Université de Montréal

Optimization of High Field Asymmetric Waveform Ion Mobility Spectrometry to enhance the comprehensiveness of mass spectrometry-based proteomic analyses

Par

Sibylle Pfammatter

Département de Chimie

Faculté des Arts et des Sciences

Thèse présentée en vue de l'obtention du grade de
Philosophiæ Doctor (Ph.D.) en chimie

Octobre 2019

© Sibylle Pfammatter, 2019

Cette thèse intitulée

Optimization of High Field Asymmetric Waveform Ion Mobility Spectrometry to enhance the comprehensiveness of mass spectrometry-based proteomic analyses

Présenté par

Sibylle Pfammatter

A été évaluée par un jury composé des personnes suivantes

Karen Waldron
Président-rapporteur

Pierre Thibault
Directeur de recherche

Pierre Chaurand
Membre du jury

Susan Abbatiello
Examinateur externe

Résumé

La grande complexité des échantillons biologiques peut compliquer l'identification des protéines et compromettre la profondeur et la couverture des analyses protéomiques utilisant la spectrométrie de masse. Des techniques de séparation permettant d'améliorer l'efficacité et la sélectivité des analyses LC-MS/MS peuvent être employées pour surmonter ces limitations. La spectrométrie de mobilité ionique différentielle, utilisant un champ électrique élevé en forme d'onde asymétrique (FAIMS), a montré des avantages significatifs dans l'amélioration de la transmission d'ions peptidiques à charges multiples, et ce, en réduisant les ions interférents. Dans ce contexte, l'objectif de cette thèse était d'explorer les capacités analytiques de FAIMS afin d'élargir à la fois la gamme dynamique de détection des protéines/peptides et la précision des mesures en protéomique quantitative par spectrométrie de masse. Pour cela, nous avons systématiquement intégré FAIMS dans des approches classiques en protéomique afin de déterminer les changements dynamiques du protéome humain en réponse à l'hyperthermie.

Nous avons d'abord étudié les avantages de FAIMS par rapport à la quantification par marquage isobare (tandem mass tag, TMT). Cette approche permet le marquage d'ions peptidiques avec différents groupements chimiques dont les masses nominales sont identiques mais différant par leur distribution respective d'isotopes stables. Les ions peptidiques marqués par TMT produisent des ions rapporteurs de masses distinctes une fois fragmentés en MS/MS. Malheureusement, la co-sélection d'ions précurseurs conduit souvent à des spectres MS/MS chimériques et une approche plus lente basée sur le MS^3 est nécessaire pour une quantification précise. Comme FAIMS améliore l'efficacité de séparation en transmettant sélectivement des ions en fonction de leur voltage de compensation (CV), nous avons obtenu moins de co-sélection de peptides. FAIMS a amélioré la quantification des peptides TMT au niveau MS^2 et a permis d'obtenir 68% plus de peptides quantifiés par rapport aux analyses LC-MS/MS classiques, fournissant ainsi un aperçu plus vaste des changements dynamiques du protéome humain en réponse au stress thermique.

De plus, nous avons étudié le marquage métabolique par incorporation d'acides aminés marqués par des isotopes stables en culture cellulaire (SILAC). Si des interférences co-éluent avec les isotopes SILAC, la quantification devient imprécise et les contreparties de SILAC peuvent être assignées de manière erronée aux ions interférants du chromatogramme, faussant ainsi le rapport SILAC. Le fractionnement post-ionisation FAIMS pourrait filtrer les ions appartenant au bruit de fond qui pourraient autrement être attribués à une paire ou à un triplet SILAC pour la quantification. Dans ce projet, FAIMS a été particulièrement bénéfique pour les espèces peu abondantes et s'est montré plus performant que le fractionnement par échange de cations (SCX). En outre, FAIMS a permis la séparation des phosphoisomères fréquemment observés dans les extraits complexes de phosphoprotéomes.

Le troisième objectif de ce travail de recherche était d'explorer la séparation de l'état de charge et la transmission améliorée de peptides fortement chargés avec FAIMS et son application à l'analyse de peptides SUMOylés. FAIMS pourrait ainsi améliorer la transmission des peptides SUMOylés triplement chargés par rapport aux peptides tryptiques usuels, lesquels sont principalement doublement chargés. Ceci permettait l'enrichissement en phase gazeuse des ions peptides SUMOylés. FAIMS est une approche alternative plus simple pour fractionner les peptides SUMOylés, ce qui réduit les pertes d'échantillon et permet de simplifier le traitement des échantillons, tout en augmentant l'efficacité de séparation de manière plus automatisée et en ajoutant un ordre de grandeur de sensibilité.

Le dernier objectif de cette thèse était d'améliorer l'instrumentation de FAIMS en le jumelant aux instruments à la fine pointe de la technologie. Avec un nouveau dispositif FAIMS, développé par nos collaborateurs chez Thermo Fisher Scientific, nous avons montré une amélioration dans la robustesse et la transmission des ions pour la nouvelle interface. Dans des expériences simples en protéomique shotgun, FAIMS a étendu la gamme dynamique d'un ordre de grandeur pour une couverture protéomique plus profonde par rapport aux analyses LC-MS/MS classiques. En outre, le fractionnement en phase gazeuse de FAIMS a généré moins d'analyses chimériques en MS^2, ce qui a permis d'obtenir plus d'identifications et une meilleure quantification. Pour ce faire, nous avons directement comparé le LC-FAIMS-MS/MS au LC-MS/MS/MS

en utilisant la sélection de précurseur synchrone (SPS) avec et sans fractionnement en phase inverse basique. Des mesures quantitatives comparables ont été obtenues pour toutes les méthodes, à l'exception du fait que FAIMS a parmi d'obtenir un nombre 2,5 fois plus grand de peptides quantifiables par rapport aux expériences sans FAIMS.

Globalement, cette thèse met en évidence certains des avantages que FAIMS peut offrir aux expériences en protéomique en améliorant à la fois l'identification et la quantification des peptides.

Mots clés: FAIMS, mobilité ionique différentielle, protéomique, spectrométrie de mass, TMT, SILAC, choc thermique, phosphorylation, SUMOylation

Summary

The high complexity of biological samples can confound protein identification and compromise the depth and coverage of mass spectrometry-based proteomic analyses. Separation techniques that provide improved peak capacity and selectivity of LC-MS/MS analyses are often sought to overcome these limitations. High-field asymmetric waveform ion mobility spectrometry (FAIMS), a differential ion mobility device, has shown significant advantages by enhancing the transmission of multiple-charged peptide ions by reducing singly-charged interferences. In this context, the goal of this thesis was to explore the analytical capabilities of FAIMS to extend both the dynamic range of proteins/peptides detection and the precision of quantitative proteomic measurements by mass spectrometry. For this, we systematically integrated FAIMS in standard workflows to monitor the dynamic changes of the human proteome in response to hyperthermia.

We first studied the merits of FAIMS to aid isobaric labeling quantification with tandem mass tags (TMT). This approach allows the labeling of peptide ions with different chemical groups of identical nominal masses but differing in their respective distribution of stable isotopes. TMT-labeled peptide ions produce reporter ions of distinct masses once fragmented by MS/MS. Unfortunately, the co-selection of precursor ions often leads to chimeric MS/MS spectra, and a slower MS^3 centric approach is needed for precise quantification. Since FAIMS improves peak capacity by selectively transmitting ions based on their compensation voltage (CV), we obtained less peptide co-selection. FAIMS improved TMT quantification at the MS^2 level and achieved 68 % more quantified peptides compared to regular LC-MS/MS, providing a deeper insight into the dynamic changes of the human proteome in response to heat stress.

Further, we investigated stable isotope labeling by amino acids in cell culture (SILAC) quantification. If interferences co-elute simultaneously with SILAC isotopomers, quantification becomes inaccurate and SILAC counterparts can be missassigned to interfering ions in the highly populated chromatogram, thus skewing the SILAC ratio. FAIMS post-ionization fractionation could filter out background ions that can otherwise

be attributed to a SILAC pair/triplet for quantification. In this work, FAIMS was especially beneficial for low abundant species and outperformed the standard strong cation exchange (SCX) fractionation workflow. In addition, FAIMS allowed the separation of phosphoisomers that are frequently observed in complex phosphoproteome extracts.

The third aim of this work explored the charge state separation and enhanced transmission of highly charged peptides with FAIMS and its application for SUMOylated peptide analysis. FAIMS could enhance the transmission of triply charged SUMOylated peptides over typical tryptic peptide that are predominantly doubly charged, by applying more negative CVs with FAIMS. This allowed for gas-phase enrichment of SUMOylated peptide ions. FAIMS is an alternate and more straightforward approach to fractionate SUMOylated peptides that reduced sample loss, avoided sample processing, while increasing peak capacity in a more automated manner and added one order of magnitude in sensitivity.

The last aim of this thesis was to improve the FAIMS instrumentation by interfacing it to the latest state-of-the-art instruments. With a new FAIMS device developed by our collaborators at Thermo Fisher Scientific, we demonstrate the robustness and the improved ion transmission for the new interface. In simple shotgun proteomics, FAIMS extended the dynamic range by one order of magnitude for deeper proteome coverage compared to regular LC-MS/MS. Moreover, fewer MS^2 chimeric scans were generated with FAIMS gas-phase fractionation, which garnered more identifications and better quantification. For this, we directly compared LC-FAIMS-MS/MS to LC-MS/MS/MS using synchronous precursor selection (SPS) with and without basic reverse phase fractionation. Comparable quantitative measurements were obtained for all methods, except that FAIMS provided a 2.5-fold increase in the number of quantifiable peptides compared with non-FAIMS experiments.

Overall, this thesis highlights some of the advantages that FAIMS can provide for proteomics experiments by improving both peptide identification and quantification.

Key words: FAIMS, differential ion mobility, Proteomics, Mass Spectrometry, TMT, SILAC, heat shock, phosphorylation, SUMOylation

Table of contents

Résumé I

Summary IV

List of tables XI

List of figures XII

List of abbreviations XVI

Acknowledgements XIX

1. Introduction 2

1.1. Proteomics 3

1.1.1. From DNA to proteins: The protein synthesis 4

1.1.1.1. Proteoforms increase proteome complexity 6

1.1.1.2. Post Translational Modifications 7

1.1.1.3. Error in protein synthesis 8

1.1.2. Eukaryotic stress responses 8

1.1.2.1. Heat stress and the response 9

1.2. Mass spectrometry-based proteomics 11

1.2.1. Bottom-up proteomics 12

1.2.1.1. Typical sample preparation in shotgun proteomics 13

1.2.1.2. Enrichment methods for Post Translation Modifications 14

1.2.2. Mass spectrometry instrumentation 18

1.2.3. Ion sources 18

1.2.3.1. Electrospray 19

1.2.4. Mass analyzer 20

1.2.4.1. Linear quadrupole 21

1.2.4.2. Linear trap quadrupole 23

1.2.4.3. Orbitrap 23
1.2.5. Tandem mass spectrometry 26
1.2.5.1. Ion dissociation mode 28
1.2.6. Hybrid mass spectrometer 29
1.2.7. Analysis mode 31
1.2.7.1. Data Dependent Acquisition (DDA) 32
1.2.7.2. Data Independent Acquisition (DIA) 34
1.2.8. Quantitative proteomics 36
1.2.8.1. Label free 37
1.2.8.2. Chemical labeling - Tandem Mass Tags 38
1.2.8.3. Metabolic labeling - stable isotope labeling by amino acids in cell culture (SILAC) 40
1.2.9. Data analysis 41
1.2.10. Sample complexity of protein digests 43
1.2.10.1. Chromatographic separation 46
1.2.10.2. Multidimensional peptide fractionation 47
1.3. Ion mobility spectrometry (IMS) 49
1.3.1. Fundamentals in IMS 49
1.3.2. Linear methods 52
1.3.2.1. Drift Tube IMS (DTIMS) 53
1.3.2.2. Travelling Wave IMS (TWIMS) 53
1.3.2.3. Trapped IMS (TIMS) 54
1.3.3. Nonlinear methods 55
1.3.4. Field Asymmetric Waveform Ion Mobility Spectrometry (FAIMS) 56
1.3.4.1. FAIMS evolution in bottom-up proteomics 59
1.4. Research objectives 62
1.5. Content of the thesis 63
1.6. References 64

2. Improvement of quantitative measurements in multiplex proteomics using high-field asymmetric waveform spectrometry (FAIMS) 85

2.1. Abstract 86

2.2. Introduction 87

2.3. Experimental Section 90

2.4. Results and Discussion 93

2.4.1. Impact of precursor co-fragmentation on protein identification 93

2.4.2. FAIMS improves accuracy and precision of quantitative measurements 96

2.4.3. Monitoring temporal changes in protein abundance following heat shock 100

2.5. Conclusion 108

2.6. Acknowledgments 110

2.7. Supplementary material 111

2.7.1. Supplementary figures 111

2.7.2. Supplementary tables 114

2.8. References 115

3. Accurate quantitative proteomic analyses using metabolic labeling and High Field Asymmetric Waveform Ion Mobility Spectrometry (FAIMS) 121

3.1. Abstract 122

3.2. Introduction 123

3.3. Experimental Section 125

3.4. Results and Discussion 128

3.4.1. FAIMS enhances the depth and precision of quantitative proteomics using SILAC 128

3.4.2. Profiling early signaling events upon heat shock 133

3.4.3. Gas phase ion fractionation enhances the resolution of isomeric phosphopeptides 139

3.5. Conclusion 141

3.6. Acknowledgments 143

3.7. Supplementary material 144

3.7.1. Supplementary figures 144
3.7.2. Supplementary tables 150
3.8. References 151
4. Gas-phase enrichment of multiply charged peptide ions by differential ion mobility extend the comprehensiveness of SUMO proteome analyses 156
4.1. Abstract 157
4.2. Introduction 158
4.3. Experimental Section 161
4.4. Results and Discussion 165
4.4.1. FAIMS improves the detection and identification of SUMOylated peptides present in complex tryptic digests 166
4.4.2. Large-scale profiling of SUMO proteome under heat shock 171
4.5. Conclusion 182
4.6. Acknowledgments 183
4.7. Supplementary material 184
4.7.1. Supplementary figures 184
4.7.2. Supplementary tables 190
4.8. References 191
5. A novel differential ion mobility device expands the depth of proteome coverage and the sensitivity of multiplex proteomic measurements 196
5.1. Abstract 197
5.2. Introduction 198
5.3. Experimental Section 202
5.4. Results 208
5.4.1. Ruggedness and stability of the FAIMS interface 208
5.4.2. Optimization of FAIMS for LC-MS/MS analyses 212

5.4.3. FAIMS extends the comprehensiveness of quantitative proteomics using isobaric labeling 216
5.4.4. Dynamic proteomics enabled the profiling of cellular response upon hyperthermia 220
5.5. Discussion 225
5.6. Acknowledgment 229
5.7. Note 229
5.8. Supplementary material 230
5.8.1. Supplementary figures 230
5.8.2. Supplementary tables 239
5.9. References 240
6. Conclusion and Perspectives 246
6.1. Conclusion 247
6.1.1. FAIMS reduces co-fragmentation and enhances identifications and quantification with isobaric labeling 247
6.1.2. FAIMS improves SILAC based quantification by reducing confounding ions 248
6.1.3. Gas phase enrichment of SUMOylated peptide ions with FAIMS 249
6.1.4. A new FAIMS device with improved speed and robustness 249
6.2. Perspectives 250
6.2.1. Short gradients with single CV fractions 250
6.2.2. Improving DIA quantification and identification with FAIMS 251
6.2.3. Expanding the depth of the proteome with sequential narrow survey scans and FAIMS 252
6.2.4. FAIMS for the identification of peptides from non-coding regions 252
6.2.5. Charge state separation for highly charged crosslink peptides 253
6.2.6. Machine learning for targeted analysis with FAIMS 253
6.3. References 254

List of tables

***Supplementary Table 2-1**: List of identified and quantified peptides from yeast and human obtained from FAIMS and non-FAIMS LC–MS/MS experiments. 114*

***Supplementary Table 2-2:** List of quantified proteins and corresponding changes in abundance upon heat shock. 114*

***Supplementary Table 2-3:** Bioinformatics and GO terms analyses of clusters showing dynamic changes in protein abundance upon heat shock. 114*

***Supplementary Table 3-1**: List of quantified SILAC tryptic peptides 150*

***Supplementary Table 3-2**: List of Phospho substrates and modification sites. 150*

***Supplementary Table 4-1**: List of identified SUMO substrates and modification sites. 190*

***Supplementary Table 5-1**: List of FAIMS LC-MS/MS experiments at CV -45V for 100 Replicates. 239*

***Supplementary Table 5-2**: List of peptide identifications of LC-MS/MS experiments with and without FAIMS for 500ng HEK293 digest/injection. 239*

***Supplementary Table 5-3**: List of protein identifications of LC-MS/MS experiments with and without FAIMS for 500ng HEK293 digest/injection. 239*

***Supplementary Table 5-4**: List of quantified dynamic Protein groups assigned to a cluster 239*

List of figures

***Figure 1-1**: Protein synthesis.* 5

***Figure 1-2**: Overview of different approaches to analyze proteins in mass spectrometry-based proteomics, showing typical workflows for top-down, middle-down and bottom-up.* 11

***Figure 1-3**: Bottom-up proteomics.* 13

***Figure 1-4:** Enzymatic cascade for the conjugation of the mature form of our in house SUMO3 mutant to target proteins.* 17

***Figure 1-5:** Quadrupole filter.* 22

***Figure 1-6**: Orbitrap mass analyzer.* 25

***Figure 1-7:** Nomenclature for peptide fragment ions* 27

***Figure 1-8:** Scheme of the hybrid Orbitrap Elite and Fusion.* 29

***Figure 1-9:** Illustration for data dependent acquisition and data independent acquisition.* 32

***Figure 1-10:** Co-elution of similar m/z ions from complex samples negatively impacts all three main quantification methods that are used in proteomics.* 45

***Figure 1-11:** Classical drift time IMS compared to DMS..* 52

***Figure 1-12:** Scheme of the classical drift tube IMS and the dynamic waveform IMS TWIMS.* 53

***Figure 1-13:** Parallel accumulation of ions with TIMS device.* 55

***Figure 1-14:** Ion separation using FAIMS.* 57

***Figure 2-1:** Impact of precursor co-selection on peptide identification for the analysis of a HEK293 tryptic digest.* 95

***Figure 2-2**: Impact of peptide ion co-selection on TMT-quantification with a two-proteome model.* 97

***Figure 2-3:** Distortion of TMT ion ratios and extent of ion contamination for the two-proteome model.* 99

***Figure 2-4:** Temporal profiling of protein abundance from HEK293 cells following heat shock treatment for up to 10 h.* 101

***Figure 2-5:** Proteome wide effects of heat shock monitored by LC-FAIMS-MS/MS.* 102

***Figure 2-6:** Bioinformatic analyses of proteins dynamically regulated upon hyperthermia.* 104

***Figure 2-7:** Summary of protein expression changes observed during heat stress.* 107

***Figure 3-1:** FAIMS extends the depth of quantitative proteome analyses using SILAC.* 129

Figure 3-2: *FAIMS improves the accuracy of quantitative measurements performed using SILAC and LC-MS/MS.* *131*

Figure 3-3: *FAIMS reduces sample complexity and facilitate the detection of low abundance peptide isotopomers.* *132*

Figure 3-4*: Quantitative phosphoproteomic analyses of HEK293 cells following heat shock.* *135*

Figure 3-5: *Separation of co-eluting phosphopeptide isomers using FAIMS.* *140*

Figure 4-1: *Gas-phase separation of SUMOylated peptides using FAIMS.* *165*

Figure 4-2: *Charge state distribution of tryptic peptides and SUMOylated peptides.* *168*

Figure 4-3: *LC-MS/MS analyses of SUMOylated peptides spiked in a HEK293 tryptic digest.* *170*

Figure 4-4: *Workflow SUMO proteome analysis of HEK293 cells following heat shock.* *172*

Figure 4-5: *SUMO proteome analyses of HEK293 cells following heat shock..* *174*

Figure 4-6: *Profiling changes in protein SUMOylation of HEK293 cells following heat shock.* *176*

Figure 4-7: *Network representation of identified SUMOylated proteins highlighting different cluster subsets for enriched protein functions.* *180*

Figure 5-1: *Architecture and separation capacity of a novel FAIMS interface.* *209*

Figure 5-2: *Reproducibility and robustness of FAIMS-Fusion LC-MS system with 100 replicate injections of HEK293 digest at a fix CV value of -45 V.* *211*

Figure 5-3: *LC-FAIMS-MS CV stepping improves coverage of the proteome compared to regular LC-MS.* *213*

Figure 5-4: *Depth and coverage of proteomic analyses.* *215*

Figure 5-5: *FAIMS improves TMT quantification of the human proteome.* *217*

Figure 5-6: *The heat shock response in HEK293 cells alters protein abundances with distinct dynamics.* *219*

Figure 5-7: *Heat stress affects several key cellular processes that impact protein homeostasis.* *222*

Supplementary Figure 2-1:*Distribution unique peptides as a function of parent ion fraction (PIF). 111*

Supplementary Figure 2-2: *Comparison of peptide and protein identification for HEK 293 tryptic digest using LC–MS/MS with and without FAIMS.....112*

Supplementary Figure 2-3: *Extent of interference with precursor ion intensity with and without FAIMS for the two-proteome (yeast and human) model.....112*

Supplementary Figure 2-4: *Influence of TMT derivatization on LC-FAIMS-MS analysis of HEK293 tryptic digest.....113*

Supplementary Figure 2-5: *Fold change median for dynamic and static proteins after heat shock. 114*

Supplementary Figure 3-1: *Physicochemical characteristics of unique and common peptide sequences identified with the various methods. 144*

Supplementary Figure 3-2: *FAIMS expands the identification and quantification of low abundance peptides with missed cleavage sites. 145*

Supplementary Figure 3-3: *FAIMS improves SILAC quantitation accuracy for large fold changes. 146*

Supplementary Figure 3-4: *FAIMS improves the precision and accuracy of quantitative measurements performed using SILAC and LC-MS/MS. 146*

Supplementary Figure 3-5: *FAIMS improves the SILAC quantitation of the lower abundance peptides. 147*

Supplementary Figure 3-6: *Distribution of observed fold change and ratio compression for LC-MS/MS experiments performed with and without FAIMS.....148*

Supplementary Figure 3-7: *Precision of phosphoproteomic measurements for LC-MS/MS performed with and without FAIMS.....149*

Supplementary Figure 3-8: *Distribution of the number of PSM identified and the relative number of missed cleavages for the different CV transmission ranges. 150*

Supplementary Figure 4-1: *SUMO3 peptide LLVHMGLLS**K**EDK at constant concentration of 2pmol/µl was infused with increased concentration of spiked in HEK293 peptides.....184*

Supplementary Figure 4-2: *SUMO3 peptide LLVHMGLLS**K**EDK at constant concentration of 2pmol/µl was infused with increased concentration of spiked in HEK293 peptides for different AGC values. 185*

Supplementary Figure 4-3: *Distribution of 3+ SUMOylated and tryptic peptides ions. Bar chart showing the distribution of 3+ peptide ions for different m/z ranges.....186*

Supplementary Figure 4-4: *Coefficient of determination r^2 for common peptides identified with and without FAIMS. 186*

Supplementary Figure 4-5: *Dynamic changes in protein SUMOylation upon heat shock. Western blot using anti SUMO2/3 antibody.* *187*

Supplementary Figure 4-6: *FAIMS extend the depth and sensitivity of SUMO proteome analyses.* *188*

Supplementary Figure 4-7: *Match between runs extend the detection of SUMOylated peptides in FAIMS experiments.* *189*

Supplementary Figure 4-8: *Role of BRCA1 in DNA damage response pathway created with Ingenuity Pathway Analysis (IPA).* *190*

Supplementary Figure 5-1: *Schematic of the new FAIMS interface.* *230*

Supplementary Figure 5-2: *Ion transmission and stability of the new FAIMS prototype by direct infusion of Angiotensin II.* *232*

Supplementary Figure 5-3: *Resolution of the FAIMS interface using a BSA tryptic digest as benchmark.* *233*

Supplementary Figure 5-4: *Reproducibility and robustness of FAIMS-Fusion LC-MS system. 100 replicate injections of 500 ng HEK293 digest at a fixed CV (CV -45 V) were continuously monitored over 5 days.* *234*

Supplementary Figure 5-5: *Reproducibility and robustness of regular LC-MS system with 27 replicate injections of HEK293 digest.* *235*

Supplementary Figure 5-6: *Optimization of FAIMS CV stepping program. HEK293 digests were analyzed with FAIMS with CV from -37 V to -93 V with 7 V steps to cover the entire peptide transmission range.* *236*

Supplementary Figure 5-7: *Comparison between SPS and FAIMS based TMT quantifications.* *237*

Supplementary Figure 5-8: *The heat shock response affects proteins with similar functions in the same demeanor.* *238*

List of abbreviations

ACN	Acetonitrile
AGC	Automatic Gain Control
APD	Advanced Peak Determination
AQUA	Absolute Quantification
AS	Alternative Splicing
ATP	Adenosine Triphosphate
BSA	Bovine Serum Albumin
CCS (Ω)	Collision Cross Section
CID	Collision Induced Dissociation
CTL	Control
CV	Compensation Voltage
Da	Dalton
DC	Direct Current
DDA	Data Dependent Acquisition
DIA	Data Independent Acquisition
DMS	Differential Mobility Spectrometry
DNA	Deoxyribonucleic Acid
DTIMS	Drift Tube Ion Mobility Spectrometry
DV	Dispersion Voltage
EFG	Electric Field Gradient
eFT	enhanced Fourier Transformation
ESI	Electrospray Ionization
ETD	Electron Transfer Dissociation
FA	Formic Acid
FAIMS	High Field Asymmetric Waveform Ion Mobility Spectrometry
FC	Fold Change
FDR	False Discovery Rate
FT	Fourier Transformation
FT-IRC	Fourier Transform–Ion Cyclotron Resonance
FWHM	Full-Width at Half-Maximum
GO	Gene Ontology
HCD	Higher-energy Collisional Dissociation
HILIC	Hydrophobic Interaction
HpH	High pH
HPLC	High Performance Liquid Chromatography
HR/AM	High Resolution/Accurate Mass
HRMS	High Resolution Mass Spectrometer
HSF	Heat Shock Factor
HSP	Heat Shock Proteins
HSR	Heat Shock Response
ICAT	Isotope-Coded Affinity Tag
IEF	Isoelectric Focusing
IMS	Ion Mobility Spectrometry
IP	Immunoprecipitation
IQR	Interquartile Range
iTRAQ	isobaric Tags for Relative and Absolute Quantification
LC	Liquid Chromatography
LFQ	Label Free Quantification
LIT	Linear Ion Trap
LLOQ	Lower Limit Of Quantitation

LOD	Limit Of Detection
LTQ	Liner Trap Quadrupole
m/z	Mass-to-Charge
MALDI	Matrix-Assisted Laser Desorption/Ionization
MCB	Main Control Board
mRNA	messenger Ribonucleic Acid
MS	Mass Spectrometry
MS/MS or MS^2	Tandem Mass Spectrometry
MS^3	Triple-stage Mass Spectrometry
NCE	Normalized Collision Energy
NHS	N-Hydroxysuccinimide
nsSNP	non-synonymous Single Nucleotide Polymorphism
ORF	Open Reading Frame
pAGC	predictive Automatic Gain Control
PIF	Precursor Ion Fraction
ppm	Part Per Million
PRM	Parallel Reaction Monitoring
PSM	Peptide Spectrum Match
PTM	Post Traditional Modifications
RF	Radio Frequency
RNA	Ribonucleic Acid
RP	Reversed-Phase
RSD	Relative Standard Deviation
S/N	Signal-to-Noise
SAE	SUMO Activating Enzyme
SAP	Single Amino Acid Polymorphism
SCoPE-MS	Single Cell ProtEomics by Mass Spectrometry
SCX	Strong Cation-Exchange
SD	Standard Deviation
SDM	Spectrum Deconvolution Method
SENP	Sentrin-specific Protease
SILAC	Stable Isotope Labeling by Amino Acids in Cell Culture
SPS	Synchronous Precursor Selection
SRIG	Stacked Ring Ion Guide
SRM	Selected Reaction Monitoring
SUMO	Small Ubiquitin-like MOdifier
TCE	Total Cell Extract
TCEP	Tris (2-Carboxyethyl) Phosphine Hydrochloride
Td	Townsends
TEAB	Triethyl Ammonium Bicarbonate
TFA	Trifluoroacetic Acid
TFS	Thermo Fisher Scientific
Th	Thomson
TIMS	Trapped Ion Mobility Spectrometry
TMT	Tandem Mass Tag
TOF	Time-Of-Flight
tRNA	Transfer Ribonucleic Acid
TWIMS	Travelling Wave Ion Mobility Spectrometry
UVPD	Ultra-Violet Photodissociation
XIC	Extracted Ion Chromatogram

Pour Eric Bonneil

Acknowledgements

''A mentor is someone who sees more talent and ability within you, than you see in yourself, and helps bring it out of you.'' A quote from Bob Proctor that accurately describes the support I obtained from my supervisor ***Pierre Thibault***. Pierre, it was a tremendous honor for me to be one of your students. I am infinitely grateful for the opportunity you gave me to join your group. With FAIMS I could explore my analytical curiosity and with it, came a great passion. But most importantly, you always encouraged me and showed me my hidden capacities. Thank you for having faith in me!

Heartfelt thanks to my family: ***Mama***, ***Papa***! Ich danku ew, dass iher mier d'Meglichkeit gigäh heid, mine Tröim va Kanada z'erfillu. Mini Bigeischtrig fer Kanada isch unändlich und ich ha hiä zMontreal miis Glick gfundu. Öi wenn die Distanz nit immer eifach isch gsi, ich ha die Ziit sehr chännu gniässu, unerandrum durch ewi Unerstüzig. Danke ***Tamara*** und ***Fabi***! Ihär siid die beschtu Gschwischterti wa mu schich nur cha vorstellu. Ich bi froh ew alli z'ha!

Thank you to all of the people that I was fortunate to interact with during my time in Quebec. In particular ***Cédric***, ***Chongyang***, ***Christina***, ***Christine***, ***Clémence***, ***Fanny***, ***Fiona***, ***Irene***, ***Jenna***, ***Peter***, ***Simon*** and ***Trent***. We had lovely times in the student office and outside of the lab and I am sure we will stay in contact. Special mention to Irene - you were the first person who made my time in Montreal memorable, you motivated me to continue with my master's and I am glad that despite the distance we still keep in contact and can talk about French manicure, haha! Christina und Christine, danke für die unzähligen Erlebnisse und die ganzen kulinarischen Entdeckungsreisen. Ich vermisse euch sehr! Simon (I know, you were not a student), as a passionate hiker, the swiss alps are waiting for you. Thank you also to all other former and present members of the lab, who made my days easier. Good luck to ***Zhaoguan*** as you continue on the project.

Thank you ***Francis McManus*** (alias Patric Deslauriers de Paris, inventor of the McManus method), you are one of the most efficient, cordial and helpful person I know! Thank you for being a full-on FAIMS supporter from the very beginning ☺! Tes niaiseries avec Éric vont me manquer beaucoup!

GRAZIE ***Matteo***, ***Eleonora***, ***Filippo***, ***Alice***, ***Alessandro***, ***Marie***, ***Davide*** and ***Iris***. I guess my Italian did not improve as much as my Italian cooking skills, but I will continue to work on it. It was a wonderful time with you and I am excited to meet you back in Italy. Grazie mille Matteo per il tuo supporto in tutti questi anni nonostante tu abbia dovuto condividere il mio tempo e amore con FAIMS ☺.

I am probably one of the luckiest student to have had the chance to be involved in a collaborative project with Thermo Fisher Scientific in San Jose, CA for the development of a new generation of FAIMS. A special thank you to the team of ***Jean-Jacques Dunyach***! Thank you ***Satendra*** and ***Derek*** for all your help, patience and time!

Finally, my greatest thanks go to ***Eric*** Alexandre ***Bonneil*** (“mais Eric”). You are the reason why I decided to continue with my graduate studies in Montreal. Your ‘’Sibylle now to 30 years life plan’’ greatly influenced my decision (and the fact, I had no other choice)! Without your endless encouragement I would not have written this thesis. I know, it was not always easy with me. But Eric, your patience and incredible kindness helped me greatly through this time. You know me better than probably anyone else and you were able to bring a smile to my face every single day. I learned immensely from your vast instrument and research knowledge and I truly enjoyed every single day that I got to work with you in the lab. Eric, this thesis is dedicated to you. I will miss you and for sure, I will never forget you! ‘’How lucky I am to have someone who makes saying goodbye so hard’! Danke Eric! I’m looking forward to your visits to the paradise on earth: THE VALAIS.

Je me souviens

CHAPTER ONE

1. Introduction

1.1. Proteomics

To understand the nature of highly dynamic and complex biological processes in organisms, different molecular and cellular aspects have to be investigated. Distinct systems-level termed ‘’omics’’ have been developed to profile changes in the abundance of deoxyribonucleic acid, DNA (Genomics), messenger ribonucleic acid, mRNA transcript (Transcriptomics), and proteins (Proteomics) to mention only a selected number of the different fields. Accordingly, in proteomics, we exclusively focus on the proteome with the aim to characterize and gather the entirety of the proteins encoded by the genome [1]. This includes not only proteins but also their post-traditional modifications (PTMs) and other variants (e.g. polymorphisms, splicing events, etc…).

Why this interest? Proteins are required in a large number of essential cellular processes where they play key roles in maintaining proper cellular function. They are involved in the synthesis and repair of DNA [2, 3], maintenance of the cellular structure [4], management of the transport of small molecules across the cell [5], catalysis of reactions as enzymes [6], interaction to communicate in-going/out-coming signals or response to external cell stimulation [7].

The interest in proteomics has grown significantly in recent years due to the sequencing of various genomes. In 2001, forty-eight years after the double helix DNA structure was introduced by Watson and Crick [8], two first drafts of the human DNA sequence were published simultaneously, containing over twenty-five thousand protein-encoding genes [9, 10]. The number of these genes varies considerably depending on the organism. For instance, the yeast genome of *Saccharomyces cerevisiae* possess approximately 6,000 protein-encoding genes [11], which is four times less than the human proteome. Furthermore, depending on the tissues or cell types, different genes are expressed, although they are from the same genome and contain identical DNA information [12]. The expression of different genes results in the synthesis of cell-intrinsic proteins with unambiguous structures and functions. It is thus important to understand their function, expression and subcellular location [13]. To understand the function of specific proteins, we must understand their activity and partner proteins with whom they interact [6, 14] or their dynamic in space and time [15]. Although we are able

to sequence the whole genome and postulate the existence of proteins based on the genomic code, the actual existence of a protein must be confirmed by other means [16]. Fortunately, high-resolution mass spectrometry (HRMS) (coupled to liquid chromatography) can achieve this aim due to its high performance, which can routinely sequence over 4,000 human proteins in a single one hour injection [17, 18].

1.1.1. From DNA to proteins: The protein synthesis

To understand the role of proteomics, we first need to address the process by which proteins are produced in the cell. In eukaryotic cells, the genetic information is stored in the form of chromosomes in the nucleus. There are 23 different chromosome pairs in human cells that are made of DNA strands wrapped tightly around the core of eight histone proteins (nucleosome) to form chromatin. Histones, among other proteins, provide a textbook example of the importance that proteins play in living organisms. Furthermore, depending on histone modifications (i.e. phosphorylation, acetylation, methylation or/and ubiquitylation [19]), the expression of certain genes can be modulated, which is the basis of epigenetics [12]. Therefore, even if the genome of two cells is identical the proteins that are synthesized may be drastically different.

The DNA sequence is further divided into different sections, including the genes that contain the information for protein synthesis. The numbers of genes on the different chromosomes ranges from 98 (chromosome X) to 2,514 (chromosome 1) genes [20]. However, around 99% of the human genome is intergenic and contains non-coding DNA [10]. Recent numbers estimate that approximately 20,000 proteins are expressed in humans [16, 21].

In the first step of protein synthesis – transcription – the DNA is unzipped by helicases to enable RNA polymerase and transcription factors to interact with the target gene(s) (***Figure 1-1***) [22]. The attachment of the RNA polymerase on the ‘’switched-on’’ gene(s) is termed the initiation. After initiation, nucleotides from one strand (template strand) of the DNA helix are read and copied with the corresponding complementary base pair nucleotides into precursor mRNA. The precursor mRNA will therefore be

identical to the complementary strand except for the replacement of thymine (T) by uracil (U) as the base pairing counterpart to adenine (A) found on the template strand. The single strand of pre-mRNA is spliced and recombined, removing the non-coding intron segments before it can be used as an instruction plan for protein synthesis. This process happens in the spliceosome machinery, located in the nucleus and forms the mature mRNA with the help of dozens of proteins [23, 24]. Afterwards, the spliced mRNA exits the nucleus and localizes to the cytoplasmic ribosomes.

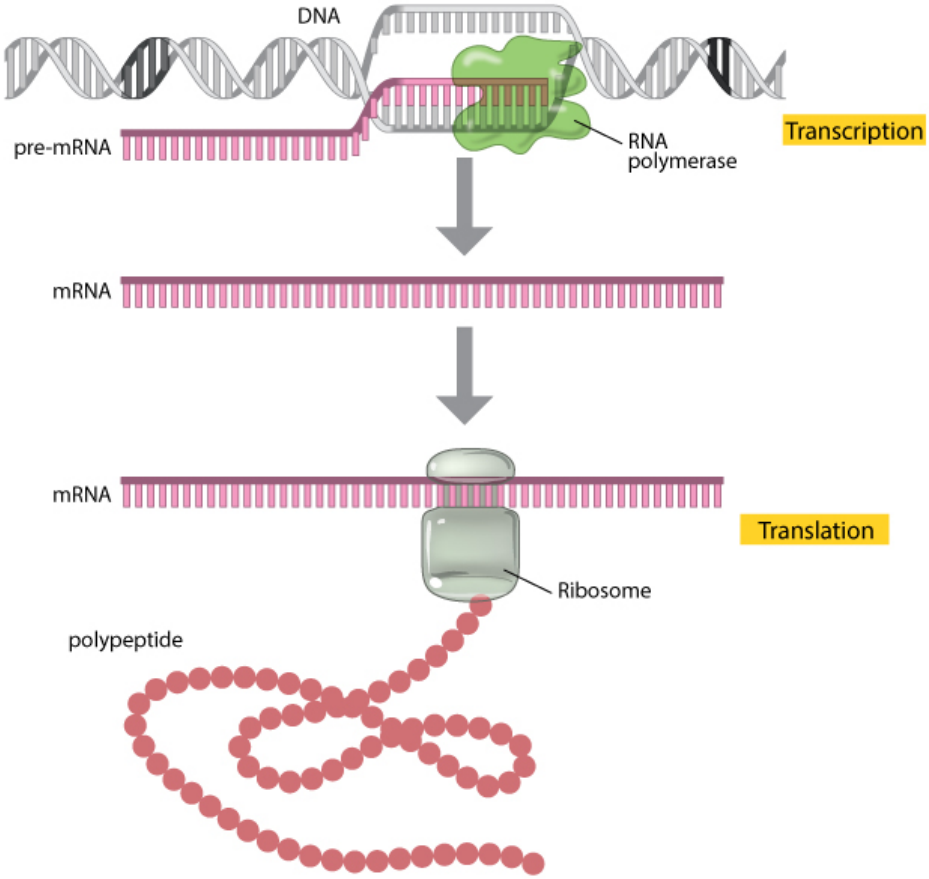


***Figure 1-1**: Protein synthesis: Only one strand of the DNA double helix is translated to mRNA. After translation in the nucleus, ribosomes in the cytoplasm use the mRNA as a template to translate proteins. From reference [25].*

In the second part of the protein synthesis – translation – proteins are assembled by the ribosomes. Briefly, during ribosome-mediated protein synthesis, the two subunits (the large and the small subunit) of the ribosomes embed into mRNA, which contains the protein template. Three adjacent nucleotides build a triplet codon. In total there are 61 different triplet combinations for the 20 standard amino acids, with AUG as initiation codon and the three remaining triplets (UAA, UAG, and UGA) as translation stop codon [26]. The mRNA segments that are translated between the start and stop codons are called open reading frames (ORFs). Per definition, only ORFs with a minimum of 100 codons are considered to be coding regions and available in reference databases such as UniProt [27, 28]. Furthermore, these mRNA instructions are decoded by transfer RNA (tRNA) adaptor molecules. The tRNA molecules have two functional sites: one site can decrypt the genetic code with the corresponding anticodon for the complementary base pair nucleotide, the other site is physically linked to one of the 20 corresponding amino acids. By translating triplet by triplet and merging the amino acids, polypeptide chains are synthesized [25]. The nascent polypeptide chain is bound to chaperons until the chain is released and folded to its tertiary structure to become a fully operational protein. The timescale for this structural self-assembling takes places within nano- to milliseconds [29].

1.1.1.1. Proteoforms increase proteome complexity

Although there are only about 20,000 protein encoding genes in the human genome, the effective number of distinct protein variations is substantially higher. One gene rarely corresponds to one specific protein: an estimated 250,000 to 1 million protein species (proteoforms) are believed to be synthesized in humans [30]. The diversity of these proteoforms [31] arises from transcriptional, post-transcriptional, and translational events. Modifications on the canonical sequence can be introduced by alternative splicing (AS), where exons from a single gene are differently included or excluded during splicing and thus result in different proteins [30, 32, 33]. Single amino acid polymorphisms (SAPs), induced by non-synonymous single nucleotide polymorphism (nsSNP) [34], change the initial genetic information by the substitution of

single nucleotides. In addition, proteins can be modified by a multitude of PTMs [35, 36]. Accordingly, up to 100 different transcripts from a single gene can occur, expanding the breath of the proteome. The abundance of the various proteoforms spans over seven orders of magnitude in the human proteome, where protein abundance ranges from one to ten million copies per cell [37]. This makes protein analysis even more difficult, considering that low abundance proteins are virtually inaccessible to mass spectrometry since only the upper five orders of magnitudes can be typically covered by regular mass spectrometers (protein distribution follows a Gaussian shape with average ten thousand copies per cell) [37, 38].

1.1.1.2. Post Translational Modifications

Compared to the transcriptome, the proteome is highly dynamic, especially in view of the nature of PTMs that can be added or removed very rapidly in the cell. The ability to capture these PTMs is important, due to their essential functions. PTMs control the biogenesis of proteins and thus their cellular function, localization, stability, and deregulation of their activity is often associated with a wide range of diseases including cancer [39]. Half of all human proteins harbor glycosylation [40]. In addition to glycosylation, phosphorylation and sulfation [41] are two of the most abundant/common protein modifications out of more than 300 different PTMs [1]. PTMs can be catalyzed by enzymes such as kinases and phosphatases that add or remove phosphate from acceptor residues, or can be from non-enzymatic origin. The hydrophilic and flexible side chain of lysine, which is positively charged under physiological pH, is the most common amino acid modified by PTMs. Lysine residues are known to undergo acetylation, methylation, ubiquitination, and SUMOylation [42]. Moreover, crosstalk between different PTMs on the same protein can dictate cell signaling events and further increase the complexity of the proteome [43].

1.1.1.3. Errors in protein synthesis

Although it is not fully understood yet, errors in protein synthesis occur frequently. This can affect the folding and the function of proteins, which can alter a cellular phenotype and be toxic to the cell. Such errors in protein synthesis are not always detrimental and can also lead to a new beneficial function [44]. The source of error is diverse and can arise from transcription, where nucleotides were skipped or misincorporated or during splicing when exons are skipped or introns inadvertently integrated [24]. Protein translation can often lead to errors either by frameshifting, premature termination or misincorporation of amino acids. Surprisingly, current estimations indicate that one out of 1,000 to 10,000 amino acids is misincorporated during translation [45], suggesting that ~15% of average-length proteins have an error in their polypeptide chain [44]. However, this does not always mean that the mature protein shows an error in the tertiary structure or function. Lastly, erroneous post-translational modifications or mistakes during protein folding are other sources of faulty proteins.

1.1.2. Eukaryotic stress responses

Errors in protein synthesis can occur under basal levels or can be promoted under certain stressful physiological conditions. Under basal conditions, cells continuously produce nascent polypeptide chains and fold them to be fully functional in their designated cell compartments. In contrast, immature proteins are mislocalized within the cell and are sent towards degradation.

When cells are exposed to acute or chronical stress, proteins can start to misfold, unfold, and aggregate and cause fluctuations to the normal cellular homeostasis [46]. Cells respond by shutting down non-essential processes, including translation, to attain a quiescent state, maintain homeostasis and reduce energy consumption. One outcome of this is the formation of stress granules that are generated to store translation initiation mRNA in order to suppress protein translation [47, 48]. In contrast, during the stress response the transcription of heat shock genes are up-regulated by heat shock factors

(HFS) [49]. Although heat shock proteins (HSP) are ubiquitously expressed under normal/basal conditions, the translation increases significantly when cells are under physical stress [49]. These heat shock proteins act as chaperones and are involved in the refolding of proteins, the removal of misfolded proteins and inhibit protein aggregation [50, 51]. The degradation of misfolded or unessential proteins begins minutes or hours after stress exposure. On the other hand, the stress-protective pathways are activated within the order of seconds to minutes. Activation can be initiated by PTMs such as phosphorylation [52]. These pathways help to protect cells against cell death by repairing the damage. Moreover, PTMs can sequester active proteins to inactive complexes [53].

1.1.2.1. Heat stress and the response

One well know stress factor is hyperthermia. Mammalian cells typically thrive at 37 °C, while an increase of only a few degrees (>4°C [51]) can stress the cell and lead to the heat shock response (HSR) [54]. Heat stress is one of the most frequently studied stimuli that leads to an imbalance in protein homeostasis. Interestingly, an abrupt thermal change does not only change protein homeostasis, but also affects the cytoskeletal structures and leads to mitochondrial damage [54]. As with other stressors, cell growth and proliferation are hindered through the storage of mRNA in stress granules [47, 48]. This highly conserved defense response is characterized by elevated production of heat shock proteins and includes Hsp100, Hsp90, Hsp70, Hsp60, and other small heat shock proteins [55]. Moreover, more than 100 human proteins are up regulated in response to heat shock, which can be regrouped into classes. The first class includes molecular chaperones that selectively stabilize specific targets [56], while other classes are involved in the metabolism, DNA/RNA repair, protein degradation, transport and detoxification, cell regulation and organizations change [54]. Proteins that are involved in protein degradation are increased during the heat shock response. This adaptation is brought upon to eliminate potentially toxic misfolded proteins from the cells (misfolded during translation or from irreversible aggregation). Further, enzymes involved in RNA and DNA damage repair are also upregulated during hyperthermia.

Cellular metabolism is also perturbed during heat stress. It is important that cells maintain an optimal physiological pH, osmotic pressure and ion concentration required for proper adenosine triphosphate (ATP) production and consumption. However, during stress ATP production is hampered [53], and cells must adapt to this low energy state. Consequently, proteins that suppress energy consumption and that rewire metabolism are activated. Transcriptions factors, including heat shock factors, which are specific key players in stress-induced transcription of proteins also increased in abundance [49]. Finally, proteins involved in cell organization, transport and detoxification of the cell are also upregulated upon heat shock.

In addition to changes in protein expression during heat shock, protein modifications are also regulated and can vary acutely in early phases to control important pathways [52]. For instance, acetylation, phosphorylation, and SUMOylation regulate the four human heat shock factors (HSF1-4). In the case of HSF1, a key player in the heat shock response [50, 57], phosphorylation and SUMOylation are crucial in its activation [58]. Importantly, the heat shock response is not only produced by heat but by other stressors and is an important pathway to maintain protein homeostasis, while increasing molecular chaperones [54]. The heat shock response can also be triggered by a wide range of stresses, including infections [59, 60]. In this context, a fever is the immune response during inflammation, infection and injury of the organism. The temporary increase of core temperature serves as a self-protection mechanism and activates elements of the heat shock response pathway [59]. Proteomics is currently the most effective tool to determine changes in protein abundance and to profile variation in PTMs in an unbiased system-level wide manner [61, 62]. Mass spectrometry-based proteomics is typically used for these experiments [37, 63] as it provides sensitivity in the femtomole range [64].

1.2. Mass spectrometry-based proteomics

Historically, proteins and their modifications were measured by low resolution 2D gels [65] and it is only recently, at the beginning of the twenty-first century, that mass spectrometry (MS)-based proteomics became inarguably the first choice for proteomics analyses. This is primarily because MS instruments became so powerful due to their high sensitivity and speed that supplies remarkable amounts of information.

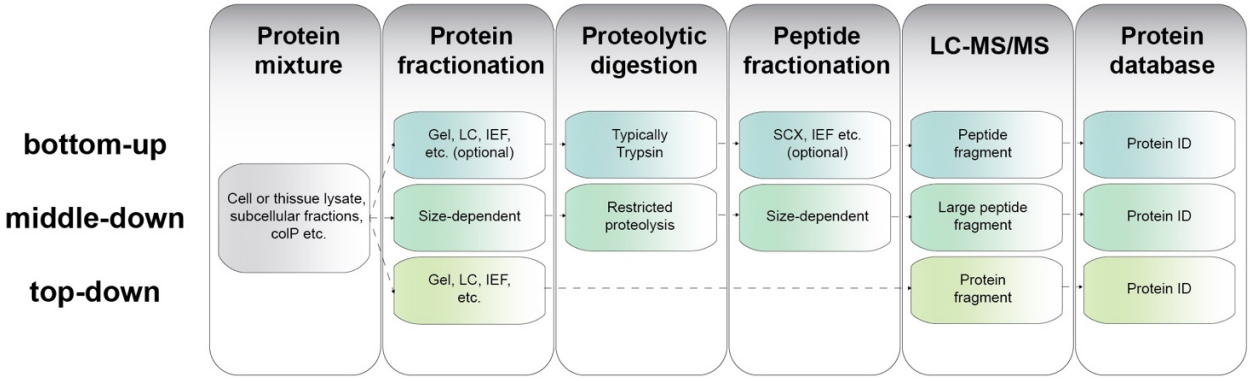


***Figure 1-2**: Overview of different approaches to analyze proteins in mass spectrometry-based proteomics, showing typical workflows for top-down, middle-down and bottom-up. Adapted from reference [66].*

High resolution accurate mass measurements can be routinely used to gather information about protein expression [67], temporal dynamic changes [68], monitor the cell cycle [69], study the turnover of the proteome [70], subcellular localization of protein [71] or tissue specific protein distribution [72]. Proteomics is also a valuable tool to study protein isoforms [73] and protein-protein interactions [74].

A typical workflow for mass spectrometry based proteomics studies requires sample preparation, protein or peptide separation, mass spectrometry analysis followed by data analysis. Depending on the nature of the experiment, there exists three strategies for proteomics: ‘’top-down’’ [75], ‘’middle-down’’ [76] or ‘’bottom-up’’ [66] proteomics for which intact proteins (M > 10 kDa), large proteins (2.5 kDa < M < 10 kDa) or small peptides (0.8 kDa < M < 3 kDa) are analyzed [77], respectively (***Figure 1-2***). Although top-down analysis is not a high-throughput method because it is limited

to single proteins or very low-complexity protein samples, the major advantage of this method is its ability to provide a greater understanding of a specific protein or group of proteins and their PTMs. It is known that assigning of shared peptide sequences to the exact source protein is impossible in bottom-up analysis. Moreover, missing peptides in the protein sequence does not allow for full sequence coverage of the protein, including PTMs. The top-down technology can overcome this problem since MS/MS spectra contain only fragments from one protein isoform [78]. Notwithstanding, the MS analysis is technically challenging and is limited to high resolution mass spectrometers equipped with different fragmentation techniques. Although one can obtain more detailed information by the top-down method, the sophisticated application is a big challenge. This explains why peptide analysis is currently the first choice in MS-based proteomics.

Compared to top-down proteomics, middle-down and bottom-up rely on enzymatic digestion. In middle-down proteomics, this digestion is restricted and depends on the type of protease, digestion time and reaction conditions. Optimal peptide lengths after proteolysis should result in ~25-100 amino acid-long chains [77]. Since the peptides are longer, tandem mass spectra are complex and make the bioinformatics analysis more difficult. Middle-down strategies provide advantages for the analyses of histone PTMs as these proteins are enriched in lysine and arginine residues, and give rise to peptides often too short for LC-MS analyses when using trypsin. However, the use of alternate enzymes in middle-down proteomics can lead to peptides of suitable size for histone PTM analyses [79, 80].

1.2.1. Bottom-up proteomics

In bottom-up or shotgun proteomics [81], peptides are obtained through complete proteolysis induced by an enzymatic digestion. Samples are from different complexity levels, but as a basic principle, the identification of proteins is obtained following liquid-chromatography separation prior to peptide sequencing by tandem MS (MS/MS). The MS method and the type of analysis depend on the individual requirements [82].

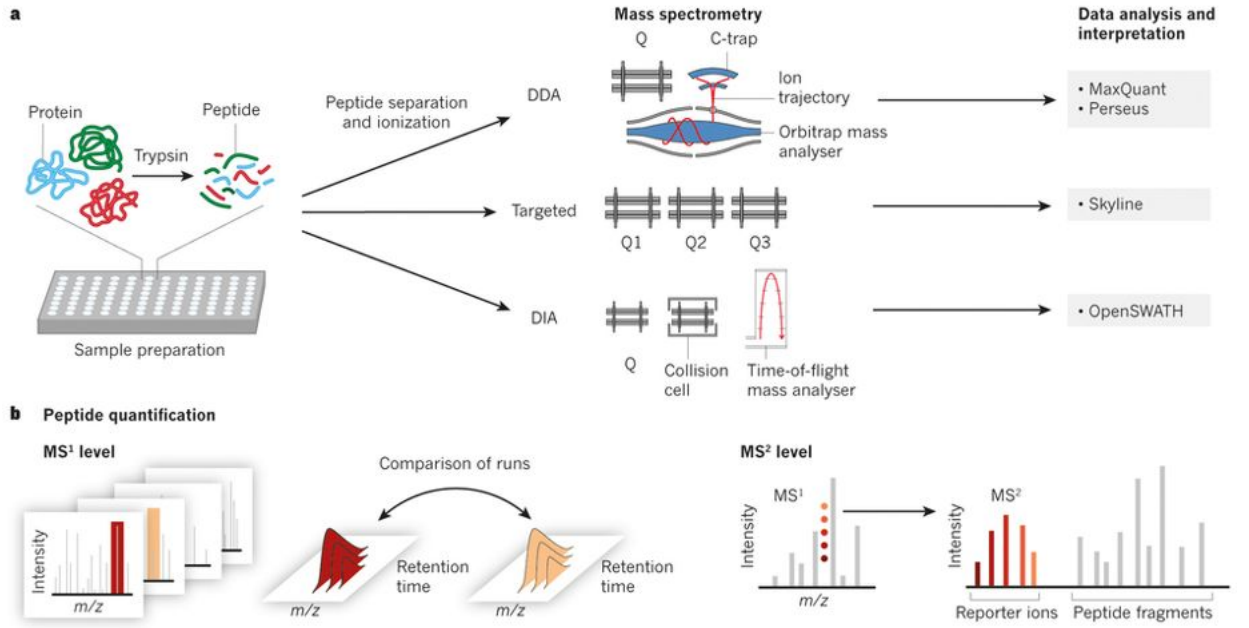


Figure 1-3: *a) In bottom-up proteomics, peptide are obtained typically by trypsinization of proteins. Complex samples are partially fractionated prior to MS analysis. Depending on the mode of acquisition, a variety of bioinformatics tools are available. b) Relative quantification strategies are based on MS or MS/MS level intensities. In label free or metabolic labeling, corresponding m/z precursors are compared. In chemical labeling, reporter ion intensities are compared in the MS^n level. From reference [82].*

1.2.1.1. Typical sample preparation in shotgun proteomics

Proteins from cell culture, body fluids, or tissue samples can be extracted by mechanical lyses of the cells under physiological conditions [83]. In the first step of the sample preparation, the disulfide bridges from the native proteins are reduced and alkylated to facilitate the subsequent proteolysis. The most widely used digestion enzyme in shotgun proteomics is trypsin, which cleaves specifically on the C-terminal side of lysine and arginine residues [84]. Tryptic digestion of proteins typically yields peptides with a charge of +2 (see ***chapter 1.2.3.1*** Electrospray) or more having a length of approximately 10 -14 amino acids [85, 86]. The quality of the MS/MS spectra is improved as a result of the charges residing on both ends of the peptide, which enhance fragment ion coverage and improve success rate of identification. However, the mandatory digestion step is time-consuming and success depends on the

proteolysis conditions [85] and the sequence motif [86]. Nevertheless, a successful digestion for MS-compatible peptides results in low percentages of missed cleavages; around 20% or less [85]. Although trypsin is the gold standard and is well established in proteomics, other enzymes are often used to improve the proteome coverage [87]. The majority (56%) of the generated tryptic peptides are shorter than 7 amino acids [88] and do not lead to unambiguous protein identifications.

In a last step, before the pool of peptides is injected into the MS instrument, samples need to be “cleaned up” to remove salt and detergents remaining from the preparation processes [89, 90]. Moreover, the tens of thousands of peptides originating from the protein digestion [84] have to be fractionated before MS analysis. Proper and optimized sample preparation is crucial for successful, efficient and reproducible analysis [91]. State-of-the-art workflows are established to increase proteome coverage and to characterize more challenging proteins such as membrane proteins [92]. However, all sample preparation techniques introduce additional sample losses and variability. Therefore, researchers often opt to reduce sample preparation steps, time and losses by using ‘’one-pot’’ sample preparation strategies [93, 94]. Typically, stage tips [89, 95], beads [92] or nanoparticles [96] are used for sample preparation, where low amounts of sample can be digested, labeled, PTM enriched and/or cleaned up and this results in the identification of upwards of ten thousand proteins [95]. Current improvements in sample preparation minimize and automate workflows in a high-throughput fashion. Notably for samples in the clinical domain, these advances simplify and increase performance, which is highly attractive for diagnostic measurements [97].

1.2.1.2. Enrichment methods for Post Translation Modifications

Post translational modifications occur at low stoichiometry and thus in the lower end of the dynamic range [37, 62]. Consequently, PTMs need to be enriched for proteomic analyses. For phosphorylation, where phosphate groups (HPO_3) are bound to either serine (Ser), threonine (Thr) or tyrosine (Tyr) residues, enrichment can be achieved by immobilized metal ion affinity chromatography (IMAC) using Fe(3+), Ga(3+), Al(3+), Zr(4+) or Ti(4+) as metal scaffolds [98]. The negatively-charged phosphate

group interacts with the positively charged metal. However, for protein samples from complex sources such as whole-cell extracts, the selectivity of phosphopeptide affinity enrichment methods is limited due to nonspecific binding that stems from the presence of highly acidic peptides [98].

Protein phosphorylation, involved in cell signaling, regulates the normal development of proteins. It is a reversible process, where kinases (over 500 in humans [99]) mediate the esterification of a phosphate group from ATP primarily to serine and threonine residues. Tyrosine phosphorylation accounts for less than 1% of the total phosphorylation events. Recent work has highlighted the phosphorylation on histidine (His) and lysine (Lys) residues, though these modifications are much less stable in acids, which are typically used for MS analysis [100]. Erroneous phosphorylation-dephosphorylation can deregulate cascades in signaling pathways of the cell cycle and trigger several diseases, including cancer [39]. Accordingly, kinase inhibitors have gained use as potential therapeutic drugs to prevent initiation and progression of cancer [99]. Localizing phosphorylation events on proteins by standard MS analysis is challenging since 13% of peptides are phosphorylated at more than one site and/or exist as positional isomers, where two peptides have the same sequence but differ by the location of the phosphate group on the acceptor amino acid [101].

Another way to enrich PTMs is by immunoprecipitation. For example, for low abundance phosphorylation events, such as those found on tyrosine residues, monoclonal antibodies can be used [102]. To specifically enrich tyrosine phosphorylation, these antibodies can be linked to beads/solid supports and afterwards removed from the bulk and washed extensively. Other PTMs can be enriched using the same antibody based approach, where the monoclonal antibodies recognize the PTM of choice. This includes the study of SUMOylation, a Small Ubiquitin-like MOdifier (SUMO) protein that mediated the stress response [103]. Interestingly, instead of small chemical groups, as in phosphorylation, the 12 kDa SUMO proteins themselves target substrate proteins on lysine residues.

SUMOylation is a reversible covalent bond created by an enzymatic cascade where SUMO proteins are linked to target proteins via an isopeptide bond (***Figure 1-4***).

In humans, four different SUMO isoforms exist, namely SUMO1-4. Interestingly, SUMO2 and SUMO3 share 97% of the polypeptide sequence, whereas sequence identity between SUMO1 and SUMO2/3 is only 47% [104, 105]. The SUMO pathway involves different enzymatic cascades required for the transfer or removal of SUMO on their protein substrates. In humans, SUMO proteins are processed by specific sentrin/small ubiquitin-like modifier-specific proteases (SENPs). These SENPs are bifunctional and cleave/remove SUMO modifications from target proteins, but simultaneously also expose the characteristic di-glycine motif by cleaving the C-terminus residues of the non-mature SUMO. Next, the activating enzyme SAE1/2 binds to the mature SUMO through a thioester bond. This thioester occurs exclusively on the catalytic cysteine of SAE2 at Cys173. In a third step, the SUMO protein from the SUMO-SAE1/2 complex is transferred to the E2 enzyme. Unlike other organisms that can harbour multiple E2 enzymes, there is a single E2 enzyme in humans called Ubc9, that processes all four SUMO paralogues. The role of Ubc9 is to transfer the activated SUMO protein to its target protein with the aid of E3 ligases. In this last step, the *epsilon*-amino group of the target lysine will link via an isopeptide bond to the carboxy group of the C-terminal diglycine of SUMO [103]. More details about SUMOylation can be found in ***chapter 4.2*** [106].

SUMOylated proteins form, after tryptic digestion, branched peptides with a long SUMO remnant of over 30 amino acids and are consequently not compatible with mass spectrometry analysis. Furthermore, SUMOylation is a highly dynamic modification and is difficult to capture. Interestingly, increased SUMOylation occurs during cell stress [107] and reduces the number of free (or non-conjugated) SUMO proteins. Therefore, most proteomic studies are performed with cells that are undergoing stress treatments [108, 109]. Proteins that are modified by SUMO2/3 are predominantly found in the nuclear part of the cell [110]. To study the third isoform of SUMO (SUMO3), a human cell line stably expressing a functional but mutated SUMO3 was created. This mutated SUMO3 contains six consecutive histidine residues at the N-terminal that provide a convenient way to enrich SUMOylated proteins using affinity chromatography on Ni^{+2} charged resin. Moreover, to facilitate the analysis by MS, two crucial Q87R and Q88N substitutions were incorporated. The replacement of glutamine 87 with arginine provides

a five amino acid long peptide remnant on the target lysine after tryptic digestion and the substitution to asparagine on position 88 is there to distinguish SUMO isoform 2 and 3 by MS [111]. An in-house antibody recognizing the NQTGG epitope (sequence of the SUMO remnant that is produced on the SUMOylated Lys residue after tryptic digestion) is used for a second enrichment by immunoprecipitation, allowing for a peptide level purification. Using this workflow, over ten thousand SUMO sites can be identified in large scale experiments [109, 112]. Such an approach is also amenable for *in vitro* SUMO assays, where the desired substrate is added with the SUMO conjugation machinery in a microfuge tube [113].

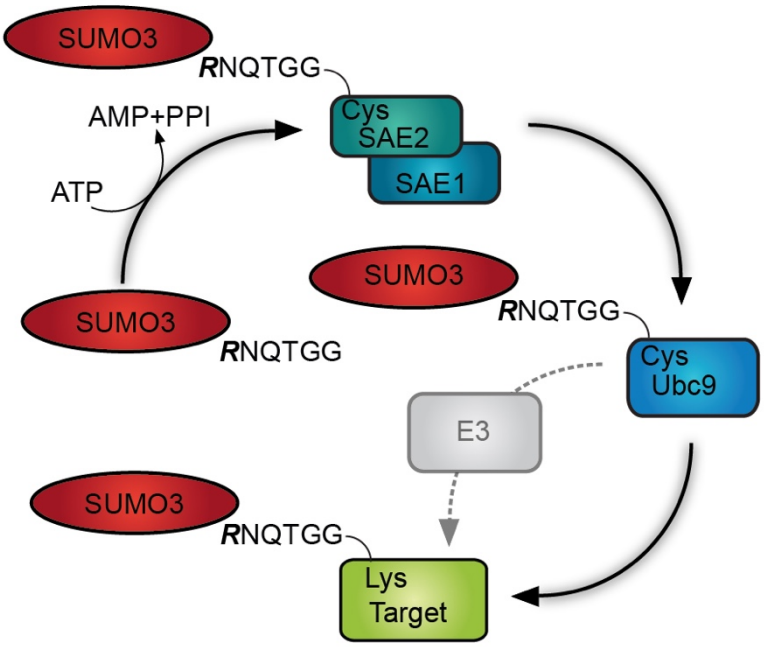


Figure 1-4: *Enzymatic cascade for the conjugation of the mature form of our in house SUMO3 mutant to target proteins. The SUMO protein is first activated by the E1 complex (SAE1/SAE2), transferred to the E2 conjugating enzyme and then transferred to the target protein by forming an isopeptide bond between SUMO and its substrates. Adapted from reference [113].*

1.2.2. Mass spectrometry instrumentation

Over the last three decades, mass spectrometry has gained wide popularity in proteomics and is used for the routine characterization of peptide composition and has replaced Edman degradation sequencing [114]. In proteomics-based mass spectrometry, amino acid sequencing can be achieved by measuring the loss of individual amino acids after dissociation from the polypeptide sequence during MS/MS fragmentation. Numerous types of ionization, ion activation methods and mass analyzers are available for peptide sequencing by MS/MS [115].

1.2.3. Ion sources

Ions can be obtained through various techniques, but the choice relies on the polarity and size of the molecule. Unlike small molecules, where harsh ionization techniques (e.g. electron impact with normally 70eV electron beam [116]) are often used, the ion source for biomolecules is, in general, a soft ionization method to form multi-protonated peptide ions $[M+nH]^{n+}$ [117]. Matrix-assisted laser desorption/ionization (MALDI) [118] and electrospray ionization (ESI) [119] are the most common ionization methods for biomolecules. These methods can overcome the challenge of analyzing large, polar, and nonvolatile molecules in mass spectrometry [117]. In MALDI, large biomolecules with molecular masses exceeding hundreds of thousands of Daltons [118] co-crystalize with the matrix. The matrix compounds are present in high excess and strongly absorb the pulsed laser radiation. Generally, Nd:YAG lasers in the ultraviolet range excite the matrix molecules, which typically are composed of aromatic moieties. Through their strong absorptivity at the excitation wavelength, localized heating sublimates the matrix and releases (desorbs) analyte molecules into gaseous form. Simultaneously, the matrix transfers protons to the peptide analytes - *''desorption /ionization''* - before the nascent ions can be accelerated through an electric field [117, 120]. During this process, analyte molecules are protected by the matrix to minimize fragmentation. Interestingly, MALDI conveniently shows a higher tolerance towards salts or other impurities, but these can still inhibit the proper co-crystallization. MALDI-

MS can be used for histological imaging of tissue slices [120]. The use of both, matrix-specific excitation molecules and high spatial resolution [121] allows for the imaging of tissue slices, due to the reduced complexity of the imaging sections. Due to the nature of the pulsed ionization, MALDI is typically coupled to a time-of-flight (TOF) analyzer [117].

1.2.3.1. Electrospray

Generally, bottom-up proteomics is conducted with electrospray ionization (ESI) [122]. Liquid chromatography, which is a necessity for bottom-up proteomics, requires an ionization technique that is compatible with molecules in solution. Pioneering work on ESI was done in the 1980s by Fenn and colleagues [119] for the ionization of biological macromolecules, although work on ESI goes back to 1745 [123]. Compared to MALDI ionization, where mainly singly charged ions are formed, ESI forms multiply charged ions. ESI ionization occurs under atmospheric pressure and is not pulsed, so it allows the continuous ionization of eluting species from the liquid chromatography separation. The outlet of the chromatographic column is connected to a narrow capillary that generates a liquid meniscus, called Taylor cone, that propagates into an aerosol spray when exposed to an electric field [122]. For proteomics, flow rates are usually in the nanoflow range (nL/min) and are referred to as nanoelectrospray (nESI) [64]. A difference in electric potential is applied between the capillary outlet and the heated MS orifice, which is usually created through a high voltage applied at the tip of the emitter. For polypeptides, positive voltages are usually applied (~2-3 kV). This is because peptides are commonly diluted in acidic aqueous solutions (the LC mobile phase) prompting the amino groups from the side-chains and N-terminus of the peptides to become protonated and therefore positively charged in solution. The tiny cationic droplets are repelled from the positive needle and are simultaneously accelerated towards the negatively charged MS interface by Coulombic forces. The solvent progressively evaporates and the charged droplets shrink to enter as solvent-free ions into the radio frequency (RF)-guided vacuum of the MS [117, 122]. Exactly how the gas phase ions are actually released from the charged liquid droplets is not clear yet, but

two different models have been proposed. One is based on ion evaporation, where more than one molecule is presumably in each droplet. When the droplet shrinks, the surface charge density exceeds the maximum Rayleigh limit and ions are liberated from the initial droplet. In the charge residue model, each droplet contains only one ion. This ion is released by solvent evaporation [122]. It should be noted that the ionization efficiency depends on the physical-chemical properties of the polypeptide composition (with enhanced ionization of hydrophobic molecules), but also on the flow rate as well as the diameter of the capillary tip [117, 122]. Low flow rates and small capillary diameters form smaller droplets with a higher surface-to-volume ratio which improves sensitivity [64, 124]. Additionally, to avoid ion suppression, it is critical to remove salts and remaining detergents during sample preparation prior to ionization [125] and thus prior to the chromatography step.

1.2.4. Mass analyzer

After ionization, the ions are guided into the mass spectrometer to be separated by the mass analyzer. Common mass analyzers include quadrupole mass filters, time-of-flight, ion traps, orbital traps (Orbitrap), and Fourier transform–ion cyclotron resonance (FT-ICR). The mass resolution depends on the type of MS analyzer, but typically resides in the range of 1,000 to 100,000 [37, 126]. The resolution is the ability to separate two species having very similar *m/z* value by one peak width measured at half-height (full-width at half-maximum (FWHM)) for the smaller of the two peaks. Resolution is defined as $m/\Delta m_{50\%}$, where m is the *m/z* ratio of the ion and $\Delta m_{50\%}$ is the FWHM [127].

All mass analyzers have different performances in terms of acquisition speed, dynamic range, mass accuracy, sensitivity and resolving power [37, 114]. Analyzers with extremely precise and high resolving power of up to one million (like FT-ICR) are slower to acquire data than those with lower resolving power. On the other hand, analyzers with high scan rates, moderate sensitivity and no upper mass limit, such as a TOF have a resolution typically less than 50,000 [126]. The high-resolution Orbitrap

analyzers have routinely been used in proteomics since their introduction and commercialization twenty years ago [128, 129]. Next to the TOF system, the Orbitrap has become the choice for proteomic analyses [130] and is commonly used for high throughput sequencing analyses.

1.2.4.1. Linear quadrupole

The linear quadrupole mass filter is the most commonly used mass analyzer. It has an efficient ion transmission, high scan speed and a modest resolution, which makes the quadrupole extremely useful [131]. Quadrupoles are continuous mass analyzers. An oscillating electric field is created through a combination of alternating RF and a static direct current (DC), applied on two hyperbolic or circular electrode rod pairs with opposite polarities. These rods run parallel and form an imaginary square in their center (***Figure 1-5a***) [132]. Ions are ejected perpendicular to the two-dimensional electric fields (the plane of the fields being in the imaginary square) and oscillate along the z-direction. On each point within the transmission axis, forces act from the x- and y-direction. These electric forces arise from each electrode pair. Presuming that the positive electrode pair is in the x-z plane and the negative pair in the y-z direction the following supply voltage is applied:

$$\pm (\boldsymbol{U} + \boldsymbol{V_{RF}}\, \boldsymbol{cos}(\boldsymbol{\omega t}))$$

Equation 1-1: *Supply voltage applied on electrode rods of a quadrupole. Each connected pair is from opposite polarity and the resulting electric field is the sum of a constant DC voltage (U) and the radio frequency, composed of the alternating voltage (V_{RF}) and the frequency ω.*

The DC voltage (U) creates an offset of the cosine wave, and the voltage for the radio frequency (V_{RF}) determines the amplitude of the function of voltages applied to two pairs of electrodes of positive and negative polarity. The voltage functions applied to each pair of electrodes are such that only selected *m/z* ions have stable trajectories and are transmitted through the quadrupole. By increasing U and V_{RF} with time, ions of

progressing *m/z* values are transmitted in turn [133]. Hence, most quadrupoles on the market work in the so-called first stability range [131] with the lowest voltage requirements.

The slope for the varying U/V_{RF} ratios determines the resolution. Increased U/V_{RF} ratio, close to the edge of the stability diagram, gives sharp peaks in the mass spectrum (see ***Figure 1-5b***) since the ion stability width is smaller [131-133]. In this regard, quadrupoles can be used to specifically transmit *m/z* of selected values. Interestingly, at constant U and V_{RF}, resolving power can be changed by the frequency rate ω [131], but typically quadrupoles are operated at ~1 MHz [133]. Their low resolution is not mass-dependent and quadrupoles are usually operated at the so-called unit resolution, meaning peaks with 1 Thomson (Th) difference are distinguishable. Regular quadrupoles are operated in the mass-to-charge range of *m/z* 1000, but can go up to *m/z* 4000. Further, quadrupoles are used either in a single configuration, in serial configuration (most commonly as a triple quadrupole) or hybrid mode [117] and can be extended to n-pols [133]. By applying only the RF voltage, broad *m/z* ranges are transmitted and are convenient as ion guides in, for example, a hexapole, but can also be transformed to collision cells or traps in MS platforms [134, 135].

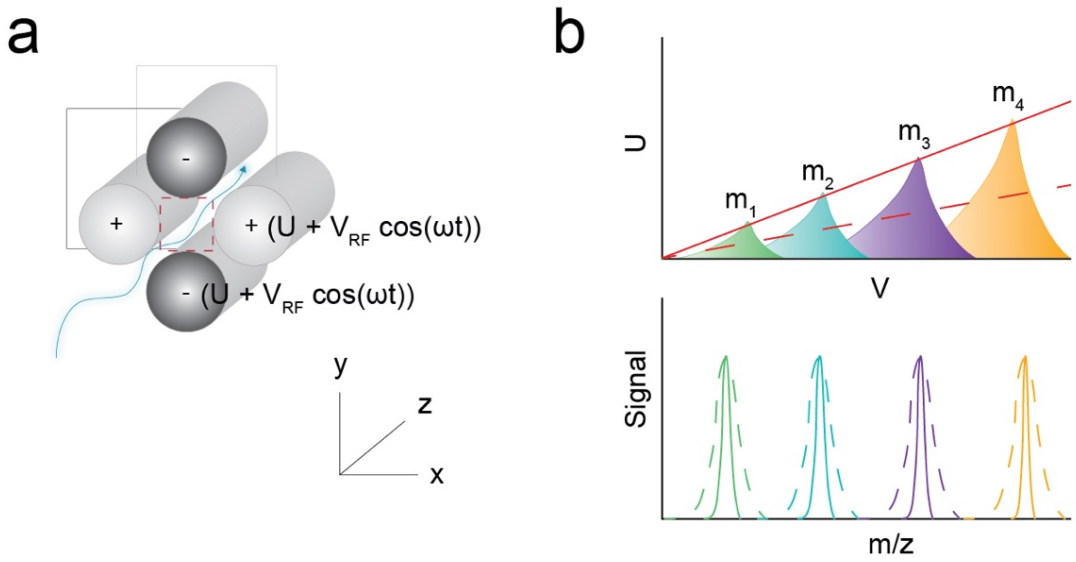


Figure 1-5: *a) Quadrupole filter for ions in an RF electrical field, generated on two electrode rods of opposite polarity. b) Transmission range is influenced by the DC and RF and determines the peak shapes.*

1.2.4.2. Linear trap quadrupole

Based on the Mathieu stability diagram, ion motion in the x-y plane in an RF quadrupole can be guided within the two-dimensional plane by homogeneous electric fields [133]. To use a quadrupole as a trap, one can simply apply at the entrance and exit of the electrode rods a stopping voltage, known as constant end cap voltage (DC end-caps), and the quadrupole can store ions. Once the kinetic ion movement along the central z-axis is null, ions must be trapped along the radial axis. The limits of the ion cloud can be outwardly defined by the applied RF field. This system is denoted as a linear ion trap (LIT). Importantly, the two-dimensional electric fields in the ion trap are not to be confused with the three-dimensional trapping type published by Wolfgang Paul [135, 136]. Accordingly, we distinguish *3D quadrupole ion trap* and *linear trap quadrupole (LTQ)* operated under a 2D electric field, which was introduced by Thermo Fisher Scientific [137]. One major advantage of the 2D linear trap is the higher ion capacity, compared to the limited volume of 3D traps, where space-charge effects arise when too many ions are contained within the trap. By simply making the electrode rods longer, more ions can be stored, increasing the injection efficiency [136]. The implications of the space-charge effect were overcome by introducing the automatic gain control (AGC), which limits the number of ions transferred to the trap [136].

After trapping and accumulation of a stable ion beam, ions are ejected through a slit along the center of one of the hyperbolic rods [137]. By increasing the RF amplitude (*resonance excitation ac voltage*), smaller *m/z* ions are resonantly excited, increasing their kinetic energy, and are successively ejected from the linear ion trap. Therefore, ions with increasing *m/z* values can be ejected sequentially by increasing the RF amplitude and detected by the electron multiplier outside the trap.

1.2.4.3. Orbitrap

In addition to the radiofrequency ion traps, proteomic experiments are typically performed on high-resolution Orbitrap instruments with mass inaccuracy below 5 ppm [127]. The Orbitrap was developed by Makarov but the physical concept dates back to

the 1920s [128]. Compared to the dynamic oscillating electrical field in the quadrupole, the Orbitrap confines ions in a logarithmic electrostatic field. Ions orbit axially back and forth along an inner, spindle-like electrode and are contained by a barrel-like outer electrode, split into two halves (***Figure 1-6***). A DC voltage is applied between the two electrodes to trap the ions.

Ions are tangentially injected to the quadrologarithmic electrical field slightly off-axis. The ion cloud is electrostatically attracted to the inner electrode, where at the beginning a voltage ramp is applied. The ion package, however, is injected with a small time delay (50-90µs) after the voltage is applied [138]. Ions get captured inside the Orbitrap and start their circular trajectory around the inner electrode. By increasing the electrical field, ions get closer to the z-axis and are electro-dynamically squeezed, resulting in a smaller radial orbit trajectory. Although the radial frequency varies dependent on the initial parameters, the ions move along the z-axis in a harmonic back and forth motion, where their harmonic frequency (rad/sec) is proportional to their *m/z* [138, 139].

$$\omega = \sqrt{\frac{k\,z}{m}}$$

Equation 1-2: *Axial frequency ω of trapped ion rings inside the Orbitrap, where k is the field curvature constant, m is the mass and z is the charge.*

The detection of the ion cloud starts when the voltage ramp is finished and the electrical field stays static. The axial motion of the ions in the orbital induces an image current (opposite current) on each half, which is registered on sensors on the symmetric Orbitrap electrodes. Since the ion motions are *m/z* dependent, the ion current for the different *m/z* species is recorded as time-domain transient. The transformed analog-to-digital signal is subsequently processed by Fourier Transformation (FT) to obtain the mass spectrum [138]. The sophisticated splitting of the outer electrodes in two mirrored halves allows to double the signal because of the "*Ping-Pong"* motion of the ion.

For Orbitraps, the resolving power is defined by the capability to distinguish between different frequencies of the phase-separated ion rings. Because ions are

oscillating in a harmonic wave on an axis in an electrical field, the resolution is inversely proportional to $m/z^{-1/2}$ according to the following relation:

$$\frac{\mathbf{m}}{\boldsymbol{\Delta m_{50\%}}} = \frac{\boldsymbol{\omega}}{\mathbf{2\Delta\omega_{50\%}}} = \frac{\mathbf{1}}{\mathbf{2\Delta\omega_{50\%}}}\sqrt{\frac{\boldsymbol{k\,z}}{\boldsymbol{m}}}$$

Equation 1-3: *Frequency resolving power* $\omega/\Delta\omega_{50\%}$ *is twice the mass resolving* $m/2\Delta m_{50\%}$ *[138].*

Acquiring ion motion in the Orbitrap for extended periods gives a higher resolution. As an example, in the Orbitrap Elite model commercialized in 2007 (the instrument used in ***chapter 2-4***), transient time for 240,000 resolution at *m/z* 400 required 786 ms [139].

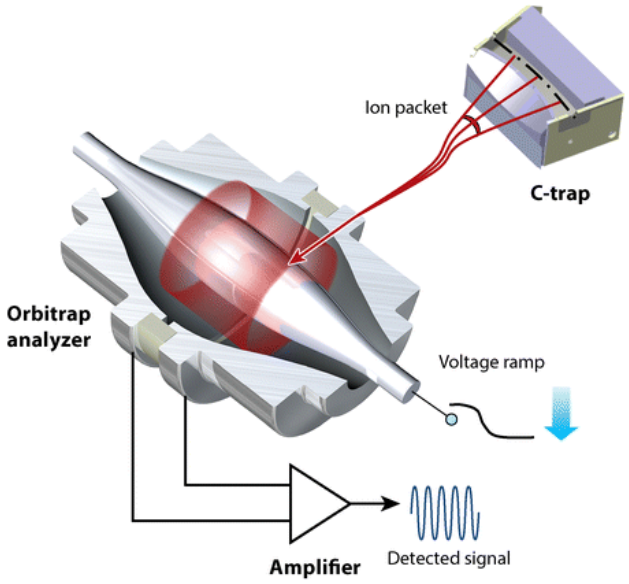


Figure 1-6*: Ion packages are pulse wise injected into the Orbitrap and electrodynamic squeezed. The axial back and forth frequency along the spindle axis induces an image current on the outer electrode halves and is Fourier Transformed to a mass spectrum. From reference [140].*

A crucial feature for the Orbitrap is the external source of the pulsed ion beam. Electrospray is a continuous source and should not actually be a compatible ionization method for the Orbitrap. Pulsed ion packages are measured one after the other and

require roughly 1s for data acquisition, storage, and processing [127]. However, ions can be externally accumulated, stored and released in a short-pulsed manner. To this end the C-trap was created. This bent RF-only trap allows the accumulation of ions based on the AGC targets, proper ion injection to the Orbitrap and minimal transfer loss [136, 141]. The C-trap is filled with a bath-gas for collisional cooling of the ions. To release ion packets, an RF voltage is quickly ramped down, since ions in an oscillating electrical field move towards the lower electric field. To form a focused ion beam, ions are electrically accelerated towards the Orbitrap.

The Orbitrap series of MS instruments has revolutionized high-resolution mass spectrometry and brought accurate and sensitive proteomic analyses to an affordable level compared to FT-ICR MS. Resolving powers of >150,000 at *m/z* 200 are achieved with precise mass accuracy for unambiguous spectral assignments. The mass-to-charge range is typically between *m/z* 200 to *m/z* 2000 [127]. Interestingly, this significant improvement in analytical performance comes from a small-sized analyzer (diameter approximately of a one euro coin) and shows high space-charge capacity [139, 141] with constant scan speeds and performance improvements in the latest generations [18]. In addition, improved computational algorithms like enhanced FT (eFT) [140, 142] or phased spectrum deconvolution methods (ΦSDM) [143] provide better resolution with shorter transient times and improved spectral quality.

1.2.5. Tandem mass spectrometry

As briefly mentioned above, structural information is obtained from peptide sequencing by tandem mass spectrometry (MS/MS). Peptide precursors are dissociated into fragments in a collision cell. Fragment ion information can then be used to unravel and annotate the sequence. Only fragments holding a charge from the initial (multi) charged precursor can be seen in the MS/MS acquisition and fragment ions are labeled according to a nomenclature first introduced by Roepstorff *et al*., and further elaborated by Biemann [144]. Assuming a doubly protonated peptide ion is split into two parts, and one charge is located on the N-terminus, the other on the C-terminus, we can observe

both complementary fragments. By default, products with a charge on the N-terminus are named with the prefix *a*, *b*, or *c*, whereas their charged C-terminus counterpart fragments correspond to *x*, *y*, or *z* (see ***Figure 1-7***) [145]. Accordingly, the sum of the complementary pairs is equal to the number of amino acids (*n*) in the intact polypeptide chain. For bottom-up proteomics, dissociation methods with specific cleavages on peptide bonds are highly deployed and result in the typically *b*-*y*-ions [66]. The mass shift between neighbor-fragments from the same fragment type corresponds to a loss of one amino acid or their PTM modified form. For ideal sequence coverage and unambiguous identification, both counterparts should be detected. Moreover, nothing prevents the peptide from breaking into several fragments to form internal fragments, or low *m/z* fingerprints from individual detached amino acids, the so-called immonium ions [146].

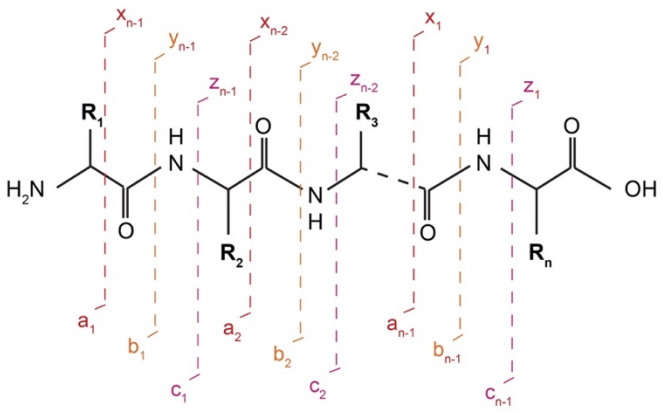


***Figure 1-7*:** Nomenclature for peptide fragment ions *for a peptide with n- amino acids. Adapted from reference* [145].

Machine learning tools like MS^{2}PIP are able to predict theoretical *in silico* intensity spectra for the most common fragments [147]. By integrating such predictor tools successfully in the database search, false discovery rates decrease and the identification improves, as shown recently in the Wilhelm lab [148]. It should be noted that amino acid composition and charge state can have a major impact on the resulting MS/MS spectra. In particular highly charged peptides result in a higher diversity of fragment ions and make the spectral assignments more difficult [149, 150].

1.2.5.1. Ion dissociation mode

Different ion activation methods are available to promote the formation of fragment ions. Fragmentation can be induced by electron interactions, absorption of photons or gas collisions [145]. Collision-induced dissociation (CID) and the Orbitrap-specific higher-energy collisional dissociation (HCD) mode are commonly used in proteomics. Distinct *b* and *y* ions are generated, where with HCD the fragmentation is further extended to some *a* and *x* ions (CO loss from *b*-type ion). Other, less used, fragmentation modes exist like electron transfer dissociation (ETD) that is compatible with larger biomolecules such as intact proteins that have higher charge states, where the release of *c* and *z* ions is from broken N–C_α bonds. Ultra-violet photodissociation (UVPD) fragmentation has been developed, where *a* and *x* ions are produced from C_α–C bond cleavages next to the *b*,*y,* and *c*,*z* ions [130, 145].

For low energy CID and higher energy HCD, the accumulation of internal energy upon collisional activation leads to the dissociation of the labile C-N peptide bonds. This occurs when the ion cloud enters the collision cell or the trap during the MS analysis. Upon collision with target gases such as nitrogen or helium, a portion of the kinetic energy of the precursor ion is converted into internal energy which is transferred to all oscillators of the activated ion resulting in the cleavage of the more labile bonds.

CID activation is the most popular fragmentation method and is routinely used in qTOF, triple quadrupoles and ion traps. The types of fragment ions produced, however, are partially instrument-dependent [151]. CID in the ion trap leads to neutral losses of ammonium, water, and CO_2 and exclusion of low *m/z* ions [145]. The missing structural information of immonium ions and small *b*,*y* ions was overcome on the Orbitrap mass analyzer with the introduction of HCD [152]. Olsen *et al*. introduced the higher-energy C-trap dissociation. Instead of using the C-trap to store ions and send them as packets to the Orbitrap, they increased the RF voltage to induce fragmentation. However, this led to lower trapping efficacy of small *m/z* ions. To solve this issue, Thermo built the next generation of Orbitrap with an octopole collision cell, which is maintained at a pressure of 5 mbar under nitrogen and is aligned next to the C-trap. RF voltage could exceed the former limitation in the C-trap [140, 152]. Despite the slower data acquisition

rate compared to low-resolution CID ion trap analyses, HCD provides complete, no mass cut-off, high-quality MS/MS spectra. The impact of the missing low *m/z* for clear peptide identification is more drastic for shorter peptides [151]. HCD is now included in the hybrid ion trap/Orbitrap instruments in addition to the Q-Exactive series [140]. Overall, the different fragmentation modes provide complimentary data that can be used together to garner a greater coverage of the proteome [145].

1.2.6. Hybrid mass spectrometer

Recent instrumental progress made it possible to combine multiple mass analyzers together to improve instrument capabilities. These hybrid instruments drastically evolved in speed, flexibility, sensitivity and merged different analyzers into one platform to combine their best attributes in one instrument.

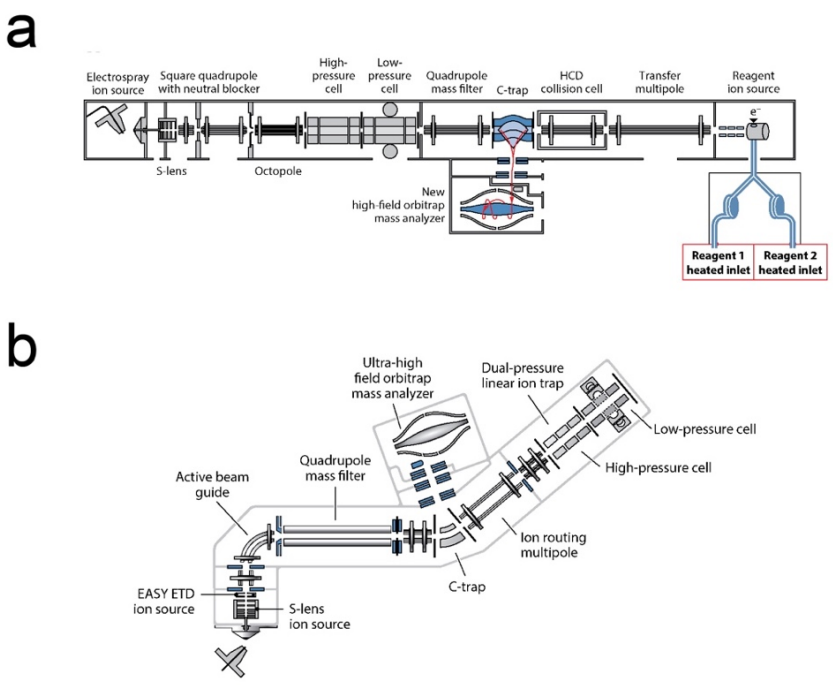


Figure 1-8: *Scheme of the hybrid Orbitrap a) Elite and b) Fusion joined with the parallel ion accumulation and detection mode. Adapted from reference [140].*

One of these MS hybrid platforms is the MS series from Thermo Scientific, where high resolution and accurate mass from the Orbitrap meets the speed and ion detection sensitivity from the linear ion trap mass analyzer [140]. In this work, we took advantage of the Elite linear ion trap Orbitrap mass spectrometer (***Figure 1-8a***), introduced nearly a decade ago [153] and the later launched Orbitrap Fusion tribrid mass spectrometer (***Figure 1-8b***) [17, 154]. Both platforms offer the flexibility to use different dissociation mechanisms (HCD, CID and ETD [155]) for enhanced reliability and comprehensiveness of the analyses.

The **Elite** Orbitrap is an optimized version of the initial commercialized hybrid ion trap-Orbitrap platforms and like its Velos predecessor [156], it has enhanced ion optics, a fast performing dual ion trap and a smaller Orbitrap. The Elite instrument is equipped with an Orbitrap with a higher field strength to gain better resolution. The improved field was obtained by making the Orbitrap smaller (diameter of the outer electrode was reduced from 30 mm to 20 mm with inner diameter of 10 mm instead of 12 mm). In combination with the eFT algorithm a resolution of 240,000 at *m/z* 400 is achieved with 768 ms transient time, which is the time needed to produce a MS spectrum from when the ion enters the Orbitrap. On the other hand, lower resolutions require less time (a resolution of 15,000 has 48 ms transient time). Scan rates for the Elite are between 4 and 10 scans/sec, depending on the acquisition mode. MS and MS/MS scans can be recorded either in the Orbitrap-Orbitrap, or the faster iontrap-iontrap ‘’low-low’’ mode or by parallelizing the high resolution/accurate mass (HR/AM) MS strategy that combines precursor analysis in the Orbitrap and the MS/MS analysis in the ion trap. The later ‘’high-low’’ mode benefits highly from the accurate mass and charge state determination in the survey scan, and from the use of the ion trap for the product scans to adjust as close as possible to the continuous source of ions from the ESI. With short full scans, injection times for MS/MS are predicted based on the precursor intensity in the prescan (predictive AGC or pAGC) [140, 153, 157]. CID fragmented ions can be measured in the ion trap, or transferred to the C-trap for eventual analysis in the Orbitrap. Ions dissociated via HCD are exclusively measured in the Orbitrap [153].

The Orbitrap **Fusion** tribrid mass spectrometer merges 3 mass analyzers in one instrument: quadrupole mass filter with hyperbolic rods for higher robustness and transmission up to 3400 Th isolation, a fast scanning dual-pressure linear ion trap, and Orbitrap. Identical to the LTQ Elite in function, the Fusion is equipped with CID and HCD. The architectural structure was improved and the linear high-low pressure ion trap was positioned behind the C-trap (Q-OT-qIT). This architecture allows more flexibility for HCD, since the HCD fragmentation products can routinely be measured in the ion trap in the Fusion [158]. The Fusion platform achieves much faster scan rates than the Elite Orbitrap and can go up to half a million resolution. MS^3 was improved with the synchronous precursor selection (SPS) option, enabling to select multiple fragment ions for the third MS spectra [17, 140, 154].

Despite high scan rates in the ion trap, high-quality spectra are generally prioritized, such as in the studies performed in chapters 2 to 5. Mainly because confident identifications aid in assigning PTM in an unambiguous fashion ***(chapters 3 and 4)***. Moreover, MS^2 based quantification strategies rely on high resolving power ***(see chapters 2 and 5)***.

1.2.7. Analysis mode

Different instrument methods are available to analyze samples, depending on the end goal. The different methods use varying MS/MS acquisition modes that are triggered in unique manners [159]. In most LC-MS/MS injections, where only a few samples need to be analyzed, individual peptides are selected in an untargeted manner, where the most intense precursor ions are selected for MS/MS analysis. If several samples that are similar in composition are to be analyzed, a combination of a more unrestrained acquisition and the use of spectral libraries can be suitable. On the other hand, if the analysis is aimed at studying specific targets, a list of precursors of interest can be used.

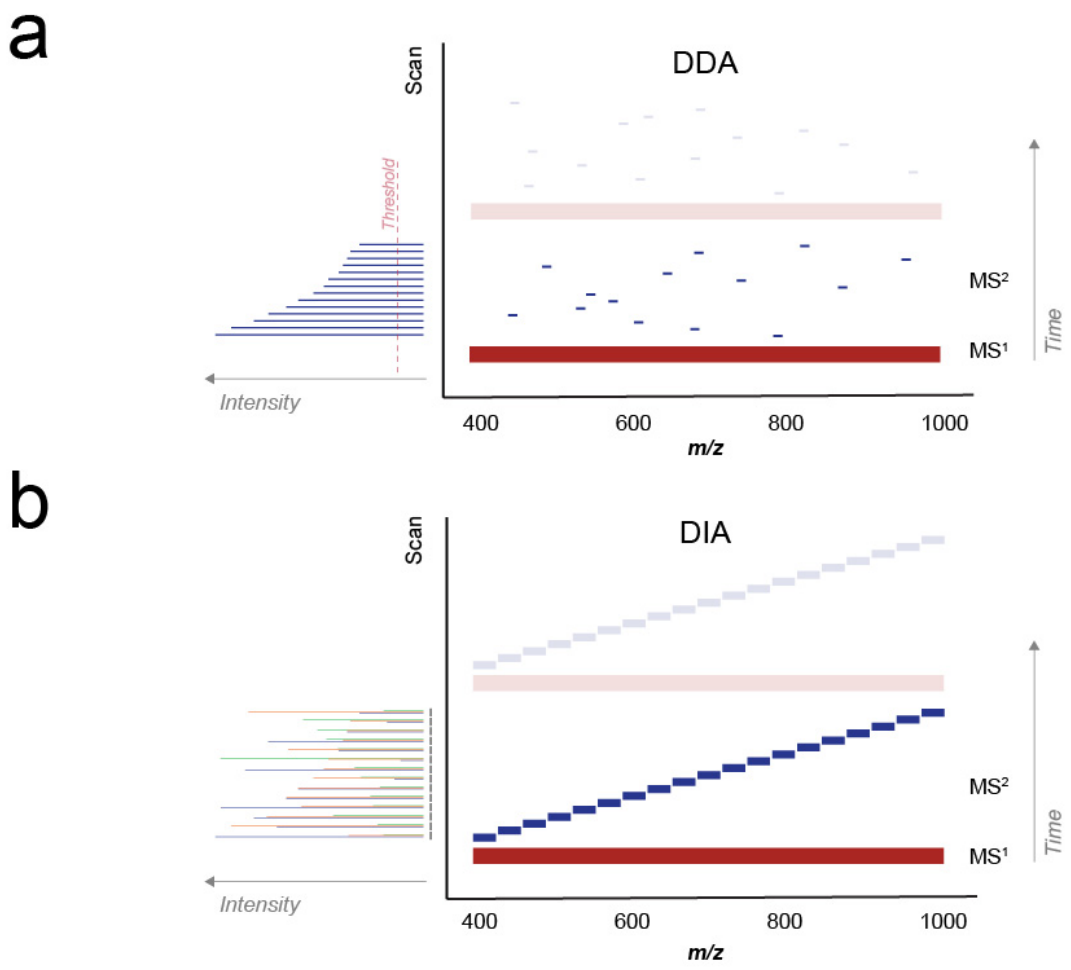


Figure 1-9: *Illustration for a) data dependent acquisition with small precursor isolation window compared to b) data independent acquisition, where multiple precursors are fragmented at once.*

1.2.7.1. Data Dependent Acquisition (DDA)

Data-dependent acquisition (DDA) is routinely used in the discovery phase and in most proteomics studies, also known as the *TopN* method [18, 160]. During one cycle, defined as a certain number of product scans after the initial survey scan, maximal *''N''* MS/MS scans can be taken. On newer instruments (Fusion and Lumos), instead of defining the number of MS/MS scans, one can also define the duty cycle time. In this so-called *Top Speed* mode [140], the instrument takes as many MS/MS scans as possible during a restricted time, but follows the given AGC target or maximal injection

time parameters. Either way, limiting the number of MS/MS scans allows a sufficient number of MS^1 or precursor during the elution profile of the precursor.

During DDA, the selection of precursor ions for MS/MS fragmentation occurs according to their intensity [66]. The instrument algorithm selects in real-time the highest abundant precursor first, followed by decreasing intensity ions. As long as the precursor is in the measured *m/z* range, there is absolutely no bias in terms of the *m/z*. However, some restrictions apply. In bottom-up proteomics, normally only precursors with charge states higher than +1, but lower than +5/+6 are selected. Reason for this is that tryptic peptides are mainly doubly or triply charged and the singly charged ions mostly emanate from chemical noise. Moreover, to prevent constant re-sequencing of previously fragmented precursors, dynamic exclusion lists are generated, and enable the acquisition of peptide ions of lower abundance [18, 161, 162]. To obtain a MS/MS spectrum of good quality, a minimal number of ions are required, but the analysis should not take too much instrument time. Therefore, a minimal precursor intensity has to be achieved before MS/MS is triggered (***Figure 1-9a***) [161]. Together with a minimal AGC target, an appropriate collision energy is needed to generate sufficient amounts of fragment ions to increase the likelihood of assigning a spectrum. Elaborate method options help enhance identification rates. Stepped or charge-dependent collisions energy [18] and automated extension of maximal injection time as far as the parallelization mode allows (Orbitrap – iontrap), are integrated into the latest Orbitrap Tribrid™ platforms.

For the subsequent fragmentation, precursor selection relies on a small isolation window around 1 Th. With this, the inclusion of isotope patterns can be ensured while limiting the selection of interfering ions and co-fragmentation [162]. Unfortunately, in regular injections without rigorous pre-fractionation, the dynamic range typically covers ~3-4 orders of magnitude and prioritizes mainly the most abundant peptides. Low abundant peptide ions are frequently selected in the isolation window along with other peptide ions resulting in chimeric spectra [162, 163]. In addition, many of the low abundant ions do not even have the chance to be selected in the enormous and dense pool of potential species. Peptides with unassigned charge states that result from overlapping isotopic envelopes are automatically excluded from the MS/MS candidate

list, despite underutilization of the MS. Significant improvements to overcome this problem were achieved in the Orbitrap platforms with the real-time advanced peak determination (APD) algorithm [164], resulting in nearly complete use of the available sequencing events.

Alas, another general consequence of the stochastic sampling is limited reproducibility between replicates [18, 165]. This can result in MS/MS spectra of varying quality for the same selected peptide ions leading to variable sequencing events [161]. High scan rates (around 40Hz) with Orbitrap MS, with a good balance between speed and quality [18, 164], improve this number considerably. In contrast, ingenious DDA acquisition strategies try specifically to avoid re-sequencing the precursors from previous runs, aiming to catch low abundant peptides that would otherwise be excluded from MS/MS sequencing [161]. In any case, a strong benefit of DDA is the universal application field. All types of analyses, from identification to quantification, can be measured in this mode. Regarding this thesis, analyses were exclusively produced in DDA mode.

1.2.7.2. Data Independent Acquisition (DIA)

Unlike DDA, MS/MS spectra obtained in data independent acquisition (DIA) arise from all precursor ions from selected *m/z* windows regardless of their intensity. Moreover, the MS^1 scan in DDA is not essential for DIA between continuous acquisitions of ion fragments. Nevertheless, it is recommended to add MS^1 scans between cycles for precursor information during data analysis [166]. Basic methods are built out of scan cycles, where the anticipated *m/z* scan window is divided into wider windows of *m/z* ~10-50 for the sequential fragment acquisitions [18, 167]. Highly complex and convoluted ion fragment spectra result from the multiplex precursor co-fragmentation. The variable isolation width [166, 168] has an impact on spectral complexity. However, the cycle time for serial fragmentation of subsequently adjacent isolation windows should be limited to generate an extensive set of fragment ions (***Figure 1-9b***). A typical cycle time of 2-4 s permits sufficient iterative MS scans to generate enough data points per peak. DIA requires an unbiased data collection for

peptides and a complete digital proteome archive, but analysis becomes sophisticated. Information needs to be extracted and matched against previously generated, high-quality spectral libraries containing ion transition information [168]. DIA was introduced in the early 2000s [169], and in the traditional analysis workflow, ion fragments and precursor elution profiles were extracted and merged together prior to running the classical database search. Many DIA variants have evolved since, with intrinsic variations stemming from data acquisition to data handling [166, 168]. One frequent approach comes from the Aebersold lab, introduced in 2012 as Sequential Windowed Acquisition of All Theoretical Fragment Ion Mass Spectra (SWATH‐MS) on the AB Sciex Triple TOF [170, 171]. Small DIA windows or SWATH, are combined with targeted data extraction of the most abundant products. The elution profile of fragments is generated [172] and becomes a target analysis in a large-scale format in a lab-independent manner [171]. Although DIA could not keep up with DDA in the early 2000s high resolution/accurate mass platforms, recent advances in MS instruments [168, 173], improved bioinformatics tools [172] and a statistic filter [174] for large-scale data sets make it now more enticing [175]. DIA outperforms DDA for short LC-MS runs in terms of reproducibility, dynamic range and proteome depth [18, 176]. However, the preparation of sample-specific, high-quality libraries is time and sample consuming and it must be carefully considered whether the pre-preparation time is appropriate with the number of samples. Also, unknown mutations and PTMs are impossible or challenging to identify by DIA. Finally, DIA is not applicable to MS^n quantification strategies for multiplexing.

1.2.8. Quantitative proteomics

In the recent years, interests in qualitative proteomics have shifted towards quantitative information. It is now routine to study what proteins are present and at what expression level [177]. Furthermore, being able to compare different conditions or samples gives important insights for the understanding of the dynamic nature of the proteome. In addition, not only protein abundance, but also PTMs changes and profiling these changes have gained much interest [178, 179]. Simply comparing absolute ion intensities between different samples is unfortunately not the ideal solution due to sample variation. Inconsistent protein extraction (subcellular dependent) and digestion efficiencies, technical fluctuation during sample preparation [94], injection and ionization, as well peptide-dependent ionization properties [180] and inconsistent instrument performance make it unfeasible to compare protein concentrations and copy numbers, respectively.

Insofar as the target is known, heavy peptides can be synthesized and spiked-in at various concentrations in the sample to be analysed. Synthetically generated peptides with stable isotopes incorporated exhibit the same physical and chemical properties in chromatography and MS, and can be used to directly quantify the endogenous light form. *Ab*solute *Qu*antification (AQUA) peptides [181] are normally added before LC-MS analysis, but one can also include proteolysis with, for instance, synthetic *concat*amers (QconCAT) proteins [182]. Simplified, for quantification, SRM/PRM transitions or extracted ion chromatograms (XIC) of light/heavy peptides provide the absolute concentration [183]. It is becoming clear that the necessary pre-knowledge, the cost-intensive nature of synthetic peptides, and high throughput limitation make it un-suitable for use on a daily basis. Relative quantification is more accessible and straightforward for discovery phases and thus the preferred option. To obtain relative protein changes between samples, one can use label-free or isotopically label-based strategies that are further divided into metabolic, chemical, or enzymatic labeling [177, 184]. In bottom-up proteomics, label free, metabolic or chemical labeling are the approaches that are most commonly used [185]. Instead of absolute

quantification, as with AQUA, a relative quantification that is expressed as x-fold change between two conditions is computed [82].

1.2.8.1. Label free

The universal, less costly and most simple approach is label free quantification (LFQ). The steady development of MS performance in terms of scan speed and enhanced peak capacity by pre-fractionation clearly promote its usefulness [177]. LFQ can be based on the spectral count. Protein copy number per cell is correlated with the spectral number and provides semi-quantitative information [186], but this relies on spectra acquisition and can be biased by several factors. In addition to spectral counting, the chromatographic elution profile can provide valuable information and can furthermore be turned into an approximatively absolute scale (see “Proteomic Ruler” [187]).

LFQ relies on comparing the MS^1 abundance (intensity or area) of the same peptide ion in different injections. Ideally, precursors are assigned based on their identifications. However, this is not always the case due to the nature of the DDA, where all the same peaks are not selected for MS/MS sequencing across several instrumental replicates [18, 165]. Alternatively, powerful algorithms assign features by first aligning the retention time across all injections prior to matching *m/z*, charge state and retention time to features and attributing it the identification that was provided from at least one injection [188, 189]. LFQ is especially attractive in preliminary proteomic studies to find potential targets during a discovery phase, but also to verify if specific experimental treatments worked. Moreover, it is pretty straightforward because no special sample preparation is needed and all types of samples can be analyzed. A good example of the power of LFQ is its use in clinical tissues samples or body fluids [190]. However, it is more a semi-quantitative analysis and serves as a rough estimate for expression differences. Further, the LFQ approach suffers from higher variations that are introduced during sample preparation and/or by instrumental fluctuations (e.g. variability in chromatography or ionization). In addition, no multiplexing can be done, meaning every sample has to be injected individually and leads to longer instrument use.

Software algorithms face many challenges: low abundant or short proteins with limited MS information, shared peptide sequences and peptides subdivided in multiple fractions lead to more challenging feature assignments as well more difficult calculations for the total signal value [191].

1.2.8.2. Chemical labeling - Tandem Mass Tags

An approach to decrease instrument time is multiplexing several samples together by chemical labeling. One of the first methods was isotope-coded affinity tag (ICAT) introduced by Gygi *et al*. [192]. The structure of the ICAT is based on a tag containing a functional group reacting with the side chain of cysteine, an isotopic linker part and biotin to isolate the labeled peptides by avidin affinity chromatography. With this method samples are analyzed in a pairwise fashion. The tags make the two samples chemically identical, so that peptides of the two conditions co-elute simultaneously in the LC. However, the heavy form of the tag is 8 Da larger and can therefore be distinguished by MS/MS.

Nowadays another popular chemical labeling method is Tandem Mass Tags (TMT). In TMT, peptides or proteins are labeled with a tag, containing three parts: (**1**) a functional group reacting with primary amines of amino acids, (**2**) a balance group to equilibrate the total weight of the tag and (**3**) a reporter part, made up of different isotopes to obtain different signals in the MS^n spectrum and to be able to distinguish the different samples [193]. Synthetic analogues to isobaric tag for relative and absolute quantitation (iTRAQ) are also available [194], where the signal from the release of the reporter ions in the tandem spectra is used. For the commercialized TMT tags from Thermo Fisher Scientific, the reactive group is an N-hydroxysuccinimide (NHS) ester, which reacts with the epsilon-amino group of lysine residues or primary amines on the N-terminal of a peptide [195]. Besides primary amines, the NHS ester has been documented to react with serine, tyrosine, and threonine carboxyl groups to form covalent bonds [196]. With TMT, one can multiplex up to eleven (recently up to sixteen with TMTpro [197]) conditions at the same time [198] using the HR/AM resolving power for tandem mass spectra, capable of distinguishing the mass difference of 0.00632 Th

between different substitutions of ^{13}C and ^{15}N atoms within the tags [199, 200]. After labeling and combining the samples together, the different taged peptides all have the same mass and appear as one peak in the full MS. In order to see the signals for the different conditions, the peptides have to be fragmented to generate the tandem mass spectra. In this step, the TMT tags will lose their low *m/z* reporter groups and one can read out the quantitative information in the tandem MS. Acquisition in the Orbitrap shows the low *m/z* reporter ions with few interfering ions beside the reporter ions in the *m/z* 126-131 range [151]. By having the quantitation done in the MS^n level, the sample and MS spectral selectivity is improved and can therefore be used with highly complex samples. Further, the abundance of the peptides in the MS^1 is the summation of all individual conditions. Therefore, the sample amount for each condition can be decreased. Or keeping the same amount per channel, the TMT sample can alternatively be pre-fractionated for MS analysis. Combined with extensive fractionation, TMT gives a deep quantitative proteome coverage [201]. The fractionation helps to simplify the spectral complexity and leads to less co-isolation and co-fragmentation. Despite the above-mentioned advantages, TMT has some drawbacks. A well-known problem is the MS^2 co-selection of isobaric, co-eluting precursors. The co-fragmentation gives chimera scans and the signal from TMT reporters are mixed and become incorrect [202]. A way to overcome this problem is to use synchronous precursor selection (SPS-MS^3) [203]. First, the MS/MS will be measured in the ion trap with CID fragmentation and permits the identification of the peptide sequence. In the following tandem MS (MS^3), the most abundant peptide fragments are re-fragmented and measured in the Orbitrap to obtain the low *m/z* reporter ratios for relative quantification. The TMT quantification becomes more accurate and precise, however, the duty cycle decreases because of the multi-level measurements [204]. Indeed, for every feature an MS^2 is needed for identification and an MS^3 is required for the quantification. Moreover, if one of the spectra is of poor quality, both spectra cannot be used since both pieces of information are vital in the analysis. To address this, a recent improvement was suggested by the group of Gygi. This method uses an on-the-fly database search for the MS^2 to select only the precursor ions for MS^3, which were effectively identified as a peptide in first tandem MS [205, 206].

Although co-fragmentation occurs, TMT harbors advantages over LFQ due to improvement in PTM quantitation precision and reduced technical variations that may arise during LC-MS injections. Extensive studies of sample preparation procedures and label amount make the TMT workflow robust and more cost-efficient [207]. Further, TMT was adapted to create the Single Cell ProtEomics by Mass Spectrometry (SCoPE-MS) method that relies on using a channel-specific amplifier [208]. Unrestricted application of TMT is attractive for clinical samples, where high throughput, but also low sample amounts are the main challenges [201, 209]. In addition, only a few values in each multiplex are missing and it gives a more complete proteome coverage for the individual samples. However, one has to consider that there is a loss of peptides/proteins in inter batch analysis due to the nature of DDA mode, where the same precursors are not all selected for fragmentation in each batch [210].

1.2.8.3. Metabolic labeling - stable isotope labeling by amino acids in cell culture (SILAC)

The first method of choice to minimize variations during sample preparation and to improve instrumental reproducibility is stable isotope labeling by amino acids in cell culture (SILAC). It was introduced almost two decades ago [211] for mammalian cell cultures, but is nowadays expended to even animal organisms that are fed with specific labeled amino acids as part of their diet [212]. It is a highly accurate and robust method with the simple idea to replace almost exclusively the two amino acids lysine and arginine with their isotopically labeled forms. Taking advantage of trypsin, the most widely used enzyme for bottom-up proteomics [63], there is always at least one lysine or arginine in a given peptide sequence. Therefore, depending on the growing conditions, one can distinguish the different forms of lysine and arginine by mass spectrometry. For the sample preparation, cells of different conditions are simply grown in medium spiked with either normal Arg0/Lys0 (light), [$^{13}C_6$] Arg6/[$^{2}H_4$] Lys4 (medium) or [$^{15}N_4{}^{13}C_6$] Arg10/[$^{15}N_2{}^{13}C_6$] Lys8 (heavy) lysine and arginine isotopes. This allows to distinguish the different SILAC isotopomer peptides inside the mass spectrometer. After a few cell doublings in SILAC medium, proteins have fully incorporated the isotopically labeled

amino acids. Normally samples can be multiplexed with up to three conditions after cell collection and are then directly combined for sample preparation [177, 213, 214]. The lysis and following sample preparations steps are performed with one sample, which significantly reduces all the variations induced by sample handling. Moreover, other key benefits include the reduced instrument time and the lower impact of instrument robustness because SILAC doublets and triplets elute at the same time and are also measured in the same survey scan. Ionization fluctuations become negligible, which is convenient for better PTM quantification on the peptide level [214]. SILAC was successfully implemented for protein turnover measurements by temporally changing the cell culture medium to another SILAC channel [215, 216]. Notwithstanding, a serious weakness of this quantitation method is its limited use to cell culture samples. However, scientists have found a solution to partially overcome this problem with a method called super SILAC [217]. In this methodology, an internal standard containing heavy proteins from a cell line culture is spiked into the proteins that are extracted from tissue sections or specimen fluids. Various samples can be compared by using the same standard and the same amount of spiked-in across samples. Unfortunately, if some proteins are not expressed in cell culture or do not contain certain sequential mutations, no quantification can be performed. One drawback of SILAC-based quantification is the instrumental resampling of the isotopomers of the same peptide that increase spectral complexity, leading to a lower proteome coverage and dynamic range. In addition, the cost of isotopically labeled amino acid that are added to the cell culture medium are expensive.

1.2.9. Data analysis

Dealing with the tremendous number of acquired spectra would be impossible without automated tools. Powerful processing is involved in the data analysis for identification and quantification. Fortunately for bottom-up proteomics, MS/MS assignment for tryptic peptides is quite predictable and thus straightforward. Characteristic fragments for HCD or CID show mainly *b*/*y* ions and diagnostic peaks, including *m/z* 147.113 (lysine) or *m/z* 175.119 (arginine) for the y_1 of the C-terminus and

also immonium ions [218]. Peptide identification strategies are divided into two sub-branches, the *in silico* database search methods and the *de novo* sequencing. *In silico* algorithms rely on genomic databases that try to attribute to each acquired MS/MS spectrum a computationally generated one that is deduced from the proteome database [145]. Frequently used *in silico* algorithms include Andromeda (integrated in MaxQuant) [219], Mascot [220] or SEQUEST [221] (the two latter are both available in Proteome Discoverer) [222]. UniProt is one of the main sources for public proteome databases [223] and is freely accessible. Tens of millions of sequences, the majority stemming from genome sequencing projects from different species or strains, are translated *in silico* into canonical proteins. These proteins are further divided into UniProtKB/Swiss-Prot and the far larger UniProtKB/TrEMBL databases. Swiss-Prot proteins are manually curated and controlled, whereas TrEMBL proteins are unreviewed entries. Current databases follow restricted translation rules, for instance small ORFs are omitted since they are often arising from so-called non-coding RNA (ncRNA). Some non-canonical proteins, however, have shown important functions in the cell [224]. Recent online sources explore the potential of alternate splicing and non-canonical proteins for the extended generation of more complete databases [225], but it is a controversial subject in terms of the actual gains in identification [226].

The second method, *de novo* algorithms, do not require the pre-knowledge of databases and show some advantages for less known proteomes. Peptide sequences are derived from the MS/MS spectra. Obviously, better mass accuracy and comprehensive fragmentation positively influence the success of peptide spectrum matches (PSM). PEAKS software [227] integrates *de novo* sequencing in the classical database search to achieve maximal identification [228].

With the help of different scoring functions, algorithms calculate the probability of the sequence match and provide a spectral score. Score ranges vary from engine to engine, but all of them consider ion patterns for sequence coverage as well as default background noise [218-220, 227]. Better matches are expressed with higher scores. In addition, statistical significance for identifications is obtained using false-discovery rates (FDR) [229], which are generally set to 1% [222]. This calculation method can be applied against a decoy database that contains reversed/shuffled sequences for

classical database search. As result, a list of true and false positive (TP or FP, respectively) matches are generated. Since the reversed or scrambled sequence from the decoy database shows lower spectral probability-scores, one can set a score cut-off for a corresponding FDR, when plotting the score distribution of PSM for decoy and target databases. In this context, 1% FDR means simply, that the proportion of false positives out of the total matches does not exceed this threshold [230].

$$FDR = \frac{FP}{FP + TP}$$

Equation 1-4: *FDR calculation with TP for true positive and FP for false positive matches, usually set at 0.01.*

Software like PEAKS and MaxQuant are high-performance and can deal with large datasets for identification. They offer mass recalibration, an improved algorithm for chimeric spectra assignment, and are able to deconvolute and assign sequences from co-fragmented peptides (second peptide search algorithm [219] in MaxQuant pipeline). In addition to their database search capabilities, both softwares are also equipped to run a labeled and label free quantification [231, 232]. Interestingly, MaxQuant was initially developed for SILAC experiments [232, 233]. In addition, they include a broad spectrum of supplementary features. For instance, MaxQuant provides advanced precursor information by calculating the quantitative fraction of co-fragmentation (precursor ion fraction or PIF) [162]. PEAKS, on the other hand, offers the additional options to identify sequential modifications [234] or non-specified PTMs [235].

1.2.10. Sample complexity of protein digests

Proteomic samples are by nature inherently complex and require high performance platforms to screen at least the most abundant species. The enormous population of proteoforms within a widespread dynamic range over orders of magnitude makes it challenging [37, 74]. Analyzing the proteolytic peptides rather than proteins

increases the complexity by another several orders of magnitudes and results in hundreds of thousands of peptides from a single proteome [236, 237], whereby PTM information should be ideally conserved. State-of-the-art MS instruments, such as the Fusion Orbitrap, are built to offer fast acquisition rates exceeding 20 Hz [17], but MS/MS identification rates are not 100%. Spectra of low quality due to low signal-to-noise, poor fragmentation, co-fragmentation, isomers or inadequate bioinformatics parameters (e.g. undefined PTMs, incomplete database, nontryptic peptides) hamper the success of peptide identification [162, 163, 238-241]. Co-selection and fragmentation occur regularly in shotgun proteomics and show a significant impact on MS/MS identification rates [162]. The proportion of chimeric spectra resulting from these co-eluting precursors with similar *m/z* are estimated to affect nearly half of all MS/MS spectra [163]. Incidentally, this also applies to non-separable isomeric peptides. Their chimeric MS/MS spectra have fragment patterns from all isomeric forms [240]. Decreasing the isolation window reduces the overall proportion of co-fragmentation [162], but also reduces the number of identifications. Increasing the isolation window above 2 Th also decreases identification rates. Predictions suggest that narrowing the isolation window to 0.2 Th would lead to 10% chimera scans, but this is currently unfeasible due to instrument limitations [238]. Additionally, chimeric scans negatively impact isobaric labeling based quantitation [242] leading to ratio suppression [202] and reduced dynamic range [185]. Although MS^3 quantitation method using SPS improves the accuracy of measurements [203], many labs still prefer using MS^2 based quantification to maximize the number of quantified peptides and proteins [18, 201, 207]. Moreover, using the APD option increases the number of chimeric scans and has a negative impact on TMT quantification, and should therefore be deactivated [243].

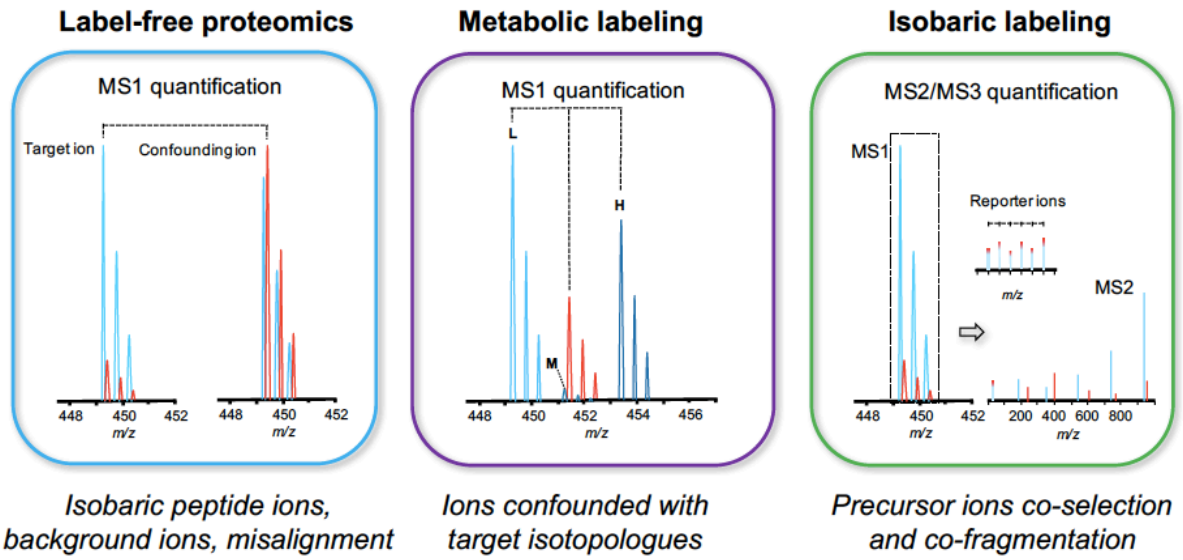


Figure 1-10: *Co-elution of similar m/z ions from complex samples negatively impacts all three main quantification methods that are used in proteomics: For the MS1 based methods (label free and SILAC) target ions can be misaligned with cofounding ions when matching ions between injections. For MS^2 based isobaric labeling, co-fragmentation result in chimera scans and reporter ratios are disordered.*

One issue in proteomics comes from the overwhelming number of features present in complex samples and the limited charge space capacity in the ion traps. Merely 1% of all ions make it to the trap, where a bias to highly abundant ions has been observed [244]. Low abundant species that are normally suppressed can be preferentially transferred with smaller *m/z* windows. Combined with label free quantification, multi-segmented acquisition can increase the number of quantified PSM by 50%, but cannot outperform sample pre-fractionation [245]. Multi-segmented acquisition benefits from well characterized reference samples as exemplified by Meier *et al*. with their *"*BoxCar*"* method that uses 16 small *m/z* windows. Indeed, with this method the authors quantified >10,000 protein in a 100 min LC elution window by aligning their elution profile to the reference sample and matching the features [244]. However, the fact that more ions are detected and are "visible" to the MS entails that more co-elution occurs from isobaric ions. Ultimately, multi-segmented acquisition methods do not really decomplexify the MS/MS spectra, and so the suboptimal peak capacity remains a hurdle.

1.2.10.1. Chromatographic separation

Shotgun proteomics includes, per definition, a high-performance liquid chromatographic separation prior to MS analysis [246]. A suitable separation of highly complex peptide mixtures is crucial and contributes significantly to a comprehensive proteome analysis. In the standard configuration, reversed-phase (RP) chromatography is coupled to the MS [66] and can accomplish great in-depth analysis in single-shot injections [247, 248]. The separation is based on the hydrophobic interaction of alkyl chains with the peptides and the affinity towards the mobile phase. The acidified water-acetonitrile gradient contains 0.1% of formic acid (FA) where ion pairing between the formate anions and positively charged peptides enhance their retention on reverse phase column and improve loading capacity [249]. Although trifluoroacetic acid (TFA) could serve as an ion pairing agent, it is less compatible with the downstream MS analysis since it lead to ion suppression during the ESI [250].

The use of C_4 and C_{18} long alkyl chains for protein and peptides separations, respectively, has been well established over the years [251]. Chromatographic performance is influenced by column length, diameter, temperature, and the properties of the stationary phase (e.g. particle size, surface, monolith or particles). Although in theory proteomic analysis would benefit from longer columns, in practice the added length creates high backpressures. Also, proteomic studies are limited to certain types of stationary phase, limiting the peak capacity of a columns [252, 253] (peak capacity defines the maximal number of features resolved in a given analytical column [254]). Frequently used columns are around 20-30 cm long and packed with 2-3 µm beads of 100-300 Å pore size [255, 256]. Low flow rates in the nanoliter/min-range ensure an efficient ionization and better sensitivity for low sample amounts, and reduce the backpressure.

With the steady increase of MS scan speed, a highly efficient separation becomes even more important [257]. Ion suppression, resulting from co-eluting peptides, can reduce the proteome depth [162]. Sharper chromatographic peaks are needed to minimization redundant resampling of the same precursor ions. Sharper peaks lead to less co-fragmentation and provides sufficient precursors, which lead to a more complete

use of the MS/MS scan rates and duty cycle [257]. Since many labs prepare in-house column, proper pressure during column packing is mandatory to maintain peak capacity. Peptides of low abundance benefit from higher packing pressure, yielding narrower peaks, thus increasing the relative intensity. With adequate packing pressure, the stationary phase is uniformly dispersed within the column and the mobile phase can traverse the column with a constant flow. As a result of proper packing, the backpressure of the LC decreases [255].

1.2.10.2. Multidimensional peptide fractionation

An alternative approach to increase peak capacity is rigorous pre-fractionation. Peptides are separated based on their physical-chemical properties prior to LC-MS, to provide more ions for the different features and thus gain better sequence coverage, PTM and isoform information, and to uncover peptides of low abundance. Initially, the most common multidimensional separation technique was strong cation-exchange (SCX) followed by RP and was referred to multidimensional protein identification technology (MudPIT). This allowed to cover peptides/proteins ranging over four orders of magnitude in abundance [258]. In SCX, the peptides are first separated based on charge. The acidified peptides are applied on a negatively charged stationary phase and subsequently eluted with increasing salt or pH fractions. Fractions containing peptides with a similar charge or isoelectric point (pI), are then separated based on their hydrophobicity in RP prior to online MS detection [249]. Peptide fractionation techniques have evolved continuously since, including the use of hydrophobic interaction (HILIC) and isoelectric focusing (IEF) technics for multidimensional separation [37, 249]. Nowadays, concatenated high-pH (HpH) RPLC in combination with ion-pairing RP is the preferred approach [259]. Compared to low pH RP, the negative charge on the C-terminus and the side chain of aspartic and glutamic acid in HpH are neutralized, eliminating the electrostatic interaction with the stationary phase. The separation is based on pure hydrophobic interaction [259]. The orthogonal separation in multidimensional fractionation is important to prevent the loss of resolution that was acquired in the first dimension [260]. Therefore, if there is a lack of orthogonality,

fractions can be concatenated to guarantee orthogonality for beneficial comprehensiveness. Although SCX shows a better orthogonality to RP than HpH RP [259], combining the early, middle and late HpH RP fractions together improves chromatographic capacity and show low fraction overlap [236, 261]. The Olsen group devised a strategy for exhaustive HpH RP fractionation without concatenating fractions by shortening the MS acquisition time with very short gradients of 15 minutes. Regardless of the lower peak capacity per run, by injecting 46 fractions they achieved peak capacity over 10,000 proteins, which represents to-date the deepest proteome coverage [236]. More extensive studies have even used an additional third dimension of separation. The 3D-LC approach provides two orders more in terms of dynamic range than the conventional 1D setup [237]. However, whilst identifications increase, the need for larger sample amounts and instrument time have significant practical implications [237]. Interestingly, the separation prior ESI is nearly always ion-pairing RP because of the better suitability in ESI ionization [236, 237, 258, 261].

1.3. Ion mobility spectrometry (IMS)

Sample preparation, systematic analysis approaches, instrumentation and artificial intelligence have evolved tremendously in the last decades to facilitate proteomic analyses. One such advance was in the field of ion mobility mass spectrometry. There are still hurdles that cannot be overcome with conventional fractionation, MS instruments, and computational approaches. Mostly, there is a need for reproducible analyses of low abundant proteoforms in a more automated fashion [239, 262]. For this reason, an enhanced interest in ion mobility for the analysis of complex biological has developed [263], as instrumentation became more user-friendly, and can be integrated into state of the art MS instruments. The main MS manufacturers, including Thermo Scientific Fisher, improved their ion mobility devices in recent years [264-267].

1.3.1. Fundamentals in IMS

In modern mass spectrometry, ions are separated based on their mass-to-charge ratio in an electric field. In comparison, the separation in classical drift ion mobility is based on shape, ionic charge rather than on mass. Thereby, high-resolution IMS can separate isomers [268, 269]. Ions are separated based on their mobility in an electrical field where a counter current flow of neutral gas is applied. This mobility can be expressed through the ion mobility coefficient K(cm^2 V^{-1} s^{-1}), which is defined by the quotient of the ions constant drift velocity (v_d, cm s^{-1}) and the electric field E (V cm^{-1}).

$$\boldsymbol{K = \frac{v_d}{E}}$$

Equation 1-5: *Ion mobility coefficient K(cm^2 V^{-1} s^{-1}) is based on the ion velocity and the electric field. The velocity is based on the drift time, d_t(ms range) and the distance of the drift tube.*

Equation 1-5 is valid, as long as the resulting velocity caused by the collision energy of the ion swarm with gas is not higher than the velocity associated with the

thermal motion at constant gas atmosphere, pressure, and temperature. If these conditions are fulfilled, then this situation is defined as thermalized ion swarms. The ion mobility coefficient is proportional to the charge state q divided by the collision cross section Ω at an effective temperature T_{eff} (sum of ion thermal energy and acquired energy in the electric field, in Kelvin) and the square root of the reduced mass μ.

$$K \propto \frac{q}{\Omega_{T_{eff}}\sqrt{\mu}}$$

Equation 1-6: *Ion mobility coefficient is dependent on the reduced mass μ, the charge $q = z \cdot e$ (product of number of charges z, and the electron charge e ,1.6022 × 10^{-19}C) and the $\Omega_{T_{eff}}$ collision cross section.*

The collision cross section (ion shape/size profile) affects friction with the gas molecules. Thus, gas number density N (particles/cubic centimeter) has to be considered. The following model was deduced to express the linear ion motion in an electric field through gas [270].

$$K = \frac{3q}{16N}\frac{(1+\alpha)}{\Omega_{T_{eff}}}\sqrt{\frac{2\pi}{\mu k T_{eff}}}$$

Equation 1-7: *Mason-Schamp [270] equation to express the relationship between collision cross section and ion mobility coefficient under thermalized conditions, correction factor α and the Boltzmann constant k.*

Ions with higher charge states show higher ion mobility than lower charge state ions with the same reduced mass and collision cross section. Moreover, velocity increases at higher temperatures at constant pressure. Likewise, the pressure shows a major impact on the system. Since the ion motion is influenced by temperature and pressure, the ion mobility coefficient K is normalized to 273 K and 760 torr to express the reduced mobility K_o:

$$K_o = K \frac{273.15}{T} \frac{P}{760}$$

Equation 1-8: *Reduced ion mobility coefficient K_o (cm^2 V^{-1} s^{-1}) is normalized to ambient pressure and zero Celsius.*

This reduced ion mobility depends on the electric field strength E (V/cm) divided by buffer gas density N, to express the reduced electric field. By definition, the reduced electric field $\frac{E}{N}$ is given by units of Townsends (Td) with 1 Td = 1 x 10^{-17} V cm^2 [271].

$$Td = \frac{E}{N} 10^{17} V\ cm^2$$

Equation 1-9: *Reduced electric field in ion mobility is expressed in Townsends (Td) units.*

Reduced mobility is constant up to 50 Td (low-field region), where ions are thermalized. Inside an intermediate-field region (~50-300 Td) the reduced mobility still increases with higher Td, but not in a linear relationship. In the subsequent adjoining high-field region, reduced mobility decreases with increasing Td. In this region, ions are not thermalized and so the electrical field has a stronger effect than the thermal energy received from the buffer gas [272]. The linear ion mobility device operates in the low field (< 30 Td, ~7,500 V/cm) [273], where ions are separated based on their different reduced mobility in a tube. In this classical IMS (***Figure 1-11a***), the electrical field attracts the ions towards the detector, and move against the drag forces arising from the drift gas based on their collision cross sections. This IMS was originally known as plasma chromatography/gaseous electrophoresis [270].

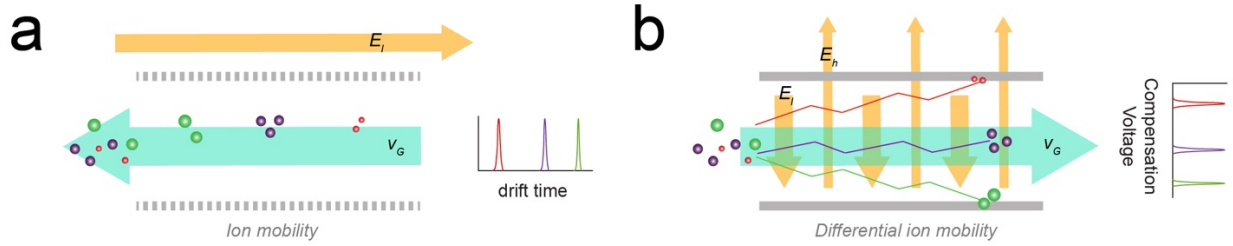


Figure 1-11: *Drift gas (turquoise) and electric field (orange) act differently in a) classical drift time IMS (K_l) compared to b) DMS (K_h/K_l). In IMS (a), all ions are temporally separated, whereas in DMS (b) ions are spatial separation and only a selected portion of the ions are transmitted. Adapted from reference [274].*

In contrast to IMS, ions can also be separated based on their difference in ion mobility in the high K_h and low field K_l, by the so-called differential ion mobility (DMS) (***Figure 1-11b***). In this circumstance the electric field acts perpendicular to the ions. Thus, ions move in towards the electrodes, but because a gas flow passes simultaneously through the DMS device, ions are pushed towards the exit and can be transmitted. High Field Asymmetric Waveform Ion Mobility Spectrometry (FAIMS), the central method in this thesis, falls in this category and will be discussed shortly in more detail. However, different separation methods evolved and are regrouped in low-field *linear methods*, and *nonlinear methods*, where ions partially reside in a high electric field [271].

1.3.2. Linear methods

The most common low-field ion mobility IMS technics are Drift Tube IMS (DTIMS), Travelling Wave IMS (TWIMS) and Trapped IMS (TIMS) [275]. The IMS separations can be achieved in static (DTIMS) or dynamic (TWIMS; TIMS) electric fields.

1.3.2.1. Drift Tube IMS (DTIMS)

DTIMS (***Figure 1-12a***) is the simplest IMS in a low electric field with drift gas collisions, meaning the collision cross section can be calculated from the Mason–Schamp equation (***Equation 1-7***) and requires no calibration. With ring electrodes along the drift tube, a uniform static electric field is generated (1 to 15 Td) and accelerates ions against the direction of the gas flow [271]. Good resolution (R > 100-200 Ω/ΔΩ) can be achieved but depends on tube length, pressure, and type of gas [276]. DTIMS must be held in the low field, but can be operated under atmospheric pressure or low pressure [275]. Since the number of ions transmitted from the ion source is relatively low (low duty cycle), newer generations have integrated ion storage tunnels to accumulate ions during the on-going IMS separation [263].

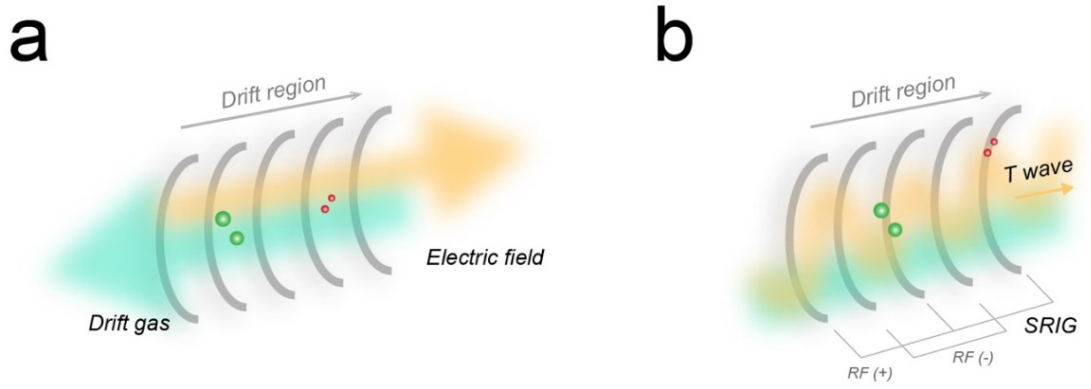


Figure 1-12: *Scheme of a) the classical drift tube IMS with static electric field and b) the dynamic waveform IMS TWIMS. Adapted from reference [263].*

1.3.2.2. Travelling Wave IMS (TWIMS)

TWIMS (***Figure 1-12b***) drift tubes operate at higher pressure (0.5 – 1.0 mbar) [277] and show greater ion transmission. Instead of a static low-field, a waveform-like electric field is created on the ring electrodes by applying opposite phase RF voltages in a time-shifted manner on adjacent ring electrodes to create traveling voltage wave (T-wave). The complete set of electrodes is referred to as a stacked ring ion guide (SRIG), and ions surf on the T-wave towards the tube. Ions of greater mobility surf along the waves, thus arrive faster to the end of the SRIG as low-mobility ions. The ions from

lower mobility roll over the waves and often slip behind the wave creating a slower axial velocity [276, 278]. To enhance the duty cycle and improve ion utilization, ions are also trapped and accumulated, providing a better sensitivity [277]. Because of the dynamic field, the correlation between K_o and Ω needs to be calibrated for a given condition (e.g. wave height and frequency, gas pressure and type). Note, the TWIMS device is filled with gas, but does not flow in the opposite direction of the drift field as for DTIMS. Still, ion shape determines the collision rate with the gas molecules. In addition, the gas impacts the resolving power of the instrument [273]. TWIMS is incorporated in the Synapt IMS/mass spectrometry system from Waters [277, 279] and temporally separates the injected ions with acquisition times of around 100 ms.

1.3.2.3. Trapped IMS (TIMS)

One novel low-field IMS system is TIMS [280] commercialized by Bruker. In TIMS, the IMS device is divided into two regions: an ion trapping tunnel for parallel ion accumulation with nearly 100% efficiency (***Figure 1-13***) [281], and an IMS tunnel that separates packets of ions. Unlike DTIMS, the drift gas in TIMS follows the same direction as the ion velocity direction. Ions are pushed by the gas in proportion to their collision cross section, whereby ions of larger Ω move faster than those of smaller Ω. Simultaneously, RF voltages create an electric field gradient (EFG) that opposes the drift gas and immobilize the ions. To elute the temporarily confined ions, the magnitude E_{t0} of the initial axial EFG can be decreased and ions with higher mobility K (charge-dependent) escape and are transmitted first [282]. The resolution is not limited by the length of the tube, but by the scan speed of the TIMS (EFG gradient release) and can exceed 250 [283]. Since IMS takes place in the hundred milliseconds range, it can be optimally coupled to fast and mass accurate qTOF instruments (~ 100 µs). The separated ions are displayed on a heat map of the ion mobility expressed as $1/K_0$ drift time *versus m/z* [265].

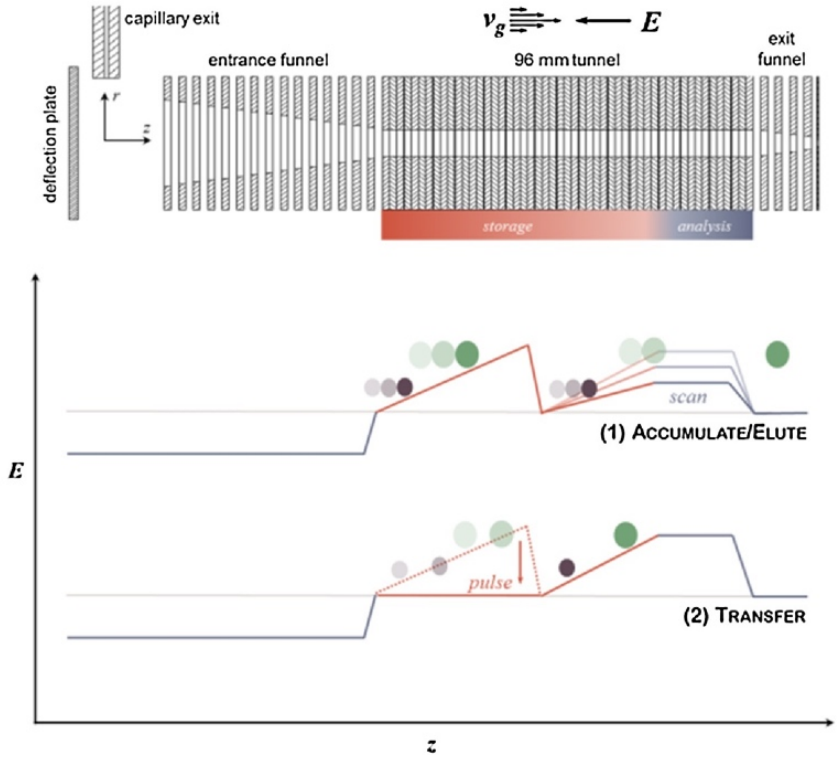


Figure 1-13: *Parallel accumulation of ions with nearly 100% duty cycle efficiency and subsequent transmission of separated ions from the TIMS device. Adapted from reference [281].*

1.3.3. Nonlinear methods

Unlike linear methods where $K \neq f(\frac{E}{N})$ with $v_d = K \times E$ is field independent, differential mobility spectrometry (DMS) or field asymmetric ion mobility spectrometry (FAIMS) lie outside the thermalized ion condition and the ion mobility principles no longer apply. The ion mobility in the high field K_h becomes field-dependent and an additional alpha function for the field strength needs to be considered:

$$\boldsymbol{K_h(^E/_N) = K(0)\left[1 + \alpha\left(\frac{E}{N}\right)\right]}$$

Equation 1-10: *Nonlinear function describing the high- field mobility K_h coefficient. With $K(0)$ being the mobility coefficient in zero field, and the nonlinear electric field dependence with alpha function $\alpha\left(\frac{E}{N}\right) = \alpha_2\left(\frac{E}{N}\right)^2 + \alpha_4\left(\frac{E}{N}\right)^4 + \cdots \alpha_{2n}\left(\frac{E}{N}\right)^{2n}$.*

1.3.4. Field Asymmetric Waveform Ion Mobility Spectrometry (FAIMS)

FAIMS is a differential ion mobility based method, where ions travel with a carrier gas and oscillate between two electrodes to which are applied low and high electric fields (***Figure 1-11b***). FAIMS acts as an ion filter, transmitting only a part of the ions. FAIMS gas-phase ion filtering is not based on mass-to-charge as is the case with the quadrupole filter, rather ions are selectively transmitted based on mass, shape, clustering between ions with the neutral gas molecules and dipole moment [274]. FAIMS can transmit selected ion populations based on their mobilities at low and high electric fields. Moreover, with FAIMS, the collision cross section cannot be determined and selective transmission parameters have to be determined empirically. Interestingly, the alpha factor (***Equation 1-10***) shows less correlation with the ion mass than low-field IMS Ω and therefore possesses a better orthogonality to MS. Therefore, FAIMS could be used in combination with IMS [275]. However, FAIMS suffers from lower resolution (~10-20) compared to IMS [284].

Notwithstanding, the ion-path is mainly controlled by the different behaviors of the ions in the high field during the application of a high/low electric field waveform. The integral for the altering time in high-field (short) and low-field (longer) must be zero (***Figure 1-14a***), if not the ion-path would be affected by the delta between the two electrical fields.

$$(\boldsymbol{E})_{\boldsymbol{h}}\boldsymbol{t}_{\boldsymbol{h}} + (\boldsymbol{E})_{\boldsymbol{l}}\boldsymbol{t}_{\boldsymbol{l}} = \mathbf{0}$$

Equation 1-11: *The sum under the curve for the high field and low field is zero. Typically the time in the low field is twice as long as for the high field and depends on* $(E)_h$ *and* $(E)_l$.

The asymmetric high/low field waveform is applied only to one electrode and the other electrode is grounded (***Figure 1-14a***). Moreover, the amplitude of the high field is called dispersion voltage (DV). Ions enter the electrodes and stray off the x-axis because of the perpendicular electric field (see purple arrows in ***Figure 1-14a***). The

offset depends on the differential ion mobility in the high and low field. If the displacement by K_l (or K) can be balanced by K_h, ions are transmitted. If this is not the case, a sub DC voltage on the asymmetric waveform electrode, known as compensation voltage (CV), refocuses ions with increased/decreased ion mobility K_h back to the center. Thereby, the average ions return back between the electrodes gap and are successfully transmitted. Basically, CV represents the difference between K_l and K_h. If the difference is bigger, higher CV voltages need to be applied to refocus the ions. However, at CV=0V, the net displacement of K_l and K_h is also zero.

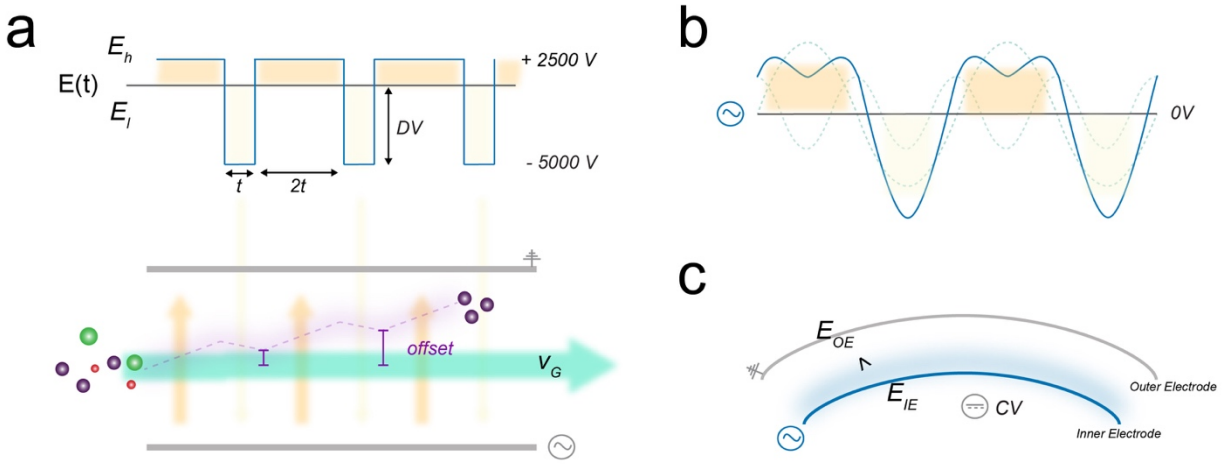


Figure 1-14: *Ion separation using FAIMS. a)The net displacement of ions along the x axis (purple bars) is due to the difference of ion mobility in the high versus low field is arising from the asymmetric waveform of the electric field and can be corrected by a DC compensation voltage on the inner electrode. b) The time dependent high/low field is the sum of two sinus waves. c) For FAIMS cylindrical electrodes, DV and CV is applied on the inner electrode. The temperature difference of the electrodes affects the electric field gradient, resolution and ion transmission. Adapted from references [274, 285].*

FAIMS was initially introduced in Russia and was further developed in North America by two separate groups. One focused on the planar FAIMS and was commercialized by SCIEX as SelexION, the other researchers at the National Research Council of Canada focused on the curve-shaped electrodes. Cylindrical FAIMS was commercialized by Ionanalytics and later acquired by Thermo Fisher Scientific (TFS) [286]. The focus of this thesis is on the current cylindrical FAIMS electrodes from TFS.

Therefore for the remainder of this thesis, the term FAIMS refers explicit to the cylindrical electrodes and excludes the planar version [287] and the micro version [288] ultraFAIMS, available from Owlstone Medical Ltd. Ultimately, all of the instruments work by the same gas phase ion separation principle, but vary in resolution (planar FAIMS), speed (ultra FAIMS) and sensitivity (cylindrical FAIMS) [289]. It was found that cylindrical electrodes can focus better ions in an atmospheric non-homogenous electric field in contrast to the planar version that requires a homogenous field (more ion diffusion), providing a greater gain in sensitivity [290, 291].

For FAIMS, the waveform shape should ideally be rectangular, as shown in ***Figure 1-14a***, but due to instrumental limitations this is not feasible [292]. Therefore, the waveform resembles a bisinusoidal wave (***Figure 1-14b***) [290]. Taking the latest FAIMS build as example (with a two times faster frequency ω as the previous FAIMS), a DV of -5000 V with a low field amplitude of 2500 V is generated by the sum of two 90°C phase-shifted sinusoidal waves of 3 MHz (amplitude 1.67 kV) and 1.5 MHz (amplitude 3.33 kV), respectively [285]. Notably, an ion transmission time of greater than 100 ms could be decreased to 40 ms in the newest FAIMS, partly because of the faster frequency waveform. This had an important impact for the integration of FAIMS with fast-scanning MS instruments [198, 266, 267].

FAIMS separation occurs between a 1.5 mm gap between the outer (grounded) and inner (asymmetric waveform) electrode (***Figure 1-14c***) using nitrogen as a carrier gas and an electric field of ~150 Td at DV -5000 V. However, a non-homogenous field of $(\frac{E}{N})_{OE} < (\frac{E}{N})_{IE}$ creates a dispersion field gradient (***Figure 1-14c***). The bigger the delta gradient, the narrower the CV ion transmission bands, the sharper the CV peaks and the better the resolution, which comes at the cost of lower sensitivity [285]. This gradient is affected by the electrode temperature. When keeping the outer electrode constant at 90°C, but changing the inner temperature (90°C, 70°C, and 50°C), the delta $\left[(\frac{E}{N})_{IE}\right]$-$\left[(\frac{E}{N})_{OE}\right]$ decreases along with decreasing of the ion transmission. Keeping the outer and inner electrodes at 90°C improves ion transmission by providing the maximum field strength due to lower gas density N, and higher local dispersion field [293]. Also, higher

DV increase the electric field and ion transmission. Under atmospheric pressure, better ion signals are achieved due to the enhanced ion focusing at higher $\frac{E}{N}$ [294]. On the other hand, if the dispersion field increases due to DV changes, the CV will shift and needs to be adjusted to transmit the desired ion population [285, 293, 295]. This is because the larger K_h increases the difference in high/low field mobility and therefore higher voltages are required to refocus ions. Nonetheless, since ion populations are selectively focused based on CV, one has to scan over the whole CV transmission range to transmit all ions in turn. Unfortunately, this is time-consuming and not high-throughput. Another issue is that proper ion transmission through the electrodes must occur under atmospheric pressure. A substantial amount of ion loss was reported at nano flow rates, as FAIMS was initially developed for larger flow rates [296]. FAIMS was only available for the older generation of Orbitraps until very recently. For more than a decade no fundamental changes were done to improve FAIMS technologies. The old FAIMS setup suffered from slower MS instrument performance than the current state-of-the-art MS platforms. Tremendous improvements in ion transmission, parallelization and scan rates were achieved in the last decade for MS instruments [18, 140, 164], but loss of ion transmission [296] and limitations in speed [297] with FAIMS hindered its wider use in proteomics.

1.3.4.1. FAIMS evolution in bottom-up proteomics

The potential use of cylindrical FAIMS for proteomics was first shown in 1999 when Purves and Guevremont coupled ESI to FAIMS [294, 298]. FAIMS could be used as a low-resolution ion filter able to separate peptides of different charge states [298]. Later, short peptides as well as cytochrome C were directly infused into FAIMS. Purves observed a separation of high – low masses based on the CV and obtained spectra of reduced complexity by reducing background from solvents and ESI ion clusters, leading to a better signal-to-noise (S/N). With these benefits, they extended the linear range of detection for leucine encephalin by 3 orders of magnitude [294]. Guevremont *et al*. showed that FAIMS could separate interfering isobaric peptides from tryptic hemoglobin digest from pig [299] and further demonstrated this application for thirteen proteins and

a complex mixture containing six-proteins [300]. The CV showed poor correlation with *m/z,* but FAIMS provided good orthogonality when combined with MS [299]. Enhanced signal-to-noise ratio of one order of magnitude in FAIMS provided a convenience improvement for CID-based MS/MS and facilitated spectral assignments compared to regular ESI-MS/MS [301]. Considering that the infusion of peptides was analyzed over a limited CV transmission range and that the sample coverage was adequate, ESI-LC-FAIMS-MS became an attractive avenue for proteomics [300]. The Guevremont group also investigated the separation of different conformers of bovine ubiquitin and found a partial correlation between protein cross sections and CV for these proteins [302, 303].

To cover more CVs in fewer injections, CV stepping was introduced within the same LC-MS run. FAIMS reproducibility and its ability to remove chemical background, in the form of singly charged ions (transmitted at less negative CV values), enhanced the peptide detection by 20%. Notably, lower abundant peptides were detected thanks to the greater S/N ratio and greatly improved the sensitivity [304]. Canterbury *et al*. also reached the same conclusion three years later, when FAIMS was used with tandem-mass-spectrometry for bottom-up proteomics [297]. FAIMS increased peak capacity without additional analysis time as for multidimensional pre-fractionation. The theoretical calculation showed that for the same peak capacity obtained with FAIMS, MudPIT would require 16 times more instrument use [297, 305]. With a *Top2* 5CV switching method and a 100 ms delay between the CV changes (FAIMS dwell time, theoretically at 70 ms in this study case), the group highlighted FAIMS parallelization abilities. However, for CV switching, less data points are obtained for the precursor elution profile. In addition, the number of CV is limited, since the time required to return to the initial CV increases the duty cycle time. The little overlap in identified peptides between the CVs results in supplementary identifications for each CV analyzed, making it a suitable online fractionation method. Nevertheless, although the separation mechanism of MS and IMS are different, the two methods are not entirely orthogonal. Moreover, a significant drop in intensity was reported due to the additional ion transfer path required with FAIMS resulting in only 10% ion transmission [297]. In the same year, FAIMS was also used for drug quantification in rat plasma with a targeted SRM approach and resulted in great accuracy, due to the enhanced S/N ratio [306, 307].

Ten years ago, our lab used FAIMS for the first time with an HR/AM Orbitrap for accurate label free quantification, including PTMs, and gained 55% PSM [308]. Since then, other groups used FAIMS in this configuration. The group of Cooper took advantage of this setup to study synthetic isobaric phosphopeptides [309, 310], and further developed this method in recent studies to aid in endogenous phosphopeptides analysis [311]. Compared to SCX, FAIMS aided in the identification of the challenging multi phosphorylated peptides, although resulting in an overall lower number of phosphosites. Similar results were noted with tryptic digests, where the CV stepping method lead to fewer identifications. Interestingly, FAIMS aided in the identification of triply charged peptides, suggesting a complementarity with SCX [312]. In addition to the benefits observed with FAIMS for phosphopeptides [309-311, 313], FAIMS possesses potential advantages for glycopeptides [314, 315].

Swearingen and colleagues have studied the use of FAIMS post ionization fractionation for shotgun proteomics. They performed 40 LC-MS/MS injections using a single CV per injection and covering 40 V in 1 V steps. The top 3 most populated CVs garnered an equal number of peptides than 3 replicate injections with conventional LC-MS/MS. While the number of identification from replicate LC-MS/MS injections reached a plateau, the number of identifications for additional CV fractions continued to grow when using LC-FAIMS-MS/MS. The overlap between FAIMS and nonFAIMS for peptides and proteins was nearly complete thanks to the exhaustive number of injections. FAIMS extended the proteomic analysis with 64% more proteins. This gain stemmed from the 10-fold increase in dynamic range achievable using FAIMS. A sheath gas for ionization was added to improve ion transmission in FAIMS, which increased the signal by roughly 5 times. Indeed, ion transmission was slightly higher than previously reported, but achieved only 1/6 of the transmission of conventional ESI-MS [316]. The most important improvement was realized when pure nitrogen was used as carrier gas instead of a mixture of nitrogen and helium and by redesigned the electrodes by decreasing the gap size between outer and inner electrodes from 2.5 mm to 1.5 mm [296, 317]. This created less gas turbulences, increased the electric field strength and a shortened the transient time/dwell time to 13.2 ms ± 3.9 ms, resulting in sharper CV peaks. These improvements enhanced peak capacity and ion transmission [296, 318].

1.4. Research objectives

The goal of this thesis was to gain further understanding of the dynamic changes of the human proteome under heat shock by integrating FAIMS in the workflow of mass spectrometry-based proteomics. More specifically, we evaluated how quantitative proteomic analyses could be improved when using FAIMS. Accordingly, we evaluated the potential applications of FAIMS in label free, metabolic and isobaric labeling. Furthermore, we also investigated if FAIMS could be used to improve the depth of phosphoproteome and SUMO proteome analyses. Lastly, as part of a collaboration with TFS, we evaluated the merits of a new FAIMS interface coupled to the Tribrid Fusion mass spectrometer to improve the comprehensiveness of label free quantitative proteomic analyses using both FAIMS MS^2 and SPS MS^3. This thesis is centered on the following four aims:

1. Evaluation of FAIMS to improve quantitative proteomics based on TMT isobaric labeling.
2. Improvement of metabolic labeling using FAIMS and its application to the profiling of protein phosphorylation upon hyperthermia.
3. Enhancement of SUMO proteome analyses in label-free quantitative proteomics.
4. Extended proteome coverage and sensitivity of multiplex proteomic analyses using a novel FAIMS interface.

1.5. Content of the thesis

CHAPTER ONE of my thesis provides a general overview of proteomics and how state of the art mass spectrometry-based methods deal with these high complexity samples. It also gives a brief overview of protein expression in normal cells or cells undergoing stress situations, because we are interested afterwards in the temporal dynamics of cells that are exposed to heat stress. A brief summary of sample preparations, instrumental analysis, and computational data treatment shows how one can capture these dynamic occurrences by LC-MS to obtain quantitative and qualitative information. Cellular responses induced by heat changes occur in two phases, early or late. Therefore, we look in chapters 2 to 5 at the heat shock response within different time frames to study effect on protein phosphorylation, SUMOylation and proteome changes. We evaluated the merits of FAIMS for different quantitative analysis strategies. ***CHAPTER TWO*** focusses on the improvements imparted by FAIMS for MS^2 based TMT quantification. ***CHAPTER THREE*** highlights the improved quantitation ability of FAIMS for SILAC experiments. ***CHAPTER FOUR*** shows how FAIMS can be used to increase the coverage of the identified SUMO proteome as well as its quantitation using a label free quantification approach. ***CHAPTER FIVE*** focuses on characterizing the new FAIMS instrument. This chapter demonstrates the improved robustness and enhancement in protein identification as well as quantification of the new FAIMS interface. Remarkably, gains could be obtained for TMT quantifications with LC-FAIMS-MS^2 compared to LC-SPS-MS^3. Finally, ***CHAPTER SIX*** provides a few reflections and conclusions that were drawn from the completed work. This chapter also gives future perspectives and what could be expected from FAIMS in the near future.

1.6. References

1. Yu, L.-R., N.A. Stewart, and T.D. Veenstra, *Chapter 8 - Proteomics: The Deciphering of the Functional Genome*, in *Essentials of Genomic and Personalized Medicine*, G.S. Ginsburg and H.F. Willard, Editors. 2010, Academic Press: San Diego. p. 89-96.
2. Burgers, P.M.J. and T.A. Kunkel, *Eukaryotic DNA Replication Fork.* Annu Rev Biochem, 2017. **86**: p. 417-438.
3. Sirbu, B.M. and D. Cortez, *DNA damage response: three levels of DNA repair regulation.* Cold Spring Harb Perspect Biol, 2013. **5**(8): p. a012724.
4. Herrmann, H. and U. Aebi, *Intermediate Filaments: Structure and Assembly.* Cold Spring Harb Perspect Biol, 2016. **8**(11).
5. Cura, A.J. and A. Carruthers, *Role of monosaccharide transport proteins in carbohydrate assimilation, distribution, metabolism, and homeostasis.* Compr Physiol, 2012. **2**(2): p. 863-914.
6. Braun, P. and A.C. Gingras, *History of protein-protein interactions: from egg-white to complex networks.* Proteomics, 2012. **12**(10): p. 1478-98.
7. Flick, K. and P. Kaiser, *Protein degradation and the stress response.* Seminars in cell & developmental biology, 2012. **23**(5): p. 515-522.
8. Watson, J.D. and F.H.C. Crick, *Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid.* Nature, 1953. **171**(4356): p. 737-738.
9. International Human Genome Sequencing, C., E.S. Lander, L.M. Linton, B. Birren, C. Nusbaum, M.C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, et al., *Initial sequencing and analysis of the human genome.* Nature, 2001. **409**: p. 860.
10. Venter, J.C., M.D. Adams, E.W. Myers, P.W. Li, R.J. Mural, G.G. Sutton, H.O. Smith, M. Yandell, C.A. Evans, R.A. Holt, et al., *The Sequence of the Human Genome.* Science, 2001. **291**(5507): p. 1304-1351.
11. Goffeau, A., B.G. Barrell, H. Bussey, R.W. Davis, B. Dujon, H. Feldmann, F. Galibert, J.D. Hoheisel, C. Jacq, M. Johnston, et al., *Life with 6000 Genes.* Science, 1996. **274**(5287): p. 546-567.
12. Romanowska, J. and A. Joshi, *From Genotype to Phenotype: Through Chromatin.* Genes, 2019. **10**(2): p. 76.
13. Drissi, R., M.L. Dubois, and F.M. Boisvert, *Proteomics methods for subcellular proteome analysis.* Febs j, 2013. **280**(22): p. 5626-34.
14. Tian, R., *Exploring intercellular signaling by proteomic approaches.* Proteomics, 2014. **14**(4-5): p. 498-512.
15. Lamond, A.I., M. Uhlen, S. Horning, A. Makarov, C.V. Robinson, L. Serrano, F.U. Hartl, W. Baumeister, A.K. Werenskiold, J.S. Andersen, et al., *Advancing Cell Biology Through Proteomics in Space and Time (PROSPECTS).* Molecular & Cellular Proteomics, 2012. **11**(3): p. O112.017731.
16. Omenn, G.S., L. Lane, C.M. Overall, F.J. Corrales, J.M. Schwenk, Y.-K. Paik, J.E. Van Eyk, S. Liu, M. Snyder, M.S. Baker, et al., *Progress on Identifying and Characterizing the Human Proteome: 2018 Metrics from the HUPO Human Proteome Project.* Journal of Proteome Research, 2018. **17**(12): p. 4031-4041.

17. Hebert, A.S., A.L. Richards, D.J. Bailey, A. Ulbrich, E.E. Coughlin, M.S. Westphall, and J.J. Coon, *The one hour yeast proteome.* Mol Cell Proteomics, 2014. **13**(1): p. 339-47.

18. Kelstrup, C.D., D.B. Bekker-Jensen, T.N. Arrey, A. Hogrebe, A. Harder, and J.V. Olsen, *Performance Evaluation of the Q Exactive HF-X for Shotgun Proteomics.* Journal of Proteome Research, 2018. **17**(1): p. 727-738.

19. Bannister, A.J. and T. Kouzarides, *Regulation of chromatin by histone modifications.* Cell Research, 2011. **21**: p. 381.

20. Sakharkar, M.K., V.T. Chow, and P. Kangueane, *Distributions of exons and introns in the human genome.* In Silico Biol, 2004. **4**(4): p. 387-93.

21. Southan, C., *Last rolls of the yoyo: Assessing the human canonical protein count.* F1000Res, 2017. **6**: p. 448.

22. Schleif, R.F., *Modulation of DNA Binding by Gene-Specific Transcription Factors.* Biochemistry, 2013. **52**(39): p. 6755-6765.

23. Matera, A.G. and Z. Wang, *A day in the life of the spliceosome.* Nat Rev Mol Cell Biol, 2014. **15**(2): p. 108-21.

24. Coltri, P.P., M.G.P. dos Santos, and G.H.G. da Silva, *Splicing and cancer: Challenges and opportunities.* Wiley Interdisciplinary Reviews: RNA. **0**(0): p. e1527.

25. Clancy, S. and W. Brown, *Translation: DNA to mRNA to Protein.* Nature Education, 2008. **1**(1): p. 101.

26. Fan, Y., C.R. Evans, and J. Ling, *Rewiring protein synthesis: From natural to synthetic amino acids.* Biochim Biophys Acta Gen Subj, 2017. **1861**(11 Pt B): p. 3024-3029.

27. Brunet, M.A., M. Brunelle, J.F. Lucier, V. Delcourt, M. Levesque, F. Grenier, S. Samandi, S. Leblanc, J.D. Aguilar, P. Dufour, et al., *OpenProt: a more comprehensive guide to explore eukaryotic coding potential and proteomes.* Nucleic Acids Res, 2019. **47**(D1): p. D403-d410.

28. Consortium, T.U., *UniProt: the universal protein knowledgebase.* Nucleic Acids Res, 2017. **45**(D1): p. D158-d169.

29. Munoz, V. and M. Cerminara, *When fast is better: protein folding fundamentals and mechanisms from ultrafast approaches.* Biochem J, 2016. **473**(17): p. 2545-59.

30. de Klerk, E. and P.A. t Hoen, *Alternative mRNA transcription, processing, and translation: insights from RNA sequencing.* Trends Genet, 2015. **31**(3): p. 128-39.

31. Smith, L.M., N.L. Kelleher, P. The Consortium for Top Down, M. Linial, D. Goodlett, P. Langridge-Smith, Y. Ah Goo, G. Safford, L. Bonilla*, G. Kruppa, et al., *Proteoform: a single term describing protein complexity.* Nature Methods, 2013. **10**: p. 186.

32. Irimia, M. and S.W. Roy, *Origin of spliceosomal introns and alternative splicing.* Cold Spring Harb Perspect Biol, 2014. **6**(6).

33. Marquez, Y., M. Hopfler, Z. Ayatollahi, A. Barta, and M. Kalyna, *Unmasking alternative splicing inside protein-coding exons defines exitrons and their role in proteome plasticity.* Genome Res, 2015. **25**(7): p. 995-1007.

34. Vegvari, A., *Mutant Proteogenomics.* Adv Exp Med Biol, 2016. **926**: p. 77-91.

35. Roth, M.J., A.J. Forbes, M.T. Boyne, 2nd, Y.B. Kim, D.E. Robinson, and N.L. Kelleher, *Precise and parallel characterization of coding polymorphisms, alternative splicing, and modifications in human proteins by mass spectrometry.* Mol Cell Proteomics, 2005. **4**(7): p. 1002-8.

36. Ponomarenko, E.A., E.V. Poverennaya, E.V. Ilgisonis, M.A. Pyatnitskiy, A.T. Kopylov, V.G. Zgoda, A.V. Lisitsa, and A.I. Archakov, *The Size of the Human Proteome: The Width and Depth.* Int J Anal Chem, 2016. **2016**: p. 7436849.

37. Mallick, P. and B. Kuster, *Proteomics: a pragmatic perspective.* Nature Biotechnology, 2010. **28**: p. 695.

38. Zubarev, R.A., *The challenge of the proteome dynamic range and its implications for in-depth proteomics.* Proteomics, 2013. **13**(5): p. 723-6.

39. Singh, V., M. Ram, R. Kumar, R. Prasad, B.K. Roy, and K.K. Singh, *Phosphorylation: Implications in Cancer.* Protein J, 2017. **36**(1): p. 1-6.

40. Wong, C.H., *Protein glycosylation: new challenges and opportunities.* J Org Chem, 2005. **70**(11): p. 4219-25.

41. Nemeth-Cawley, J.F., S. Karnik, and J.C. Rouse, *Analysis of sulfated peptides using positive electrospray ionization tandem mass spectrometry.* J Mass Spectrom, 2001. **36**(12): p. 1301-11.

42. Azevedo, C. and A. Saiardi, *Why always lysine? The ongoing tale of one of the most modified amino acids.* Adv Biol Regul, 2016. **60**: p. 144-150.

43. Vu, L.D., K. Gevaert, and I. De Smet, *Protein Language: Post-Translational Modifications Talking to Each Other.* Trends in Plant Science, 2018. **23**(12): p. 1068-1080.

44. Allan Drummond, D. and C.O. Wilke, *The evolutionary consequences of erroneous protein synthesis.* Nature Reviews Genetics, 2009. **10**: p. 715.

45. Ogle, J.M. and V. Ramakrishnan, *Structural insights into translational fidelity.* Annu Rev Biochem, 2005. **74**: p. 129-77.

46. Morimoto, R.I., *Regulation of the heat shock transcriptional response: cross talk between a family of heat shock factors, molecular chaperones, and negative regulators.* Genes Dev, 1998. **12**(24): p. 3788-96.

47. Advani, V.M. and P. Ivanov, *Translational Control under Stress: Reshaping the Translatome.* Bioessays, 2019. **41**(5): p. e1900009.

48. Buchan, J.R. and R. Parker, *Eukaryotic Stress Granules: The Ins and Outs of Translation.* Molecular Cell, 2009. **36**(6): p. 932-941.

49. Åkerfelt, M., R.I. Morimoto, and L. Sistonen, *Heat shock factors: integrators of cell stress, development and lifespan.* Nature Reviews Molecular Cell Biology, 2010. **11**: p. 545.

50. Voellmy, R. and F. Boellmann, *Chaperone Regulation of the Heat Shock Protein Response*, in *Molecular Aspects of the Stress Response: Chaperones, Membranes and Networks*, P. Csermely and L. Vígh, Editors. 2007, Springer New York: New York, NY. p. 89-99.

51. Feder, M.E. and G.E. Hofmann, *HEAT-SHOCK PROTEINS, MOLECULAR CHAPERONES, AND THE STRESS RESPONSE: Evolutionary and Ecological Physiology.* Annual Review of Physiology, 1999. **61**(1): p. 243-282.

52. Li, W., A.N. Reeb, B. Lin, P. Subramanian, E.E. Fey, C.R. Knoverek, R.L. French, E.H. Bigio, and Y.M. Ayala, *Heat Shock-induced Phosphorylation of TAR DNA-binding Protein 43 (TDP-43) by MAPK/ERK Kinase Regulates TDP-43 Function.* J Biol Chem, 2017. **292**(12): p. 5089-5100.

53. Rabouille, C. and S. Alberti, *Cell adaptation upon stress: the emerging role of membrane-less compartments.* Curr Opin Cell Biol, 2017. **47**: p. 34-42.

54. Richter, K., M. Haslbeck, and J. Buchner, *The Heat Shock Response: Life on the Verge of Death.* Molecular Cell, 2010. **40**(2): p. 253-266.

55. Haslbeck, M. and E. Vierling, *A first line of stress defense: small heat shock proteins and their function in protein homeostasis.* J Mol Biol, 2015. **427**(7): p. 1537-48.

56. Ellis, R.J., *The molecular chaperone concept.* Seminars in cell biology, 1990. **1**(1): p. 1-9.

57. Vihervaara, A. and L. Sistonen, *HSF1 at a glance.* Journal of Cell Science, 2014. **127**(2): p. 261-266.

58. Xu, Y.-M., D.-Y. Huang, J.-F. Chiu, and A.T.Y. Lau, *Post-Translational Modification of Human Heat Shock Factors and Their Functions: A Recent Update by Proteomic Approach.* Journal of Proteome Research, 2012. **11**(5): p. 2625-2634.

59. Singh, I.S. and J.D. Hasday, *Fever, hyperthermia and the heat shock response.* Int J Hyperthermia, 2013. **29**(5): p. 423-35.

60. Morimoto, R., *Cells in stress: transcriptional activation of heat shock genes.* Science, 1993. **259**(5100): p. 1409-1410.

61. Olsen, J.V., B. Blagoev, F. Gnad, B. Macek, C. Kumar, P. Mortensen, and M. Mann, *Global, In Vivo, and Site-Specific Phosphorylation Dynamics in Signaling Networks.* Cell, 2006. **127**(3): p. 635-648.

62. Olsen, J.V. and M. Mann, *Status of Large-scale Analysis of Post-translational Modifications by Mass Spectrometry.* Molecular & Cellular Proteomics, 2013. **12**(12): p. 3444-3452.

63. Aebersold, R. and M. Mann, *Mass spectrometry-based proteomics.* Nature, 2003. **422**(6928): p. 198-207.

64. Wilm, M., A. Shevchenko, T. Houthaeve, S. Breit, L. Schweigerer, T. Fotsis, and M. Mann, *Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry.* Nature, 1996. **379**(6564): p. 466-9.

65. Dunn, M.J. and A.H.M. Burghes, *Chapter 4 - High Resolution Two-dimensional Polyacrylamide-gel Electrophoresis*, in *Gel Electrophoresis of Proteins*, M.J. Dunn, Editor. 1986, Butterworth-Heinemann. p. 203-261.

66. Zhang, Y., B.R. Fonslow, B. Shan, M.-C. Baek, and J.R. Yates, *Protein Analysis by Shotgun/Bottom-up Proteomics.* Chemical Reviews, 2013. **113**(4): p. 2343-2394.

67. Sharma, K., S. Schmitt, C.G. Bergner, S. Tyanova, N. Kannaiyan, N. Manrique-Hoyos, K. Kongi, L. Cantuti, U.-K. Hanisch, M.-A. Philips, et al., *Cell type– and brain region–resolved mouse brain proteome.* Nature Neuroscience, 2015. **18**: p. 1819.

68. Kanshin, E., L.P. Bergeron-Sandoval, S.S. Isik, P. Thibault, and S.W. Michnick, *A cell-signaling network temporally resolves specific versus promiscuous phosphorylation.* Cell Rep, 2015. **10**(7): p. 1202-14.

69. Ly, T., A. Endo, and A.I. Lamond, *Proteomic analysis of the response to cell cycle arrests in human myeloid leukemia cells.* Elife, 2015. **4**.

70. McShane, E., C. Sin, H. Zauber, J.N. Wells, N. Donnelly, X. Wang, J. Hou, W. Chen, Z. Storchova, J.A. Marsh, et al., *Kinetic Analysis of Protein Stability Reveals Age-Dependent Degradation.* Cell, 2016. **167**(3): p. 803-815.e21.

71. Itzhak, D.N., S. Tyanova, J. Cox, and G.H. Borner, *Global, quantitative and dynamic mapping of protein subcellular localization.* Elife, 2016. **5**.

72. Wang, D., B. Eraslan, T. Wieland, B. Hallstrom, T. Hopf, D.P. Zolg, J. Zecha, A. Asplund, L.H. Li, C. Meng, et al., *A deep proteome and transcriptome abundance atlas of 29 healthy human tissues.* Mol Syst Biol, 2019. **15**(2): p. e8503.

73. Deshmukh, A.S., M. Murgia, N. Nagaraj, J.T. Treebak, J. Cox, and M. Mann, *Deep proteomics of mouse skeletal muscle enables quantitation of protein isoforms, metabolic pathways, and transcription factors.* Mol Cell Proteomics, 2015. **14**(4): p. 841-53.

74. Larance, M. and A.I. Lamond, *Multidimensional proteomics for cell biology.* Nat Rev Mol Cell Biol, 2015. **16**(5): p. 269-80.

75. Catherman, A.D., O.S. Skinner, and N.L. Kelleher, *Top Down proteomics: facts and perspectives.* Biochemical and biophysical research communications, 2014. **445**(4): p. 683-693.

76. Wu, C., J.C. Tran, L. Zamdborg, K.R. Durbin, M. Li, D.R. Ahlf, B.P. Early, P.M. Thomas, J.V. Sweedler, and N.L. Kelleher, *A protease for 'middle-down' proteomics.* Nature Methods, 2012. **9**: p. 822.

77. Pandeswari, P.B. and V. Sabareesh, *Middle-down approach: a choice to sequence and characterize proteins/proteomes by mass spectrometry.* RSC Advances, 2019. **9**(1): p. 313-344.

78. Tran, J.C., L. Zamdborg, D.R. Ahlf, J.E. Lee, A.D. Catherman, K.R. Durbin, J.D. Tipton, A. Vellaichamy, J.F. Kellie, M. Li, et al., *Mapping intact protein isoforms in discovery mode using top-down proteomics.* Nature, 2011. **480**(7376): p. 254-8.

79. Sidoli, S. and B.A. Garcia, *Middle-down proteomics: a still unexploited resource for chromatin biology.* Expert Rev Proteomics, 2017. **14**(7): p. 617-626.

80. Sidoli, S., V. Schwämmle, C. Ruminowicz, T.A. Hansen, X. Wu, K. Helin, and O.N. Jensen, *Middle-down hybrid chromatography/tandem mass spectrometry workflow for characterization of combinatorial post-translational modifications in histones.* PROTEOMICS, 2014. **14**(19): p. 2200-2211.

81. Link, A.J., J. Eng, D.M. Schieltz, E. Carmack, G.J. Mize, D.R. Morris, B.M. Garvik, and J.R. Yates, 3rd, *Direct analysis of protein complexes using mass spectrometry.* Nat Biotechnol, 1999. **17**(7): p. 676-82.

82. Aebersold, R. and M. Mann, *Mass-spectrometric exploration of proteome structure and function.* Nature, 2016. **537**: p. 347.

83. Winter, D. and H. Steen, *Optimization of cell lysis and protein digestion protocols for the analysis of HeLa S3 cells by LC-MS/MS.* Proteomics, 2011. **11**(24): p. 4726-4730.

84. Tsiatsiani, L. and A.J. Heck, *Proteomics beyond trypsin.* Febs j, 2015. **282**(14): p. 2612-26.

85. Burkhart, J.M., C. Schumbrutzki, S. Wortelkamp, A. Sickmann, and R.P. Zahedi, *Systematic and quantitative comparison of digest efficiency and specificity reveals the impact of trypsin quality on MS-based proteomics.* J Proteomics, 2012. **75**(4): p. 1454-62.

86. Šlechtová, T., M. Gilar, K. Kalíková, and E. Tesařová, *Insight into Trypsin Miscleavage: Comparison of Kinetic Constants of Problematic Peptide Sequences.* Analytical Chemistry, 2015. **87**(15): p. 7636-7643.

87. Trevisiol, S., D. Ayoub, A. Lesur, L. Ancheva, S. Gallien, and B. Domon, *The use of proteases complementary to trypsin to probe isoforms and modifications.* Proteomics, 2016. **16**(5): p. 715-28.

88. Swaney, D.L., C.D. Wenger, and J.J. Coon, *Value of using multiple proteases for large-scale mass spectrometry-based proteomics.* J Proteome Res, 2010. **9**(3): p. 1323-9.

89. Rappsilber, J., M. Mann, and Y. Ishihama, *Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips.* Nature Protocols, 2007. **2**: p. 1896.

90. Tubaon, R.M., P.R. Haddad, and J.P. Quirino, *Sample Clean-up Strategies for ESI Mass Spectrometry Applications in Bottom-up Proteomics: Trends from 2012 to 2016.* Proteomics, 2017. **17**(20).

91. Rogers, J.C. and R.D. Bomgarden, *Sample Preparation for Mass Spectrometry-Based Proteomics; from Proteomes to Peptides.* Adv Exp Med Biol, 2016. **919**: p. 43-62.

92. Hughes, C.S., S. Foehr, D.A. Garfield, E.E. Furlong, L.M. Steinmetz, and J. Krijgsveld, *Ultrasensitive proteome analysis using paramagnetic bead technology.* Molecular Systems Biology, 2014. **10**(10): p. 757.

93. Mayne, J., Z. Ning, X. Zhang, A.E. Starr, R. Chen, S. Deeke, C.-K. Chiang, B. Xu, M. Wen, K. Cheng, et al., *Bottom-Up Proteomics (2013–2015): Keeping up in the Era of Systems Biology.* Analytical Chemistry, 2016. **88**(1): p. 95-121.

94. Kanshin, E., M. Tyers, and P. Thibault, *Sample Collection Method Bias Effects in Quantitative Phosphoproteomics.* Journal of Proteome Research, 2015. **14**(7): p. 2998-3004.

95. Kulak, N.A., G. Pichler, I. Paron, N. Nagaraj, and M. Mann, *Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells.* Nature Methods, 2014. **11**: p. 319.

96. Martins, G., J. Fernandez-Lodeiro, J. Djafari, C. Lodeiro, J.L. Capelo, and H.M. Santos, *Label-free protein quantification after ultrafast digestion of complex proteomes using ultrasonic energy and immobilized-trypsin magnetic nanoparticles.* Talanta, 2019. **196**: p. 262-270.

97. Geyer, Philipp E., Nils A. Kulak, G. Pichler, Lesca M. Holdt, D. Teupser, and M. Mann, *Plasma Proteome Profiling to Assess Human Health and Disease.* Cell Systems, 2016. **2**(3): p. 185-195.

98. Thingholm, T.E. and M.R. Larsen, *Phosphopeptide Enrichment by Immobilized Metal Affinity Chromatography.* Methods Mol Biol, 2016. **1355**: p. 123-33.

99. Bhullar, K.S., N.O. Lagaron, E.M. McGowan, I. Parmar, A. Jha, B.P. Hubbard, and H.P.V. Rupasinghe, *Kinase-targeted cancer therapies: progress, challenges and future directions.* Mol Cancer, 2018. **17**(1): p. 48.

100. Day, E.K., N.G. Sosale, and M.J. Lazzara, *Cell signaling regulation by protein phosphorylation: a multivariate, heterogeneous, and context-dependent process.* Curr Opin Biotechnol, 2016. **40**: p. 185-192.

101. Searle, B.C., R.T. Lawrence, M.J. MacCoss, and J. Villén, *Thesaurus: quantifying phosphopeptide positional isomers.* Nature Methods, 2019. **16**(8): p. 703-706.

102. Mithoe, S.C. and F.L. Menke, *Phosphopeptide immuno-affinity enrichment to enhance detection of tyrosine phosphorylation in plants.* Methods Mol Biol, 2015. **1306**: p. 135-46.

103. Wilson, V.G., *Introduction to Sumoylation.* Adv Exp Med Biol, 2017. **963**: p. 1-12.

104. Eifler, K. and A.C.O. Vertegaal, *Mapping the SUMOylated landscape.* The FEBS Journal, 2015. **282**(19): p. 3669-3680.

105. Flotho, A. and F. Melchior, *Sumoylation: A Regulatory Protein Modification in Health and Disease.* Annual Review of Biochemistry, 2013. **82**(1): p. 357-385.

106. Pfammatter, S., E. Bonneil, F.P. McManus, and P. Thibault, *Gas-Phase Enrichment of Multiply Charged Peptide Ions by Differential Ion Mobility Extend the Comprehensiveness of SUMO Proteome Analyses.* J Am Soc Mass Spectrom, 2018. **29**(6): p. 1111-1124.

107. Častorálová, M., D. Březinová, M. Švéda, J. Lipov, T. Ruml, and Z. Knejzlík, *SUMO-2/3 conjugates accumulating under heat shock or MG132 treatment result largely from new protein synthesis.* Biochimica et Biophysica Acta (BBA) - Molecular Cell Research, 2012. **1823**(4): p. 911-919.

108. Hendriks, I.A., D. Lyon, D. Su, N.H. Skotte, J.A. Daniel, L.J. Jensen, and M.L. Nielsen, *Site-specific characterization of endogenous SUMOylation across species and organs.* Nature Communications, 2018. **9**(1): p. 2456.

109. McManus, F.P., F. Lamoliatte, and P. Thibault, *Identification of cross talk between SUMOylation and ubiquitylation using a sequential peptide immunopurification approach.* Nat Protoc, 2017. **12**(11): p. 2342-2358.

110. Hendriks, I.A. and A.C. Vertegaal, *A comprehensive compilation of SUMO proteomics.* Nat Rev Mol Cell Biol, 2016. **17**(9): p. 581-95.

111. Galisson, F., L. Mahrouche, M. Courcelles, E. Bonneil, S. Meloche, M.K. Chelbi-Alix, and P. Thibault, *A novel proteomics approach to identify SUMOylated proteins and their modification sites in human cells.* Mol Cell Proteomics, 2011. **10**(2): p. M110.004796.

112. Lamoliatte, F., D. Caron, C. Durette, L. Mahrouche, M.A. Maroui, O. Caron-Lizotte, E. Bonneil, M.K. Chelbi-Alix, and P. Thibault, *Large-scale analysis of lysine SUMOylation by SUMO remnant immunoaffinity profiling.* Nat Commun, 2014. **5**: p. 5409.

113. McManus, F.P., C.D. Altamirano, and P. Thibault, *In vitro assay to determine SUMOylation sites on protein substrates.* Nature Protocols, 2016. **11**: p. 387.

114. Domon, B. and R. Aebersold, *Mass Spectrometry and Protein Analysis.* Science, 2006. **312**(5771): p. 212-217.

115. Haag, A.M., *Mass Analyzers and Mass Spectrometers*, in *Modern Proteomics – Sample Preparation, Analysis and Practical Applications*, H. Mirzaei and M. Carrasco, Editors. 2016, Springer International Publishing: Cham. p. 157-169.

116. Dagan, S. and A. Amirav, *Electron impact mass spectrometry of alkanes in supersonic molecular beams.* Journal of the American Society for Mass Spectrometry, 1995. **6**(2): p. 120-131.

117. Mano, N. and J. Goto, *Biomedical and biological mass spectrometry.* Anal Sci, 2003. **19**(1): p. 3-14.

118. Karas, M. and F. Hillenkamp, *Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons.* Analytical Chemistry, 1988. **60**(20): p. 2299-2301.

119. Fenn, J., M. Mann, C. Meng, S. Wong, and C. Whitehouse, *Electrospray ionization for mass spectrometry of large biomolecules.* Science, 1989. **246**(4926): p. 64-71.

120. Dreisewerd, K., *Recent methodological advances in MALDI mass spectrometry.* Anal Bioanal Chem, 2014. **406**(9-10): p. 2261-78.

121. Feenstra, A.D., M.E. Dueñas, and Y.J. Lee, *Five Micron High Resolution MALDI Mass Spectrometry Imaging with Simple, Interchangeable, Multi-Resolution Optical System.* Journal of The American Society for Mass Spectrometry, 2017. **28**(3): p. 434-442.

122. Wilm, M., *Principles of Electrospray Ionization.* Molecular & Cellular Proteomics, 2011. **10**(7): p. M111.009407.

123. Hendricks, C.D. and J. Hogan, *Investigation of the charge-to-mass ratio of electrically sprayed liquid particles.* AIAA Journal, 1965. **3**(2): p. 296-301.

124. Wilm, M. and M. Mann, *Analytical Properties of the Nanoelectrospray Ion Source.* Analytical Chemistry, 1996. **68**(1): p. 1-8.

125. Annesley, T.M., *Ion Suppression in Mass Spectrometry.* Clinical Chemistry, 2003. **49**(7): p. 1041-1044.

126. Han, X., A. Aslanian, and J.R. Yates, *Mass spectrometry for proteomics.* Current Opinion in Chemical Biology, 2008. **12**(5): p. 483-490.

127. Marshall, A.G. and C.L. Hendrickson, *High-Resolution Mass Spectrometers.* Annual Review of Analytical Chemistry, 2008. **1**(1): p. 579-599.

128. Makarov, A., *Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis.* Analytical Chemistry, 2000. **72**(6): p. 1156-1162.

129. Hu, Q., R.J. Noll, H. Li, A. Makarov, M. Hardman, and R. Graham Cooks, *The Orbitrap: a new mass spectrometer.* J Mass Spectrom, 2005. **40**(4): p. 430-43.

130. Yates, J.R., 3rd, *Recent technical advances in proteomics.* F1000Res, 2019. **8**.

131. Landais, B., C. Beaugrand, L. Capron-Dukan, M. Sablier, G. Simonneau, and C. Rolando, *Varying the radio frequency: a new scanning mode for quadrupole analyzers.* Rapid Communications in Mass Spectrometry, 1998. **12**(6): p. 302-306.

132. March, R.E.a.T., J. F. , *Theory of Quadrupole Instruments*, in *Quadrupole Ion Trap Mass Spectrometry*, J.D. Winefordner, Editor. 2005, John Wiley & sons, inc.: Wiley-Interscience. p. 34-72.

133. Douglas, D.J., *Linear quadrupoles in mass spectrometry.* Mass Spectrometry Reviews, 2009. **28**(6): p. 937-960.

134. Douglas, D.J., *Applications of collision dynamics in quadrupole mass spectrometry.* Journal of the American Society for Mass Spectrometry, 1998. **9**(2): p. 101-113.

135. Douglas, D.J., A.J. Frank, and D. Mao, *Linear ion traps in mass spectrometry.* Mass Spectrometry Reviews, 2005. **24**(1): p. 1-29.

136. Nolting, D., R. Malek, and A. Makarov, *Ion traps in modern mass spectrometry.* Mass Spectrometry Reviews, 2019. **38**(2): p. 150-168.

137. Schwartz, J.C., M.W. Senko, and J.E.P. Syka, *A two-dimensional quadrupole ion trap mass spectrometer.* Journal of the American Society for Mass Spectrometry, 2002. **13**(6): p. 659-669.

138. Perry, R.H., R.G. Cooks, and R.J. Noll, *Orbitrap mass spectrometry: Instrumentation, ion motion and applications.* Mass Spectrometry Reviews, 2008. **27**(6): p. 661-699.

139. Zubarev, R.A. and A. Makarov, *Orbitrap Mass Spectrometry.* Analytical Chemistry, 2013. **85**(11): p. 5288-5296.

140. Eliuk, S. and A. Makarov, *Evolution of Orbitrap Mass Spectrometry Instrumentation.* Annual Review of Analytical Chemistry, 2015. **8**(1): p. 61-80.

141. Makarov, A., E. Denisov, A. Kholomeev, W. Balschun, O. Lange, K. Strupat, and S. Horning, *Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer.* Anal Chem, 2006. **78**(7): p. 2113-20.

142. Michalski, A., E. Damoc, J.-P. Hauschild, O. Lange, A. Wieghaus, A. Makarov, N. Nagaraj, J. Cox, M. Mann, and S. Horning, *Mass Spectrometry-based Proteomics Using Q Exactive, a*

High-performance Benchtop Quadrupole Orbitrap Mass Spectrometer. Molecular & Cellular Proteomics, 2011. **10**(9): p. M111.011015.

143. Grinfeld, D., K. Aizikov, A. Kreutzmann, E. Damoc, and A. Makarov, *Phase-Constrained Spectrum Deconvolution for Fourier Transform Mass Spectrometry.* Analytical Chemistry, 2017. **89**(2): p. 1202-1211.

144. Biemann, K., *Contributions of mass spectrometry to peptide and protein structure.* Biomedical & Environmental Mass Spectrometry, 1988. **16**(1-12): p. 99-111.

145. Brodbelt, J.S., *Ion Activation Methods for Peptides and Proteins.* Analytical Chemistry, 2016. **88**(1): p. 30-51.

146. Biemann, K., *Mass Spectrometry of Peptides and Proteins.* Annual Review of Biochemistry, 1992. **61**(1): p. 977-1010.

147. Degroeve, S., D. Maddelein, and L. Martens, *MS2PIP prediction server: compute and visualize MS2 peak intensity predictions for CID and HCD fragmentation.* Nucleic Acids Res, 2015. **43**(W1): p. W326-30.

148. Gessulat, S., T. Schmidt, D.P. Zolg, P. Samaras, K. Schnatbaum, J. Zerweck, T. Knaute, J. Rechenberger, B. Delanghe, A. Huhmer, et al., *Prosit: proteome-wide prediction of peptide tandem mass spectra by deep learning.* Nature Methods, 2019. **16**(6): p. 509-518.

149. Tang, X.J., P. Thibault, and R.K. Boyd, *Fragmentation reactions of multiply-protonated peptides and implications for sequencing by tandem mass spectrometry with low-energy collision-induced dissociation.* Anal Chem, 1993. **65**(20): p. 2824-34.

150. Sonsmann, G., A. Römer, and D. Schomburg, *Investigation of the influence of charge derivatization on the fragmentation of multiply protonated peptides.* Journal of the American Society for Mass Spectrometry, 2002. **13**(1): p. 47-58.

151. Michalski, A., N. Neuhauser, J. Cox, and M. Mann, *A Systematic Investigation into the Nature of Tryptic HCD Spectra.* Journal of Proteome Research, 2012. **11**(11): p. 5479-5491.

152. Olsen, J.V., B. Macek, O. Lange, A. Makarov, S. Horning, and M. Mann, *Higher-energy C-trap dissociation for peptide modification analysis.* Nat Methods, 2007. **4**(9): p. 709-12.

153. Michalski, A., E. Damoc, O. Lange, E. Denisov, D. Nolting, M. Müller, R. Viner, J. Schwartz, P. Remes, M. Belford, et al., *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, 2012. **11**(3): p. O111.013698.

154. Senko, M.W., P.M. Remes, J.D. Canterbury, R. Mathur, Q. Song, S.M. Eliuk, C. Mullen, L. Earley, M. Hardman, J.D. Blethrow, et al., *Novel Parallelized Quadrupole/Linear Ion Trap/Orbitrap Tribrid Mass Spectrometer Improving Proteome Coverage and Peptide Identification Rates.* Analytical Chemistry, 2013. **85**(24): p. 11710-11714.

155. Riley, N.M., C. Mullen, C.R. Weisbrod, S. Sharma, M.W. Senko, V. Zabrouskov, M.S. Westphall, J.E.P. Syka, and J.J. Coon, *Enhanced Dissociation of Intact Proteins with High Capacity Electron Transfer Dissociation.* Journal of the American Society for Mass Spectrometry, 2016. **27**(3): p. 520-531.

156. Olsen, J.V., J.C. Schwartz, J. Griep-Raming, M.L. Nielsen, E. Damoc, E. Denisov, O. Lange, P. Remes, D. Taylor, M. Splendore, et al., *A dual pressure linear ion trap Orbitrap instrument with very high sequencing speed.* Mol Cell Proteomics, 2009. **8**(12): p. 2759-69.

157. Kalli, A., G.T. Smith, M.J. Sweredoski, and S. Hess, *Evaluation and Optimization of Mass Spectrometric Settings during Data-dependent Acquisition Mode: Focus on LTQ-Orbitrap Mass Analyzers.* Journal of Proteome Research, 2013. **12**(7): p. 3071-3086.

158. Ferries, S., S. Perkins, P.J. Brownridge, A. Campbell, P.A. Eyers, A.R. Jones, and C.E. Eyers, *Evaluation of Parameters for Confident Phosphorylation Site Localization Using an Orbitrap Fusion Tribrid Mass Spectrometer.* Journal of Proteome Research, 2017. **16**(9): p. 3448-3459.

159. Kockmann, T., C. Trachsel, C. Panse, A. Wahlander, N. Selevsek, J. Grossmann, W.E. Wolski, and R. Schlapbach, *Targeted proteomics coming of age – SRM, PRM and DIA performance evaluated from a core facility perspective.* PROTEOMICS, 2016. **16**(15-16): p. 2183-2192.

160. Olsen, J.V., J.C. Schwartz, J. Griep-Raming, M.L. Nielsen, E. Damoc, E. Denisov, O. Lange, P. Remes, D. Taylor, M. Splendore, et al., *A Dual Pressure Linear Ion Trap Orbitrap Instrument with Very High Sequencing Speed.* Molecular & Cellular Proteomics, 2009. **8**(12): p. 2759-2769.

161. Kreimer, S., M.E. Belov, W.F. Danielson, L.I. Levitsky, M.V. Gorshkov, B.L. Karger, and A.R. Ivanov, *Advanced Precursor Ion Selection Algorithms for Increased Depth of Bottom-Up Proteomic Profiling.* Journal of Proteome Research, 2016. **15**(10): p. 3563-3573.

162. Michalski, A., J. Cox, and M. Mann, *More than 100,000 Detectable Peptide Species Elute in Single Shotgun Proteomics Runs but the Majority is Inaccessible to Data-Dependent LC−MS/MS.* Journal of Proteome Research, 2011. **10**(4): p. 1785-1793.

163. Houel, S., R. Abernathy, K. Renganathan, K. Meyer-Arendt, N.G. Ahn, and W.M. Old, *Quantifying the Impact of Chimera MS/MS Spectra on Peptide Identification in Large-Scale Proteomics Studies.* Journal of Proteome Research, 2010. **9**(8): p. 4152-4160.

164. Hebert, A.S., C. Thoing, N.M. Riley, N.W. Kwiecien, E. Shiskova, R. Huguet, H.L. Cardasis, A. Kuehn, S. Eliuk, V. Zabrouskov, et al., *Improved Precursor Characterization for Data-Dependent Mass Spectrometry.* Anal Chem, 2018. **90**(3): p. 2333-2340.

165. Berg, M., A. Parbel, H. Pettersen, D. Fenyö, and L. Björkesten, *Reproducibility of LC-MS-based protein identification.* Journal of Experimental Botany, 2006. **57**(7): p. 1509-1514.

166. Ludwig, C., L. Gillet, G. Rosenberger, S. Amon, B.C. Collins, and R. Aebersold, *Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial.* Molecular Systems Biology, 2018. **14**(8): p. e8126.

167. Venable, J.D., M.-Q. Dong, J. Wohlschlegel, A. Dillin, and J.R. Yates, *Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra.* Nature Methods, 2004. **1**(1): p. 39-45.

168. Bilbao, A., E. Varesio, J. Luban, C. Strambio-De-Castillia, G. Hopfgartner, M. Müller, and F. Lisacek, *Processing strategies and software solutions for data-independent acquisition in mass spectrometry.* PROTEOMICS, 2015. **15**(5-6): p. 964-980.

169. Masselon, C., G.A. Anderson, R. Harkewicz, J.E. Bruce, L. Pasa-Tolic, and R.D. Smith, *Accurate mass multiplexed tandem mass spectrometry for high-throughput polypeptide identification from mixtures.* Anal Chem, 2000. **72**(8): p. 1918-24.

170. Gillet, L.C., P. Navarro, S. Tate, H. Röst, N. Selevsek, L. Reiter, R. Bonner, and R. Aebersold, *Targeted Data Extraction of the MS/MS Spectra Generated by Data-independent Acquisition: A New Concept for Consistent and Accurate Proteome Analysis.* Molecular & Cellular Proteomics, 2012. **11**(6): p. O111.016717.

171. Collins, B.C., C.L. Hunter, Y. Liu, B. Schilling, G. Rosenberger, S.L. Bader, D.W. Chan, B.W. Gibson, A.-C. Gingras, J.M. Held, et al., *Multi-laboratory assessment of reproducibility, qualitative and quantitative performance of SWATH-mass spectrometry.* Nature Communications, 2017. **8**(1): p. 291.

172. Röst, H.L., G. Rosenberger, P. Navarro, L. Gillet, S.M. Miladinović, O.T. Schubert, W. Wolski, B.C. Collins, J. Malmström, L. Malmström, et al., *OpenSWATH enables automated, targeted analysis of data-independent acquisition MS data.* Nature Biotechnology, 2014. **32**: p. 219.

173. Meier, F., A.-D. Brunner, M. Frank, A. Ha, E. Voytik, S. Kaspar-Schoenefeld, M. Lubeck, O. Raether, R. Aebersold, B.C. Collins, et al., *Parallel accumulation – serial fragmentation combined with data-independent acquisition (diaPASEF): Bottom-up proteomics with near optimal ion usage.* bioRxiv, 2019: p. 656207.

174. Rosenberger, G., I. Bludau, U. Schmitt, M. Heusel, C.L. Hunter, Y. Liu, M.J. MacCoss, B.X. MacLean, A.I. Nesvizhskii, P.G.A. Pedrioli, et al., *Statistical control of peptide and protein error rates in large-scale targeted data-independent acquisition analyses.* Nat Methods, 2017. **14**(9): p. 921-927.

175. Larochelle, S., *Bottoms-up!* Nature Methods, 2017. **14**: p. 764.

176. Reubsaet, L., M.J. Sweredoski, and A. Moradian, *Data-Independent Acquisition for the Orbitrap Q Exactive HF: A Tutorial.* Journal of Proteome Research, 2019. **18**(3): p. 803-813.

177. Bantscheff, M., S. Lemeer, M.M. Savitski, and B. Kuster, *Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present.* Analytical and Bioanalytical Chemistry, 2012. **404**(4): p. 939-965.

178. Kanshin, E., P. Kubiniok, Y. Thattikota, D. D'Amours, and P. Thibault, *Phosphoproteome dynamics of Saccharomyces cerevisiae under heat shock and cold stress.* Mol Syst Biol, 2015. **11**(6): p. 813.

179. Kubiniok, P., H. Lavoie, M. Therrien, and P. Thibault, *Time-resolved Phosphoproteome Analysis of Paradoxical RAF Activation Reveals Novel Targets of ERK.* Mol Cell Proteomics, 2017. **16**(4): p. 663-679.

180. Cech, N.B. and C.G. Enke, *Relating Electrospray Ionization Response to Nonpolar Character of Small Peptides.* Analytical Chemistry, 2000. **72**(13): p. 2717-2723.

181. Gerber, S.A., J. Rush, O. Stemman, M.W. Kirschner, and S.P. Gygi, *Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS.* Proc Natl Acad Sci U S A, 2003. **100**(12): p. 6940-5.

182. Pratt, J.M., D.M. Simpson, M.K. Doherty, J. Rivers, S.J. Gaskell, and R.J. Beynon, *Multiplexed absolute quantification for proteomics using concatenated signature peptides encoded by QconCAT genes.* Nature Protocols, 2006. **1**(2): p. 1029-1043.

183. Brun, V., C. Masselon, J. Garin, and A. Dupuis, *Isotope dilution strategies for absolute quantitative proteomics.* Journal of Proteomics, 2009. **72**(5): p. 740-749.

184. Gevaert, K., F. Impens, B. Ghesquiere, P. Van Damme, A. Lambrechts, and J. Vandekerckhove, *Stable isotopic labeling in proteomics.* Proteomics, 2008. **8**(23-24): p. 4873-85.

185. Hogrebe, A., L. von Stechow, D.B. Bekker-Jensen, B.T. Weinert, C.D. Kelstrup, and J.V. Olsen, *Benchmarking common quantification strategies for large-scale phosphoproteomics.* Nature Communications, 2018. **9**(1): p. 1045.

186. Liu, H., R.G. Sadygov, and J.R. Yates, *A Model for Random Sampling and Estimation of Relative Protein Abundance in Shotgun Proteomics.* Analytical Chemistry, 2004. **76**(14): p. 4193-4201.

187. Wiśniewski, J.R., M.Y. Hein, J. Cox, and M. Mann, *A "Proteomic Ruler" for Protein Copy Number and Concentration Estimation without Spike-in Standards.* Molecular & Cellular Proteomics, 2014. **13**(12): p. 3497-3506.

188. Bateman, N.W., S.P. Goulding, N.J. Shulman, A.K. Gadok, K.K. Szumlinski, M.J. MacCoss, and C.C. Wu, *Maximizing Peptide Identification Events in Proteomic Workflows Using Data-Dependent Acquisition (DDA).* Molecular & Cellular Proteomics, 2014. **13**(1): p. 329-338.

189. Cox, J., M.Y. Hein, C.A. Luber, I. Paron, N. Nagaraj, and M. Mann, *Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ.* Mol Cell Proteomics, 2014. **13**(9): p. 2513-26.

190. Megger, D.A., T. Bracht, H.E. Meyer, and B. Sitek, *Label-free quantification in clinical proteomics.* Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics, 2013. **1834**(8): p. 1581-1590.

191. Al Shweiki, M.H.D.R., S. Mönchgesang, P. Majovsky, D. Thieme, D. Trutschel, and W. Hoehenwarter, *Assessment of Label-Free Quantification in Discovery Proteomics and Impact of Technological Factors and Natural Variability of Protein Abundance.* Journal of Proteome Research, 2017. **16**(4): p. 1410-1424.

192. Gygi, S.P., B. Rist, S.A. Gerber, F. Turecek, M.H. Gelb, and R. Aebersold, *Quantitative analysis of complex protein mixtures using isotope-coded affinity tags.* Nat Biotechnol, 1999. **17**(10): p. 994-9.

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

194. Wiese, S., K.A. Reidegeld, H.E. Meyer, and B. Warscheid, *Protein labeling by iTRAQ: A new tool for quantitative mass spectrometry in proteome research.* PROTEOMICS, 2007. **7**(3): p. 340-350.

195. Koniev, O. and A. Wagner, *Developments and recent advancements in the field of endogenous amino acid selective bond forming reactions for bioconjugation.* Chemical Society Reviews, 2015. **44**(15): p. 5495-5551.

196. Kalkhof, S. and A. Sinz, *Chances and pitfalls of chemical cross-linking with amine-reactive N-hydroxysuccinimide esters.* Analytical and Bioanalytical Chemistry, 2008. **392**(1): p. 305-312.

197. Inc., T.F.S. *TMTpro Label Reagents—higher multiplex quantitation for up to 16 samples*. Next generation of TMT labeling reagents for higher throughput and greater quantitative accuracy for multiplexed proteome analysis [Product Brochure] 2019; Available from: https://assets.thermofisher.com/TFS-Assets/BID/brochures/tmtpro-label-reagents-higher-multiplex-quantitation-brochure.pdf.

198. Schweppe, D.K., S. Prasad, M.W. Belford, J. Navarrete-Perea, D.J. Bailey, R. Huguet, M.P. Jedrychowski, R. Rad, G. McAlister, S.E. Abbatiello, et al., *Characterization and Optimization of Multiplexed Quantitative Analyses Using High-Field Asymmetric-Waveform Ion Mobility Mass Spectrometry.* Anal Chem, 2019. **91**(6): p. 4010-4016.

199. Li, H., J. Han, J. Pan, T. Liu, C.E. Parker, and C.H. Borchers, *Current trends in quantitative proteomics – an update.* Journal of Mass Spectrometry, 2017. **52**(5): p. 319-341.

200. Kelstrup, C.D., K. Aizikov, T.S. Batth, A. Kreutzman, D. Grinfeld, O. Lange, D. Mourad, A.A. Makarov, and J.V. Olsen, *Limits for Resolving Isobaric Tandem Mass Tag Reporter Ions Using Phase-Constrained Spectrum Deconvolution.* Journal of Proteome Research, 2018. **17**(11): p. 4008-4016.

201. Mertins, P., L.C. Tang, K. Krug, D.J. Clark, M.A. Gritsenko, L. Chen, K.R. Clauser, T.R. Clauss, P. Shah, M.A. Gillette, et al., *Reproducible workflow for multiplexed deep-scale*

proteome and phosphoproteome analysis of tumor tissues by liquid chromatography-mass spectrometry. Nat Protoc, 2018. **13**(7): p. 1632-1661.

202. Savitski, M.M., T. Mathieson, N. Zinn, G. Sweetman, C. Doce, I. Becher, F. Pachl, B. Kuster, and M. Bantscheff, *Measuring and Managing Ratio Compression for Accurate iTRAQ/TMT Quantification.* Journal of Proteome Research, 2013. **12**(8): p. 3586-3598.

203. Ting, L., R. Rad, S.P. Gygi, and W. Haas, *MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics.* Nature Methods, 2011. **8**: p. 937.

204. Erickson, B.K., M.P. Jedrychowski, G.C. McAlister, R.A. Everley, R. Kunz, and S.P. Gygi, *Evaluating multiplexed quantitative phosphopeptide analysis on a hybrid quadrupole mass filter/linear ion trap/orbitrap mass spectrometer.* Anal Chem, 2015. **87**(2): p. 1241-9.

205. Erickson, B.K., J. Mintseris, D.K. Schweppe, J. Navarrete-Perea, A.R. Erickson, D.P. Nusinow, J.A. Paulo, and S.P. Gygi, *Active Instrument Engagement Combined with a Real-Time Database Search for Improved Performance of Sample Multiplexing Workflows.* J Proteome Res, 2019. **18**(3): p. 1299-1306.

206. Schweppe, D.K., J.K. Eng, D. Bailey, R. Rad, Q. Yu, J. Navarrete-Perea, E.L. Huttlin, B.K. Erickson, J.A. Paulo, and S.P. Gygi, *Full-featured, real-time database searching platform enables fast and accurate multiplexed quantitative proteomics.* bioRxiv, 2019: p. 668533.

207. Zecha, J., S. Satpathy, T. Kanashova, S.C. Avanessian, M.H. Kane, K.R. Clauser, P. Mertins, S.A. Carr, and B. Kuster, *TMT Labeling for the Masses: A Robust and Cost-efficient, In-solution Labeling Approach.* Mol Cell Proteomics, 2019. **18**(7): p. 1468-1478.

208. Budnik, B., E. Levy, G. Harmange, and N. Slavov, *SCoPE-MS: mass spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation.* Genome Biology, 2018. **19**(1): p. 161.

209. Johansson, H.J., F. Socciarelli, N.M. Vacanti, M.H. Haugen, Y. Zhu, I. Siavelis, A. Fernandez-Woodbridge, M.R. Aure, B. Sennblad, M. Vesterlund, et al., *Breast cancer quantitative proteome and proteogenomic landscape.* Nature Communications, 2019. **10**(1): p. 1600.

210. Brenes, A., J.L. Hukelmann, D. Bensaddek, and A.I. Lamond, *Multi-batch TMT reveals false positives, batch effects and missing values.* Molecular & Cellular Proteomics, 2019: p. mcp.RA119.001472.

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

212. Zanivan, S., M. Krueger, and M. Mann, *In vivo quantitative proteomics: the SILAC mouse.* Methods Mol Biol, 2012. **757**: p. 435-50.

213. Chen, X., S. Wei, Y. Ji, X. Guo, and F. Yang, *Quantitative proteomics using SILAC: Principles, applications, and developments.* PROTEOMICS, 2015. **15**(18): p. 3175-3192.

214. Hoedt, E., G. Zhang, and T.A. Neubert, *Stable isotope labeling by amino acids in cell culture (SILAC) for quantitative proteomics.* Adv Exp Med Biol, 2014. **806**: p. 93-106.

215. Boisvert, F.-M., Y. Ahmad, M. Gierliński, F. Charrière, D. Lamont, M. Scott, G. Barton, and A.I. Lamond, *A Quantitative Spatial Proteomics Analysis of Proteome Turnover in Human Cells.* Molecular & Cellular Proteomics, 2012. **11**(3): p. M111.011429.

216. Snider, J., D. Wang, D.F. Bogenhagen, and J.D. Haley, *Pulse SILAC Approaches to the Measurement of Cellular Dynamics.* Adv Exp Med Biol, 2019. **1140**: p. 575-583.

217. Geiger, T., J. Cox, P. Ostasiewicz, J.R. Wisniewski, and M. Mann, *Super-SILAC mix for quantitative proteomics of human tumor tissue.* Nature Methods, 2010. **7**: p. 383.

218. Ma, B. and R. Johnson, *De Novo Sequencing and Homology Searching.* Molecular & Cellular Proteomics, 2012. **11**(2): p. O111.014902.

219. Cox, J., N. Neuhauser, A. Michalski, R.A. Scheltema, J.V. Olsen, and M. Mann, *Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment.* Journal of Proteome Research, 2011. **10**(4): p. 1794-1805.

220. Perkins, D.N., D.J.C. Pappin, D.M. Creasy, and J.S. Cottrell, *Probability-based protein identification by searching sequence databases using mass spectrometry data.* Electrophoresis, 1999. **20**(18): p. 3551-3567.

221. Eng, J.K., A.L. McCormack, and J.R. Yates, *An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database.* Journal of the American Society for Mass Spectrometry, 1994. **5**(11): p. 976-989.

222. Paulo, J.A., *Practical and Efficient Searching in Proteomics: A Cross Engine Comparison.* Webmedcentral, 2013. **4**(10).

223. Consortium, T.U., *UniProt: a hub for protein information.* Nucleic Acids Research, 2014. **43**(D1): p. D204-D212.

224. Giambruno, R., M. Mihailovich, and T. Bonaldi, *Mass Spectrometry-Based Proteomics to Unveil the Non-coding RNA World.* Frontiers in Molecular Biosciences, 2018. **5**(90).

225. Brunet, M.A., M. Brunelle, J.-F. Lucier, V. Delcourt, M. Levesque, F. Grenier, S. Samandi, S. Leblanc, J.-D. Aguilar, P. Dufour, et al., *OpenProt: a more comprehensive guide to explore eukaryotic coding potential and proteomes.* Nucleic Acids Research, 2018. **47**(D1): p. D403-D410.

226. Tress, M.L., F. Abascal, and A. Valencia, *Alternative Splicing May Not Be the Key to Proteome Complexity.* Trends in Biochemical Sciences, 2017. **42**(2): p. 98-110.

227. Ma, B., K. Zhang, and C. Liang, *An effective algorithm for peptide de novo sequencing from MS/MS spectra.* Journal of Computer and System Sciences, 2005. **70**(3): p. 418-430.

228. Zhang, J., L. Xin, B. Shan, W. Chen, M. Xie, D. Yuen, W. Zhang, Z. Zhang, G.A. Lajoie, and B. Ma, *PEAKS DB: de novo sequencing assisted database search for sensitive and accurate peptide identification.* Mol Cell Proteomics, 2012. **11**(4): p. M111.010587.

229. Elias, J.E. and S.P. Gygi, *Target-decoy search strategy for mass spectrometry-based proteomics.* Methods Mol Biol, 2010. **604**: p. 55-71.

230. Aggarwal, S. and A.K. Yadav, *False Discovery Rate Estimation in Proteomics.* Methods Mol Biol, 2016. **1362**: p. 119-28.

231. Cox, J., M.Y. Hein, C.A. Luber, I. Paron, N. Nagaraj, and M. Mann, *Accurate Proteome-wide Label-free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ.* Molecular & Cellular Proteomics, 2014. **13**(9): p. 2513-2526.

232. Cox, J. and M. Mann, *MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification.* Nature Biotechnology, 2008. **26**: p. 1367.

233. Cox, J., I. Matic, M. Hilger, N. Nagaraj, M. Selbach, J.V. Olsen, and M. Mann, *A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics.* Nat Protoc, 2009. **4**(5): p. 698-705.

234. Han, Y., B. Ma, and K. Zhang, *SPIDER: software for protein identification from sequence tags with de novo sequencing error.* J Bioinform Comput Biol, 2005. **3**(3): p. 697-716.

235. Han, X., L. He, L. Xin, B. Shan, and B. Ma, *PeaksPTM: Mass spectrometry-based identification of peptides with unspecified modifications.* J Proteome Res, 2011. **10**(7): p. 2930-6.

236. Bekker-Jensen, D.B., C.D. Kelstrup, T.S. Batth, S.C. Larsen, C. Haldrup, J.B. Bramsen, K.D. Sørensen, S. Høyer, T.F. Ørntoft, C.L. Andersen, et al., *An Optimized Shotgun Strategy for the Rapid Generation of Comprehensive Human Proteomes.* Cell Systems, 2017. **4**(6): p. 587-599.e4.

237. Spicer, V., P. Ezzati, H. Neustaeter, R.C. Beavis, J.A. Wilkins, and O.V. Krokhin, *3D HPLC-MS with Reversed-Phase Separation Functionality in All Three Dimensions for Large-Scale Bottom-Up Proteomics and Peptide Retention Data Collection.* Analytical Chemistry, 2016. **88**(5): p. 2847-2855.

238. Zhang, B., M. Pirmoradian, A. Chernobrovkin, and R.A. Zubarev, *DeMix Workflow for Efficient Identification of Cofragmented Peptides in High Resolution Data-dependent Tandem Mass Spectrometry.* Molecular & Cellular Proteomics, 2014. **13**(11): p. 3211-3223.

239. Bonneil, E., S. Pfammatter, and P. Thibault, *Enhancement of mass spectrometry performance for proteomic analyses using high-field asymmetric waveform ion mobility spectrometry (FAIMS).* Journal of Mass Spectrometry, 2015. **50**(11): p. 1181-1195.

240. Abshiru, N., O. Caron-Lizotte, R.E. Rajan, A. Jamai, C. Pomies, A. Verreault, and P. Thibault, *Discovery of protein acetylation patterns by deconvolution of peptide isomer mass spectra.* Nature Communications, 2015. **6**: p. 8648.

241. Courcelles, M., G. Bridon, S. Lemieux, and P. Thibault, *Occurrence and detection of phosphopeptide isomers in large-scale phosphoproteomics experiments.* J Proteome Res, 2012. **11**(7): p. 3753-65.

242. Ow, S.Y., M. Salim, J. Noirel, C. Evans, I. Rehman, and P.C. Wright, *iTRAQ Underestimation in Simple and Complex Mixtures: "The Good, the Bad and the Ugly".* Journal of Proteome Research, 2009. **8**(11): p. 5347-5355.

243. Myers, S.A., S. Klaeger, S. Satpathy, R. Viner, J. Choi, J. Rogers, K. Clauser, N.D. Udeshi, and S.A. Carr, *Evaluation of Advanced Precursor Determination for Tandem Mass Tag (TMT)-Based Quantitative Proteomics across Instrument Platforms.* Journal of Proteome Research, 2019. **18**(1): p. 542-547.

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

245. Vincent, C.E., G.K. Potts, A. Ulbrich, M.S. Westphall, J.A. Atwood, 3rd, J.J. Coon, and D.B. Weatherly, *Segmentation of precursor mass range using "tiling" approach increases peptide identifications for MS1-based label-free quantification.* Analytical chemistry, 2013. **85**(5): p. 2825-2832.

246. Murray Kermit, K., K. Boyd Robert, N. Eberlin Marcos, G.J. Langley, L. Li, and Y. Naito, *Definitions of terms relating to mass spectrometry (IUPAC Recommendations 2013)*, in *Pure and Applied Chemistry*. 2013. p. 1515.

247. Pirmoradian, M., H. Budamgunta, K. Chingin, B. Zhang, J. Astorga-Wells, and R.A. Zubarev, *Rapid and Deep Human Proteome Analysis by Single-dimension Shotgun Proteomics.* Molecular & Cellular Proteomics, 2013. **12**(11): p. 3330-3338.

248. Thakur, S.S., T. Geiger, B. Chatterjee, P. Bandilla, F. Fröhlich, J. Cox, and M. Mann, *Deep and Highly Sensitive Proteome Coverage by LC-MS/MS Without Prefractionation.* Molecular & Cellular Proteomics, 2011. **10**(8): p. M110.003699.

249. Manadas, B., V.M. Mendes, J. English, and M.J. Dunn, *Peptide fractionation in proteomics approaches.* Expert Review of Proteomics, 2010. **7**(5): p. 655-663.

250. Chakraborty, A.B. and S.J. Berger, *Optimization of reversed-phase peptide liquid chromatography ultraviolet mass spectrometry analyses using an automated blending methodology.* J Biomol Tech, 2005. **16**(4): p. 327-35.

251. Neverova, I. and J.E. Van Eyk, *Role of chromatographic techniques in proteomic analysis.* Journal of Chromatography B, 2005. **815**(1): p. 51-63.

252. Iwasaki, M., N. Sugiyama, N. Tanaka, and Y. Ishihama, *Human proteome analysis by using reversed phase monolithic silica capillary columns with enhanced sensitivity.* Journal of Chromatography A, 2012. **1228**: p. 292-297.

253. Grinias, K.M., J.M. Godinho, E.G. Franklin, J.T. Stobaugh, and J.W. Jorgenson, *Development of a 45kpsi ultrahigh pressure liquid chromatography instrument for gradient separations of peptides using long microcapillary columns and sub-2mum particles.* J Chromatogr A, 2016. **1469**: p. 60-67.

254. Giddings, J.C., *Maximum number of components resolvable by gel filtration and other elution chromatographic methods.* Analytical Chemistry, 1967. **39**(8): p. 1027-1028.

255. Shishkova, E., A.S. Hebert, M.S. Westphall, and J.J. Coon, *Ultra-High Pressure (>30,000 psi) Packing of Capillary Columns Enhancing Depth of Shotgun Proteomic Analyses.* Analytical Chemistry, 2018. **90**(19): p. 11503-11508.

256. Wilson, S.R., T. Vehus, H.S. Berg, and E. Lundanes, *Nano-LC in proteomics: recent advances and approaches.* Bioanalysis, 2015. **7**(14): p. 1799-815.

257. Shishkova, E., Alexander S. Hebert, and Joshua J. Coon, *Now, More Than Ever, Proteomics Needs Better Chromatography.* Cell Systems, 2016. **3**(4): p. 321-324.

258. Wolters, D.A., M.P. Washburn, and J.R. Yates, *An Automated Multidimensional Protein Identification Technology for Shotgun Proteomics.* Analytical Chemistry, 2001. **73**(23): p. 5683-5690.

259. Delmotte, N., M. Lasaosa, A. Tholey, E. Heinzle, and C.G. Huber, *Two-Dimensional Reversed-Phase × Ion-Pair Reversed-Phase HPLC: An Alternative Approach to High-Resolution Peptide Separation for Shotgun Proteome Analysis.* Journal of Proteome Research, 2007. **6**(11): p. 4363-4373.

260. Fournier, M.L., J.M. Gilmore, S.A. Martin-Brown, and M.P. Washburn, *Multidimensional Separations-Based Shotgun Proteomics.* Chemical Reviews, 2007. **107**(8): p. 3654-3686.

261. Wang, Y., F. Yang, M.A. Gritsenko, Y. Wang, T. Clauss, T. Liu, Y. Shen, M.E. Monroe, D. Lopez-Ferrer, T. Reno, et al., *Reversed-phase chromatography with multiple fraction concatenation strategy for proteome profiling of human MCF10A cells.* Proteomics, 2011. **11**(10): p. 2019-2026.

262. Winter, D.L., M.R. Wilkins, and W.A. Donald, *Differential Ion Mobility-Mass Spectrometry for Detailed Analysis of the Proteome.* Trends Biotechnol, 2019. **37**(2): p. 198-213.

263. Lanucara, F., S.W. Holman, C.J. Gray, and C.E. Eyers, *The power of ion mobility-mass spectrometry for structural characterization and the study of conformational dynamics.* Nature Chemistry, 2014. **6**: p. 281.

264. Meier, F., A.D. Brunner, S. Koch, H. Koch, M. Lubeck, M. Krause, N. Goedecke, J. Decker, T. Kosinski, M.A. Park, et al., *Online Parallel Accumulation-Serial Fragmentation (PASEF) with a Novel Trapped Ion Mobility Mass Spectrometer.* Mol Cell Proteomics, 2018. **17**(12): p. 2534-2545.

265. Meier, F., S. Beck, N. Grassl, M. Lubeck, M.A. Park, O. Raether, and M. Mann, *Parallel Accumulation-Serial Fragmentation (PASEF): Multiplying Sequencing Speed and Sensitivity by Synchronized Scans in a Trapped Ion Mobility Device.* J Proteome Res, 2015. **14**(12): p. 5378-87.

266. Pfammatter, S., E. Bonneil, F.P. McManus, S. Prasad, D.J. Bailey, M. Belford, J.J. Dunyach, and P. Thibault, *A Novel Differential Ion Mobility Device Expands the Depth of Proteome Coverage and the Sensitivity of Multiplex Proteomic Measurements.* Mol Cell Proteomics, 2018. **17**(10): p. 2051-2067.

267. Hebert, A.S., S. Prasad, M.W. Belford, D.J. Bailey, G.C. McAlister, S.E. Abbatiello, R. Huguet, E.R. Wouters, J.J. Dunyach, D.R. Brademan, et al., *Comprehensive Single-Shot Proteomics with FAIMS on a Hybrid Orbitrap Mass Spectrometer.* Anal Chem, 2018. **90**(15): p. 9529-9537.

268. Kaszycki, J.L., A. La Rotta, B. Colsch, F. Fenaille, C. Dauly, A. Kamleh, and C. Wu, *Separation of biologically relevant isomers on an Orbitrap mass spectrometer using high-resolution drift tube ion mobility and varied drift gas mixtures.* Rapid Communications in Mass Spectrometry, 2019. **33**(S2): p. 3-10.

269. Gray, C.J., B. Thomas, R. Upton, L.G. Migas, C.E. Eyers, P.E. Barran, and S.L. Flitsch, *Applications of ion mobility mass spectrometry for high throughput, high resolution glycan analysis.* Biochim Biophys Acta, 2016. **1860**(8): p. 1688-709.

270. Revercomb, H.E. and E.A. Mason, *Theory of plasma chromatography/gaseous electrophoresis. Review.* Analytical Chemistry, 1975. **47**(7): p. 970-983.

271. Gabelica, V., A.A. Shvartsburg, C. Afonso, P. Barran, J.L.P. Benesch, C. Bleiholder, M.T. Bowers, A. Bilbao, M.F. Bush, J.L. Campbell, et al., *Recommendations for reporting ion mobility Mass Spectrometry measurements.* Mass Spectrom Rev, 2019. **38**(3): p. 291-320.

272. Eiceman, G.A., Z. Karpas, and H.H. Hill Jr, *Ion mobility spectrometry*. 2013: CRC press.

273. Cumeras, R., E. Figueras, C.E. Davis, J.I. Baumbach, and I. Gràcia, *Review on ion mobility spectrometry. Part 1: current instrumentation.* The Analyst, 2015. **140**(5): p. 1376-1390.

274. Winter, D.L., M.R. Wilkins, and W.A. Donald, *Differential Ion Mobility–Mass Spectrometry for Detailed Analysis of the Proteome.* Trends in Biotechnology, 2019. **37**(2): p. 198-213.

275. Kirk, A.T., A. Bohnhorst, C.R. Raddatz, M. Allers, and S. Zimmermann, *Ultra-high-resolution ion mobility spectrometry-current instrumentation, limitations, and future developments.* Anal Bioanal Chem, 2019.

276. Baker, E.S., B.H. Clowers, F. Li, K. Tang, A.V. Tolmachev, D.C. Prior, M.E. Belov, and R.D. Smith, *Ion mobility spectrometry-mass spectrometry performance using electrodynamic ion funnels and elevated drift gas pressures.* Journal of the American Society for Mass Spectrometry, 2007. **18**(7): p. 1176-1187.

277. Smith, D.P., T.W. Knapman, I. Campuzano, R.W. Malham, J.T. Berryman, S.E. Radford, and A.E. Ashcroft, *Deciphering Drift Time Measurements from Travelling Wave Ion Mobility Spectrometry-Mass Spectrometry Studies.* European Journal of Mass Spectrometry, 2009. **15**(2): p. 113-130.

278. Giles, K., S.D. Pringle, K.R. Worthington, D. Little, J.L. Wildgoose, and R.H. Bateman, *Applications of a travelling wave-based radio-frequency-only stacked ring ion guide.* Rapid Communications in Mass Spectrometry, 2004. **18**(20): p. 2401-2414.

279. Shvartsburg, A.A. and R.D. Smith, *Fundamentals of traveling wave ion mobility spectrometry.* Anal Chem, 2008. **80**(24): p. 9689-99.

280. Fernandez-Lima, F.A., D.A. Kaplan, and M.A. Park, *Note: Integration of trapped ion mobility spectrometry with mass spectrometry.* Review of Scientific Instruments, 2011. **82**(12): p. 126106.

281. Silveira, J.A., M.E. Ridgeway, F.H. Laukien, M. Mann, and M.A. Park, *Parallel accumulation for 100% duty cycle trapped ion mobility-mass spectrometry.* International Journal of Mass Spectrometry, 2017. **413**: p. 168-175.

282. Silveira, J.A., K. Michelmann, M.E. Ridgeway, and M.A. Park, *Fundamentals of Trapped Ion Mobility Spectrometry Part II: Fluid Dynamics.* J Am Soc Mass Spectrom, 2016. **27**(4): p. 585-95.

283. Silveira, J.A., M.E. Ridgeway, and M.A. Park, *High Resolution Trapped Ion Mobility Spectrometery of Peptides.* Analytical Chemistry, 2014. **86**(12): p. 5624-5627.

284. Shvartsburg, A.A., F. Li, K. Tang, and R.D. Smith, *High-Resolution Field Asymmetric Waveform Ion Mobility Spectrometry Using New Planar Geometry Analyzers.* Analytical Chemistry, 2006. **78**(11): p. 3706-3714.

285. Michael Belford, S.P., Jean-Jacques Dunyach, *Characterization of FAIMS Waveform With Regards to Amplitude, Frequency, Phase, and Electrode Temperature*, in *Poster Note 64438*, S.J. Thermo Fisher Scientific, CA, Editor. 2015, Thermo Fisher Scientific, San Jose, CA.

286. Kolakowski, B.M. and Z. Mester, *Review of applications of high-field asymmetric waveform ion mobility spectrometry (FAIMS) and differential mobility spectrometry (DMS).* Analyst, 2007. **132**(9): p. 842-64.

287. Shvartsburg, A.A., F. Li, K. Tang, and R.D. Smith, *High-resolution field asymmetric waveform ion mobility spectrometry using new planar geometry analyzers.* Anal Chem, 2006. **78**(11): p. 3706-14.

288. Shvartsburg, A.A., K. Tang, R.D. Smith, M. Holden, M. Rush, A. Thompson, and D. Toutoungi, *Ultrafast Differential Ion Mobility Spectrometry at Extreme Electric Fields Coupled to Mass Spectrometry.* Analytical Chemistry, 2009. **81**(19): p. 8048-8053.

289. Swearingen, K.E. and R.L. Moritz, *High-field asymmetric waveform ion mobility spectrometry for mass spectrometry-based proteomics.* Expert review of proteomics, 2012. **9**(5): p. 505-517.

290. Guevremont, R. and R.W. Purves, *Atmospheric pressure ion focusing in a high-field asymmetric waveform ion mobility spectrometer.* Review of Scientific Instruments, 1999. **70**(2): p. 1370-1383.

291. Krylov, E.V., *A method of reducing diffusion losses in a drift spectrometer.* Technical Physics, 1999. **44**(1): p. 113-116.

292. Cumeras, R., E. Figueras, C.E. Davis, J.I. Baumbach, and I. Gràcia, *Review on Ion Mobility Spectrometry. Part 2: hyphenated methods and effects of experimental parameters.* Analyst, 2015. **140**(5): p. 1391-1410.

293. Barnett, D.A., M. Belford, J.-J. Dunyach, and R.W. Purves, *Characterization of a temperature-Controlled FAIMS system.* Journal of the American Society for Mass Spectrometry, 2007. **18**(9): p. 1653-1663.

294. Purves, R.W. and R. Guevremont, *Electrospray Ionization High-Field Asymmetric Waveform Ion Mobility Spectrometry−Mass Spectrometry.* Analytical Chemistry, 1999. **71**(13): p. 2346-2357.

295. Guevremont, R. and R. Purves, *Comparison of experimental and calculated peak shapes for three cylindrical geometry FAIMS prototypes of differing electrode diameters.* Journal of the American Society for Mass Spectrometry, 2005. **16**(3): p. 349-362.

296. Prasad, S., M.W. Belford, J.J. Dunyach, and R.W. Purves, *On an aerodynamic mechanism to enhance ion transmission and sensitivity of FAIMS for nano-electrospray ionization-mass spectrometry.* J Am Soc Mass Spectrom, 2014. **25**(12): p. 2143-53.

297. Canterbury, J.D., X. Yi, M.R. Hoopmann, and M.J. MacCoss, *Assessing the dynamic range and peak capacity of nanoflow LC-FAIMS-MS on an ion trap mass spectrometer for proteomics.* Anal Chem, 2008. **80**(18): p. 6888-97.

298. Guevremont, R. and R.W. Purves, *High field asymmetric waveform ion mobility spectrometry-mass spectrometry: an investigation of leucine enkephalin ions produced by electrospray ionization.* J Am Soc Mass Spectrom, 1999. **10**(6): p. 492-501.

299. Guevremont, R., D.A. Barnett, R.W. Purves, and J. Vandermey, *Analysis of a tryptic digest of pig hemoglobin using ESI-FAIMS-MS.* Anal Chem, 2000. **72**(19): p. 4577-84.

300. Barnett, D.A., B. Ells, R. Guevremont, and R.W. Purves, *Application of ESI-FAIMS-MS to the analysis of tryptic peptides.* J Am Soc Mass Spectrom, 2002. **13**(11): p. 1282-91.

301. Barnett, D.A., L. Ding, B. Ells, R.W. Purves, and R. Guevremont, *Tandem mass spectra of tryptic peptides at signal-to-background ratios approaching unity using electrospray ionization high-field asymmetric waveform ion mobility spectrometry/hybrid quadrupole time-of-flight mass spectrometry.* Rapid Communications in Mass Spectrometry, 2002. **16**(7): p. 676-680.

302. Purves, R.W., D.A. Barnett, and R. Guevremont, *Separation of protein conformers using electrospray-high field asymmetric waveform ion mobility spectrometry-mass spectrometry.* International Journal of Mass Spectrometry, 2000. **197**(1): p. 163-177.

303. Purves, R.W., D.A. Barnett, B. Ells, and R. Guevremont, *Investigation of bovine ubiquitin conformers separated by high-field asymmetric waveform ion mobility spectrometry: cross section measurements using energy-loss experiments with a triple quadrupole mass spectrometer.* J Am Soc Mass Spectrom, 2000. **11**(8): p. 738-45.

304. Venne, K., E. Bonneil, K. Eng, and P. Thibault, *Improvement in Peptide Detection for Proteomics Analyses Using NanoLC−MS and High-Field Asymmetry Waveform Ion Mobility Mass Spectrometry.* Analytical Chemistry, 2005. **77**(7): p. 2176-2186.

305. Wolters, D.A., M.P. Washburn, and J.R. Yates, 3rd, *An automated multidimensional protein identification technology for shotgun proteomics.* Anal Chem, 2001. **73**(23): p. 5683-90.

306. Xia, Y.-Q., S.T. Wu, and M. Jemal, *LC-FAIMS-MS/MS for Quantification of a Peptide in Plasma and Evaluation of FAIMS Global Selectivity from Plasma Components.* Analytical Chemistry, 2008. **80**(18): p. 7137-7143.

307. Klaassen, T., S. Szwandt, J.T. Kapron, and A. Roemer, *Validated quantitation method for a peptide in rat serum using liquid chromatography/high-field asymmetric waveform ion mobility spectrometry.* Rapid Commun Mass Spectrom, 2009. **23**(15): p. 2301-6.

308. Saba, J., E. Bonneil, C. Pomiès, K. Eng, and P. Thibault, *Enhanced Sensitivity in Proteomics Experiments Using FAIMS Coupled with a Hybrid Linear Ion Trap/Orbitrap Mass Spectrometer.* Journal of Proteome Research, 2009. **8**(7): p. 3355-3366.

309. Xuan, Y., A.J. Creese, J.A. Horner, and H.J. Cooper, *High-field asymmetric waveform ion mobility spectrometry (FAIMS) coupled with high-resolution electron transfer dissociation mass spectrometry for the analysis of isobaric phosphopeptides.* Rapid Communications in Mass Spectrometry, 2009. **23**(13): p. 1963-1969.

310. Creese, A.J., J. Smart, and H.J. Cooper, *Large-Scale Analysis of Peptide Sequence Variants: The Case for High-Field Asymmetric Waveform Ion Mobility Spectrometry.* Analytical Chemistry, 2013. **85**(10): p. 4836-4843.

311. Zhao, H., D.L. Cunningham, A.J. Creese, J.K. Heath, and H.J. Cooper, *FAIMS and Phosphoproteomics of Fibroblast Growth Factor Signaling: Enhanced Identification of Multiply Phosphorylated Peptides.* J Proteome Res, 2015. **14**(12): p. 5077-87.

312. Creese, A.J., N.J. Shimwell, K.P. Larkins, J.K. Heath, and H.J. Cooper, *Probing the complementarity of FAIMS and strong cation exchange chromatography in shotgun proteomics.* J Am Soc Mass Spectrom, 2013. **24**(3): p. 431-43.

313. Bridon, G., E. Bonneil, T. Muratore-Schroeder, O. Caron-Lizotte, and P. Thibault, *Improvement of phosphoproteome analyses using FAIMS and decision tree fragmentation. application to the insulin signaling pathway in Drosophila melanogaster S2 cells.* J Proteome Res, 2012. **11**(2): p. 927-40.

314. Creese, A.J. and H.J. Cooper, *Separation and Identification of Isomeric Glycopeptides by High Field Asymmetric Waveform Ion Mobility Spectrometry.* Analytical Chemistry, 2012. **84**(5): p. 2597-2601.

315. Ulasi, G.N., A.J. Creese, S.X. Hui, C.W. Penn, and H.J. Cooper, *Comprehensive mapping of O-glycosylation in flagellin from Campylobacter jejuni 11168: A multienzyme differential ion mobility mass spectrometry approach.* Proteomics, 2015. **15**(16): p. 2733-45.

316. Swearingen, K.E., M.R. Hoopmann, R.S. Johnson, R.A. Saleem, J.D. Aitchison, and R.L. Moritz, *Nanospray FAIMS Fractionation Provides Significant Increases in Proteome Coverage of Unfractionated Complex Protein Digests.* Molecular & Cellular Proteomics, 2012. **11**(4): p. M111.014985.

317. Barnett, D.A. and R.J. Ouellette, *Elimination of the helium requirement in high-field asymmetric waveform ion mobility spectrometry (FAIMS): beneficial effects of decreasing the analyzer gap width on peptide analysis.* Rapid Communications in Mass Spectrometry, 2011. **25**(14): p. 1959-1971.

318. Swearingen, K.E., J.M. Winget, M.R. Hoopmann, U. Kusebauch, and R.L. Moritz, *Decreased Gap Width in a Cylindrical High-Field Asymmetric Waveform Ion Mobility Spectrometry Device Improves Protein Discovery.* Anal Chem, 2015. **87**(24): p. 12230-7.

CHAPTER TWO

2. Improvement of quantitative measurements in multiplex proteomics using high-field asymmetric waveform spectrometry (FAIMS)

Sibylle Pfammatter[1,2], Eric Bonneil[1], and Pierre Thibault[1,2]*

[1]Institute for Research in Immunology and Cancer, Université de Montréal, C.P. 6128, Succursale centre-ville, Montréal, Québec, H3C 3J7, Canada.

[2]Department of Chemistry, Université de Montréal, C.P. 6128, Succursale centre-ville, Montréal, Québec, H3C 3J7, Canada.

Published:

Journal of Proteome Research, 2016, 15 (12), pp 4653–4665

Reprinted with permission from JOURNAL OF PROTEOME RESEARCH.

Copyright (2019) American Chemical Society

Author contributions:

Sibylle Pfammatter designed, executed and analyzed all experiments, generated all figures and wrote the paper. Eric Bonneil (MS platform manager) helped to design the project, supported the data acquisition and analysis, and contributed to the draft of the manuscript. Pierre Thibault (supervisor) managed the project. All authors approved the content and submission of the paper.

2.1. Abstract

Quantitative proteomics using isobaric reagents tandem mass tags (TMT) or isobaric tags for relative and absolute quantitation (iTRAQ) provide a convenient approach to compare changes in protein abundances across multiple samples. However, the analysis of complex protein digests by isobaric labeling can be undermined by the relative large proportion of co-selected peptide ions that lead to distorted reporter ion ratios and affect the accuracy and precision of quantitative measurements. Here, we investigated the use of high-field asymmetric waveform ion mobility spectrometry (FAIMS) in proteomic experiments to reduce sample complexity and improve protein quantification using TMT isobaric labeling. LC-FAIMS-MS/MS analyses of human and yeast protein digests led to significant reduction of interfering ions which increased the number of quantifiable peptides by up to 68 % while significantly improving the accuracy of abundance measurements compared to conventional LC-MS/MS. The improvement of quantitative measurements using FAIMS is further demonstrated for the temporal profiling of protein abundance of HEK293 cells following heat shock treatment.

2.2. Introduction

Recent advances in mass spectrometry (MS)–based proteomics have facilitated the large-scale identification of proteins to unravel the dynamic nature of the proteome and the expanding repertoire of protein modifications [1]. The improvements in speed, sensitivity, and resolution of MS instruments enable the analysis of the whole proteome of an organism such as *Saccharomyces cerevisiae* in just over an hour [2]. Furthermore, the dynamic range of MS instruments makes it possible to identify and quantify thousands of peptides in a single experiment to profile the changes in protein abundance across different biological conditions [3, 4]. When combined with affinity chromatography, these advances also facilitated the monitoring of cell signaling events with second to minute temporal resolution [5-7].

Quantitative proteomics can be performed using three different methods: label-free quantification [8, 9], metabolic labeling with stable isotope labeling by amino acids in cell culture (SILAC) [10-12], NeuCode SILAC [13], and stable-isotope labeling using chemical reagents that are covalently attached *in vitro* such as dimethyl-labeling [14], tandem mass tags (TMTs) [15] and isobaric tags for relative and absolute quantification (iTRAQ) [16]. Each of these approaches has its own merits and limitations for quantitative proteomics. SILAC methods provide unmatched accuracy and robustness as most variations attributed to sample-handling and MS analyses are minimized. However, this approach is mostly limited to cell cultures since numerous samples including clinical specimens cannot be incorporated by metabolic labeling. In contrast, label-free quantification is in principle applicable to any kind of sample, and can be used to analyze a large number of samples compared to labeling methods that are limited by the availability of isotopically labeled amino acids or reagents [17]. Label-free also provides higher proteome coverage for identification, but its quantification performance in terms of reproducibility and number of quantifiable peptides is usually lower than labeling-based approaches [18]. While both metabolic and chemical labeling can provide accurate, precise, and reproducible quantification of many proteins, only chemical labeling offers the capability to multiplex sample analysis up to 10-plex, which simplifies experimental design and reduces overall analysis time.

Quantitative proteomics via isobaric chemical tags provide an interesting avenue to enhance sample throughput while maintaining quantification performances. Chemical labeling enables parallel quantification through the monitoring of reporter fragment ions originating from isobaric precursor ions of different samples. Although elegant by its simplicity, this approach has been limited by the co-isolation and co-fragmentation of interfering ions within the target peptide isolation window that results in distorted reporter ion ratio. The impact of interfering ions on quantitative measurements has been reported by several groups [19-21]. Different strategies have been implemented to alleviate this issue. For instance, proton-transfer ion-ion reactions have been used to reduce the charge state of precursor ions, remove interferences, and improve accuracy of quantitative measurements [22]. Ting *et al.* demonstrated that triple-stage mass spectrometry (MS^3) almost completely eliminates interfering ions in complex samples [23]. While this approach successfully restored the quantitative performance of isobaric tagging, it initially suffered from impaired sensitivity due to the selection of a single MS^2 ion for MS^3 fragmentation. A similar method called Multinotch MS^3 was introduced to select and co-isolate two or more MS^2 product ions by the application of attenuated ion selection waveforms [24]. Those product ions are fragmented to generate several second-generation fragment ions including isobaric tags to improve quantitative accuracy and sensitivity of the MS experiment up to n-fold, where n is the number of MS fragments selected and simultaneously isolated [24]. Wuhr *et al.* performed quantification based on the fragment ion cluster that carries most of the TMT mass balance portion, an approach that also reduced interference as ions used for quantification are dependent on the precursor mass and charge state [25]. Although these strategies have significant analytical merits, they do extend duty cycle and require a hybrid ion trap instrument to perform multi-stage fragmentation.

In an effort to improve the accuracy of isobaric labeling, we investigated the analytical merits of high field asymmetric waveform ion mobility spectrometry (FAIMS), a form of ion mobility that exploits ion separation at low and high electric fields [26, 27]. In FAIMS, ions are transported by a carrier gas in a gap between two electrodes to which is applied a high voltage asymmetric waveform. Ions of different mobilities are transmitted in turn by scanning the compensation voltage (CV) voltage [28]. Ion

prefractionation using FAIMS enables the removal of singly charged ions and interfering isobaric species, and can be easily interfaced to MS instruments for targeted and data dependent MS/MS acquisition. Previous studies also reported an increase in sensitivity of cylindrical FAIMS at higher dispersion voltages through enhanced ion focusing [29, 30]. FAIMS has provided significant benefits by improving the depth of proteomics analyses and reducing the extent of precursor ions co-fragmentation that lead to chimeric MS/MS spectra and impede the identification of low abundance peptide ions [31-37]. FAIMS is also complementary to chromatography as it can separate positional isomers [38-41]. FAIMS has been successfully used for peptide quantitation in rat serum [42, 43]. The capability of FAIMS to reduce the contribution of interfering ions enabled a 18-fold improvement in the signal to noise ratio (S/N) with the lower limit of quantitation (LLOQ) and slightly better accuracy and precision. We have also successfully used FAIMS for large-scale proteomics quantitation enabling us to quantify differential regulation of low abundance proteins such as integral membrane proteins [34] and phosphoproteins [44]. In the present study, we evaluate the benefits of FAIMS to enhance the accuracy and precision of measurements using TMT-based quantitative proteomics. We also demonstrate its application to profile the temporal changes in protein abundance of HEK293 cells following heat shock to gain biological insights on cell response essential to adaptation and survival.

2.3. Experimental Section

Cell culture. Yeast strain D504 was grown in Yeast Extract Peptone Dextrose Medium (Sigma-Aldrich) until OD_{600} = 0.8. HEK293 strain was grown in Dulbecco's Modified Eagles Medium (high glucose, GE Healthcare HyClone) supplemented with 10% fetal bovine serum (Seradigm VWR Life Science), 1% Penicillin/Streptomycin Solution (Gibco Thermo Scientific) and 1% L-Glutamine (Gibco Thermo Scientific) and cultured at 37 °C in a 5% CO_2 constant atmosphere. For the heat shock treatment, the medium was removed and replaced by fresh medium at 43°C. Flasks with cultures were placed in an incubator at 43°C. Cells were collected at 0, 2, 4, 6, 8 and 10 h, washed twice with cold PBS (GE Healthcare HyClone) and then stored overnight at -30°C. After incubation with phosphatase and protease (Sigma-Aldrich) inhibitors in 50 mM Tris 8M Urea (both VWR Life Science), cells were mechanically lysed with short sonicate pulses. After centrifugation, supernatants were transferred in a new tube and protein concentration was determined by Bradford assay (Bio-Rad).

Protein digestion and TMT labeling. Protein pellets (500 µg) were resuspended in 100 mM triethyl ammonium bicarbonate (TEAB) (Sigma-Aldrich) with 5 mM Tris (2-Carboxyethyl) phosphine Hydrochloride (TCEP) (Thermo Scientific) and vortexed for 1 hr at 55°C. After addition of 200 µL of 50 mM of chloroacetamide (Sigma-Aldrich) in 100 mM of TEAB, the samples were vortexed for 1 hr at room temperature. Samples were digested overnight with trypsin at 37°C (Promega) (enzyme/protein w/w ratio of 1:25) and dried down in a speed-vac. Samples were reconstituted at 1 µg/µL with 100 mM TEAB. The TMT (Thermo Fisher Scientific) reagents were dissolved in 40 µL of anhydrous acetonitrile (Sigma-Aldrich) and added to 100 µg of peptides. After incubation overnight at room temperature, the reaction was quenched by adding 8 µL of 5% w/v hydroxylamine (Thermo Fisher Scientific). Samples were desalted on HLB columns (Oasis, Waters), dried down and reconstituted in formic acid (EMD Millipore) 0.2% at 1 µg/µL. Following labeling, yeast aliquots were mixed at ratio of 0:2.5:4:10:4:2.5, and HEK293 extracts were mixed at 4:6:4:0:0:0 (*Figure 2-1A*). Those two samples were then mixed at a 1:1 w/w ratio. For the heat shock treatment, peptide

digests from separate time points were labeled with TMT reagents, and mixed at a 1:1:1:1:1:1 weight ratio.

Mass spectrometry. LC-MS/MS analyses were performed on a nano-LC 2D pump (Eksigent) coupled to a LTQ-Orbitrap Elite hybrid mass spectrometer (Thermo Scientific). Tryptic peptides were loaded on a 360 µm ID × 4 mm, C_{18} trap column prior to separation on a 150 µm ID x 20 cm LC column (Jupiter C_{18}, 3 µm, 300 Å, Phenomenex). For the 2-proteome (yeast/human) and HEK293 heat shock experiments peptides were separated on a Optiguard SCX trap column, 5 µm, 300Å, 0.5 ID × 23 mm (Optimize Technologies) and eluted on-line to a 360 µm ID × 4 mm, C_{18} trap column prior to separation on a 150 µm ID x 20 cm LC column (Jupiter C_{18}, 3 µm, 300 Å, Phenomenex). Tryptic digests were loaded on the SCX trap and sequentially eluted using salt plugs of 0, 250, 500, 750, 1000, and 2000 mM ammonium acetate, pH 3.5. For 1DLC and 2DLC analyses, peptides were separated on the analytical column using a linear gradient of 5–40% acetonitrile (0.2% formic acid) in 53 min with a flow rate of 600 nL/min. MS survey scan were acquired at a resolution of 60,000 while tandem mass spectra were acquired in either CID mode for non TMT labeled sample or HCD mode at a resolution of 15,000 for TMT-labeled samples. For LC-MS/MS acquisition each duty cycle consists of 1 survey scan and 12 tandem mass spectra. Isolation window widths was set to 2 Th except for the precursor intensity fraction evaluation (*Figure 2-1*) where analyses were performed with isolation window widths of 0.5, 1, 2, 4, 8, 16, 20 Th. Charge states of 2 and higher were selected for MS/MS.

FAIMS. The FAIMS interface was coupled to the LTQ-Orbitrap Elite via a nanoelectropray source. The FAIMS electrodes are cylindrical with a 1.5 mm gap and were operated at a dispersion voltage (DV) of -5000 V with a 2.5 L/min flow of nitrogen and a CV dwell time of 50 ms as described here [45]. We did not use the usual mixture of N_2/He (50/50) because of the corona discharge occurring with the 33kV/cm electric field provided by the 1.5 mm gap. The temperature of the inner and outer electrodes was set to 70 and 90^{o}C, respectively. For CV stepping, the CV transmission range was determined by infusion of a 8-protein tryptic digest at 200 fmol/µl. For the experiment on *Figure 2-1* we stepped from CV -34V to -42V with 2V increments. For the 2-proteome experiment, the duty cycle was made of one MS followed by 5 MS/MS for 2 CV values.

In this experiment we stepped from CV -31V to -63 V with -2.5V increments. For the heat shock experiment, the duty cycle was made of one MS followed by 3 MS/MS for 3 CV values. In this experiment we stepped from CV -32V to -74 V with -3V increments.

Data processing. MS data were analyzed using the Xcalibur software (version 2.1). Database searches were performed using PEAKS 7.0 search engine (Bioinformatics Solutions Inc.). Searches were conducted using either a Uniprot human database or merged Uniprot databases of yeast and human containing 74,508 entries and 146,661 entries respectively (Uniprot release 1302S). The error window for precursor and fragment ion mass values was set to 10 ppm and 0.01 Da, respectively. The number of allowed missed cleavage sites for trypsin was set to 2 and phosphorylation (STY), oxidation (M), deamidation (NQ) and carbamidomethylation (C) were all selected as variable modifications. TMT-labeled peptide amino terminus and TMT-labeled lysine (+ 229.163 Da) were also set as variable modifications. The false discovery rate (FDR) for peptide was set to 1% by applying the target-decoy strategy [46]. TMT reporter ion isotopic distributions were corrected based on the lot product data sheet (Thermo Fisher Scientific). Clustering of dynamic profiles of 2 technical FAIMS replicates from the heat shock dataset was performed with R (http://www.r-project.org/) with the Mfuzz package [47] with additional filtering (MS/MS spectra with less than 5 out of 6 TMT reporter ions were discarded) as described previously [6]. The protein-protein network was built in STRING (http://string-db.org, version 10), and experimental predictions of high confidence (0.700) were transferred in Cytoscape v3.2.1 for network visualization. Gene ontology enrichment for Biological Processes was performed on the human proteome using the Database for Annotation, Visualization and Integrated Discovery (DAVID Bioinformatics Resources 6.7) [48].

2.4. Results and Discussion

2.4.1. Impact of precursor co-fragmentation on protein identification

Large-scale proteomics analyses typically rely on two different LC-MS/MS acquisition strategies namely data-dependent acquisition (DDA) and data-independent acquisition (DIA) [49], while targeted acquisition methods traditionally used selected reaction monitoring (SRM) [50] or, more recently, parallel reaction monitoring (PRM) [51]. The generation of fragment ions from these different approaches requires the selection of appropriate precursor ion isolation windows that vary from *m/z* 0.7 to 20. In view of the inherent sample complexity found in proteomic analyses, these isolation windows inevitably lead to co-fragmentation of precursor ions that share overlapping elution time of *m/z* ranges. In DDA experiments, the relative proportion of MS/MS spectra that show precursor ion co-fragmentation can be as high as 50% [52]. The impact of chimeric MS/MS spectra reduces the number of peptide identifications by increasing false negatives (i.e. peptides rejected due to poor scores) with modest increase in false positives (i.e., by increasing score thresholds at a given false discovery rate) [53]. Furthermore, precursor co-fragmentation can also affect targeted proteomics analyses when fragment ions used for quantitative measurements arise from different peptides.

For isobaric labeling, precursor co-fragmentation can impact both protein identification and quantification. Recent reports on TMT-labeled peptides indicated that FAIMS significantly decreased the extent of co-fragmentation in large-scale proteomic analyses by transmitting isobaric precursor ions at different CV values [31]. To determine the extent of co-fragmentation and its impact on protein identification, we first compared LC-MS/MS analyses of a HEK293 tryptic digest performed with and without FAIMS for different isolation windows (***Figure 2-1***). We evaluated the proportion of ion current in the isolation window associated to the target ion using the precursor ion fraction (PIF) available in Maxquant [54]. A PIF value approaching one indicates that most of the ion current comes from the precursor of interest and contains a low level of contaminating ions. ***Figure 2-1A*** shows the PIF values for the doubly protonated ion of the tryptic peptide DAGTIAGLNVLR at *m/z* 600.34 for isolation windows of *m/z* 0.7, 2

and 20 typically used for PRM, DDA and DIA experiments, respectively. PIF values ranged from 0.15 to 0.74 without FAIMS and 0.67 to 0.99 for FAIMS with higher values observed for decreasing isolation windows. These results suggest that LC-FAIMS-MS/MS experiments reduced the proportion of precursor co-isolation by up to 4-fold compared to non-FAIMS experiments.

Next, we evaluated the systematic changes in PIF values for a higher number of isolation windows ranging from *m/z* 0.5 to 20 (***Figure 2-1B***). A total of five injections were obtained for LC-MS/MS experiments performed with and without FAIMS. As observed, FAIMS consistently yielded 30-120% higher PIF values, even for isolation windows as small as *m/z* 0.5. To evaluate the impact of PIF values on the extent of precursor ion co-fragmentation, we determined the distribution of unique peptide identifications for different PIF values and isolation windows (***Supplementary Figure 2-1***). No significant change in the distribution of unique peptide identification is observed for isolation windows below *m/z* 2.0 when using FAIMS with median PIF values higher than 0.75. In contrast, experiments performed without FAIMS led to a 30-40% reduction in number of identifications for isolation windows as small as 0.7. A higher number of identification was observed for an isolation window of *m/z* 2.0 in both FAIMS and non-FAIMS experiments (***Figure 2-1C***). For this isolation window, FAIMS led to the identification of 10,257 unique peptides compared to 6,379 peptides for non-FAIMS. As expected, the number of identifications gradually decreased for larger *m/z* windows due to the increasing number of precursor ions co-fragmentation. The proportion of identified MS/MS spectra over the range of *m/z* 2-20 also decreased from 42 to 6 %, and from 30 to 6 % for FAIMS and non-FAIMS, respectively. We also noted a decrease in the number of identifications for isolation windows below *m/z* 2 due to lower ion transmission and sensitivity for both experimental setups. FAIMS also provided a greater increase in the number of unique peptide identification across CV steps irrespective of *m/z* window compared to repeated injections (***Supplementary Figure 2-2***). Accordingly, an isolation window of 2.0 was selected for subsequent LC-MS/MS experiments.

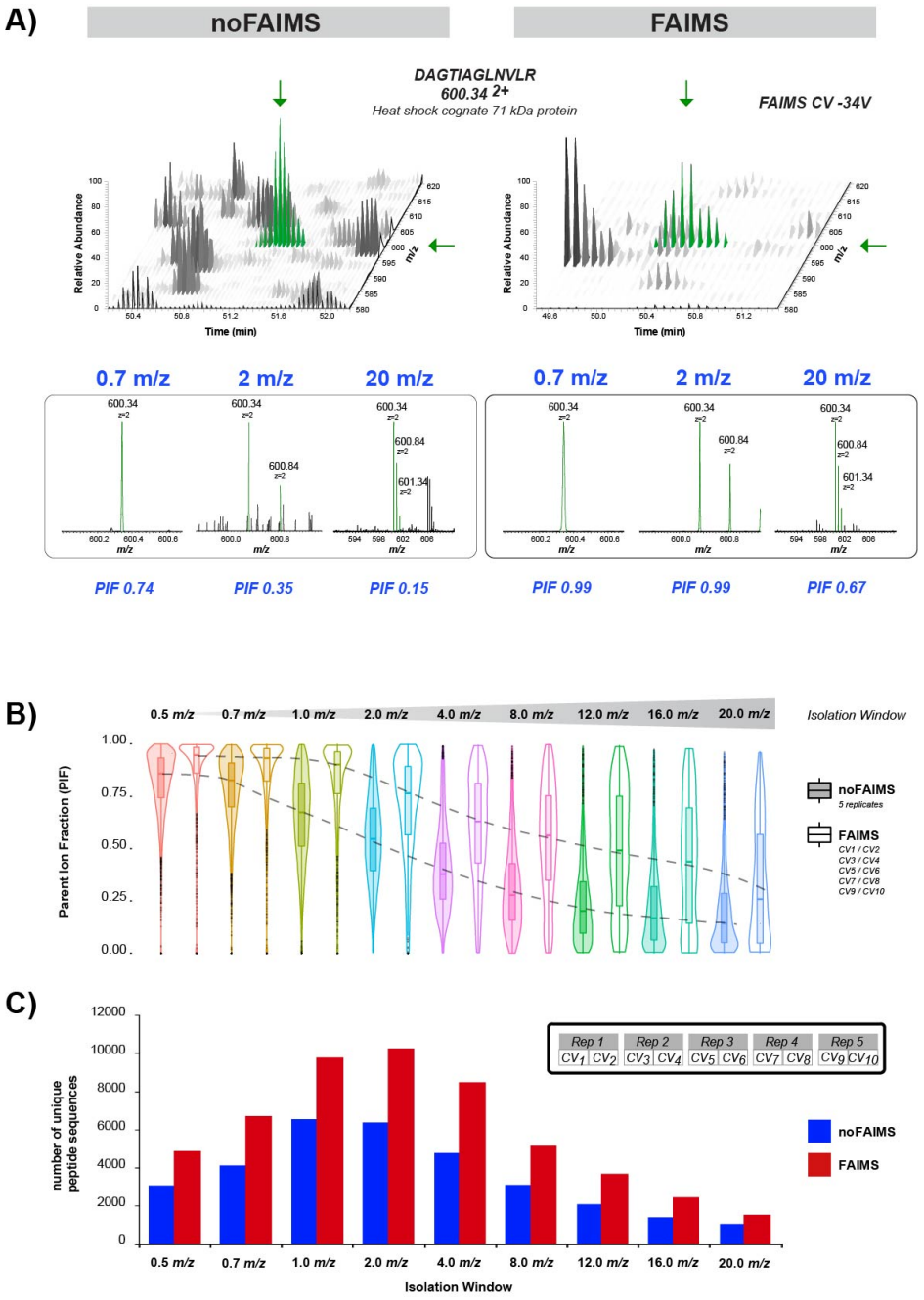


Figure 2-1: *Impact of precursor co-selection on peptide identification for the analysis of a HEK293 tryptic digest. (A) Comparison of peptide maps and mass spectra with FAIMS (right panel) and without FAIMS (left panel). Co-selection of isobaric precursors occurring with the increase of the isolation window leads to a decrease of the precursor intensity fraction (PIF). (B) Distribution of PIF values with and without FAIMS for peptide ions isolated with different isolation windows. The dotted lines represent the median values while the contour traces around the box plot represent the frequency distribution of PIF values. (C) Numbers of identified peptides obtained with and without FAIMS for each isolation window. 5 replicates were performed. For each replicate, the duty cycle was made of one MS followed by 5 MS/MS scans for 2 CV values. The CV values chosen were from -34V to -42V with a -2V increment.*

2.4.2. FAIMS improves accuracy and precision of quantitative measurements

To determine the extent of co-fragmentation and its impact on quantitative measurements using isobaric labeling, we created a series of samples based on a two-proteome model [23]. We first digested yeast proteins with trypsin, labeled separate aliquots with TMT reagents, and mixed those aliquots at ratio of 0:1:1.6:4:1.5:1 (***Figure 2-2A***). We also digested human HEK293 cell lysate with trypsin, labeled three aliquots, and mixed them at ratio of 1:1.5:1:0:0:0. We then combined yeast and human labeled sample at a 1:1 ratio. Yeast peptides were not labeled with TMT-126 while channels TMTC-129 to TMTC-131 were not used for human peptides. This labeling scheme facilitated the identification of co-fragmentation arising from peptides of each species. An example of the LC-MS/MS analysis of the combined sample is shown in ***Figure 2-2B*** where two precursors at *m/z* 695.4 and 695.9 were co-selected for fragmentation. Although the peptide at *m/z* 695.9 has been identified as EITALAPSTMK from human cDNA FLJ58286, a distorted TMT reporter ion ratio of 1:2.7:1.8:3.1:2.0:1.2 was observed for TMT 126-131 instead of 1:1.5:1:0:0:0. The fold change ratios observed here were generally overestimated by 80%. The presence of reporter ions in channel TMT-129-131 confirmed the co-fragmentation of a yeast precursor ion at *m/z* 695.4. In cases where the mixing ratio is unknown, it would be impossible to distinguish which fraction of the reporter ions originated from the targeted and contaminating ions. When the same LC-MS/MS analysis was performed using FAIMS, these peptides were transmitted at CV -41 V for human EITALAPSTMK and CV -35V for yeast VGGAVESEVGEK peptides (***Figure 2-2C***). The reporter ion ratio for EITALAPSTMK was 1:1.8:1.2:0.2:0:0.05, in close agreement we the expected ratio of 1:1.5:1:0:0:0 (20% variation). Similarly, the ion at *m/z* 695.4 was identified as peptide VGGASEVEVGEK from yeast HSP60 with a reporter ion ratio of 0.1:1.1:1.5:4.5:1.9:1.0 (expected ratio of 0:1:1.6:4:1.5:1) corresponding to an 11 % overestimation of ion ratio, on average.

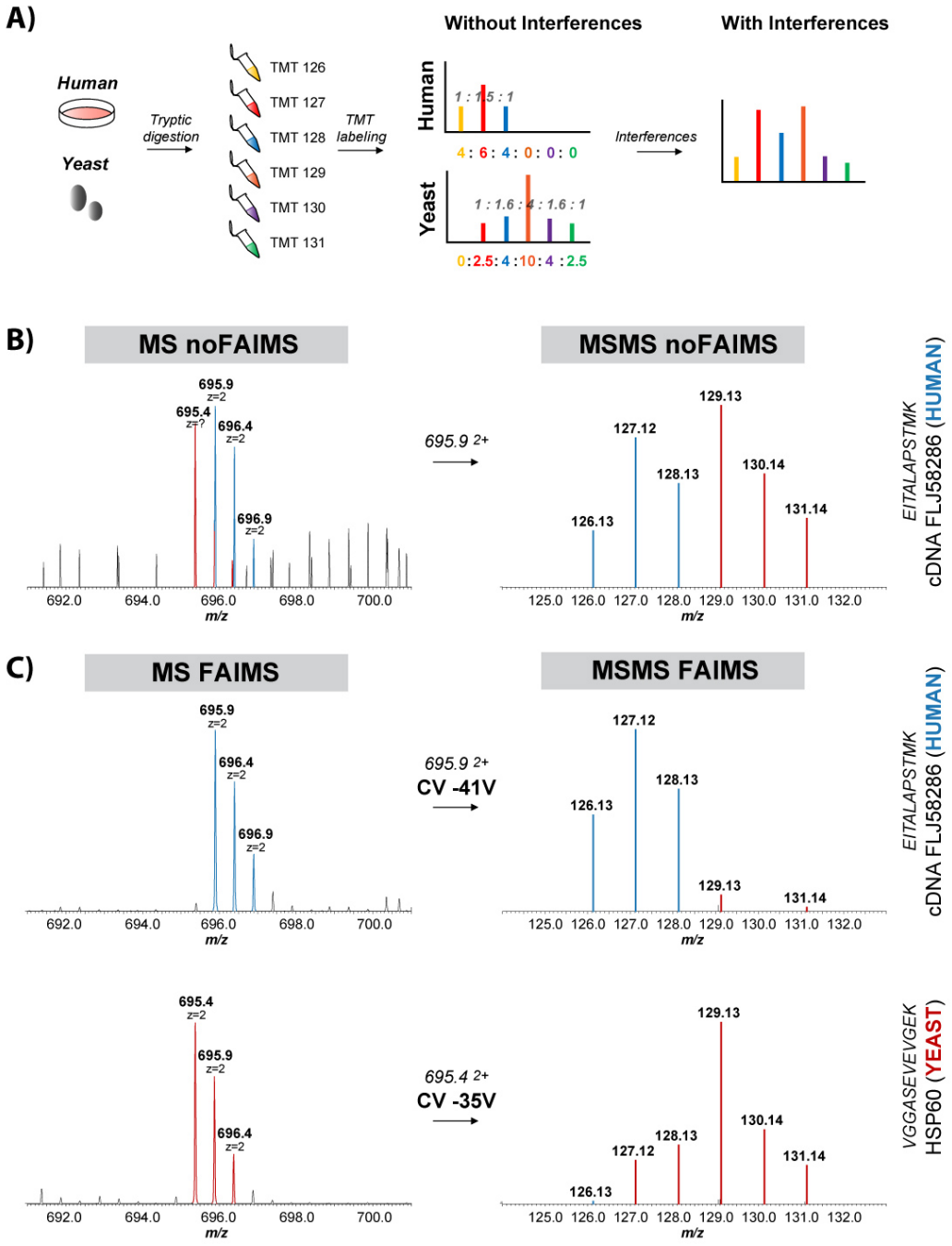


Figure 2-2: *Impact of peptide ion co-selection on TMT-quantification with a two-proteome model. (A) A human proteome and a yeast proteome are digested and labeled with TMT tags at ratios 1:1.5:1:0:0:0 (human) and 0:1:1.6:4:1.5:1 (yeast). (B) zoomed survey scan for mixed peptide ions at m/z 695.9^{2+} with the corresponding TMT-reporter fragment ions. (C) FAIMS separation of human peptide at m/z 695.9^{2+} and yeast peptide at m/z 695.4^{2+} with the corresponding reporter fragment ions.*

The list of identified and quantified peptides from yeast and human obtained from FAIMS and non-FAIMS LC-MS/MS experiments is presented in ***Supplementary Table 2-1***. A global analysis of all TMT ion ratios obtained for yeast and human peptides is shown in ***Figure 2-3*** for LC-MS/MS analyses performed with and without FAIMS. MS/MS spectra that did not show reporter ion intensities in channels *m/z* 126-128 for human peptides and in channels *m/z* 127 to 131 for yeast peptides were filtered out. These results indicated that a higher number of identifications for yeast and human peptides was obtained using FAIMS (1,871 human and 1,527 yeast peptides corresponding to 549 human and 291 yeast proteins) compared to non-FAIMS (1,327 human and 1,410 yeast peptides corresponding to 365 human and 245 yeast proteins) experiments. A decrease of background level and occurrence of co-fragmentation led to a higher identification of peptides, consistent with an earlier report [31]. Boxplots of reporter ion intensity ratios are presented in the upper panel of ***Figure 2-3A*** along with the expected ratios shown with dotted lines. Normalization was performed using TMT-126 for human peptides and TMT-131 for yeast peptides. Generally, we observed that quantitative measurements obtained using FAIMS were more precise with interquartile ranges of 0.1 to 0.9 compared to 0.2 to 1.2 for non-FAIMS experiments. The improved precision observed with FAIMS is also reflected by the narrower distribution of ion ratios in the box plots shown in ***Figure 2-3A***. Co-fragmentation observed using FAIMS did not significantly affect the ratio of human peptides even when the extent of contamination is significant (median of 0.8, 0.3 and 0.2 for TMT-129, -130 and -131, respectively). Interestingly, contamination from yeast is almost absent in human peptides when using FAIMS (median of 0.1, 0.05 and 0 for TMT-129, -130 and -131 respectively), and is possibly accounted for by the lower sample complexity of the yeast digest. Overall, FAIMS provided more accurate quantitative measurements with an average overestimation of 14 % compared to 48 % for non-FAIMS experiments. We next examined the fold change distribution of TMT reporter ratios for both FAIMS and non-FAIMS experiments (***Figure 2-3B***). We noted that the distribution of fold change for human peptides was almost superimposable in both experiments, although more quantifiable peptides and reduced distortion of TMT ratios were obtained using FAIMS (***Figure 2-3B***, upper panel). The effect of co-fragmentation on TMT reporter ion ratios

was more pronounced for yeast peptides without FAIMS with median ratios of 1.3:2.4:2.0:3.7:1.2:1 and 0.4:1.4:1.5:3.8:1.2:1 with FAIMS (***Figure 2-3A*** and ***Figure 2-3B*** lower panels). The distribution of fold change was centered along the expected ratios for the FAIMS experiments (dotted lines). We observed a more significant distortion of fold change ratios in the non-FAIMS experiments, especially for TMT-127 and TMT-128. This observation is possibly accounted for by the larger proportion of human peptide interferences in TMT-127 and TMT-128 channels. We noted up to 4-fold decrease in the proportion of contaminated MS/MS spectra with FAIMS compared to non-FAIMS experiments (***Supplementary Figure 2-3***). Altogether, these results suggest that FAIMS significantly reduced the extent of co-fragmentation, thereby providing up to 40% more quantifiable peptides with higher accuracy of ion ratio measurements.

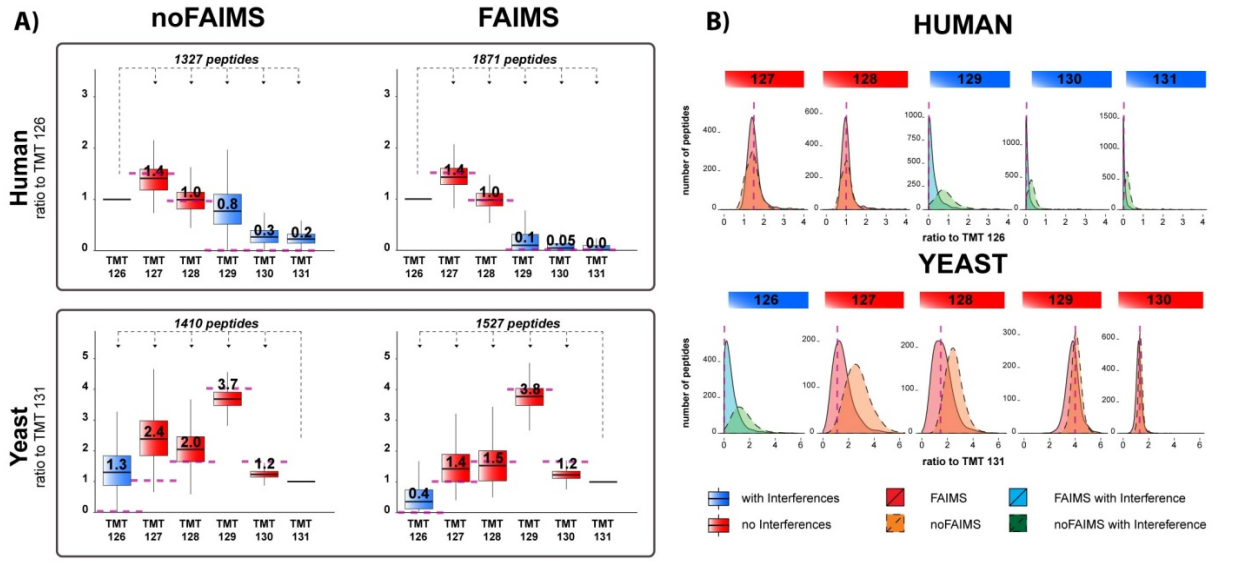


Figure 2-3: *Distortion of TMT ion ratios and extent of ion contamination for the two-proteome model. (A) Box plot and (B) frequency distribution of TMT reporter ion ratios normalized using TMT 126 and TMT 131 for for human and yeast peptide ions, respectively.*

An earlier contribution reported that peptide labeling with TMT reagents can introduce additional protonation sites thereby changing the charge state distribution compared to native tryptic peptides [55]. We observed the same phenomenon when we compared the charge distribution of a native and TMT labeled HEK293 tryptic digest

(***Supplementary Figure 2-4A***). Indeed, we noted a higher proportion of triply charged ions following TMT derivatization (46 %) compared to native tryptic peptides that showed a prominent distribution of doubly charged precursor ions (51%). Previous LC-MS experiments performed using traveling wave ion mobility separation (TWIMS) indicated that the transmission of multiply charged ions 2+, 3+ and 4+ overlapped more significantly, suggesting that labeled peptides tend to have similar mobilities in spite of their different charge states [56]. We observed a similar trend when we compared LC-FAIMS-MS analyses of a HEK293 tryptic digest with and without TMT labeling (***Supplementary Figure 2-4B-D***).

2.4.3. Monitoring temporal changes in protein abundance following heat shock

To evaluate the analytical merits of FAIMS and TMT labeling in quantitative proteomics, we profiled the changes in protein abundance of HEK293 cells exposed to heat shock. HEK293 cells were subjected to heat shock treatment for time periods up to 10 h, and samples were collected every 2 h. Cells were harvested, lysed in denaturing buffer, and proteins were extracted and digested with trypsin prior to TMT labeling. Tryptic peptides from different time points were labeled with separate TMT reagents, then combined together in equal ratio and analyzed by on-line 2D-LC-MS/MS with and without FAIMS (***Figure 2-4A***). Altogether, we identified 8,088 unique peptides from 1,582 proteins across all time points. FAIMS identified 68 % more peptides (26% more proteins) than non-FAIMS and more than 87 % of all identified peptides, and approximately 40% of these identifications were common to both approaches (***Figure 2-4B***). More than 73 % of all proteins were identified with two or more peptides when using FAIMS compared to 58% without FAIMS. Accordingly, the FAIMS data provided more comprehensive proteome coverage and was selected for further analyses.

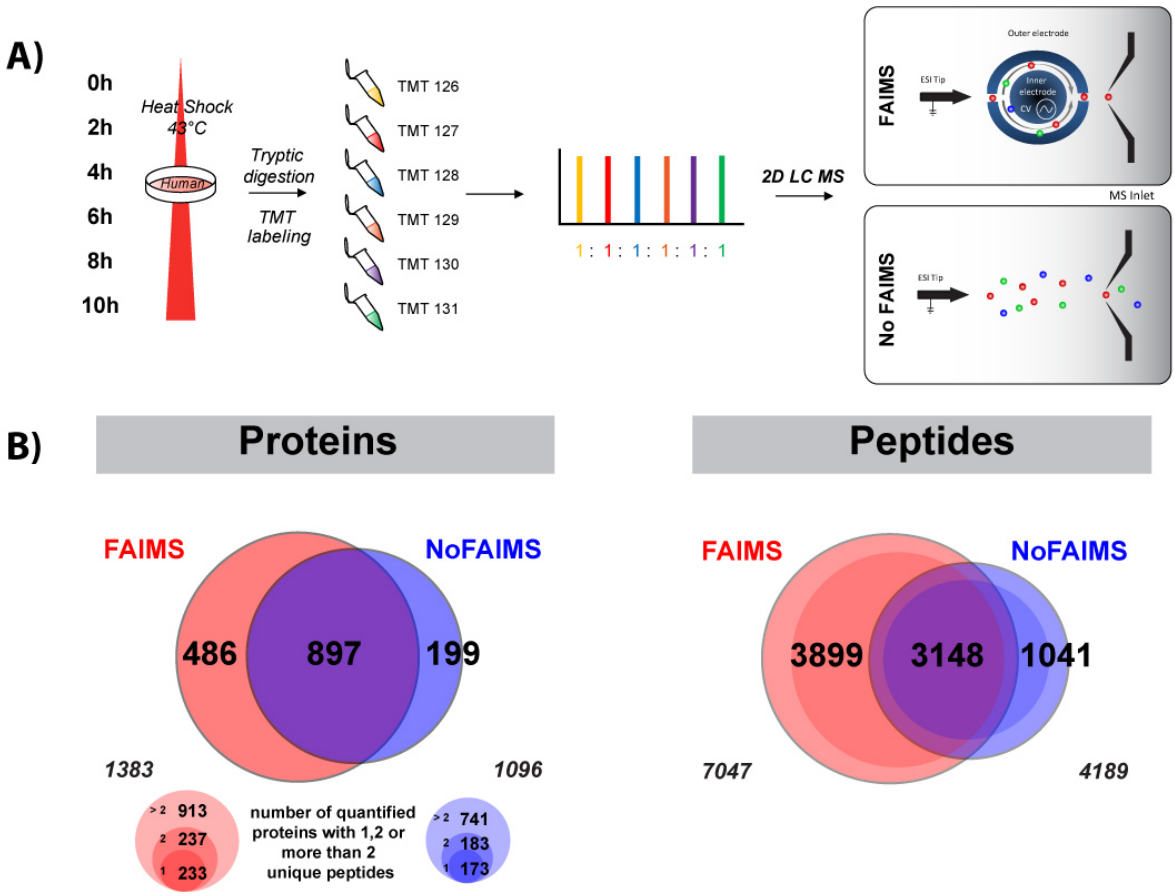


Figure 2-4: *Temporal profiling of protein abundance from HEK293 cells following heat shock treatment for up to 10 h. (A) Schematic representation of the experimental design. (B) Venn diagrams comparing the number of protein and peptide identification with and without FAIMS along with the distribution of proteins identified with 1, 2 or more than 2 unique peptides with and without FAIMS.*

Next, we determined the overall changes in protein abundance in response to heat shock by calculating the fold change (FC) ratio between conditions (***Figure 2-5A***). The global effect of the heat shock on the proteome was evaluated from the width of the $\log_2$ FC distribution of all proteins for a given time point, and the width of this distribution is represented as the interquartile range (IQR). A global change in protein abundance is reflected by a widening of the FC distribution with time (***Figure 2-5B***). A progressive increase in FC ratio of protein abundance was observed during the first 6h upon heat shock, and reached a plateau at 8 h suggesting a gradual adaptation of cells to their environment.

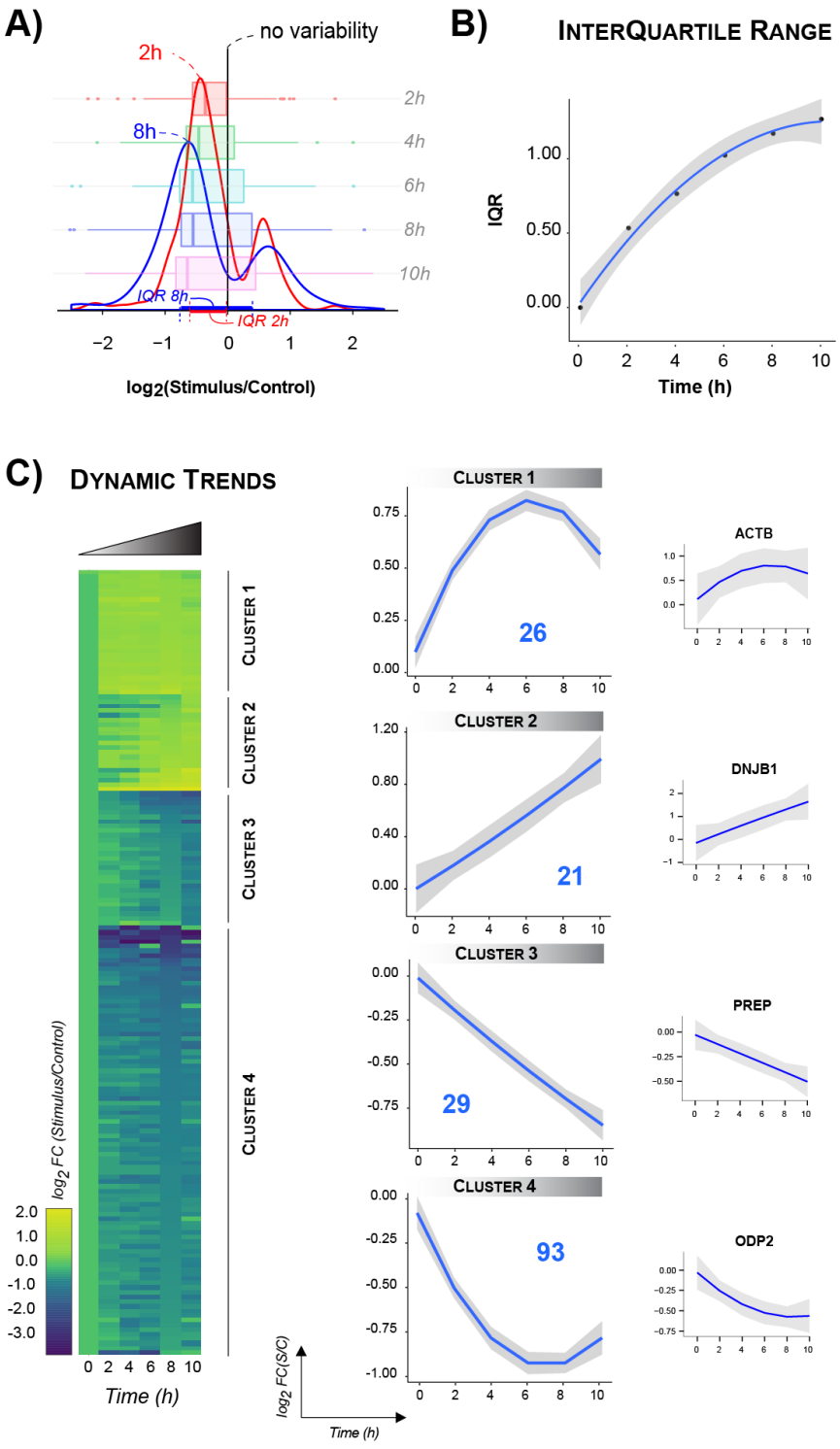


Figure 2-5: *Proteome wide effects of heat shock monitored by LC-FAIMS-MS/MS. (A) Distribution of Interquartile range (IQR) of Log_2 Fold change for different duration of heat shock treatment. (B) Variation of IQR with time. (C) Fuzzy c-means clustering of regulated proteins with 4 distinct profiles with representative examples for each cluster.*

In order to define relevant changes in protein abundance in our dataset, we first performed fitting of all kinetic profiles with a polynomial model, selecting only those with a coefficient of determination R^2 > 0.7 and a $|\log_2 FC|>0.6$ which enabled the identification of kinetic profiles of low variance. We identified 169 proteins that showed distinct changes in protein abundance (increase or decrease) with time. These proteins were regrouped into 4 clusters with unimodal distribution showing monotonous (clusters 2 and 3) and adaptation-like (clusters 1 and 4) profiles (***Figure 2-5C***). Representative examples are also shown for each cluster. For convenience, we hereafter defined these proteins as “dynamic” to distinguish them from the remaining “static” proteins. Dynamic proteins showed significant changes in abundance where maximum $|\log_2 FC|$ measurements were up to 10-fold higher than static proteins (***Supplementary Figure 2-5***). The list of quantified proteins and their corresponding changes in protein abundance is reported in ***Supplementary Table 2-2***.

Heat shock was previously reported to induce different cellular response including protein denaturation and aggregation, formation of mRNA storing granules [57] [58] and transcription of heat-shock response genes [59]. To define functional groups associated with dynamic proteins of each cluster, we performed gene ontology (GO) enrichment analyses against the whole human proteome using DAVID bioinformatic resources (***Figure 2-6A, Supplementary Table 2-3***). We identified several protein groups associated with the classical heat shock response involving physiological and metabolic adaptation to stress, and protein aggregation and sequestration**.** For instance**,** Cluster 2 regroups 21 proteins showing a regular increase in abundance, and includes 8 chaperones and heat shock proteins (HSPs) such as HSP70, HSP90 and DNAJB1 (HSP40). Their over expression upon heat shock was reported previously [60], and is necessary to prevent misfolded proteins and maintain proteostasis [61]. This cluster also comprises other chaperone-associated proteins including BAG family molecular chaperone regulator 3 (BAG3), a protein that interacts with the proteasome and modulate its activity [62].

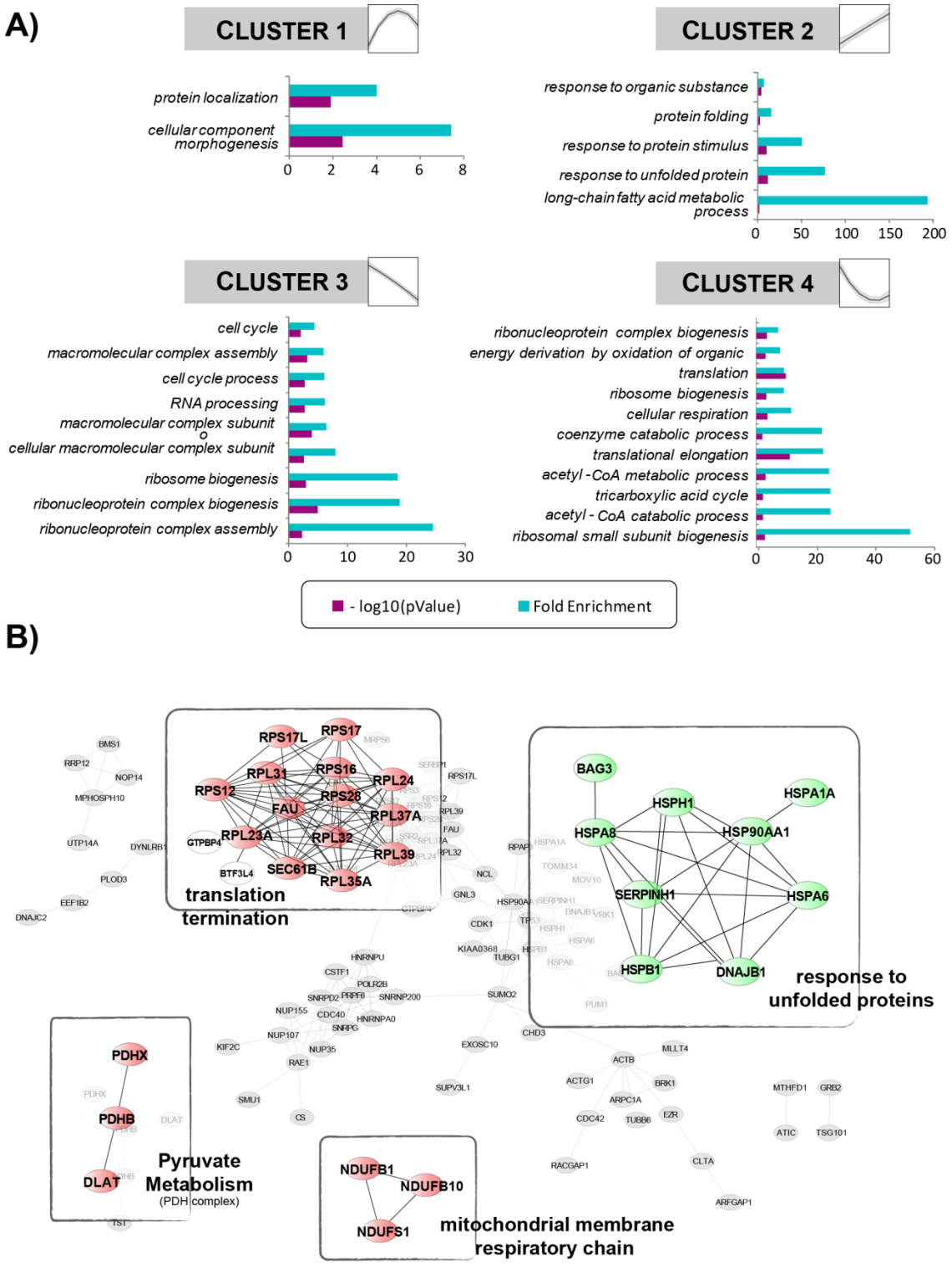


Figure 2-6: *Bioinformatic analyses of proteins dynamically regulated upon hyperthermia. (A) GO enrichment analysis for individual cluster. (B) Interaction network using the STRING database selecting only proteins with high stringency (score > 0.7) interactions inferred from experimental data.*

While newly synthesized proteins are particularly sensitive to heat-induced misfolding, mature proteins can also aggregate in response to heat shock by forming protein/poly(A)-RNA structures called stress granules. These subcellular structures contain a diverse group of proteins that include stalled initiation complexes, mRNA transcripts, small ribosomal subunits, and different eukaryotic initiation factors [63]. We identified several putative components of stress granules in cluster 1 (26 proteins) including several eukaryotic initiation factors and small ribosomal subunits that all showed a progressive increase in abundance and peaked at 4-6 h following heat shock. The abundance profiles observed for these proteins are consistent with the short lived stress granules and their roles in reprogramming mRNA translation during cellular stress [64].

The remaining two clusters regroup proteins that are downregulated upon heat shock. For example, cluster 3 contains several ribosomal proteins involved in the elongation and termination translation processes along with proteins associated with ribosome biogenesis. The decrease in ribosomal proteins is consistent with a reduced protein synthesis and the accumulation of aggregated translation factors taking place during heat shock [58]. We also noted the decreased abundance of cyclin-dependent kinase 1 (CDK1), serine/threonine-protein kinase VRK1, and several metabolic kinases (e.g. GUK1, PFKM). Inactivation of CDK1 was previously reported to restrain cell cycle progression during the early yeast response to heat shock [6], and decreased levels of CDK1 observed under prolonged thermal stress might be necessary to maintain cell arrest.

Cluster 4 regroups the largest number of proteins, several of which are associated with metabolic and catabolic processes. We identified members of the pyruvate dehydrogenase complex (e.g. PDHX, PDHB, DLAT) together with other mitochondrial enzymes involved in the tricarboxylic acid (TCA) cycle (e.g. MDH2, CS, SUCLG2) that all showed a decreased in abundance in the first 6 h upon treatment. Heat stress was previously shown to trigger the overproduction of reactive oxygen species that cause mitochondrial protein denaturation of several protein complexes including the pyruvate dehydrogenase (PDH) complex, subunits of the ATP synthase, and enzymes of the TCA cycle [65]. Unfolded proteins typically form insoluble

aggregates that can accumulate beyond the capacity of the chaperone repair systems, a situation that ultimately leads to impaired mitochondria and selective degradation through mitophagy [66]. Accordingly, irreversibly oxidized and/or denatured proteins arising through heat shock could be degraded by mitochondrial proteases or mitophagy.

A subset of the proteins identified in the present study are regrouped into a network of interacting partners representing 92 of the 169 dynamic proteins (***Figure 2-6B***). This network identifies several interacting proteins that were upregulated upon hyperthermia such as heat shock and chaperone proteins. Downregulated proteins comprise members of the acetyl-CoA biosynthetic process (e.g. PDH and TCA complexes) and mitochondrial membrane respiratory chain (NDUFB10, NDUFS1, NDUFB1) together with several small ribosomal subunits.

Our study provided several lines of evidence indicating that cells exposed to hyperthermia accumulate protein aggregates, decelerate protein synthesis and catabolic activities, and mount a transcriptional program resulting in the upregulation of HSP proteins (***Figure 2-7***). The latter proteins are the classic targets of the heat shock response, and act as molecular chaperones to facilitate folding, transport and degradation of proteins. While HSP proteins showed a constant increased in abundance over the first 8 h upon heat shock, our quantitative proteomics analysis also suggested the formation of stress granules and the accumulation of several eukaryotic initiation factors, small ribosomal subunits and Poly(A) binding proteins that regulates mRNA stability and protein translation. Heat stress also affected the cytoskeletal organization of the cell beyond the unfolding of individual proteins. For example, we noted the decreased abundance of vimentin, tubulin and actin related proteins that are part of the ATP-binding component of the Arp2/3 complex involved in regulation of actin polymerization. These results are consistent with morphological changes leading to the accumulation of actin aggregates [67], looser microtubule meshwork [68], and collapse of intermediate filaments [68, 69] previously observed in other cell lines. Finally, our quantitative proteomics analyses also revealed that heat shock affected mitochondrial activities as reflected from the downregulation of protein complexes involved in the biosynthesis of acetyl-CoA and subunits of the membrane respiratory chain NADH dehydrogenase.

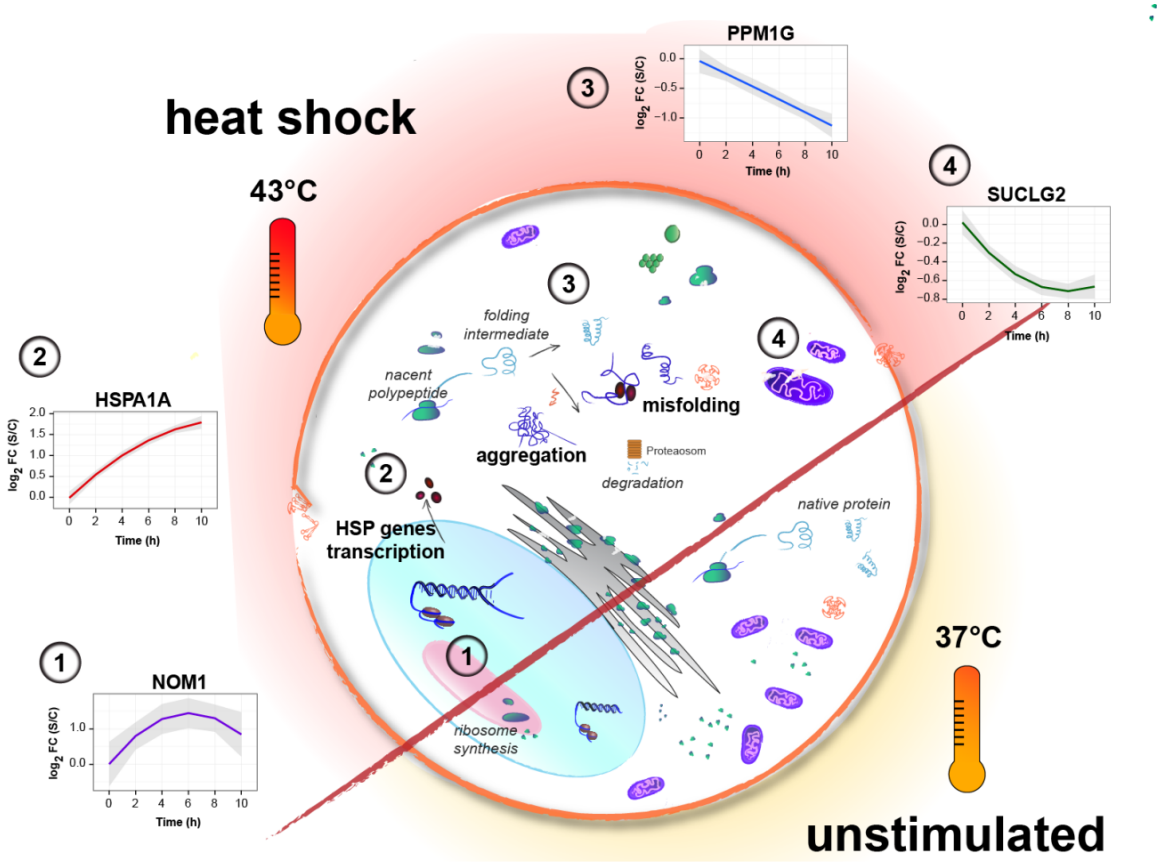


Figure 2-7: *Summary of protein expression changes observed during heat stress.*

2.5. Conclusion

The high sample complexity of protein digests has profound impacts on the number of detectable and quantifiable peptides when using quantitative proteomic strategies based on isotopic chemical labeling. While chromatographic fractionation can reduce sample complexity, the dynamic range of quantification in LC-MS/MS experiments is limited by the extent of background ions that mitigate the detection of low abundance peptide ions. In this context, FAIMS can provide more comprehensive quantitative analyses by separating peptide ion populations based on their mobilities at low and high electric fields to reduce the extent of interfering ions and further improve the detection and quantification of peptide ions.

The increased sensitivity of FAIMS provides several advantages in quantitative proteomic studies. First, the reduction of interferences in LC-MS/MS experiments facilitates the detection and identification of low abundance peptide ions while reducing the occurrence of chimeric MS/MS spectra by 40-60%. The gain in peptide identification also increased the number of quantifiable peptides by up to 68% compared to non-FAIMS experiments. Second, the decrease in confounding ions provided more accurate quantitative measurements using isobaric chemical labeling. LC-MS/MS analyses of yeast and human protein digests mixed in predetermined proportions indicated that FAIMS significantly reduced the ratio compression observed with TMT labeling, and fold change measurements were on average within 14 % of the calculated values compared to 48 % for non-FAIMS experiments. Third, the reduced contribution of interfering ions observed using FAIMS led to more precise quantitative measurements with inter quartile range values between 0.1 to 1.0 compared to 0.2 and 1.2 for conventional LC-MS/MS analyses.

The improved performances of FAIMS in quantitative proteomics were also evaluated for the temporal profiling of protein abundance in HEK293 cells during the first 10 h following heat shock. Consistent with previous LC-MS/MS analyses, we observed that FAIMS enabled the identification of 68% more peptides compared to non-FAIMS experiments. The enhanced proteome coverage obtainable using FAIMS enabled the profiling of 1,383 proteins, of which 169 were dynamically regulated upon

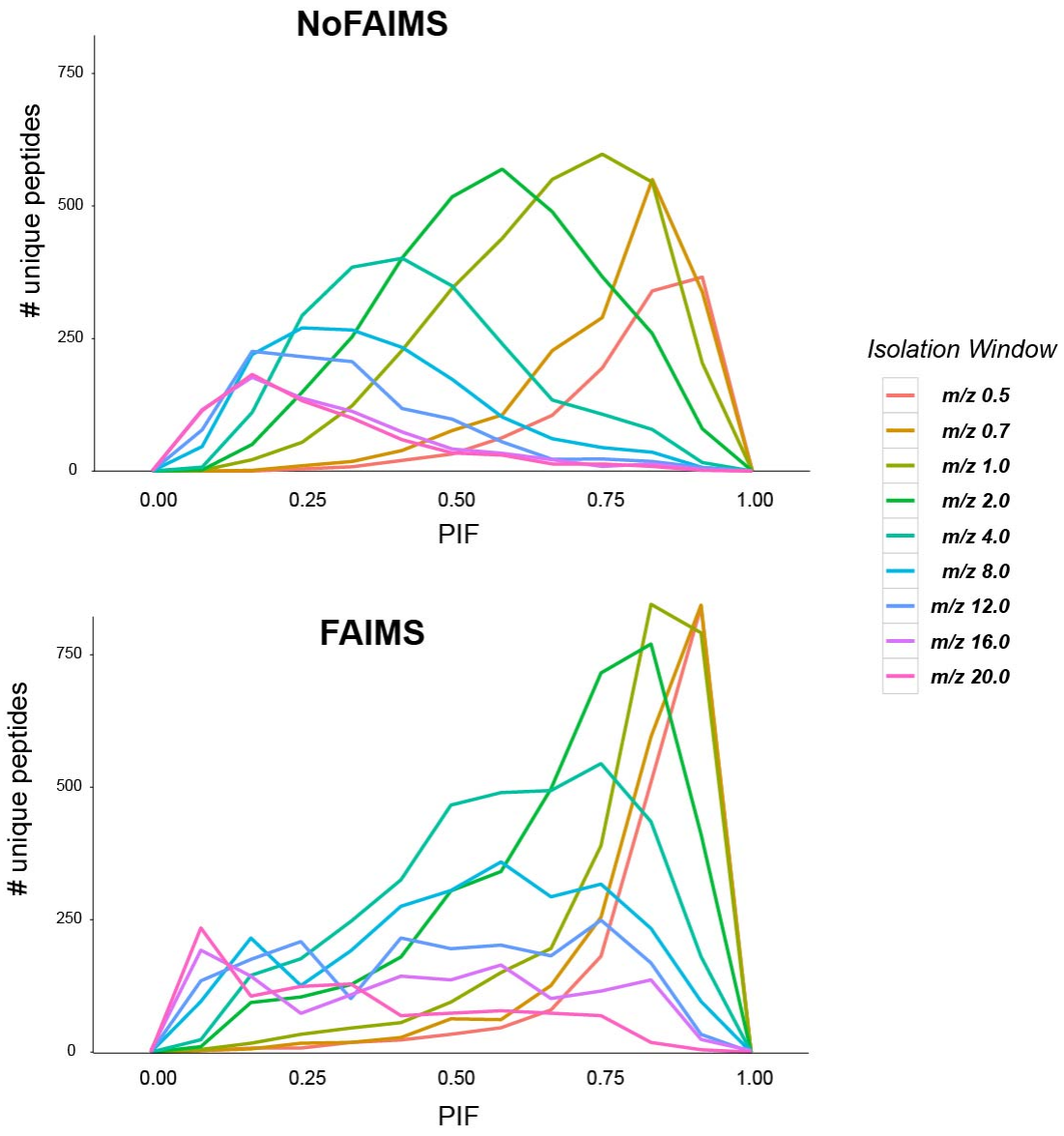
hyperthermia. These analyses confirmed the upregulation of several chaperone and heat shock proteins (e.g. HSP70, HSP90 and DNAJB1) and the unsuspected downregulation of mitochondrial enzymes involved in metabolic and catabolic processed (PDH and TCA complexes). We anticipate that the advantages of FAIMS in terms of increased proteome coverage and improved precision and accuracy of quantitative measurements using isobaric chemical labeling will also have similar analytical benefits for other forms of quantitative proteomics.

2.6. Acknowledgments

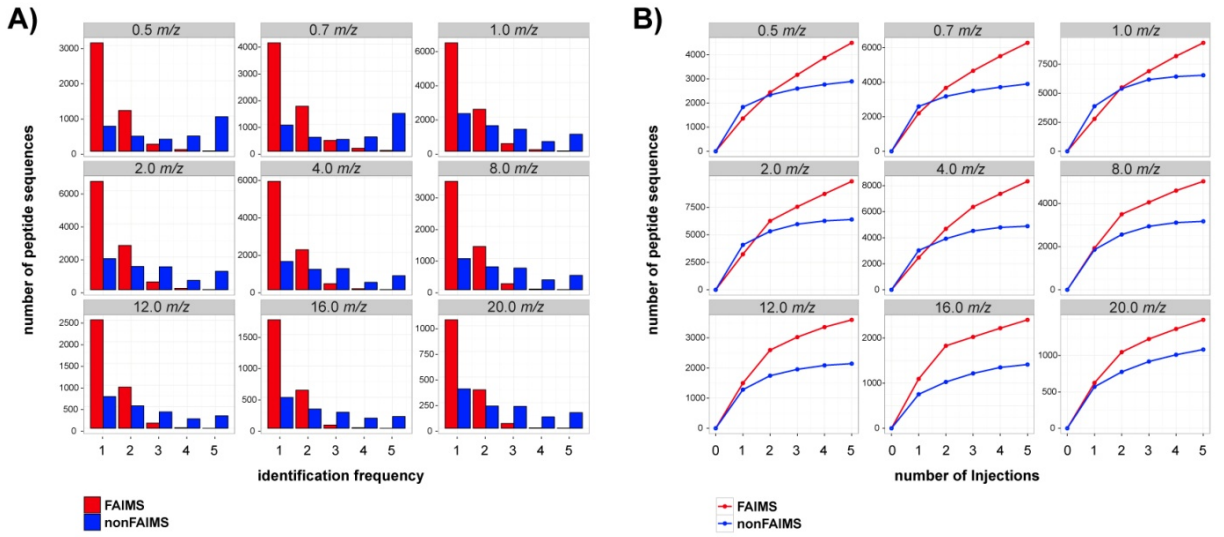
This work was carried out with financial support from the Natural Sciences and Engineering Research Council (NSERC 311598). The authors thank Jean-Jacques Dunyach and Michael Belford (Thermo Fisher) for valuable help and assistance with the FAIMS interface. The Institute for Research in Immunology and Cancer (IRIC) receives IRICoR, the Canadian Foundation for Innovation, and the Fonds de Recherche du Québec - Santé (FRQS). IRIC proteomics facility is a Node of the Canadian Genomic Innovation Network and is supported by the Canadian Government through Genome Canada.

2.7. Supplementary material

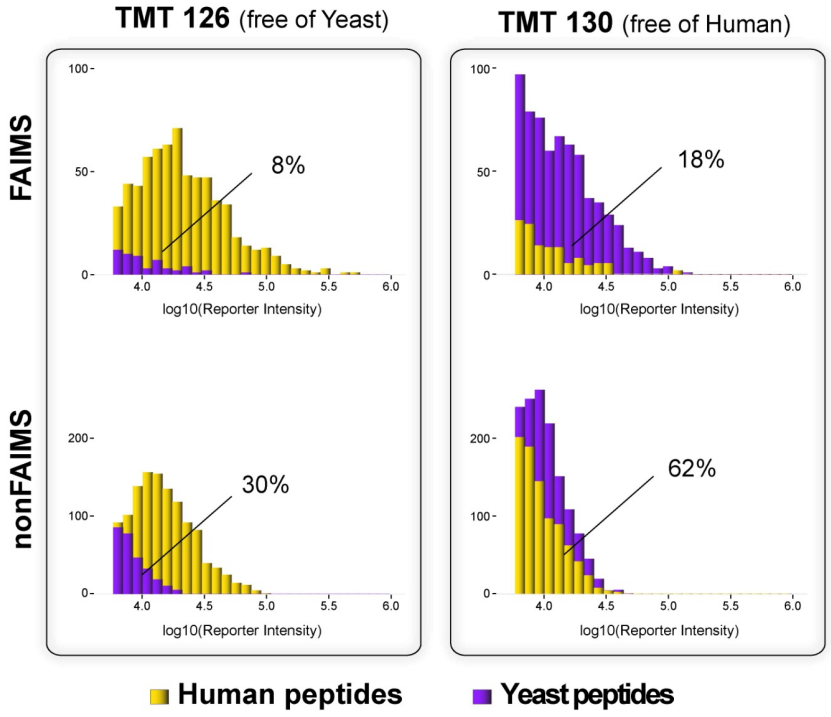
2.7.1. Supplementary figures



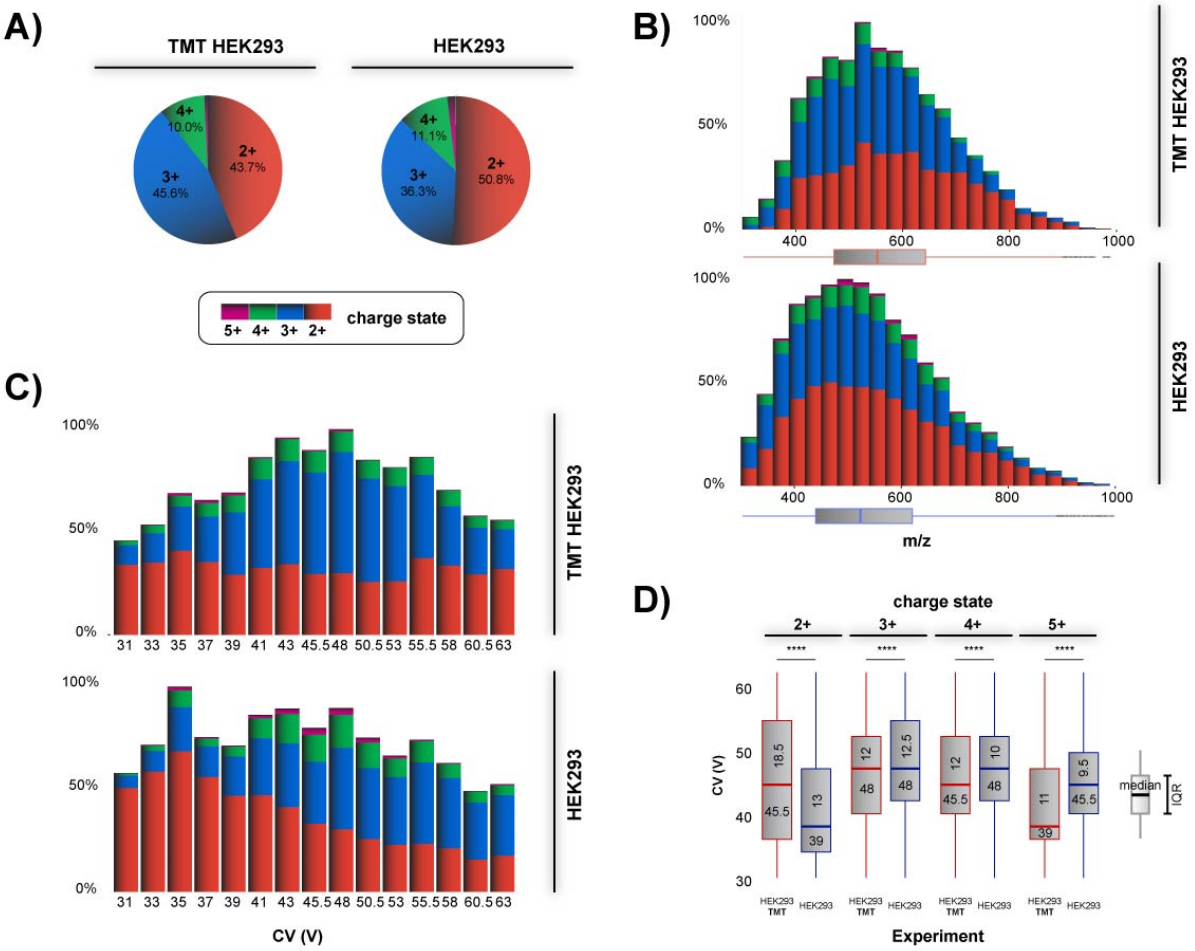
Supplementary Figure 2-1:*Distribution unique peptides as a function of parent ion fraction (PIF).*



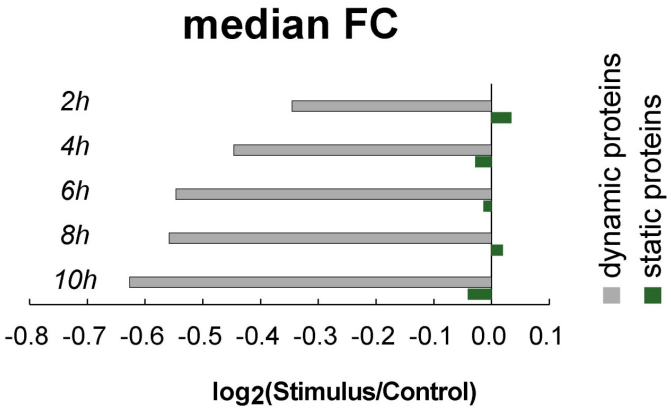
Supplementary Figure 2-2: *Comparison of peptide and protein identification for HEK 293 tryptic digest using LC–MS/MS with and without FAIMS. (A) Peptide identification frequency representing the cumulative number of identifications per replicate bin for FAIMS and non FAIMS experiments across 5 replicates for all the tested isolation window widths. (B) Increase of unique peptide identification with replicate LC–MS/MS analyses with and without FAIMS.*



Supplementary Figure 2-3: *Extent of interference with precursor ion intensity with and without FAIMS for the two-proteome (yeast and human) model.*



Supplementary Figure 2-4: *Influence of TMT derivatization on LC-FAIMS-MS analysis of HEK293 tryptic digest. (A) Charge state distribution for HEK 293 tryptic peptides analyzed by LC-FAIMS-MS/MS. (B) m/z distribution for HEK 293 tryptic peptides analyzed by LC-FAIMS-MS. (C) Number and distribution of HEK 293 tryptic peptides and charge state across replicate LC–FAIMS -MS/MS analyses. (D) Boxplot distribution of transmission CV for all charge state with p-value* $\leqslant$ *0.0001 (****).*



Supplementary Figure 2-5: *Fold change median for dynamic (grey) and static (green) proteins after heat shock.*

2.7.2. Supplementary tables

Supplementary Table 2-1*: List of identified and quantified peptides from yeast and human obtained from FAIMS and non-FAIMS LC–MS/MS experiments (CD-ROM).*

Supplementary Table 2-2: *List of quantified proteins and corresponding changes in abundance upon heat shock (CD-ROM).*

Supplementary Table 2-3: *Bioinformatics and GO terms analyses of clusters showing dynamic changes in protein abundance upon heat shock (CD-ROM).*

2.8. References

1. Larance, M. and A.I. Lamond, *Multidimensional proteomics for cell biology.* Nat Rev Mol Cell Biol, 2015. **16**(5): p. 269-80.
2. Hebert, A.S., A.L. Richards, D.J. Bailey, A. Ulbrich, E.E. Coughlin, M.S. Westphall, and J.J. Coon, *The one hour yeast proteome.* Mol Cell Proteomics, 2014. **13**(1): p. 339-47.
3. Bantscheff, M., M. Schirle, G. Sweetman, J. Rick, and B. Kuster, *Quantitative mass spectrometry in proteomics: a critical review.* Anal Bioanal Chem, 2007. **389**(4): p. 1017-31.
4. Ong, S.E. and M. Mann, *Mass spectrometry-based proteomics turns quantitative.* Nat Chem Biol, 2005. **1**(5): p. 252-62.
5. Kanshin, E., L.P. Bergeron-Sandoval, S.S. Isik, P. Thibault, and S.W. Michnick, *A cell-signaling network temporally resolves specific versus promiscuous phosphorylation.* Cell Rep, 2015. **10**(7): p. 1202-1214.
6. Kanshin, E., P. Kubiniok, Y. Thattikota, D. D'Amours, and P. Thibault, *Phosphoproteome dynamics of Saccharomyces cerevisiae under heat shock and cold stress.* Mol. Syst. Biol., 2015. **11**(6): p. 813.
7. Olsen, J.V., B. Blagoev, F. Gnad, B. Macek, C. Kumar, P. Mortensen, and M. Mann, *Global, in vivo, and site-specific phosphorylation dynamics in signaling networks.* Cell, 2006. **127**(3): p. 635-48.
8. Bantscheff, M., S. Lemeer, M.M. Savitski, and B. Kuster, *Quantitative mass spectrometry in proteomics: critical review update from 2007 to present.* Anal. Bioanal. Chem., 2012. **404**: p. 939-965.
9. Cox, J., M.Y. Hein, C.A. Luber, I. Paron, N. Nagaraj, and M. Mann, *Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ.* Mol. Cell Proteomics, 2014. **13**(9): p. 2513-2526.
10. Hilger, M. and M. Mann, *Triple SILAC to determine stimulus specific interactions in the Wnt pathway.* J. Proteome Res., 2012. **11**(2): p. 982-994.
11. Ong, S.E., B. Blagoev, I. Kratchmarova, D.B. Kristensen, H. Steen, A. Pandey, and M. Mann, *Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression in proteomics.* Mol. Cell Proteomics, 2002. **1**(5): p. 376-386.
12. Schwanhäusser, B., M. Gossen, G. Dittmar, and M. Selbach, *Global analysis of cellular protein translation by pulsed SILAC.* Proteomics, 2009. **9**(1): p. 205-209.
13. Hebert, A.S., A.E. Merrill, D.J. Bailey, A.J. Still, M.S. Westphall, E.R. Strieter, D.J. Paglianni, and J.J. Coon, *Neutron-encoded mass signatures for multiplexed quantification.* Nat. Methods, 2013. **10**(4): p. 332-334.
14. Boersema, P.J., R. Raijmakers, S. Lemeer, S. Mohammed, and A.J. Heck, *Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics.* Nat. Methods, 2009. **4**(4): p. 484-494.
15. Thomson, A., J. Schafer, K. Kuhn, S. Kienle, J. Schwarz, G. Schmidt, T. Neumann, and C. Hamon, *Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MSMS.* Anal. Chem., 2003. **75**: p. 1895-1904.

16. Ross, P.L., Y.N. Huang, J.N. Marchese, B. Williamson, K. Parker, S. Hattan, N. Khainovski, S. Pillai, S. Dey, S. Daniels, et al., *Multiplexed protein quantitation in S. Cerevisiae using amine-reactive isobaric tagging reagents.* Mol. Cell Proteomics, 2004. **3**(12): p. 1154-1169.

17. Hein, M.Y., N.C. hubner, I. Poser, J. Cox, N. Nagaraj, Y. Toyoda, I.A. gak, I. Weisswange, J. Mansfeld, F. Buchholz, et al., *A human interactome in three quantitative dimensions organized by stoichiometries and abundances.* Cell, 2015. **163**: p. 712-723.

18. Li, Z., R.M. Adams, K. Chourey, G.B. Hurst, R.L. Hettich, and C. Pan, *Systematic comparison of label-free, metabolic labeling, and isobaric chemical labeling for quantitative proteomics on LTQ Orbitrap Velos.* J Proteome Res, 2012. **11**(3): p. 1582-90.

19. Bantscheff, M., M. Boesche, D. Eberhard, T. Matthieson, G. Sweetman, and B. Kuster, *Robust and sensitive iTRAQ quantification on an LTQ Orbitrap mass spectrometer.* Mol Cell Proteomics, 2008. **7**(9): p. 1702-13.

20. Ow, S.Y., M. Salim, J. Noirel, C. Evans, I. Rehman, and P.C. Wright, *iTRAQ underestimation in simple and complex mixtures: "the good, the bad and the ugly".* J Proteome Res, 2009. **8**(11): p. 5347-55.

21. Shirran, S.L. and C.H. Botting, *A comparison of the accuracy of iTRAQ quantification by nLC-ESI MSMS and nLC-MALDI MSMS methods.* J Proteomics, 2010. **73**(7): p. 1391-403.

22. Wenger C., Lee M.V., Hebert A.S., McAlister G.C., P. D.H., W. M.S., and C. J.J., *Gas-Phase purification enables accurate multiplexed proteome quantification with isobaric tagging.* Nat. Methods, 2011. **8**(11): p. 933-935.

23. Ting, L., R. Rad, S.P. Gygi, and W. Haas, *MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics.* Nat. Methods, 2011. **8**(11): p. 938-940.

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

25. Wuhr, M., W. Haas, M. G.C., L. Peshkin, R. Rad, M.W. Kirschner, and S.P. Gygi, *Accurate multiplexed proteomics at the MS2 level using the complement reporter ion cluster.* Anal. Chem., 2012. **84**(21): p. 9214-9221.

26. Guevremont, R., *High-field asymmetric waveform ion mobility spectrometry: a new tool for mass spectrometry.* J. Chromatogr. A, 2004. **1058**(1-2): p. 3-19.

27. Purves, R.W. and R. Guevremont, *Electrospray ionization high-field asymmetric waveform ion mobility spectrometry-mass spectrometry.* Anal. Chem., 1999. **71**(13): p. 2346-2357.

28. Barnett, D.A., B. Ells, R. Guevremont, and R.W. Purves, *Application of ESI-FAIMS-MS to the analysis of tryptic peptides.* J. Am. Soc. Mass Spectrom., 2002. **13**: p. 1282-1291.

29. Guevremont, R. and R.W. Purves, *Atmospheric pressure ion focusing in a high-field asymmetric waveform ion mobility spectrometer.* Rev. Sci. Instrum., 1999. **70**: p. 1370-1383.

30. Shvartsburg, A.A., K. Tang, and R.D. Smith, *Modeling the resolution and sensitivity of FAIMS analyses.* J Am Soc Mass Spectrom, 2004. **15**(10): p. 1487-98.

31. Bonneil, E., S. Pfammatter, and P. Thibault, *Enhancement of mass spectrometry performances for proteomics analyses using High-Field asymmetric waveform spectrometry (FAIMS).* J. Mass Spectrom., 2015. **50**(11): p. 1181-95.

32. Canterbury, J.D., X. Yi, M.R. Hoopmann, and M.J. McCoss, *Assessing the dynamic range and peak capacity of nanoflow LC-FAIMS-MS on an ion trap mass spectrometer for proteomics.* Anal. Chem., 2008. **80**(18): p. 6888-6897.

33. Creese, A.J., N.J. Shimwell, K.P.B. Larkins, J.K. Heath, and H.J. Cooper, *Probing the complementarity of FAIMS and strong cation exchange chromatography in shotgun proteomics.* J. Am. Soc. Mass Spectrom., 2013. **24**: p. 431-443.

34. Saba, J., E. Bonneil, C. Pomiès, K. Eng, and P. Thibault, *Enhanced sensitivity in proteomics experiments using FAIMS coupled with a hybrid linear ion trap/orbitrap mass analyzer.* J. Proteome Res., 2009. **8**: p. 3355-3366.

35. Swearingen, K.E., M.R. Hoopmann, R.S. Johnson, R.A. Saleem, J.D. Aitchinson, and R.L. Moritz, *Nanospray FAIMS fractionation provides significant increases in proteome coverage of unfractionated complex protein digests.* Mol. Cellular Proteomics, 2012. **11**(4): p. M111.014985.

36. Venne, K., E. Bonneil, K. Eng, and P. Thibault, *Improvement in peptide detection for proteomics analyses using nanoLC-MS and high-field asymmetry waveform ion mobility mass spectrometry.* Anal. Chem., 2005. **77**(7): p. 2176-2186.

37. Xuan, Y., A.J. Creese, J.A. Horner, and H.J. Cooper, *High-field asymmetric waveform ion mobility spectrometry (FAIMS) coupled with high-resolution electron transfer dissociation mass spectrometry for the analysis of isobaric phosphopeptides.* Rapid Commun. Mass Spectrom., 2009. **23**: p. 1963-1969.

38. Courcelles, M., G. Bridon, S. Lemieux, and P. Thibault, *Occurrence and detection of phosphopeptide isomers in large-scale phosphoproteomics experiments.* J. Proteome Res., 2012. **11**: p. 3753-3765.

39. Shvartsburg, A.A., A.J. Creese, R.D. Smith, and H.J. Cooper, *Separation of peptide isomers with variant modified sites by high-resolution differential ion mobility spectrometry.* Anal. Chem., 2010. **82**(19): p. 8327-8334.

40. Shvartsburg, A.A., D. Singer, R.D. Smith, and R. Hoffmann, *Ion mobility separation of isomeric phosphopeptides from a protein with variant modification of adjacent residues.* Anal. Chem., 2011. **83**: p. 5078-5085.

41. Shvartsburg, A.A., Y. Zheng, R.D. Smith, and N.L. Kelleher, *Separation of variant methylated histone tails by differential ion mobility.* Anal. Chem., 2012. **84**(15): p. 6317-6320.

42. Klaassen, T., S. Szwandt, J.T. Kapron, and A. Roemer, *Validated quantitation method for a peptide in rat serum using liquid chromatography/high-field asymmetric waveform ion mobility spectrometry.* Rapid Commun Mass Spectrom, 2009. **23**(15): p. 2301-6.

43. Xia, Y.Q., S.T. Wu, and M. Jemal, *LC-FAIMS-MS/MS for quantification of a peptide in plasma and evaluation of FAIMS global selectivity from plasma components.* Anal Chem, 2008. **80**(18): p. 7137-43.

44. Bridon, G., E. Bonneil, T. Muratore-Schroeder, O. Caron-Lizotte, and P. Thibault, *Improvement of phosphoproteome analyses using FAIMS and decision tree fragmentation. application to the insulin signaling pathway in Drosophila melanogaster S2 cells.* J Proteome Res, 2012. **11**(2): p. 927-40.

45. Prasad, S., M.W. Belford, J.J. Dunyach, and R.W. Purves, *On an aerodynamic mechanism to enhance ion transmission and sensitivity of FAIMS for nano-electrospray ionization-mass spectrometry.* J Am Soc Mass Spectrom, 2014. **25**(12): p. 2143-53.

46. Elias, J.E. and S.P. Gygi, *Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry.* Nat Methods, 2007. **4**(3): p. 207-14.

47. Kumar, L. and E.F. M, *Mfuzz: a software package for soft clustering of microarray data.* Bioinformation, 2007. **2**(1): p. 5-7.

48. Huang, D.W., B.T. Sherman, Q. Tan, J.R. Collins, W.G. Alvord, J. Roayaei, R. Stephens, M.W. Baseler, H.C. Lane, and R.A. Lempicki, *The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists.* Genome Biol, 2007. **8**(9): p. R183.

49. Chapman, J.D., D.R. Goodlett, and C.D. Masselon, *Multiplexed and data-independent tandem mass spectrometry for global proteome profiling.* Mass Spectrom Rev, 2014. **33**(6): p. 452-70.

50. Domon, B. and R. Aebersold, *Options and considerations when selecting a quantitative proteomics strategy.* Nat Biotechnol, 2010. **28**(7): p. 710-21.

51. Gallien, S., A. Bourmaud, S.Y. Kim, and B. Domon, *Technical considerations for large-scale parallel reaction monitoring analysis.* J Proteomics, 2014. **100**: p. 147-59.

52. Luethy, R., D.E. Kessner, J.E. Katz, B. MacLean, R. Grothe, K. Kani, V. Faca, S. Pitteri, S. Hanash, D.B. Agus, et al., *Precursor-ion mass re-estimation improves peptide identification on hybrid instruments.* J. Proteome Res., 2008. **7**: p. 4031-4039.

53. Houel, S., R. Abernathy, K. Renganathan, K. Meyer-Arendt, N.G. Ahn, and W.M. Old, *Quantifying the impact of chimera Ms/Ms spectra on peptide identification in large-scale proteomics studies.* J. Proteome Res., 2010. **9**(8): p. 4152-4160.

54. Cox, J. and M. Mann, *MaxQuant enbables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification.* Nat. Biotechnol., 2008. **26**(12): p. 1367-1372.

55. Thingholm, T.E., G. Palmisano, F. Kjeldsen, and M.R. Larsen, *Undesirable charge-enhancement of isobaric tagged phosphoppetides lead to reduced identification efficiency.* J. Prot. Res., 2010. **9**(8): p. 4045-4052.

56. Shliaha, P.V., R. Jukes-Jones, A. Christoforou, J. Fox, C. Hughes, J. Langridge, K. Cain, and K.S. Lilley, *Additional precursor purification in isobaric mass tagging experiments by traveling wave ion mobility separation (TWIMS).* J Proteome Res, 2014. **13**(7): p. 3360-9.

57. Cherkasov, V., S. Hofmann, S. Druffel-Augustin, A. Mogk, J. Tyedmers, G. Stoecklin, and B. Bukau, *Coordination of translational control and protein homeostatsis during severe heat stress.* Curr. Biol., 2013. **23**: p. 2452-2462.

58. Wallace, E.W.J., J.L. Kear-Scott, E.V. Pilipenko, M.H. Schwartz, P.R. laskowski, A.E. Rojek, C.D. Katanski, J.A. Riback, M.F. Dion, A.M. Franks, et al., *Reversible specific active aggregates of endogenous proteins assemble upon heat stress.* Cell, 2015. **162**(6): p. 1286-1298.

59. Richter, K., M. haslbeck, and J. Buchner, *The heat shock response: life on the edge of death.* Mol. Cell, 2010. **40**: p. 253-266.

60. Jolly, C. and R.I. Morimoto, *Role of heat shock response and molecular chaperone in oncogenesis and cell death.* J. Nat. Cancer Res. Institute, 2000. **92**(19): p. 1564-1572.

61. Kim, Y.E., M.S. Hipp, A. Bracher, M. Hayer-Hartl, and F.U. Hartl, *Molecular chaperone functions in protein folding and proteostasis.* Annu Rev Biochem, 2013. **82**: p. 323-55.

62. Chen, Y., L.N. Yang, L. Cheng, S. Tu, S.J. Guo, H.Y. Le, Q. Xiong, R. Mo, C.Y. Li, J.S. Jeong, et al., *Bcl2-associated athanogene 3 interactome analysis reveals a new role in modulating proteasome activity.* Mol Cell Proteomics, 2013. **12**(10): p. 2804-19.

63. Anderson, P. and N. Kedersha, *Stress granules: the Tao of RNA triage.* Trends Biochem Sci, 2008. **33**(3): p. 141-50.

64. Buchan, J.R. and R. Parker, *Eukaryotic stress granules: the ins and outs of translation.* Mol Cell, 2009. **36**(6): p. 932-41.

65. Slimen, I.B., T. Najar, A. Ghram, H. Dabbebi, M. Ben Mrad, and M. Abdrabbah, *Reactive oxygen species, heat stress and oxidative-induced mitochondrial damage. A review.* Int J Hyperthermia, 2014. **30**(7): p. 513-23.

66. Lionaki, E. and N. Tavernarakis, *Oxidative stress and mitochondrial protein quality control in aging.* J Proteomics, 2013. **92**: p. 181-94.

67. Huang, S.H., K.J. Yang, J.C. Wu, K.J. Chang, and S.M. Wang, *Effects of hyperthermia on the cytoskeleton and focal adhesion proteins in a human thyroid carcinoma cell line.* J Cell Biochem, 1999. **75**(2): p. 327-37.

68. Pawlik, A., J.M. Nowak, D. Grzanka, L. Gackowska, J. Michalkiewicz, and A. Grzanka, *Hyperthermia induces cytoskeletal alterations and mitotic catastrophe in p53-deficient H1299 lung cancer cells.* Acta Histochem, 2013. **115**(1): p. 8-15.

69. Wang, T.T., A.S. Chiang, J.J. Chu, T.J. Cheng, T.M. Chen, and Y.K. Lai, *Concomitant alterations in distribution of 70 kDa heat shock proteins, cytoskeleton and organelles in heat shocked 9L cells.* Int J Biochem Cell Biol, 1998. **30**(6): p. 745-59.

CHAPTER THREE

3. Accurate quantitative proteomic analyses using metabolic labeling and High Field Asymmetric Waveform Ion Mobility Spectrometry (FAIMS)

Sibylle Pfammatter[1,2], Eric Bonneil[1], Francis P. McManus[1], and Pierre Thibault[1,2]*

[1]Institute for Research in Immunology and Cancer, Université de Montréal, C.P. 6128, Succursale centre-ville, Montréal, Québec, H3C 3J7, Canada.

[2]Department of Chemistry, Université de Montréal, C.P. 6128, Succursale centre-ville, Montréal, Québec, H3C 3J7, Canada.

Published:

Journal of Proteome Research, 2019, 18 (5), pp 2129–2138

Reprinted with permission from JOURNAL OF PROTEOME RESEARCH.

Copyright © (2019) American Chemical Society

Author contributions:

Sibylle Pfammatter designed, executed and analyzed all experiments, generated all figures and wrote the paper. Eric Bonneil (MS platform manager) helped to design the project, supported the data acquisition and analysis, and contributed to the draft of the manuscript. Francis P. McManus (lab manager) contributed to the data interpretation of the biological processes and the preparation of the manuscript. Pierre Thibault (supervisor) managed the project. All authors approved the content and submission of the paper.

3.1. Abstract

Stable isotope labeling by amino acids in cell culture (SILAC) is routinely used to profile changes in protein and peptide abundance across different experimental paradigms. As with other quantitative proteomic approaches, the detection of peptide isotopomers can be limited by the presence of interference ions that ultimately affect the quality of quantitative measurements. Here, we evaluate high field asymmetric waveform ion mobility spectrometry (FAIMS) to improve the accuracy and dynamic range of quantitative proteomic analyses using SILAC. We compared quantitative measurements for tryptic digests of isotopically labeled protein extracts mixed in different ratios using LC-MS/MS with and without FAIMS. To further reduce sample complexity, we also examined the improvement in quantitative measurements when combining strong cation exchange (SCX) fractionation prior to LC-MS/MS analyses. Using the same amount of sample consumed, analyses performed using FAIMS provided more than 30% and 200% increase in the number of quantifiable peptides compared LC-MS/MS performed with and without SCX fractionation, respectively. Furthermore, FAIMS reduced the occurrence of interfering isobaric ions and improved the accuracy of quantitative measurements. We leveraged the application of FAIMS in phosphoproteomic analyses to profile dynamic changes in protein phosphorylation in HEK293 cells subjected to heat shock for periods up to 20 min. In addition to the enhanced phosphoproteomic coverage, FAIMS also provided the ability to separate phosphopeptide isomers that often co-elute and can be misassigned in conventional LC-MS/MS experiments.

3.2. Introduction

Global proteomic experiments rely on robust identification and quantitation of proteins. This quantitation is essential to further understand the functions of proteins and their dynamic expression changes [1, 2]. Multiple strategies and workflows have been developed to quantify protein abundance by mass spectrometry (MS) and to profile their corresponding modifications [3, 4]. Some of these strategies rely on isotopic labeling in the form of chemical functionalization [5-7] or metabolic incorporation [8]. One such method that has garnered a wide audience is stable isotope labeling by amino acids in culture (SILAC) [9]. SILAC experiments consist of providing the cultured cells with isotopically labeled amino acids (e.g. arginine and lysine) directly in the media that are incorporated directly into the growing polypeptide during translation. Cells can be labeled with light, medium or heavy amino acids, allowing the quantification of three conditions simultaneously in a single LC-MS analysis. The isotopically labeled cells are pooled after collection and proteins are extracted and digested with trypsin. Since its inception in 2002 by Ong *et al.* [10, 11], SILAC has been widely used for quantitative proteomics [8, 12, 13]. The primary advantage of this method is its robustness, including the reduced variability during technical sample preparation and LC-MS measurement due to the early combination of the different conditions. Nevertheless, SILAC applications are primarily limited to cell culture samples to facilitate the comparison of up to three conditions simultaneously. For this reason, clinical samples such a tissue biopsies are not amenable to such a method. In addition, it suffers from the bias of increased mass spectral complexity since the quantification is performed in the survey scan, where resequencing of redundant peptides impact the overall proteome coverage. With increasing sample complexity, a diminished number of identification is expected due to variable abundance of peptide isotopologues [10]. Furthermore, the accuracy of the SILAC pair/triplet ratios can be distorted due to co-eluting isobaric peptides and interfering ions. In an effort to improve the depth and accuracy of proteomic analyses, complex protein digests are often fractionated using strong cation exchange (SCX) [14] or high pH reverse phase [15] chromatography prior to LC-MS analyses. Fractionation, whether performed offline or online, require additional analysis time and sample handling steps that can lead to sample loss and lower reproducibility.

We therefore sought to examine alternative approaches to reduce sample complexity without the drawbacks of additional instrument time or chromatographic fractionation. We surmised that high-field asymmetric waveform ion mobility spectrometry (FAIMS) could improve the dynamic range and precision of quantitative proteomic analyses based on preliminary experiments using metabolic labeling [16]. This technique has improved greatly in recent years [17] and its versatility has been demonstrated. Several groups have shown its application for proteins [18-20], peptides [21-24] as well as post-translational modifications [25-28]. In FAIMS, ions are separated between an inner and outer electrodes to which is applied an alternating high and low electric field. Ions experience a drift towards one of the electrodes based on their difference in mobility between high and low electric fields. To transmit ions to the mass spectrometer, a compensation voltage (CV) is applied. The CV value at which ions are transmitted is specific to each ion and facilitates the separation of ion populations while simultaneously reducing mass spectral complexity [29].

In this work, we evaluated the analytical merits of FAIMS in quantitative proteomic analyses performed using SILAC. We compared the depth and accuracy of quantitative measurements using LC-MS/MS with and without FAIMS for protein digests mixed in different ratios. We also investigated the use of metabolic labeling and FAIMS to profile the dynamic changes in protein phosphorylation of HEK293 cells exposed to heat shock.

3.3. Experimental Section

Cell culture and sample preparation. For SILAC experiments human Embryonic Kidney 293 cells (HEK 293) were grown in Dulbecco's Modified Eagle Medium (DMEM, Thermo Fisher Scientific) containing 10% heat inactivated dialyzed fetal bovine serum (Wisent), 1% L-Glutamine (Gibco) and 1% Penicillin/Streptavidin (Gibco). Cells were grown in medium supplemented with either Arg0/Lys0 (light), Arg6/Lys4 (medium) or Arg10/Lys8 (heavy) lysine and arginine stable isotopes (Silantes, Germany) and were cultured at 37°C and 5% CO_2. Light, medium and heavy labeled cells for different ratio experiments were collected simultaneously and washed twice with cold Phosphate Buffer Saline (PBS) and centrifuged at 215 g for 10 minutes. Cell pellets were snap frozen in liquid nitrogen before lysis in a buffer composed of 8M Urea, 50mM HEPES, 75mM NaCl at pH 8.5 followed by 3 x 5 sec sonication bursts. After centrifugation (13,000 rpm, 10 min), supernatants were transferred to a new tube and protein concentration determined by Bradford assay. SILAC channels were combined in at 13.5 : 5.4 : 1 (wt:wt:wt) ratio and diluted to 1 M Urea with 50 mM ammonium bicarbonate. Samples were reduced with 5 mM Tris(2-carboxyethyl)phosphine hydrochloride (TCEP-HCl, Thermo Fisher Scientific) and alkylated with 10 mM 2-Chloroacetamide (Aldrich). Digestion was performed overnight with Trypsin (1:50 enzyme:peptide ratio, Promega) Peptides were desalted on Oasis HLB cartridges (Waters) according to the manufacturer's instructions, dried down and reconstituted in 0.2% formic acid (FA) for MS analysis.

For heat shock experiments, the media was removed from labeled cells and replaced with preheated media at 43°C and incubation at 43°C, 5% CO_2 for 10 or 20 min. Untreated and stressed cells were collected at the same time, washed twice with PBS, pelleted and frozen in liquid nitrogen until reduction, alkylation and trypsin digestion as indicated above.

Phosphopeptide enrichment. In-house TiO_2 spin tips were packed with two C_{18} Empore membrane (3M) plugs and filled with 5 mg of TiO_2 beads (Canadian Life Science). Before sample loading (1 mg of peptide per tip), spin tips were equilibrated with 150 µL of loading buffer (250 mM lactic acid, 70% acetonitrile, ACN and 3%

trifluoroacetic acid, TFA). After sample loading, nonspecific binding peptides were removed by washing the tips with 100 µL of loading buffer, followed by 100 µL Q/N buffer (125 mM asparagine, 125 mM glutamine in 70% ACN and 3% TFA) and 100 µL of wash buffer (70% ACN and 3% TFA). Phosphopeptides were eluted with 200 µL of 1% ammonium hydroxide, pH 10.5 and 100 µL of washing buffer. After HLB desalting, phosphopeptides were frozen in liquid nitrogen, dried in a Speedvac and stored at −80°C prior to LC-MS analysis.

LC-MS/MS analyses. All samples were separated on a nano-LC-ultra 2D system (Eksigent) coupled to a LTQ-Orbitrap Elite mass spectrometer (Thermo Fisher Scientific). Peptides were loaded on a 360 µm ID × 5 mm pre-column (Jupiter C_{18}, 3 µm, 300 Å, Phenomenex) and separated on a 150 µm ID x 20 cm in house packed C_{18} column. Peptides were separated using a linear gradient of 5-30 % ACN (0.2% FA) in 60 min at a flow rate of 600 nL/min. For strong cation exchange (SCX) chromatography, samples were loaded on a Optiguard SCX trap column, 5 µm, 300 Å, 0.5 ID × 23 mm (Optimize Technologies) and eluted online with salt plug fractions of ammonium acetate buffer, pH 3.5. Unless indicated, SCX fractionation was performed on 24 µg of tryptic digest and peptides were eluted in 12 salt-fractions ranging from 0 to 2000 mM (in 200 mM steps) followed by a final fraction with 2500 mM ammonium acetate. For phosphoproteomic experiments, tryptic digests from 1.6 mg of protein extract were fractionated by SCX into 5 SCX fractions (250, 500, 750, 1000 and 2000 mM) prior to LC-MS/MS analyses. For MS analyses, survey scans were acquired from 300 to 1200 *m/z* with 60,000 resolution at *m/z* 400. MS/MS spectra were acquired either in the ion trap or in the Orbitrap at a resolution of 15,000 with a 2.0 Th isolation width. Normalized collision energy (NCE) for collision-induced dissociation (CID) was set at 35 whereas higher energy collisional dissociation (HCD) spectra used a NCE of 28.

FAIMS. For differential ion mobility separation, FAIMS was interfaced to a nanoelectrospray source coupled to a LTQ-Orbitrap Elite (Thermo Fisher Scientific). Pure nitrogen was used at a flow rate of 2.3 L/min and the FAIMS device comprised two cylindrical electrodes spaced by a 1.5 mm gap [30]. The temperature of the inner and outer electrodes were set to 70°C and 90°C, respectively. For LC-FAIMS-MS/MS

analyses, the dispersion voltage (DV) was set at -5000V and the compensation voltage was stepped in 3V increments from -30V to -63V unless otherwise indicated.

Data Analysis. RAW files from the Xcalibur software (version 2.1) were searched with MaxQuant (version 1.5.3.8) against the Human UniProt Database (February 17, 2016), including the reversed database (74,508 entries). The RAW output files from FAIMS experiments acquired with CV stepping were first split into individual mzXML files (each CV step was converted to a different mzXML) with an in-house Python script prior to their analysis with MaxQuant. Trypsin/P was selected as a specific enzyme with a maximum of three missed cleavages and a minimum peptide length of seven amino acids. Methionine oxidation, N-terminal acetylation, deamidation (NQ) were set as variable modifications and carbamidomethylation (C) was set as a fixed modification. Arg10 and Lys8 were included in the search for the heavy labeled SILAC samples and Arg6 and Lys4 for medium labeled samples. The precursor ion tolerance was set at 20 ppm and fragment ion tolerance to 0.3 Da (ion trap CID) or 0.01 Da (Orbitrap HCD).

Raw Data Repository. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [31] partner repository with the dataset identifier PXD012924.

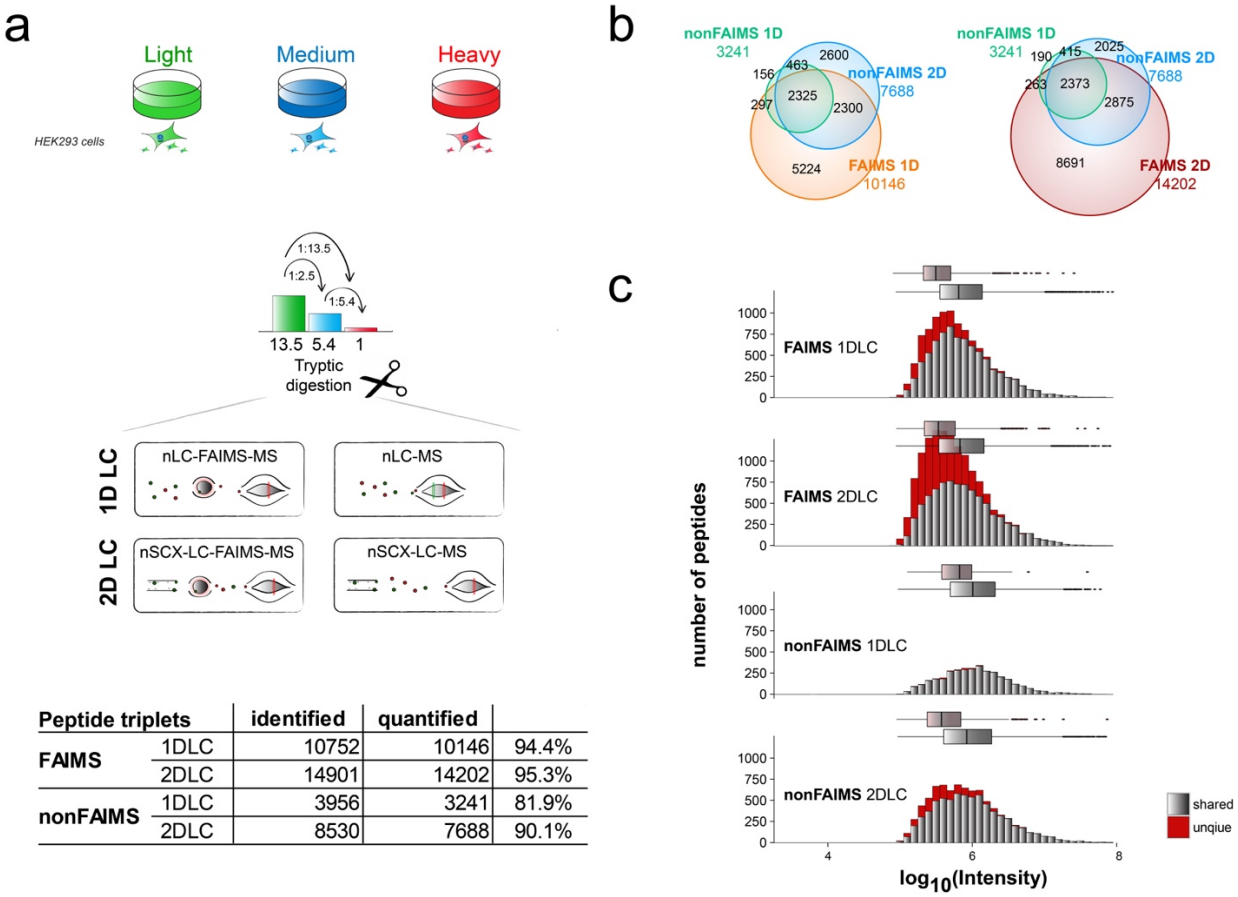
3.4. Results and Discussion

3.4.1. FAIMS enhances the depth and precision of quantitative proteomics using SILAC

In a previous investigation, we reported that FAIMS reduces the variability in quantitative measurements performed using LC-MS/MS when isotopically labeled peptides are mixed in equal proportions [16]. To provide an in depth analysis of the benefits of FAIMS for quantitative proteomics using metabolic labeling, we evaluated the precision accuracy and dynamic range of quantification for LC-MS/MS analyses performed with and without ion exchange fractionation to reduce sample complexity. We cultured HEK293 cells in light, medium and heavy media that contained isotopically labeled lysine and arginine amino acids and mixed the corresponding cell extracts in the proportion of 13.5 : 5.4 : 2.5, respectively (***Figure 3-1a***). Following tryptic digestion, samples were analyzed by LC-MS/MS with and without FAIMS. For FAIMS analyses, we stepped the CV values from -30V to -63V in 3V increments. This range enabled the transmission of the densest population of multiply charged ions, thereby reducing the contribution of singly charged ions [27]. LC-FAIMS-MS/MS were performed at individual CV and the same number of injections and amount of HEK293 digests were used for each set of experiments.

In total, we identified 10,752 and 3,956 unique peptides, of which 10,146 and 3,241 peptides were quantified in LC-MS/MS experiments performed with and without FAIMS, respectively (***Supplementary Table 3-1***). To further improve the comprehensiveness of proteome analysis, we performed SCX fractionation prior to LC-MS/MS analysis and obtained up to 12 salt fractions ranging from 0 to 2.5 M ammonium acetate. SCX fractionation enabled a two-fold enhancement in peptide identification with a total of 8,530 unique peptides (***Figure 3-1a***), and led to a higher proportion of quantifiable peptides (90.1%, 7,688 peptide triplets) compared to LC-MS/MS (81.9%, 3,142 peptide triplets). However, in spite of the increased proteome coverage available with SCX, the number of quantifiable peptides were significantly lower than those achieved with LC-FAIMS-MS/MS alone (94.4%, 10,146 peptide triplets). Overall, FAIMS-LC-MS/MS provided 32% and 210 % increase in the number of quantifiable

peptides compare to LC-MS/MS with and without SCX fractionation, respectively. It is noteworthy that the combination of SCX with LC-FAIMS-MS/MS led to a further increase in peptide identification with 14,202 peptide triplets for the same sample loading (***Figure 3-1a***).



Peptide triplets		identified	quantified	
FAIMS	1DLC	10752	10146	94.4%
	2DLC	14901	14202	95.3%
nonFAIMS	1DLC	3956	3241	81.9%
	2DLC	8530	7688	90.1%

Figure 3-1: *FAIMS extends the depth of quantitative proteome analyses using SILAC. a) Experimental overview: HEK293 proteins grown in light, medium and heavy SILAC media were combined together at a ratio 13.5 : 5.4 : 1.0. Samples were analyzed by LC-MS/MS (1DLC) with and without FAIMS. Fractionation using SCX (2DLC) was used to reduce sample complexity prior to LC-MS/MS analyses. The number of identified and quantified peptides for the various experiments are listed. b) Venn diagrams showing the overlap of peptides identified between FAIMS and non-FAIMS experiments. c) Bar graph showing the number of identified peptides as function of their intensity. Grey bars depict peptides identified by both strategies while red bars depict peptides identified by FAIMS or non-FAIMS only (unique).*

Gas phase ion fractionation using FAIMS is complementary to chromatography separation as indicated by the overlap in the Venn diagram of ***Figure 3-1b***. Out of a total of 13,365 unique peptide triplets, 76 %, 58%, and 24% were identified with LC-FAIMS-MS/MS, LC-MS/MS with and without SCX, respectively. To understand the differences between peptide populations identified with or without FAIMS, we compared the distribution of peptide intensities along with peptides uniquely detected by each method. As reported before [32] and evidenced in the intensity distribution in ***Figure 3-1c***, peptides exclusively detected with or without FAIMS correspond to peptides of low intensity. Peptides exclusive to non-FAIMS analyses are also of low abundance, but were not quantified with FAIMS, possibly due to their optimal transmission requiring CV values outside of the range under investigation.

To determine the nature of common and unique peptide sequences in greater detail, we first looked at the charge state distribution (***Supplementary Figure 3-1a***). Peptides that were uniquely identified by FAIMS showed a greater proportion of triply charged ions than those identified by non-FAIMS, consistent with a previous report by Creese *et al*. [33] This is partly explained by the ability of FAIMS to reduce suppression effects that undermine the transmission of multiply-charged peptide ions when approaching the ion capacity defined by the automatic gain control (AGC). Indeed, a previous study found that the distribution of multiply-charged ions from synthetic peptides spiked into a tryptic digest is generally skewed toward lower charge state ions when the concentration of the digest is progressively increased [27]. On the other hand, FAIMS does not show a transmission bias in terms of peptide length (***Supplementary Figure 3-1b***) or *m/z* range (***Supplementary Figure 3-1c***). As stated earlier, peptides uniquely detected with or without FAIMS were of lower intensity. These peptides also displayed lower scores relative to common peptides (***Supplementary Figure 3-1d***). Interestingly, we noted that FAIMS identified a higher number of peptides with missed cleavages. This observation is likely associated to the higher charge state of the corresponding peptides which favors their identification by FAIMS when present at low abundance in the protein digest (***Supplementary Figure 3-2a***). Despite the missed cleavages, these peptides show the same trends as peptides without missed cleavages

(***Supplementary Figure 3-2b***) and enhance the comprehensiveness and quantification of FAIMS analyses.

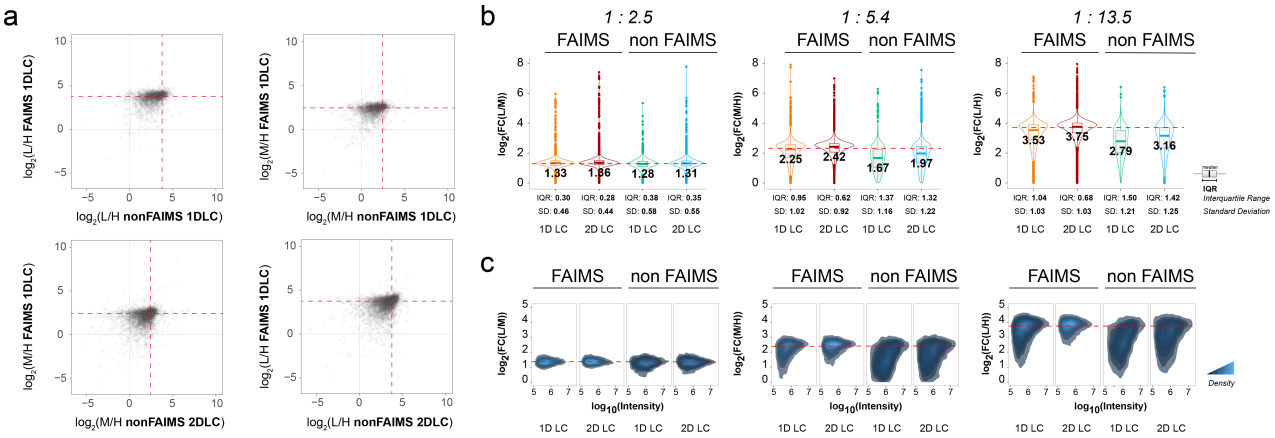


Figure 3-2: *FAIMS improves the accuracy of quantitative measurements performed using SILAC and LC-MS/MS. a) Scatterplot comparing the accuracy for fold changes 1.0 : 13.5 (upper panels) and 1.0 : 5.4 (bottom panels) for 1D LC (left panels) and 2D LC (right panels) methods. Red lines indicate the expected fold change ratio. b) Violin plots showing the distribution of fold change measurements obtained for each method. c) Density map comparing the intensity and the fold change observed for all quantified peptides using all four methods. The dotted red lines represent the expected fold change.*

Next, we evaluated the precision and accuracy of fold change measurements from experiments described in ***Figure 3-1***. Values observed with and without FAIMS are displayed in ***Figure 3-2a*** for all quantified peptides according to each expected ratios (shown as dotted red lines on the corresponding plots). For a predicted fold change of 2.5 ($\log_2$ L/M: 1.32), LC-MS/MS experiments performed with and without FAIMS gave similar ratios with a uniform distribution centered on a $\log_2$ L/M of 1.33 and 1.28, respectively (***Figure 3-2b***). However, for higher predicted ratios ($\log_2$ M/H: 2.43 and $\log_2$ L/H: 3.75) the values observed for LC-MS/MS showed higher ratio compression with $\log_2$ M/H: 1.67 and $\log_2$ L/H: 2.79, compared to those obtained for LC-FAIMS-MS/MS experiments ($\log_2$ M/H: 2.25 and $\log_2$ L/H: 3.53). This underestimation was partly alleviated using SCX fractionation prior to LC-MS/MS (***Figure 3-2b***). Also, the distribution becomes wider and more asymmetric for higher ratios, and the extent of ratio compression is more significant with increasing SILAC ratios without ion mobility separation (***Figure 3-2a, Figure 3-2b***). We also noted that standard deviation (SD)

obtained for these measurements were generally 10% lower for FAIMS compared to non-FAIMS experiments (***Figure 3-2b***). The underestimation of fold change measurements was noted previously by Hogrebe *et al.* for common quantification strategies in large-scale phosphoproteomic experiments including SILAC [4]. The extent of ratio compression is directly related to peptide ion intensity (***Figure 3-2c***). At higher fold change (middle and right plots in ***Figure 3-2c***), low abundance ions show more pronounced underestimation that is more significant in the non-FAIMS experiments. Interestingly, LC-FAIMS-MS/MS with prior SCX fractionation alleviates this issue, and the distribution of fold change measurements is closer to the expected ratios (***Figure 3-2b, Supplementary Figure 3-3***). It is noteworthy that similar observations were also obtained in a separate set of experiments where cell extracts labeled with different stable isotopes were mixed in a proportion ranging from 1.7 to 17.0 (***Supplementary Figure 3-4***).

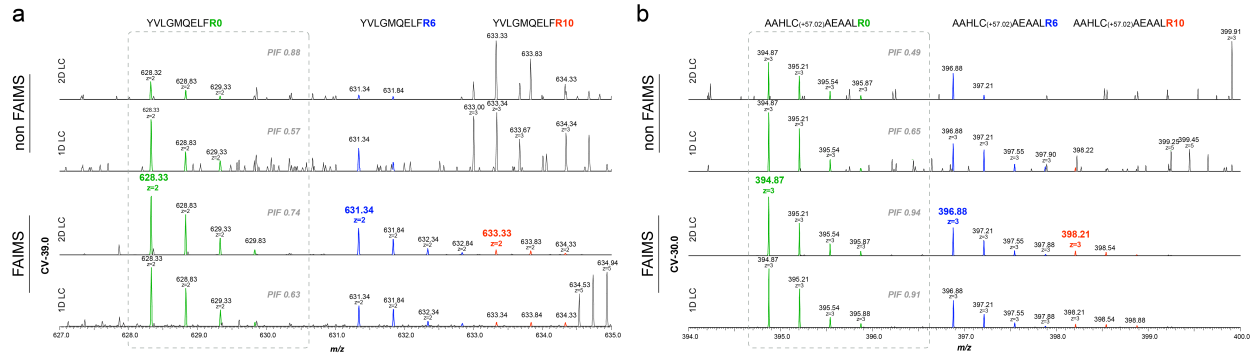


Figure 3-3: *FAIMS reduces sample complexity and facilitate the detection of low abundance peptide isotopomers. a) Representative mass spectra for the doubly charged peptide YVLGMQELFR showing the light (m/z 628.33), medium (m/z 631.34) and heavy (m/z 633.33) isotopomers. FAIMS reduced mass spectral complexity and facilitated the quantification of the heavy peptide isotopomer. b) Representative mass spectra for the triply charged peptide AAHLCAEAAL highlighting the light (m/z 394.87), medium (m/z 396.88) and heavy (m/z 398.21) isotopomers. The detection of all peptide isotopomers was achieved using FAIMS whereas conventional LC-MS analyses performed with or without SCX pre-fractionation still showed interfering peptide ions masking the heavy isotopomers. The precursor ion fraction (PIF) that defines the proportion of target peptide ion intensity in the MS/MS selection window is shown for each light peptide isotopomer.*

The ability of FAIMS to reduce the contribution of interfering ions proximal to peptide isotopomers improves the detection of low abundance ions that may be underrepresented in complex survey scans. Increased ratio compression was correlated

with a lower base peak fraction, where the reduced proportion of precursor ions relative to the base peak in the corresponding survey scan resulted in more significant underestimation of fold change ratios (***Supplementary Figure 3-5***). Another source of ratio compression stems from the data analysis software that confound interfering ions with peptide isotopomers, as exemplified in ***Figure 3-3*** for the heavy labeled peptide ions. However with FAIMS, the reduction of interfering ions improves the sensitivity and detection of the heavy peptide isotopomers. For the peptide YVLGMQELFR in ***Figure 3-3a***, the spectra without FAIMS for either 1D or 2DLC show intense interfering ions contributing to quantification errors. Missing values for low abundance isotopomers are a common source of error in SILAC experiments. Frequently, one peptide isotopomer is missing due to abundant neighboring ions in LC-MS/MS, but can be efficiently detected using LC-FAIMS-MS/MS (***Figure 3-3b***). The proportion of quantification errors increases with the fold change ratio. As noted in ***Figure 3-2c***, fold change ratios showed a progressive underestimation with decreasing peptide intensity. While most peptide ions quantified with and without FAIMS fall within ± 3 SDs of the expected ratio at low fold change values (L:M 2.5:1), the proportion of peaks estimated to be correctly quantified at high fold change values (L/H 13.5:1) decreased to 65 % and 37 % for measurements obtained with and without FAIMS, respectively (***Supplementary Figure 3-6***). SCX fractionation reduced the occurrence of quantification errors resulting in 78 % and 48% of peptides ions showing the expected fold change ratios for LC-MS/MS experiments performed with and without FAIMS, respectively (***Supplementary Figure 3-6***).

3.4.2. Profiling early signaling events upon heat shock

The benefits of FAIMS for phosphopeptide analysis were previously reported for the identification of multiphosphorylated peptides [34] and for the separation of phosphopeptide isomers [26]. Here, we used SILAC to profile the dynamic changes of protein phosphorylation in HEK293 cells upon heat shock. Control cells and cells exposed to heat shock for 10 and 20 min, were cultured in light, medium and heavy SILAC media in biological triplicate. Protein extracts from the treated cells were mixed

together and compared to control cells (***Figure 3-4a***). After trypsin digestion and phosphopeptide enrichment on TiO_2, peptides were sequentially eluted from a SCX column with five ammonium acetate fractions prior to LC-MS/MS analyses. Each fraction was analyzed with a one-hour gradient. For FAIMS, four different CV stepping ranges (-28V/-32V/-36V, -40V/-44V/-48V, -52V/-56V/-60V and -64V/-68V/-72V) were employed, where 3 different CVs were investigated for each injection. In total, we identified 3956 phosphosites with a localization confidence > 0.75, of which 3,370 and 2,316 phosphosites were identified with and without FAIMS, respectively (***Supplementary Table 3-2***). This corresponds to a gain of 45% more quantified phosphorylated sites with FAIMS compared to conventional LC-MS/MS analysis (***Figure 3-4b***). It is noteworthy that the relative standard deviations of fold change measurements progressively increased with decreasing phosphopeptide abundance, but were comparable for LC-MS/MS experiments performed with and without FAIMS (***Supplementary Figure 3-7***).

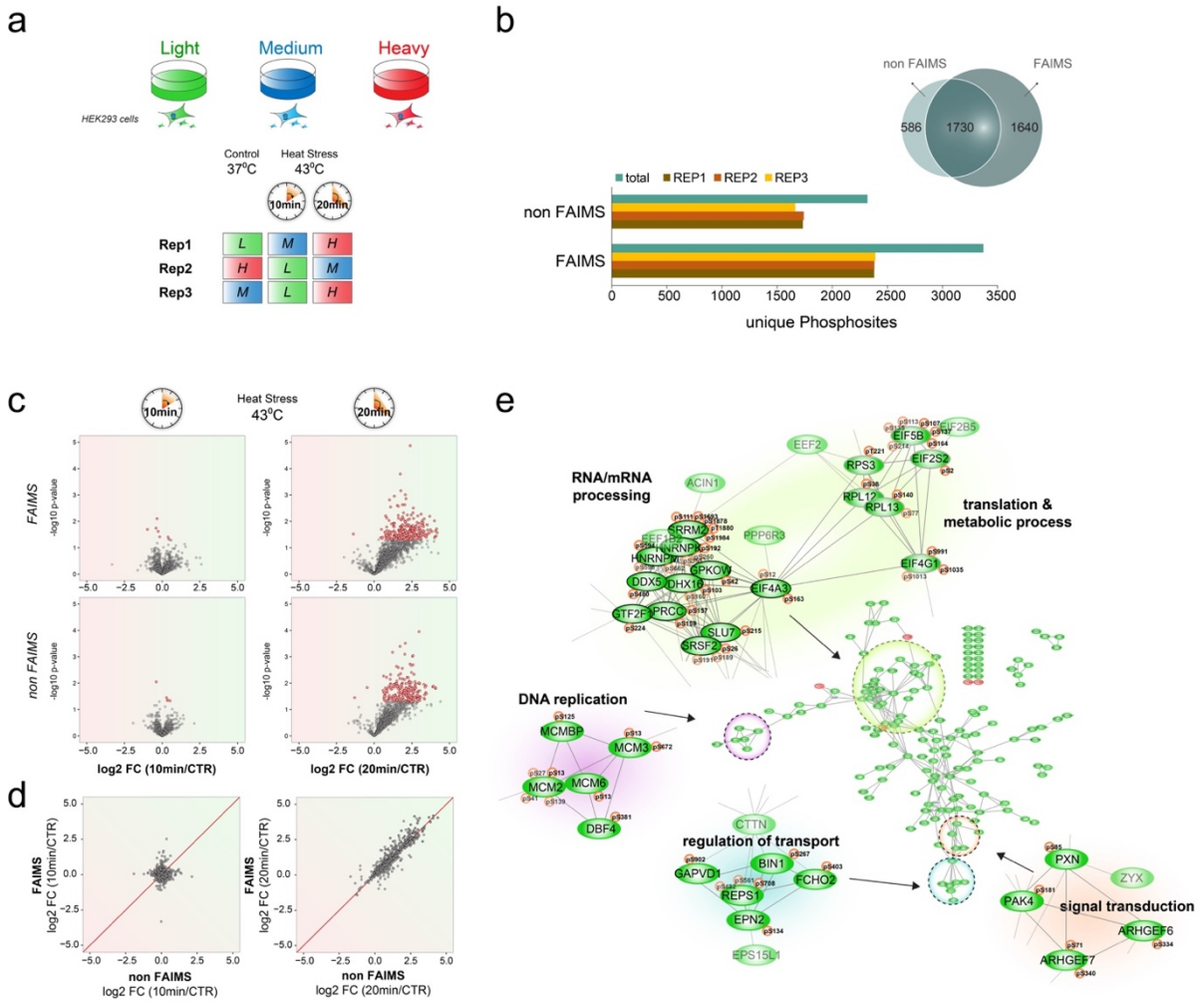


Figure 3-4*: Quantitative phosphoproteomic analyses of HEK293 cells following heat shock. a) Workflow and labeling scheme for biological replicates of HEK293 cells exposed to hyperthermia for 0, 10 and 20 min. b) Barplots of the number of identified phosphosites for replicate LC-MS/MS analyses performed with and without FAIMS along with a Venn diagram showing the overlap between conditions. c) Volcano plots showing the change in abundance of phosphopeptides after ten and twenty min of heat stress. Red data points represent significantly regulated phosphopeptides based on three biological replicates with p-value < 0.05. d) Scatterplot of phosphopeptides quantified in LC-MS/MS experiments with and without FAIMS for 10 min (left) and 20 min (right) following heat stress. e) STRING network analysis for the significantly regulated phosphoproteins using the combined dataset.*

As shown in ***Supplementary Figure 3-8***, our FAIMS program composed of 4 injections and CV stepping covered the better part of the CV transmission range of the phosphoproteome. At both ends of the CV range (i.e. -28V and -76V) the number of phosphopeptides identified is significantly lower. A large fraction of identified phosphopeptides show one or more missed cleavages. This is not surprising

considering that trypsin activity close to phosphorylated amino acids is reduced and these long peptides still remain challenging to identify [35]. Overall, a missed cleavage appears in more than half of identified phosphopeptides. Furthermore, the proportion of missed cleavages increases at higher CVs, where the lower peptide density favour their detection (***Supplementary Figure 3-8***).

Short-term exposure to moderate heat shock has been reported to induce a complex cellular response that eventually lead to a transient cell cycle arrest at G1/S transition in both yeast and mammalian cells [36]. In response to heat shock several housekeeping processes are down regulated including *de novo* protein synthesis to avoid increased misfolding and aggregation, and to maintain cell viability [37]. Significant changes in the cellular phosphoproteome were detected 20 min after heat shock where ~10% of the identified phosphopeptides (260 and 185 phosphosites with and without FAIMS, respectively) showed more than 2-fold increase in abundance (***Figure 3-4d***).

Our phosphoproteome analyses identified several kinases involved in cell cycle regulation and cell proliferation that were differentially phosphorylated upon hyperthermia including cyclin-dependent kinase 1 (CDK1), mitogen-activated protein kinase kinase kinase 2 (MAP3K2), RAF proto-oncogene serine/threonine-protein kinase (RAF1), and the serine/threonine-protein kinase mTOR (MTOR). Consistent with previous reports on the yeast ortholog CDC28, we observed an increase phosphorylation of CDK1 at the inhibitory sites T14 and Y15 following heat shock [1, 38]. The decrease activity of CDK1 observed during the heat shock response delays cell cycle progression, and allow cells to adapt to hyperthermia [1]. We observed the increased phosphorylation of MTOR at S1261, a site previously reported to promote mTORC1-mediated substrate phosphorylation to regulate cell growth and size [39]. Interestingly, we also noted the increase phosphorylation of MTOR substrate eukaryotic translation initiation factor 4E-binding protein 1 (EIF4EBP1) at S65 and T70, where hyperphosphorylation abolish binding to EIF4E [40].

To gain further biological insights from the dynamic phosphorylation events taking place during heat shock, the complete phosphoproteomic dataset was merged

and the proteins with significantly regulated phosphosites were clustered in a network (***Figure 3-4e***). Gene ontology analysis showed that changes in phosphorylation status resided on proteins involved in several biological process, including RNA/mRNA processing, translation and metabolic processes, DNA replication, regulation of transport and signal transduction.

The Serine/arginine repetitive matrix protein 2 (SRRM2), which is a key component of the spliceosome and is involved in the first catalytic step of splicing showed 5 upregulated phosphosites [41]. Although the exact role of the phosphorylation on SRRM2 is unknown, it has been shown to be modulated during the DNA damage response, whereby the global phosphorylation on this protein is decreased, suggesting that the phosphorylation of SRRM2 directly affects global RNA splicing [42]. This notion is also evidenced by the changes in the phosphorylation state of SRRM2 during the HIV response, where SRRM2 phosphorylation dynamics is linked to HIV capabilities to alter host cell alternative splicing to favour viral replication and release [43]. In addition to splicing, phosphorylation dynamics appear to play another role on RNA/mRNA processing by altering the rate of RNA transcription. We noted an increase in the phosphorylation of GTF2F1 during the heat stress. GTF2F1 is a constituent of the general transcription initiating factor II F (TIIF) that is involved in RNA transcription by recruiting RNA polymerase II to the initiation complex. The phosphorylation of GTF2F1 is proposed to inhibit transcription by reducing its ability to promote the recruitment of RNA polymerase II through transcriptional processes [44].

The heat shock caused a general increase in the phosphorylation of several proteins that regulate protein translation. We noted that different eukaryotic initiation factors (eIFs) were hyperphosphorylated under hyperthermia including EIF2S2, EIF4A3, EIF4G1 and EIF5B. eIFs play an integral part in protein synthesis since they regulate the initial phase of the protein translation process. Little is known about the phosphorylation of eIFs and the functions that they may impose on the cell. However, the phosphorylation of EIF4G, a protein whose role involves recognizing and binding the mRNA cap region to promote its association with the ribosome, promotes its interaction with EIF4A and favours the dissociation of this complex from the RNA, thereby reducing protein synthesis at the translational step [45]. Interestingly, only EIF5B S164 was

regulated acutely during cell stimulation. Although the function of this phosphorylation event is unknown we can surmise that it should impede its function considering the biological context and the fact that cells adapt to a heat shock by reducing protein synthesis [37]. Since EIF5B is responsible for the association of the initiator methionine into the initiation complex at the p-site in the later stage of the initiation process, its regulation would prompt the halting of protein translation in the final step of the initiation process. This would allow the cell to halt protein synthesis, even those events whose initiation complex formation and processing are virtually complete.

Three ribosomal proteins were hyperphosphorylated following heat shock (RPL12, RPL13 and RPS3). Notably, we observed an increase in the phosphorylation of RPS3 at T221, an event that has been shown to inhibit the association of RPS3 into the ribosomal complex [46]. We also observed the hyperphosphorylation of three proteins (MCM2, MCM3 and MCM6) from the minichromosome maintenance protein complex (MCM) as well as two adaptor proteins that regulate the MCM (DBF4 and MCMBP). The MCM complex is a group of proteins (MCM2-7) that make up a DNA helicase essential for DNA replication and correct cell cycle progression [47]. These proteins are themselves targets of cell cycle checkpoints [48]. Typically, the phosphorylation status of MCM proteins are known to fluctuate throughout the cell cycle since this modification is thought to orchestrate the dynamics of complex assembly and activity. Interestingly, the phosphorylation of MCM2 at S13, which was increased by 4-fold upon heat shock, is not regulated during normal cell cycle progression suggesting another role for this phosphorylation event [49]. Whether this phosphorylation is unique to heat shock or is more general and applies to other cellular challenges remains to be discovered.

Clearly heat stress perturbs the biology of the system and alters several key biological processes that are essential for proper cell proliferation. Overall, 407 phosphorylation sites on 330 proteins were altered in response to heat shock. Since a reduction in protein synthesis is known to occur during the heat shock response it is not surprising that we observed a phospho-dependent regulation of transcription (GTF2F1) [44] and protein translation through the regulation of the initiation complex formation (EIF4G) [45] and ribosomal activity (RPS3) [46]. We also noted changes in phosphorylation on proteins that are involved in other biological function that could

greatly impede normal growth and proliferation, such as regulation of the RNA splicing machinery (SRRM2) [42] and modulation of DNA replication through regulation of the minichromosome maintenance protein complex (MCM2, MCM3 and MCM6) [47] .

3.4.3. Gas phase ion fractionation enhances the resolution of isomeric phosphopeptides

The large-scale phosphoproteomic analyses by LC-FAIMS-MS/MS also facilitated the separation of isomeric phosphopeptides that often co-elute and lead to chimeric tandem mass spectra under conventional MS experiments. Previous reports have documented the benefits of FAIMS for the separation of isomeric peptides [26, 32, 34]. Here, we demonstrate the capability of FAIMS to separate isomeric phosphopeptides following heat shock based on their distinct CVs.

Overall, 42 isomeric phosphopeptides were identified at well separated CV values (> 4V), and examples of the corresponding peptides are listed in ***Figure 3-5a***. The MS/MS spectra of two phosphoisomers AGGPA**pT**PLSPTR and AGGPATPL**pS**PTR of Lamin-B2 are presented in ***Figure 3-5b*** and are separated at CV values of -56 and -40 V, respectively. The corresponding MS/MS spectra show distinct y-ions that enabled high localization confidence assignment. Another example of this separation is the monophosphorylated peptide AQSLVI**S**PPAP**S**PR from the source antigen KI-67 (***Figure 3-5c***). The two phosphoisomers with phosphorylation either on S579 or S584 were identified with FAIMS, while non-FAIMS only identified the phosphorylated S584. With LC-MS/MS, both peptides elute simultaneously leading to a chimeric MS/MS spectrum. Shared fragments between the two isomers are highlighted in purple, whereas unique fragments for pS579 and pS584 are highlighted in blue and red, respectively. FAIMS enables the separation of the phosphorylated peptides AQSLVISPPAP**pS**PR at CV: -28V and AQSLVI**pS**PPAPSPR at CV: -32V, and the corresponding MS/MS spectra contain fragments specific to each isoform. All separated isoforms obtained from our FAIMS dataset had, on average, a localization confidences of 94%. Moreover, the enhanced sensitivity of FAIMS facilitated the identification of peptides comprising two or more phosphorylation sites. Our LC-MS/MS analyses

performed with and without FAIMS identified 384 and 271 multi-phosphorylated peptides, respectively. The identification of the corresponding peptides is typically challenging due to suppression effects that can affect their detection [34].

a

Gene names	Peptide	Site	Localization	CV	Peptide	Site	Localization	CV
AFF4	ELLSPLSEPDDRYPLIVK	703	0.999	-60	ELLSPLSEPDDRYPLIVK	706	0.992	-52
ATP2B1	IEDSEPHIPLIDDTDAEDDAPTKR	1119	1	-44	IEDSEPHIPLIDDTDAEDDAPTKR	1129	1	-52
BCLAF1	NTPSQHSHSIQHSPER	255, 266	1, 0.964	-60	NTPSQHSHSIQHSPER	262, 266	0.972, 0.999	-52
CBX8	RQDSDLVQCGVTSPSSAEATGK	256	1	-56	RQDSDLVQCGVTSPSSAEATGK	265	0.854	-64
CTNNB1	RTSMGGTQQQFVEGVR	552	0.939	-68	RTSMGGTQQQFVEGVR	556	0.78	-72
DSG2	GSSSASIVK	700	0.909	-72	GSSSASIVK	701	0.947	-68
EPB41L3	GISQTNLITTVTPEKK	478	0.997	-44	GISQTNLITTVTPEKK	487	0.978	-40
ESYT2	SSSSLLASPGHISVK	145	0.965	-72	SSSSLLASPGHISVK	146	0.84	-60
FAM169A	AVDSSSEEIEVEVPVVDRR	635	0.965	-44	AVDSSSEEIEVEVPVVDRR	636	0.787	-40
FOXK1	EEAPASPLRPLYPQISPLK	213	1	-64	EEAPASPLRPLYPQISPLK	223	0.998	-60
IWS1	HQASDSENEEPPKPR	196	0.892	-52	HQASDSENEEPPKPR	198	0.987	-56
KMT2A	RHSTSSLSPQR	2158	0.813	-64	RHSTSSLSPQR	2163	0.999	-60
LEO1	IYSSDSDEGSEEDKAQR	550	0.989	-68	IYSSDSDEGSEEDKAQR	554	1	-48
LMNB1	AGGPTTPLSPTR	20	1	-44	AGGPTTPLSPTR	23	0.808	-36
LMNB2	AGGPATPLSPTR	14	1	-56	AGGPATPLSPTR	17	0.816	-40
MDC1	GPGAPGLAHLQESQAGSDTDVEEGKAP	372	0.795	-56	GPGAPGLAHLQESQAGSDTDVEEGK	376	0.94	-48
MKI67	AQSLVISPPAPSPR	579	0.939	-32	AQSLVISPPAPSPR	584	0.999	-28
CAPH2	ESRSPQQSAALPR	260	0.891	-44	ESRSPQQSAALPR	262	0.832	-40
RANBP2	SALSPSKSPAK	2276	1	-64	SALSPSKSPAK	2280	0.992	-32
SRRM2	GEFSASPMLK	1122	0.968	-48	GEFSAS(1)PMLK	1124	1	-40
SRRM2	ELSNSPLRENSFGSPLEFR	1318, 1320	0.791,0.813	-64	ELSNSPLRENSFGSPLEFR	1320, 1329	0.998,0.999	-56

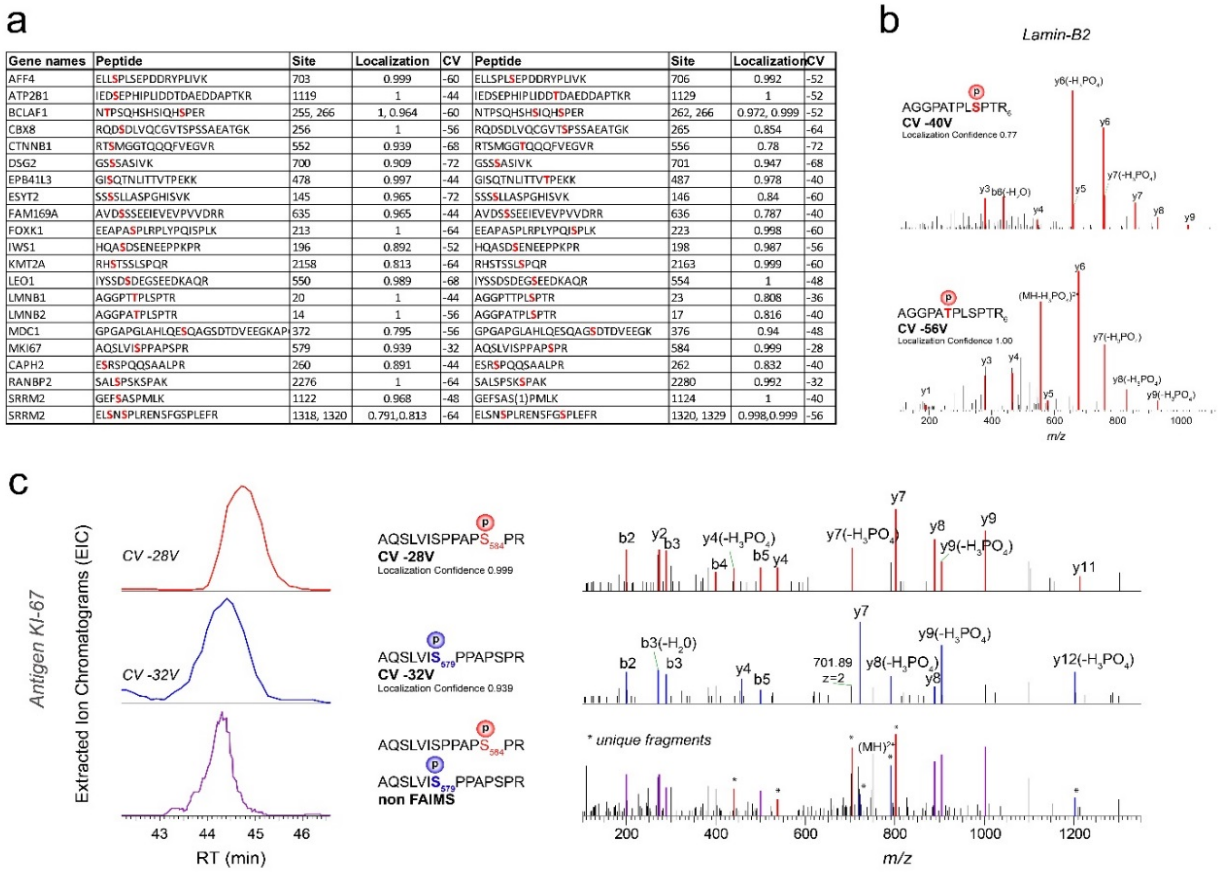


Figure 3-5: *Separation of co-eluting phosphopeptide isomers using FAIMS. a) List of 42 phosphopeptide isomers separated with ΔCV >4V. b) MS/MS spectra for the two phosphopeptides AGGPApTPLSPTR (CV -40V) and AGGPATPLpSPTR (CV -56V) of Lamin-B2. c) Ion chromatograms of phosphopeptide isomers from MKI67 in typical LC-MS (bottom) and LC-FAIMS-MS (upper two profiles). FAIMS resolved the phosphopeptide isomers AQSLVISPPAP**pS**PR (CV -28V) and AQSLVI**pS**PPAPSPR (CV -32V) allowing for the unambiguous identification of co-eluting isomers from their tandem mass spectra.*

3.5. Conclusion

Sample complexity can undermine the ability to unambiguously detect and quantify low abundance proteins by mass spectrometry. This is particularly true for metabolic labeling experiments where several isotopologues of the same peptides are present at different abundances, and can remain undetected or misassigned. In this context, FAIMS can be advantageously exploited to reduce sample complexity and the occurrence of confounding ions, thus providing several advantages in quantitative proteomics using SILAC. First, on-line gas fractionation using FAIMS extends the dynamic range of peptide detection and reduces suppression effects that typically affect the accuracy and comprehensiveness of quantitative measurements. In a head-to-head comparison of SILAC protein digest mixtures analyzed by LC-MS/MS with and without FAIMS, we found that FAIMS enhanced the number of quantifiable peptides by at least three-fold (10,146 vs 3,241 peptide triplets). While SCX fractionation provided a two-fold gain in quantifiable peptides (7,688 peptide triplets), these results were 32% lower than those achieved with LC-FAIMS-MS/MS alone. Importantly, the ability of FAIMS to reduce the occurrence of conflicting ions improved the accuracy of quantitative measurements, especially for higher predicted fold change ratios. For example, predicted fold change ratio of 13.5 were underestimated by almost 50% in LC-MS/MS experiments compared to only 15% when using FAIMS.

Second, FAIMS also improves the comprehensiveness of phosphoproteomic analyses and facilitates the profiling of phosphorylation dynamics. We combined metabolic labeling and phosphopeptide enrichment to determine the changes in early signaling events of HEK293 cells exposed to hyperthermia. The quantitative profiling of protein phosphorylation at 10 and 20 min following heat shock provided valuable information on the temporal changes in kinase and phosphatase activities and their downstream substrates in response to heat stress. In particular, these analyses enabled the identification of several important proteins regulating protein synthesis, translation, ribosomal activity, RNA splicing and DNA replication. Changes in protein phosphorylation were also consistent with a transient G1/S cell arrest that allow cells to temporally adjust to this environmental stress. Finally, FAIMS enabled the separation of

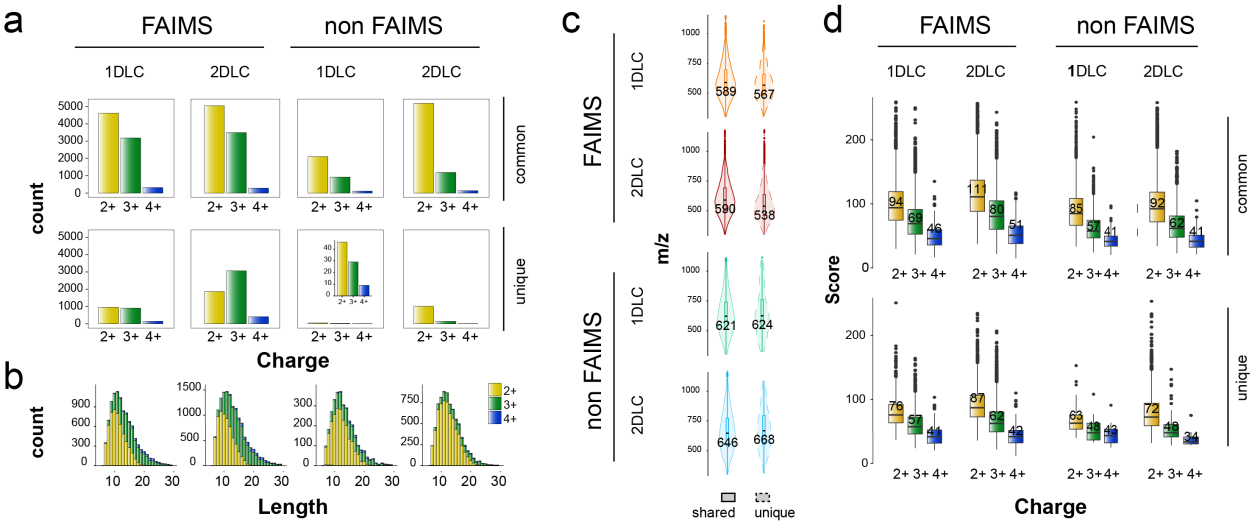
phosphopeptide isomers that shared the same amino acid sequence but differ by the location of their modification sites. Indeed, approximately two percent of all phosphopeptides corresponded to co-eluting isomers that were undistinguishable by LC-MS/MS but could be resolved using FAIMS. This unique capability of FAIMS is of obvious relevance in phosphoproteomics to correctly assign sites regulated by specific cell stimulus for subsequent follow-up experiments.

3.6. Acknowledgments

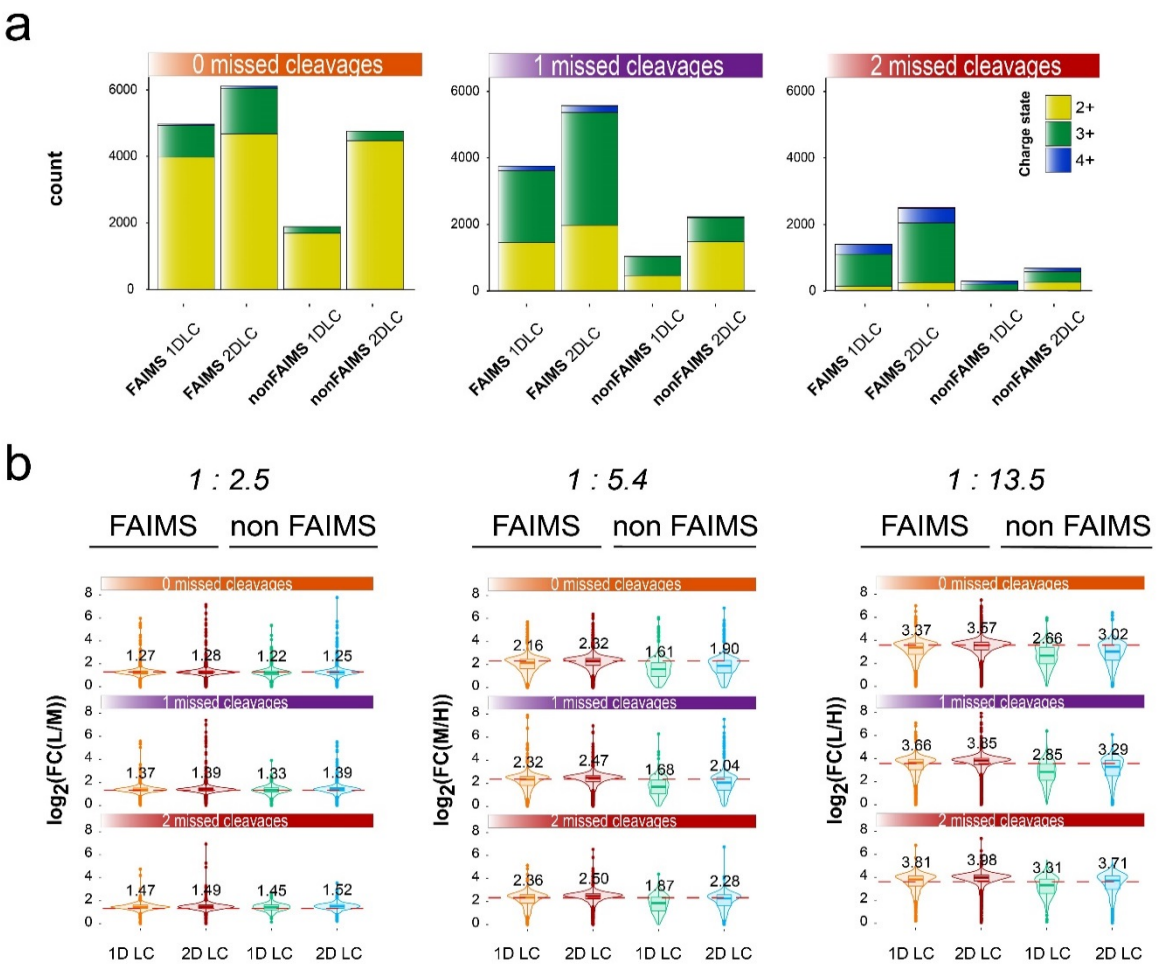
This work was carried out with financial support from the Natural Sciences and Engineering Research Council (NSERC 311598) and the Genomic Applications Partnership Program (GAPP) of Genome Canada. The authors thank Jean-Jacques Dunyach, Michael Belford and Satendra Prasad (Thermo Fisher Scientific) for valuable help and assistance with the FAIMS interface. The Institute for Research in Immunology and Cancer (IRIC) receives infrastructure support from IRICoR, the Canadian Foundation for Innovation, and the Fonds de Recherche du Québec - Santé (FRQS). IRIC proteomics facility is a Genomics Technology platform funded in part by the Canadian Government through Genome Canada.

3.7. Supplementary material

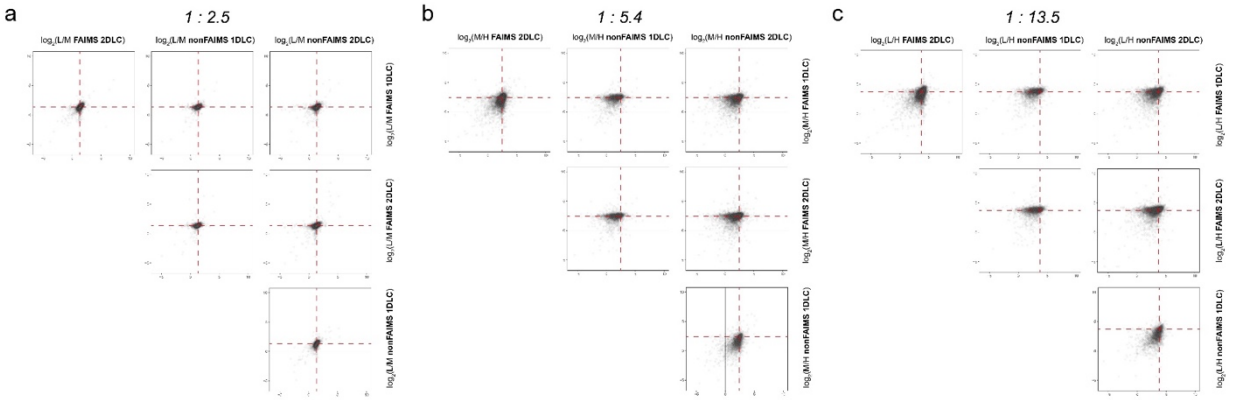
3.7.1. Supplementary figures



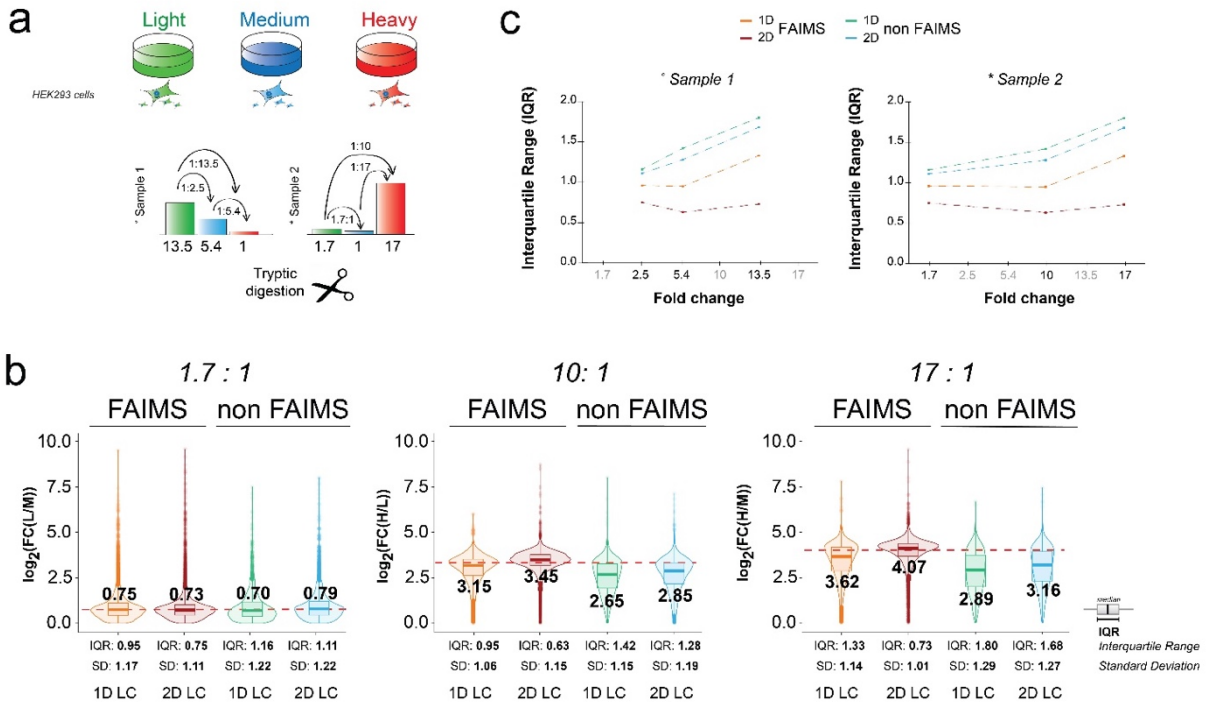
Supplementary Figure 3-1: *Physicochemical characteristics of unique and common peptide sequences identified with the various methods. a) Charge state distribution shows that FAIMS favours the identification of 3+ peptides. b) The distribution of peptide length in amino acids and its dependence on charge state c) Distribution of peptides according to m/z values (c) and Maxquant score (d).*



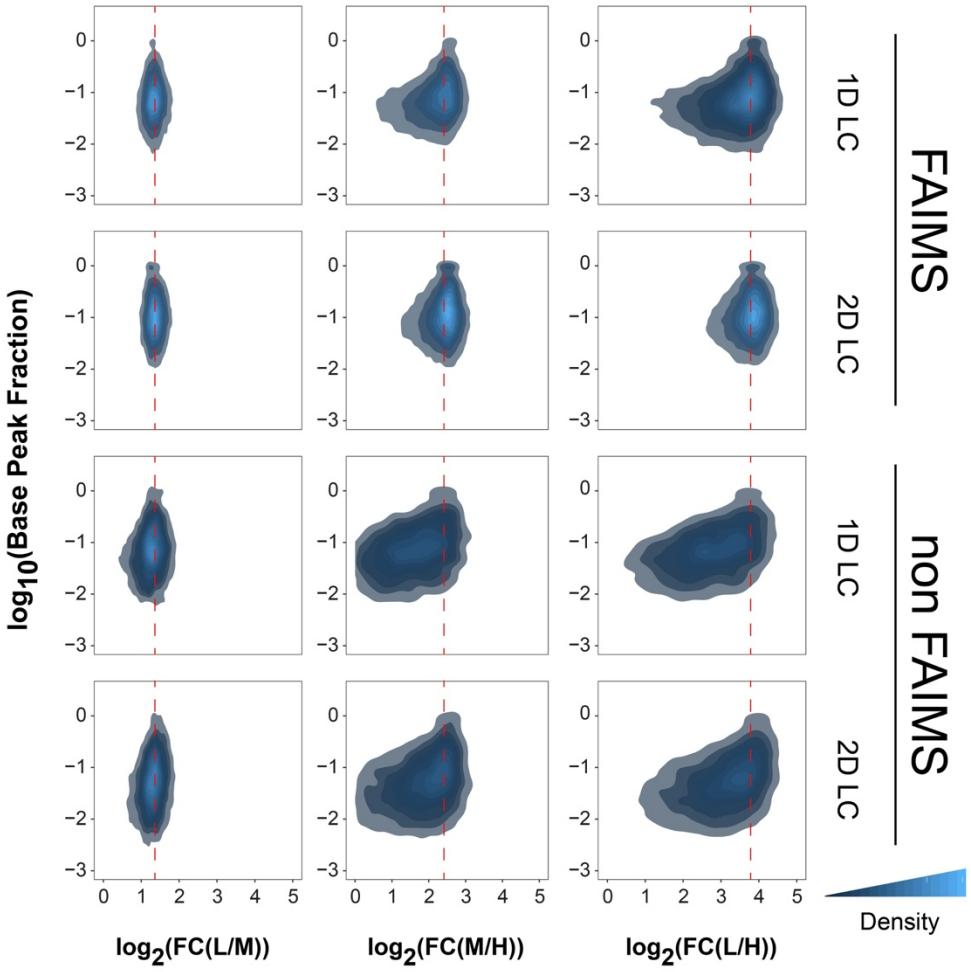
Supplementary Figure 3-2: *FAIMS expands the identification and quantification of low abundance peptides with missed cleavage sites. a) The charge state distribution of the peptides increases as a function of the number of missed cleavages. b) Missed cleaved peptides show the same ratio distribution as peptides without missed cleavages and can therefore be used for reliable quantification.*



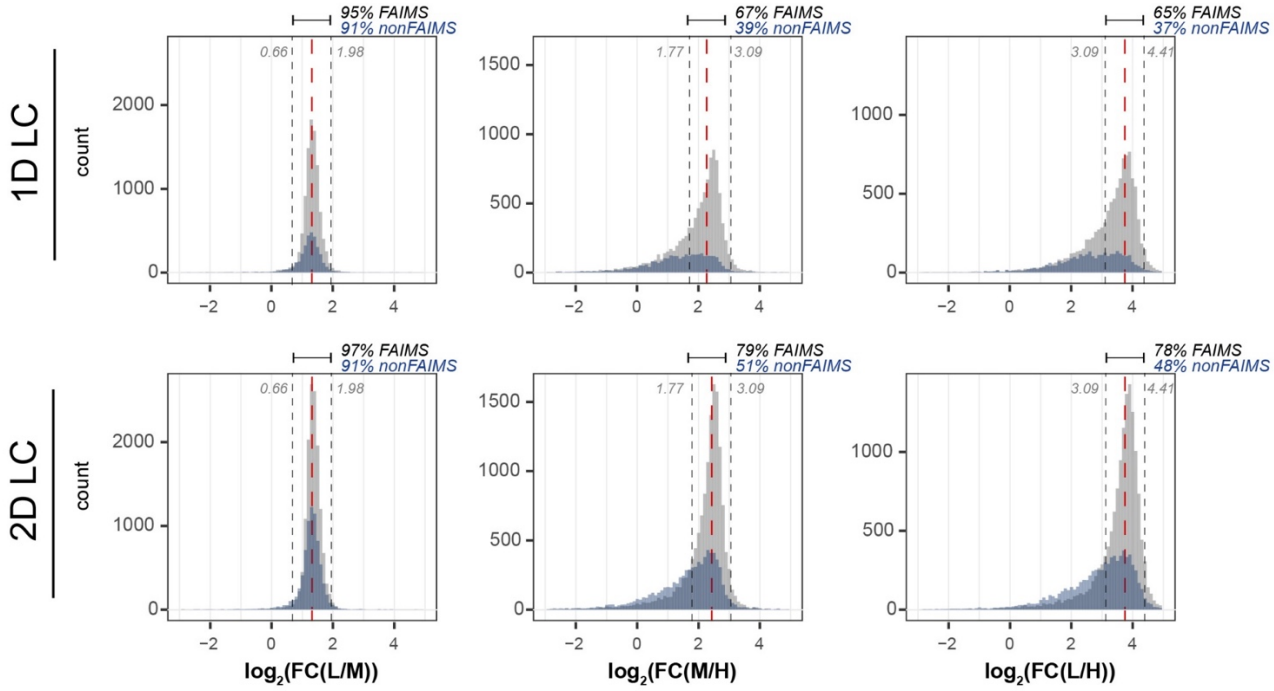
Supplementary Figure 3-3: *FAIMS improves SILAC quantitation accuracy for large fold changes. a) Ratio 1 : 2.5 for light to medium labeled peptides, b) 1 : 5.4 for medium to heavy labeled peptides and c) 1 : 13.5 for light to heavy peptides.*



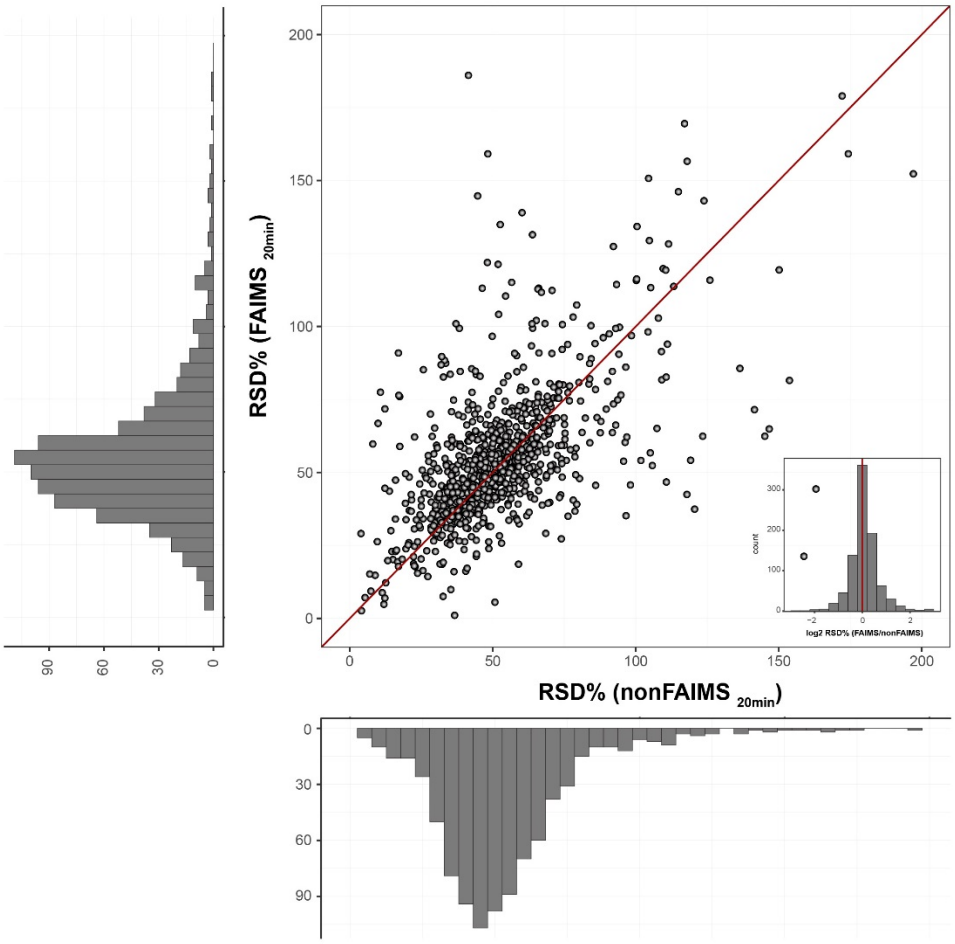
Supplementary Figure 3-4: *FAIMS improves the precision and accuracy of quantitative measurements performed using SILAC and LC-MS/MS. a) Mixing scheme showing that cell extracts with grown with different stable isotopes were mixed in different proportions to obtain fold change ratios 2.5: 5.4: 13.5 (experiment 1) and 1.7: 10.0: 17.0 (experiment 2). Violin plots comparing the observed fold change values for experiment 2 using LC-MS/MS with (2D) and without (1D) SCX fractionation for FAIMS and non-FAIMS experiments. Red lines indicate the expected fold change ratio. SD indicates standard deviation for each method. c) Variation of interquartile range (IQR) vs fold change ratio for experiments 1 and 2 using different methods.*



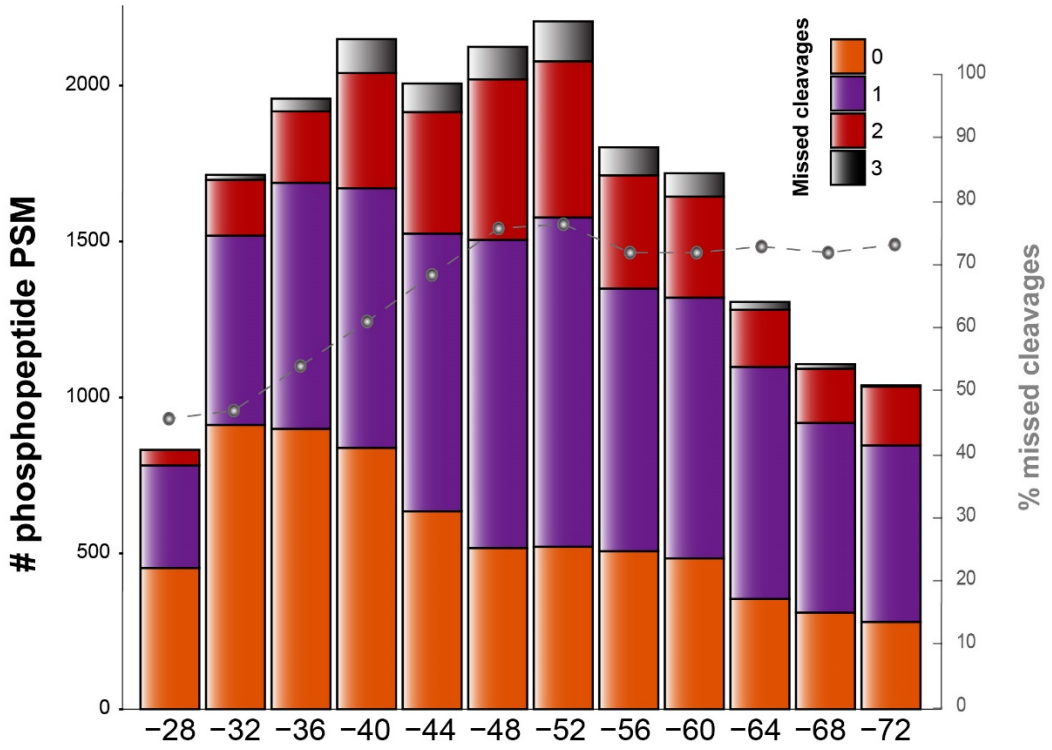
Supplementary Figure 3-5: *FAIMS improves the SILAC quantitation of the lower abundance peptides. Density map of the log_2 fold change versus log_{10} Base Peak Fraction, where the red dotted line depicts the expected fold change of 2.5: 5.4: 13.5.*



Supplementary Figure 3-6: *Distribution of observed fold change and ratio compression for LC-MS/MS experiments performed with and without FAIMS. The distribution is shown for experiment performed with (bottom) and without (top) SCX fractionation. Dotted black lines shows the boundary for +/- 3 SDs. The standard deviation correspond to 0.44 for all experiments and represents the lowest standard deviation obtained for the FAIMS 2D-LC experiment at ratio M:L 1:2.5. The proportion of peptides ions that fall within these boundaries are shown above each plot for FAIMS and non-FAIMS experiments.*



Supplementary Figure 3-7: *Precision of phosphoproteomic measurements for LC-MS/MS performed with and without FAIMS. Scatterplot comparing the relative standard deviation for phosphopeptides common to both FAIMS and non-FAIMS experiments, taken from data shown in Figure 3-4c for 20 min heat shock. Histograms for all phosphopeptides on the x and y axes are displayed for non-FAIMS and FAIMS experiments, respectively.*



Supplementary Figure 3-8: *Distribution of the number of PSM identified and the relative number of missed cleavages for the different CV transmission ranges. Barplot of PSM and CV shows that most of the PSM are in the central region of the CV scan range and fewer PSM are at the distal regions, suggesting that the selected CV range was properly selected. Dotted line shows the percentage of missed cleavages.*

3.7.2. Supplementary tables

Supplementary Table 3-1*: List of quantified SILAC tryptic peptides (CD-ROM)*

Supplementary Table 3-2*: List of Phospho substrates and modification sites (CD-ROM).*

3.8. References

1. Kanshin, E., P. Kubiniok, Y. Thattikota, D. D'Amours, and P. Thibault, *Phosphoproteome dynamics of Saccharomyces cerevisiae under heat shock and cold stress.* Mol Syst Biol, 2015. **11**(6): p. 813.
2. Kubiniok, P., H. Lavoie, M. Therrien, and P. Thibault, *Time-resolved phosphoproteome analysis of paradoxical RAF activation reveals novel targets of ERK.* Mol Cell Proteomics, 2017. **16**(4): p. 663-679.
3. Ankney, J.A., A. Muneer, and X. Chen, *Relative and Absolute Quantitation in Mass Spectrometry–Based Proteomics.* Annual Review of Analytical Chemistry, 2018. **11**(1): p. 49-77.
4. Hogrebe, A., L. von Stechow, D.B. Bekker-Jensen, B.T. Weinert, C.D. Kelstrup, and J.V. Olsen, *Benchmarking common quantification strategies for large-scale phosphoproteomics.* Nat Commun, 2018. **9**(1): p. 1045.
5. Dayon, L., A. Hainard, V. Licker, N. Turck, K. Kuhn, D.F. Hochstrasser, P.R. Burkhard, and J.-C. Sanchez, *Relative Quantification of Proteins in Human Cerebrospinal Fluids by MS/MS Using 6-Plex Isobaric Tags.* Analytical Chemistry, 2008. **80**(8): p. 2921-2931.
6. Beck, F., J.M. Burkhart, J. Geiger, R.P. Zahedi, and A. Sickmann, *Robust workflow for iTRAQ-based peptide and protein quantification.* Methods Mol Biol, 2012. **893**: p. 101-13.
7. Rauniyar, N. and J.R. Yates, *Isobaric Labeling-Based Relative Quantification in Shotgun Proteomics.* Journal of Proteome Research, 2014. **13**(12): p. 5293-5309.
8. Chen, X., S. Wei, Y. Ji, X. Guo, and F. Yang, *Quantitative proteomics using SILAC: Principles, applications, and developments.* Proteomics, 2015. **15**(18): p. 3175-92.
9. Wasinger, V.C., M. Zeng, and Y. Yau, *Current Status and Advances in Quantitative Proteomic Mass Spectrometry.* International Journal of Proteomics, 2013. **2013**: p. 12.
10. Mann, M., *Fifteen years of Stable Isotope Labeling by Amino Acids in Cell Culture (SILAC).* Methods Mol Biol, 2014. **1188**: p. 1-7.
11. Ong, S.E., B. Blagoev, I. Kratchmarova, D.B. Kristensen, H. Steen, A. Pandey, and M. Mann, *Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics.* Mol Cell Proteomics, 2002. **1**(5): p. 376-86.
12. Terzi, F. and S. Cambridge, *Chapter Three - An Overview of Advanced SILAC-Labeling Strategies for Quantitative Proteomics*, in *Methods in Enzymology*, K.S. Arun, Editor. 2017, Academic Press. p. 29-47.
13. Hoedt, E., G. Zhang, and T.A. Neubert, *Stable isotope labeling by amino acids in cell culture (SILAC) for quantitative proteomics.* Adv Exp Med Biol, 2014. **806**: p. 93-106.
14. Betancourt, L.H., P.-J. De Bock, A. Staes, E. Timmerman, Y. Perez-Riverol, A. Sanchez, V. Besada, L.J. Gonzalez, J. Vandekerckhove, and K. Gevaert, *SCX charge state selective separation of tryptic peptides combined with 2D-RP-HPLC allows for detailed proteome mapping.* Journal of Proteomics, 2013. **91**: p. 164-171.
15. Yang, F., Y. Shen, D.G. Camp, and R.D. Smith, *High pH reversed-phase chromatography with fraction concatenation as an alternative to strong-cation exchange chromatography for two-dimensional proteomic analysis.* Expert Review of Proteomics, 2012. **9**(2): p. 129-134.
16. Bonneil, E., S. Pfammatter, and P. Thibault, *Enhancement of mass spectrometry performance for proteomic analyses using high-field asymmetric waveform ion mobility spectrometry (FAIMS).* J Mass Spectrom, 2015. **50**(11): p. 1181-95.

17. Schneider, B.B., E.G. Nazarov, F. Londry, P. Vouros, and T.R. Covey, *Differential mobility spectrometry/mass spectrometry history, theory, design optimization, simulations, and applications.* Mass Spectrom Rev, 2016. **35**(6): p. 687-737.

18. Sarsby, J., R.L. Griffiths, A.M. Race, J. Bunch, E.C. Randall, A.J. Creese, and H.J. Cooper, *Liquid Extraction Surface Analysis Mass Spectrometry Coupled with Field Asymmetric Waveform Ion Mobility Spectrometry for Analysis of Intact Proteins from Biological Substrates.* Anal Chem, 2015. **87**(13): p. 6794-800.

19. Purves, R.W., D.A. Barnett, B. Ells, and R. Guevremont, *Elongated conformers of charge states +11 to +15 of bovine ubiquitin studied using ESI-FAIMS-MS.* Journal of the American Society for Mass Spectrometry, 2001. **12**(8): p. 894-901.

20. Cooper, H.J., *To What Extent is FAIMS Beneficial in the Analysis of Proteins?* J Am Soc Mass Spectrom, 2016. **27**(4): p. 566-77.

21. Barnett, D.A., B. Ells, R. Guevremont, and R.W. Purves, *Application of ESI-FAIMS-MS to the analysis of tryptic peptides.* Journal of the American Society for Mass Spectrometry, 2002. **13**(11): p. 1282-1291.

22. Saba, J., E. Bonneil, C. Pomies, K. Eng, and P. Thibault, *Enhanced sensitivity in proteomics experiments using FAIMS coupled with a hybrid linear ion trap/Orbitrap mass spectrometer.* J Proteome Res, 2009. **8**(7): p. 3355-66.

23. Swearingen, K.E., M.R. Hoopmann, R.S. Johnson, R.A. Saleem, J.D. Aitchison, and R.L. Moritz, *Nanospray FAIMS fractionation provides significant increases in proteome coverage of unfractionated complex protein digests.* Mol Cell Proteomics, 2012. **11**(4): p. M111.014985.

24. Venne, K., E. Bonneil, K. Eng, and P. Thibault, *Improvement in peptide detection for proteomics analyses using NanoLC-MS and high-field asymmetry waveform ion mobility mass spectrometry.* Anal Chem, 2005. **77**(7): p. 2176-86.

25. Baird, M.A. and A.A. Shvartsburg, *Localization of Post-Translational Modifications in Peptide Mixtures via High-Resolution Differential Ion Mobility Separations Followed by Electron Transfer Dissociation.* J Am Soc Mass Spectrom, 2016. **27**(12): p. 2064-2070.

26. Bridon, G., E. Bonneil, T. Muratore-Schroeder, O. Caron-Lizotte, and P. Thibault, *Improvement of phosphoproteome analyses using FAIMS and decision tree fragmentation. application to the insulin signaling pathway in Drosophila melanogaster S2 cells.* J Proteome Res, 2012. **11**(2): p. 927-40.

27. Pfammatter, S., E. Bonneil, F.P. McManus, and P. Thibault, *Gas-Phase Enrichment of Multiply Charged Peptide Ions by Differential Ion Mobility Extend the Comprehensiveness of SUMO Proteome Analyses.* Journal of The American Society for Mass Spectrometry, 2018. **29**(6): p. 1111-1124.

28. Shliaha, P.V., M.A. Baird, M.M. Nielsen, V. Gorshkov, A.P. Bowman, J.L. Kaszycki, O.N. Jensen, and A.A. Shvartsburg, *Characterization of Complete Histone Tail Proteoforms Using Differential Ion Mobility Spectrometry.* Anal Chem, 2017. **89**(10): p. 5461-5466.

29. Guevremont, R., *High-field asymmetric waveform ion mobility spectrometry: A new tool for mass spectrometry.* Journal of Chromatography A, 2004. **1058**(1): p. 3-19.

30. Prasad, S., M.W. Belford, J.-J. Dunyach, and R.W. Purves, *On an Aerodynamic Mechanism to Enhance Ion Transmission and Sensitivity of FAIMS for Nano-Electrospray Ionization-Mass Spectrometry.* Journal of The American Society for Mass Spectrometry, 2014. **25**(12): p. 2143-2153.

31. Perez-Riverol, Y., A. Csordas, J. Bai, M. Bernal-Llinares, S. Hewapathirana, D.J. Kundu, A. Inuganti, J. Griss, G. Mayer, M. Eisenacher, et al., *The PRIDE database and related tools and resources in 2019: improving support for quantification data.* Nucleic Acids Res, 2019. **47**(D1): p. D442-D450.

32. Bonneil, E., S. Pfammatter, and P. Thibault, *Enhancement of mass spectrometry performance for proteomic analyses using high-field asymmetric waveform ion mobility spectrometry (FAIMS).* Journal of Mass Spectrometry, 2015. **50**(11): p. 1181-1195.

33. Creese, A.J., N.J. Shimwell, K.P. Larkins, J.K. Heath, and H.J. Cooper, *Probing the complementarity of FAIMS and strong cation exchange chromatography in shotgun proteomics.* J Am Soc Mass Spectrom, 2013. **24**(3): p. 431-43.

34. Zhao, H., D.L. Cunningham, A.J. Creese, J.K. Heath, and H.J. Cooper, *FAIMS and Phosphoproteomics of Fibroblast Growth Factor Signaling: Enhanced Identification of Multiply Phosphorylated Peptides.* Journal of Proteome Research, 2015. **14**(12): p. 5077-5087.

35. Solari, F.A., M. Dell'Aica, A. Sickmann, and R.P. Zahedi, *Why phosphoproteomics is still a challenge.* Molecular BioSystems, 2015. **11**(6): p. 1487-1493.

36. Kuhl, N.M. and L. Rensing, *Heat shock effects on cell cycle progression.* Cell Mol Life Sci, 2000. **57**(3): p. 450-63.

37. Duncan, R.F. and J.W. Hershey, *Protein synthesis and protein phosphorylation during heat stress, recovery, and adaptation.* J Cell Biol, 1989. **109**(4 Pt 1): p. 1467-81.

38. Keaton, M.A., E.S. Bardes, A.R. Marquitz, C.D. Freel, T.R. Zyla, J. Rudolph, and D.J. Lew, *Differential susceptibility of yeast S and M phase CDK complexes to inhibitory tyrosine phosphorylation.* Curr Biol, 2007. **17**(14): p. 1181-9.

39. Acosta-Jaquez, H.A., J.A. Keller, K.G. Foster, B. Ekim, G.A. Soliman, E.P. Feener, B.A. Ballif, and D.C. Fingar, *Site-specific mTOR phosphorylation promotes mTORC1-mediated signaling and cell growth.* Mol Cell Biol, 2009. **29**(15): p. 4308-24.

40. Gingras, A.C., S.P. Gygi, B. Raught, R.D. Polakiewicz, R.T. Abraham, M.F. Hoekstra, R. Aebersold, and N. Sonenberg, *Regulation of 4E-BP1 phosphorylation: a novel two-step mechanism.* Genes Dev, 1999. **13**(11): p. 1422-37.

41. Blencowe, B.J., R. Issner, J.A. Nickerson, and P.A. Sharp, *A coactivator of pre-mRNA splicing.* Genes Dev, 1998. **12**(7): p. 996-1009.

42. Bennetzen, M.V., D.H. Larsen, J. Bunkenborg, J. Bartek, J. Lukas, and J.S. Andersen, *Site-specific phosphorylation dynamics of the nuclear proteome during the DNA damage response.* Mol Cell Proteomics, 2010. **9**(6): p. 1314-23.

43. Wojcechowskyj, J.A., C.A. Didigu, J.Y. Lee, N.F. Parrish, R. Sinha, B.H. Hahn, F.D. Bushman, S.T. Jensen, S.H. Seeholzer, and R.W. Doms, *Quantitative phosphoproteomics reveals extensive cellular reprogramming during HIV-1 entry.* Cell Host Microbe, 2013. **13**(5): p. 613-623.

44. Rossignol, M., A. Keriel, A. Staub, and J.M. Egly, *Kinase activity and phosphorylation of the largest subunit of TFIIF transcription factor.* J Biol Chem, 1999. **274**(32): p. 22387-92.

45. Dobrikov, M.I., M. Shveygert, M.C. Brown, and M. Gromeier, *Mitotic phosphorylation of eukaryotic initiation factor 4G1 (eIF4G1) at Ser1232 by Cdk1:cyclin B inhibits eIF4A helicase complex binding with RNA.* Mol Cell Biol, 2014. **34**(3): p. 439-51.

46. Kim, T.S., H.D. Kim, and J. Kim, *PKCdelta-dependent functional switch of rpS3 between translation and DNA repair.* Biochim Biophys Acta, 2009. **1793**(2): p. 395-405.

47. Forsburg, S.L., *Eukaryotic MCM proteins: beyond replication initiation.* Microbiol Mol Biol Rev, 2004. **68**(1): p. 109-31.

48. Cortez, D., G. Glick, and S.J. Elledge, *Minichromosome maintenance proteins are direct targets of the ATM and ATR checkpoint kinases.* Proc Natl Acad Sci U S A, 2004. **101**(27): p. 10078-83.

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

CHAPTER FOUR

4. Gas-phase enrichment of multiply charged peptide ions by differential ion mobility extend the comprehensiveness of SUMO proteome analyses

Sibylle Pfammatter[1,2], Eric Bonneil[1], Francis P. McManus[1], and Pierre Thibault[1,2]*

[1]Institute for Research in Immunology and Cancer, Université de Montréal, C.P. 6128, Succursale centre-ville, Montréal, Québec, H3C 3J7, Canada.

[2]Department of Chemistry, Université de Montréal, C.P. 6128, Succursale centre-ville, Montréal, Québec, H3C 3J7, Canada.

Published:

Journal of the American Society for Mass Spectrometry. 2018 Jun;29(6):1111-1124

Reprinted with permission from JOURNAL OF THE AMERICAN SOCIETY FOR MASS SPECTROMETRY.

Copyright © (2019) Springer Science+Business Media

Author contributions:

S.P. and E.B. contributed equally to this work.

Sibylle Pfammatter designed, executed and analyzed all experiments, generated all figures and wrote the paper. Eric Bonneil (MS platform manager) helped to define the project, supported the data acquisition and analysis and contributed to the drafting of the manuscript. Francis P. McManus (lab manager) contributed to interpreting the data of the biological processes and the preparation of the manuscript. Pierre Thibault (supervisor) managed the project. All authors approved the content and submission of the paper.

4.1. Abstract

The small ubiquitin like modifier (SUMO) is a member of the family of ubiquitin-like modifiers (UBLs) and is involved in important cellular processes, including DNA damage response, meiosis and cellular trafficking. The large-scale identification of SUMO peptides in a site-specific manner is challenging not only because of the low abundance and dynamic nature of this modification, but also due to the branched structure of the corresponding peptides that further complicate their identification using conventional search engines. Here, we exploited the unusual structure of SUMO peptides to facilitate their separation by high field asymmetric waveform ion mobility spectrometry (FAIMS) and increase the coverage of SUMO proteome analysis. Upon trypsin digestion, branched peptides contain a SUMO remnant side chain and predominantly form triply-protonated ions that facilitate their gas-phase separation using FAIMS. We evaluated the mobility characteristics of synthetic SUMO peptides and further demonstrated the application of FAIMS to profile the changes in protein SUMOylation of HEK293 cells following heat shock, a condition known to affect this modification. FAIMS typically provided a 10-fold improvement of detection limit of SUMO peptides, and enabled a 36% increase in SUMO proteome coverage compared to the same LC-MS/MS analyses performed without FAIMS.

4.2. Introduction

The small ubiquitin-like modifier (SUMO) is a member of the larger family of ubiquitin-like modifiers (UBLs) that share structural and evolutionary relationships with ubiquitin [1, 2]. Protein SUMOylation is involved in numerous cellular pathways in both the nucleus and the cytoplasm including DNA replication, DNA damage response, cell division, repression of transcription, nuclear trafficking and receptor internalization at the plasma membrane [3-5]. This modification is conjugated to its substrate lysine amino group via an enzyme cascade that resembles ubiquitylation and comprises E1 activating, E2 conjugating and E3 ligating enzymes [6, 7]. SUMOylated lysine residues were first reported to lie in the consensus motif ψKxE/D where ψ is a large hydrophobic residue, K is the modified lysine, and x is any residue preceding the glutamate (E) or aspartate (D) residues. However, other consensus sequences were reported such as a phospho-dependent sequence, reverse consensus and non-consensus regions [7]. SUMO is also one of few UBLs that exist as multiple paralogs in human cells and SUMO1-3 are ubiquitously expressed in all tissues. While SUMO2 and SUMO3 shared 97% sequence identity, they are only ~50% identical to SUMO1 [8]. SUMO-2/3 have the ability to form polySUMO chains by covalent binding via the lysine residue at the N terminus consensus motif KxE/D [9]. On the other hand, SUMO-1 is unable to polymerize readily since it lacks the consensus site and is thought to act as a polySUMO chain terminator.

The identification of SUMOylation sites by mass spectrometry (MS) is challenging due to the relatively low abundance of SUMOylated proteins, the highly dynamic nature of the modification, and the relatively long SUMO remnant appended on lysine residues following tryptic digestion (e.g. up to 32 amino acids for SUMO2/3) that further complicate the interpretation of MS/MS spectra. To circumvent those issues, different approaches that used affinity purification of SUMO mutant proteins were proposed. Our group generated a functional 6xHis-SUMO3-Q87R/Q88N mutant [10] that enables the enrichment of SUMOylated proteins on a nickel-nitrilotriacetic acid (Ni-NTA) column. A five amino acid SUMO remnant is left on the substrate upon trypsin digestion, and the corresponding peptides can be immunoprecipitated using an antibody

that recognizes the NQTGG epitope [11, 12]. Similar enrichment strategies involving the SUMO3 T90K mutant and the anti-diglycine antibody [13], or a His10-Lys-deficient SUMO mutant with Ni-NTA enrichment [14] have also enabled the identification of thousands of SUMO sites. More recently, a multi-step approach combining free lysine acetylation followed by SENP2-cleavage of the fully acetylated SUMO2 enabled the identification of 751 putative wild-type SUMO-2 conjugation sites [15].

The presence of a side chain on SUMOylated peptides confer unusual structural features compared to other types of modifications such as acetylation, methylation, or phosphorylation. Indeed, the corresponding branched peptides typically expose a side-chain amino group that provides an additional protonation site, and SUMOylated peptides have a propensity to form abundant triply-protonated ions in LC-MS/MS experiments [16]. Furthermore, the presence of the side chain sequence also contributes to changes in the ion mobility of the peptide ion compared to its linear counterpart. These structural features were advantageously exploited in traveling wave ion mobility mass spectrometry to separate SUMO peptides from the typically smaller and lesser charged linear peptides [17].

Other forms of ion mobility such as high field asymmetric waveform ion mobility spectrometry (FAIMS) can be used to facilitate the identification of SUMOylated peptides in large-scale SUMO proteome analyses. In the context of proteomic analyses, FAIMS provides additional advantages as it can separate and accumulate multiply-charged ions from co-eluting ions and background contaminants thereby improving MS sensitivity [18-20]. In FAIMS, ions are entrained by a carrier gas between two electrodes to which is applied a high voltage asymmetric waveform. Ions are separated in the gas-phase based on their difference in mobility at low and high electric fields [21]. Ion selection is achieved by applying a compensation voltage (CV) that is superimposed to the waveform and enables the transmission of specific ions at their corresponding CV values. Previous reports have successfully combined FAIMS in LC-MS/MS analyses to expand the depth and comprehensiveness of proteomic analyses [22-25]. FAIMS also provides additional advantages in proteomics by separating phosphopeptides and peptide isomers [26-28], and by reducing the extent of precursor co-fragmentation to improve quantitative measurements of multiplex proteomics [29].

In the present study, we examined the mobility characteristics of SUMOylated peptides separated by FAIMS, and evaluated the analytic merits of this gas phase fractionation method when combined to LC-MS/MS. We further demonstrate the application of LC-FAIMS-MS/MS to expand the coverage and the dynamic range of the SUMO proteome of HEK293 cells exposed to heat shock.

4.3. Experimental Section

Cell culture and Heat Shock. HEK293 wild type cells and HEK293 cell line stably expressing 6xHis-SUMO-3-Q87R-Q88N (SUMO3m) [11] were cultured at 37 °C in a 5% CO_2 constant atmosphere. Dulbecco's Modified Eagles Medium (high glucose, GE Healthcare HyClone) was supplemented with 10% fetal bovine serum (Seradigm VWR Life Science), 1% Penicillin/Streptomycin Solution (Gibcon) and 1% L-Glutamine (Gibcon). For HEK293 SUMO3m, 0.5 mg/ml Neomycine (Gibcon) was added during cell culture. Cells were collected and washed twice with PBS (GE Healthcare HyClone) before mechanical lyses (2 x 10s sonicate pulses) in 50 mM Tris in 8M Urea (Bio Basic). Protein concentration was determined by Bradford assay (Bio Rad).

For heat stress experiments, HEK293 SUMO3m cells were grown in 15 cm petri plates to near confluency (0.7-0.8), the media was removed and replaced by media prewarmed at 43°C without Neomycine. Cell dishes were placed in incubators for 60 min at 43°C and 5% CO_2. Control and heat shock-treated cells were collected simultaneously and washed twice with 37°C PBS. Cell pellets were frozen in liquid nitrogen and stored at -30°C until further processing.

Synthetic peptides. SUMOylated synthetic peptides were obtained from JPT Peptide Technologies GmbH (Berlin, Germany). Synthesis details were reported previous by our group [16]. For linearity experiment, each synthetic peptide was spiked at concentrations of 0.02 to 1 pmol in 760 ng of HEK293 protein digests. For HEK293 digests, 1 mg of proteins was resuspended in 1 mL of 50 mM ammonium bicarbonate (Sigma-Aldrich) and reduced for 30 min at 37°C with 5 mM Tris (2-Carboxyethyl) phosphine Hydrochloride (TCEP) (ThermoFisher Scientific). Alkylation reactions were performed for 30 min at room temperature with 10 mM 2-chloroacetamide (Sigma-Aldrich). Proteins were digested overnight at 37°C with trypsin at a 1:50 trypsin:protein (wt/wt) (Promega, Madison, WI). Samples were dried down in a speed-vacuum prior to their reconstitution in 0.2% formic acid at a final concentration of 1 mg/mL.

SUMO peptide enrichment. SUMOylated peptides from control or heat shock treated cells were enriched according to the protocol described previously [12]. Briefly, cell pellets were lysed in 5ml buffer A (6 M Guanine, 0.1 M NaH_2PO_4, 0.01 M Tris-HCl

pH 8, 0.01 M Imidazole, 0.02 M 2-chloroacetamide, 0.01 M β-mercaptoethanol) followed by three short sonication pulses of 5 seconds to shear the DNA. Protein concentrations were determined by Bradford assay. SUMOylated proteins were purified by nickel affinity chromatography through the 6xHis tag on SUMO3m. 70 mg of protein from each condition (control or heath shock treated cells) were immobilized on Ni-NTA Agarose (Qiagen) beads. For each sample, 1.5 mL of equilibrated Ni-NTA agarose beads (3 mL of Ni-NTA slurry that was previously washed four time with 10 mL buffer A) were incubated overnight with 70 mg of total cell lysate at 4°C. The SUMO bound beads were washed once with 10 mL buffer A, five times with 10 mL buffer B (8 M Urea, 0.1 M NaH_2PO_4, 0.01 M Tris-HCl pH 6.3, 0.01 M Imidazole, 0.01 M β-mercaptoethanol), twice with 50 mM Ammonium bicarbonate and finally resuspended in 3.5 mL 50 mM Ammonium bicarbonate for quantification by Bradford assay. Proteins were digested on beads overnight at 37°C with trypsin (Promega) at a 1:50 trypsin:protein (wt/wt) ratio. Peptides were acidified by adding TFA (Sigma-Aldrich) to final concentration of 1% and desalted on Oasis HLB columns (Waters). Peptides were dried down using a speed-vacuum.

SUMOylated peptides were immuno purified (IP) from the Ni-NTA enriched material using the anti-K(NQTGG) monoclonal antibody that was crosslinked to protein A magnetic beads [30]. Briefly, the peptides were resuspended in 1 mL of PBS and incubated with the anti-K(NQTGG) bound beads for 1 h at 4 °C at a 1:200 antibody:protein starting material (wt:wt). The immuno complexes were washed three times with 1 mL of PBS, twice with 1 mL of 0.1 x PBS and once with 1 mL of water. The SUMOylated peptides were eluted from the beads with 500 µL 0.2 % formic acid and dried down in a speed vacuum. The peptides were reconstituted in 4 % formic acid in water for MS analysis (6 mg of starting material per injection).

Immunoblot analysis. Total cell extracts (10 µg) were prepared in Laemmli buffer (0.06 M Tris-Cl pH 6.8, 2% SDS, 10% glycerol, 5% β-mercaptoethanol, 0.01% bromophenol blue) and resolved on a precast gel (4-12% Bis-Tris) (Criterion XT). Proteins were transferred from the gel onto a nitrocellulose membrane at 100 mA for 16 h. The membrane was blocked with 5% non-fat milk and incubated with the primary antibody (SUMO2/3 or Histone H3, cell Signaling). Bands were visualized using the

ECL chemiluminescence detection methodology while using an anti-rabbit horseradish peroxidase-conjugated secondary antibody.

Mass spectrometry. LC-MS/MS analyses were performed on a nano-LC 2D pump (Eksigent, Dublin, CA) interfaced to a LTQ-Orbitrap Elite hybrid mass spectrometer (Thermo Fisher Science, San Jose, CA). On-line peptide separation was performed on a Optiguard SCX trap column, 5 µm, 300Å, 0.5 ID × 23 mm (Optimize Technologies, Oregon City, OR) followed by a 360 µm ID × 4 mm, C_{18} trap column before separation on a in house packed 150 µm ID x 20 cm LC column (Jupiter C_{18}, 3 µm, 300 Å, Phenomenex, Torrance, CA). For two-dimensional chromatography, peptides were loaded with 0.2% formic acid to the SCX cartridge and eluted using increasing concentration of ammonium acetate (250, 500 , 750, 1000 and 2000 mM) at pH 3.5. Reversed phase separation was performed with a linear gradient of 5-40% acetonitrile (0.2% formic acid) at a flow rate of 600 nL/min. For synthetic peptides, HEK293 and HEK293 SUMO3m digests, the separation was achieved with 56 min or 106 min long gradient, respectively, before re-equilibrating the column for 14 min at 5% acetonitrile (0.2% formic acid). MS/MS scans were triggered using a top 12 method, MS scans were acquired in the Orbitrap at a resolution of 60,000 whereas the resolution for MS/MS spectra was at 15,000. Full MS were scanned between 300 -1200 *m/z* and the normalized collision energy for high-energy collision-induced dissociation fragmentation was set at 30. Automatic Gain Control for full MS and MS/MS were set at 10^6 and $5x10^4$, respectively with exclusion times of 45 s (without FAIMS) or 20 s (with FAIMS). The maximum injection time for the full MS was 1 s, whereas for MS/MS the maximum injection times were set to 300 ms.

FAIMS. For LC-FAIMS-MS/MS measurements, the Ion Max Source (Thermo Fisher Science, San Jose, CA) was replaced with a FAIMS interface (Thermo Fisher Science, San Jose, CA). Experiments were conducted on a FAIMS device with a curved well ion inlet [31] with two cylindrical electrodes with a gap width of 1.5 mm between inner and outer electrodes at 90°C and 100°C, respectively. Dispersion voltage (DV) was held at -5000 V (ED -3,333V/cm). The FAIMS interface used nitrogen as a carrier gas at a flow rate of 2.3 L/min. For electrospray ionization, a cylindrical stainless steel tube was connected to the analytical column outlet and mounted at a 60° angle. MS

parameters for LC-FAIMS-MS/MS measurements were similar to those described above, except that the Top 12 method (1 MS survey scan followed by a maximum of 12 tandem MS^2) was replaced with 3 CV step scan method (1 MS survey scan followed by 3 tandem MS^2 for each CV step). For synthetic peptides, the CV range from -34V to -85V (-227 V/cm to -567 V/cm) was covered, for large scale experiment where cells were exposed to heat stress, 4 injections covered the range from CV -40V to CV -73V (-267 V/cm to -487 V/cm) using 3V (20 V/cm) increments.

Data processing. Raw data acquired with Xcalibur software (version 2.2) were either processed with MaxQuant version 1.5.3.8 (http://www.maxquant.org) or Mascot Daemon (Matrix Science, version 2.4.0). SUMOylated peptides isolated by immunoprecipitations were searched using MaxQuant and the Uniprot Human database (release March 3, 2015). FAIMS Raw files were transformed with an inHouse pyton script to mzXML files containing only one CV per file. Search parameters were as follows: Trypsin/P as specific enzyme, Oxidation (M) – Phospho (STY) – Deamidation (NQ) – GlyGly (K) and SUMO3: NQTGG (K) as variable modification. The tolerance for MS and MS/MS were set to 25 ppm and 20 ppm for the first search and 6 ppm and 10 ppm for the main search. MS/MS identification were filtered with an FDR of 0.1%. For network analysis, gene names were searched against the whole human proteome using string version 10.5 (https://string-db.org). Experimentally determined interactors with high confidence (0.700) were visualized in Cytoscape (Version 3.5.1) and clustered with the MCODE application (degree Cutoff 2, Node Score Cutoff 0.1, K-Core 2 and Max.Depth 10). Gene ontology enrichment analysis of the extracted clusters was performed with BiNGO with a significance level of 0.05.

4.4. Results and Discussion

The strategy to identify SUMOylated proteins and their modified sites relies on HEK293 cells stably expressing a functional SUMO3 mutant that contains an N-terminal 6xHis tag and a C-terminus cleavage site that leaves a 5 amino-acid SUMO remnant (NQTGG) on the lysine acceptor site upon tryptic digestion [10, 11]. While this protocol provides enrichment level up to 50 %, the identification of SUMOylated peptides in a progressively more complex population of low abundance peptides still represent a sizable analytical problem. In this context, we surmised that the ability of FAIMS to separate ions based on their charge states and their differences in mobility at low and high electric fields may provide a unique advantage to facilitate the detection of low abundance SUMOylated peptides. We reasoned that the branched structure of SUMOylated peptides and the presence of an extra N-terminus group confer distinctive features that favors their separation from linear tryptic peptides (***Figure 4-1a***). Consequently, we evaluated the distribution of SUMOylated peptides across CV values to determine conditions under which target peptide ions can be transmitted for enhanced SUMO proteome coverage.

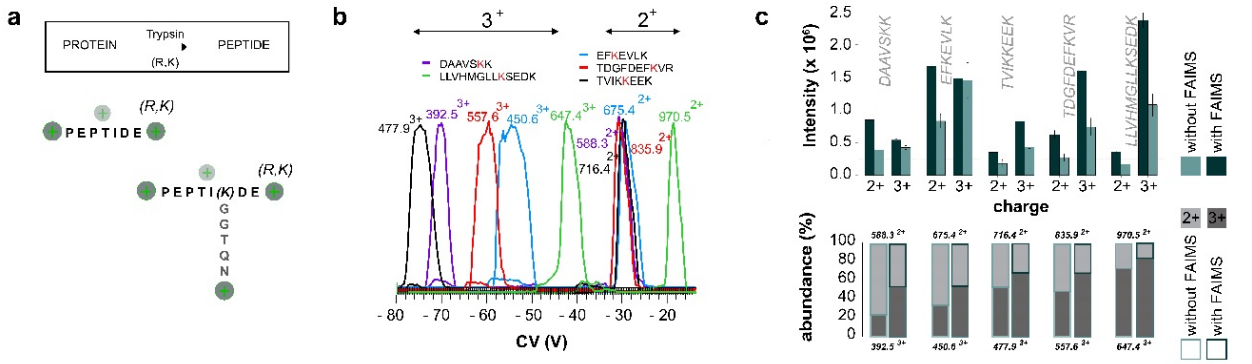


Figure 4-1: *Gas-phase separation of SUMOylated peptides using FAIMS. (a) Schematic representation of regular tryptic peptides and peptides modified by SUMO3 with remnant NQTGG. (b) Infusion of 5 SUMO peptides separated between two FAIMS electrodes with a gap of 1.5 mm. (c) Intensity and abundance distribution for the 5 SUMO peptides. Conditions: N_2 carrier gas 2.3 L/min, DV: -5000V.*

4.4.1. FAIMS improves the detection and identification of SUMOylated peptides present in complex tryptic digests

In preliminary experiments, we infused a mixture of 5 synthetic SUMOylated peptides at 100 nM with and without FAIMS. To increase both ion transmission and resolution, we used FAIMS electrodes with a narrow gap of 1.5 mm that provides higher resolution and lower gas flow turbulences compared to previous generation devices [31]. This FAIMS device provides a larger range of CV distribution extending over more than 60 V and uses nitrogen only as carrier gas [32]. A plot of the signal intensities versus CV values is shown in ***Figure 4-1b***, and highlights the ability of FAIMS to separate triply charged SUMOylated peptides from their doubly charged counterparts. As indicated, the triply charged ions are transmitted at higher CV values between -40 and -80 V whereas doubly charged ions were observed between -16 and -32 V. Also, the triply charged peptide ions were resolved from each other compared to doubly charged ions that were all transmitted at similar CVs with limited separation except for peptide LLVHMGLLKSEDK (bold type indicated the modified residue). A comparison of the intensity of the SUMOylated peptides analyzed with and without FAIMS, revealed that higher transmission was generally obtained when using gas-phase ion fractionation with FAIMS (***Figure 4-1c***). This observation might reflect the selective enrichment of target ions for a fix ion trapping capacity when using FAIMS. Interestingly, we noted that the relative proportion of doubly- and triply-charged peptide ions varied between experiments, and that the latter ions showed a higher transmission when using FAIMS. The transmission bias of doubly charged SUMOylated peptides observed without FAIMS might be explained by suppression of higher charge states that could take place when approaching the ion capacity defined by the automatic gain control (AGC).

To investigate the impact of sample complexity on the identification of multiply-charged SUMOylated peptide ions, we infused the synthetic SUMO peptide LLVHMGLLKSEDK at a constant concentration of 2 µM with increasing amounts of HEK293 tryptic digest (***Supplementary Figure 4-1***). Without any HEK293 digest, the quadruply-charged peptide ion was the most abundant charge state (57% relative abundance compared with 2+ ion at 5%). However the relative proportion of this ion was

progressively reduced with increasing contribution of HEK293 digest, and represented only 31% when spiked in 4 ug of HEK293 digest, whereas its doubly-charged ion increased to 33%. The discrimination of higher charge states with increasing sample complexity was also observed at different AGC values, but was more significant at AGC 10^4 compared to an AGC of 10^7 (***Supplementary Figure 4-2***). These results confirm that the distribution and intensity of multiply-charged SUMO peptide ions is affected by increasing sample complexity, a situation that may impede their successful identification when present at low abundance in cell digests.

To determine if suppression effects undermine the detection of SUMOylated peptides, we performed LC-MS/MS experiments with and without FAIMS on a mixture of 154 synthetic SUMO peptides and a tryptic digest of HEK293 cells. To ensure that we maximize the number of identification, we performed triplicate LC-MS/MS experiments for synthetic SUMO peptides and 12 replicates for the HEK293 digest. The same samples were analyzed by LC-FAIMS-MS/MS except that the number of CV acquired was scaled according to sample complexity (i.e., injection at each CV for HEK293 digest, and every 3V increments for synthetic SUMO peptides). For the HEK293 digest, LC-MS/MS performed with and without FAIMS enabled the identification of 14,442 and 6,103 unique peptides, with an overlap of 4,499 peptides. Out of these common peptides only 1,173 and 993 peptides were identified as doubly and triply-charged peptide ions in FAIMS and non FAIMS experiments, respectively. The LC-MS/MS analyses of synthetic SUMOylated peptides with and without FAIMS identified 118 and 144 peptides, of which 108 were common. However, only 15 and 25 of these common peptides displayed both doubly and triply charged ions with and without FAIMS, respectively. Next, we determined the intensity ratio of doubly and triply-charged ions for all identified peptides. The boxplot distribution of the corresponding ratios indicates that HEK293 tryptic peptides analyzed without FAIMS show an almost equal intensity of doubly and triply charged ions whereas the latter ions were more abundant when using FAIMS (***Figure 4-2a***). This intensity bias with and without FAIMS was also observed for SUMOylated peptides where the intensity of triply-charged ions was more intense than their doubly-charged counterparts, and suggest that the underrepresentation of multiply-charged ions is correlated with increasing sample complexity.

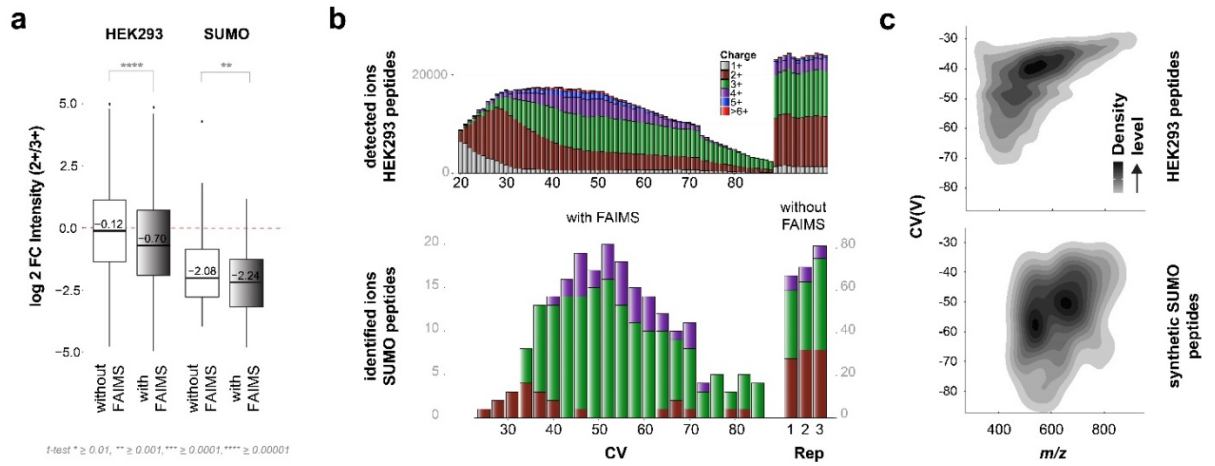


Figure 4-2: *Charge state distribution of tryptic peptides and SUMOylated peptides. (a) Distribution of intensity ratios between doubly- and triply protonated peptide ions from a HEK293 tryptic digest and SUMOylated peptides. (b) Bar chart representation of charge states vs CV values for tryptic peptides from HEK293 cells and SUMOylated peptides. (c) Density plot showing the distribution of m/z vs. CV for tryptic peptides from HEK293 cells and SUMOylated peptides.*

The ability of FAIMS to separate ion populations into different CV regions enables the selective transmission of triply charged ions with limited contribution of lower charge state ions. This is illustrated in ***Figure 4-2b*** that compares the distribution of multiply-charged ions identified with FAIMS according to their CV values. As indicated, a larger proportion of triply-charged peptides, including SUMOylated peptides, are transmitted between -37V and -70V, in a region where reduced contribution of doubly-charged ions is typically observed. While triply-charged peptides ions from HEK293 and synthetic SUMOylated peptides are transmitted across a similar range of CV, the branched structure and the charge distribution of the latter peptides confer notable differences in mobility that can be advantageously exploited to facilitate their separation from the bulk of tryptic peptides. ***Figure 4-2c*** shows the density plots of *m/z* distribution vs. CV values for HEK293 and synthetic SUMOylated peptides. As anticipated, SUMOylated peptides are clearly distinct from the larger population of tryptic peptides based on their distribution of *m/z* and CV values. A more detailed examination of the data also revealed that 3+ ions from SUMO peptides were transmitted at lower CV values than 3+ tryptic peptides over the *m/z* range examined. However, the difference in mean CV values at which SUMO and tryptic peptide 3+ ions

were transmitted gradually decreased with increasing *m/z* values (***Supplementary Figure 4-3***). These observations suggest that conformational changes imparted by branched SUMO peptides have more important effects on changes in ion mobility at high and low electric field than tryptic peptide ions of similar masses. Therefore, the selection of appropriate CV values could favor the transmission of SUMOylated peptides to enhance their identification in large-scale SUMO proteome analyses.

We next evaluated the linearity and limits of detection (LOD) for the analysis of a mixture of SUMOylated peptides spiked at levels of 20 fmol to 2 pmol in a tryptic digest of HEK293 proteins (760 ng) using LC-MS/MS with and without FAIMS (***Figure 4-3a***). Triplicate measurements were obtained for all experiments. For LC-FAIMS/MS/MS, 3 CV values were selected per injection to cover the range of-34V to -80V in 3V increments. In total we identified 107 synthetic SUMOylated peptides common to each experiment corresponding to 69.5% of the total number of spiked peptides, and the distribution of the linear regression coefficients of determination (r^2) are shown in ***Figure 4-3b***. Similar distribution of r^2 values were obtained for LC-MS/MS experiments performed with and without FAIMS. Examination of all linear regression analyses indicates that the CV stepping using in LC-FAIMS-MS/MS has no substantial effect on the linear regression (***Supplementary Figure 4-2***). However, analyses performed using FAIMS provided significant gains in sensitivity that were reflected with lower LOD values. An example of this is shown in ***Figure 4-3c*** for the triply-protonated peptide ion at *m/z* 576.65 corresponding to the synthetic SUMO peptide INEILSNALKR spiked at a level of 200 fmol. It is noteworthy that the LC-MS/MS analysis performed without FAIMS showed the co-elution of another peptide ion at *m/z* 577.32 which gave rise to a chimeric tandem mass spectrum, and prevented the non-ambiguous identification of the corresponding SUMOylated peptide. In contrast, the same analysis performed using FAIMS showed a distinct signal at *m/z* 576.65 with a clear improvement in signal to noise that translated into a ~10-fold improvement in LOD. A comparison of the LOD values obtained for all detected SUMOylated peptides is shown in ***Figure 4-3d***. Of the 107 SUMOylated peptides quantified using LC-MS/MS, 100 showed an improvement of LOD that ranged from 2 to 228 fold when using FAIMS. Marked improvements of sensitivity were noted for SUMO peptides of lower molecular mass that are observed in

more densely populated region of the mass spectrum where co-elution of abundant peptide ions with similar *m/z* values can affect peak detectability (***Figure 4-3c***).

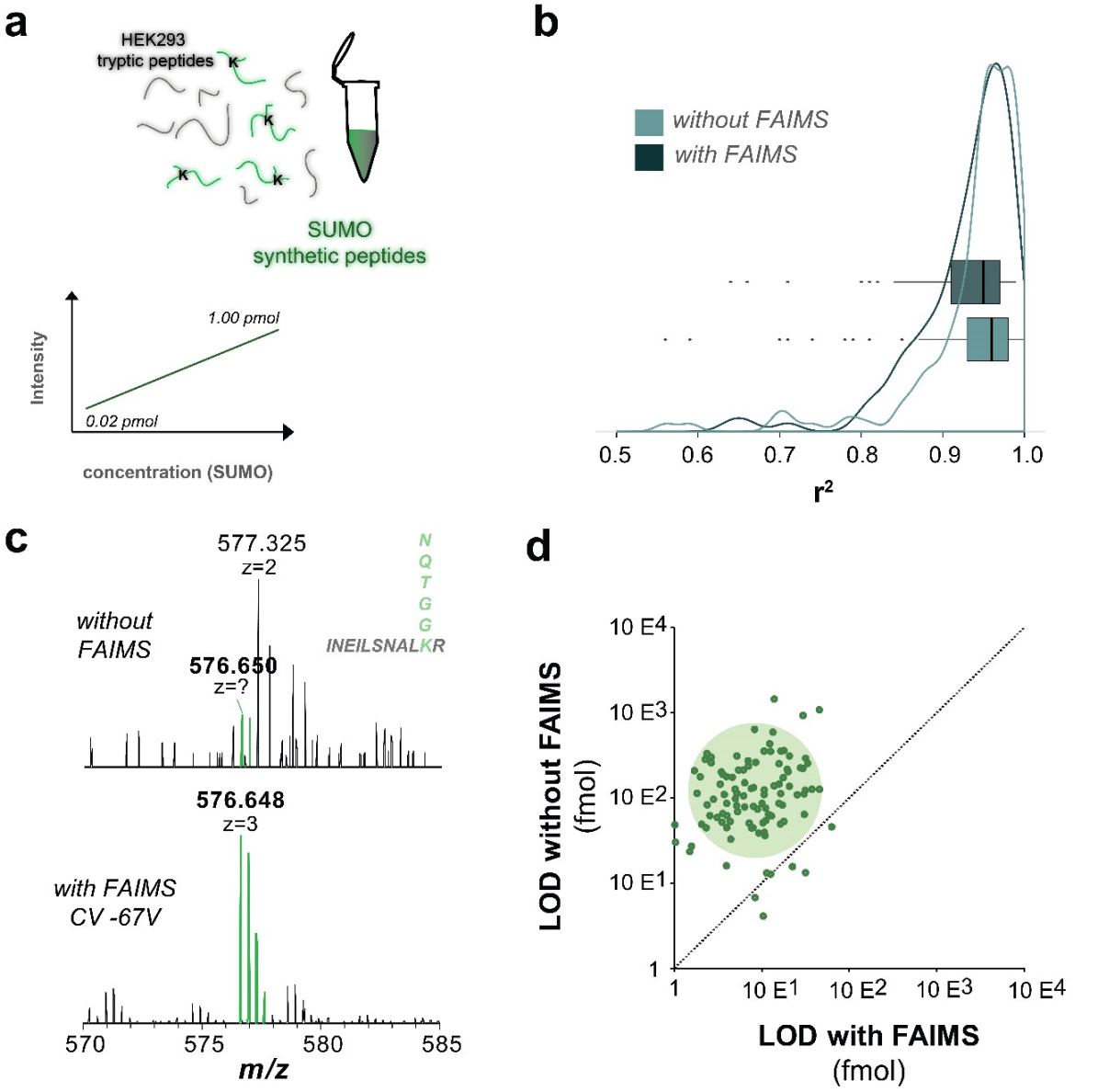


Figure 4-3: *LC-MS/MS analyses of SUMOylated peptides spiked in a HEK293 tryptic digest. (a) Synthetic SUMOylated peptides were spiked at concentration ranging from 0.02 to 1.00 pmol in 760 ng of a HEK293 tryptic digest. (b) Distribution of r^2 values for all identified SUMOylated peptides analyzed by LC-MS/MS with and without FAIMS. (c) Narrow section of the mass spectrum showing m/z 576.650^{3+} corresponding to the peptide INEILSNA**K**R with and without FAIMS. (d) Scatter plot comparing the limit of detection of SUMOylated peptides detected in both LC-MS/MS experiments.*

4.4.2. Large-scale profiling of SUMO proteome under heat shock

To evaluate the analytical merits of FAIMS for the identification of SUMOylated proteins in a site-specific manner, we profiled the changes in protein SUMOylation of HEK293 SUMO3m cells upon heat shock. Protein SUMOylation is an important modification regulating the cellular defense against hyperthermic cytotoxicity. For example, several transcription factors are known to be SUMOylated in response to heat shock, a modification that impedes their functions via several mechanisms including reduced nuclear entry, inhibition of DNA binding, and recruitment of transcriptional repressor [33]. Global changes in protein SUMOylation is rapidly observed upon heat shock treatment [34]. Accordingly, we performed label-free quantitative proteomic analyses on HEK293 SUMO3m cells subjected to 43 °C heat stress for one hour (***Figure 4-4a***). Increase in global protein SUMOylation upon heat shock was monitored by western blot, as depicted in ***Figure 4-4b***. We selected this incubation period based on the progressive increase in protein SUMOylation that reached a plateau after one hour (***Supplementary Figure 4-5***). Control (CTL) and heat shock treated HEK293 SUMO3m cells were harvested and lysed in denaturing buffer. SUMOylated proteins were then enriched on a Ni-NTA column prior to trypsin digestion (***Figure 4-4c***). SUMOylated tryptic peptides were purified by immunoaffinity using a custom antibody that recognizes the five amino acid SUMO3m remnant epitope [12]. The CTL and heat shock samples were then fractionated by on-line SCX and analyzed by LC-MS/MS or LC-FAIMS-MS/MS on the Orbitrap Elite mass spectrometer (***Figure 4-4c***). A total of 6 replicate injections were conducted on each SCX fractions analyzed by LC-MS/MS, while a single injection at different CV values was performed for individual SCX fraction analyzed by LC-FAIMS-MS/MS.

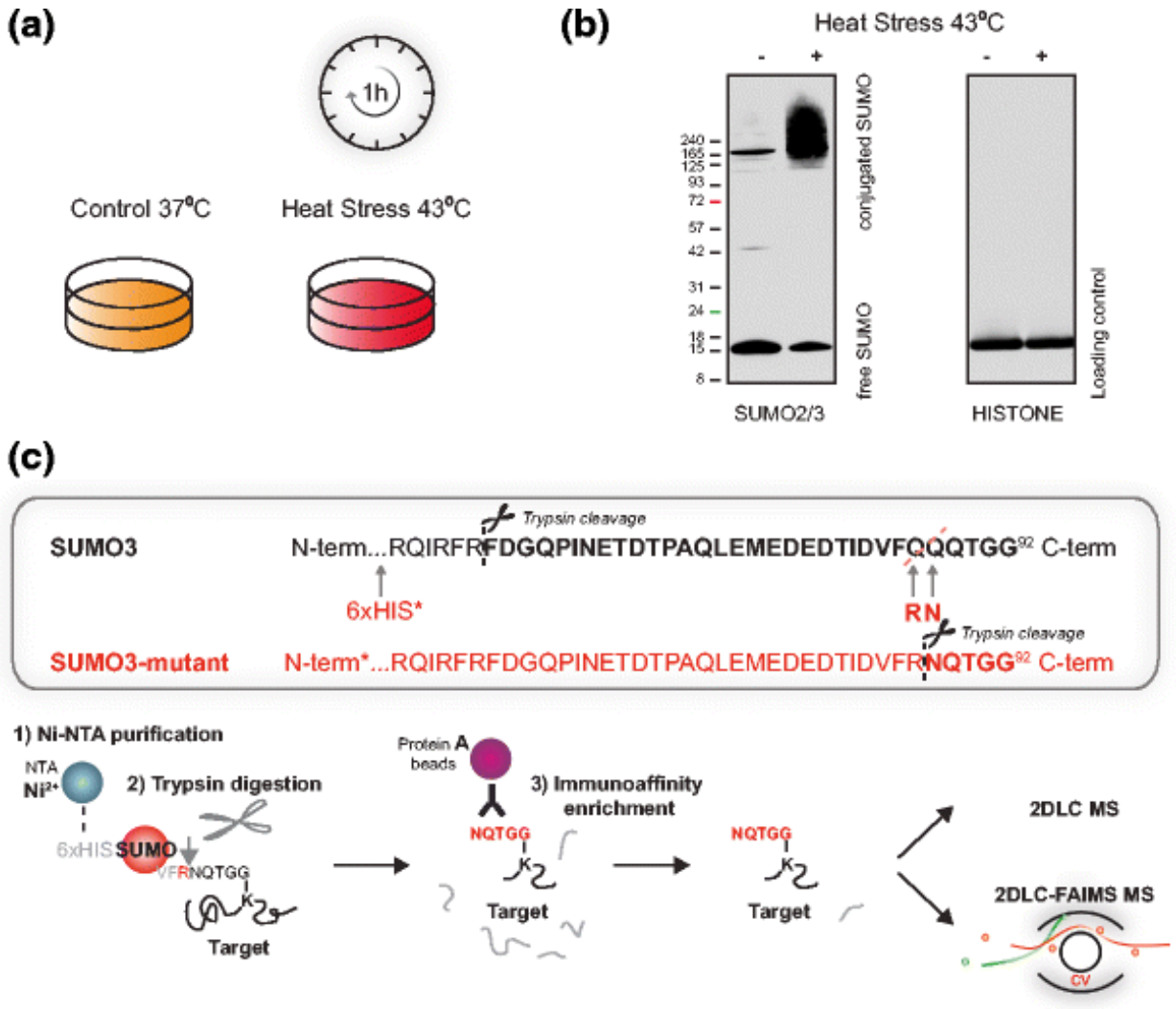


Figure 4-4: *SUMO proteome analysis of HEK293 cells following heat shock. (a) HEK293 cells overexpressing SUMO3m were cultured at 37°C. Control cells were kept at 37°C. For heat shock experiments the culture media was removed and replaced with preheated media and incubated for one hour at 43°C. (b) Western blot of control and heat shock-treated HEK293 cells using a SUMO2/3 specific antibody. The western blot of histone H3 is shown as a loading control. (c) Overview of the immunoaffinity purification and LC-MS/MS analyses of SUMOylated peptides. Total cell lysates are subjected to a NiNTA column to enrich SUMOylated proteins before tryptic digestion. Peptides containing the SUMO3m remnant are enriched using the anti-K-(NQTGG) antibody prior to their analyses using on-line SCX fractionation and LC-MS/MS with and without FAIMS.*

The large-scale SUMO proteome analyses of HEK293 SUMO3m cells identified a total of 3,658 unique SUMO sites on 1,433 protein substrates (***Figure 4-5a, Supplementary Table 4-1***). FAIMS enabled a 36 % increase in SUMO proteome

coverage compared to the conventional LC-MS/MS approach with 2,820 SUMO sites (1,254 SUMOylated proteins) and 2,078 SUMO sites (951 SUMOylated proteins) for FAIMS and non FAIMS experiments, respectively. These experiments also highlighted that the number of SUMO sites identified by both approaches represent only ~1/3 of all identified SUMO sites. In an effort to understand the reasons behind these differences, we compared the distribution of charge states for SUMOylated peptides identified in FAIMS and non FAIMS experiments. ***Supplementary Figure 4-6a*** indicates that SUMOylated peptides uniquely identified in either FAIMS or non FAIMS experiments were predominantly triply-charged ions. To determine if SUMOylated peptides unique to non FAIMS experiments were filtered out using FAIMS due to the CV steps selected, we used the “match between runs” function of Maxquant. This function enables the time and *m/z* alignment of peptide ions and the correlation of their intensities in a pairwise fashion for peptides identified in at least one of the two experiments. The correlation of detected peptides using “match between runs” indicated that 79.5 % of all SUMOylated peptides were present in both FAIMS and non FAIMS experiments (***Figure 4-5b***). These analyses also revealed that unique peptides originally identified without FAIMS were detected in FAIMS experiments, but that the converse was not necessarily true (***Supplementary Figure 4-7***). Indeed, SUMOylated peptides identified without FAIMS were detected in FAIMS experiments as a distinct isotopic cluster at an intensity level approaching the threshold for triggering MS/MS acquisition. However, SUMOylated peptides uniquely identified with FAIMS showed either interfering ions or were of too low intensity to be detected in non-FAIMS experiments. Interestingly, we only observed a marginal increase in the number of SUMOylated peptides identified beyond three replicate injection of the SCX fraction without FAIMS, whereas a progressive gain in new identification was observed using FAIMS across different CV steps (***Figure 4-5b*** and ***Supplementary Figure 4-6b***).

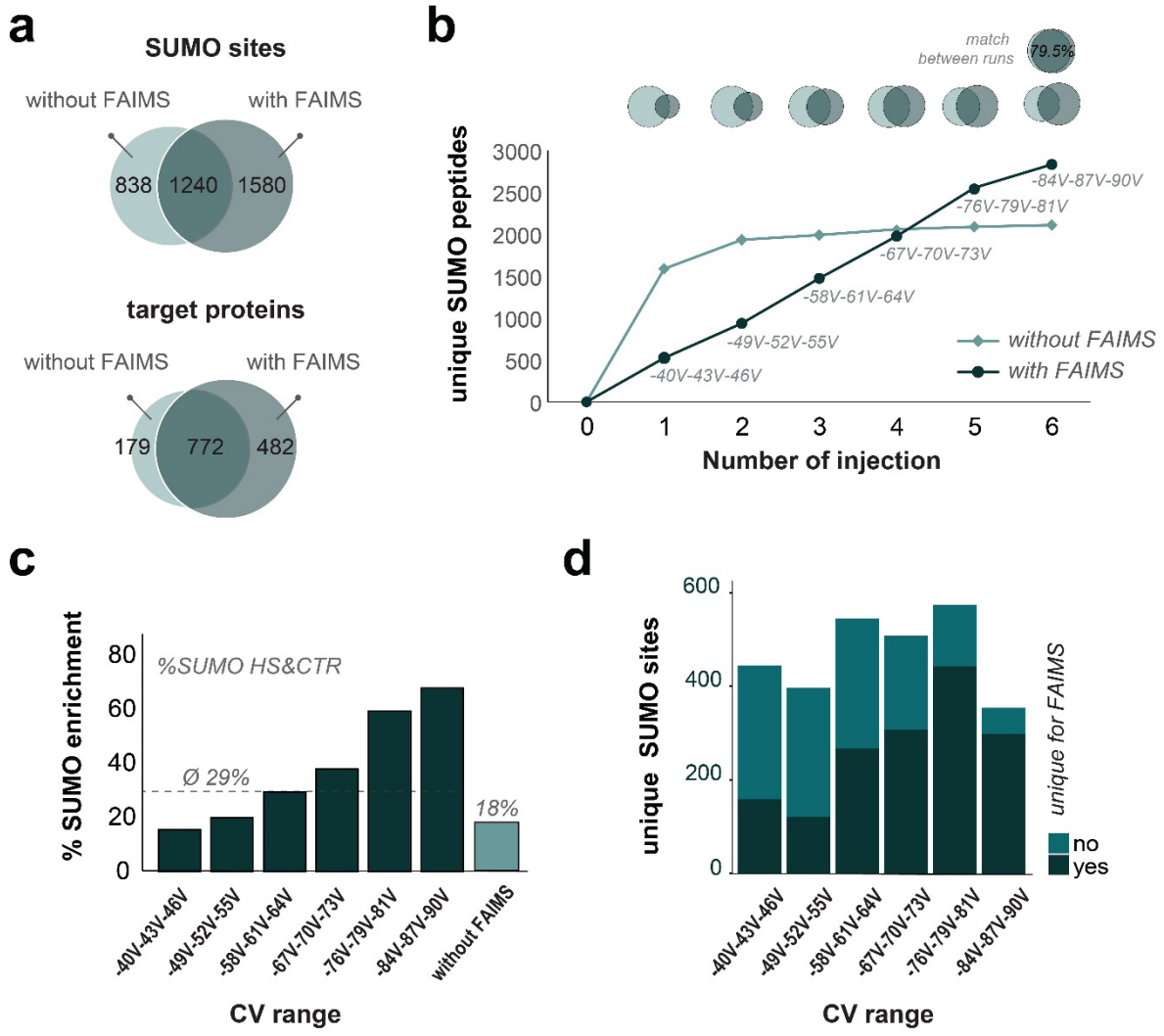


Figure 4-5: *SUMO proteome analyses of HEK293 cells following heat shock. (a) Venn diagram showing the distribution of SUMOylation sites and SUMOylated proteins identified with and without FAIMS. (b) Sum of unique SUMOylation sites identified after each replicate. (c) Enrichment of SUMOylated peptides across different ranges of CV values. (d) Distribution of unique SUMOylation sites as a function of CV values.*

The increased number of identified SUMOylated peptides using FAIMS is reminiscent of observations made during the large scale analysis of tryptic peptides [35], where sample complexity limits the progressive increase of new identification for replicate injections performed without FAIMS. In this context, gas phase fractionation using FAIMS provides sufficient resolution to prevent the resampling of peptides across CV steps. Importantly, FAIMS enables the enrichment of SUMOylated peptides by

favoring their selective transmission in a CV range that minimizes the contribution of interfering peptides ions (***Supplementary Figure 4-6e***).

Next, we compared the enrichment levels of SUMO peptides for FAIMS and non-FAIMS experiments (***Figure 4-5c***). These analyses indicated that the proportion of SUMOylated peptides in non FAIMS experiments represented only 18% of all peptides identified. In contrast, FAIMS provided enrichment levels up to 68% as observed for high CV values (-84V-87V-90V). Interestingly, the increase in enrichment level correlates well with the progressive increase in the number of unique SUMOylated peptides observed with FAIMS for higher CVs (***Figure 4-5d***).

The gas phase fractionation capability of FAIMS reduced sample complexity and enabled more comprehensive SUMO proteome analysis. To evaluate the extent of sample complexity, we determined the precursor intensity fraction (PIF) for all precursor ions acquired in FAIMS and non FAIMS experiments (***Supplementary Figure 4-6c***). The PIF represents the proportion of ion current in the isolation window associated with the target ion and is available in Maxquant [36]. A PIF value approaching 1 indicates that most of the ion current comes from the precursor of interest and contains a low proportion of co-isolated ions. When comparing the distribution of PIF values, we noted that FAIMS reduced significantly the proportion of contaminating precursor ions which decreased the occurrence of co-fragmentation and provided higher Andromeda scores (***Supplementary Figure 4-6c, Supplementary Figure 4-6d***). The intensity distribution of peptides identified only by FAIMS (***Supplementary Figure 4-6f***) highlights the ability of FAIMS to identify lower intensity peptides by reducing background level, consistent with that reported previously in large scale proteomic experiments [35]. We noted that the intensity of the SUMO peptides common to both FAIMS and non FAIMS was 30% lower in the former experiment (***Supplementary Figure 4-6g***), suggesting that the distribution of selected CV steps and slower duty cycle resulted in reduced ion transmission.

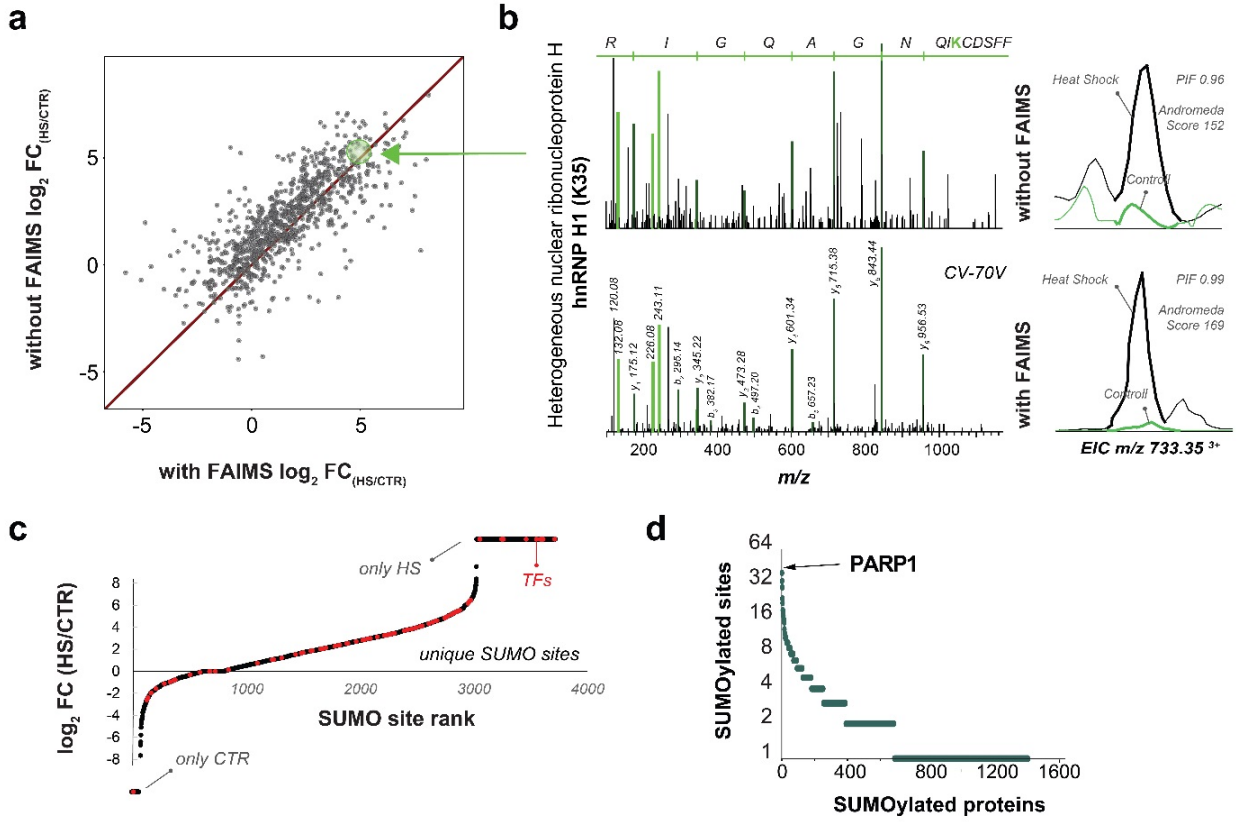


Figure 4-6: *Profiling changes in protein SUMOylation of HEK293 cells following heat shock. (a) Comparison of fold change measurements for common SUMOylated peptides identified with and without FAIMS. (b) MS/MS spectra for FFSCD**K**IQNGAQGIR and the corresponding extracted ion chromatogram for LC-MS/MS analyses with (bottom) and without FAIMS (top). (c) Distribution of fold change and SUMOylation sites rank plot showing sites for transcription factors in red. (d) Number of SUMOylated sites per target proteins. Poly [ADP-ribose] polymerase 1 (PARP-1) shows the highest number of SUMOylation with 40 different sites.*

To profile the changes in protein SUMOylation upon heat shock, we used a label free quantitative proteomic approach and compared the extent of quantifiable peptides/proteins by LC-MS/MS with and without FAIMS. We quantified 2,889 unique SUMO sites on 1,235 protein substrates, of which 887 were common to both FAIMS and non-FAIMS experiments. A comparison of the fold changes observed for the corresponding SUMOylated peptides indicated that the majority of quantified peptides displayed comparable changes in abundance between experiments (***Figure 4-6a***). These analyses revealed a global increase in protein SUMOylation upon one hour heat shock treatment, consistent with the immunoblots shown in ***Figure 4-4b***. However, peptides quantified using FAIMS displayed higher PIF values (median: 0.95) compared

to those from non-FAIMS LC-MS/MS experiments (median: 0.84). An example is shown in ***Figure 4-6b*** for the quantification of the SUMOylated peptide FFSCDKIQNGAQGIR from the heterogeneous nuclear ribonucleoprotein H1. This peptide was identified in FAIMS and non-FAIMS experiments with similar changes in abundance upon heat shock (~15-fold), though the lower PIF values observed in conventional LC-MS/MS resulted in a mixed MS/MS spectrum with a lower identification score (***Figure 4-6b***).

Overall, we observed that approximately two third of the SUMOylation sites were upregulated in response to the heat shock with both approaches (***Figure 4-6c***). Heat shock typically results in an accumulation of misfolded proteins and elevated levels of protein aggregates. To maintain their viability during the heat stress, cells inhibit the transcription and translation machinery whereby several transcription factors (TFs) are SUMOylated [37-39]. TFs are known to be an important target of SUMOylation, predominantly resulting in suppression of transcription and regulation of binding of TFs to chromatin [40]. In contrast, heat shock factors 1 and 2 (HSF1 and HSF2) which play a central role in the transcriptional activation of the heat shock response [37, 41] are activated, and promote the expression of heat shock proteins that act as molecular chaperone to facilitate the refolding of damaged/misfolded proteins [42]. Our SUMO proteome analyses revealed that SUMOylation on K126 and K131 of HSF1 and K2, K82, K139, K151 of HSF2 showed more than two-fold increase in SUMOylation upon heat shock. Likewise, most TFs showed a high increase in SUMOylation following heat shock (***Figure 4-6c***). A specific example is the multifunctional general transcription factor II-I (GTF2I), a protein known to play a role in cell cycle regulation and the DNA damage response pathway. In total, 34 SUMO sites were identified on this protein, half of which were observed only after heat stress. Although the increase in the SUMOylation of GTF2I has been observed in response to various stressors the function of its SUMOylation is still elusive [43].

Cells can also survive the heat stress by other means, such as promoting the degradation of misfolded proteins via the ubiquitin-proteasome machinery. Earlier reports indicated that promyelocytic leukaemia protein (PML) recognizes misfolded nuclear proteins and initiates their SUMO2/3 conjugation [44]. This suggests that SUMOylation of specific proteins by SUMO2/3 is essential for their subsequent

conjugation to ubiquitin via the ubiquitin E3 activity of the SUMO targeted ubiquitin ligase RNF4, leading to the proteosomal degradation of their SUMOylated targets [45]. Our high confidence dataset (0.1% FDR) contains 3,716 unique SUMO sites from 1,564 target proteins, 46% with multiple SUMO sites (≥2 SUMO 3 sites per protein). For example, multi-SUMOylation was detected on proteins involved in the SUMOylation pathway (SUMO1, SUMO2, SUMO3, PIAS1, PIAS4, UBA2, PML, TRIM28, SMARCA4, SMC5, SMC6, NSMCE2, BRCA1, BLM WRN, PARP1, RANBP2), where TOPORS, PIAS1, PIAS4 and NSMCE2 are SUMO E3 ligases. Among all SUMO targets, 263 contained four or more SUMO sites, including poly(ADP-ribose) polymerase (PARP1), a protein known to be SUMOylated [46] (***Figure 4-6d***). PARP1 was previously known to be multi-SUMOylated, and our experiments identified 40 SUMOylation sites on PARP1, the majority of which were strongly up-regulated after heat shock. More than 12 percent of the amino acid sequence of PARP1 is composed of lysine residues and harbor more than 20 putative SUMO consensus motifs, including K203 and K249 residues located within its DNA-binding domain, K486 and K512 are found in its auto-modification domain and K798 is in its catalytic domain [46]. PARP1 which is the most abundant member of the PARP family [47], plays an important role in DNA repair where it recognizes DNA damage and initiates the repair response for single or double strand breaks [48]. Furthermore, PARP1 is modified by several PTM such as phosphorylation, acetylation, ubiquitination in addition to its SUMOylation [47]. In our dataset, 21 out of the 40 SUMOylation sites identified on PARP1 showed more than 2-fold change in abundance upon heat stress. Martin *et al.* reported that multi or polySUMOylation of PARP1 was initiated by PIAS4 [46]. Moreover, the SUMO targeted ubiquitin ligase RNF4 recognizes SUMOylated PARP1 and mediates its proteasome degradation. This cascade promotes the heat shock response through the activation of the HSP70.1 promoter [46].

The SUMOylated proteins were subjected to a network analysis. Clusters of tightly interacting proteins were extracted from this network for gene ontology analysis [11, 15]. Ribosomes that are involved in translation elongation form the largest cluster group, together with proteins involved in ribosome biogenesis. FAIMS extended the coverage of the ribosome biogenesis cluster and identified nucleolus proteins DCAF13,

WDR36, WDR46 and NOL6 (***Figure 4-7***). Moreover, a deeper coverage of the proteasome complex was achieved with FAIMS (PSMD8, PSMA4, PSMA5, PSMA7, SHFM1, UBLCP1 and BAP1 were identified with FAIMS whereas PSMD11 and PSMD12 were unique to non FAIMS). Interestingly, our group showed that the SUMOylation of the proteasome is greatly induced during an MG132 treatment, which lead to the discovery that the SUMOylated proteasome is shuttled to the PML nuclear bodies through its SUMO interacting motif, thereby promoting the degradation of SUMOylated proteins found in the PML nuclear bodies [12]. This hearkens back to the notion that cells can overcome heat stress by degrading misfolded proteins, in this case in a SUMO dependent manner. Moreover BAP1, which was found only in the FAIMS dataset, encodes the ubiquitin carboxy-terminal hydrolase BRCA1 associated protein-1. BRCA1 is a SUMO-regulated ubiquitin ligase (SRUbL), where SUMOylation of BRCA1 promotes its ubiquitinylation activity [49]. BRCA1 is a co-factor of GTF2I [43] and is involved in the DNA damage response (***Supplementary Figure 4-8***). Interestingly, the SUMOylation of several proteins involved in the DNA damage response pathway were upregulated during heat shock, and is consistent with previous reports indicating that hyperthermia induces DNA double-strand breaks (DSBs) in mammalian cells [50, 51].

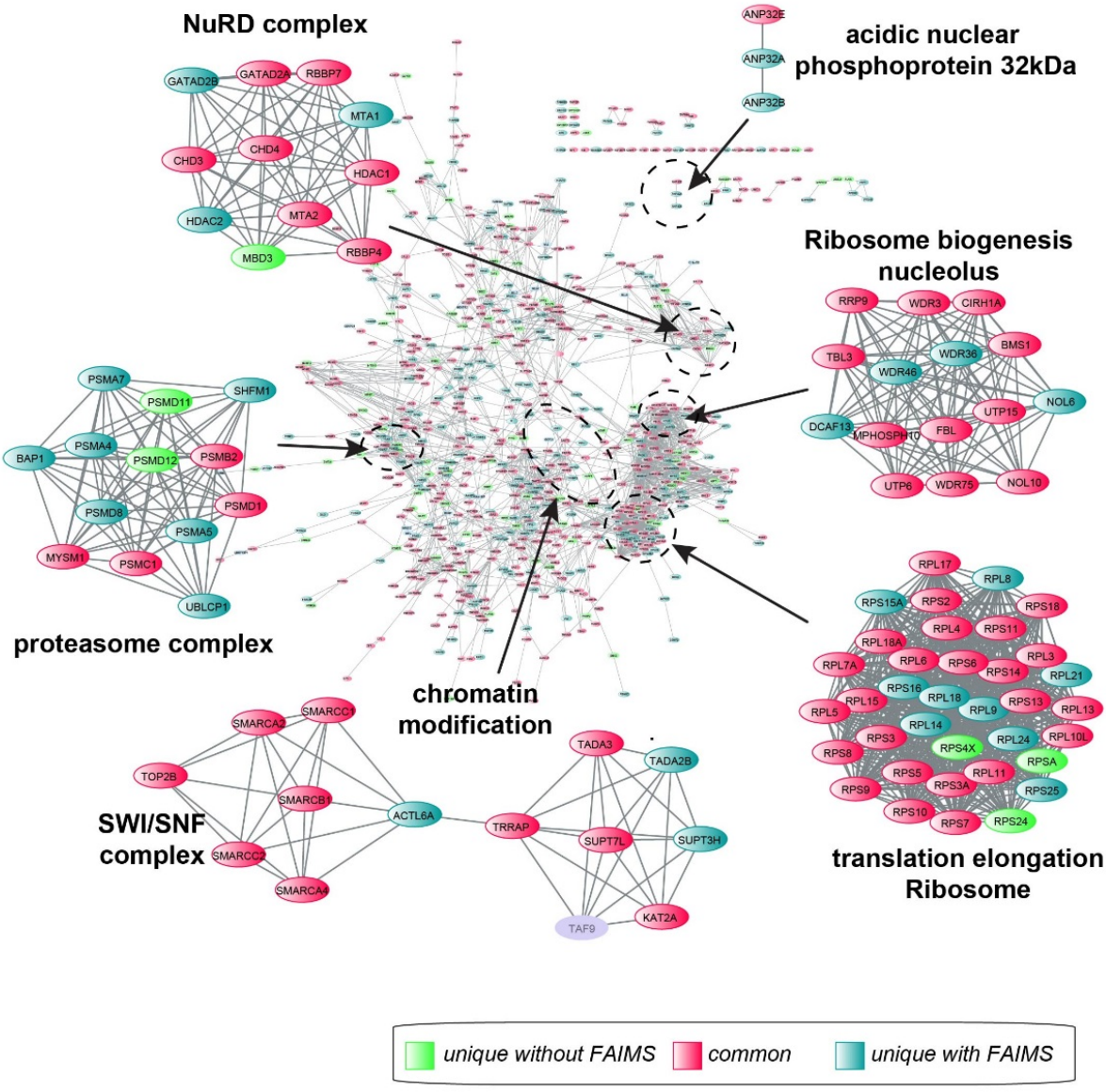


Figure 4-7: *Network representation of identified SUMOylated proteins highlighting different cluster subsets for enriched protein functions. Identification common (red), unique for FAIMS (blue) or without FAIMS (green) are shown.*

Proteins involved in chromatin modification/organization through the SWI/SNF chromatin remodeling complex also showed an increase in SUMOylation upon heat shock. Several modification sites were located near or within DNA binding domain and could affect the transcriptional activities. For example, we noted that upon heat shock

the transcriptional activator BRG1 (SMARCA4) was SUMOylated at residue K711 proximal to the helicase ATP binding domain, and at K1131 within the helicase C-terminal domain. The SUMOylation of SMARCA4 at these residues could impede its helicase activity. Similarly, we identified the increase in SUMOylation of several members of the NuRD complex that control chromatin remodeling through the use of histone deacetylases (HDACs). For example, HDAC1 was SUMOylated at 10 residues including K66, K412, K467 and K476 that showed more than 4-fold up-regulation. Previous studies reported that the SUMOylation of HDAC1 is required to repress transcriptional activities at defined promoters [52, 53]. However, it is still unclear if SUMOylation directly impedes HDAC activity or mediate the recruitment of other chromatin repressors.

4.5. Conclusion

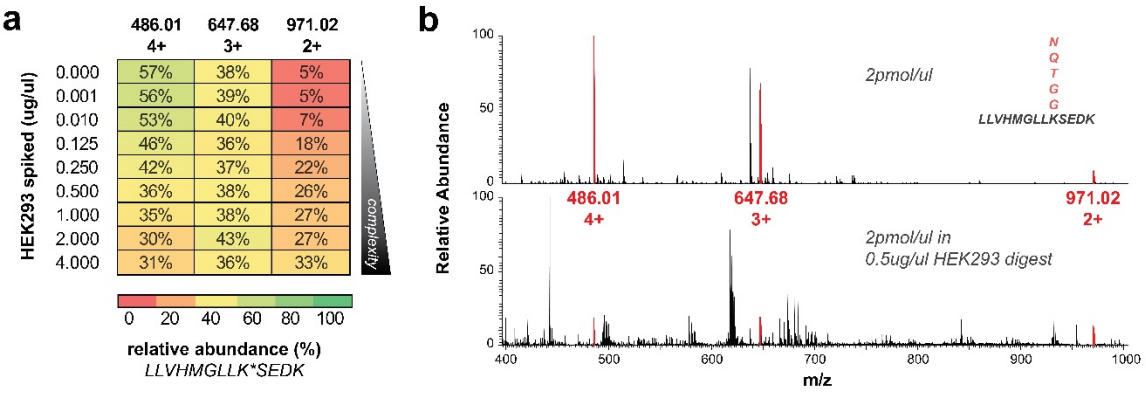
This study shows that gas phase ion fractionation using FAIMS can provide a unique advantage in SUMO proteome analyses by favoring the transmission of branched SUMOylated peptides that can be underrepresented in complex tryptic digests even after immunoaffinity enrichment. When compared to tryptic peptides, SUMOylated peptides show a higher propensity to form triply-charged ions that can be selectively transmitted by FAIMS at CV values where fewer linear tryptic peptides are present. The reduced sample complexity observed at the corresponding CV values confer two important advantages in proteomics. First, it improves the detection and identification of low abundance peptides such as those modified by SUMOylation as shown in the present study. Gas phase ion fractionation using FAIMS also facilitates the detection of multiply-charged ions that could be suppressed when present in a complex population of co-eluting peptides. The reduced contribution of co-eluting ions and contaminants observed when using FAIMS significantly improved the quality of the MS/MS spectra and led to a 36 % improvement in SUMO proteome coverage compared to non FAIMS experiments. Secondly, the reduced sample complexity also enabled the quantification of a larger number of SUMOylated peptides and provided more accurate measurements of changes in protein SUMOylation. While the present study used label-free quantitative proteomics, marked improvements in precision of quantitative measurements were previously reported for multiplex isobaric labeling when using FAIMS. We anticipated that the availability of FAIMS on new generation of mass spectrometers that provide higher resolution and acquisition rates will significantly expand the depth and coverage of proteome analyses and will facilitate the profiling of low abundance protein modifications that remain inaccessible with current LC-MS/MS strategies.

4.6. Acknowledgments

SP is the recipient of a Swiss national science foundation scholarship (P1SKP3-168335). This work was carried out with financial support from the Natural Sciences and Engineering Research Council (NSERC 311598) and the Genomic Applications Partnership Program (GAPP) of Genome Canada. The authors thank Jean-Jacques Dunyach, Michael Belford and Satendra Prasad (Thermo Fisher Scientific) for valuable help and assistance with the FAIMS interface. The Institute for Research in Immunology and Cancer (IRIC) receives infrastructure support from IRICoR, the Canadian Foundation for Innovation, and the Fonds de Recherche du Québec - Santé (FRQS). IRIC proteomics facility is a Genomics Technology platform funded in part by the Canadian Government through Genome Canada.

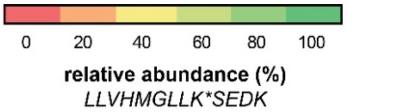
4.7. Supplementary material

4.7.1. Supplementary figures

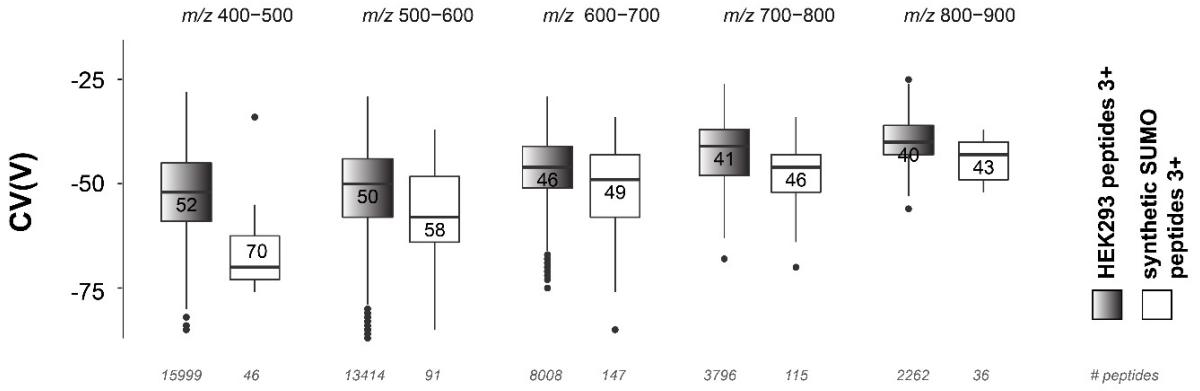


Supplementary Figure 4-1: *SUMO3 peptide LLVHMGLLS**K**EDK at constant concentration of 2pmol/µl was infused with increased concentration of spiked in HEK293 peptides. (a) Relative abundance of charge states 2+, 3+ and 4+ for the peptide. (b) MS spectra for SUMO peptides without HEK293 peptides (upper panel) and with 0.5µg/µl (lower panel).*

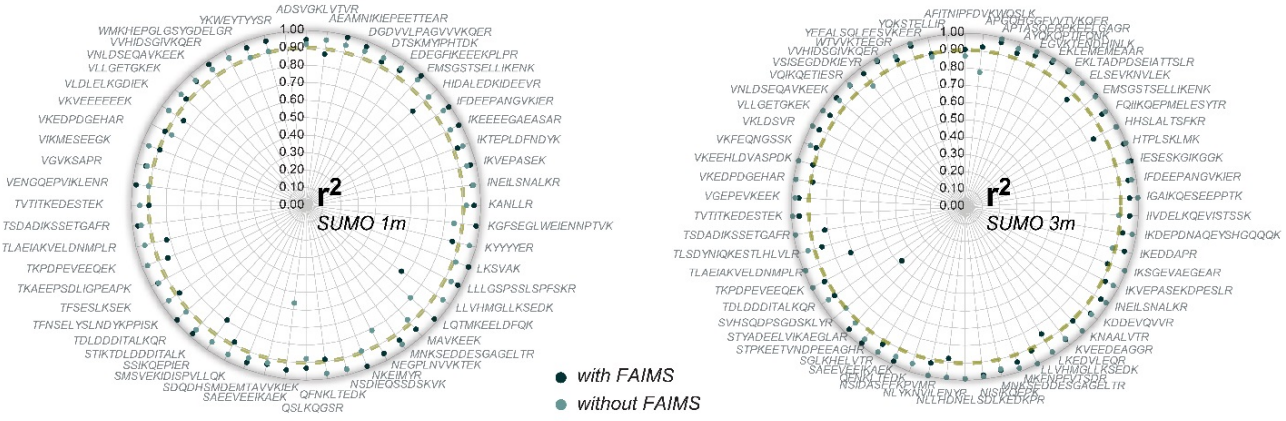
AGC MS	HEK293 ug/ul	486.01 4+	647.68 3+	971.02 2+
1E+04	0.000	52.6%	43.8%	3.6%
5E+04		52.0%	44.4%	3.6%
1E+05		51.7%	44.6%	3.7%
5E+05		51.1%	45.2%	3.7%
1E+06		51.8%	44.9%	3.3%
5E+06		54.4%	43.4%	2.2%
1E+07		57.2%	41.2%	1.7%
1E+04	0.001	51.9%	44.6%	3.5%
5E+04		51.1%	45.5%	3.4%
1E+05		50.7%	45.8%	3.5%
5E+05		54.6%	42.7%	2.8%
1E+06		54.9%	42.6%	2.5%
5E+06		57.5%	41.1%	1.5%
1E+07		58.8%	40.0%	1.2%
1E+04	0.010	43.0%	51.3%	5.7%
5E+04		42.6%	51.7%	5.7%
1E+05		42.5%	51.7%	5.8%
5E+05		42.5%	51.7%	5.8%
1E+06		43.5%	51.4%	5.0%
5E+06		45.8%	49.7%	4.5%
1E+07		48.3%	48.1%	3.6%
1E+04	0.125	42.3%	50.9%	6.8%
5E+04		41.9%	51.5%	6.6%
1E+05		41.7%	51.9%	6.5%
5E+05		42.3%	51.3%	6.4%
1E+06		43.2%	51.4%	5.4%
5E+06		45.4%	50.4%	4.2%
1E+07		47.6%	49.0%	3.3%
1E+04	0.250	14.8%	51.8%	33.4%
5E+04		14.2%	51.1%	34.7%
1E+05		14.5%	53.2%	32.3%
5E+05		15.2%	56.7%	28.1%
1E+06		16.5%	59.1%	24.4%
5E+06		19.2%	62.9%	17.9%
1E+07		19.9%	63.6%	16.5%
1E+04	0.500	1.4%	40.2%	58.4%
5E+04		1.0%	42.2%	56.9%
1E+05		1.2%	40.7%	58.1%
5E+05		1.1%	42.9%	56.1%
1E+06		1.4%	46.6%	52.0%
5E+06		1.7%	61.7%	36.7%
1E+07		1.6%	68.8%	29.6%
1E+04	1.000	0.0%	24.6%	75.4%
5E+04		0.5%	27.8%	71.7%
1E+05		0.4%	27.4%	72.2%
5E+05		0.8%	32.8%	66.4%
1E+06		1.9%	38.9%	59.1%
5E+06		4.6%	51.1%	44.3%
1E+07		3.7%	45.3%	50.9%
1E+04	2.000	0.0%	24.3%	75.7%
5E+04		0.0%	25.5%	74.5%
1E+05		0.3%	26.0%	73.8%
5E+05		0.2%	25.1%	74.7%
1E+06		0.3%	26.0%	73.8%
5E+06		0.7%	40.1%	59.2%
1E+07		1.6%	36.7%	61.7%
1E+04	4.000	0.0%	1.6%	98.4%
5E+04		0.0%	4.1%	95.9%
1E+05		0.0%	8.3%	91.7%
5E+05		0.0%	7.0%	93.0%
1E+06		0.0%	7.1%	92.9%
5E+06		0.0%	6.7%	93.3%
1E+07		0.0%	6.2%	93.8%



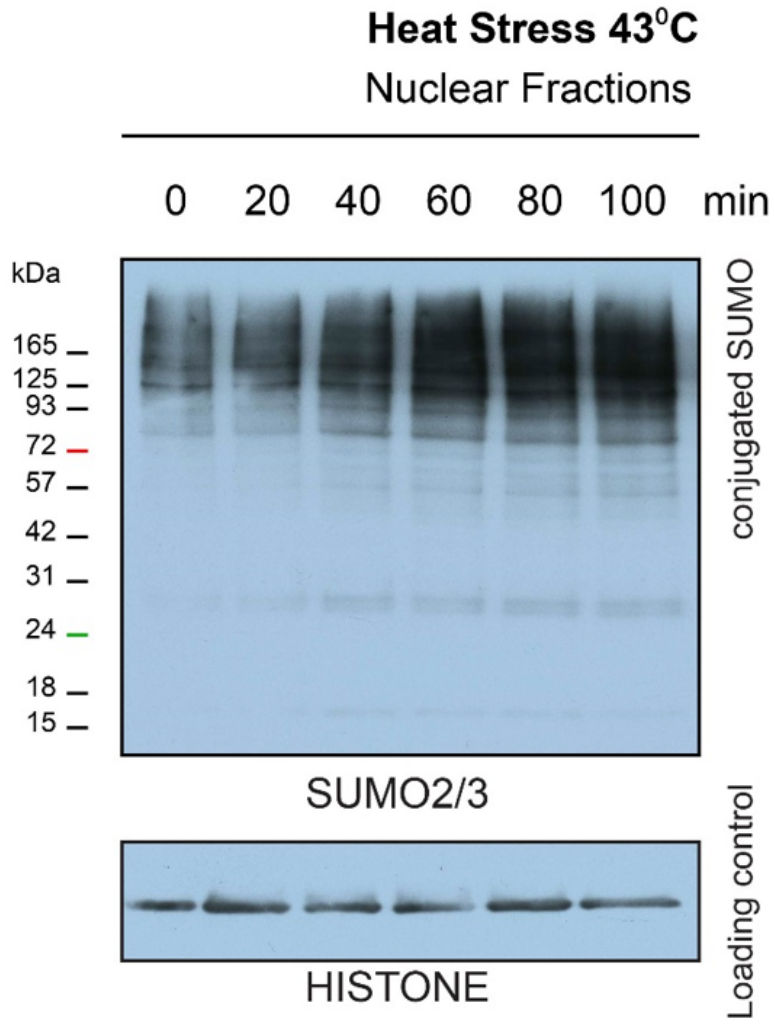
Supplementary Figure 4-2: *SUMO3 peptide LLVHMGLLS**K**EDK at constant concentration of 2pmol/µl was infused with increased concentration of spiked in HEK293 peptides for different AGC values. The relative abundance of charge states 2+, 3+ and 4+ for the peptide is reported for each concentration and AGC values.*



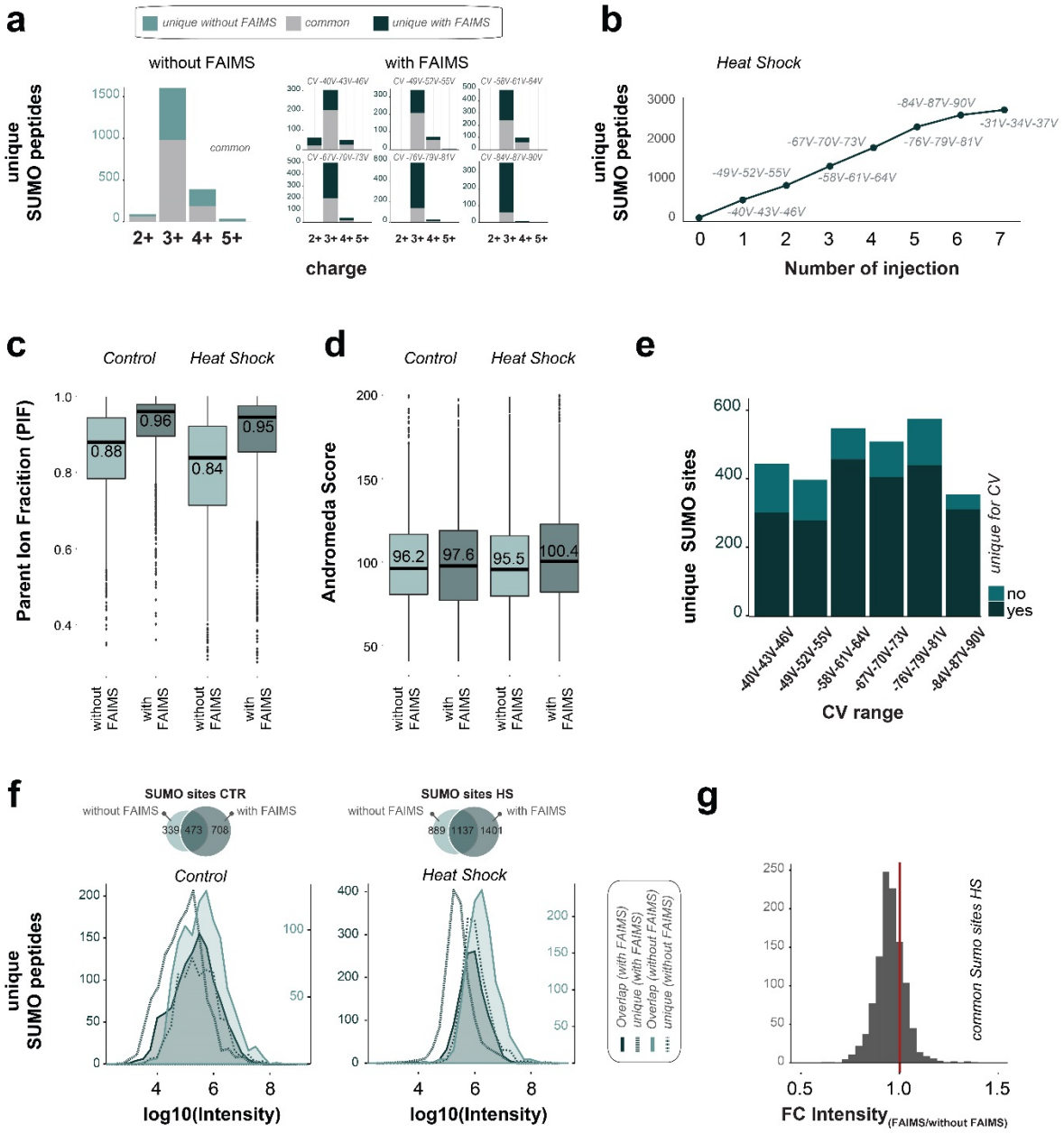
Supplementary Figure 4-3: *Distribution of 3+ SUMOylated and tryptic peptides ions. Bar chart showing the distribution of 3+ peptide ions for different m/z ranges. SUMOylated 3+ peptide ions are consistently transmitted at lower CV values than 3+ tryptic peptide ions.*



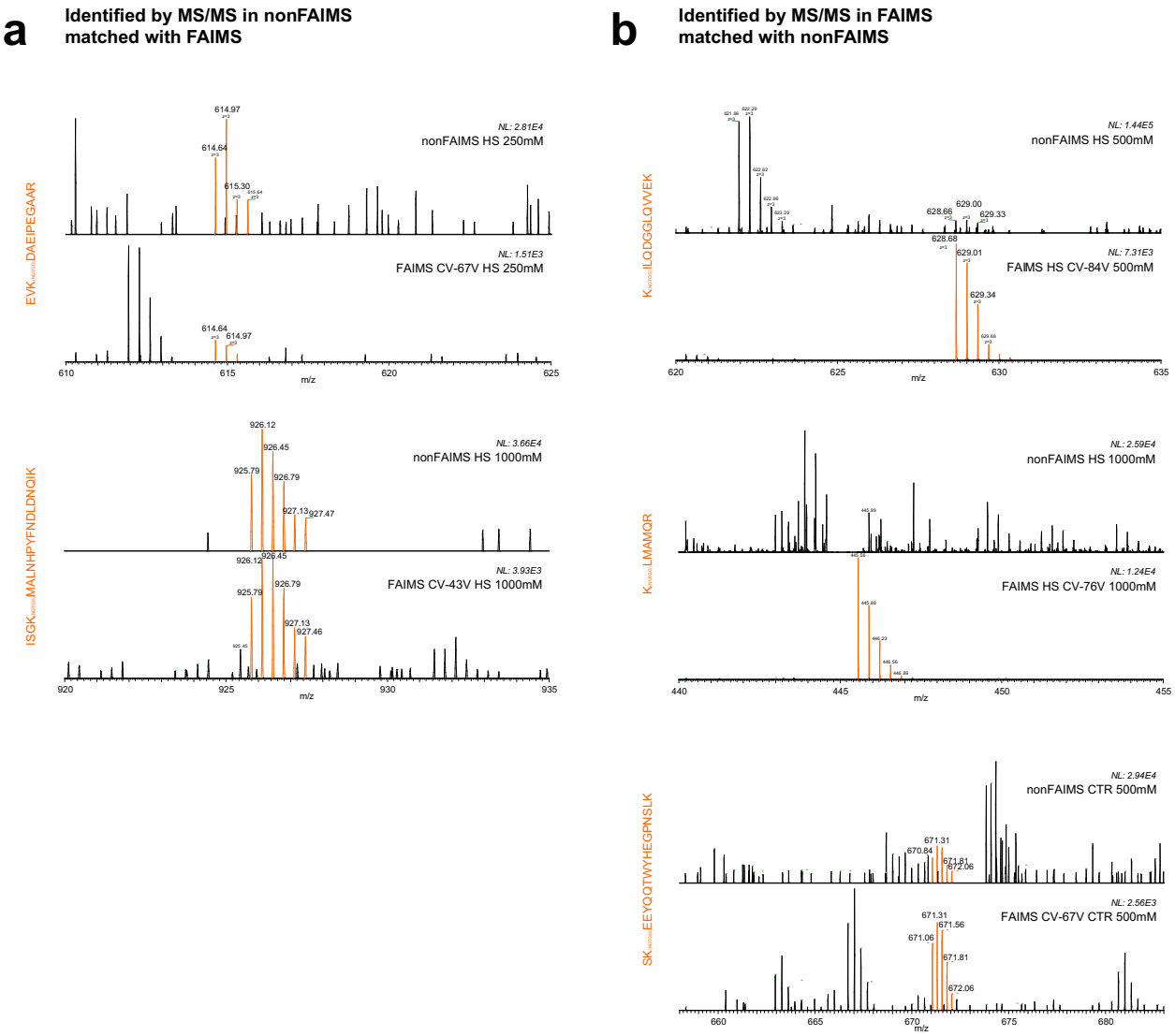
Supplementary Figure 4-4: *Coefficient of determination r^2 for common peptides identified with and without FAIMS. Peptides modified with SUMO1m and SUMO3m paralogs are indicated.*



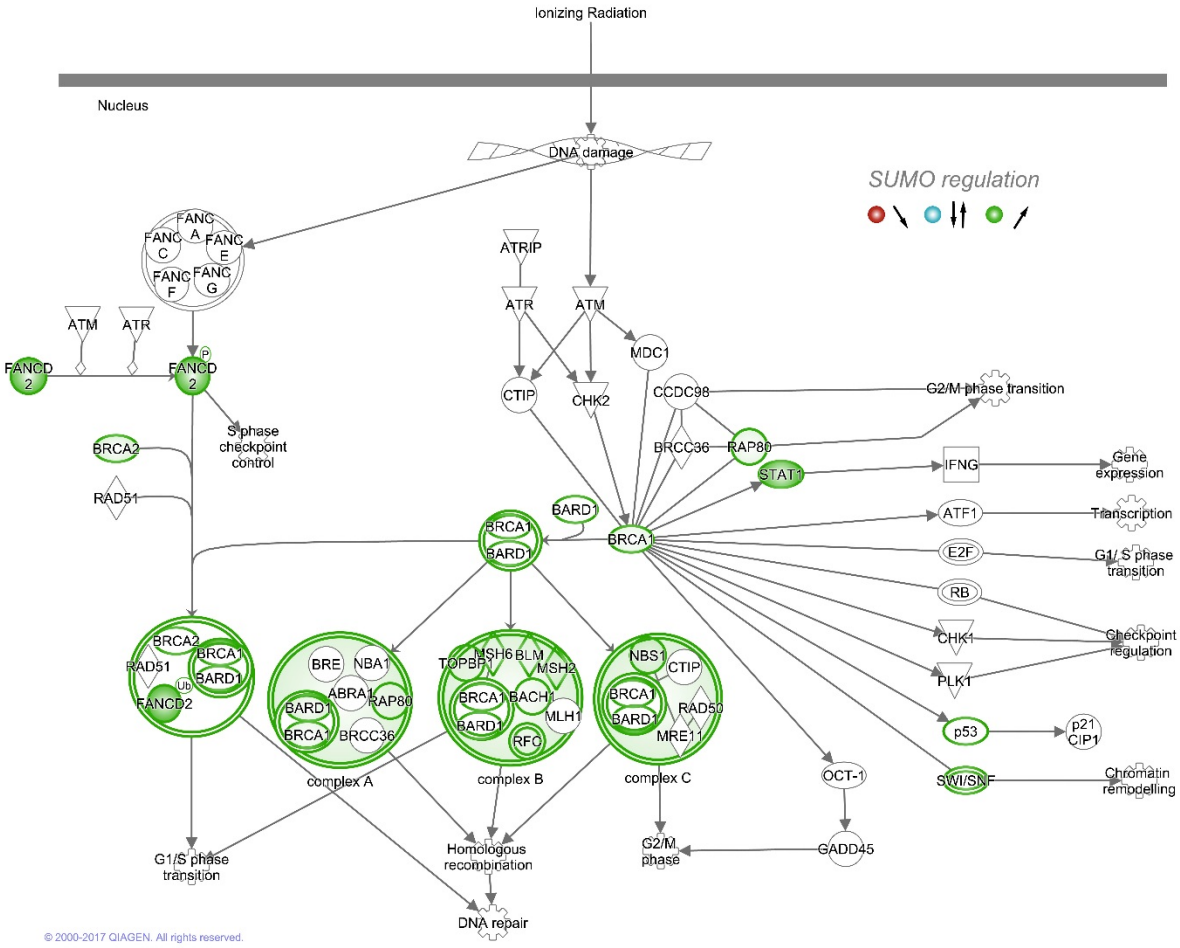
Supplementary Figure 4-5: *Dynamic changes in protein SUMOylation upon heat shock. Western blot using anti SUMO2/3 antibody shows the steady increase in protein SUMOylation following heat shock (43°C) for period ranging from 0 to 100min. Western blot using anti H3 antibody is shown as loading control.*



Supplementary Figure 4-6: *FAIMS extend the depth and sensitivity of SUMO proteome analyses. (a) Charge distribution for SUMO sites identified for large scale heat shock experiment. (b) Cumulative number of unique SUMO sites over separate injections of stepped CV ranges extending from -30 to -80 V. Boxplots showing the distribution of (c) precursor ion fraction (PIF) and (d) Andromeda scores for SUMOylated peptides from control and heat shock treated HEK293 cells obtained with and without FAIMS. (e) CV distribution for unique SUMO sites with FAIMS. (f) Intensity distribution for unique SUMOylated peptides that are common (solid line) or unique (dotted line) for FAIMS and non FAIMS experiments. (g) Fold change of intensity for SUMOylated peptides common to FAIMS and non FAIMS experiments.*



Supplementary Figure 4-7: *Match between runs extend the detection of SUMOylated peptides in FAIMS experiments. a) Representative examples of extracted mass spectra showing SUMOylated peptides identified in non-FAIMS experiments, but detected as a distinct isotopic profile of low abundance in FAIMS experiments. b) Representative examples of extracted mass spectra showing SUMOylated peptides identified in FAIMS experiments, but not detected in non-FAIMS experiments due to low ion abundance or presence of interfering peaks mitigating peptide detection. The sequence of identified SUMOylated peptides is shown at the left of the extracted MS spectra.*



Supplementary Figure 4-8: *Role of BRCA1 in DNA damage response pathway created with Ingenuity Pathway Analysis (IPA).*

4.7.2. Supplementary tables

Supplementary Table 4-1*: List of identified SUMO substrates and modification sites (CD-ROM).*

4.8. References

1. Streich, F.C., Jr. and C.D. Lima, *Structural and functional insights to ubiquitin-like protein conjugation.* Annu Rev Biophys, 2014. **43**: p. 357-79.
2. van der Veen, A.G. and H.L. Ploegh, *Ubiquitin-like proteins.* Annu Rev Biochem, 2012. **81**: p. 323-57.
3. Dou, H., C. Huang, T. Van Nguyen, L.S. Lu, and E.T. Yeh, *SUMOylation and de-SUMOylation in response to DNA damage.* FEBS letters, 2011. **585**: p. 2891-2896.
4. Niskanen, E.A., M. Malinen, P. Sutinen, S. Toropainen, V. Paakinaho, A. Vihervaara, J. Joutsen, M.U. Kaikkonen, L. Sistonen, and J.J. Palvimo, *Global SUMOylation on active chromatin is an acute heat stress response restricting transcription.* Genome Biol., 2015. **16**: p. 153-172.
5. Tatham, M.H., I. Matic, M. Mann, and R.T. Hay, *Comparative Proteomic Analysis Identifies a Role for SUMO in Protein Quality Control.* Sci. Signal, 2011. **4**(178): p. 1-.
6. Eiffer, K. and A.C.O. Vertegaal, *Mapping the sumoylated laadscape.* FEBS J., 2015. **282**(19): p. 3669-3680.
7. Hay, R.T., *Decoding the SUMO signal.* Biochem Soc Trans., 2013. **41**(2): p. 463-473.
8. Hay, R.T., *SUMO: A history of modification.* Mol. Cell., 2005. **18**: p. 1-12.
9. Tatham, M.H., E.G. Jaffray, O.A. Vaughan, J.M. desterro, C.H. Bottino, J.H. Naismith, and R.T. Hay, *Polymeric chains of SUMO-2 and SUMO-3 are conjugated to protein substrates by SAE1/SAE2 and Ubc9.* J. Biol. Chem., 2001. **276**(38): p. 35368-35374.
10. Galisson, F., L. Mahrouche, M. Courcelles, E. Bonneil, S. Meloche, M.K. Chelbi-Alix, and P. Thibault, *A novel proteomics approach to identify SUMOylated proteins and their modification sites in human cells.* Mol. Cell. Proteomics, 2011. **10**(2).
11. Lamoliatte, F., D. Caron, C. Durette, L. Mahrouche, M.A. Maroui, O. Caron-Lizotte, E. Bonneil, M.K. Chelbi-Alix, and P. Thibault, *Large-scale analysis of lysine SUMOylation by SUMO remnant immunoaffinity profiling.* Nat. Commun., 2014. **5**: p. 5409-5420.
12. Lamoliatte, F., F.P. McManus, G. Maarifi, M.K. Chelbi-Alix, and P. Thibault, *Uncovering the SUMOylation and Ubiquitylation crosstalk in Human Cells Using Sequential Peptide Immunopurification.* Nat. Commun., 2017.
13. Tammsalu, T., I. Matic, E.G. Jaffray, A.F. Ibrahim, M.H. Tatham, and R.T. Hay, *Proteome-wide identification of SUMO2 modification sites.* Sci Signal, 2014. **7**(323).
14. Hendriks, I.A., R.C. D'Souza, B. Yang, M. Verlaan-de Vries, M. Mann, and A.C. Vertegaal, *Uncovering global SUMOylation signaling networks in a site-specific manner.* Nat Struct Mol Biol., 2014. **21**(10): p. 927-936.
15. Hendriks, I.A., R.C. D'Souza, J.-G. Chang, M. Mann, and A.C.O. Vertegaal, *System-wide identification of wild-type SUMO-2 conjugation sites. .* Nat. Commun., 2015. **6**: p. 7289-7305.
16. Lamoliatte, F., E. Bonneil, C. Durette, O. Caron-Lizotte, D. Wildemann, J. zerweck, H. Wenshuck, and P. Thibault, *Targeted identification of SUMOylation sites in human proteins using affinity enrichment and paralog-specific reporter ions.* Mol. Cell. Proteomics, 2013. **12**(9): p. 2536-2550.
17. Dumont, Q., D.L. Donaldson, and W.P. Griffith, *Screening method for isopeptides from Small Ubiquitin-Related Modifier-Conjugated proteins by ion mobility mass spectrometry.* Anal. Chem., 2011. **83**: p. 9638-9642.

18. Barnett, D.A., L. Ding, B. Ells, R.W. Purves, and R. Guevremont, *Tandem mass spectra of tryptic peptides at signal-to-background ratios approaching unity using electrospray ionization high-field asymmetric waveform ion mobility spectrometry/hybrid quadrupole time-of-flight mass spectrometry.* Rapid Commun. Mass Spectrom., 2002. **16**: p. 676-680.

19. Barnett, D.A., B. Ells, R. Guevremont, and R.W. Purves, *Application of ESI-FAIMS-MS to the analysis of tryptic peptides.* J. Am. Soc. Mass Spectrom., 2002. **13**: p. 1282-1291.

20. Purves, R.W. and R. Guevremont, *Electrospray ionization high-field asymmetric waveform ion mobility spectrometry-mass spectrometry.* Anal. Chem., 1999. **71**(13): p. 2346-2357.

21. Guevremont, R., *High-field asymmetric waveform ion mobility spectrometry: a new tool for mass spectrometry.* J. Chromatogr. A, 2004. **1058**(1-2): p. 3-19.

22. Creese, A.J., N.J. Shimwell, K.P.B. Larkins, J.K. Heath, and H.J. Cooper, *Probing the complementarity of FAIMS and strong cation exchange chromatography in shotgun proteomics.* J. Am. Soc. Mass Spectrom., 2013. **24**: p. 431-443.

23. Saba, J., E. Bonneil, C. Pomiès, K. Eng, and P. Thibault, *Enhanced sensitivity in proteomics experiments using FAIMS coupled with a hybrid linear ion trap/orbitrap mass analyzer.* J. Proteome Res., 2009. **8**: p. 3355-3366.

24. Swearingen, K.E., M.R. Hoopmann, R.S. Johnson, R.A. Saleem, J.D. Aitchinson, and R.L. Moritz, *Nanospray FAIMS fractionation provides significant increases in proteome coverage of unfractionated complex protein digests.* Mol. Cellular Proteomics, 2012.

25. Xuan, Y., A.J. Creese, J.A. Horner, and H.J. Cooper, *High-field asymmetric waveform ion mobility spectrometry (FAIMS) coupled with high-resolution electron transfer dissociation mass spectrometry for the analysis of isobaric phosphopeptides.* Rapid Commun. Mass Spectrom., 2009. **23**: p. 1963-1969.

26. Bridon, G., E. Bonneil, T. Muratore-Schroeder, O. Caron-Lizotte, and P. Thibault, *Improvement of phosphoproteome analyses using FAIMS and decision tree fragmentation. application to the insulin signaling pathway in Drosophila melanogaster S2 cells.* J Proteome Res, 2012. **11**(2): p. 927-40.

27. Creese, A.J., J. Smart, and H.J. Cooper, *Large-scale analysis of peptide sequence variants: the case for high-field asymmetric waveform ion mobility spectrometry.* Anal Chem, 2013. **85**(10): p. 4836-43.

28. Zhao, H., D.L. Cunningham, A.J. Creese, J.K. Heath, and H.J. Cooper, *FAIMS and Phosphoproteomics of Fibroblast Growth Factor Signaling: Enhanced Identification of Multiply Phosphorylated Peptides.* J Proteome Res, 2015. **14**(12): p. 5077-87.

29. Pfammatter, S., E. Bonneil, and P. Thibault, *Improvement of quantitative measurements in multiplex proteomics using high-field asymmetric waveform spectrometry.* J Proteome Res., 2016. **15**(12): p. 4653-4665.

30. McManus, F.P., F. Lamoliatte, and P. Thibault, *Identification of cross talk between SUMOylation and ubiquitylation using a sequential peptide immunopurification approach.* Nat Protoc, 2017. **12**(11): p. 2342-2358.

31. Prasad, S., M.W. Belford, J.J. Dunyach, and R.W. Purves, *On an aerodynamic mechanism to enhance ion transmission and sensitivity of FAIMS for nano-electrospray mass spectrometry.* J. Am. Soc. Mass Spectrom., 2014. **25**(12): p. 2143-2153.

32. Barnett, D.A. and R.J. Ouellette, *Elimination of the helium requirement in high-field asymmetric waveform ion mobility spectrometry (FAIMS): beneficial effects of decreasing the analyzer gap width on peptide analysis.* Rapid Commun. Mass Spectrom., 2011. **25**: p. 1959-1971.

33. Enserink, J.M., *Sumo and the cellular stress response.* Cell Div, 2015. **10**: p. 4.

34. Golebiowski, F., I. Matic, M.H. Tatham, C. Cole, Y. Yin, A. Nakamura, J. Cox, G.J. Barton, M. Mann, and R.T. Hay, *System-wide changes to SUMO modifications in response to heat shock.* Sci Signal, 2009. **2**(72): p. ra24.

35. Bonneil, E., S. Pfammatter, and P. Thibault, *Enhancement of mass spectrometry performances for proteomics analyses using High-Field asymmetric waveform spectrometry (FAIMS).* J. Mass Spectrom., 2015. **50**(11): p. 1181-1195.

36. Tyanova, S., T. Temu, and J. Cox, *The MaxQuant computational platform for mass spectrometry-based shotgun proteomics.* Nat. Protocols, 2016. **11**(12): p. 2301-2319.

37. Gill, G., *Post-translational modification by the small ubiquitin-related modifier SUMO has big effects on transcription factor activity.* Curr Opin Genet Dev, 2003. **13**(2): p. 108-13.

38. Seo, J. and K.J. Lee, *Post-translational modifications and their biological functions: proteomic analysis and systematic approaches.* J Biochem Mol Biol, 2004. **37**(1): p. 35-44.

39. Filtz, T.M., W.K. Vogel, and M. Leid, *Regulation of transcription factor activity by interconnected post-translational modifications.* Trends in Pharmacological Sciences, 2014. **35**(2): p. 76-85.

40. Rosonina, E., A. Akhter, Y. Dou, J. Babu, and V.S. Sri Theivakadadcham, *Regulation of transcription factors by sumoylation.* Transcription, 2017: p. e1311829.

41. Hong, Y., R. Rogers, M.J. Matunis, C.N. Mayhew, M.L. Goodson, O.K. Park-Sarge, and K.D. Sarge, *Regulation of heat shock transcription factor 1 by stress-induced SUMO-1 modification.* J Biol Chem, 2001. **276**(43): p. 40263-7.

42. Fulda, S., A.M. Gorman, O. Hori, and A. Samali, *Cellular Stress Responses: Cell Survival and Cell Death.* International Journal of Cell Biology, 2010. **2010**: p. 23.

43. Roy, A.L., *Biochemistry and biology of the inducible multifunctional transcription factor TFII-I: 10 years later.* Gene, 2012. **492**(1): p. 32-41.

44. Chu, Y. and X. Yang, *SUMO E3 ligase activity of TRIM proteins.* Oncogene, 2011. **30**(9): p. 1108-16.

45. Guo, L., B.I. Giasson, A. Glavis-Bloom, M.D. Brewer, J. Shorter, A.D. Gitler, and X. Yang, *A cellular system that degrades misfolded proteins and protects against neurodegeneration.* Mol Cell, 2014. **55**(1): p. 15-30.

46. Martin, N., K. Schwamborn, V. Schreiber, A. Werner, C. Guillier, X.D. Zhang, O. Bischof, J.S. Seeler, and A. Dejean, *PARP-1 transcriptional activity is regulated by sumoylation upon heat shock.* Embo j, 2009. **28**(22): p. 3534-48.

47. Luo, X. and W.L. Kraus, *On PAR with PARP: cellular stress signaling through poly(ADP-ribose) and PARP-1.* Genes Dev, 2012. **26**(5): p. 417-32.

48. Woodhouse, B.C. and G.L. Dianov, *Poly ADP-ribose polymerase-1: an international molecule of mystery.* DNA Repair (Amst), 2008. **7**(7): p. 1077-86.

49. Morris, J.R., C. Boutell, M. Keppler, R. Densham, D. Weekes, A. Alamshah, L. Butler, Y. Galanty, L. Pangon, T. Kiuchi, et al., *The SUMO modification pathway is involved in the BRCA1 response to genotoxic stress.* Nature, 2009. **462**(7275): p. 886-890.

50. Takahashi, A., H. Matsumoto, K. Nagayama, M. Kitano, S. Hirose, H. Tanaka, E. Mori, N. Yamakawa, J. Yasumoto, K. Yuki, et al., *Evidence for the involvement of double-strand breaks in heat-induced cell killing.* Cancer Res, 2004. **64**(24): p. 8839-45.

51. Takahashi, A., E. Mori, G.I. Somakos, K. Ohnishi, and T. Ohnishi, *Heat induces gammaH2AX foci formation in mammalian cells.* Mutat Res, 2008. **656**(1-2): p. 88-92.

52. Cheng, J., D. Wang, Z. Wang, and E.T. Yeh, *SENP1 enhances androgen receptor-dependent transcription through desumoylation of histone deacetylase 1.* Mol Cell Biol, 2004. **24**(13): p. 6021-8.

53. David, G., M.A. Neptune, and R.A. DePinho, *SUMO-1 modification of histone deacetylase 1 (HDAC1) modulates its biological activities.* J Biol Chem, 2002. **277**(26): p. 23658-63.

CHAPTER FIVE

5. A novel differential ion mobility device expands the depth of proteome coverage and the sensitivity of multiplex proteomic measurements

Sibylle Pfammatter[1,2], Eric Bonneil[1], Francis P. McManus[1], Satendra Prasad[3], Derek J. Bailey[3], Michael Belford[3], Jean-Jacques Dunyach[3], and Pierre Thibault[1,2]*

[1]Institute for Research in Immunology and Cancer, Université de Montréal, C.P. 6128, Succursale centre-ville, Montréal, Québec, H3C 3J7, Canada.

[2]Department of Chemistry, Université de Montréal, C.P. 6128, Succursale centre-ville, Montréal, Québec, H3C 3J7, Canada.

[3]Thermo Fisher Scientific, San Jose, California 95134, United States

Published:

Journal of Molecular and Cellular Proteomics. 2018 October;17(10):2051-2067

Reprinted with permission from JOURNAL OF MOLECULAR AND CELLULAR PROTEOMICS.

Copyright © (2019) The American Society for Biochemistry and Molecular Biology

Author contributions:

Sibylle Pfammatter designed, executed and analyzed all experiments, generated all figures and wrote the first draft of the paper. Eric Bonneil (MS platform manager) helped to define the project, supported the data acquisition and analysis and contributed to the drafting of the manuscript. Francis P. McManus (lab manager) contributed to interpreting the data of biological processes and the preparation of the manuscript. Pierre Thibault (supervisor) managed the project and wrote the paper. Jean-Jacques Dunyach, Satendra Prasad, Derek Baily, and Michael Belford (all from Thermo Fisher Scientific) conducted preliminary testing and benchmarking FAIMS experiments. All authors approved the content and submission of the paper.

5.1. Abstract

The depth of proteomic analyses is often limited by the overwhelming proportion of confounding background ions that compromise the identification and quantification of low abundance peptides. To alleviate these limitations, we present a new high field asymmetric waveform ion mobility spectrometry (FAIMS) interface that can be coupled to the Orbitrap Tribrid mass spectrometers. The interface provides several advantages over previous generations of FAIMS devices including ease of operation, robustness, and high ion transmission. Replicate LC-FAIMS-MS/MS analyses (N=100) of HEK293 protein digests showed stable ion current over extended time periods with uniform peptide identification on more than 10,000 distinct peptides. For complex tryptic digest analyses, the coupling of FAIMS to LC-MS/MS enabled a 30 % gain in unique peptide identification compared to non FAIMS experiments. Improvement in sensitivity facilitated the identification of low abundance peptides, and extended the limit of detection by almost an order of magnitude. The reduction in chimeric MS/MS spectra using FAIMS also improved the precision and the number of quantifiable peptides when using isobaric labeling with tandem mass tag (TMT) 10-plex reagent. We compared quantitative proteomic measurements for LC-MS/MS analyses performed using synchronous precursor selection (SPS) and LC-FAIMS-MS/MS to profile the temporal changes in protein abundance of HEK293 cells following heat shock for periods up to 9 h. FAIMS provided 2.5-fold increase in the number of quantifiable peptides compared to non-FAIMS experiments (30,848 peptides from 2,646 proteins for FAIMS vs. 12,400 peptides from 1,229 proteins with SPS). Altogether, the enhancement in ion transmission and duty cycle of the new FAIMS interface extended the depth and comprehensiveness of proteomic analyses and improved the precision of quantitative measurements.

5.2. Introduction

Over the past decade, significant improvements in sensitivity, resolution and scan speed of mass spectrometers have extended the depth of proteomic analyses. Peptide identification by mass spectrometry (MS) relies on the detection of precursor ions with predefined charge state and their sequential selection for fragmentation as they elute from the chromatographic column [1, 2]. For low sample complexity, most capillary LC-MS systems are capable of acquiring at least one product ion spectrum for every precursor ion detected.

However, large-scale proteomic analyses represent a sizable analytical challenge in view of the overwhelming peptide density, the occurrence of chimeric tandem mass spectra from co-eluting isobaric ions, and the difficulty of sequencing low abundance peptide ions that are often confounded with background noise. These limitations not only affect the dynamic range of peptide identification and the depth of proteome analyses, but can also compromise quantitative measurements. For label free quantitation (LFQ) the desired precursor ion must be selected in every injection and assigned based on the MS/MS spectra to quantify the peptide between different samples. Congested MS spectra hinder the selection of the precursor ion of choice between the various samples due to the acquisition of MS/MS spectra on other ions in the chromatography window. This difficulty can be partly addressed using spectral alignment and “match between run“ that correlate relevant features based on recalibrated peptide coordinates across LC-MS runs [3]. Quantitation approaches such as label free and metabolic labeling that rely on the correlation of peptide abundances from the survey scan can be affected by the occurrence of confounding ions that share similar ion coordinates (*m/z*, charge and retention time). Furthermore, the presence of co-eluting interfering ions of similar *m/z* values can also affect the accuracy and precision of multiplex quantitative proteomics when using isobaric peptide labeling (e.g. TMT or iTRAQ). Indeed, the co-isolation and co-fragmentation of interfering ions results in distorted reporter ion ratios arising from isobaric tags of the target peptide and interfering ions that cannot be distinguish from each other [4, 5]. The resolution of

conflicting ions often requires multi-stage ion fragmentation and synchronous precursor selection (SPS-MS^3) [6, 7].

To tackle sample complexity and alleviate issues associated with interfering ions, several approaches based on peptide fractionation, improved MS resolution and peak assignments were proposed. For example, two dimensional chromatography fractionation [8, 9] and increased reverse phase column length [10, 11] were reported to enhance peak capacity for proteome analyses. Improvements in quadrupole performance through higher resolution and scanning rate have enabled more efficient isolation and transmission of target ions with narrow isolation windows in turn reducing chimeric tandem mass spectra [12]. Newer generation of MS instruments that combine high resolution of survey scan with fast scanning rate of MS/MS acquisition are progressively bridging the gap that exists between detectable and identifiable peptide ions [13, 14]. The higher resolution of MS instruments also highlights the presence of convoluted isobaric ions with unassigned charge states that prevent their selection for MS/MS acquisition. The recent development of algorithms that facilitate advanced peak determination (APD) can rapidly and iteratively annotate several isotope envelopes in the same *m/z* window, and can enhanced peptide identification by more than 40 % in a single run [15]. While APD can distinguish precursors with overlapping isotope envelopes, its application for quantitative proteomics using isobaric peptide labeling is mitigated by the occurrence of chimeric MS/MS spectra and often requires SPS-MS^3.

These remarkable technological advances have also unveiled the depth of sample complexity that exists in cell extracts as we improve sensitivity and sample peptide ions of lower abundance. While improvements in MS resolution and peak assignment algorithms have expanded the dynamic range of peptide detection, the relative proportion of identifiable and quantifiable peptides remains significantly small [13]. This is particularly true for ion trap instruments where the ion population is constrained to minimize space-charging effects for optimal ion isolation and analysis [16]. Further enhancement of proteome coverage thus requires alternate approaches that can enhance the chromatographic peak capacity of current LC-MS systems while simultaneously enabling the identification of peptides over several orders of magnitude in intensity.

In this context, gas phase ion fractionation using ion mobility spectrometry (IMS) opens new perspectives to overcome sample complexity of proteomic samples. Previous contributions highlighted a broad range of application of IMS to the separation of peptides and proteins including the study of protein conformation dynamics [17-21]. IMS can be divide into two broad categories that regroup drift-tube ion mobility spectrometry (DTIMS) and differential mobility spectrometry (DMS), the latter being mostly represented by high-field asymmetric waveform ion mobility spectrometry (FAIMS). In DTIMS, ions are separated in a drift tube filled with an inert background gas based on their mobilities in a low but constant electric field (E) [22]. For appropriate separation, ions are introduced as a narrow pulse into the drift tube, and resolution up to 200 can be achieved using DTIMS [23]. In FAIMS, ions are separated based on their differences in mobility when subjected to low and high electric fields [24, 25]. A carrier gas entrains ions between two electrodes to which is applied an asymmetric dispersion voltage (DV) that progressively deflects ions towards one of the electrodes. To counter this drift and allow ions to move across the device, a compensation voltage (CV) is applied, and ions are transmitted in turn by scanning the CV [26]. When compared to DTIMS, FAIMS with cylindrical electrodes has several advantages for large scale proteomic studies. First, FAIMS operates at atmospheric pressure and facilitates the separation of co-eluting isobaric interferences including undesired singly charged contaminant ions. Second, FAIMS does not need periodic pulsing of ions to achieve separation and resolution which confers a significant advantage in terms of sensitivity compared to DTIMS. Third, the electric field gradient in the cylindrical FAIMS enables ion focusing and provides a further gain in sensitivity compared to planar FAIMS where the electric field is homogenous [27]. Finally, FAIMS ion transit time is typically less than 50 ms and can be parallelized with the LC-MS duty cycle. Previous studies have reported the benefits of FAIMS to improve proteome coverage and to reduce the extent of co-fragmentation that impedes the identification of low-abundance peptide ions [28-33]. The separation capability of FAIMS also facilitates the resolution of isomeric peptides, including histone variants and isomeric phosphopeptides [34-43]. Furthermore, the reduction of peptide co-elution and co-fragmentation observed with FAIMS can

significantly improve the accuracy and the comprehensiveness of multiplex proteomic analyses [44].

Despite these advantages, a wider acceptance of FAIMS particularly among the recent MS instruments has been relatively limited. This can be explained by the fact that previous generation FAIMS devices were attenuating ion signal by up to an order of magnitude and limiting its usefulness only to applications suffering from overwhelming high chemical background [32]. Another deterrent from routine use of FAIMS has been the long transit time of ions through the ion separation gap leading to ~100 ms delay between CV switching. Poor sensitivity and long CV switch time combined with limited ruggedness and complexity of operation were seen as key drawbacks of the older generation FAIMS for use with modern MS. Here, we introduce a novel FAIMS device that addresses these limitations, and demonstrate its application for proteomics analyses on a Tribrid Orbitrap mass spectrometer. We also compare its analytical performances in multiplexed proteomic experiments with that of SPS-MS^3 to profile the temporal changes in protein abundance of HEK293 cells during heat shock.

5.3. Experimental Section

Direct infusion. Angiotensin II human (Sigma-Aldrich, Oakville, ON, Canada) for direct infusion was prepared in 50 % in methanol (Fisher Scientific, Whitby, ON, Canada) with 0.2 % formic acid (FA; Fisher Scientific, Whitby, ON, Canada). Bovine Serum Albumin (BSA; Bioshop, Burlington, ON, Canada) was dissolved in 50 mM ammonium bicarbonate (Sigma-Aldrich, Oakville, ON, Canada), reduced with 5 mM tris(2-carboxyethyl)phosphine (TCEP; Thermo Scientific, Logan, UT) and alkylated with 10 mM 2-chloroacetamide (Sigma-Aldrich, Oakville, ON, Canada) prior to digestion with trypsin (Promega, Madison, WI, USA) at a ratio 1:50 (wt/wt) overnight. The sample was desalted on an Oasis HLB extraction cartridge (Waters, Milford, MA): The HLB cartridge was first conditioned with 1 mL of 50 % MeOH/0.2 % FA and 2 mL 0.2 % FA. The acidified sample was applied and the cartridge washed with 1 mL 0.2 % FA. Peptides were eluted with 1 mL 50 % MeOH/0.2 % FA and dried down. For direct infusions, the protein mix was diluted to 0.4 µg/µL in 50 % MeOH/0.2 % FA.

Cell culture. HEK293 cells were cultured at 37 °C in a 5 % CO_2 constant atmosphere unless indicated otherwise. Dulbecco's Modified Eagles Medium (Fisher Scientific, Whitby, ON, Canada) was supplemented with 10 % fetal bovine serum (Wisent Inc., St-Bruno, QC, Canada), 1 % Penicillin/Streptomycin Solution (Fisher Scientific, Whitby, ON, Canada) and 1 % L-Glutamine (Fisher Scientific, Whitby, ON, Canada). For heat stress experiments, HEK293 cells were grown in 15 cm petri plates to near confluency (70-80 %), the media was removed and replaced by media pre-warmed at 43 °C. The petri dishes were placed in a 43 °C heated incubator and collected at 1 h intervals for a total of 9 h. Control and heat shock-treated cells were collected simultaneously and washed twice with 37 °C PBS. PBS-washed HEK293 cell pellets were mechanical lysed (3 x 10s sonicate pulses) in a buffer composed of 8 M Urea (Fisher Scientific, Whitby, ON, Canada), 50 mM HEPES (Bio Basic Inc., Markham, ON, Canada) and 75 mM sodium chloride (Fisher Scientific, Whitby, ON, Canada), pH 8.2. The total cell extracts (TCE) were centrifuged for 5 min at 17,000 g to remove cellular debris. The protein concentration of the supernatants were determined by Bradford assay (Bio Rad, Mississauga, ON, Canada).

Tryptic digestion. The TCE were reduced with TCEP (5 mM final concentration) for 30 minutes and alkylated with 2-chloroacetamide (10 mM final concentration) for 30 min at 37 °C in the dark. HEK293 TCE were diluted to 1 M Urea with 50 mM ammonium bicarbonate. Trypsin was added to each sample at a 1:50 (trypsin:protein) ratio and incubated overnight at 37 °C. The reaction was quenched by acidifying the samples with formic acid to pH 3. Samples were desalted on HLB and dissolved in 0.2 % FA for LC-MS analyses. Peptide concentrations were determined using the Thermo Scientific Pierce Quantitative Colorimetric Peptide assay (Thermo Scientific, Logan, UT).

Isobaric peptide labeling. For tandem mass tag (TMT; Thermo Scientific, Logan, UT) labeling, 100 µg of peptides from each time point were dissolved in 100 uL 200 mM HEPES pH 8.2 and mixed with 41 µl anhydrous acetonitrile (ACN; Thermo Scientific, Logan, UT) containing 0.2 mg of individual TMT channel for one hour. The reaction was quenched for 15 minutes with 8 µl of 5 % hydroxylamine (Thermo Scientific, Logan, UT). The ten TMT channels were mixed together in equal volumes and desalted on HLB. The combined TMT samples were dissolved in 0.2 % aqueous FA. An aliquot of 500 ng of this mixture (50 ng of each channel) was injected for each LC-MS/MS analysis.

Mass Spectrometry. For chromatographic reversed-phase separation, an in-house packed 150 µm ID x 25 cm capillary LC column (C_{18}, 3 µm, 300 Å; Phenomenex, Torrance, CA) with eluent A (0.2 % FA in water) and eluent B (0.2 % FA in ACN) was operated through an Easy nLC 1000 Thermo system. Samples were loaded on column and separated at a flow rate of 600 nL/min using a linear gradient starting from 5 % eluent B to 32 % over 125 min, a 10 min hold at 90 % eluent B prior to a 10 min analytical column equilibration with 5 % eluent B. A one-hour gradient was employed for the instrumental reproducibility experiments using 100 replicate injections. For electrospray ionisation, a nanospray Flex ion source (Thermo Scientific, Logan, UT) at 320 ˚C was coupled to an Orbitrap Fusion Tribrid (Thermo Scientific, San Jose, CA) controlled by XCalibur. Spray voltage was set to 3000 V and 2800 V with and without FAIMS, respectively. Full MS (range from *m/z* 300 to *m/z* 1200) in the Orbitrap were acquired at 60 k resolution followed by a top speed MS^2 acquisition of 3 sec at 15 k. Maximal injection time for full MS was 50 ms with an AGC of $5e^5$. Maximal injection time

for MS^2 was 100 ms with an AGC of $2e^4$. For MS^2 triggering, only charge states 2-6 were selected with normalized HCD collision energy activation of 30 % and an isolation window 1.6 Th. For the TMT 10-plex analysis with FAIMS, the MS^2 approach described above was used with minor modifications. The MS^2 resolution was set at 50 k with a maximum injection time of 120 ms and a normalized collision energy of 35 % [45]. For the SPS-MS^3 method, the parameters for the MS scan were the same as for the regular method (scan range *m/z* 300-1200, Orbitrap resolution 60 k, AGC $5e^5$ and maximum injection time 50 ms) followed by a 3 sec top speed approach for MS^2 in the ion trap (Isolation window 0.7 Th, CID at 35 % collision energy, turbo scan rate mode, AGC $1e^4$ with maximal injection time of 50 ms) followed by the selection of 10 synchronous precursor ions for MS^3 acquisition (scan range *m/z* 100-500, Orbitrap resolution of 50 k, AGC of $2e^4$, maximum injection time of 100 ms, 10 notches, isolation window of 2.0 Th and a collision energy of up to 65 % using 5 % stepping).

FAIMS. The new FAIMS interface was re-engineered with the goal of reducing the footprint, improving sensitivity, speeding up CV switch time, and removing helium from the carrier gas. A schematic of the FAIMS interface is presented in ***Supplementary Figure 5-1*** with additional details on the Main Control Board (MCB), adapter flange and electrode assembly. Briefly, FAIMS inner and outer electrodes were separated by a 1.5 mm gap and can be independently heated between room temperature to 100 °C. For this study, the inner and outer electrodes were heated to a common temperature of 100 °C to maximize ion transmission. The inlet on the outer electrode was modified to improve gas dynamics for ion entrainment as described previously by Prasad *et al.* [46]. Nitrogen (N_2) was used as the temperature control gas (12 L/min) and user FAIMS carrier gas (1.6 L/min). The Dispersion Voltage (DV) was set to -5000 V with a 3 MHz frequency for the high electric field. The FAIMS transit time was 40 ms. For CV scanning injections, an individual scan experiment was defined using CV Scan Tool in the Tune User Interface Software. MS parameters for every scan experiment were the same as explained in the previous paragraph.

Data analysis and visualization. RAW files were searched with PEAKS engine (Bioinformatics Solutions Inc., Version 8.5) against the Uniprot Human database (http://www.uniprot.org/) release 2016_02 (17-Feb-2016) including the reversed

database with 74,508 entries. The precursor tolerance was set to 12 ppm and fragment tolerance to 0.015 Da (Orbitrap) and 0.5 Da (Iontrap). The maximum number of missed cleavages for trypsin was set to 3. Deamidation (NQ), Oxidation (M) and Carbamidometylation were set as variable modifications with a maximum of 3 modifications per peptide. For TMT quantification, the TMT 10-plex was added as a fixed modification while adding Phosphorylation (STY) and Methylation (C-term) as extra variable modifications. General false discovery rate (FDR) for peptides was at 1.0 % with decoy removal. Furthermore, only unique peptides within a mass error tolerance of 0.01 Da of the reporter ion and with a 1 % FDR were considered for TMT quantification. Proteins were quantified with ≥ 2 unique peptides presenting reporters in at least 7 channels (Intensity Threshold > 500). TMT dynamic ratios were normalized to the control time point (TMT 126). Coefficient of determination (R^2) for polynomial order up to n=4 were calculated and protein groups with $R^2 \geq 0.75$ were selected for further processing using fuzzy clustering, see Experimental Design and Statistical Rationale for further details [47]. Protein groups were classified into 4 different dynamic profiles. R Studio Version 1.1.414 (https://www.rstudio.com/) was used for data processing and visualization. RawMeat (VAST Scientific, Version 2.1) was used to extract information about the duty cycle metrics for the various LC-MS injections.

Protein interaction networks were constructed with the STRING database using experimentally verified interaction sources only with a medium confidence (0.400) [48] using proteins whose abundances changed in response to the challenge and fit one of the four dynamic profiles described in the results section. The resulting interacting proteins were visualized with Cytoscape version 3.4.0.[49]. Gene Ontology (GO) enrichment analyses were performed and visualized using the BiNGO plugin for Cytoscape [50].

Experimental Design and Statistical Rationale. All regular LC-MS/MS or LC-FAIMS-MS/MS optimization and benchmarking experiments were conducted on the same HEK293 digest to directly compared datasets. The tests for system reproducibility were conducted with N=100 for the LC-FAIMS-MS experiments and N=27 for the regular LC-MS experiments. 27 replicates without FAIMS were conducted to limit instrument use while providing enough replicates to directly compare the LC-MS/MS

configuration to 3 instrumental replicates with LC-FAIMS-MS/MS for the 9 individual CV injections.

For the experiments that compared 3 injections for LC-MS versus LC-FAIMS-MS with HEK293 digests (*Figure 5-3* and *Figure 5-4*), 3 instrumental replicates were conducted without FAIMS, while the 3-stepped CV combination was used for the LC-FAIMS-MS (CV -37 V/-44 V/-51 V, CV -58 V/-65 V and CV -72 V/-79 V/-86 V/-93 V) providing one instrumental replicate, thereby normalizing the instrument usage time between both analysis methods.

The quantitative kinetic analysis of the changes in the proteome upon heat stress were conducted with a single biological replicate per time point (*Figure 5-5, Figure 5-6* and *Figure 5-7*). One 15 cm petri was used for each time point by removing the petri from the 43 °C incubator at the indicated times and processing the sample as indicated above. The TMT10 126 labeled peptide sample was used as an internal non-stimulated control while the other TMT10 tags were used to label peptides from cells that were subjected to the heat shock. A different TMT10 reagent was used to label peptides from the various time points with a 1 h resolution. The 10 TMT labeled samples were combined in an equimolar fashion prior to their analysis by mass spectrometry. For the regular LC-MS experiment the samples was injected 3 times to obtain 3 instrumental replicates. For the LC-FAIMS-MS experiment, the same sample that was used for the regular LC-MS experiment was injected 3 times using the 3-stepped CV combination (CV -37 V/-44 V/-51 V, CV -58 V/-65 V and CV -72 V/-79 V/-86 V/-93 V) providing one instrumental replicate, thereby normalizing the instrument usage time between both analysis methods.

We selected protein groups that contained at least 2 unique peptides and that were quantified (FC = heat shock/control) in at least 7 of the 9 time points. Each profile was fitted to a polynomial with orders of up to n=4. As a first criteria, proteins were selected as “dynamic proteins” if their kinetic profiles fit a polynomial curve with a R^2 value ≥ 0.75. Due to the low proportion and amplitude of the changes in protein levels with respect to the proteome as a whole an FDR value based cutoff on the dynamic profiles by generating a randomized dataset was not possible. Instead an R^2 cutoff

value ≥ 0.75 had to be employed. This cutoff value was chosen since this threshold has been shown to correspond to an FDR of ~5% in large scale proteomic analyses [51]. These “dynamic proteins” were subsequently fit to 4 different dynamic trends based on their c-means fuzzy clustering [52]. “Dynamic proteins” assigned to a cluster with membership ≥ 0.9 confidence were retained as “clustered proteins”.

5.4. Results

5.4.1. Ruggedness and stability of the FAIMS interface

This is the first report of a novel FAIMS interface that is compatible with the newest generation of Orbitrap Fusion Tribrid™ mass spectrometers. An exploded view of the FAIMS device and its coupling to the Orbitrap Fusion instrument are depicted in ***Figure 5-1a***. One advantage of the new FAIMS system is its small footprint that makes interfacing FAIMS to the MS as trivial as interfacing an ion source to the MS. The FAIMS electrodes can be disassembled/assembled for maintenance within a few minutes without interrupting the MS vacuum.

Benchmark experiments were performed to determine ion distribution and evaluate the stability and ruggedness of the FAIMS interface using infusion and LC-MS experiments. A solution of 2 µM of Angiotensin II (*m/z* 523.78^{2+}) was infused at a 1 µL/min flow rate via the nanospray Flex ion source while maintaining identical MS parameters between FAIMS and non-FAIMS experiments. Intensity of Angiotensin II at *m/z* 523.78^{2+} monitored at CV = -78V was $6.0e^{7}$ with FAIMS and $8.1e^{7}$ without FAIMS, corresponding to an ion transmission of 74 % (***Supplementary Figure 5-2a***). The comparison of MS spectra acquired with and without FAIMS also indicate clear visual differences between the two configurations. Experiments performed without FAIMS display different ion populations including singly charged ions and interfering ions (e.g. solvent clusters, convoluted isotopic profiles). In contrast, the FAIMS MS spectrum only showed the transmission of the doubly-protonated peptide ion with no contribution of interfering ions. Under these conditions, the width at half height for Angiotensin II is about 8 V when using an inner and outer electrode temperature of 100 ˚C. The prototype was evaluated for short term stability using the same Angiotensin II solution infused over a period of 24 h. The robustness of CV and DV voltages as well the temperatures of the inner and outer electrode were evaluated. ***Supplementary Figure 5-2b*** shows that the relative standard deviation (RSD) for the signal of Angiotensin II was 5.3 %, highlighting the stability of the spray over this period. The plots for CV, DV and electrode stability over the course of 24 h further emphasize the robustness of the FAIMS interface for sustained operation.

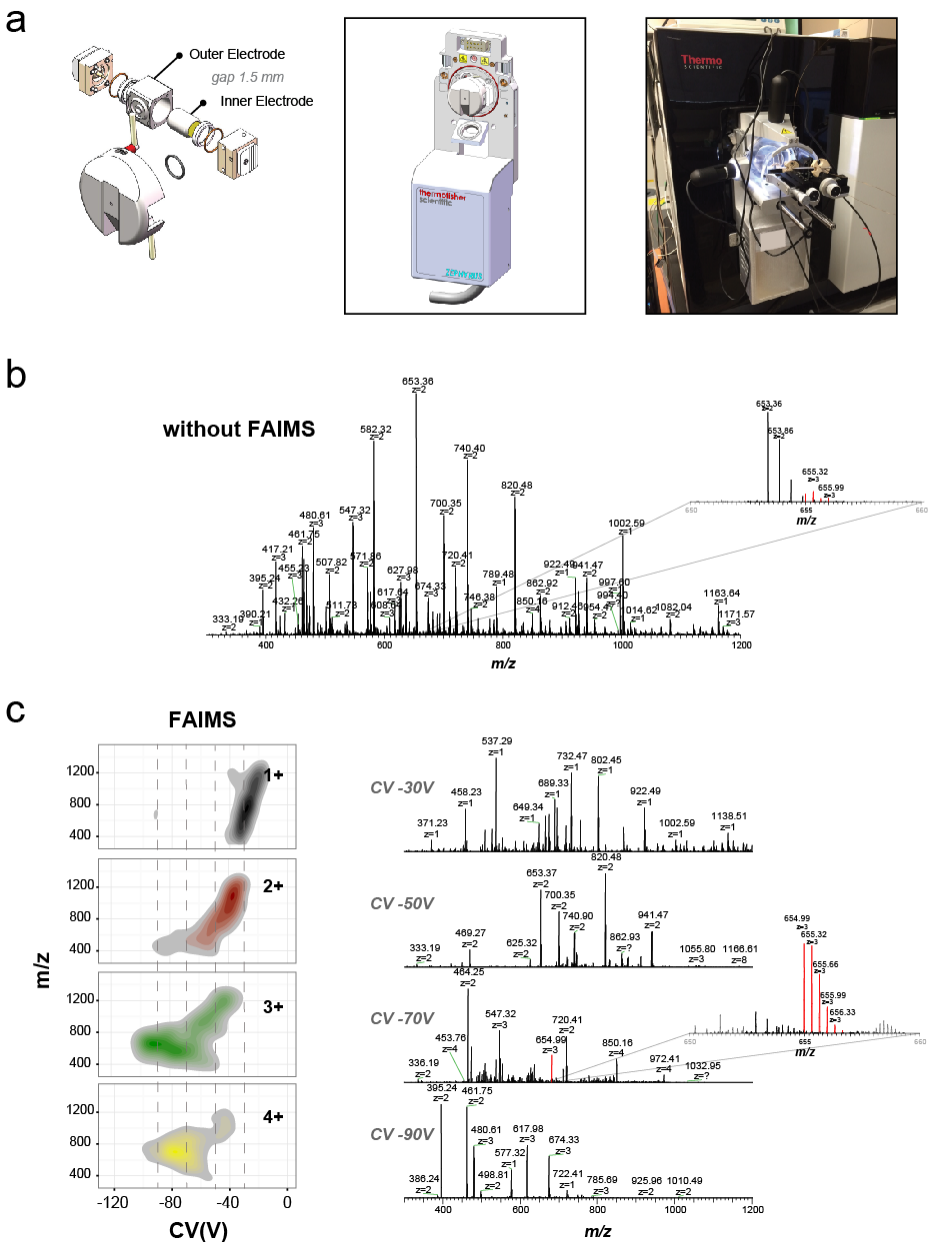


Figure 5-1: *Architecture and separation capacity of a novel FAIMS interface. (a) Scheme of the new FAIMS components with the narrowed electrode gap of 1.5 mm (left). 3D rendered representation of the FAIMS unit (middle) with the clip in installation mechanism that allows the FAIMS unit to be installed like an ion source on the FUSION Orbitrap front end (right). (b) Direct infusion of BSA tryptic digest by standard ESI-MS showing the ion at m/z 654.99 in background noise of the instrument. (c) Direct infusion of BSA tryptic digest by ESI-FAIMS-MS by switching the CV in 2 V steps from -8 V up to -98 V at DV -5000 V. Left, charge state distribution of ions as a function of the CV transmission range. Right, examples of MS spectra for the BSA digest infusion at CV -30 V, CV -50 V, CV -70 V and CV -90 V with the improved detection of the ion at m/z 654.99 highlighted.*

Next, we determined the distribution of peptide ions with and without FAIMS for a tryptic digest of Bovine Serum Albumin (BSA) infused using the nanospray Flex ion source (***Figure 5-1b and Figure 5-1c***). A heat map representation of ion features as a function of their charge state is shown in ***Figure 5-1c*** along with representative MS spectra taken at CV -30 V, -50 V, -70 V and -90 V. The data revealed that gas phase fractionation of the peptide ion current with FAIMS enabled the detection of more low abundance peptide ions that are typically underrepresented with conventional electrospray. This is highlighted with the triply charged peptide ion at *m/z* 654.99^{3+}, which shows improved coverage of the isotopic cluster with FAIMS at CV -70 V. These results also show the selectivity of FAIMS for gas phase ion fractionation and the enrichment of specific charge state ion populations. The analyses of the CV distribution of all observed BSA peptide ions indicated that the median width at half peak height was 8 V (***Supplementary Figure 5-3a***). The distribution of ions across the range of CV steps can also be used to determine the optimal CV values for maximal transmission of specific multiply-charged ions and to reduce the contribution of singly-charge interferences (***Supplementary Figure 5-3b***).

We evaluated the robustness of the new FAIMS interface over repeated LC-MS/MS analyses (CV= -45 V) of 500 ng injections of a HEK293 tryptic digest taken over a 120 h consecutive period (N=100). Peptides were separated with a 70 min gradient and individually searched against the Human database using the PEAKS search engine. The peptide intensities from the various injections were obtained using the label free quantification option in the PEAKS software. The replicates were normalized based on the total ion current and the representation of all the identified peptides present in at least 98 out of 100 replicates were used to create the heatmap shown in ***Supplementary Figure 5-4a***. The chromatographic separation for the 100 replicates was reproducible and constant over the various replicates as exemplified by the total ion chromatograms of replicates 1, 50 and 100 (***Figure 5-2a***). Similar experiments were performed for LC-MS/MS analysis without FAIMS for 27 consecutive injections in the corresponding results are displayed in the heat map of ***Supplementary Figure 5-5a***. In the FAIMS experiments, 4,427 out of a total of 11,964 unique peptide sequences were present in 98 % of all replicates. The reproducibility of peptide detection across

replicates can be obtained from ***Supplementary Table 5-1***. The reproducibility of ion transmission of the FAIMS interface is also shown in in ***Supplementary Figure 5-4a***, where the intensity of common peptides is highly reproducible across the different runs. The RSD for the intensity of each peptide had a median deviation of 18.7 % between replicates for FAIMS (in ***Supplementary Figure 5-4b***), compared to a median deviation of 22 % for non FAIMS experiments (***Supplementary Figure 5-5b***). As expected, we observed an increasing variation in RSD ranging from 9-35 % for decreasing ion intensity (***Supplementary Figure 5-3c***). LC-MS/MS experiments performed without FAIMS showed RSD values ranging from 16-30 % (***Supplementary Figure 5-5c***). The Pearson correlation coefficients were extracted for each pair of replicates as a metric for replicate reproducibility. All replicates showed coefficient values greater than 0.91 for FAIMS (***Figure 5-2b***), similar to that observed for LC-MS/MS experiments performed without FAIMS (***Supplementary Figure 5-5d***). The number of peptide spectrum matches (PSM), unique peptide sequences and protein groups were consistent across the entire duration of the LC-FAIMS-MS/MS analyses (***Figure 5-2c***). After 5 days of constant usage the electrodes showed some trace of accumulation, but without a loss of sensitivity. This accumulation of debris indicates that the FAIMS electrodes can serve as a neutral blocker that prevents particulate matter from entering the orifice of the MS and could extend the sensitivity and lifetime of the instrument. The FAIMS device can be operated for more than a month without having to take the assembly apart for cleaning.

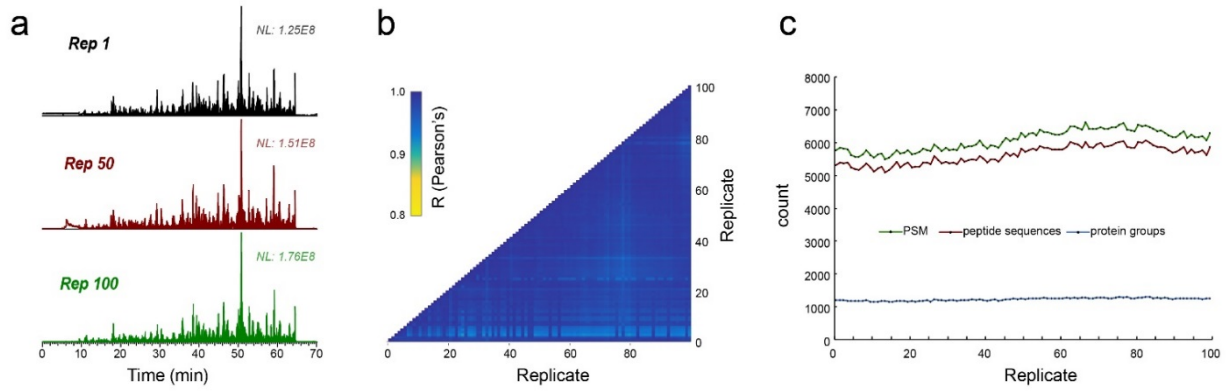


Figure 5-2: *Reproducibility and robustness of FAIMS-Fusion LC-MS system with 100 replicate injections of HEK293 digest at a fix CV value of -45 V. (a) Ion chromatogram for replicate number 1, 50 and 100. (b) Heat map for Pearson correlation coefficients between all 100*

replicate injections. (c) Number of PSM, unique peptide sequences and protein groups for all 100 replicates.

5.4.2. Optimization of FAIMS for LC-MS/MS analyses

To optimize the selection of CV steps for maximal use of MS duty cycle, we first performed LC-FAIMS-MS/MS analyses at individual CVs ranging from -37 V up to -93 V in 7 V steps. The 7 V step was selected based on the distribution of peak width observed for the infusion of BSA tryptic peptides as well as the range of CVs that needed to be covered (***Supplementary Figure 5-3a*** and ***Supplementary Figure 5-3b***). Each LC-FAIMS-MS/MS was performed on a 500 ng injection of a HEK293 tryptic digest using a 2 h gradient. The distribution of acquired MS/MS spectra per survey scan is shown in ***Supplementary Figure 5-6a***, and indicates that peripheral CVs were less populated as evidenced from their lower number of MS/MS acquisition events. To make greater use of the MS duty cycle, the number of injections was reduced by combining different CVs into 3 separate runs. CV combinations (run1: -37 V/-44 V/-51 V, run2: -58 V/-65 V and run3: -72 V/-79 V/-86 V/-93 V) were selected to maximize the number of acquisition events with an average of 15-20 MS/MS spectra triggered per survey scan. The 9 CVs were further combined into two runs (CV -37 V/-44 V/-51 V/ -58 V and CV -65 V/-72 V/-79 V/-86 V/-93 V), which reached the maximum duty cycle. The distribution of the number of unique peptides per run for the 9 different CV steps (***Supplementary Figure 5-6b***) shows that the outermost CV generated fewer identifications, supporting the notion that these ion chromatograms were not densely populated in peptide ions. In contrast, the number of unique peptides obtained for runs that combined multiple CVs indicated that the number of MS/MS events were uniformly distributed for each run, suggesting that the CV combinations were optimal. When analyzing each CV separately (9 runs), the overlap in common peptides ranged from 0.3 % to 17.5 %, whereas the maximum peptide overlap obtained for 3 and 2 runs were 10.7 % and 10.9 %, respectively (***Supplementary Figure 5-6c***).

We evaluated the reproducibility of the peptide identification for separate CV steps and combined within 2 or 3 individual LC-FAIMS-MS/MS analyses. Triplicate analyses performed for 9 separate CV steps led to the identification of 59,044 unique

peptides on average with an overlap of 37,662 common peptides (***Supplementary Figure 5-6d***). The combination of CV steps within 2 or 3 separate LC-FAIMS-MS/MS analyses enabled an average identification of 29,890 and 38,540 unique peptides with an overlap of 17,353 and 21,456 common peptides, respectively. While LC-FAIMS-MS/MS runs performed at individual CV step provided a greater depth of proteome analysis, the combination of CV steps in three separate runs identified ~60 % of the corresponding peptides, and offered a reasonable compromise to maximize sample utilization.

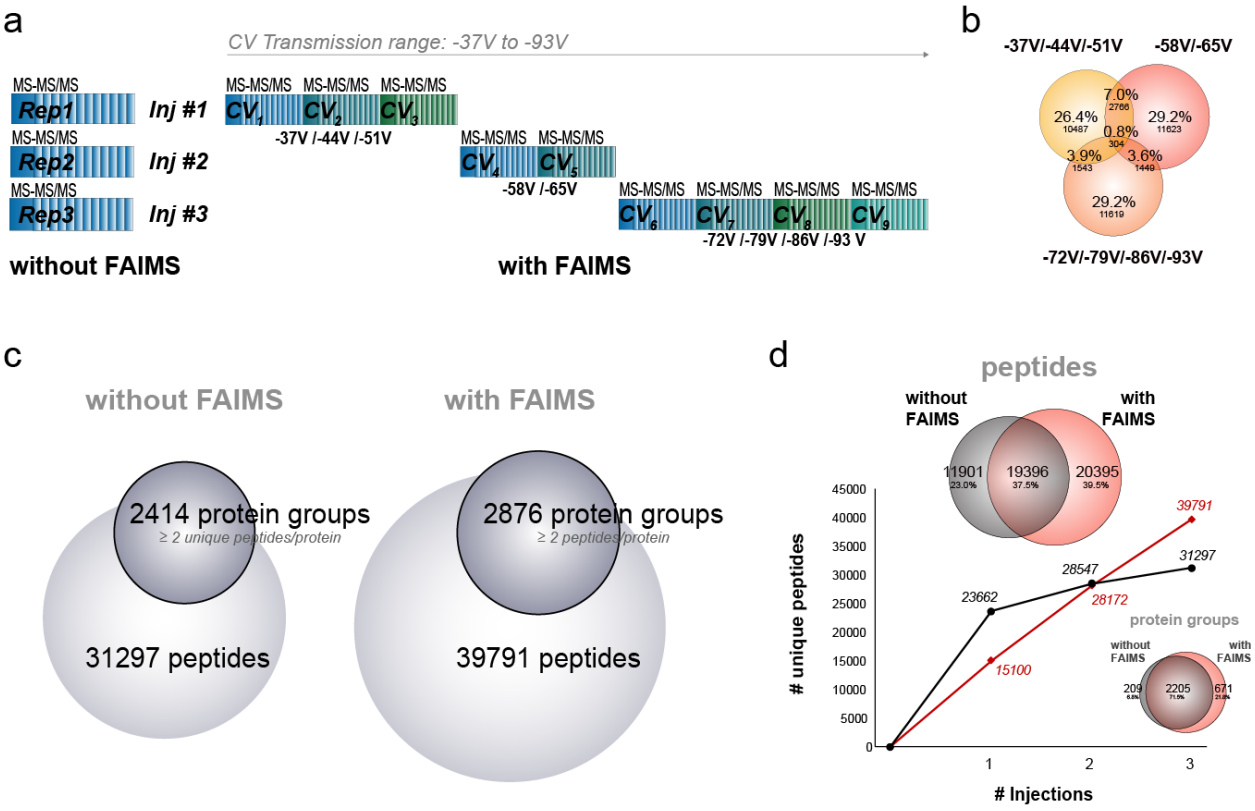


Figure 5-3: *LC-FAIMS-MS CV stepping improves coverage of the proteome compared to regular LC-MS. (a) Overview of the 3 injection LC-FAIMS-MS CV stepping program with optimal CV stepping of CV -37 V/-44 V/-51 V, CV -58 V/-65 V and CV -72 V/-79 V/-86 V/-93 V. The CV stepping cycle consist of a full MS and a top 3 sec MS/MS acquisition at each CV. (b) Venn diagram representation of the overlap in peptide identifications between the 3 CV stepping injections. (c) Number of identified peptides and protein groups with at least 2 unique peptides per protein for the 3 CV stepping injections with FAIMS or 3 repeat injections without FAIMS. (d) Cumulative number of unique peptides identified as a function of repeat injections without FAIMS (black) or CV stepping program with FAIMS (red). Overlap in peptide and protein identifications between the two methods are depicted in Venn diagrams above and below the curves, respectively.*

Next, we compared the 3-injection combined CV stepping method (2-4 CVs per run) with conventional LC-MS/MS analysis. Three replicates of 500 ng HEK293 tryptic digest were analysed by LC-MS/MS with and without FAIMS according to the experimental workflow shown in ***Figure 5-3a***. LC-MS/MS analyses performed with and without FAIMS enabled the identification of 39,791 and 31,297 and unique peptides, respectively. The acquisition of different peptide ion populations across CVs was clearly evident in FAIMS as only 15.3 % of the total number of identified peptides appeared in more than one injection (***Figure 5-3b***). The gain of 27 % in peptide identification with FAIMS also translated in a 19 % increase in the number of protein groups containing at least 2 unique peptides (***Figure 5-3c***). The cumulative number of unique identification obtained for each injection is shown in ***Figure 5-3d***, and indicate that a comparable number of identification was obtained in both configurations after 2 injections. Beyond this point the FAIMS setup outperformed the standard LC-MS/MS approach. Despite the 38 % overlap in identified peptides in the two configurations the overlap in protein groups was excellent with 72 % of protein groups being shared. It should be noted that with FAIMS no replicate experiments were performed, which would add further identifications (***Supplementary Figure 5-6d*** for 2 – 3 and 9 injection methods).

Next, we investigated the proportion of detected versus identified peptide ions across the range of precursor intensities for LC-MS/MS performed with and without FAIMS. These analyses indicated that FAIMS enabled a 50 % gain in the detection of new features (***Figure 5-4a***). From the detected ions, 33 % and 29 % were identified with and without FAIMS, respectively. Interestingly, peptides that were common to both FAIMS and conventional LC-MS/MS followed similar distributions with median log10 intensities of 6.3 and 6.2, respectively. As shown in ***Figure 5-4b*** an average transmission of 78 % was observed for LC-FAIMS-MS, which correlates well with the transmission that was obtained earlier with Angiotensin II. The peptides that were identified only with FAIMS (median log10 intensity of 5.7) were less intense ions than those identified only without FAIMS (median log10 intensity of 6.0). This gain in sensitivity favors the identification of lower intensity features (***Figure 5-4c***).

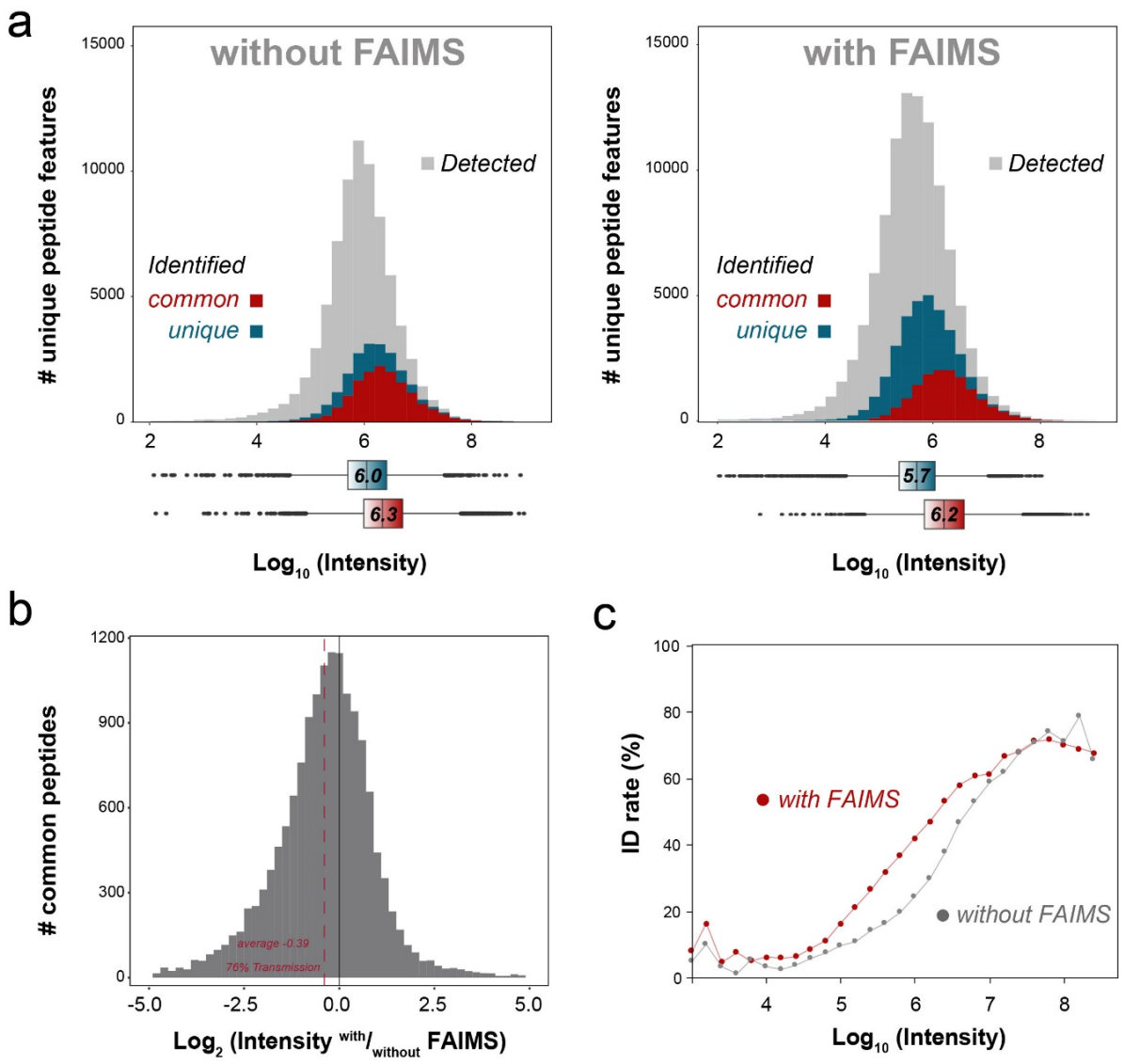
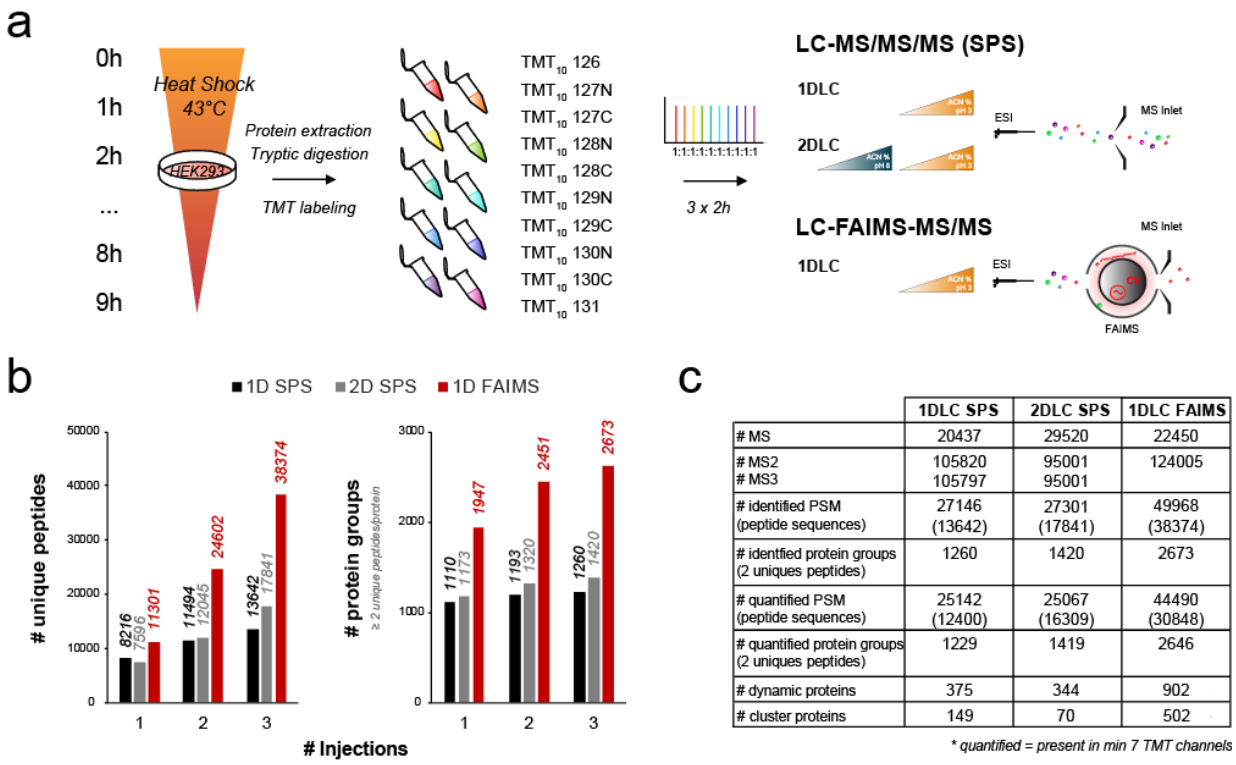


Figure 5-4: *Depth and coverage of proteomic analyses. (a) Intensity distribution for detected peptide ions (grey), peptides identified by both with and without FAIMS (red) and peptides identified only (turquoise) without FAIMS (left) or with FAIMS (right). The corresponding boxplots for the common and unique peptides for both methods are represented below the plots with the same colour scheme. (b) FAIMS peptide transmission rates for common peptides identified by both methods with median transmission of 76 %. (c) Plots for peptide identification rates based on precursor intensity without FAIMS (grey) and with FAIMS (red).*

5.4.3. FAIMS extends the comprehensiveness of quantitative proteomics using isobaric labeling

The ability of FAIMS to reduce spectral complexity and precursor co-selection was advantageously exploited on previous generations of Orbitrap mass spectrometers to improve the accuracy and precision of quantitative measurements in proteomic analyses [28, 44]. In the present study, we compared the analytical benefits of FAIMS for multiplex proteomics using isobaric labeling with that of the synchronous precursor selection SPS-MS^3 method which is commonly used to reduce peptide co-fragmentation and distorted reporter ion ratios [53, 54]. Accordingly, we profiled the temporal change in protein abundance of HEK293 cells upon heat shock in increments of 1 h, up to a total period of 9 h (***Figure 5-5a***). After protein extraction and tryptic digestion, peptides from each of the ten time points were labeled individually with the 10-plex TMT reagent and then combined prior to MS analysis. To directly compare LC-MS/MS with FAIMS or SPS-MS^3, 500 ng of the TMT labeled sample was analyzed using 3 injections, similar to that described previously in ***Figure 5-3a***. In addition, we also compared FAIMS with SPS-MS^3 for 3 concatenated fractions obtained from off-line high pH/basic reverse phase (bRP) fractionation of 1.5 µg of the TMT labeled sample, as this is the most common method for achieving depth and avoiding/decreasing interference when using isobaric labeling (***Figure 5-5a***). Results obtained for each experiment are presented in ***Figure 5-5b*** where a progressive increase in the number of peptide identification is observed for each successive injection. Without FAIMS, replicate injections lead to modest gains in identification with a total of 13,642 peptide sequences (27,146 PSM). A ~30% increase in the number of identification (17,841 peptide sequences) was achieved using bRP fractionation. In comparison, LC-FAIMS-MS/MS analyses enabled up to 3-fold increase in the number of identified peptides compared to SPS-MS^3 (***Figure 5-5c***). A major advantage of using FAIMS for TMT quantification is the increased precursor ion purity that enables high resolution peptide sequencing and reporter ion quantification from the same MS^2 spectrum. Although more dependent scan events were registered with SPS-MS^3, only half are MS^2 events that provide meaningful peptide identification. Indeed, for SPS-MS^3, two spectra must be acquired for identification and quantification,

where identifications occur at the MS^2 level, while quantification of the reporter ions is achieved using the MS^3 spectra.



	1DLC SPS	2DLC SPS	1DLC FAIMS
# MS	20437	29520	22450
# MS2 # MS3	105820 105797	95001 95001	124005
# identified PSM (peptide sequences)	27146 (13642)	27301 (17841)	49968 (38374)
# identfied protein groups (2 uniques peptides)	1260	1420	2673
# quantified PSM (peptide sequences)	25142 (12400)	25067 (16309)	44490 (30848)
# quantified protein groups (2 uniques peptides)	1229	1419	2646
# dynamic proteins	375	344	902
# cluster proteins	149	70	502

** quantified = present in min 7 TMT channels*

Figure 5-5: *FAIMS improves TMT quantification of the human proteome. (a) TMT labeling and quantification workflow. HEK293 cells were exposed to a 43 ˚C heat stress for up to 9 h in 1 h increments. At each time point, cells were collected, lysed, digested with trypsin and labeled with a different TMT_{10} channel. The ten TMT-labeled samples were combined in an equimolar amount and analyzed by 2 h reversed phase LC gradient. Triplicate injections were performed for the SPS method with and without high pH/basic reverse phase fractionation (bRP) while three different CV combinations (CV -37 V/-44 V/-51 V, CV -58 V/-65 V and CV -72 V/-79 V/-86 V/-93 V) were used with FAIMS. (b) Cumulative number of unique peptides and protein groups identified as a function of repeat injections for 1D SPS (black), 2D SPS (grey) or CV stepping program with FAIMS (red). (c) Summary table comparing MS analysis parameters between SPS and FAIMS.*

To determine the number of relevant kinetic profiles that showed dynamic changes in protein abundance, we selected protein groups that contained at least 2 unique peptides and that were observed in at least 7 time points. For convenience, we compared kinetic profiles for LC-MS/MS analyses performed using FAIMS and high pH reverse phase combined with SPS-MS^3. These analyses enabled the identification of 1,260 and 2,673 protein groups, of which 1,229 and 2,646 protein groups were

quantified using SPS-MS^3 and FAIMS, respectively. Dynamic profiles were determined by computing the coefficient of determination (R^2) for polynomial functions with orders of up to n=4. As a first criteria, proteins were selected as "dynamic proteins" if their kinetic profiles fit a polynomial curve with a R^2 value ≥0.75. Secondly, the "dynamic proteins" were subsequently fit to 4 different dynamic trends based on their fuzzy clustering. "Dynamic proteins" assigned to a cluster with membership ≥0.9 confidence were retained as "clustered proteins" and resulted in 149 and 502 profiles for SPS-MS^3 and FAIMS, respectively. A list of quantified dynamic Protein groups (cluster membership > 0.9, with at least 2 unique peptides and TMT reporters present in ≥ 7 TMT channels) is provided in ***Supplementary Table 5-4***. Next we grouped these profiles into 4 dynamic trends corresponding to 1) upregulated early, 2) upregulated late, 3) down regulated early and 4) down regulated late (***Figure 5-6a***). The grey lines in ***Figure 5-6a*** represents the relative fold changes for all proteins that fit the trend, while the blue curve represents the smoothed average $\log_2$ fold changes for all clustered proteins from that group. As illustrated by the heat maps in ***Figure 5-6b***, the overall changes in protein abundances observed by SPS-MS^3 and FAIMS followed similar trends. More importantly, proteins that were quantified by both SPS-MS^3 and FAIMS showed comparable fold changes, as illustrated in the representative plots for HSPA1A and LRPPRC in ***Figure 5-6c***. A comparison of fold change observed for common peptides and proteins is provided in ***Supplementary Figure 5-7***, and highlight the correlation of protein quantification measurements between SPS-MS^3 and FAIMS.

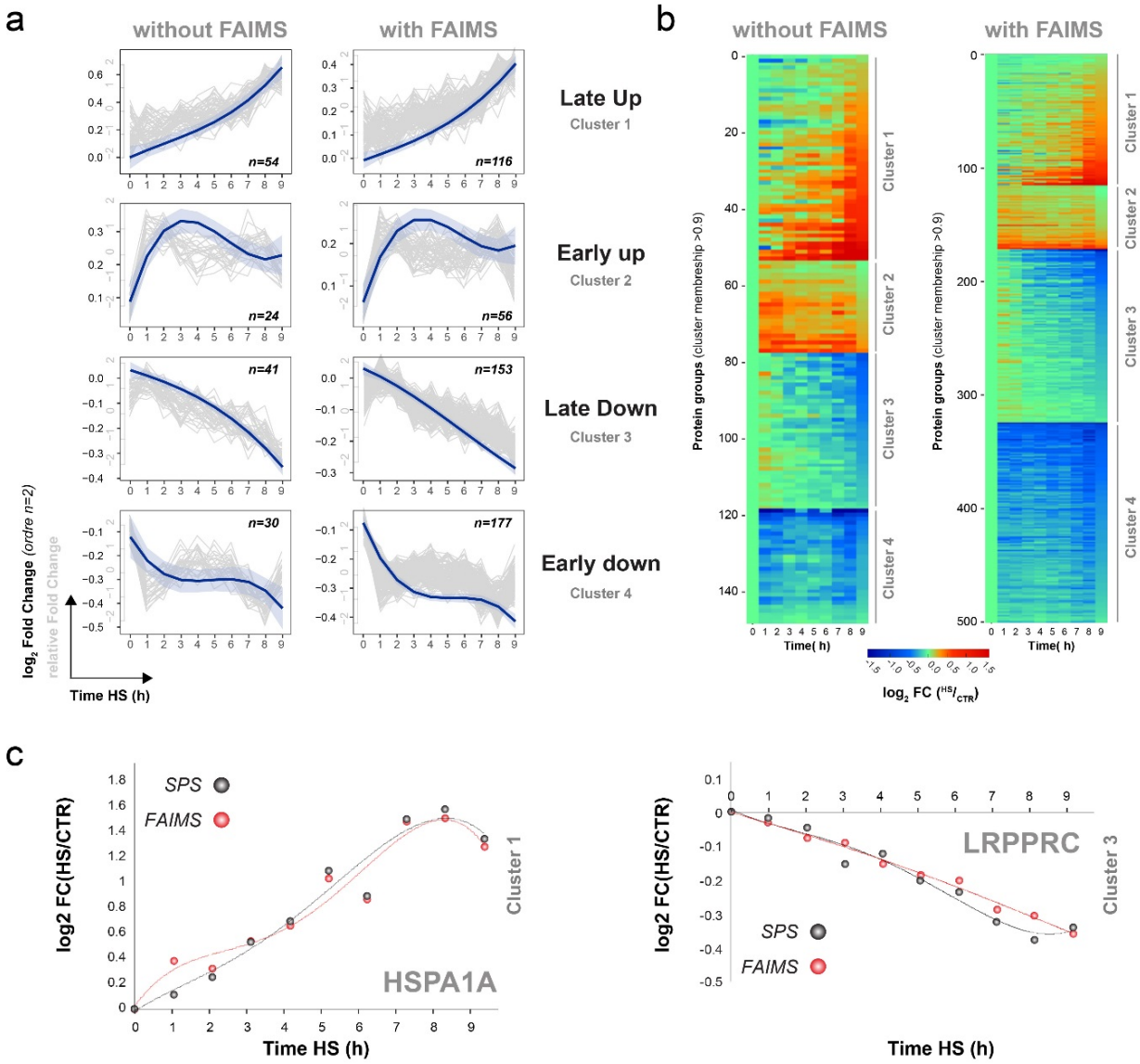


Figure 5-6: *The heat shock response in HEK293 cells alters protein abundances with distinct dynamics. (a) Dynamic clusters for heat shock regulated proteins without FAIMS (left) and with FAIMS (right). The grey lines show the relative fold changes of the individual proteins with high membership (≥ 0.9), the blue lines the average fold changes of all the proteins in the corresponding cluster. (b) Heat map for all proteins in the clusters from (a). (c) Representative dynamic profile of HSPA1A (assigned to a late up regulation) and LRPPRC (assigned to a late up downregulation) for SPS (grey) and FAIMS analysis (red), highlighting virtually identical profiles/quantifications for both acquisition methods.*

5.4.4. Dynamic proteomics enabled the profiling of cellular response upon hyperthermia

The protein network shown in ***Figure 5-7*** provides an overview of the heat shock response in the HEK293 cells. Proteins whose abundances vary in response to the heat shock stimulus were grouped together. The majority of proteins depicted in the network are known to interact with multiple partners whose abundance also varies during the heat stress. Proteins that are grouped into specific GO terms (outlined shapes) have a tendency to alter their abundance in the same demeanor except for proteins from the spliceosomal complex.

Using TMT 10-plex to follow the kinetics of the heat shock process allowed for a greater understanding of the underlying dynamics. Interestingly, the acute heat shock response was primarily mediated by the down regulation of proteins. Several proteins involved in chromatin remodeling were rapidly decreased in abundance upon heat shock, suggesting that chromatin remodeling is an important adaptation response to hyperthermia. Chromatin remodeling is taking place in the early phase of the heat stress prior to upregulation of heat shock proteins (response to unfolded protein GO term), and has been reported previously [55]. Indeed, all regulated proteins that are involved in the response to unfolded protein (DNAJA1, DNAJB1, DNAJB4, HSP90AA1, HSPA1A, HSPA6, HSPA8, HSPE1 and HSPH1) were upregulated in the later stages of the heat shock response.

The heat shock response is believed to be initiated in part by HSF1 [56]. HSF1 is translocated to the nucleus upon heat stress to promote the transcription of heat shock proteins to favour the refolding of proteins that may have denatured as a result of the stress [57]. The efficient transcription of the heat shock proteins requires chromatin remodeling as exemplified by the delay of the transcription of the HSP70 gene by nucleosome formation [55]. This is in line with the rapid change that was observed with chromatin associated proteins followed by the subsequent induction of proteins involved in the cellular response to unfolded proteins (***Supplementary Figure 5-8***). It is well documented that upon heat stress proteins can denature and promote their aggregation or miss-folding. Induction of heat shock proteins favours their refolding or prevents

denatured protein from aggregating, which can have detrimental effects on the cell. An increase in abundance of several heat shock proteins was observed, including the HSP90 protein HSP90AA1. This chaperon not only prevents the aggregation and favours refolding of proteins but also serves to inactivate HSF1 [58]. The upregulation of HSP90AA1 in the later stages of the heat stress could serve as a feedback loop during the process to inhibit the induction of more heat shock proteins beyond a set threshold.

Proteins that are localized in the mitochondria decreased rapidly in response to heat shock. A rapid down regulation in proteins involved in the electron transport chain (ETC) was observed upon hyperthermia consistent with earlier reports [59]. Proteins that are part of complex I (NDUFA2, NDUFA4, NDUFA5, NDUFA7 and NDUFS3), complex II (SDHA and SDHB), complex III (UQCRC1) and complex IV (COX4I1, COX5A and COX6B1) were all down regulated. Moreover, a depletion in the levels of the mitochondrial ribosomes was observed throughout the heat shock process (acute and delayed). The rapid depletion of the ETC proteins from Complexes I-IV during the heat shock is known to lead to a rapid and drastic drop in ATP levels in the cell. This process is coupled to the release of reactive oxygen species (ROS) [60]. These observations are congruent with the notion that hyperthermia promotes the disassembly of the mitochondrial wall, resulting in the release of its content.

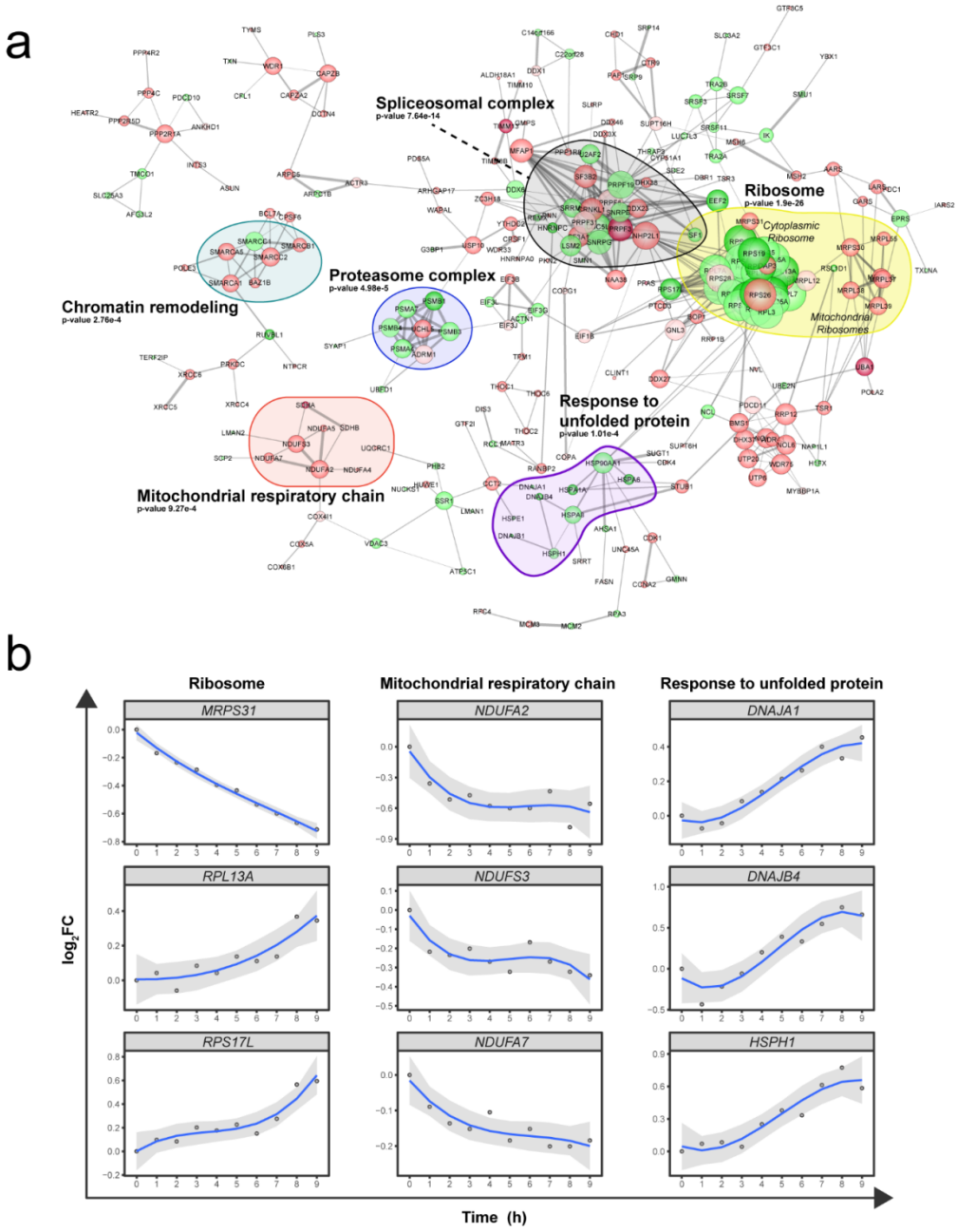


Figure 5-7: *Heat stress affects several key cellular processes that impact protein homeostasis. (a) Interaction network for upregulated proteins (green) and down regulated proteins (red) based on the clustering shown in Figure 5-6a. The darkest shades of green and red correspond to proteins that were quantified by SPS and FAIMS, the intermediate shades to proteins that were quantified by LC-FAIMS-MS only and the lightest shades to proteins that were quantified by SPS only. Proteins that belong to enriched GO-terms are outlined by colored shapes. (b) Examples of kinetic profiles for proteins associated with the ribosome, the mitochondrial respiratory chain, and the response to unfolded proteins.*

Prolonged heat stress causes protein miss-folding and requires the removal of impaired and non-functional proteins. As a result, 5 of the 14 core proteins that make up the 20S proteasome catalytic complex were upregulated in the later stages of the heat shock. In contrast, the two proteasomal proteins that were not part of the 20S proteasome subunit (ADRM1 and UCHL5) were downregulated in response to the stress (***Supplementary Figure 5-8***). ADRM1, a component of the 19S proteasome complex, interacts and activates the deubiquitinase protein UCHL5 [61]. ADRM1 favours the association of UCHL5 to the proteasome to promote the removal of poly Lys-48 linked ubiquitin from proteins that are targeted to the proteasome as a means to reduce their aberrant proteolysis. Overall, the levels of proteasome in the cell are increased and output of this machinery is augmented at the potential cost of error by eliminating the proof reading capability of UCHL5 to alleviate cytotoxic protein aggregates.

Interestingly, only components of the spliceosomal complex were increased in abundance in the acute adaptation response to the heat stress. Heat stress has been shown to adversely affect the splicing machinery in HeLa cells [62]. Due to the temperature sensitivity of the splicing machinery, the low temperatures that can inhibit this process do not affect transcription. This leads to the buildup of mRNA precursors that can erroneously be translated into toxic proteins. The buildup of mRNA precursors can be partially mitigated by increasing the copy number of the splicing machinery in order to keep up with the downstream premRNA production.

In contrast to the mitochondrial ribosomal content that was reduced, the heat shock led to a late increase in non-mitochondrial ribosomal content (***Supplementary Figure 5-8***). Of the 28 non-mitochondrial ribosomal proteins that were quantified, only RPS26 was downregulated in an acute manner. Interestingly, RPS26 is involved in the recognition of the Kozak sequence [63]. Indeed, ribosomes that lack RPS26 have a greater propensity to translate proteins that do not start with canonical Kozak leader sequences. More importantly, proteins from the stress response pathway tend to begin with non-Kozak sequences. These findings indicate that both increasing the level of ribosomes in the cell while altering the ribosomal complex by eliminating RPS26 favours the heat shock response by promoting the translation of key proteins.

Overall, the early decrease in proteins involved in cell-cell adhesion, ATP-dependent chromatin remodeling and mitochondrial proteins, and the rapid increase in proteins involved in mRNA splicing during the stress are presumably the catalysts that prompt the upregulation of non-mitochondrial ribosomal proteins, the cellular response to unfolded proteins and the down regulation of the mitochondrial ribosome in the later phase of the challenge. Clearly, the heat shock response includes a number of cellular pathways that result in the regulation of the toxic effects of protein aggregates, alteration in protein synthesis, increased catabolic activities and targeted translation of HSP gene products.

5.5. Discussion

This report described the analytical benefits of a re-engineered FAIMS device that can be interfaced on the latest generation of Orbitrap Tribrid instruments and can provide significant advantages for proteomic analyses. FAIMS devices were previously available on earlier generations of mass spectrometers but had several drawbacks in terms of ease of operation, ruggedness and sensitivity that prevented their wider acceptance. Indeed, the installation and operation of previous FAIMS devices were tedious due to their large sizes and the use of non-standard ionization sources that were not equipped with camera to assist spray optimization. The use of capillary LC-MS/MS system operating at nanoflow was not straightforward with previous interface due to the spray instability arising from the high flow rate of gas flow exiting the FAIMS device. Also, the entry of droplets into the FAIMS device led to DV instability that prevented the use of LC-FAIMS-MS/MS over extended time periods. However, the more compact footprint of the new FAIMS unit facilitates its installation on the newest generation of Orbitrap Tribrid mass analyzers and can be assembled, mounted and operated within a few minutes. While LC-FAIMS-MS/MS experiments described here were performed with 10 cm x 150 µm I.D. capillary columns and a flow rate of 600 nL/min, smaller columns (e.g. 75 µm I.D.) operating at lower flow rates can also be used equally well with the new FAIMS interface. The FAIMS interface can be used in combination with standard ionization sources equipped with camera to rapidly optimize the position of the electrospray emitter. The stability of the new FAIMS device is vastly improved and can be used for time periods extending over 120 h without significant losses in sensitivity. The use of FAIMS in LC-MS/MS analyses also prevented the direct entry of droplets and contaminants in the transfer tube and S-lens assembly, and can extend the continuous operation of the mass spectrometer.

The ion transmission through the FAIMS device was significantly improved and led to ~80 % ion transmission for both direct infusion and LC-FAIMS-MS/MS analyses. This is in stark contrast with previous generation of FAIMS units that provided transmission efficiency of 13-77 % [64]. Several improvements in the electrode assembly contributed to this enhanced transmission. First, nitrogen flow was supplied to

the FAIMS inlet to maximize Coandã effects and prevent ions from preferentially striking the inner electrode upon entry [65]. Second, the gap space between the outer and inner electrode of the unit was reduced from 2.5 mm to 1.5 mm [64, 66]. This in addition to the smaller FAIMS electrodes translated into an increased electric field and a decreased transit time of ions into the FAIMS device from 130 ms to 40 ms to improve the duty cycle and the number of MS/MS spectra acquired per cycle.

For proteomics applications, the ability of FAIMS to fractionate ions based on their gas phase differential mobility between high and low electric fields lead to significant benefits in reducing the inherent complexity of cell extracts and improving the depth of proteome coverage. Indeed, FAIMS not only reduces the proportion of singly-charged contaminating ions, but also separates overlapping multiply-charge peptide ions of similar *m/z* values that lead to chimeric MS/MS spectra. This results in the selection of precursor ions of higher purity and the detection of lower abundance peptide ions. To maximize the acquisition of MS/MS spectra, multiple CV steps can be combined in a single LC-FAIMS-MS/MS run. The CV stepping method can be easily modified depending on the desired depth of the proteome needed. We optimized the CV stepping method where the entire CV range can be covered with 2-3 injections, which provides a 30 % gain in the number of unique peptide identifications compared to traditional LC-MS/MS analyses for the same number of injections (***Figure 5-3*** and ***Figure 5-4***). It is noteworthy that parallel acquisition of ion trap MS/MS spectra available on the Orbitrap Tribrid instrument can improve the duty cycle and proteome coverage in LC-FAIMS-MS/MS experiments and can facilitate the incorporation of multiple CVs for single shot proteomics.

The reduced proportion of mixed precursor ions in FAIMS experiments also has important advantages for any type of quantitative proteomics workflows, and more specifically for isobaric labeling where both peptide sequencing and abundance measurements are obtained from the MS/MS spectra. Different strategies including SPS-MS^3 were devised to reduce reporter ions suppression and chimeric MS/MS spectra that undermine the precision and comprehensiveness of quantitative measurements using isobaric labeling. Here, we evaluated the benefits of LC-MS/MS analysis performed using SPS-MS^3 and FAIMS to profile the temporal changes in

protein abundances using TMT 10-plex reagents. These analyses indicated that FAIMS led to a 2.5-fold increase in the number of quantified peptides compared to SPS-MS^3. Since the differential ion mobility reduces the occurrence of co-selection and co-fragmentation, there is a substantial reduction in ratio compression with more accurate TMT quantification [44]. Since the TMT quantification can be performed at the MS^2 level with FAIMS, the number of scan events leading to peptide sequencing and quantification is higher than that achievable with SPS-MS^3 for the same sample amount and analysis time. In addition, the HCD spectra acquired in the Orbitrap increases the identification success rate for FAIMS compared to the CID ion trap spectra in SPS-MS^3.

The improved duty cycle imparted by FAIMS for TMT-based quantification over SPS-MS^3 helped to increase the coverage of the dynamic protein profiles upon heat shock (502 vs 149 for FAIMS and SPS, respectively). The added proteins that were quantified as a result of using FAIMS revealed how vast the effects of heat shock is on the cell by monitoring the dynamics of the heat shock response in greater detail, as evidenced in ***Figure 5-7a***. Indeed, several processes were altered in response to the stimulus and with varying kinetics. Moreover, the processes that were affected by the threat of the heat shock were varied and independent from each other, highlighting the global effect of the heat shock response. Cells subjected to heat stress are known to adapt to this challenge by altering several key cellular processes, which include but are not limited to: aggregation of ribosomal proteins, alterations in the RNA splicing machinery [67], disruption of the cytoskeleton structure, mitochondrial dysfunction [68], and increased production of heat shock proteins [69, 70]. In HEK293 cells, the order in which these events occur with regards to each other is not known. The dynamics of these events holds great value to better understand how cells can adapt to stress, since the heat shock response is not unique to mild heat stress but is also shared with adaptation to oxidative stress and exposure to heavy metals [71]. Indeed the thermotolerance that is imparted to cells by heat stress has been shown to increase cellular resistance to anticancer drugs and trophic factor withdrawal, indicating that the heat shock response is versatile and robust [71].

By following the changes in the proteome mediated by hyperthermia in a time resolved manner we expanded the comprehensiveness of the cellular response. We

found that proteins that shared similar functions had a tendency to alter in abundance with similar dynamics. Seminal work by the Savitski group, which followed protein turnover using a SILAC quantification methodology in several primary cells highlighted that proteins that make up complexes have similar half-lives, suggesting that complexes as a whole degrade together [72]. Our results are consistent with this earlier study, where the proteins that share the same GO term in the network not only possess similar functions but interact readily with one another, explaining their shared dynamics.

This work highlights several advantages of the new FAIMS interface with the most valuable being the ability of the electrodes to perform gas phase fractionation of the complex ion population from an ESI within a 40 ms period. An important outcome from such a fractionation mechanism was the enrichment of multiply charged species that facilitated a deeper sampling of precursors in the MS survey scan that could have otherwise remained un-sampled. Due to its novelty and equal or superior performance over condensed phase fractionation, FAIMS devices will gain wider acceptance as an important technology in the field of MS-based proteomics. There are possibilities to improve duty cycle usage by invoking the parallelization that is accessible on the Orbitrap Tribrid mass spectrometers. Ideally, the instrument duty cycle can be improved to cover the whole CV transmission range with a single injection. This would allow reduction in the instrument usage time by 2-3 fold. During the review of this manuscript, this application was demonstrated on the Orbitrap Tribrid LUMOS mass spectrometer, and enabled increased depth for single shot proteome analyses [73]. The application of FAIMS on the new Tribrid MS instruments in the field of proteomics is in infancy and has the potential to flourish in the coming years.

5.6. Acknowledgment

The authors wish to thank Philip Remes and Jesse D. Canterbury at Thermo Fisher Scientific in San Jose (CA) for technical assistance, and Susan Abbatiello for constructive criticism and valuable comments during the review of this manuscript. S.Pf is the recipient of a Swiss national science foundation scholarship (P1SKP3-168335). This work was carried out with financial support from the Natural Sciences and Engineering Research Council (NSERC 311598) and the Genomic Applications Partnership Program (GAPP) of Genome Canada. The Institute for Research in Immunology and Cancer (IRIC) receives infrastructure support from IRICoR, the Canadian Foundation for Innovation, and the Fonds de Recherche du Québec - Santé (FRQS). IRIC proteomics facility is a Genomics Technology platform funded in part by the Canadian Government through Genome Canada.

5.7. Note

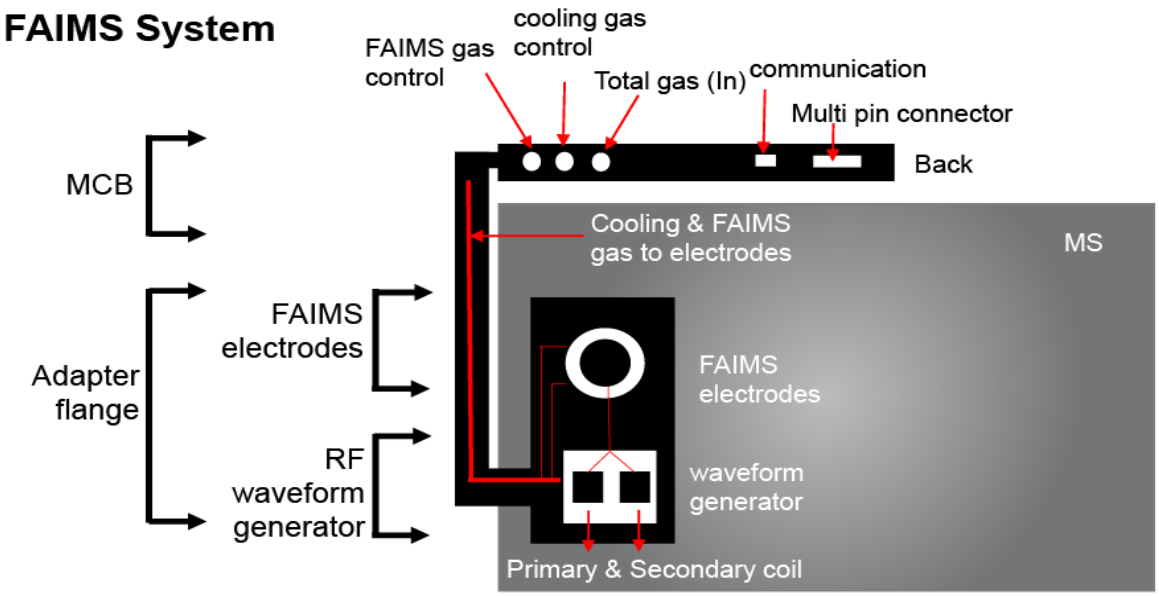
The authors declare the following competing financial interest(s): M.B., S.P., D.B., and J.J.D., are employees at Thermo Fisher Scientific, which develops and distributes MS instruments including FAIMS device.

The MS proteomics data have been deposited to the ProteomeXchange Consortium via the massIVE partner repository which can be accessed at the following location ftp://MSV000082293@massive.ucsd.edu with the following password grthibault.

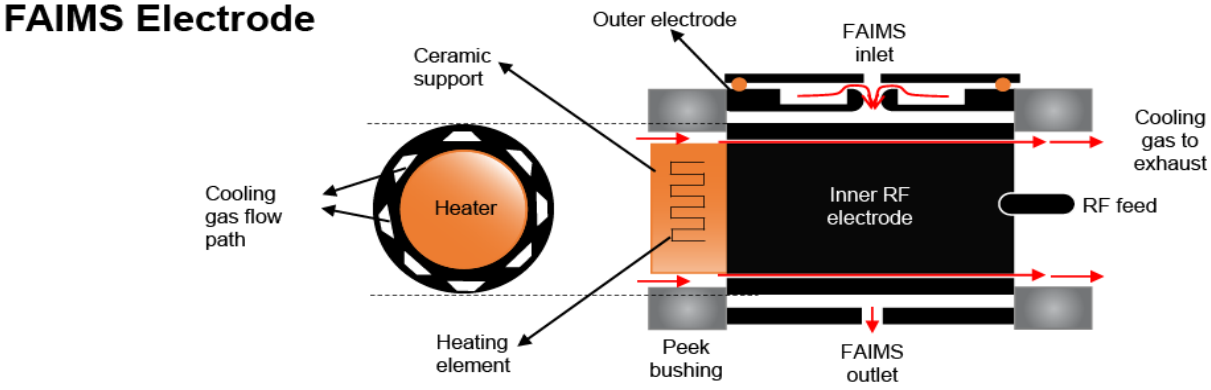
5.8. Supplementary material

5.8.1. Supplementary figures

a

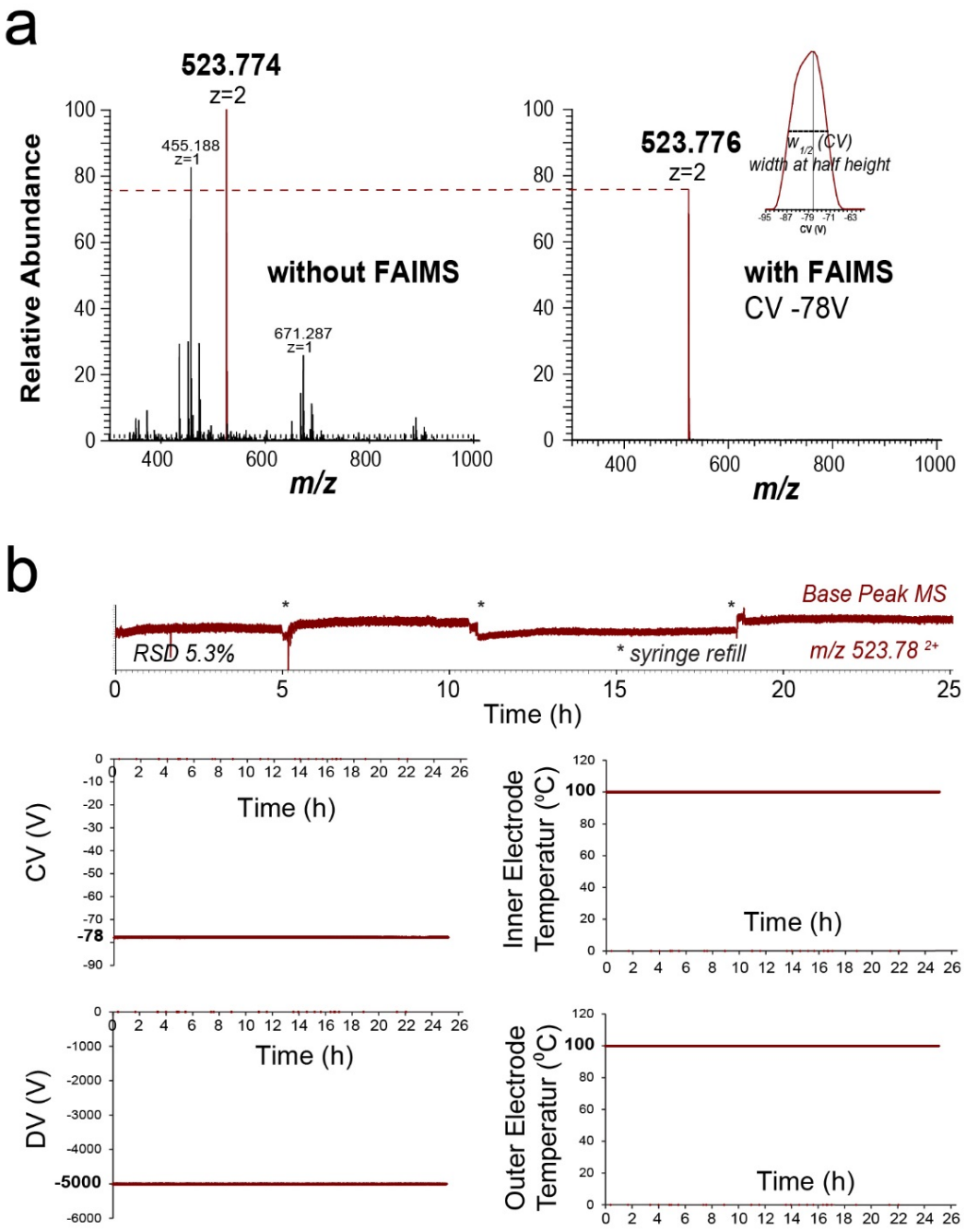


b

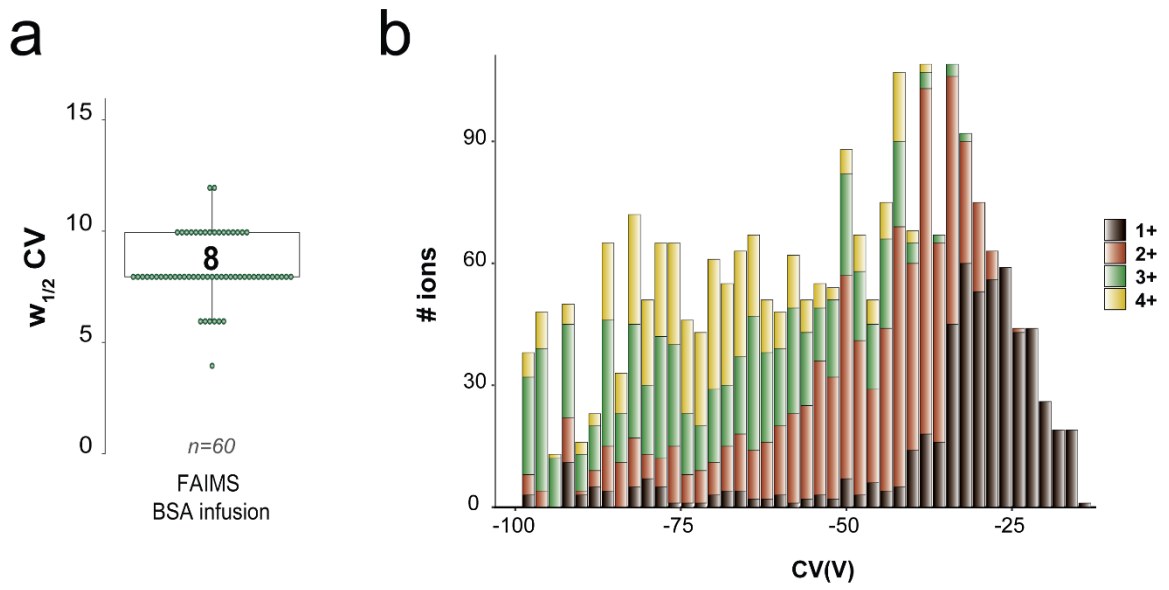


Supplementary Figure 5-1: *Schematic of the new FAIMS interface. a) FAIMS interface with Main Control Board (MCB) and an adapter flange that include a RF waveform generator and FAIMS electrodes. The FAIMS electrode sub assembly contain an inner electrode, outer electrode, entrance plate, and three heaters: one for inner electrode and two for outer electrode. The MCB provides multiple DC drive voltages, gas controls, FAIMS electrode heaters and*

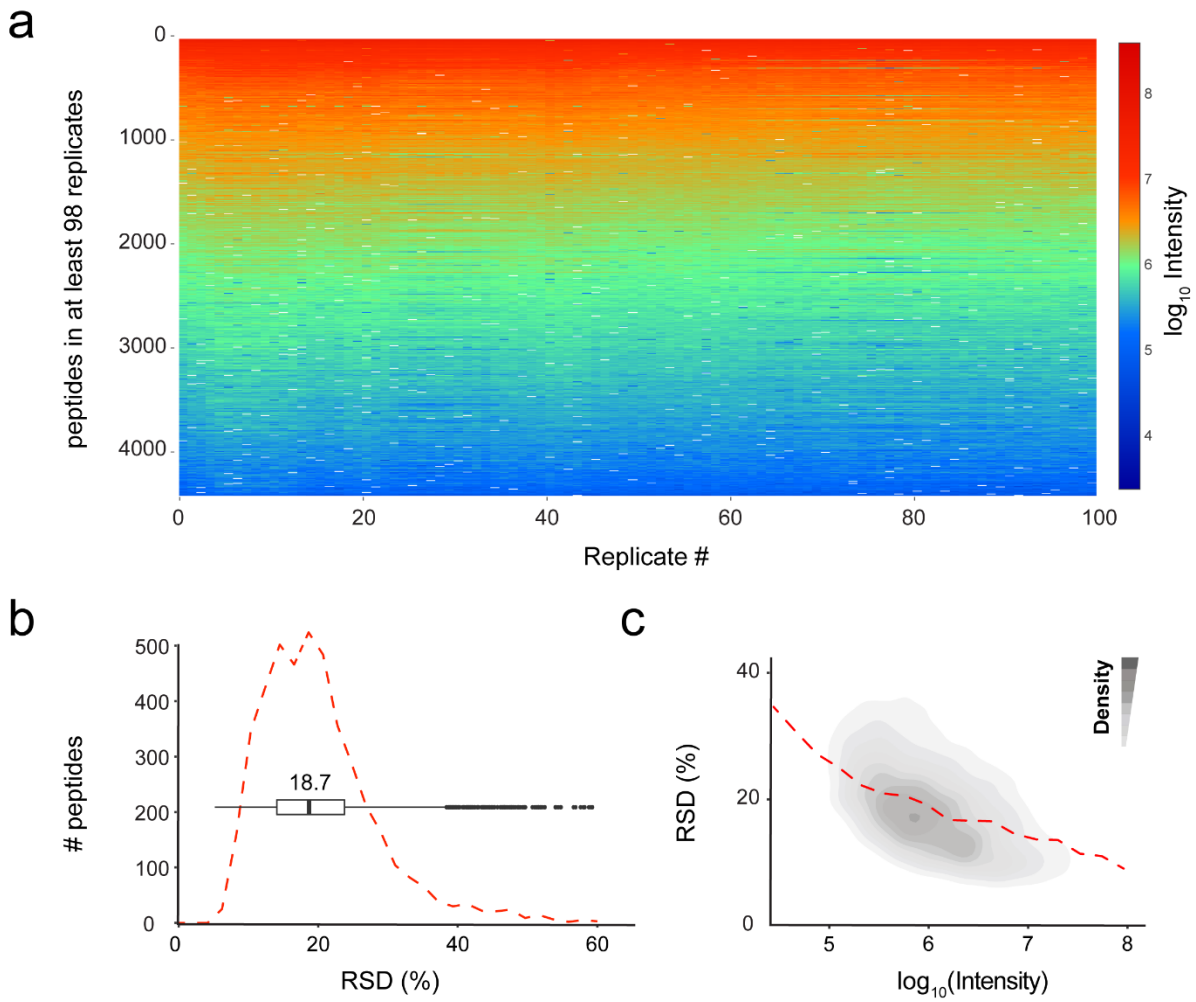
temperature sensors, and a separate supply drives the primary and secondary RF coils to produce +/-5000V ion separation voltage. Other DC voltages in the MCB include voltage for CV and Entrance Plate. An industrial grade N_2 supply is delivered into two separate gas valves used to supply the FAIMS ion transport gas and the electrodes' cooling gas. All DC voltages and two ¼" Teflon tubing gas lines (ion transport and cooling gas) are routed through a "side arm" cable structure. Gas and DC connections to the electrodes were concealed in the adapter flange eliminating the need for user intervention to connect/disconnect the DC voltage and gas lines when alternating between with and without FAIMS experiments. A single task of detaching the adapter flange from the MS restores the system to a "without FAIMS" state and a reverse step restores the MS system to a "with FAIMS" state within seconds. b) Cross section of the FAIMS electrode assembly. A heater block is pressed against one end of the inner electrode to promote conduction of heat transfer. The inner electrode can also be rapidly cooled by disabling the heater and forcing air flow through multiple narrow slits bored in the inner electrode. The cooling air flow is separated from the ion transport flow using peek bushings fitted with vacuum sealing O rings. The ion transport flow is delivered using a separate supply line into an air gap between the outer electrode and the entrance plate. The entire assembly in Supplementary Figure 5-1B is pressed into a cap shaped entrance plate that is easily retractable by releasing tool free latches and sliding the assembly from the adapter flange.



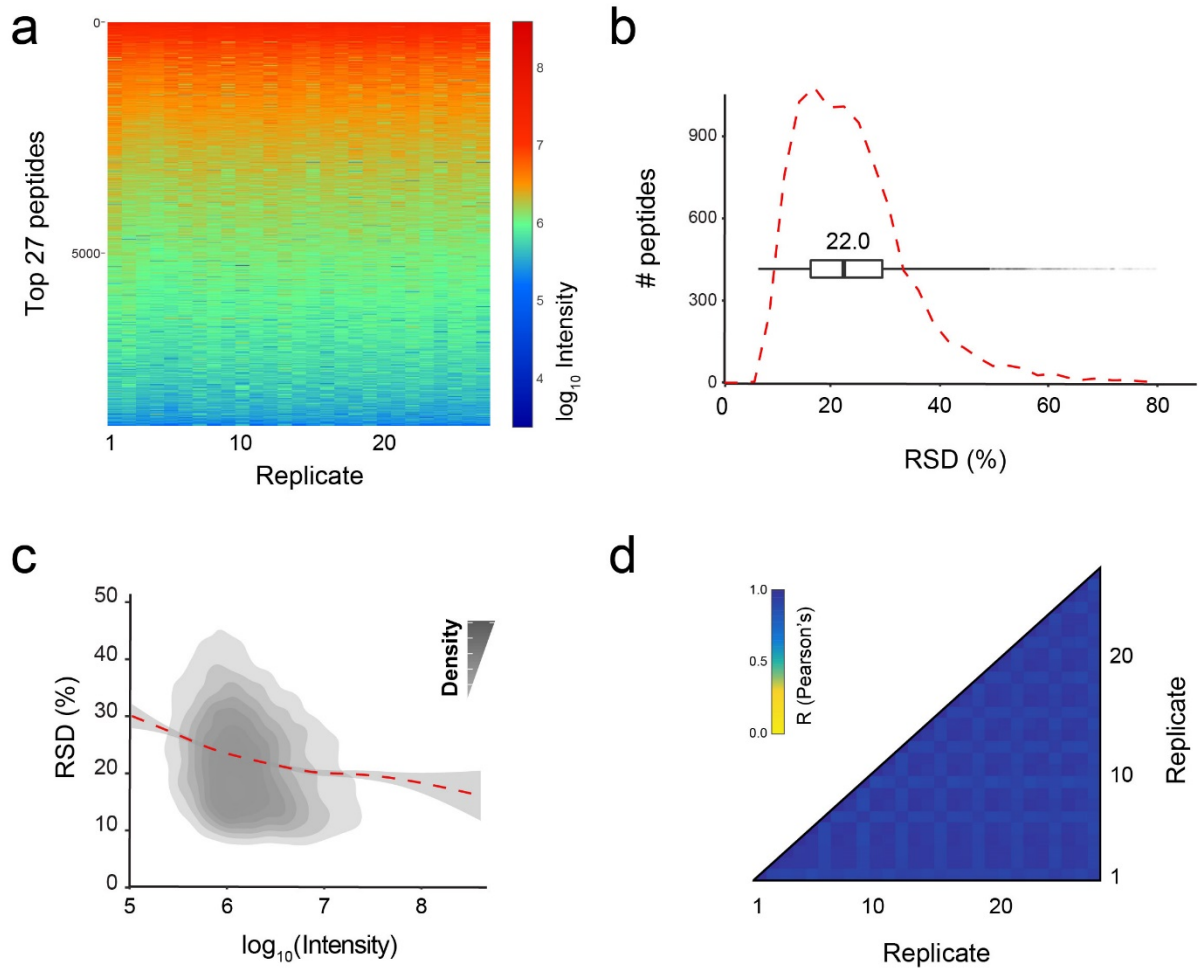
Supplementary Figure 5-2: *Ion transmission and stability of the new FAIMS prototype by direct infusion of Angiotensin II. (a) Angiotensin II peptide m/z 523.77^{2+} infusion with (left) regular ESI-MS and (right) ESI-FAIMS-MS at the optimal transmission CV of -78V, showing an ion transmission of almost 80 %. (b) Top, Angiotensin II infusion over 24 h with relative standard deviation of signal intensity of 5.3 %. Deviations in the CV voltage (-78V), DV voltage (-5000V) and both inner and outer electrode temperatures (100 ˚C) over the course of 24 h.*



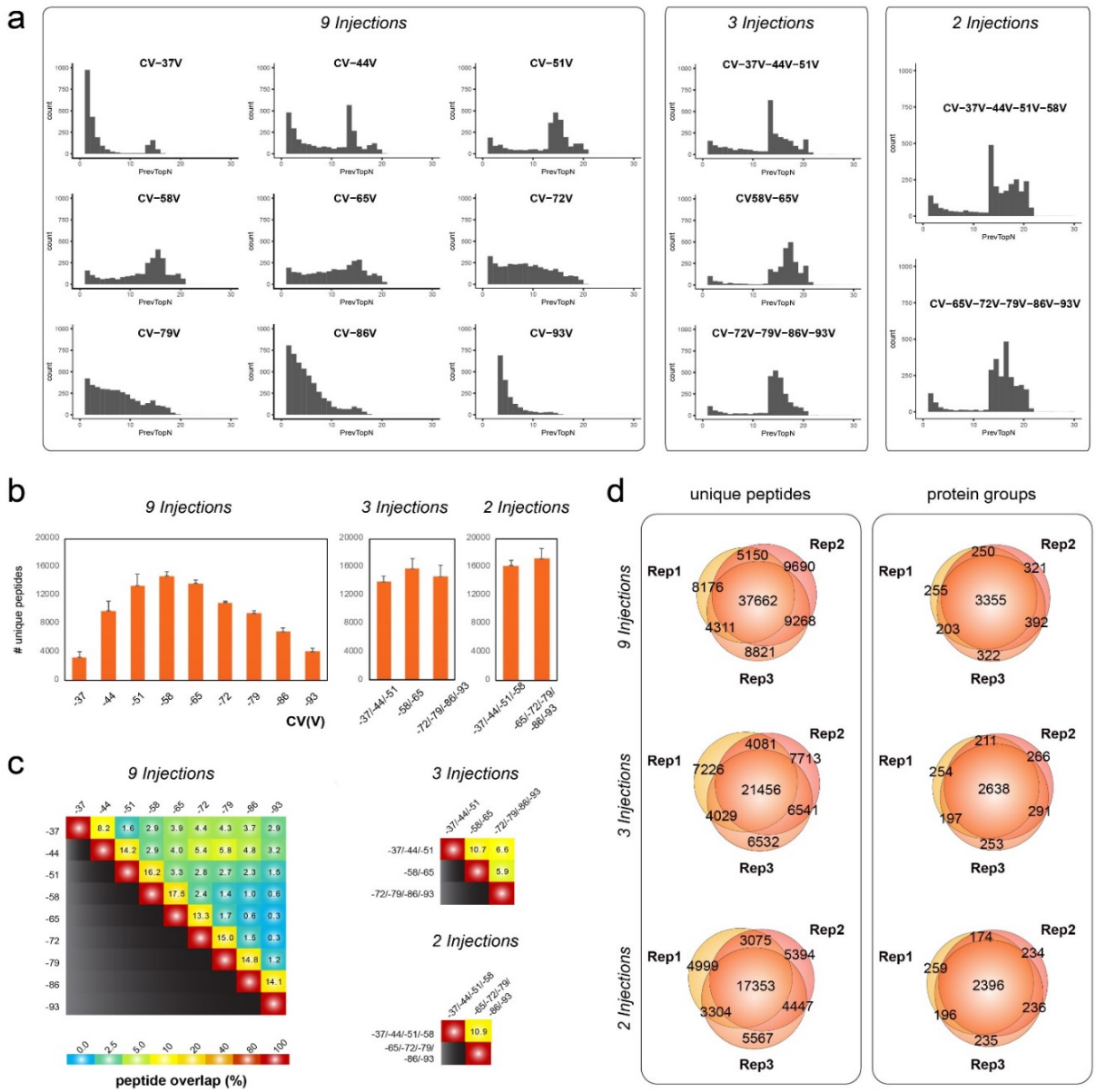
Supplementary Figure 5-3: *Resolution of the FAIMS interface using a BSA tryptic digest as benchmark. (a) Boxplot with data spread of the ions from a BSA tryptic digest over the FAIMS CV transmission range showing a median resolution (width at half height) of 8 V. Each green circle represents a different BSA tryptic peptide. (b) Distributions of ion charge states as a function of the FAIMS CV investigated.*



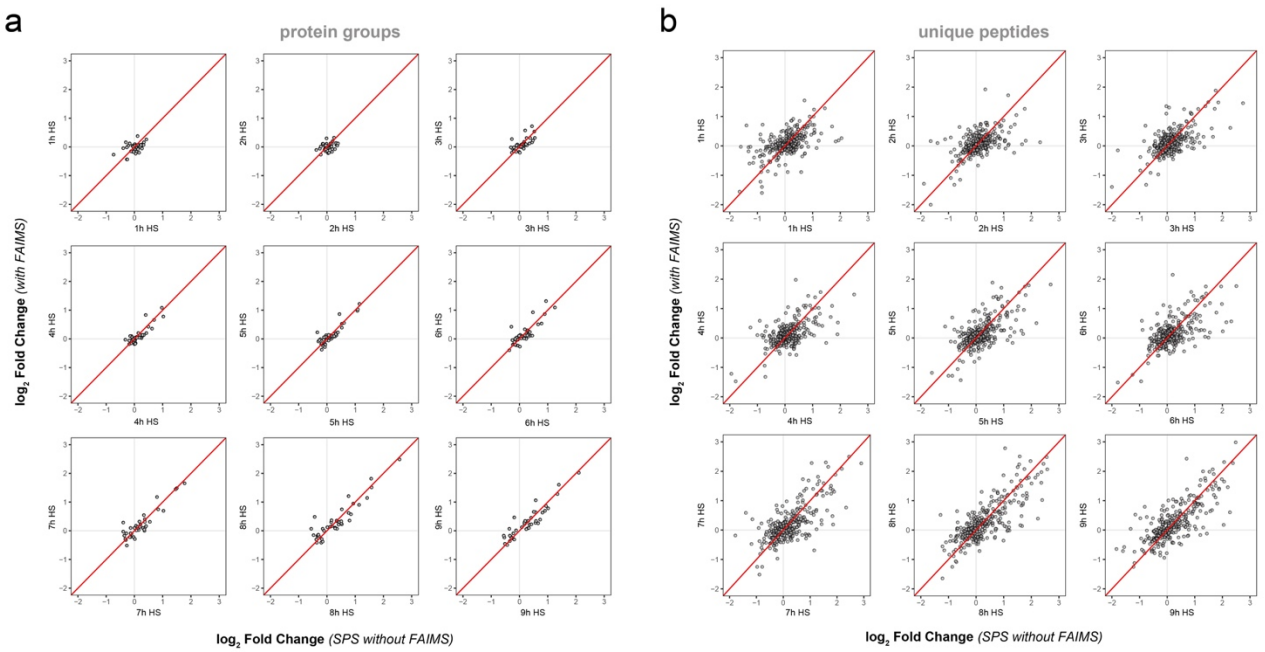
Supplementary Figure 5-4: *Reproducibility and robustness of FAIMS-Fusion LC-MS system. 100 replicate injections of 500 ng HEK293 digest at a fixed CV (CV -45 V) were continuously monitored over 5 days. (a) Heat map representation of the intensity of the 4,427 peptides that were present in 98 % of the injections across the 100 replicates. (b) Relative standard deviation distribution for the corresponding peptides. (c) Plot of the relative standard deviation as a function of ion intensity.*



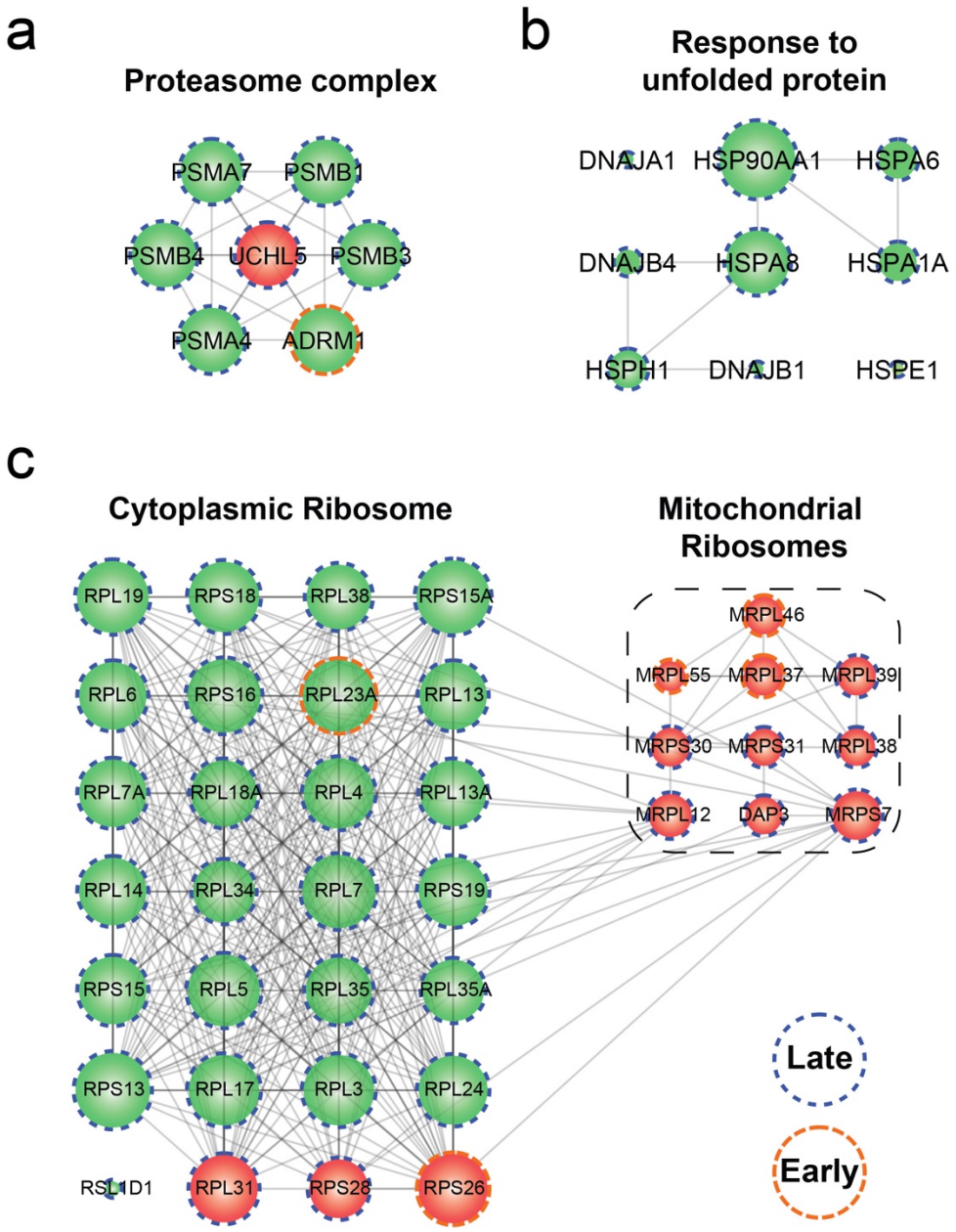
Supplementary Figure 5-5: *Reproducibility and robustness of regular LC-MS system with 27 replicate injections of HEK293 digest. (a) Heat map representation of the intensity of the 9,225 peptides (total of 42,511 peptides) that were present in all 27 replicate injections. (b) Relative standard deviation distribution for the common peptides from the 27 replicate injection with a median of 22.0 %. (c) Plot of RSD as a function of log_{10} Intensity for all common peptides. (d) Heat map for Pearson correlation coefficients between all 27 replicate injections.*



Supplementary Figure 5-6: *Optimization of FAIMS CV stepping program. HEK293 digests were analyzed with FAIMS with CV from -37 V to -93 V with 7 V steps to cover the entire peptide transmission range. (a) Duty cycle usage was monitored by following the number of MS^2 triggered per survey scan for (left) 9 injections at individual CV, (middle) 3 injections combining CV-37 V/-44 V/-51 V, CV-58 V/-65 V and CV-72 V/-79 V/-86 V/-93 V and (right) 2 injections combining CV-37 V/-44 V/-51 V/-58 V and CV-65 V/-72 V/-79 V/-86 V/-93 V. (b) Number of unique peptide sequences identified per injection for (left) individual CV, (middle) 3 injection combination and (right) 2 injection combination. (c) Heat map representation of the overlap in peptide identification between the various injections. (d) Overlap of three technical replicates for (top) individual CV, (middle) 3 injection combination and (bottom) 2 injection combination. The overlap for the number of (left) unique peptides and (right) protein groups (≥2 unique peptides/protein) are depicted in the Venn diagrams.*



Supplementary Figure 5-7: *Comparison between SPS and FAIMS based TMT quantifications. Scatterplot representations for the common dynamic (a) proteins and (b) peptides that changed in abundance during the heat shock response in HEK293 cells. Scatterplots correspond to each of the 9 time points investigated.*



Supplementary Figure 5-8: *The heat shock response affects proteins with similar functions in the same demeanor. Sub networks were extracted from the protein network shown in Figure 5-7a based on their enriched GO-term. Proteins outlined in blue display a delayed response (>1 h) and proteins outlined in orange displayed acute responses (<1 h). Red nodes depict proteins whose abundances decreased in response to the heat shock, while green nodes depict proteins whose abundances increased. Proteins from the same sub network changed abundance with similar dynamics. (a) Proteasome complex sub network. (b) Response to unfolded protein sub network. (c) Ribosome sub network.*

5.8.2. Supplementary tables

***Supplementary Table 5-1**: List of FAIMS LC-MS/MS experiments at CV -45V for 100 Replicates (CD-ROM).*

***Supplementary Table 5-2**: List of peptide identifications of LC-MS/MS experiments with and without FAIMS for 500ng HEK293 digest/injection (CD-ROM).*

***Supplementary Table 5-3**: List of protein identifications of LC-MS/MS experiments with and without FAIMS for 500ng HEK293 digest/injection (CD-ROM).*

***Supplementary Table 5-4**: List of quantified dynamic Protein groups assigned to a cluster with a membership 0.9. These proteins have at least 2 unique peptides and TMT reporters present in at least 7 TMT channels (CD-ROM).*

5.9. References

1. Beck, S., A. Michalski, O. Raether, M. Lubeck, S. Kaspar, N. Goedecke, C. Baessmann, D. Hornburg, F.P. Meier, I. Paron, et al., *The Impact 2 a very high resolution qTOF for deep shotgun proteomics.* Mol. Cell. Proteomics, 2015. **14**(7): p. 2014-2029.
2. Eliuk, S. and A. Makarov, *Evolution of the orbitrap mass spectrometry instrumentation.* Annu. Rev. Anal. Chem., 2015. **8**: p. 61-80.
3. Cox, J., M.Y. Hein, C.A. Luber, I. Paron, N. Nagaraj, and M. Mann, *Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ.* Mol Cell Proteomics, 2014. **13**(9): p. 2513-26.
4. Karp, N.A., W. Huber, P.G. Sadowski, P.D. Charles, S.V. Hester, and K.S. Lilley, *Addressing accuracy and precision issues in iTRAQ quantitation.* Mol Cell Proteomics, 2010. **9**(9): p. 1885-97.
5. Ow, S.Y., M. Salim, J. Noirel, C. Evans, I. Rehman, and P.C. Wright, *iTRAQ underestimation in simple and complex mixtures: "the good, the bad and the ugly".* J Proteome Res, 2009. **8**(11): p. 5347-55.
6. McAlister, G.C., E.L. Huttlin, W. Haas, L. Ting, M.P. Jedrychowski, J.C. Rogers, K. Kuhn, I. Pike, R.A. Grothe, J.D. Blethrow, et al., *Increasing the multiplexing capacity of TMTs using reporter ion isotopologues with isobaric masses.* Anal. Chem., 2012. **84**(17): p. 7469-7478.
7. McAlister, G.C., D.P. Nusinow, M.P. Jedrychowski, M. Wuhr, E.L. Huttlin, B.K. Erickson, R. Rad, W. Haas, and S.P. Gygi, *MultiNotch MS3 enables accurate, sensitive, and multiplexed detection of differential expression across cancer cell line proteomes.* Anal. Chem., 2014. **86**(14): p. 7150-8.
8. Bekker-Jensen, D.B., C.D. Kelstrup, T.S. Batth, S.C. Larsen, C. Haldrup, J.B. Bramsen, K.D. Sorensen, S. Hoyer, T.F. Orntoft, C.L. Andersen, et al., *An Optimized Shotgun Strategy for the Rapid Generation of Comprehensive Human Proteomes.* Cell Syst, 2017. **4**(6): p. 587-599 e4.
9. Kelstrup, C.D., R.R. Jersie-Christensen, T.S. Batth, T.N. Arrey, A. Kuehn, M. Kellmann, and J.V. Olsen, *Rapid and deep proteomes by faster sequencing on a benchtop quadrupole ultra-high-field Orbitrap mass spectrometer.* J Proteome Res, 2014. **13**(12): p. 6187-95.
10. Iwasaki, M., N. Sugiyama, N. Tanaka, and Y. Ishihama, *Human proteome analysis by using reversed-phase monolithic silica capillary column with enhanced sensitivity.* J. Chromatogr. A, 2012. **1228**: p. 292-297.
11. Nagaraj, N., N.A. Kulak, J. Cox, N. Neuhauser, K. Mayr, O. Hoerning, O. Vorm, and M. Mann, *System-wide perturbation analysis with nearly complete coverage of the yeast proteome by single-shot ultra HPLC runs on a bench top orbitrap.* Mol. Cell. Proteomics, 2012. **11**(3): p. 3722-3733.
12. Scheltema, R.A., J.P. Hauschild, O. Lange, D. Hornburg, E. Denisov, E. Damoc, A. Kuehn, A. Makarov, and M. Mann, *The Q Exactive HF, a benchtop mass spectrometer with a pre-filter, high-performance quadrupole and an ultra-high-field orbitrap analyzer.* Mol. Cell. Proteomics, 2014. **13**(12): p. 3698-3708.
13. Michalski, A., J. Cox, and M. Mann, *More than 100,000 detectable peptide species elute in single shotgun proteomics runs but the majority is inaccessible to data-dependent LC-MS/MS.* J Proteome Res, 2011. **10**(4): p. 1785-93.
14. Shishkova, E., A.S. Hebert, and J.J. Coon, *Now, More Than Ever, Proteomics Needs Better Chromatography.* Cell Syst, 2016. **3**(4): p. 321-324.

15. Hebert, A.S., C. Thoing, N.M. Riley, N.W. Kwiecien, E. Shiskova, R. Huguet, H.L. Cardasis, A. Khuen, S. Eliuk, V. Zabrouskov, et al., *Improved precursor characterization for data-dependent mass spectrometry.* Anal. Chem., 2017. **90**(3): p. 2323-2340.

16. O'Brien, J.J., J.D. O'Connell, J.A. Paulo, S. Thakurta, C.M. Rose, M.P. Weekes, E.L. Huttlin, and S.P. Gygi, *Compositional Proteomics: Effects of Spatial Constraints on Protein Quantification Utilizing Isobaric Tags.* J Proteome Res, 2018. **17**(1): p. 590-599.

17. Cumeras, R., E. Figueras, C.E. Davis, J.I. Baumbach, and I. Gracia, *Review on ion mobility spectrometry. Part 2: hyphenated methods and effects of experimental parameters.* Analyst, 2015. **140**(5): p. 1391-410.

18. Lapthorn, C., F. Pullen, and B.Z. Chowdhry, *Ion mobility spectrometry-mass spectrometry (IMS-MS) of small molecules: separating and assigning structures to ions.* Mass Spectrom Rev, 2013. **32**(1): p. 43-71.

19. Schneider, B.B., E.G. Nazarov, F. Londry, P. Vouros, and T.R. Covey, *Differential mobility spectrometry/mass spectrometry history, theory, design optimization, simulations, and applications.* Mass Spectrom Rev, 2016. **35**(6): p. 687-737.

20. Swearingen, K.E. and R.L. Moritz, *High-field asymmetric waveform ion mobility spectrometry for mass spectrometry-based proteomics.* Expert Rev Proteomics, 2012. **9**(5): p. 505-17.

21. Wyttenbach, T., N.A. Pierson, D.E. Clemmer, and M.T. Bowers, *Ion mobility analysis of molecular dynamics.* Annu Rev Phys Chem, 2014. **65**: p. 175-96.

22. Ruotolo, B.T., J.A. McLean, K.J. Killig, and D.H. Russell, *The influence and utility of varying field strength for the separation of tryptic peptides by ion mobility-mass spectrometry.* J. Am. Soc. Mass Spectrom., 2004. **16**: p. 158-165.

23. Harvey, S.R., C.E. McPhee, and P.E. Barran, *Ion mobility mass spectrometry for peptide analysis.* Methods, 2011. **54**(11): p. 454-461.

24. Guevremont, R., *High-field asymmetric waveform ion mobility spectrometry: a new tool for mass spectrometry.* J. Chromatogr. A, 2004. **1058**(1-2): p. 3-19.

25. Purves, R.W. and R. Guevremont, *Electrospray ionization high-field asymmetric waveform ion mobility spectrometry-mass spectrometry.* Anal. Chem., 1999. **71**(13): p. 2346-2357.

26. Barnett, D.A., B. Ells, R. Guevremont, and R.W. Purves, *Application of ESI-FAIMS-MS to the analysis of tryptic peptides.* J. Am. Soc. Mass Spectrom., 2002. **13**: p. 1282-1291.

27. Shvartsburg, A.A., K. Tang, and R.D. Smith, *Modeling the resolution and sensitivity of FAIMS analyses.* J. Am. Soc. Mass Spectrom., 2004. **15**(10): p. 1487-1498.

28. Bonneil, E., S. Pfammatter, and P. Thibault, *Enhancement of mass spectrometry performance for proteomic analyses using high-field asymmetric waveform ion mobility spectrometry (FAIMS).* J Mass Spectrom, 2015. **50**(11): p. 1181-95.

29. Canterbury, J.D., X. Yi, M.R. Hoopmann, and M.J. MacCoss, *Assessing the dynamic range and peak capacity of nanoflow LC-FAIMS-MS on an ion trap mass spectrometer for proteomics.* Anal Chem, 2008. **80**(18): p. 6888-97.

30. Creese, A.J., N.J. Shimwell, K.P. Larkins, J.K. Heath, and H.J. Cooper, *Probing the complementarity of FAIMS and strong cation exchange chromatography in shotgun proteomics.* J Am Soc Mass Spectrom, 2013. **24**(3): p. 431-43.

31. Saba, J., E. Bonneil, C. Pomies, K. Eng, and P. Thibault, *Enhanced sensitivity in proteomics experiments using FAIMS coupled with a hybrid linear ion trap/Orbitrap mass spectrometer.* J Proteome Res, 2009. **8**(7): p. 3355-66.

32. Swearingen, K.E., M.R. Hoopmann, R.S. Johnson, R.A. Saleem, J.D. Aitchison, and R.L. Moritz, *Nanospray FAIMS fractionation provides significant increases in proteome coverage of unfractionated complex protein digests.* Mol Cell Proteomics, 2012. **11**(4): p. M111 014985.

33. Venne, K., E. Bonneil, K. Eng, and P. Thibault, *Improvement in peptide detection for proteomics analyses using NanoLC-MS and high-field asymmetry waveform ion mobility mass spectrometry.* Anal Chem, 2005. **77**(7): p. 2176-86.

34. Baird, M.A. and A.A. Shvartsburg, *Localization of Post-Translational Modifications in Peptide Mixtures via High-Resolution Differential Ion Mobility Separations Followed by Electron Transfer Dissociation.* J Am Soc Mass Spectrom, 2016. **27**(12): p. 2064-2070.

35. Garabedian, A., M.A. Baird, J. Porter, K. Jeanne Dit Fouque, P.V. Shliaha, O.N. Jensen, T.D. Williams, F. Fernandez-Lima, and A.A. Shvartsburg, *Linear and Differential Ion Mobility Separations of Middle-Down Proteoforms.* Anal Chem, 2018. **90**(4): p. 2918-2925.

36. Kaszycki, J.L., A.P. Bowman, and A.A. Shvartsburg, *Ion Mobility Separation of Peptide Isotopomers.* J Am Soc Mass Spectrom, 2016. **27**(5): p. 795-9.

37. Shliaha, P.V., M.A. Baird, M.M. Nielsen, V. Gorshkov, A.P. Bowman, J.L. Kaszycki, O.N. Jensen, and A.A. Shvartsburg, *Characterization of Complete Histone Tail Proteoforms Using Differential Ion Mobility Spectrometry.* Anal Chem, 2017. **89**(10): p. 5461-5466.

38. Shvartsburg, A.A., A.J. Creese, R.D. Smith, and H.J. Cooper, *Separation of peptide isomers with variant modified sites by high-resolution differential ion mobility spectrometry.* Anal Chem, 2010. **82**(19): p. 8327-34.

39. Shvartsburg, A.A., A. Haris, R. Andrzejewski, A. Entwistle, and R. Giles, *Differential Ion Mobility Separations in the Low-Pressure Regime.* Anal Chem, 2018. **90**(1): p. 936-943.

40. Shvartsburg, A.A., T.A. Seim, W.F. Danielson, R. Norheim, R.J. Moore, G.A. Anderson, and R.D. Smith, *High-definition differential ion mobility spectrometry with resolving power up to 500.* J Am Soc Mass Spectrom, 2013. **24**(1): p. 109-14.

41. Shvartsburg, A.A., D. Singer, R.D. Smith, and R. Hoffmann, *Ion mobility separation of isomeric phosphopeptides from a protein with variant modification of adjacent residues.* Anal Chem, 2011. **83**(13): p. 5078-85.

42. Shvartsburg, A.A., Y. Zheng, R.D. Smith, and N.L. Kelleher, *Separation of variant methylated histone tails by differential ion mobility.* Anal Chem, 2012. **84**(15): p. 6317-20.

43. Shvartsburg, A.A., Y. Zheng, R.D. Smith, and N.L. Kelleher, *Ion mobility separation of variant histone tails extending to the "middle-down" range.* Anal Chem, 2012. **84**(10): p. 4271-6.

44. Pfammatter, S., E. Bonneil, and P. Thibault, *Improvement of Quantitative Measurements in Multiplex Proteomics Using High-Field Asymmetric Waveform Spectrometry.* J Proteome Res, 2016. **15**(12): p. 4653-4665.

45. Keshishian, H., M. Burgess, H. Specht, L. Wallace, K. Clauser, M. Gillette, and S. Carr, *Quantitative, multiplexed workflow for deep analysis of human blood plasma and biomarker discovery by mass spectrometry.* Nat. Protoc., 2017. **12**(8): p. 1683-1701.

46. Prasad, S., M.W. Belford, J.J. Dunyach, and R.W. Purves, *On an aerodynamic mechanism to enhance ion transmission and sensitivity of FAIMS for nano-electrospray ionization-mass spectrometry.* J. Am. Soc. Mass Spectrom., 2014. **25**(12): p. 2143-2153.

47. Schwammie, V. and O.N. Jensen, *A simple and fast method to determine the parameters for fuzzy c‐means cluster analysis.* Bioinformatics, 2010. **26**: p. 2841-2848.

48. Franceschini, A., D. Szklarczyk, S. Frankild, M. Kuhn, M. Simonovic, A. Roth, J. Lin, P. Minguez, P. Bork, C. von Mering, et al., *STRING v9.1: protein-protein interaction networks, with increased coverage and integration.* Nucleic Acids Res, 2013. **41**(Database issue): p. D808-15.

49. Smoot, M.E., K. Ono, J. Ruscheinski, P.L. Wang, and T. Ideker, *Cytoscape 2.8: new features for data integration and network visualization.* Bioinformatics, 2011. **27**(3): p. 431-2.

50. Maere, S., K. Heymans, and M. Kuiper, *BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks.* Bioinformatics, 2005. **21**(16): p. 3448-9.

51. Kubiniok, P., H. Lavoie, M. Therrien, and P. Thibault, *Time-resolved Phosphoproteome Analysis of Paradoxical RAF Activation Reveals Novel Targets of ERK.* Mol Cell Proteomics, 2017. **16**(4): p. 663-679.

52. Nock, R. and F. Nielsen, *On weighting clustering.* IEEE Trans Pattern Anal Mach Intell, 2006. **28**(8): p. 1223-35.

53. McAlister, G.C., E.L. Huttlin, W. Haas, L. Ting, M.P. Jedrychowski, J.C. Rogers, K. Kuhn, I. Pike, R.A. Grothe, J.D. Blethrow, et al., *Increasing the multiplexing capacity of TMTs using reporter ion isotopologues with isobaric masses.* Anal Chem, 2012. **84**(17): p. 7469-78.

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

55. Brown, S.A., A.N. Imbalzano, and R.E. Kingston, *Activator-dependent regulation of transcriptional pausing on nucleosomal templates.* Genes Dev, 1996. **10**(12): p. 1479-90.

56. de Nadal, E., G. Ammerer, and F. Posas, *Controlling gene expression in response to stress.* Nat Rev Genet, 2011. **12**(12): p. 833-45.

57. Anckar, J. and L. Sistonen, *Regulation of HSF1 function in the heat stress response: implications in aging and disease.* Annu Rev Biochem, 2011. **80**: p. 1089-115.

58. Zou, J., Y. Guo, T. Guettouche, D.F. Smith, and R. Voellmy, *Repression of heat shock transcription factor HSF1 activation by HSP90 (HSP90 complex) that forms a stress-sensitive complex with HSF1.* Cell, 1998. **94**(4): p. 471-80.

59. Patriarca, E.J. and B. Maresca, *Acquired thermotolerance following heat shock protein synthesis prevents impairment of mitochondrial ATPase activity at elevated temperatures in Saccharomyces cerevisiae.* Exp Cell Res, 1990. **190**(1): p. 57-64.

60. Kikusato, M., H. Yoshida, K. Furukawa, and M. Toyomizu, *Effect of heat stress-induced production of mitochondrial reactive oxygen species on NADPH oxidase and heme oxygenase-1 mRNA levels in avian muscle cells.* J Therm Biol, 2015. **52**: p. 8-13.

61. Yao, T., L. Song, W. Xu, G.N. DeMartino, L. Florens, S.K. Swanson, M.P. Washburn, R.C. Conaway, J.W. Conaway, and R.E. Cohen, *Proteasome recruitment and activation of the Uch37 deubiquitinating enzyme by Adrm1.* Nat Cell Biol, 2006. **8**(9): p. 994-1002.

62. Bond, U., *Heat shock but not other stress inducers leads to the disruption of a sub-set of snRNPs and inhibition of in vitro splicing in HeLa cells.* EMBO J, 1988. **7**(11): p. 3509-18.

63. Ferretti, M.B., H. Ghalei, E.A. Ward, E.L. Potts, and K. Karbstein, *Rps26 directs mRNA-specific translation by recognition of Kozak sequence elements.* Nat Struct Mol Biol, 2017. **24**(9): p. 700-707.

64. Barnett, D.A. and R.J. Ouellette, *Elimination of the helium requirement in high-field asymmetric waveform ion mobility spectrometry (FAIMS): beneficial effects of decreasing the analyzer gap width on peptide analysis.* Rapid Commun Mass Spectrom, 2011. **25**(14): p. 1959-71.

65. Prasad, S., M.W. Belford, J.J. Dunyach, and R.W. Purves, *On an aerodynamic mechanism to enhance ion transmission and sensitivity of FAIMS for nano-electrospray ionization-mass spectrometry.* J Am Soc Mass Spectrom, 2014. **25**(12): p. 2143-53.

66. Swearingen, K.E., J.M. Winget, M.R. Hoopmann, U. Kusebauch, and R.L. Moritz, *Decreased Gap Width in a Cylindrical High-Field Asymmetric Waveform Ion Mobility Spectrometry Device Improves Protein Discovery.* Anal Chem, 2015. **87**(24): p. 12230-7.

67. Shukla, R.R., Z. Dominski, T. Zwierzynski, and R. Kole, *Inactivation of splicing factors in HeLa cells subjected to heat shock.* J Biol Chem, 1990. **265**(33): p. 20377-83.

68. Welch, W.J. and J.P. Suhan, *Morphological study of the mammalian stress response: characterization of changes in cytoplasmic organelles, cytoskeleton, and nucleoli, and appearance of intranuclear actin filaments in rat fibroblasts after heat-shock treatment.* J Cell Biol, 1985. **101**(4): p. 1198-211.

69. Richter, K., M. Haslbeck, and J. Buchner, *The heat shock response: life on the verge of death.* Mol Cell, 2010. **40**(2): p. 253-66.

70. Tabuchi, Y., I. Takasaki, S. Wada, Q.L. Zhao, T. Hori, T. Nomura, K. Ohtsuka, and T. Kondo, *Genes and genetic networks responsive to mild hyperthermia in human lymphoma U937 cells.* Int J Hyperthermia, 2008. **24**(8): p. 613-22.

71. Fulda, S., A.M. Gorman, O. Hori, and A. Samali, *Cellular stress responses: cell survival and cell death.* Int J Cell Biol, 2010. **2010**: p. 214074.

72. Mathieson, T., H. Franken, J. Kosinski, N. Kurzawa, N. Zinn, G. Sweetman, D. Poeckel, V.S. Ratnu, M. Schramm, I. Becher, et al., *Systematic analysis of protein turnover in primary cells.* Nat Commun, 2018. **9**(1): p. 689.

73. Hebert, A.S., S. Prasad, M.W. Belford, D.J. Bailey, G.C. McAlister, S.E. Abbatiello, R. Huguet, E.R. Wouters, J.J. Dunyach, D.R. Brademan, et al., *Rapid and Comprehensive Single Shot Proteomics with LC-FAIMS-MS/MS.* Anal Chem, 2018.

CHAPTER SIX

6. Conclusion and Perspectives

6.1. Conclusion

This thesis aimed at deciphering the analytical impact of FAIMS in large scale proteomic experiments. To this end, we monitored changes in the human proteome in response to hyperthermia and compared results stemming from standard LC-MS workflows to LC-FAIMS-MS. FAIMS improved the quantitation of the large scale experiments for label free, TMT and SILAC approaches by eliminating interfering ions that skew the results when using the standard workflows. A new FAIMS device that can be installed on new state-of-the-art Orbitrap instruments was evaluated. Indeed, this new instrument was robust, reproducible and much faster than its predecessor and holds much potential for the future of proteomics.

6.1.1. FAIMS reduces co-fragmentation and enhances identifications and quantification with isobaric labeling

The high complexity of proteomics samples hampers their analysis by MS. In ***CHAPTER TWO****,* the goal was to show the merits of FAIMS for TMT labeling, where co-isolation and co-fragmentation negatively impact the quantification. Since FAIMS post-ionization fractionation reduces spectral complexity, we started first by widening the isolation window from smaller windows used in DDA up to DIA-like mode windows. FAIMS always outperformed conventional LC-MS setups with higher PIF averages (between 30% to 120% higher). Little co-fragmentation occurred with isolation windows below 1.0 Th with FAIMS, where only 10% of the precursors were contaminated with interfering ions (compared to 30% without FAIMS). Regular LC-MS as well as LC-FAIMS-MS achieved the highest number of identifications with the isolation window of 2.0 Th, since sufficient isotope patterns were included for fragmentation. Interestingly, FAIMS reduced the proportion of chimeric scans with only 25% of scans having an impurity, compared to 50% without FAIMS, which corresponds to the previously reported proportion of chimeric scans in regular LC-MS [1]. In addition, the higher spectral quality afforded with FAIMS lead to a gain of over 50% more identifications, which was consistent for all isolation windows tested. This higher identification rate

expanded the dynamic range, where more low abundant peptides were accessible. In addition, performing CV fractions instead of replicate injections (as is the case with regular LC-MS) prevented resequencing of the same peptides in DDA mode [2]. The lower co-fragmentation observed with FAIMS significantly benefits accuracy and precision for TMT based quantification. Indeed, in regular MS^2 mode FAIMS drastically reduced ratio compression and interferences. LC-FAIMS-MS was employed to expand our understanding of the temporal changes that occur to the global proteome when cells are exposed to elevated temperatures. Proteins could be classified according their temporal behavior after heat shock to obtain early and late up/down regulations. In ***CHAPTER FIVE*** a similar TMT quantification with FAIMS-MS^2 was benchmarked against SPS-MS^3 for 1D-LC as well as 2D-LC setups. The FAIMS-MS^2 approach was clearly faster and more efficient since the added ion filtering provided by FAIMS eliminates the need for MS^3 scans, providing shorter duty cycles than the SPS-MS^3 method. Since the quantification and the identification are obtained in a single high-resolution scan in FAIMS-MS^2, the quality of the identification scan (MS^2) is improved since it is acquired in the Orbitrap, unlike the SPS-MS^3 approach.

6.1.2. FAIMS improves SILAC based quantification by reducing confounding ions

The second aim of this work was to evaluate the value of FAIMS for improving quantification using metabolic labeling. In metabolic labeling based quantitation, multiplexing samples doubles or triples the number of peptides present in the sample pool and in the MS survey scan. Improved peak capacity is essential for adequate coverage of the proteome, which generally relies on a second chromatographic separation. Therefore, we used SCX and combined it directly with FAIMS gas phase fractionation. In ***CHAPTER THREE***, by varying SILAC mixing ratios for tryptic human peptides over one order of magnitude FAIMS showed proper fold changes in accordance to the mixing ratios, whereas without FAIMS only small fold changes resulted in accurate and precise quantitation. Once again, the low abundance peptides

greatly benefited from FAIMS by enhancing the precision and accuracy of the analysis through sample decomplexification that lead to less confounding ions.

6.1.3. Gas phase enrichment of SUMOylated peptide ions with FAIMS

In ***CHAPTER FOUR*** the merits of FAIMS for the analysis of protein SUMOylation were investigated. FAIMS was able to separate the branched, mainly triply charged SUMO peptides from the rest of regular tryptic peptides that are typically doubly charged. FAIMS charge state segregations also showed different ion mobilities for the branched SUMOylated peptides compared to linear peptides. At more negative CV values, SUMOylated peptides could be selectively enriched. By including FAIMS in the SUMO proteome analysis, we showed for the first time a new approach to enrich this low abundant and dynamic modification in the gas phase. FAIMS was used to monitor the changes in protein SUMOylation in response to heat stress, which showed that globally protein SUMOylation increased during hyperthermia. In addition, gas phase fractionation using FAIMS provided greater reproducibility than other 2D-LC prefractionation methods that are typically used in proteomics. Also, FAIMS increased the sensitivity of the analysis by one order of magnitude due to a reduction in background ions and an improved signal-to-noise.

6.1.4. A new FAIMS device with improved speed and robustness

In ***CHAPTER FIVE***, we benchmarked the new FAIMS device that can be mounted on the new generation of Orbitrap instruments. The new FAIMS increases the depth of proteome coverage by roughly one order of magnitude over conventional LC-MS workflows. Three regular LC-MS injections with FAIMS provided a depth of proteome that spans 4 orders of magnitude in quantification. This increase is in part due to the decomplexification of the MS data, which provides improved survey and MS^2 scans. The major advantage with the new FAIMS over its predecessor is its vastly improved ion transmission time that is imparted by the smaller electrode gap and higher

DV frequency. LC-FAIMS-MS provided similar gains in identification as other fractionation methods [3], where improved coverage is generally obtained by injecting the various fractions on the MS instrument, which requires extended MS time. Moreover, LC-FAIMS-MS requires less sample than traditional offline fractionation [4]. The improved ion transmission with the new FAIMS and the coupling to the Fusion platform significantly improved the analysis. The enhanced ion transmission with the new FAIMS is especially beneficial on the newer Orbitrap instruments with higher scanning rates [5]. An added benefit of using FAIMS is the reduced instrument downtime that is needed to clean the instrument since FAIMS acts as a filter that inhibits contaminants from entering the MS orifice.

Overall, FAIMS improves peak capacity, reduces noise from contaminating ions, extends the dynamic range of the proteome and provides reproducible and automated post-ionization fractionation. With the new FAIMS platform, one can finally use this ion mobility device in a daily manner for MS analyses without suffering from substantial transmission loss and MS speed limitation as in the past.

6.2. Perspectives

We have just begun to understand the potential of the new FAIMS platform since it has been on the market for a relative short time [6-8]. It is already clear that FAIMS shows great performance for TMT quantifications [7, 8]. This could be promising for applications involving single cell proteomics [9], where sensitivity and reduced noise clearly improved the quality of the data. In addition, lower spectral complexity imparted by FAIMS facilitates charge state assignment and reduces precursor co-fragmentation that can compromise quantitative proteomics using metabolic and isobaric labeling which cannot be addressed with APD algorithm [10].

6.2.1. Short gradients with single CV fractions

Using FAIMS with short LC gradients may hold some promise. To achieve this, a short LC-MS gradient run is performed for every single CV under investigation, where

no CV stepping occurs during the injection. This method can be seen as analogous to fractionation methods where LC-FAIMS-MS is performed for each individual CV [3]. Since FAIMS can greatly decomplexify the samples, short LC gradients may be sufficient to obtain a reasonable coverage of the proteome. The duty cycles will be more efficient since there would be no CV switching mid-run, which can consume a substantial amount of time. Faster duty cycles correlate with more identifications as well as better elution profiles since more survey scans can be acquired for the precursor ions. Evidently, single CVs would also increase reproducibility between CV replicates, compared to CV switching injections. However, such a workflow would demand more material than the CV switching approach [6] because reduction/splitting of sample could reduce the overall identifications and would lower the ultimate gain. Since FAIMS favors the identification of low abundance peptides compared to other pre-fractionation workflows [6], the short-gradient with single-CV fractions may still demand less material than alternative fractionation methods. This is especially important because it means less sample processing. The electrode temperature could be optimized to vary the FAIMS resolution [11] to either increase or decrease the number of CV fractions required, in accordance with sample complexity.

6.2.2. Improving DIA quantification and identification with FAIMS

Although DIA [12] provides a greater coverage of the proteome and increases the reproducibility for peptide ions across runs, the dense ion population that makes up the large MS^2 isolation windows makes the data analysis more challenging. FAIMS could be used in DIA applications to reduce the level of contaminant ions as well as to decomplexify the MS^2 spectra due to the increased peak capacity, providing an improved identification and quantification. Reducing the ion population for DIA spectra, which we showed in ***CHAPTER TWO*** with the PIF values [13], helps the bioinformatics search. Furthermore, it may also allow for wider isolation windows to be used, increasing the efficiency of the method.

6.2.3. Expanding the depth of the proteome with sequential narrow survey scans and FAIMS

Instead of using libraries to match precursors, as is the case with the BoxCar approach that was described by Mathias Mann's group [14], one could simply inject narrow sequential *m/z* windows for the survey scans followed by a DDA approach for MS^2 based identification. FAIMS could enhance this sequential isolation window workflow by providing better ion transmission for low abundant ions with enhanced signal-to-noise, thus reducing the proportion of chimeric scans to increase quantification accuracy while at the same time expanding the comprehensiveness of the proteome coverage. A major advantage of such a method is its ability to obtain PSMs for each injection, while other methods have a tendency to rely on matching features between injections, which can often lead to erroneous quantification. Preliminary tests on the Q-Exactive™ platform showed that FAIMS could reduce the number of chimeric scans by more than half, which translated into a better proteome coverage compared with regular LC-MS.

6.2.4. FAIMS for the identification of peptides from non-coding regions

As MS analyses are become more and more sensitive, researchers have begun looking into peptides that stem from non-coding regions of the genome [15]. These peptides are expressed at very low levels and require enhanced sensitivity and high quality spectra with sufficient diagnostic peaks for proper assignment. High quality spectra are especially important because the corresponding databases are exponentially larger than typical databases, increasing the score threshold for a set FDR [16]. FAIMS differential ion mobility could aid in the discovery of the non-coding proteome. In the same vein, FAIMS technology should aid in detecting errors in protein synthesis [17] or other sequential mutations.

6.2.5. Charge state separation for highly charged cross-linked peptides

Chemical cross-linking based MS analysis has become a popular choice to study protein-protein interactions [18]. Cross-linking based MS analysis not only provides definite proof of a protein-protein interaction but also the location of the interaction interface. After protein digestion of cross-linked interacting proteins, two peptides are tethered through a linker, forming branched species. These peptides possess higher charge states due to the two N-termini and two C-terminal lysine/arginine residues. For this reason these peptides are often enriched by pre-fractionating the samples with SCX or other charge based fractionation methodologies [19]. Analogous to the multiply-charged and branched SUMOylated peptides, there is a great potential for FAIMS identification of cross-linked peptides where selectively transmitting cross-linked peptides, while removing linear tryptic peptides by optimizing the CV range, could enhance the analysis. Preliminary results from our lab suggest that cross-linked peptides are transmitted at more negative CV regions, similar to what was already observed for SUMO [20]. Similar results were also reported at this year's ASMS conference by our collaborator Thermo Fisher Scientific.

6.2.6. Machine learning for targeted analysis with FAIMS

Currently, the correct CV for a selected peptide ion must be determined empirically. In many cases this is not a problem, but for low amounts of material this may be problematic. Indeed, determining the proper CV for a set peptide can be time- and material-consuming. Although the collision cross section of a peptide cannot currently be predicted accurately through mathematical models, there is a clear correlation between charge, *m/z* and the CV [21]. As more and more data using FAIMS is accumulated with time, machine learning approaches will probably be capable of predicting CVs for peptide ions. This will increase the efficiency of FAIMS based targeted proteomics.

6.3. References

1. Houel, S., R. Abernathy, K. Renganathan, K. Meyer-Arendt, N.G. Ahn, and W.M. Old, *Quantifying the Impact of Chimera MS/MS Spectra on Peptide Identification in Large-Scale Proteomics Studies.* Journal of Proteome Research, 2010. **9**(8): p. 4152-4160.
2. Canterbury, J.D., X. Yi, M.R. Hoopmann, and M.J. MacCoss, *Assessing the dynamic range and peak capacity of nanoflow LC-FAIMS-MS on an ion trap mass spectrometer for proteomics.* Anal Chem, 2008. **80**(18): p. 6888-97.
3. Bekker-Jensen, D.B., C.D. Kelstrup, T.S. Batth, S.C. Larsen, C. Haldrup, J.B. Bramsen, K.D. Sørensen, S. Høyer, T.F. Ørntoft, C.L. Andersen, et al., *An Optimized Shotgun Strategy for the Rapid Generation of Comprehensive Human Proteomes.* Cell Systems, 2017. **4**(6): p. 587-599.e4.
4. Spicer, V., P. Ezzati, H. Neustaeter, R.C. Beavis, J.A. Wilkins, and O.V. Krokhin, *3D HPLC-MS with Reversed-Phase Separation Functionality in All Three Dimensions for Large-Scale Bottom-Up Proteomics and Peptide Retention Data Collection.* Analytical Chemistry, 2016. **88**(5): p. 2847-2855.
5. Shishkova, E., Alexander S. Hebert, and Joshua J. Coon, *Now, More Than Ever, Proteomics Needs Better Chromatography.* Cell Systems, 2016. **3**(4): p. 321-324.
6. Hebert, A.S., S. Prasad, M.W. Belford, D.J. Bailey, G.C. McAlister, S.E. Abbatiello, R. Huguet, E.R. Wouters, J.J. Dunyach, D.R. Brademan, et al., *Comprehensive Single-Shot Proteomics with FAIMS on a Hybrid Orbitrap Mass Spectrometer.* Anal Chem, 2018. **90**(15): p. 9529-9537.
7. Pfammatter, S., E. Bonneil, F.P. McManus, S. Prasad, D.J. Bailey, M. Belford, J.J. Dunyach, and P. Thibault, *A Novel Differential Ion Mobility Device Expands the Depth of Proteome Coverage and the Sensitivity of Multiplex Proteomic Measurements.* Mol Cell Proteomics, 2018. **17**(10): p. 2051-2067.
8. Schweppe, D.K., S. Prasad, M.W. Belford, J. Navarrete-Perea, D.J. Bailey, R. Huguet, M.P. Jedrychowski, R. Rad, G. McAlister, S.E. Abbatiello, et al., *Characterization and Optimization of Multiplexed Quantitative Analyses Using High-Field Asymmetric-Waveform Ion Mobility Mass Spectrometry.* Anal Chem, 2019. **91**(6): p. 4010-4016.
9. Budnik, B., E. Levy, G. Harmange, and N. Slavov, *SCoPE-MS: mass spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation.* Genome Biology, 2018. **19**(1): p. 161.
10. Myers, S.A., S. Klaeger, S. Satpathy, R. Viner, J. Choi, J. Rogers, K. Clauser, N.D. Udeshi, and S.A. Carr, *Evaluation of Advanced Precursor Determination for Tandem Mass Tag (TMT)-Based Quantitative Proteomics across Instrument Platforms.* Journal of Proteome Research, 2019. **18**(1): p. 542-547.
11. Michael Belford, S.P., Jean-Jacques Dunyach, *Characterization of FAIMS Waveform With Regards to Amplitude, Frequency, Phase, and Electrode Temperature*, in *Poster Note 64438*, S.J. Thermo Fisher Scientific, CA, Editor. 2015, Thermo Fisher Scientific, San Jose, CA.
12. Ludwig, C., L. Gillet, G. Rosenberger, S. Amon, B.C. Collins, and R. Aebersold, *Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial.* Molecular Systems Biology, 2018. **14**(8): p. e8126.
13. Pfammatter, S., E. Bonneil, and P. Thibault, *Improvement of Quantitative Measurements in Multiplex Proteomics Using High-Field Asymmetric Waveform Spectrometry.* J Proteome Res, 2016. **15**(12): p. 4653-4665.

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

15. Giambruno, R., M. Mihailovich, and T. Bonaldi, *Mass Spectrometry-Based Proteomics to Unveil the Non-coding RNA World.* Frontiers in Molecular Biosciences, 2018. **5**(90).

16. Brunet, M.A., M. Brunelle, J.-F. Lucier, V. Delcourt, M. Levesque, F. Grenier, S. Samandi, S. Leblanc, J.-D. Aguilar, P. Dufour, et al., *OpenProt: a more comprehensive guide to explore eukaryotic coding potential and proteomes.* Nucleic Acids Research, 2018. **47**(D1): p. D403-D410.

17. Allan Drummond, D. and C.O. Wilke, *The evolutionary consequences of erroneous protein synthesis.* Nature Reviews Genetics, 2009. **10**: p. 715.

18. Iacobucci, C., C. Piotrowski, R. Aebersold, B.C. Amaral, P. Andrews, K. Bernfur, C. Borchers, N.I. Brodie, J.E. Bruce, Y. Cao, et al., *First Community-Wide, Comparative Cross-Linking Mass Spectrometry Study.* Analytical Chemistry, 2019. **91**(11): p. 6953-6961.

19. Tinnefeld, V., A.S. Venne, A. Sickmann, and R.P. Zahedi, *Enrichment of Cross-Linked Peptides Using Charge-Based Fractional Diagonal Chromatography (ChaFRADIC).* J Proteome Res, 2017. **16**(2): p. 459-469.

20. Pfammatter, S., E. Bonneil, F.P. McManus, and P. Thibault, *Gas-Phase Enrichment of Multiply Charged Peptide Ions by Differential Ion Mobility Extend the Comprehensiveness of SUMO Proteome Analyses.* J Am Soc Mass Spectrom, 2018. **29**(6): p. 1111-1124.

21. Venne, K., E. Bonneil, K. Eng, and P. Thibault, *Improvement in Peptide Detection for Proteomics Analyses Using NanoLC−MS and High-Field Asymmetry Waveform Ion Mobility Mass Spectrometry.* Analytical Chemistry, 2005. **77**(7): p. 2176-2186.