Filière Technologies du vivant

Orientation Chimie analytique

Diplôme 2014

Sibylle Pfammatter

Développement de la mobilité ionique pour la détection des sites de sumoylation des protéines

Professeur
Franka Kalman

Expert
Pierre Thibault, Ph.D.

Date de la remise du rapport
03.11.2014

☐ FSI ☒ FTV	Année académique / *Studienjahr* **2013/2014**	No TD / *Nr. DA* **ca/2014/38**
Mandant / *Auftraggeber* ☐ HES—SO Valais-Wallis ☐ Industrie ☒ Ecole hôte	Etudiant / *Student* **Sibylle Pfammatter**	Lieu d'exécution / *Ausführungsort* ☐ HES—SO Valais-Wallis ☐ Industrie ☒ Ecole hôte
Professeur / *Dozent* **Franka Kalman**	Expert / *Experte* *(données complètes)* **Pierre Thibault, Ph.D.** Professeur et chaire de recherche du Canada en Protéomique Tel: (514) 343-6910 Fax: (514) 343-6843 Courriel: pierre.thibault@umontreal.ca Site Web: http://www.iric.ca/Recherche/Chercheurs/Thibault_P.html **Département de chimie et Institut de Recherche en Immunologie et Cancérologie (IRIC)** **Université de Montreal** C.P. 6128, Station Centre-ville Montreal, QC H3C 3J7	
Travail confidentiel / *vertrauliche Arbeit* ☐ oui / ja [1] ☒ non / nein		

Titre / *Titel*

« Développement de la mobilité ionique pour la détection des sites de sumoylation des protéines »

Description et Objectifs / *Beschreibung und Ziele*

La sumoylation est une modification protéique similaire à l'ubiquitination impliquée dans différents processus cellulaires, incluant la régulation du trafic intracellulaire, la réparation de l'ADN et la réplication, et la signalisation cellulaire. Nous avons créé une souche de cellules humaine HEK293 facilitant la détection des sites de sumoylation (Mol Cell. Proteomics, 10, M110.004796 (2011)). Suite à la digestion trypsique des protéines sumoylées, les peptides correspondant ont un embranchement de cinq acides aminés sur la chaine latérale de la lysine modifiée et ont tendance à former des ions multiplement charges (n >2) lors de l'ionisation par electrospray (Mol. Cell. Proteomics, 12, 2536 (2013)). Tirant profit de cette observation, ce projet portera sur l'utilisation de la mobilité ionique différentielle (J. Proteome Res., 8, 3355–3366 (2009); J. Proteome Res., 11, 927 (2012)) afin d'identifier de façon sélective les peptides sumoylées dans des extraits cellulaires complexes par LC-MS/MS. Suite à l'optimisation avec des peptides synthétiques, cette approche sera appliqué à l'étude des protéines sumoylées en réponse au dommage a l'ADN. Le dommage à l'ADN est très fréquent chez les cellules de mammifères (> 60000 fois par jour) et donne lieu à des modifications (dommage oxydatif, cassure simple ou double brin) suite à des réactions métaboliques et enzymatiques. Le maintien de l'intégrité génomique est essentiel à la survie de la cellule et la coordination des différents mécanismes de reconnaissance, signalisation et réparation de l'ADN nécessite entre autre la sumoylation des protéines. En plus des développements analytiques, ce projet permettra d'obtenir une meilleure compréhension de la sumoylation et des différents intervenants moléculaires impliqués dans la réparation de l'ADN. Dans le cadre de ce projet, le stagiaire utilisera différentes approches à la fine pointe de la technologie en de pointe en chimie des protéines, spectrométrie de masse et protéomique.

[1] *Par sa signature, l'étudiant-e s'engage à respecter strictement la directive et le caractère confidentiel du travail de diplôme qui lui est confié et des informations mises à sa disposition.*
Durch seine Unterschrift verpflichtet sich der Student, die Richtlinie einzuhalten sowie die Vertraulichkeit der Diplomarbeit und der dafür zur Verfügung gestellten Informationen zu wahren.

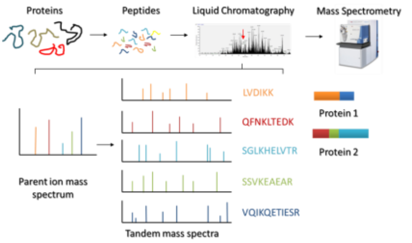


Figure 1: *Vue d`ensemble de la proteomique. Après digestion, les peptides sont analyses par LC-MS/MS ou ils sont fragmentes pour en déduire leur séquence.*

Identification de sumoylation des protéines

Diplômant/e Sibylle Pfammatter

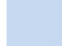
Objectif du projet

La SUMOylation est une modification post-traductionnelle des protéines. La détection par chromatographie liquide couplée à de la spectrométrie de masse est rendue difficile par leur faible abondance et par la faible présence de substrat.

Méthodes | Expériences | Résultats

L'utilisation de High-Field Assymetric Waveform Ion Mobility (FAIMS) résulte en une diminution significative du bruit de fond en LC-MS. Grâce au fractionnement effectué, il est possible d'identifier des peptides de faible abondance par LC-MS/MS. La performance de FAIMS a été testée en comparant des analyses de digestats trypsiques de lysat de cellules HEK293 avec et sans FAIMS par LC-MS/MS. 8260 peptides ont été identifiés (1165 protéines) avec FAIMS et seulement 4820 peptides (734 protéines) sans FAIMS. Quand un nombre représentatif de peptides synthétiques SUMO ont été ajoutés dans un digestat de HEK293, les analyses nano-LC MS ont montré une limite de détection 18 fois plus faible lorsque FAIMS y est couplé. Finalement, la sumoylation est une des premières réponses cellulaires au stress thermique. Donc pour augmenter la concentration de protéines sumoylées, nous avons effectué un choc thermique (43°C) sur des cellules HEK293 exprimant SUMO3 mutante. Un enrichissement de SUMOlylation dans la fraction nucléaire a été observé après 60 min par immunobuvardage. Une analyse LC-FAIMS-MS/MS a donné une augmentation de 2 fois du nombre de peptides SUMOylés identifiés après choc thermique.

Travail de diplôme | édition 2014 |



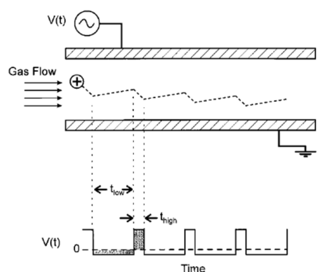
Filière
Technologie du vivant

Domaine d'application
Chimie analytique

Professeur responsable
Prof. Kàlmàn Franka
Franka.Kalman@hevs.ch

Partenaire
Prof. Thibault Pierre
pierre.thibault@umontreal.ca

Institute for Research in Immunology and Cancer (IRIC) Montreal



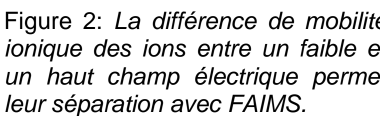


Figure 2: *La différence de mobilité ionique des ions entre un faible et un haut champ électrique permet leur séparation avec FAIMS.*

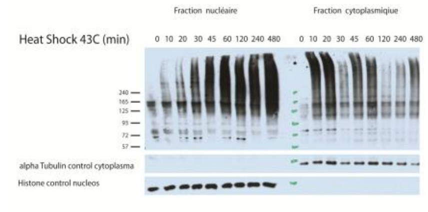


Figure 3: *Réponse au choc thermique à 43°C par immunobuvardage contre les anticorps SUMO 2/3.*

Délais / *Termine*

Attribution du thème / *Ausgabe des Auftrags:*
28.04.2014

Fin des travaux de diplôme dans les laboratoires de l'école / *Ende der Diplomarbeiten in den Labors der Schule:*
-

Remise du rapport / *Abgabe des Schlussberichts:*
Au plus tard le 03.11.2014 | 12h00

Exposition publique / *Ausstellung Diplomarbeiten:*
-

Défense orale / *Mündliche Verteidigung:*
à définir

Signature ou visa / *Unterschrift oder Visum*

Responsable de la filière
Leiter des Studiengangs: .... [1] Etudiant/*Student*: ....

Rapport reçu le / *Schlussbericht erhalten am* .... Visa du secrétariat / *Visum des Sekretariats* ....

Ce rapport est l'original remis par l'étudiant.
Il n'a pas été corrigé et peut donc contenir des inexactitudes ou des erreurs.

University of Applied Sciences Western Switzerland

Identification of SUMOylation Sites using High-Field Asymmetric Waveform Ion Mobility (FAIMS) coupled to a Linear Ion Trap/Orbitrap Mass Spectrometer

provided by

Sibylle Pfammatter

from Eischoll (CH)

Institute for Research in Immunology and Cancer (IRIC)

scientific mentoring by

Pierre Thibault, Ph.D.

Principal Investigator, Proteomics and Bioanalytical Mass Spectrometry research unit, IRIC

Eric Bonneil, Ph.D.

Platform Manager, Proteomics and Bioanalytical Mass Spectrometry research unit, IRIC

Our greatest weakness lies in giving up. The most certain way to succeed is always to try just one more time.

[Thomas A. Edison]

I conducted my Bachelor Thesis at the Institute of Research in Immunology and Cancer (IRIC) of the University Montreal. This institute has 29 different research teams with complementary disciplines dedicated to research in immunology and cancer. My principal Investigator was Prof. Dr. Thibault Pierre, who leads the Proteomics and Bioanalytical Mass Spectrometry platform. His unit focuses on the development and application of methods with approaches in bioanalytical chemistry, protein chemistry, biochemistry and cell biology.

Acknowledgments

First, I thank my supervisor of the internship, **Prof. Thibault Pierre** for providing the research support during my stay at IRIC. I always received generous assistance and warm encouragement.

A special thanks goes to **Dr. Bonneil Eric**, my mentor for the project. Without his guidance my bachelor thesis would not have been possible. I am particularly grateful for all his patience in the last 6 months and I have greatly benefited from his comments and suggestions. I have learned a lot and improved my analytical skills, particularly with FAIMS, through his extensive knowledge and experience in mass spectrometry.

My sincere thanks also go to **Frédéric Lamoliatte** and **Dr. Francis McManus** for their help in the SUMO project, and I would also like to extend my thanks to our bioinformatician **Olivier Caron-Lizotte**. I am most grateful for the assistance given by all of the other collaborators at IRIC.

A special thanks also to **Dr. Kàlmàn Franka** and **Dr. Piantini Umberto** from the University of Applied Sciences Sion and of course to **my family** who gave me the opportunity to do this internship abroad.

And last but not the least I would like to offer my special thanks to my dearest colleagues **Christina, Christine** and **Irene** for their nice welcome, their advice and all the great time and experiences which we shared together.

SUMMARY

SUMOylation is a post-translational modification of proteins that is involved in different cellular processes including mitosis, DNA replication and transcription. However its detection by liquid chromatography coupled to mass spectrometry (LC-MS) can be challenging due to its low abundance and dynamic regulation. High Field Asymmetric Waveform Ion Mobility Spectrometry (FAIMS) helps in eliminating background ions which allows to improve the limit of detection. FAIMS separation is based on the differential mobility of ions between low to high electric field. In a preliminary experiment performed on a tryptic digest of HEK293 proteins, we identified 8260 peptides (1165 proteins) with FAIMS compared to 4820 peptides (734 proteins) without FAIMS. The limit of detection was 12 fmol with FAIMS for synthetic SUMO peptides spiked in a complex tryptic digest and 183 fmol without FAIMS. A LC-MS analysis with FAIMS of a mutant SUMO3 HEK293 cell line resulted in a 3-fold increase in the identification of SUMO peptides.

ABBREVIATIONS

BSA	Bovine serum albumin
CID	Collision-induced dissociation
CV	Compensation Voltage
DC	Direct current
DTT	Dithiothreitol
ECL	Enhanced chemiluminescence
EDT	Electron-transfer dissociation
ESI	Electrospray ionization
FAIMS	High-Field asymmetric waveform Ion mobility spectrometry
FDR	False discovery rate
HCD	Higher-energy C-trap dissociation
HEK293	Human embryonic kidney 293
HRP	Horseradish peroxidase
IE	Inner electrode
LC	Liquid chromatography
LTQ	Linear trap quadrupole
MALDI	Matrix-assisted laser desorption/ionization
MGF	Mascot generic format
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
Ni NTA	Nickel nitrilotriacetic acid
OE	Outer electrode
PBS	Phosphate buffered saline
PTM	Post-translational modification
RF	Radio frequency
rpm	Revolutions per minute
SDS PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SENP	Sentrin/small ubiquitin-like modifier-specific protease
SIM	SUMO-interacting motif
SPE	Solid phase extraction
SUMO	Small ubiquitin-like modifier
TBS	Tris buffered saline
TCEP	Tris(2-carboxyethyl)phosphine
TIC	Total ion chromatogram

1. Introduction 1

1.1. The objective of the work 2

1.2. SUMOylation 3

1.2.1. Enzymatic cascade 3

1.2.2. SUMO Motif 4

1.2.1. SUMOylation in response to stress 5

1.3. Mass Spectrometry 6

1.3.1. Identification of SUMO Targets by Mass Spectrometry 6

1.3.2. Proteomics 6

1.3.2.1. Nano LC-MS 7

1.3.2.1. SUMO Mutant 9

1.3.2.2. Remove SUMO tags by MS/MS Edition 10

1.3.1. FAIMS 11

1.3.1.1. Separation 12

1.3.1.2. Faims for target identification of SUMOylated sites 13

2. Materials and Methods 15

2.1. Material 16

2.1.1. Equipment 16

2.1.2. Chemicals 16

2.1.3. Proteins and Peptides 16

2.1.4. Cell Culture and Western blot 17

2.2. Methods 18

2.2.1. Cell Culture 18

2.2.1.1. Cell culture 18

2.2.1.2. Heat shock 18

2.2.1.3. Cell Collection 18

2.2.2. Protein extraction and purification 19

2.2.2.1. Fractionation of Nuclear/Cytoplasmic fractions 20

2.2.2.2. Nucleus Lysis 20

2.2.2.1. Protein Assay 20

2.2.2.2. Ni-NTA Purification 21

2.2.3. Western Blotting 21

2.2.3.1. SDS PAGE ... 22
2.2.3.2. Protein Transfer ... 23
2.2.3.3. Incubation with Antibody ... 23
2.2.3.4. Detection with enhanced chemiluminescence ... 24
2.2.4. In Solution Digestion ... 25
2.2.5. Desalting ... 25
2.3. Mass Spectrometry Analysis ... 25
2.3.1. Sample Preparation ... 25
2.3.2. LC Separation System ... 25
2.3.2.1. 1D LC ... 25
2.3.2.2. 2D LC ... 26
2.3.2.3. Column ... 26
2.3.3. Mass Analyzer ... 27
2.3.3.1. FAIMS ... 27
2.3.3.2. System Test solution: Promix ... 28
2.4. Software ... 29
2.5. Data Processing ... 30
2.5.1. Clustering with ProteoProfile ... 30
2.6. Comparison FAIMS vs NoFAIMS ... 31
2.7. Linear Dynamic Range of Detection and Quantitation for SUMO 1 and SUMO 3 Peptides ... 32
3. Results and Discussion ... 33
3.1. Optimization of FAIMS Interface ... 34
3.1.1. Gas Flow ... 34
3.1.2. Temperature ... 35
3.1.3. Dispersions Voltage ... 37
3.1.4. Acetonitrile ... 37
3.2. Comparison of LC-MS with and without FAIMS ... 38
3.3. Enhancement of sensitivity with FAIMS ... 45
3.4. Heat Shock HEK293-SUMO 3 ... 47
3.4.1. Western blot Analysis ... 47
3.4.2. Mass Spectrometry Analysis ... 48

4. Conclusion 49

5. Literature 51

6. Annex 55

6.1. UniProt Data's SUMO proteins 56

6.2. Limit of detection 56

6.3. Heat Shock 59

6.4. SUMO 1 and SUMO 3 Synthetics 60

6.5. CAPILLARY COLUMN PACKING 64

6.6. PROMIX 08/07/2014 66

6.7. Group Professor Dr.Pierre Thibault 67

1. Introduction

1.1. THE OBJECTIVE OF THE WORK

The aim of this project is to identify small ubiquitin-like modifier (SUMO)-modifed substrates and their modification sites using ion mobility and mass spectrometry. SUMOylation is a dynamic process, found in all common species, and is essential for eukaryotic life. Till this day, the exact outcomes and dynamics of this modification are not completely understood. For a better understanding of this post-translational modification (PTM), we aimed to identify SUMO target sites using mass spectrometry. However, there are several aspects that make this investigation difficult. First, this modification is highly dynamic and reversible. Secondly, cell extracts are highly complex and SUMOylated proteins are found at low concentrations with variable stoichiometry of modification site (1). Third, the sequences of SUMOylated peptides are branched, which makes the MS/MS spectra complex and hard to interpret. To improve sensitivity, nano LC-ESI-FAIMS-MS can be used. High-Field Asymmetric Waveform Ion Mobility Spectrometry (FAIMS) reduces background noise and allows for selective measurement of peptide groups in function of their charge state (2). This work comprises the following three aims:

1. Determine the distribution of charge state as of function of the compensation voltage (CV, parameter FAIMS) for synthetic SUMO peptides and tryptic peptides from Human embryonic kidney 293 (HEK293) total cell lysate.
2. Determine the limit of detection of synthetic SUMO peptides with and without FAIMS.
3. Determine the dynamic change in SUMOylation during heat shock treatment for protein extracts isolated from a mutant HEK293 SUMO3 cell line.

1.2. SUMOylation

SUMOylation regulated the activity, stability and the interaction of hundreds of proteins in different pathways. There are several cellular functions where the SUMO proteins are important. Some of these functions include: gene transcription, cell cycle and DNA repair (3). We can find SUMOylated proteins most frequently in the nucleus, where the SUMO enzymes are enriched (1). Additionally, free and conjugated SUMO is also present in the cytoplasm. All eukaryotes express at least one SUMO precursor protein. In humans we have 4 different SUMO isoforms, but only SUMO 1, SUMO 2 and SUMO 3 are ubiquitously expressed. SUMO 2 and SUMO 3 share 97% sequence identity and only differ by three amino acid residues. Antibodies cannot distinguish between these two isoforms. This work primarily focused on SUMO 3. An important feature of SUMO 1, 2 and 3 is that they all have the same glycine-glycine motif at the C-terminus.

```
         10         20         30         40         50         60         70         80         90        100         human
MSDQEAKPST EDLGDKKEGE YIKLKVIGQD SSEIHFKVKM TTHLKKLKES YCQRQGVPMN SLRFLFEGQR IADNHTPKEL GMEEEDVIEV YQEQTGGHST V       SUMO 1
MADEKPKEGV KTENNDHINL KVAGQDGSVV QFKIKRHTPL SKLMKAYCER QGLSMRQIRF RFDGQPINET DTPAQLEMED EDTIDVFQQQ TGGVY             SUMO 2
MSEEKPKEGV KTENDHINLK VAGQDGSVVQ FKIKRHTPLS KLMKAYCERQ GLSMRQIRFR FDGQPINETD TPAQLEMEDE DTIDVFQQQT GGVPESSLAG HSF    SUMO 3
```

Figure 1: The different sequences of the human protein SUMO 1, SUMO 2 and SUMO 3. SUMO 1 contains 101 amino acids; terminal cleavage with trypsin gives a peptide with 23 amino acids (ELGMEEEDVIEVYQEQTGGHSTV). SUMO 2 and SUMO 3 are distinguishable by their 3 different amino acids at the beginning of the peptide sequence; here marked in yellow. SUMO 2 and SUMO 3 consist of 95 and 103 amino acids, respectively. The last trypsin cleavage gives the peptides FDGQPINETDTPAQLEMEDEDTIDVFQQQTGGVY for SUMO 2 (34 amino acids long) and FDGQPINETDTPAQLEMEDEDTIDVFQQQTGGVPESSLAGHSF for SUMO 3 (43 amino acids). The di-glycine motif can be seen close to the end of the sequence in all three SUMO proteins.

1.2.1. Enzymatic cascade

SUMOylation of proteins is a reversible process. First, SENP proteases (sentrin/small ubiquitin-like modifier-specific protease) deconjugate SUMO modifications from target proteins and cleave terminal C-teminus residues of free SUMO to expose its di-glycine motif. This cleavage is required for the activation of SUMO and for the conjugation of SUMO to lysine residues of the target protein. Free SUMO are first transferred to the E1 activating enzyme (SAE1/SAE2) via a thioester bond. It is then transferred to the enzyme E2 (UBC9) to form an isopeptide bond. Finally, E3 ligases can bind specific substrate to catalyze a nucleophile attack of the target lysine once interacting to UBC9. As previously mentioned, SENP proteases can activate the SUMO protein as well as catalyze the deSUMOylation of proteins (1). If they activate the SUMO precursor protein, they are called endopeptidase because they break the peptide

bond. If they remove the SUMO modification from the target, the SENP is referred to as an isopeptidase.

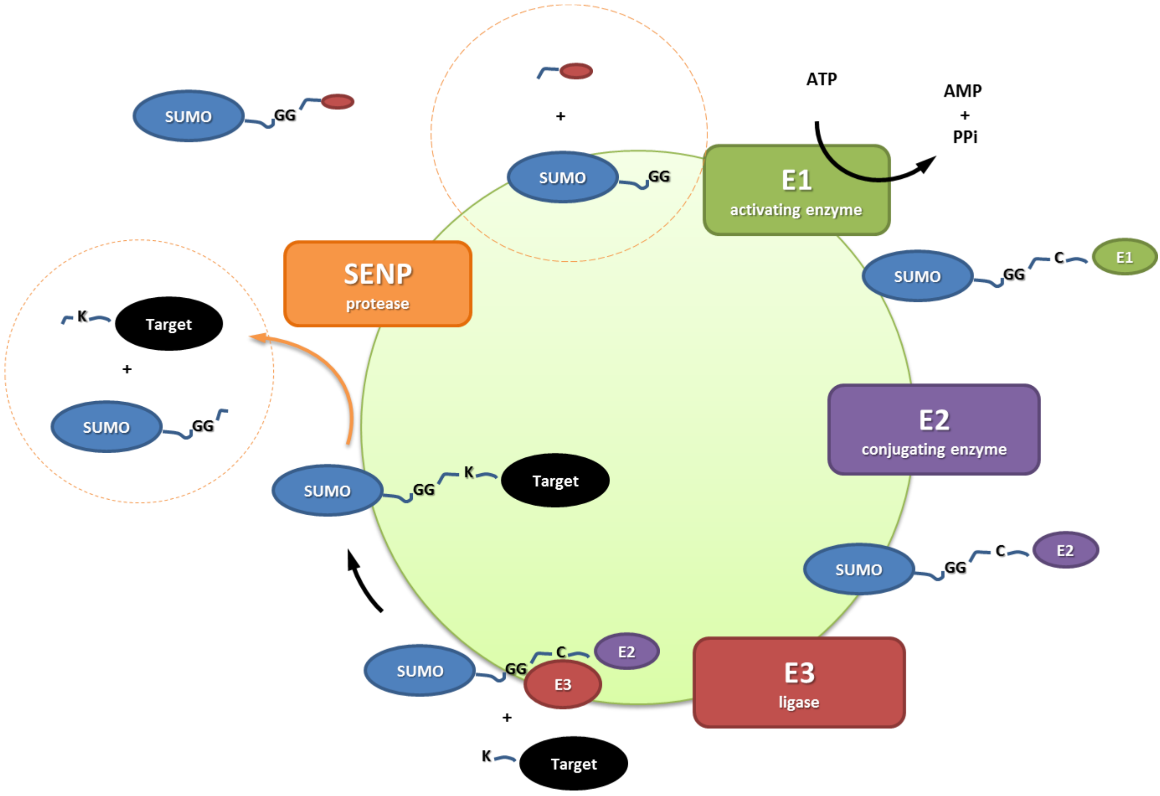


Figure 2 There are several enzymes involved in SUMOylation. First the SENP activates the SUMO precursor protein, then it is bound by the activating enzyme E1 and E2 and then conjugated to the target protein by a isopeptide link with the help of E3. The SUMO conjugation is a reversible process and the modification can be removed by a SENP protease.

1.2.2. SUMO Motif

SUMO attaches covalently to target proteins *via* an isopeptide bond. This process takes place between the C-terminal glycine of the mature SUMO sequence and the ε-amino group of the lysine residue from target proteins. Although there are different kinds of target motifs, the most common consensus motif is ψKXE/D. Where ψ is a large hydrophobic amino acid, K is the lysine which forms the isopeptide link with the SUMO, X is any amino acid, and the last residue is either an aspartate or glutamate residue. This is a reversible process where isopeptidase SENP can cleave the isopeptide bond. Moreover, SUMO can also bind non-covalently *via* a SUMO-interacting motif (SIM) (4).

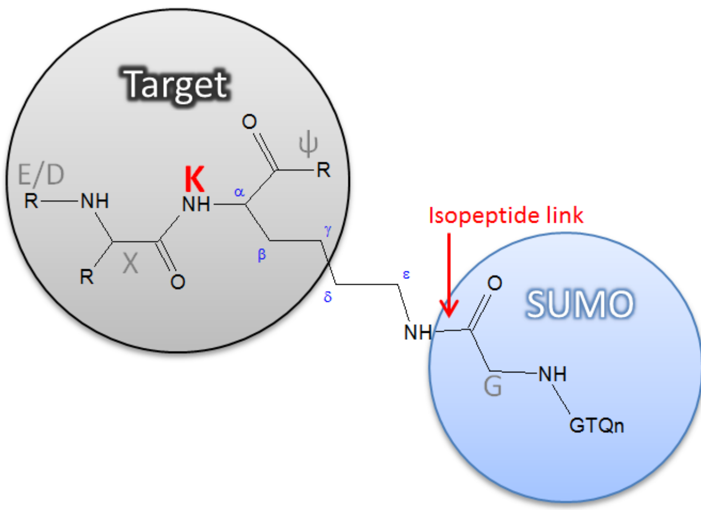


Figure 3: Isopeptide bond between the target and the SUMO peptide

1.2.1. SUMOYLATION IN RESPONSE TO STRESS

Protein SUMOylation is regulated by different environmental changes including viral infection, proteasome inhibition, and heat shock (5) (6). An accumulation of SUMO 2/3 conjugates are detected under heat shock as well as treatments with ethanol, salt or oxidation. Similar results are observed when cells are treated with the protease inhibitor MG132.

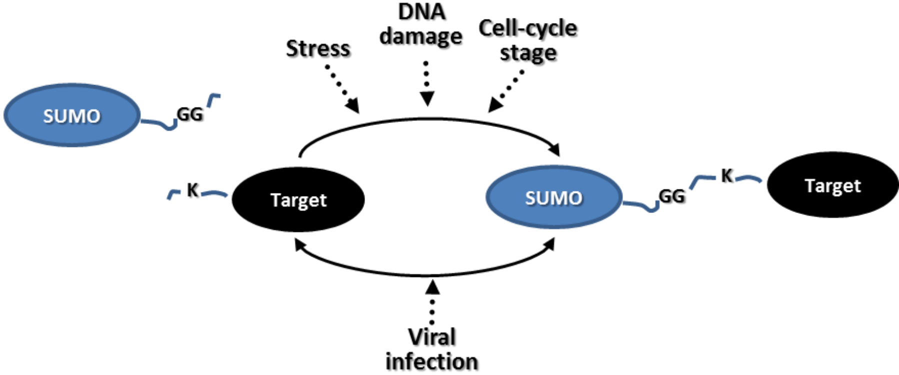


Figure 4: The overexpression of SUMOylation can be caused by stress or DNA damage.

1.3. MASS SPECTROMETRY

1.3.1. IDENTIFICATION OF SUMO TARGETS BY MASS SPECTROMETRY

SUMO target recognition is very difficult. The consensus motif ψKXE/D is very short and might be found frequently in mass spectrometry analysis. The occurrence of this motif does not necessarily imply that the corresponding lysine residue will be necessarily SUMOylated (7). To get a better understanding of the proteome, we can apply mass spectrometry based proteomics (8). With help of mass spectrometry we can identify modifications at a protein level.

1.3.2. PROTEOMICS

Proteomics allows us to have a better look at the different process that take place in cells. For instance, we can quantify proteins that are up or down regulated in diseases. In fact, proteomics is split in two phases: discovery and the targeted proteomics. In the discovery phase, we perform a global large-scale identification of peptides to determine which one are regulated. In targeted proteomics we are looking for specific peptides or modifications in proteins. To get the most comprehensive structural characterization of modified proteins, the Bottom-up method is used. The proteins are most often digested with the serine protease trypsin. Trypsin is mostly used because it cleaves specifically at the C terminus of lysine(K) and arginine(R), which both have basic residues found at the carboxyl terminus (9) (10). Another advantage is the relative high frequency of lysine and arginine in protein sequences. Also, the presence of lysine or arginine residues can be identified in tandem mass spectra by characteristic fragment ions at m/z 147.11 and 175.11.

Protein quantification using mass spectrometry provides a deeper insight into the dynamics of organisms. The following figure is a brief explanation of the workflow in proteomics: proteins are digested to peptides, then the corresponding samples are separated by liquid chromatography prior to their analysis on the mass spectrometer. Bioinformatics tools then allow us to sequence and identify peptides for the acquired tandem mass spectra.

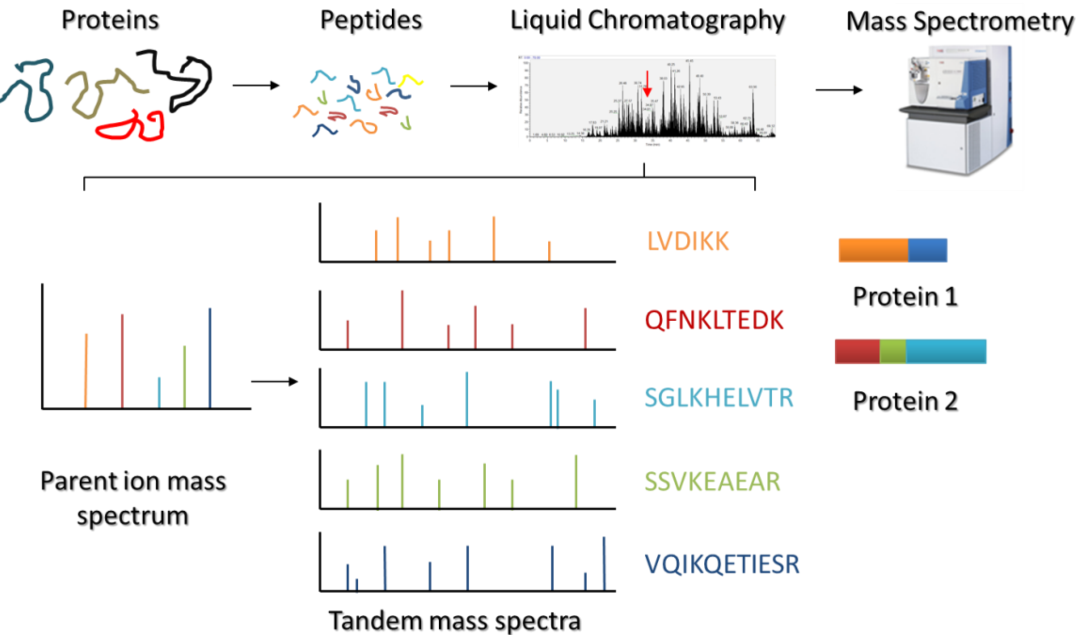


Figure 5: Basic overview of proteomics. After tryptic digestion, we get peptides that are analyzed in a LC-MS system. With the help of the mass spectrometer we produce tandem mass spectra. These are fingerprints of each peptide. We can identify peptides with the help of bioinformatics tools. Then we try to reconstruct the peptides sequences and assign them to their respectful proteins.

1.3.2.1. NANO LC-MS

In proteomics nanoscale liquid chromatography coupled to tandem mass spectrometry is used and it has a very high sensitivity and specificity. After separation on the liquid based system, the sample is transferred into the mass spectrometer *via* an electrospray source. Electrospray ionisation (ESI) and matrix-assisted laser desorption/ionization (MALDI) are common methods used to ionized and volatize the peptides in complex samples (8). Usually peptide amounts present in digest at present at sub μM concentrations, and, low flow rates and small column diameters are typically to increase sensitivity. This improves at the same time the ionization efficiency of electrospray by forming smaller droplets (11). The spray is generated in an electric field between the end of the chromatogram system and the entrance of the mass analyzer and gives rise to the formation of a Taylor cone. LC effluent emerging from the emitter is subjected to the applied electric field and is nebulized into small droplets that undergo subsequent fission events and solvent evaporation to form droplets of smaller sizes. Ions contained within these smaller droplets undergo increasing electrostatic repulsion and are eventually expelled and analyzed by the mass spectrometer.

For this work we employed the Orbitrap Elite from Thermo. The instrument is composed of a combination of a linear ion trap and orbitrap which provide great advantages: robustness and sensitivity by the ion trap and high resolution and accuracy by the orbitrap (8).

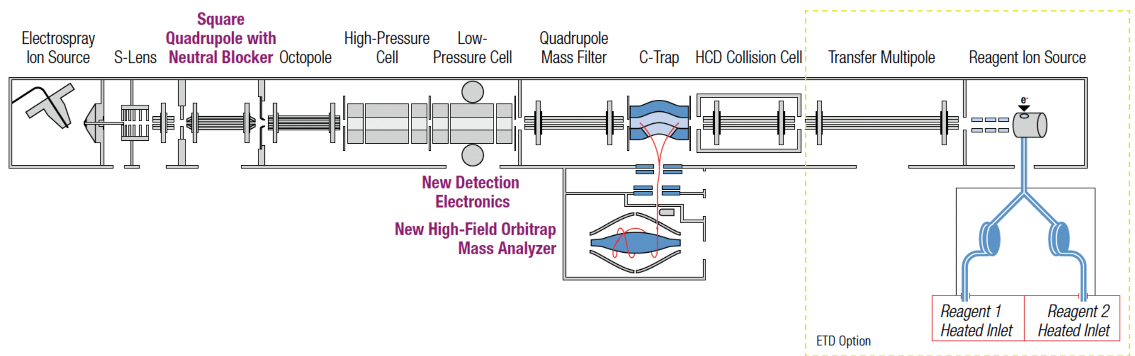


Figure 6: Ions enter the mass analyzer in the form of ion gas. In the first step they are trapped in the linear ion trap before they transferred to the orbitrap (12)

The gaseous peptides enter the mass analyzer and the ion beam is focused through a set of stainless steel apertures to which an RF voltage is applied called the S-Lens. This S-lens has improved transmission of ions by a factor of 5 to 10 compare to previous design which did not have the S-lens. In MS mode, ions are travelling through the dual pressure LTQ ion trap and are stored in the C-trap to prepare them for the injection in the Orbitrap. The orbitrap mass analyzer is composed of a spindle-like central electrode and a barrel-like outer electrode. A DC voltage is applied between the two axially symmetric electrodes, resulting in a quadro-logarithmic field. Ions are accelerated from the C-trap through an optical deflection system and are trapped radially around a central spindle electrode. The motion of the ions in the orbitrap involves axial oscillations along and rotation around the central electrode. The rotation frequency of the ions is dependent on the charge to mass ratio as described in Eq.1. This frequency is determined with the image current given by the ion movement around the central electrode. In MS/MS mode, two options are available. For both, precursor selection and isolation is performed in the LTQ ion trap. Then, the precursor can be either fragmented and analyzed in the LTQ ion trap (collision induced dissociation CID) or fragmented in the HCD collision cell. In this latter case, the ions are sent back to the C-trap and finally to the Orbitrap to be analyzed (higher-energy collisional dissociation HCD). CID is very fast but provides low mass accuracy on the fragments (± 0.5 ppm). HCD is slower than CID but provide a much better accuracy on the fragments (± 3 ppm).

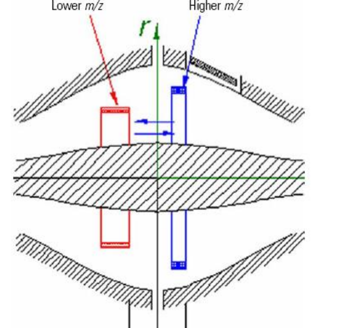


Figure 7: Ions rotate around an axis with different frequency (12)

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

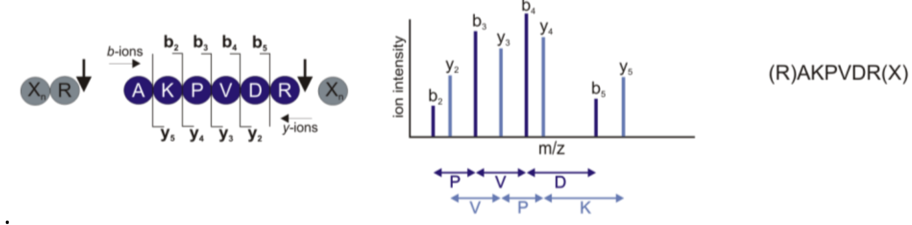
Equation 1: Frequency of the ions in the orbitrap

1.3.2.1. SUMO MUTANT

To facilitate the identification of SUMO proteins by mass spectrometry, a HEK293 cell line with a SUMO 3 mutant was used (13). The mutant SUMO 3 resembles the wild-type except for a few modifications. There is a His_6 tag on the N terminus which allows for selective enrichment on a Ni-NTA column. There is also a strategic replacement of glutamine 87 with a lysine. Digestion by trypsin provides us with a 5 amino acid long peptide on the target lysine instead of 32 amino acids. A Q88N substitution was also introduced to distinguish SUMO 2 and SUMO 3 modifications by mass spectrometry.

Figure 8: Created SUMO 3 mutant in human cell HEK293 (13). On the N terminus there are six consecutive histidine residues and there is a lysine at position 87 instead of a glutamine.

Unfortunately bioinformatic applications are designed to analyze linear peptides and the interpretation of MS/MS spectra is compromised by the occurrence of branched peptides. Conventional search against a database with Mascot leads to low score identifications. The following are examples of MS/MS spectrums of a linear and a branched peptide.



.

Figure 9: MS/MS spectre for linear peptide AKPVDR (3)

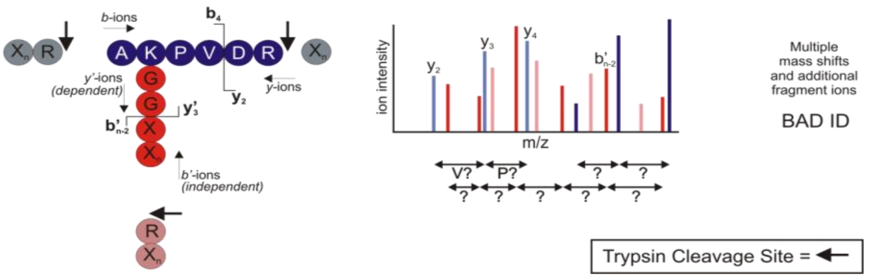


Figure 10: Same peptide but this time SUMOylated (branched peptide). We get a complex MS/MS spectre and bioinformatic tools interpret them as false (3)

Previous experiments by Thibault's lab members also showed that the score in HCD mode for SUMO is better than CID or ETD (3). This was achieved using a representative number of synthetic SUMOylated peptides (7-21 amino acids long) to test the three different fragmentation methods and tabulating the occurrence of specific SUMO fragments. They showed that there was a higher frequency with HCD than with the CID or ETD (3). Moreover, they created a novel bioinformatic tool called MS/MS edition to increase the identification score by Mascot for SUMO modified peptides.

1.3.2.2. REMOVE SUMO TAGS BY MS/MS EDITION

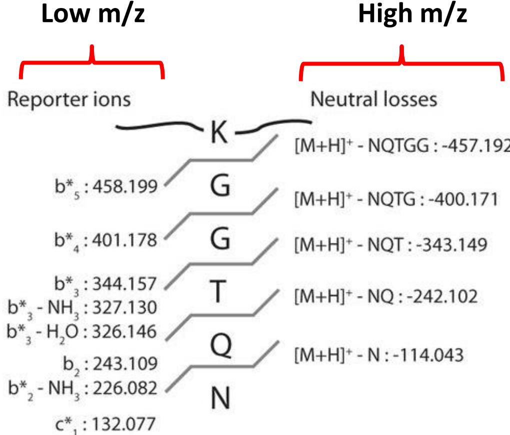


Figure 11: Specific SUMO 3 fragments from the HEK293 SUMO 3 mutant (3)

The following explanation refers to the SUMO 3 mutant tryptic peptide that comprise the pentapeptide NQTGG linked to the modified lysine residue *via* an isopeptide bond. In a typical case we see the fragments of the reporter ions at a low m/z value and the neutral losses at high m/z value. These fragments are characteristic of the branched part of the peptide and disturb the software. With an in house software MS/MS edition we can edit the MS/MS spectra and convert it into a ”linearized” version more easily interpreted by conventional search engines.

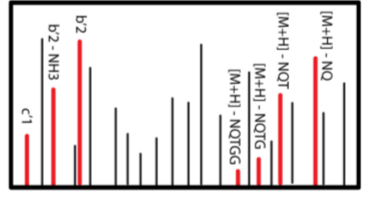


Figure 12: Specific SUMO 3 fragments. At low m/z value we expect to see the fragments of the reporter ions like the loss of N, NQ, NQT … At high m/z we see the neutral losses.

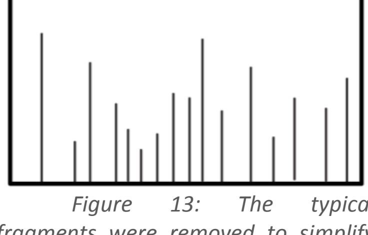
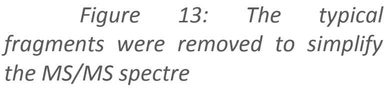


Figure 13: The typical fragments were removed to simplify the MS/MS spectre



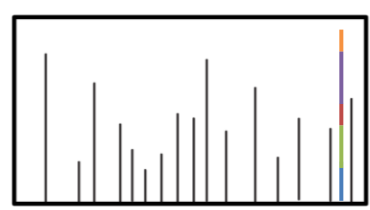


Figure 14: The removed mass will be added to the peak to get the right mass

After linearization of the MS/MS spectre as described above, using the common sequencing and data base softwares is less confusing and the identification scores is much higher. With help of the SUMO 3 mutant and the mgf editing using MS/MS edition we are able to use proteomic mass spectrometry to successfully identification SUMOion sites in the proteome.

1.3.1. FAIMS

FAIMS adds an additional dimension of separation and can be applied as a novel approach for the identification of SUMOylated target sites. In proteomics the samples are complex and a one-dimension separation is usually not enough due to matrix complexity (14). We thus have to find a way to improve selectivity. In addition, high background noise and low abundance of some peptides makes it difficult to detect them. Characterization of the entire complement of proteins is never feasible with a normal separation system since instrument performances are limited. With FAIMS we can reduce background noise and select for transmitted ions based on their different mobility in high and low electric field (15). If we scan under various conditions, we can detect different kinds of ions and the fractionating step in the gas phase allows for the separation of singly charged peptides from multi charged peptides.

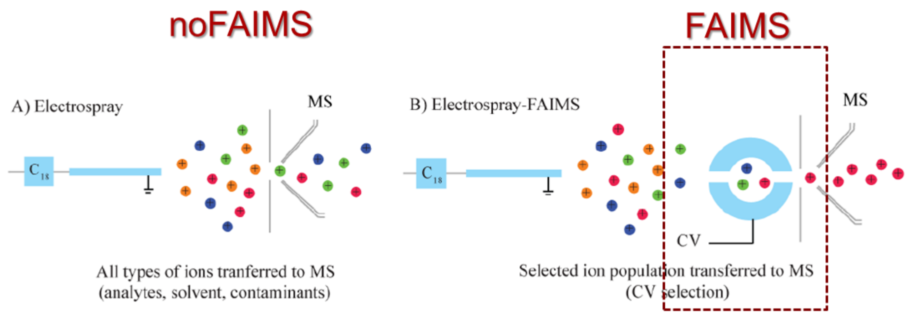


Figure 15: Overview of the basic differences between LC-MS and LC-FAIMS-MS. With LC-MS ions pass directly after ionization in the mass spectrometer. With FAIMS we transmit only ions with the desired mobility to the mass analyzer (16)

1.3.1.1. SEPARATION

The separation in FAIMS takes place between 2 electrodes, which produce an electric field, and an inert carrier gas that helps to mobilise the ions forward to the mass analyzer. The electric field is applied *via* an asymmetric waveform (V(t)) and during one cycle this waveform integral voltage-time product is equal to zero ($V_{high}t_{high} + V_{low}t_{low} = 0$). The maximal amplitude for the applied cycle voltage is called the dispersion voltage.

The ions have different drift velocity and so not all of them need the same requirements to migrate to the electrodes. The traveled distance to the electrode can be calculated with equation 2. Where K is the ion mobility which is independent on the low electric field, this means the velocity in the low electric field is proportional to the field strength. This is not the case in high electric field. In high electric field the behavior of ions depends on the electric field thus we speak about K_h. If K and K_h are identical, the traveled distance is the same and the ion is back to its initial position after one waveform cycle. If $K_h > K$ the ions displace themselves in direction to the mass spectrum.

$$d = v \cdot t = K \cdot E \cdot t$$

Equation 2: traveled distance of ions in an electric field

In addition to the dispersion voltage, a compensation voltage (CV) is applied on the lower electrode to prevent ions from drifting towards the electrode. In the case where ions touch the electrode they become neutral and do not pass through the electrodes to the mass analyzer.

The compensation voltage value is specific for each compound and allows us to regulate the separation mechanisms. In short, the separation is achieved because of the various types of ions travelling to different distances (17).

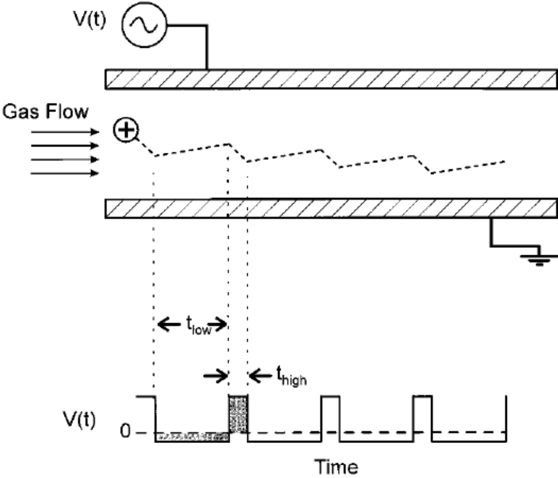


Figure 16: Scheme depicting ion separation in FAIMS. The behaviour of different ions in the electric field vary and allows for separation.

1.3.1.2. FAIMS FOR TARGET IDENTIFICATION OF SUMOYLATED SITES

Digestion by trypsin has different advantages. One of them is the positive charge of lysine/arginine after tryptic digestion. With the SUMO modification we have also a positive charge arising from the N-terminus amino group of the SUMO chain. The possibility of getting a multiply charged peptide is high and we can make use of this in ion mobility spectroscopy. FAIMS also provides the possibility to increase the signal to noise ratios by adding a new dimension without increasing the analysis time. Given that less than 1% of the peptides are SUMOylated eliminating background noise is a necessity. The following is an example of an MS spectrum without and with FAIMS to show the decrease in background noise.

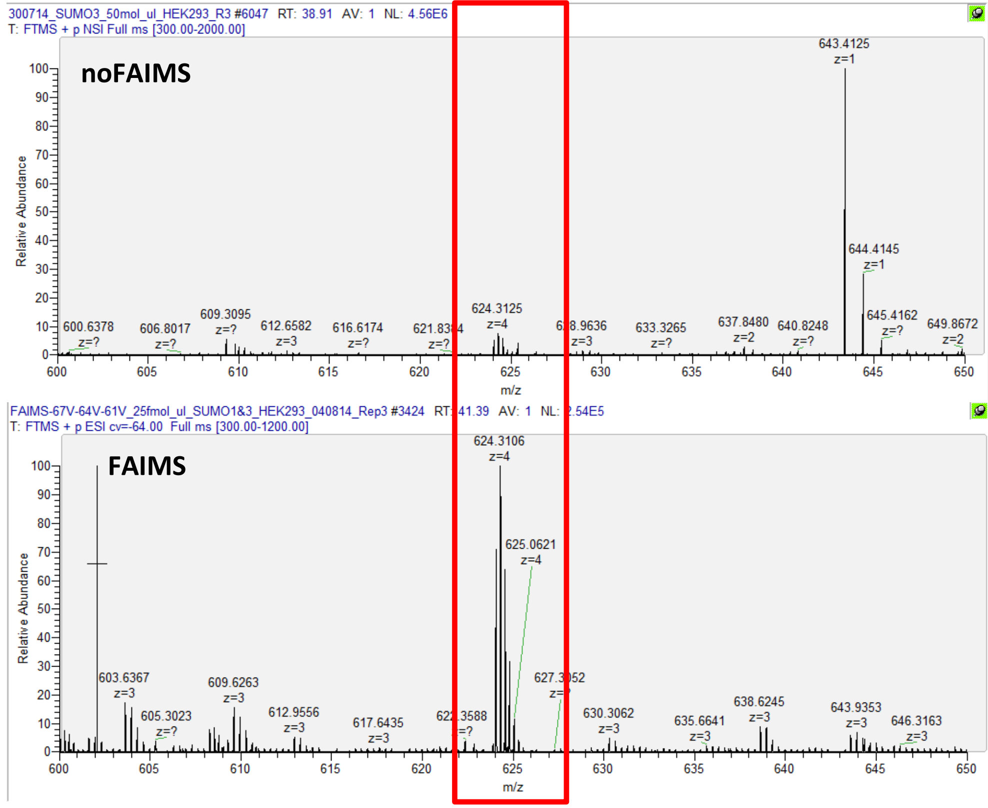


*Figure 17: The full MS spectrum for the SUMO 3 modified peptide NLLHDNELSDL**K**EDKPR (4+) m/z 624.31 at 400 fmol. Above is a spectrum with the normal nano-LC-MS system, below is with nano-LC-FAIMS-MS. The signal to noise ratio is increased 5-fold when using FAIMS.*

2. Materials and Methods

2.1. MATERIAL

In the following tables all of the laboratory equipment and materials that were used.

2.1.1. EQUIPMENT

Table 1: Labor equipment which was used for the different manipulations in the laboratory

Speed Vac	Thermo SPD131DDA-115
Incubator	Eppendorf Thermomixer R
Centrifuge	Sorvall Legend RT 20057916
UV CO_2 Incubator	Panasonic UV Safe cell MCO-19AIC-PA
Steril Gard III Advannced Hood	Baker Company SG-403A
Power Pac Universal	Bio Rad 164-5070
Xray cassette	Fisher FBXC810
Digital Dry Bath	Labnet D1200
Multiscan spectrum plate reader	Thermo electron corp.

2.1.2. CHEMICALS

Table 2: Chemical list for the different manipulations in the laboratory

Formic acid (98%)	EMD FX0440-7
Acetonitrile HPLC grade	Fisher Scientific A998-4
Trypsin modified (sequencing grade)	PROMEGA V511A
tris(2-carboxyethyl) phosphine (TCEP)	Themro Scientific 77720
Ammonnium Bicarbonate (98-101%)	Bio Basic VWR CA99501-001
2-Chloroacetamide	SIGMA C0267
Methanol	VWR CA-MX0475
Glycine	VWR CA-EM4840
Tris base	VWR CA-EM9230
Imidazole (99.0%)	WVR CA99501-844
Sodiumphosphate monobasic monohydrate	Sigma S9638
Guanidine Hydrochloride	Bio Shop GUA003.1
SDS	Sigma L3771

2.1.3. PROTEINS AND PEPTIDES

Table 3: List of proteins and peptides used for different manipulations

	Supplier	**Use**
Alcohol Dehydrogenase from *Saccharomyces cerevisiae*	Sigma A7011, LOT# 041M7354V	Promix solution
Bovine Serum Albumin	BioShop ALB00150, LOT# 3927998	Promix solution
Carbonic Anhydrase from bovine erythrocytes	Sigma C3934, LOT# 081M1626V	Promix solution

Cytochrome c from *Saccharomyces cerevisiae*	Sigma C2436, LOT# 037K7008V	Promix solution
Enolase from *Saccharomyces cerevisiae*	Sigma E6126, LOT| 097K7690V	Promix solution
Hemoglobin human	Sigma H7379, LOT# SLBC0633V	Promix solution
Phosphorylase b from rabbit muscle	Sigma P6635, LOT# 110M7675V	Promix solution
Transferrin human	Sigma T3309, LOT| 0L194488	Promix solution
Angiotensin human acetate salt hydrate	Sigma A9650	FAIMS spray position
SUMO synthetic (see annex 6.4)	JPT Peptide Technologie GmbH 16034 PO#U629407_2004655	Linearity
SUMO 1 'GGTQE'	N-terminus amine **51 peptides** C terminus lysine acid **45 peptides** C terminus arginine acid	
SUMO 3 'GGTQN'	N-terminus amine **40 peptides** C terminus lysine acid **56 peptides** C terminus arginine acid	

2.1.4. CELL CULTURE AND WESTERN BLOT

Table 4: Substance and equipment used for the cell culture of HEK293 cell line

	Supplier	**Use**
Dulbecco's modified eagle medium (DMEM)	Thermo Scientific SH30081.02	cell culture
Fetal bovin serum (FBS) deactivated filtred	SERADIGM 1400-500	cell culture
Phosphate buffered saline (PBS)	HyClone SH30258.01	cell culture
Genectin	Gibco 10131-027	cell culture
Antibody SUMO 2	Invitro LOT 394064A	Western Blotting
Antibody α-Tubulin	Cell Signaling # 2144S LOT 4	Western Blotting
Antibody Histone H3	Cell Signaling # 9715S LOT 20	Western Blotting
Anti Rabbit IgG (H+L) HRP	Merck AP307P	Western Blotting
Blotting paper	VWR 28298-020	Western Blotting
Tween 20	Sigma P1379 P1379	Western Blotting
Nitrocellulose membrane	Pall DR-058711112E	Western Blotting
Full blue Xray films	Cie univ de rayons X du Canada, 1081728	Western Blotting
96 well plate	VWR CA62406-343	Bradford
Ni-NTA Agarose	QIAGEN 1018240	Purification
NuPAGE (Mes SDS Running Buffer)	Novex NP0002	Gel electrophoresis

Criterion XT Precast Gel (4-12% Bis-Tris)	Bio Rad 345-0124	Gel electrophoresis
Protein Assay Dye Reagent Concentrate	Bio Rad 500-0006	Bradford
Amersham ECL Western Blotting Detection Reagents	GE Healthcare RPN2235 and RPN2106	Detection Western Blot
BLUeye Prestained Protein Ladder	FroggBio PM007-0500 (10-245 kDa)	Gel electrophoresis

2.2. METHODS

2.2.1. CELL CULTURE

2.2.1.1. CELL CULTURE

For culture of HEK293 WT and HEK293 SUMO3 cells, Dulbecco's Modified Eagles Medium (high glucose) supplemented with 10% fetal bovine serum, 1% Penicillin/Streptomycin Solution and 1% L-Glutamine was used. The cells were grown at 37 °C in a 5% CO_2 incubator. 750µg/ml Geneticin was added to maintain the selection of HEK293 SUMO3 cells.

2.2.1.2. HEAT SHOCK

To study the effect of stress on protein SUMOylation, the cells were treated with a heat shock at 43°C. We added preheated medium (43°C) to the cells and incubated them at 43°C in a 5% CO_2 incubator for different periods of time. For the large scale study, the cells were treated with a 60 minute heat shock.

2.2.1.3. CELL COLLECTION

The medium is removed and the cells washed with PBS. The cells are subsequently collected in fresh PBS. The cells are then centrifuged for 10 minutes at 1000 rpm. The PBS (supernatant) is removed to isolate the cell pellet. The pellet is washed again with PBS and centrifuged for 10 minutes at 1000 rpm. The PBS is removed and the cell pellet is stored at -20°C for least for 12 h.

2.2.2. PROTEIN EXTRACTION AND PURIFICATION

Figure 18 shows the workflow for the treatment of the mutant HEK293 SUMO3 cell line. After inducing stress by heat shock, a nucleus – cytoplasm fractionation is done. Afterwards the purification on Ni NTA beads of the 6 x His-tagged proteins is done. For the mass spectrometry analysis the proteins are digested by trypsin.

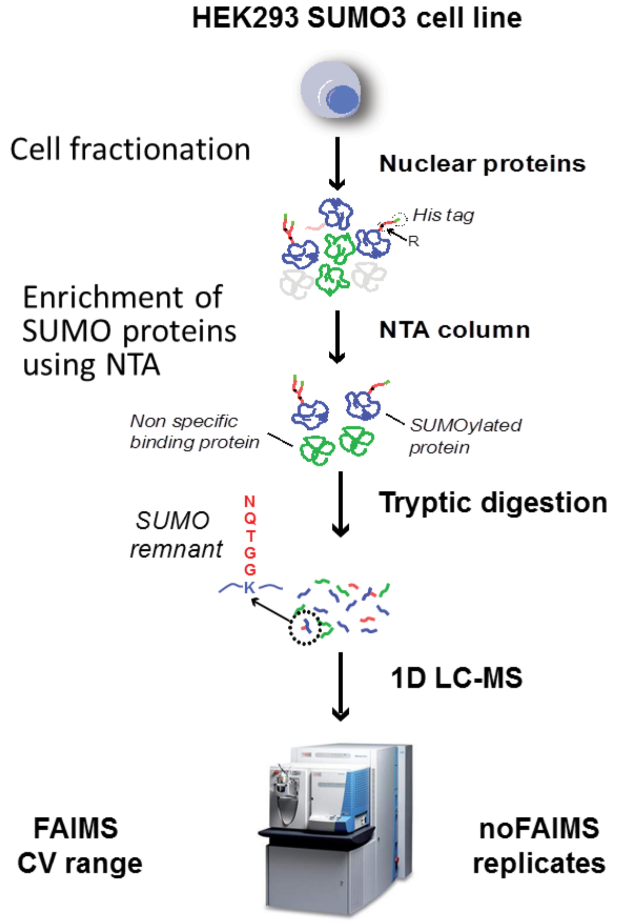


Figure 18: Brief overview of the workflow for the heat shock treatment (3). In the first step, the nucleus and cytoplasm are separated. After NiNTA purifications we digest with trypsin and measure on the LC-MS.

LSB

10 mM	Tris pH 7.6	stock solution
1.5 mM	magnesium chloride	
20 mM	2-Chloroacetamide	
100 X	Protease Inhibitor	add fresh
100 X	Phosphatase Inhibitor	

BUFFER A

8 M	Urea	stock solution
100 mM	Sodiumphosphate monobasic monohydrate	
10 mM	Imidazole	
10 mM	Tris-HCl pH 8	
20 mM	2-Chloroacetamide	add fresh
10 mM	β-Mercaptoethanol	

BUFFER B

8 M	Urea	stock solution
100 mM	Sodiumphosphate monobasic monohydrate	
20 mM	Imidazole	
10 mM	Tris-HCl pH 6.3	
10 mM	β-Mercaptoethanol	add fresh

2.2.2.1. FRACTIONATION OF NUCLEAR/CYTOPLASMIC FRACTIONS

The cell pellet is dissolved carefully in LSB buffer, and if necessary, it is vortexed slightly. Incubate with the phosphatase and protease (for SENP) inhibitor for at least 30 minutes with shaking (4°C). After, centrifuge for 15 minutes at 215g. The supernatant is the **cytoplasmic fraction**. Wash the nucleus pellet with LSB buffer and centrifuge again at 215g for 15 minutes. Remove the wash LSB solution and add buffer A to the nuclear pellet. Till this point, all manipulations were performed on ice.

2.2.2.2. NUCLEUS LYSIS

After adding 8M Urea (buffer A), the proteins are denatured. Vortex vigorously the pellet solution and sonicate 2 times for 10 seconds to shear the DNA. Centrifuge 15 minutes at 16000 g and transfer the supernatant to a fresh tube and label as **nuclear fraction**.

2.2.2.1. PROTEIN ASSAY

A Bradfords assay is used to quantitate the protein content of the various fractionated samples. The calibration curve includes standard solutions with concentrations of 0, 5, 25, 50, 125, 250, 500, 750, 1000 µg/mL of bovine serum albumin. For measurements, we take 96 well microplates and scan the absorbance at 595 nm. According to the Lambert Beer's law.

$$A = \varepsilon \cdot l \cdot c = \varepsilon \cdot l \cdot \frac{n}{V} = \varepsilon \cdot \cancel{l} \cdot \frac{n}{s \cdot \cancel{l}} = \varepsilon \cdot \frac{n}{s}$$

Equation 3: Lambert Beer law. Sample volume is negligible.

where

A = absorbance [-]
ε = molar extinction coefficient [M^{-1} cm^{-1}]
l = pathlength; height well [cm]
c = molar concentration [M^{-1}]
n = amount of substance [mmol]
V = volume well [cm^3]
s = surface well [cm^2]

Consequently, the volume of the sample has no influence when the incident beam is perpendicular to the plane of the 96-well plate. By default, 190 µl Bradford reagent and 5-10 µl sample solution is taken for the direct preparation in each well of the micro plate.

2.2.2.2. Ni-NTA Purification

For the Ni-NTA purification we add 1 mL of Ni-NTA beads to 50 mg of proteins. In practical terms, that means we have to take 2 mL of Ni-NTA Agarose beads since the Ni-NTA solution comes as 50% beads and 50% ethanol mixture. The final calculated volume of Ni-NTA beads is centrifuged for 1 min at 215 g. The supernatant is removed. Wash the beads 4 times with buffer A. The sample is added to the equilibrated beads and incubated overnight at room temperature. Centrifuge for 1 min at 215 g and remove the flowthrough (keep it for western blot analysis). Wash beads 1 time with buffer A, 4 times with buffer B and 2 times with 50 mM ammonium bicarbonate. To remove the proteins from the beads for Western Blot, use 200 mM of Imidazole buffer (with 20 mM Tris adjust at pH 7.0) and incubate for 10-20 min at room temperature.

2.2.3. Western Blotting

Figure 19: Western blot Workflow. We first run a SDS PAGE, transfer the proteins on a nitrocellulose membrane and incubate the membrane with primary and secondary antibodies. ECL is added to the membrane and the emission is captured on Xray films.

TBS 0.1% TWEEN (TBST)

2.42 g	Tris-HCl
8 g	NaCl
1 ml	Tween 20
1000 ml	Water
	adjust pH to 8 with HCl

TRANSFER BUFFER SOLUTION

2.9 g	Tris-HCl
14.5 g	Glycine
200 ml	Methanol
800 ml	Water

LAEMMLI BUFFER 3X:

2.4 ml	1M Tris-HCl pH6.8
3 ml	20% SDS
3 ml	Glycerol 100%
1.6 ml	β-mercapthoethanol
6 mg	Bromophenol blue
	Add to 10 ml with water

2.2.3.1. SDS PAGE

Take 10 µg of protein for each sample and reconstitute in 1X Laemmli buffer. Heat the samples for 10 minutes at 95°C on a heat block. Before loading the gel, vortex the sample. Load the samples into each well and deposit 5 µl for the protein molecular weight marker. Run the gel in the appropriate apparatus for approximately 50 minutes at 150V or until the tracking dye reaches 1 cm from the bottom of the gel.

REMARK: As previously mentioned, if you load Ni NTA samples, use 200mM Imidazole buffer (20mM Tris, adjust at pH 7.0) instead of Laemmli buffer. This is to elute the proteins from the beads.

2.2.3.2. Protein Transfer

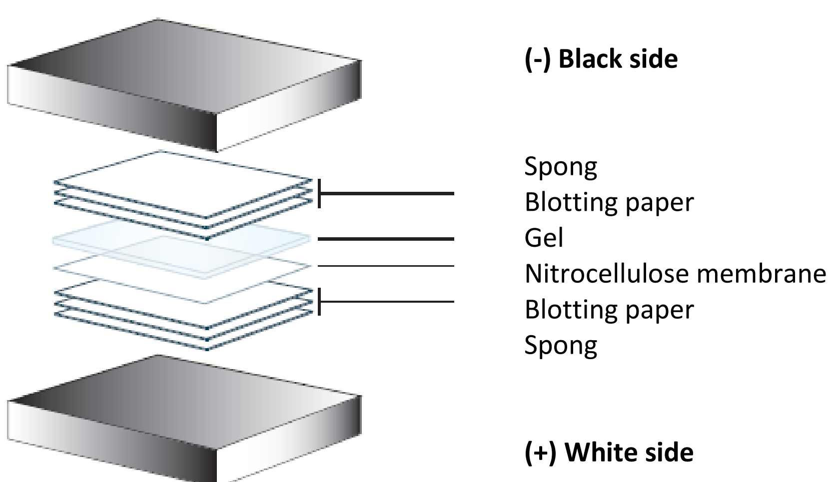


Figure 20: Schema of the order to prepare the wet transfer of the proteins from the gel to the membrane (18)

Remove the gel from the electrophoresis unit. Cut the gel, membrane and blotting papers to the desired size and equilibrate all of them in the transfer buffer solution. Then, pack the transfer casette without any airbubbles in the order that is showed in figure 20 .Transfer proteins for 90 minutes at 600 mA. Incubate the membrane for 10-15 min in Ponceau S reagent to verify the integrity of the transfer. The Ponceau stain is removed from the membrane with several water washes.

2.2.3.3. Incubation with Antibody

The membrane is blocked with a 5% milk in TBST prior to probing with the primary antibody. The membrane is then cut in different parts for the loading control. For the nucleus we cut bellow around 24 kDa and incubate it with the anti-Histone H3 antibody. We expect to detect a band of approximately 17 kDa. The recommended dilution of the antibody for western blotting is 1:2000 in 1% milk TBST (19). This antibody is an isotype of rabbit IgG. For the cytoplasm blot we cut above around 70 kDa is incubated with an antibody blotting against α-tubulin, which is also an isotype of rabbit IgG. The band should be visible at 50 kDa and the recommended dilution ratio is 1:1000 (20). For the incubation with antibody against SUMO 2 we take the whole membrane for blotting. Like the other two antibodies, it is an isotype of rabbit IgG. Blots are incubated overnight at 4°C or for 2 hours at room temperature.

In the next step, the primary antibody solutions are removed and the blots washed 2 times quickly and two times for 20 min with TBST. The membrane can now be incubated under gentle agitation for 1 hour at room temperature with the secondary antibody. The secondary antibody is an anti-rabbit IgG, HRP-linked antibody. After incubation, the blots are washed as previously explained. The wash steps are important to keep the back ground noise as low as possible.

2.2.3.4. DETECTION WITH ENHANCED CHEMILUMINESCENCE

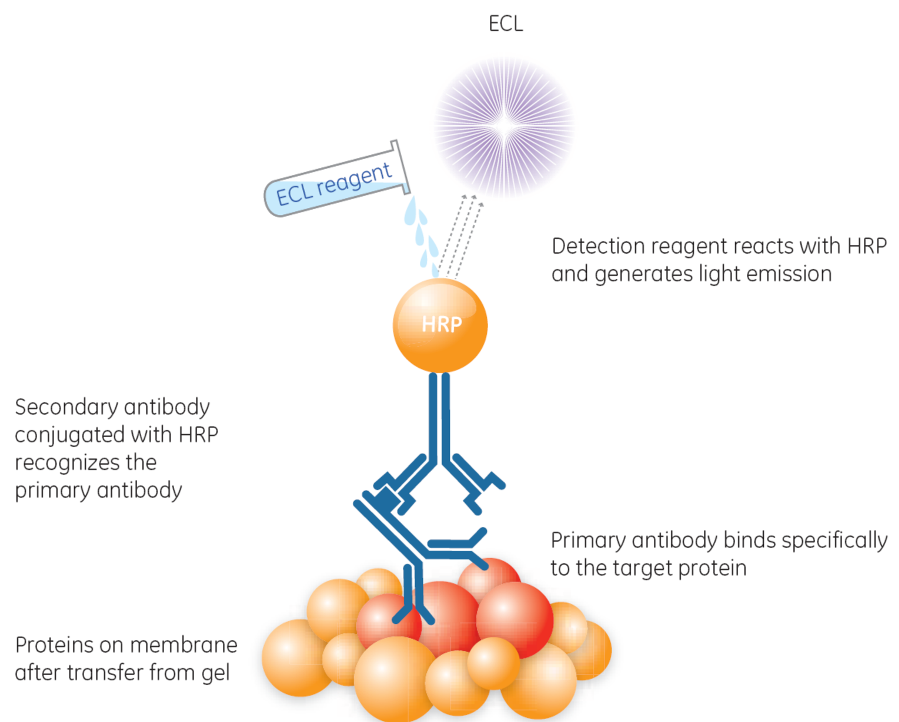


Figure 21: The principle of detection in Western blotting (18). Primary antibody binds to the target protein. Followed with a secondary antibody containing HRP (enzyme horseradish peroxidase). The enzyme linked antibody with HRP reacts with the ECL reagent and resulting emission detected.

Moist the membrane with sufficient ECL developing solution. For the histone blot, use the normal sensitive reagent, and for α-tubulin and SUMO blots, use the highly sensitive solution or a mixture of both. Put the membrane in a vinyl protector which is taped into an X-ray cassette. Expose the film for different lengths of time to develop the film.

2.2.4. In Solution Digestion

Dissolve the protein sample in 50 mM NH_4HCO_3 to a known concentration (the following amounts correlate for 1 mg of proteins.) Add 10 µl of TCEP to the sample to reduce the disulfide bonds. Incubate for 30 minutes at 37°C. Then, add 100 µl of 5 mg/ml chloroacetamide to alkylate the free sulfhydryl groups from the cysteine residues. Incubate for 30 minutes at 37°C. The unracted chloroacetamide is quenched with 10 uL of DTT at 500 mM. The digestion of the reduced and alkylated protein is done by adding 20 µg of trypsin (1:50 ratio to protein). Incubate for 4 h or overnight at 37°C.

2.2.5. Desalting

Formic acid is added to a 1%final concentration. Condition the HLB extraction cartridge 2 times with 50 v/v% methanol (0.1% FA) and 3 times with Millipore water. Load the sample, wash with Millipore water and elute with 50 v/v% methanol (0.1% FA). The sample is then lyophilized to dryness.

2.3. Mass Spectrometry Analysis

First, the samples were analyzed after trypsin digestion on the LTQ-Orbitrap Elite mass analyzer coupled to an Eksigent nanoLC sytem and a nano ESI source.

2.3.1. Sample Preparation

In general, 1 µg of digested protein dissolved in 0.1% formic acid was injected per replicate/CV.

2.3.2. LC Separation System

Prior to separation on the analytical column, the sample is passed through a pre column with the same stationary phase as in the analytical column. The pre column serves to desalt the sample and to remove the coomassie staining. The prevalent length for the pre column is 50 mm and for the analytical column 15 cm.

2.3.2.1. 1D LC

The peptides are eluted from the pre and analytical columns with a gradient of 0-40% acetonitrile in 0.2% formic acid. The flow for the nano LC system rate is 600nL/min. The following tables describe the gradient for general separation.

Table 5: standard gradient for a 70 min run method

Time [min]	0.2% formic acid in water [%]	0.2% formic acid in acetonitrile [%]
0	95	5
56	60	40
57	40	60
58	40	60
59	95	5
70	95	5

Table 6: standard gradient for a 120 min run method

Time [min]	0.2% formic acid in water [%]	0.2% formic acid in acetonitrile [%]
0	95	5
106	60	40
107	40	60
108	40	60
109	95	5
120	95	5

2.3.2.2. 2D LC

The separation of digested proteins can be performed with two-dimensional chromatography. Digested proteins were loaded first on a SCX trap column and then eluted by a sequential ammonium acetate elution (0 – 250 – 500 – 750 – 1000 and 2000 mM) at pH 3.5. The cleaning step of salts in a C18 pre column and analytical column separation remain the same as for the 1D LC systems.

2.3.2.3. COLUMN

LC nano columns are packed in house under Nitrogen (see annex 6.5)

Capillary: ID 52 µm ex Polymicro Technologies (Part #106815017)

ID 150 µm ex Polymicro Technologies (Part #1068150024)

REVERSED PHASE

Stationary Phase: Jupiter 3µ C18 300 Å ex phenomenex 04A-4263

Membrane : SPE Disk EMPORE 2215-C18 (Octadecyl) 47mm ex 3M

CATION EXCHANGE Cartridge

Stationary Phase: Opti guard 5µ SCX 300 Å ex

2.3.3. MASS ANALYZER

Thermo Scientific™ **Orbitrap Elite** mass spectrometer. HCD activation mode with normalized collision energy 30% for SUMO peptides and 35% for common peptides in positive mode. 10^6 ions were accumulated in the Orbitrap before they were passed over to the linear ion trap. The resolutions for the MS is 60000 at 400 m/z and 15000 for the MS/MS. A minimal signal of 10000 is required.

2.3.3.1. FAIMS

Thermo Scientific™ high-field asymmetric waveform ion mobility spectrometry (**FAIMS**) has an interface with a 1.5 mm gap and a curved wall ion inlet between the electrodes. Spray position is diagonal and nitrogen is used as the gas. The temperature in the inner electrode is 70°Cand 90°C for the outer electrode. A V_{high} –of 5000V and V_{low} of +1000V were used as dispersion voltages.

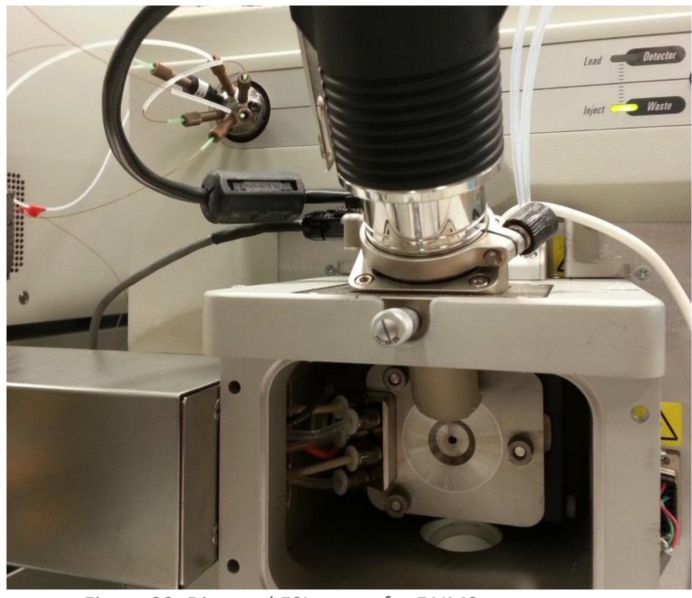


Figure 22: Diagonal ESI source for FAIMS

Before injecting with the FAIMS ESI source, we had to find the optimal position of the spray. Usually, we fix the right position with help of the cameras. But unfortunately, this is not possible with this kind of source. To overcome this we infused a solution of angiotensin (0.2 µg/µL in 50% MeOH in 0.2% formic acid) to manoeuvre them into the optimum diagonal position. We scanned through a CV range to determine the optimal CV range for Angiotensin (peptide observed at an m/z of 450.24 in a 2+ state). We found the optimum had a compensation voltage of about -50V. When this step is done, the position can be adapted.

2.3.3.2. SYSTEM TEST SOLUTION: PROMIX

To test and control the LC-ESI-MS system, we use a solution, called Promix. The Promix is a complex tryptic digested solution, composed of 8 proteins from different species (see protocol 6.6). The corresponding LC-MS analysis is used as a quality check test to ensure that operational system performance is achieved. This includes a defined value of the signal intensity, the control of specific precursor ions and their retention time as well as a comparison to the LC chromatograph of previous promix runs.

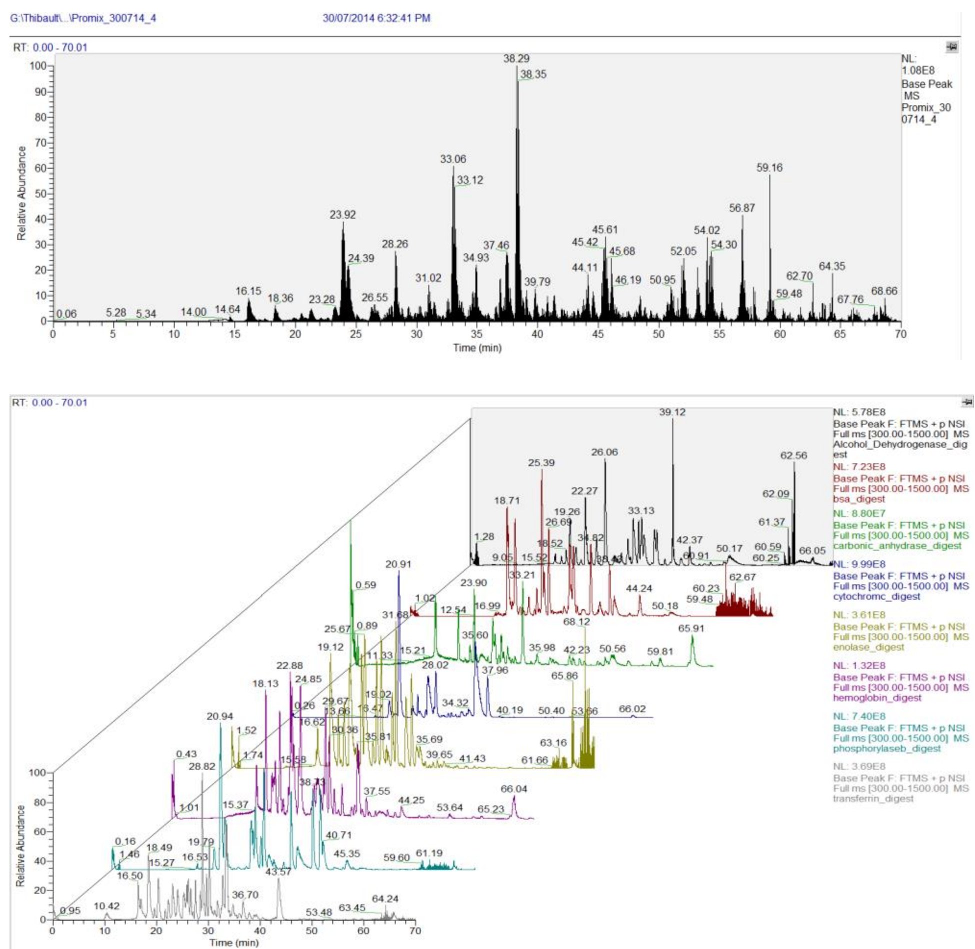


Figure 23: Above example for a LC chromatogram of the Promix solution (July 2014) of 8 tryptic digested proteins. The sample was injected on the Thermo Scientific Orbitrap Elite mass spectrometer. Above are the chromatograms of each digested peptide.

In proteomics, the amount of sample to analyse is usually low. To prevent losing sample due to instrument problems and to have optimal conditions during analysis, the promix is useful.

2.4. Software

RAW data's were acquired with XCalibur. For database searches, clusterings and validations of the results, different software's were used. The following table shows all of them.

Table 7: List of different bioinformatic software's

Instrument control	**Xcalibur**	Thermo (version 2.2, 2011.08.11)	instrument control and data analysis
Data processing	**ProteoProfile**	InHouse software IRIC (version 3.2014.05.23)	Extract peptides from RAW file (peptide card , sds.file)Clustering, remove SUMO tag, recalibrate RAW files
	PEAKS™7	Bioinformatics Solution Inc.(7)	De novo sequencing – database search – PTM Identification
	MASCOT Daemon	matrixscience	Search algorithm for Mass Spectral proteomics peaklists
	Scaffold	Proteome Software (4.3.4)	Visualize and validate complex MS/MS

2.5. Data Processing

We convert .Raw Data files from XCalibur Software to Mascot generic format (MGF) by Mascot Distiller. The database UniprotHuman (37275 sequences) is used with a mass tolerance of ± 15 ppm and MS/MS tolerance ± 0.02 Da. The reason for this relative high tolerance value is due to the fact that we have no set mass to control the calibration. With help of the inHouse software MS/MS editions (part of the ProteoProfile package) we can edit the mgf files and remove the SUMO tags from the MS/MS to aid in SUMO peptide identification. For the identification of the SUMOylated peptides we take the mgf.edited files gereated from MS/MS edition and run the Mascot search again against the UniprotHuman database. For variable modifications we look for GlyGly (K), deamidated (NQ) and SUMO 3 (K). The maximal number of missed cleavages by trypsin was set to 3. In some case a Peaks search with the .Raw files is launched in parallel to the Mascot search. The parameters are the same. To validate the results of the heat shock we set the FDR at 2.0% and we take only the spectra with at least 3 fragments (done manually). For SUMO synthetic peptides we are using a database with the 96 SUMO 1 as well the 96 SUMO 3 peptides.

Figure 24: The general procedure to process the data for SUMO site identification.

A search against a database compares the experimental spectra with theoretical spectra. It they are equivalent the peptide is assigned, if the experimental spectre is more or less the same, it is unassigned. If the sequence is the opposite, it is a DECOY peptide. The false discovery rate (FDR) is the ratio between the decoy peptides and the assigned peptides.

2.5.1. Clustering with ProteoProfile

To compare between different fractions, conditions or replicates, the software ProteoProfile allow us to create peptides cards (plot retention time to m/z plus charge state) and clustering. For the peptide cards, a threshold is set. Threshold is the value of the signal where we can detect peptides and not background. For FAIMS, measures on the Elite Orbitrap the Threshold is in general set at 1000, for noFAIMS it is set at 5000. The reason for the lower threshold with FAIMS is due to the significantly reduced background.

2.6. COMPARISON FAIMS VS NOFAIMS

Basically, complex samples are challenging for mass spectrometry. Background noise makes life complicated and the discovery of a whole proteome impossible. To demonstrate the performance of FAIMS, we injected a tryptic solution of the total cell lysate from HEK293.

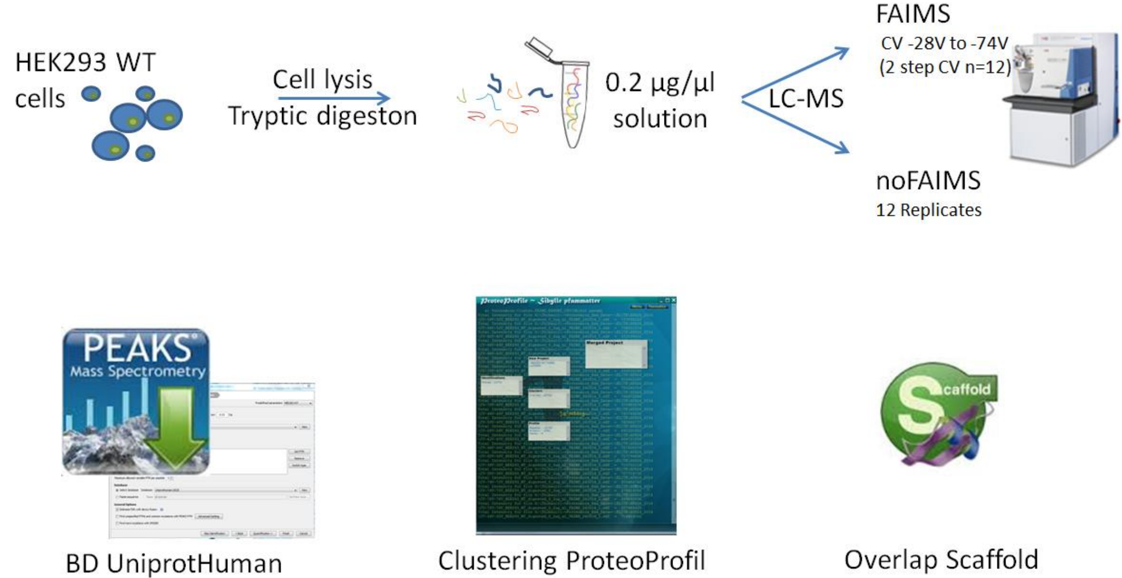


Figure 25: Workflow to demonstrate the performance of FAIMS compared to noFAIMS. The parameter of the mass analyzer were the same. For the validation of the data we used Peaks and for the data base search we used UniproHuman. After this we create the peptides cards as well the clustering with ProteoProfile. To get the overlap of the two conditions we used the Scaffold software.

To identify HEK293 peptides, we injected our solution on the Orbitrap Elite and instead of replicates, we scanned a range of 12 different CV voltages (-28V to -78V in two steps CV) with FAIMS. The common nano LC-MS analysis (noFAIMS) was done with 12 replicates. With this method, we used a TOP5 approach (5 MS/MS) for each CV step (FAIMS) or a TOP12 approach for the replicate (noFAIMS). The amount of protein injected was 1 ug per injections. To compare the two methods, we analyzed at the detected, identified and quantified peptides. Clustering without a database search gave us the peptides that were detected (peptide cards). If we add the identification files in the cluster step, we get the list of identified peptides. For the frequency of quantified peptides we set the limit for the signal to noise ratio at 3. For the false discovery rate (FDR) we took only the unique peptide sequences that were lower than 1.0%.

2.7. Linear Dynamic Range of Detection and Quantitation for SUMO 1 and SUMO 3 Peptides

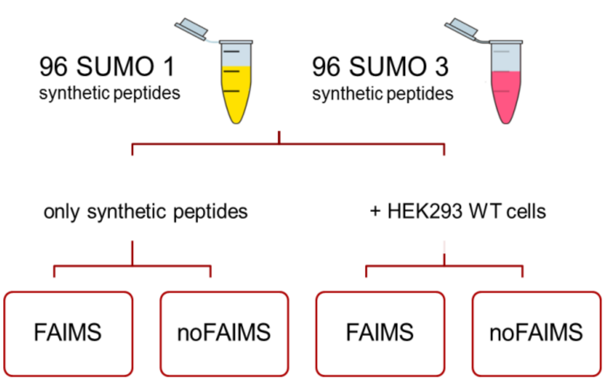


Figure 26: Overview for the 4 different conditions to determine the LOD for SUMO 1 and SUMO 3 peptides with FAIMS and noFAIMS

To determine the limit of detection (LOD) of SUMO peptides, a number of representative SUMO 1 and SUMO 3 synthetic peptides (96 for each isotype) were used. With the synthetic SUMO peptides, we prepared a serial dilution to prepare solutions with final injection concentrations of 20 – 100 – 200 – 400 – 1000 – 2000 fmol. The pure peptides were used for the first injections, and for the second injection set they were spiked in 1 µg of HEK293 cell lysate. The measurements were acquired with FAIMS and noFAIMS. Each condition were conducted in triplicates. For the FAIMS we stepped over the CV range -34V to -85V by acquiring 3 CV per MS (-40V-37V-34V | -49V-46V-43V | -58V-55V-52V | -67V-64V-61V | -76V-73V-70V | -85V-82V-85V).

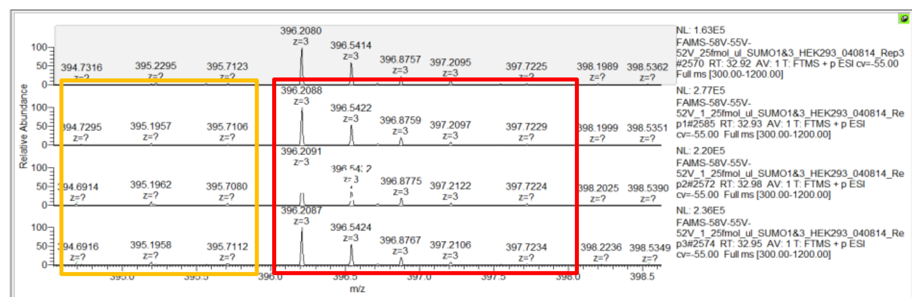


Figure 27: Example of how the limit of detection was manually determined

After clustering for the linearity curves, we looked at each spectrum manually to determine the signal to noise ratio. The procedure was as follows: The average of the intensity (Full MS) for the triplicate was calculated for the lowest concentration point in which the SUMO peptide was detected. The signal of the noise was taken before the signal of the precursor ion of the peptide.

3. Results and Discussion

3.1. Optimization of FAIMS Interface

With FAIMS we have different parameters to optimize. There are no general rules that we can follow to get the best conditions. To optimize the instrument parameters for SUMO peptides, we tested different parameters like the gas flow, temperature of the electrodes and the dispersions voltage.

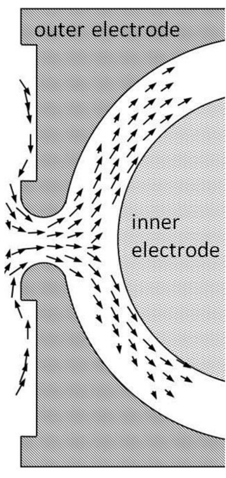


Figure 28: Shema for gas flow at the inlet of the outer and inner electrode

The figure on the left shows a sheme of the FAIMS interface. A gas flow passes through the outer and inner electrode. The electrode are individually adjusted to heat the electrodes. The temperature has to be stable because of the influence by the gas intensity (21). The dispersion voltages is the peak maximum for one waveform cycle and it is applied on the two electrodes. The compensation voltages is applied on the inner electrode to prevent the ions from drifting onto the electrodes.

3.1.1. Gas Flow

Angiotensin infusion was used to determine the optimal gas flow. Moreover, the stability of the signal can fluctuate as well. If we change only the flow parameter we observe a higher Intensity for Angiotensin at 1.5 L/min but the signal is not stable. If the flow is at 2.5 L/min the intensity is lower but the signal more stable over time. The best conditions we obtained was with a 2.0 L/min gas flow. 2.0 L/min has proven to work well since the signal is stable and the intensity is high.

Above all, we are interested in the best parameters for SUMO peptides. To this effect, we infused a solution with 5 synthetic SUMO 3 peptides (200 fmol/µl). The temperature of the inner electrode was 70°C and the outer electrode 90°. As usual, the dispersion voltage was set at -5000V and the ionisation of the spray was set at 3.60 kV. The signal intensity for each individual ion was taken on the optimal CV. All of the SUMO 3 peptides can appear in different charge states.

Table 8: Follow list shows the m/z value of the 5 peptides as 2+ and 3+ precursor ions.

392.54^{3+}	588.30^{2+}	450.57^{3+}	675.35^{2+}	477.93^{3+}	716.40^{2+}	557.60^{3+}	835.89^{2+}	647.39^{3+}	970.50^{2+}

The gas flow rates of 1.5 – 2.0 – 2.5 – 3.0 and 3.5 L/min were tested. There seems to be an increase in intensity for the triple charge peptides at higher flow rate and the opposite for the doubly charged peptides. However this difference does not really have an impact. Out of curiosity and to confirm this observation we did the same with a digested solution of BSA and obtained the same trend.

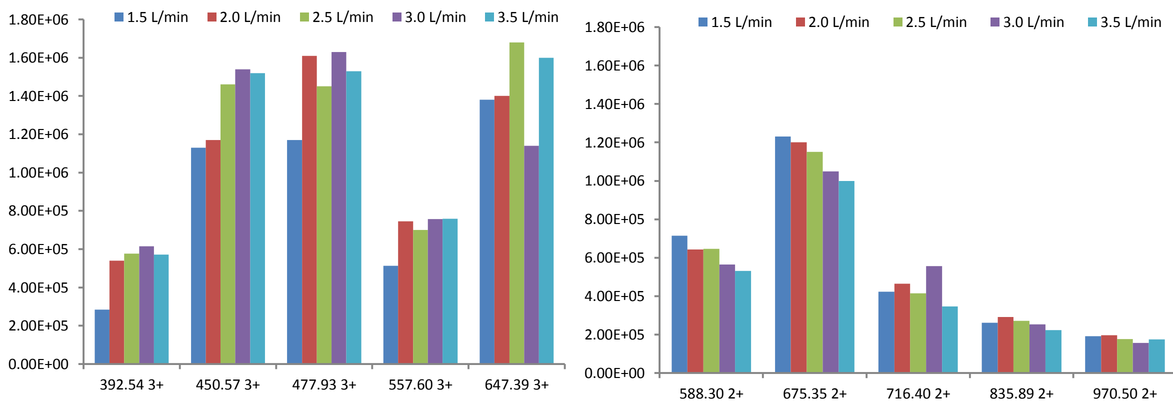
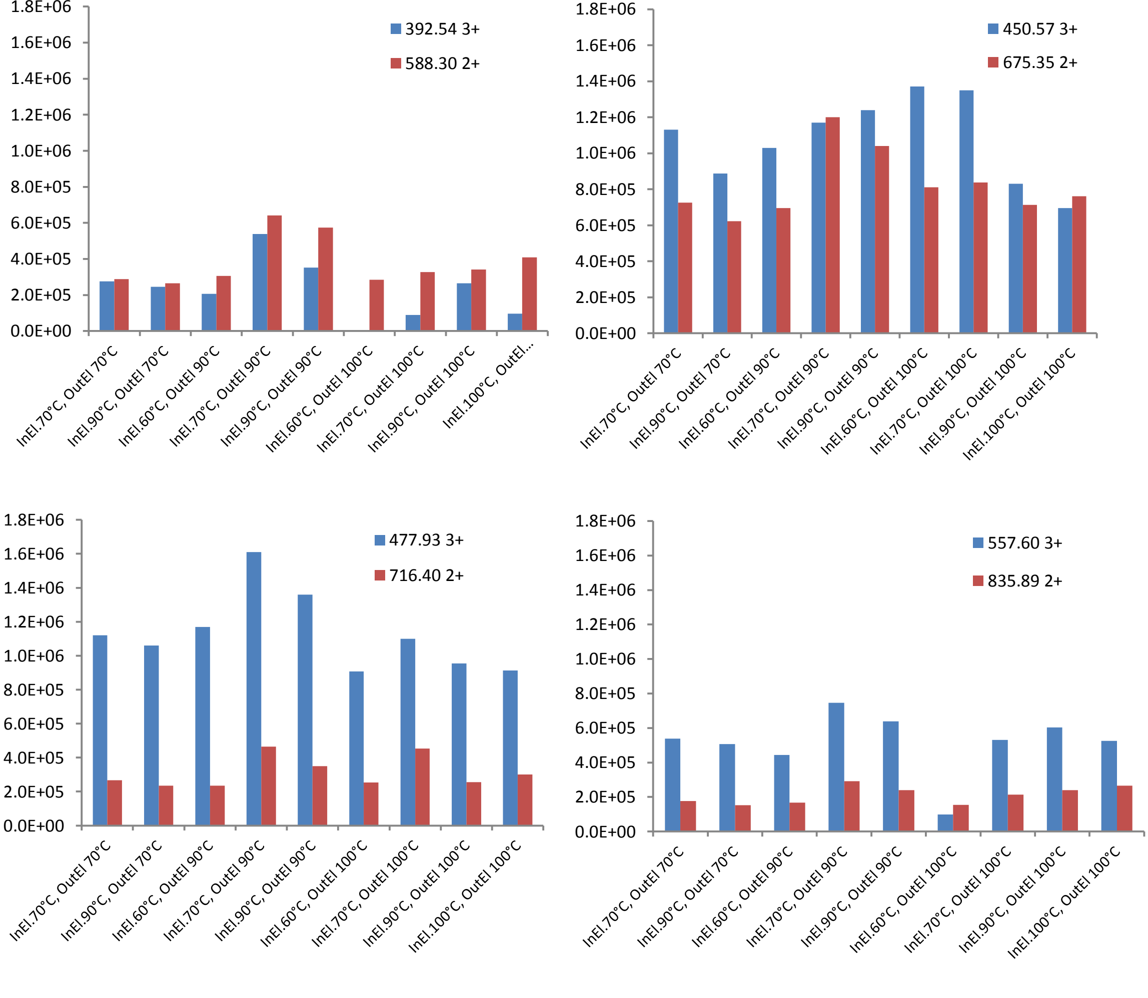


Figure 29: Intensity for the 5 SUMO3 peptides for different gas flow rate. Left the ion has charge 3, right as double charge ion.

3.1.2. Temperature

As mentioned for the flow rate tests, we were using the standard parameter and changed only the temperature of the electrodes. For the parameter experiments we were scanning over a CV range from -95V to -15V. In addition to the intensity of the ions, resolution between peptides is another important aspect. If the CV values for the optimal intensity are the same, the peptides co-elute. When we tested different temperature combinations we observed this effect more for doubly charged peptides. A better resolution can be achieved if the temperature is increased between the electrodes.

To illustrate the effect of the temperature on the intensity we plotted the different temperature combinations against the ion intensity.



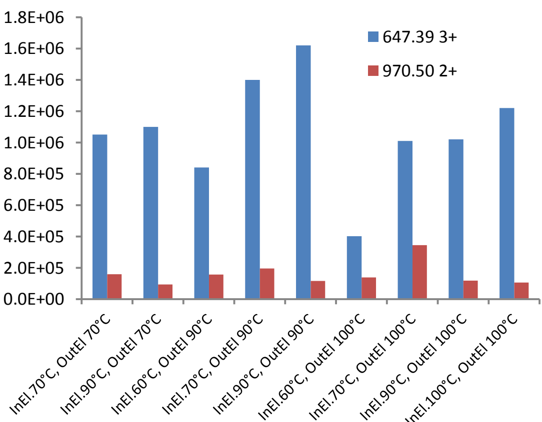


Figure 30: The 5 different intensity profiles for different electrode temperatures

We get similar profiles for each SUMO3 peptide and there is no significant difference. However, it seems that the intensity for the triply charged ions in FAIMS is better. It becomes apparent that the ratio of the 3+ to 2+ ion increase with higher m/z value.

The result tells us that the combination of IE 70°C – OE 90°C is a good choice. The Intensity is more or less the same for all peptides compared to other combinations and we also

have good transmission for the 3+ and 2+ ions. Figure 30 shows a better transmission of the triply charged form of the peptide than for the corresponding doubly charged ions. To summarize, FAIMS has a better transmission and separation of ions for charges states > 2+ (delta CV is larger).

3.1.3. DISPERSIONS VOLTAGE

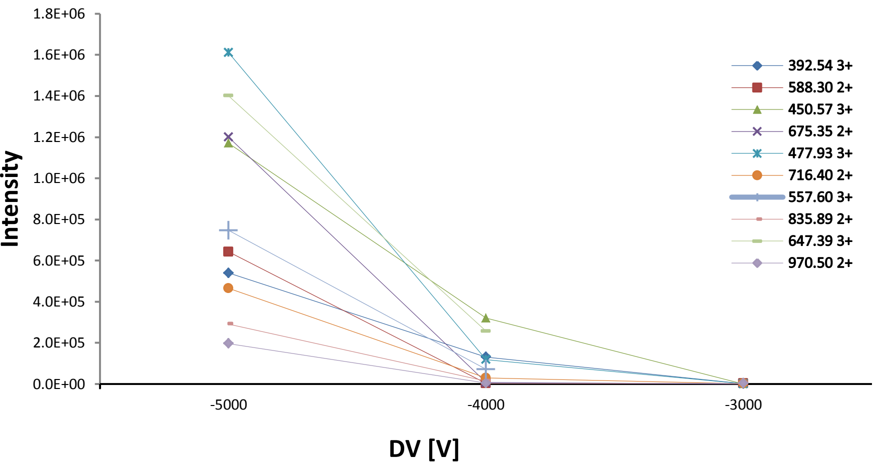


Figure 31: The influence of the DV voltage shows a huge loss of intensity at lower voltages. The results of the graph bellows were obtained at gas flow 2.0L/min and IE70°C – OE 90°C.

Different dispersion voltages show that the best DV is at -5000V. The transmission of ions is most effective when the amplitude of the waveform cycle is at its maximum.

3.1.4. ACETONITRILE

Usually, FAIMS is conducted with inert nitrogen gas. It has been reported that gas modifiers like acetonitrile or ethanol improve sensitivity by increasing signal intensity (22). This is due to the cluster effect between the peptides in the gas phase. Even if we couldn't add modifiers to our gas source, we were using different concentrations of acetonitrile for the dilution (only infusion) but we noticed no significant difference. This proves that the sole addition of organic solvent is not enough for declustering peptides in the gas phase.

3.2. Comparison of LC-MS with and without FAIMS

Equal amount of digested tryptic HEK293 total lysate were injected with FAIMS (12 CV steps) and without FAIMS (12 replicate).

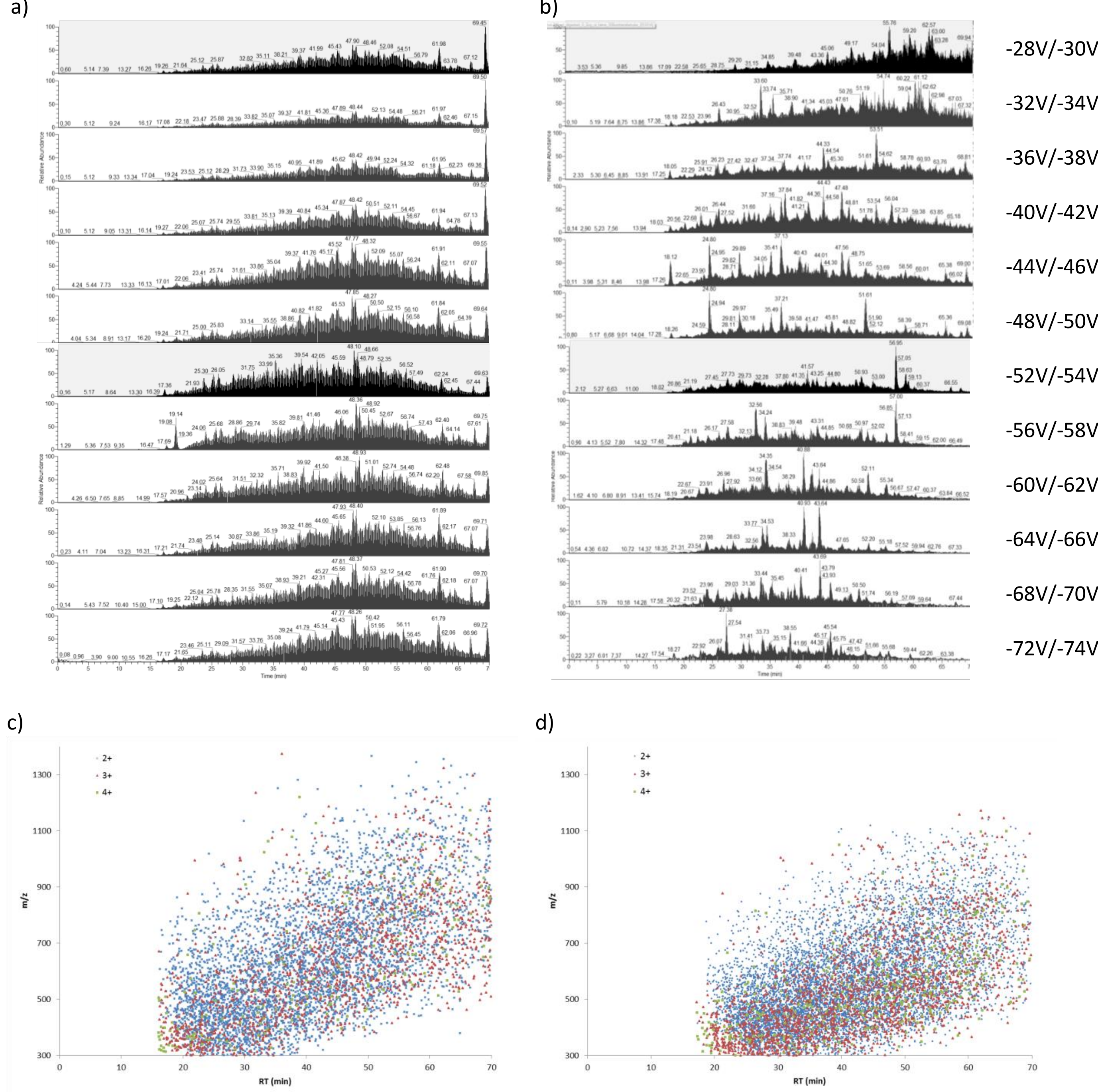


Figure 32: TIC LC-MS Chromatogram without FAIMS (a) for the 12 replicates and with FAIMS (b) for the 12 CV steps. The peptide map of RT versus m/z for all identified peptide ions without FAIMS (c) and with FAIMS (d).

As observed in figure 32, the total ion chromatograms without FAIMS (a) are all similar for the different replicates. With FAIMS (b) we see a change of transmitted ions for the different CV ranges. By comparing figure 32 c) and d) we can see a denser peptide ion population for the CV stepping with FAIMS located in the lower m/z range. This indicates a higher charge state for peptide ions as supported with the greater amounts of the 3+ and 4+ peptide ions for FAIMS. The following table summarizes the number of identified peptide ions in function of their charge.

Table 9: Summary identified peptides ions in function of their charge

Ion charge	**2+**	**3+**	**4+**
without FAIMS	4162	1392	207
with FAIMS	7623	1894	380

With FAIMS we have a markedly higher score on identified peptides ions compared to no FAIMS. Not only do we detect a large number of doubly charged peptides, but also a significant increase of triply charged peptides. For the 4+ population we have noted a decrease of 46% without FAIMS. To have an idea of the intensity range in which we detect these peptide ions, we can look at the distribution of the ions as a function of their signal intensity.

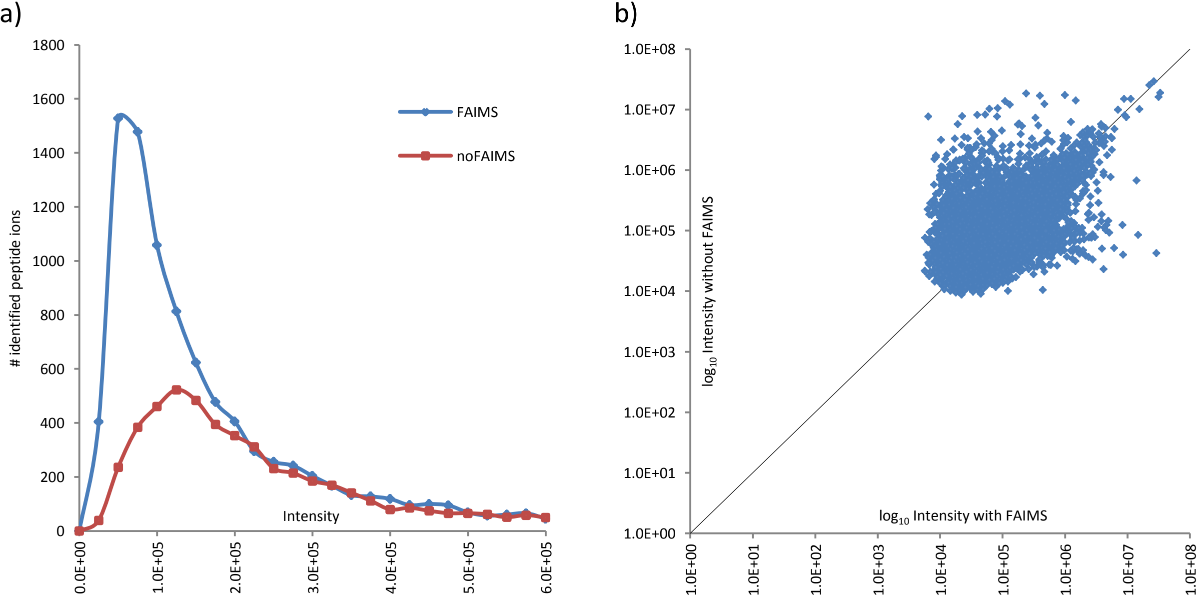


Figure 33: Left plot shows frequency of identified peptide ions belonging to a population of tryptic peptides of HEK293 total cell lysate in function of their intensity with FAIMS and without FAIMS (a). On the right site the intensity of the common peptides with FAIMS and without FAIMS are plotted (b).

The peptides with identical sequence but different PTM modification and charge state are counted individually. As can be seen in figure 33, FAIMS identified more peptide ions, especially peptides of low intensity. Indeed we are able to identify a larger population of ions with low abundance because of the increase of the signal to noise ratio. Furthermore we have no loss of intensity, since 88% of all common peptide ions record less than 30% intensity variation. The scatter plot shows that a major part of all ions are lined up on the diagonal. In this area the intensity with FAIMS and without FAIMS are equal.

When we look at the number of new unique peptides after each injection (figure 34 a), we notice a two-fold higher score for identified peptides with FAIMS after 12 injections. The right graph below shows in how many injections we identified the same peptide ion.

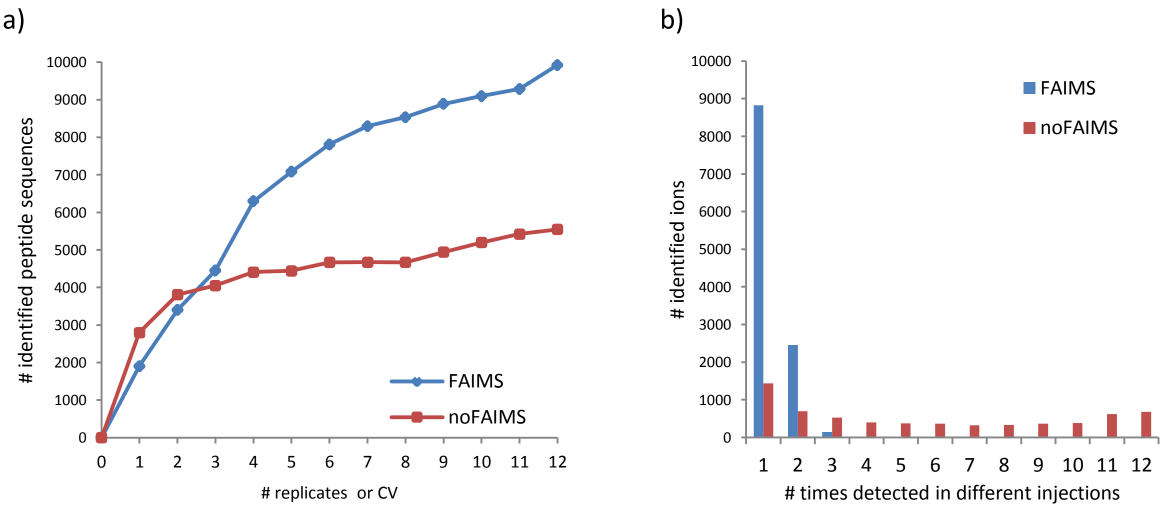


Figure 34: Left (a): number of unique peptide after each injection or CV range. Right (b): How many times we identified the same precursor ion

It is noteworthy that FAIMS transmits about 50% more peptides. Additionally, with FAIMS, starting from the third injection, we identify more peptides than without FAIMS. While we observe without FAIMS a plateau after the fourth replicate, it seems that for FAIMS the slope is positive even after the twelfth CV range. Figure 34 b shows that the same transmitted peptides are not detected more than 3 times in different CV ranges for FAIMS. The majority of them are identified only once while we recorded multiple identification without FAIMS.

Finally, in this experiment we found 8260 unique peptides corresponding to 1165 proteins with the 12 CV scanning method of FAIMS and 4820 peptides (734 proteins) for replicate

injections without FAIMS. The number of protein and peptides shown in the figure below include the 1.0% false discovery rate and do not differ between the different charged states.

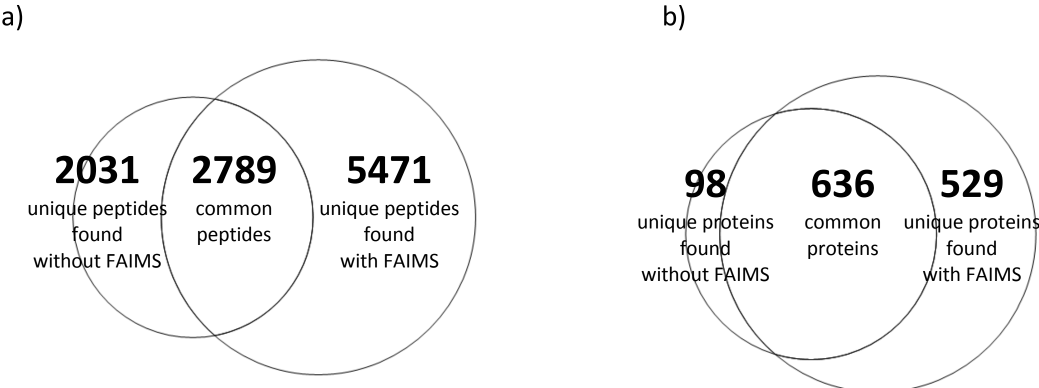


Figure 35: Venn diagram showing identified unique and common peptides (a) as well the proteins (b) resulting from a tryptic digest of HEK293 protein extracts. Analysis was done without FAIMS (12 replicates) and with FAIMS (12 CV steps between -28V and -74V). FDR for proteins and peptides are at 1%.

The non-overlapping areas of the identified common peptides correspond to the unique peptides detected with FAIMS and without FAIMS. These are 34% and 58% of all peptides for FAIMS and without FAIMS, respectively. The overlapping area for FAIMS is only a small fraction of the total and this highlights its performance at different CV ranges. FAIMS is able to transmit more peptide ions to increase peptide identification. To identify peptides, bioinformatics tools compare the MS/MS spectrum with the database spectrum. Because of the low signal to noise ratio of precursor ions, many MS/MS spectra without FAIMS were not acquired. The peptide identification is missing. But we also observe unique peptides upon analysis with FAIMS. There are various reasons for the absence of peptides in FAIMS. First, the transmission of an ion is over 3V (23), but we are not able to detect all ions at the peak top and the tandem mass spectrum cannot be recorded because of the low intensity of the signal. Then it can be that the cycle of the orbitrap is saturated with ions and no more ions are transmitted. This may explain the unique peptides found only without FAIMS. Moreover we were scanning only over a limited CV range from -28V to -74V. Ions with other CV were not transmitted. For the proteins, 42% of the total identified proteins were identified with FAIMS compared to the 8% identified without FAIMS. Even if we miss some peptides with FAIMS, we cover almost all proteins without FAIMS. Furthermore we are able to identify much more.

When we are looking on a MS spectrum of a unique peptide identified only by FAIMS, we observe a significant increase of the signal to noise ratio. The following is an example of a spectrum for the triply charged peptide RDQDNMQAELNR corresponding to the precursor ion m/z 497.23 $^{3+}$. This peptide ion was identified only with FAIMS.

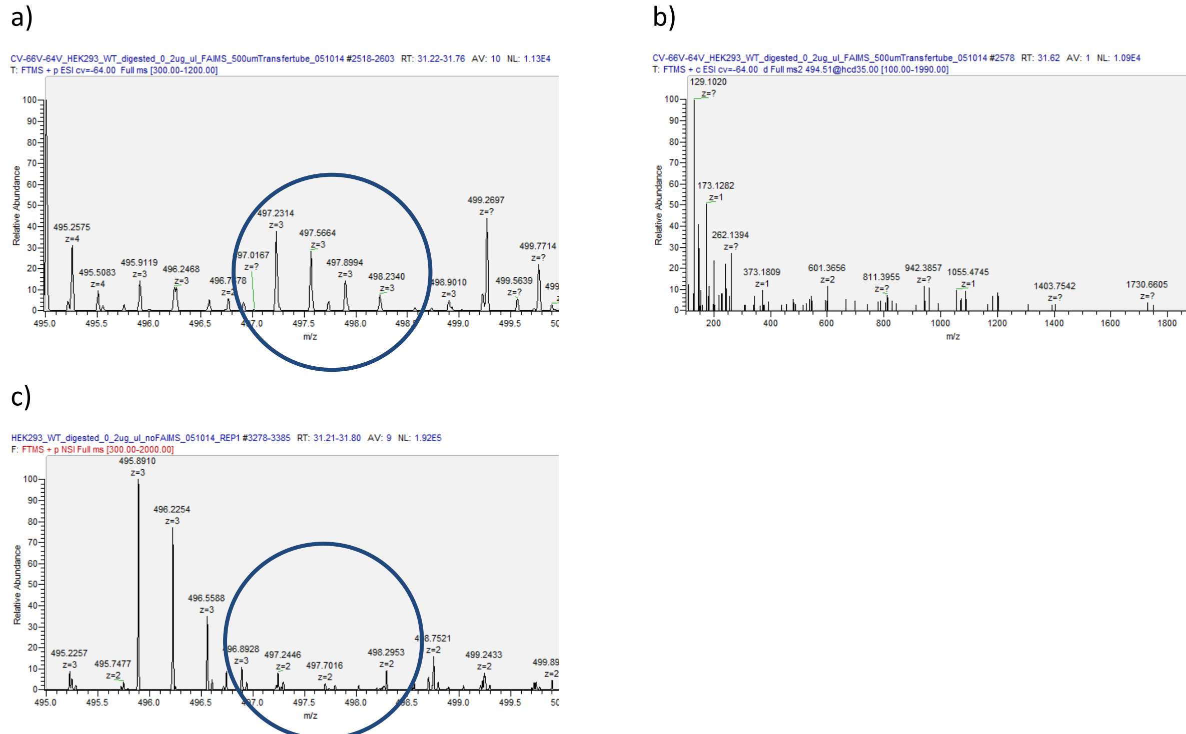


Figure 36: MS of precursor ion m/z 497.23 $^{3+}$ with FAIMS (a) and without FAIMS (c). The right spectrum (c) corresponds to the MS/MS fragmentation of RDQDNMQAELNR with FAIMS.

It is clear that without FAIMS the signal of the precursor ion m/z 497.23 $^{3+}$is undermined by the higher abundance of other precursor ions, while in FAIMS analysis a higher S/N ratio is obtained. These results demonstrate the capacity and efficiency obtainable with FAIMS. The number of identified peptides ions is significantly higher because the samples were fractionated and simplified. The signal to noise ratios increase and more ions can be transmitted by FAIMS. This allows for an improvement in large-scale proteomics studies.

The comparison nano-LC-FAIMS-MS/MS with nano-LC-MS/MS for SUMO target was done with synthetic SUMO peptides that were spiked in total HEK293 cell lysates (see 2.7).

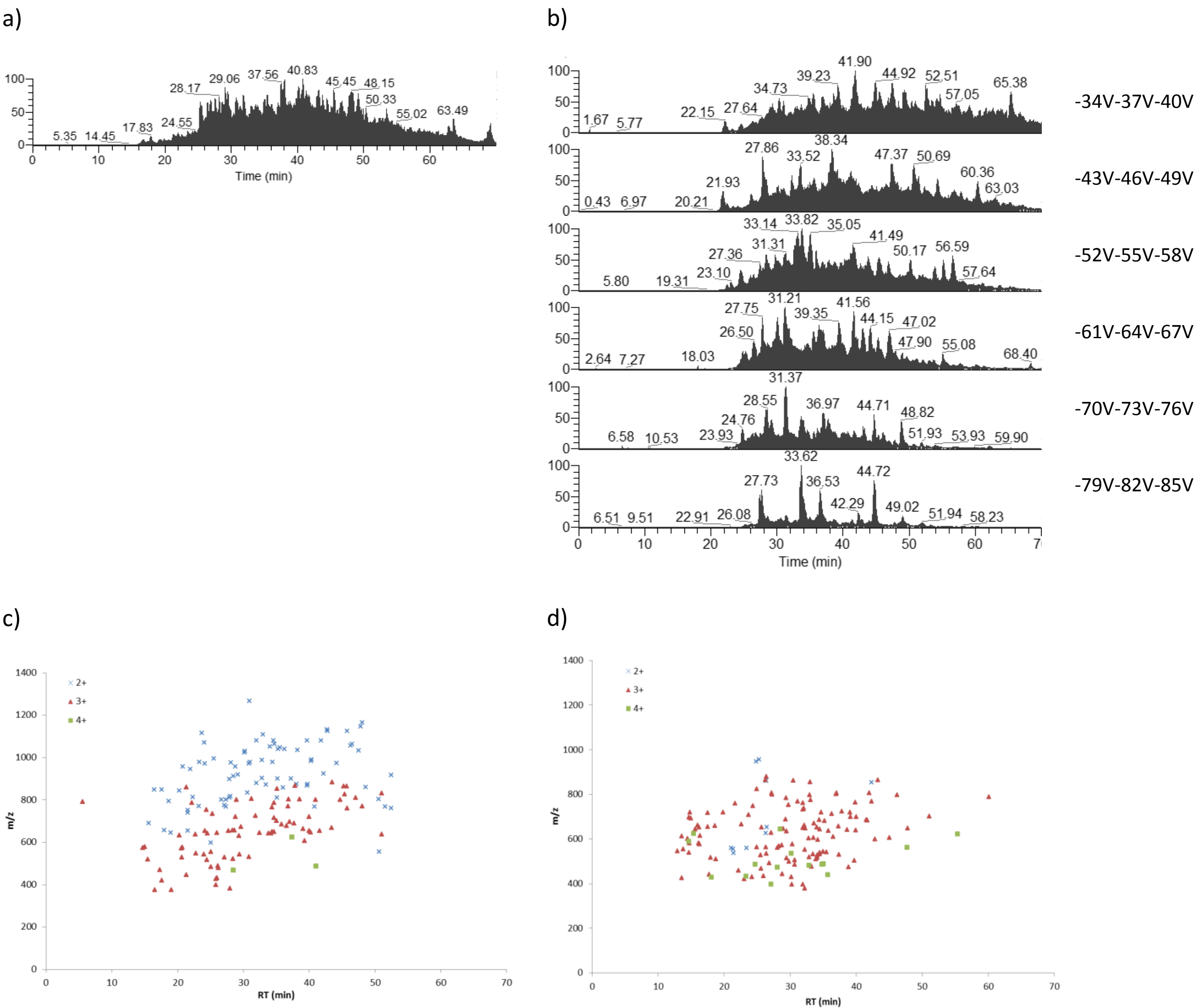


Figure 37: TIC LC-MS Chromatogram for 4 pmol synthetic SUMO peptides spiked in HEK293 without FAIMS (a) and with FAIMS at 6 different CV ranges(b). Below the peptide map of RT versus m/z for all identified peptide ions without FAIMS (c) and with FAIMS (d).

The total ion chromatogram for FAIMS application shows the same tendency as shown previously. Interestingly, we identified a strikingly larger population of all SUMO peptides with FAIMS at the charge state 3+ in peptide maps. This is not the case without FAIMS, where there were huge populations on 2+ SUMO peptides ions with a higher m/z value.

If we look at the CV range of the identified SUMO precursors compared to the tryptic digest of HEK293, we can observe a shift in the measured CV range. The histograms below distinguish between the charge states of identified peptide ions.

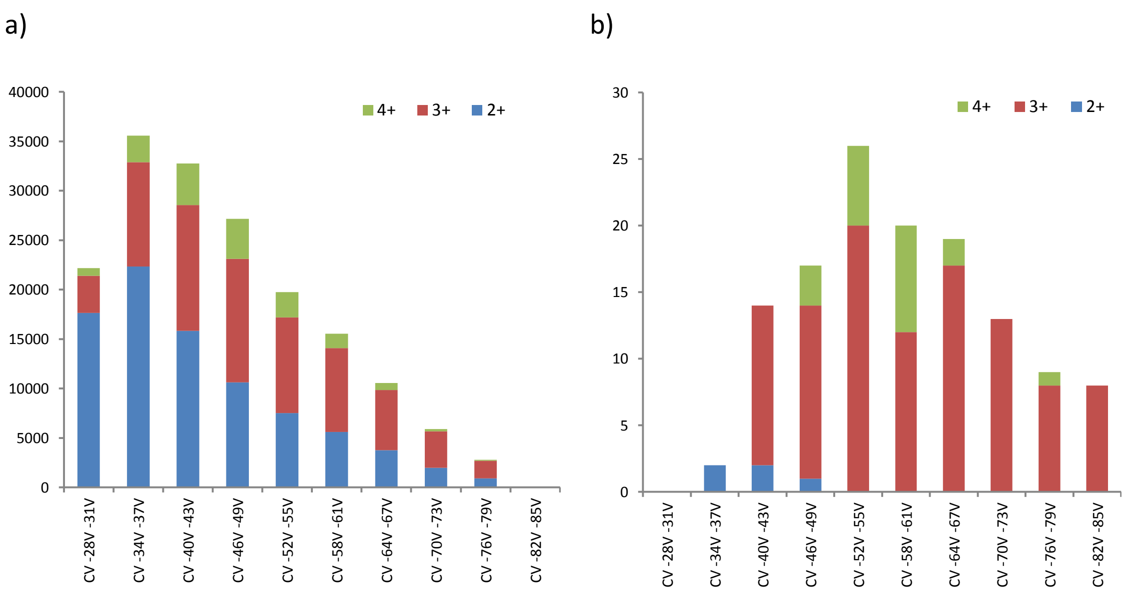


Figure 38: Stacked histogram for tryptic population compared to synthetic SUMO peptides in a compensations voltage range from -28V to -85V.

The SUMO peptides are idenified at a higher CV range. This allows us to identify SUMO targeted peptides. We can separate the low abundance SUMO peptides from the complex HEK293 population without an additional handling step. Extra sample preparation steps could induce loss and increase the risk of contamination. It was also observed that tryptic peptide ions have lower charged states while SUMO peptides were mostly triply charged. The CV distribution helped in the selection of peptides and facilitated the idenfication of SUMOylated sites when a characteristic CV range was used. A CV range of -64V and higher allows for better identification of SUMO targets.

3.3. Enhancement of sensitivity with FAIMS

As indicated, the aim of this study is to improve the sensitivity for the identification of SUMO targets. We were looking for the ideal limit of detection for identifying synthetic SUMO peptides with and without FAIMS. For this, we injected different concentrations of synthetic SUMO peptides to determine the linear dynamic range. The following example illustrates the synthetic peptide TSDADIKSSETGAFR conjugated to either SUMO 1 or SUMO 3.

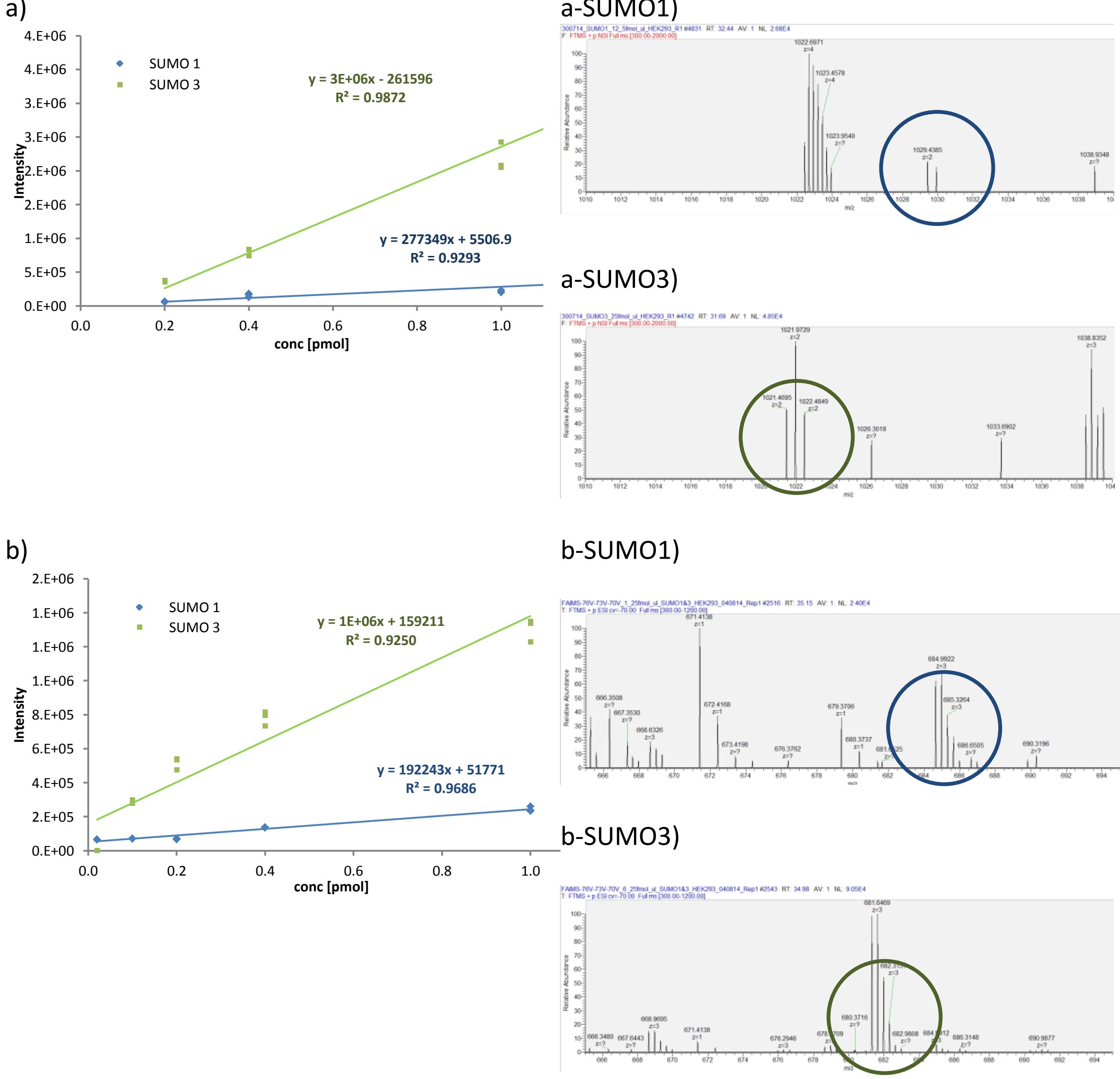


Figure 39: The peptide TSDADIKSSETGAFR was synthetized with the modification SUMO 1 (residue EQTGG) and SUMO 3 (residue NQTGG). The linearity curve without (a) and with (b) FAIMS are on the left site. On the right side the MS spectres of the precursor ions correspond to the lowest concentration of detected ions. Without FAIMS we detected SUMO 1 and SUMO 3 at m/z 1029.4385 $^{2+}$ and m/z 1021.9729 $^{2+}$ respectively. With FAIMS we detected m/z 684.9922 $^{3+}$ and 681.6469 $^{3+}$ precursor ions.

With and without FAIMS we get a linear range, even if the SUMO 1 peptide shows a lower intensity. Without FAIMS the precursor ions were detected as doubly charged ions, while with FAIMS they were detected as triply charged ions. However, although the MS spectra at higher m/z for 1029.4385 2+ and 1021.9729 2+ are less complex, we have an increase of signal intensity compared to the neighbouring ions with FAIMS for the triply charged ions 684.9922 3+ and 681.6469 3+. The above figure shows a limit of detection of 217 fmol for SUMO 1 (S/N 1 @ 100 fmol) and 213 fmol for SUMO 3 (S/N 3 @ 200 fmol) without FAIMS. With FAIMS the LOD are 19 fmol for SUMO 1 (S/N 3 @ 20 fmol) and 5 fmol for SUMO 3 (S/N 55 @ 100 fmol). Thus it appears that FAIMS is able to detect low abundance peptides and the linear range starts at low concentrations.

To summarize all obtained results, LOD values without FAIMS are plotted against FAIMS.

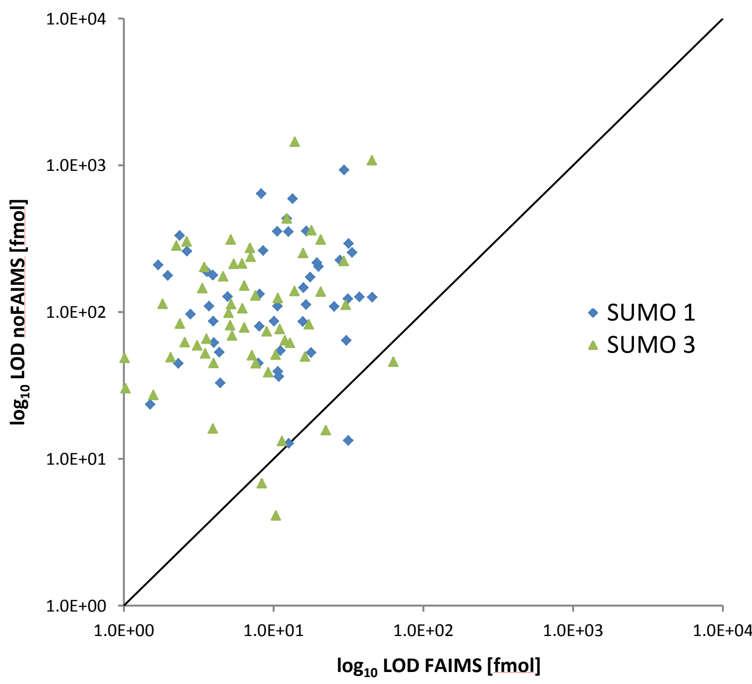


Figure 40: Limit of detection obtained from synthetic SUMO peptides spiked in digested HEK293 total lysat. LOD values obtained by FAIMS are plotted against LOD without FAIMS

Most often, we get a lower limit of detection with FAIMS. FAIMS compared to no FAIMS is on average 18 times more sensitive. This was expected from the literature (2) and our results confirm it. For FAIMS we were at an average limit of detection of 12 fmol compared to 183 fmol without FAIMS. Data obtained from this experiment and previous experiments, help us in improving the detection of SUMO target sites in samples.

3.4. Heat Shock HEK293-SUMO 3

80 million cells per condition were cultured for the large-scale experiment. This corresponds to 32 mg proteins (16 mg nucleus and 16 mg cytoplasm). After Ni-NTA purification, 100 µg of nuclear proteins were purified.

3.4.1. Western blot Analysis

If antibodies are available, western blot is a favored method to validate and confirm the results before injecting on the mass spectrometer. In our case, we observed the changes caused by the heat shock and monitored the changes in the enrichment of SUMO peptides. For the heat shock, we wanted to investigate the trend of sumoylation over a time range. In the paper from Knejzlik and Co. maximal accumulation of SUMOylation was found to be after 45 min at 42°C (24). In our case, the heat shock of HEK293 mutant SUMO3 cells was done at 43°C.

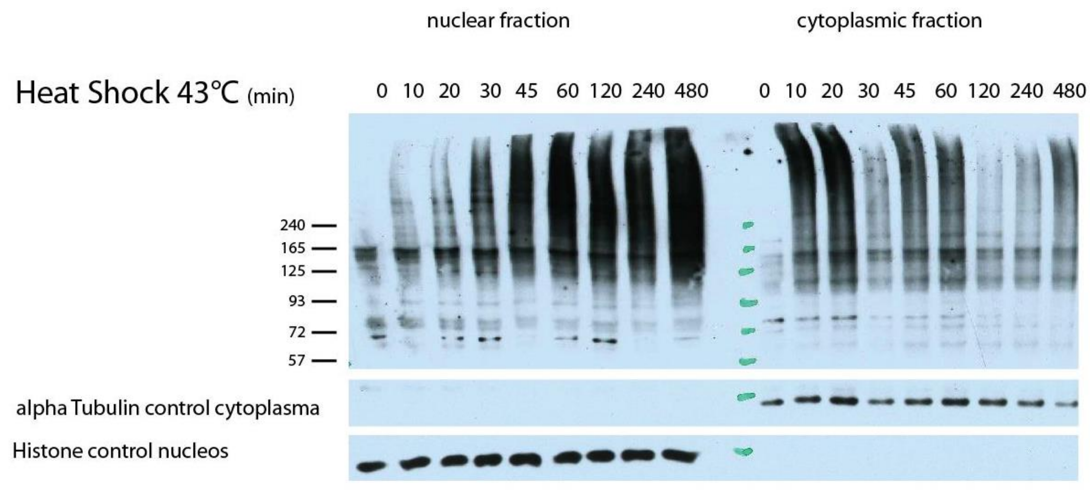


Figure 41: The effects of heat shock (43°C) on HEK293 SUMO3 cells over an 8hr period.

In the first experiment we were looking for cell response to heat in the form of SUMOylation. To study this, time points between 0 to 8 hours were taken. The trend shows a maximum of response after 60 minutes as well as at 8 hours in the nucleus. For the cytoplasm, an early accumulation followed by a decrease in SUMOylated proteins can be observed. For the loading control, we got similar signals for alpha tubulin and histone in all fractions and this shows that the nucleus and the cytoplasm were efficiently separated. We can also see that the SUMO peptides are more enriched in the nucleus. For Ni-NTA purification (6.3) a heat shock of

one hour was used. Increased SUMOylation was at both 45 and 60 minutes. During the mass spectrometry analysis, we injected only the nuclear fraction.

3.4.2. Mass Spectrometry Analysis

Based on the previously mentioned results, 4 replicates were injected without FAIMS. After 4 replicates found no new peptides. With FAIMS the CV range from -36V to -82V was over scanned. The samples were prepared as descripted in paragraph 2.2 of the methods section. For the LC-MS/MS analysis, a 20 cm long analytical column with a C18 stationary phase was used with a gradient performed of 120 min. As shown by the Western Blot (point 6.3), the amount of SUMOylated peptides is weak.

Table 10: Identified SUMO3ylated peptides for heat shock experiment

FAIMS		noFAIMS	
0 min	**60 min**	**0 min**	**60 min**
APGVANK**K**IHLVVLVKPSGAR *EGV**K**TENDHINLK* KKGAMEELE**K**ALSCPGQPSNCVTIPR LQE**K**LSPPYSSPQEFAQDVGR M**K**TENLKKR	ANEKPTEEVKTENNNHINL**K**VAGQDGSVVQFK ATMHL**K**QEVTPR *EGV**K**TENDHINLK* KDDGENA**K**PIKK KDQLIT**K**CNEIESHIIK K**K**EPPKELR KVI**K**MESEEGKEAR L**K**SAQCGSSIIYHCKHK		ANE**K**PTEEVKTENNNHINLK IVHQSLNIMNSFSQKV**K**IQQIRSLSEDVR *KVI**K**MESEEGKEAR* NTELKKLQIFGAGP**K**VVGLAMGTK

Finally, we identified 12 SUMO sites in total for FAIMS (5151 unique peptides) and 4 without FAIMS (4045 unique peptides). A higher signal to noise ratio for the identified peptides with FAIMS was observed. The separation was done only in a 1D LC system. There were several reasons for this. In a recent published paper from Vertegaal and Co. (25) they used a 1D LC method for the identification of SUMOylation sites. In addition, during the injection of the Promix solution in a 1D LC system compared with the 2D LC system, we lost a lot of transmission (the intensity was 2times less with 2D LC than 1D LC). Unfortunately we can only load 1 µg proteins on a 1D LC system compared to up to 3 µg on the 2D LC system. The fact that SUMO proteins are present less than 1% in a cell (26) limits the success of direct identification from complex mixtures without effective enrichment of SUMO peptides, like with immunoprecipitation. The low number of SUMOylated proteins is also due to the fact that we did only used one biological replicate as well as only one technical replicate. The probability of identifying more SUMO peptides with different replicates increases significantly after one replicate. However, we could only show that the identification of SUMO peptides was higher for nanoLC-FAIMS-MS than for nanoLC-MS.

4. Conclusion

Preliminary experiments using FAIMS with 5 synthetic SUMOylated peptides showed an optimized parameter for a better transmission and resolution for the triply charged peptides at a higher CV range as well as a more stable signal. FAIMS benefits from this high performance and allows for protein identifications without the loss of any signal intensity. We also showed that the number of synthetic SUMO peptides had limit of detection 15 times lower with FAIMS than without FAIMS. Moreover, the triply charged state for SUMOylated peptides is the most favourable form. The specific CV range distribution of SUMO peptides facilitate their detection when present in complex tryptic digests of HEK293 cell extracts. Finally, we demonstrated that there is an accumulation of SUMOylation in cells that undergo a heat shock treatment and we were able to identify twice the amount of SUMOylated peptides with FAIMS than without FAIMS.

5. Literature

1. **Flotho A, Melchior F.** *Sumoylation: A Regulatory Protein Modification in Health and Disease.* Annu. Rev. Biochem. (2013) 82:357-385.

2. **Creese A, Shimwell N, Larkins K, Heath K, Cooper H.** *Probing the Complementarity of FAIMS and Strong Cation Exchange Chromatography in Shotgun Proteomics.* J.Am. Soc. Mass Spectrom. (2013) 24: 431-443.

3. **Lamoliatte F, Bonneil E, Caron-Lizotte O, Wildemann D, ZerwecknJ, Wenshuk H, Thibault P.** *Targeted Identification of SUMOylation Sites in Human Proteins Using Affinity Enrichment and Paralog-specific Reporter Ions.* Molecular & Cellular Proteomics (2013) 12, 1-15.

4. **O, Kerscher.** *What's your function? New insights through SUMO interacting motifs.* EMBO Rep 8, 2007, 550-555.

5. **Wilkinson K, Henley J.** *Mechanisms, regulation and consequences of protein SUMOylation.* Biochem. J. 2010, 438, 133-145.

6. **Saitoh H, Hinchey J.** *Functional Heterogeneity of Small Ubiquitin-related Protein Modifiers SUMO-1 versus SUMO-2/3.* J.Biol.Chem.2000, 275: 6252-6258.

7. **Matic I, Schimmel J, Hendriks I, van Santen M, van de Rijke F, van Dam H, Gnad F, Mann M, Veregaal A.** *Site-Specific Identification of SUMO-2 Targets in Cells Reveals an Inverted SUMOylation Motif and a Hydrophobic Cluster SUMOylation Motif.* Molecular Cell 39, 641-652, August 27, 2010.

8. **Aebersold R., Mann M.** *Mass spectrometry-based proteomics.* Natur Vol 422 (13 march 2003), p.198-207.

9. **Hustoft H, Malerod H, Wilson S, Reubsaet L, Lundanes E, Greibrokk T.** *A Critical Review of Trypsin Digestion for LC-MS Based Proteomics.* ISBN 978-953-51-0070-6, Published: February 24, 2012, capter 4.

10. **J, Wilkinson.** *Fragmentation of Polypeptides by Enymic Methods.* Practical Protein Chemistry, A Practical Approach, IRL Press Oxford (1986).

11. **Fröhlich T, Arnold G.** *A newcomer's guide to nano-liquid-chromatography of peptides.* Methods Mol Biol. 2009;564.

12. **Thermo Fisher Scientific Inc.** *Product Specifications Orbitrap Elite - High-Field Orbitrap Hybrid Mass Spectrometer.* 2011 PS30229_E 05/11G.

13. **Galisson F, Mahrouche L, Courcelles M, Bonneil E, Meloche S, Chelbi-Alix M, Thibault P.** *A Novel Proteomics Approach to Identifiy SUMOylated Proteins and Their Modification Sites in Human Cells.* Molecular and Cellular Proteomics 10, 2011, 1-15.

14. **Reinders J, Lewandrowski U,Moebius J,Wagner Y, Sickmann A.** *Challenges in mass spectrometry-based proteomics.* Proteomics 2004, 4, 3686–3703.

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

16. **G, Bridon.** *Étude de la voie de signalisation de l'insuline chez la drosophile par une approche phosphoprotéomique.* Thèse présentée à la Faculté des études supérieures et postdoctorales en vue de l'obtention du grade de Doctorat, avril 2012.

17. **Purves R, Guevremont R.** *Electrospray Ionization High-Field Asymmetric Waveform Ion Mobility Spectrometry - Mass Spectrometry.* Anal. Chem. 1999, 71, 2346-2357.

18. **Sciences, Handbooks from GE Healthcare Life.** *Western Blotting Principles and Methods.* 28-9998-97 AC, Figure 6.2 p.74.

19. **Cell Signaling Technology, Inc.** *Histone H3 Antibody #9715.* Datasheet rev. 06/02/11.

20. **Cell Signaling Technology Inc.** *α-Tubulin Antibody #2144.* Data sheet rev. 1/15/10.

21. **Thermo Scientific.** *FAIMS Operator's Manual.* 70111-97134 Revision B June 2007.

22. **Purves R, Ozog A, Ambrose S, Prasad S, Belford M, Dunyach J.** *Using Gas Modifiers to Significantly Improve Sensitivity and Selectivity in a Cylindrical FAIMS Device.* J. Am. Soc. Mass Spectrom. (2014), DOI: 10.1007/s13361-014-0878-z.

23. **Venne K, Bonneil E, Eng K, Thibault P.** *Improvement in Peptide Detection for Proteomics Analyses Using NanoLC-MS and High-Field Asymmetry Waveform Ion Mobility Mass Spectrometry.* Anal.Chem. 2005, 77,2176-2186.

24. **Castoralava and Brezinova D, Sveda M, Lipov J, Ruml T, Knejzlik Z.** *SUMO-2/3 conjugates accumulating under heat shock or MG132 treatment result largely from new protein synthesis.* Biochemica et Biophysica Acta 1823 (2012) 911-919.

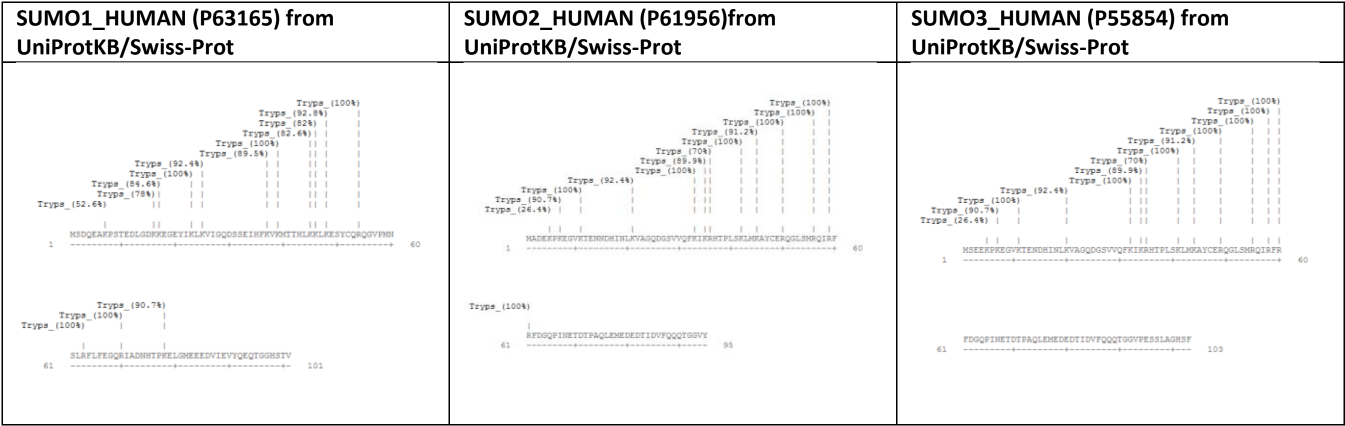
25. **Hendriks I, Souza R, Yang B, Verlaan-de Vries M, Mann M, Vertegaal A.** *Uncoverin global SUMOylation signaling networks in a site-specific manner.* nature structural & molecular biology 14 September 2014.

26. **J, Wohlschlegel.** *Identification of SUMO-Conjugated Proteins and their SUMO Attachment Sites Using Proteomic Mass Spectrometry.* Methods in Molecular Biology: SUMO Protocols (2009), Vol 497.

6. Annex

6.1. UniProt Data's SUMO proteins

Table 11: Cleave sites for SUMO isoforms by tryptic digestion



6.2. Limit of detection

Table 12: List of LOD values for SUMO spiked in HEK293

	CV FAIMS	SignalBackground MSn	Signal MSn conc	conc pmol	LOD fmol	LOQ fmol	*factor*	SignalBackground MSn noFAIMS	Signal MSn conc	conc pmol	LOD fmol	LOQ fmol
AFITNIPFDVKWQSLK	-49	2.3E+03	1.1E+04	0.02	**12**	**41**	*35*	1.6E+04	4.3E+04	0.40	**434**	**1445**
APGQHGGFVVTVKQER	-49	3.1E+03	6.7E+04	0.10	**14**	**46**	*10*	1.1E+04	2.4E+04	0.10	**139**	**464**
APTASQERPKEELGAGR	-43	1.7E+03	3.3E+04	0.02	**3**	**10**	*19*	1.1E+04	5.5E+04	0.10	**59**	**197**
AYQKQPTIFQNK	-73	1.4E+03	2.8E+03	0.02	**30**	**101**	*4*	1.1E+04	5.8E+04	0.20	**112**	**374**
EGVKTENDHINLK	-64	6.8E+02	2.4E+03	0.02	**17**	**112**	*5*	1.0E+04	3.7E+04	0.10	**83**	**275**
EKLEMEMEAAR	-73	1.2E+03	6.3E+03	0.02	**12**	**40**	*5*	9.7E+03	9.1E+04	0.20	**64**	**214**
EKLTADPDSEIATTSLR	-49	1.8E+03	1.3E+04	0.02	**8**	**28**	*1*	2.0E+04	8.9E+05	0.10	**7**	**23**
ELSEVKNVLEK	-67	1.7E+03	1.3E+04	0.02	**8**	**25**	*6*	1.0E+04	6.9E+04	0.10	**45**	**149**
EMSGSTSELLIKENK	-40	2.1E+03	4.0E+04	0.10	**16**	**53**	*16*	1.6E+04	3.9E+04	0.20	**253**	**842**
FQIIKQEPMELESYTR	-46	2.8E+03	4.8E+04	0.02	**3**	**11**	*59*	2.5E+04	7.3E+04	0.20	**202**	**674**
HHSLALTSFKR	-67	3.0E+03	8.2E+04	0.10	**11**	**37**	*7*	1.7E+04	6.8E+04	0.10	**76**	**255**
HTPLSKLMK	-49	2.9E+03	2.8E+04	0.20	**63**	**210**	*1*	9.0E+03	5.9E+04	0.10	**46**	**153**
IESESKGIKGGK	-52	1.8E+03	1.4E+05	0.02	**1**	**3**	*228*	4.4E+03	7.4E+03	0.10	**179**	**596**
IFDEEPANGVKIER	-43	1.8E+03	6.0E+03	0.02	**18**	**60**	*20*	1.8E+04	3.1E+04	0.20	**361**	**1203**
IGAIKQESEEPPTK	-61	1.9E+03	2.2E+04	0.02	**5**	**17**	*60*	1.5E+04	2.8E+04	0.20	**312**	**1041**
IIVDELKQEVISTSSK	-52	1.9E+03	5.5E+03	0.02	**21**	**69**	*15*	2.0E+04	3.9E+04	0.20	**312**	**1039**
IKDEPDNAQEYSHGQQQK	-52	1.3E+03	1.0E+04	0.02	**8**	**25**	*17*	9.8E+03	2.3E+04	0.10	**130**	**433**
IKEDDAPR	-70	1.3E+03	3.9E+04	0.02	**2**	**7**	*24*	6.0E+03	3.7E+04	0.10	**49**	**164**
IKSGEVAEGEAR	-76	1.2E+03	7.1E+04	0.02	**1**	**3**	*48*	1.1E+04	6.9E+04	0.10	**49**	**162**
IKVEPASEKDPESLR	-58	1.8E+03	4.5E+04	0.02	**2**	**8**	*35*	1.4E+04	5.1E+04	0.10	**83**	**278**
INEILSNALKR	-67	2.6E+03	5.7E+04	0.10	**14**	**46**	*104*	3.7E+04	1.5E+05	2.00	**1449**	**4830**
KDDEVQVVR	-64	1.8E+03	5.9E+04	0.10	**9**	**30**	*8*	6.5E+03	2.6E+04	0.10	**74**	**246**
KNAALVTR	-79	5.6E+02	9.4E+03	0.02	**4**	**12**	*18*	6.4E+03	2.9E+04	0.10	**66**	**219**
KVEEDEAGGR	-85	1.9E+03	7.3E+04	0.02	**2**	**5**	*17*	3.8E+03	4.2E+04	0.10	**27**	**91**
LKEDVLEQR	-73	1.6E+03	2.4E+04	0.02	**4**	**13**	*4*	9.7E+03	1.8E+05	0.10	**16**	**54**
LLVHMGLLKSEDK	-43	1.7E+03	1.6E+04	0.02	**6**	**21**	*12*	1.1E+04	8.6E+04	0.20	**79**	**262**
MKFNPFVTSDR	-85	1.4E+03	3.3E+04	0.02	**3**	**8**	*24*	1.8E+04	8.7E+04	0.10	**62**	**208**
MNKSEDDESGAGELTR	-55	1.6E+03	9.0E+03	0.02	**11**	**36**	*12*	1.1E+04	2.8E+04	0.10	**124**	**414**
NISIKQEPK	-67	1.6E+04	1.6E+05	0.10	**29**	**98**	*8*	1.1E+04	1.5E+05	1.00	**222**	**741**
NLLHDNELSDLKEDKPR	-64	1.7E+03	4.5E+03	0.02	**22**	**75**	*1*	1.2E+04	2.3E+05	0.10	**16**	**52**
NLYKNVILENYR	-52	1.7E+03	9.7E+03	0.02	**10**	**35**	*0*	1.5E+04	2.2E+05	0.02	**4**	**14**
NSIDASEEKPVMR	-55	1.4E+03	4.0E+03	0.02	**21**	**69**	*7*	8.9E+03	3.9E+04	0.20	**138**	**459**
QFNKLTEDK	-73	1.3E+03	6.3E+03	0.02	**13**	**43**	*5*	1.4E+04	6.6E+04	0.10	**62**	**206**
SAEEVEEIKAEK	-52	1.8E+03	2.1E+04	0.02	**5**	**17**	*22*	1.5E+04	3.9E+04	0.10	**113**	**376**
SGLKHELVTR	-67	1.5E+03	1.7E+04	0.02	**5**	**18**	*13*	8.3E+03	3.6E+04	0.10	**69**	**231**
STPKEETVNDPEEAGHR	-46	2.1E+03	2.0E+04	0.02	**6**	**21**	*24*	8.9E+03	1.8E+04	0.10	**151**	**504**
STYADEELVIKAEGLAR	-34	2.3E+03	3.1E+04	0.20	**45**	**151**	*24*	2.2E+04	6.0E+04	1.00	**1084**	**3613**
SVHSQDPSGDSKLYR	-52	2.2E+03	1.9E+04	0.02	**7**	**23**	*39*	9.4E+03	2.1E+04	0.20	**274**	**913**

TDLDDDITALKQR	-49	2.1E+03	2.7E+04	0.02	5	15	38	1.7E+04	2.9E+04	0.10	175	584
TKPDPEVEEQEK	-73	1.2E+03	7.0E+03	0.02	10	34	5	9.9E+03	5.8E+04	0.10	51	171
TLAEIAKVELDNMPLR	-43	1.9E+03	2.3E+04	0.02	5	17	20	3.0E+04	9.2E+04	0.10	99	329
TLSDYNIQKESTLHLVLR	-58	2.7E+03	1.2E+05	0.10	7	23	34	2.4E+04	6.0E+04	0.20	238	794
TSDADIKSSETGAFR	-70	1.7E+03	9.6E+04	0.10	5	18	39	1.6E+04	4.4E+04	0.20	213	711
TVTITKEDESTEK	-67	1.7E+03	1.5E+04	0.02	7	24	7	8.4E+03	5.0E+04	0.10	51	169
VGEPEVKEEK	-64	1.3E+03	2.3E+04	0.02	4	12	15	1.3E+04	7.7E+04	0.10	52	174
VKEDPDGEHAR	-52	3.3E+03	3.2E+04	0.02	6	21	35	3.6E+03	9.9E+03	0.20	214	714
VKEEHLDVASPDK	-76	6.6E+03	1.2E+05	0.10	16	54	3	1.1E+04	6.5E+04	0.10	50	165
VKFEQNGSSK	-85	6.6E+02	1.8E+04	0.02	2	7	126	9.1E+03	3.9E+04	0.40	284	946
VKLDSVR	-85	2.3E+03	2.2E+04	0.02	6	21	17	1.3E+04	7.3E+04	0.20	106	354
VLLGETGKEK	-73	1.9E+03	6.4E+04	0.02	2	6	63	1.7E+04	8.9E+04	0.20	114	379
VNLDSEQAVKEEK	-52	1.9E+03	2.2E+04	0.02	5	17	16	1.5E+04	5.5E+04	0.10	81	272
VQIKQETIESR	-70	1.3E+03	2.4E+04	0.02	3	11	44	1.0E+04	2.1E+04	0.10	146	485
VSISEGDDKIEYR	-55	1.9E+03	1.2E+04	0.02	9	31	4	1.0E+04	7.9E+04	0.10	39	130
VVHIDSGIVKQER	-55	2.7E+03	4.1E+04	0.02	4	13	11	1.3E+04	8.4E+04	0.10	45	150
WTVVKTEEGR	-67	2.0E+03	4.5E+04	0.02	3	9	115	2.9E+04	2.8E+04	0.10	302	1008
YEEALSQLEESVKEER	-40	2.1E+03	1.1E+04	0.02	11	38	1	1.5E+04	6.8E+04	0.02	13	44
YQKSTELLIR	-85	8.6E+02	5.0E+04	0.02	1	3	29	1.2E+04	1.2E+05	0.10	30	101
ADSVGKLVTVR	-58	1.8E+03	5.4E+04	0.02	2	7	27.0	2.3E+04	7.8E+04	0.20	177	591
AEAMNIKIEPEETTEAR	-52	2.2E+03	1.4E+04	0.10	46	152	0.8	2.9E+04	1.4E+05	0.20	127	422
APPNVKNEGPLNVVK	-70	2.7E+03	1.0E+04	0.02	16	53	2.8	2.1E+04	8.6E+04	0.20	147	489
DGDVVLPAGVVVKQER	-43	1.8E+03	3.3E+04	0.10	17	55	6.5	2.7E+04	4.6E+04	0.20	356	1188
DTSKMYIPHTDK	-61	4.2E+03	5.0E+04	0.10	25	85	1.3	2.3E+04	1.3E+05	0.20	109	365
EDEGFIKEEEKPLPR	-52	2.4E+03	3.6E+04	0.10	20	67	3.1	2.1E+04	6.1E+04	0.20	205	682
EMSGSTSELLIKENK	-40	2.1E+03	1.2E+04	0.02	11	36	1.1	3.1E+04	4.6E+05	0.20	40	132
HIDALEDKIDEEVR	-61	3.0E+03	1.7E+04	0.02	11	35	10.1	3.3E+04	5.5E+04	0.20	356	1185
IFDEEPANGVKIER	-43	1.7E+03	5.8E+03	0.02	18	60	0.9	2.1E+04	2.3E+05	0.20	53	177
IKEEEEGAEASAR	-34	7.9E+02	7.6E+03	0.10	31	105	1.2	5.5E+03	1.3E+04	0.10	123	410
IKTEPLDFNDYK	-64	2.1E+03	1.2E+04	0.02	11	35	3.1	4.2E+04	2.3E+05	0.20	110	366
IKVEPASEK	-58	8.5E+03	6.3E+04	0.02	8	27	4.9	5.6E+03	1.3E+04	0.10	133	442
INEILSNALKR	-70	1.4E+03	4.2E+04	0.10	10	34	2.6	4.8E+04	3.3E+05	0.20	87	290
KANLLR	-55	1.1E+04	2.4E+05	0.02	3	9	29.6	1.6E+04	7.2E+04	0.40	260	866
KGFSEGLWEIENNPTVK	-46	3.5E+03	5.4E+04	0.02	4	13	13.7	1.8E+04	6.0E+04	0.20	179	597
KYYYYER	-85	1.4E+03	5.2E+03	0.02	16	55	2.1	5.4E+03	2.9E+04	0.20	113	375
LKSVAK	-46	1.5E+03	2.3E+04	0.02	4	13	6.6	4.9E+03	3.4E+04	0.20	87	289
LLLGSPSSLSPFSKR	-49	1.6E+03	7.7E+03	0.02	13	42	0.3	1.4E+04	6.4E+04	0.02	13	43
LLVHMGLLKSEDK	-43	2.5E+03	1.3E+06	0.02	0.1	0.4	13.6	2.0E+04	2.4E+05	0.02	5	17
LQTMKEELDFQK	-52	2.6E+03	1.4E+04	0.02	11	36	1.0	1.7E+04	2.8E+04	0.02	36	121
MAVKEEK	-70	1.2E+03	1.6E+04	0.02	4	14	3.7	8.3E+03	4.7E+04	0.10	53	178
MNKSEDDESGAGELTR	-55	1.4E+03	2.6E+03	0.02	33	111	2.3	2.1E+04	4.9E+04	0.20	255	851
NEGPLNVVKTEK	-52	4.2E+03	5.1E+04	0.02	5	16	7.7	1.6E+04	7.4E+04	0.20	127	425
NKEIMYR	-85	2.2E+04	4.8E+04	0.02	28	92	2.4	1.3E+04	3.5E+04	0.20	226	754
NSDIEQSSDSKVK	-34	1.7E+03	1.4E+04	0.10	37	125	1.0	6.2E+03	1.5E+04	0.10	127	423
QFNKLTEDK	-79	2.6E+03	9.9E+04	0.10	8	26	1.7	8.0E+03	5.3E+04	0.10	45	150
QSLKQGSR	-76	1.3E+03	5.0E+04	0.10	8	27	3.0	8.8E+03	3.3E+04	0.10	80	267
SAEEVEEIKAEK	-49	2.3E+03	3.5E+04	0.02	4	13	4.7	8.6E+03	4.2E+04	0.10	62	207
SDQDHSMDEMTAVVKIEK	-49	1.5E+03	2.9E+03	0.02	32	106	2.8	2.2E+04	4.5E+04	0.20	293	978
SMSVEKIDISPVLLQK	-61	2.5E+03	2.4E+04	0.10	32	105	0.1	1.6E+04	7.0E+04	0.02	13	45
SSIKQEPIER	-64	1.4E+03	3.6E+04	0.02	2	8	42.3	1.5E+04	2.7E+04	0.20	333	1109
STIKTDLDDDITALK	-61	1.6E+03	7.0E+03	0.02	13	45	13.3	3.3E+04	3.3E+04	0.20	593	1976
TDLDDDITALKQR	-49	2.1E+03	1.0E+04	0.02	13	42	8.4	1.0E+04	1.8E+04	0.20	353	1175
TFNSELYSLNDYKPPISK	-43	2.0E+03	5.4E+04	0.10	11	37	1.5	2.7E+04	1.5E+05	0.10	55	182
TFSESLKSEK	82	5.7E+03	9.6E+04	0.02	4	12	15.8	2.0E+04	6.2E+04	0.20	189	629
TKAEEPSDLIGPEAPK	-52	1.9E+03	2.6E+04	0.02	4	15	2.2	1.7E+04	3.1E+05	0.20	33	110
TKPDPEVEEQEK	-67	1.8E+03	6.5E+04	0.02	2	6	37.2	1.2E+04	3.4E+04	0.20	210	700
TLAEIAKVELDNMPLR	-46							2.4E+04	3.5E+04	1.00	2017	6723
TSDADIKSSETGAFR	-70	1.8E+03	5.4E+03	0.02	19	65	3.3	2.3E+03	3.2E+03	0.10	217	724
TVTITKEDESTEK	-64	1.3E+03	2.9E+04	0.02	3	9	10.4	2.8E+03	8.8E+03	0.10	97	323
VENGQEPVIKLENR	-49	1.8E+03	2.9E+04	0.02	4	12	8.9	1.3E+03	3.5E+03	0.10	110	367
VGVKSAPR	-67	5.5E+03	1.4E+05	0.02	2	8	5.8	4.8E+03	3.2E+04	0.10	45	149
VIKMESEEGK	-76	1.4E+03	9.6E+04	0.02	1	3	9.8	3.1E+03	3.2E+04	0.10	29	96
VKEDPDGEHAR	-49	1.6E+03	7.2E+05	0.02	0.1	0.4	17.9	1.0E+04	7.6E+04	0.02	8	27
VKVEEEEEEK	-79	1.7E+03	6.9E+04	0.02	2	5	4.7	7.3E+03	9.3E+04	0.10	24	78
VLDLELKGDIEK	-73	9.5E+02	1.9E+03	0.02	30	99	9.4	2.4E+04	7.7E+04	1.00	932	3107
VLLGETGKEK	-61	8.7E+03	1.7E+05	0.10	16	52	1.7	4.5E+03	1.6E+04	0.10	86	288
VNLDSEQAVKEEK	-52	1.5E+03	1.1E+04	0.02	9	28	9.3	1.2E+04	5.4E+04	0.40	263	875
VVHIDSGIVKQER	-40	3.4E+03	3.3E+04	0.10	31	102	0.6	9.9E+03	4.6E+04	0.10	64	214
WMKHEPGLGSYGDELGR	-58	1.6E+03	1.2E+04	0.02	8	28	23.3	3.2E+04	5.9E+04	0.40	641	2137
YKWEYTYYSR	-58	2.2E+03	7.3E+03	0.02	18	59	3.0	1.2E-04	2.1E+04	0.10	173	578

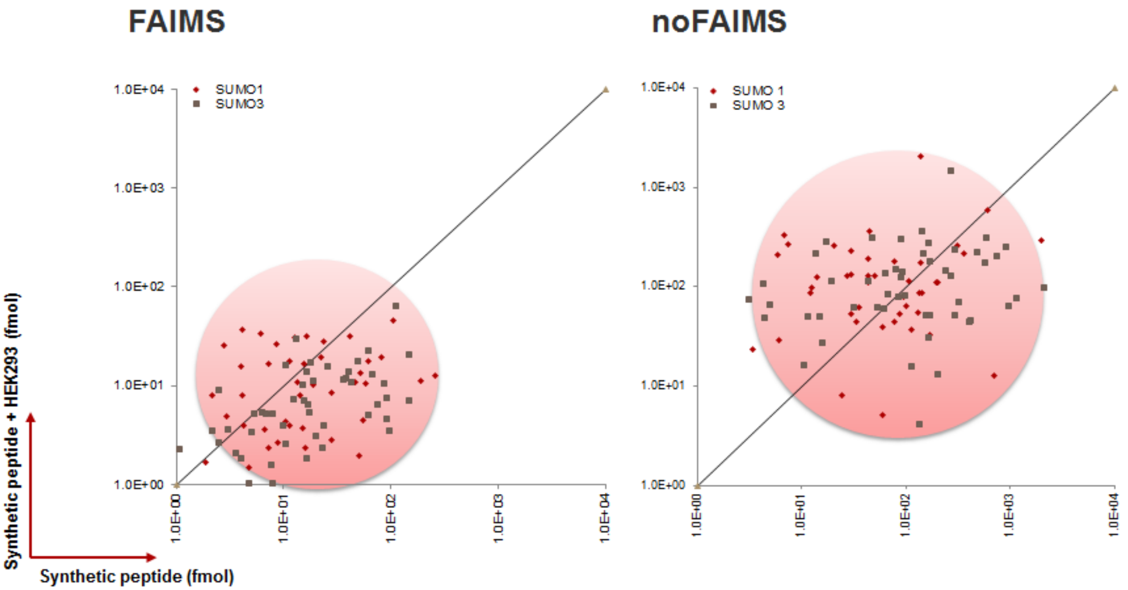


Figure 42: Comparison LOD for pure synthetic SUMO peptides or when spiked in HEK293 cell lysat.

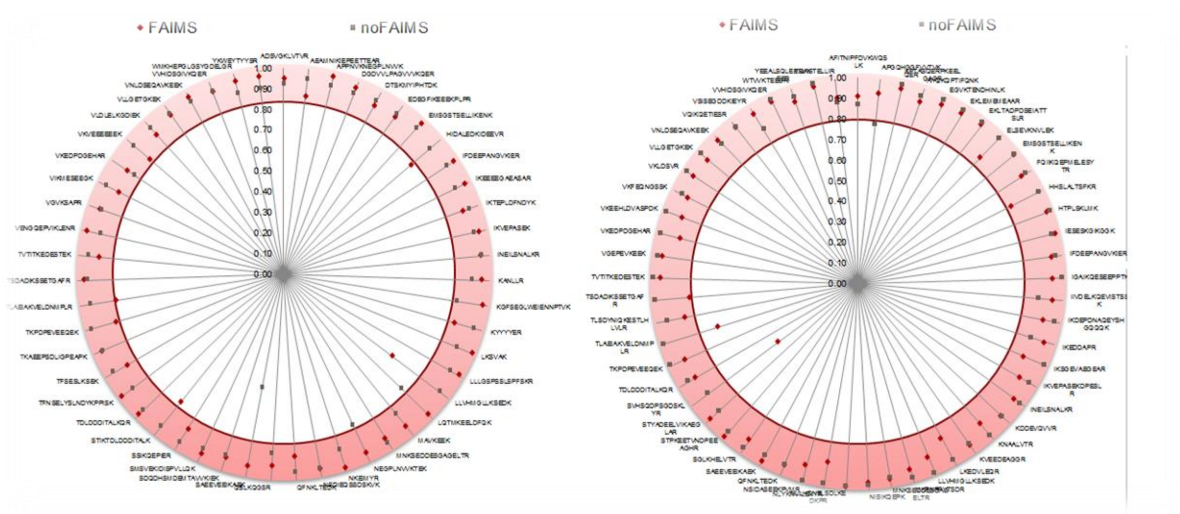


Figure 43: R^2 values of linearity for the common SUMO 1 (left) and SUMO 3 (right) peptides

6.3. Heat Shock

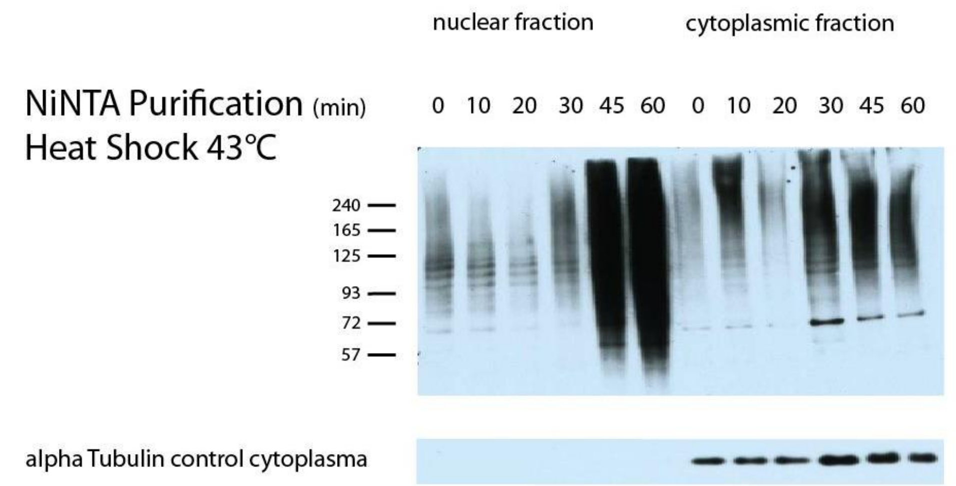


Figure 44: Western blot after purification on NiNTA beads. The heat shock was performed on the time points 0 - 10 - 20 -30 -45 - 60 minutes at 43 °C

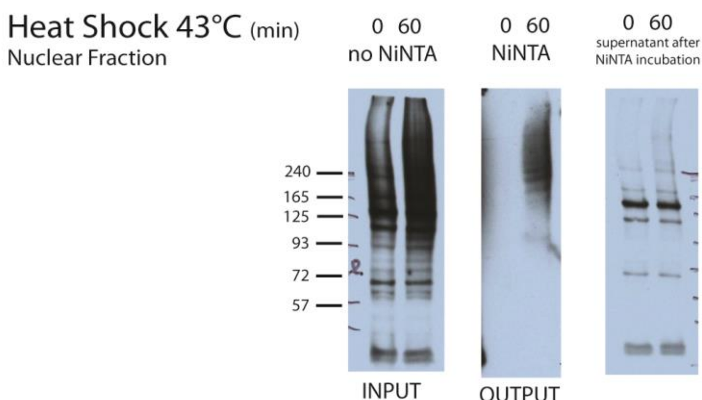


Figure 45: Western Blot against antibody SUMO 2. After NiNTA enrichment the response on the western blot is less. There is no loss of SUMOs on the pulldown after NiNTA incubation.

6.4. SUMO 1 AND SUMO 3 SYNTHETICS

Table 13: List of the 96 SUMO 1 peptides synthetic par PJG Berlin

Sequences synthesized	C-terminus	Comments	Length w/o C-term	Synthesis-#	Plate-#	Row	Column	Amount/Well [nmol]	sequence	Peptide SUMO	Peptide
ALPPEVaVEGP	K	a=Lys(EQTGG)	13	2438_4a	2438_4a Lys	A	1	60	ALPPEVKVEGPK	1735.92	1263.731
APPNVaNEGPLNVV	K	a=Lys(EQTGG)	16	2438_4a	2438_4a Lys	A	2	60	APPNVKNEGPLNVVK	2048.08	1575.885
DCPIaEE	K	a=Lys(EQTGG)	9	2438_4a	2438_4a Lys	A	3	60	DCPIKEEK	1433.66	961.4659
DGIEPMWEDEaN	K	a=Lys(EQTGG)	14	2438_4a	2438_4a Lys	A	4	60	DGIEPMWEDEKNK	2062.90	1590.71
DGLa	K	a=Lys(EQTGG)	6	2438_4a	2438_4a Lys	A	5	60	DGLKK	1032.53	560.3402
DaSS	K	a=Lys(EQTGG)	6	2438_4a	2438_4a Lys	A	6	60	DKSSK	1036.49	564.2988
DTSaMYIPHTD	K	a=Lys(EQTGG)	13	2438_4a	2438_4a Lys	A	7	60	DTSKMYIPHTDK	1907.88	1435.689
EDFIELCTPGVIaQE	K	a=Lys(EQTGG)	17	2438_4a	2438_4a Lys	A	8	60	EDFIELCTPGVIKQEK	2321.13	1848.941
EMSGSTSELLIaEN	K	a=Lys(EQTGG)	16	2438_4a	2438_4a Lys	A	9	60	EMSGSTSELLIKENK	2138.03	1665.836
GGSLPaVEA	K	a=Lys(EQTGG)	11	2438_4a	2438_4a Lys	A	10	60	GGSLPKVEAK	1457.76	985.5677
GVaEEV	K	a=Lys(EQTGG)	8	2438_4a	2438_4a Lys	A	11	60	GVKEEVK	1260.64	788.4512
IaTEPLDFNDY	K	a=Lys(EQTGG)	13	2438_4a	2438_4a Lys	A	12	60	IKTEPLDFNDYK	1954.94	1482.748
IaVEPASE	K	a=Lys(EQTGG)	10	2438_4a	2438_4a Lys	B	1	60	IKVEPASEK	1472.76	1000.567
IMTPSVaVE	K	a=Lys(EQTGG)	11	2438_4a	2438_4a Lys	B	2	60	IMTPSVKVEK	1603.84	1131.644
aGFSEGLWEIENNPTV	K	a=Lys(EQTGG)	18	2438_4a	2438_4a Lys	B	3	60	KGFSEGLWEIENNPTVK	2420.17	1947.981
aPDP	K	a=Lys(EQTGG)	6	2438_4a	2438_4a Lys	B	4	60	KPDPK	1056.53	584.3402
aYLI	K	a=Lys(EQTGG)	6	2438_4a	2438_4a Lys	B	5	60	KYLIK	1136.63	664.4392
LaIEPSSNWDMTGYGSHS	K	a=Lys(EQTGG)	20	2438_4a	2438_4a Lys	B	6	60	LKIEPSSNWDMTGYGSHSK	2609.19	2137.002
LaSVA	K	a=Lys(EQTGG)	7	2438_4a	2438_4a Lys	B	7	60	LKSVAK	1117.62	645.4294
LaTEPEEVSIEDSAQSDL	K	a=Lys(EQTGG)	20	2438_4a	2438_4a Lys	B	8	60	LKTEPEEVSIEDSAQSDLK	2590.24	2118.045
LLVHMGLLaSED	K	a=Lys(EQTGG)	14	2438_4a	2438_4a Lys	B	9	60	LLVHMGLLKSEDK	1955.03	1482.835
LQTMaEELDFQ	K	a=Lys(EQTGG)	13	2438_4a	2438_4a Lys	B	10	60	LQTMKEELDFQK	1981.95	1509.762
LVDIa	K	a=Lys(EQTGG)	7	2438_4a	2438_4a Lys	B	11	60	LVDIKK	1187.66	715.4713
MALa	K	a=Lys(EQTGG)	6	2438_4a	2438_4a Lys	B	12	60	MALKK	1062.56	590.3694
MAVaEE	K	a=Lys(EQTGG)	8	2438_4a	2438_4a Lys	C	1	60	MAVKEEK	1306.63	834.439
NEGPLNVVaTE	K	a=Lys(EQTGG)	13	2438_4a	2438_4a Lys	C	2	60	NEGPLNVVKTEK	1799.91	1327.722
NGSVa	K	a=Lys(EQTGG)	7	2438_4a	2438_4a Lys	C	3	60	NGSVKK	1104.56	632.3726
NSDIEQSSDSaV	K	a=Lys(EQTGG)	14	2438_4a	2438_4a Lys	C	4	60	NSDIEQSSDSKVK	1908.88	1436.686
QFNaLTED	K	a=Lys(EQTGG)	10	2438_4a	2438_4a Lys	C	5	60	QFNKLTEDK	1594.77	1122.579
QQLQQAEEALVAaQEVID	K	a=Lys(EQTGG)	20	2438_4a	2438_4a Lys	C	6	60	QQLQQAEEALVAKQEVIDK	2640.35	2168.156
SAEEVEEIaAE	K	a=Lys(EQTGG)	13	2438_4a	2438_4a Lys	C	7	60	SAEEVEEIKAEK	1833.87	1361.68
SDQDHSMDEMTAVVaIE	K	a=Lys(EQTGG)	19	2438_4a	2438_4a Lys	C	8	60	SDQDHSMDEMTAVVKIEK	2535.13	2062.942
SaPIPIMPASPQ	K	a=Lys(EQTGG)	14	2438_4a	2438_4a Lys	C	9	60	SKPIPIMPASPQK	1865.98	1393.787
SMSVEaIDISPVLLQ	K	a=Lys(EQTGG)	17	2438_4a	2438_4a Lys	C	10	60	SMSVEKIDISPVLLQK	2259.19	1786.998
STIaTDLDDDITAL	K	a=Lys(EQTGG)	16	2438_4a	2438_4a Lys	C	11	60	STIKTDLDDDITALK	2121.06	1648.864
TFNSELYSLNDYaPPIS	K	a=Lys(EQTGG)	19	2438_4a	2438_4a Lys	C	12	60	TFNSELYSLNDYKPPISK	2588.25	2116.06
TFSESLaSE	K	a=Lys(EQTGG)	11	2438_4a	2438_4a Lys	D	1	60	TFSESLKSEK	1627.78	1155.589
TaAEEPSDLIGPEAP	K	a=Lys(EQTGG)	17	2438_4a	2438_4a Lys	D	2	60	TKAEEPSDLIGPEAPK	2154.06	1681.864
TaPDPEVEEQE	K	a=Lys(EQTGG)	13	2438_4a	2438_4a Lys	D	3	60	TKPDPEVEEQEK	1900.88	1428.685
TPaEAIEGTYID	K	a=Lys(EQTGG)	14	2438_4a	2438_4a Lys	D	4	60	TPKEAIEGTYIDK	1936.95	1464.758
TVFDNFFIa	K	a=Lys(EQTGG)	11	2438_4a	2438_4a Lys	D	5	60	TVFDNFFIKK	1730.87	1258.683
TVTITaEDESTE	K	a=Lys(EQTGG)	14	2438_4a	2438_4a Lys	D	6	60	TVTITKEDESTEK	1952.93	1480.738
VIaMESEEG	K	a=Lys(EQTGG)	11	2438_4a	2438_4a Lys	D	7	60	VIKMESEEGK	1621.77	1149.582
VaTEDE	K	a=Lys(EQTGG)	8	2438_4a	2438_4a Lys	D	8	60	VKTEDEK	1320.63	848.436
VaVEEEEEE	K	a=Lys(EQTGG)	11	2438_4a	2438_4a Lys	D	9	60	VKVEEEEEEK	1719.79	1247.6
VaVE	K	a=Lys(EQTGG)	6	2438_4a	2438_4a Lys	D	10	60	VKVEK	1074.58	602.3872

VaYT	K	a=Lys(EQTGG)	6	2438_4a	2438_4a Lys	D	11	60	VKYTK	1110.58	638.3872
VLDLELaGDIE	K	a=Lys(EQTGG)	13	2438_4a	2438_4a Lys	D	12	60	VLDLELKGDIEK	1843.96	1371.773
VLLGETGaE	K	a=Lys(EQTGG)	11	2438_4a	2438_4a Lys	E	1	60	VLLGETGKEK	1545.81	1073.62
VNLDSEQAVaEE	K	a=Lys(EQTGG)	14	2438_4a	2438_4a Lys	E	2	60	VNLDSEQAVKEEK	1960.95	1488.754
VTTHPLAaD	K	a=Lys(EQTGG)	11	2438_4a	2438_4a Lys	E	3	60	VTTHPLAKDK	1581.82	1109.631
AAPGAEFAPNa	R	a=Lys(EQTGG)	13	2438_4b	2438_4b Arg	A	1	52	AAPGAEFAPNKR	1700.84	1228.643
ADSVGaLVTV	R	a=Lys(EQTGG)	12	2438_4b	2438_4b Arg	A	2	52	ADSVGKLVTVR	1616.86	1144.668
AEAMNIaIEPEETTEA	R	a=Lys(EQTGG)	18	2438_4b	2438_4b Arg	A	3	52	AEAMNIKIEPEETTEAR	2404.13	1931.938
AaGP	R	a=Lys(EQTGG)	6	2438_4b	2438_4b Arg	A	4	52	AKGPR	1000.52	528.3253
AQADIYaADFQAE	R	a=Lys(EQTGG)	15	2438_4b	2438_4b Arg	A	5	52	AQADIYKADFQAER	2097.98	1625.792
CPNEGCLHaMEL	R	a=Lys(EQTGG)	14	2438_4b	2438_4b Arg	A	6	52	CPNEGCLHKMELR	2001.89	1529.702
DGDVVLPAGVVVaQE	R	a=Lys(EQTGG)	17	2438_4b	2438_4b Arg	A	7	52	DGDVVLPAGVVVKQER	2153.12	1680.928
DIaQEGETAICSEMTAD	R	a=Lys(EQTGG)	19	2438_4b	2438_4b Arg	A	8	52	DIKQEGETAICSEMTADR	2469.09	1996.895
EDEGFIaEEEKPLP	R	a=Lys(EQTGG)	16	2438_4b	2438_4b Arg	A	9	52	EDEGFIKEEEKPLPR	2288.10	1815.912
GCGVVaFESPEVAE	R	a=Lys(EQTGG)	16	2438_4b	2438_4b Arg	A	10	52	GCGVVKFESPEVAER	2078.98	1606.789
GSTSFLETLaTE	R	a=Lys(EQTGG)	14	2438_4b	2438_4b Arg	A	11	52	GSTSFLETLKTER	1940.96	1468.764
GTGYIaTELISVSEVHPS	R	a=Lys(EQTGG)	20	2438_4b	2438_4b Arg	A	12	52	GTGYIKTELISVSEVHPSR	2545.29	2073.098
HIDALEDaIDEEV	R	a=Lys(EQTGG)	15	2438_4b	2438_4b Arg	B	1	52	HIDALEDKIDEEVR	2154.03	1681.839
HLaHL	R	a=Lys(EQTGG)	7	2438_4b	2438_4b Arg	B	2	52	HLKHLR	1275.69	803.4999
IAQYa	R	a=Lys(EQTGG)	7	2438_4b	2438_4b Arg	B	3	52	IAQYKR	1250.65	778.457
IFDEEPANGVaIE	R	a=Lys(EQTGG)	15	2438_4b	2438_4b Arg	B	4	52	IFDEEPANGVKIER	2089.02	1616.828
IFVGGIaEDTEEHHL	R	a=Lys(EQTGG)	17	2438_4b	2438_4b Arg	B	5	52	IFVGGIKEDTEEHHLR	2352.16	1879.966
IaADPDGPEAQAEACSGE	R	a=Lys(EQTGG)	20	2438_4b	2438_4b Arg	B	6	52	IKADPDGPEAQAEACSGER	2416.07	1943.876
IaEEEEGAEASA	R	a=Lys(EQTGG)	14	2438_4b	2438_4b Arg	B	7	52	IKEEEEGAEASAR	1890.87	1418.676
IaVE	R	a=Lys(EQTGG)	6	2438_4b	2438_4b Arg	B	8	52	IKVER	1116.60	644.409
INEILSNALa	R	a=Lys(EQTGG)	12	2438_4b	2438_4b Arg	B	9	52	INEILSNALKR	1742.94	1270.748
aANLL	R	a=Lys(EQTGG)	7	2438_4b	2438_4b Arg	B	10	52	KANLLR	1186.65	714.4621
aYYYYE	R	a=Lys(EQTGG)	8	2438_4b	2438_4b Arg	B	11	52	KYYYYER	1556.70	1084.51
LLLGSPSSLSPFSa	R	a=Lys(EQTGG)	16	2438_4b	2438_4b Arg	B	12	52	LLLGSPSSLSPFSKR	2061.10	1588.906
LLQaPL	R	a=Lys(EQTGG)	8	2438_4b	2438_4b Arg	C	1	52	LLQKPLR	1339.77	867.5774
LTDNIaYEDCED	R	a=Lys(EQTGG)	14	2438_4b	2438_4b Arg	C	2	52	LTDNIKYEDCEDR	2085.90	1613.711
LTEDaADVQSIIGLQ	R	a=Lys(EQTGG)	17	2438_4b	2438_4b Arg	C	3	52	LTEDKADVQSIIGLQR	2258.16	1785.971
MNaSEDDESGAGELT	R	a=Lys(EQTGG)	17	2438_4b	2438_4b Arg	C	4	52	MNKSEDDESGAGELTR	2210.95	1738.755
NaEIMY	R	a=Lys(EQTGG)	8	2438_4b	2438_4b Arg	C	5	52	NKEIMYR	1425.68	953.4873
QATSISSETaNTL	R	a=Lys(EQTGG)	15	2438_4b	2438_4b Arg	C	6	52	QATSISSETKNTLR	2007.99	1535.802
QQEGFaGTFPDA	R	a=Lys(EQTGG)	14	2438_4b	2438_4b Arg	C	7	52	QQEGFKGTFPDAR	1952.91	1480.718
QSLaQGS	R	a=Lys(EQTGG)	9	2438_4b	2438_4b Arg	C	8	52	QSLKQGSR	1375.69	903.5007
SSIaQEPIE	R	a=Lys(EQTGG)	11	2438_4b	2438_4b Arg	C	9	52	SSIKQEPIER	1658.83	1186.643
SSVaVEAEAS	R	a=Lys(EQTGG)	12	2438_4b	2438_4b Arg	C	10	52	SSVKVEAEASR	1634.80	1162.606
TDLDDDITALaQ	R	a=Lys(EQTGG)	14	2438_4b	2438_4b Arg	C	11	52	TDLDDDITALKQR	1975.96	1503.765
TLAEIAaVELDNMPL	R	a=Lys(EQTGG)	17	2438_4b	2438_4b Arg	C	12	52	TLAEIAKVELDNMPLR	2285.18	1812.989
TSDADIaSSETGAF	R	a=Lys(EQTGG)	16	2438_4b	2438_4b Arg	D	1	52	TSDADIKSSETGAFR	2056.94	1584.75
VENGQEPVIaLEN	R	a=Lys(EQTGG)	15	2438_4b	2438_4b Arg	D	2	52	VENGQEPVIKLENR	2097.06	1624.865
VGVaSAP	R	a=Lys(EQTGG)	9	2438_4b	2438_4b Arg	D	3	52	VGVKSAPR	1285.69	813.4941
VaEDPDGEHA	R	a=Lys(EQTGG)	12	2438_4b	2438_4b Arg	D	4	52	VKEDPDGEHAR	1724.78	1252.592
VaEEPPSPPQSP	R	a=Lys(EQTGG)	14	2438_4b	2438_4b Arg	D	5	52	VKEEPPSPPQSPR	1919.95	1447.754
VaLDSV	R	a=Lys(EQTGG)	8	2438_4b	2438_4b Arg	D	6	52	VKLDSVR	1288.69	816.4938
VVHIDSGIVaQE	R	a=Lys(EQTGG)	14	2438_4b	2438_4b Arg	D	7	52	VVHIDSGIVKQER	1952.02	1479.828
WMaHEPGLGSYGDELG	R	a=Lys(EQTGG)	18	2438_4b	2438_4b Arg	D	8	52	WMKHEPGLGSYGDELGR	2404.10	1931.907
YaWEYTYYS	R	a=Lys(EQTGG)	11	2438_4b	2438_4b Arg	D	9	52	YKWEYTYYSR	1930.86	1458.669

Table 14: List of the 96 SUMO 3 peptides synthetic par PJG Berlin

Sequences synthesized	C-terminus	Comments	Length w/o C-term	Synthesis-#	Plate-#	Row	Column	Amount/Well [nmol]	sequence	Peptide SUMO	Peptide
ADILEDrDG	K	r=Lys(NQTGG)	11	2438_5a	2438_5a Lys	A	1	60	ADILEDKDGK	1560.75	1103.5579
AFITNIPFDVrWQSL	K	r=Lys(NQTGG)	17	2438_5a	2438_5a Lys	A	2	60	AFITNIPFDVKWQSLK	2364.23	1907.0425
AYQrQPTIFQN	K	r=Lys(NQTGG)	13	2438_5a	2438_5a Lys	A	3	60	AYQKQPTIFQNK	1922.97	1465.7798
DGrYSQVLANGLDN	K	r=Lys(NQTGG)	16	2438_5a	2438_5a Lys	A	4	60	DGKYSQVLANGLDNK	2079.01	1621.818
EGVrTENDHINL	K	r=Lys(NQTGG)	14	2438_5a	2438_5a Lys	A	5	60	EGVKTENDHINLK	1953.96	1496.7703
ELSEVrNVLE	K	r=Lys(NQTGG)	12	2438_5a	2438_5a Lys	A	6	60	ELSEVKNVLEK	1744.91	1287.7155
EMSGSTSELLIrEN	K	r=Lys(NQTGG)	16	2438_5a	2438_5a Lys	A	7	60	EMSGSTSELLIKENK	2123.03	1665.8364
FrQEPEDELPEAPP	K	r=Lys(NQTGG)	16	2438_5a	2438_5a Lys	A	8	60	FKQEPEDELPEAPPK	2211.06	1753.8643
GrYKEETIE	K	r=Lys(NQTGG)	11	2438_5a	2438_5a Lys	A	9	60	GKYKEETIEK	1681.84	1224.647
HTPLSrLM	K	r=Lys(NQTGG)	10	2438_5a	2438_5a Lys	A	10	60	HTPLSKLMK	1511.80	1054.6078
IESESrGIKGG	K	r=Lys(NQTGG)	13	2438_5a	2438_5a Lys	A	11	60	IESESKGIKGGK	1689.88	1232.6845
IGAIrQESEEPPT	K	r=Lys(NQTGG)	15	2438_5a	2438_5a Lys	A	12	60	IGAIKQESEEPPTK	1984.00	1526.8061
IIVDELrQEVISTSS	K	r=Lys(NQTGG)	17	2438_5a	2438_5a Lys	B	1	60	IIVDELKQEVISTSSK	2246.19	1788.9953
IrDEPDNAQEYSHGQQQ	K	r=Lys(NQTGG)	19	2438_5a	2438_5a Lys	B	2	60	IKDEPDNAQEYSHGQQQK	2572.17	2114.9738
ITAFrME	K	r=Lys(NQTGG)	9	2438_5a	2438_5a Lys	B	3	60	ITAFKMEK	1424.72	967.5281
ITPEEArY	K	r=Lys(NQTGG)	10	2438_5a	2438_5a Lys	B	4	60	ITPEEAKYK	1535.77	1078.5779
LrVIGQDSSEIHF	K	r=Lys(NQTGG)	15	2438_5a	2438_5a Lys	B	5	60	LKVIGQDSSEIHFK	2058.06	1600.8693
LLVHMGLLrSED	K	r=Lys(NQTGG)	14	2438_5a	2438_5a Lys	B	6	60	LLVHMGLLKSEDK	1940.03	1482.8349
NDQNNSDTrISETETL	K	r=Lys(NQTGG)	18	2438_5a	2438_5a Lys	B	7	60	NDQNNSDTKISETETLK	2394.10	1936.9094
NISIrQEP	K	r=Lys(NQTGG)	10	2438_5a	2438_5a Lys	B	8	60	NISIKQEPK	1513.80	1056.6048
NLrQNLSELEQLNENL	K	r=Lys(NQTGG)	18	2438_5a	2438_5a Lys	B	9	60	NLKQNLSELEQLNENLK	2484.27	2027.0768
NQTILr	K	r=Lys(NQTGG)	8	2438_5a	2438_5a Lys	B	10	60	NQTILKK	1301.72	844.5251
QFNrLTED	K	r=Lys(NQTGG)	10	2438_5a	2438_5a Lys	B	11	60	QFNKLTEDK	1579.77	1122.579
QSLVMLQGNrLDITTE	K	r=Lys(NQTGG)	18	2438_5a	2438_5a Lys	B	12	60	QSLVMLQGNKLDITTEK	2375.22	1918.0314
RPrVEYSEEEL	K	r=Lys(NQTGG)	13	2438_5a	2438_5a Lys	C	1	60	RPKVEYSEEELK	1963.97	1506.7798
SAEEVEEIrAE	K	r=Lys(NQTGG)	13	2438_5a	2438_5a Lys	C	2	60	SAEEVEEIKAEK	1818.87	1361.6795
STIrTDLDDDITAL	K	r=Lys(NQTGG)	16	2438_5a	2438_5a Lys	C	3	60	STIKTDLDDDITALK	2106.06	1648.864
TFSESLrSE	K	r=Lys(NQTGG)	11	2438_5a	2438_5a Lys	C	4	60	TFSESLKSEK	1612.78	1155.5892
TrAEEPSDLIGPEAP	K	r=Lys(NQTGG)	17	2438_5a	2438_5a Lys	C	5	60	TKAEEPSDLIGPEAPK	2139.06	1681.8643
TrPDPEVEEQE	K	r=Lys(NQTGG)	13	2438_5a	2438_5a Lys	C	6	60	TKPDPEVEEQEK	1885.88	1428.6853
TPrEAIEGTYID	K	r=Lys(NQTGG)	14	2438_5a	2438_5a Lys	C	7	60	TPKEAIEGTYIDK	1921.95	1464.7581
TVFDNFFIr	K	r=Lys(NQTGG)	11	2438_5a	2438_5a Lys	C	8	60	TVFDNFFIKK	1715.88	1258.683
TVTITrEDESTE	K	r=Lys(NQTGG)	14	2438_5a	2438_5a Lys	C	9	60	TVTITKEDESTEK	1937.93	1480.7377
VAGGArE	K	r=Lys(NQTGG)	9	2438_5a	2438_5a Lys	C	10	60	VAGGAKEK	1216.63	759.4359
VGEPEVrEE	K	r=Lys(NQTGG)	11	2438_5a	2438_5a Lys	C	11	60	VGEPEVKEEK	1600.78	1143.5892
VrEEHLDVASPD	K	r=Lys(NQTGG)	14	2438_5a	2438_5a Lys	C	12	60	VKEEHLDVASPDK	1923.94	1466.7485
VrFEQNGSS	K	r=Lys(NQTGG)	11	2438_5a	2438_5a Lys	D	1	60	VKFEQNGSSK	1580.77	1123.5742
VLLGETGrE	K	r=Lys(NQTGG)	11	2438_5a	2438_5a Lys	D	2	60	VLLGETGKEK	1530.81	1073.6201
VNLDSEQAVrEE	K	r=Lys(NQTGG)	14	2438_5a	2438_5a Lys	D	3	60	VNLDSEQAVKEEK	1945.95	1488.754
YrEETIE	K	r=Lys(NQTGG)	9	2438_5a	2438_5a Lys	D	4	60	YKEETIEK	1496.72	1039.5306
ADSLLAVVr	R	r=Lys(NQTGG)	11	2438_5b	2438_5b Arg	A	1	52	ADSLLAVVKR	1528.84	1071.6521
AEAMNIrIEPEETTEA	R	r=Lys(NQTGG)	18	2438_5b	2438_5b Arg	A	2	52	AEAMNIKIEPEETTEAR	2389.13	1931.9379
APGQHGGFVVTVrQE	R	r=Lys(NQTGG)	17	2438_5b	2438_5b Arg	A	3	52	APGQHGGFVVTVKQER	2167.10	1709.9082
APTASQERPrEELGAG	R	r=Lys(NQTGG)	18	2438_5b	2438_5b Arg	A	4	52	APTASQERPKEELGAGR	2254.12	1796.925
ATGDETGArVE	R	r=Lys(NQTGG)	13	2438_5b	2438_5b Arg	A	5	52	ATGDETGAKVER	1690.80	1233.607
ATMHLrQEVTP	R	r=Lys(NQTGG)	13	2438_5b	2438_5b Arg	A	6	52	ATMHLKQEVTPR	1867.94	1410.7522
EEDAErAVIDLNN	R	r=Lys(NQTGG)	15	2438_5b	2438_5b Arg	A	7	52	EEDAEKAVIDLNNR	2072.98	1615.7922
EIAQDFrTDL	R	r=Lys(NQTGG)	12	2438_5b	2438_5b Arg	A	8	52	EIAQDFKTDLR	1792.88	1335.6903

ErLEMEMEAA	R	r=Lys(NQTGG)	12	2438_5b	2438_5b Arg	A	9	52	EKLEMEMEAAR	1793.82	1336.6235
ErLTADPDSEIATTSL	R	r=Lys(NQTGG)	18	2438_5b	2438_5b Arg	A	10	52	EKLTADPDSEIATTSLR	2304.13	1846.9393
FQIIrQEPMELESYT	R	r=Lys(NQTGG)	17	2438_5b	2438_5b Arg	A	11	52	FQIIKQEPMELESYTR	2469.21	2012.0157
GPTLGASAPVTVr	R	r=Lys(NQTGG)	15	2438_5b	2438_5b Arg	A	12	52	GPTLGASAPVTVKR	1810.98	1353.7849
GTAGLLEQWLr	R	r=Lys(NQTGG)	13	2438_5b	2438_5b Arg	B	1	52	GTAGLLEQWLKR	1828.97	1371.7743
GVPDArI	R	r=Lys(NQTGG)	9	2438_5b	2438_5b Arg	B	2	52	GVPDAKIR	1312.70	855.5047
GYFEYIEENrYS	R	r=Lys(NQTGG)	14	2438_5b	2438_5b Arg	B	3	52	GYFEYIEENKYSR	2154.97	1697.7806
HHSLALTSFr	R	r=Lys(NQTGG)	12	2438_5b	2438_5b Arg	B	4	52	HHSLALTSFKR	1753.91	1296.7171
IEEIrDFLLTA	R	r=Lys(NQTGG)	13	2438_5b	2438_5b Arg	B	5	52	IEEIKDFLLTAR	1905.01	1447.8155
IFDEEPANGVrIE	R	r=Lys(NQTGG)	15	2438_5b	2438_5b Arg	B	6	52	IFDEEPANGVKIER	2074.02	1616.8279
IrEDDAP	R	r=Lys(NQTGG)	9	2438_5b	2438_5b Arg	B	7	52	IKEDDAPR	1400.68	943.4843
IrIEPGIEPQ	R	r=Lys(NQTGG)	12	2438_5b	2438_5b Arg	B	8	52	IKIEPGIEPQR	1736.93	1279.7369
IrSGEVAEGEA	R	r=Lys(NQTGG)	13	2438_5b	2438_5b Arg	B	9	52	IKSGEVAEGEAR	1702.84	1245.6434
IrVEPASEKDPESL	R	r=Lys(NQTGG)	16	2438_5b	2438_5b Arg	B	10	52	IKVEPASEKDPESLR	2155.10	1697.9068
INEILSNALr	R	r=Lys(NQTGG)	12	2438_5b	2438_5b Arg	B	11	52	INEILSNALKR	1727.94	1270.7478
rAEAGAGSATEFQF	R	r=Lys(NQTGG)	16	2438_5b	2438_5b Arg	B	12	52	KAEAGAGSATEFQFR	2026.96	1569.7656
rDDEVQVV	R	r=Lys(NQTGG)	10	2438_5b	2438_5b Arg	C	1	52	KDDEVQVVR	1544.77	1087.5742
rEETVEDEIDV	R	r=Lys(NQTGG)	13	2438_5b	2438_5b Arg	C	2	52	KEETVEDEIDVR	1918.90	1461.7067
rNAALVT	R	r=Lys(NQTGG)	9	2438_5b	2438_5b Arg	C	3	52	KNAALVTR	1329.72	872.5312
rVEEDEAGG	R	r=Lys(NQTGG)	11	2438_5b	2438_5b Arg	C	4	52	KVEEDEAGGR	1546.71	1089.5171
LrEDVLEQ	R	r=Lys(NQTGG)	10	2438_5b	2438_5b Arg	C	5	52	LKEDVLEQR	1586.81	1129.6212
LrTEEGEIDYSAEEGEN	R	r=Lys(NQTGG)	19	2438_5b	2438_5b Arg	C	6	52	LKTEEGEIDYSAEEGENR	2526.12	2068.9305
LTEDrADVQSIIGLQ	R	r=Lys(NQTGG)	17	2438_5b	2438_5b Arg	C	7	52	LTEDKADVQSIIGLQR	2243.16	1785.9705
MrFNPFVTSD	R	r=Lys(NQTGG)	12	2438_5b	2438_5b Arg	C	8	52	MKFNPFVTSDR	1798.85	1341.662
MNrSEDDESGAGELT	R	r=Lys(NQTGG)	17	2438_5b	2438_5b Arg	C	9	52	MNKSEDDESGAGELTR	2195.95	1738.7548
NLLHDNELSDLrEDKP	R	r=Lys(NQTGG)	18	2438_5b	2438_5b Arg	C	10	52	NLLHDNELSDLKEDKPR	2493.23	2036.0407
NLYrNVILENY	R	r=Lys(NQTGG)	13	2438_5b	2438_5b Arg	C	11	52	NLYKNVILENYR	1996.02	1538.8326
NSIDASEErPVM	R	r=Lys(NQTGG)	14	2438_5b	2438_5b Arg	C	12	52	NSIDASEEKPVMR	1932.91	1475.7159
QLATrAA	R	r=Lys(NQTGG)	9	2438_5b	2438_5b Arg	D	1	52	QLATKAAR	1315.71	858.5156
QQEGFrGTFPDA	R	r=Lys(NQTGG)	14	2438_5b	2438_5b Arg	D	2	52	QQEGFKGTFPDAR	1937.91	1480.7179
SGLrHELVT	R	r=Lys(NQTGG)	11	2438_5b	2438_5b Arg	D	3	52	SGLKHELVTR	1596.85	1139.6531
SSVrVEAEAS	R	r=Lys(NQTGG)	12	2438_5b	2438_5b Arg	D	4	52	SSVKVEAEASR	1619.80	1162.6062
STPrEETVNDPEEAGH	R	r=Lys(NQTGG)	18	2438_5b	2438_5b Arg	D	5	52	STPKEETVNDPEEAGHR	2353.07	1895.873
STYADEELVIrAEGLA	R	r=Lys(NQTGG)	18	2438_5b	2438_5b Arg	D	6	52	STYADEELVIKAEGLAR	2322.16	1864.9651
SVHSQDPSGDSrLY	R	r=Lys(NQTGG)	16	2438_5b	2438_5b Arg	D	7	52	SVHSQDPSGDSKLYR	2133.00	1675.8034
TDLDDDITALrQ	R	r=Lys(NQTGG)	14	2438_5b	2438_5b Arg	D	8	52	TDLDDDITALKQR	1960.96	1503.7649
TLAEIArVELDNMPL	R	r=Lys(NQTGG)	17	2438_5b	2438_5b Arg	D	9	52	TLAEIAKVELDNMPLR	2270.18	1812.9888
TLSDYNIQrESTLHLVL	R	r=Lys(NQTGG)	19	2438_5b	2438_5b Arg	D	10	52	TLSDYNIQKESTLHLVLR	2587.35	2130.1553
TSDADIrSSETGAF	R	r=Lys(NQTGG)	16	2438_5b	2438_5b Arg	D	11	52	TSDADIKSSETGAFR	2041.94	1584.75
VrEDPDGEHA	R	r=Lys(NQTGG)	12	2438_5b	2438_5b Arg	D	12	52	VKEDPDGEHAR	1709.78	1252.5917
VrLDSV	R	r=Lys(NQTGG)	8	2438_5b	2438_5b Arg	E	1	52	VKLDSVR	1273.69	816.4938
VQIrQETIES	R	r=Lys(NQTGG)	12	2438_5b	2438_5b Arg	E	2	52	VQIKQETIESR	1787.92	1330.7325
VSISEGDDrIEY	R	r=Lys(NQTGG)	14	2438_5b	2438_5b Arg	E	3	52	VSISEGDDKIEYR	1967.93	1510.7384
VVHIDSGIVrQE	R	r=Lys(NQTGG)	14	2438_5b	2438_5b Arg	E	4	52	VVHIDSGIVKQER	1937.02	1479.8278
WTVVrTEEG	R	r=Lys(NQTGG)	11	2438_5b	2438_5b Arg	E	5	52	WTVVKTEEGR	1661.82	1204.6321
YEEALSQLEESVrEE	R	r=Lys(NQTGG)	17	2438_5b	2438_5b Arg	E	6	52	YEEALSQLEESVKEER	2396.12	1938.9291
YrWEYTYYS	R	r=Lys(NQTGG)	11	2438_5b	2438_5b Arg	E	7	52	YKWEYTYYSR	1915.86	1458.6688
YQrSTELLI	R	r=Lys(NQTGG)	11	2438_5b	2438_5b Arg	E	8	52	YQKSTELLIR	1708.64	1251.4518

6.5. CAPILLARY COLUMN PACKING

Introduction

This protocol uses a nitrogen tank and an empty HPLC column as a reservoir.

Protocol

1-Preparation of the precolumn

a- Cut a 3-cm piece of TFE tubing (0.3 mm ID × 1.58 mm OD Sigma cat#58702).

b- Cut a 5-cm piece of 50 um ID (50 um ID 375 um OD Molex TSP050375).

c- Place a piece of SDB membrane at the end of the piece of teflon tubing and push it inside with the piece of capillary until it is at 2-cm of the TFE tubing inlet.

d- Put a piece of 150 um ID capillary (150 um ID 375 um OD Molex TSP150375) at the other end of the Teflon tubing. Just slightly push it inside the Teflon tubing (3mm).

e- Connect the other end of the 150 um id capillary to the empty HPLC column with a ferrule.

f- Add stationary phase (tip of a spatula) to 1 ml of ACN in an eppendorf tube. Mix thoroughly.

g- Put the slurry packing in the empty HPLC column and connect it to the N2 tank.

h- Increase the pressure to 50-100 psi and pack the TFE tubing over 5mm. Close the N2 tank and let the setup go back to ambient pressure.

i- Take the HPLC column off the tank. Remove the TFE tubing from the piece of 150 um-id capillary.

j- Place a piece of SDB membrane at the free end of the piece of teflon tubing and push it inside with the piece of capillary until it reaches the stationary phase.

k- The precolumn is ready to use.

2-Column packing

a- Cut a 2-cm piece of TFE tubing (0.3 mm ID × 1.58 mm OD Sigma cat#58702).

b- Cut a 5-cm piece of 50 um ID (50 um ID 375 um OD Molex TSP050375).

c- Place a piece of SDB membrane at the end of the piece of teflon tubing and push it inside with the piece of capillary until it is at 1-cm of the TFE tubing inlet.

d- Put a piece of 17-cm of 150 um ID capillary (150 um ID 375 um OD Molex TSP150375) at the other end of the Teflon tubing. Just push it inside the Teflon tubing until it reaches the membrane.

e- Connect the other end of the 150-um id capillary to the HPLC column.

f- Put the remaining slurry packing used for the precolumn in the empty HPLC column and connect it to the N2 tank.

g- Increase the pressure to 100 psi. The column will start to pack. Raise the pressure slowly during the column packing (until 1000 psi) until 15 cm are packed.

l- Close the N2 tank and let the setup go back to ambient pressure.

m- Take the column off the HPLC. You can connect the column with a piece of Teflon tubing with a piece of membrane in it to any piece of capillary connected to a load/inject valve

Hints and pitfalls

1- If you raise the pressure too fast, the column will not pack correctly.

2- DO NOT CONNECT the column inlet directly on the valve without any frit since some stationary phase might get into the valve when the valve switches.

3- You can help packing with a dremmel by 'dremmeling' the outlet of the HPLC column.

6.6. PROMIX 08/07/2014

a- 50 mM Ammoniumbicarbonat ex BioBasic VWR #CA99501 (395.5mg/100 ml Millipore water store at 4°C)

b- Weight in a tube 1mg protein and dissolve in 1ml 50mM Ammoniumbicarbonat

mg/ml			Intensity 100ng digested
1.03	Alcohol Dehydrogenase from *Saccharomyces cerevisiae*	Sigma A7011, LOT# 041M7354V	4.2E 8
1.30	Bovine Serum Albumin	BioShop ALB00150, LOT# 3927998	7.2 E8
1.18	Carbonic Anhydrase from bovine erythrocytes	Sigma C3934, LOT# 081M1626V	5.2 E7
0.99	Cytochrome c from *Saccharomyces cerevisiae*	Sigma C2436, LOT# 037K7008V	1.0 E9
1.02	Enolase from *Saccharomyces cerevisiae*	Sigma E6126, LOT| 097K7690V	3.6 E8
1.15	Hemoglobin human	Sigma H7379, LOT# SLBC0633V	1.3 E8
1.24	Phosphorylase b from rabbit muscle	Sigma P6635, LOT# 110M7675V	7.4 E8
1.06	Transferrin human	Sigma T3309, LOT| 0L194488	3.7 E8

c- Add 30 ul TCEP sol ex Thermo 77720 (LOT # 0L194488). Incubate for 30min at 37°C, 450 rmp

d- Add 100ul of 5mg/ml Chloroacetamide solution [6.87mg 2-Chloroacetamide ex Sigma C0267/1.4 ml 50mM Ammoniumbicarbonat]. Incubate for 30min at 37°C, 450 rmp

e- Add 20ug Trypsin (modified Trypsin ex Promega V511A LOT # 97191) dissolved in 100 ul 50mM Ammoniumbicarbonat. Incubate over night at 37°C, 450 rmp

f- For 100ng peptide: 100 times dilution of the peptide solution with 0.2% Formic acid:

g- 5ul peptide solution + 495ul 0.2% FA (pH ~ 3). 1000ng/ul → 10ng/ul; Injection Volume: 10 ul

Digested peptide solution		
175 ul	Alcohol Dehydrogenase from *Saccharomyces cerevisiae*	In a 50ml Volumen flask and fill up with 0.2% FA to the mark.
105 ul	Bovine Serum Albumin	
850 ul	Carbonic Anhydrase from bovine erythrocytes	
75 ul	Cytochrome c from *Saccharomyces cerevisiae*	
210 ul	Enolase from *Saccharomyces cerevisiae*	
575 ul	Hemoglobin human	
95 ul	Phosphorylase b from rabbit muscle	
205 ul	Transferrin human	

500 ul aliquots, dryed with the speed vac.

6.7. Group Professor Dr.Pierre Thibault

Figure 46: The photo shows (from right): Anja Roderbrock, Prof. Dr. Pierre Thibault, Sibylle Pfammatter, Dr. Francis McManus, Irene Campoy Mocayo, Nebiyu Abshiru, Christine Desroches, Peter Kubiniok, Christelle Pomies, Dr. Eric Bonneil, Dr. Christina Bell, Dr. Evgeny Kanshin, Olivier Caron-Lizotte and Frederic Lamoliatte.