates, the experiment 1-4 indicates that detection of differentially regulated phosphopeptides is substantially higher with ERLIC (+80 %). This is most likely attributed to the improved recovery and the resulting higher signal intensities (section 7.1). Especially upon correction for multiple testing (Benjamini-Hochberg), ERLIC allowed the identification of considerably more regulated phosphopeptides as the lower p-values (Student's t-test) have higher chances to pass the correction.

Another reason for improved identification of regulated phosphopeptides might be the flexibility of ERLIC to adjust the fraction size and hence, adjust the complexity of the sample for LC-MS to better control interference. In contrast, reducing complexity after TiO_2 would require an additional HPLC- or SPE-step for robust fractionation which introduces additional sample loss. As an alternative, a sequential elution from TiO_2-beads is possible, but requires careful optimization to achieve a homogenous fractionation. In either case, the problem of poor recovery of certain phosphopeptides persists and could probably only be avoided at the expense of reduced specificity as this would require to remove the displacer from the binding buffer.

7.1.2 Complementarity of ERLIC and TiO_2 and future potential

Overlaps of ERLIC and TiO_2 were extremely low (< 5 % for detection of differentially regulated phosphopeptides) highlighting the frequently observed and well-exploited complementarity of enrichment methods[113, 126, 228-230] (extensively reviewed for instance by Loroch et al. or Pagel & Loroch et al.).[14, 25] Hence, for experiments with very limited sample material such as biopsies (≤ 10 µg protein), ERLIC might be the first choice, but might be complemented by TiO_2 if additional sample material is available (≤ 100 µg). In case of higher amounts of starting material (> 100 µg), when lower recovery can be tolerated, TiO_2 might be the first choice for a robust quantification with excellent ID-rates – especially when combined with extensive fractionation via SCX,[119, 231] HILIC[104] or high-pH-RP.[96] Alternatively, ERLIC might be exploited for fractionation after TiO_2 even with volatile salt-buffers. Even though volatile salt buffers have been demonstrated to exhibit lower selectivity for phosphopeptide enrichment,[89] fractionation would probably not be negatively affected, as the majority of non-modified peptides are already depleted. Notably, ERLIC has already been successfully described for fractionation of different samples.[84, 232] Vice-versa, in case of a TiO_2-ERLIC workflow, the high percentage of TFA from the TiO_2 loading and washing buffers might interfere with the HILIC mechanism in ERLIC and lead to additional retention of positively-charged analytes (basic peptides). Hence, a careful removal of

residual TFA might be necessary. Vice-versa, ERLIC fractions might be also subjected to a simple TiO_2 enrichment to further increase specificity, which would be especially interesting for early fractions, which comprise a high share of co-eluting non-modified peptides. In such cases, TiO_2 might not even require a displacer as phosphorylated peptides are already enriched. Notably, even the Na-methylphosphonate from the ERLIC buffer A might suffice as displacer to boost specificity.

In summary, even though TiO_2 exhibits higher identification rates, identification of differentially regulated phosphopeptides with minute amount of samples seems more promising with ERLIC (~10-50 µg). However, if additional sample material is available (~50-500 µg), TiO_2 might be used to complement the findings with ERLIC. Only if large amounts of sample are available (> 500 µg), and lower quantitative recovery can be tolerated, TiO_2 might be the method of choice because of the excellent ID rates.

7.1.3 ERLIC for targeted phosphoproteomics

In comparison to discovery approaches, aiming at large-scale identification and quantification of as many phosphorylation events as possible, a targeted analyses aims at a precise quantification of a pre-known selection of phosphorylation events. In a very well-established targeted assay, a maximum of 100 selected protein phosphorylations might be quantified spanning a dynamic range of several orders of magnitude usually down to the amol-range. For such assays, the phosphopeptide-enrichment method might be selected according to several aspects (I) the quantitative recovery of the target phosphopeptides of lowest signal intensities (II) the complexity of the resulting sample subjected to LC-MS to reduce interference of co-eluting isobaric ions and (III) the potential for processing larger sample cohorts.

For targeted quantification of EGLN1 S125 phosphorylation (~10 amol / µg digest) in colon cancer patients (section 6.5), ERLIC was chosen for two reasons. First, because of the higher estimated recovery based on the low amino acid composition value of -3 (section 6.3.). Second, because screening

of in-house datasets (predominantly with HeLa cells) revealed that this phosphopeptide has been already identified in ERLIC-based large-scale discovery approaches, but never with TiO_2. Using ERLIC, the target phosphopeptide was reproducibly purified from 20 patient tissue samples in consecutive HPLC runs. Reproducible processing with TiO_2 would have been comparably challenging, because of the consecutive loading, washing and elution steps, which render TiO_2 with multiple samples extremely labour intense. A well-established TiO_2-protocol in 96-well format would represented a powerful alternative – but only, if quantitative recovery is sufficient. For decisions whether to apply ERLIC, large-scale phosphopeptide RT-libraries might be generated to evaluate if the specific RT of a target phosphopeptide allows for appreciable purification. This is especially important for phosphopeptides of low signal intensities when isobaric and co-eluting peptides (which quench a target signal or lead to interference) need to be widely depleted. However, in ERLIC the complexity of fractions roughly correlates with retention time and can therefore be estimated. In contrast, TiO_2-eluates exhibits a constant complexity (without additional fractionation). Hence, if a phosphopeptide elutes early in ERLIC, it might be rather purified by TiO_2. In contrast, ERLIC represents the method of choice for well-retained ones.

7.2 Proteomics and phosphoproteomics analysis of cytomegalovirus-infected fibroblasts

This large-scale study represents the first comprehensive, time-resolved characterization of the proteome and phosphoproteome during the course of MCMV replication in fibroblast cells. The combinatorial exploitation of ERLIC and high-pH RP, as economic strategy with minimal sample loss, allowed an appreciable depth of analysis (complete time courses for 6,618 proteins and 2,311 phosphopeptides) Hence, at an estimated presence of ~10,000 proteins in a cell, the study comprises quantitative information for nearly 70 % of the proteome and qualitative information for almost 90 %. As the analysis was successfully conducted with comparably low sample amount (120 µg per condition),

the workflow has excellent potential for future biomedical and clinical studies with restricted sample material such as mouse or patient tissue (usually << 1 mg protein). The dataset gives deep insights into viral protein expression pattern, viral protein phosphorylation and global re-organization of the host cell during the course of replication and serves as repository for evaluation of MCMV as mouse model for HCMV. The depth of the proteome analysis allows for a systematic comparison with other large-scale datasets previously published.[94, 185, 190] In the next sections, results are discussed with major focus on (I) temporal profiling of viral protein expression and protein phosphorylation (II) systematic comparison with HCMV data from literature, (III) alteration of signalling and cellular functions of the host cell and (IV) unveiling potential key-markers and processes which can be used as starting points for future functional studies and as drug targets.

7.2.1 Analysis of viral protein expression and phosphorylation

The dataset provides evidence for the translation of more than 100 viral proteins during the course of MCMV infection. It comprises temporal profiles for 95 proteins (irrespective of missing data points) and complete time courses for 63 proteins (data points from 3 biological replicates at every time point). Based on these time courses, proteins were clustered into IE, E, LL and TL, but it is important to note that this classification does not fully match to the traditional classification based on assays, which detect transcript levels relative to cellular events. This might partially attributed to the insufficient temporal resolution of this study (0, 4, 8, 24 hpi), which did not allow to differentiate between onset of immediate early and early genes (which were both already detected at 4 hpi). On the other hand, protein level detection was somewhat expected to yield a different classification, as transcript level detection does not reflect translational control and protein degradation. However, these mechanisms might be investigated in more detail by integration of transcriptomics and proteomics.

Profiles of the 4 clusters in the present study (Figure 38) revealed high degree of similarity to profiles of the clusters described by Weekes et al. using HCMV-infected human fibroblasts (HFF cells, Figure 54), which were structured by k-means clustering in either four or five cluster. The first cluster of the Weekes et al. data set (TP1) showed an early maximum at 12-16 hpi, followed by decrease in protein expression at later time points, which is similar to the IE-cluster with a maximum at 4-8 h. The second cluster (TP2) showed a slightly later maximum of expression at 16-24 hpi with a constant protein expression afterwards as observed for the MCMV E cluster with a maximum at 8 hpi. The two remaining clusters (TP3 and TP4) have a clear later onset (48 hpi and 72 hpi) similar to the LL and TL cluster with a dominant onset between 8-24 hpi and 24 hpi. Notably, the HCMV protein expression is approximately 3-4 times slower, according to the decelerated replication cycle,[158] but temporal profiles exhibit an appreciable level of homology (24 hpi in MCMV ≈ 72 hpi in HCMV).

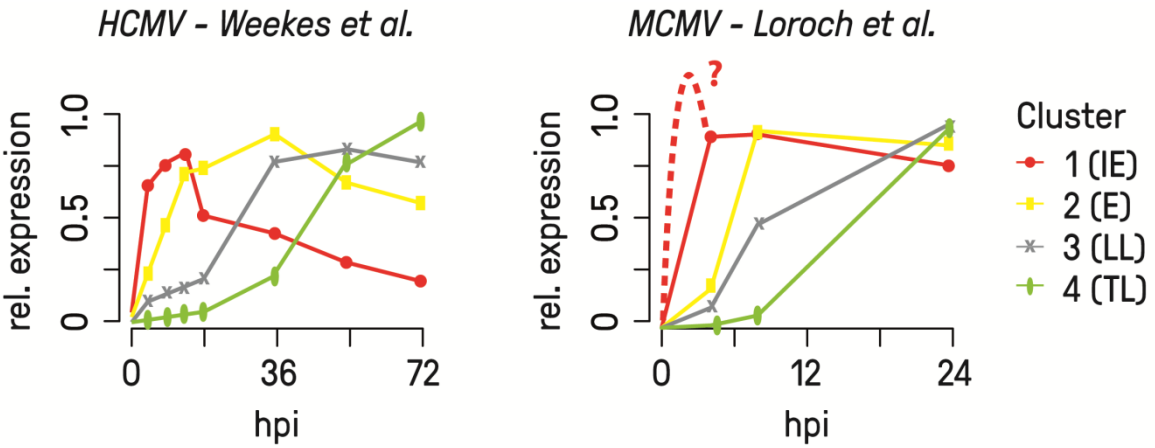


Figure 54: Temporal profiles of viral protein expression based on the dataset from Weekes et al. (HCMV) in comparison to the present dataset (Loroch *et al.*). The first cluster (immediate early) of HCMV proteins (immediate early) exhibits a rapid decrease after the initial maximum which might not have been resolved in this study (red dashed line) without time points between 0 and 4 hpi. Besides, all clusters show highly comparable profiles when considering the 3-fold decelerated replication course of HCMV.

IE proteins: Surprisingly, the well-characterized IE1-IE3 proteins clustered with the serine/threonine kinase M97, which was not reported to be of high abundance at early time points in previous studies with respect to transcript levels.[233] Weekes et al. reported the HCMV homologue UL97 rather as late, but also their dataset display appreciable UL97 levels at the earliest time point (6 hpi). Notably,

Oberstein et al. reported UL97 to be of early kinetic.[190] Interestingly, numerous phosphorylations at the heavily phosphorylated N-terminus of M97 were found drastically regulated between 8 and 24 hpi (S2/S13, S107/T111, T6, S86) and might represent key regulator events of M97 activity or function. The dataset of Oberstein et al. displays similar dynamic phosphorylation of UL97, but not in such close proximity on the N-terminus.[190] As the viral protein kinase (M97/UL97) has a central role in all herpes viruses, modulation via phosphorylation (potentially auto-phosphorylation)[234] might be a central regulatory instance of activity or specificity. Besides, the other proteins in the IE cluster, m31, m84, m152 and m164, are partially involved in immune evasion and exhibit potential for developing vaccines.[235-237] Notably, in comparison to other proteins in the IE cluster (such as IE1, IE3 and m84), m31, m152 and m164 were found less expressed at 4 hpi and therefor might be rather of early kinetic. A comparison with the dataset from Weekes et al. was not possible, as m152 and m164 have no sequence homologue in HCMV[238] and UL31 (M31 homologue) was not found.

E proteins: The US22-family members, m25.1, m139, m140, m141, m143 were detected in the E cluster with a maximum at 8 hpi. This is in good accordance to the literature as US22 gene product have been reported to exhibit early kinetics even though transcripts of m142 and m143 follow an IE kinetic.[157, 239] Interestingly, the US22 family member m142, which forms a complex with m143 to bind dsRNA and prevent protein kinase R-mediated synthesis shut-off,[204] exhibited a late kinetic with a clear maximum of expression at 24 hpi. In addition, multiple phosphorylation sites were detected on m143 close to the C-terminus (sites S425 and S503), indicating a regulatory function, but no sites were detected on m142. Notably, the human cytomegalovirus dsRNA binding proteins IRS1 and TRS1 (HCMV) were found with one phosphorylation site each in the dataset of Oberstein et al..[190]

LL and TL proteins: M36, another US22 family member and major suppressor of cell death by interacting with procaspase 8,[240] was found in the LL cluster with a clear maximum at 24 hpi. Similar kinetics were observed for M45, known to block Bax-mediated apoptosis signalling, and M38.5 an

inhibitor of RIP3-mediated necroptosis.[158] These expression profiles indicate increasing necessity to maintain cell survival at late time points. Notably, the HCMV homologue UL36[241] and UL37[242] in the dataset of Weekes *et al.* were found with comparably early maxima (24 hpi) followed by a decrease in protein expression. Either additional key-player complement the action of UL36 and UL37 in HCMV at later time points or both together are very potent cell death-inhibitors (compared to M36, M45 and M38.5 in orchestration with m41.1 - not detected in this study)[158] and facilitate complete shut-off already at earlier time points. Notably, m45 was detected with two upregulated phosphopeptides at 24 hpi (S209, S209/T214), indicating modulation of function during the course of infection. This phosphorylation might be connected to M97, as UL45 phosphorylation (S123/S124) appears to be UL97-mediated (Oberstein et al.).[190]

The late clusters display some discrepancies to the traditional classification. For instance, M44 and M57 are involved in viral DNA replication rendering them not true late in the traditional sense. However, in the proteomics dataset by Weekes et al., the HCMV homologues UL44 and UL57 were assigned to the latest cluster as well, as protein expression profiles clearly matched to true late proteins. The LL cluster predominantly comprised proteins known to be involved in immune-evasion such as m04,[243] m18,[244] m27[245] and viral DNA-synthesis such as m54,[246] m56.[247] In addition, a few virion-associated proteins, M25, m69 (tegument) and m166.5 (potentially) were in the LL-cluster, even though most tegument proteins (M32, M83, M99) and capsid proteins (M85, M86)[248] were found in the TL-cluster. These results are in good accordance with the dataset from Weekes et al., describing the homologue HCMV tegument proteins UL32[248] and UL83[249] to have a later kinetic than UL54,[250] UL56,[247] UL69.[251] Notably, all virion-associated proteins in the LL cluster exhibited an appreciable degree of phosphorylation (15 sites each on m25 and m69 and 2 sites on m166.5), whereas only a few phosphorylation sites were detected on tegument and capsid proteins in the TL cluster (3 sites on M32 and 1 site on M85). This allows to speculate that LL kinetic might favour phosphorylation, as most kinases might be already too

extensively downregulated at later time points. The Oberstein et al. dataset displays one UL97-dependent phosphorylation on the HCMV homologues UL25 and UL69. However, both proteins including M25 from this dataset were detected with differential tyrosine phosphorylation rendering an indirect phosphorylation by a host kinase more likely than from UL97/M97 (serine/threonine kinase).[190] Interestingly, M25 and M69 were detected with aberrant phosphorylation levels at 24 hpi, hence, they represent interesting targets for future functional studies on virion assembly.

In summary, temporal profiling gave appreciable insights into the orchestration of viral protein expression. The dataset shows a high degree of similarity to HCMV proteins in human fibroblasts based on the large-scale dataset by Weekes et al., but showed discrepancies to the traditional classification. Differential phosphorylation analysis of proteins such as M97 or m143 (but not its interactor m142) revealed new targets for functional studies. Numerous viral protein phosphorylation events might be connected to M97 activity, indicated by comparable UL97-mediated phosphorylation events in the Oberstein et al. dataset. As phosphorylated regions (Figure 40) are likely to represent key-regulatory elements of viral protein activity, subsequent point mutation or truncation experiments would allow a more detailed investigation.

7.2.2 Analysis of differential protein expression and protein phosphorylation in MCMV-infected fibroblasts

Global proteomics revealed substantial regulation of the host cell proteome including a wide variety of altered protein phosphorylation events. Volcano plots (Figure 41, Figure 42), interaction networks (Figure 51) and pathway maps (Figure 47-Figure 53) displayed regulation of entire functional networks and signalling pathways. It is important to note, that clustering of the data in conjunction with functional enrichment and pathway mapping revealed pathways to be systemically altered, even though only a minor fraction of proteins would have passed the fold-change cut-off. Hence, data-analysis with respect

to fold-change cut-offs (usually in the range of 2-fold, or with respect to overall variance) might not always be favourable if the perturbation of a biological system affects entire pathways but with moderate regulation. Hence, the data analysis strategy chosen in this work proofed quite beneficial – though as well more labour-intense.

The complex data analysis in this study revealed that altered protein expression and aberrant phosphorylation affected numerous cellular resource and signalling pathways. This is quite likely as herpesviruses substantially reorganize host cells.[129] Among the most important processes (based on this data) are MAPK/ERK- (Figure 47), TNF- (Figure 48), cell cycle- (Figure 50) and apoptosis signalling (Figure 49) and further DNA-replication, protein synthesis and vesicular transport. Noteworthy, also JAK/STAT (Jak1, Stat1, Stat2, Stat3, Stat6) signalling was found widely abrogated (Figure 46), as systematically described by Weekes *et al.* for HCMV[94], but was not further investigated in this work.

Based on the time resolution of this experiment, manipulation of the host cell was divided into phases similar to the viral protein expression clusters (Figure 55). After 4 hpi, when IE proteins and partially E proteins were detected, the first early responders Ifi-204 and calpain-5 were found regulated. In parallel, but more pronounced at 8 hpi, aberrant kinase phosphorylation levels were detected including MAPK pathway members (Tao1, Rps6ka4 syn Msk2). At 8 hpi, MAPK and TNFα signalling appeared systemically downregulated on the protein level (Figure 47, Figure 48, Figure 56). Interestingly, an inhibitor of MAPK signalling, Dusp3 was detected slightly upregulated, emphasizing a tailored shutdown program. This shutdown on the protein level most likely represents the virus' attempt to inhibit host defence, which is mediated via phosphorylation cascades. In parallel, the splicing machinery and ribosomal proteins were detected with considerably higher levels (Figure 51), indicating that transcription and translation is required at this time point for synthesis of TL and partially LL proteins. The vast of nascent proteins requires substantially enhanced protein processing capabilities. Accordingly, a substantial fraction of ER proteins was detected upregulated, including chaperones (Calr,

Hyou1, Dnajb11), translocation proteins (Sec61, Ssr3) and OST proteins (Rpn1, Rpn2). As the concomitant ER stress threatens the ongoing protein synthesis and even host cell survival (UPR-mediated synthesis block or apoptosis),[179, 252, 253] members of the UPR were detected downregulated (PERK and ATF6, Figure 52). Ultimately, as the final step of viral replication, anterograde vesicle transport became upregulated (Snap23, Vamp3 and Bet1), which is most likely attributed to protein trafficking and virion assembly and release. In contrast, the retrograde transport system appeared somewhat downregulated (Use1, Sec21, Stx8), which allows to speculate that this system is dispensable for viral replication (Figure 53).

The next sections discuss this series of observation in more detail including a systematic comparison with two datasets on HCMV-infected fibroblasts: the time-resolved proteome analysis by Weekes et al.[94] and a time-resolved phosphoproteomics analysis for deciphering UL97-mediated phosphorylation events by Oberstein et al.[190]

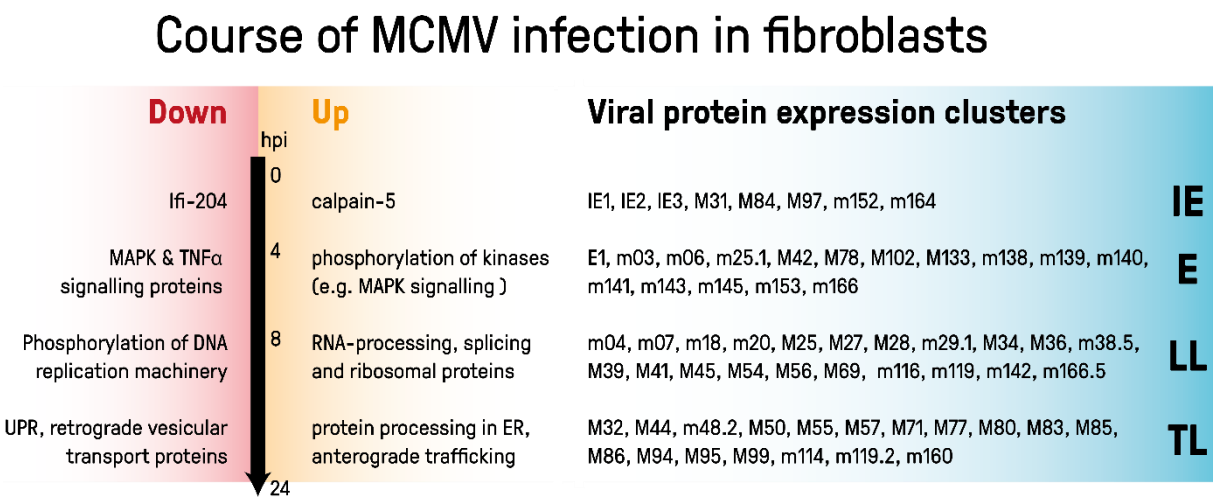


Figure 55: Alterations of cellular signalling pathways and cellular functions during the time course of MCMV infection relative to viral protein expression.

7.2.2.1 Abrogation of Cellular signalling via MAPK-, TNF- and Caspases

MAPK signalling: Proteins involved in MAPK signalling were considerably downregulated in all the three branches: ERK1/2-, JNK/p38 and ERK5. The most diminished expression levels were detected for Mapk8, Mapk9, Mapk14, Map2k4, Map4k4, Map4k5, Mapkapk2 with an average downregulation of ~2-fold. Most notably, the MAPK inhibitor Dusp3 was not downregulated (Figure 47), pointing towards a tailored abrogation of this pathway by suppression of (I) signalling proteins and (II) a concomitant upregulation of a key-inhibitor. These findings are in good accordance with the HCMV dataset from Weekes *et al.,* describing MAPKs systemically downregulated as well (Figure 56), albeit with a higher degree of variation (most likely due to ratio distortion – a well-known problem in quantitative experiments with isobaric labels[70]). In contrast to MCMV, the MAPK inhibitor Dusp3 was downregulated in the HCMV dataset, but Dusp14 exhibited a slight upregulation and Dusp22 was stable. Both proteins are known to inhibit MAPK signalling by dephosphorylation of ERK1/2, but were suspected to dephosphorylate p38 (*syn.* Mapk14) as well. In contrast, Dusp3 has only been described to dephosphorylate ERK1/2 (*syn.* Mapk3 and Mapk1) but not p38[217]. This finding is interesting since ERK1/2 did not exhibit a significant downregulation in both datasets emphasizing inhibition via Dusp3, Dusp14 and Dusp22. Interestingly, p38 was only well-downregulated in MCMV (but not in HCMV, Figure 56) rendering additional inhibition unnecessary. Hence, Dusp14 or Dusp22 might inhibit p38 in HCMV by dephosphorylation whereas inhibition in MCMV might be simply realized by downregulation. Unfortunately, Dusp14 and Dusp22 were not found in the MCMV dataset leaving the question of species-specific orchestration of Dusp isoforms unexplored.

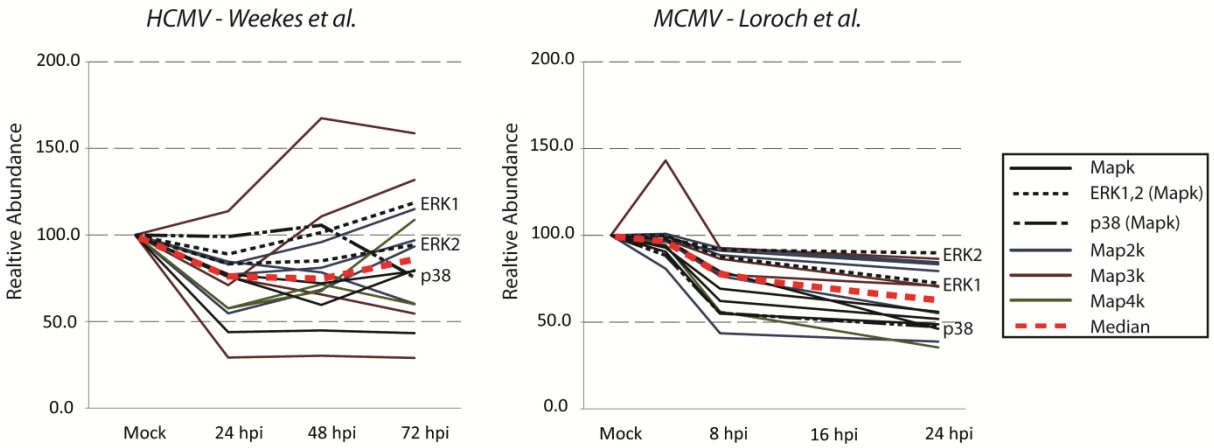


Figure 56: Systemic downregulation of MAPKs in HCMV (based on the dataset of Weekes et al.) and MCMV. In both systems, ERK1 and ERK2 are barely downregulated. However inhibition might be mediated by dephosphorylation via Dusp3 (MCMV) or Dusp14 (HCMV, see text). In contrast, p38 showed only pronounced downregulation in MCMV and might be inhibited by Dusp22 in HCMV (see text).

Noteworthy, Dusp3 has already been reported as upregulated in cervix carcinoma.[254] In conjunction with diminished kinase levels, this might be a biomarker candidate for acute CMV infection. However, attenuated MAPK signalling might be just a short-term effect since *in vivo* experiments from Tang-Feldman et al., indicate MAPK/ERK genes to be considerably upregulated in aortas of MCMV-infected mice 2.5 months post-infection.[255] It might be possible that upregulation *in vivo* reflects the host-response to infection (maybe via TNFα)[256] and experiments on a shorter timescale just reveal the active suppression by CMV. In comparison, also the HSV-1 virus has been shown to induce MAPK pathway activation, as demonstrated via p38/JNK phosphorylation assays.[257, 258] It is also quite likely that aberrant phosphorylation patterns within MAPK signalling (predominantly at 8 hpi) are connected to the host's activation since no UL97-dependent MAPK phosphorylation events were found in the dataset from Oberstein et al..[190] Another highly interesting target is Mapk7 (*syn.* ERK5, T723 2-fold up), which represents an alternative to the classical MAPK pathway. Signalling via the Mapk7 branch is thus far known to be involved in proliferation, differentiation and cell survival (for instance in response to hyperosmolarity, H_2O_2 and EGF),[259] hence, involvement in infection appears likely.

TNFα signalling: Further data examination revealed a comprehensive downreguslation of the entire superordinated TNFα signalling pathway including the downstream MAPK-, PI3K-Akt-, NF-κB- and necroptosis pathways (Figure 48). As central players, the RAC-alpha serine/threonine-protein kinases 1 and 2 (Akt1, Akt2) were downregulated 1.5- and 2-fold and the Pi3k isoforms r2 (Pi3kr2) was downregulated 2-fold, (isoform Pi3kr1 only 1.3-fold). Although Akt/Pi3k has been described as activated upon virus infection,[260, 261] no differential phosphorylation was detected in this study. However, downregulation of Akt1, Akt2, Pi3kr1 and Pi3kr2 upon HCMV infection was found in the dataset of Weekes *et al.* as well. Since Pi3K inhibition has been recently shown to favour CMV reactivation from latency,[262] a systemic downregulation might be favourable for CMV replication in general. In addition, both datasets showed downregulation of the "rapid-acting" transcription factor NF-κB (~2.5-fold) and the associated NF-κB essential modulator (Nemo, ~2-fold) as already described for MCMV[207] and HCMV,[263] complementing the picture of extensively attenuated TNFα signalling. Notably, the NF-κB repressing factor (NKRF) was only downregulated in MCMV (1.5-fold) but 2-fold upregulated in HCMV, rendering it a target for studying species-specificity.

Programmed cell death: A closer look on apoptosis signalling revealed diverse regulation within pathways downstream of MAPK and TNFα signalling. Diablo, Cathepsin D and Cathepsin L1 exhibited elevated expression levels in both MCMV and HCMV. In contrast, all caspases, known to be major signal transducers in apoptosis, were slightly or moderately downregulated in HCMV (up to 4-fold), and slightly downregulated in MCMV (Figure 57). As a fundamental difference, the effector Caspase Casp7 was only upregulated in MCMV rendering its activity state as highly interesting for future studies (cleaved or pro-caspase). However, it might be possible that suppression of upstream Casp8 leads to an accumulation of the inactive pro-caspase 8 resulting in elevated Casp7 levels.[158] Another explanation might be that Casp8, Cathepsin D, Cathepsin L1 and diablo are actively upregulated by the host in order to maintain apoptosis signaling[264, 265] to compensate suppression by MCMV. However, the major key-

players of apoptosis signaling Bad, Bax, Bak and Bid were not found to be altered in MCMV, but Bad was detected with elevated phosphorylation levels (S155, 3-fold), which might play an additional role in the complex modulation of apoptosis signaling.

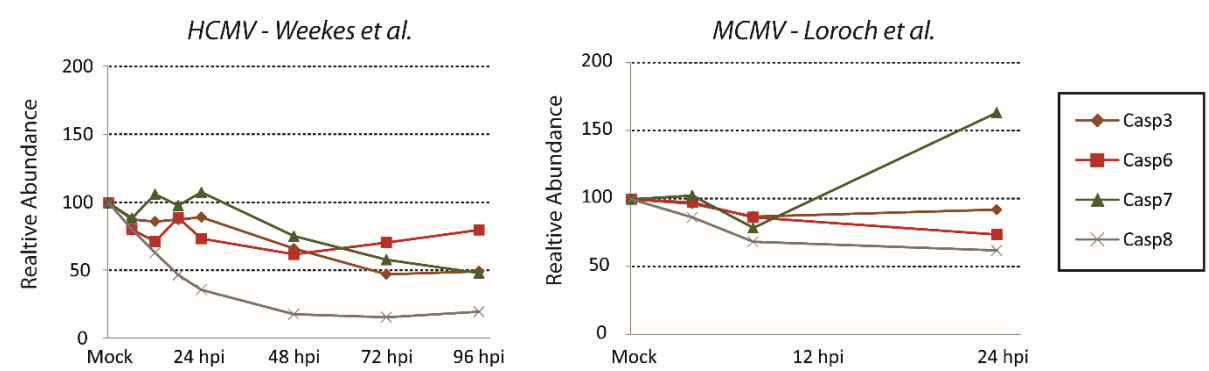


Figure 57: Caspases were systematically downregulated in both systems. As only exception, Casp7 was clearly upregulated in MCMV at 24 hpi.

7.2.2.2 Manipulation of cell cycle signalling and shutdown of DNA-replication

Cell cycle regulation: Among proteins involved in cell cycle signalling, considerable differences were observed between HCMV and MCMV (Figure 58). CycA2, Cdk2 and Kip1 were among the most upregulated proteins in MCMV, whereas only CDK7 was found upregulated in HCMV. Other key-players were fairly down (Cdk1, Cdk2, Cdk4, Cdk6, CycA2, CycB1, CycD1) or not regulated (Kip1). The observed differences are not surprising since HCMV and MCMV are well-described to display cell cycle-specific regulation of gene transcription during lytic and latent infection: for HCMV, it has been shown that CycA2 has a strong negative influence on lytic HCMV infection, whereas MCMV even favours CycA2 activity,[140, 219] hence, the observed upregulation is not surprising. However, the concomitant upregulation of Kip1 (*syn.* Cdkn1b), known to block cyclin E-CDK2 and cyclin D-CDK4,

might be a novel interesting mechanism to understand species specific downregulation of cyclines and CDKs upon MCMV infection.

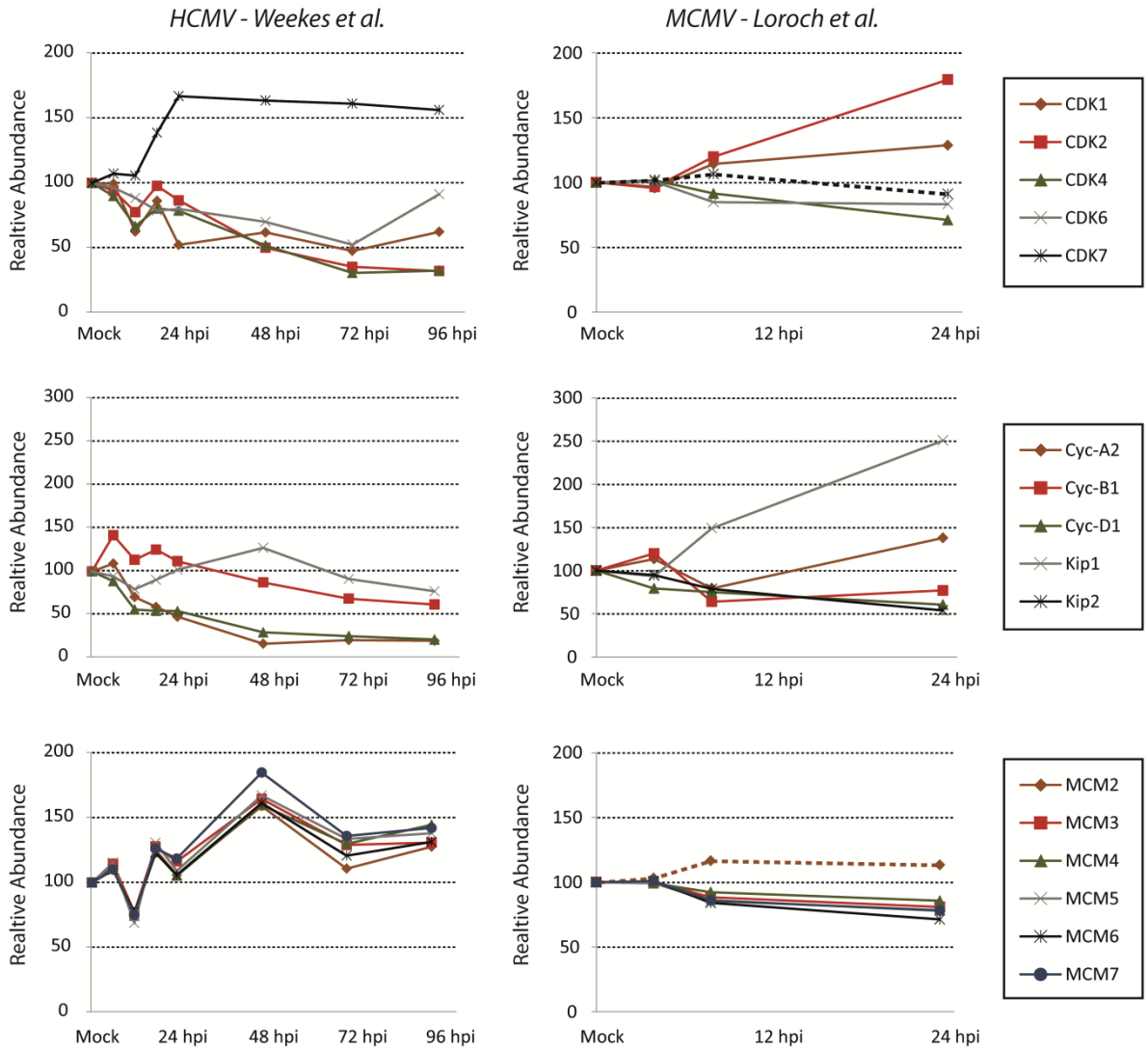


Figure 58: Comparison of cell cycle protein levels upon infection with HCMV (Weekes *et al.*) or MCMV reveals fundamental differences for CDKs and cyclines, perhaps connected to differences in the abundance of minichromosome maintenance proteins. Dashed lines represent proteins with ambiguous quantitative data.

DNA replication: Most notably, all six minichromosome maintenance proteins (MCM2-7) were upregulated in HCMV, which is in accordance with previous studies in human fibroblasts (Figure 58).[190, 266, 267] In contrast, no regulation was observed upon MCMV infection, which might be attributed to the differences in cycline and CDK manipulation. These results render MCM proteins as interesting targets

for future studies, as these complexes might not only be involved in DNA-replication shut down, but also in viral DNA/RNA replication as shown for influenza.[268] Notably, a complex pattern of differential phosphorylation has been detected on MCM2 (S26/S27) and MCM3 (S672/S681, T729/S732) maybe involved in the organization of the MCM protein complex. However, it remains unclear whether these phosphorylation events are connected to altered CDK activities[269] or rather viral protein kinase activities, as shown for HCMV (by Oberstein et al.)[190] and also for Epstein-Barr virus replication (Kudoh et al.).[270]

The extensive manipulation of cyclines and CDKs are in good accordance with the observed downregulation of a wide variety of phosphorylations on proteins annotated with the functional terms "DNA replication" and "DNA replication initiation" (especially detected 8 hpi). For instance, three topoisomerase isoforms Top1 (S10, 2.5-fold down), Top2a (S1211 and T1245, 3-fold and 2-fold down) and Top2b (S1568, 1.5-fold down) exhibited pronounced dephosphorylation, which might reflect a shutdown of the cellular replication machinery. However, as early studies from the 90's have described human cytomegalovirus to induce topoisomerase expression (in terms of viral DNA maintenance, shown for latent HCMV infection),[271, 272] dephosphorylation might be connected to a change of substrate specificity towards viral DNA replication. According to network analyses (data not shown), Top1a and b are connected to a cluster of proteins involved in DNA-replication and DNA-repair, including Mdc1 (S929 and T1480, 2-fold down) and ligase I (Lig1, S51, 1.5-fold down). Notably, as no corresponding dephosphorylation events where found for HCMV by Oberstein et al.,[190] global dephosphorylation might not be mediated by M97 kinase activity. However, it might be attributed to MCMV-mediated inhibition of host DNA-synthesis[273] and recruitment of Polymerase 2A, as shown by interaction studies in yeast.[274]

7.2.2.3 Reorganization of the host cell into a viral protein synthesis machinery

Splicing: The host cell displayed comprehensive upregulation on all levels of protein expression, namely transcription, translation, protein processing and transport. A group comprising Y-Box protein 1 (Ybx1), the small nuclear ribonucleoprotein 70 (Snrnp70), the ribonucleic acid binding protein S1 (Rnps1) and the RNA binding-motif protein (Rbmx) known to be involved in splicing, formed a cluster with the general transcription factor IIF (Gtf2f), which is known to recruit RNA polymerase II into the initiation complex to promote transcription (Figure 46 A). Since it is known that certain MCMV mRNAs undergo alternative splicing, especially E transcripts,[147] splicing machinery manipulation at a comparably early time point might be essential for MCMV to generate mature transcripts. It might as well be connected to activation of host-cell resources, such as protein processing or trafficking, which are required for viral replication. Notably, the proteins Snrnp70, Rnps1 and Rbmx were as well clearly upregulated in the HCMV dataset of Weekes et al. pointing towards a well-conserved manipulation.

Numerous proteins annotated with “RNA-splicing” and “RNA-processing” displayed aberrant phosphorylation events. The DNA-directed RNA polymerase II subunit RPB1 (Polr2a) was differentially phosphorylated either on S1913 and S1920 or 1927 and 1934 (1.5-fold up, indistinguishable due to sequence repetition). Polr2a is known to be recruited by HCMV for viral gene transcription[275] and was demonstrated as crucial for an efficient virus replication in host cells.[153, 276] In addition, a variety of Polr2a associated proteins were found differentially phosphorylated such as Splicing factor 3B subunit 1 (Sf3b1, S194 1.5-fold up), Serine/arginine repetitive matrix protein 1 (Srrm1, T600/S602, S624/S626 via doubly phosphorylated peptides, 1.5-fold up; S1151, S1216, S2224, 1.5-fold down) and Nuclear cap-binding protein subunit 1 (Ncbp1, 2-fold down). Notably, no corresponding phosphorylation was found in the Oberstein et al. dataset, although a large share of splicing-related proteins were described as differentially phosphorylated in an UL97-dependent way.[190]

This discrepancy might be either attributed to species-specific differences or to the probabilistic nature of MS-based peptide detection. However, even though the underlying mechanisms are unknown, differential phosphorylation of Polr2a and various other splicing-related proteins is certainly connected to a switch of the host cell into a viral replication machinery.

Ribosomes: As anticipated, numerous ribosomal proteins were found systemically upregulated, such as S27a (Rps27a), L39 (Rpl39), S24 (Rps24), L23a (Rpl23a), S14 (Rps14), S15 (Rps15), S18 (Rps18), L27 (Rpl27), L35 (Rpl35), L32 (Rpl32) indicating enhanced translation (Figure 51). In addition, the eukaryotic translation initiation factor - subunits 1 and 3, namely Eif1ax and Eif3j2, have been found slightly upregulated (1.5-fold) and reflect further optimization towards translation of viral mRNA. Notably, an essential role of Eif3 for viral protein expression was shown for hepatitis C virus[277, 278] and has further been described for herpes simplex virus.[279] In addition, the eukaryotic translation initiation factors 4 gamma 2 (Eif4g2, T507, 1.5-fold up) and 3 subunit B (Eif3b, S68, 1.5-fold down and S79, 1.5-fold up) exhibited slightly elevated phosphorylation levels. Both factors are involved in the initiation of viral replication (section 6.6.4) and reported to be proteolytically cleaved upon viral infection.[280, 281] Furthermore, the Elongation factor 1-delta (Eef1d) was detected with elevated S118 and S119 phosphorylation levels (1.5-fold up) indicating orchestration of the translation machinery towards optimal viral protein synthesis.[282] Notably, the dataset of Oberstein et al. did not display comparable UL97-mediated regulations, even though non-regulated phosphorylation sites were detected on Eif2A, 2B, 2C and 2G. [190] Therefore, involvement of the MCMV homologue M97 is rather unlikely.

Protein processing in ER and UPR: Since global upregulation of protein synthesis is connected to substantial ER stress, concomitant upregulation of the protein processing machinery can be anticipated as well as concomitant suppression of the UPR.[179, 252, 253] Accordingly, a substantial fraction of proteins involved in trafficking and glycosylation was upregulated such as Sec61b, Ssr3, (*syn* Trap) and the two isoforms of the OST complex Rpn1 and Rpn2. A concomitant upregulation was observed for the

chaperones Calr (*syn* Crt) and hypoxia up-regulated protein 1 (Hyou1 *syn* Nef) and the co-chaperon Dnajb11 (*syn* Hsp40). Systematic comparison with the dataset from Weekes et al. [94] revealed similar profiles in HCMV indicating that manipulation of the protein processing machinery is well-conserved among CMVs. Noteworthy, although numerous proteins displayed less pronounced upregulation by ~1.5-fold, partially below the cut-offs, the systemic upregulation of almost every involved protein emphasizes a critical effect on the level of cellular function.

Additionally, numerous chaperones connected to "protein processing in ER" displayed aberrant phosphorylation, such as Hsp90b1 (*syn* Grp94 S306, 3-fold down) and Hspa5 (*syn*. Grp78 S650, 2-fold down). Hsp90b1 is involved in processing and transportation of secreted proteins and might be phosphorylated by the extracellular serine/threonine protein kinase Fam20c.[283] It has been suspected to assist folding for hepatitis B virus[284] and to form transient complexes with UL55 (glycoprotein B) of HCMV.[285] Similarly, Hspa5 is associated with protein complex assembly in the ER, assists in correct folding and degradation of misfolded proteins and has been connected to HCMV virion assembly.[286] Hence, these aberrant phosphorylation events might be a key to understand viral protein synthesis and virion assembly during the lytic cycle of CMV.

A 2-fold downregulation was observed for two key-players connected to the UPR, namely ATF6b and Eif2ak4 (*syn.* PERK, but not isoform Eif2ak2). Since CMV infection is known to activate chaperones via the UPR,[287] this downregulation most likely links to the block of UPR-mediated synthesis shutdown and further might link to suppression of UPR-mediated apoptosis, the cell's radical consequence to extreme protein accumulation.[288]

Vesicular transport: Ultimately, as final step of protein synthesis and concomitant virion assembly, the vesicular transport machinery exhibited a slight, but systemic upregulation of nearly all key-players (Figure 53). The most pronounced regulation was detected for Snap23 (~2-fold), which has been

connected with membrane budding necessary for CMV virion production[289] and Vamp3 (~1.5-fold up), which is involved in fusion with the endocytic recycling compartment and the Golgi.[290] Though Snap23 and Vamp3 have been shown to be supressed by HCMV's miRNAs,[291] this upregulation might be connected to the increasing demands for protein trafficking and membrane budding upon virion assembly. Notably, in the HCMV data of Weekes *et al.*, Snap23 appeared rather stable, but Vamp3 and Snap29, both essential mediator for membrane fusion, were considerably upregulated. Similarly, the proteins Stx6, Stx7, Stx18 and the Golgi SNAP receptor complex 2 (Gosr2) were upregulated in both datasets indicating a conserved mechanism for enhancing the host cell's transport capabilities between ER, Golgi and endosomes.

In addition, considerable regulations were detected regarding the phosphorylation patterns of Snap23 (S110, 2-fold down), Stx4 (S117, 2-fold up), Vamp4 (S90, 1.5-fold) and Stx7 (S126/S129 and S129, ~2-fold down) indicating complementary fine-tuning of the vesicular transport machinery. Interestingly, Oberstein et al. as well detected phosphorylation on proteins involved in vesicular transport but not as regulated, hence, involvement of UL97 is rather unlikely. However, modifications in this pathway are perhaps connected to the final steps for virion assembly and release – as they represent the logical consequence of enhanced protein synthesis and virion assembly.

8. Conclusion & Outlook

In the present work, ERLIC has proven to be highly beneficial for phosphoproteomics with low sample amounts, which is a critical requirement for application in clinical and biomedical research. It proofed especially beneficial for phosphopeptides which are short or of low AACV, as those are unlikely to be detected using TiO_2. A next step would be a detailed comparison of ERLIC with IMAC, another widely used enrichment technique, which became (again) increasingly popular.[114, 115, 292] As frequently demonstrated[104, 126, 231] and also extensively reviewed (by Loroch et al.),[14, 25] this study demonstrates

again that no universal method for phosphoproteomics has been described so far, as all techniques exhibit substantial complementarity. Hence, only careful evaluation can help answering the question which workflow to apply for a successful analysis of a given sample. In case of targeted phosphoproteomics, aiming at detection of a few selected phosphopeptides, large libraries might be employed to estimate the probabilities for detection and recovery rates with each method. Machine learning algorithms might be used to compute chances for good recovery, as they have proven quite beneficial for HPLC retention time prediction.[293-295] However, as such algorithms require large training datasets, such approaches critically depend on systematic characterization of techniques in terms of quantitative recovery. In this respect, it is rather disconcerting that studies for addressing quantitative recovery are comparably seldom in the proteomics field, especially when considering that sample processing is a main source of technical variation and a major problem for applying proteomics in routine diagnostics.

The discovery experiment on MCMV-infection demonstrated that ERLIC enables in-depth characterization of biomedical samples. It allowed monitoring viral protein expression and phosphorylation as well as manipulation of the host-cell. The substantial difference between protein clusters based on temporal protein expression profiles (this study) and the traditional clustering in the field of virology highlights the benefit of LC-MS in the fields of virology. The precise quantification of LC-MS appears beneficial for monitoring the majority of viral proteins (over multiple conditions) as biochemical detection requires highly multiplexed immunoassays and depends on the availability of antibodies. As a drawback, bottom-up proteomics exhibits limitations in the detection of splicing events if digestion does not yield the specific surrogate peptides for each isoform. Hence, integration with quantitative transcriptomics appears quite promising and has the potential to address translation control, another instance of regulation, which cannot be investigated solely by transcriptomics or proteomics. Notably, the combination of proteomics and phosphoproteomics in this study unveiled a series of

differentially phosphorylated targets (viral and host proteins), which might be in future investigated for their specific role. This is of actual clinical relevance when considering the increased drug resistance of HCMV to the four licenced drugs – which all target the viral polymerase.[296, 297] Hence, drugs for tailored inhibition of phosphorylation events to attenuate or even supress viral replication appear promising. As hundreds of clinical kinase drugs are available,[298] the search for a selective viral kinase inhibitor (as already ongoing)[297] or even host kinase inhibitors might be a good starting point towards more successful therapies. In summary, proteomics and phosphoproteomics in the field of virology, especially in conjunction with complementarity high-content screening methods, can help to unveil critical mechanisms of infection and novel targets for related fields such as immunology and pharmacology.

9. References

1. Winkler, H., *Verbreitung und Ursache der Parthenogenesis im Pflanzen- und Tierreiche*. 1920.

2. Wasinger, V. C.; Cordwell, S. J.; Cerpa-Poljak, A.; Yan, J. X.; Gooley, A. A.; Wilkins, M. R.; Duncan, M. W.; Harris, R.; Williams, K. L.; Humphery-Smith, I., Progress with gene-product mapping of the Mollicutes: Mycoplasma genitalium. *Electrophoresis* **1995,** *16* (7), 1090-4.

3. O'Farrell, P. H., High resolution two-dimensional electrophoresis of proteins. *J Biol Chem* **1975,** *250* (10), 4007-21.

4. 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), 231-43.

5. Steen, H.; Mann, M., The ABC's (and XYZ's) of peptide sequencing. *Nature reviews. Molecular cell biology* **2004,** *5* (9), 699-711.

6. Nagaraj, N.; Wisniewski, J. R.; Geiger, T.; Cox, J.; Kircher, M.; Kelso, J.; Paabo, S.; Mann, M., Deep proteome and transcriptome mapping of a human cancer cell line. *Molecular systems biology* **2011,** *7*, 548.

7. Wang, D.; Eraslan, B.; Wieland, T.; Hallstrom, B.; Hopf, T.; Zolg, D. P.; Zecha, J.; Asplund, A.; Li, L. H.; Meng, C.; Frejno, M.; Schmidt, T.; Schnatbaum, K.; Wilhelm, M.; Ponten, F.; Uhlen, M.; Gagneur, J.; Hahne, H.; Kuster, B., A deep proteome and transcriptome abundance atlas of 29 healthy human tissues. *Molecular systems biology* **2019,** *15* (2), e8503.

8. Jungblut, P. R.; Holzhutter, H. G.; Apweiler, R.; Schluter, H., The speciation of the proteome. *Chemistry Central journal* **2008,** *2*, 16.

9. Schluter, H.; Jungblut, P. R., Special issue on protein species. *Amino acids* **2011,** *41* (2), 217-8.

10. Kelleher, N. L., A cell-based approach to the human proteome project. *Journal of the American Society for Mass Spectrometry* **2012,** *23* (10), 1617-24.

11. Sielaff, M.; Kuharev, J.; Bohn, T.; Hahlbrock, J.; Bopp, T.; Tenzer, S.; Distler, U., Evaluation of FASP, SP3, and iST Protocols for Proteomic Sample Preparation in the Low Microgram Range. *Journal of proteome research* **2017,** *16* (11), 4060-4072.

12. Burkhart, 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. *Blood* **2012,** *120* (15), e73-82.

13. Geyer, P. E.; Holdt, L. M.; Teupser, D.; Mann, M., Revisiting biomarker discovery by plasma proteomics. *Molecular systems biology* **2017,** *13* (9), 942.

14. Pagel, O.; Loroch, S.; Sickmann, A.; Zahedi, R. P., Current strategies and findings in clinically relevant post-translational modification-specific proteomics. *Expert review of proteomics* **2015,** *12* (3), 235-53.

15. Meier, F.; Brunner, A. D.; Koch, S.; Koch, H.; Lubeck, M.; Krause, M.; Goedecke, N.; Decker, J.; Kosinski, T.; Park, M. A.; Bache, N.; Hoerning, O.; Cox, J.; Rather, O.; Mann, M., Online Parallel Accumulation-Serial Fragmentation (PASEF) with a Novel Trapped Ion Mobility Mass Spectrometer. *Molecular & cellular proteomics : MCP* **2018,** *17* (12), 2534-2545.

16. Masselon, C.; Anderson, G. A.; Harkewicz, R.; Bruce, J. E.; Pasa-Tolic, L.; Smith, R. D., Accurate mass multiplexed tandem mass spectrometry for high-throughput polypeptide identification from mixtures. *Analytical chemistry* **2000,** *72* (8), 1918-24.

17. Chapman, J. D.; Goodlett, D. R.; Masselon, C. D., Multiplexed and data-independent tandem mass spectrometry for global proteome profiling. *Mass spectrometry reviews* **2014,** *33* (6), 452-70.

18. Meier, F.; Geyer, P. E.; Virreira Winter, S.; Cox, J.; Mann, M., BoxCar acquisition method enables single-shot proteomics at a depth of 10,000 proteins in 100 minutes. *Nature methods* **2018,** *15* (6), 440-448.

19. Ting, Y. S.; Egertson, J. D.; Payne, S. H.; Kim, S.; MacLean, B.; Kall, L.; Aebersold, R.; Smith, R. D.; Noble, W. S.; MacCoss, M. J., Peptide-Centric Proteome Analysis: An Alternative Strategy for the Analysis of Tandem Mass Spectrometry Data. *Molecular & cellular proteomics : MCP* **2015,** *14* (9), 2301-7.

20. Kim, M. S.; Pinto, S. M.; Getnet, D.; Nirujogi, R. S.; Manda, S. S.; Chaerkady, R.; Madugundu, A. K.; Kelkar, D. S.; Isserlin, R.; Jain, S.; Thomas, J. K.; Muthusamy, B.; Leal-Rojas, P.; Kumar, P.; Sahasrabuddhe, N. A.; Balakrishnan, L.; Advani, J.; George, B.; Renuse, S.; Selvan, L. D.; Patil, A. H.; Nanjappa, V.; Radhakrishnan, A.; Prasad, S.; Subbannayya, T.; Raju, R.; Kumar, M.; Sreenivasamurthy, S. K.; Marimuthu, A.; Sathe, G. J.; Chavan, S.; Datta, K. K.; Subbannayya, Y.; Sahu, A.; Yelamanchi, S. D.; Jayaram, S.; Rajagopalan, P.; Sharma, J.; Murthy, K. R.; Syed, N.; Goel, R.; Khan, A. A.; Ahmad, S.; Dey, G.; Mudgal, K.; Chatterjee, A.; Huang, T. C.; Zhong, J.; Wu, X.; Shaw, P. G.; Freed, D.; Zahari, M. S.; Mukherjee, K. K.; Shankar, S.; Mahadevan, A.; Lam, H.; Mitchell, C. J.; Shankar, S. K.; Satishchandra, P.; Schroeder, J. T.; Sirdeshmukh, R.; Maitra, A.; Leach, S. D.; Drake, C. G.; Halushka, M. K.; Prasad, T. S.; Hruban, R. H.; Kerr, C. L.; Bader, G. D.; Iacobuzio-Donahue, C. A.; Gowda, H.; Pandey, A., A draft map of the human proteome. *Nature* **2014,** *509* (7502), 575-81.

21. Wilhelm, M.; Schlegl, J.; Hahne, H.; Gholami, A. M.; Lieberenz, M.; Savitski, M. M.; Ziegler, E.; Butzmann, L.; Gessulat, S.; Marx, H.; Mathieson, T.; Lemeer, S.; Schnatbaum, K.; Reimer, U.; Wenschuh, H.; Mollenhauer, M.; Slotta-Huspenina, J.; Boese, J. H.; Bantscheff, M.; Gerstmair, A.; Faerber, F.; Kuster, B., Mass-spectrometry-based draft of the human proteome. *Nature* **2014,** *509* (7502), 582-7.

22. An integrated encyclopedia of DNA elements in the human genome. *Nature* **2012,** *489* (7414), 57-74.

23. Serang, O.; Kall, L., Solution to Statistical Challenges in Proteomics Is More Statistics, Not Less. *Journal of proteome research* **2015,** *14* (10), 4099-103.

24. Sharma, K.; D'Souza, R. C.; Tyanova, S.; Schaab, C.; Wisniewski, J. R.; Cox, J.; Mann, M., Ultradeep human phosphoproteome reveals a distinct regulatory nature of Tyr and Ser/Thr-based signaling. *Cell reports* **2014,** *8* (5), 1583-94.

25. Loroch, S.; Dickhut, C.; Zahedi, R. P.; Sickmann, A., Phosphoproteomics--more than meets the eye. *Electrophoresis* **2013,** *34* (11), 1483-92.

26. Hunter, T.; Sefton, B. M., Transforming gene product of Rous sarcoma virus phosphorylates tyrosine. *Proceedings of the National Academy of Sciences of the United States of America* **1980,** *77* (3), 1311-5.

27. Batth, T. S.; Papetti, M.; Pfeiffer, A.; Tollenaere, M. A. X.; Francavilla, C.; Olsen, J. V., Large-Scale Phosphoproteomics Reveals Shp-2 Phosphatase-Dependent Regulators of Pdgf Receptor Signaling. *Cell reports* **2018,** *22* (10), 2784-2796.

28. Zhou, H.; Di Palma, S.; Preisinger, C.; Peng, M.; Polat, A. N.; Heck, A. J.; Mohammed, S., Toward a comprehensive characterization of a human cancer cell phosphoproteome. *Journal of proteome research* **2013,** *12* (1), 260-71.

29. Hunt, D. F.; Henderson, R. A.; Shabanowitz, J.; Sakaguchi, K.; Michel, H.; Sevilir, N.; Cox, A. L.; Appella, E.; Engelhard, V. H., Characterization of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry. *Science* **1992,** *255* (5049), 1261-3.

30. Perkins, D. N.; Pappin, D. J.; Creasy, D. M.; Cottrell, J. S., Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **1999,** *20* (18), 3551-67.

31. MacCoss, M. J.; Wu, C. C.; Yates, J. R., 3rd, Probability-based validation of protein identifications using a modified SEQUEST algorithm. *Analytical chemistry* **2002,** *74* (21), 5593-9.

32. Olsen, J. V.; Ong, S. E.; Mann, M., Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Molecular & cellular proteomics : MCP* **2004,** *3* (6), 608-14.

33. Roepstorff, P.; Fohlman, J., Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomedical mass spectrometry* **1984,** *11* (11), 601.

34. Dorfer, V.; Pichler, P.; Stranzl, T.; Stadlmann, J.; Taus, T.; Winkler, S.; Mechtler, K., MS Amanda, a universal identification algorithm optimized for high accuracy tandem mass spectra. *Journal of proteome research* **2014,** *13* (8), 3679-84.

35. Vaudel, M.; Barsnes, H.; Berven, F. S.; Sickmann, A.; Martens, L., SearchGUI: An open-source graphical user interface for simultaneous OMSSA and X!Tandem searches. *Proteomics* **2011,** *11* (5), 996-9.

36. Craig, R.; Beavis, R. C., A method for reducing the time required to match protein sequences with tandem mass spectra. *Rapid communications in mass spectrometry : RCM* **2003,** *17* (20), 2310-6.

37. Keller, A.; Nesvizhskii, A. I.; Kolker, E.; Aebersold, R., Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Analytical chemistry* **2002,** *74* (20), 5383-92.

38. Nesvizhskii, A. I.; Keller, A.; Kolker, E.; Aebersold, R., A statistical model for identifying proteins by tandem mass spectrometry. *Analytical chemistry* **2003,** *75* (17), 4646-58.

39. Kelstrup, C. D.; Jersie-Christensen, R. R.; Batth, T. S.; Arrey, T. N.; Kuehn, A.; Kellmann, M.; Olsen, J. V., Rapid and deep proteomes by faster sequencing on a benchtop quadrupole ultra-high-field Orbitrap mass spectrometer. *Journal of proteome research* **2014,** *13* (12), 6187-95.

40. Kelstrup, C. D.; Bekker-Jensen, D. B.; Arrey, T. N.; Hogrebe, A.; Harder, A.; Olsen, J. V., Performance Evaluation of the Q Exactive HF-X for Shotgun Proteomics. *Journal of proteome research* **2018,** *17* (1), 727-738.

41. Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M., Electrospray ionization for mass spectrometry of large biomolecules. *Science* **1989,** *246* (4926), 64-71.

42. Wilm, M.; Mann, M., Analytical properties of the nanoelectrospray ion source. *Analytical chemistry* **1996,** *68* (1), 1-8.

43. Wilm, M., Principles of electrospray ionization. *Molecular & cellular proteomics : MCP* **2011,** *10* (7), M111 009407.

44. Banks, J. F., High-sensitivity peptide mapping using packed-capillary liquid chromatography and electrospray ionization mass spectrometry. *Journal of Chromatography A* **1996,** *743* (1), 99-104.

45. Mitulović, G.; Smoluch, M.; Chervet, J.-P.; Steinmacher, I.; Kungl, A.; Mechtler, K., An improved method for tracking and reducing the void volume in nano HPLC–MS with micro trapping columns. *Analytical and bioanalytical chemistry* **2003,** *376* (7), 946-951.

46. Shen, Y.; Zhao, R.; Berger, S. J.; Anderson, G. A.; Rodriguez, N.; Smith, R. D., High-efficiency nanoscale liquid chromatography coupled on-line with mass spectrometry using nanoelectrospray ionization for proteomics. *Analytical chemistry* **2002,** *74* (16), 4235-49.

47. Kocher, T.; Pichler, P.; De Pra, M.; Rieux, L.; Swart, R.; Mechtler, K., Development and performance evaluation of an ultralow flow nanoliquid chromatography-tandem mass spectrometry set-up. *Proteomics* **2014,** *14* (17-18), 1999-2007.

48. Brice, R. W.; Zhang, X.; Colon, L. A., Fused-core, sub-2 microm packings, and monolithic HPLC columns: a comparative evaluation. *Journal of separation science* **2009,** *32* (15-16), 2723-31.

49. Hayes, R.; Ahmed, A.; Edge, T.; Zhang, H., Core-shell particles: preparation, fundamentals and applications in high performance liquid chromatography. *Journal of chromatography. A* **2014,** *1357*, 36-52.

50. Oberacher, H.; Premstaller, A.; Huber, C. G., Characterization of some physical and chromatographic properties of monolithic poly(styrene-co-divinylbenzene) columns. *Journal of chromatography. A* **2004,** *1030* (1-2), 201-8.

51. Burkhart, 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. *Journal of proteomics* **2012,** *75* (4), 1454-62.

52. Michalski, A.; Damoc, E.; Hauschild, J. P.; Lange, O.; Wieghaus, A.; Makarov, A.; Nagaraj, N.; Cox, J.; Mann, M.; Horning, S., Mass spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole Orbitrap mass spectrometer. *Molecular & cellular proteomics : MCP* **2011,** *10* (9), M111 011015.

53. Yates, J. R.; Cociorva, D.; Liao, L.; Zabrouskov, V., Performance of a linear ion trap-Orbitrap hybrid for peptide analysis. *Analytical chemistry* **2006,** *78* (2), 493-500.

54. Michalski, A.; Damoc, E.; Lange, O.; Denisov, E.; Nolting, D.; Muller, M.; Viner, R.; Schwartz, J.; Remes, P.; Belford, M.; Dunyach, J. J.; Cox, J.; Horning, S.; Mann, M.; Makarov, A., Ultra high resolution linear ion trap Orbitrap mass spectrometer (Orbitrap Elite) facilitates top down LC MS/MS and versatile peptide fragmentation modes. *Molecular & cellular proteomics : MCP* **2012,** *11* (3), O111 013698.

55. Senko, M. W.; Remes, P. M.; Canterbury, J. D.; Mathur, R.; Song, Q.; Eliuk, S. M.; Mullen, C.; Earley, L.; Hardman, M.; Blethrow, J. D.; Bui, H.; Specht, A.; Lange, O.; Denisov, E.; Makarov, A.; Horning, S.; Zabrouskov, V., Novel parallelized quadrupole/linear ion trap/Orbitrap tribrid mass spectrometer improving proteome coverage and peptide identification rates. *Analytical chemistry* **2013,** *85* (24), 11710-4.

56. Hebert, A. S.; Richards, A. L.; Bailey, D. J.; Ulbrich, A.; Coughlin, E. E.; Westphall, M. S.; Coon, J. J., The one hour yeast proteome. *Molecular & cellular proteomics : MCP* **2014,** *13* (1), 339-47.

57. Espadas, G.; Borras, E.; Chiva, C.; Sabido, E., Evaluation of different peptide fragmentation types and mass analyzers in data-dependent methods using an Orbitrap Fusion Lumos Tribrid mass spectrometer. *Proteomics* **2017,** *17* (9).

58. Stark, A. K.; Meyer, C.; Kraehling, T.; Jestel, G.; Marggraf, U.; Schilling, M.; Janasek, D.; Franzke, J., Electronic coupling and scaling effects during dielectric barrier electrospray ionization. *Analytical and bioanalytical chemistry* **2011,** *400* (2), 561-9.

59. Olsen, J. V.; Macek, B.; Lange, O.; Makarov, A.; Horning, S.; Mann, M., Higher-energy C-trap dissociation for peptide modification analysis. *Nature methods* **2007,** *4* (9), 709-12.

60. Frese, C. K.; Altelaar, A. F.; Hennrich, M. L.; Nolting, D.; Zeller, M.; Griep-Raming, J.; Heck, A. J.; Mohammed, S., Improved peptide identification by targeted fragmentation using CID, HCD and ETD on an LTQ-Orbitrap Velos. *Journal of proteome research* **2011,** *10* (5), 2377-88.

61. Olsen, J. V.; Schwartz, J. C.; Griep-Raming, J.; Nielsen, M. L.; Damoc, E.; Denisov, E.; Lange, O.; Remes, P.; Taylor, D.; Splendore, M.; Wouters, E. R.; Senko, M.; Makarov, A.; Mann, M.; Horning, S., A dual pressure linear ion trap Orbitrap instrument with very high sequencing speed. *Molecular & cellular proteomics : MCP* **2009,** *8* (12), 2759-69.

62. Kocher, T.; Pichler, P.; Swart, R.; Mechtler, K., Analysis of protein mixtures from whole-cell extracts by single-run nanoLC-MS/MS using ultralong gradients. *Nature protocols* **2012,** *7* (5), 882-90.

63. Hsu, J. L.; Huang, S. Y.; Chow, N. H.; Chen, S. H., Stable-isotope dimethyl labeling for quantitative proteomics. *Analytical chemistry* **2003,** *75* (24), 6843-52.

64. Ong, S. E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; Steen, H.; Pandey, A.; Mann, M., Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Molecular & cellular proteomics : MCP* **2002,** *1* (5), 376-86.

65. Thomas, P. J.; Boller, A. J.; Satagopan, S.; Tabita, F. R.; Cavanaugh, C. M.; Scott, K. M., Isotope discrimination by form IC RubisCO from Ralstonia eutropha and Rhodobacter sphaeroides, metabolically versatile members of 'Proteobacteria' from aquatic and soil habitats. *Environmental microbiology* **2019,** *21* (1), 72-80.

66. Kennedy, B. V.; Krouse, H. R., Isotope fractionation by plants and animals: implications for nutrition research. *Canadian journal of physiology and pharmacology* **1990,** *68* (7), 960-72.

67. Thompson, A.; Schafer, J.; Kuhn, K.; Kienle, S.; Schwarz, J.; Schmidt, G.; Neumann, T.; Johnstone, R.; Mohammed, A. K.; Hamon, C., Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Analytical chemistry* **2003,** *75* (8), 1895-904.

68. Ross, P. L.; Huang, Y. N.; Marchese, J. N.; Williamson, B.; Parker, K.; Hattan, S.; Khainovski, N.; Pillai, S.; Dey, S.; Daniels, S.; Purkayastha, S.; Juhasz, P.; Martin, S.; Bartlet-Jones, M.; He, F.; Jacobson, A.; Pappin, D. J., Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Molecular & cellular proteomics : MCP* **2004,** *3* (12), 1154-69.

69. Erickson, B. K.; Jedrychowski, M. P.; McAlister, G. C.; Everley, R. A.; Kunz, R. C.; Gygi, S. P., Evaluating multiplexed quantitative phosphopeptide analysis on a hybrid quadrupole mass filter/linear ion trap/Orbitrap mass spectrometer. *Analytical chemistry* **2014**.

70. McAlister, G. C.; Nusinow, D. P.; Jedrychowski, M. P.; Wuhr, M.; Huttlin, E. L.; Erickson, B. K.; Rad, R.; Haas, W.; Gygi, S. P., MultiNotch MS3 enables accurate, sensitive, and multiplexed detection of differential expression across cancer cell line proteomes. *Analytical chemistry* **2014,** *86* (14), 7150-8.

71. Kocher, T.; Pichler, P.; Schutzbier, M.; Stingl, C.; Kaul, A.; Teucher, N.; Hasenfuss, G.; Penninger, J. M.; Mechtler, K., High precision quantitative proteomics using iTRAQ on an LTQ

Orbitrap: a new mass spectrometric method combining the benefits of all. *Journal of proteome research* **2009,** *8* (10), 4743-52.

72. Cao, Z.; Evans, A. R.; Robinson, R. A., MS(3)-based quantitative proteomics using pulsed-Q dissociation. *Rapid communications in mass spectrometry : RCM* **2015,** *29* (11), 1025-30.

73. Yost, R. A.; Enke, C. G., Triple quadrupole mass spectrometry for direct mixture analysis and structure elucidation. *Analytical chemistry* **1979,** *51* (12), 1251-64.

74. Gallien, S.; Peterman, S.; Kiyonami, R.; Souady, J.; Duriez, E.; Schoen, A.; Domon, B., Highly multiplexed targeted proteomics using precise control of peptide retention time. *Proteomics* **2012,** *12* (8), 1122-33.

75. Aebersold, R., Constellations in a cellular universe. *Nature* **2003,** *422* (6928), 115-6.

76. Gerber, S. A.; Rush, J.; Stemman, O.; Kirschner, M. W.; Gygi, S. P., Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proceedings of the National Academy of Sciences of the United States of America* **2003,** *100* (12), 6940-5.

77. Li, H.; Han, J.; Pan, J.; Liu, T.; Parker, C. E.; Borchers, C. H., Current trends in quantitative proteomics - an update. *Journal of mass spectrometry : JMS* **2017,** *52* (5), 319-341.

78. Link, A. J.; Eng, J.; Schieltz, D. M.; Carmack, E.; Mize, G. J.; Morris, D. R.; Garvik, B. M.; Yates, J. R., 3rd, Direct analysis of protein complexes using mass spectrometry. *Nature biotechnology* **1999,** *17* (7), 676-82.

79. Kwiatkowski, M.; Krosser, D.; Wurlitzer, M.; Steffen, P.; Barcaru, A.; Krisp, C.; Horvatovich, P.; Bischoff, R.; Schluter, H., Application of Displacement Chromatography to Online Two-Dimensional Liquid Chromatography Coupled to Tandem Mass Spectrometry Improves Peptide Separation Efficiency and Detectability for the Analysis of Complex Proteomes. *Analytical chemistry* **2018,** *90* (16), 9951-9958.

80. Opiteck, G. J.; Lewis, K. C.; Jorgenson, J. W.; Anderegg, R. J., Comprehensive on-line LC/LC/MS of proteins. *Analytical chemistry* **1997,** *69* (8), 1518-24.

81. Trusch, M.; Tillack, K.; Kwiatkowski, M.; Bertsch, A.; Ahrends, R.; Kohlbacher, O.; Martin, R.; Sospedra, M.; Schluter, H., Displacement chromatography as first separating step in online two-dimensional liquid chromatography coupled to mass spectrometry analysis of a complex protein sample--the proteome of neutrophils. *Journal of chromatography. A* **2012,** *1232*, 288-94.

82. Gilar, M.; Olivova, P.; Daly, A. E.; Gebler, J. C., Orthogonality of separation in two-dimensional liquid chromatography. *Analytical chemistry* **2005,** *77* (19), 6426-34.

83. Boersema, P. J.; Mohammed, S.; Heck, A. J., Hydrophilic interaction liquid chromatography (HILIC) in proteomics. *Analytical and bioanalytical chemistry* **2008,** *391* (1), 151-9.

84. Loroch, S.; Schommartz, T.; Brune, W.; Zahedi, R. P.; Sickmann, A., Multidimensional electrostatic repulsion-hydrophilic interaction chromatography (ERLIC) for quantitative analysis of the proteome and phosphoproteome in clinical and biomedical research. *Biochimica et biophysica acta* **2015,** *1854* (5), 460-468.

85. Giddings, J. C., Concepts and comparisons in multidimensional separation. *Journal of High Resolution Chromatography* **1987,** *10* (5), 319-323.

86. Ficarro, S. B.; Zhang, Y.; Carrasco-Alfonso, M. J.; Garg, B.; Adelmant, G.; Webber, J. T.; Luckey, C. J.; Marto, J. A., Online nanoflow multidimensional fractionation for high efficiency phosphopeptide analysis. *Molecular & cellular proteomics : MCP* **2011,** *10* (11), O111 011064.

87. Betancourt, L. H.; De Bock, P. J.; Staes, A.; Timmerman, E.; Perez-Riverol, Y.; Sanchez, A.; Besada, V.; Gonzalez, L. J.; Vandekerckhove, J.; Gevaert, K., SCX charge state selective separation of tryptic peptides combined with 2D-RP-HPLC allows for detailed proteome mapping. *Journal of proteomics* **2013,** *91*, 164-71.

88. Peng, J.; Elias, J. E.; Thoreen, C. C.; Licklider, L. J.; Gygi, S. P., Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. *Journal of proteome research* **2003,** *2* (1), 43-50.

89. Alpert, A. J.; Hudecz, O.; Mechtler, K., Anion-exchange chromatography of phosphopeptides: weak anion exchange versus strong anion exchange and anion-exchange chromatography versus electrostatic repulsion-hydrophilic interaction chromatography. *Analytical chemistry* **2015,** *87* (9), 4704-11.

90. Hao, P.; Qian, J.; Ren, Y.; Sze, S. K., Electrostatic repulsion-hydrophilic interaction chromatography (ERLIC) versus strong cation exchange (SCX) for fractionation of iTRAQ-labeled peptides. *Journal of proteome research* **2011,** *10* (12), 5568-74.

91. Gilar, M.; Olivova, P.; Daly, A. E.; Gebler, J. C., Two-dimensional separation of peptides using RP-RP-HPLC system with different pH in first and second separation dimensions. *Journal of separation science* **2005,** *28* (14), 1694-703.

92. Wang, Y.; Yang, F.; Gritsenko, M. A.; Wang, Y.; Clauss, T.; Liu, T.; Shen, Y.; Monroe, M. E.; Lopez-Ferrer, D.; Reno, T.; Moore, R. J.; Klemke, R. L.; Camp, D. G., 2nd; Smith, R. D., Reversed-phase chromatography with multiple fraction concatenation strategy for proteome profiling of human MCF10A cells. *Proteomics* **2011,** *11* (10), 2019-26.

93. Bekker-Jensen, D. B.; Kelstrup, C. D.; Batth, T. S.; Larsen, S. C.; Haldrup, C.; Bramsen, J. B.; Sorensen, K. D.; Hoyer, S.; Orntoft, T. F.; Andersen, C. L.; Nielsen, M. L.; Olsen, J. V., An Optimized Shotgun Strategy for the Rapid Generation of Comprehensive Human Proteomes. *Cell systems* **2017,** *4* (6), 587-599 e4.

94. Weekes, M. P.; Tomasec, P.; Huttlin, E. L.; Fielding, C. A.; Nusinow, D.; Stanton, R. J.; Wang, E. C.; Aicheler, R.; Murrell, I.; Wilkinson, G. W.; Lehner, P. J.; Gygi, S. P., Quantitative temporal viromics: an approach to investigate host-pathogen interaction. *Cell* **2014,** *157* (6), 1460-72.

95. Batth, T. S.; Olsen, J. V., Offline High pH Reversed-Phase Peptide Fractionation for Deep Phosphoproteome Coverage. *Methods in molecular biology (Clifton, N.J.)* **2016,** *1355*, 179-92.

96. Batth, T. S.; Francavilla, C.; Olsen, J. V., Off-line high-pH reversed-phase fractionation for in-depth phosphoproteomics. *Journal of proteome research* **2014,** *13* (12), 6176-86.

97. Trinidad, J. C.; Specht, C. G.; Thalhammer, A.; Schoepfer, R.; Burlingame, A. L., Comprehensive identification of phosphorylation sites in postsynaptic density preparations. *Molecular & cellular proteomics : MCP* **2006,** *5* (5), 914-22.

98. Villen, J.; Gygi, S. P., The SCX/IMAC enrichment approach for global phosphorylation analysis by mass spectrometry. *Nature protocols* **2008,** *3* (10), 1630-8.

99. Mohammed, S.; Heck, A., Jr., Strong cation exchange (SCX) based analytical methods for the targeted analysis of protein post-translational modifications. *Current opinion in biotechnology* **2011,** *22* (1), 9-16.

100. Hennrich, M. L.; Groenewold, V.; Kops, G. J.; Heck, A. J.; Mohammed, S., Improving depth in phosphoproteomics by using a strong cation exchange-weak anion exchange-reversed phase multidimensional separation approach. *Analytical chemistry* **2011,** *83* (18), 7137-43.

101. McNulty, D. E.; Annan, R. S., Hydrophilic interaction chromatography reduces the complexity of the phosphoproteome and improves global phosphopeptide isolation and detection. *Molecular & cellular proteomics : MCP* **2008,** *7* (5), 971-80.

102. Mysling, S.; Palmisano, G.; Hojrup, P.; Thaysen-Andersen, M., Utilizing ion-pairing hydrophilic interaction chromatography solid phase extraction for efficient glycopeptide enrichment in glycoproteomics. *Analytical chemistry* **2010,** *82* (13), 5598-609.

103. Engholm-Keller, K.; Hansen, T. A.; Palmisano, G.; Larsen, M. R., Multidimensional strategy for sensitive phosphoproteomics incorporating protein prefractionation combined with SIMAC, HILIC, and TiO(2) chromatography applied to proximal EGF signaling. *Journal of proteome research* **2011,** *10* (12), 5383-97.

104. Engholm-Keller, K.; Birck, P.; Storling, J.; Pociot, F.; Mandrup-Poulsen, T.; Larsen, M. R., TiSH--a robust and sensitive global phosphoproteomics strategy employing a combination of TiO2, SIMAC, and HILIC. *Journal of proteomics* **2012,** *75* (18), 5749-61.

105. Rush, J.; Moritz, A.; Lee, K. A.; Guo, A.; Goss, V. L.; Spek, E. J.; Zhang, H.; Zha, X. M.; Polakiewicz, R. D.; Comb, M. J., Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. *Nature biotechnology* **2005,** *23* (1), 94-101.

106. Brill, L. M.; Salomon, A. R.; Ficarro, S. B.; Mukherji, M.; Stettler-Gill, M.; Peters, E. C., Robust phosphoproteomic profiling of tyrosine phosphorylation sites from human T cells using immobilized metal affinity chromatography and tandem mass spectrometry. *Analytical chemistry* **2004,** *76* (10), 2763-72.

107. Posewitz, M. C.; Tempst, P., Immobilized gallium(III) affinity chromatography of phosphopeptides. *Analytical chemistry* **1999,** *71* (14), 2883-92.

108. Pinkse, M. W.; Uitto, P. M.; Hilhorst, M. J.; Ooms, B.; Heck, A. J., Selective isolation at the femtomole level of phosphopeptides from proteolytic digests using 2D-NanoLC-ESI-MS/MS and titanium oxide precolumns. *Analytical chemistry* **2004,** *76* (14), 3935-43.

109. Larsen, M. R.; Thingholm, T. E.; Jensen, O. N.; Roepstorff, P.; Jorgensen, T. J., Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Molecular & cellular proteomics : MCP* **2005,** *4* (7), 873-86.

110. Eyrich, B.; Sickmann, A.; Zahedi, R. P., Catch me if you can: mass spectrometry-based phosphoproteomics and quantification strategies. *Proteomics* **2011,** *11* (4), 554-70.

111. Grimsrud, P. A.; Swaney, D. L.; Wenger, C. D.; Beauchene, N. A.; Coon, J. J., Phosphoproteomics for the masses. *ACS chemical biology* **2010,** *5* (1), 105-19.

112. Ficarro, S. B.; McCleland, M. L.; Stukenberg, P. T.; Burke, D. J.; Ross, M. M.; Shabanowitz, J.; Hunt, D. F.; White, F. M., Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae. *Nature biotechnology* **2002,** *20* (3), 301-5.

113. Tsai, C. F.; Hsu, C. C.; Hung, J. N.; Wang, Y. T.; Choong, W. K.; Zeng, M. Y.; Lin, P. Y.; Hong, R. W.; Sung, T. Y.; Chen, Y. J., Sequential Phosphoproteomic Enrichment through Complementary Metal-Directed Immobilized Metal Ion Affinity Chromatography. *Analytical chemistry* **2014,** *86* (1), 685-93.

114. de Graaf, E. L.; Giansanti, P.; Altelaar, A. F.; Heck, A. J., Single-step enrichment by Ti4+-IMAC and label-free quantitation enables in-depth monitoring of phosphorylation dynamics with high reproducibility and temporal resolution. *Molecular & cellular proteomics : MCP* **2014,** *13* (9), 2426-34.

115. Giansanti, P.; Aye, T. T.; van den Toorn, H.; Peng, M.; van Breukelen, B.; Heck, A. J., An Augmented Multiple-Protease-Based Human Phosphopeptide Atlas. *Cell reports* **2015,** *11* (11), 1834-43.

116. Pinkse, M. W.; Mohammed, S.; Gouw, J. W.; van Breukelen, B.; Vos, H. R.; Heck, A. J., Highly robust, automated, and sensitive online TiO2-based phosphoproteomics applied to study endogenous phosphorylation in Drosophila melanogaster. *Journal of proteome research* **2008,** *7* (2), 687-97.

117. Engholm-Keller, K.; Larsen, M. R., Titanium dioxide as chemo-affinity chromatographic sorbent of biomolecular compounds--applications in acidic modification-specific proteomics. *Journal of proteomics* **2011,** *75* (2), 317-28.

118. Palmisano, G.; Parker, B. L.; Engholm-Keller, K.; Lendal, S. E.; Kulej, K.; Schulz, M.; Schwammle, V.; Graham, M. E.; Saxtorph, H.; Cordwell, S. J.; Larsen, M. R., A novel method for the simultaneous enrichment, identification, and quantification of phosphopeptides and sialylated glycopeptides applied to a temporal profile of mouse brain development. *Molecular & cellular proteomics : MCP* **2012,** *11* (11), 1191-202.

119. Wu, J.; Shakey, Q.; Liu, W.; Schuller, A.; Follettie, M. T., Global profiling of phosphopeptides by titania affinity enrichment. *Journal of proteome research* **2007,** *6* (12), 4684-9.

120. Beck, F.; Geiger, J.; Gambaryan, S.; Veit, J.; Vaudel, M.; Nollau, P.; Kohlbacher, O.; Martens, L.; Walter, U.; Sickmann, A.; Zahedi, R. P., Time-resolved characterization of cAMP/PKA-dependent signaling reveals that platelet inhibition is a concerted process involving multiple signaling pathways. *Blood* **2014,** *123* (5), e1-e10.

121. Alpert, A. J., Electrostatic repulsion hydrophilic interaction chromatography for isocratic separation of charged solutes and selective isolation of phosphopeptides. *Analytical chemistry* **2008,** *80* (1), 62-76.

122. Hao, P.; Guo, T.; Sze, S. K., Simultaneous analysis of proteome, phospho- and glycoproteome of rat kidney tissue with electrostatic repulsion hydrophilic interaction chromatography. *PloS one* **2011,** *6* (2), e16884.

123. Chien, K. Y.; Liu, H. C.; Goshe, M. B., Development and application of a phosphoproteomic method using electrostatic repulsion-hydrophilic interaction chromatography (ERLIC), IMAC, and LC-MS/MS analysis to study Marek's Disease Virus infection. *Journal of proteome research* **2011,** *10* (9), 4041-53.

124. Loroch, S.; Zahedi, R. P.; Sickmann, A., Highly sensitive phosphoproteomics by tailoring solid-phase extraction to electrostatic repulsion-hydrophilic interaction chromatography. *Analytical chemistry* **2015,** *87* (3), 1596-604.

125. Zarei, M.; Sprenger, A.; Gretzmeier, C.; Dengjel, J., Combinatorial use of electrostatic repulsion-hydrophilic interaction chromatography (ERLIC) and strong cation exchange (SCX) chromatography for in-depth phosphoproteome analysis. *Journal of proteome research* **2012,** *11* (8), 4269-76.

126. Gan, C. S.; Guo, T.; Zhang, H.; Lim, S. K.; Sze, S. K., A comparative study of electrostatic repulsion-hydrophilic interaction chromatography (ERLIC) versus SCX-IMAC-based methods for phosphopeptide isolation/enrichment. *Journal of proteome research* **2008,** *7* (11), 4869-77.

127. Zhang, H.; Guo, T.; Li, X.; Datta, A.; Park, J. E.; Yang, J.; Lim, S. K.; Tam, J. P.; Sze, S. K., Simultaneous characterization of glyco- and phosphoproteomes of mouse brain membrane proteome

with electrostatic repulsion hydrophilic interaction chromatography. *Molecular & cellular proteomics : MCP* **2010,** *9* (4), 635-47.

128. Zarei, M.; Sprenger, A.; Gretzmeier, C.; Dengjel, J., Rapid combinatorial ERLIC-SCX solid-phase extraction for in-depth phosphoproteome analysis. *Journal of proteome research* **2013,** *12* (12), 5989-95.

129. Fields, B. N.; Knipe, D. M.; Howley, P. M., *Fields virology*. Wolters Kluwer Health/Lippincott Williams & Wilkins: Philadelphia, 2013.

130. Vossen, M. T.; Westerhout, E. M.; Soderberg-Naucler, C.; Wiertz, E. J., Viral immune evasion: a masterpiece of evolution. *Immunogenetics* **2002,** *54* (8), 527-42.

131. Perry, R. H.; Cooks, R. G.; Noll, R. J., Orbitrap mass spectrometry: instrumentation, ion motion and applications. *Mass spectrometry reviews* **2008,** *27* (6), 661-99.

132. Goodrum, F., Human Cytomegalovirus Latency: Approaching the Gordian Knot. *Annual review of virology* **2016,** *3* (1), 333-357.

133. Boppana, S. B.; Ross, S. A.; Fowler, K. B., Congenital cytomegalovirus infection: clinical outcome. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **2013,** *57 Suppl 4*, S178-81.

134. Geder, L.; Sanford, E. J.; Rohner, T. J.; Rapp, F., Cytomegalovirus and cancer of the prostate: in vitro transformation of human cells. *Cancer treatment reports* **1977,** *61* (2), 139-46.

135. Harkins, L.; Volk, A. L.; Samanta, M.; Mikolaenko, I.; Britt, W. J.; Bland, K. I.; Cobbs, C. S., Specific localisation of human cytomegalovirus nucleic acids and proteins in human colorectal cancer. *Lancet (London, England)* **2002,** *360* (9345), 1557-63.

136. Epstein, S. E.; Speir, E.; Zhou, Y. F.; Guetta, E.; Leon, M.; Finkel, T., The role of infection in restenosis and atherosclerosis: focus on cytomegalovirus. *Lancet (London, England)* **1996,** *348 Suppl 1*, s13-7.

137. Liashkovich, I.; Hafezi, W.; Kuhn, J. M.; Oberleithner, H.; Shahin, V., Nuclear delivery mechanism of herpes simplex virus type 1 genome. *Journal of molecular recognition : JMR* **2011,** *24* (3), 414-21.

138. Mocarski Jr, E. S., Comparative analysis of herpesvirus-common proteins. In *Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis*, Arvin, A.; Campadelli-Fiume, G.; Mocarski, E.; Moore, P. S.; Roizman, B.; Whitley, R.; Yamanishi, K., Eds. Cambridge University Press: Cambridge, 2007.

139. Mettenleiter, T. C.; Klupp, B. G.; Granzow, H., Herpesvirus assembly: an update. *Virus research* **2009,** *143* (2), 222-34.

140. Ostermann, E.; Loroch, S.; Qian, Z.; Sickmann, A.; Wiebusch, L.; Brune, W., Activation of E2F-dependent transcription by the mouse cytomegalovirus M117 protein affects the viral host range. *PLoS pathogens* **2018,** *14* (12), e1007481.

141. Wiebusch, L.; Neuwirth, A.; Grabenhenrich, L.; Voigt, S.; Hagemeier, C., Cell cycle-independent expression of immediate-early gene 3 results in G1 and G2 arrest in murine cytomegalovirus-infected cells. *Journal of virology* **2008,** *82* (20), 10188-98.

142. Hudson, J. B., The murine cytomegalovirus as a model for the study of viral pathogenesis and persistent infections. *Archives of virology* **1979,** *62* (1), 1-29.

143. Brune, W.; Hengel, H.; Koszinowski, U. H., A mouse model for cytomegalovirus infection. *Current protocols in immunology / edited by John E. Coligan ... [et al.]* **2001,** *Chapter 19*, Unit 19 7.

144. Hummel, M.; Abecassis, M. M., A model for reactivation of CMV from latency. *Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology* **2002,** *25 Suppl 2*, S123-36.

145. Almanan, M.; Raynor, J.; Sholl, A.; Wang, M.; Chougnet, C.; Cardin, R. D.; Hildeman, D. A., Tissue-specific control of latent CMV reactivation by regulatory T cells. *PLoS pathogens* **2017,** *13* (8), e1006507.

146. Rattay, S.; Trilling, M.; Megger, D. A.; Sitek, B.; Meyer, H. E.; Hengel, H.; Le-Trilling, V. T., The Canonical Immediate Early 3 Gene Product pIE611 of Mouse Cytomegalovirus Is Dispensable for Viral Replication but Mediates Transcriptional and Posttranscriptional Regulation of Viral Gene Products. *Journal of virology* **2015,** *89* (16), 8590-8.

147. Schommartz, T.; Loroch, S.; Alawi, M.; Grundhoff, A.; Sickmann, A.; Brune, W., Functional Dissection of an Alternatively Spliced Herpesvirus Gene by Splice Site Mutagenesis. *Journal of virology* **2016,** *90* (9), 4626-36.

148. Ciocco-Schmitt, G. M.; Karabekian, Z.; Godfrey, E. W.; Stenberg, R. M.; Campbell, A. E.; Kerry, J. A., Identification and characterization of novel murine cytomegalovirus M112-113 (e1) gene products. *Virology* **2002,** *294* (1), 199-208.

149. Kerry, J. A.; Priddy, M. A.; Jervey, T. Y.; Kohler, C. P.; Staley, T. L.; Vanson, C. D.; Jones, T. R.; Iskenderian, A. C.; Anders, D. G.; Stenberg, R. M., Multiple regulatory events influence human cytomegalovirus DNA polymerase (UL54) expression during viral infection. *Journal of virology* **1996,** *70* (1), 373-82.

150. Iskenderian, A. C.; Huang, L.; Reilly, A.; Stenberg, R. M.; Anders, D. G., Four of eleven loci required for transient complementation of human cytomegalovirus DNA replication cooperate to activate expression of replication genes. *Journal of virology* **1996,** *70* (1), 383-92.

151. Kattenhorn, L. M.; Mills, R.; Wagner, M.; Lomsadze, A.; Makeev, V.; Borodovsky, M.; Ploegh, H. L.; Kessler, B. M., Identification of proteins associated with murine cytomegalovirus virions. *Journal of virology* **2004,** *78* (20), 11187-97.

152. Chapa, T. J.; Perng, Y. C.; French, A. R.; Yu, D., Murine cytomegalovirus protein pM92 is a conserved regulator of viral late gene expression. *Journal of virology* **2014,** *88* (1), 131-42.

153. Chapa, T. J.; Yu, D.; French, A. R., Proteomic and Phylogenetic Coevolution Analyses of pM79 and pM92 Identify Interactions with RNAPII and Delineate the MCMV Late Transcription Complex. *The Journal of general virology* **2016**.

154. Pan, D.; Han, T.; Tang, S.; Xu, W.; Bao, Q.; Sun, Y.; Xuan, B.; Qian, Z., Murine Cytomegalovirus Protein pM91 Interacts with pM79 and Is Critical for Viral Late Gene Expression. *Journal of virology* **2018,** *92* (18).

155. Heilingloh, C. S.; Grosche, L.; Kummer, M.; Muhl-Zurbes, P.; Kamm, L.; Scherer, M.; Latzko, M.; Stamminger, T.; Steinkasserer, A., The Major Immediate-Early Protein IE2 of Human Cytomegalovirus Is Sufficient to Induce Proteasomal Degradation of CD83 on Mature Dendritic Cells. *Frontiers in microbiology* **2017,** *8*, 119.

156. Isomura, H.; Stinski, M. F., Coordination of late gene transcription of human cytomegalovirus with viral DNA synthesis: recombinant viruses as potential therapeutic vaccine candidates. *Expert opinion on therapeutic targets* **2013,** *17* (2), 157-66.

157. Hanson, L. K.; Dalton, B. L.; Cageao, L. F.; Brock, R. E.; Slater, J. S.; Kerry, J. A.; Campbell, A. E., Characterization and regulation of essential murine cytomegalovirus genes m142 and m143. *Virology* **2005,** *334* (2), 166-77.

158. Handke, W.; Krause, E.; Brune, W., Live or let die: manipulation of cellular suicide programs by murine cytomegalovirus. *Medical microbiology and immunology* **2012,** *201* (4), 475-86.

159. Barnhart, B. C.; Alappat, E. C.; Peter, M. E., The CD95 type I/type II model. *Seminars in immunology* **2003,** *15* (3), 185-93.

160. Scaffidi, C.; Fulda, S.; Srinivasan, A.; Friesen, C.; Li, F.; Tomaselli, K. J.; Debatin, K. M.; Krammer, P. H.; Peter, M. E., Two CD95 (APO-1/Fas) signaling pathways. *The EMBO journal* **1998,** *17* (6), 1675-87.

161. Galluzzi, L.; Vitale, I.; Abrams, J. M.; Alnemri, E. S.; Baehrecke, E. H.; Blagosklonny, M. V.; Dawson, T. M.; Dawson, V. L.; El-Deiry, W. S.; Fulda, S.; Gottlieb, E.; Green, D. R.; Hengartner, M. O.; Kepp, O.; Knight, R. A.; Kumar, S.; Lipton, S. A.; Lu, X.; Madeo, F.; Malorni, W.; Mehlen, P.; Nunez, G.; Peter, M. E.; Piacentini, M.; Rubinsztein, D. C.; Shi, Y.; Simon, H. U.; Vandenabeele, P.; White, E.; Yuan, J.; Zhivotovsky, B.; Melino, G.; Kroemer, G., Molecular definitions of cell death subroutines: recommendations of the Nomenclature Committee on Cell Death 2012. *Cell death and differentiation* **2012,** *19* (1), 107-20.

162. Kroemer, G.; Galluzzi, L.; Vandenabeele, P.; Abrams, J.; Alnemri, E. S.; Baehrecke, E. H.; Blagosklonny, M. V.; El-Deiry, W. S.; Golstein, P.; Green, D. R.; Hengartner, M.; Knight, R. A.; Kumar, S.; Lipton, S. A.; Malorni, W.; Nunez, G.; Peter, M. E.; Tschopp, J.; Yuan, J.; Piacentini, M.; Zhivotovsky, B.; Melino, G., Classification of cell death: recommendations of the Nomenclature Committee on Cell Death 2009. *Cell death and differentiation* **2009,** *16* (1), 3-11.

163. Fliss, P. M.; Brune, W., Prevention of cellular suicide by cytomegaloviruses. *Viruses* **2012,** *4* (10), 1928-49.

164. Terhune, S.; Torigoi, E.; Moorman, N.; Silva, M.; Qian, Z.; Shenk, T.; Yu, D., Human cytomegalovirus UL38 protein blocks apoptosis. *Journal of virology* **2007,** *81* (7), 3109-23.

165. Goldmacher, V. S.; Bartle, L. M.; Skaletskaya, A.; Dionne, C. A.; Kedersha, N. L.; Vater, C. A.; Han, J. W.; Lutz, R. J.; Watanabe, S.; Cahir McFarland, E. D.; Kieff, E. D.; Mocarski, E. S.; Chittenden, T., A cytomegalovirus-encoded mitochondria-localized inhibitor of apoptosis structurally unrelated to Bcl-2. *Proceedings of the National Academy of Sciences of the United States of America* **1999,** *96* (22), 12536-41.

166. Boujrad, H.; Gubkina, O.; Robert, N.; Krantic, S.; Susin, S. A., AIF-mediated programmed necrosis: a highly regulated way to die. *Cell cycle (Georgetown, Tex.)* **2007,** *6* (21), 2612-9.

167. Zhang, J.; Yang, Y.; He, W.; Sun, L., Necrosome core machinery: MLKL. *Cellular and molecular life sciences : CMLS* **2016,** *73* (11-12), 2153-63.

168. Upton, J. W.; Kaiser, W. J.; Mocarski, E. S., Virus inhibition of RIP3-dependent necrosis. *Cell host & microbe* **2010,** *7* (4), 302-13.

169. Rebsamen, M.; Heinz, L. X.; Meylan, E.; Michallet, M. C.; Schroder, K.; Hofmann, K.; Vazquez, J.; Benedict, C. A.; Tschopp, J., DAI/ZBP1 recruits RIP1 and RIP3 through RIP homotypic interaction motifs to activate NF-kappaB. *EMBO reports* **2009,** *10* (8), 916-22.

170. Unterholzner, L., The interferon response to intracellular DNA: why so many receptors? *Immunobiology* **2013,** *218* (11), 1312-21.

171. Mocarski, E. S.; Guo, H.; Kaiser, W. J., Necroptosis: The Trojan horse in cell autonomous antiviral host defense. *Virology* **2015,** *479-480*, 160-6.

172. Jault, F. M.; Jault, J. M.; Ruchti, F.; Fortunato, E. A.; Clark, C.; Corbeil, J.; Richman, D. D.; Spector, D. H., Cytomegalovirus infection induces high levels of cyclins, phosphorylated Rb, and p53, leading to cell cycle arrest. *Journal of virology* **1995,** *69* (11), 6697-704.

173. Roy, S.; He, R.; Kapoor, A.; Forman, M.; Mazzone, J. R.; Posner, G. H.; Arav-Boger, R., Inhibition of human cytomegalovirus replication by artemisinins: effects mediated through cell cycle modulation. *Antimicrobial agents and chemotherapy* **2015,** *59* (7), 3870-9.

174. Sanchez, V.; McElroy, A. K.; Yen, J.; Tamrakar, S.; Clark, C. L.; Schwartz, R. A.; Spector, D. H., Cyclin-dependent kinase activity is required at early times for accurate processing and accumulation of the human cytomegalovirus UL122-123 and UL37 immediate-early transcripts and at later times for virus production. *Journal of virology* **2004,** *78* (20), 11219-32.

175. Zydek, M.; Hagemeier, C.; Wiebusch, L., Cyclin-dependent kinase activity controls the onset of the HCMV lytic cycle. *PLoS pathogens* **2010,** *6* (9), e1001096.

176. Drori, A.; Messerle, M.; Brune, W.; Tirosh, B., Lack of XBP-1 impedes murine cytomegalovirus gene expression. *PloS one* **2014,** *9* (10), e110942.

177. Kaspari, M.; Tavalai, N.; Stamminger, T.; Zimmermann, A.; Schilf, R.; Bogner, E., Proteasome inhibitor MG132 blocks viral DNA replication and assembly of human cytomegalovirus. *FEBS letters* **2008,** *582* (5), 666-72.

178. Prosch, S.; Priemer, C.; Hoflich, C.; Liebenthaf, C.; Babel, N.; Kruger, D. H.; Volk, H. D., Proteasome inhibitors: a novel tool to suppress human cytomegalovirus replication and virus-induced immune modulation. *Antiviral therapy* **2003,** *8* (6), 555-67.

179. Isler, J. A.; Skalet, A. H.; Alwine, J. C., Human cytomegalovirus infection activates and regulates the unfolded protein response. *Journal of virology* **2005,** *79* (11), 6890-9.

180. Roy, S.; Arav-Boger, R., New cell-signaling pathways for controlling cytomegalovirus replication. *American journal of transplantation : official journal of the American Society of Transplantation and the American Society of Transplant Surgeons* **2014,** *14* (6), 1249-58.

181. Sung, H.; Schleiss, M. R., Update on the current status of cytomegalovirus vaccines. *Expert review of vaccines* **2010,** *9* (11), 1303-14.

182. Cannon, M. J.; Schmid, D. S.; Hyde, T. B., Review of cytomegalovirus seroprevalence and demographic characteristics associated with infection. *Reviews in medical virology* **2010,** *20* (4), 202-13.

183. Hill, A. B., The immune response to CMV infection and vaccination in mice, monkeys and humans: recent developments. *Current opinion in virology* **2018,** *28*, 161-166.

184. Lurain, N. S.; Chou, S., Antiviral drug resistance of human cytomegalovirus. *Clinical microbiology reviews* **2010,** *23* (4), 689-712.

185. Jean Beltran, P. M.; Mathias, R. A.; Cristea, I. M., A Portrait of the Human Organelle Proteome In Space and Time during Cytomegalovirus Infection. *Cell systems* **2016,** *3* (4), 361-373 e6.

186. Arend, K. C.; Lenarcic, E. M.; Vincent, H. A.; Rashid, N.; Lazear, E.; McDonald, I. M.; Gilbert, T. S.; East, M. P.; Herring, L. E.; Johnson, G. L.; Graves, L. M.; Moorman, N. J., Kinome Profiling Identifies Druggable Targets for Novel Human Cytomegalovirus (HCMV) Antivirals. *Molecular & cellular proteomics : MCP* **2017,** *16* (4 suppl 1), S263-S276.

187. John, S.; Yuzhakov, O.; Woods, A.; Deterling, J.; Hassett, K.; Shaw, C. A.; Ciaramella, G., Multi-antigenic human cytomegalovirus mRNA vaccines that elicit potent humoral and cell-mediated immunity. *Vaccine* **2018,** *36* (12), 1689-1699.

188. Yunis, J.; Farrell, H. E.; Bruce, K.; Lawler, C.; Sidenius, S.; Wyer, O.; Davis-Poynter, N.; Stevenson, P. G., Murine cytomegalovirus degrades MHC class II to colonize the salivary glands. *PLoS pathogens* **2018,** *14* (2), e1006905.

189. Lanoix, J.; Durette, C.; Courcelles, M.; Cossette, E.; Comtois-Marotte, S.; Hardy, M. P.; Cote, C.; Perreault, C.; Thibault, P., Comparison of the MHC I Immunopeptidome Repertoire of B-Cell Lymphoblasts using Two Isolation Methods. *Proteomics* **2018**.

190. Oberstein, A.; Perlman, D. H.; Shenk, T.; Terry, L. J., Human cytomegalovirus pUL97 kinase induces global changes in the infected cell phosphoproteome. *Proteomics* **2015,** *15* (12), 2006-22.

191. Olsen, J. V.; de Godoy, L. M.; Li, G.; Macek, B.; Mortensen, P.; Pesch, R.; Makarov, A.; Lange, O.; Horning, S.; Mann, M., Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. *Molecular & cellular proteomics : MCP* **2005,** *4* (12), 2010-21.

192. Burkhart, J. M.; Premsler, T.; Sickmann, A., Quality control of nano-LC-MS systems using stable isotope-coded peptides. *Proteomics* **2011,** *11* (6), 1049-57.

193. Rawlinson, W. D.; Farrell, H. E.; Barrell, B. G., Analysis of the complete DNA sequence of murine cytomegalovirus. *Journal of virology* **1996,** *70* (12), 8833-49.

194. Taus, T.; Kocher, T.; Pichler, P.; Paschke, C.; Schmidt, A.; Henrich, C.; Mechtler, K., Universal and confident phosphorylation site localization using phosphoRS. *Journal of proteome research* **2011,** *10* (12), 5354-62.

195. Schwammle, V.; Leon, I. R.; Jensen, O. N., Assessment and improvement of statistical tools for comparative proteomics analysis of sparse data sets with few experimental replicates. *Journal of proteome research* **2013,** *12* (9), 3874-83.

196. Wettenhall, J. M.; Smyth, G. K., limmaGUI: a graphical user interface for linear modeling of microarray data. *Bioinformatics (Oxford, England)* **2004,** *20* (18), 3705-6.

197. Huang da, W.; Sherman, B. T.; Lempicki, R. A., Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols* **2009,** *4* (1), 44-57.

198. Ogata, H.; Goto, S.; Sato, K.; Fujibuchi, W.; Bono, H.; Kanehisa, M., KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic acids research* **1999,** *27* (1), 29-34.

199. Luo, W.; Brouwer, C., Pathview: an R/Bioconductor package for pathway-based data integration and visualization. *Bioinformatics (Oxford, England)* **2013,** *29* (14), 1830-1.

200. Gonczarowska-Jorge, H.; Loroch, S.; Dell'Aica, M.; Sickmann, A.; Roos, A.; Zahedi, R. P., Quantifying Missing (Phospho)Proteome Regions with the Broad-Specificity Protease Subtilisin. *Analytical chemistry* **2017,** *89* (24), 13137-13145.

201. Adzhubei, A. A.; Sternberg, M. J.; Makarov, A. A., Polyproline-II helix in proteins: structure and function. *Journal of molecular biology* **2013,** *425* (12), 2100-32.

202. Di Conza, G.; Trusso Cafarello, S.; Loroch, S.; Mennerich, D.; Deschoemaeker, S.; Di Matteo, M.; Ehling, M.; Gevaert, K.; Prenen, H.; Zahedi, R. P.; Sickmann, A.; Kietzmann, T.; Moretti, F.; Mazzone, M., The mTOR and PP2A Pathways Regulate PHD2 Phosphorylation to Fine-Tune HIF1alpha Levels and Colorectal Cancer Cell Survival under Hypoxia. *Cell reports* **2017,** *18* (7), 1699-1712.

203. Gill, R. B.; James, S. H.; Prichard, M. N., Human cytomegalovirus UL97 kinase alters the accumulation of CDK1. *The Journal of general virology* **2012,** *93* (Pt 8), 1743-55.

204. Ostermann, E.; Warnecke, G.; Waibler, Z.; Brune, W., Knockout of the Host Resistance Gene Pkr Fully Restores Replication of Murine Cytomegalovirus m142 and m143 Mutants In Vivo. *Journal of virology* **2015,** *90* (2), 1144-7.

205. Sharp, E. L.; Davis-Poynter, N. J.; Farrell, H. E., Analysis of the subcellular trafficking properties of murine cytomegalovirus M78, a 7 transmembrane receptor homologue. *The Journal of general virology* **2009,** *90* (Pt 1), 59-68.

206. Wu, C. A.; Carlson, M. E.; Henry, S. C.; Shanley, J. D., The murine cytomegalovirus M25 open reading frame encodes a component of the tegument. *Virology* **1999,** *262* (2), 265-76.

207. Krause, E.; de Graaf, M.; Fliss, P. M.; Dolken, L.; Brune, W., Murine cytomegalovirus virion-associated protein M45 mediates rapid NF-kappaB activation after infection. *Journal of virology* **2014,** *88* (17), 9963-75.

208. Singh, R.; Brewer, M. K.; Mashburn, C. B.; Lou, D.; Bondada, V.; Graham, B.; Geddes, J. W., Calpain 5 is highly expressed in the central nervous system (CNS), carries dual nuclear localization signals, and is associated with nuclear promyelocytic leukemia protein bodies. *J Biol Chem* **2014,** *289* (28), 19383-94.

209. Arthur, J. S.; Cohen, P., MSK1 is required for CREB phosphorylation in response to mitogens in mouse embryonic stem cells. *FEBS letters* **2000,** *482* (1-2), 44-8.

210. Duncan, E. A.; Anest, V.; Cogswell, P.; Baldwin, A. S., The kinases MSK1 and MSK2 are required for epidermal growth factor-induced, but not tumor necrosis factor-induced, histone H3 Ser10 phosphorylation. *J Biol Chem* **2006,** *281* (18), 12521-5.

211. Xing, J.; Ginty, D. D.; Greenberg, M. E., Coupling of the RAS-MAPK pathway to gene activation by RSK2, a growth factor-regulated CREB kinase. *Science* **1996,** *273* (5277), 959-63.

212. Peng, C.; Cho, Y. Y.; Zhu, F.; Xu, Y. M.; Wen, W.; Ma, W. Y.; Bode, A. M.; Dong, Z., RSK2 mediates NF-{kappa}B activity through the phosphorylation of IkappaBalpha in the TNF-R1 pathway. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **2010,** *24* (9), 3490-9.

213. Yamashiro, S.; Yamakita, Y.; Totsukawa, G.; Goto, H.; Kaibuchi, K.; Ito, M.; Hartshorne, D. J.; Matsumura, F., Myosin phosphatase-targeting subunit 1 regulates mitosis by antagonizing polo-like kinase 1. *Developmental cell* **2008,** *14* (5), 787-97.

214. Carnegie, G. K.; Sleeman, J. E.; Morrice, N.; Hastie, C. J.; Peggie, M. W.; Philp, A.; Lamond, A. I.; Cohen, P. T., Protein phosphatase 4 interacts with the Survival of Motor Neurons complex and enhances the temporal localisation of snRNPs. *Journal of cell science* **2003,** *116* (Pt 10), 1905-13.

215. Chowdhury, D.; Xu, X.; Zhong, X.; Ahmed, F.; Zhong, J.; Liao, J.; Dykxhoorn, D. M.; Weinstock, D. M.; Pfeifer, G. P.; Lieberman, J., A PP4-phosphatase complex dephosphorylates gamma-H2AX generated during DNA replication. *Molecular cell* **2008,** *31* (1), 33-46.

216. Andreu-Vieyra, C. V.; Chen, R.; Agno, J. E.; Glaser, S.; Anastassiadis, K.; Stewart, A. F.; Matzuk, M. M., MLL2 is required in oocytes for bulk histone 3 lysine 4 trimethylation and transcriptional silencing. *PLoS biology* **2010,** *8* (8).

217. Patterson, K. I.; Brummer, T.; O'Brien, P. M.; Daly, R. J., Dual-specificity phosphatases: critical regulators with diverse cellular targets. *The Biochemical journal* **2009,** *418* (3), 475-89.

218. Marcinowski, L.; Lidschreiber, M.; Windhager, L.; Rieder, M.; Bosse, J. B.; Radle, B.; Bonfert, T.; Gyory, I.; de Graaf, M.; Prazeres da Costa, O.; Rosenstiel, P.; Friedel, C. C.; Zimmer, R.; Ruzsics, Z.; Dolken, L., Real-time transcriptional profiling of cellular and viral gene expression during lytic cytomegalovirus infection. *PLoS pathogens* **2012,** *8* (9), e1002908.

219. Oduro, J. D.; Uecker, R.; Hagemeier, C.; Wiebusch, L., Inhibition of human cytomegalovirus immediate-early gene expression by cyclin A2-dependent kinase activity. *Journal of virology* **2012,** *86* (17), 9369-83.

220. Rao, H.; Sastry, A., Recognition of specific ubiquitin conjugates is important for the proteolytic functions of the ubiquitin-associated domain proteins Dsk2 and Rad23. *J Biol Chem* **2002,** *277* (14), 11691-5.

221. Subramaniam, V. N.; Loh, E.; Horstmann, H.; Habermann, A.; Xu, Y.; Coe, J.; Griffiths, G.; Hong, W., Preferential association of syntaxin 8 with the early endosome. *Journal of cell science* **2000,** *113 (Pt 6)*, 997-1008.

222. Belgareh-Touze, N.; Corral-Debrinski, M.; Launhardt, H.; Galan, J. M.; Munder, T.; Le Panse, S.; Haguenauer-Tsapis, R., Yeast functional analysis: identification of two essential genes involved in ER to Golgi trafficking. *Traffic (Copenhagen, Denmark)* **2003,** *4* (9), 607-17.

223. Lewis, M. J.; Rayner, J. C.; Pelham, H. R., A novel SNARE complex implicated in vesicle fusion with the endoplasmic reticulum. *The EMBO journal* **1997,** *16* (11), 3017-24.

224. Wakabayashi, M.; Kyono, Y.; Sugiyama, N.; Ishihama, Y., Extended Coverage of Singly and Multiply Phosphorylated Peptides from a Single Titanium Dioxide Microcolumn. *Analytical chemistry* **2015,** *87* (20), 10213-21.

225. Kanshin, E.; Michnick, S. W.; Thibault, P., Displacement of N/Q-rich peptides on TiO2 beads enhances the depth and coverage of yeast phosphoproteome analyses. *Journal of proteome research* **2013,** *12* (6), 2905-13.

226. Eriksson, A. I.; Edwards, K.; Agmo Hernandez, V., Cooperative adsorption behavior of phosphopeptides on TiO leads to biased enrichment, detection and quantification. *The Analyst* **2014**.

227. Thingholm, T. E.; Jorgensen, T. J.; Jensen, O. N.; Larsen, M. R., Highly selective enrichment of phosphorylated peptides using titanium dioxide. *Nature protocols* **2006,** *1* (4), 1929-35.

228. Thingholm, T.; Jensen, O.; Larsen, M., Enrichment and Separation of Mono- and Multiply Phosphorylated Peptides Using Sequential Elution from IMAC Prior to Mass Spectrometric Analysis. In *Phospho-Proteomics*, Graauw, M., Ed. Humana Press: 2009; Vol. 527, pp 67-78.

229. Zhang, G.; Neubert, T. A., Comparison of three quantitative phosphoproteomic strategies to study receptor tyrosine kinase signaling. *Journal of proteome research* **2011,** *10* (12), 5454-62.

230. Di Palma, S.; Zoumaro-Djayoon, A.; Peng, M.; Post, H.; Preisinger, C.; Munoz, J.; Heck, A. J., Finding the same needles in the haystack? A comparison of phosphotyrosine peptides enriched by immuno-affinity precipitation and metal-based affinity chromatography. *Journal of proteomics* **2013,** *91*, 331-7.

231. Zarei, M.; Sprenger, A.; Metzger, F.; Gretzmeier, C.; Dengjel, J., Comparison of ERLIC-TiO2, HILIC-TiO2, and SCX-TiO2 for global phosphoproteomics approaches. *Journal of proteome research* **2011,** *10* (8), 3474-83.

232. Hao, P.; Ren, Y.; Dutta, B.; Sze, S. K., Comparative evaluation of electrostatic repulsion-hydrophilic interaction chromatography (ERLIC) and high-pH reversed phase (Hp-RP) chromatography in profiling of rat kidney proteome. *Journal of proteomics* **2013,** *82*, 254-62.

233. Rawlinson, W. D.; Zeng, F.; Farrell, H. E.; Cunningham, A. L.; Scalzo, A. A.; Booth, T. W.; Scott, G. M., The murine cytomegalovirus (MCMV) homolog of the HCMV phosphotransferase (UL97(pk)) gene. *Virology* **1997,** *233* (2), 358-63.

234. He, Z.; He, Y. S.; Kim, Y.; Chu, L.; Ohmstede, C.; Biron, K. K.; Coen, D. M., The human cytomegalovirus UL97 protein is a protein kinase that autophosphorylates on serines and threonines. *Journal of virology* **1997,** *71* (1), 405-11.

235. Ye, M.; Morello, C. S.; Spector, D. H., Multiple epitopes in the murine cytomegalovirus early gene product M84 are efficiently presented in infected primary macrophages and contribute to strong CD8+-T-lymphocyte responses and protection following DNA immunization. *Journal of virology* **2004,** *78* (20), 11233-45.

236. Fink, A.; Renzaho, A.; Reddehase, M. J.; Lemmermann, N. A., The p36 isoform of murine cytomegalovirus m152 protein suffices for mediating innate and adaptive immune evasion. *Viruses* **2013,** *5* (12), 3171-91.

237. Seckert, C. K.; Schader, S. I.; Ebert, S.; Thomas, D.; Freitag, K.; Renzaho, A.; Podlech, J.; Reddehase, M. J.; Holtappels, R., Antigen-presenting cells of haematopoietic origin prime cytomegalovirus-specific CD8 T-cells but are not sufficient for driving memory inflation during viral latency. *The Journal of general virology* **2011,** *92* (Pt 9), 1994-2005.

238. Stempel, M.; Chan, B.; Juranic Lisnic, V.; Krmpotic, A.; Hartung, J.; Paludan, S. R.; Fullbrunn, N.; Lemmermann, N. A.; Brinkmann, M. M., The herpesviral antagonist m152 reveals differential activation of STING-dependent IRF and NF-kappaB signaling and STING's dual role during MCMV infection. *The EMBO journal* **2019,** *38* (5).

239. Menard, C.; Wagner, M.; Ruzsics, Z.; Holak, K.; Brune, W.; Campbell, A. E.; Koszinowski, U. H., Role of murine cytomegalovirus US22 gene family members in replication in macrophages. *Journal of virology* **2003,** *77* (10), 5557-70.

240. McCormick, A. L.; Skaletskaya, A.; Barry, P. A.; Mocarski, E. S.; Goldmacher, V. S., Differential function and expression of the viral inhibitor of caspase 8-induced apoptosis (vICA) and the viral mitochondria-localized inhibitor of apoptosis (vMIA) cell death suppressors conserved in primate and rodent cytomegaloviruses. *Virology* **2003,** *316* (2), 221-33.

241. Skaletskaya, A.; Bartle, L. M.; Chittenden, T.; McCormick, A. L.; Mocarski, E. S.; Goldmacher, V. S., A cytomegalovirus-encoded inhibitor of apoptosis that suppresses caspase-8 activation. *Proceedings of the National Academy of Sciences of the United States of America* **2001,** *98* (14), 7829-34.

242. Bozidis, P.; Williamson, C. D.; Colberg-Poley, A. M., Mitochondrial and secretory human cytomegalovirus UL37 proteins traffic into mitochondrion-associated membranes of human cells. *Journal of virology* **2008,** *82* (6), 2715-26.

243. Kavanagh, D. G.; Gold, M. C.; Wagner, M.; Koszinowski, U. H.; Hill, A. B., The multiple immune-evasion genes of murine cytomegalovirus are not redundant: m4 and m152 inhibit antigen presentation in a complementary and cooperative fashion. *The Journal of experimental medicine* **2001,** *194* (7), 967-78.

244. Holtappels, R.; Grzimek, N. K.; Thomas, D.; Reddehase, M. J., Early gene m18, a novel player in the immune response to murine cytomegalovirus. *The Journal of general virology* **2002,** *83* (Pt 2), 311-6.

245. Doring, M.; Lessin, I.; Frenz, T.; Spanier, J.; Kessler, A.; Tegtmeyer, P.; Dag, F.; Thiel, N.; Trilling, M.; Lienenklaus, S.; Weiss, S.; Scheu, S.; Messerle, M.; Cicin-Sain, L.; Hengel, H.; Kalinke,

U., M27 expressed by cytomegalovirus counteracts effective type I interferon induction of myeloid cells but not of plasmacytoid dendritic cells. *Journal of virology* **2014,** *88* (23), 13638-50.

246. Morello, C. S.; Kelley, L. A.; Munks, M. W.; Hill, A. B.; Spector, D. H., DNA immunization using highly conserved murine cytomegalovirus genes encoding homologs of human cytomegalovirus UL54 (DNA polymerase) and UL105 (helicase) elicits strong CD8 T-cell responses and is protective against systemic challenge. *Journal of virology* **2007,** *81* (14), 7766-75.

247. Wang, J. B.; McVoy, M. A., Mutagenesis of the murine cytomegalovirus M56 terminase gene. *The Journal of general virology* **2008,** *89* (Pt 11), 2864-8.

248. Morello, C. S.; Cranmer, L. D.; Spector, D. H., Suppression of murine cytomegalovirus (MCMV) replication with a DNA vaccine encoding MCMV M84 (a homolog of human cytomegalovirus pp65). *Journal of virology* **2000,** *74* (8), 3696-708.

249. Morello, C. S.; Cranmer, L. D.; Spector, D. H., In vivo replication, latency, and immunogenicity of murine cytomegalovirus mutants with deletions in the M83 and M84 genes, the putative homologs of human cytomegalovirus pp65 (UL83). *Journal of virology* **1999,** *73* (9), 7678-93.

250. Zarrouk, K.; Piret, J.; Boivin, G., Herpesvirus DNA polymerases: Structures, functions and inhibitors. *Virus research* **2017**.

251. Kronemann, D.; Hagemeier, S. R.; Cygnar, D.; Phillips, S.; Bresnahan, W. A., Binding of the human cytomegalovirus (HCMV) tegument protein UL69 to UAP56/URH49 is not required for efficient replication of HCMV. *Journal of virology* **2010,** *84* (18), 9649-54.

252. Yoshida, H., ER stress and diseases. *The FEBS journal* **2007,** *274* (3), 630-58.

253. Schroder, M.; Kaufman, R. J., The mammalian unfolded protein response. *Annu Rev Biochem* **2005,** *74*, 739-89.

254. Henkens, R.; Delvenne, P.; Arafa, M.; Moutschen, M.; Zeddou, M.; Tautz, L.; Boniver, J.; Mustelin, T.; Rahmouni, S., Cervix carcinoma is associated with an up-regulation and nuclear localization of the dual-specificity protein phosphatase VHR. *BMC cancer* **2008,** *8*, 147.

255. Tang-Feldman, Y. J.; Lochhead, S. R.; Lochhead, G. R.; Yu, C.; George, M.; Villablanca, A. C.; Pomeroy, C., Murine cytomegalovirus (MCMV) infection upregulates P38 MAP kinase in aortas of Apo E KO mice: a molecular mechanism for MCMV-induced acceleration of atherosclerosis. *Journal of cardiovascular translational research* **2013,** *6* (1), 54-64.

256. Chaudhuri, S.; Lowen, B.; Chan, G.; Davey, A.; Riddell, M.; Guilbert, L. J., Human cytomegalovirus interacts with toll-like receptor 2 and CD14 on syncytiotrophoblasts to stimulate expression of TNFalpha mRNA and apoptosis. *Placenta* **2009,** *30* (11), 994-1001.

257. Huang, H.; Chan, H.; Wang, Y. Y.; Ouyang, D. Y.; Zheng, Y. T.; Tam, S. C., Trichosanthin suppresses the elevation of p38 MAPK, and Bcl-2 induced by HSV-1 infection in Vero cells. *Life sciences* **2006,** *79* (13), 1287-92.

258. Zachos, G.; Clements, B.; Conner, J., Herpes simplex virus type 1 infection stimulates p38/c-Jun N-terminal mitogen-activated protein kinase pathways and activates transcription factor AP-1. *J Biol Chem* **1999,** *274* (8), 5097-103.

259. Nithianandarajah-Jones, G. N.; Wilm, B.; Goldring, C. E.; Muller, J.; Cross, M. J., ERK5: structure, regulation and function. *Cellular signalling* **2012,** *24* (11), 2187-96.

260. Basha, W.; Kitagawa, R.; Uhara, M.; Imazu, H.; Uechi, K.; Tanaka, J., Geldanamycin, a potent and specific inhibitor of Hsp90, inhibits gene expression and replication of human cytomegalovirus. *Antiviral chemistry & chemotherapy* **2005,** *16* (2), 135-46.

261. Chan, G.; Nogalski, M. T.; Bentz, G. L.; Smith, M. S.; Parmater, A.; Yurochko, A. D., PI3K-dependent upregulation of Mcl-1 by human cytomegalovirus is mediated by epidermal growth factor receptor and inhibits apoptosis in short-lived monocytes. *Journal of immunology (Baltimore, Md. : 1950)* **2010,** *184* (6), 3213-22.

262. Buehler, J.; Zeltzer, S.; Reitsma, J.; Petrucelli, A.; Umashankar, M.; Rak, M.; Zagallo, P.; Schroeder, J.; Terhune, S.; Goodrum, F., Opposing Regulation of the EGF Receptor: A Molecular Switch Controlling Cytomegalovirus Latency and Replication. *PLoS pathogens* **2016,** *12* (5), e1005655.

263. Mathers, C.; Schafer, X.; Martinez-Sobrido, L.; Munger, J., The human cytomegalovirus UL26 protein antagonizes NF-kappaB activation. *Journal of virology* **2014,** *88* (24), 14289-300.

264. McNeish, I. A.; Bell, S.; McKay, T.; Tenev, T.; Marani, M.; Lemoine, N. R., Expression of Smac/DIABLO in ovarian carcinoma cells induces apoptosis via a caspase-9-mediated pathway. *Experimental cell research* **2003,** *286* (2), 186-98.

265. Andoniou, C. E.; Fleming, P.; Sutton, V. R.; Trapani, J. A.; Degli-Esposti, M. A., Cathepsin C limits acute viral infection independently of NK cell and CD8+ T-cell cytolytic function. *Immunology and cell biology* **2011,** *89* (4), 540-8.

266. Wiebusch, L.; Uecker, R.; Hagemeier, C., Human cytomegalovirus prevents replication licensing by inhibiting MCM loading onto chromatin. *EMBO reports* **2003,** *4* (1), 42-6.

267. Biswas, N.; Sanchez, V.; Spector, D. H., Human cytomegalovirus infection leads to accumulation of geminin and inhibition of the licensing of cellular DNA replication. *Journal of virology* **2003,** *77* (4), 2369-76.

268. Kawaguchi, A.; Nagata, K., De novo replication of the influenza virus RNA genome is regulated by DNA replicative helicase, MCM. *The EMBO journal* **2007,** *26* (21), 4566-75.

269. Montagnoli, A.; Valsasina, B.; Brotherton, D.; Troiani, S.; Rainoldi, S.; Tenca, P.; Molinari, A.; Santocanale, C., Identification of Mcm2 phosphorylation sites by S-phase-regulating kinases. *J Biol Chem* **2006,** *281* (15), 10281-90.

270. Kudoh, A.; Daikoku, T.; Ishimi, Y.; Kawaguchi, Y.; Shirata, N.; Iwahori, S.; Isomura, H.; Tsurumi, T., Phosphorylation of MCM4 at sites inactivating DNA helicase activity of the MCM4-MCM6-MCM7 complex during Epstein-Barr virus productive replication. *Journal of virology* **2006,** *80* (20), 10064-72.

271. Benson, J. D.; Huang, E. S., Human cytomegalovirus induces expression of cellular topoisomerase II. *Journal of virology* **1990,** *64* (1), 9-15.

272. Tarrant-Elorza, M.; Rossetto, C. C.; Pari, G. S., Maintenance and replication of the human cytomegalovirus genome during latency. *Cell host & microbe* **2014,** *16* (1), 43-54.

273. Poole, E.; Bain, M.; Teague, L.; Takei, Y.; Laskey, R.; Sinclair, J., The cellular protein MCM3AP is required for inhibition of cellular DNA synthesis by the IE86 protein of human cytomegalovirus. *PloS one* **2012,** *7* (10), e45686.

274. Gambus, A.; van Deursen, F.; Polychronopoulos, D.; Foltman, M.; Jones, R. C.; Edmondson, R. D.; Calzada, A.; Labib, K., A key role for Ctf4 in coupling the MCM2-7 helicase to DNA polymerase alpha within the eukaryotic replisome. *The EMBO journal* **2009,** *28* (19), 2992-3004.

275. Tamrakar, S.; Kapasi, A. J.; Spector, D. H., Human cytomegalovirus infection induces specific hyperphosphorylation of the carboxyl-terminal domain of the large subunit of RNA polymerase II that is associated with changes in the abundance, activity, and localization of cdk9 and cdk7. *Journal of virology* **2005,** *79* (24), 15477-93.

276. Perng, Y. C.; Campbell, J. A.; Lenschow, D. J.; Yu, D., Human cytomegalovirus pUL79 is an elongation factor of RNA polymerase II for viral gene transcription. *PLoS pathogens* **2014,** *10* (8), e1004350.

277. Sun, C.; Querol-Audi, J.; Mortimer, S. A.; Arias-Palomo, E.; Doudna, J. A.; Nogales, E.; Cate, J. H., Two RNA-binding motifs in eIF3 direct HCV IRES-dependent translation. *Nucleic acids research* **2013,** *41* (15), 7512-21.

278. Hashem, Y.; des Georges, A.; Dhote, V.; Langlois, R.; Liao, H. Y.; Grassucci, R. A.; Pestova, T. V.; Hellen, C. U.; Frank, J., Hepatitis-C-virus-like internal ribosome entry sites displace eIF3 to gain access to the 40S subunit. *Nature* **2013,** *503* (7477), 539-43.

279. Cheshenko, N.; Trepanier, J. B.; Segarra, T. J.; Fuller, A. O.; Herold, B. C., HSV usurps eukaryotic initiation factor 3 subunit M for viral protein translation: novel prevention target. *PloS one* **2010,** *5* (7), e11829.

280. Gradi, A.; Svitkin, Y. V.; Sommergruber, W.; Imataka, H.; Morino, S.; Skern, T.; Sonenberg, N., Human rhinovirus 2A proteinase cleavage sites in eukaryotic initiation factors (eIF) 4GI and eIF4GII are different. *Journal of virology* **2003,** *77* (8), 5026-9.

281. Rodriguez Pulido, M.; Serrano, P.; Saiz, M.; Martinez-Salas, E., Foot-and-mouth disease virus infection induces proteolytic cleavage of PTB, eIF3a,b, and PABP RNA-binding proteins. *Virology* **2007,** *364* (2), 466-74.

282. McKinney, C.; Perez, C.; Mohr, I., Poly(A) binding protein abundance regulates eukaryotic translation initiation factor 4F assembly in human cytomegalovirus-infected cells. *Proceedings of the National Academy of Sciences of the United States of America* **2012,** *109* (15), 5627-32.

283. Tagliabracci, V. S.; Wiley, S. E.; Guo, X.; Kinch, L. N.; Durrant, E.; Wen, J.; Xiao, J.; Cui, J.; Nguyen, K. B.; Engel, J. L.; Coon, J. J.; Grishin, N.; Pinna, L. A.; Pagliarini, D. J.; Dixon, J. E., A Single Kinase Generates the Majority of the Secreted Phosphoproteome. *Cell* **2015,** *161* (7), 1619-32.

284. Shin, H. J.; Kim, S. S.; Cho, Y. H.; Lee, S. G.; Rho, H. M., Host cell proteins binding to the encapsidation signal epsilon in hepatitis B virus RNA. *Archives of virology* **2002,** *147* (3), 471-91.

285. Zheng, Z.; Maidji, E.; Tugizov, S.; Pereira, L., Mutations in the carboxyl-terminal hydrophobic sequence of human cytomegalovirus glycoprotein B alter transport and protein chaperone binding. *Journal of virology* **1996,** *70* (11), 8029-40.

286. Buchkovich, N. J.; Maguire, T. G.; Yu, Y.; Paton, A. W.; Paton, J. C.; Alwine, J. C., Human cytomegalovirus specifically controls the levels of the endoplasmic reticulum chaperone BiP/GRP78, which is required for virion assembly. *Journal of virology* **2008,** *82* (1), 31-9.

287. Qian, Z.; Xuan, B.; Chapa, T. J.; Gualberto, N.; Yu, D., Murine cytomegalovirus targets transcription factor ATF4 to exploit the unfolded-protein response. *Journal of virology* **2012,** *86* (12), 6712-23.

288. Yang, J.; Liu, H.; Li, L.; Liu, H.; Shi, W.; Yuan, X.; Wu, L., Structural Insights into IRE1 Functions in the Unfolded Protein Response. *Current medicinal chemistry* **2016,** *23* (41), 4706-4716.

289. Liu, S. T.; Sharon-Friling, R.; Ivanova, P.; Milne, S. B.; Myers, D. S.; Rabinowitz, J. D.; Brown, H. A.; Shenk, T., Synaptic vesicle-like lipidome of human cytomegalovirus virions reveals a role for SNARE machinery in virion egress. *Proceedings of the National Academy of Sciences of the United States of America* **2011,** *108* (31), 12869-74.

290. Jovic, M.; Kean, M. J.; Dubankova, A.; Boura, E.; Gingras, A. C.; Brill, J. A.; Balla, T., Endosomal sorting of VAMP3 is regulated by PI4K2A. *Journal of cell science* **2014,** *127* (Pt 17), 3745-56.

291. Hook, L. M.; Grey, F.; Grabski, R.; Tirabassi, R.; Doyle, T.; Hancock, M.; Landais, I.; Jeng, S.; McWeeney, S.; Britt, W.; Nelson, J. A., Cytomegalovirus miRNAs target secretory pathway genes to facilitate formation of the virion assembly compartment and reduce cytokine secretion. *Cell host & microbe* **2014,** *15* (3), 363-73.

292. Post, H.; Penning, R.; Fitzpatrick, M. A.; Garrigues, L. B.; Wu, W.; MacGillavry, H. D.; Hoogenraad, C. C.; Heck, A. J.; Altelaar, A. F., Robust, Sensitive, and Automated Phosphopeptide Enrichment Optimized for Low Sample Amounts Applied to Primary Hippocampal Neurons. *Journal of proteome research* **2017,** *16* (2), 728-737.

293. Badgett, M. J.; Boyes, B.; Orlando, R., Peptide retention prediction using hydrophilic interaction liquid chromatography coupled to mass spectrometry. *Journal of chromatography. A* **2018,** *1537*, 58-65.

294. Maboudi Afkham, H.; Qiu, X.; The, M.; Kall, L., Uncertainty estimation of predictions of peptides' chromatographic retention times in shotgun proteomics. *Bioinformatics (Oxford, England)* **2017,** *33* (4), 508-513.

295. Moruz, L.; Staes, A.; Foster, J. M.; Hatzou, M.; Timmerman, E.; Martens, L.; Kall, L., Chromatographic retention time prediction for posttranslationally modified peptides. *Proteomics* **2012,** *12* (8), 1151-9.

296. Ahmed, A., Antiviral treatment of cytomegalovirus infection. *Infectious disorders drug targets* **2011,** *11* (5), 475-503.

297. Piret, J.; Boivin, G., Clinical development of letermovir and maribavir: Overview of human cytomegalovirus drug resistance. *Antiviral research* **2019,** *163*, 91-105.

298. Klaeger, S.; Heinzlmeir, S.; Wilhelm, M.; Polzer, H.; Vick, B.; Koenig, P. A.; Reinecke, M.; Ruprecht, B.; Petzoldt, S.; Meng, C.; Zecha, J.; Reiter, K.; Qiao, H.; Helm, D.; Koch, H.; Schoof, M.; Canevari, G.; Casale, E.; Depaolini, S. R.; Feuchtinger, A.; Wu, Z.; Schmidt, T.; Rueckert, L.; Becker, W.; Huenges, J.; Garz, A. K.; Gohlke, B. O.; Zolg, D. P.; Kayser, G.; Vooder, T.; Preissner, R.; Hahne, H.; Tonisson, N.; Kramer, K.; Gotze, K.; Bassermann, F.; Schlegl, J.; Ehrlich, H. C.; Aiche, S.; Walch, A.; Greif, P. A.; Schneider, S.; Felder, E. R.; Ruland, J.; Medard, G.; Jeremias, I.; Spiekermann, K.; Kuster, B., The target landscape of clinical kinase drugs. *Science* **2017,** *358* (6367).

10. Appendix

10.1. Regulated viral phosphopeptides

				0 hpi			4 hpi			8 hpi			24 hpi		
Entry	Sequence / phosphosite positions	Sites Protein	Entry	1	2	3	1	2	3	1	2	3	1	2	3
IE3	HHPFTPPSAKR/5, 8	T295, S298	P29832|VIE3	NA	NA	NA	NA	NA	NA	0.9	0.2	1.3	4.5	1.8	3.8
IE3	HHPFTPPSAK/5, 8	T295, S298	P29832|VIE3	NA	NA	NA	NA	NA	NA	NA	0.4	1.1	NA	1.7	2.9
M116	ASSPQPGAAPAAPANTEPAR/3	S300	A8E1M7|A8E1M7	0.0	0.0	0.0	1.3	0.0	1.1	2.0	2.2	2.0	0.7	0.8	0.7
M116	ASSPQPGAAPAAPANTEPARDASTELSSK/16	T313	A8E1M7|A8E1M7	NA	NA	NA	NA	NA	NA	0.0	NA	1.1	1.5	NA	25
M116	ASSPQPGAAPAAPANTEPAR/2	S299	A8E1M7|A8E1M7	NA	0.0	NA	NA	0.0	NA	2.1	2.3	2.1	0.7	0.8	0.7
M143	EKPPPPSPEAGR/7	S435	A8E1Q8|A8E1Q8	0.0	0.0	0.0	0.0	0.0	0.3	0.9	1.0	1.5	1.1	1.5	1.6
M143	EKPPPPSPEAGRYDTDQSPAAVAR/7	S435	A8E1Q8|A8E1Q8	0.0	0.0	0.0	0.5	0.0	0.4	1.0	0.8	1.2	1.3	1.4	1.5
M143	FKEKPPPPSPEAGR/9	S435	A8E1Q8|A8E1Q8	0.0	0.0	0.0	0.0	0.0	0.4	1.4	0.9	NA	2.1	1.4	NA
M143	YTTDDSPPPAPPSTPNSTGSSPSPRPIPAPR/6	S503	A8E1Q8|A8E1Q8	0.0	NA	NA	1.1	NA	NA	0.6	0.5	0.7	1.6	1.4	1.9
M143	FKEKPPPPSPEAGRYDTDQSPAAVAR/9	S435	A8E1Q8|A8E1Q8	0.0	0.0	0.0	0.5	0.0	0.4	2.1	0.9	1.1	3.0	1.5	1.6
M160	SANRDSESSEEDEEEEDEDDTTSK/8, 9	S109, S110	A8E1S8|A8E1S8	NA	NA	NA	NA	NA	NA	0.5	1.1	0.3	1.5	1.6	1.2
M18	SPTNQTSTTTTPSSSPTPK/15	S724	A8E1E9|A8E1E9	0.0	0.0	NA	0.0	0.0	NA	0.5	0.6	NA	1.4	1.6	NA
M18	SPEAQQQSDDDGHGDGATDDRR/8	S820	A8E1E9|A8E1E9	0.0	NA	0.0	2.5	NA	0.9	0.6	0.6	0.8	1.3	1.6	1.6
M18	SPTNQTSTTTTPSSSPTPKR/15	S724	A8E1E9|A8E1E9	NA	NA	NA	NA	NA	NA	0.7	1.0	0.7	1.8	2.6	2.0
M18	SCGSFEGEPSNDQKR/4	S602	A8E1E9|A8E1E9	NA	NA	11	NA	NA	>99	5.2	3.3	2.1	2.2	1.3	0.9
M25	RHSQDDDFTYADPADVR/3	S185	A8E1F6|A8E1F6	NA	0.0	NA	NA	0.0	NA	64	0.7	0.6	>99	1.1	1.2
M25	RHSQDDDFTYADPADVR/3, 10	S185, Y192	A8E1F6|A8E1F6	NA	NA	NA	NA	NA	NA	0.8	0.7	0.6	1.5	1.3	1.4
M25	RSSTPPPSSSPPPPAAGGPK/3, 4, 10	S908, T909, S915	NCBI71|61660143	0.0	NA	NA	0.0	NA	NA	0.4	0.0	0.6	2.0	1.3	1.4
M25	KAQETAPAAAAGSGAR/13	S329	A8E1F6|A8E1F6	NA	NA	NA	NA	NA	NA	0.7	1.6	0.4	1.9	44	1.0
M25	NDSGSVSPVTLFDEDGYAIIPDPPTSR/17	Y240	A8E1F6|A8E1F6	NA	NA	NA	NA	NA	NA	2.4	1.8	1.5	1.2	0.8	0.8
M25	RSSTPPPSSSPPPPAAGGPK/8, 10	S913, S915	NCBI71|61660143	NA	NA	NA	NA	NA	NA	NA	0.6	0.5	NA	1.3	1.1
M25	SSTPPPSSSPPPPAAGGPK/9	S915	NCBI71|61660143	NA	NA	NA	NA	NA	NA	NA	1.6	0.8	NA	79	33
M31	SASFGGLKSPR/9	S157	A8E1E2|A8E1E2	0.0	0.0	0.0	0.7	0.0	0.5	1.0	1.0	1.2	1.3	1.4	1.3
M31	MEAPETERASPPSSR/10	S10	NCBI62|61660153	0.0	>99	0.0	0.8	>99	0.7	1.2	1.0	1.5	0.7	0.8	0.9
M31	STSVTPPPSPK/5	T205	NCBI62|61660153	0.0	NA	NA	0.0	NA	NA	0.5	0.9	NA	3.2	3.2	NA
M31	STSVTPPPSPK/5, 9	T205, S209	NCBI62|61660153	0.0	0.0	0.0	0.1	0.0	0.1	0.4	0.8	0.4	3.1	3.0	2.6
M31	SAPASVIR/1	S160	A8E1E2|A8E1E2	0.0	0.0	NA	0.0	0.0	NA	0.7	0.6	0.7	2.4	2.7	2.4
M31	SVSAIKEPAPAPQIR/3	S87	A8E1E2|A8E1E2	NA	NA	4.1	NA	NA	>99	1.2	1.1	1.5	0.5	0.6	0.6
M31	STSVTPPPSPK/3, 5	S203, T205	NCBI62|61660153	NA	NA	NA	NA	NA	NA	NA	1.0	92	NA	>99	>99
M31	WTSACNSPPSSSGHPVRR/3, 7	S230, S234	NCBI62|61660153	NA	NA	NA	NA	NA	NA	0.3	0.4	0.2	1.9	5.5	2.9
M34	STPTFDSDPFSPGR/7	S650	A8E1F9|A8E1F9	NA	NA	NA	NA	NA	NA	1.9	1.1	2.3	0.7	0.8	1.0
M41	ELGGSTPSLGDNVSSEGVIFVGDHAR/8, 14	S60, S66	NCBI55|61660162	NA	NA	NA	NA	NA	NA	1.9	0.0	43.1	4.7	2.1	62
M42	YETVFLNDTATTR/1	Y23	A8E1G8|A8E1G8	NA	NA	NA	NA	NA	NA	0.6	NA	2.3	0.3	NA	1.0
M45	RPSPSPSTASTAPCDEGSSPR/3	S209	A8E1H1|A8E1H1	NA	NA	NA	NA	NA	NA	0.7	0.4	0.8	1.3	1.3	1.4

				0 hpi			4 hpi			8 hpi			24 hpi		
Entry	Sequence / phosphosite positions	Sites Protein	Entry	1	2	3	1	2	3	1	2	3	1	2	3
M45	RPSPSPSTASTAPCDEGSSPR/3, 8	S209, T214	A8E1H1|A8E1H1	NA	NA	NA	NA	NA	NA	0.7	1.1	1.1	1.3	2.4	1.8
M69	ELATAGDEDEGVRSPGECSVATR/14	S670	A8E1H5|A8E1H5	NA	0.0	0.0	NA	0.0	0.1	0.7	1.1	0.8	1.2	1.3	1.4
M69	RSATTSPEVEPAPPSR/4, 6	T476, S478	A8E1H5|A8E1H5	0.0	0.0	NA	14	1.1	NA	1.3	1.0	1.1	0.6	0.7	0.8
M78	TGPPPSLTEIAMTPVVDSQQK/8	T407	A8E1J7|A8E1J7	NA	NA	NA	NA	NA	NA	NA	0.7	3.2	NA	1.9	>99
M97	SVELTPPRSDGSVGFAPVVVPPAPR/5	T6	A8E1L5|A8E1L5	0.0	0.0	0.0	0.8	0.0	0.6	2.2	1.2	1.2	1.7	0.9	0.9
M97	TAADVVSDSPK/9	S86	A8E1L5|A8E1L5	NA	0.0	NA	NA	0.0	NA	1.9	0.7	22	2.6	3.8	55
M97	KTAADVVSDSPKR/10	S86	A8E1L5|A8E1L5	NA	NA	NA	NA	NA	NA	1.9	0.9	0.9	2.9	1.6	1.2
M97	TAADVVSDSPKR/9	S86	A8E1L5|A8E1L5	NA	NA	NA	NA	NA	NA	1.4	1.2	1.3	2.9	2.2	1.9
M97	RTDDSSTAGEDGYTHCVHSCACTPGER/19, 23	S107, T111	A8E1L5|A8E1L5	NA	NA	NA	NA	NA	NA	1.1	0.7	NA	2.7	2.5	NA
M97	SVELTPPRSDGSVGFAPVVVPPAPR/1, 12	S2, S13	A8E1L5|A8E1L5	NA	NA	NA	NA	NA	NA	NA	0.9	0.1	NA	30	2.5

10.2. Phosphorylation sites of viral proteins

Protein	# Phospho	Sites	Identifier
E1	3	S9, S195, S293	P00001|E1P33
IE3	4	T295, S298, S320, S344	P29832|VIE3
m02/m06	1	S120	A8E1I7|A8E1I7
M18	20	T268, S457, T581, S584, S599, S602, S710, T712, T715, T719, T720, S723, S724, T726, S820, S836, S877, S920, S925, T926	A8E1E9|A8E1E9
M25	15	S185, T191, Y192, S230, Y240, Y270, S278, S329, T373, S490, S907, S908, T909, S913, S915	A8E1F6|A8E1F6
M28	1	S102	A8E1D9|A8E1D9
M31	18	T6, S10, S13, S14, S52, S85, S87, S102, S157, S160, S164, S168, S169, S203, T205, S209, S230, S234	A8E1E2|A8E1E2
M32	3	T457, S477, S701	A8E1E3|A8E1E3
M34	5	T630, S631, S650, S654, T725	A8E1F9|A8E1F9
M38	2	S295, S337	A8E1G3|A8E1G3
M38.5	1	S63	A8E1G4|A8E1G4
M41	2	S60, S66	NCBI55|61660162
M42	1	Y23	A8E1G8|A8E1G8
M44	7	S345, S347, Y380, T384, S388, S390, S393	A2Q6M6|A2Q6M6
M45	8	S171, T197, S209, S211, T214, S224, S225, S718	A8E1H1|A8E1H1
M48.2	7	S7, S10, S18, S19, S23, S30, S34	A8E1C6|A8E1C6
M50	5	S179, S235, S238, S242, S243	A8E1C8|A8E1C8
M53	5	S54, S56, T80, T81, S92	A8E1D1|A8E1D1
M55	3	T157, Y158, S791	A8E1D3|A8E1D3
M57	1	S470	A2Q6N0|A2Q6N0
M69	15	S71, S72, S75, S474, T476, T477, S478, S620, S623, S625, S670, S675, T678, S801, T803	A8E1H5|A8E1H5

Protein	# Phospho	Sites	Identifier
M70	1	S837	A2Q6M9|A2Q6M9
M78	5	S382, S384, T400, S405, T407	A8E1J7|A8E1J7
M80	3	S483, S487, S536	A8E1J9|A8E1J9
M82	4	T425, S430, S431, T448	NCBI37|61660193
M85	1	T3	A8E1K4|A8E1K4
M90	2	S298, S302	A8E1K8|A8E1K8
M95	4	S66, T118, S120, T143	A8E1L3|A8E1L3
M96	1	S124	A8E1L4|A8E1L4
M97	12	S2, T6, S10, S13, S65, S67, S68, S69, S84, S86, S107, T111	A8E1L5|A8E1L5
M104	1	S674	A2Q6M4|A2Q6M4
M105	2	T644, S645	A8E1M2|A8E1M2
M116	4	S299, S300, T313, S320	A8E1M7|A8E1M7
M119.2	3	S102, T109, S111	A8E1N3|A8E1N3
M139	2	S49, S80	A8E1Q4|A8E1Q4
M141	1	S441	A8E1Q6|A8E1Q6
M143	9	S281, S435, S446, Y498, S503, S510, T511, S514, S520	A8E1Q8|A8E1Q8
M160	3	S107, S109, S110	A8E1S8|A8E1S8
M162	3	S124, S131, T138	A8E1T0|A8E1T0
M166	2	Y374, T382	A8E1T4|A8E1T4
M166.5	2	S331, Y422	A8E1T5|A8E1T5

10.3. Viral protein sequences for phosphorylation site mapping

(125 characters per full line)

M18(PG0048_A8E1E9;NCBI77):
MADTGHGTRTRGSRVVRAPVVPVDMPSDGSAAAEPVGGNNSTARSQRQQRQQQQSRTHTQQQQQQQQEQQQVPTVSVVTTATTATVFDSSAATGWIRVSGTPFGSFVPTPPLPQIQSPAPTSWTT
LLTPAPSPTPPHAPVFLGYAPHLGWPHASAFQPLPGPAYATFQPFPWLPAAYPTINPLPLDTSVYDFSQPTLPPPPTPTQFILPFGIVSPFGSGSAVPIATMGAPPTPVPTTLIASSAVPPAPPP
TPAGNVPILPRPERETETPPERPPRCSVRPPGRMTPVAGRSRSPVARGKKKAASNNNLLNGGPVETTTGSRKAAGGVPPRQKAASAASSSSSASSSGPSRGRIISVSDPNLGAVSSSSSQMWATA
AVVQPFGVVDQFAPGQQGQQQHSHHTQVQQQQMHDQQYQQHTRHRVTGSGRSGGSGRAGMSHSSVGGGGGEDGSCDNLFLGSPISSDAKTNVSFSSSPLLSLLLDSVIDSGTFDDSVVGESGTAG
EDATRDGRGGGGGGNVGLGGDDNTNEAPPLLGGVAGEVVGVESEPSGSALRVVGGDDAPSLVPSSGPFEESGGCPRSRSSTIASDSTTTDSESKWALRSCGSFEGEPSNDQKRSGIERASPARVG
FRPIGATPSDASSRETTPERRGRKSRLVDSREPSSSDSSRERPADDTANIPSRGGPGGSSATEMPATRLRRSLPPSSRSDSTSRSPTNQTSTTTTPSSSPTPKRSRGRRTTVTATCDPIVPAPIS
RRRQSLDGRPSAATFPESSAARLGGNPRPVENGVITRDRAKRDGHPLSKPLHTPPRKKKPPRSPEAQQQSDDDGHGDGATDDRRSSSDSGADLPRKTRRDHKISVDEGTQKKHKTRSPSKGNERA
KSPAAMTAEDEPKDTRGPGFVQRNLRRVASRVAREAAKRPLLPPSPLRPSTPQQQLVNPNNSLSRHQTEDTGLEGGSEAVVSAAPDTTSGGEEEEVDEIEAAEALSDALALLKGPGDEDDPLGDA
VVPDGMTGVARVVEVVSDYDVECFAEILNDEIGGPLWWDD
M25(PG0046_A8E1F6;NCBI71):
MSQFVQHVADRGLGVVHQRLARGVEPRADFPRRIHLRIRTLPRDGDVQLYKITRAGVCVPAVLFPVPARPFRPRTPVREPPLTRTMNRRSSKDRRMFVTDDSSDDDDDDVMIMDPPETTSSSSSA
LATAAGIRGVPHAPASNSATAAAASESTYKPLSIPSEELNGEEEERDEEDMSRDGPRRHSQDDDFTYADPADVRLRAMMGNRYGGQSRSAATAGAASRNDSGSVSPVTLFDEDGYAIIPDPPTSR
DDSRHVVVDDDDDGDYDSHYGVMTVAPSPPKLPRKSRPSTKKSAEEKQSSTAGRSRGRSTARRTPKKAQETAPAAAAGSGARQKQRQQQQQQPPRRQSYHPPPDYPPPPPPVQATVSRPLPRTPN
ANDDDDDDDNDEPGPSNTRRGKTPCRRVDHTENNHLYETPISATAMVIDIEDDEDEETGGADDASIVVEDDDEEEENDCEEICDGEEEPAAAAAAAAASSSTPHRTQPLPVPPSSPRITRELGFL
PGVVSGQDARFIAACLHHSHAPQVDIINTCYPMPPYTLDALSEPVLTKKALRCAGVLRPVIKLAILVNYYCVGIGRLARARALSKDLMTPPRIDTLRRRLEGLLPQQTSPSPPMCLRVLGRLNIT
AAQHKASCETIDQLMKPMQERERRRQKTQCAQLFRSKNLLFSPPRFTREGAKTLYMRNIKILNSDEEDTTLNLVMTLNPHPTREDVLNDAIFCLSLGNFVYNFSRALEELRGMIRCQFEDLTETL
YAAYYQCPIMRDDYRVLCSEVANEITSPREDGQGLSALCRRSLAFARRCYNEGVFFSPSYVKYLIKCAAMEEAGFEGYSLESAARSLANPDIFRPLPDESSARRMLRRTIHFVRVDGTPSSSRQI
PTTHIPTHANYELFLQASRMIVPQQQQSRRSSTPPPSSSSPPPPAAGGPKYSKRTFL
M25_b(PG0046_NCBI71):
MSQFVQHVADRGLGVVHQRLARGVEPRADFPRRIHLRIRTLPRDGDVQLYKITRAGVCVPAVLFPVPARPFRPRTPVREPPLTRTMNRRSSKDRRMFVTDDSSDDDDDDVMIMDPPETTSSSSSA
LATAAGIRGVPHAPASNSATAAAASESTYKPLSIPSEELNGEEEERDEEDMSRDGPRRHSQDDDFTYADPADVRLRAMMGNRYGGQSRSAATAGAASRNDSGSVSPVTLFDEDGYAIIPDPPTSR
DDSRHVVVDDDDDGDYDSHYGVMTVAPSPPKLPRKSRPSTKKSAEEKQSSTAGRSRGRSTARRTPKKAQETAPAAAAGSGARQKQRQQQQQQPPRRQSYHPPPDYPPPPPPVQATVSRPLPRTPN
ANDDDDDDDNDEPGPSNTRRGKTPCRRVDHTENNHLYETPISATAMVIDIEDDEDEETGGAADDASIVVEDDDEEEENDCEEICDGEEEPAAAAAAAAASSSTPHRTQPLPVPPSSPRITRELGF
LPGVVSGQDARFIAACLHHSHAPQVDIINTCYPMPPYTLDALSEPVLTKKALRCAGVLRPVIKLAILVNYYCVGIGRLARARALSKDLMTPPRIETLRRRLEGLLPQQTSPSPPMCLRVLGRLNI
TAAQHKASCDTIDQLMKPMQERERRRQKTQCAQLFRSKNLLFSPPRFTREGAKTLYMRNIKILNSDEEDTTLNLVMTLNPHPTREDVLNDAIFCLSLGNFVYNFSRALEELRGMIRCQFEDLTET
LYAAYYQCPIMRDDYRVLCSEVANEITSPREDGQGLSALCRRSLAFARRCYNEGVFFSPSYVKYLIKCAAMEEAGFEGYSLESAARSLANPDIFRPLPDESSARRMLRRTIHFVRVDGTPSSSRQ
IPTTHIPTHANYELFLQASRMIVPQQQQSRRSSTPPPSSSPPPPAAGGPKYSKRTFL
M28(A8E1D9_A8E1D9):
MSLEAVGMRIISDDDLDEVFLRDDEQGDPVADDIPAAAAAVDPAPTPPTPPIPPVAAFSEMVIQPTVPVPTSVASKTPVVPVNVVTGTAQRVDLDSGSRGTSVDNFHARGELLRCARACNGSSFD
EVKAFVKRNTGRRFGLDHMADITLRIGDADNICPTFDVEKLLGYRCSERSSEMTVVGRIVFGARSHGDLIAVYTGGDRCYVHSQDTDLLYIVSERGLVDLLSREGHRHIYEMFDAPPTTGDDVEV

PSCLLPLTELRDLESIECFVRDRTGLSTFRQASNLRFMGEFFGYFMVGDEESLRLGQVVPGRVFFYLRRAGYRVIGRSEMQFIVLCNEKLEVFVLLGGARVLKVANTIAGFLRDRLRINLQPYRR
VFARNGVAERTVCVGQVVNFYCEIDYVVPGGADFAKWLALRAPSGIAETCVPTLG

M31(PG0041_A8E1E2;NCBI61):

MEAPESERASPPSSRRPDADNISRLSTTSSRERSRSRSVSRDHHRDKSRSVSVPKPYKDRYRPPESRPSSSSSHKNAETPRERRSVSAIKEPAPAPQIRSCSVNIGGSSASNSAAVTESGCLKRS
APDAAVSMQAKQRKPDTTTTPEKSASFGGLKSPRSAPASVIRSSPPQRKNPDATHYRATIEDLRRLQNILDPHHAVDERDTTAFSEAHPAGLCGPLAALRQAVDLSVQQPSIIVGAPSSSVALNA
GDFHAQVVMLYGSQATVLDTAAESMRDTCNSMIAYTNSIDASCPDTRVKLDRTERRHVSNYRYRLALSASQYNEVVPSELRARYYKGAIHFATAVIGTMFRTSDELTYTWTTLLVPRVSNIPGGQ
FLLCSLKGNAAVCQPTVTTGGPHTILLAHDPSNVPPDNRPVRLELSAFFFVPYVAPDFGVRMLTLEDAAIRASGISFGIMTPVRRTGGVCIGLFSALSWTPRTIMNENGVSTTSYVAWFECTQEN
SYSLCKYWRPYRKVTYECGDFEAGINVESMYVVDGGRRLIGTLVVGSDTTTNTPSLSPELMPSMIVLRFELCNSRHREPTIVFSTNPTLHLTWPSESPDDPITMYAPYDLCFREGKCAVFMDIRY
AKAADNRCFFISGVPGEYRFHTGLTVWRPQSPIRLSLISPTPSLHICRGTPIARLFFMHSTEGNIYQHNERTATRVIKTGDRINLHLGDLRLPYQNFTTSDEL

M31_b(PG0041_NCBI62):

MEAPETERASPPSSRRPDADNISRLSTTSSRERSRSRSVSRDHHRDKSRSVSVPKPYKDRYRPPESRPSSSSSHKNAETPRERRSVSAIKEPAPAPQIRSCSVNIGGSSVSNSAAVTESGCLKRS
AQTPPSPCRRSSENRTRRPPQRRAPVSAASRAPGRHQLPSSGAAHRRGRTPTRRITVPLSKTCVVSRTSSIRIMRSTSVTPPPSPKLIPQGCADPWRLSAKRWTSACNSPPSSSGHPVRRWRSMP
EISMLRSSCCTAHRPPSWTRPQSLCAIPATA

M31_c(PG0041_A8E1E2;NCBI62):

MEAPESERASPPSSRRPDADNISRLSTTSSRERSRSRSVSRDHHRDKSRSVSVPKPYKDRYRPPESRPSSSSSHKNAETPRERRSVSAIKEPAPAPQIRSCSVNIGGSSASNSAAVTESGCLKRS
APDAAVSMQAKQRKPDTTTTPEKSASFGGLKSPRSAPASVIRSSPPQRKNPDATHYRATIEDLRRLQNILDPHHAVDERDTTAFSEAHPAGLCGPLAALRQAVDLSVQQPSIIVGAPSSSVALNA
GDFHAQVVMLYGSQATVLDTAAESMRDTCNSMIAYTNSIDASCPDTRVKLDRTERRHVSNYRYRLALSASQYNEVVPSELRARYYKGAIHFATAVIGTMFRTSDELTYTWTTLLVPRVSNIPGGQ
FLLCSLKGNAAVCQPTVTTGGPHTILLAHDPSNVPPDNRPVRLELSAFFFVPYVAPDFGVRMLTLEDAAIRASGISFGIMTPVRRTGGVCIGLFSALSWTPRTIMNENGVSTTSYVAWFECTQEN
SYSLCKYWRPYRKVTYECGDFEAGINVESMYVVDGGRRLIGTLVVGSDTTTNTPSLSPELMPSMIVLRFELCNSRHREPTIVFSTNPTLHLTWPSESPDDPITMYAPYDLCFREGKCAVFMDIRY
AKAADNRCFFISGVPGEYRFHTGLTVWRPQSPIRLSLISPTPSLHICRGTPIARLFFMHSTEGNIYQHNERTATRVIKTGDRINLHLGDLRLPYQNFTTSDEL

M31_d(PG0041_A8E1E2):

MEAPESERASPPSSRRPDADNISRLSTTSSRERSRSRSVSRDHHRDKSRSVSVPKPYKDRYRPPESRPSSSSSHKNAETPRERRSVSAIKEPAPAPQIRSCSVNIGGSSASNSAAVTESGCLKRS
APDAAVSMQAKQRKPDTTTTPEKSASFGGLKSPRSAPASVIRSSPPQRKNPDATHYRATIEDLRRLQNILDPHHAVDERDTTAFSEAHPAGLCGPLAALRQAVDLSVQQPSIIVGAPSSSVALNA
GDFHAQVVMLYGSQATVLDTAAESMRDTCNSMIAYTNSIDASCPDTRVKLDRTERRHVSNYRYRLALSASQYNEVVPSELRARYYKGAIHFATAVIGTMFRTSDELTYTWTTLLVPRVSNIPGGQ
FLLCSLKGNAAVCQPTVTTGGPHTILLAHDPSNVPPDNRPVRLELSAFFFVPYVAPDFGVRMLTLEDAAIRASGISFGIMTPVRRTGGVCIGLFSALSWTPRTIMNENGVSTTSYVAWFECTQEN
SYSLCKYWRPYRKVTYECGDFEAGINVESMYVVDGGRRLIGTLVVGSDTTTNTPSLSPELMPSMIVLRFELCNSRHREPTIVFSTNPTLHLTWPSESPDDPITMYAPYDLCFREGKCAVFMDIRY
AKAADNRCFFISGVPGEYRFHTGLTVWRPQSPIRLSLISPTPSLHICRGTPIARLFFMHSTEGNIYQHNERTATRVIKTGDRINLHLGDLRLPYQNFTTSDEL

M32(PG0040_A8E1E3;NCBI60):

MSARGRAAGDDGRQAELMATLGFVRLSKSSVGKVKKFLNNLYDLKSINLCRHPRVIAECRGTDLSRETQLYNEMVLWLRYHEKLTARRPGHLPLLTRIRQDYDKLFGFVAARPELCGFDGLTEVN
VFDDAVYGDDYVPRVDVFLRGLEDLARCLCAQGPDKPARAVIMGFINMRAEEVNRLMDNVRDAAERVLVYEVLDVRDPLNEDPSVLVHNRLVYLCRLAYAISKSWQTLSHMCLDRINSLRRRLIL
AFHDRPAFARVYARNALERPVDGTTAYNLLRRLEEDFLLFRNALRWGDPDWGLESELESEGDNSDAGSDLDLEDEEDDDDDDDGGGPGGHDDESGGNRTPDPGMSLHDDTGIANTCLIGGDDEDC
GGGSGRCLEFDPDSERCFGVKMVNGRALRWWNPTGIMVDNEAAVWIDEHGRVMDKPPPKELRKASSDDGGNKKNPPPKKNVTPPVSGSNSVGGGVQTPSTASGKRTGKKKEGGGGYLLRSRSTDD
DEVRKMKKDGTIDDRADRELKMALQKARESTADSDLSTILPRTEPLRKVAFVGDPVAAFGDTVRTTSSSKGFDDGPFTTQGASVLLPPPLGGPGSTLTLPPDLPDLSGVLADEDVQDSRYGKIPK
SRTKKHPTFPENYKQRPPPHNKDDEYYWDETGDNMPVGEDGGGVLEDLRKGLEGIDLKTGGGGSLQPPLSQQFAGSPFAGSDGDGGGLVKKSSSSH

M34(PG0039_A8E1F9;NCBI59):

METASASRGVGATDRTFGRSGHPRPGADPRSRRPPAAEDTRTDGGNGRHRHRRRKRSRSDNNNNEGVRGAFIKRPRSFSKSSSSSTRDQTTTKTTSSSSGPAGTPHAPRPRHSSESTTYRRPDHP
FDRHHPPHHRHSSSSSSQPHHPPQHVSSAPPIHHLERRRSDPVRAQRRYSDREQQHQQQQQGGGGNGRKESGGGSRHEPPTQPPPSHHRETTMPSSAPSKIVRYEAAHRYTMAEGFIRSVSGTYK
MMNRAEDRKRMLSNQLPSCIKSKSIANSFIFCTSSDGDILESEIASMRRHQKHITDFSKLDRAIHGVSCRIRKHPISPNLSEQQVAMIRAIRIIAISFNRITYVARVKHYCDKDSRFSNYLRDQL
TKRCSEGSRLNLGIRRFISSVNVEKNRDLCMILVGMLCQTPHMWARSIRLLCRLKIFYQNVLIKMFADEKIDLRDVFELQYHSTGYKIQSQIRQYTSSAFVLNDAVGTVVNMIRQRQSSSSGGAM
IPGTVAALPPDPHHHHLHHHSSSSSSHHHHHHPHHHHSPHRFEDDGVVGDEAMVVDQSGFPSFSPPKTPIIREDILDEYIFYPRPGTLSDVSSDTDSPRHHRDTDSDTGRSSVASRSSRTPPFAE
AAEETSDLEEEDGEETAKSTPTFDSDPFSPGRGGVVGGEEEEEEEENNKRASSSKSSSRGGSSTCSSGSSSGGEGDSSDSSTGGSRSHTGSPTLSDRGNTPPLPPPPPLRHRPPPPPVASSPSPP
PPPLHHLPPPIDSIGSHDMMTMGGEEEGMMMIVDDLDAFHHQAQQFHHPYHHYSPSPHELGYFSPNEINGSPPKNFISVHVGRSTPPGAAAAAAAAQKHSECINILL

M38(A8E1G3_A8E1G3):

MAVSTTTIVPNTERMFANAMDRESAIRAAKLVQALETSTYTFVRTLRECTIQKMEITHSSGLVISVCDSSYEHLTIKEPALSWIRRDFCPLSGPLMIIFGVAEEWRFASNRPSRQVIFIIGRDGV
VMAYDRGVMFYICPSLQDFWTADIVFEHDNSIFPSPLRRYVKQLNIDLKDLMGFYNKLRLQRVIIESKEKRQKSGTVLPLKCNRLMRILISAASAIQQGSIPPLFSDRATLHQNVDMGFLEMYWK
SRRYVPQQSAQNGVVRRNGTIILESTPRGSQSIQSTPCIDLAAASPPIARRAIEGPPMPMTVTLISSPPMFRVPVNPVPGGGGVPSSPVPVDLSVRGPAPGVHVLFGPVDEEDENDDNAFNSVRF
AVSPPRDHTEPPCIDLTLDSSEDQQEQPMDLTVNGSPPVALSYNGCIQVASVPRKRPAPEDSRVLEDGSVGAKITKTEECFEQKERGDAQLWGPAVMTTSIVLGAEGAPRVCDRASSDGPDT

M38.5(A8E1G4_A8E1G4):

MESVRRPFFGFVSERPALLFASGALGFVAGWILSNWLTRRRDSQNCSGSSCDRPAGDVTDDGSAVYVGLRRRRFSWGSRERRFAEETANELIRVLRDIRRDCQQQLQQQQEQQQQEKQQQQQPQP
QSRPQSQRQEAGDGGVDNDNRAEHREDVRQRDGSGVRHPGGETGTGFGDEHLYVCEDFEGVYDSEDGDYTF

M41(PG0036_NCBI55):

MGDDDRRGDDGGVYGSGSADGEKIAAPANLSSRPPRRPNLRRLGGLFTMPYRELGGSTPSLGDNVSSEGVIFVGDHARLELFPVDANGVAIPARPPRRVGTVAKLCNGVRKKITPDVVYASVLAV
ICIVILIVVIIDR

M42(A8E1G8_A8E1G8):

MTSSEPEHRLPTYLEAVGENARYETVFLNDTATTRLVVTNLRQQADDDPPPSYESLFGNNDERRARDNSSLRAALRGENSISISVSSSTASTTSSSTEVAARALAADRARTASERLRRTLSLERM
DKEFQWTAITVGVCVSLLVILVLVITAVVTGRGGGGDD

M44(PG0035_A2Q6M6;NCBI54):

MEGGRKVREHEPPTLAFRLKSYKTAIQQLRCVVRSLKENTTVSFLPTPALIVQTVKNQFIAKIVFNSSCLYITDKSFSAKTINNSIPLLGNLMYMTSSRDLTKFTVQDTSDLSAKVCMSAPDYNM
EFSSACVHNQDIIRETGDSAARVDLDSAVVGELIRWIAPNIRPKRNSKKQSTSSSTVQITLHANPPTVKFSLGCNSELEFTASNRIAFHEVKNLRITVQAKNLHQALCNCVVTKLACTLRVMTDH
ETMLYVASKNANFTIENFLSEEPFVRGDVGFDRMPVANSNNYQNSSSSAGDDFAACVDQVIDNCTKKHERVSRKAGGGGGGGVVVNDDHQGGGGSGKDNKYDQHKITSFMVSKGAVGGGAGGGGS
DRGGYFNDTKEESDSEDSVTFEYTPNTKKQKCAA

M44_b(PG0035_A2Q6M6):

MEGGRKVREHEPPTLAFRLKSYKTAIQQLRCVVRSLKENTTVSFLPTPALIVQTVKNQFIAKIVFNSSCLYITDKSFSAKTINNSIPLLGNLMYMTSSRDLTKFTVQDTSDLSAKVCMSAPDYNM
EFSSACVHNQDIIRETGDSAARVDLDSAVVGELIRWIAPNIRPKRNSKKQSTSSSTVQITLHANPPTVKFSLGCNSELEFTASNRIAFHEVKNLRITVQAKNLHQALCNCVVTKLACTLRVMTDH
ETMLYVASKNANFTIENFLSEEPFVRGDVGFDRMPVANSNNYQNSSSSAGDDFAACVDQVIDNCTKKHERVSRKAGGGGGGGVVVNDDHQGGGGSGKDNKYDQHKITSFMVSKGAVGGGAGGGGS
DRGGYFNDTKEESDSEDSVTFEYTPNTKKQKCAA

M44_c(PG0035_NCBI54):

MEGGRKVREHEPPTLAFRLKSYKTAIQQLRCVVRSLKENTTVSFLPTPALIVQTVKNQFIAKIVFNSSCLYITDKSFSAKTINNSIPLLGNLMYMTSSRDLTKFTVQDTSDLSAKVCMSAPDYNM
EFSSACVHNQDIIRETGDSAARVDLDSAVVGELIRWIAPNIRPKRNSKKQSTSSSTVQITLHANPPTVKFSLGCNSELEFTASNRIAFHEVKNLRITVQAKNLHQALCNCVVTKLACTLRVMTDH
ETMLYVASKNANFTIENFLSEEPFVRGDVGFDRMPVANSNNYQNSSSSAGDDFAACVDQVIDNCTKKHERVSRKAGGGGGGGGVVVNDDHQGGGGSGKDNKYDQHKITSFMVSKGAVGGGAGGGGS
GSDRGGYFNDTKEESDSEDSVTFEYTPNTKKQKCAA

M45(PG0002_A8E1H1;NCBI53;Q06A28):

MDRQPKVYSDPDNGFFFLDVPMPDDGQGGQQTATTAAGGAFGVGGGHSVPYVRIMNGVSGIQIGNHNAMSIASCWSPSYTDRRRRSYPKTATNAAADRVAAAVSAANAAVNAAAAAAAAGGGGGA
NLLAAAVTCANQRGCCGGNGGHSLPPTRAPKTNATAAAAPAVAVASNAKSDNNHANAASGAGSAAATPAATTSAAAAVENRRPSPSPSTASTAPCDEGSSPRHHRPSHVSVGTQATPSTPIPIPA
PRCSTGQQQQQPQAKKLKPAKADPLLYAATMPPPASVTTAAAAAVAPESESSPAASAPPAAAAMATGGDDEDQSSFSFVSDDVLGEFEDLRIAGLPVRDEMRPPTPTMTVIPVSRPFRAGRDSGR
DALFDDAVESVRCYCHGILGNSRFCALVNEKCSEPAKERMARIRRYAADVTRCGPLALYTAIVSSANRLIQTDPSCDLDLAECYVETASKRNAVPLSAFYRDCDRLRDAVAAFFKTYGMVVDAMA
QRITERVGPALGRGLYSTVVMMDRCGNSFQGREETPISVFARVAAALAVECEVDGGVSYKILSSKPVDAAQAFDAFLSALCSFAIIPSPRVLAYAGFGGSNPIFDAVSYRAQFYSAESTINGTLH
DICDMVTNGLSVSVSAADLGGDIVASLHILGQQCKALRPYARFKTVLRIYFDIWSVDALKIFSFILDVGREYEGLMAFAVNTPRIFWDRYLDSSGDKMWLMFARREAAALCGLDLKSFRNVYEKM
ERDGRSAITVSPWWAVCQLDACVARGNTAVVFPHNVKSMIPENIGRPAVCGPGVSVVSGGFVGCTPIHELCINLENCVLEGAAVESSVDVVLGLGCRFSFKALESLVRDAVVLGNLLIDMTVRTN
AYGAGKLLTLYRDLHIGVVGFHAVMNRLGQKFADMESYDLNQRIAEFIYYTAVRASVDLCMAGADPFPKFPKSLYAAGRFYPDLFDDDERGPRRMTKEFLEKLREDVVKHGIRNASFITGCSADE
AANLAGTTPGFWPRRDNVFLEQTPLMMTPTKDQMLDECVRSVKIEPHRLHEEDLSCLGENRPVELPVLNSRLRQISKESATVAVRRGRSAPFYDDSDDEDEVACSETGWTVSTDAVIKMCVDRQP
FVDHAQSLPVAIGFGGSSVELARHLRRGNALGLSVGVYKCSMPPSVNYR

M45_b(PG0002_NCBI53;Q06A28):

MDRQPKVYSDPDNGFFFLDVPMPDDGQGGQQTATTAAGGAFGVGGGHSVPYVRIMNGVSGIQIGNHNAMSIASCWSPSYTDRRRRSYPKTATNAAADRVAAAVSAANAAVNAAAAAAAAGGGGGA
NLLAAAVTCANQRGCCGGNGGHSLPPTRMPKTNATAAAAPAVAGASNAKSDNNHANATSGAGSAAATPAATTPAATAVENRRPSPSPSTASTAPCDEGSSPRHHRPSHVSVGTQATPSTPIPIPA
PRCSTGQQQQQPQAKKLKPAKADPLLYAATMPPPASVTTAAAAAVAPESESSPAASAPPAAAAMATGGDDEDQSSFSFVSDDVLGEFEDLRIAGLPVRDEMRPPTPTMTVIPVSRPFRAGRDSGA
TPCLTTPSSPCAATATASSATADSAPSSTRSAPNPPRSAWLASAATPRT

M45_c(PG0002_A8E1H1;Q06A28):

MDRQPKVYSDPDNGFFFLDVPMPDDGQGGQQTATTAAGGAFGVGGGHSVPYVRIMNGVSGIQIGNHNAMSIASCWSPSYTDRRRRSYPKTATNAAADRVAAAVSAANAAVNAAAAAAAAGGGGGA
NLLAAAVTCANQRGCCGGNGGHSLPPTRAPKTNATAAAAPAVAVASNAKSDNNHANAASGAGSAAATPAATTSAAAAVENRRPSPSPSTASTAPCDEGSSPRHHRPSHVSVGTQATPSTPIPIPA
PRCSTGQQQQQPQAKKLKPAKADPLLYAATMPPPASVTTAAAAAVAPESESSPAASAPPAAAAMATGGDDEDQSSFSFVSDDVLGEFEDLRIAGLPVRDEMRPPTPTMTVIPVSRPFRAGRDSGR
DALFDDAVESVRCYCHGILGNSRFCALVNEKCSEPAKERMARIRRYAADVTRCGPLALYTAIVSSANRLIQTDPSCDLDLAECYVETASKRNAVPLSAFYRDCDRLRDAVAAFFKTYGMVVDAMA
QRITERVGPALGRGLYSTVVMMDRCGNSFQGREETPISVFARVAAALAVECEVDGGVSYKILSSKPVDAAQAFDAFLSALCSFAIIPSPRVLAYAGFGGSNPIFDAVSYRAQFYSAESTINGTLH

DICDMVTNGLSVSVSAADLGGDIVASLHILGQQCKALRPYARFKTVLRIYFDIWSVDALKIFSFILDVGREYEGLMAFAVNTPRIFWDRYLDSSGDKMWLMFARREAAALCGLDLKSFRNVYEKM
ERDGRSAITVSPWWAVCQLDACVARGNTAVVFPHNVKSMIPENIGRPAVCGPGVSVVSGGFVGCTPIHELCINLENCVLEGAAVESSVDVVLGLGCRFSFKALESLVRDAVVLGNLLIDMTVRTN
AYGAGKLLTLYRDLHIGVVGFHAVMNRLGQKFADMESYDLNQRIAEFIYYTAVRASVDLCMAGADPFPKFPKSLYAAGRFYPDLFDDDERGPRRMTKEFLEKLREDVVKHGIRNASFITGCSADE
AANLAGTTPGFWPRRDNVFLEQTPLMMTPTKDQMLDECVRSVKIEPHRLHEEDLSCLGENRPVELPVLNSRLRQISKESATVAVRRGRSAPFYDDSDDEDEVACSETGWTVSTDAVIKMCVDRQP
FVDHAQSLPVAIGFGGSSVELARHLRRGNALGLSVGVYKCSMPPSVNYR

M48.2(A8E1C6_A8E1C6):

MSTNVSSAASGGGSSGGSSGASSGGGGGGSGGSSKKEEERRKQFGANVLNLAPAMVAQPVISTMIPKYMKMGGHEDKLAYQLDLLRMLSIAKKATVIQ

M50(PG0033_A8E1C8;NCBI51):

MEIDKNVGADLISNTRRILRLDENELRITDTALICKNPNYSLCDAMLTTDIVYPVEYLLSYWECRSGRTACFVFKNTGCRVSLSCYIGFPERLKDLKRVCDFNFLSVNEALVVTLADIERIKPCD
KGVLTNCVVRKSNSGMSYNIEVVAFGPDNEAEYQALLRDIYARRMTSVPTDCGSLICRRARCLAAAPPRRPPPPPPPGQRWGSLRKHGPVLTRRYAGGGGAAKNQPAAASPTSTSTSSPAAPSRD
QDQTQRPPPAGDTNVTAAETTYSERTISFLTRHANAIHCAPILAAAIALVLLWLLYWHAARSAGHP

M53(PG0032_A8E1D1;NCBI50):

MFRSPEGEERDAADREEEEGGEARRRSRMMMSPRRVKRARHRPAGSGLRTPLRSPSACRCSSPSPERQWQQRRRAEKRSTTPTDPPPPPKRSAASAAAGAAPESEYLNVKLSELHDVFQRHPDLE
QKYLKIMKLPITGKESIRLPFDFKSHRQHTCLDLSPYGNDQVSRSACTTCKETTRLPTASDSMVAFINQTSNVMKHRKFYFGFRKNMELLKMAANQPQLFQIYYIVQSCVQEIVPLIYYDREMAH
MQLIFEKETVHIPSQCIEQILTVAKDAYGVSLDIAHQRITLTARCLRLESSSLRIDVLMLQRKVDELEIPNETNEKFESYSL

M53_b(PG0032_NCBI50):

MFRSPEGEERDAADREEEEGGEARRRSRMMMSPRRVKRARHRPAGSGLRTPLRSPSACRCSSPSPERQWQQRRRAEKRSTTPTDPPPPPKRSAASAAAGAAAPESEYLNVKLSELHDVFQRHPDL
EQKYLKIMKLPITGKESIRLPFDFKSHRQHTCLDLSPYGNDQVSRSACTTCKETTRLPTASDSMVAFINQTSNVMKHRKFYFGFRKNMELLKMAANQPQLFQIYYIVQSCVQEIVPLIYYDREMA
HMQLIFEKETVHIPSQCIEQILTVAKDAYGVSLDIAHQRITLTARCLRLESSSLRIDVLMLQRKVDELEIPNETNEKFESYSL

M55(PG0031_A8E1D3;NCBI49;P27171):

MVVYKENIVAYTFEVITYHKDAIFQRSYADTTTNYFLGTSVTKMAFPIWELDEVNRNNRCYSAASRILNGEVYVAYHEDSYRNYTMVLVEDDYRSKNSKRYVTTKSRYHKGAWTWRYTESCNMNC
VVVVTKARSNTPYEFFVLSSGEVVEISPFYNGENSEPFEEDTRNFWIRKNYTMKTYFGELAAPKKVVPLMAFLEREDMTIGWEIFPKQNVTCDWKKWQTVSRAIRTDTNTSYHFVSKSLTATFVA
SKRKIDYNTTTEGKNYNTFRCVYDEFVEEVNRVFEDEYNETHVKDGELEMYRTTGGLIVLWQGLKAKSLHNLEKFAALNNVSVATASPPVTTAATENGTTVRSRRKRSFDNLDDVVTDISYAQLQ
FTYDVLKDYINDALRNIMDAWCRDQKRTAEMLKELSKINPSNILSAIYERPVTAKLAGDVIAMSECVKVDQSSVKVLKDMRIFQDGKVVNCYSRPLVVFQFINSTKLESGQLGENNEIMLGTFRT
ENCDTNSRKIFVVGTVGYEYRDYRFRNVTSLEHIDLVDTLIGLDIEPLENTDFKVLELYSKGELRASNVFSLDEIMREYNSQKQHIRTLSAKVNDNTPSYLLGLDTFMQGLGVAGKGIGVAIGAV
GGAVSSVVNAVTGFLTNPFGGFTTILLVIGVLAVVYLIFTRQRSAAARPVEYFFPYATQTAVQYAPPGGAHGGLESGPPGAPGLHRRVNAGGSDDSGKAWTSDKKGLERTYTEQDALLILRALKQ
LDDSQRNEKAQQKATRLPTGILDRLKGNDTSGYQRLPAEDSDFEY

M55_b(PG0031_P27171):

MSRRNERGCRSSSWYAMSTALAVTIWCLLACTSEVIAAASTPGTTPKAKTDTSSETASAETETATSGAATGKKEATPTQASKITGTTIVPFVNETEDMVSVDIDKYPYRVCMAVSTDLVRFGKSI
DCINHTPKTPVQEGIMVVYKENIVAYTFEVITYHKDAIFQRSYADTTTNYFLGTSVTKMAFPIWELDEVNRNNRCYSAASRILNGEVYVAYHEDSYRNYTMVLVEDDYRSKNSKRYVTTKSRYHK
GAWTWRYTESCNMNCVVVVTKARSNTPYEFFVLSSGEVVEISPFYNGENSEPFEEDTRNFWIRKNYTMKTYFGELAAPKKVVPLMAFLEREDMTIGWEIFPKQNVTCDWKKWQTVSRAIRTDTNT
GYHFVSKGLTATFVASKRKIDYNTTTEGKNYNTFRCVYDEFVEEVNRVFEDEYNETHVKDGELEMYRTTGGLIVLWQGLKAKSLHNLEKFAALNNVSVATASPPVTTAATENGTTVRSRRKRSFD
NLDDVVTDISYAQLQFTYDVLKDYINDALRNIMDAWCRDQKRTAEMLKELSKINPSNILSAIYERPVTAKLAGDVIAMSECVKVDQSSVKVLKDMRIFQDGKVVNCYSRPLVVFQFINSTKLESG
QLGENNEIMLGTFRTENCDTNSRKIFVVGTVGYEYRDYRFRNVTSLEHIDLVDTLIGLDIEPLENTDFKVLELYSKGELRASNVFSLDEIMREYNSQKQHIRTLSAKVNDNTPSYLLGLDTFMQG
LGVAGKGIGVAIGAVGGAVSSVVNAVTGFLTNPFGGFTTILLVIGVLAVVYLIFTRPVEYFFPYATQTAVQYAPPGGAHGGLESGPPGAPGLHRRVNAGGSDDSGKAWTSDKKGLERTYTEQDAL
LILRALKQLDDSQRNEKAQQKATRLPTGILDRLKGNDTSGYQRLPAEDSDFEY

M57(PG0060_A2Q6N0;A8E1D5;P30672):

MADDDLSSLAPVAPAVWMFFLKKTRELADIVAAMSLCDKATPVVIAPLLIDLTVDRDFCGAVRTPMSTYEGGVLTKVTSFCPFAFFFHNTDEILDVVEDHGDVVHLCDDARRRFGVQAFSPLANR
DRTDVDVLCDELGIAPAEYTGHVVCGNGLKELLYAGQLIPCPEEAVKVQVGAVDGVKVPLYPYTLFSGGADAAHADGASAAVACDDPWVLEHGFYDPALSEALFYFMFTSWGQSLRVCETSRLIE
AGLQQFVEDTQQTVKLTPFKKYHGYTSQKLTAVERDQLMTVDAVCSELAFSYASIYLDSVYEFSTASNFLEWPLVKNAKTHADLLDNLRDFQLHLAKHIAALIFSSNSILYQTRIVFVPSAGKGA
NSNPSAQDSLLKSIRFFNGLTGMYDDILNDAKKTIRFEGAVGRDEKYSPHHLAYFCGTSPQLFSTLMWFFNRMSIYSTGVTSGDTVFSHIVNAGSKLCGACGGRCCHTCYATSFIRVNTRLPGIP
KQIKKEPVVVTLLSRAFADADLLGNYGKRYGLESREAGDGGGGAGGRTDEVAAGPPAGGASGLNFVSVDRMKYLGQVLDYCKKNSLIDAITGEDIINVRSKRDFVATVTALNQTIDDAVCRFAM
DVRRSGHGRDEISGSTQSFNLDLSPYATAFSPVLSFQYYRTMFSIIQNLALINAASYVVDNPLTTAQISKWVTLHFQSICGAFGTTPLKKGFLNVKDTKNLKSVEFERIMDFRSFQETGRYRKIS
TEIKSCKMSVQSLKSCRIKNRPISKTPQSSVFFKKGALQRKNPIKGCLSFLLFRCHEKLFPACGLSCLEFWQRVLQNSLPRSVNVGKVEDFDNLVRFLLTVTDDYDESDVVDIQPDCLLSYVENR
FHNKFLYMFGFRDYMSTIQGMSTRLTPQNHSQFPCLLKDAPKFVSIAEYVLHFKKMKLDGVKAPQVATITREPVLKNVFDGRSLVSVSFAVEKYSSSMGTRDVFQFGQIGYYVGSGVDRSLNTGS
MGTQDYRFMRYRYIIATKLVDVLIRRSRRENVMYDADVVRSRVLAALDSTGLDVDPELAAIAELMEGRDEGDIPEIDDILFYVDQQEYIARSMYRKMRSLAERGVTDFSLASLREATATNATAAG
SAAGGGGSATEGGGGGAAADESGPMYDFSALFSRRDEAEDVNAGLINGDDVRGDDEFELPSKRSRL

M69(PG0062_A8E1H5;Q69154):

MLRTGVKRRLGPFAGYDEDDAATGGVSRRSKYSQQQSQHYYYGHNQSSYRDSGASHPNWKRNAHLMPPPLSSPSSPPPQYDKNIAALTHLNKKLDCLGPDDLECLKAMIRIREARAQGRRPEPSS
APSILESSLVSSNNSNNNTTLSLGGGGGGDYHRQTSPDIRDYTTGSLGLCMFPMDLPDPIKLLENRYTDNDRHAPAVVTHDELINTNYLLLFRKHFDALPPEELRVLVQDRTFAINNAPSLDVVA
AMADENLTYVKFHRVHNLPVNPKDLYMSTLGLIKYATFNKLNLGELSCLLDSPGGGGSDREYHILRQIANKPASPCRKGGSSAAAAASFDVLRRPPLSFKHPLQQALALIASFARIVGVIRRRSL
RHSGPFFIRDFDDTGATDSYRCGMISELIFDYLRGHRCQNEICRVKLKKLLQPYTSTLFFCAYNNTRKHPNGTYRREGRQKRRAPDATPNIPRLAYRRSATTSPEVEPAPPSRMTSSSPRVDSRG
GGGDRRGDSSSTSSNHHRHHTRRARTRSTHDSSSSGSRRRSSATDGRRSRRGSRRGEAQRESNGHHSSKSPSTVSSTTVHGQNGARGDSAPSRKSQQSQQQPETTSKESSKTAAMPPPPSPCSPS
PASRERRPSKSPSSSPRPHDPPSGEPADAEKELATAGDEDEGVRSPGECSVATRRGSSADESSDSSSSSSDSSSSDEEESDVEDCRELDLQSKRLEEALEERCERDFEADDEEFAEPIEEDDLH
CSLDMEEDIEDEPLDPETESVWTASVTPLAAPPSIRILDHEPGDAEEEEESDTDFYDETDQPLNKRIHLRSATPTDDVIMECDLSYSEMDSD

M70(PG0029_A2Q6M9;A8E1H7;NCBI45;Q69153):

MTVVLFATEYDTPNIVVNMLSETPTEHHLFPLMIKYKPSNRIEFVLQTQRCPDSTRVRPVFICDARRLSLSEYVSTNTPLPARVICAGIDADATRELYEHLFDRKKDETGHDEENGSAGGDLFSD
LTSTLKCLVHYNRSAILRYLNNTFLSPTSPSWFLSTYGTHEGTLILTMSYYLFERQYSTIQTTRDYTKCFTADPGRNLFTYINMRDFMATMNGSRFRKQTARFAAFAKARNARDRRELEYVDAKI
NAFREESRLAADSCVYYVYLAYRTALCREKFLQYCEHTAYDKNLPDDQQCAAEENYLGRSLDAELISIMNTYFSVEGYFGSYIHVDRAKLSPPHSYRGYDWNTEADTMVGYSSTATNLAISLRKL
NSTCESLFSPLPPTLMGLLKLCASDRYVPRAEKSRKRTSGGREKEDETRVCRRNYLLNDTSRPIGPMPVFRVEMPEKRHVFCAVSAENWTRRLLPKDLMKNLPSEYVSDECLTDAVWLREDIAAS
CEVGEQLYRTRHEMFNENLPVFNFVGDVDLKLREDLQGLSRQEVFDLCRALRRTLIGAWRHLFPEVDPDSHPVFFFKSACPQNAAGAADEAMLYGGGGYDEDDDPRPEHAAAMVDYGDAVRRPPF
CVCRRKLGLRVIIPFFPPRTAAIGAQTLKRLAGILDHTLCLDRDLVCKLNAISHPGECFDTGIYSHGRSIRMPLMYKLDEASGLILHSRLNPIFIVPAGYRDRPAEFVLQQLCPQNLTHHGRPPRR
DGSADQLTEVVLHITDRACADSDGNFLQSRARRAMSRRRLPLGPLLRAHLSLESGQSAPSLPTLVGRGGGGEGGASSDYEEEERAVGSDEEEDDDDVENLQAFARRIAWPALLRHTRNHYREEVQQ
QLEAATVFTAVGRTCVAVKRGLYGRARDFSCLAREHYTRQETVQVFLDIRGDQRRNVWATLWSRCFTRRCNSNAKQTHLSLKISLPSQ

M78(PG0024_A8E1J7;NCBI40):

MPTSSCAVDYSYPEVDAEHLSFIQTYNILVTSTFAVVGLISIVLILGSLILSMWRKVWHRRNGGARIYVFSLFLAEFMYVASFVAIWLVRVVCPSKVTEPYCRLAIMVNGMSEGAVSFFYLYLTL
DRLCLVGAVTEKKTIFNTSAGKDLVGVIWVTVVAWISAFIVGAPALTSQVMRDRAGFLPVCEMQGPDDMPYVLMHVTAVYVVPFLIIISRLFEMRYSARDGTWHFVSGASIYYVMYMIIVVPIII
CRVVISFGEIGHSLSQTTLQYVDLISTAGWQGRGIAVAFACDVVVKYVEAAGELHASVDGSDDRRSLCADLPSCIYPVGTTLVKSTFGFFLKGAKTMLRLCRRVSPFGIQKDALKSPAAMKALEG
AAYDNKSFSPDEDDDDKIPTVCEKTGPPPSLTEIAMTPVVDSQQKMTSLPMSVFYSDGGEKEGVQGDEGGHIIPIPPSDGVTDDVSALLPTSSSVV

M80(PG0022_A8E1J9;NCBI38):

MTGDAARAPDAGSMIYVGGFLTLYDEDPQDERLRLPRDVVARELRRAAAGGPAVPLNININHDESSTVGTVRLFDAEAGLFCLGRLSSPAFLGIVEKAAGKSKLVARGPAKGLEADPVVEYLSAGFP
ALSLSSFSPDAVAAAAADADTSENSGEEAEGQPRRQTTDSGGFFRHVSLCGLGRRRGTLAVYGRDRDWIVGRFAALTPDERAEIARVDDAALGGWDGDADPFGSDSYGLLASTVDDGYIAERLCR
LRYDKRLLGLQSKETYVKASELPAEADDGPEPQSIRRDASRDSEERAAGREEMAQSSHGLTPAAVSVPGTANPAASAFPADCVYLSRDALMSILAAAAKQNAVPGALTSPQQALPQVPSYYGMPP
DGVQYHLPPPPPPPPSHHRGGGGTFDPPLPHGGYGPPYHHPDAYRGGYHHPDRDPRGGVPYEGWYRPRYDPAGDDHPSYNNRRGDRYRADRPPQQQPLYRGERNRRRSPPDSDDDDDDDEDLEA
GERTGGKRTRQRGSADSGRKRRRRGAAPDDDGGDLSLPGERGYPKRTAGDHHQSAPPASRTDEFGEVRATLNEIRKDISQIRAAARAEGNGAREDAASVGSSDQKCAAPPPGVTEMMASEPPAGG
TVVARMALDPAVAAATGHTAGLLTAGKLVNASCEPTPMEVGEPSGGGTSRKGGEASMLEVNKRMFVSLLNKME

M82(PG0021_A8E1K0;NCBI37):

MAEEFNFLTADPDKPGSTFVHGWCTNLSLRVTKGPLFRGQETKYLNTGLSLATSVPAITCLAQRLPQGRTPSALSVPFTVLSDPEDFINVVLPVRNMSTFTRNSAETQDQIFVTAFAIPLPIVPI
NEYPLFRQHPNFKTDFEQEAHQTTAEELNDSHLIKITHRRMRWVHSKRIQFGTHCTTITATLPRSVPANTYVAAQPKAISDPMVRLDSSILIPRPSETAHHKLRIQLVFDGHAENNGKKVNKPPT
SLTVHILLHKDAITFALRHSPYKSLQRFDAFDNGYKVFCPRKIESRKGLTWTITIDNGYSCDGRYTALFFPKYSPQLDIHVCRWIETRTLNICITAMANVTIPVGMELGRVHFFPNRLTQELVDL
LTPVGIINKESMALSTSQYECKLFVPVGSETYRPGPEPPQDADQLPPATTPRRMSSADEQQLPEDQFAQLSLNTPRPVGRELPPRRTSADERLFAQMDADNVPSDQEEDDSDEGEVFIDDQDFDE
DRNWVYNAVPDSYEEEAAAQRQLLTLQPHDLIRIRAGRLATLTFTMYTPFINEGVYFNEAMGVDKVAVQLTTKGRYGARPMPSSLPRLKTAIIPPHLQP

M82_b(PG0021_NCBI37):MAEEFNFLTADPDKPGSTFVHGWCTNLSLRVTKGPLFRGQETKYLNTGLSLATSVPAITCLAQRLPQGRTPSALSVPFTVLSDPEDFINVVLPVRNMSTFTRNS
AETQDQIFVTAFAIPLPIVPINEYPLFRQHPNFKTDFEQEAHQTTAEELNDSHLIKITHRRMRWVHSKRIQFGTHCTTITATLPRSVPANTYVAAQPKAISDPMVRLDSSILIPRPSETAHHKLR
IQLVFDGHAENNGKKVNKPPTSLTVHILLHKDAITFALRHSPYKSLQRFDAFDNGYKVFCPRKIESRKGLTWTITIDNGYSCDGRYTALFFPKYSPQLDIHVCRWIETRTLNICITAMANVTIPV
GMELGRVHFFPNRLTQELVDLLTPVGIINKESMALSTSQYECKLFVPVGSETYRPGPEPPQDADQLPPATTPRRMSSADEQLPEDQFAQLSLNTPRPVGRELPPRRTSADERLFAQMDADNVPSD
QEEDDSDEGEVFIDDQDFDEDRNWVYNAVPDSYEEEAAAQRQLLTLQPHDLIRIRAGRLATLTFTMYTPFINEGVYFNEAMGVDKVAVQLTTKGRYGARPMPSSLPRLKTAIIPPHLQP

M85(A8E1K4_A8E1K4):

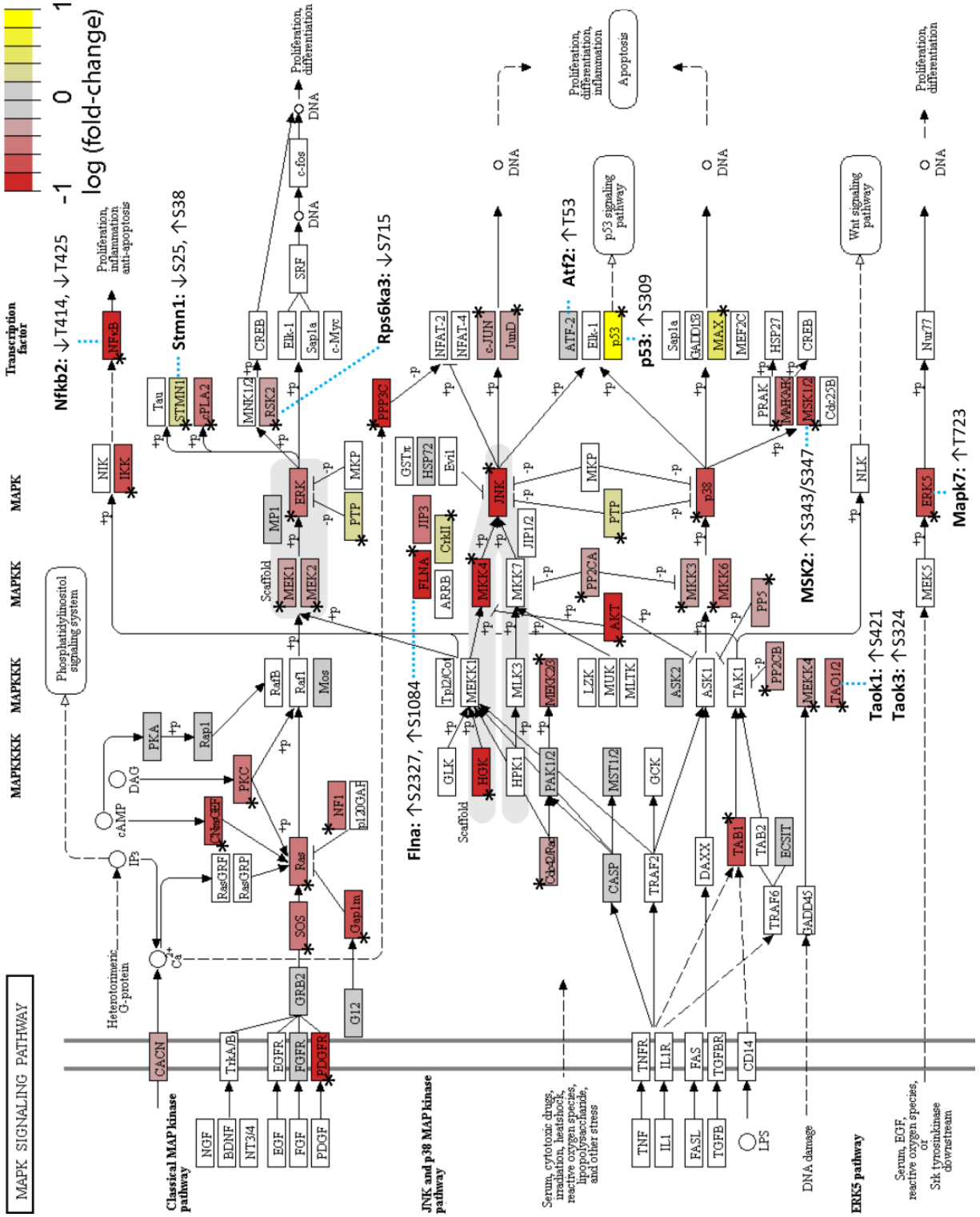
METTVLVTFEQRLTTGDVGKLSRLIGAVIPIPYRHHLLGSSQVGLDAVVKDKTRDYSRMRARMREMTLTIMRRVEGNQMILGVPTHGQCYTIRNTGPVSWEKGDVLTTLPPVFSGEVTGLVSVSD
WDLVLPWIVPMALATEINQRMMMLALLSLDRSHEEVRAATAQLRVVRYRDATLTLPEITIDDTVLIDMRNVCISLSMIANLSSEVTLAYVRKLALEDSNMLLMKCQEILGRRMPQVGVGAGSSGD
RNDPPARSRTNYNITPTEELNKLTALFVMIRQITDVISEQPAFLVCDVSPDDKSALCIYKG

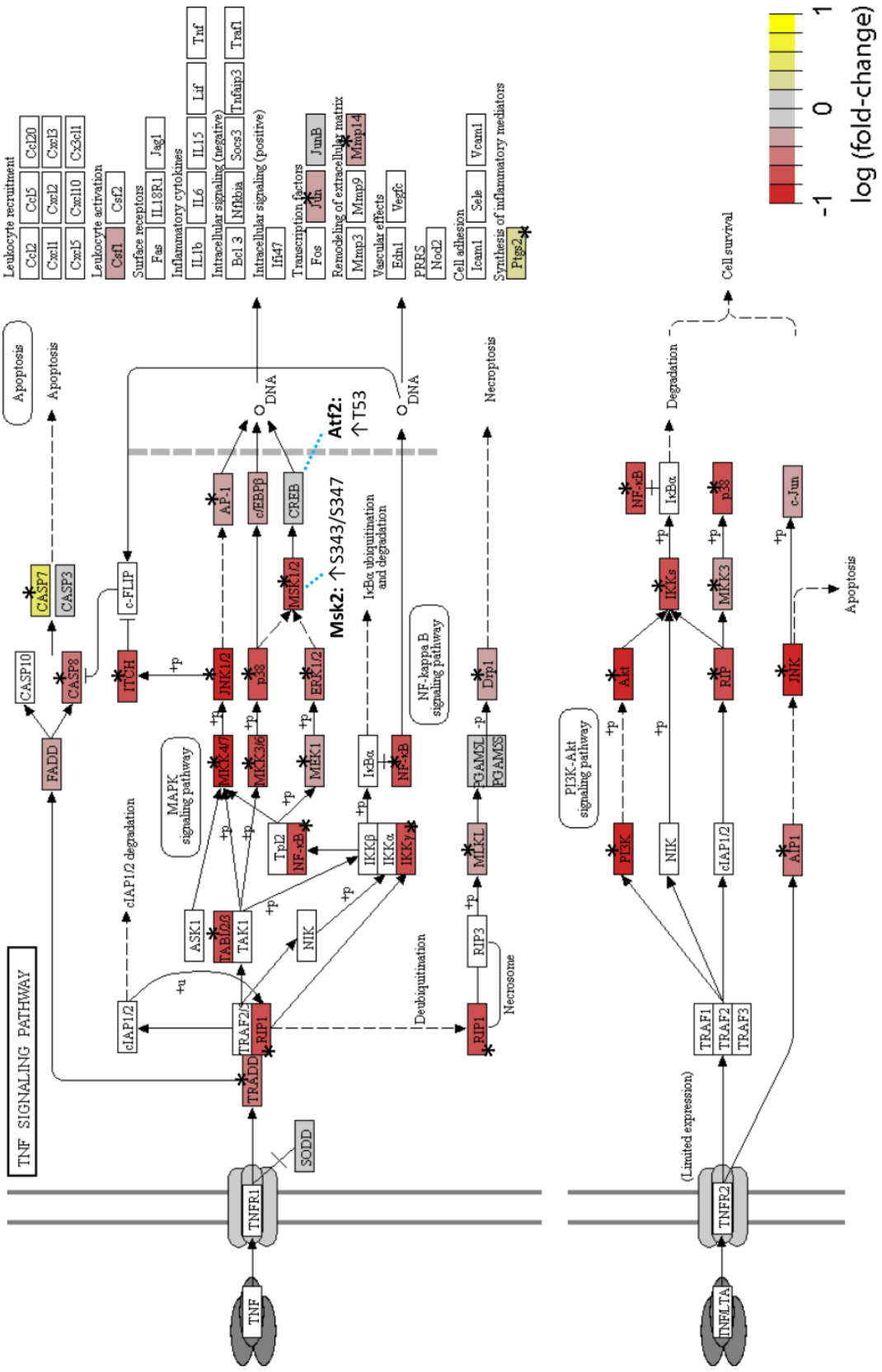
M90(A8E1K8_A8E1K8):

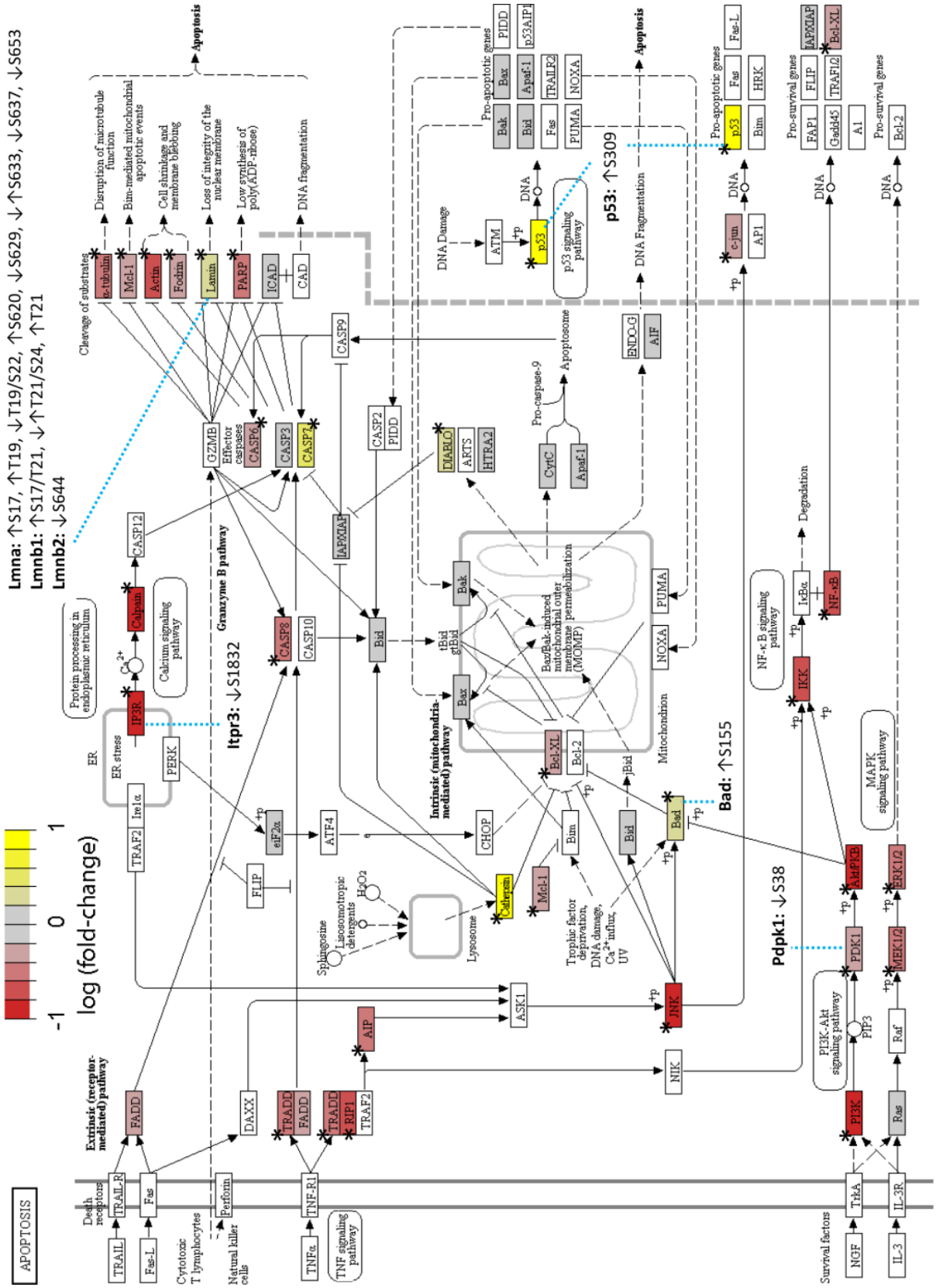
MDSARRNTSSRVSCSMPAFLSSPRITFMLLGVLLCFALASGGTADLYTWIQATPSNITSPSDVFTMNVTISATDSDRSGRINASTFAYVTEAEQSQNEADAPEYDTASKQIHTTFAPVLFNWFRV
TGKNNTARFRTVLFILTCNETHSSLLMNSPPRFTAMRIDLDNTNATQPLYLQMFSKRQCMQKIEPIITKRSHVGIRRILSKTCLTYRRAADPEVQLAEGTVGREPQEAYTFLHVLIAFLFCIGTG
LIVCLYAILRPYYQKKSANAFPKKKKNRDDKATLWPRKPIFTEEYYDSPPGSIKKPYAVVNPGFMYKD
M95(A8E1L3_A8E1L3):
MATAAGGVYAGGEDQQQQRPRATDVSTALCDVEALAAVEEGRVSEADVNRYREAVDAALIACEASSPRDRFRLIETAGGNFLLVTNALPKDRTEQQPPCVLEGSGRASSNNNYEGIGTPSAGSGN
AFDGLLALERGTSGGGLTATVPSAPGYVAKSVNTLSYDGRLLSNSYVLYTKEQLRKSLSPDKRAIVERILRFVDTPGILDHNNVQDVEAVLWLLFCGPQSVCQNPTCFGRDRECEVSYPVLLPPV
FYDPITDYSAYINLAELYVYVWYRNYDFDSEPTRCYELGTVAMDRVKKTLQSVRQRFSDRSVPVWPVSSRTCVFCALYNQNRVCLDLAKSDVDVTSYSPIIIKDCRDAATNVTLSHVLPGQRVAS
LFPVYDIGTLLRALCDSNDGEERRKRMRETIDSALSTTDDAV
M96(A8E1L4_A8E1L4):
MGSRRPARSESALLKDAMRADLEKRQKRFFVEAFGPDHPLVRIQGIRTSDAKTRASIRECAQSTREVAAAIAVGKAQMRSGGVSTTGKDSSSWDRVHRDADEFADALAEMRDVAEELRSVFTSSD
SEEV
M97(PG0016_A8E1L5;NCBI32):
MSVELTPPRSDGSVGFAPVVVPPAPRKPLRRRAVSDLEKLYKVKRRLVFGADDGAVDNDTSNNNSGSSSTTSRSRRKTAADVVSDSPKRTDDSSTAGEDGYTHCVHSCACTPGERHLLCCELVSI
GDSVSVARCPLCSLGISTTYLSRGCCRGRSKVTGGDEDEEDEDEEENSQDEDRDEEEAASASSSGGLEWSDDSNSALSWSDENIIISPFPGLKCYVTTFEDIRQPVLLETGSAYLPVYVPYDESF
CRNRCLERGGDDDDERDATLIGKGSFGQVWRLSDKKTALKAAASESINETLLTVWISGVVRSRAQDAGYRGELDDSVYCNILVATGSCLRHNLVSFASFDRDLYNYRGWHYAGLASYRRAFSGIA
DALRFLNLRCGVGHFDVTPMNVLINYDRADDRQIARAVICDFSLSQCHTEGTTGHCVVVFQQTKTVRALPKSAYYLTDIYHPAFKPLMLQKLCAIEPRKQFPKPSANRFCVSDLCALGHVAAFCL
VRVLDERGQLKVRSTSEDALFGVARKTCDALARHSVDEVANFCSLLITRQLAYTATLLGSDDMREPMARLCDYFETVSDKDAPDRFRSVYKRARREIDGSYMVRLLLAASETEDGRYLLDNIRAT
CLMVDSEDLDVDPYKIFP
M104(A2Q6M4_A2Q6M4):
MWRNQSLYRDSRENRFKASDLTRSTIRSIFEADDIFRTKMLSYLDNPPEPPSDPLFPTDSSLDLFSMINGTEGACIGQTIHQILRDPSVFRKQIFYAMMRFLLNGISVGELSTAWASHRRRFVRA
EDEGGAALEQFEIWADALKHTIVDSIAALLEKLIYTYAADDRYCRYVDWIVSVGVVPIAEVRTAEREKAVDAAQRRFLAEVAECHLLDRPDPLRARTLRACVSALMTREVPNVPDIRIHRLKSNG
HIECFSGKRRLKRFIYAEPTILEEERLILTTPLARIRYERKRHNELRIHKKICQLLNTNPIKVVTTSRHEMNTKRIVELMEKRDRQVDAKTSIVKFLLNVSDSKSKIGLEDSVESFLQDLTPSVD
QARLLPSRAPLIQPAPSGSGAQDIRELFRRQVIRCLEDQIQDHVEEIENLKLLNKTWESKTRELRDALDRYESEGRRGRGPPAFDLQTLDTVNALRRVQGLPTAPVTVDDNRVVCNSFFSQFVPD
ERESDERLSRLWEQEYFRCFKFRRNVTNQGAEDSISYSNYTIERVLLPFLTAVIEFPMLDAIPEEYLFLSLSELANVIYETSKLQRYTDYIRYRETIRVQAFLEREQATAAAAAAAGAATAAAPS
ERIGRAPGQVSGPPTKIRRLDETTPGTANYRPQQKTIVTTTPIGLPPPSSPAPEVSPRFRSPQQKLETLRDRNVQHLNG
M105(A8E1M2_A8E1M2):
MEKRSSDESVGNKGSDGGSGGLSRYDNIFVLNMSSASKIERIVDRVKSLALKRFSRESLYKDWFRHMLDPCAGLVAPELGDDGSSEGGGNAAMIVGDRELARRPPFLPFSCLLITGTAGAGKTSS
VQVLAANLDCVITGSTVISSQALSSALNRSRSAQIKTIFRTFGFNSRHVALADRVHLRRRDDVAFDGDVDPICQQQWRDLSTYWPVVSDIAIRALDGGKGRKDTDDLCRSNIIVIDECGVILRHM
LHVVVFFYYFYNALNDSELYRQRAAPCIVCVGSPTQSEALESRYDHRTQNRDVQRGMDVLSALISDPVLSEYCDVAHNWVMFINNKRCLDLEFGDLLKHIEFGLPLKSEHVEYLDRFVRPAGLIR
DPAHAIDVTRLFISHAEVKRYFTALHDRVRIYSQHLIFEVPVYCVLNNSAFHEYCASMCTGEPTPRPETWFRKNLARISNYSQFTDHNLSEDIQVEELAQSCGGGGAGGDDDGFDLEEEMINETL
LTCRITFIRDSAVGVTAKTKACVVGYTGTFDDFAEILQKDLFIERTPCEQAVYAYSLISGLLFSAMYLFYSSPLTTPEILRDLSEIPLPDIPTLVIGANGGDGARDSDDNDEYEEDLEGGGCFDG
VSGGGSGNGGGEKYRRRLTSDDEDDFYDLSYVDRGRQPEPPQLQPPPQPQPQPRLTMSAALPPRQIDEEISDVEMLCYSDIYTDKFFLKYSIPPPVSSISFEEIVHIYTIFRDIFLARYRIMQKH
TKGAFGKTRLVTYNRRNVWRRKNCEIESQTGSFVGMLTFVSPSNNYVLEGFTNHNVFIMDAERNRIHRRILEKGLPRLIVRDACGFLLILDYNVSKFSDVIDGKSVHICTMVDYGVTSRMAMTIA
KSQGIGLESVAIDFGDNPKNLKMSQIYVGISRVVDPDRLVLNTNPVRNTYENNTFITPFIRRALQNKDTTLVF
M116(PG0015_A8E1M7;NCBI29):
MFRVGALLCVAYAVTAVYITPDVKHLQTHSPGGGRAQNSVTSKQPSIPKQPKLDRGLREVQSDPGTSQPAANANSSDAAPEKSTMSSPNPAQTAQAGKNGNETTVAAPAPPAAPAPAAASASAAS
TRAETPAVAAAAAATTAAKKRVSTILSRASDGTKILTPTDGTKILIYASSSSDSSDISSDDSDSETNSSSSEPSNSSLPPGSETEPSSSSSYSSSSSSSSSPSSSDSEDEPSSSSSSSTSSDGA
SSDSEAPSSDSGSESIGSTSPKSESESSDSEKDQEKNKKTSTPASTRASSPQPGAAPAAPANTEPARDASTELSSKMLSASIVDKPTKLADAAKTTPQQVTNATGGDDALVVLTVGSITNSPLSD
NRSTQSATTVQTESGHAGLQKETPKKSTSLSTLKRAGRQTRELQQSSETATITDASRVPSTVTAVPPTSIPYLTTLTTAPPPIPSTSTSDVDNNTNTIITAITTKDPATTAVPSRASLTVTPLLS
SSSSTISFSSFPSKPQTQLSVTQNATNKDNEWINSTESVSSPSAAFRTQVVTTTTTTNPTITTFPGDLCTRRYCPAKYGPFVLSAFTRVTFSRAECASDVAIFNDSLIIDPGLGASVFLTIYKRL
SLSNRMIRLASELTEYTYTPADFI
M119.2(A8E1N3_A8E1N3):
MESSLKSNDTIRTPQPSSIAIFNLVMELTSTDLLTYGLISCILLALGFLSGILMAVAAMRAPHPNLFLHKLKKLLGLTSYQLYTICDHAPQTPEEKKKNKKSNDHGDGTESAEDDFMYIPTK
M139(A8E1Q4_A8E1Q4):
MWSLRGAQRRDLRFDGQDDEWCLAYEDCMDQKFGEYGGVDERSLPPIRSPVYPPRSFINDQFSFGLPLDSTDRTAGGRESCGAMAESDLFIDRAALDGTGRSDAGSRDSESDSDFDMESDSDLSD
DGDAAGAYVYGSPWKTSAVTVSRKAVEKYEKKKLAQRLYSLTMRDFKSFHLALSFPPGQIAAAVVHYLGKPLVLRYPENWYLIPQRRSEIRVLRGVDLSKYLCCGDTAGELTPLGVCAARSPDLH
FRETPCALVMDERGRFFLYDAESDGLYYAARNIDQLARRGLSLCEPVYRDGGAVVSMPKPKTLVRKIVSAAVVGLENVAATATAFRGSTIALRDPVSGRRETFQVFGASDLKRKPPFSRMDDTTY
TLVREYITFRLAEAWTVIGAVGEYRDDGFVFEVSTVVLVGARGTVYGFCLLSNDVFRIAEDISVFFKRGGVSGGAAVPNRFDRGARGELRLERPPLCPHRDEDRPPIATSLASQEIARRDLEDWY
RWRLGARSRLQDGVALSNVSEAKRQLTHPAKGIPVSIVTAESYHPRGDPEEAFESLSGQTRPYRYPRVCDASLSCSDRDFYRRYVRAMDMFDGGKEKSEREIMFIRAERVGSLQKTGTPLKLPPL
VRATQAGSTSNNMETGSDR
M141(PG0009_A8E1Q6;NCBI15):
MERPVTANAISAYKRVIDAFSRLVATGGDQSARRRAIERLDGKVIPLGFPQDWYVVCGSGKSCWSVLTDRANALLCCEEALIPLGLCGRTSDPYDTRGLRGYGVSLVLSGEYGRVYVYTASDDTL
HLVARNLEELGRFGISRRSFAYKEDMQRSVPRADLVRLRYIASCDVPRGYPGMELLVVFSDPNADAATSVGYTRFVDGKFVRLTTPGYGTSALFLAKDPGELRGIWPFRGMCDPEFKRCWNAVSD
ALHCSWSLIGVLGYRPGGLPDAEFRADHVIVADKFGAVYSLLVYSRAGLTRRVADSVAEFFQMGMLKRVFSSSFEFRQRERRKERLESMPRCPHWPEQKFKYIEENCRDRFDSPTDDEDYVPADC
MAEHYAWLVQSTVKVHSCRPWDLDDPNVAPRALVTNADTGDPMKTYWDRLESAIAKVKLTAAATASDVESSGGRYGEGPYVYKLSPVPPAYFALTCHETESMTMEEVYACRRRKFFCDTIVPFYR
ARPGPSET
M143(PG0007_A8E1Q8):
MSWVTGDPAWRPEALIRPADREQILRDFKSFFLTQTSERSLKKTVKRESGTKLCVGYPPGWWLAVVPRGELVEMDGRDLRPILPTGEWVVVLGEIRSPMVQHTSLYLAMGKESRMFVYSAEEDAI
ALAANDLDEFSRIGLLTVEFIYRVPLSIPLKKDPVWDAIGLCVTGYDLNRHLSTYRNSVIELRTPGKQETNPLILLDRLELCCKYWPLVAMDKSRIDGLIRYASRKMSSRWYTLGLVGVYRDTGV
FHGAQLLVFDDSGSIFYLSLISGELWRLADSAAELRRMGLLKIFTAGRRVDRDLAGSVRLEPPPDQELWFHSRPAATHLCVDVAPPGEAQLAQQFAWISRPGRAEILSEEETAVADAFLRLSRLK
VTVTEENSSRVLCTHVAAPPEPRTRRLVSALLRDEDCWRPESADKADIVPRFKEKPPPPSPEAGRYDTDQSPAAVARRRAHAASSQTGVATRHHAPILEPPRSLRSEAEGDCHECLRERQYRYTT
DDSPPPAPPSTPNSTGSSPSPRPIPAPRRRRPPPPRDERATDA
M143_b(PG0007_A8E1Q8;NCBI13):
MSWVTGDPAWRPEALIRPADREQILRDFKSFFLTQTSERSLKKTVKRESGTKLCVGYPPGWWLAVVPRGELVEMDGRDLRPILPTGEWVVVLGEIRSPMVQHTSLYLAMGKESRMFVYSAEEDAI
ALAANDLDEFSRIGLLTVEFIYRVPLSIPLKKDPVWDAIGLCVTGYDLNRHLSTYRNSVIELRTPGKQETNPLILLDRLELCCKYWPLVAMDKSRIDGLIRYASRKMSSRWYTLGLVGVYRDTGV
FHGAQLLVFDDSGSIFYLSLISGELWRLADSAAELRRMGLLKIFTAGRRVDRDLAGSVRLEPPPDQELWFHSRPAATHLCVDVAPPGEAQLAQQFAWISRPGRAEILSEEETAVADAFLRLSRLK
VTVTEENSSRVLCTHVAAPPEPRTRRLVSALLRDEDCWRPESADKADIVPRFKEKPPPPSPEAGRYDTDQSPAAVARRRAHAASSQTGVATRHHAPILEPPRSLRSEAEGDCHECLRERQYRYTT
DDSPPPAPPSTPNSTGSSPSPRPIPAPRRRRPPPPRDERATDA
M160(A8E1S8_A8E1S8):
MRLITKLCLLLGYLIQINNADYQVHLTAKDPDVNSLEYTCTSSGSDGFENVQGFWKILCTGNKTETILATFGGYSGGFQRKHPYVLSGSSTTTMDIHLPSRSANRDSESSEEDEEEEDEDDTTSK
SGPSSTLKMKPLCDGLISCKLGYNSTDVTGRSSALPVLGPLITTHTRNRTRTGQTLGVQMSCRPQRMCGTYNVTWYHSSNTITTIGRATYWNESYSDGCNTTWGSEKLKWDHDGSVQVNRSLWTD
EFEYAYCVSCKVETCGAAVWSMVCDSGYTAYYSSGHRIQPSSWQQMLLVFLGVFLITR
M162(A8E1T0_A8E1T0):
MYLYNNTTYVPADNYSDHTVLATTSQILTSTVVASSALGGPVIENVTTGGRGGGGRVHILAYPSNGLGGPPITTTNGLGSVIYTSGSYTALAVALSLVILCLAASLLFALIYAHGTFDLRRSQSE
TLNRDSGEFDPGTPLQLRAVLRTIPESEMVESNE
M166(PG0003_A8E1T4;NCBI05):
MALPAIPQQLSPVARLRAASFSVLLCLWCLAWLVSRPALVDCSKPKKTTTTAAALGSVTQEERPRPEPIKIFLYDTQDQNVIQFSWSGRHHDVRSGFNVSYRNHDDQTVYLGQLDWRGFVRYRDN
SDLYVGDALIFHGPPLTDDTVDVYASGWLKLPPQTVGEIVIKGYGQTAVVTVLPPAQVALLQSTDRDTLAPFLFRCLANIRGSTPLQPTLMRWWHGPYPLVSIEWSENNPVEGRVRWEQLPDGDR
TYYVNAYTGDIHVFARSTIQRLACLQCTMQADKVATSSRLSCPARAKTIALPFDFGQDTTVTPTERRHMDILHTVLIVITWFVAGLGILALGIVLIGNALALLRSCCFPSRSPGAGGKKGGPAYE
ILVNEETA
M166.5(A8E1T5_A8E1T5):
MDHPAQRRRGLRLGERDLERIAAALAEGEPGLDRRGLSWRRPPRFFVPTSALSEWLEARCRQQKKWWKQESTDSDPEASEEDEGVLGAFVDRVDRKFRGQPGVCRVQLVFRGLRGARALVKLGRS
VARKLGCAMPASTRRPPPAFESFGSVRLGLVVKRVSGGYCIVSARVLGTSDPVDLQNGDGDLGTWFKQLLVRASPAAGDDGKRQQQNNHRPSDPKTKSGPPPVVRDRNGVRSDGASGPGLGRGDP
LRFMHGHSTETAGGNHDGVSFSSSFQETPRRFQTVDPKRLRVERDGGGGSRSAVRSGGHGGPDADPRASSSSRVGQVPQQSPLRASSCACGRGCLDRRSLCQGASPSGDLEKSEVWKFLASIRHD
PSRGSGTISASAPAAGVQRQRTPAADTNDPGWRFQQFLNSYRQSPGYVRPASLSSSASGVSRGSSAGPRS
m02/m06(PG0054_A8E1I7;NCBI87):
MPSWSDTLTMDTTARGSTRSRPVQLLALVTVLASMTFQMGESLIPIIPDFSSFMMNPLPMPQIMPPSTNETKETKESYVKTEEPIVGCNVSRTEINRLKNQMKKIPNTFKCFKKDGVRTSLDMQT
TGEKRFACEIPNNVYVNATWYVHWVVGKIAASVSPIVYFTSTTSSPPTLDGNMHPFYRRKIVTAANGFKVDEKTGDITVARSNASLADSVRCRLIVCLWTKNDSISDLPDDDPQMKNMSGVIKLP
DYSGPDTLLTVPFDYAAWRQRMRTEMEEPSRRRRQLLLVISVIASLLWLAVGAMLFYTYGREPLARLLKRYGKRLAAVRIPADGKDQSLTSPLLTK
E1(PG0064_P00001;P00002;P00003;P00004):
MAAPDRRGSPIVSRRRYFTFRNSSRTLHQNVTRMFDLRQWTYESARVLDCTNGEGRTENWGPGWLCATIMQSPDSGSSSGSAAQGCMSLDITSDDGPEKVMNMFHRGSIVCNKTVSSVATPPGSD
GTGASLLTLVADGSLLQVILVEHSPGAHRAEADDVAESVGNAAVAAAVQSATSMGSSSSVGGAGGGGFNSSEERRRDKDSSKSHDDDRRKRFELSNGTLHRERTSGSSGSSGGGGGSGSGGSSGL
STKQKERSRRLEEECK

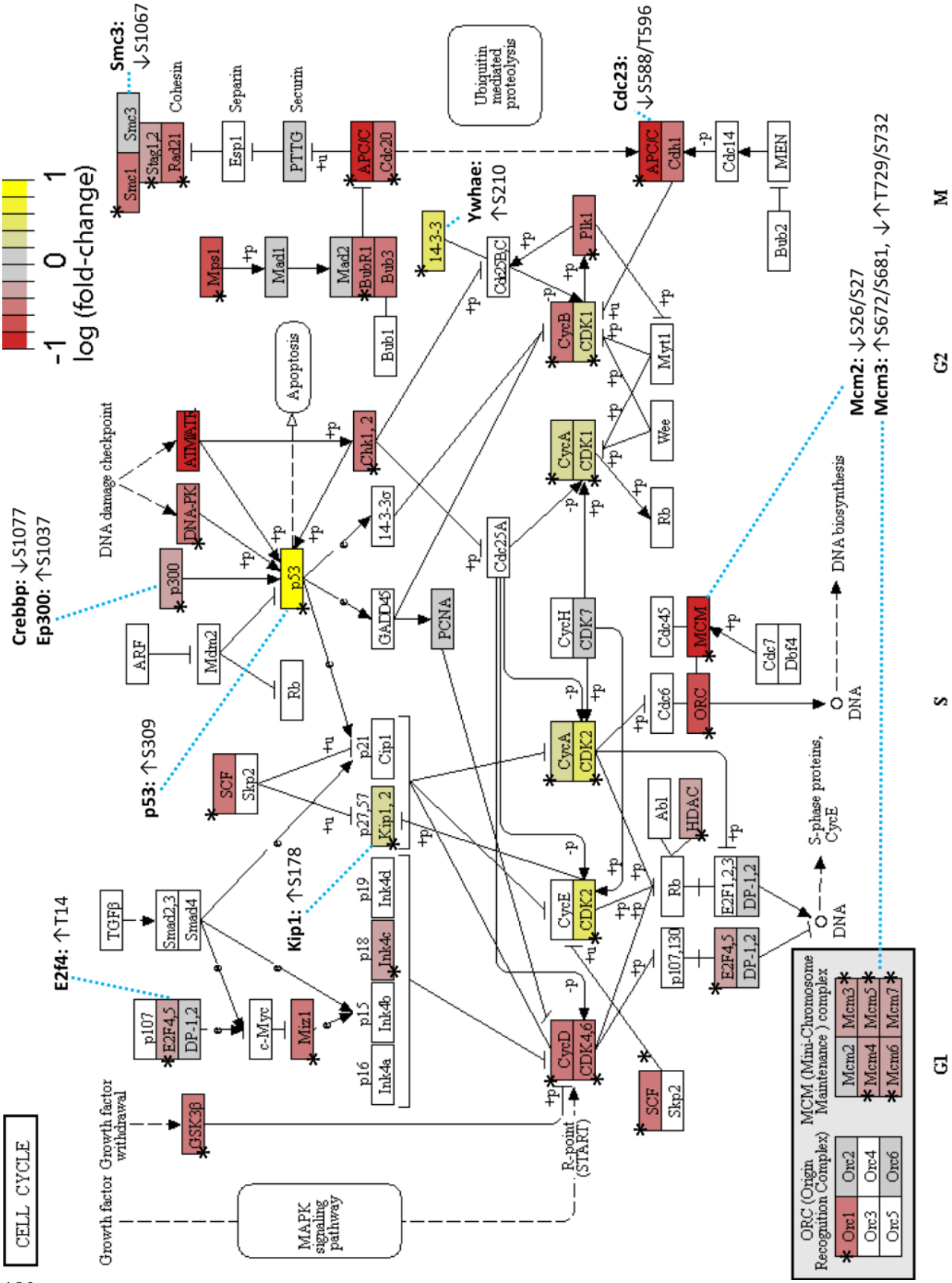
E1_b(PG0064_P00002):
MAAPDRRGSPIVSRRRYFTFRNSSRTLHQNVTRMFDLRQWTYESARVLDCTNGEGRTENWGPGWLCATIMQSPDSGSSSGSAAQGCMSLDITSDDGPEKVMNMFHRGSIVCNKTVSSVATPPGSD
GTGASLLTLVADGSLLQVILVEHSPGAHRAEADDVAESVGNAAVAAAVQSATSMGSSSSVGGAGGGGFNSSEERRRDKDSSKSHDDDRRKRFELSNGTLHRERTSGSSGSSGGGGGSGGGSSGL
STKQKERSRRLEEECSPRSGGEPKRQKTHHDPRPERDLEPPRSSTTVDGGQQPAF
IE3(P29832_P29832):
MEPAAPSCNMIMIADQASVNAHGRHLDENRVYPSDKVPAHVANKILESGTETVRCDLTLEDMLGDYEYDDPTEEEKILMDRIADHVGNDNSDMAIKHAAGPEMNIDIVTAAASMSGITTASHSAL
PIRRPNTPDFRRHSEKKAKKHKNKQRQRLDDVVAKLHKPSESEADEDFIFQQPQEEEYQEEQEPQVISSSPQHFDERSPSRSSSDHGGHSGGNSRDGQFYLSSGSEDEDDEDDERVDSGYRGSSR
SEDSSRYRPHGGSHSSRSSIKSSGSGSSRHHHKRKAVPERHHPFTPPSAKRYAAAAPSSRYECPVRVDSSDSDDAPTMVQRGTLKIKSFRPPSSGSNSNKHSSSSGGSTSSSHKKQQQQQAPSKK
PVMSSGSDKIRDMVDRTAGGYVAPNAHKKCREDKSRKYPARALEYKNLPFRPQSPQYLLGKAIQFCKEETVHDKFIMLFYTRSQDVRKAVDETRARMGMRPNLSISCPFMTEHTKPINHSRETID
RTSAACTAGTQAVWDMEERRGQKCVPRTSDYRSMIIQAANPPDFLGAVKTCLHLSQVFPKQVCMRLCSITGGLNPLPIYEETVSSYVNAQFEADDISHHEDESGEYESDCE

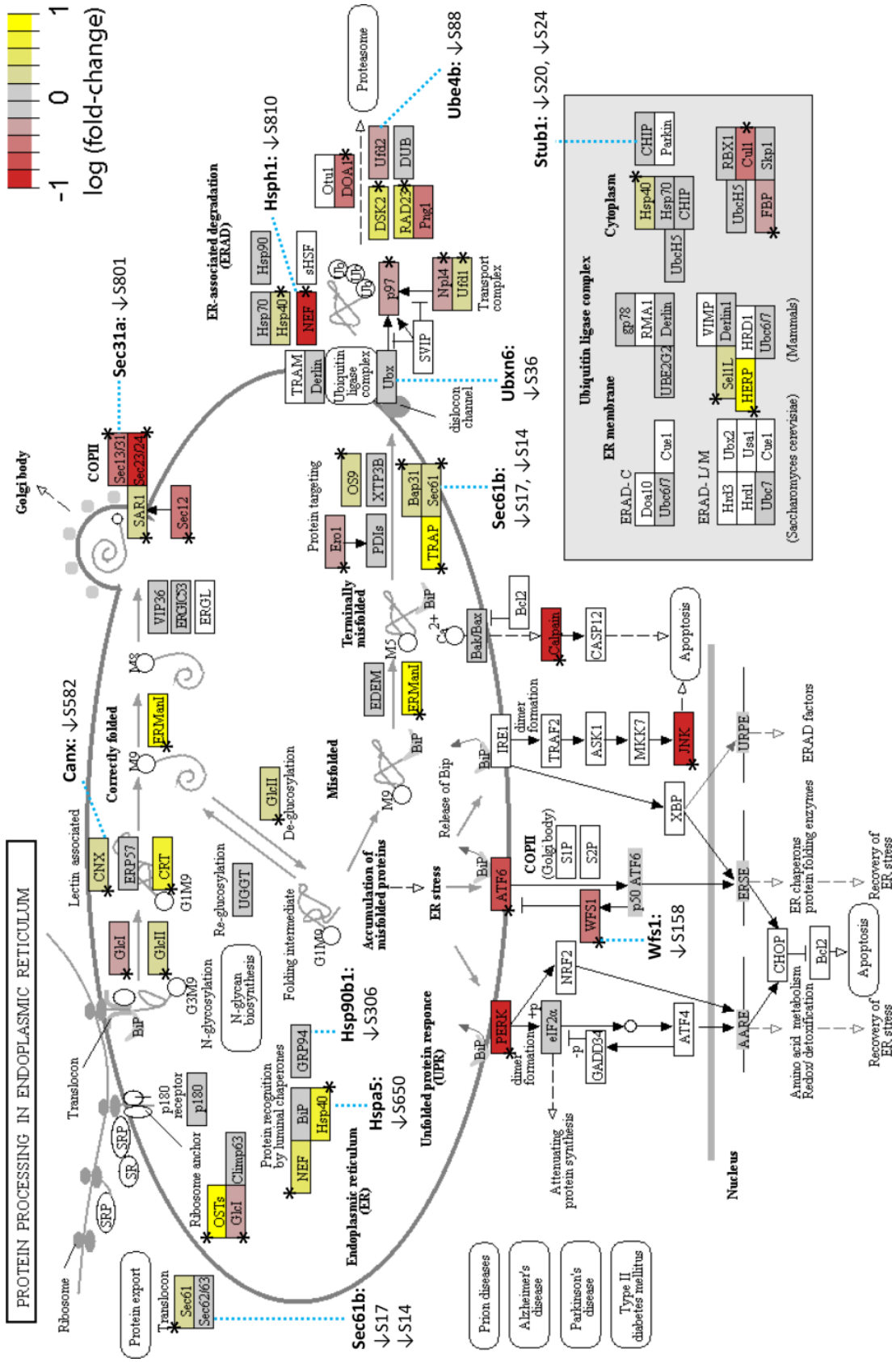
10.4. KEGG pathway maps (enlarged)

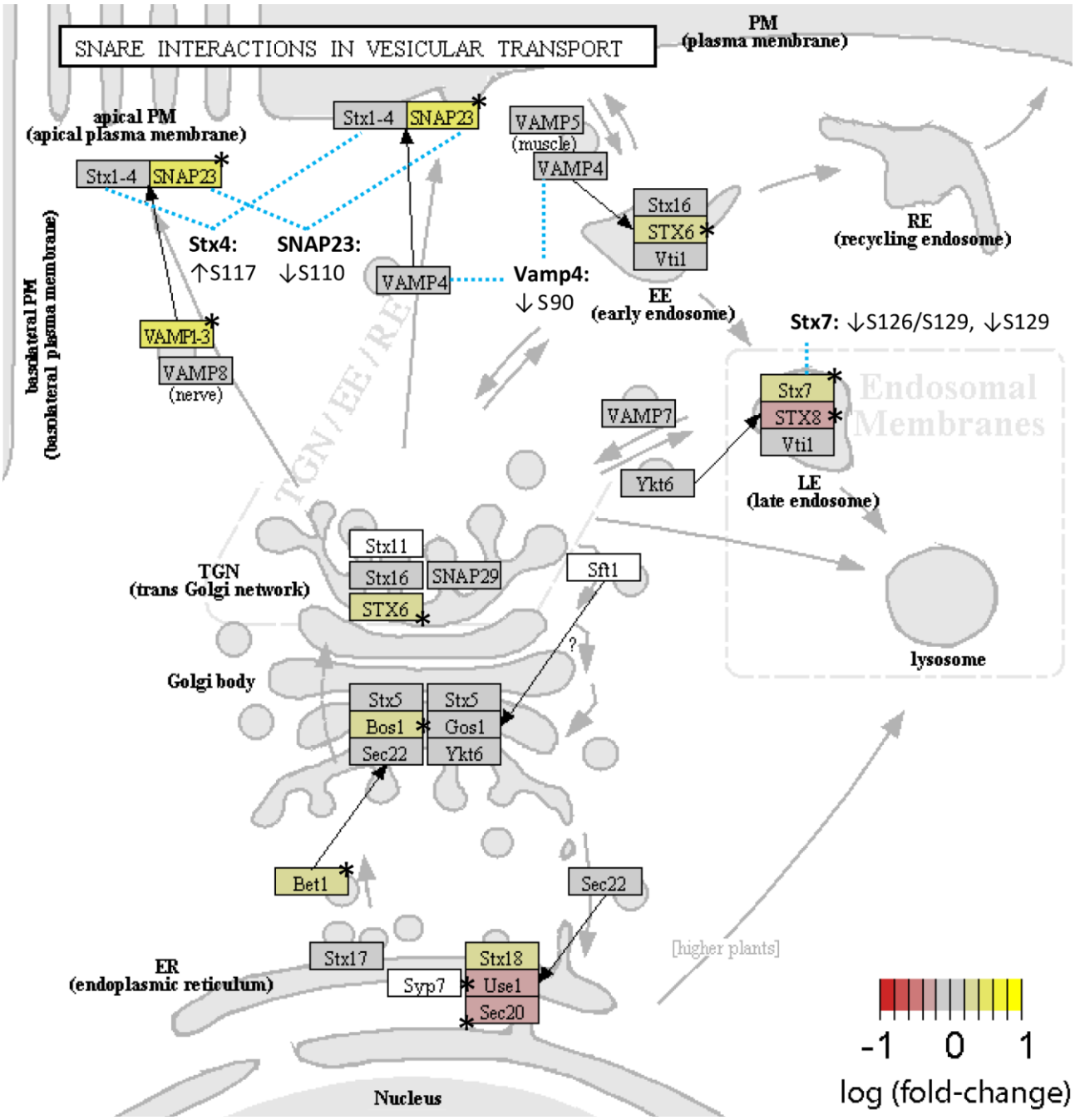



TNF SIGNALING PATHWAY
Leukocyte recruitment
Ccl2
Ccl5
Ccl20
Cxcl1
Cxcl2
Cxcl3
Cxcl5
Cxcl10
Cx3cl1
Leukocyte activation
Csf1
Csf2
Surface receptors
Fas
IL18R1
Jag1
Inflammatory cytokines
IL1b
IL6
IL15
Lif
Tnf
Intracellular signaling (negative)
Bcl 3
Nfkbia
Socs3
Tnfaip3
Traf1
Intracellular signaling (positive)
Ifi47
Transcription factors
Fos
Jun
JunB
Remodeling of extracellular matrix
Mmp3
Mmp9
Mmp14
Vascular effects
Edn1
Vegfc
PRRS
Nod2
Cell adhesion
Icam1
Sele
Vcam1
Synthesis of inflammatory mediators
Ptgs2
log (fold-change)
Apoptosis
CASP10
CASP8
CASP7
CASP3
FADD
c-FLIP
ITCH
cIAP1/2 degradation
cIAP1/2
TRADD
TRAF2/5
RIP1
TNFR1
SODD
TNF
ASK1
TAB1/2/3
TAK1
NIK
MAPK signaling pathway
MKK4/7
MKK3/6
MEK1/2
JNK1/2
p38
ERK1/2
Tpl2
NF-κB
AP-1
c/EBPβ
CREB
MSK1/2
Msk2: ↑S343/S347
Atf2: ↑T53
DNA
IκBα ubiquitination and degradation
IκBα
IKKβ
IKKα
IKKγ
NF-kappa B signaling pathway
Deubiquitination
RIP3
Necrosome
MLKL
PGAM5L
PGAM5S
Drp1
Necroptosis
Cell survival
Degradation
NF-κB
IκBα
IKKε
Akt
PI3K
PI3K-Akt signaling pathway
TRAF1
TRAF2
TRAF3
cIAP1/2
NIK
(Limited expression)
TNFR2
TNFSF1A
MKK3
RIP
p38
JNK
AIP1
c-Jun
Apoptosis


Lmna: ↑S17, ↑T19, ↓T19/S22, ↑S620, ↓S629, ↓↑S633, ↓S637, ↓S653
Lmnb1: ↑S17/T21, ↓↑T21/S24, ↑T21
Lmnb2: ↓S644
log (fold-change)
-1
0
1
APOPTOSIS
Itpr3: ↓S1832
p53: ↑S309
Bad: ↑S155
Pdpk1: ↓S38
Apoptosis
Extrinsic (receptor-mediated) pathway
Intrinsic (mitochondria-mediated) pathway
Granzyme B pathway
ER stress
Protein processing in endoplasmic reticulum
Calcium signaling pathway
TNF signaling pathway
p53 signaling pathway
NF-κB signaling pathway
MAPK signaling pathway
PI3K-Akt signaling pathway
Degradation
Mitochondrion
Lysosome
Cytotoxic T lymphocytes
Natural killer cells
Survival factors
Death receptors
Pro-apoptotic genes
Pro-survival genes
Cleavage of substrates
Disruption of microtubule function
Bim-mediated mitochondrial apoptotic events
Cell shrinkage and membrane blebbing
Loss of integrity of the nuclear membrane
Low synthesis of poly(ADP-ribose)
DNA fragmentation
DNA Damage
DNA Fragmentation
Baz/Bak-induced mitochondrial outer membrane permeabilization (MOMP)
Tropic factor deprivation, DNA damage, Ca2+ influx, UV
Sphingosine
Lisosomotropic detergents
H2O2
TRAIL
Fas-L
TNFα
NGF
IL-3
TRAIL-R
Fas
TNF-R1
TrkA
IL-3R
FADD
DAXX
TRADD
RIP1
TRAF2
AIP
ASK1
JNK
NIK
PI3K
Ras
PIP3
PDK1
Raf
MEK1/2
ERK1/2
AktPKB
IKK
IκBα
NF-κB
AP1
c-jun
p53
ATM
DNA
PUMA
NOXA
Bak
Bax
Bid
tBid
gBid
jBid
Bad
Bcl-2
Bcl-XL
Bim
Bmf
Mcl-1
Calpain
Ca2+
IP3R
PERK
Ire1α
TRAF2
eIF2α
ATF4
CHOP
FLIP
CASP12
CASP8
CASP10
CASP9
CASP2
PIDD
CASP6
CASP3
CASP7
GZMB
Effector caspases
IAPXIAP
DIABLO
ARTS
HTRA2
Pro-caspase-9
Apoptosome
CytC
Apaf-1
ENDO-G
AIF
Pro-apoptotic genes
PIDD
p53AIP1
Bax
Apaf-1
Bid
TRAILR2
Fas
PUMA
NOXA
Pro-apoptotic genes
Fas-L
Fas
HRK
Bim
Pro-survival genes
IAPXIAP
FLIP
TRAF1/2
Bcl-XL
Pro-survival genes
FAP1
Gadd45
A1
Bcl-2
α-tubulin
Mcl-1
Actin
Fodrin
Lamin
PARP
ICAD
CAD


CELL CYCLE
log (fold-change)
-1
0
1
Smc3: ↓S1067
Cdc23: ↓S588/T596
Crebbp: ↓S1077
Ep300: ↑S1037
p53: ↑S309
E2f4: ↑T14
Kip1: ↑S178
Ywhae: ↑S210
Mcm2: ↓S26/S27
Mcm3: ↑S672/S681, ↓↑T729/S732
Cohesin
Separin
Securin
Ubiquitin mediated proteolysis
DNA damage checkpoint
Apoptosis
Growth factor withdrawal
MAPK signaling pathway
R-point (START)
DNA biosynthesis
S-phase proteins, CycE
DNA
ORC (Origin Recognition Complex)
MCM (Mini-Chromosome Maintenance) complex
G1
S
G2
M


PROTEIN PROCESSING IN ENDOPLASMIC RETICULUM
log (fold-change)
-1
0
1
Sec31a: ↓S801
Canx: ↓S582
Hsph1: ↓S810
Ube4b: ↓S88
Stub1: ↓S20, ↓S24
Ubxn6: ↓S36
Sec61b: ↓S17, ↓S14
Hsp90b1: ↓S306
Hspa5: ↓S650
Wfs1: ↓S158
Sec61b: ↓S17 ↓S14
Golgi body
COPII
ER-associated degradation (ERAD)
Ubiquitin ligase complex
ER membrane
Cytoplasm
(Saccharomyces cerevisiae)
(Mammals)
Protein export
Translocon
Ribosome
Lectin associated
Correctly folded
Misfolded
Terminally misfolded
Protein targeting
Folding intermediate
Accumulation of misfolded proteins
ER stress
Unfolded protein responce (UPR)
Endoplasmic reticulum (ER)
Protein recognition by luminal chaperones
Ribosome anchor
N-glycosylation
Re-glucosylation
De-glucosylation
Release of Bip
dimer formation
Attenuating protein synthesis
Amino acid metabolism
Redox detoxification
ER chaperons protein folding enzymes
ERAD factors
Recovery of ER stress
Nucleus
Transport complex
dislocon channel
Prion diseases
Alzheimer's disease
Parkinson's disease
Type II diabetes mellitus
Proteasome
Apoptosis


SNARE INTERACTIONS IN VESICULAR TRANSPORT
PM
(plasma membrane)
apical PM
(apical plasma membrane)
basolateral PM
(basolateral plasma membrane)
Stx1-4
SNAP23
VAMP5
(muscle)
VAMP4
Stx16
STX6
Vti1
EE
(early endosome)
RE
(recycling endosome)
Stx4:
↑S117
SNAP23:
↓S110
Vamp4:
↓S90
VAMP1-3
VAMP8
(nerve)
Stx7: ↓S126/S129, ↓S129
Stx7
STX8
Vti1
LE
(late endosome)
Endosomal
Membranes
VAMP7
Ykt6
lysosome
TGN
(trans Golgi network)
Stx11
Stx16
SNAP29
STX6
Sft1
Golgi body
Stx5
Bos1
Sec22
Gos1
Ykt6
Bet1
Sec22
ER
(endoplasmic reticulum)
Stx17
Syp7
Stx18
Use1
Sec20
[higher plants]
Nucleus
-1
0
1
log (fold-change)

10.5. Hazard and precautionary statements of chemicals (GHS)

Substance	GHS-pictogram	Hazard statements	Precautionary statements
Acetone	GHS02, GHS07	H225-H319-H336	P210-P233-P261-P280-P303+P361+P353-P370+P378
Acetonitrile, ACN	GHS02	H: 225-332-302-312-319	P210+P233+P240+P241+P242+P243+P261+P264+P270+P271+P280
Ammonium acetate	none	none	none
Ammonium bicarbonate, ABC	GHS07	H302	P301+P312-P330
Ammonium hydroxide, NH_4OH	GHS05, GHS07, GHS09	H302-H314-H335-H400	P273-P280-P303+P361+P353-P304+P340+P310-P305+P351+P338-P391
Arginine	GHS07	H319	P280-P264-P305+P351+P338-P337+P313
$^{13}C_6$-Arginine	GHS07	H319	P280-P264-P305+P351+P338-P337+P313
$^{13}C_6$-$^{15}N_4$-Arginine	GHS07	H319	P280-P264-P305+P351+P338-P337+P313
Benzonase nuclease	none	none	none
Calcium chloride, $CaCl_2$	GHS07	H319	P305+P351+P338
Complete Mini (protease inhibitors)	GHS05	H314 -H318	P260-P264-P280-P301+P330+P331-P303+P361+P353-P304+P340-P310-P305+P351+P338-P363-P405-P501
Dithiothreitol, DTT	GHS07	H302-H412	P264-P270-P273-P301+P312+P330-P501
D-MEM high glucose (SILAC-Medium)	none	none	none
D-MEM low glucose	none	none	none
Ethanol, EtOH	GHS02	H225,H319	P210+P233+P240+P241+P242+P243+P264+P280
Ethylenediaminetetraacetic acid, EDTA (for cell culture)	GHS07	H319	P305+P351+P338
Fetal calf serum, FCS – dialyzed (for SILAC medium)	none	none	none
Fetal calf serum, FCS for HeLa S3	none	none	none
Fetal bovine serum, FCS for NIH-3T3	none	none	none
Formic Acid, FA	GHS02, GHS05, GHS06	H226-H302-H314-H331-EUH071	P210-P260-P280-P304+P340+P310-P305+P351+P338-P370+P378
Glutamine	none	none	none
Glycolic acid	GHS05, GHS07	H314-H332	P260-P280-P301+P330+P331-P303+P361+P353-P304+P340+P310-P305+P351+P338
Guanidinium hydrochlorid, GuHCl	GHS07	H302+H332-H315-H319	P261-P301+P312+P330-P305+P351+P338
Iodoacetamide, IAA	GHS06, GHS08	H301-H317-H334	P261-P280-P301+P310-P342+P311
Lysine	none	none	none
$^{13}C_6$-$^{15}N_2$-Lysine	none	none	none
$^{2}H_4$-Lysine	none	none	none
Methylphosphonic acid	GHS05	H314	P260-P280-P303+P361+P353-P304+P340+P310-P305+P351+P338

Substance	GHS-pictogram	Hazard statements	Precautionary statements
Phosphate-buffered saline, PBS	none	none	none
Phosphoric acid, H_3PO_4 (HPLC grade)	GHS05	H290-H314	P280-P305+P351+P338-P310
PhosSTOP (phosphatase inhibitors)	GHS07	H302	P264-P270-P301+P330+P331-P501
Potassium chloride, KCl (SupraPur)	none	none	none
Potassium dihydrogen phosphate, KH_2PO_4 (SupraPur)	none	none	none
Proline	none	none	none
Sodium dodecyl sulfate, SDS	GHS02, GHS05, GHS07	H228-H302+H332-H315-H318-H335-H412	P210-P261-P280-P301+P312+P330-P305+P351+P338+P310-P370+P378
Sodium hydroxide, NaOH (HPLC grade)	GHS05	H290-H314	P260-P280-P303+P361+P353-P304+P340+P310-P305+P351+P338
Subtilisin (P5380)	GHS05, GHS07, GHS08, GHS09	H302-H315-H318-H334-H335-H410	P261-P273-P280-P305+P351+P338-P342+P311
Titanium dioxide, TiO_2	GHS08	H319-H351	P202-P264-P280-P305+P351+P338-P337+P313-P308+P313-P405-P501
Triethylamine, TEA	GHS02, GHS05, GHS06	H225-H302-H311+H331-H314-H335	P210-P261-P280-P303+P361+P353-P305+P351+P338-P370+P378
Trifluoroacetic acid, TFA	GHS05, GHS07	H314-H332-H412	P273-P280-P305+P351+P338-P310
Trypsin (sequencing grade, modified)	GHS07, GHS08	H315-H319-H334-H335	P280-P302+P352-P305+P351+P338-P304+P341-P342+P311
Trypsin T1426 (from bovine pancreas)	GHS07, GHS08	H315-H319-H334-H335	P280-P302+P352-P305+P351+P338-P304+P341-P342+P311

11. Danksagung

An dieser Stelle möchte ich Personen danken, ohne deren Mithilfe die Anfertigung dieser Promotionsschrift niemals zustande gekommen wäre:

Zunächst danke ich Prof. Dr. Albert Sickmann, meinem Betreuer, Mentor und Lehrer, der die Mühen nicht scheute einem Biologen die Prinzipien der Analytischen Chemie zu vermitteln und der mir mit seiner Fähigkeit zur kritischen Durchleuchtung einen wissenschaftlichen Zugang zur Bioanalytik ermöglichte. Die Gespräche auf intellektueller und persönlicher Ebene werden mir immer als bereichernd und konstruktiv in Erinnerung bleiben.

Ich danke Prof. Dr. Wolfram Brune, meinem zweiten Betreuer aus Hamburg, der mir die Faszination für die Virologie beibrachte. Eine sinnvolle und systematische Auswertung von biologischen Experimenten lernte ich durch unsere eingehenden Diskussionen der Ergebnisse.

Herzlichen Dank an Prof. Dr. Hartmut Schlüter, der sich die Zeit nimmt ein zweites Gutachten zu schreiben. Ebenfalls geht ein herzlicher Dank an Prof. Dr. Thomas Dobner, Prof. Dr. Nicole Fischer und Dr. Maria Riedner, die als Mitglieder der Kommission die Arbeit eingehend prüfen.

Großer Dank an Dr. Tim Schommartz (HPI, Hamburg) und Prof. Dr. Massimiliano Mazzone (VIB, Leuven) für die lange und erfolgreiche Kooperationsarbeit und die Durchführung der Validierungsexperimente.

Vielen Dank an die Mitarbeiter des ISAS für die gute Stimmung! Ein besonderer Dank geht an Dr. Robert Ahrends, Dr. Olga Shevchuk, Oliver Pagel, Bernhard Blank-Landeshammer, Dr. Yvonne Reinders, Ingo Feldmann und Dr. Hans Frieder Schött (ISAS Dortmund) für die kritischen und wertvollen Diskussionen.

Von ganzem Herzen danke ich meiner Frau Fiorella und meinem Sohn Benjamin für die stete Fröhlichkeit, die Hilfsbereitschaft und den Rückhalt! Weiterhin danke ich meinen Eltern und meiner ganzen Familie für die Unterstützung während des gesamten Studiums.

12. Eidesstattliche Versicherung

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

Datum, Unterschrift