Method establishment for the analysis of complex biological samples using

nLC-MALDI MS/MS

DISSERTATION

ZUR ERLANGUNG DES DOKTORGRADES

der Naturwissenschaften

vorgelegt beim Fachbereich Chemie (14)

der Johann Wolfgang Goethe – Universität

in Frankfurt am Main

von

Tabiwang Ndipanquang Arrey

aus Mamfe (Cameroon)

Frankfurt (2010)

(D 30)

vom Fachbereich Chemie Biochemie, Pharmazie (14) der

Johann Wolfgang Goethe Universität als Dissertation angenommen.

Dekan:	**Prof. Dr. Dieter Steinhilber**
Gutachter:	**Prof. Dr. Michael. Karas**
	Prof. Dr Rolf Marschalek

Datum der Disputation:

Contents

Figures

Tables

This is dedicated to my parents, Perry Wayne Ndip Arrey and Christina Arrah Egbe Arrey, who always gave me their loving support and encouragement; to my brothers and sisters, Linda Ebanga Arrey, Nchenge N. Arrey, Arrey N. Arrey Ndip N. Arrey, EgbeTakang N. Arrey for their love and support throughout the years.

Acknowledgements

First of all I would like to thank my supervisor Prof. Dr. Michael Karas not only for giving me the opportunity to do what I was longing for but also for his trust, fruitful discussions and the unconditional support.

I kindly thank Prof. Dr. Rolf Marschalek and members of his group especially for the constructive discussion and the successful collaboration. Furthermore, I would like to thank Dr. Karlheinz Grillitsch from the group of Prof. Günther Dauma and Anja Rosenow for the good co-oporation.

The support of my colleagues and of newly made friends made my time here pleasurable so I would give special thanks to Tobias Beckhaus Dr. Sandra Bornemann, Dr. Markus Persike, Dr. Benjamin Rietschel, Dominic Bäumlisberger and Marion Rohmer for the encouragement, constructive criticism, beneficial discussion and for the beautiful time we had after working hours. Furthermore, I will thank Dr. Sabrina Baltruschat, Dr. Tri Hoang, Dr.Sandra Hofmann, Thorsten Jaskolla, Dr. Björn Meyer, Malte Schürken, Dr. Markus Weidlich, Benjamin Müller, Anne Arnold, Jan Gorka Stavroula Markoutsa, Dimitrios Papasotiriou Tobias Schorge, Andreas Münch, Dorota Urbanek, Juliane Rietschel, Dragana Cubrilovic and Walter Eckhard for the good working atmosphere.

In addition, special thanks to Frau Dr. Ute Bahr not only for the help in instrumental analytic and semester practical but also for willingness to explain the functions and application of different analytical instrument.

Finally I would like to thank my parents, Perry-Wayne Ndip Arrey and Christina Arrah Egbe Arrey, my siblings Linda Ebanga Arrey, Nchenge N. Arrey, Arrey N. Arrey Ndip N.

Arrey, EgbeTakang N. Arrey, and my Niece /Nephew, Linda Arrah Arrey and Renen-Elken Abunaw Ashu, my brother-in-law, Frederick Ebot Ashu and my friends especially Ako Franklin (frankieboy), Francis Njinjo, Chi Willibroad for their unending love, support, encouragement, and patience. Each day I give thanks for my good fortune of having such a wonderful family and friends. My mother and father's unending strength and determination are only a few of the qualities that I try to emulate.

Glossary

1-D LC:	One dimensional RPLC technique
2D-SDS-PAGE:	Two dimensional SDS-PAGE
3-HPA:	3-hydroxy picolinic acid
A:	Alanine
ACN:	Acetonitrile
ACTH:	Adrenocorticotropin
AF4:	MLL translocation partner on chromosome 4q21
AF9:	MLL translocation partner on chromosome 9p22

AML:	Acute myeloid leukemia, a clonal malignancy of granulocyte progenitors
Bp:	Base pair
BPC:	Base Peak-Chromatogram
BSA:	Bovine serum albumin
C:	Cysteine
C18:	Reversed phase
CA:	Carrier ampholytes
CB:	Corynebacterium glutamicum membranes
CE:	Capillary electrophoresis
CGC:	Corynebacterium glutamicum-Cytosol
CGM:	Corynebacterium glutamicum-Membranene
CID:	Collision-induced dissociation
ClCCA:	4-chloro-α-cyanocinnamic acid
CNBr:	Cyanogen bromide
C-Trap:	Curved-Trap

D:	Aspartatic acid
Da:	Dalton
DB:	Database
DHB:	2, 5-dihydroxybenzoic acid
DNA:	Deoxyribonucleic Acid
DTT:	1,4-Dithiothreitol
E:	electric field
E:	Glutamic acid
ECD:	Electron capture dissociation
ENL:	‘Eleven nineteen leukemia’, an MLL translocation partner on 19p13.3
ER:	Endoplasmic Reticulum
ESI:	Electrospray ionization
ETD:	Electron transfer dissociation
FA:	Formic acid
FFE:	Free flow electrophoresis

FFF:	Field flow fractionation
FTICR:	Fourier transform ion cyclotron resonance
Gln:	Glutamine
Glu:	Glutamic acid
Gly:	Glycine
GRAVY:	Grand average of hydropathy
H:	Histidine
HABA:	2-(4-hydroxylphenylazo) benzoic acid
CHCA:	α-cyano-4-hydroxycinnamic acid
HCD:	Higher Collisional Dissociation
HILIC:	Hydrophilic Interaction Liquid Chromatography
His:	Histidine
hppK:	High-pH-proteinase K
I:	Isoleucine
IAA:	N-Iodoacetamide

ICAT:	Isotope Coded Affinity Tag
ICAT:	Isotope coded affinitt tag
ICPL:	Isotope Coded Protein Label
ICPL:	Isotope coded protein label
ID:	Internal diameter
IEC:	Ion-exchange chromatography
IEF:	Isoelectric focusing
Ile:	Isoleucine
IMP:	Integral membrane protein
IPG	Immobilized pH gel strips
IT:	Ion Trap
IT:	Ion trap
iTRAQ:	Isobaric tag for relative quantification
K:	Lysine
kDa:	Kilodalton

L:	Length
L:	Leucine
LC:	Liquid chromatography
Leu:	Leucine
LIT:	Linear Ion Trap
LP:	Lipid particles
Lys:	Lysine
m/z:	Mass-to-Charge
m:	Ion mass
M:	Methionine
MALDI:	Matrix-assisted laser desorption/ionization
MCP:	Multichannel plate Detector
MeOH:	Methanol
MEP:	Membrane-embedded peptide
Met:	Methionine

mg:	Milligram
MGF:	Mascot generic format
min:	Minute
mM:	Millimole
MOWSE:	Molecular Weight Score
MS/MS:	Tandem mass spectrometry
MS:	Mass Spectrometry
MudPIT:	Multidimensional protein identification technology
N_2- laser:	Nitrogen laser
NA:	Nicotinic acid or 3-pyridincarboxylic acid
NCBI:	National Center for Biotechnology Information
Nd:YAG:	Neodymium-doped yttrium aluminium garnet; Nd:$Y_3Al_5O_{12}$
ne:	Charge of the ion
NHS:	N-hydroxysuccinimide
nLC:	Nano Liquid chromstography

Offgel-IEF:	Offgel- Isoelectric focusing
OGE:	Off-gel-Isoelectric focusing
o-TOF:	orthogonal Time of flight analyser
PA:	Picolinic acid
p*I*:	Isoelectric point
PM:	Halobacterium salinarium purple membranes
PMF:	Peptide mass fingerprinting
ppm:	Parts per million
PSD:	Post Source Decay
PTM:	Post-translational modification
Q:	Glutamine
Q:	ion charge
QIT:	Quadrupole ion trap
R:	Arginine
RP-HPLC:	Reverse phase high pressure liquid chromatography

S/N:	Signal-to-noise ratio
SA:	3,5-dimethoxy-4-hydroxycinnamic acid (sinapinic acid)
SDS-PAGE:	Sodium dodecyl sulfatepolyacrylamide gelelectrophoresis
SEC:	Size-exclusion chromatography
SET:	Domain histone methyltransferase domain found in Sur(var)3-9, enhancerof zeste, trithorax and other proteins located at C terminus of MLL
SGBS:	Human Simpson-Golabi-Behmel syndrome
SILAC:	Stable Isotope Labelling with Amino Acids in Cell Culture
SILAC:	Stable isotope labelling by amino acids in cell culture
STE:	steryl esters
t:	ions flight time
TAG:	Triacylglycerols
t-AML: prior	Acute myeloid leukemia arising as a result of chemotherapy for a
TCA:	Trichloroacetic acid

TFA:	Trifluroacetic acid
THAP:	2,4,6-trihydroxyacetophonon
TiO_2:	Titanium dioxide
TMD:	Transmembrane domain, specifically a transmembrane
TMT:	Tandem mass Tag
TOF:	Time of flight
TRIS:	2-Amino-2-(Hydoxymethyl)-Propane-1,3-Diol
UV:	Ultraviolet
$v \times B$:	vector cross product of the ion velocity and the magnetic field
V:	Potential
v:	Velocity of magnitude
YPD:	Yeast strains grown in Glucose rich media
YPO:	Yeast strains grown in oleic acid media

1. Proteomics

Though individual genes and proteins or small clusters of related components or specific biochemical pathways have widely been studied for decades, the term “proteome” was first coined in the early 1990s by Wilkins and colleagues (Wilkins *et al.*, 1995) to describe the protein complement of the genome. It was later expanded to include protein identification studies by Patterson and Aebersold who defined proteomics, “as a systematic study of the many and diverse properties of proteins in a parallel manner with the aim of providing detailed descriptions of the structure, function and control of biological systems in health and disease” (Patterson & Aebersold, 2003). Different cellular states result in different gene regulations (Cox & Mann, 2007). This leads to the expression of many different proteins, involved in different activities, such as cell-cycle control, signal transduction and transcriptional-factor regulation. While the genome of a certain cell is stable, the proteome is subjected to continuous changes. Each gene product does not result in only one protein entity of the cell, due to activities of the cell as a result of co-translational and posttranslational modifications such as phoyphorylation or glycosylation (Liebler, 2002). The extent to which these proteins are modified depends on the regulatory mechanism within the cell and environmental factors. As a consequence, many proteins and their isoforms are generated and the detection and differentiation, even with modern technological innovations, are still difficult. Furthermore, the wide dynamic range of protein concentrations constitutes proteomics studies challenging. Since proteins represent the major part of the cell machinery, the proteome study of sequenced organisms delivers information about their occurrence and abundance and is essential for the understanding of protein activities and regulation.

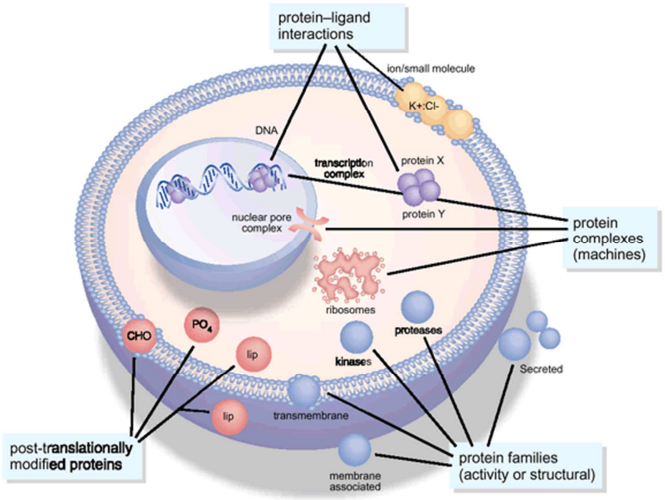

Figure 1 A section through a eukaryotic cell (Patterson & Aebersold, 2003).

1.1 Tools for proteomics research

Driven by the goal to understand the different biochemical functions of cells and the increasing change in gene expression levels, the study of the whole proteome has gained the attention of many research groups all over the globe. The increased innovations in proteomics research have been based on the following:

(i) The constantly improving separation technology, which provides two advantages: reduced sample complexity and permits the differentiation in expression level.

(ii) Mass spectrometric instrumentations

(iii) A collection of emerging software tools to process MS and MS/MS data

The separation technologies can be classified in two categories: gel-based and gel-free systems.

1.1.2 Gel-based separation methods

1.1.2.1 Sodium dodecyl sulfatepolyacrylamide gelelectrophoresis (SDS-PAGE)

Traditionally, SDS-PAGE, the most widely used electrophoretic techniques, has been used for protein separation according to molecular weight (O´farrel, 1975). The degree of acrylamide/bis-acrylamide cross-linking depends on the characteristic of the protein of interest. A lower degree of cross-linking permits an easier passage of large proteins through the gel (Gygi *et al.*, 1997). The proteins are initially dissolved in SDS and loaded on gel pockets. Upon application of a voltage, the protein-SDS complexes migrate through the gel and are resolved into bands. The separated proteins are then visualised by conventional staining techniques such as fluorescent dyes, coomassie and silver stains, which also determine the detection sensitivity. Due to its limited resolution each band can still contain not only iso-forms of a particular proteins but also isobaric proteins. This serves as a limitation of this technique; nevertheless it is still an important tool for proteomic analysis.

1.1.2.2 Isoelectric focusing (IEF)

Another technique, which has widely been implemented for separating proteins, is isoelectric focusing (IEF). In this technique, proteins are separated according to their isoelectric point (p*I*), since the net charge of a protein is determined by the pH gradient of their local environment (Righetti *et al.,* 1988; Egen *et al*., 1988). When subjected to a medium with a linear pH gradient, proteins move towards the electrode with the opposite charge. During migration through the pH gradient the protein either picks up or loses protons. As it does, its net charge and mobility decrease. The protein slows down and eventually arrive at the point in the pH gradient equalling its p*I* and stop migrating. If a protein at its p*I* should diffuse to a region of lower pH, it will become protonated and forced towards the cathode by the electric field. On the other hand, if it diffuses into a pH higher than its p*I*, the protein will become negatively charged and will be driven towards the anode.

1.1.2.3 Two dimensional SDS-PAGE (2D-SDS-PAGE)

An alternative to SDS-PAGE or IEF is the combination of both methods. Therefore, the separation is based on two physiochemical properties of proteins (2D-SDS-PAGE). In this approach, a protein mixture is separated according to charge (p*I*) by isoelectric focusing (IEF) in the first dimension and according to size (Mr) by SDS-PAGE in the second dimension. Visualisation is also achieved by conventional staining techniques as in the SDS-PAGE approach. Compared to conventional IEF, the introduction of the immobilized pH (IPG) gel strips significantly improved the reproducibility in the first dimension (Görg *et al., 1988;* Görg *et al.,* 2000*;* Görg *et al.,* 2004). This method has a unique capacity for the resolution of complex mixtures of proteins, permitting the simultaneous analysis of hundreds or even thousands of proteins. Although this method remains a standard tool for proteomic research, it has some limitations. The main disadvantage being the limited amount of sample capacity and detection sensitivity, thus, only relatively abundant proteins would be identified, especially when analysing samples from a whole cell lysate (Gygi *et al., 2000;* Pedersen *et al.,* 2003). Secondly, despite making it compatible to this class of proteins, the analysis of insoluble and hydrophobic proteins is still a major problem (Görg *et al.,* 2004). Last but not least, proteins with extreme p*I* (below 3 and above 10) or extreme molecular weight are usually excluded from separation (*Lam,* 2007).

1.1.2.4 Offgel- IEF

Protein separation according to p*I* in gels has been proven to be very successful. However, staining, excision, digesting and extraction from the gel prior to MS are still tedious. New technological advancements in this field are based on keeping the separated sample in solution. Different instruments such as the multi-compartment electrolyzers (Righetti *et al.,* 1989; Herber *et al.,* 2000; Pedersen *et al.,* 2003) or the Rotofor system (Wall *et al*., 2000) are commercially available. The most promising system is the Off-gel-isoelectric focusing from Agilent (Ros *et al.,* 2002) (see figure 2). This method exploits the high resolution and separation power of the immobilized pH gradient (IPG)-strip (Cargile *et al.,* 2004, *Cargile et al.,* 2005). The separated peptides and proteins are present in solution making recovery for LC/MS analysis much easier. Contrary to conventional electrophoretic separation techniques, isoelectric focusing (IEF)

offers the highest resolution, due to the inherent nature of the focusing process: it is a dynamic process resulting from the constant equilibrium between diffusion and migration (Lam *et al*., 2007). Offgel-IEF thus, combines separation and high sample amount, a useful feature for preparative or semi-preparative purposes.

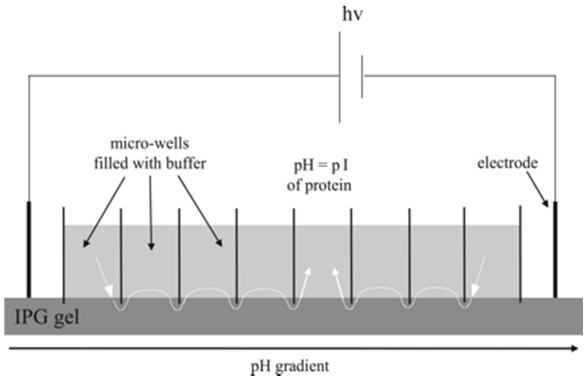

Figure 2 Schematic representation of the Offgel-IEF (Hoert et al., 2006).

1.1.3. Gel-free methods

With the improvement of the fast separation techniques, traditional 2-D-SDS-PAGE methods are increasingly being substituted by gel-free MS-based strategies. A driving force behind the success of gel-free methods is that they exhibit higher sensitivity, high resolving power and higher degree of automation. Most frequently reversed phase high pressure liquid chromatography (RP-HPLC), on C_{18} materials coupled with tandem MS. In these strategies, the proteins are converted to peptides with the use of a proteolytic enzyme or a chemical reagent. The advantage of this method over the conventional 1D or 2D- SDS-PAGE method is that peptides are easily subjected to MS and easier to keep in solution (Stalder *et al*., 2008). However, using this approach makes it difficult to monitor the distribution of intact protein isoforms and protein fragments (Stalder *et al*., 2008). Due to the availability of improved stationary phase chemistry,

chromatographic approaches have greatly enhanced the understanding of biological systems. A wide variety of stationary phase materials are commercially available. Depending on the mode of separation the can be classified into:

1.1.3.1 Ion-exchange chromatography (IEC)

Ion Exchange Chromatography relies on charge-charge interactions between the sample and the stationary phase. It can be subdivided into cation exchange chromatography and anion exchange chromatography. In cation exchange chromatography, positively charged molecules are more retained due to the negatively charged stationary phase. On other hand, negatively charged analytes are more retained in the anion exchange chromatography due to its positively charged surface. Elution of the loaded samples is achieved by steadily increasing the ionic strength of the mobile phase solution (Small, 1989; Weiss & Weiss, 2005).

1.1.3.2 Size-exclusion chromatography (SEC)

SEC, also known as gel filtration, discriminates between molecular species on the basis of size, due to differential permeation into matrices of controlled porosity. Large molecules, which cannot penetrate into the pores pass un-retained through the column and are eluted first. The small molecules, which penetrate the pores are eluted last.

1.2.2.4 Affinity Chromatography

Affinity chromatography is based on the bioaffinity of a protein for a specific ligand coupled to a solid support. The immobilised ligand interacts only with protein that can selectively bind to it, while the unretained proteins are eluted. Hence, this technique offers high selectivity, high resolution and usually high capacity for the protein sample of interest (Voet and Voet, 1995).

1.1.3.3 Normal-phase chromatography

Normal or adsorption chromatography separates analytes based on adsorption to a stationary phase by polarity. The stationary phase is strongly polar in nature (e.g., silica gel). Polar biomolecules are thus retained on the polar surface of the column longer than less polar materials. This separation technique has gained less acceptance due to lack of reproducibility of retention times as water or protic organic solvents change the hydration state of the silica chromatographic media (Snyder & Dolan, 2006).

A special case of the normal phase is the Hydrophilic Interaction Liquid Chromatography (HILIC). The stationary phase contains polar materials such as cyano, amino or diol. The advantage of this form of chromatography over the normal chromatography is that the solvent used is highly organic and miscible with water (Alpert *et al.,* 1990). During separation, the analyte is partitioned between a liquid/liquid layer form between the stationary and the mobile phase. Furthermore, hydrogen donor interactions between neutral polar species as well as weak electrostatic mechanisms under high organic solvent conditions also contribute to the separation mechanism (Alpert *et al.,* 1990). Compared to the conventional normal phase separation, this technique is increasingly gaining acceptance in proteomic workflows, especially for the separation of hydrophilic biomolecules such as glycopeptides or phosphopeptides.

1.1.3.4 Reversed phase chromatography.

Reverse phase high-performance liquid chromatography (RP-HPLC) involves the separation of molecules on the basis of hydrophobicity. The separation depends on the hydrophobic binding of the solute molecules from the mobile phase to the hydrophobic ligands attached to the stationary phase. The solutes are eluted in order of increasing molecular hydrophobicity. Elution is preceded either by isocratic condition, where the concentration of the organic solvent remains constant or by gradient elution, whereby the amount of organic solvent is increased over a period of time. RP-HPLC is an established technique for the analysis of peptides because of a number of reasons including (i) the excellent resolution that can be achieved under a wide range of chromatographic conditions for closely related molecules as well as structurally distinct molecules, (ii) the experimental ease with which chromatographic selectivity can be manipulated through changes in mobile phase characteristics, (iii) the generally high recovery hence, high productivity, and (iv) the excellent

reproducibility of repetitive separation carried out over a long period of time, which is caused partly by the stability of the sorbent material under a wide range of mobile phase condition (reviewed in Aguilar, 2004).

The stationary phase material is mostly silica based sorbent. The porous nature of these materials enable the use of high linear flow velocities, thus rapid analysis time is possible. The silica materials are chemically modified by derivatised silane bearing an n-alkyl hydrophobic ligand. The most commonly used are C-18 for peptide separation and C-8 and C-4 for protein separations. Phenyl and cyanopropyl ligands have also been used to modify silanol groups. The chemical immobilisation of the silica surface results in approximately half of the surface silanol groups being modified. The sorbents are therefore subjected to further silanisation with small reactive silane to produce an end-capped packing material. The type of n-alkyl ligands used can signifantly influence the retention of peptides and proteins. Although the detailed molecular basis of the effect of ligands structure is not fully understood, a number of factors including the relative hydrophobicity, the ligand chain length, flexibility, and the degree of exposure of surface silanol play an important role in the retention process.

1.1.3.5 LC-MALDI Spotting

The coupling of MALDI to mass spectrometers with MS/MS capability enabled the analyses of complex samples. This resulted in the interfacing of MALDI with separation techniques, such as HPLC. Most of these interfaces differ in the method in which the eluate gets to the MALDI target. In either way, both the eluate and the matrix are co-mixed on a MicroTee or at the needle tip through a 3 duct system and spotted on a MALDI target (see figure 3). Deposition on the MALDI target is achieved through techniques such as blotting methods, which involve contact between the eluate droplet and the MALDI plate(Griffin *et al*., 2001; Bodnar *et al*., 2003; Parker *et al*., 2004; Tegeler *et al*., 2004), pulse electric field depositors that use a pulsed electric field to eject droplets (Ericson *et al*., 2003; Young *et al*., 2006), piezoelectric spotting, which dislodge a series of tiny droplets at a high-frequency (Miliotis *et al*., 2000; Ekstrom *et al*., 2001; Mirgorodskaya *et al*., 2005) or heated droplet interfaces, that use thermal energy to evaporate solvents in the hanging droplets to concentrate the eluate and produce smaller droplets for deposition (Zhang *et al*., 2004). In spite the different deposition methods, all share a commonality in that the eluate from the HPLC is mixed with a matrix and co-crystallised with a matrix on a metal target.

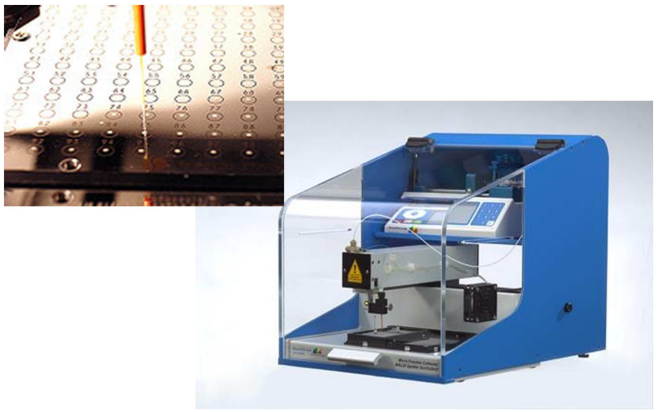

Figure 3 MALDI spotter (Sunchrom)

11.3 Two dimensional separation techniques

Despite the high separation power offered by RP-HPLC (Lin *et al.*, 1999*; Washburn et al., 2001;* Cargile *et al.*, 2004*;* Cargile *et al.*, 2005; Speers & Wu, 2007*;*) for the separation of peptides or proteins, some samples possess complexity beyond the separation capacity of a one-dimensional technique (Opiteck *et al.*, 1997). An effective approach to reduce the complexity of biological samples is the combination of orthogonal separation techniques. In multi-dimensional separation strategies, samples are separated in stages using different physiochemical properties. Two-dimensional separation methods (2-D), the most common form of multidimensional strategies, have been shown to have high resolving power and to improve the number of identified proteins (Peng *et al.*, 2003). Different forms of separation techniques such as SCX/RP-HPLC, Offgel-IEF/RP-HPLC or RP/-HPLC have been coupled (Simpson *et al.*, 2000*;* Hörth*. et al*., 2006).

2 Introduction

2.1 Mass spectrometry

A wide range of new mass spectrometry-based analytical platforms and experimental strategies are available for protein identification, characterisation and quantification (Domon *et al.*, 2006). Generally, a typical mass spectrometer consists of an ion source, which converts samples to ions, a mass analyzer, which sorts the ions according to their mass-to-charge ratio (m/z) and the detector that measures the amount of ions present (see figure 4). Two different types of instruments are typically used for proteomic analyses: instruments equipped with a Matrix assisted laser desorption/ionisation (MALDI) (Karas *et al.*, 1985) and those equipped with electrospray ionisation (ESI) (Fenn *et al.*, 1989) sources. Both constitute "soft" ionisation techniques that enable the transfer of intact proteins, peptides and nucleic acids into the gas phase without fragmentation.

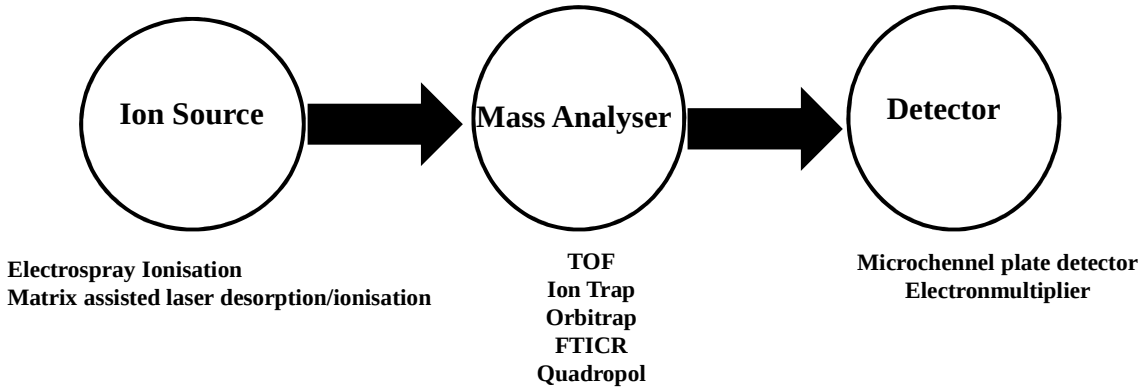

Figure 4 A general representation of a mass spectrometer.

2.1.1 Electrospray ionisation (ESI)

In electrospray ionisation, a spray of highly charged droplets containing the analytes is created at atmospheric pressure in the presence of an electric field. The aerosol is sampled into the first vacuum stage of a mass spectrometer through a nozzle or a capillary, which can be heated to aids further solvent evaporation from the charged droplets. As the liquid in these droplets evaporates, they shrink, until the ionic charge repulsion is sufficiently intense to overcome the surface tension holding the droplet together, thus resulting to an explosion into smaller drops (Columbic explosion). This process repeats itself until the solvent is completely evaporated and the droplets have split up to the point that each contains a single, charged molecule (see figure 5).

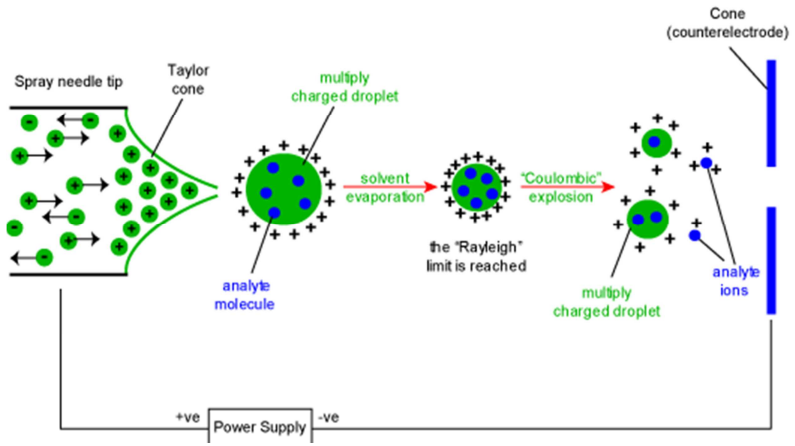

Figure 5 Mechanism of ion formation in ESI (University of Bristol).

2.1.2 Matrix assisted Laser Desorption/Ionisation (MALDI)

Laser desorption and ionisation was first applied for the mass determination of small organic compounds (Cotter, *1984*). In 1987 Tanaka et al. reported the use of UV-laser desorption for the analysis of proteins up to 34 kDa after co-deposition with ultra fine metal

powder in liquid glycerol (Tanaka *et al.*, 1987). The extension of the mass range for proteins to 100 kDa and other large biomolecules using a large excess of nicotinic acid as matrix by Karas and Hillenkam marked the rebirth of UV-laser for the ionisation/desorption (Karas & Hillenkamp, 1988; Karas *et al*, 1989; Karas *et al.*1989; Hillenkamp *et al*. 1991). In this technique, the analytes are co-crystallised with a matrix and the ionisation is triggered by a laser beam. Laser ion sources used in MALDI differ in specific technical details but all are oblige to the same principles of operation. For example, they all use pulse lasers with pulse durations of 1-200 ns. Most commonly used lasers are N_2-laser emitting at 337 nm or Nd-YAG lasers, whose emission of 1064 nm has been frequency-tripled to 355 nm (Yariv, 1989). MALDI matrices constitute small molecular compounds. For a compound to be qualified as a matrix, it must possess certain requirements such as, absorption of laser energy, they must be vacuum stable and be able to embed and isolate analytes, promote analyte ionization, finally causing desorption of analytes upon irradiation with a laser. The most commonly used matrices are α-cyano-4-hydroxycinnamic acid (HCCA), 2, 5-dihydroxybenzoic acid (DHB) and 3, 5-dimethoxy-4-hydroxycinnamic acid (sinapinic acid). Recently, a new rationally designed matrix, 4-chloro-α-cyanocinnamic acid (Cl-CCA) with enhanced signal intensities, a much more uniform response to peptides irrespective of their amino acid composition and a low detection limit for (phospho)-peptides compared to HCCA (Jaskolla *et al.*, 2008) has been introduced. Since its introduction, the use of various interfaces, laser and matrices have continuously been exploited.

Table 1 List of MALDI matrices, their shortcuts and area of application.

Compound	**Shortcuts**	**Applications**
2,5-dihydroxy benzoic acid (Strupat *et al*., 1991)	DHB	Peptides, Proteins, Nucleotides, Oligonucleotides, Oligosaccharide
3,5-dimethoxy-4-hydroxycinnamic acid (Beavis *et al*.,	SA	Peptides, Proteins, Glycopeptides

1989*)*		
4-hydroxy-3-methoxycinnamic acid or ferulic acid (Beavis *et al*., 1989*)*	FA	Proteins, Peptides
α-cyano-4-hydroxycinnamic acid (Beavis *et al*., 1992*)*	CHCA	Peptides, Lipids, Nucleotides
Picolinic acid (Tang *et al*., 1994)	PA	Oligonucleotides
3-hydroxy picolinic acid (Wu *et al*., 1993)	3-HPA	Oligonucleotides, Peptides
4-chloro-α-cyanocinnamic acid (Cl-CCA) (jaskolla *et al.*, 2008)	ClCCA	Peptides
2,4-di-Flouro-α-cyanocinnamic acid (di-FCCA) (jaskolla *et al*., 2010)	di-FCCA	Peptides
2,4,6-trihydroxyacetophonon (Pieles *et al.*, 1993)	THAP	Peptides, Oligonucleotides

2.1.2.1 Mass Analyzers

Mass analyzers separate ions according to their mass-to-charge ratio (m/z). Mass analyzers such as quadrupole (Q), ion trap (quadrupole ion trap, QIT; linear ion trap, LIT or LTQ), time-of-flight (TOF) mass analyzer, and Fourier-transform ion cyclotron resonance (FTICR) mass analyzer have been used for proteomics studies. These analyzers vary in their physical principles and analytical performance. Nevertheless, they are characterized by a number of parameters, such as mass accuracy, mass resolution, accessible mass range, sensitivity (limit of detection) and the signal-to-noise-ratio.

2.1.2.2 Time of flight (TOF) Analyzer

The most commonly used analyzer for MALDI is the TOF. It is particularly suited for MALDI sources because they generate only a very short pulse of ions, which then fly down the flight tube and hit the detector. The TOF-analyzers were first described by Stephen W.E in 1946. The prototype of modern TOF instruments was built by Wiley and McLaren in 1955. The extent to which the TOF is used in modern MS can be attributed to the introduction of MALDI. Theoretically, at the start of the flight tube, all ions have the same kinetic energy, since they all experience the same potential difference during acceleration. However, as a result of the MALDI process, they exhibit different initial velocities. The ions reach the detector in order of increasing mass (see figure 6). Their arrival time at the detector is proportional to their mass. The m/z of the ions relate to the flight times t by the following equations:

$$mv^2/2 = zV;\ v = \sqrt{2V/m} \quad \text{.....(1)}$$

The m/z of the ions is related to the flight times t by the following equations:

$$t = l/v = l\sqrt{m/2zV} \quad \text{..... (2)}$$

Where l is the length of the flight tube, v is the velocity of the ion of mass m and charge z, and V is the acceleration potential.

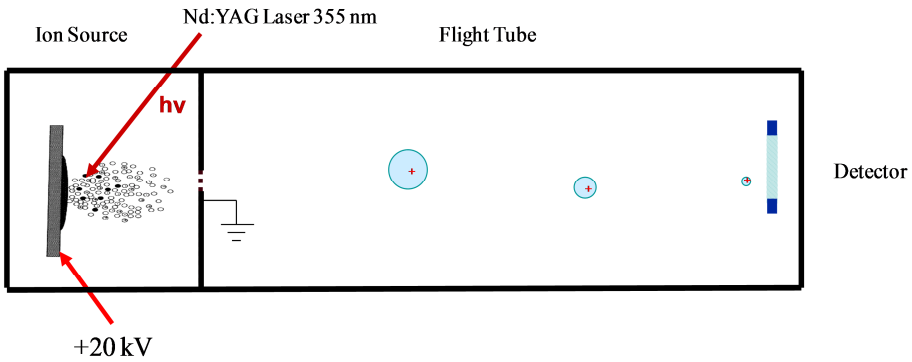

Figure 6 Schematic representation of a linear MALDI-TOF instrument.

In its early years, MALDI-TOF instruments were operated with low-acceleration-voltage reflectron instruments or in the linear mode. Despite the successes in analysing biomolecules, these instruments were faced with insufficient mass resolution resulting from flight time variations of ions of the same m/z. The ionisation process adds a certain amount of initial kinetic energy/initial velocity to the molecules before acceleration. Finally, the different spatial positions in the source, from where the ions are formed, lead to varying values of l (Equation 2) and thus to flight time variations. Delayed extraction was introduced on linear-TOF instruments to compensate for the velocity distributions of ions. By delayed extraction, a mass dependent delay time is built between the ionization and the acceleration in the flight tube. This is achieved by keeping the electric field between the sample and the first electrode at zero or to a minimum value. After this time frame (usually a few 100 nanoseconds), the electric field is switched on to extract the ions to the flight tube. During this time, the faster ions will move towards the first electrode and upon switching on the electric field, they experience less potential difference than the slower ones, thus compensating their higher initial velocity and thereby increasing resolution (reviewed in Hillenkamp & Katalinić, 2007).

On the other hand, the reflectrons compensate energy distribution by means of a series of ring electrode with linear voltage gradient to decelerate and turn around the ions (see figure 7). Due to different energy distribution, ions of the same m/z will travel different path lengths in the reflectron. Ions with greater kinetic energy arrive at the reflectron first but penetrate deeper into the electrode then those of low energy, thus increasing their flight time. Both low and high energy-ions are then re-focused to hit the detector at the same time, thereby improving the resolution power of the instrument. With these implementations, high resolutions (> 20000) and high mass accuracies (5-20 ppm) can be achieved. Modern TOF instruments commonly employ these two techniques to enhance resolution.

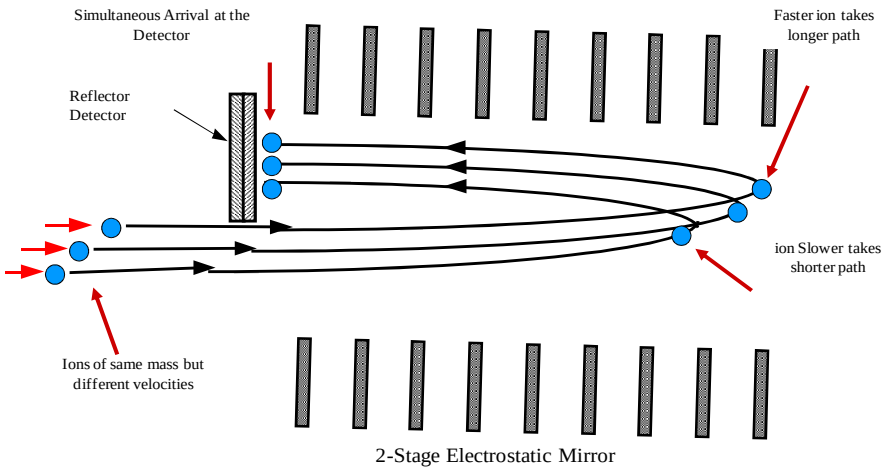

Figure 7 Velocity focusing in reflector mode (Training material ABI).

A more recent modification of MALDI instrumentation is the introduction of the orthogonal TOF mass analyzer with improved mass accuracy (< 10 ppm) and higher resolving power. In this instrument, ions are generated as usual, but by extraction, a voltage pulse is used to deflect the ions sideways so that the ion beam path is deflected perpendicular to the original direction of motion. Due to the elimination

of the high axial velocity distribution of the plume from the arrival time of the ions, a high resolving and a high mass accuracy are achieved (reviewed in Hillenkamp & Katalinić, 2007).

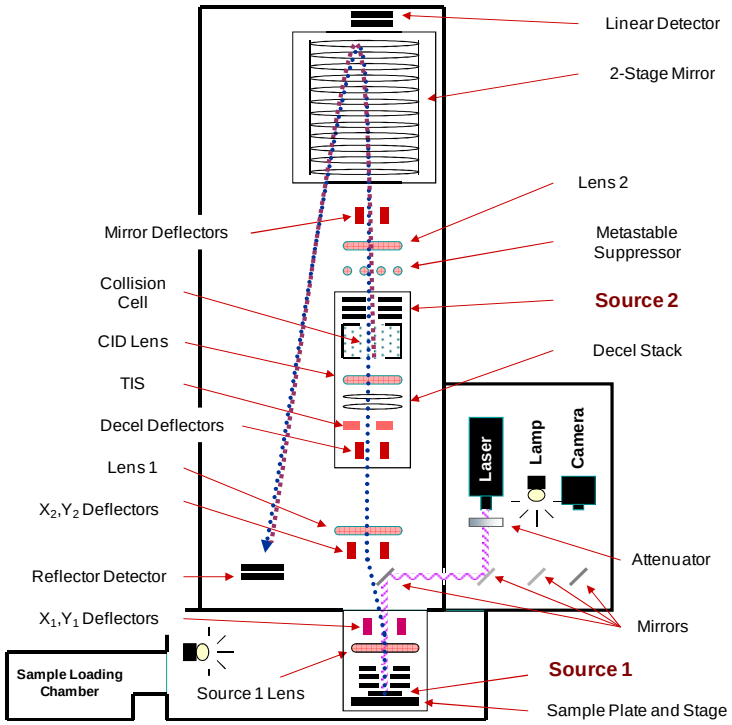

Figure 8 Schematic representation of a modern 4800 MALDI TOF/TOF™ Analyzer (Training material ABI)

2.1.2.3 Other detectors

The MALDI-LTQ-Orbitrap constitutes a recently introduced instrumentation in which MALDI is implemented as ion source. In this instrument, ions are trapped and confined by application of appropriate radio frequency (RF) and direct current (DC) voltages with

their final position maintained within the centre of the ion trap. These ions are then ejected succeedingly from the linear ion trap, whereby ions of interest are detected by conventional secondary electron multipliers. For high mass accuracy measurements, the ions are ejected from the linear trap to the C-Trap, where they are captured, cooled by collision with nitrogen gas and are squeezed into a smaller cloud. The ion packet is then ejected towards the Orbitrap analyzer, where the voltage on the central electrode increases and forces the ion packets into circling around the electrode. As a result the ions start oscillating along the central electrode (forwards and backwards). The image current is recorded on the outer split electrode. The signals are amplified and transformed by Fourier transformation into a mass spectrum (Makarov *et al*., 1999, 2000; Hardman& Makarov, 2003; Hu *et al*., 2005). Futhermore, 'Hybrid' instruments have been designed to combine the capabilities of different mass analyzers and include the MALDI QqTOF (Shevchenko *et al*., 2000) or MALDI-FT-ICR (Stoop *et al.*, 2008).

2.1.3 Tandem TOF Mass spectrometer

For detailed structural analysis fragment spectra are generated from the peptides precursor masses. Proteins and peptides contain C- and N-terminal amino acids. Cleavage of peptides at the peptide bonds yields fragments containing N- and the C- termini. As proposed by Roepstorff and Fohlmann (Roepstorff and Fohlmann, 1984), peptide-bond fragments carrying the N-termini from the original peptide are termed b-ions and y-ions for those containing the C-termini (see figure 9). Depending on the amino-acid sequence fragment ions of the type x-, z-, a-, c-may show up and are promoted by higher energy collisions. In TOF analyzers, fragmentation can be achieved either by making use of metastable decay or by externally inducing fragmentation. The decay, which occurs as a result of the induction of excess internal energy on a molecule from the MALDI process, can be classified into in-source and post-source decay. In in-source decay, c-and z type fragment ions formed during desorption/ionisation processes are detected with low intensities. Fragment ions, which result from decays in the field free region of the flight tube, are known as post source decay (PSD) products. They drift along the flight tube with the same velocity as the precursor ions, but except for fragment ions with masses close to the precursor are poorly focused on the reflector detector, therefore MALDI PSD analyses in simple reflectron instruments have never become a widely applied tool. This

problem was solved in today´s TOT/TOF instruments by extending the time for PSD in a low-acceleration voltage first linear TOF and than applying post-acceleration to a high final ion energy for fragment-ion analysis in the second reflectron TOF instrument.
The fragmentation pattern in PSD favors backbone cleavages producing predominantly a_n, b_n b_{n-17} y_{n-17} and y_n type fragment ions with very little (if any) side chain specific cleavages. When introducing a collision gas (N_2, Ar) into TOF/TOF systems, collisions are in the intermediate energy range and boost the yield of immonium ions $[H_2N{=}CH\text{-}R]^+$ from some individual amino acids in the peptide. These ions can be observed in the low mass region of the spectrum and assist in the sequence assignment by confirming the presence of an amino acid.

The MALDI TOF/TOF approach thus differs from conventional MS/MS techniques used for ESI, where mostly induced collisions of mass selected precursor ions with air or neutral gas molecules are used to obtain fragmentation in proteomic studies (CID) (Shukla & Futrell, 2000). This process yields high intensity a_n, b_n, and y_n type fragment ions, however, in contrast to MALDI doubly or more highly charged ions are typically used as precursors if available. Furthermore, in ESI, peptides have been fragmented using techniques such as Electron transfer dissociation (ETD) or electron capture dissociation (ECD). ETD and ECD produce predominately c and z type ions (Zubarev *et al.*, 1998*;* Zubarev *et al.*, 2003*;* Syka *et al.*, 2004*)*, but these new fragmentations techniques are restricted to multiply charged ions and thus exclusively used for ESI.

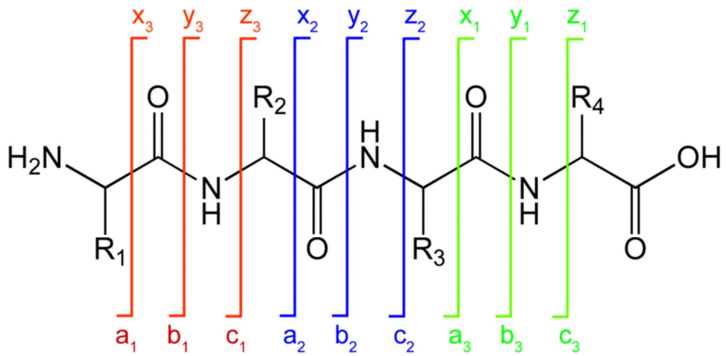

Figure 9 Normenclature of Peptide fragmentation as decribed by Roepstorff and Fohlman (1984).

2.2 Protein Identification

Traditionally, protein identification was performed by N-terminal analysis using the Edman degradation (Edman, 1950). During this process, the amino-terminal residue was labeled and cleaved from the peptide without disrupting the peptide bonds between other amino acid residues. The major limitations of this method were the difficulty to analyse samples with multiple posttranslational modifications (PTMs), its finite sensitivity as well as the high analysis time, preventing high throughput analyses. Additionally, peptides over 50 to 60 amino acids could not be fully sequenced due to incomplete cyclical derivatation (Liu *et al.*, 2007). With the evolution of mass spectrometric technologies, protein identification has been achieved using:

2.2.1 Bottom-up approach

Bottom up protein identification is the most widely used approach in proteomics. Prior to MS and or MS/MS, proteins are cleaved into peptides with a specific protease. Their masses are then measured and compared with generated peptide masses of known protein sequences (Pappin *et al.*, 1993) using database search engines such as Mascot (Perkins *et al.*, 1999) and Sequest (Eng *et al.*, 1994). There

are two common approaches for protein identification using the bottom-up approach, peptide mass fingerprinting and tandem MS (MS/MS) or the shotgun method.

2.2.1.1 Protein identification by peptide mass fingerprinting (PMF)

In peptide mass fingerprinting, peptide masses obtained from an MS measurement are compared to calculated peptide masses generated by "in silico" cleavage of protein in the database. This method is limited to single proteins or simple mixtures of proteins (Pappin *et al.,* 1993). Furthermore, a considerable number of peptides with high mass accuracy are required for protein identification. This method is advantageous for a fast screening of unknown samples and most of all; it can be performed with the same instrumentation used for MS/MS.

2.2.1.2 Shotgun protein/ Tandem MS identification

In the shotgun or tandem MS method, protein digestion is performed without any prefractionation/separation and peptides are separated either by one dimensional method or by multidimensional chromatography followed by tandem mass spectrometric analysis. Tandem mass spectra are collected for as many peptides as possible, and the results are then searched by an algorithmic comparison via SEQUEST (Eng *et al.,* 1994) or MASCOT (Perkins *et al*., 1999) against a database of proteins derived from genomic sequencing to identify the peptides. For proteins which are not found in the database, the amino acid sequence or stretches can be deduced from the fragment ion spectrum (Mann & Wilm, 1994; Hunt *et al.,* 1986). Though the shotgun approach seems simple, it has some limitations: The digestion of complex protein samples results in greatly increased complexity of the generated peptide mixture therefore, highly sensitive and efficient separation is a necessity. In addition, not all peptides resulting from the digestion of a protein(s) can be observed or correctly identified with MS/MS analysis, especially those with diverse or unexpected modifications (Han *et al*., 2008). Information is also lost upon the conversion of intact proteins into a mixture of peptides, which can lead to incorrect identifications (Han *et al*., 2008). Furthermore, the limited dynamic range of mass spectrometric analysis only allows for the peptides present at high relative abundance to

be preferentially sampled, while information regarding the proteins represented as low abundance peptides in the complex mixture is commonly not obtained (Han *et al*., 2008).

2.2.2 Top-Down protein identification

In top-down proteomics, intact proteins are directly fragmented without prior digestion. The top-down strategy holds two major promises, the potential access to the complete protein sequence and the ability to locate and characterise PTMs. In addition, the time-consuming protein digestion required for bottom-up methods is eliminated (Wehr *et al.*, 2006). However, the application of the top-down approach within proteomics schemes would require efficient and reproducible protein fractionation and clean-up by e.g. HPLC to produce pure isolated protein samples. These would then be ionised by ESI and selected multiply charged protein ions would then be subjected to ECD or ETD and possibly CID to acquire comprehensive fragmentations data (Wehr *et al.*, 2006). However, these techniques require long ion accumulation, activation, and detection time, thus are termed low-efficiency processes and the fragmentation of proteins is not well understood (Reid & Scott, 2002) yet.

2.2.3 HPLC; On- and Off-line coupling (ESI, MALDI)

With the establishment of electrospray ionisation (ESI) (Whitehouse *et al.*, 1985) and matrix assisted laser desorption/ionization (MALDI) (Karas & Hillenkamp, 1988), routine analysis of large biomolecules became possible. Since ESI is a technique that generates ions directly from a liquid sample, the benefit of a combination with HPLC was obvious and technically realised almost immediately. However, the birth of high-throughput analysis was only possible in early 1989, when Hunt and his colleagues demonstrated that a triple quadrupole instrument could be used to fragment the multiply charged molecular ions produced by ESI (Hunt *et al.*, 1989). Shortly afterwards, Covey et al. demonstrated the use of tandem MS in combination with online HPLC to obtain peptide sequence information from a complex proteins mixture after tryptic digest (Covey *et al.* 1991). Contrary to ESI, the identification of proteins by MALDI relied on the peptide mass fingerprint approach (Yates *et al.*, 1993*;* James *et al.*, 1993). This was due to the fact that MALDI instruments were initially coupled to time-of-flight (TOF) mass analyzer (Zhen *et al.*, 2004). These systems were originally not capable of generating

MS/MS data and even after the discovery that peptides could be sequenced by post-source decay (PSD) (Kaufmann *et al.*, 1993), MALDI-TOF MS systems were not as capable of routinely yielding peptide sequence information as ESI-MS/MS systems (Zhen *et al.*, 2004). Coupling of fast spotting devices with RP-HPLC and the combination of MALDI with Tandem MS mass spectrometers (Fountain *et al.*, 1994; Li *et al.*, 1994; *Doroshenko &* Cotter, 1995; Medzihradszky *et al.* 2000*;* Loboda *et al.*, 2000) initiated and established high-throughput instrumentations using TOF as mass analyzers.

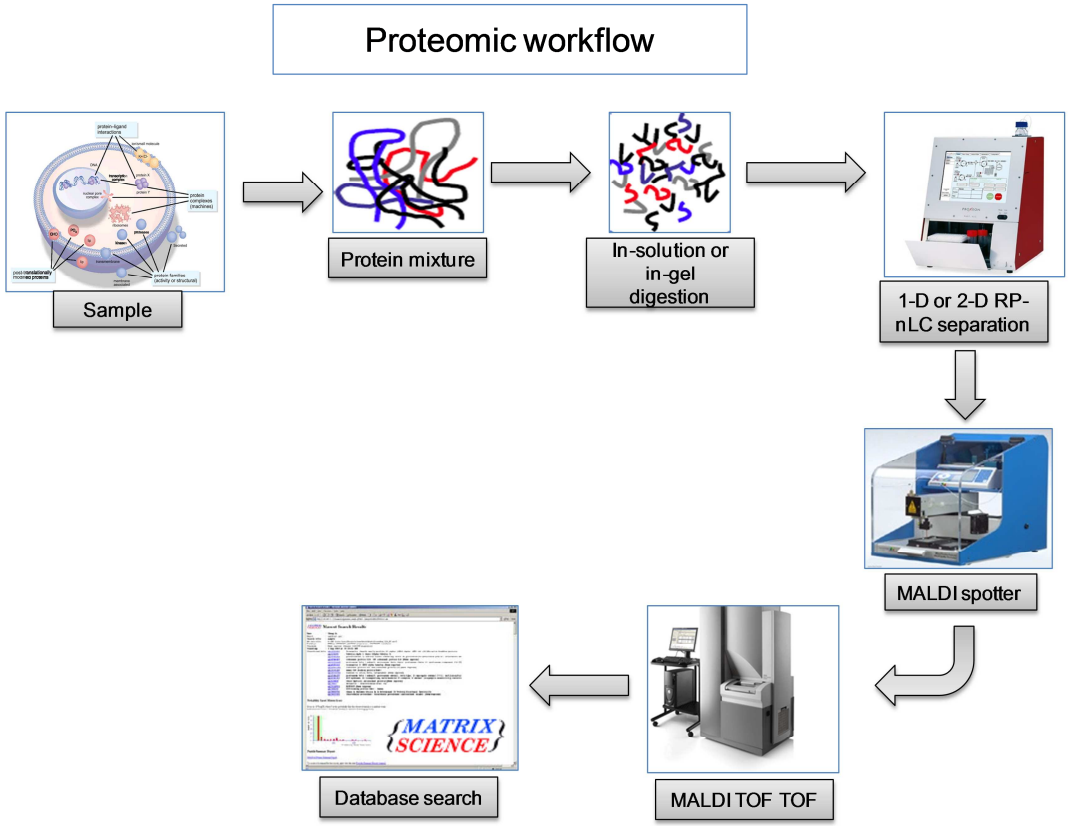

Figure 10 Workflow employed for the analysis of proteome samples.

3 Proteomics application

For the evaluation of the established proteomics workflow, different biological samples were analyzed. These samples present different challenges for the nLC-MALDI-MS/MS workflow. The success of this method was determined by the number of proteins and peptides identified. All samples were digested in-solution either with trypsin and or elastase. The samples were analysed using the Agilent 1100 series and the nLC Easy from Proxeon, using a 150 min gradient. With the exception of the elastase and tryptic digestion of *Halobacterium salinarium* and *Corynebacterium glutamicum*, all separations were carried out at 40 °C.

3.1 Lipid particle proteome

Lipid particles, lipid droplets or oil bodies are the lipid storage organelles of all organisms. In vertebrate animals, the most abundant energy reserve is stored as triacylglycerol in the lipid droplets of adipocytes (Brasaemle *et al., 2004*). Other types of cells contain tiny lipid droplets that store primarily cholesterol esters, which serve as a reservoir of cholesterol for the synthesis and maintenance of membranes. Steroidogenic cells of the adrenal cortex, testes, and ovaries use stored cholesterol additionally as a source of substrate for steroid hormone synthesis. Little is known about the mechanisms that control the flux of neutral lipids into and out of lipid droplets in any type of cell, but it is clear that the processes that control lipid traffic in adipocytes are central to the regulation of whole body energy metabolism (Brasaemle *et al., 2004*). Lipid particles have also been associated with many diseases such as atherosclerosis, diabetes or obesity (Beller *et al.,* 2008). Furthermore, due to their dynamism, they participate in several cellular processes and interact with various other cellular compartments (Beller *et al.,* 2008). Lipid droplets are composed of a highly hydrophobic core formed from neutral lipids, which represent more than 95 % of lipid droplet consisting of the non-polar lipids triacylglycerols (TAG) and steryl esters (STE) (Athenstaedt *et al.,* 1999). The surface of the lipid particle consists of a monolayer of phospholipids, which protects the highly hydrophobic interior from the cellular environment (Tauchi-Sato *et al.,* 2002). Specific sets of proteins, the lipoproteins, are embedded on the membrane monolayer of phospholipids (see figure 11).

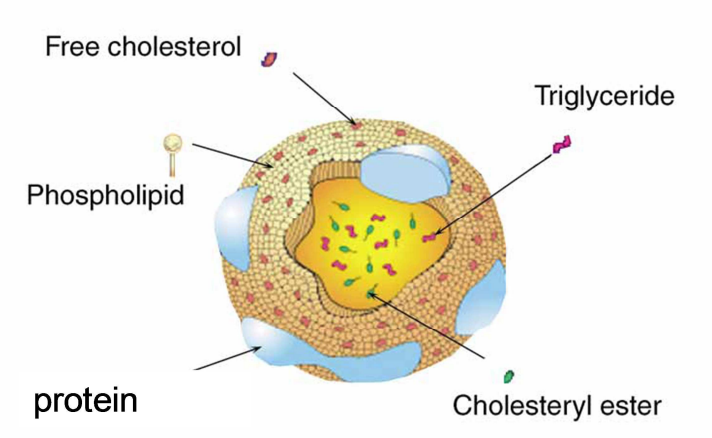

Figure 11 Schematic representation of a lipoprotein *(Kostner, 2002).*

The protein composition of lipid droplets from Chinese hamster ovary fibroblasts (*Liu et al.,* 2004*),* cultured human HuH7 hepatoma cells (Fujimoto *et al.,* 2004*),* cultured human A431 epithelial cells (*Umlauf. et al.,* 2004), and mouse mammary glands (*Wu et al.,* 2000*)* have been studied. These studies revealed that lipid droplet-associated proteins include enzymes involved in many aspects of lipid metabolism. In mammals, lipoproteins such as perilipin, adipophilin, hormone-sensitive lipase (HSL), adipose TAG lipase (ATGL), TIP47 (Tail-Interacting Protein 47 kDa), PAT-proteins, S3-12 and OXPAT have been identified (Goodman *et al.,* 2008*;* Olofsson*et al.,* 2008*;* Thiele & Spandl, 2008*;* Farese Jr. &Walther, 2009*;* Goodman *et al.*,2009*;* Bickel*et al.,* 2009*;* Robenek*et al.,* 2009*;*Yamaguchi & Osumi, 2009; Ohsaki*et al.,* 2009).

Though the yeast proteome has often been studied, only a limited number of yeast lipoproteins have been identified. Lipid particle proteins involved in ergosterol biosynthetic pathway such as serol-Δ^{24}-methyltransferase (Erg6p) (**Leber** *et al*., 1994) and Erg7p (**Leber** *et al.,* 1998) reveals a significant contribution to yeast cellular sterol formation (Zweytick *et al.,* 2000). Besides these, enzymes associated with activation of fatty acid such as Faa1p (long-chain fatty acid CoA ligase 1), Faa4p (long chain fatty acid CoA ligase 4) and Fat1p

(very long-chain fatty acid activator) have been reported. Faa1p and Faa4p are responsible for 99 % of cellular myristoyl-CoA and palmitoyl-CoA synthetase activity (Smid *et al.,* 1995; Watkins *et al.* 1998) and have also been suggested to stimulate the uptake of exogenous fatty acids (Knoll *et al.,* 1994). Furthermore, the identification of 1-acyldihydroxyacetonephosphate reductase, Ayr1p (Athenstaedt & Daum, 2000), which catalyses the formation of lysophosphatidic acid from 1-acyldihydroxyacetonephosphate through dihydroxyacetonephosphate pathway, associated its involvement in phosphatidic acid biosynthesis. Of all the lipoproteins identified, only the gene product of YDL193w is essential for cell growth (Choi & Martin, 1999). Through localisation studies, it has been demonstrated that some of these lipoproteins such as Erg1p, Slc1p, Gat1p or Ayr1p are located both in the lipid particles and endoplasmic reticulum (**Athenstaedt & Daum,** 1997; **Leber** *et al.,* 1998; Athenstaedt *et al.,* 1999; *Zweytick et al.,* 2000).

3.2 Interaction of Mixed linage leukemia complexes

All informations necessary for the assembly and functions of a cell are present within the DNA but realised only through the expression of the RNAs and proteins encoded by the blueprint. The manner in which cells response to extracellular changes or how they associate to form multi-cellular organisation and how individual cells work has been a primary interest of many researchers. Protein-protein interactions are imminent to virtually every cellular process. They control a large number of cellular processes such as the interaction of protein kineases, protein phosphatases, glycosyl transferase, acyl transferase, prototeases etc., with their substrate proteins. Such protein-modifying enzymes encompasse a large number of protein-protein interactions in the cell and regulate all sort of fundamental processes such as cell growth, cell cycle, metabolic pathways and signal transduction (Phizicky & Stanley 1995). Understanding protein-protein interactions should facilitate the investigation of intracellular signalling pathways, modelling of protein complex structures, targeted-drug design and gain insights into various diseases and biochemical processes.

One of these processes, which have not fully been understood, is the development of cancer. Leukemia, a severe hematological disorder originating from the bone marrow, is characterised by abnormal increase of blood cells, usually leukocytes (white blood cells). It results from somatic mutation in the DNA. Certain mutations produce leukemia by activating oncogenes or deactivating tumor suppressor genes

and thereby disrupting the regulation of cell death, differentiation or division. These mutations may occur spontaneously or as a result of exposure to radiation or carcinogenic substances and are likely to be influenced by genetic factors (Wiernik, 2001; Robinette *et al*., 2001; Stass *et al*., 2000). Clinically, leukemia can be subdivided in two main groups; acute leukemia, which is characterised by the rapid increase of immature blood cells, and chronic leukemia, which is distinguished by the excessive amounts of relatively mature but still abnormal white blood cells. Depending on the type of blood cells affected it can be further divided into lymphoblastic or lymphocytic leukemias and myeloid or myelogenous leukemias. Acute lymphoblastic leukemia (ALL) originates from progenitor cells able to differentiate into the B or T-lineage. Acute myeloid leukemia affects progenitors able to differnetiate into macrophages or granulocytes. In both cases, normal red blood cells and platelets are dramatically reduced. When combined, these two groups result in 4 different categories; Acute Myelogenous Leukemia (AML), Acute Lymphocytic Leukemia (ALL), Chronic Myelogenous Leukemia (CML) and Chronic Lymphocytic Leukemia (CLL).

Chromosomal rearrangements involving the *MLL* (*MLL1, ALL1, TRX,* and *HTRX*) gene, including balanced and unbalanced translocations, inversions, insertions, and a partial tandem duplication, have been associated with a heterogeneous group of lymphoid, myeloid, and mixed lineage leukemias (MLL) (for review see Meyer *et al*., 2006; Meyer *et al*., 2009). MLL is a histone methyltransferase considered to be a positive global regulator of gene transcription (Hess *et al*., 2004). This protein belongs to the group of histone-modifying enzymes and is involved in the epigenetic maintenance of transcriptional memory (Guenther *et al*., 2005). The *MLL* gene is located at chromosome 11q23, consists of 37 exons (Nilson *et al*., 1996; Rasio *et al.*, 1996; Marschalek *et al*., 1997) and encodes a protein of 4.005 amino acid residues with an estimated molecular mass of 435 kDa. Most *MLL* gene rearrangements map to an 8.3-kb breakpoint cluster region (bcr) and result in production of two chimeric oncoproteins (see figure 12; reviewed in Meyer *et al*., 2006; Meyer *et al*., 2009) by fusion of the amino terminal portion of MLL with the carboxy terminal portion of a partner gene, and *vice versa*, by fusing the N-terminal portion of a partner gene to the C-terminal portion of the MLL protein. Although the molecular basis for the oncogenic activity of MLL fusion proteins is incompletely understood, recent reports (Milne *et al*., 2002; Harper *et al*., 2008) demonstrate that MLL can mediate high levels of homeobox genes *HOXC8, HOXA7* and *HOXA9* expression through the binding of the MLL fusion protein to these *HOX* gene promoters.

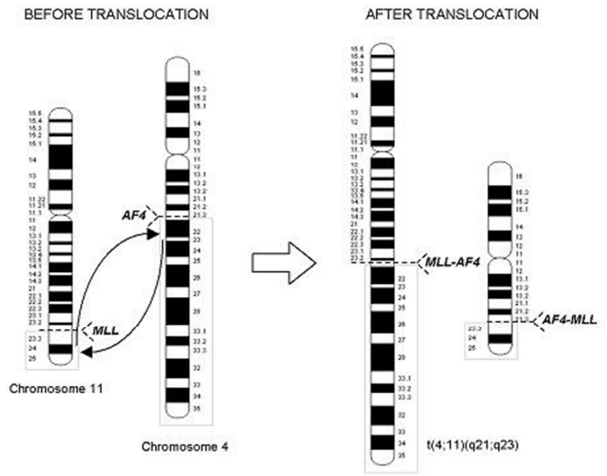

Figure 12 **Schematic representation of chromosomal translocation** t(4;11)(q21;q23), fusing the N-terminal region of the *MLL* gene on chromosome 11 to the C-terminal region of the *AF4* gene on chromosome 4, and *vice versa* (source http://www.erasmusmc.nl/alkg-cs/242905/243025/figure1).

To date approximately 60 different 11q23 chromosomal partners have been identified (Eguchi *et al.*, 2003; Meyer et al., 2006; Meyer *et al*., 2009). However, about 6 of these fusion genes (*ALL1* fused gene from chromosome 4, 6, 9 and 10, (*AF4, AF9, AF10, AF6),* eleven nineteen leukemia (*ENL*), and eleven nineteen lysin rich leukemia gene *(ELL)*) are responsible for 85 % of all translocations. The *AF4* (or *FEL*) gene is the most common *MLL* fusion partner. In all, 95 % of t(4;11)-associated leukemia is classified as ALL, but shows some lineage infidelity, typically expressing surface markers of the myeloid lienage ($CD13^+$, $CD33^+$). Therefore, some of these patients are characterised as Mixed Linage Leukemia. Just like the AF4 other homologous family members of these genes such as *AF9* or *ENL*

encode a serine/proline-rich nuclear protein with the ability to activate gene transcription (Nakamura *et al.,* 1993; Domer *et al.,* 1993; Prasad *et al.,* 1995).

Protein or genes in cells do not act in isolation. Their properties are controlled and modified through their interaction with neighboring proteins or peptides and other molecular components. Therefore, the biochemical functions of the proteins encoded by the MLL fusion partners are also influenced by the proteins with which they interact. Several interaction proteins of the MLL fusion partners have been identified (Bursen *et al.,* 2004). For example, it has been shown that the AF4 protein binds also to the AF9 and ENL (Erfurt *et al*., 2004; Zeisig *et al*., 2005), two other MLL translocation partners, as well as cyclin-dependent kinase 9 (CDK9) and Cyclin T1, a heterodimer that regulates transcriptional elongation (Estable *et al*., 2002; Benedikt et al., 2010).

Advances in the understanding of the association of proteins in a cellular system emerged primarily from two types of experiments. The first method involves the immunoprecipitation of the protein of interest with any associated proteins. The protein in the resulting immune-complex is then separated on an SDS-PAGE, transferred to a membrane and probed with antibodies suspected as partners of the target protein. The second major approach is the yeast two-hybrid system. In this method, detection of the interaction partner is obtained indirectly. Each gene that code for the protein of interest is fused to a transcription factor and the pair of hybrid genes is expressed in yeast *Saccharomyces cerevisiae* (Field & Song, 1989). A positive interaction allows the reconstitution of a functional transcription factor, and thus the expression of a selectable marker. One advantage of this system is that the target proteins are immediately identified through sequence of the cognate cDNA. Another advantage is that this procedure can be automated so that very large number of proteins can be screened for their ability to interact in the yeast two-hybrid assay. Another technological advance in the study of protein interaction was the improvement of mass spectrometry. In this approach, the proteins of interest are immuno- or affinity-precipitated, digested and analysed on a mass spectrometer. The protein of interest can also be resolved on SDS-PAGE, stained, digested, extracted and measured. In a relatively straightforward approach, the proteins resulting from a specific isolation of interaction partners are digested separated on a RP-HPLC and either online or offline connected to a mass spectrometer for MS and MS/MS data acquisition. In this approach, identification and quantification can be achieved in a single run. However, it is limited by the fact that peptides do not immediately reveal

whether associated proteins interact directly or indirectly. Other techniques such as cross-linking, affinity blotting, immunoprecipitation, Protein Affinity Chromatography or Bait and Reserve bait have also been used in protein-protein interaction studies.

3.3 Membrane Proteomics

Protein molecules that are attached to or associated with the membrane of cells or organelle are termed membrane proteins. More than half of all proteins interact with the membranes, therefore they play a vital role in numerous cellular functions (**Shama *et al.*,** 2007) such as signal transduction, cellular adhersion, ion transport and drug resistance (Ramus *et al.*, 2006). Based on the nature of the membrane-protein, they can be classified into two broad categories-integral (intrinsic) and peripheral (extrinsic) (see figure13).

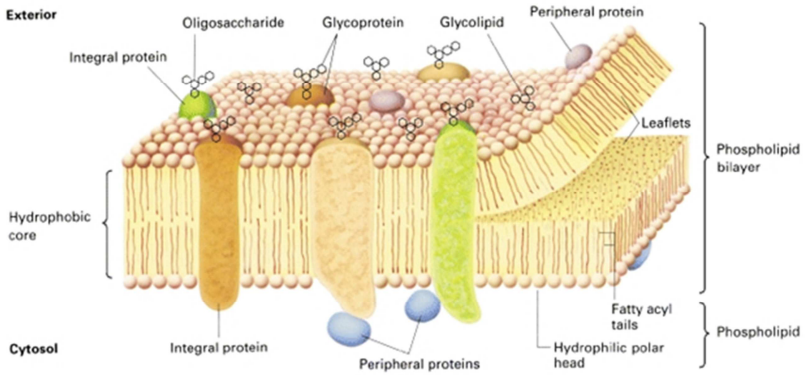

Figure 13 Schematic diagram of typical membrane proteins in a biological membrane. (Lodish *et al.*, 1999).

Peripheral membrane proteins or extrinsic proteins are bound to the membrane indirectly by interactions with integral membrane proteins or directly by interactions with lipid polar head groups. For example, the regulatory protein subunit of many transmembrane

receptors and ion channels may be described as peripheral membrane proteins (Berk *et al.*, 2000). On the contrary, integral membrane proteins (IMPs) or intrinsic proteins are permanently attached to the membrane. Due to their location, they connect a host of cellular processes, such as intercellular communication, vesicle trafficking, ion transport, protein translocation/integration, and propagation of signalling cascades (Ott *et al.*, 2002, , Torres *et al.*, 2003*;* Zhou *et al.*, 2004). As a result, they form the major protein class for drug targets (**Fischer *et al.*,** 2006). These proteins have one or more membrane-spanning domains as well as domains, from four to several hundred residues long, extending into the aqueous medium on each side of the phospholipid bilayer (Berk *et al.*, 2000). According to their structure they can be classified either as α-helical or as β-barrel proteins (**Fischer *et al.*,** 2006).

β-Barrel proteins, or porins, exist in the outer membranes of Gram-negative bacteria, chloroplasts, and mitochondria, where they regulate membrane integrity and allow for the passive influx/efflux of small molecules (Speers &Wu, 2007). The β-strands possess both lipophilic and hydrophilic characters, therefore the overall hydropathy of these IMPs are similar to those of soluble proteins (von Heijne *et al*., 1996; Santoni *et al*., 2000; Chen *et al*., 2002). Compared to their α-helical IMPs counterparts, these proteins do not present much analytical challenges (Eichacker *et al.*, 2004) in proteomics schemes. Approximately 20-25 % of all open reading frames (ORFs) in most genomes code for integral membrane proteins (von Heijne *et al*., 1998*;* Krogh *et al.*, 2000). The α-helix bundle can be subdivided into: Bitopic or single-pass IMPs with one transmembrane domain connecting 2 globular domains (Hurwitz *et al.*, 2006). These proteins often act as cell surface makers, receptors or adhesion factors with the cytoplasmic domains operating in cellular signalling pathways or in contact with the cytoskeleton (Speers & Wu, 2007). In contrast, polytopic IMPs have multiple transmembrane domains (TMDs) that span the membrane multiple times such that different regions of the protein are exposed on opposite sides of the membrane (Bowie *et al.*, 1997). Bound to the membrane surface via non-covalent interactions with phospholipid head groups or membrane-embedded proteins are the membrane-associated proteins. These proteins can readily be solubilised by treatment with high-pH or high-salt buffers, thus do not generally present the same analytical challenge as IMPs. Last but not least, the monotopic or membrane-anchored proteins, which are not transmembrane proteins but are rather bounded to the membrane bilayer by a lipid anchor. These proteins can be dissociated upon cleavage of the anchor with phospholipases and they have the hydrophilic characteristics of soluble proteins.

Since integral membrane proteins are composed of both hydrophilic and hydrophobic regions and behave more like lipids than proteins due to their association with lipid bilayer in the membrane, extraction, solubilisation and characterisation by mass spectral analysis have been a major challenge (Santoni *et al.,* 2000). Prior to proteolytic digest and mass spectrometric analysis, gel based electrophoretic methods such as two-dimensional polyacrylamide gel electrophoresis have frequently been used to separate membrane proteins. Unfortunately, most of the membrane proteins are not solubilised in non-detergent isoelectric focusing sample buffers and those solubilised tend to precipitate at their isoelectric point (Santoni *et al.,* 2000; Wu *et al., 2003*). Due to improved stationary phase chemistry, the advances in chromatographic methods, and the possibility of coupling either online or offline with high resolution mass spectrometry, gel free methods are increasingly been used in most research laboratories. Moreover, in these techniques, the tedious process of excision, digestion extraction and gel-to-gel reproducibility, which are associated with gel based electrophoretic methods, are avoided.

Despite some success, the analysis of membrane proteins still presents some major challenges, such as solubilisation in solvents, which are mass spectrometry compatible (Santoni *et al*., 2000) or their low abundance. Therefore, prefractionation and/or enrichment steps of the membrane fractions are a necessity **(**Fischer *et al.,* 2006)**.** Purification/enrichment steps such as centrifugation, density gradient centrifugation using sucrose, sorbitol, ficoll, or percoll (Goo *et al*., 2003; Klein *et al.,* 2005) using high-ionic-strength or high-pH buffers (sodium chloride, potassium chloride, sodium bromide, and potassium bromide) with subsequent sonication to remove membrane associated proteins (Pedersen *et al*., 2003; Schluesener *et al.,* 2005) etc. are commonly used in most laboratories. For the analysis of a more complex plasma membrane of higher eukaryotes, the colloidal silica and aqueous-polymer two-phase partitioning have widely been used (Speers & Wu, 2007). Protein precipitation in methanol/chloroform, cold acetone or trichloroacetic acid (TCA) has also been used to purify/enrich membrane proteins (Wessel *et al*., 1984; le Maria *et al*., 1993; Mastro *et al*., 1999; Carboni *et al*., 2002; Borner *et al*., 2005). Furthermore, due to the hydrophobic nature of certain domains of IMPs, exposure to aqueous solvents results in aggregation, adsorption and precipitation. This leads to sample loss and interference of enzymatic access during digestion. Therefore, caution is necessary when choosing solvents for solubilisation and denaturing of IMPs since this might affect separation, digestion and mass spectrometric efficiencies.

The analysis of integral membrane proteins can be restricted only to the soluble domains of the proteins. In this case the soluble domain is cleaved either with a chemical reagent or with proteolytic enzymes. The generated peptides are then separated and analysed (Wu *et al.,* 2003). To improve sequence coverage both the soluble and the hydrophobic domains are necessary. In most cases, 2 step digestion procedures can be used to analyse both domains. In the first step, the soluble domain is shaved either with a proteolytic enzyme such as Lys-C (Nielsen, *et al.,*2005*;* Bihan, *et al.,* 2006), Proteinase K (Wu *et al., 2003*) or trypsin (Rodriguez-Ortega *et al.,*2006) or a chemical cleavage reagent (cyanogen bromide) (Washburn *et al*.,2001). The intact bilayer with the embedded proteins is solubilised and after extraction of the hydrophobic domains, both soluble and hydrophobic domains are either combined, digested and analysed or separately digested and analysed (Wu *et al.,* 2003; Goshe *et al.,* 2003; Blonder *et al.,*2004; Blonder *et al.,* 2004; Wu *et al.,* 2006; Fischer *et al.,* 2006). To avoid limitations of the digestion efficiency by only addressing specific cleavages sites in the 2 step approach and to improve sequence coverage, less specific enzymes have been employed, also in combination with specific ones or chemical reagents. Furthermore, for high solubility, mostly high organic solvent content were used for the solubilisation in the second step. Unfortunately, no optimal approach has been developed; therefore there is still a need for simpler and more efficient methods of extraction and analysis of membrane proteins.

Table 2 Summary of Solubilisation/Digestion techniques used for the analysis of integral membrane proteins (Speers & Wu 2007).

solubilisation/digestion techniques	**comments**
soluble domains (membrane shaving)	
High pH proteinase K (hppK)	Non specific protease, not reliant on specific cleavage sites in soluble domains, high-pH opens vesicles all owing access to both sides of membrane

Lys-C	Requires Lys in soluble domains
Trypsin	Requires Lys or Arg in soluble domains
embedded and soluble domains	
hppK-CNBr/FA	Combines benefits of non specific protease shaving with targeted TMD solubilisation/analysis, gives ca 97% IMPs and ca 68 % TMD peptides
60 % MeOH/trypsin	Efficient solubilisation strategy for targeting both soluble and TM domains, gives ca 45 % IMPs
trypsin/hppK-60 % MeOH/trypsin	Combines benefits of non specific proteases having with good solubilisation of TMD regions, gives ca 40 % IMPs
trypsin-60% MeOH/trypsin/chymotrypsin	Some enrichment afforded by shaving, good solubilisation and use of orthogonal enzymes for TMD regions, gives 20−50 % IMPs, targets hydrophobic peptides

3.3.1 Quantification of membrane proteins

Due to the importance of membrane proteins, not only the identification, but also the quantification of their presence would improve the understanding of expression level in cells. Two dimensional SDS-PAGE has been used for the differential comparison of cell states (Tonge, *et al.*, 2001). However, this technique has been shown to have poor resolving power with highly hydrophobic proteins (Santoni *et al.*, 2000). With the introduction of high resolution mass spectrometers and successful combination with high resolution separation techniques, the identification and characterisation were the main goal of most researchers. In the past decade, not only the identification and characterisation of proteins but also the determination of the absolute or relative abundance has become a major task in proteomics. Several quantification methods have therefore been developed, whereby shotgun-proteomics MS-based techniques are more dominant. Shotgun proteomics MS-based quantification techniques can be divided into 2 major groups; label-free and stable isotope labelling. Label-free methods are based on quantitation of peak areas/intensities (Scheurer *et al.* 2005; Foster *et al.*, 2005; Wang *et al.*, 2007) or the number of MS/MS spectra for a particular peptide. Stable isotope labelling involves the incorporation of heavy atoms into proteins/peptides by metabolic means or post-processing chemical modification. At the protein level, labelling commonly involves the modification of nucleophilic amino acids (mostly Cys or Lys) with isotopically labelled tags either after protein isolation (but before digestion) such as Isotopic Coded Affinity Tag (ICAT) (Gygi *et al.*, 1999) or already on the cellular level, i.e Isotope Coded Protein Label (ICPL) or Stable isotope labelling with amino acids in cell culture (SILAC) (Ong *et al.*, 2002). Peptide based relative quantification such as Tandem Mass Tag (TMT) (see figure 14) (Thompson *et al.*, 2003), isobaric tag for relative and absolute quantitation (iTRAQ) (Ros *et al.*, 2003) focus on lysine site chains and N-terminal amino acids. As a further option, enzyme-mediated incorporation of ^{18}O (Yao *et al.*, 2001) enriched water into peptide C-termini may be employed. In any case, during nLC-MS/MS analysis, isotopologic peptides bearing heavy and light tags elute simultaneously but seperate in the mass spectrometer and the relative ion peak areas can be quantified, giving an indirect value of relative protein abundance (Spears & Wu *et al.*, 2006). Despite the advances in relative quantification studies, quantitative studies for membrane proteins may be handicaped by the reduced number of peptides generated from IMPs using standard solubilisation/digestion techniques. Therefore, for the quantification of membrane proteins, these protocols have to be optimised.

Absolute quantification on the other hand enables a direct comparison of the absolute protein content of both cellular states. Different methods such as Isotope dilution (Barr et al. 1996), absolute quantification using deuterium-labeled peptide (Barnidge *et al*., 2003), Absolute quantification of Proteins (AQUA) (Gerber *et al*. 2003), QconCAT (Beynon *et al.*, 2005) have been used to quantify the absolute protein amount in a sample. Most of the absolute quantitation methods in proteomics rely on the well-established principles of stable isotope dilution techniques, using tryptic peptides as surrogate analytes for the proteins of interest. The standard peptides might be chemically synthesed or by the design of a DNA construct that is transcribed and translated into a protein concatamer (Beynon *et al*., 2005). Therefore, absolute quantification is necessary because it offers a direct comparison of samples analysed at different times, using different instrumental platforms and within different laboratories, using a universal stable isotope-labeled internal standard of known concentration.

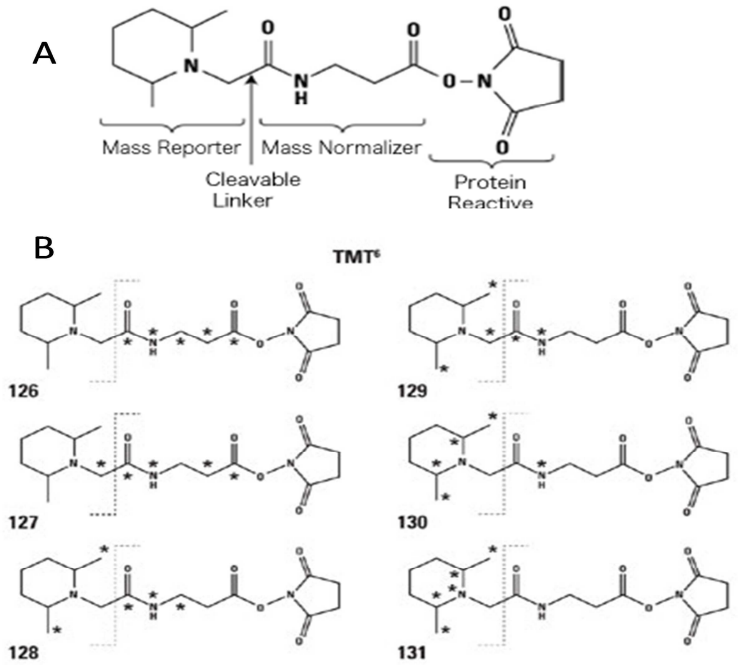

Figure 14 Structure of tandem mass tag reagent. A. TMT0 has no isotopic substitutions and is used for method development. B. A six-plex of isobaric mass labels each with five isotopic substitutions per tag is used (Thermo-scientific).

3.4 Protein phosphorylation

Covalent modifications that occur either during or after ribosomal assembly of proteins are termed post-translational modifications. These modifications usually change the properties of a protein and can determine its activity state, localisation, turnover and the interaction with other proteins (Mann & Jensen, 2003). Posttranslational modifications such as phosphorylation, acetylation, methylation, glycosylation or deamidation have been in the focus of recent research. Reversible phosphorylation of proteins is one of the

most studied post-translation modifications due to its role of controlling protein function and directing the physiological responses of cells and their environment (Blackburn & Goshe, 2009). A fundamental understanding of biological processes and signaling networks at the molecular level requires the analysis and characterisation of these phosphorylated proteins (Larsen *et al.*, 2005). At least one-third of all proteins are thought to contain covalently bound phosphate on serine, threonine and tyrosine residues, with phosphoserine being the most abundant, at some time-point (Morandell *et al*., 2006). Hunter and coworkers determined that the relative abundances of phosphotyrosine (pY), phosphothreonine (pT), and phosphoserine (pS) in normally growing cells to be 0.05 %, 10 %, and 90 % respectively (Hunter & Sefton, 1980). The low abundance of phosphotyrosine and phosphothreonine was assumed to be that the modification of these residues occurs mostly on low abundant proteins. Furthermore, it was presumed that pY is less stable in phosphoamino acid analysis than pS/pT. Therefore, pY sites would be underrepresented in the phosphoamino acid analysis (Olsen *et al*., 2006).

Despite recent developments in proteomics methodology, especially in MS, there is no ultimate method for the identification of protein phosphorylation. This is presumably due to the facts that most signaling proteins are not abundantly expressed and the stoichiometry of phosphorylation can be quite low (Morandell *et al*., 2006). Protein phosphorylations have been characterised mainly using biochemical and mass-spectrometry based methods. The classical biochemical approaches include proteins isolated from cells via immunoprecipitation with monoclonal antibodies. In these methods, the phosphoproteins are often monitored by using [^{32}P] ATP to characterise kinase activity and potential substrates. The degree of phosphorylation can be measured by phosphoimaging based on the incorporation of a ^{32}P-radiolabel, immunoblotting using antibodies recognising amino acid residues modified with a phosphate moiety, and phosphorylation specific staining using Pro-Q Diamond. Yet, these methods do not provide detailed information regarding the phosphorylation or associated phosphorylation site or stoichiometries (Blackburn & Goshe, 2008). Furthermore, the ^{32}P labelling has been combined with Edman degradation (Yan *et al*., 1998; van der Geer *et al*., 1999; Haystead *et al*., 1999). However, this method did not only have long analysis time but also exhibited a low dynamic range and low abundance phospho-proteins escape detection, therefore is not suited to complex samples (Larsen *et al*., 2005).

Mass-spectrometry based strategies are currently the method of choice to detect protein phosphorylation, since they provide information on the position and degree of phosphorylation (Mann *et al*., 2002). In these methods, phosphoproteins are digested with a

specific protease and a subsequent analysis of the peptide pool is carried out by MS and tandem mass spectrometry to detect a mass increment of 80 Da per phosphate group. Despite many advances in this field, detection and identification of phosphorylated species is challenging because of low abundances, therefore enrichment techniques are typically employed prior to analysis. Widely used is Immobilised metal affinity chromatography (IMAC). It is based on the affinity of the negatively charged phosphate group to metals like Fe^{3+}, Ga^{3+}, Zr^{4+}, Al^{3+} (Hochuli *et al*., 1987; Posewitz & Tempst, 1999; Nuhse *et al*., 2003). An alternative to IMAC for the selective enrichment of phosphorylated peptides prior to MALDI-MS or liquid chromatography tandem MS is Titanium dioxide (TiO_2) (Ikeguchi & Nakamura, 1997; Pinkse *et al*., 2004; Sano & Nakamura, 2004; Larsen *et al*., 2005). This metal oxide resin show preferential binding of phosphopetides to non-phosphorylated acidic peptides at acidic pH. Both techniques are vital for phosphopeptide analysis but have their advantages and limitations. Compared to IMAC, TiO_2 is more compatible with a number of reagents commonly used in the solubilisation and digestion of proteins (Jensen & Larsen, 2007). IMAC is more specific for multi-phosphorylated peptides while TiO_2 is to mono-phosphorylated peptides (Ficarro *et al*., 2002; Jensen & Larsen, 2007). Since the metal ions are not covalently bound to the substrate, leaching has been reported during the enrichment process on IMAC (Hochuli *et al*., 1987). The selective enrichment of large amounts of acidic non-phosphorylated peptides seems to be a limitation shared by both techniques.

To minimise the binding of acidic non-phosphorylated peptides on IMAC and TiO_2, chemical and non-chemical modification of peptides has been reported. One of these methods is based on blocking the acidic residues by O-methyl esterification (Ficarro *et al*., 2002; Moser & White, 2006; Ndassa *et al*., 2006), but O-methyl esterification often causes partial deamidation and subsequent methylation of Asn and Gln residues (Stewart *et al*., 2001; Larsen *et. al*., 2005). These by-products further complicate MS analysis and data interpretation. Furthermore, since esterification is a water-free reaction, the aqueous solvents must be completely evaporated. Adsorptive losses of the sample to the walls of the reaction tubes can not be completely ruled out, leading to decrease in analyte sensitivity. In another method, Adamczyk et al. focused on the conversion of phosphoserine and phosphothreonine residues to S-(2-mercaptoethyl) cysteinyl or -methyl-S-(2-mercaptoethyl) cysteinyl residues by -elimination/1,2-ethanedithiol addition, followed by reversible biotinylation of the modified proteins (Adamczyk *et. al*., 2001). However, this strategy suffers from several drawbacks such as lack of reproducibility, sensitivity and introduction of unwanted side reactions (McLachlin & Chait, 2003). A more prominent procedure to

minimise unspecific binding of acidic non-phosphorylated peptides without introducing unwanted side products is the exploitation of the concurrent binding with small carbonic acids such as 2,5-DHB, salicylic acid, phosphoric acid, phthalic acid, benzoic acid or glycolic acid during loading or washing steps (Larsen *et al*., 2005). Under specific experimental conditions, these acids interact more strongly with TiO_2 or IMAC beads than aliphatic carboxylic acids, but show a weaker interaction compared to the phosphate moieties (Larsen *et al*., 2005, Seeley *et al*., 2005). Observation from infrared spectroscopic studies have shown that substituted aromatic carboxylic acids (such as salicylic acid and phthalic acid) coordinate strongly to the surface of TiO_2, whereas mono-functional carboxylic acids (including benzoic acid and acetic acid) only interact very weakly with TiO_2 (Dobson *et al*., 2000). The trapped peptides are released from the TiO_2 beads by increasing the pH of the eluting solvent, usually ammonium hydroxide or ammonium bicarbonate (pH 9-10.5).

After enrichment and elution from the TiO_2 or IMAC beads, HPLC-ESI- tandem MS has been used as the method of choice for the global profiling of phosphopeptides. With the introduction of Electron Transfer dissociation (ETD) and Electron capture dissociation (ECD) fragmentation techniques, which occur along the peptide backbone in a sequence independent manner and preserve PTMs, the large scale analysis and annotation of the right phosphorylation sites is steadily increasing. Despite the advantages of MALDI over ESI such as higher (but still moderate) tolerance to detergents and salts, low sample consumption, high sensitivity and a non-chromatographic dependent analysis, it has been used mostly for manual phosphopeptide analysis. This is partially due to the limited choice of matrices for automation, given that the matrix affects the signal intensities and signal/ratio of phosphopeptides and fragmentation. DHB remains the dominant matrix for phosphopeptides analysis since its introduction as MALDI matrix (Strupat *et al*., 1991) and its subsequent optimisation with respect to analyte sensitivity using 1 % phosphoric acid (Kjellström & Jensen, 2004). DHB crystallises into thick long needle-like crystals at the edges of the spot. This distinct characteristic of DHB enables the spatial separation of analytes from the salts (in the middle of the spot). Intense signals are, in most cases, obtained only at the crystalline rim resulting in the generation of “hot spots”. This makes manual measurements superior, but time-consuming and difficult to reproduce (Strupat *et al*., 1991). In contrast, during automatic measurements a number of spectra are acquired from random positions on the spot. Due to the inhomogeneous morphology, intense and weak signals are averaged resulting in only moderate signal intensities. Therefore, automatic nLC MALDI MS/MS analysis of phosphopeptides using DHB in 1 % phosphoric acid becomes difficult. Furthermore, Also, DHB has a low absorption efficiency at 355

nm (Nd:YAG laser) (ε_{355nm} = 2.100 l mol^{-1} cm^{-1} (Tang, Allman, RCM), thus, high laser fluences are necessary for measurements. DHB has a low fragmentation yield because of its soft matrix characteristics (Glückman *et al*., 1999). Other matrices and matrix combination such as, CHCA, CHCA/diammoniumhydrogen citrate, 2,5-DHB/CHCA, THAP with ammoniumcitrate, THAP/ DHAP (Asara & Alison; 1999; Yang *et al*., 2004; Dunn *et al*., 2006) have been used for the analysis of phosphopetides.

Recent methods using MALDI plates coated with specific materials (TiO_2, $Fe^{3+}$and ZrO_2) for the enrichment and detection of low picomol amounts of phosphopeptides have been reported (Blacken *et al*., 2007; Torta *et al*., 2009; Hoang *et al*., 2010). For example, MALDI plates coated with poly (2-hydroxyethyl methacrylate) (PHEMA) brushes are derivatized with Fe (III)-nitrilotriacetate (NTA) complexes (MS) (Dunn *et al*., 2008) or immobilized zirconium on a phosphonate-terminated self-assembled monolayer (SAM) for specific phosphopeptide capture and direct analysis by MALDI MS (Hoang *et al*., 2010).

4 Objectives

Unlike gene products, the proteome is subjected to continuous changes as a result of environmental influences and cell activities. This results in the generation of huge numbers of proteins, isoforms and modified proteins. In addition, the large dynamic range of proteins makes the study of the whole proteome very challenging. Traditionally, electrophoresis separation using 1D SDS-PAGE or 2D-SDS-PAGE has been the method of choice for the study of protein biochemistry. With the establishment of electrospray ionisation and matrix assisted laser desorption/ionisation including the subsequent combination with gel-free and automation of separation technologies the identification, characterisation and quantification of proteins has become feasible. Since only a fraction of the total peptide population of a given protein is identified, information such varying types and locations of posttranslational modifications (PTMs) may be lost in shotgun proteomics. Other limitations such as sample generation, dynamic range, MS compatible solvents for protein solubilisation, required column/column material and the perfect enzymes for protein digestion still present challenges in the field of proteomics. However, this technique has enhanced our understanding of the proteome.

The aim of this study was to establish a nLC-MALDI MS/MS based methodology for the analysis of proteomic samples. The first section of this study was based on the method establishment. After the optimisation, the potential of this method was demonstrated by analysing a membrane protein sample using a less specific enzyme, elastase. Furthermore, the applicability of the established method was demonstrated by the analysis of protein-protein interaction, and the influence of the lipoprotein content from yeast cells grown in two different media was investigated. In the second section of this study, the method was extended to characterise posttranslational modifications in proteins using a biphasic pre-column. Finally, the nLC-MALDI MS/MS method was combined with an optimised Offgel-IEF separation technique for the analysis and quantification of proteomics samples.This shows that the established method is reproducible, very flexible and with small modifications, it can be used to answer many biological questions.

5 Methods and Material

5.1 Material

2,5-Dihydroxybenzoic acid (DHB), Bruker, Germany

Acetone, Carl Roth GmbH& Co, Germany

Acetonitrile, Carl Roth GmbH & Co., Germany

Agilent 1100 series HPLC, Agilent, Germany

Alcohol dehydrogenase 1-Saccharomyces cerevisiae, Sigma-Aldrich, Germany

Alpha-amylase, Bacillus amyloliquefaciens, Sigma-Aldrich, Germany

Alpha-S1-casein - Bos Taurus, Sigma-Aldrich, Germany

Alpha-S2-casein, Bos Taurus (Bovine), Sigma-Aldrich, Germany.

Ammonium bicarbonate, Carl Roth GmbH&Co, Germany

Ammonium hydrogencarbonate (extra pure)

Angiotensin II, Sigma, Germany

Auxillary pump, kdScientific

Beta-casein - Bos Taurus (Bovine), Sigma-Aldrich, Germany

Bradford-based protein assay kit, Bio-Rad, Germany

Carbonic anhydrase 2 -Bos Taurus (Bovine), Sigma-Aldrich, Germany

Column Loader, Proxeon, Denmark

Dichlormethane, Carl Roth GmbH& Co, Germany

Dithiotreitol, Carl Roth GmbH & Co, Germany

Easy nLC, Proxeon, Denmark

Elastase-1, Sus scrofa (Pig), Roche, Germany

Ethanol, Carl Roth GmbH & Co, Germany

Formamide, Sigma-Aldrich, Germany

Formic acid, Sigma- Germany

Fused silica capillary, Klaus Ziemer, Germany

Glu1-Fibrinopeptide B, Sigma-Aldrich, Germany

Hemoglobin subunit alpha - Homo sapiens, Sigma-Aldrich, Germany

Hemoglobin subunit beta - Homo sapiens, Sigma-Aldrich, Germany

Iodacetamide, Carl Roth GmbH&Co, Germany

Isopropanol, Carl Roth GmbH& Co, Germany

Methanol, Carl Roth GmbH&Co, Germany

MicroTee, Upchurch Scientific, USA

Ovalbumin-Gallus gallus, Sigma-Aldrich, Germany

Peptide Mass calibration kit, Bruker Daltonics, Germany

Phosphoric acid (85%)

Phosphoric acid, Sigma-Aldrich, Germany

Pipette, Eppendorf, Germany

Potassium silicate, PQ Corporation The Netherlands

Precast immobilized pH gradient Ready Strips™ IPG and the carrier ampholytes Bio-Lyte 3-10 (Bio-Rad Laboratories (Hercules, CA, USA).

Separation column, 75 µm × 150 mm column (Zorbax C18 300 Å 3.5 µm) Agilent, LC-Packings and Vydac); 75 µm × 150 mm column (C18 300 Å 3.5 µm)- LC-Packings ; 75 µm × 150 mm column (C18 300 Å 3.5 µm)- Vydac

Sequazyme™ Mass Standards Kit, Applied Biosystems, Germany

Sequencing grade trypsin, Sigma-Aldrich, Germany

Serotransferrin- Homo sapiens, Sigma-Aldrich, Germany

Serum albumin - Homo sapiens, Sigma-Aldrich, Germany

Serum albumin -Bos Taurus, Sigma-Aldrich, Germany

Spotter, Sunchrom, Germany

Titanosphere 5 µm, GL Sciences Inc Japan

Trifluoroacetic acid, TFA, ≥99.5 % (Carl Roth GmbH & Co., Germany).

Ubiquitin- Homo sapiens, Sigma-Aldrich, Germany

Ultrapure water (MilliQ^{TM} Bedford, MA, USA).

X-Bridge™ BEH 180 C_{18} 300 Å 3.5 µm, Waters Germany

α-cyano-4-hydroxycinnamic acid (CHCA) Bruker, Germany

All solvent were of HPLC quality and primium purity for all chemical except otherwise mention

5.2 Methods

5.2.1 Expression plasmids, cell culture and transfection experiments (AF4/ENL/AF4-MLL)

Protein expression, cell culture and transfection experiments were performed in Prof. Rolf Marschalek group by Anne Benedik and Sabrina Baltruschat according to (Baltruschat, 2009; Benedik, 2009).

5.2.2 Affinity purification (AF4/ENL/AF4-MLL)

Protein expression, cell culture and transfection experiments were performed in Prof. Rolf Marschalek group by Anne Benedik and Sabrina Baltruschat according to (Baltruschat, 2009; Benedik, 2009).

5.2.3 Immunoprecipitation experiments (AF4/ENL/AF4-MLL)

Protein expression, cell culture and transfection experiments were performed in Prof. Rolf Marschalek group by Anne Benedik and Sabrina Baltruschat according to (Baltruschat, 2009; Benedik, 2009).

5.2.4 Yeast strains and culture conditions

Yeast strains used in this study are listed in Table 1. Yeast cells were grown at 30 °C in rich media containing 1 % yeast extract, 2 % peptone, and 2 % glucose (YPD) or 0.3 % yeast extract, 0.5 % peptone, 0.1 % glucose, 0.5 % KH_2PO_4 and 0.1 % oleic acid (YPO). To solubilise fatty acids in YPO, 0.2 % Tween 80 was added to the media (Grillitsch *et al.*, 2010).

5.2.5 Isolation of Lipid particles from the yeast culture

The isolation of the plasma membrane samples and the lipoproteins were performed by Karlheinz Grillitsch in the group of Prof. Günther Daum according to published procedures (Serrano *et al* 1988; Leber *et al*., 1994; Zinser *et al*., 1995).

5.2.6 Protein quantification

Proteins from isolated lipid particles (LP) and homogenates were precipitated with 10%. Trichloroacetic acid (TCA). Proteins were quantified by the method of Lowry *et al*. (Lowry *et al*., 1951) using bovine serum albumin as standard.

5.2.7 Delipidation

The isolated lipid particleswere incubated with 2 volumes of diethyl ether with repeated vigorous shaking. After a high speed centrifugation step using a table top centrifuge the extracted non-polar lipids were with drawn and the remaining trace of diethyl ether was removed under a stream of nitrogen. The proteins were then precipitated with trichloroacetic acid, and the resulting pellets were dissolved in 25 mM ammonium bicarbonate.

5.2.8 Cell culture (Human Simpson-Golabi-Behmel syndrome)

Human Simpson-Golabi-Behmel syndrome (SGBS) cells were obtained from Prof. Dr. M. Wabitsch (University of Ulm, Germany) and were cultured according to published reports (Wabitsch *et al.,* 2001).

5.2.9 Induction of differentiation (Human Simpson-Golabi-Behmel syndrome)

Differentiation was induced by Anja Rosenow as described by Rosenow et al. (Rosenow *et al*., 2010).

5.2.10 Protein sample preparation (Human Simpson-Golabi-Behmel syndrome)

To discriminate between secreted proteins and cell-leakage proteins a similar BFA blocking strategy was used as previously described by Wang et al. (Wang *et al.*, 2004).

5.2.11 In-solution Proteolytic Digest

5.2.11.1 Digestion of Purple Membrane (Elastase/Trypsin)

The proteolytic treatment represents a modified version of the specific integral membrane peptide level enrichment (SIMPLE) as described in (Fischer *et al*., 2006). One hundred micrograms of PM were washed using 25 mM NH_4HCO_3 buffer (pH 8.0). After a 2 min centrifugation at 10,000 g the supernatant was removed and the pellet was gently re-suspended in methanol. The sample was sonicated for 20 min and the methanol was diluted to 60 % using 25 mM NH_4HCO_3 buffer. The proteolytic digest was started by adding either 2 µg of trypsin or 10 µg of elastase. The elastase digest was performed overnight at ambient temperature and the tryptic digestion at 37 °C. To prevent excessive binding of the membrane slurry to the tubes, the whole digest process was performed without any external shaking.

5.2.11.2 Digestion of CB membrane (Elastase/Trypsin)

To remove loosely attached cytoplasmic contaminants from the CB membrane, an additional digestive step was performed before the actual proteolytic treatment. Two hundred micrograms of washed membranes were digested with elastase in a 1:20 protein/enzyme (w/w) ratio in 25 mM NH_4HCO_3 for 12 hours at 30 °C. Then the supernatant was discarded after centrifugation and the pellete were re-suspended in methanol. After a 20 min sonication, the methanol was diluted with 25 mM NH_4HCO_3 buffer to 60 % and the same digestion procedure as described above, was performed.

5.2.11.3 Tryptic digest (Lipoprotein. Adipocztes)

100 µg of pellates were dissolved in 1 ml 25 mM NH_4HCO_3 solution in an Eppendorf tube. The disulfide bridges were reduced in 45 mM DTT for 1 hour at 60 °C and shaking in a Thermomixer (Eppendorf). The solution was allowed to cool to room temperature, and cysteines were alkylated in the presence of 100 mM iodoacetamide (IAA) for 45 min in the dark at room temperature. To avoid subsequent alkylation of the trypsin, the reaction was quenched after 45 min by adding an additional 12.5 µl of 45 mM DTT and

incubating for another 45 min at room temperature. Trypsin solution was added to the reduced and alkylated samples to obtain anenzyme/proteinratio of 1:50 (w/w) and the solution was incubated overnight at 37 °C. Digestion was stopped by addition of 1 µl of a 10 % TFA solution.

5.2.11.4 Tryptic digest (AF4, AF4-MLL and ENL protein complex)

For the in-solution digest 60 µl of each elution the sample was concentrated using vivaspin 500 centrifugal concentrator (Sartorius, Germany) and the buffer was replaced with 25 mM NH_4HCO_3 buffer (pH 8.0). After reducing the samples with DTT and alkylation with IAA, the proteolytic digest was started by adding 0.5 µg trypsin (Sigma) and incubated for 24 hours at 37 °C. The activity of the digestive enzyme was stopped by adding 1 µl 10 % TFA.

5.2.11.5 Tryptic/elastase digests of standard proteins

1 mg of each protein was dissolved in 1 ml 25 mM NH_4HCO_3 solution in an Eppendorf tube.100 µg of this solution was reduced by adding 45 mM DTT and incubating for 1 hour at 60 °C and shaking in a Thermomixer comfort (Eppendorf). The solution was allowed to cool to room temperature, and cysteines were alkylated with 100 mM iodoacetamide for 45 min in the dark at room temperature. 12 µl 45 mM DTT was then added and incubated for another 45 min at room temperature. Enzyme solutions were added to the reduced and alkylated proteins so that the enzyme/protein ratio of 1:50 (w/w) for trypsin and 1:30 (w/w) for elastase by weight was obtained. The tryptic sample was incubated overnight at 37 °C and elastase at ambient temperature. Digestion was stopped by addition of 1 µl of a 10 % TFA solution.

5.2.12 Reduction, Alkylation, Digestion and TMT-labelling of yeast plasma membrane

The depleted plasma membrane samples (YPO and YPD) were dissolved in 60 % methanol/TEAB, 0.5 M, (pH 8.2). Each sample was separated in three equal portions. After reduction and alkylation of disulfide bridges as described above, elastase digestion (Enzyme/protein ratio of 1/20) was performed overnight at room temperature. The peptide pool was then dried down.

The three YPD samples were labeled by TMT with reporters at m/z 126.1, 127.1, and 128.1 and the three YPO samples by TMT with reporters at m/z 129.1, 130.1, and 131.1 in 40.2 µL of CH3-CN each. The derivitisation was carried out under gentle agitation. During this process the N-termini and lysine residues were labelled. After 60 min of reaction at RT, 8 µL of 5 % hydroxylamine (w/v) was added in each tube and mixed for 15 min. The six samples were pooled in a new tube and diluted with 20 % methanol/1 % IPG ampholyte buffer to 2 ml. Pre-fractionion on an Offgel-IEF was performed as described above (section 4.2.9).

5.2.13 Construction of frit for pre-column and analytical column

88 µl potassium silicate and 16 µl Formamide were mixed in an eppendorf tube and vortexed for 1 min. One end of a fused silica capillary (8 cm, 100 µm ID) was immersed into this mixture and placed in an oven for 60 min at 110 °C with the coated end pointing downwards. After the polymerisation, the frit end is cut to about 4 mm and excess formamide was washed with isopropanol.

5.2.14 Pre-column loading

The column loader with a PTFE adapter was placed on a magnetic stirrer. Slurry of the C_{18} material was prepared in a 2 ml glass bottle using dichloromethane and inserted in the column loader. The capillary (5 cm, 100 µm ID) was then immersed into the column loader containing C_{18} material via a seal with the frit end pointing upwards and screwed. Under pressure of about 65 bar, 4 cm of C_{18} material was loaded into the column. The capillary was then removed and further compressed on the nLC with 100 µl solvent A (0.1 % TFA in 8 % ACN) at a flow rate of 5 µL/min.

5.2.15 Biphasic pre-column loading

For the biphasic pre-column, approximately 3.5 cm long C_{18} material preloaded column (8 cm, 100 µm ID) was inserted in the column loader containing TiO_2 slurry (5µm) and the same procedure was repeated. After loading approximately 3.5 cm of the TiO_2 material, the column was removed from the column loader and connected to the nLC and flushed with 50 µl solvent A at a flow rate of 2 µl/min. After a constant column pressure was obtained, the biphasic pre-column was shaped in other to reduce potential dead volume.

5.2.16 Analytical column

For the analytical column, fused silica capillary with 100 µm internal diameter and a length of 120 mm were filled with C_{18} materials at a pressure of 100-250 bar. To estimate the efficiency of the in-house packed columns, four standard peptides were separated with a short gradient and the UV spectrum (figure 15 A) was compared to that obtained from the same peptides using a commercial column (75 µm x 150 mm, C_{18} 3.5, 100 Å, LC packings Pepmap™ 100) (figure 15 B). Apart from a small retention time shift, both columns had almost the same peak width, thus it can be concluded that the in-house packed column had the same separation efficiency as the commercial column, but they showed a higher loading capacity and higher flow rates can be used for the separation.

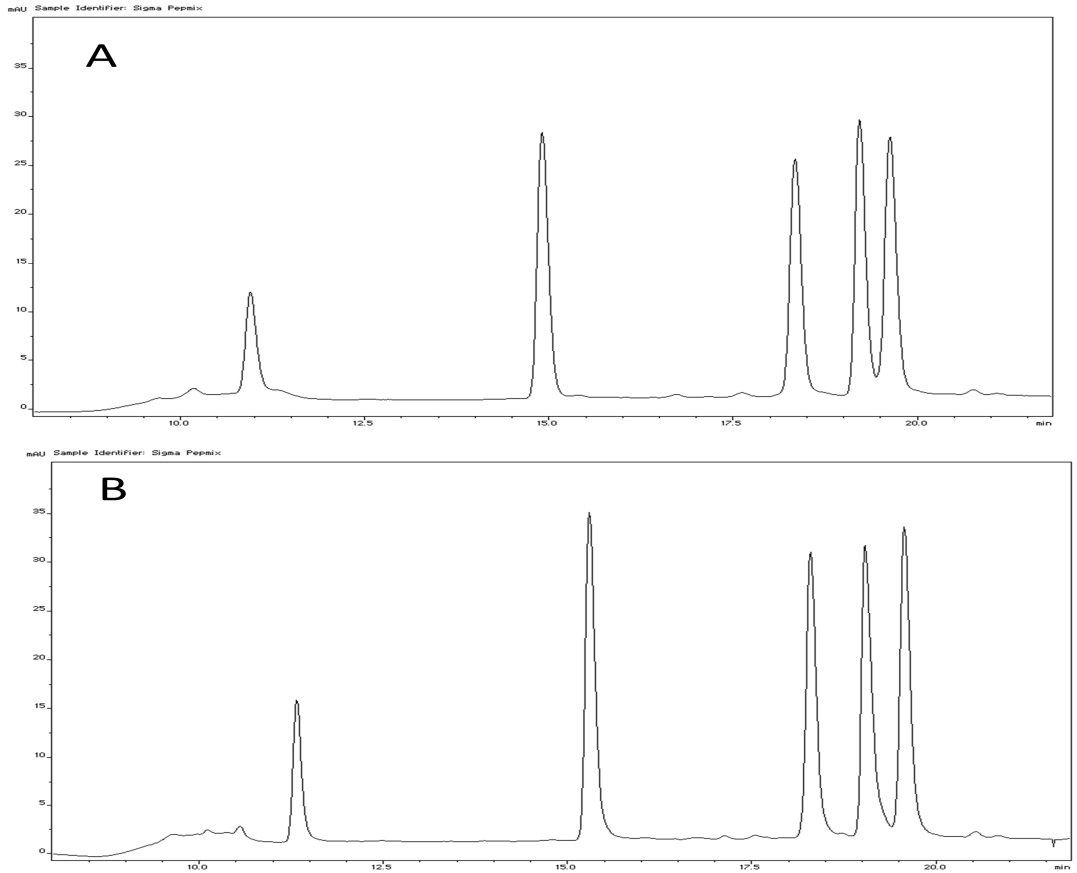

Figure 15 A shows a chromatogram of 4 standard peptides separated on the in-house packed column. B show a chromatogram of the same standards separated using 75 µm x 150 mm, C_{18} 3.5, 100 Å, LC packings Dionex Pepmap™ 100 column.

5.2.17 Nano-LC

The RP-HPLC separations were performed using 2 HPLC systems: the 1100 series from Agilent and the Easy-nLC from Proxeon.

5.2.17.1 Agilent 1100 series

The 1100 series consist of a capillary pump, nano-pump, a thermostatic autosampler, solvent degasser both on the nano- and capillary-pumps, a spotter and a 10-port-Micro valve. The sample was pulled from the autosampler and directly loaded to the pre-column using solvent A (0.1 % TFA in 8 % ACN) for on-line desalting via the capillary pump running at a flow rate of 20 µL/min for 5 min. After loading, the 10-port valve is the switched to the nano-pump and the sample is back-eluted to the analytical column for the separation step using a flow rate of 300 nl/min with increasing ACN concentration (solvent B 0.1 % TFA in 95 % ACN). Separation was performed on a 75 µm × 150 mm column (Agilent, LC-Packings and Vydac).

5.2.17.2 Proxeon Easy nLC

The Easy-nLC from Proxeon consists of 2 solvent pumps (A and B) and 1 sample pump, a refrigerated autosampler and a spotter. The sample was loaded to an in-house packed pre-column (X-Bridge™ BEH 180 C_{18} 300 Å 3.5 µm, 100 µm x 30 mm) or a 100 µm x 60 mm biphasic pre-column containing C_{18} and TiO_2 materials (Titansphere 5 µm) and desalted with 2 times the injected volume using solvent A (0.1 % TFA in 8 % ACN). Separation was performed on thermostatic (40 ^{o}C), 75 µm × 150 mm column (LC-Packings) or a custom made C_{18} column (X-Bridge™ BEH 180 C_{18} 300 Å 3.5 µm, 100 µm x 120 mm) at flow rates betwæn 300-400 nl with increasing ACN concentration (Solvent B 0.1 % TFA in 95 % ACN).

5.2.18 nLC gradient

The following gradients were used (table 3). Though some slight modifications were made to suit certain samples, the gradient slopes were maintained.

Table 3 Gradient profile used for method optimisation and separation.

Gradient 1		**Gradient 2**		**Gradient 3**	
Time(min)	**% B**	**Time (min)**	**% B**	**Time (min)**	**% B**
5	8	5	8	5	8
65	90	65	45	8	15
75	90	68	90	125	45
80	8	75	90	135	90
		80	8	145	90
				150	8
				155	8

5.2.19 Matrix for manual measurement

6 mg of each matrix, CHCA, DiFCCA and Chloro-CCA were dissolved in 1 ml 70 % acetonitrile/0.1 TFA. The matrices were then diluted so that each had a total concentration of 15 mM. The matrices and matrix mixtures containing ammoniumdihydrogen phosphate were prepared so that the final concentration of each matrix and ammoniumdihydrogen phosphate was 15 mM and 3 mM respectively. For the DHB matrix, 10 mg was dissolveed in 33 % Acetonitrile/ 1 % H_3PO_4.

5.2.20 Matrix for LC-MALDI spotting

The effluent from the nLC was pre-mixed with a matrix from an auxiliary pump (flow rate, 1.2 µl/min) via a microTee and spotted every 20 seconds on a MALDI 123 x 81 mm Opti-TOF™ metal target. This solution contained 15-20 mM CHCA in 70 % ACN/0.1 % TFA, spiked with 20 fmol final amount of [Glu^1]-Fibrinopeptide B per spot for internal calibration.

5.2.21 Isoelectric focusing (IEF) fractionation (Offgel-IEF)

Isoelectric focusing (IEF) was carried out on an Agilent Technologies 3100 OFFGEL fractionator device (Böblingen, Germany) using the 12-well setup on 13 cm pI 3-10 linear gradient IPG strips. Prior to sample loading, the IPG gelstrips in each well were re-hydratedwith 20 µl of focusing buffer. The sample was then diluted in a solvent buffer containing 20 % methanol (MeOH) and 1 % IPG buffer as final concentrations. To prevent the outer chambers from running dry through excessive water transport during focusing, additional 100 µl 1 % IPG solution was added to fractions 1 and 12 and 50 µl to fractions 2 and 11. The samplewas focused with a maximum current of 50 µA and typical voltages ranging from 500 to 4000 V until 20 kVh was reached. After focusing the fractions were removed and the wells were incubated at room temperature with methanol for another 15 min. The incubated and focused fractions were then combined and vacuum centrifuged to about 10-12 µl.

5..2.22 MALDI MS/MS

Mass spectra were acquired using an Applied Biosystems/MDS Sciex 4800 TOF/TOF™ Analyzer. The instrument is equipped with a Nd:YAG laser, emitting at 355 nm with a repetition of 200 Hz. All spectra were acquired in the positive reflector mode between 700 and 4500 *m/z* with fixed laser intensity. A total of 750 laser shots per spot were accumulated. External calibration was performed using an 8-point plate model with peptides from the Sequazyme™ Peptide Mass Standards Kit, Applied Biosystems. In order to obtain a higher precursor mass accuracy for subsequent MS/MS experiments, the MS spectra were internally recalibrated to the spiked [Glu^1]-Fibrinopeptide B. Fragmentation was performed with collision energy of 1 kV using air as collision gas at a pressure of approximately 1 x 10^{-6} Torr, a source voltage of 8 kV, a collision cell voltage of 7 kV and an accelerating voltage of 15 kV in the second source. To reduce sample consumption during measurement, stop conditions for MS/MS were defined. A minimal number of 15 peaks above 45 S/N with at least 12 accumulated sub-spectra, a minimum of 1250 and maximum of 2500 laser shots were recorded. To avoid unnecessary multiple selections of identical precursor, MS/MS precursor selection was carried out via the instrument's job-wide interpretation software. Depending on sample, between 6 and 10 precursors per spot with a minimum signal-to-noise-ratio of 100 were selected for fragmentation. Potential matrix signals were removed from precursor selection, by excluding all masses in the range from 700 to 1400 *m/z* obtained from measuring spots without analytes.

5.2.23 Database searches

Mascot Generic Format (MGF) files were retrieved from each nLC-MALDI MS/MS run using a built-in Peaks2Mascot (software 4000 explorer) feature, exporting up to 65 peaks per MS/MS spectrum, each requiring a minimum signal-to-noise of 5. The MGF files were processed using the Mascot™ database search engine v2.2.03 (Matrix Science Ltd., UK). The following settings were used: Enzyme, trypsin or none, allowed missed cleavages for trypsin 3 and 9 when no enzyme (None) definition was used; fixed modification: carboxymethylation of cysteine; variable modification: oxidation of methionin; MS precursor mass tolerance was set to between 30-50 ppm and MS/MS mass tolerance was 0.5 Da. In addition to the above settings, phosphorylation (STY) was set as a variable setting when

analysing phosphopeptides and amino-terminal pyroglutamate formation was considered as variable modification in all purple membrane database searches. Depending on the sample, the search was performed with a custom *Halobacterium* database generated from SwissProt/TrEMBL containing 2490 entries as at November 30th 2007, a *Corynebacterium glutamicum* ATCC 13032 Bielefeld database containing 4008 sequences was provided by Jörn Kalinowski (41), DB-Standard containing 24 proteins, Swissprot and NCBI database. False positives were estimated by the internal Mascot™ decoy database search function.

Table 4 Mascot search parameters

	Yeast Cells	Human fat cells	Standard proteins	Standard proteins	AF4, ENL	CB	PM	Yeast Cells
Taxonomy filter	Saccharomyces Cerevisiae	Homo sapiens	All entries	All entries	Homo sapiens	All entries	All entries	All entries
Enzyme	Trypsin	Trypsin	Trypsin	None	Trypsin	None/Trypsin	None/Trypsin	None
Max missed cleavages	3	3	3	9	3	3	3	9
Fixed modifications	Carbamidomethyl (C)	Carbamidomethyl (C)	Carbamidomethyl (C)	Carbamidomethyl (C)	Carbamidomethyl (C)	Carbamidomethyl (C)	Carbamidomethyl (C)	TMT6plex (K),TMT6plex (N-term),
Variable modifications	Oxidation (M)	Oxidation (M)	Oxidation (M)	Oxidation (M)	Oxidation (M)	Gln->pyro-Glu (N-term Q),Oxidation (M)	Gln->pyro-Glu (N-term Q),Oxidation (M)	Oxidation (M)
Peptide mass tolerance	50 ppm	50 ppm	50 ppm	50 ppm	50 ppm	50 ppm	50 ppm	50ppm
Fragment mass tolerance	0.3 Da	0.3 Da	0.3 Da	0.3 Da	0.3 Da	0.3 Da	0.3 Da	0.3 Da
Mass values	Monoisotopic	Monoisotopic	Monoisotopic	Monoisotopic	Monoisotopic	Monoisotopic	Monoisotopic	Monoisotopic
Instrument type	MALDI-TOF-TOF	MALDI-TOF-TOF	MALDI-TOF-TOF	MALDI-TOF-TOF	MALDI-TOF-TOF	MALDI-TOF-TOF	MALDI-TOF-TOF	MALDI-TOF-TOF
Total sequences	6717	466739	24	24	466739	3059	2490	6912
Total residues	3019518	165389953	7388	7388	165389953	955253	701243	3167490
Sequences after taxonomy filter	6717	20402	24	24	20402	3059	2490	6912
Database	SGD	SwissProt	DB-standard	DB-standard	SwissProt	Cg-ELA1_PIG	HALSA-ELA1_PIG	Yeast_Swissprot
Quantifcation	none	none	none	none	none	none	none	yes

5.2.24 Result interpretation and data analysis

All statistic analyses were based on peptides having Mascot™ MS/MS ions scores exceeding the "identity or extensive homology threshold" ($p < 0.05$). In the case of multiple fragmentations of identical precursors, due to recurrence in repetitive runs, only data from the highest scoring peptide were kept.

The hydrophobicity of each peptide was calculated using the GRAVY-Scores principles from Kyte and Doolittle (1982).
For the calculation of the isoelectric points, detail information of each amino acid from Nelson and Cooks *Lehninger Principles of Biochemistry 3rd Ed.* (2001) were taken into consideration.

Transmembrane regions of proteins were predicted with the *TMHMM 2.0*-Algorithmus (Krogh *et al.*, 2001) and displayed with the TMRPres2D (Spyropoulos *et al*., 2004) software.

5.2.25 Analysis of secreted protein candidates (Human Simpson-Golabi-Behmel syndrome).

For verification of secreted protein candidates a SwissProt analysis of amino acid sequences with SignalP and SecretomeP was performed. SignalP 3.0 predicts the presence and location of signal peptide cleavage sites in amino acid sequences. SecretomeP 2.0 predicts non-classical, in particular not signal-peptide-triggered protein secretion. In this process the non-classically secreted proteins should obtain an NN-score above the threshold of 0.5, but not at the same time be predicted to contain a signal peptide.

Table 5 List of phosphopeptides used for method optimisation

Peptide Sequence	Monoisotopic Mass
VDNIRpSAT	955.42
DpSFDDAFKA	1095.39

DRVYIHPF	1126.34
HYQPpYAPPR	1208.52
TWTpLCGTPEY	1250.46
SDLESQLAQpSR	1313.56
KLLpSFTSNDILR	1485.76
VRFPSpHFSSDLK	1499.69
FQpSEEQQQTEDELQDK	2060.82
VADPDHDHTGFLpTEpYVATR	2303.99

6 Results and discussion

6.1 Method Optimisation

6.1.1 Optimisation of spot deposition

The establishment of instruments capable of accurately depositing nano-liter on a target paved the way for the direct analysis of complex proteomic samples using MALDI. With the combination of RP-HPLC, the LC separation was then preserved on the MALDI target. Decoupling the separation from the MS and MS/MS analysis consequently provides the opportunity for the researcher to independently optimise the separation and the MS and MS/MS performance, improving the overall quality of the result. During this process, the eluent from the LC is mixed with a matrix and spotted on a target. The rate of deposition is independent of the flow rate and the chromatographic resolution. Though the deposition on a MALDI target preserves the LC chromatogram, the chromatographic peak width is affected. Compared to LC-ESI, where the chromatographic peak widths of most peptides are less than 60 sec, the same peptides exhibit peak width of over 60 sec in MALDI experiment.

In MALDI spotting peak broadening can result from sample mixing with the matrix in the microTee or from longer and large transfer capillaries after the analytical column. Furthermore, capillary carry-over from one spot to the next during deposition and spotting frequency can also result in peak tailing. To determine the source of the peak broadening in our system, 13 standard peptide and three different concentrations were separated and spotted every 15, 20 and 30 sec at three different concentrations (250, 100 and 33 fmol) using an 80 min gradient. The flow rate from the LC and the matrix pump were kept constant. After MS acquisition, the distributions of the peptides in different fractions were plotted against the signal-to-noise ratio of total number of spots for each gradient (see figure 16). Figure 16 shows a graphic representation of 2 standard peptides with 20 sec spotting interval. The peptides were eluted in more than 10 and at least 6 spots for the 15 and 20 sec spotting interval. An extensive tailing was observed for all peptides used.

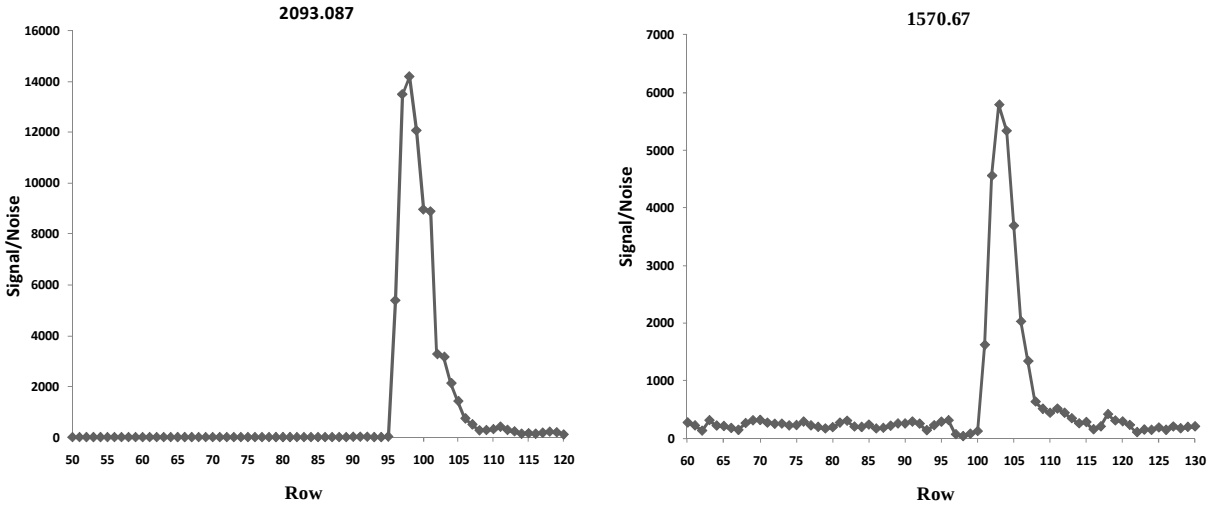

Figure 16 Signal-to-noise plot of 2 peptides (100 fmol). The spots were deposited every 20 sec on a MALDI target and a 50 cm transfer capillary with 100 µm internal diameter after the analytical column.

Tailing has a general disadvantage that the peptide is distributed in over several spots. Hence, the sample is diluted, thus leading to a decrease in sensitivity. For a complex sample, where different dynamic range of peptides concentration exist, this would result in possible peptide suppression by more abundant peptides, which might prevent selection of less abundant ones for MS/MS analysis.
With 30 sec spotting interval, only a minimal change was observed in the number of spots where the peptides were eluted (see figure 17). Though the number of spots decreases, the elution time of the peptides were still over 2 min.

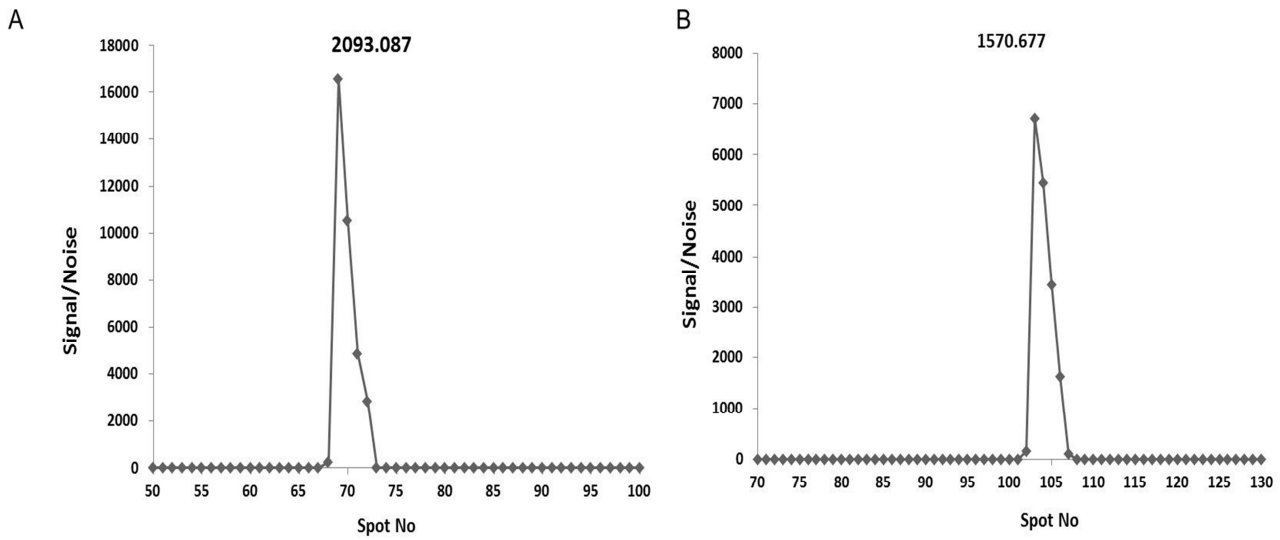

Figure 17 30 sec spotting frequency using the same 50 cm transfer line (100 µm) and concentration (100 fmol).

Since the peptides were still eluted in more than 3 fractions, the transfer capillary with internal diameter (ID) of 100 µm was replaced with a 15 cm fused silica capillary with internal diameter of 30 µm. The same peptide mixture was separated and spotted every 20 sec on a MALDI target. The same interpretation procedues were repeated using the same 2 peptides. These modifications greatly reduced the number of spots in which the peptides were detected (figure 18). The peptides appeared in at most 3 fractions, thus the peak maxima are exclusively in the second fraction. However, the best scenario would be cases were peptides are present in only 2 spots.

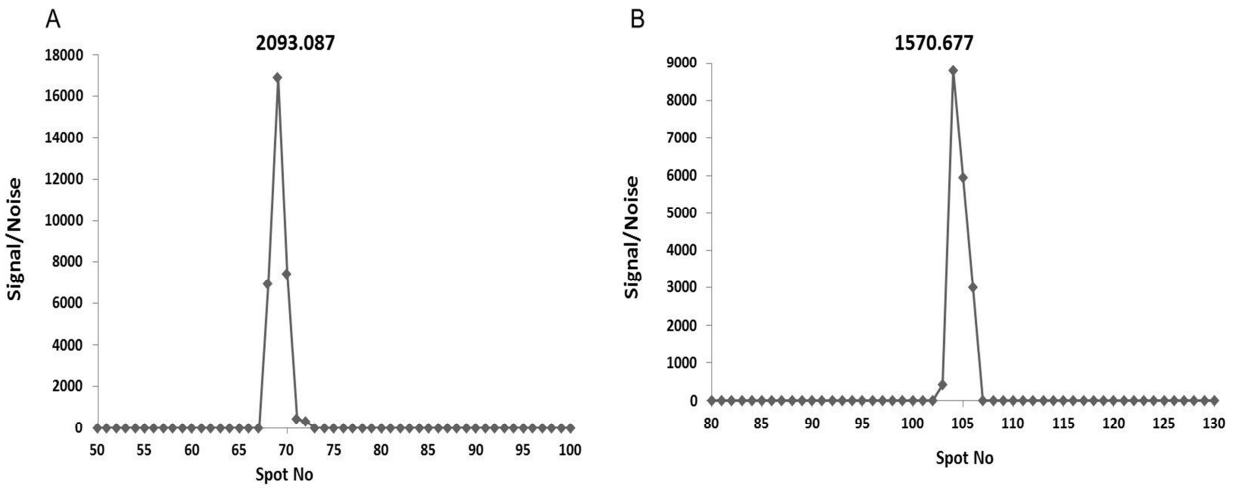

Figure 18 Signal-to-noise plots of 2 peptides. The spots were deposited every 20 sec on a MALDI target and a 50 cm transfer capillary with 30 µm internal diameter after the analytical column.

In LC-MALDI with automatic deposition dervices, it is assumed that the peak broadening could result from mixing of the matrix and the separated sample in the microTee. So the microTee was replaced with a Y connector. This connector is constructed in a way that both the matrix and the eluent from the nLC, coming from 2 different capillaries flow into one large capillary that is conneted to the needle (see figure 19). Compared to the MicroTee, the Y-connector should improve laminar flow, thus resulting in less turbulence when mixing. It was therefore expected that a Y-connector would improve the peak width, tailing and the number of fractions peptides elute.

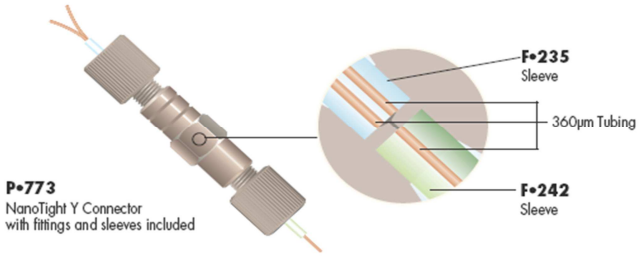

Figure 19 NanoTight Y-Connector

However, experiments using this Y-connector resulted in unexpected tailing and carryover (see figure 20). Also installing the Y-connector to the spotter was a technical challenge. For this reasons it was not used in further experiments. Experiments without a mixing chamber did not further improve the results.

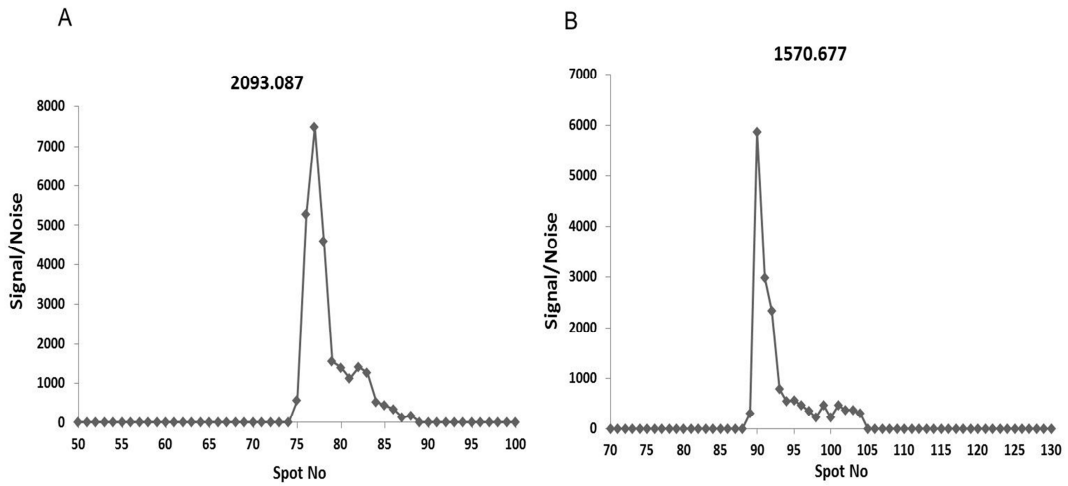

Figure 20 Extracted ion chromtograpy of 2 peptides witt the y-connector as mixing chamber.

The use of a transfer capillary with smaller inner diameter and shorter distance between column and MALDI target reduced the number of spots in which peptides are eluted. However, peptides were still eluted in 3 spots. Probably, spotting frequency of 30 sec or more would further reduce the number of spots in which peptides are eluted. In general, a compromise has to be made with respect to spotting time. Spotting frequency of 30 sec and above would still result to complex spots and thereby higher chances of losing less abundant peptides. On the other hand, spotting frequences below 20 sec might lead to diluted samples, thus lower intensities. Therefore, for further experiments during this study, a spotting frequency of 20 sec was used.

6.1.2 Matrix concentration for automatic spotting

For LC-MALDI-MS/MS experiments it is necessary to have enough matrix material on the target for a considerable amount of MS and MS/MS experiments and at the same time to keep the matrix-to-analyte ratio optimal. For high-throughput experiments, usually

more than 5 precursors are chosen for fragmentation. Therefore, it is important that enough matrixis present on each spot. Matrix concentrations ranging from 2 mg/ml up to saturated matrix solution have been used for manual crystallization. For automatic deposition concentrations above 4 mg/ml result in crystal formation on the needle, especially when the gradient starts approaching high ACN concentration. This resulted in extensive carryover and in some case, spots containing no matrix (see figure 21). Concentrations below 2 mg/ml work well, but due to ablation during MS and MS/MS, the number of accumulated spectra is limited. Matrix concentrations between 3 and 4 mg/ml were determined to be optimal for our proteomics analysis and 1:4 (v/v) samples: matrix ratio was used for all analysis.

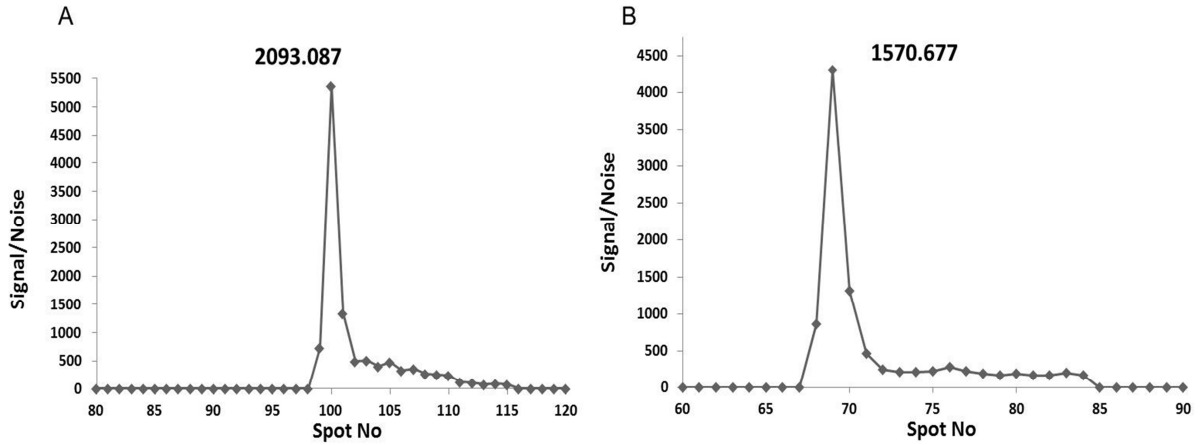

Figure 21 Peak tailing as a result of crystall formation on the needle tip

6.1.3 TOF/TOF method optimization

Modern MALDI-TOF/TOF instruments are capable of acquiring MS/MS data with significantly improved automation. Owing to the large number of spectra, which are generated in shotgun proteomics, manual interpretation of each spectrum is not possible. Therefore a high quality of the mass spectra is required to correctly match peptide sequences. However, for optimum spectral quality, operational parameters and instrumental settings must still be determined. For fragmentation purposes, the 4800 Proteomics Analyzer MALDI-TOF/TOF is equipped with two default MS/MS acquisition methods, the 1 kV MS/MS and the 2 kV MS/MS method. Both default methods have the same settings but differ in their collision energy. To determine the best settings for high throughput analysis, these methods were optimised with standard peptides. MS/MS spectra were acquired using air as a collision gas with either a 1 kV or 2 kV collision energy settings and fixed number of laser shots (2500).

The high collision energy used for the 2 kV MS/MS method resulted in the generation of high-intensity internal fragment and immonium ions in the lower mass region of the spectrum but low intensity fragment ions in the high mass region as shown in figure 22 A. Unlike the 2 kV method, the 1 kV generates a smoother spectrum, with y and b fragment ions ranging from low to the high mass region of the spectra. Thes pectra generated by the 1 kVmethod are easy to interpret (figure 22 B). Consequently, the 1 kV MS/MS method was adopted for subsequent experiments.

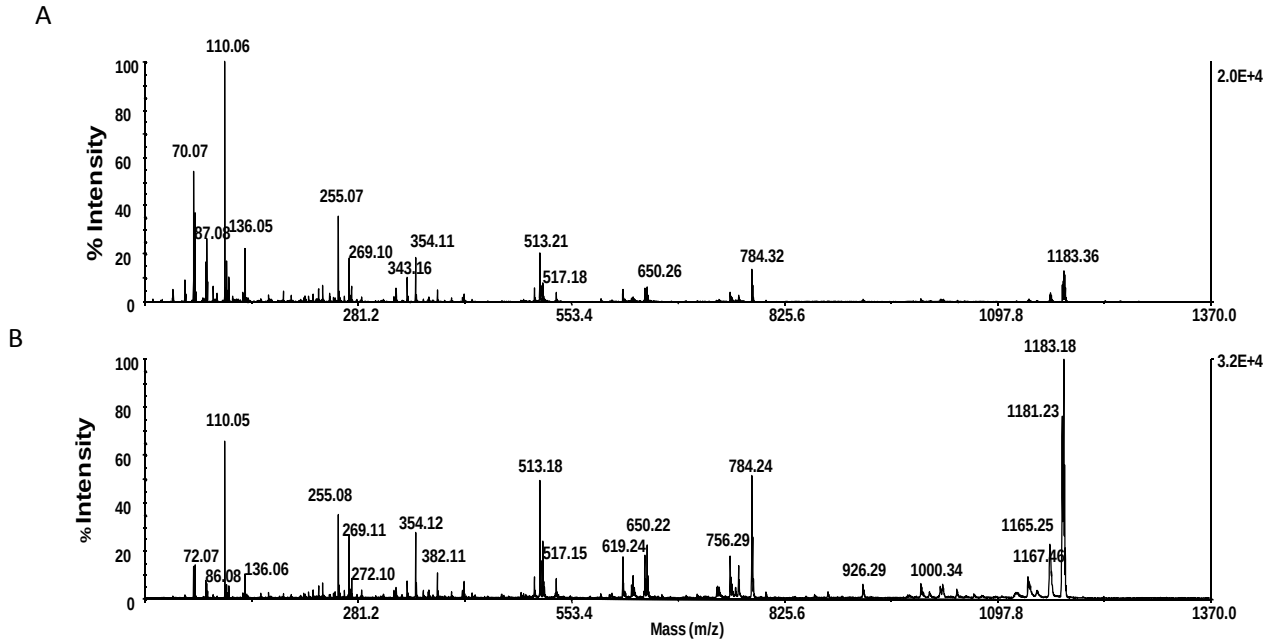

Figure 22 Fragment spectra of Angiotensin II obtain using both fragmentation methods. The spectrum below was obtained with the 1 kV and that above with the 2 kV MS/MS mode.

Matrix/analyte consumption during MS and MS/MS acquisition is often a problem in high-throughput analysis. To minimize the amount of sample consumption during MS/MS, certain threshold conditions were introduced. These conditions were set so that acquisition would stop when a number of fragment ions (15 fragment ions) having a required minimum signal-to-noise ratio (S/N of at least 30) were present in the accumulated MS/MS spectrum, in the region from *m*/*z* 100 to 90 % of the precursor mass. With these conditions a minimum of 1250 (25 sub-spectra accumulated from 50 laser shots each) and a maximum of 2500 shots were allowed for each MS/MS spectrum. The intensities of the fragment ions generated with the fixed-shot method (2500) were compared to those of the

stop-condition method using the applied standard peptides. A total of 10 replicates were performed on one spot. The fixed-shot MS/MS method resulted in decreased intensity because of increasing number of laser shots. In contrast, the intensities of the same fragments with the stop-conditions were relatively constant (figure 23 A and B). On average, fewer laser shots were fired (an average of 1250 per spectrum) and, consequently, less sample was consumed.

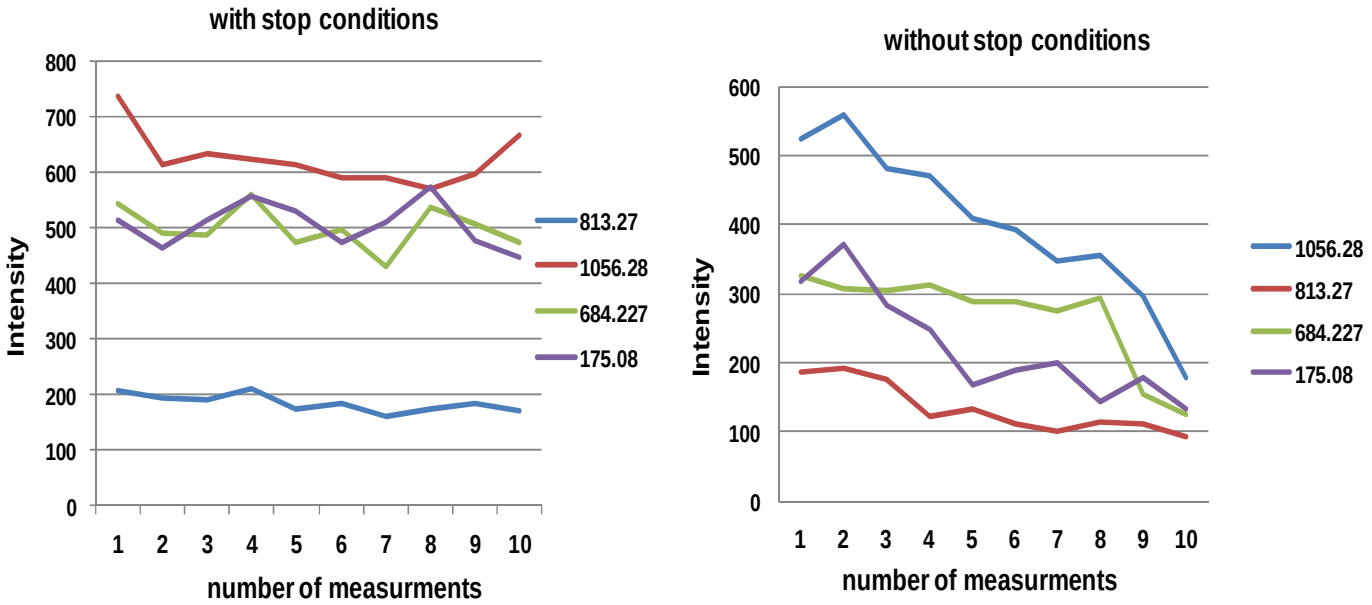

Figure 23 MS/MS method optimisation using the standard peptide with m/z 1570.677. (A) A plot of the ion intensity of 4 fragment generated using the stop-conditions. (B) Shows the signal intensities of the same four fragmentions obtained without stop conditions.

In shotgun proteomics analysis, identification is obtained by determination of a partial sequence from a series of fragment ions whose mass difference matched amino acid residues and given in the form of a peptide score. The peptide score is derived by comparing the precursor mass and number of fragment ions, which are matched to theoretical fragment ions. These scores determine how good a peptide sequence fits. Owing to the large number of spectra generated in shotgun proteomics, manual interpretation of each spectrum is no longer adequate. Therefore, a constant high spectral quality is required to correctly match peptide sequences. The quality of MS/MS spectra depends on the instrument, fragmentation principle and the collision energy used for fragmentation. Since a complete fragment series is not required for peptide identification and the most prominent fragment ions in most peptides are readily formed, defined methods with MS/MS threshold criteria can be introduced. Hence, reducing the rate of sample/analyte consumption would improve signal stability and subsequently peptides scores. On the other hand, extreme higher collision energy yields more internal fragments, which are disadvantageous for peptide identification. As a result, the stop-condition method was adapted for further measurement. With this method, enough spectra to permit peptide identification were acquired and the matrix/sample consumption due to ablation was minimised.

6.1.4 Influence of mobile phase gradients on peptide identification

Given the ion suppression phenomenon associated with MS-based (Cech*et al.*, 2000*;* Kratzer *et al.*, 1998) analyses of complex mixtures, better peptide separation prior to MS and MS/MS analysis is critical for identifying low-abundance proteins. RPLC separation can be performed either in isocratic or gradient system. Unlike gas chromatography, where the separation is exclusively dependent on temperature, RPLC separation depends on the gradient system. Peptides are separated almost exclusively using various forms of gradients; linear, exponential or step gradients, with the linear gradient being the most widely applied. Regardless of which gradient system (see figure 24), the efficiency of separation highly depends on the instrumentation.

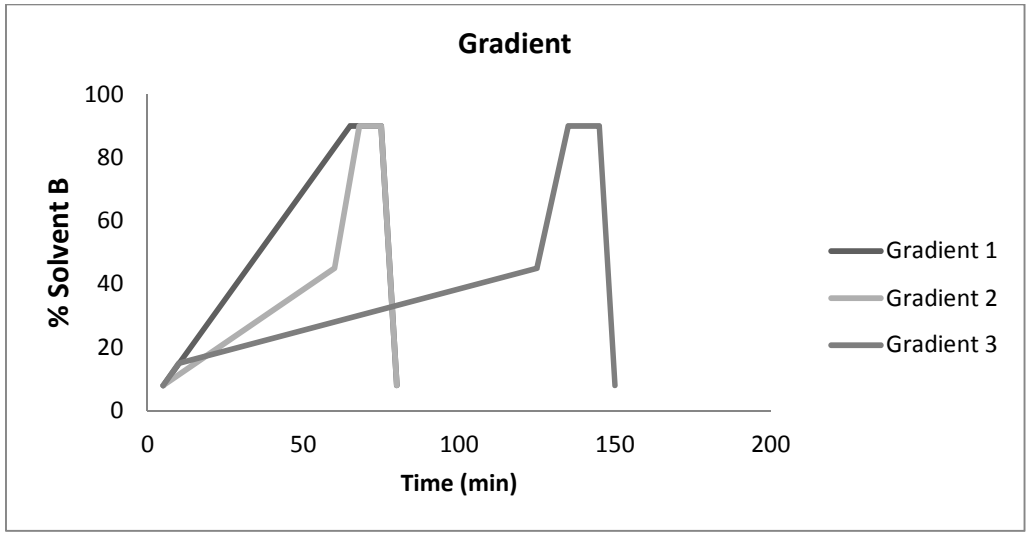

Figure 24 Gradient profiles used for the separation.

To establish a suitable gradient for proteomic analysis, a mixture of 21 proteins was digested with elastase and separated with the gradients shown above. To estimate the ability of gradient systems to separate abundant from less abundant proteins, the concentration of Lipase and Troponin were 10-fold less than the others. While for gradient 1 and 2 the total separation times were kept constant and the gradient slopes were varied, the slope and separation time for gradient 3 was varied. Column and pre-column equilibration was performed before each run and 3 replicates for each gradient were used for the data analysis. The optimised 1 kV MS/MS method was used for MS/MS acquisition and Mascot generic files (MGF) for each run were generated and searched against a standard database containing 24 proteins. For data analysis only peptides with ion score above the identity threshold ($p < 0.05$) and present in at least 2 runs were taken into consideration. The number of proteins and unique peptides identified were used to determine the efficiency of separation.

A total number of 350, 499 and 734 unique peptides were identified for gradient 1, 2 and 3 respectively (see figure 25). For the identified proteins, a gain of at least 10 % and in some cases up to 40 % additional peptides was observed from gradient 1 to 3. Also, the individual peptide score, protein score and the sequence coverage were in most cases higher for gradient 3 than for 1 and 2. The change in gradient profile leads to a 2-fold increase in peptide identification. Whilst 35.6 % of the identified peptides appeared on all gradients, 1.5, 3.4 and 59.5 % were found only in gradient 1, 2 and 3 respectively (see figure 26).

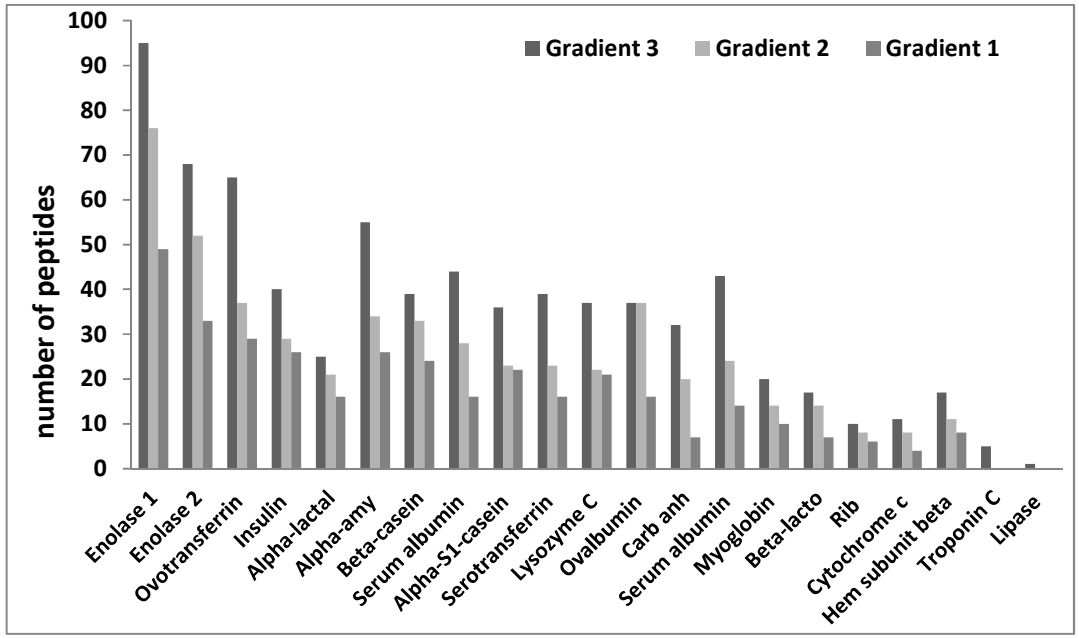

Figure 25 A comparisons of peptides identified by gradients 1, 2 and 3.

With the exception of lipase and Troponin all other proteins were identified with gradient 1 and 2. This indicates that with a faster increase in acetonitrile (ACN) concentration, peptides are not properly separated from each other, thereby inducing a greater possibility of loosing less abundant proteins. The difference in identified proteins is a result of the peptide elution profile. The retention times of peptides are relatively close to each other. The vast majority of peptides elute between 20-40 % ACN. Due to the fact that proteomic samples are complex and the complexity can be exponentially magnified if the whole sample is digested without any pre-fractionations, a small change in acetonitrile concentration in this region (20-40 %) result in the elution of many peptides. After elution, interaction between peptide and the hydrophobic groups of the stationary phase is negligible. No matter how efficient a mass spectrometer is (resolution, sensitivity and duty cycle); if the eluted peptides are poorly resolved, only peptides from the most abundant proteins or easily ionised peptides would be detected. Reducing the number of eluting peptides by introducing a shallower gradient between 20 and 40 % ACN would therefore reduce the amount of co-eluting peptides, thus reducing the degree of ion suppression. Thereby, increasing the number of identified peptides, which in turn improves sequence coverage, identification and characterisation of proteins.

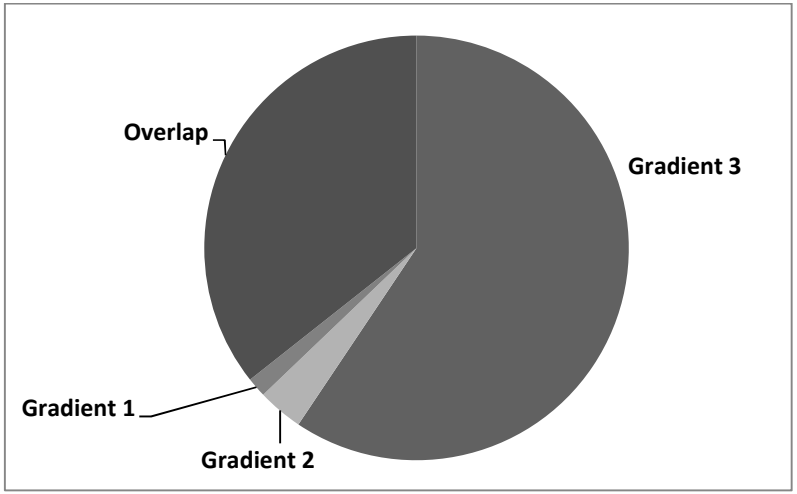

Figure 26 Peptide overlap between the 3 gradients.

This result suggests that for complex samples, shallower gradient separations are generally better than fast gradient separation, since they turn to prolong retention time, thus providing better resolution. For a less complex sample or a fast sample screening, a fast gradient would be enough, but when dealing with complex samples, shallower gradients are necessary. But it should be noted that at one point, no matter how shallow the gradient is, little or no additional information would be acquired. In such cases, more than one dimension of separation techniques would be required for sample analysis. The shallow gradient was adopted for subsequent LC experiments of complex samples and faster gradient for screening and less complex samples.

6.1.5 Improved mass accuracy

The confidence of peptide identification can be enhanced by accurately measuring the precursor mass. The importance of a constant high mass accuracy in characterising peptides has been demonstrated in numerous studies (**Haas *et al.*, 2006**). Though TOF/TOF instruments provide high mass accuracy, they are limited in the fact that the mass accuracy varies over the target. Therefore, poorly calibrated data might lead to loss of potential positive identification and increase the number of false positive identifications. To investigate mass accuracy, 10 standard peptides were digested separated and analysed. In the first triplicate experiment, only the default calibration was used for MS and MS/MS acquisition.

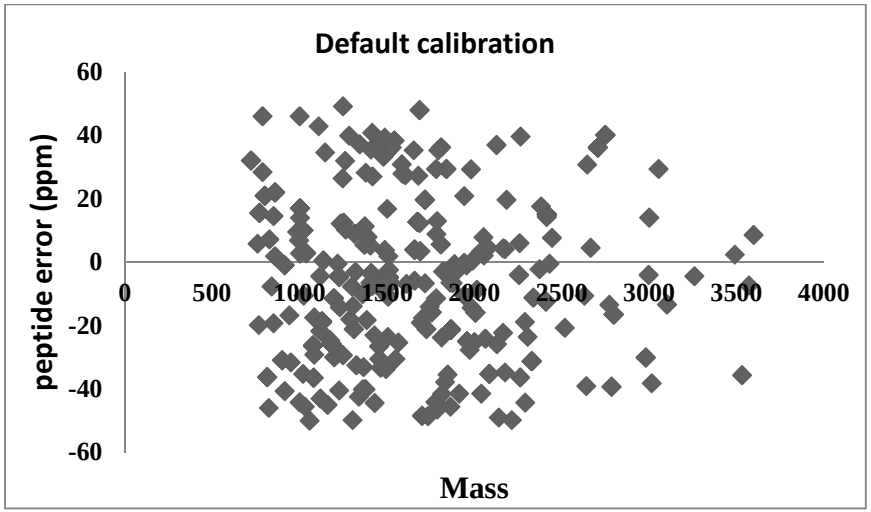

Figure 27 Peptide error distribution (±50 ppm) observed for the identified peptides during LC-MALDI experiments using the default calibration.

While the mass accuracy of the MS/MS mode is relatively constant, the precursor mass accuracy all over the target is not constant especially using different spot sets. The average mass accuracy observed for the identified peptides during LC-MALDI experiments was ±50 ppm (see figure 27). This mass accuracy might be sufficient for a tryptic digest of standard protein. However, for less specific enzymes, since in most cases the enzyme specificity is left out during database search, a higher mass accuracy is required. Peptides of known masses are commonly used to improve the mass accuracy. Glu-fibrinogen peptide, 1570.677 Da is commonly used in MALDI analysis as internal standard because of its m/z ratio, which is in the middle of the mass range (700-4000 Da). This peptide is mixed in the matrix so that each spot contains an equal amount of this peptide. It is either used directly as an internal standard during measurement or in combination with the default calibration. Therefore, to achieve a high mass accuracy Glu-fibrinogen peptide (Glu-fib) was added to the matrix. MS acquisition was acquired using the default calibration and prior to MS/MS acquisition, the precursor masses are recalibrated

internally to the Glu-fib. Though more peptides with mass accuracy between ±15 were identified, only peptides identified in both triplicate runs were used for a detailed comparison. Figure 28 shows a plot of the deviation from the calculated mass of the unique peptides identified. However, the concentration of the internal standard in each spot is very critical. Higher concentration would lead to a saturation of the detector, thus leading to a negative influence of the calibration. On the other hand, at lower concentrations, this peptide might not be detected in some spots due to the presence of highly concentrated peptides from the proteins.

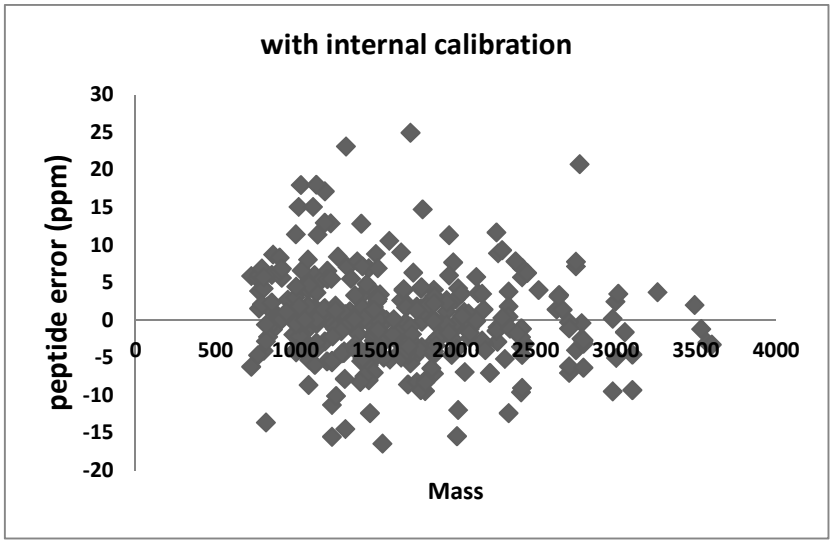

Figure 28 Peptide error tolerance after internal calibration using Glu-fibronogen peptide.

6.1.6 Effect of temperature on LC Separation

Unlike Gas Chromatography, temperature effect is not that significant in reversed phase chromatography. However, operating RP-HPLC at higher than ambient temperatures have been shown to improve peak shapes and enable faster run times (Vanhoenacker & Sandra, 2007). Pre-heated columns and mobile phases results in lower solvent viscosity, which in turn influences the flow profile and the diffusion characteristics (Vanhoenacker & Sandra, 2007). Different heating systems have been introduced and some have proven to be better than others. Independent of the heating system, it is always necessary to determine the temperature at which the HPLC system works best. To determine the optimal temperature for our proteomics workflow, 9 standard proteins were digested with Trypsin and Elastase. The samples were loaded onto a 150 µm x 35 mm C_{18} pre-column. The separation was performed on a 75 µm x 150mm C_{18} column using an 80 min gradient. The column temperature was set between room temperature and 60 °C (20, 30, 40 and 60 °C). In order to obtain the required temperature, the column was left in the column heater for 60 min before the next separation. For each temperature, three replicates were used for the data analysis and a total of 100 fmol for each of the standard protein was injected on the pre-column. After separation, the eluent was pre-mixed with a matrix and spotted every 20 sec on a metal target.

6.1.7 Effect of temperature on column back pressure

The primary gain of elevated column temperature on LC separation was the resultant decrease in backpressure as shown in figure 29. With the increase in temperature from 20 to 30 °C, a pressure drop of approximately 10 % is observed. From 35 through 60 °C the change in pressure is no longer significant, since it becomes more stable. At this point any increase in temperature would probably result in little or no change in backpressure. Temperatures above 60 °C would result in a drop of backpressure due to possible column bleeding. Therefore, for our proteomics setup temperatures between 35-40 °C would be optimal, since there is no significant change in backpressure.The decrease in backpressure comes as a result of a decrease in solvent viscosity. Reduced backpressure facilitates the use of longer columns or columns with smaller stationary-phase resin to improve resolution at higher flow rates without hampering efficiency.

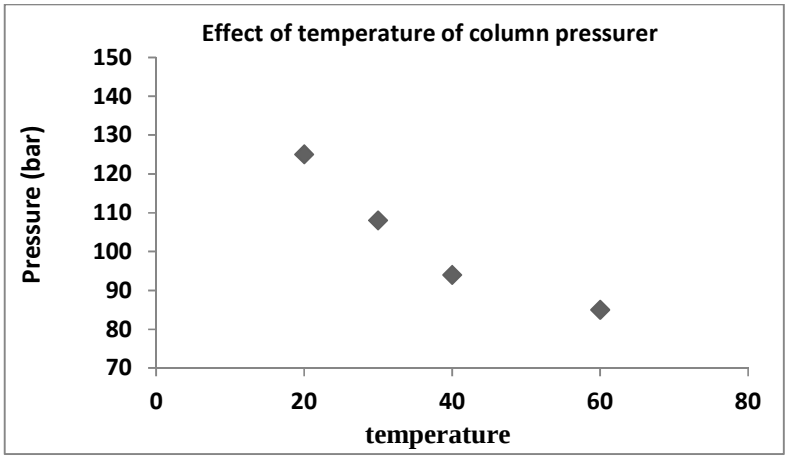

Figure 29 Shows the column back pressure at different temperatures at a flow rate of 400nl/min.

6.1.8 Effect of temperature on proteins and peptides identification

To determine the optimal temperature for our proteomics workflow, 9 standard proteins were digested with Trypsin and Elastase. The samples were loaded onto a 150 µm x 35 mm C_{18} pre-column. The separation was performed on a 75 µm x 150 mm C_{18} column using an 80 min gradient. The column temperature was set between ambient temperature and 60 °C (20, 30, 40 and 60 °C). In order to obtain the required temperature, the column was left in the column heater for 60 min before the next separation. For each temperature, three replicates were used for data analysis and a total of 100 fmol for each of the standard proteins was injected onto the pre-column. After separation, the eluent was pre-mixed with a matrix and spotted every 20 sec onto a metal target. A Mascot search against the DB-standard database resulted in the identification of all the standard proteins. Furthermore, an additional protein, Ubiquitin, which was not present in the protein list, was identified at elevated temperatures (30, 40 and 60 °C). This protein resulted from a posttranslational modification of any of the standard proteins. The effect of temperature on identification can be clearly observed at the peptide level. All

proteins were identified with different number of unique peptides. Figure 30 shows the total number of unique peptides that were identified. Between 20 and 40 °C a 33.6 % and 21 % increase in unique peptides was observed for elastase and trypsin respectively. Only a minimal increase was observed between 20 and 30 °C. At 60 °C, the number of identified peptides dropped but was still higher than for 20 °C. According to the manufacturer's recommendations, the temperature of the column should not exceed 60 °C due to irrevocable damage of the column. This could be a possible reason for the drop in unique peptides. For both proteolytic enzymes, the peptide separation at a column temperature of 40 °C leads to the identification of highest number of peptides. A complete list of the proteins and peptides identified is seen in the supplementary table 1.

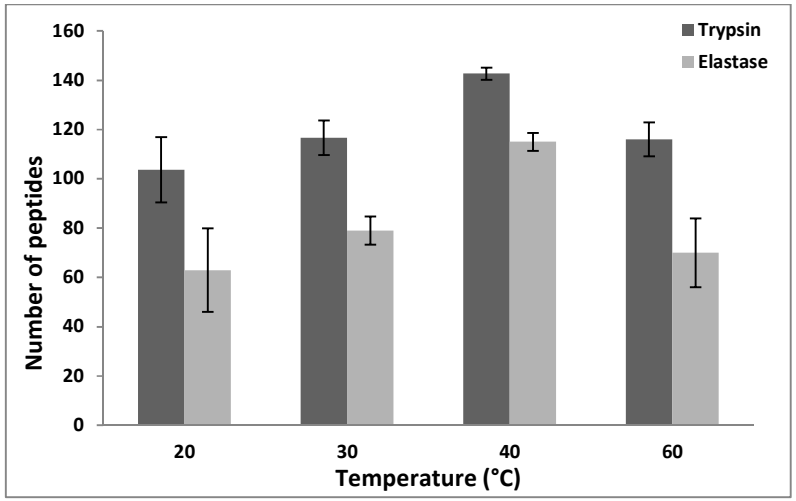

Figure 30 Total number of peptides that were identified using different column temperatures for tryptic and elastase digest of 13 standard proteins.

Furthermore, peptide overlap was determined. In general 34.8 % of peptides were found in all temperatures, 4.6 %, 15.13 %, 2.6 % and 7.2 % were found only in 60 °C, 40 °C 30 °C and 20 °C (see figure 31 A and B). Moreover, the intensities, signal-to-noise and the subsequent Mascot score for most of the shared peptides were higher at elevated temperature.

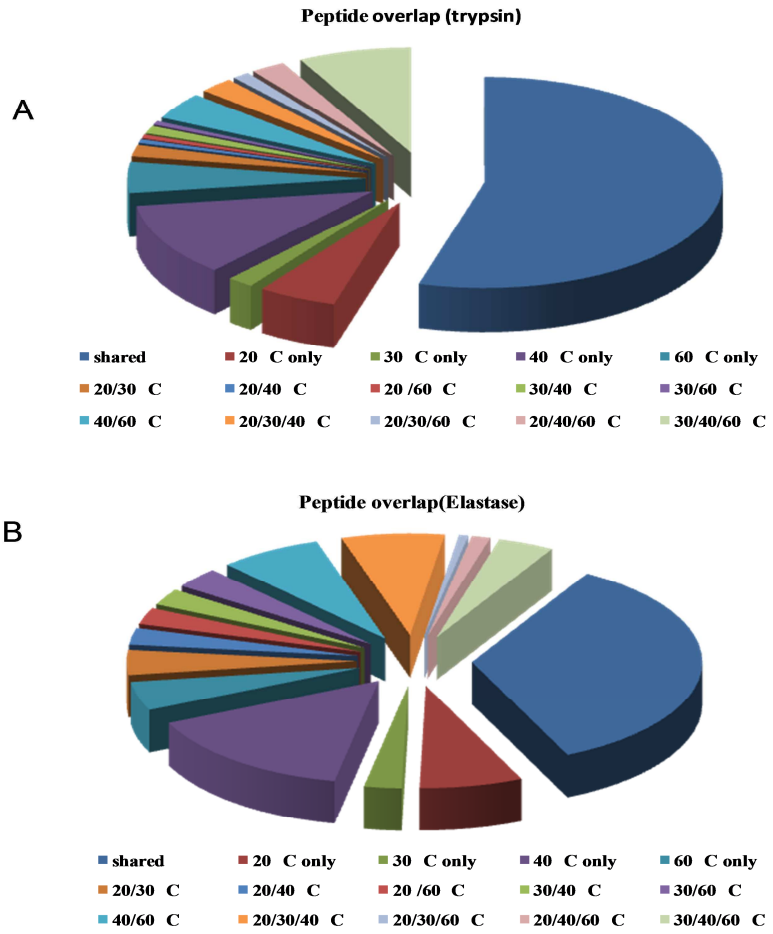

Figure 31 A summary of the peptide overlap for both tryptic (A) and elastase (B) digestion.

6.1.9 Effect of temperature on repeatability

Run-to-run repeatability is essential for shotgun proteomics studies especially when monitoring dynamic changes in living organism. In these studies, identified proteins, which are regulated or not present at all in one state or the other can be consider as potential biomarkers candidates. Therefore, repeatability of the analyses is a critical determination of the ability to distinguish true differences between sample types. Table 6 shows the run-to-run repeatability for a tryptic digest of 13 standard proteins. Higher temperature did not only improve run-to-run repeatability, but also more peptides were identified. At higher temperature, reproducible triplicate run were easily obtained, at least 5 runs were necessary to obtain reproducible triplicate runs at room temperature. Reproducibility is impaired by long inadequate equilibration time between runs, fluctuating column pressure/flow rate, fluctuating temperature and irreproducible mixing of both solvent A and B. Flow sensors have been built before and after the mixing chamber to monitor the flow rate and LC separations are now being performed at high temperature. Higher temperatures result in decrease solvent viscosity, which in turn leads to shorter column re-equilibration time, thus, improvement in run-to-run reproducibility.

Table 6 The effects of column temperature on run-to-run reproducibility and the number of peptides identified from a tryptic digestion of 13 standard proteins Proteins identified with more then 1 peptides is represented with protein > 1 and proteins identified with only 1 peptide with “one hit wonders” (OHW). All gradient runs were performed on a nLC column from Dionex (75 μm × 150 mm column).

	40 °C_1	**40 °C_2**	**40 °C_3**	**Average**
Proteins	13	13	13	13
Peptides	140	143	145	14 2.67

Proteins >1	12	12	12	12
OHW	1	1	1	0
false positive rate	2.99	3.19	3.32	3.20
	30 °C_ 1	**30 °C_2**	**30 °C_3**	**Average**
Proteins	13	13	13	13
Peptides	110	124	116	116.67
Proteins >1	9	9	9	9
OHW	3	3	3	3
false positive rate	2.70	3.67	2.23	20.87
	RT °C_ 1	**RT °C_2**	**RT °C_3**	**Average**

Proteins	11	11	11	11
Peptides	89	107	115	103.67
Proteins >1	8	7	8	7.60
OHW	3	4	3	3.30
false positive rate	2.97	3.17	3.55	3.22

From the number of peptides identified in triplicate runs at elevated temperatures (see table 3), it can be assumed that a higher repeatability is most likely to be achieved at elevated temperatures. However, assessing the run-to-run overlap amongst the same temperature reveals that only about 64 % of the identified peptides appeared in all runs (see figure 32).

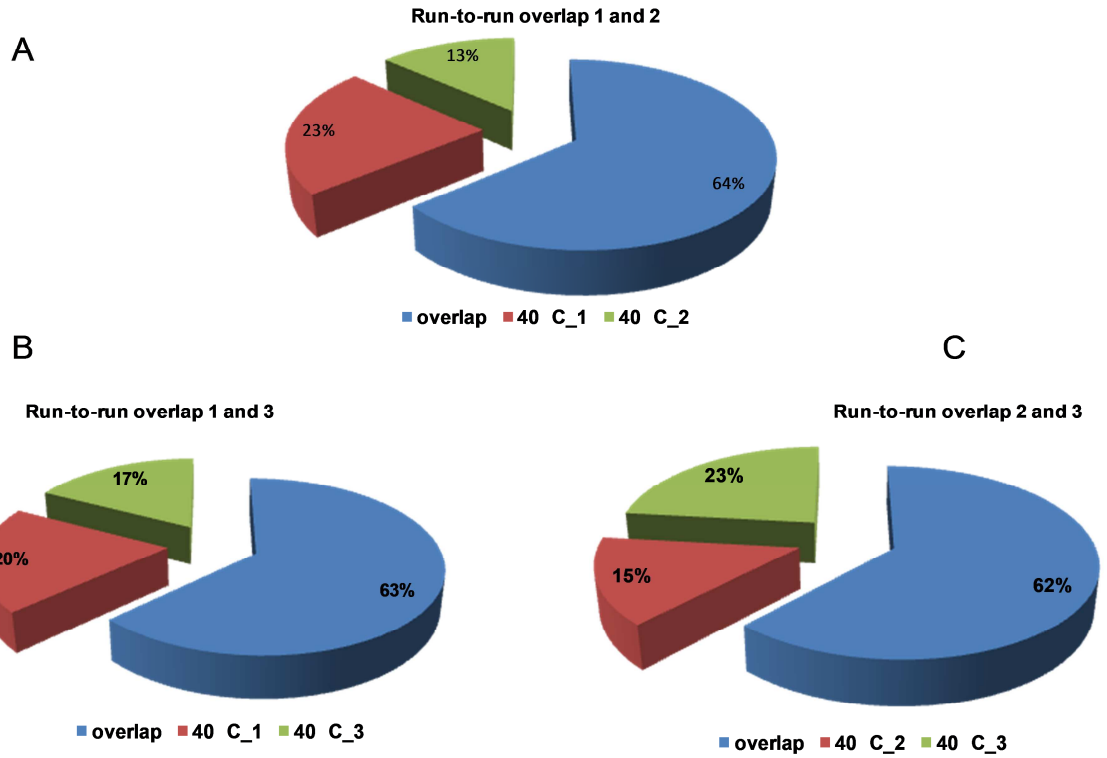

Figure 32 Run-to-run overlap.

This shows that no matter how stable a system is running, there are still alterations, which influence peptide signal-to-noise ratios and their elution times. This might influence their selection for fragmentation and the fragmentation itself. Poor fragmentation could cause misidentification or poor scoring and therefore these peptides can be omited during protein identification. For example, this serum albumin peptides RRHPEYAV or alpha S1 casein peptide YFYPELFRQ were identified by all temperature experiments but with

different ion scores (see table 7). In general, signature peptides are easily detected and their measurements are readily repeatable. Therefore, to improve sequence coverage or the identification of low abundant peptides it is necessary to perform more than 3 replicate runs and every run should be included in the data interpretation process.

Table 7 Peptide ion score for two peptides identified in triplicate run for all experiment performed. Though the peptide identification was reproducible, the scores were different.

	Temperature	peptide score_R1	peptide score_R2	peptide score_R3	average
RRHPEYAV	60	26	21.85	27.5	25.12
	40	24	32.65	38.6	31.75
	30	35.85	28.86	25.3	30.00
	RT	16.99	33.26	25.4	25.22
YFYPELFRQ	60	31.29	29.49	27.3	29.36
	40	35.22	30.3	40.2	35.24
	30	26.46	22.42	25.3	24.73
	RT	30.2	32.44	29.7	30.78

6.1.10 Conclusion

Compared to gas chromatography, the change in column temperature has a much lower effect on the degree of retention and resolution, since the heat transfer of solute between mobile and stationary phase is much smaller (Vanhoenacker & Sandra, 2008). Nevertheless, an increase in temperature led to a decrease in backpressure as a result of decrease solvent viscosity. This permits the use of longer columns and higher flow rates without compromising efficiency. Furthermore, the effect of a temperature increase was more pronounced in the number of peptides identified. A total increase of unique peptides of 33.6 % for elastase and 21 % for trypsin was achieved. Speers et al. also reported an increase in the number of identified peptides at elevated temperature (Spears *et al.*, 2008). In addition, elevated temperatures resulted in shorter re-equilibration time thus, an improvement in the run-to-run reproducibility. However, despite the run-to-run repeatability only about 60 % of the peptides observed in one LC-MALDI TOF/TOF analysis would be identified in a second run. In general, while low abundant peptides or poorly detectable peptides are less readily repeatedly identified, signature peptides are easily detected and their analyses are readily repeatable. Since only peptides that appeared in more than one run are considered as true positive, care should be taken during data interpretation of replicate runs. This is due to the fact that, some of these peptides (with high ion scores) that appear in only one run can be more informative than those which appear in 2 runs but with very low ion score. Due to the advantages of elevated temperature on peptides and protein identification, elevated temperatures were used for our proteomics setup, except otherwise mentioned.

Despite their advantages, elevated temperatures also carries some disadvantages such as: (i) the decomposition of the sample and solvent evaporation at elevated temperatures, (ii) a temperature gradient along the column might lead to tailing and (iii) the solubility of silica greatly increases at elevated temperature (Meyer, 1999). To avoid a temperature gradient, which might result in peak broadening and the loss of hydrophobic peptides most LC separations are carried out at elevated temperatures.

6.2 Proteins from yeast lipid particles and human fat cells

6.2.1 Optimisation of the sample preparation procedure

After cell culture lipid particles were isolated as described above and proteins were precipitated with trichloroacetic acid. The resultant pellets were dissolved in ammonium bicarbonate (AMBIC), digested with trypsin and 5 µg of each sample was loaded on the pre-column. The samples were separatred with a 150 min linear gradient at 40 °C. The Mascot search was performed using parameters decribed above against a SGD derived database. This resulted in the identification of 111 peptides from 35 proteins (see table 8). The most abundant protein, ERG6 with many accessible proteolytic sites was identified with only 5 peptides and sequence coverage of 31 %. In general, an average of 3 peptides was matched to each protein. The reasons for these poor results can be attributed to the following: Prior to sample loading on the nLC, the digested samples were concentrated. This resulted in a highly viscous solution probably due to the high content of lipids and phospholipids. This made sample pick up together with loading onto the pre-column difficult. Furthermore, high backpressure, poor resolution and subsequent pre-column clogging were observed during separation on the analytical column.

To overcome these limitations, lipids and phospholipids were extracted with diethyl ether prior to protein precipitation. After precipitation, proteins were dissolved in AMBIC, digested, concentrated and loaded on the pre-column. Sample picking, loading onto the pre-column and separation resulted in a decrease backpressure, thus phospholipids and peptides were successfully extracted from the proteins. A database search resulted in the identification of 101 proteins, an increase of more than 200 %. In addition, the number of identified peptides increased by approximately 300 % (table 8). For example, 31 peptides were matched to the ERG6, which was identified with only 5 peptides before the optimization. A complete list of the identified proteins and number of peptides matched to each protein is seen in the supplementary table 2 (table S2 A B and C).

Table 8 Number of peptides and proteins identified from lipid particles before and after method optimisation. Proteins identified with more then 1 protein are shown with protein>1 peptide and those with just 1 peptide as “one hit wonders” (OHW).

	Before	After
Proteins	35	101
Peptides	111	437
Proteins >1 Peptide	17	50
OHW	18	51

6.2.2 Effect of carbon source on lipoprotein composition

In order to assess the changes that occur due to different growth media proteins extracted from YPD and YPO grown cells were delipidated, precipitated, digested, separated and analysed. A database search of the MALDI MS/MS data against the SGD database resulted in the identification of 124 proteins from lipid particles grown in YPD medium, and 117 proteins isolated from yeast cell grown on oleic acid (YPO) (see Table 9). However, more than 50 % of the proteins were identified with just one peptide.A summary of the identified proteins and their functions is seen in figure 33.

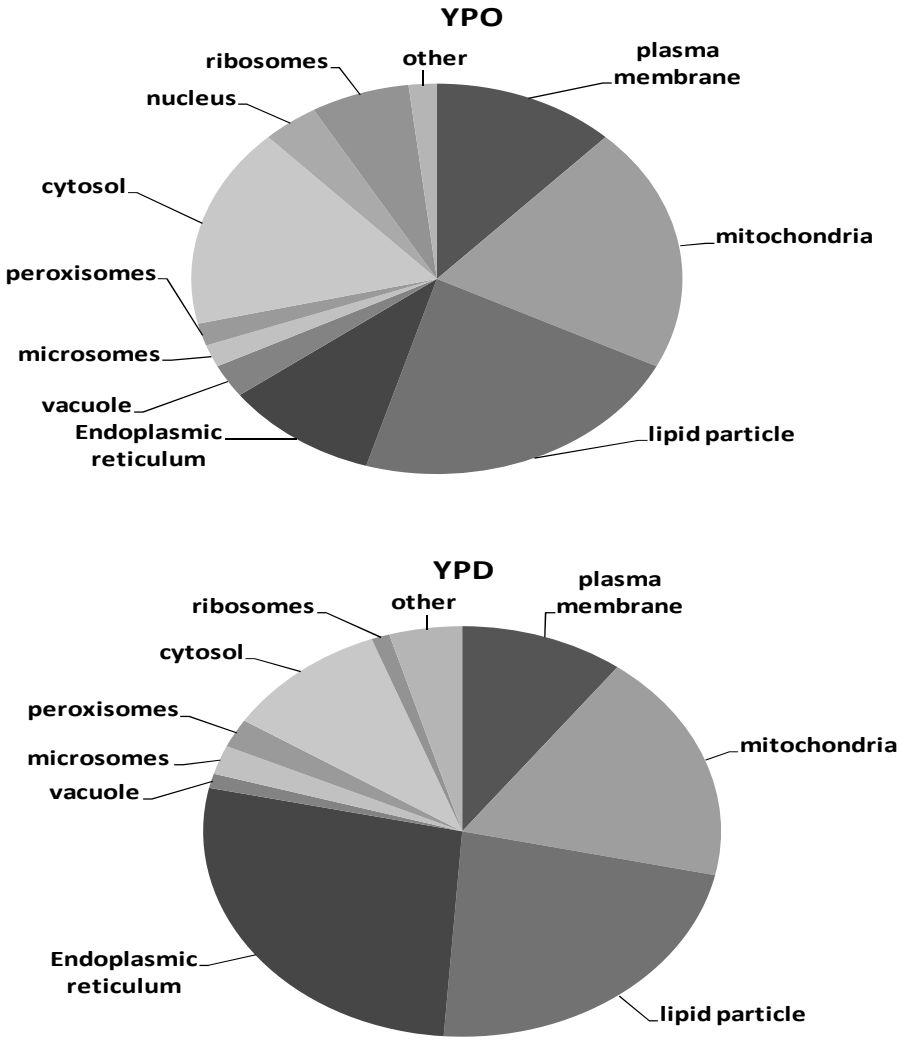

Figure 33 Assignment to subcellular fractions of LP proteins from wild type cells grown on glucose (YPD) or oleate (YPO).

Table 9 Summary of the number of peptides and proteins identified from yeast cell lipid particles grown on glucose (YPD) and oleic acid (YPO).

Proteins identified with more then 1 protein are shown with protein >1 peptide and those with just 1 peptide as “one hit wonders” (OHW).

	YPO	YPD
No. proteins	124	117
No. Peptides	384	354
Proteins > 1 Peptide	51	48
OHW	73	69

Though the peptides used for the verification exhibited ion scores which exceeded the identity threshold and appeared in more than one experiment, only proteins, ascertained by more than one peptide, were used for a detailed comparison. Thus, the number of identified proteins for both samples, YPD and YPO dropped to 48 and 53 proteins respectively. Of these proteins, 34.67 % appeared in both YPO and YPD samples, 36.00 % only in YPO and 29.33 % solely in YPD (see figure 34).

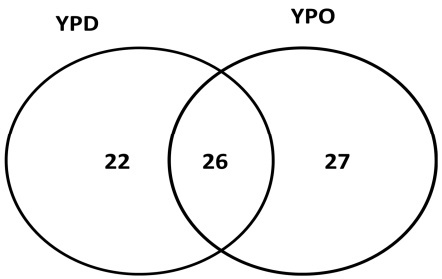

Figure 34 Protein overlap from yeast strains grown in glucose (YPD) and those grown on oleoic acid (YPO).

Among these proteins, 26 have been described to be located in the lipid particles. Nevertheless, it should be mentioned that hardly any protein is known to be explicitly localized to a particular lipid particle. For example Erg6p, one of the abundant lipid particle proteins (Athenstaedt *et al.*, 1999), is localised in the endoplasmic reticulum (ER) (McCammon *et al.*, 1984) and mitochondria (Sickmann *et al.*, 2003). Another example is the long chain fatty acyl-CoA synthetase Faa1p, which is localised in the lipid particles (Athenstaedt *et al.*, 1999), mitochondria (Sickmann *et al.*, 2003) and the plasma membrane (Delom *et al.*, 2006). This characteristic property of lipid particle proteins complicates a direct assignment of putative new proteins to this cellular compartment. Therefore, data are to be interpreted with caution. However, six proteins, Cpr5p, Gtt1p, Osh4p, Ubx2p, Vps66p, and Ypt7p detected in both samples have never before been attributed to *Saccharomyces cerevisiae* lipid particles. Two of these proteins, Osh4p and Vps66, were previously found to be localised to the lipid particles of *Yarrowia lipolytica* (Athenstaedt *et al.*, 2006). Preliminary results employing GFP-fluorescence microscopy showed a formation of distinct punctuate structures in *Saccharomyces cerevisiae*. Gtt1p, Osh4p, Vps66p and Ypt7p formed distinct punctuate structures when grown to stationary phase on glucose containing media. Dye merging and cell fractionation experiments need to be performed in order to verify whether or not these structures really overlap with lipid particle.Gene ontology annotations of the novel putative lipid-particle proteins revealed that they covered a broad spectrum of biological processes under both growth conditions (see figure 35).

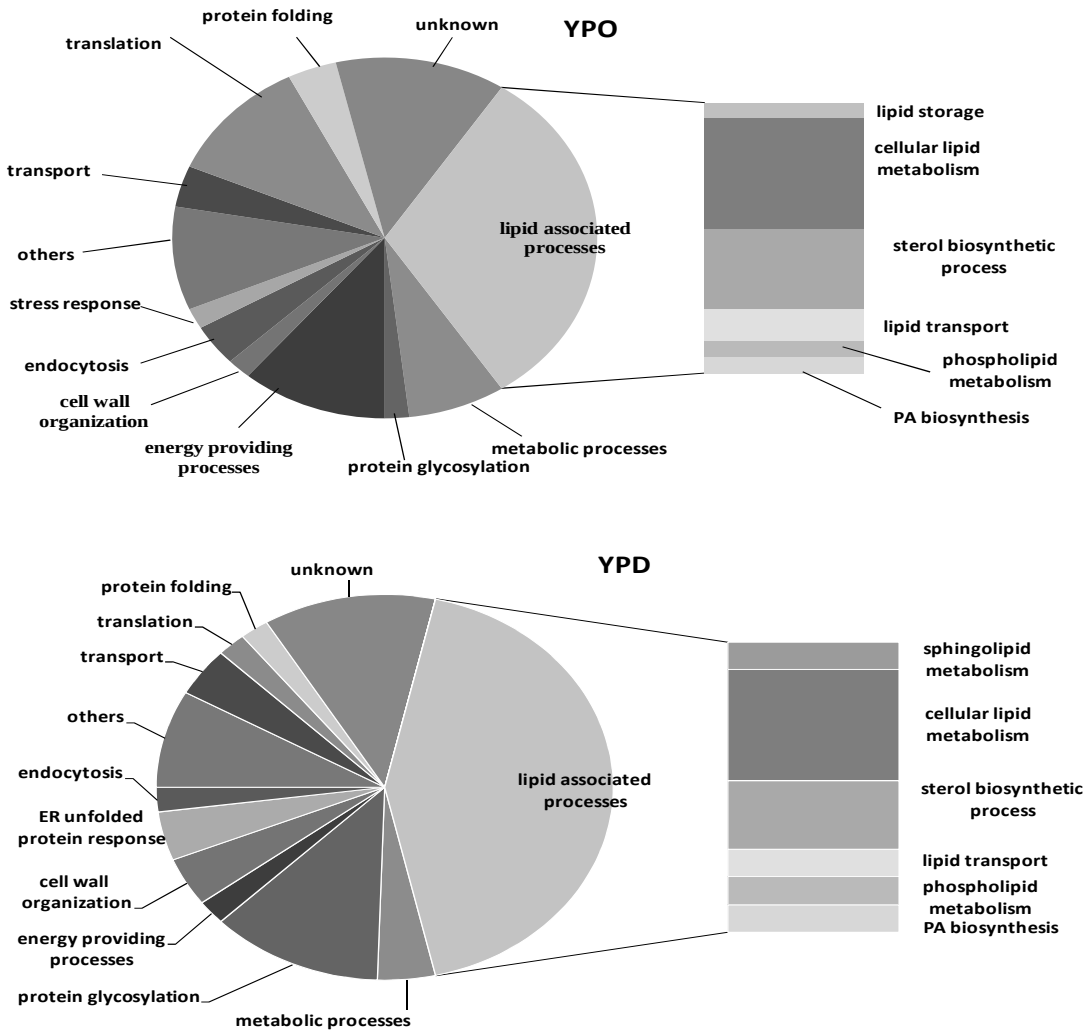

Figure 35 Assignment of newly found proteins of lipid particle from wild type cells grown on glucose (YPD) or oleate (YPO) to biological processes.

6.2.3 Human fat cells

This project was done in collaboration with Anja Rosenow, Freek Bouwman, Jean-Paul Noben, Martin Wabitsch, Edwin C.M. Mariman and Johan Renes.

6.2.3.1 Protein identification

In order to gain an insight into the human-specific secretion profiles during preadipocyte differentiation, proteins from cultured preadipocytes and adipocytes human cell strain derived from a Simpson Golabi Behmel syndrome (SGBS) patient were purified, delipidated, buffer changed, digested with trypsin and 10 µg of each sample was loaded on the pre-column. Separation was performed with the same conditions as described above and analysed. The search against the Swissprot database containing 20402 sequences after taxonomy filter (homo Sapiens), as of 2009-12-16, resulted in the identification of 92 and 97 proteins from preadipocytes and adipocytes, respectively (supplemental table S 2D and E). . Of these proteins 42 were detected with preadipocytes and 68 were detected with adipocytes, of which 34 proteins were detected with both cell types (figure 37). The motifs of the identified proteins were further confirmed using SignalP and/or SecretomeP, commonly used methods for prediction of classically secreted proteins (Bendtsen *et al*., 2004) (see figure 36). Based on these software packages, 76 of the identified proteins were predicted to be secreted proteinsProteins with a functional role in the complement system such as clusterin, complement C1r subcomponent, complement C1s subcomponent, complement C3, complement factor D, and plasma protease C1 inhibitor were found mostly in the adipocytes. The same results were reported before (Wang *et al*., 2004; Zvonic*et al*., 2007; Chiellini *et al*., 2008). In addition, adipocyte biomarkers such as PAI-1, PEDF, MIF, interleukin-25, adipsin, and RBP-4 as well as other important adipocyte secretion proteins, such as metalloproteinase inhibitors (TIMP's), 72 kDa type IV collagenase (MMP-2) and IBP's were identified. Furthermore, extracellular matrix proteins, such as collagens, basement membrane-specific heparan sulfate proteoglycan core protein (HSPG), SPARC, transforming growth factor-beta-induced protein ig-h3 (beta-igH3), and thrombospondin 1 and 2 were ascertained. The identification of proteins relevant for adipocytes such as

fatty acid binding proteins, fatty acid synthase, perilipin 4, and enoyl-CoA hyratase, which are correlated with lipid metabolism, indicates that proteins with a biological function in the complement systems are mainly secreted by mature adipocytes.

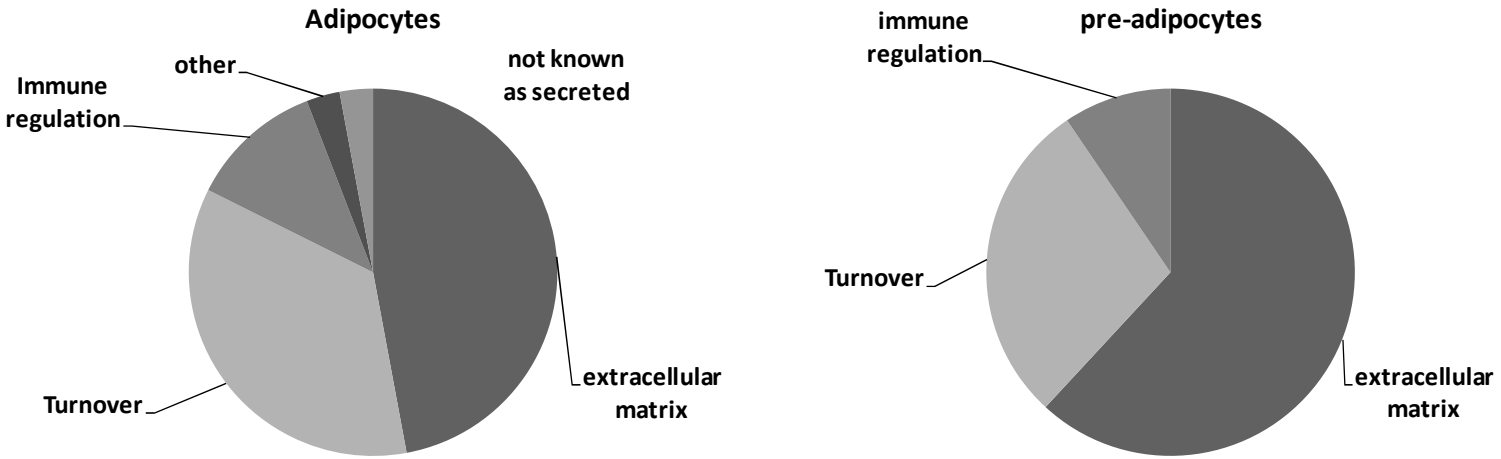

Figure 36 Identified secreted proteins classified accoding to their functions.

Table 10 Summary of the number of peptides and proteins identified from pre-adypocytes/adipocytes tissue. Proteins identified with more then 1 peptide are shownwith protein > 1 and those identified with only 1 peptide with “one hit wonders” (OHW).

	Adipocytes	Pre-adipocytes
No. proteins	97	92
No. Peptides	290	374
Proteins > 1 Peptide	44	47
OHW	53	45

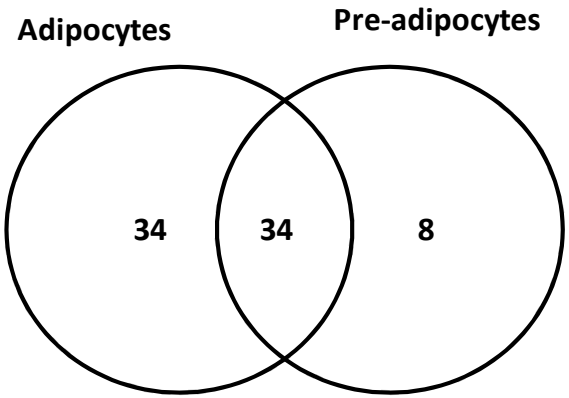

Figure 37 Overlap of identified secreted proteins from adipocye and pre-adipocytes.

6.2.3.2 Validation of the identified secretome

The identified secretome was compared to adipocyte proteins from the currently known secretome of human visceral (Alvarez-Llamas *et al*., 2007; Roelofsen *et al*., 2009), subcutaneous (Sell *et al*., 2006; Zvonic *et al*., 2007; Chiellini *et al*., 2008) and rodent adipose tissue (Kratchmarova *et al*., 2002; Wang *et al*.,2004; Chen. *et al*., 2005; Aoki *et al*., 2007; Lim. *et al*., 2008; Molina *et al*., 2009) (see supplemental table 2G). Proteins, such as annexin 5 and cyclophilin A, which are primarily known as cellular proteins, were also identified as secreted proteins. However, in published reports the function of annexin 5 is usually referred to as unknown. Wang et al. (Wang *et al*., 2009) documented that annexins are related to stress response, whereas others reports assume that annexin 5 may play a role in inhibition of blood coagulation by competing for phosphatidylserines (Urso *et al*., 2001; Rand *et al*., 2008). Phosphatidylserines are part of cell membranes and during stress situations in the cells, phosphatidylserine switches to the outside of the membrane (Fink *et al*., 2005). Since annexin 5 is predicted to interact with phosphatidylserine, it may pass the membrane and be detected as a secreted protein. Yet, there is evidence that annexin 5 might be a non-classical secreted protein, since the SecretomeP NN-score is 0.503, which is higher than the threshold value of 0.5. On the other hand, SignalP and SecretomeP analyses of Cyclophilin A were negative, which explains the SwissProt classification as a cytoplasmic protein. However, there are several reports in which cyclophilin A is described as a secreted protein by a vesicular pathway (Sherry *et al*., 1992; Jin *et al*., 2000; Jin *et al*., 2004; Suzuki *et al*., 2006) which complies with our present findings.

Of the 78 secreted proteins, 8(lysyl oxidase homolog 2, alpha-2-HS-glycoprotein, glia-derived nexin, insulin-like growth factor-binding protein (IBP) 3, and IBP-5, sushi repeat-containing protein, immunoglobulin superfamily containing leucine-rich repeat protein, and coiled-coil domain-containing protein 96) have not been reported before. Furthermore, 6 of these 8 proteins have not been reported as adipocyte-secreted before and therefore they can be described as novel (pre)adipocytes-secreted proteins (see table 11). Sushi repeat-containing protein has signal peptide cleavage sites, which allow classification as a secreted protein. Its function is rather unclear, but it seems to be involved in cell adhesion and migration (Tanaka *et al*., 2009). Coiled-coil domain-containing protein 96 and acyl-CoA-binding protein showed no evidence of a signal peptide cleavage site but reached high NN-scores on SecretomeP analysis. The function

of coiled-coil domain-containing protein 96 with a NN-score of 0.622 is still unknown. Acyl-CoA-binding protein may be an intracellular carrier of acyl-CoA esters. It occurs in 3 isoforms, but could not be distinguished. All are ubiquitously expressed. In addition, isoform 2 is highly expressed in liver and adipose tissue and isoform 3 is strongly expressed in adipose tissue and heart. The NN-scores are 0.546 (isoform 1), 0.553 (isoform 2) and 0.758 (isoform 3), indicating secretory properties. Further, Kinseth et al. (Kinseth *et al*. 2007) classified acyl-CoA-binding protein as a non-classical secreted protein. Altogether, the provided data indicate that acyl-CoA-binding protein is a genuine secreted protein. Correlations between this protein, obesity and diabetes suggest that acyl-CoA-binding protein is an adipokine (Zhou *et al*., 2008; Ejaz *et al*., 2009). In contrast, IBP-3 is already known as a secreted protein from human subcutaneous adipocytes and 3T3-L1 cells (Hossner *et al*., 1997; Wabitsch *et al*., 2000). IBP-5 mRNA levels are shown in human subcutaneous adipose tissue but the proteinis only described as an adipocyte-secreted protein from pig adipose tissue (Hausman *et al*., 2006; Baxter *et al*., 2009; Kallio *et al*., 2009). In addition, 13 proteins have been reported for rodents, but not for human adipocytes (see table 11).

Table 11 Proteins that have never been associated to the adipocye tissues (human visceral, human subcutaneous and rodent). A complete list of all identified proteins is seen in the supplementary Table S2

Accession No	protein name	Average no. of peptides (pre-adipocyte)	Average no. of peptides (adipocyte)
P07093	Glia-derived nexin (GDN) (Protease nexin I) (Protease inhibitor 7)	5	1
P17936	Insulin-like growth factor-binding protein 3 (IBP-3)	6	1, 1
P24593	Insulin-like growth factor-binding protein 5 (IBP-5)	not found	2, 2
O14498	Immunoglobulin superfamily containing leucine-rich repeat protein	3	not found
P78539	Sushi repeat-containing protein SRPX	not found	1
Q9Y4K0	Lysyl oxidase homolog 2 (Lysyl oxidase-like protein 2)	10	not found

Table 12 Summary of all secreted proteins, which have been identified only for rodent proteomic adipocyte studies.

Accession No	protein name	Average no. of peptides (preadipocyte)	Average no. of peptides (adipocyte)
P35555	Fibrillin-1 (FBN1)	11	4
P09382	Galectin-1 (Lectin galactoside-binding soluble 1)	not found	4, 3
P28300	Protein-lysine 6-oxidase (Lysyl oxidase)	6	not found
P05452	Tetranectin (TN) (Plasminogen kringle 4-binding protein)	not found	1
Q92626	Peroxidasin homolog (Vascular peroxidase 1)	1	1
Q9H293	Interleukin-25 (IL-25)	1	not found
P14174	Macrophage migration inhibitory factor (MIF)	1	1, 1
P02458	Collagen alpha-1(II) chain	11	2, 1
P02462	Collagen alpha-1(IV) chain	not found	2, 3
P05997	Collagen alpha-2(V) chain	12	4, 3
P29279	Connective tissue growth factor (Hypertrophic chondrocyte-specific protein 24)	not found	1
Q9H293	Interleukin-25 (IL-25)	1	not found

P14174	Macrophage migration inhibitory factor (MIF)	1	1, 1
P25940	Collagen alpha-3(V) chain	not found	3

6.2.4 Summary

During the past decade, published results suggest that lipid particles in yeast cells as well as other cell types do not only serve as a lipid storage compartments, but also play asignificant role in a number of other functions. For example, lipid particles harbour enzymes responsible for lipid metabolism, thus contributing actively to the formation of cellular components (Czabany *et al*., 2008). The majority of proteins identified such as Erg-proteins, lipases or fatty acid activating proteins are involved in lipid metabolism. While most of the proteins involved in this process were identified in YPD, this number decreases in the YPO sample. On the other hand, remarkable increases in proteins involved in the processes of translation were found in lipid particles extracted from yeast cells grown in oleic acid medium. Furthermore, other clusters of proteins appeared to be strongly affected by the change of carbon source. For example, proteins involved in endoplasmic reticulum (ER) unfolded response were only found in lipid particles from cells grown on glucose, and the large portion of proteins involved in protein glycosylation was drastically reduced in cells grown on YPO. In contrast, a large number of proteins involved in translation and energy providing processes were found in lipid particles from cells cultivated on oleic acid. Furthermore, membrane and membrane-associated proteins and proteins with unknown biological functions were identified.

In addition, the identification of certain proteins showed a correlation between plant and yeast cells, which has also been reported before (Athenstaedt *et al*., 2006). Also, proteins from other regions of the cell, such as endoplasmic reticulum, Golgi, Nucleus, cysotolic and peroxisomal proteins were identified. However, it was demonstrated that lipid particles and some other organelles such as the

endoplasmic reticulum share a set of certain proteins. Even though cross-contamination studies demonstrated that the samples were of high purity, it cannot be completly ruled out that some of the identified proteins were contaminants from other organelles. Nevertheless, six novel proteins were identified and were confirmed by localisation studies in the lipid particles.

Adipokines play important roles in homeostasis and metabolism. Adipocyte differentiation leads to a change in adipokine secretion profile which is probably involved in disruption of homeostasis. The study of human cell stain derived from a Simpson Golabi Behmel syndrome (SGBS) patient resulted in the identification of 6 novel (pre)adipocytes-secreted proteins (Dahlman *et al*., 2004; Lavebratt *et al*., 2005).

Alpha-2HS-glycoprotein is correlated to adipocytes and obesity. Genetic studies suggest that alpha-2HS-glycoprotein is involved in insulin action by adipocytes and in the regulation of body composition and insulin sensitivity. Lysyl oxidase homolog 2, another protein that is identified here may play a role in remodeling of adipocyte structure during adipogenesis. Furthermore, lysyl oxidase homolog 2 is associated with extracellular matrix remodeling during fibrotic disorders (Decitre *et al*., 1998). Glia-derived nexin belongs to the serine protease inhibitor superfamily and is usually related to neurobiological processes (Silverman *et al*., 2001; Gettins *et al*., 2002). Known adipokines such as plasminogen activator inhibitor 1 (PAI-1) and pigment epithelium-derived factor (PEDF) are also members of the serine protease inhibitor superfamily (Wang *et al*., 2004; Chiellini *et al*., 2008).

Immunoglobulin superfamily containing leucine-rich repeat protein (LRR) and an immunoglobulin-like domain (IgC2), which is mostly expressed in the nervous systems. However, immunoglobulin superfamily containing leucine-rich repeat protein is not only expressed in the nervous system but in various other tissues like heart, skeletal muscle, testis, and retina. Its exact functions in these tissues remain unclear (Wu *et al*., 2006; Nagasawa *et al.*, 1997). Similar to immunoglobulin superfamily containing leucine-rich repeat protein the functions and the exact role of sushi repeat-containing protein and coiled-coil domain-containing protein 96 remains unknown. Further investigations of these proteins are necessary to distinguish their functional properties.

Insulin-like growth factor-binding protein 3 (IBP-3) and Insulin-like growth factor-binding protein 5 (IBP-5) are member of the insulin-like growth factor binding protein (IGFBP) family and encodes a protein with an IGFBP domain and a thyroglobulin type-I domain. These proteins form a ternary complex with insulin-like growth factor acid-labile subunit (IGFALS) and either insulin-like

growth factor (IGF) I or II. In this form, it circulates in the plasma, prolonging the half-life of IGFs and altering their interaction with cell surface receptors. In addition, the identification of relevant adipocyte biomarkers shows that the SGBS cells can be used as a reliable model for the study of obesity related dieseses.

The combination of nLC and MALDI MS/MS proved to be a powerful tool for comprehensive identification of lipoproteins. Since these proteins are embedded in a high concentration of natural lipids, their extraction and purification required a number of sample preparation steps. Even though delipidation enhanced the identification of more lipoproteins and adipocytes-secreted proteins, it cannot be ruled out that some proteins were lost during this additional step, especially proteins which are already under-represented. Poor detection of some proteins can also be attributed to the low abundance of these proteins or insufficient number of cleavage sites for the proteolytic enzyme used in these experiments.

6.3 MLL Interaction partners

6.3.1 AF4/ENL and AF4•MLL complex

For the study of protein-protein interactions within AF4, ENL and AF4•MLL protein complexes, the group of Prof. Marschalek purified 10^9 transfected 293T cells by using Strep-tag affinity columns. The analysis of the purified AF4•MLL protein complexes resulted in the identification of 1223 unique peptides originating from 411 proteins, majority of these being mitochondrial and cytosolic proteins. A more detailed interpretation of the identified proteins revealed that only a handful of these proteins are associated with the AF4 complex (see table 13). Most of the identified proteins have been reported in numerous publications (e.g. AF9, ENL, AF10, DOT1L, CDK9 or CyclinT1). However, the association of the proteins NFkB, BRD4, DDX6, AF5q31 and NPM1 (Nucleophosmin) within the AF4 complex has never been reported (see table 13).

On the other hand, 461 peptides from 80 proteins were identified from the purified AF4•MLL protein complex. Several identified proteins (CDK9, Cyclin T1, ASH2L, WDR5, RBBP5, DPY30, HCFC1, HCFC2, CREBBP) were assumed to be associated with the AF4•MLL complex. In additon, the presence of CARM1 and DOT1L were confirmed by immunoprecipitation studies. The identification of the AF4•MLL complex could not be performed directly since this protein is not present in the database. However, C-terminal peptides from the MLL-protein (131) and the N-terminal peptide of the AF4-protein (21) and 23 C-terminal peptides of the AF4 protein indicated the presence of the AF4 wild type in the AF4•MLL protein complex. A complete list of all identified proteins is seen in supplementary table S4.

Table 13 Summary of the relevant interaction partner identified from the AF4-complex.

Represented is protein accession number (AC), molecular weight in kilodalton MW/k Da), the isoelectric points of the proteins (p*I*) and the number of peptide matched to each protein. All novel proteins are underlined.

	AF4 Complex			
AC	**Protein**	**peptide**	**MW/k Da)**	**p*I***
P51825	AF4(AFF1)	106	132.14	9.26
Q03111	ENL	22	62.42	8.75
P50750	CDK9	12	43.15	8.97
O60563	CCNT(CyclinT1)	11	81.03	8.9
P26196	DDX6	9	54.78	8.85
Q9UHB7	AF5q31(AFF4)	7	127.78	9.33
P06748	NPM1(Nucleophosmin)	6	32.73	4.64
Q9C005	DPY-30	1	11.24	4.84
O60583	Cyclin-T2 (CycT2)	1	81.49	9.04
	AF4•MLL			
AC	**Protein**	**peptides**	**MW/k Da)**	**p*I***

Q03164	MLL(HRX)	131	436.04	9.22
P51825	AF4(AFF1)	44	132.14	9.26
P51610	HCF1	14	210.7	7.49
Q15291	RbBP5	4	59.72	4.92
P61964	WDR5	4	37.14	8.54
P26196	DDX6	4	54.78	8.85
Q9UBL3	Ash2L	4	69.31	5.45
O60563	CCNT1(CyclinT1)	3	81.03	8.9
P50750	CDK9	2	43.15	8.87
Q9C005	DPY-30	2	11.24	4.84
Q9Y5Z7	HCF2	2	87.69	8.75
P06748	NPM1(Nucleophosmin)	1	32.73	4.64

6.3.2 Negative control

To assess the quality of these results, control experiments were performed. The negative control resulted to the identification of 128 peptides originating from 20 proteins. The proteins interacting with MLL•AF4 and the AF4 protein complex were compared to those

identified in the negative control. None of these interaction proteins were found in the control experiments, which substantiates the reliability of not only the novel interaction proteins but also the other interaction partners.

6.3.3 ENL-complex

A structural and functional analysis revealed that the AF4/AF5q31/AF10 binding domain in ENL coincided with the C-terminus that is essential for transformation by MLL•ENL (Zeisig *et al.*, 2005). Interaction partners such as AF4, CDK9, CCNT1, RNA-Polymerase II, transcription factor NFκB and RING1 have been identified. The analysis of an in-solution digest of the MLL•ENLcomplex, resulted in the identification of 339 peptides from 94 proteins. However, 56 % of the identified proteins were “one hit wonders”.Despite numerous nLC-MALDI TOF/TOF analyses, only the AF4 protein could be truely identified within the ENL complex. A reason for the poor results could be the over-expression of the ENL, which may lead to the suppression of interaction partners. Also the additional sample preparation step (buffer exchange) may probably result in the loss of possible interaction partners.

6.3.4 Summary

The human t(4;11) translocation is one of the most frequent MLL translocations currently known. As a result of the translocation, two chimeric proteins are formed, either fusing the N-terminal part of the AF4 protein with the C-terminal portion of MLL or *vice versa* yielding the MLL•AF4 fusion protein. The identification of fusion proteins and their interaction partners has provided a deeper understanding of leukemogenesis. MALDI-MS has long been used for the identification of interaction partners in protein-protein interaction studies. In this case, the samples are separated on a gel and stained; the gel bands are excised, digested, extracted and measured. Coupled with the low resolution power of SDS-PAGE, the process of staining, excision and extraction is tedious and might lead to sample loss. nLC-MALDI-TOF/TOF provides a high-resolution and high-throughput approach for the analyses of complex samples.

In total, 17 known interaction partners (see table 13) from the AF4 and AF4•MLL complexes were identified. The functions of the known interaction partners identified in this study and the role they play in leukemogenesis have been clearly demonstrated in numerous

studies. Furthermore, several novel binding proteins of the AF4 protein (DDX6, NPM1, NFkB, BRD4, CARM1, NSD1) and the AF4•MLL fusion protein complex (DDX6, NPM1, NFkB, HEXIM1, CARM1, DOT1L) were identified. The association of all novel proteins was confirmed by immunoprecipitation studies (Benedikt *et al*, 2010).

Lin et al. showed that Nucleophosmin-1 (NPM1) binds to NFkB1/RELA. NFkB1/NPM1 heterodimers modulate the binding of the E2F transcription factor 1(E2F1) and proline-rich protein BstNI subfamily (pRB) in order to transcriptionally activate the E2F1 promoter (Lin *et al.*, 2006). NPM1 initiates the proteasomal degradation of hexamethylene bis-acetamide inducible 1 (HEXIM1) and thereby may influence the positive transcription elongation factor b (P-TEFb) kinase activity (Gurumurthy *et al.*, 2008). Furthermore, NPM1 was shown to be frequently mutated in AML patients. These mutations change the expression pattern of specific miRNAs (Cazzaniga *et al.*, 2005; Garzon *et al.*, 2008) and influence the outcome of leukemia patients (Grummitt *et al.*, 2008; Baldus & Bullinger, 2008) thereby, indicating a role of NPM1 for human leukemia.

Cho et al. showed that in human MLL, Dumpy-30 (DPY-30) binds directly to SET1/ASH2L histone methyltransferase complex which is a component of the SET domain-containing Histone methyltransferase (HMTs) (Cho *et al*., 2007). As part of this complex, it is involved in methylation and dimethylation at Lys-4 to trimethylation of histone H3. H3 Lys-4 methylation represents a specific tag for epigenetic transcriptional activation.

DDX6 is found in the processing bodies (P-bodies), which are localised in the cytoplasm. It belongs to the DEAD box proteins, which are putative RNA helicases and are characterised by the conserved motif Asp-Glu-Ala-Asp (DEAD). They are involved in a number of cellular processes including alteration of RNA secondary structure such as translation initiation, nuclear and mitochondrial splicing and ribosome and spliceosome assembly. Recently Chu et al. showed that DDX6 is involved in the processing of miRNA (Chu *et al*., 2006).A role of DDX6 in cell apoptosis, cell growth regulations and leukemogenesis has been suggested (Akao *et al.,* 1998; Navarro *et al.,* 2001). As a result of their natural locations, there is some doubt about the interaction of this protein with the AF4 protein complex. While the AF4 exerts its effect in the cell nucleus, DDX6 is located in the cytosol. However, the interaction with AF4 was confirmed by immunoprecipitation reaction (Benedikt *et al*., 2009). For a more detailed understanding of the interaction with AF4, more functional studies are required.

Though the nLC-MALDI MS/MS approach provides high sensitivity for the analysis of protein-protein interactions, there are still some limitations. The real challenge in mapping protein-protein interactions lies on cells cultivation and sample preparation for further nLC-MALDI MS/MS analysis. Moreover, the interaction partners are sometimes present in a small amount. Therefore, high starting concentration is necessary to increase the chance of identifying less abundant proteins. However, the amount of highly abundant proteins will also be increased. Proteolytic digests would result in the generation of large amounts of peptides from the over-expressed or abundant proteins. This may result in subsequent precursor suppression during MS measurement by huge amounts of peptides from the abundant proteins. Therefore, more than one separation technique would be necessary to deal with this complexity. Furthermore, pre-purification on affinity columns, in order to eliminate abundant or unwanted proteins and, sample concentration, may lead to the loss of important proteins as a result of excessive binding to the tubes surface.

Despite the use of affinity columns for the purification of the multi-protein complexes, high amounts of cytosolic and mitochondrial proteins were still identified. Also, it cannot be ruled out that some of the interaction partners are not in close proximity or that they are interacting at different time points. Determining the point in which particular proteins interact with a protein complex is a challenging task. The complete interaction network and path ways are not yet known and it cannot be ruled out that some of the proteins identified, which could be of interest, were not taken into consideration. However, the proteins identified in this study will hopefully provide valuable information to assist future studies aimed at identification and characterisation of high-molecular-weight protein complexes involved in leukemogenesis.

6.4 Result and Discussion: Membrane proteins

This project was done in collaboration with Benjamin Rietschel. Detail results and discussion are present in his Thesis (Rietschel, 2010). However, a part of this result is presented here to demonstrate the suitability of the established nLC-MALDI-TOF/TOF method.

6.4.1 Protein identification

The use of less-specific enzymes in membrane proteomics significantly improves the recovery of integral membrane proteins. Elastase, an example of a less-specific protease of the Serine-protease family S1, has been used in phosphoproteomic studies (Schlosser *et al.*, 2001). Its use for membrane proteomics has not been reported, probably because little was known about the enzyme specificity. However, in phosphoproteomics it has been shown to generate short and often overlapping peptides, which are in the mass range of modern mass spectrometers. Due its broad specificity cleaving at small non-charged amino acids, it can be assumed that this enzyme would be suitable for the analysis of membrane proteins. To evaluate the suitability of elastase in membrane proteomic, *Halobacterium salinarium* purple and *Corynebacterium glutamicum* membranes were digested in 60 % methanol, separated using a 150 min linear gradient at ambient temperature, and analysed with the established nLC-MALDI-TOF/TOF method. Extracted peak list from the Halobacterium salinarium purple membrane were searched against a Halobacterium database, generated from SwissProt/TrEMBL containing 2490 entries. A total of 677 peptides were matched to 143 proteins. Of the 143 proteins, 58 % were identified only with one matching peptide. A search against the Corynebacterium glutamicum ATCC 13032 Bielefeld database containing 3059 sequences yielded a total of 1097 peptides corresponding to 236 proteins. Of these proteins 53 % were identified with only a single peptide. The advantage of the established method over others is that solubilisation, extraction and digestion is performed with the same solvent, the same enzyme and, most of all, this solvent is compactable with RP-HPLC separation and MS (Washburn *et al*., 2001; Wu .*et al*., 2003; Nielsen *et al*., 2005; Bihan *et al*., 2006; Rodriguez-Ortega *et al*., 2006). The large amount of generated peptides indicates the suitability of the enzyme reaction at high organic content.

A tryptic digest of Halobacterium salinarium purple membrane resulted in the identification of 217 peptides originating from 122 proteins, 21 less than for elastase (see table 14). From the 216 proteins identified using both proteases, only 49 were present in both result lists. A total of 94 appeared only in the elastase list and 73 only in the trypsin list. Another major difference between both proteases is the number of peptides matched to each protein. More unique peptides from the elastase digest were matched to individual proteins.

For the tryptic digest of the Corynebacterium glutamicum membrane, 455 peptides were matched to 199 proteins. Again, fewer peptides (< 10) were matched to individual proteins. For example, while succinate dehydrogenase flavoprotein subunit [ATCC 13032] was identified with only 19 tryptic peptides, 57 elastase peptides were matched to the same protein, an increase of about 75 %. Of the 320 proteins identified, only 25 % were identified using both proteases. In general, a comparable number of proteins were identified using one or the other enzyme. The diversity of identified proteins shows the complementary nature of both proteases in protein identification. See supplementary table S3 for a complete list of proteins identified.

Table 14 List of total numbers of unique peptides, proteins, proteins with more than one peptide (proteins > 1), and proteins identified with only one peptide, “one hit wonders” (OHW) and the number of transmembrane peptides (TM-peptides) identified in both samples analysed

	PM(Elastase)	PM(Trypsin)	CB(Elastase)	CB(Trypsin)
Peptides	677	217	1097	455
Proteins	143	122	236	199
OHW	83	74	125	107
Proteins >1 peptide	60	48	111	92
TM-peptides	174	20	149	-

6.4.2 Transmembrane coverage.

Because IMPs often have a significant portion of their sequence embedded in the membrane, the probability of identifying these proteins decreases, if the method is restricted only to soluble domains. A means to determine the success of a membrane proteomic strategy is the identification of IMPs, especially from proteins containing multiple TMDs. These TMDs are known for the problems encountered during solubilisation and separation. A reliable and common tool for predicting the location of TMD of proteins is by using TMHMM 2.0. The analysis of the identified proteins from both samples (143 PM and 236 CB) predicted that 44.06 % of these proteins exhibit at least one TMD. Within the 44.06 % of the integral membrane proteins identified, approximately 81.44 % contained at least 2 TM domains. In a control study using trypsin, in which 316 proteins were identified, less than 46.52 % of the proteins were predicted to have at least one or more TMD domain and 63.94 % of these 46.52 % exhibited more then one TMD. However, most of the identification resulting from the tryptic digest was achieved over the hydrophilic region. Elastase, on the other hand, generated higher number of membrane-spanning peptides (Rietschel *et al*., 2009). For example, whereas 89 TMD peptides from bacteriodopsin were identified with the elastese digestion, only one TMD peptide was identified when trypsin was used as a proteolytic enzyme (see figure 38). Another example is observed with the small membrane protein, porin. In the tryptic digestion, this protein was identified with only one peptide. When using elastase, the protein was identified with 8 peptides from a single helix. Furthermore, prominent respiratory chain proteins, such as cytochrome b, ATP synthase subunits a, c and cytochrome c oxidase I, were identified with a higher number of TMD peptides than for a typsin digest. The first 75 proteins identified from the elastase result list (see supplementary table S3) were membrane proteins. In total, 174 peptides of predicted TMD were identified by the elastase approach, in contrast to only 20 peptides in the tryptic digest experiment. Furthermore, the increased sequence coverage and transmembrane peptides identified with just one enzyme show the superiority of elastase over other methods in which two enzymes were combined for membrane proteomic studies (Washburn *et al*., 2001; Wu .*et al*., 2003; Nielsen *et al*., 2005; Bihan *et al.*, 2006; Rodriguez-Ortega *et al*., 2006;). In addition, the good solubilisation of the membrane proteins in 60 % methanol, which was first reported by Blonder et al. (Blonder *et al*., 2006), improved the assessibility of elastase to the TMDs. Zhang et al. also

showed that a higher number of integral membrane proteins were identified when solubilised in 60 % methanol than with SDS (Zhang *et al.*, 2007).

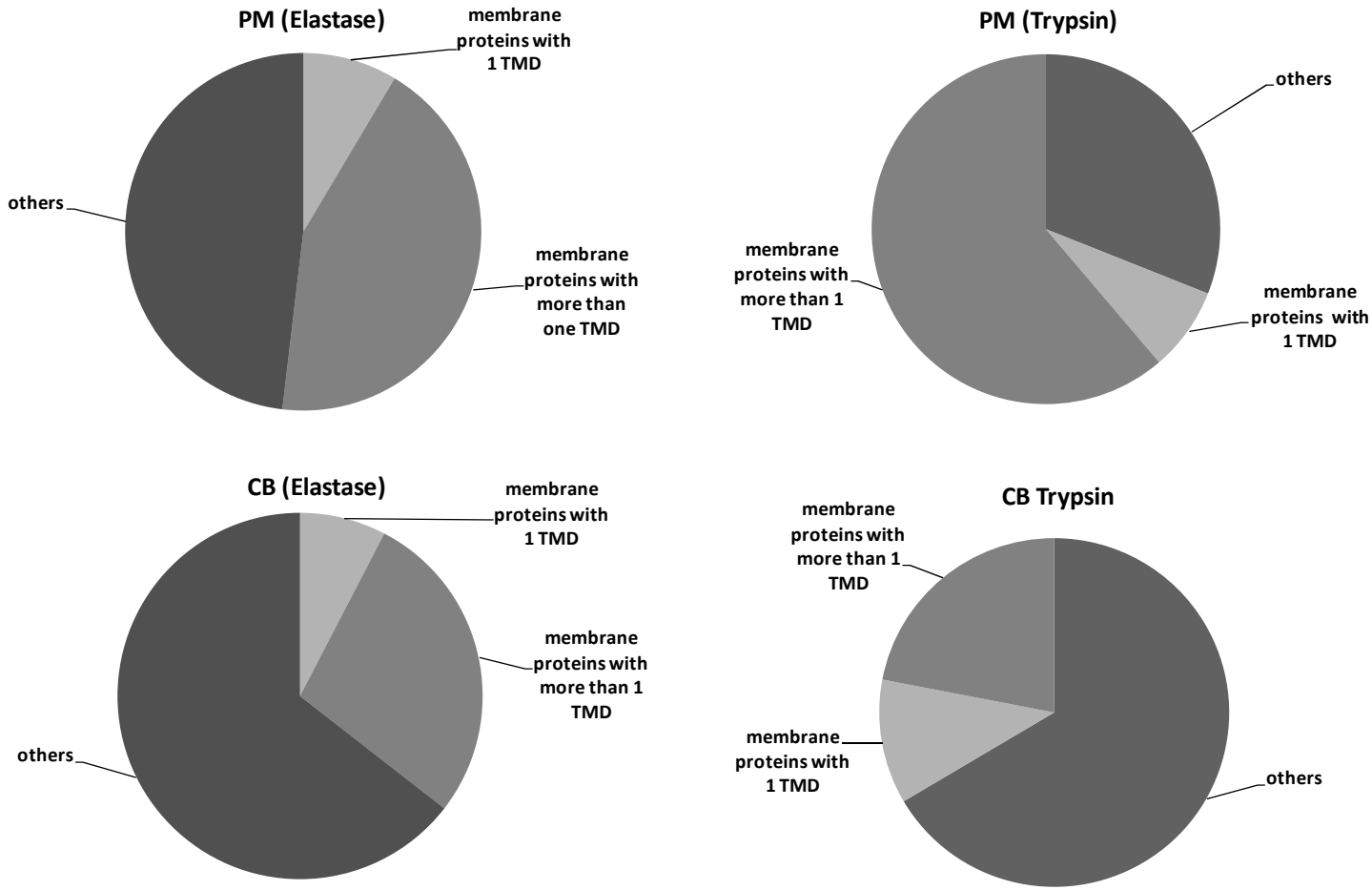

Figure 38 The number of membrane proteins, with one or more TMD identified for both PM and CB samples using elastase and trypsin as proteolytic enzymes.

6.4.3 Fragmentation behaviour of elastase generated peptides

Fragmentation plays an important role in shotgun proteomics especially when analysing uncharacterised proteins or in search of post translational modifications (PTM). Fragmentation methods and the collision energy used for fragmentation are known to largely change the characteristics of peptide fragmentation (Arnold *et al.*, 2006). Typically under CID conditions, tryptic peptides generate a number of different fragment ion series, wherein N-terminal (b-ions) and C-terminal (y-ions) fragment ions dominate the tandem mass spectra (Roepstorff *et al.*, 1984; Johnson *et al.*, 1988; Steen & Mann, 2004). The fragment ions from 7000 precursor masses revealed that nearly 29 % of the generated ions belong to the group of the internal fragments and their associated neutral losses. The rest is distributed among C-terminal ions (y, y0, y*) and the N-terminal type (b, b0, b*, A) ions (see table 15). A reason for the increase in internal ions and a decrease in b and y-ion series is probably due to the lack of charge localisation at the C- and N-terminus of peptides generated by less specific enzymes. The increase in internal fragment ions formation led to difficulty in spectral interpretation. However, by modifying the peptides with a charge-directing moiety at the N-terminal such as nicotinyl group, TMT or iTRAQ that produces extended b-ions series many unwanted internal fragment ions can be suppressed (Bendez *et al*., 2009).

Table 15 An overview of generated fragment ions

Fragment-Ionen-types	Number	%
b-Ion (N-term sequencing ions)	46.969	5.9
b^0-Ion (neutral-loss H_2O)	11.968	1.5
b*-Ion (neutral-lossNH_3)	13.227	1.7
a-Ion(neutral-lossCO)	26.818	3,4
y-Ion (C-term sequencing ions)	38.695	4.9

y^0-Ion (neutral-lossH_2O)	12.083	1,5
y*-Ion (neutral-lossNH_3)	13.203	1.7
i-Ion (Immonium-Ion)	52.575	6.7
int-Ion (internal fragment-Ion)	95.193	12.0
int^0-Ion (neutral-lossH_2O)	37.297	4,7
int*-Ion (natural lost NH_3)	39.606	5.0
$int^{\dagger}$ (natural lost CO)	58.275	7.4
Unmatched fragment	344.486	43.6
Sum	**790.395**	**100**

6.4.4 Summary

Until recently less specific enzymes did not play a relvant role for the analysis of membrane proteins. However as shown above, they improve the coverage in the hydrophobic domains of these proteins especially when used with a suitable solvent such as methanol, which is compatible to mass spectrometry. In this study, elastase was combined with the the solubilisation effect of methanol for the analysis of membrane proteins. The suitability of this method was evaluated using *Halobacterium salinarium* purple and *Corynebacterium glutamicum* membranes. Approximately 85 % of the proteins identified using this method were predicted by the α-helical transmembrane domain (TMD) prediction algorithms, TMHMM 2.0, to have at least one TMD. Approximately 60 % of these

proteins have at least 2 TMD. In a related study using trypsin as a proteolytic enzyme, only about 18 % of the identified proteins had one or more TMD. Compared to the tryptic digest, more transmembrane peptides were identified with the elastase digestion. These results confirmed the increased importance of less specific enzymes and the potentials of elastase for membrane protein analysis. The success of this enzyme for the analysis of membrane proteins is as a result of the generation of a huge amount of peptides. However, as shown by the results presented above, care should be taken when working with elastase or less specific enzymes in general, since common database search engines are not optimal for these types of enzymes. Since the enzyme specificity is omitted, a large number of peptides from different species can be matched to a precursor. Hence, the rate of false positive discovery will increase. This can be compensated by improving the mass accuracy of the precursors either by internal calibration and/or high resolution mass spectrometer such as the Orbitrap. Furthermore, elastase and other less specific enzymes leads to the generation of peptides, which lack C- or N-terminal charged amino acids. As a result, extended b- and y-ion series are prohibited during fragmentation. Poor fragmentation may influence the search result. However, extended b- and y- ion series can be again furthered by derivatisation using tandem mass tag (Bäumlisberger *et al*., 2010).

Despite their high importance and a considerable effort in method development, current approaches aiming to obtain a comprehensive coverage of the membrane proteome are far from being satisfactory. Even though the potential of elastase has been shown for phosphopeptides analysis, it has found limited application in proteomics, presumably because of its largely unspecific cleavage. However, Rietschel *et al*. showed that this enzyme cleaves preferentially at 6 amino acids (AVLIST), which render it very suitable for complex membrane proteome analysis (Rietschel *et al*., 2009). The limitations encountered by trypsin, such as lack of tryptic cleavages sites, can be overcome by using elastase. Hence, the implementation of this enzyme in database search engines would be of great help not only to membrane protein analysis but also for proteins with a low number of tryptic cleavage sites. All in all, elastase proved to be superior to trypsin in the analysis of membrane proteins. In conclusion, a combination of both enzymes will further improve the outcome of proteomic studies.

7 Two Dimensional Separation (Optimisation of the Offgel-IEF for its use as a first dimension in a two dimensional separation plateform)

7.1 Proteins and peptides identification.

To evaluate the separation efficiency of the Offgel-IEF as a first dimension for a multiple dimensional separation, 200 µg of purple membrane (PM) and *Corynebacterium glutamicum* membrane (CGM) pellates were digested with elastase. The digested samples were dissolved with 1 % IPG buffer to a total volume of 1.8 ml. 150 µl of this sample solution was loaded in each well of the 12-well setup placed on a 13 cm long IPG-gel strip with a linear pH gradient ranging from 3 to 10 and focused with a maximum current of 50 µA for 24 hours. Prior to nLC separation and MALDI TOF/TOF analysis, each fraction was vacuum concentrated. During sample concentration, the amount of glycerol, commonly included (10 % Glycerol) to all chambers to prevent electroosmotic peptide-water co-migration, is increased, thus, resulting in a viscous solution. This highly viscous solution makes sample pickup difficult and loading on the nLC pre-column resulted in clogging, which in turn increases backpressure. To circumvent this problem, glycerol was totally substituted with additional buffer solution to wells 1, 2, 11 and 12. After focusing, each fraction was concentrated, separated and analysed. The MS and MS/MS data were converted to Mascot generic files, combined and searched against the *Corynebacterium* database and complete *Halobacterium salinarium* subset of NCBI containing, 5320 sequences (3rd of December 2008). In total, 866 unique peptides originating from 147 proteins from the PM and 470 unique peptides from 149 proteins of the CGM sample were identified (see table 16). A complete list of all peptides and proteins identified is shown in Table S5 in the Supporting Information. Using TMHMM, α-helical TMD prediction algorithms 2.1, 54.4 % (80) of the identified proteins from the purple membrane and 29.5 % of the proteins from the CGM sample were predicted to have one or more TMDs (see figure 39).

Table 16 OGE fractionation results**:**

Overview of Off-Gel IEF nLC-MALDI-MS/MS experiments for elastase digest of PM and CGM samples displaying the total number of identified proteins and peptides, False discovery rate, the number of proteins identified with just one peptide (“one hit wonders”) and the peptide number/sequence coverage of the BR Protein.

	PM	CGM
peptides	866	470
proteins	147	149
protein identified with 1 peptide	64	87
false positive rate [%]	0.81	0.77
BR peptides	81	**
BR sequence coverage [%]	71.5	**

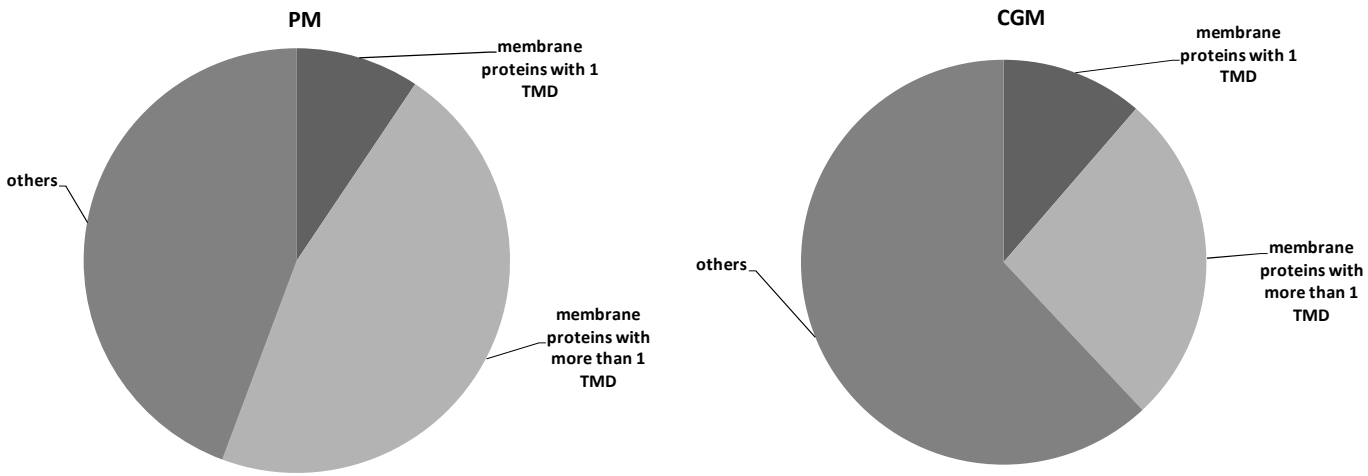

Figure 39 Membrane proteins identified from PM and CGM digested with elastase.

In general, most of the peptides were present in the acidic fractions, especially in the PM sample, which contains approx. 90 % BR and minor cytosolic proteins (see figure 40). The number of observed peptides dropped to a minimum value at fraction 10. This fraction represents peptides with theoretical pI values between 8 and 9. These isoelectric point values can only be generated by unrealistic amino acid combinations on the peptide level. The same separation behavior was observed with tryptic peptides analysed with the 12 and 24-well setup (Heller *et al*., 2005; Hört *et al*., 2006; Lam *et al*., 2007; Chenau *et al*., 2008).

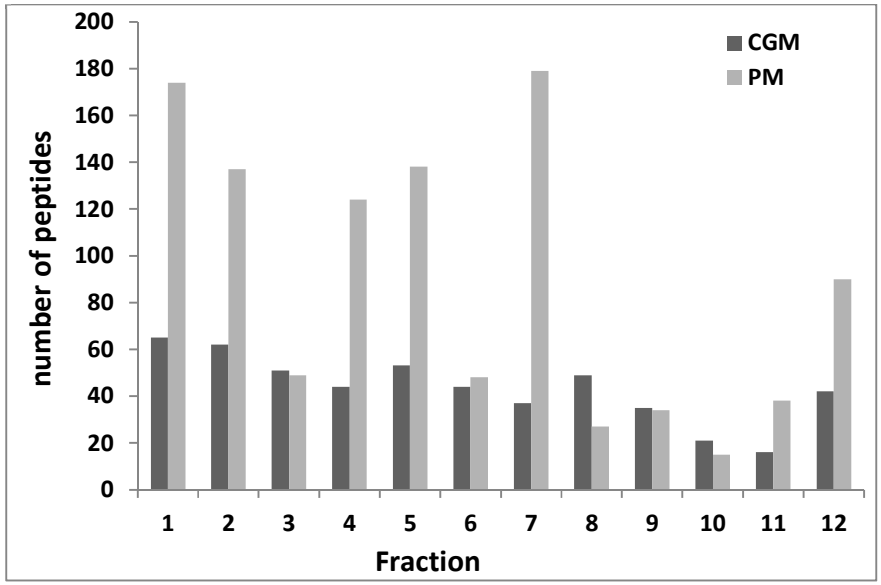

Figure 40 Total number of peptides identified in each fraction; the light gray bars represent peptides from the PM and CGM samples.

7.2 Peptide Carryover

Since it was not possible to estimate the separation efficiency by comparing samples focused in the presence and absence of glycerol, the peptide carryover to an adjacent fraction was used as a criterion. Optimally, for equimolar concentration, all peptides are expected to appear in just one or at most 2 fractions. Complex biological samples exhibit a wide dynamic range of protein concentration. Highly abundant proteins would result to the generation of large amounts of signature peptides, thus a higher risk to be recovered in several wells, and peptides from low abundant proteins in one well (Lam, 2008). Nevertheless, for both samples, approximately 85 % of all peptides were identified in just one fraction and about 95 % in a maximum of two fractions (see figure 41). Only few peptides were

observed in more fractions (see figure 41). Peptides which were focused in 3 or more fractions originated from the proteolytic enzyme through autolysis and from bacteriorhodopsin. These peptides are well-known for their high abundance in elastase BR digests and originate from the highly accessible N-terminus and second loop/helix C regions (Rietschel *et al.*, 2009).

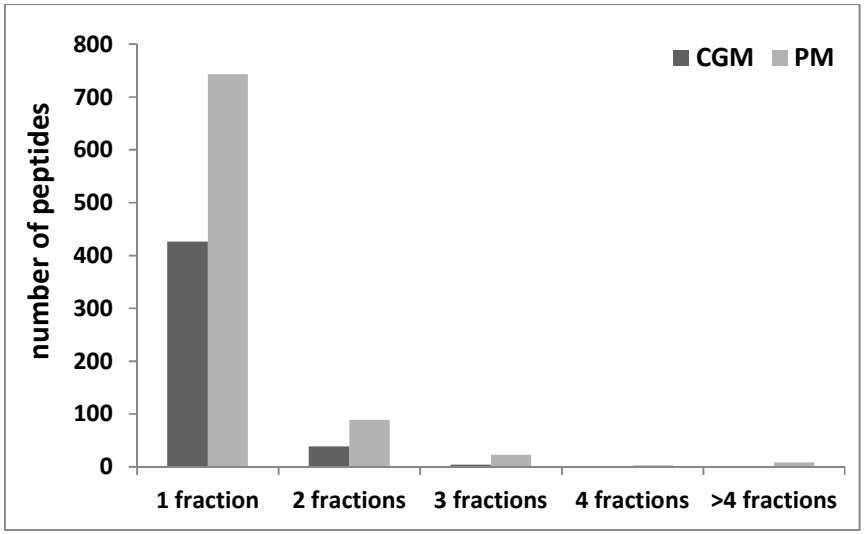

Figure 41 Number of fractions in which each distinct peptide was found.

7.3 Standard deviation

The OGE separates peptides according to their p*I*, so the distribution of peptides in the different fractions can be used to determine the separation efficiency.The average p*I* should increase by 0.58 pH units per fraction, because a 13 cm long IPG gel strip with a linear pH gradient ranging from 3 to 10 was used. Slightly lower actual values of 0.53 and 0.56 for both samples respectively were observed: The average pH increase follows a linear trend except in the neutral pH range around fractions 6 and 7, where no substantial pH

difference is observed (see figure 42). IEF seems to be rather limited for neutral peptides, because most of these peptides do not carry more than the two functional charged groups at their termini. Also, increased p*I* variances in the alkaline fractions are attributed to the lower peptide numbers used to determine these statistics (see figure 42).

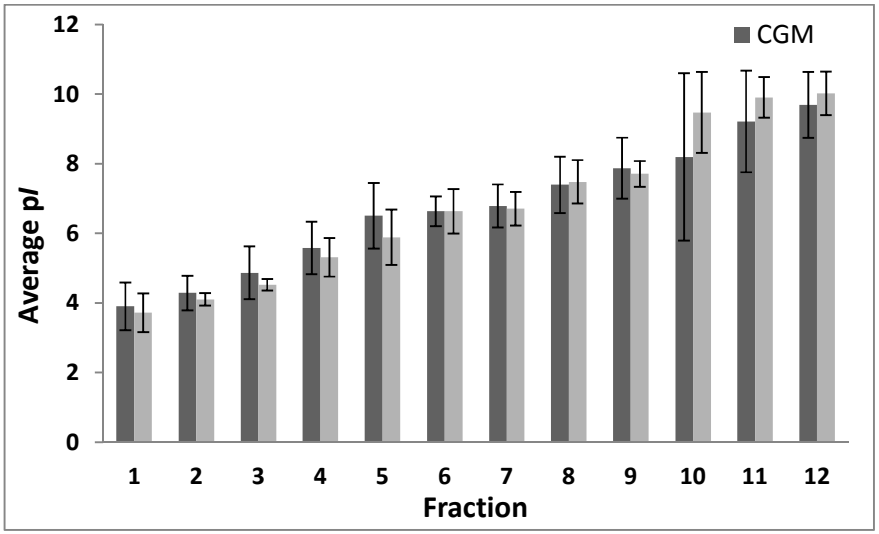

Figure 42 Calculated average p*I* values for all peptides identified in each fraction for both samples.

Generally, pH deviations can also be influenced by the position of the well on the gel strip. In cases where the well is exactly between pH gradients, peptides would appear in these 2 fractions. Furthermore, commercially available strips vary slightly in relative length and position of the immobilised gradient. These variations can also result in pH deviations or affect other factors such as reproducibility (Hubner *et al*., 2008).

7.4 p*I* distribution

Besides the increased separation power, IEF techniques also provide additional physiochemical information about each peptide, i.e. its isoelectric point (Cargile *et al*., 2004; Cargile *et al*., 2004; Heller *et al*., 2005; Uwaje*et al*., 2007). During focusing, peptides migrate through the IPG-strip and settle in the well at which their net charge is zero (p*I*). Detailed information of each amino acid from Nelson and Cooks *Lehninger Principles of Biochemistry 3rd Ed.* (2001) were used for the p*I* calculation. The calculated p*I* distribution of each well shows an uneven distribution across the pH scale.

An overall view of the deviation of the predicted p*I* values of every identified peptide spectrum from the average p*I* calculated for each fraction is shown in figure 42. The majority of the peptides were found in the p*I* range from 3.6 to 4.7 (fractions 1-2). These fractions display homogeneous clusters within the appropriate areas as shown in the peptide mass vs. calculated p*I* scatter plots (see figures 43). The fractions become more diffuse in the neutral region, followed by a cluster of His-dominated peptides and an obligatory lack of peptides having p*I* values between 8 and 9. In this case, the number of identified peptides dropped to a minimum. The last two fractions, carrying strictly Lys & Arg-based peptides, display significant overlap caused by loss of fractionation efficiency at this end of the IPG strip.

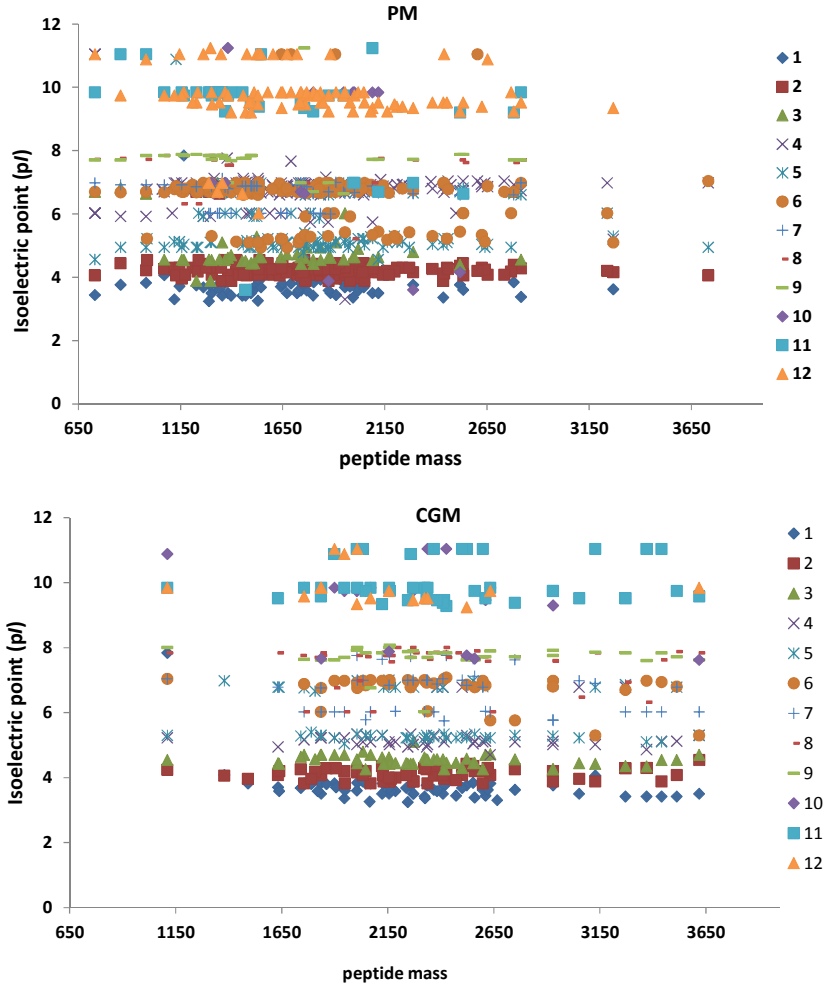

Figure 43 Peptide p*I* distribution. Scatter plot of the calculated pI value of every peptide against its mass for the PM sample.

7.5 Using the p*I* of peptides to manually validate search results

A major advantage of the OGF is its ability to predict the pI values allowing validation of the identified peptides (Cargile *et al.*, 2005). This can be achieved by calculating the p*I* of the identified peptides from each fraction, average fraction pH value and the standard deviations (Stdev). These values allow setting a tolerance window and eliminating peptides deviating too far from the average value. For a 13 cm linear gradient strip, pH 3-10, the pH should increase by a value of 0.58 units per cm. Peptides focused in this fraction, even when using a greater standard deviation, corresponding to a pH difference between 2 wells should not be above 1.2 units per cm. For example, the first peptide (HSIAAQGGVNSA) in fraction 1, pH region 3.5-4, of the result list of both samples (see table S5C) has a calculated p*I* value of 7.4. This peptide has a total ion score (70.1), which is above the homology and identified score (33/35). Taking the standard deviation into account, this peptide should not appear in this fraction, thus it can therefore be considered a false positive identification. Also the peptide (FRALLL) identified in fraction 5 (pH range 5.5-6.8) with a p*I* value of 11.04 should in the case of a p*I* filter be considered as a false positive identification (see figure 44). This is only possible if the focusing efficiency has been validated. Figure 4 shows a plot of the p*I* of peptides identified in fraction 2. The peptides in circle are above the standard deviation and thus, can be eliminated from the list. However, p*I* filtering has not yet been introduced or established in modern search engines. A reason for this could be the difficulty in predicting physiochemical properties, especially if the peptides are modified (Heller *et al.*, 2005). For example, there are many different formulas (EMBOSS, DTASelect, Solomon, Sillero, Rodwell, Patrickios) for the calculation of the p*I* values and each delivers different results as shown in figure 2 for 12 peptides chosen from each well (see figure 45). The theoretical values were obtained by adding the average pH increase, 0.58, to the pH gradient of the strip (3-10).

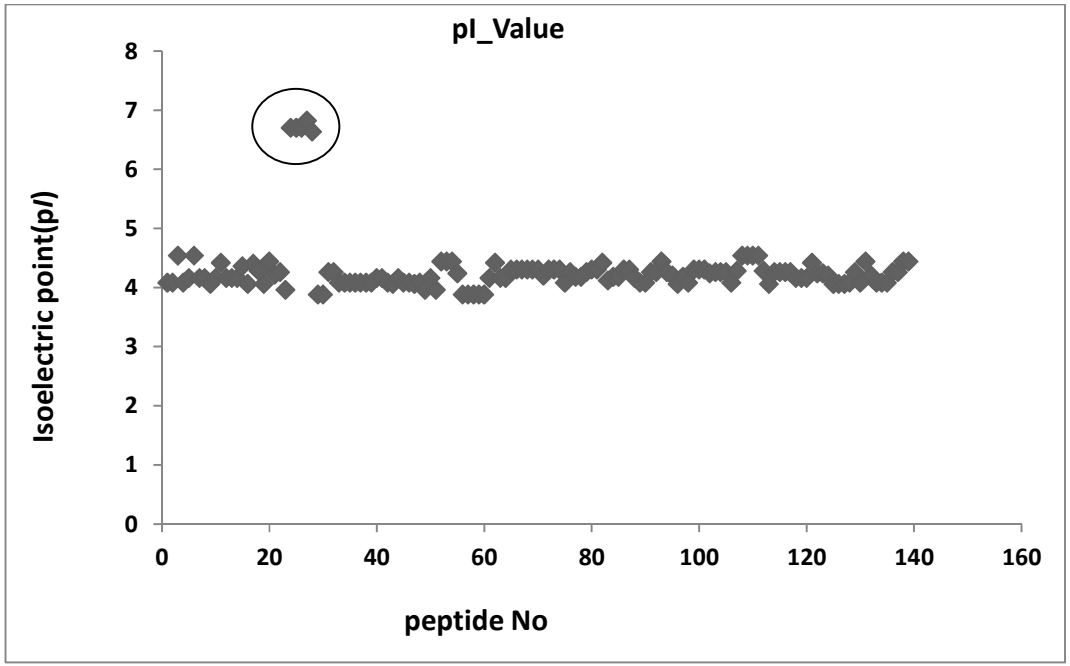

Figure 44 A density plot of pI for each peptide identified in fraction 2; encirculated peptides indicate false positives (see text).

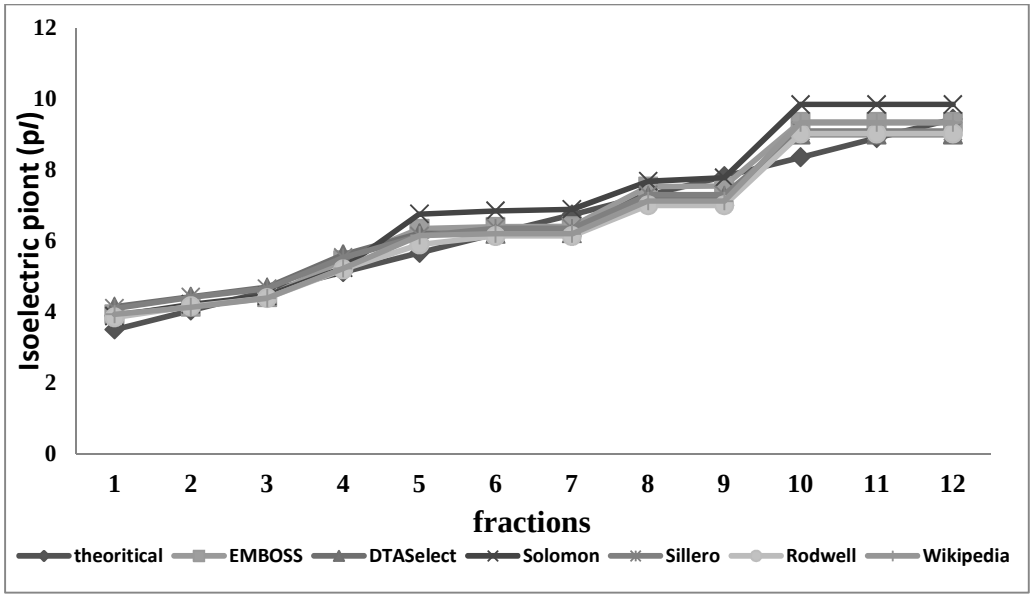

Figure 45 pI values calculated for 12 peptides chosen from the 12 wells using 6 different formulas.

7.6 Suitability of the Offgel-IEF peptide fractionators for relative quantification

One of the major advantages of TMTsixplett is that parallel proteomic analysis of six different samples or technical replicates of samples can simultanenously be achieved, therefore reducing the total analysis time. Hence multi-dimensional separation techniques are required to reduce sample complexity. To determine the suitability of relative quantative technique, such as TMT with the optimized glycerol free offgel-IEF protocol, membrane proteins extracted from yeast cell grown in 2 different media, glucose (YPD) and oleic acid (YPO) were analysed. In order to obtain technical replicates, each sample was separated in 3 fractions, digested and labelled with TMT 6-plex. All the 6 fractions were combined, focused, each of the 12 IEF fractions were separated on a nLC and analyzed with the MALDI TOF TOF 5800 (AB SCIEX).

Due to modifications of N-terminal amino acids and/or lysines, a more basic moiety is introduced into the peptide sequence. Therefore, it would be expected that these peptides may be focused differently from unlabelled peptides, probably with a shift towards the basic region. Nevertheless, most of the peptides identified were focused in the acidic regions (1-5). Moving towards the basic region, the number of identified peptides decreases. Lengqvist et al. also showed that peptides labeled with iTRAQ experienced only a minimal pH shift (0.17-0.22) when the pI-strip covering a smaller acidic range (3.4-4.9) was used (Lengqvist *et al*., 2007). For a wide pH range (3-10) this shift is negligible. If this is true, the separation of unlabeled peptides should show the same behavior as their labeled counterparts. An in silico digest of the same proteins identified in this experiment was performed, and the distribution of the calculated pI was estimated. Except for fraction 11 and 12, the pI distribution plot of the labeled and the unlabelled in silico digest fits fairly for all fractions (figure 46 A). However, the general results show that the change in peptide p*I* through TMT has a negligible influence when focusing on a wide pH range IPG-strip. While the fraction containing more peptides had a lower standard deviation, higher standard deviations were observed for fractions with a lower number of peptides. Since 84.4 % of all peptides identified were located in a single fraction (see figure 46 B.) and 96 % of peptides in one or two, it can be concluded that the separation efficiency was not compromised by the basic (TMT) residue. These results are not different from those obtained from labelled and unlabelled peptides (Chenau *et al*., 2008). Therefore, the glycerol free protocol can be used in comparative quantitative proteomics studies without restriction.

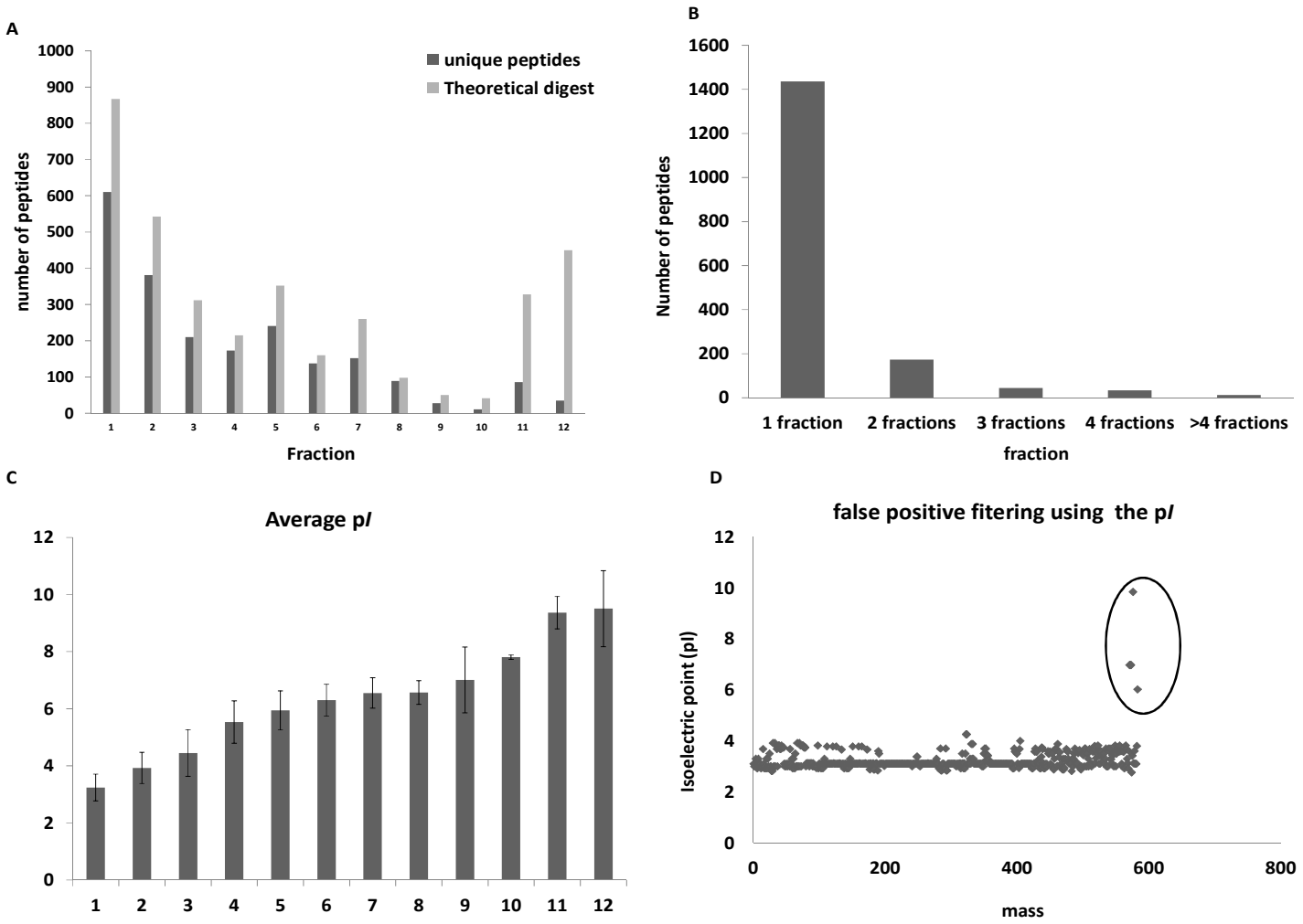

Figure 46 (A) Most of the peptides were identified in the acidic region. (B) About 85 % of the peptides were focused only in one fraction and 96 % in 2 fractions. (C) The average standard deviation is larger for fraction with less number of identified peptides. (D) Shows that the peptide p*I* can be used as filtering citeria to minimize false positive identification.

7.8 Protein identification

Because tandem-mass-tag reagents target all peptides for labelling, multiple peptides from the same protein can be detected. Protein identification was achieved by peptides for which individual ion scores were above the 95 % confidence threshold ($p < 0.05$). In general, a total of 323 unique proteins were identified from 1756 peptides. A total of 200 proteins (62 %) were identified by two or more of the significant peptides. However, only 22.5 % were membrane proteins. Nevertheless 85.9 % of these membrane proteins carry more than one transmembrane dormain (see Table S5D in the Supporting Information).

Most of the peptides identified exhibited mass-to-charge ratios between 1000 and 2000 (see figure 47). A reason for the identification of large number of peptides in the region comes as a result of the basic TMT. This effect is more predominant for smaller and medium-sized peptides then for large peptides. An increase in basicity leads to enhanced signal intensities and signal/noise ratio. Therefore, these peptides would be more likely chosen for fragmentation (Baeumlisberger *et al.*, 2010). These results show that TMT does not only aid in the relative quantification, but also enhances fragmentation and improves identification.

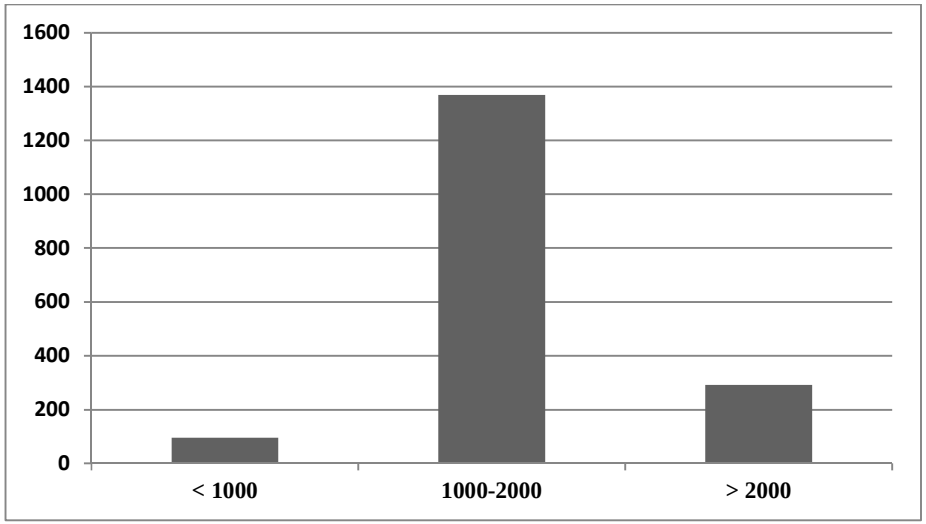

Figure 47 Molecular weight distribution of peptides matched to proteins from elastase digest of YPO/YPD samples.

7.8 Protein quantification

If cells of the same cell-lines are cultivated in different media or exposed to different environmental conditions, particular genes might be prone to regulation. The extent of regulation is dependent on the gene and the conditions. For yeast cells grown in glucose and in oleic acid this changes can be monitored by comparing the relative intensities of reporters ions (see figure 48) generated from peptides labeled with a tandem mass tag containing 6 tags at m/z = 126.1, 127.1, 128.1, 129.1, 130.1, and 131.1. For the analysis of changes in growth media, a threshold value of 2-fold in either direction is considered as a significant regulation. In these studies, a total of 199 proteins having two or more significant scoring peptides were relatively quantified. The most prominent being the Plasma membrane ATPase and the Hexose transporters (HXT). In general 9 (HXT1, HXT2 HXT3 HXT4 HXT5 HXT6 HXT7 HXT9 and HXT11) of the 17 (HXT1 to HXT17*)* hexose transporters were identified and quantified. These proteins transport glucose through the plasma membrane by

two kinetically distinct systems, a glucose-repressible high-affinity system and a constitutive low-affinity system. Since the yeast cell primarily uses glucose as the main carbon and energy source, efficient metabolism of glucose can be used as an environmental stimulus that regulates the quantity, types, and activity of glucose transporters, both at the transcriptional and post-translational levels (Özcan & Mark Johnston, 1999). Hxt2, Hxt6, and Hxt7 are responsible for transporting glucose in cases of low or the absence of fermentable carbon source (Liang *et al*., 1996; Boles *et al*., 1997; Reifenberger *et al*., 1997). The results obtained showed that these proteins and others were slightly up-regulated. Citrate synthase, a mitochondrial protein was the most up-regulated protein. The presence of this protein in high amount is a result of contamination, since it was difficult to isolate membrane proteins from yeast cells grown in oleic acid. A complete list of all the quantified proteins is found in the supplementary sheet.

To determine the distribution of quantification errors, the ratio of the tags was calculated (see figure 49). Since the protein expression is not expected to change within the samples after protein harvest, a triplicate comparison should be 1:1:1. The standard deviation of the measured ratios was below 20 % for all three experiments.

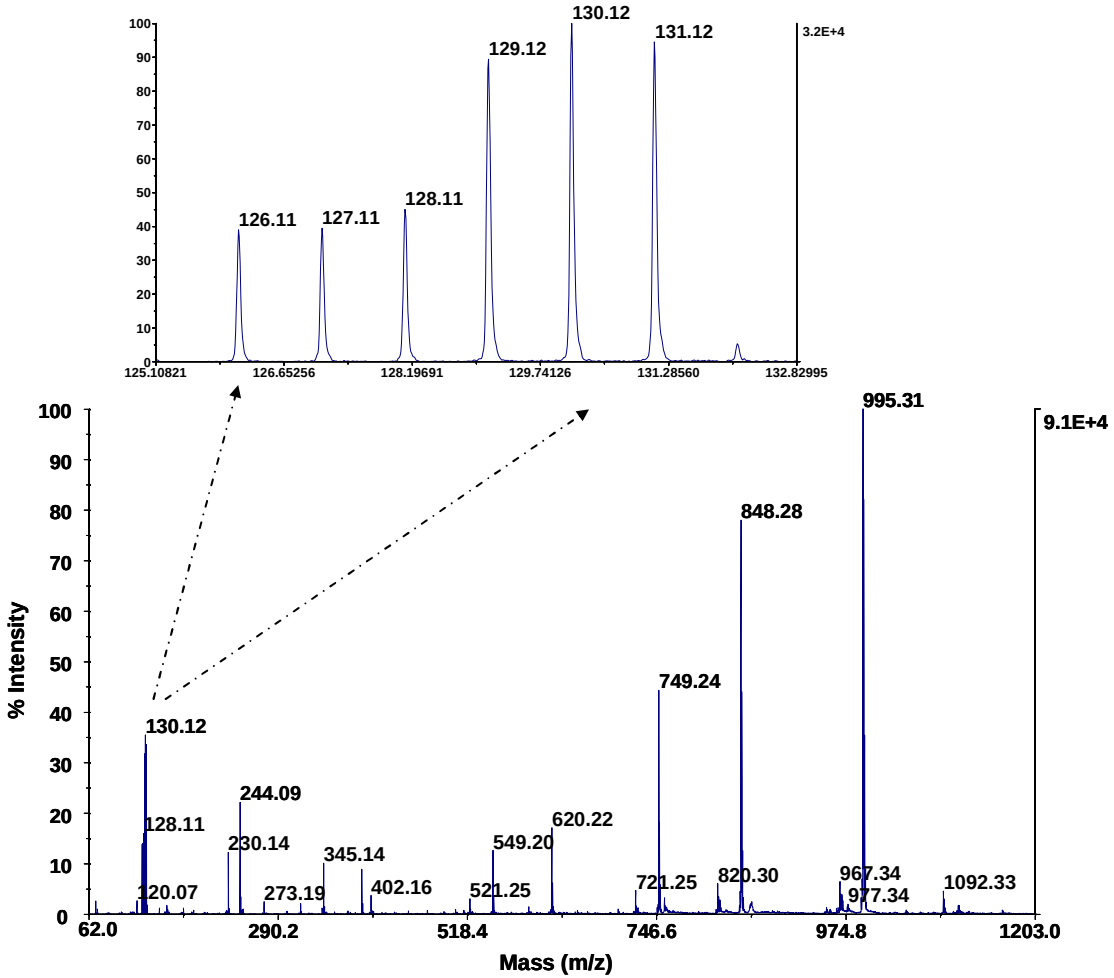

Figure 48 Fragment spectrum of peptide at m/z = 1238.73 Da. The zoom of the reporter ion region that facilitates the peptide relative quantification through the abundance of the reporter ions at m/z = 126.1, 127.1, 128.1, 129.1, 130.1 and 131.1

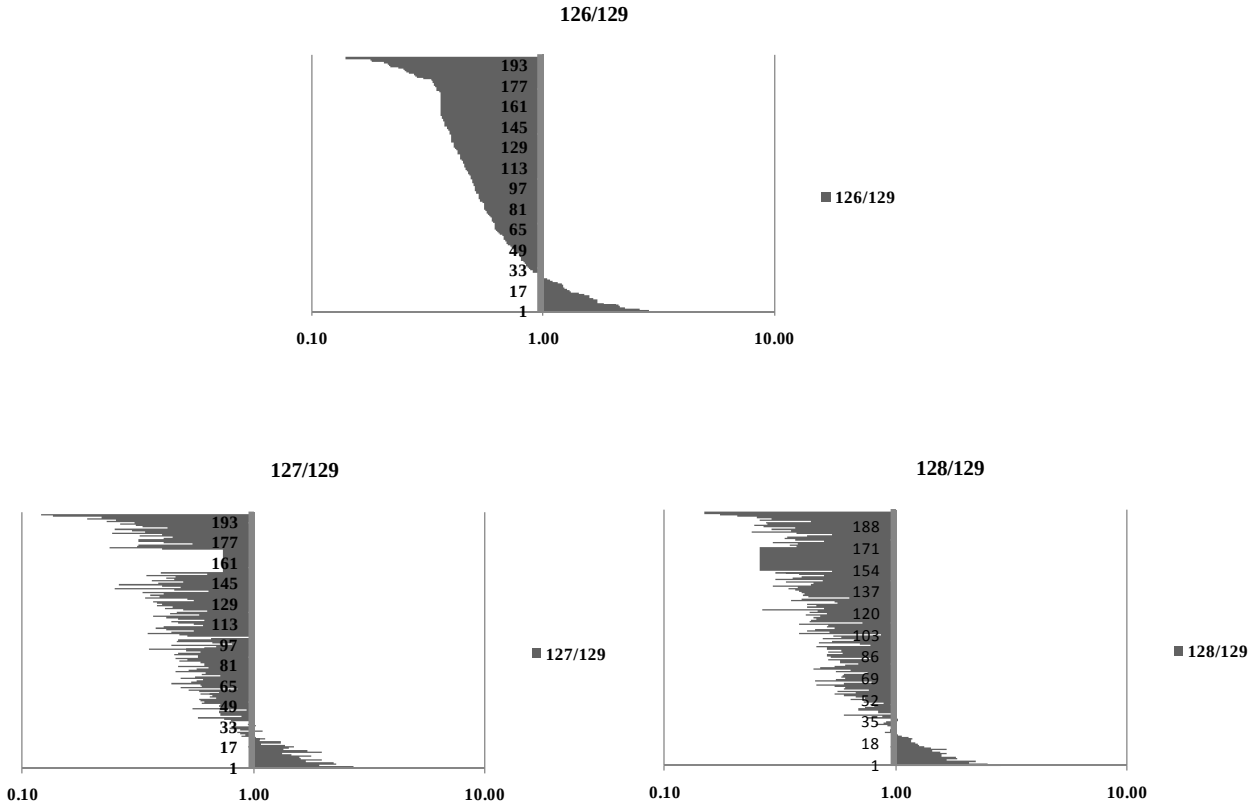

Figure 49 Summary of all quantified proteins. Most of the proteins in the yeast cell grown in oleic acid were down regulated.

7.9 Conclusion

In many published reports using the manufacturer´s separation protocol (20 % Glycerol) approximately 85-90 % of all peptides were identified in 1 or 2 fractions (Heller *et al.*, 2005; Hörth *et al.*, 2006; Lam *et al*., 2007; Chenau *et al*., 2008). Using the glycerol free protocol, approximately 90-95 % of all peptides were identified in one or two fractions, with very little carryover. Therefore, neither the absence of glycerol nor the increased volume in the outer chambers interfered with the Offgel-IEF fractionation process. However, the modification prevented the wells from running dry, which would normally happen due to the electroosmotic peptide-water co-migration. In addition, due to the solubilising effect of methanol, no precipitation was observed during the entire fractionation process. Furthermore, each fraction was directly injected onto the pre-column after vacuum concentration without resulting in clogging and high back pressure. Compared to conventional 2 D separations methods, Offgel-IEF permits the separation of large amounts of samples. However, in spite of the high resolving power, the separation efficiency of Offgel IEF in the neutral and alkaline p*I* region is still far from being optimal. A reduction of the electro-osmotic water flow should overcome this limitation. Strategies such as using a 10 % sorbitol gradient, and adding organic solvents or methyl cellulose, have been suggested (Hubner *et al*., 2008). Unfortunately, sorbitol and methyl cellulose have a higher density than glycerol and would lead to the same problems. Even though the use of additional focusing buffer, in place of glycerol, prevented the outer chambers from running dry, the focusing quality in the critical p*I* regions was not significantly improved. The water migration is suspected to be a major reason for the lower focusing quality in the 24-well format (Hubner *et al*., 2008). Furthermore, the lower consistency of p*I* prediction algorithms in the basic pH range implies that the accuracy ofthe p*I* prediction is not better than in the acidic pH range (Cargile *et al., 2004*). Although there is still much to be done about the focusing efficiency in the neutural and alkaline p*I* regions, an optimised protocol was established using direct sample injection, separation via nLC without column clogging and increased backpressure and mass analysis by MALDI-TOF/TOF.

Detection and quantification of differences in the protein profiles of cells, tissues or body fluids of different origins or states is increasingly being recognised as a key objective of proteomics research (Ji *et al*.**,** 2005). Techniques that are largely based on stable isotope protein or peptide labelling and automated tandem mass spectrometry are increasingly being applied in quantitative proteomic

studies. Shotgun proteomics experiments involving TMT reagents have been successful in identifying and quantifying proteins across a variety of prokaryotic and eukaryotic samples. One of the major advantages of TMTsixplex is that parallel proteomic analysis of six different samples can be achieved, reducing the total analysis time. However, the compatibility with different separation techniques such as the Offgel-IEF has not been assessed. The applicability of the Offgel-IEF glycerol free protocol was assessed for relative quantitation studies. Generally TMT labelling induces a small p*I*-shift compared to the native peptide. Nevertheless, the results obtained from in silico digest of the same identified proteins shows that, this difference is relatively small and can only be observed when using IPG strips with very narrow pH ranges. For a large pH range like the one used in these experiments (3-10), this shift is negligible. The results presented here show that the focusing efficiency for peptides is not altered by the TMT derivatisation. Therefore, this method would be suitable for the inclusion in quantitative studies using Isobaric Tags. Furthermore, the introduction of TMT to the peptide backbone significantly improved signal-to-noise, signal intensity in MS and fragmentation yield and quality. An improved signal-to-noise ratio with increased signal intensity in MALDI-TOF/TOF of isobarically tagged peptides resulted in detection of a greater number of small and medium-size peptides. This has also been described for experiments using iTRAQ isobaric tags (Aggarwal *et al*., 2006). However, care should be taken when dealing with proteins identified using short and medium size peptides, since they may fit perfectly to many other proteins.

The inclusion of Offgel electrophoresis into our proteomic workflow offers several advantages: (i) the p*I* of the peptides which can be used to further substantiate the search result, thereby reducing the false positive rate (Tan *et al., 2002,)*, (ii) the peptides are recovered in solution and can directly be loaded on the nLC for further purification and (iii) it offers the possibility to start with large amounts of samples, thus increasing the probability of identifying low abundant proteins. Unfortunately, the IEF separation takes at least 24 hours and each fraction has to be further analysed. This increases the total analysis time to over 2 days.

8.1.1 Results and Discussion

In order to facilitate high-throughput phosphopeptide analysis using MALDI, co- matrices and matrix additives were used to improve the analyte sensitivity of existing matrices, such as α-cyano-4-hydroxycinnamic acid (CHCA), 4-chloro-a-cyanocinnamic acid (ClCCA (Jaskolla *et al.*, 2008) 2,4-Difluoro-α-cyanocinnamic (Di-FCCA) (Jaskolla *et al.*, 2009; Teuber *et al.*, 2010) as well as matrix combinations were used for the optimisation. For this purpose, 5 phosphopeptides were used to assess the analyte sensitivity (see Table S7 in the Supporting Information). Their sensitivities were compared to those obtained using the standard matrix, DHB in 1 % H_3PO_4. Furthermore, additives in MALDI-MS analysis have been used to improve matrix characteristics, such as obtaining a homogenous preparation, clean-up from salts contaminant and reduction of cationisation (Billeci & Stults, 1993; Pieles et al., 1993; Glückmann et al., 1999). Nabetani et al. showed that the addition of 3 mM $NH_4H_2PO_4$ as additive to CHCA enhanced the signal intensities of phosphopeptides in positive- and negative- ion mode (Nabetani *et al.*, 2006); this ammonium salt was used as matrix additive for all other matrices and matrix mixtures.

8.1.2 Crystal morphology and analyte stability

Despite the advances in matrix development such as optimising its proton affinity, the crystal morphology is hard to predict. However, matrix homogenity can be be influenced by additional factors such as co-matrices, solvents or additives. Homogeneous and robust crystallisation, as e.g. observed with CHCA is always of advantage. This assures a stable ion signal over an extended period of time. Signals of stable ion intensities from random positions on a spot are a necessity for high throughput MS and MS/MS analysis. After analyte co-crystallisation with the matrices/matrix mixtures, the crystal morphology of all matrices and matrix mixtures was observed using a microscope (see figure 50). With the exception of DHB, the crystal morphology of the matrices and matrix mixtures were similar to those of CHCA. The addition of 3 mM $NH_4H_2PO_4$ showed no negative influence on the crystallisation. The extent to which the

analytes were distributed on the matrix/matrix mixture can only be determined by randomly measuring the intensity of the analytes in each matrix (see figure 51).

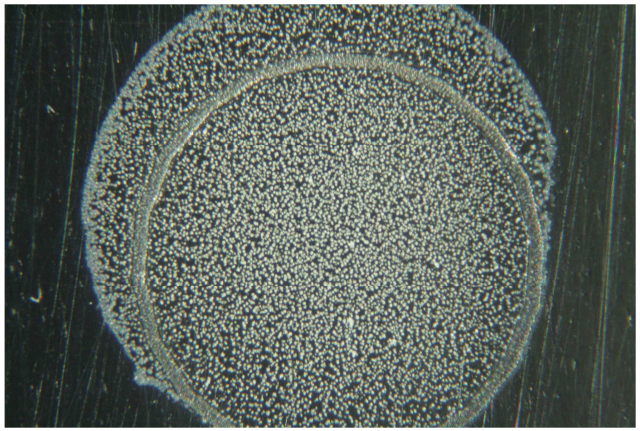

Figure 50 Crystal morphology of the FCCA/CHCA 3:1 with 3mM $NH_4H_2PO_4$.

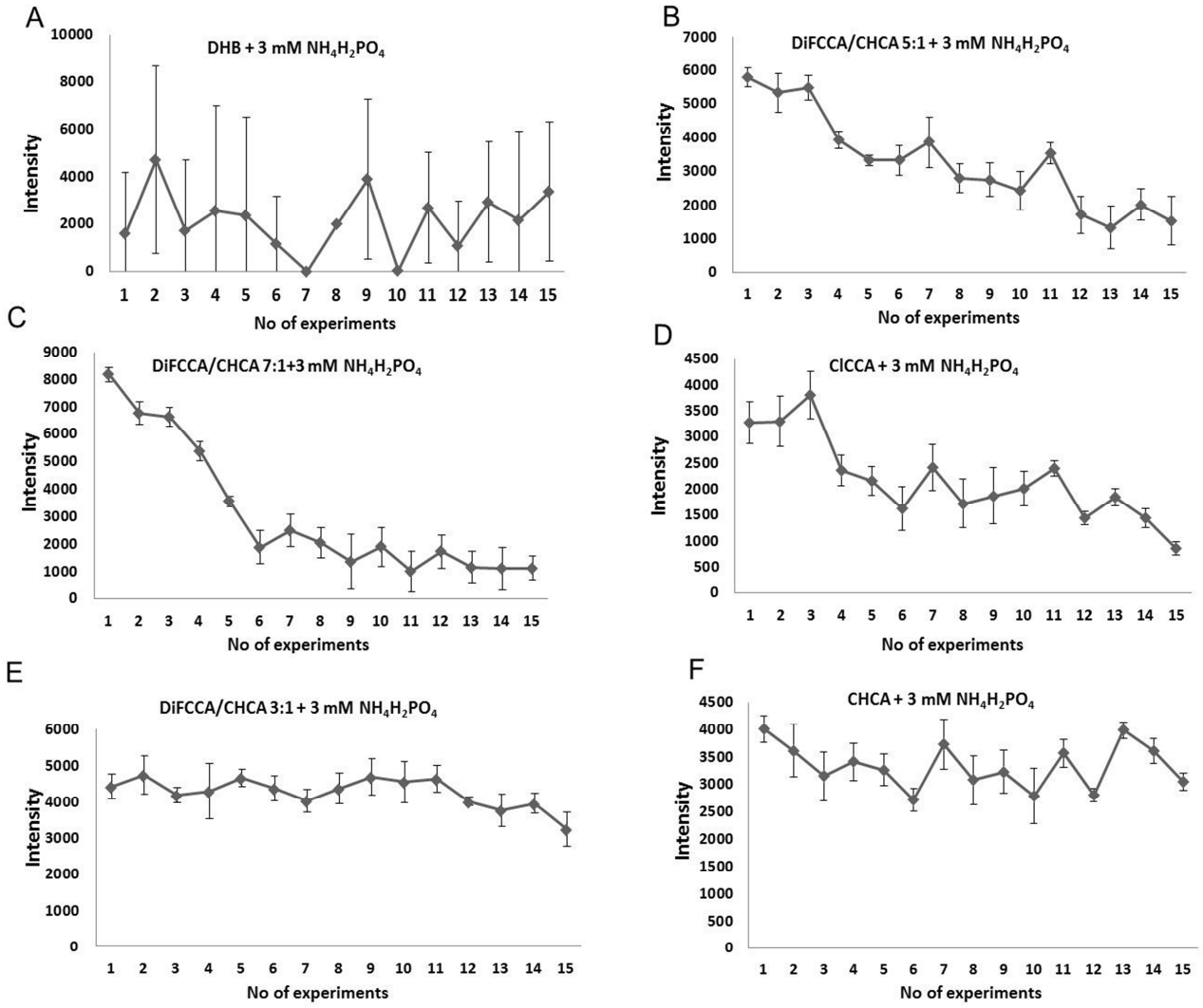

Figure 51 Signal stability, ion signal intensity of the $[M + H]^+$ of phosphorylated Angiotensin I with an increasing number of laser pulses. 15 spectra were acquired as a sum of 1000 shot per spectra.

Figure 51 shows the ion stability for phosphorylated Angiotensin I (100 fmol spotted). The intensities were obtained from 15 spectra using each matrix/ matrix mixture recorded over the course of 15 000 laser pulses (a summation of 1000 single-shots) per spectra. The signal intensity of the analytes in the matrix mixtures FCCA/CHCA 5:1- 9:1 and ClCCA rapidly decreases with increasing number of spectra acquired. This could be attributed to the high sample/matrix consumption during ablation as a result of the high laser fluence. The signal intensity of the analytes in DHB reflects its inhomogeneous crystal morphology. This matrix showed the highest position-to-position variation of all the matrices and matrix mixtures. Whenever the laser hits a hot spot, the highest signal intensity is registered. This explains the superiority of the matrix in manual measurements, since only spectra from the hot spots would be accumulated. However, the irregular signal intensity limits its use for high-throughput measurements. On the other hand, the signal intensity of FCCA/CHCA 3:1 and CHCA showed relatively stable signal intensities. This is particularly important, since more than 5 precursors are usually selected for fragmentation.

8.1.3 Analyte sensitivity

During MALDI ionisation the matrix transfers its charge to the analyte molecules. Therefore, the reduction of proton affinity of a matrix would enhance analyte intensity.The incorporation of chlorine and fluorine into the aromatic system of CHCA, resulted in low proton affinities of the resultant matrices, which in turn led to lower detection limits for analytes (Jaskolla *et al*., 2008, Teuber *et al*., 2010). However, the poor absorbance at 355 nm, ε_{ClCCA}= 140 l/M^{-1}cm-1, ε_{diFCCA} = 13 l/M^{-1}cm-1, is a limitation necessitating a high laser fluence. CHCA, which possesses good crystal morphology and is structurally related to these molecules, could be added in order to improve the absorbance while preserving their sensitivity. It was shown for ClCCA that upon crystallisation in combination with CHCA, improved absorbance is observed. However, the matrix compound competes in protonation process, leading to a decrease in analyte sensitivity (Jaskolla *et al*., 2008). On the other hand, depending on the amount of CHCA, the combination with DiFCCA improved the overall absorbance and sensitivity. Though the absorbance was not directly measured, the lower laser fluences used for acquisition of the mass spectra showed that the absorbance was enhanced. In general, each MS acquisition method was optimised for each matrix or matrix

mixture. The laser fluences varied from 4000 au (CHCA +3 mM $NH_4H_2PO_4$) to 5500 au (FCCA/CHCA 9:1 + 3 mM $NH_4H_2PO_4$). The frequency-tripled Nd-YAG laser on the 4800 MALDI Analyzer has a maximum laser insity of 7000 au.

High analyte signal intensities for MS measurement using 3 different analyte concentrations (25, 100 and 250 fmol) were obtained with FCCA/CHCA 5:1- 9:1+3 mM $NH_4H_2PO_4$, ClCCA+3 mM $NH_4H_2PO_4$, FCCA/CHCA 3:1+3 mM $NH_4H_2PO_4$ and CHCA+ 3 mM $NH_4H_2PO_4$ respectively (see figure 52 A and B). Compared to the same matrices and matrix mixtures without 3 mM $NH_4H_2PO_4$, the signal quality is substantially increased (see figure 53 A and B). A summary of the analyte intensities for selected matrices is represented in the figure 54. The improvement in analyte intensities can be attributed to the lower proton affinity of the DiFCCA matrix allowing for a more efficient proton transfer from the $[DiFCCA + H]^+$ matrix ions to the analyte and the improved absorbance from the CHCA matrix. Furthermore, it can be assumed that the poor detection of phosphopeptides in MALDI is a result of poor incorporation of the analytes into the matrix because of high phosphate group affinity to metal surface. When phosphopeptides are spotted on the metal target, the phosphate groups bind strongly to the target surface, thereby preventing efficient incorporation into the matrix. The addition of phosphate to matrix solutions probably saturates the metal surfaces, therefore increasing the incorporation efficiency of subsequently added phosphate analytes into the matrices. For this reason, it can also be assumed that the better sensitivity of DHB in 1 % phosphoric acid is due to the saturation of the target surface with phosphoric acid or inorganic phosphate. However, the addition of 3 mM $NH_4H_2PO_4$ led to a slight increase in required laser fluence for all investigated matrices/matrix mixtures.

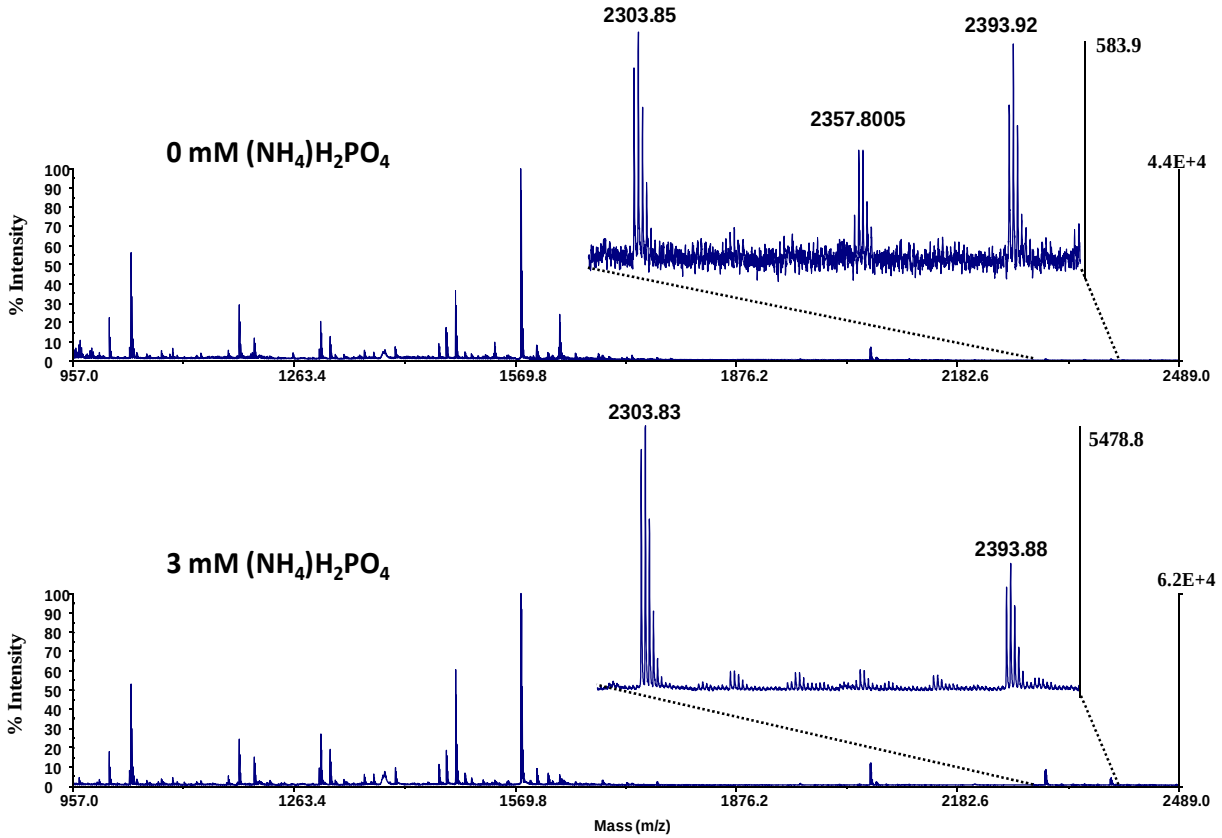

Figure 52 Mass spectra showing the effect of $(NH_4)H_2PO_4$ on the phosphopeptide signal at m/z 2303.94.

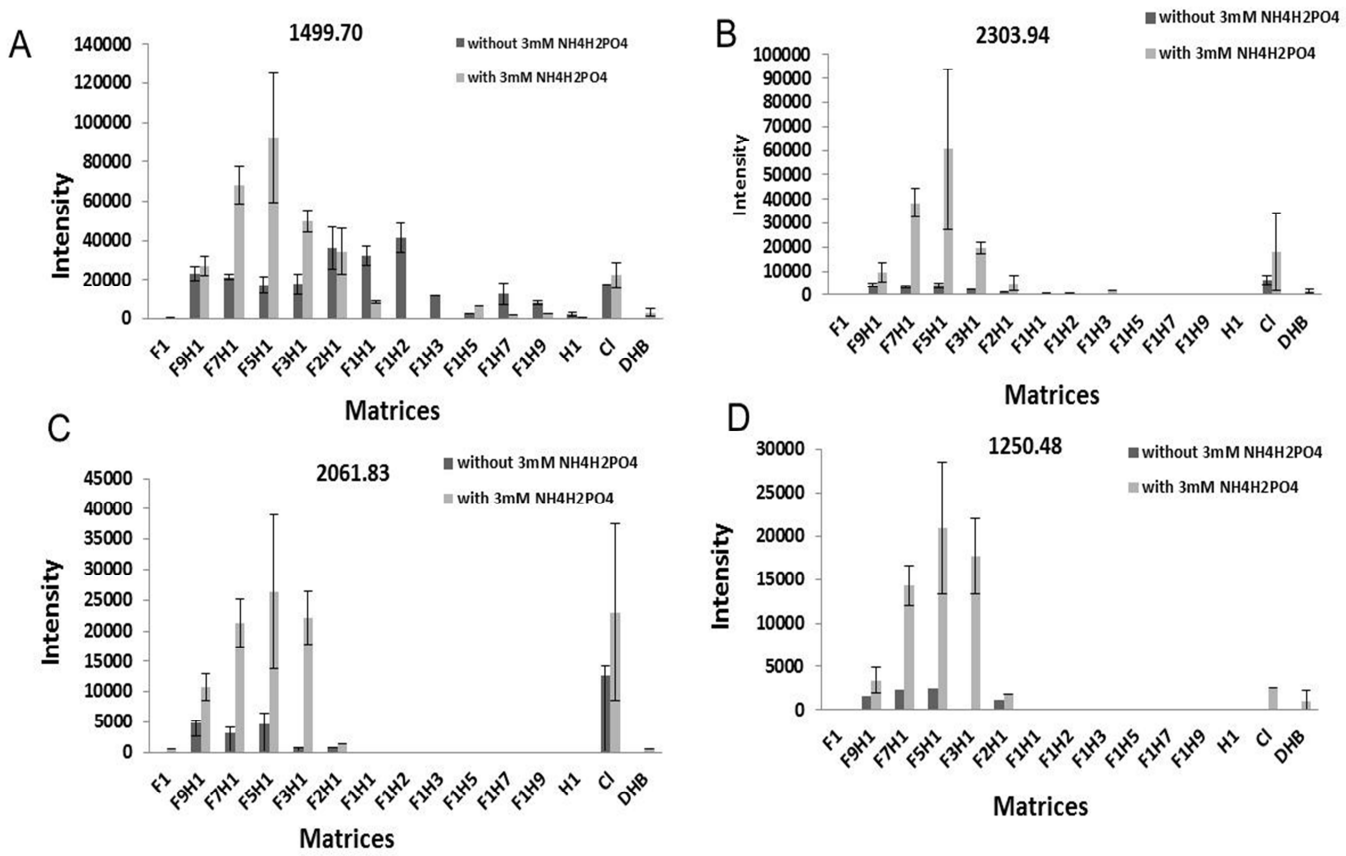

Figure 53 Summary of the effect of (NH4)H2PO4 on matrices and matrix mixtures for 25 fmol of a (A) di-phosphorylated (VADPDHDHTGFLpTEpYVATR) and (B) a mono-phosphorylated peptide (VRFPpSHFSSDLK).

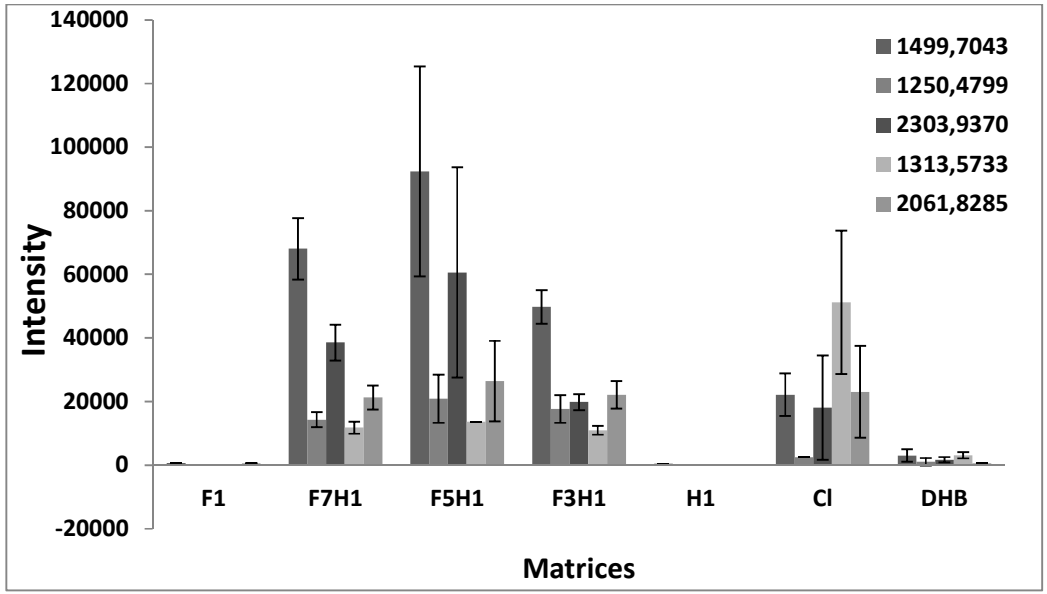

Figure 54 Comparison of the analyte ion intensities of different matrices. 25 fmol of each peptide was used DHB was dissolve in 1% phosphoric acid. All the other matrices and matrix mixture contained 3 mM $NH_4H_2PO_4$

8.1.4 MALDI MS/MS

To map modifications in a protein, additional information is always needed for the correct identification and localization of these modifications, especially if more than one posttranslational modification (PTM) is present. MS/MS is a prerequisite for the identification of unknown phosphorylation sites in a protein. Peptide fragmentation in MALDI TOF/TOF is influenced by many factors that determine the internal energy of analytes such as matrix starting velocities, the applied acceleration voltages, the collision energy and the laser energy/fluence used. To determine the extent to which the matrices and matrix mixtures affect the fragmentation, 100 fmol of each test peptide was mixed with matrices and matrix mixtures that yielded the highest intensities MS mode. The MS/MS acquisition method was

optimised for each matrix/matrix mixture. Due to the high laser fluence (approximately 6000 au) observed during MS/MS method optimisation, the FCCA/CHCA 7:1 with 3mM $NH_4H_2PO_4$ mixture was omitted. Fragment ion mass spectra obtained using FCCA/CHCA 5:1 see figure 55 C), ClCCA, DHB in 1 % H_3PO_4 (see figure 55 A) were of low resolution, low intensity and poor signal-to-noise ratio (S/N). Furthermore, there were less intense immonium ion signals in the lower mass region, which aid in the identification of the amino acids present in the sequence.

On the other hand, fragment ions generated using FCCA/CHCA 3:1 (see figure 55 B), were very similar to those generated with CHCA. However, in most cases, they exhibited higher intensities and S/N ratios. For some peptides, almost a complete y-ions series could be identified. Additionally, lower laser fluence (4900 au) was required for the FCCA/CHCA 3:1 mixture to acquire high quality spectra. Higher laser fluences led to increased ablation, and thus poor spectrum quality. Increased ablation is a limitation for high-throughput MALDI-based proteomic analysis, since it is desirable to acquire a minimum number of MS/MS spectra per spot (typically $\geq$ 5). Furthermore, higher laser fluences lead to an increase of unwanted fragmentation during MS and MS/MS, thereby reducing the quality of the mass spectra. A summary of the MS/MS scores for 6 phosphopeptides as calculated by the Biotools software is shown in figure 56.

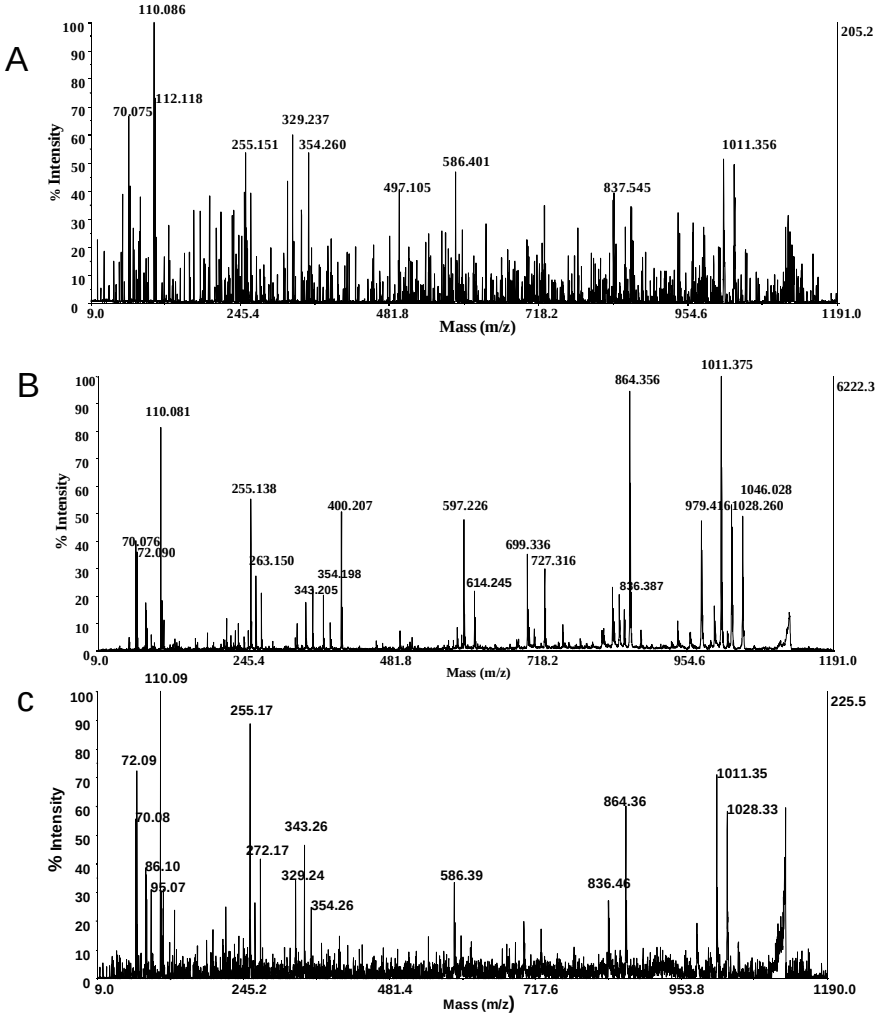

Figure 55 MS/MS spectrum of the peptide with m/z =1126.3 Da obtained using 3 matrices. (A) spectrum was obtained using DHB, (B) with the matrix mixture FCCA/CHCA 3:1 + 3 mM $NH_4H_2PO_4$ and (C) was obtained with the FCCA/CHCA 5:1 + 3 mM $NH_4H_2PO_4$.

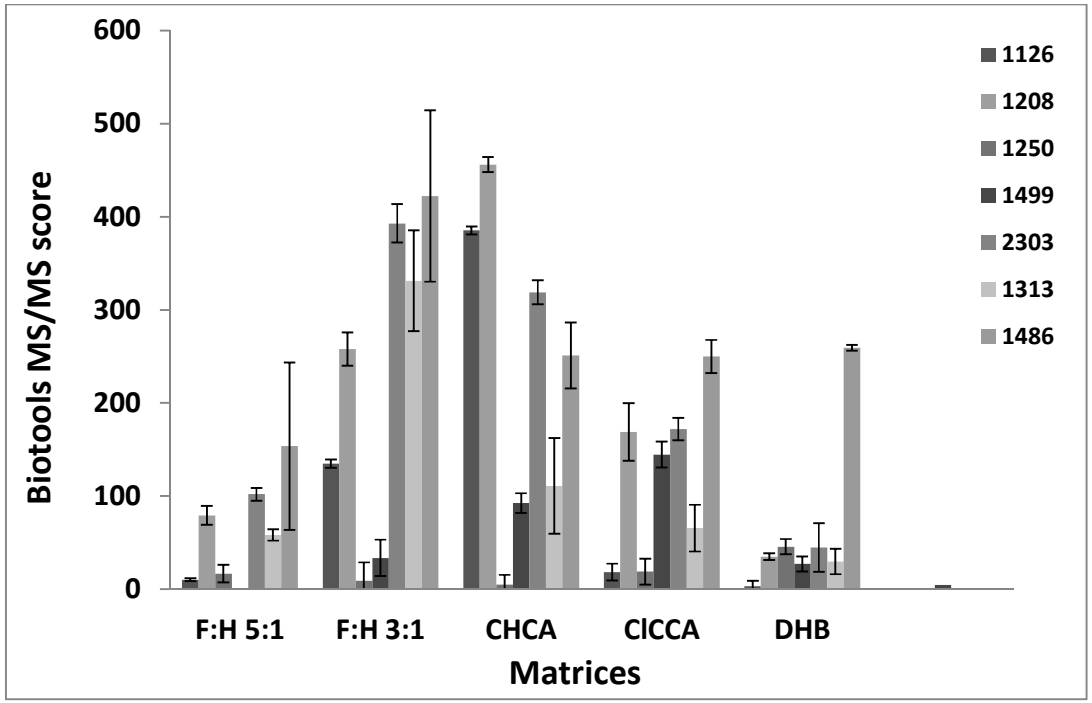

Figure 56 Comparison of the fragmentation patterns using Biotools 2.2 (Bruker Daltonics, Bremen). DHB was dissolve in 1% phosphoric acid. All the other matrices and matrix mixture contained 3 mM $NH_4H_2PO_4$.

Based on MS results, it can be concluded that FCCA/CHCA 7:1, FCCA/CHCA 5:1, ClCCA, FCCA/CHCA 3:1 with 3 mM $NH_4H_2PO_4$ can be used in the positive ion mode for MS measurements of phosphopeptides. These matrices have more homogeneous morphologies than the standard matrix DHB. Moreover, the signal intensities of these peptides are higher than those obtained using the standard matrix. In MS/MS experiments, high intensities fragment ions were obtained with FCCA/CHCA 3:1 + 3 mM $NH_4H_2PO_4$ at relatively lower laser fluence than with DHB in 1 % H_3PO_4, FCCA/CHCA 5:1 + 3 mM $NH_4H_2PO_4$ and ClCCA + 3 mM $NH_4H_2PO_4$.

Therefore, for high-throughput nLC-MALDI-TOF/TOF analysis of phophopeptides FCCA/CHCA 3:1 + 3 mM $NH_4H_2PO_4$ was chosen for further analysis.

8.2.1 Biphasic pre-column

The establishment of a matrix, with improved sensitivity compared to that of the standard matrix DHB, allowed us to establish a high-throughput MALDI method for the analysis of phosphopeptides. This method is based on the online enrichment and separation of both the flow through and the enriched phosphopeptides using a biphasic pre-column containing TiO_2/C_{18} material. The enrichment of phosphopeptides on TiO_2 has shown to improve their signal intensities. Most of these enrichment processes are carried out in offline mode. Pinkse et al. introduced an automatic method for the online enrichment and analysis of phosphopeptides via LC-ESI MS/MS. In this method, three short columns of C_{18}, TiO_2 and C_{18} material respectively, were combined. This strategy is particularly advantageous because column switching is not necessary and it allows a full characterisation of the flow-through (non-phosphorylated peptides) and TiO_2-eluate (phosphorylated peptides). For our experiments, both C_{18} and TiO_2 were packed in one column (biphasic pre-column) as shown in figure 57. This procedure, which is similar to multidimensional protein identification technology (MudPIT), reduced the number of connections and possible void volumes in the system. Digested samples are loaded onto the biphasic pre-column, phosphopeptides are trapped on the TiO_2 beads and the non phosphorylated peptides on the C_{18} materials. During the first gradient, the peptides trapped on the C_{18} materials are separated on the analytical column (in-house packed 120 cm × 100 µm C_{18}), mixed with CHCA and spotted on a MALDI target. Prior to the second gradient, phosphopeptides trapped on the TiO_2 materials are eluted to the C_{18} material using ammonium bicarbonate (pH 10) injected via the sample loop. The phosphopeptides are subsequently separated in the second gradient using the same analytical column, mixed with the novel matrix mixture and spotted on a MALDI target. This online enrichment procedure facilitates high reproducibility, good enrichment efficiency, and the analysis of both flow through. Furthermore, the use of 2 different matrices for the analysis is possible. In addition, due to the high pH range stability of the C_{18} materials, it is possible to analyse many samples using the same biphasic pre-column.

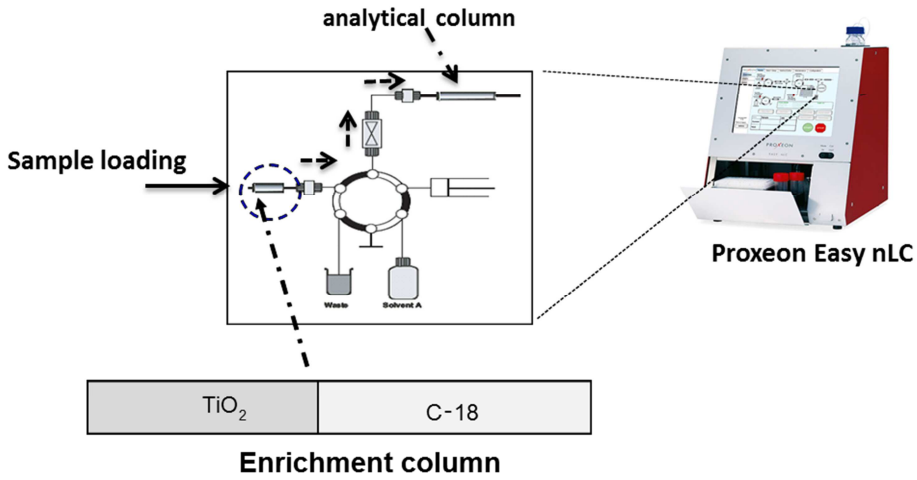

Figure 57 Sample loading setup.

8.2.2 Evaluation of non-phosphorylated acidic peptides excluders

A limitation to the efficiency of TiO_2 enrichment strategies is given by the non-specific binding of primarily multiply acidic residues (i.e., Glu and Asp) in peptides. In an attempt to reduce the rate of nonspecific binding, DHB, glycolic acid and TFA were included in the sample and loaded on the biphasic pre-column. Using 4 standard phosphopeptides in BSA digest of 20-fold higher concentration as a semi-complex background, the loading and eluting conditions were optimised. Even though, in the case of DHB, the biphasic pre-column was washed with four times the injected sample volume, concentrations above 60 mg/ml resulted in extensive tailing on the first gradient (see figure 58). This in turn resulted in inhomogeneous crystallisation of the matrix (figure 50), leading to irreproducible measurements and an increase in the required laser fluence. With concentrations below 40 mg DHB/ml, only a few spots (10 spots) were affected. Also, the overall intensities of the eluted phosphopeptides decreased, when

using DHB in the enrichment buffer, an effect, which has also been reported by several groups (Sugiyama *et al*., 2007; Bodenmiller *et al.*, 2007).

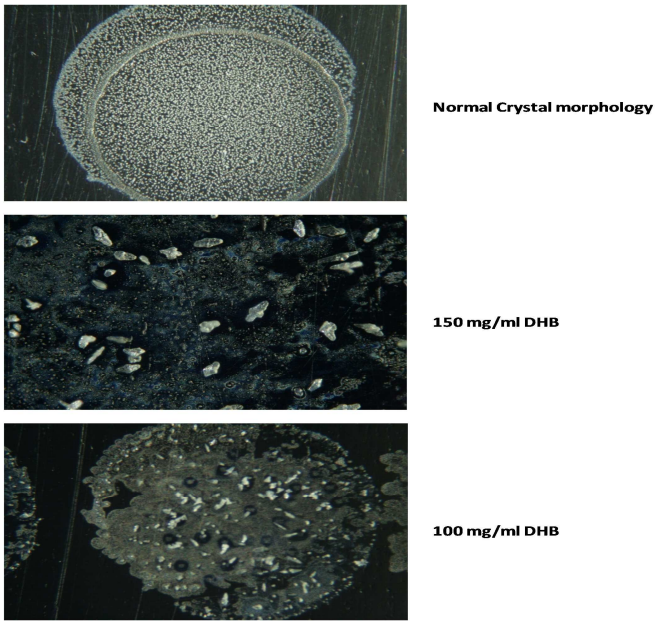

Figure 58 The effect of different concentrations of DHB in the loading solution on the crystal morphology for the FCCA/CHCA 3:1 with 3mM $NH_4H_2PO_4$ matrix mixture.

On the other hand, 1 M glycolic acid and 5 % TFA showed little or no effect on the matrix crystallisation. In general, all three acid excluders reduced unspecific binding, thus increasing the intensity ratio of the phosphopeptides. 5 % TFA showed no negative influence on the crystallisation, less non-phosphorylated peptides were identified (see figure 59 A), yielded the highest intensities and signal-to-noise ratios of most of the phosphorylated peptides (see figure 59).

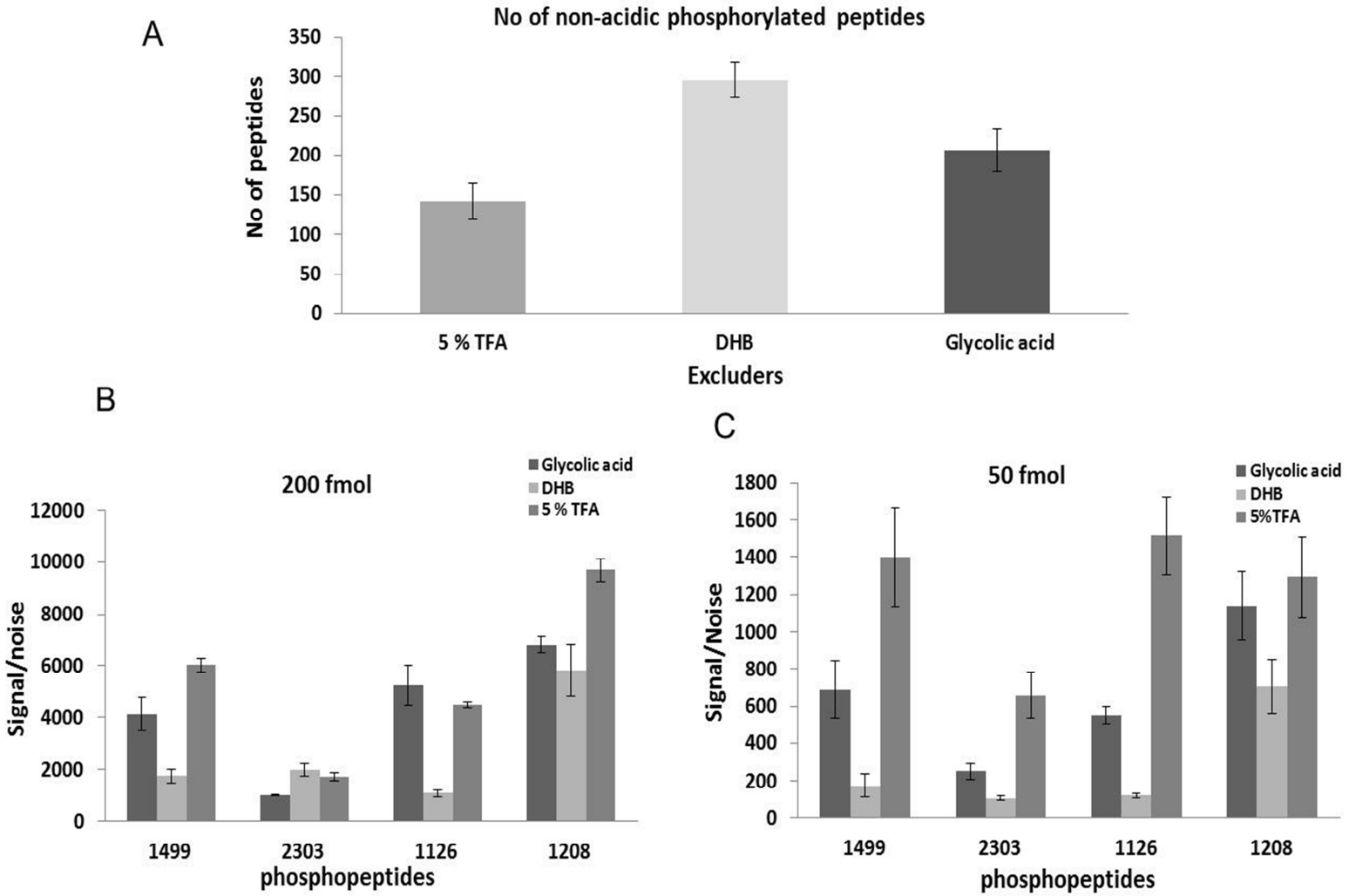

Figure 59 A shows the number of acidic non-phosphorylated peptides identified. B and C show Comparison of the signal intensity of individual phosphopeptides after loading with Glycolic acid, DHB and 5 % TFA.

To assess the affinity and capacity of the titanium beads, between 250 fmol and 50 fmol of standard phosphopeptides in a 20-fold excess of BSA digest were loaded on the column. After separation and precursor mass measurement, the intensities of the distribution of the phosphopeptides in both flow through (gradient 1 and 2) were plotted against the spots collected during the case of both gradients. Figure 60 shows that even at 250 fmol all the phosphopeptides were effectively trapped on the TiO_2 biphasic pre-column, indicating the high selectivity of TiO_2 in trapping phosphopeptides. All phosphopeptides were identified in the second gradient, however with varying intensities and S/N ratios.

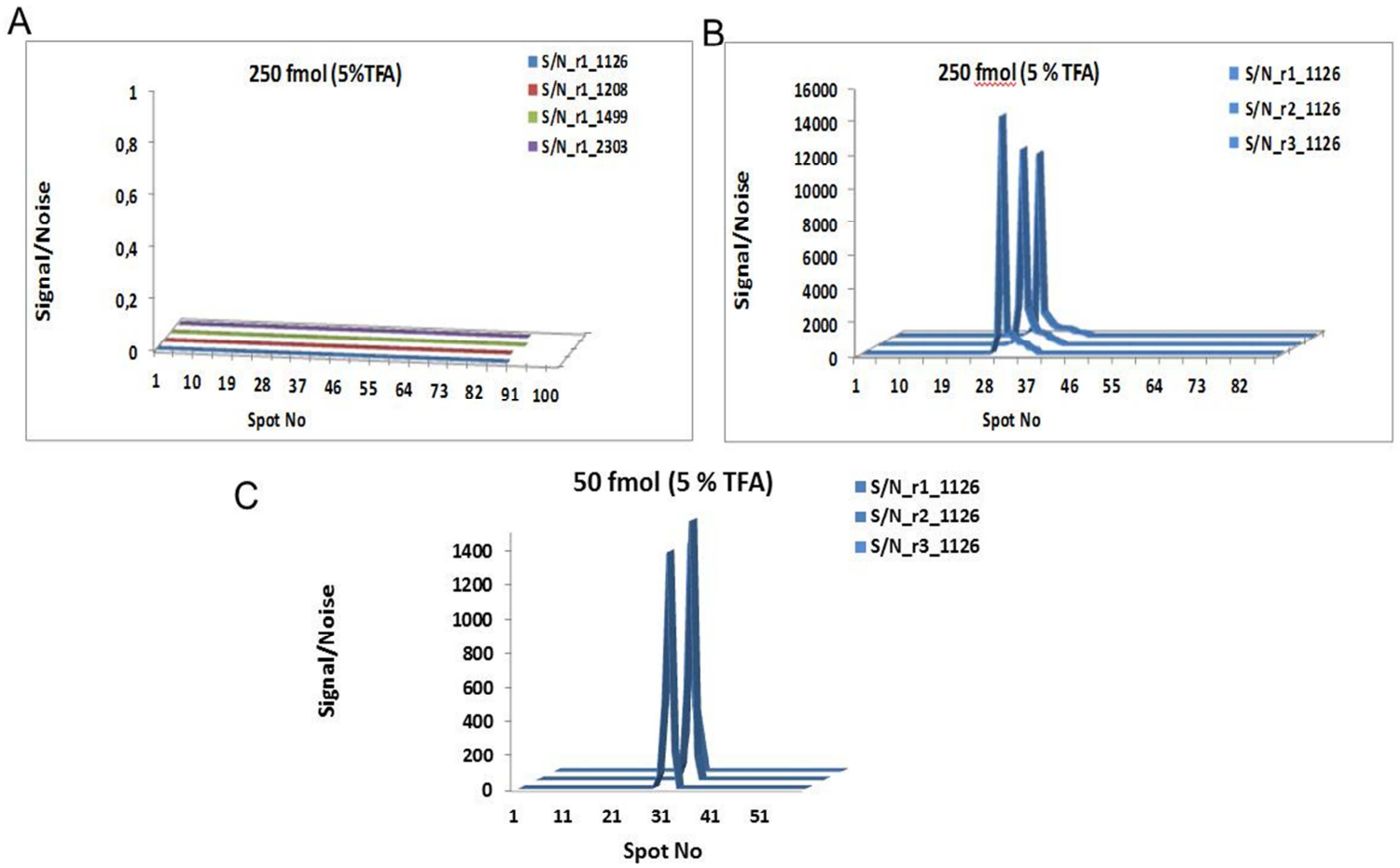

Figure 60 Plot of the Signal/Noise against spot number of the phosphopeptide loaded on the biphasic pre-column and separated on the EASY-nLC. (A) Shows the flow through during sample loading. (B and C) show reproducible base peak chromatogram of the phosphopeptides trappped on the TiO_2 beads.

8.2.3 Evaluation of the online enrichment using α/β-casein

To evaluate the applicability of the matrix mixture and the biphasic pre-column, a tryptic digest of α/β-casein, was analysed. A direct analysis of 500 fmol of these proteins, using 5 % TFA as acid excluder resulted in the detection of 14 phosphopeptides representing 8 different phosphorylation sites. The MGF files were searched against the DB-standard database containing 24 proteins. These peptides were considered significant only if they appeared in at least 2 of the triplicate MS and MS/MS experiments and result in Mascot scores above the identity threshold ($p < 0.5$) (see supplementry Table 6). Table 17 shows a summary of all phosphopeptides identified. The identification of the phosphopeptide, YLGEYLIVPNpSAEER (1832.83 Da) (see figure 61) was done manually. This peptide represents a new sequence variant of the alpha-S1 casein in the region 104–119 and was first described by Larsen and his group (Larsen *et al*., 2005). In addition, a semi-tryptic phosphopeptide RELEELNVPGEIVEpSL (1905.86 Da) from β-casein representing amino acids 16-31 was identified. The identification of these peptides was only possible because of their isolation from the non-phosphorylated peptides. Out of the 14 phosphopeptides identified from α-S1casein, only 2 peptides (YKVPQLEIVPNpSAEER and VPQLEIVPNpSAEER) were detected in the first and second gradient. This indicates that, no matter how selective the TiO_2 is, some peptides might still escape enrichment while some might not be eluted from the beads. Though 5 % TFA significantly reduced the amount of acidic non-phosphorylated peptides on the TiO_2 material, it did not completely prohibit the binding of all acidic non-phosphorylated peptides.

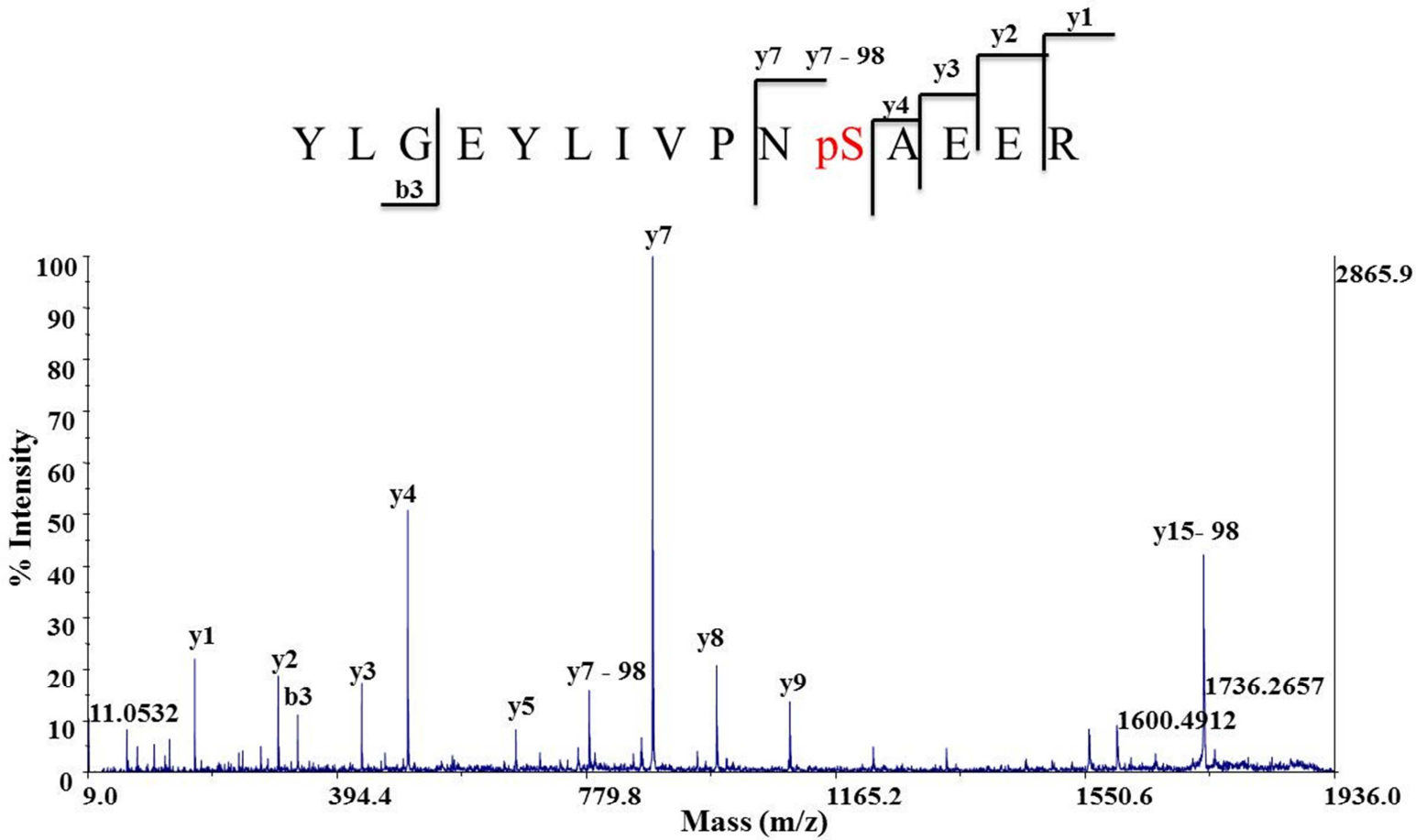

Figure 61 A new sequence variant of the alpha-S1 casein in the region 104–119 (YLGEYLIVPN**pS**AEER, 1832.83 Da). This peptides is not present in Mascot, thus was identified manually.

Table 17 Overview of the identified phosphopeptides generated from α/β-casein.

*phosphopeptides found in both the first and the second gradient

§peptide not identified by mascot search

Peptides	M+H	Ion score
DIGSESTEDQAMEDIK (α-S1)	1927.64	36
YKVPQLEIVPNSAEER (α-S1)	1951.93*	101
YKVPQLEIVPNSAEERLHSMK (α-S1)	2548.27	74
VPQLEIVPNSAEER (α-S1)	1660.79*	104
YKVPQLEIVPNSAEERLHSMK (α-S1)	2564.25	24
VPQLEIVPNSAEERLHSMK (α-S1)	2273.07	23
IEKFQSEEQQQTEDELQDK (β)	2431.98	53
FQSEEQQQTEDELQDK (β)	2061.81	68
NAVPITPTLNREQLSTSEENSKK (α-S1)	2716.15	25
RELEELNVPGEIVESLSSSEESSITR (β)	3122.3	63
IEKFQSEEQQQTEDELQDK (β)	2432.1	42
YLGEYLIVPNSAEER (α-S1)	1832.83§	-
RELEELNVPGEIVESL (α)	1905.86	87

8.2.4 ENL protein complex

To demonstrate that this method can be used for the study of phosphorylation sites in a real sample, affinity purified ENL complex was digested with trypsin. The peptide mixture was then loaded onto the custom-made biphasic pre-column, separation and analyses were performed as described above. Three replicate runs were performed; the generated Mascot generic format files were combined and searched against a Swissprot database. For data interpretation, only peptides that appeared in at least 2 analyses with Mascot score above the homology threshold ($p > 0.3$) were taken into consideration. A total of 352 peptides were matched to 82 proteins (see supplementary excel sheet 1)

Due to the frequent loss of phosphate groups within the peptide sequence in MS/MS experiments, assigning phosphorylation sites is rather difficult. Phosphorylation sites of peptides containing just one possible modifications site can easily be obtained from the Mascot search results. However, difficulties arise if the sequence contains more than one possible modification site. Therefore, the phosphorylation sites of the identified peptides were manually verified. For this purpose, the data explorer (AB Sciex) was used to generate theoretical fragment ion mass spectra from the sequence with the assumed phosphorylation site. These fragments were then used to label the fragment of the spectrum in order to determine the best possible fit. For example the phosphopeptide LSFSDSESDNSADSSLPSR with m/z 2080 Da has 8 serine residues. The loss of 98 Da during fragmentation from phosphoserine results in an intensive signal and is usually an indicator of these phosphopeptides. However, DeGore and Qin (DeGore & Qin, 1998) showed that in some phosphoserine-containing peptides, the 98 Da loss is not dominant, nevertheless still indicates the present of a phosphate group in the peptide backbone (see figure 62). An extended y- ion series is observed, but the series partially disappears after y-15. The lack of b- and y-type ions with a deficit of 98 Da in the region between y1-y15 (see figure 62) indicated that the serine residues in this section of the sequence were not modified. This section of the sequence contains 6 serine residues (-DSESDNSADSSLPSR). On the other hand the presence of the b7-98 was indicative of a phosphorylated serine in this section of the spectrum. A zoom between the regions 1740-1878 Da revealed a poorly resolved fragment, which represent the y17-ion, with a 98 Da deficit. The loss of H_3PO_4 from the y17 ion indicated that ser2 is most likely carrying the phosphate group. Another example is the phosphopeptide, ETKLESTSPK, with m/z 1198.55. This

peptide was identified as a peptide phosphorylated at serine and it possess 2 serines, ser6 and ser8, which can be modified. The presence of y3-98 or b6-98 would give a clear hint of the modified phosphoserine. Considering ser8 as the phosphorylation site, the absence of the y3or y4 and the subsequent loss of a phosphate group indicated that that ser8 was not phosphorylated (see figure 63 A). On the other hand, assuming that ser6 is phosphorylated, the presence of the b6-98 ions is indicative of a phosphorylated serine (see figure 63 B). Therefore it can be concluded that ser6 is phosphorylated

In general, 352 peptides were uniquely identified, of which 18 were phosphopeptides, covering 15 different phosphorylation sites. 13 of these phosphopeptides, containing 9 phosphorylations site were matched to the ENL protein. 9 phosphorylations sites have been predicted for this protein but never been identified. Six of the 9 predicted and 3 novel phosphorylation sites were identified. All identified peptides were phosphorylated at the serine residues and amongst these only one was a di-phosphorylated. A list of all identified phosphopeptides is presented in table 18.

The function and the role of the phosphoserine modification on ENL have not fully been investigated but it is known that Eleven-nineteen-leukemia protein (ENL) and 3 other proteins, AF4, AF9, and AF10 are the four most frequent fusion protein partners of Mixed-Lineage Leukaemia (*MLL*) gene in pediatric-, adult- and therapy-related acute leukemias (Meyer *et al*., 2006). In a recent publication, Bitoun et al. (Bitoun *et al*., 2007) demonstrated that AF4 functions as a positive regulator of Pol II transcription elongation factor b (P-TEFb) kinase and in complex with MLL fusion partners AF9, ENL and AF10, it acts as a mediator of histone H3-K79 methylation by recruiting Dot1 to elongating Pol II. Furthermore, it was mentioned that these pathways are interconnected and tightly regulated by the P-TEFb-dependent phosphorylation of AF4, AF9 and ENL, which controls their transactivation activity and/or protein stability.

In summary, the identification of these phosphorylation sites shows that the method presented here works well for real sample.

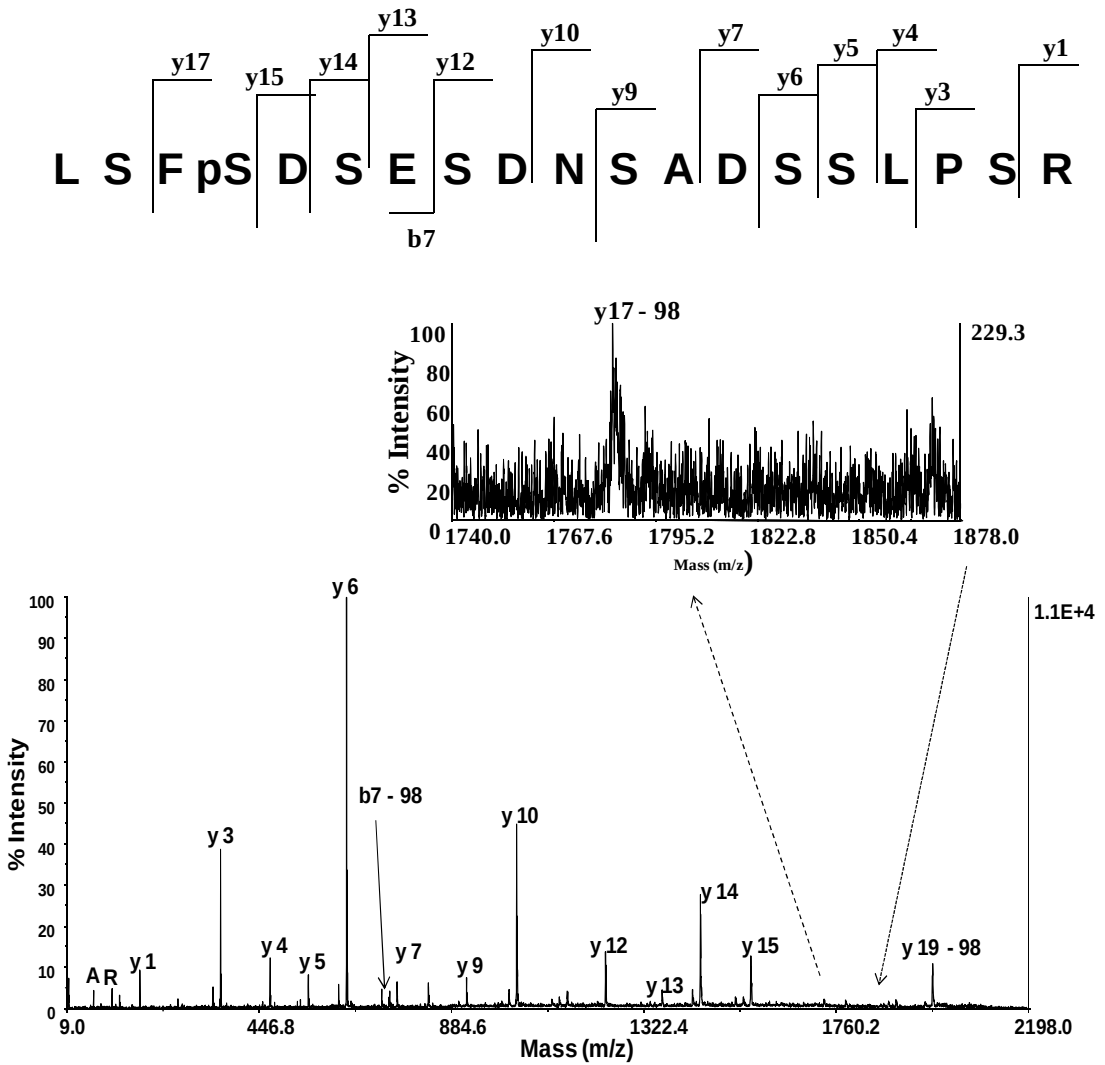

Figure 62 Fragment spectrum of the peptide with mass 2080.743 (LSFpSDSESDNSADSSLPSR).

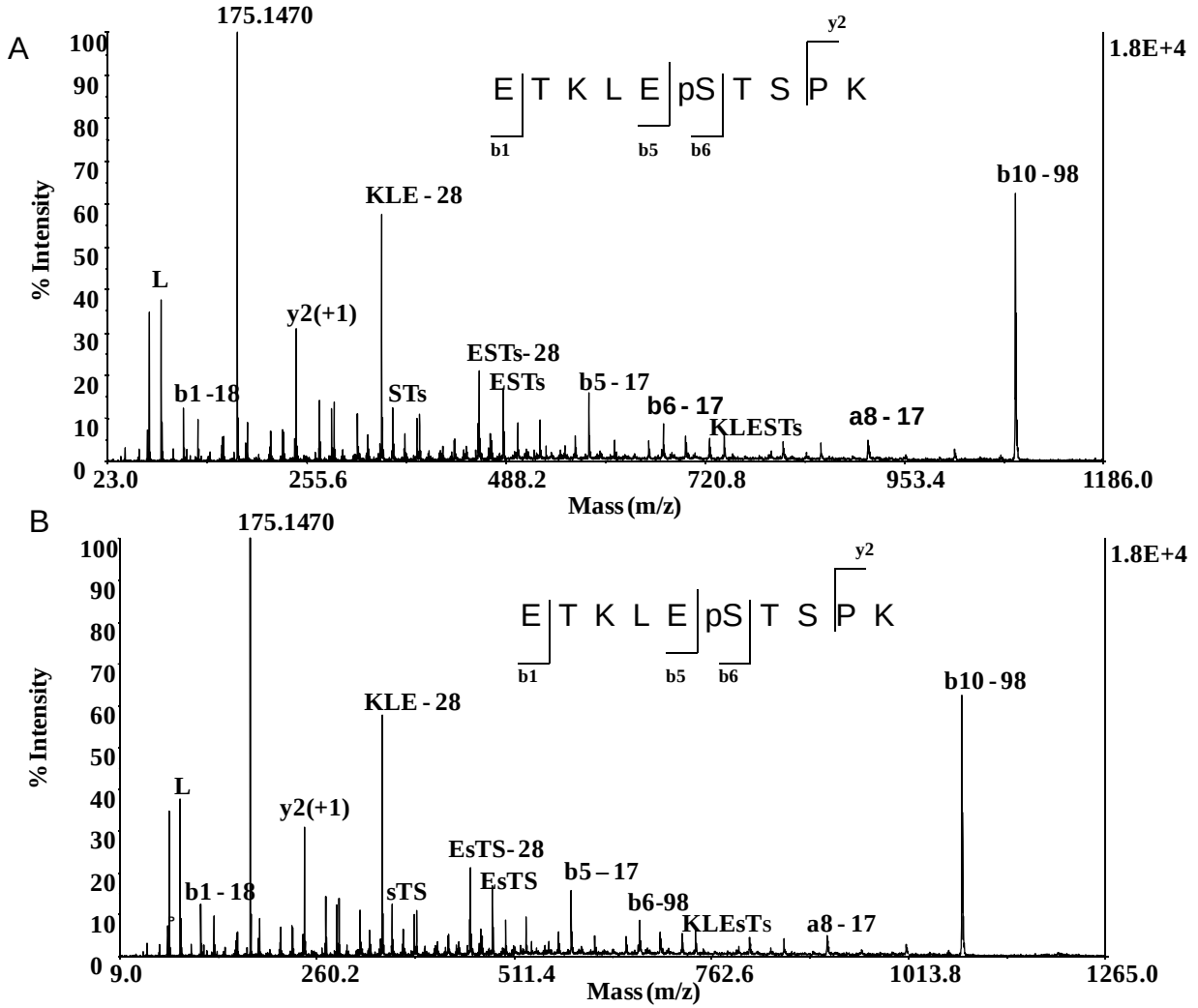

Figure 63 Fragment spectrum of the peptide with m/z of 1197.54 (ETKLEpSTSPK).

Table 18 Phosphopeptides identified. All novel phosphorylation sites are underlined.

Mass	Peptide sequence	Peptide score	Peptide start	Peptide end
852,35	SAPGT**pS**PR	37,87	310	317
1199,52	ETKLEST**pS**PK	46,52	260	269
1228,44	**pS**PESCSKPEK	54,81	475	484
1384,57	R**pS**PESCSKPEK	54,01	474	484
1405,67	RPATAD**pS**PKPSAK	69,02	286	298
1783,86	VSGRR**pS**PESCSKPEK	44,04	470	484
1863,75	ALEVEESN**pS**EDEASFK	135,56	353	368
1943,69	ALEVEEpSN**pS**EDEASFK	74,63	353	368
1984,96	LEST**pS**PKGGPPPPPPPPPR	113,45	263	281
1991,85	KALEVEESNSEDEA**pS**FK	88,48	352	368
2080,85	LSF**pS**DSESDNSADSSLPSR	142,7	437	455
2343,09	ETKLEST**pS**PKGGPPPPPPPPPR	58,49	260	281
3571,59	LSF**pS**DSESDNSADSSLPSREPPPPQKPPPPNSK	66,74	437	469

8.3 Conclusion

Many cellular processes are regulated by reversible phosphorylation of proteins on serine, threonine, and tyrosine residues. A better understanding of these biological signalling networks requires global identification of phosphoproteins and their phosphorylation sites. Although the analysis of protein phosphorylation has been a rapidly evolving field, the goal of phosphoproteome analysis for cells and tissues remains an analytical challenge. MALDI is one of the most commonly used MS-based methods for phosphoproteomics. However, its use for high-throughput phosphorproteomics has been limited due to the lack of optimal matrices. DHB has proven to be the matrix of choice in this field. Nevertheless, as a result of the inhomogenous morphology of this matrix, automation has been difficult for instruments lacking a crystal positioning system. It was shown that crystallisation can be positively affected with co-matrices or matrix additives, moreover, the sensitivity of matrices for phosphopeptide analyses can be enhanced, while preserving the crystal morphology. Compared to DHB, which is known to exhibit high sensitivities for phosphopeptides, the combination of DiFCCA and CHCA in a 3:1 ratio in 3 mM $NH_4H_2PO_4$ led to an increase in the intensities of phosphopeptides. This matrix mixture exhibited higher absorbance and more uniform morphology, therefore more reproducible analyte intensities with lower standard deviations and less laser fluences. In addition, this matrix mixture showed better fragmentation.

Based on this novel matrix mixture, a high-throughput MALDI-based method for the analysis of phosphorpeptides was established. This method is based on trapping and analysing both the flow through during enrichment and the enriched phosphopeptide using a custom-packed biphasic pre-column. The applicability of this method was tested using standard phosphopeptides spiked to a 20-fold excess of a tryptic BSA digest, a tryptic digest α/β-casein and an affinity-purified ENL protein complex. The advantage of this biphasic column is the ability to analyse both flow throughs without column switching and that the pre-column is ready to be used immediately after each gradient. The use of 5 % TFA greatly enhances the exclusion of acidic non-phosphorylated peptides. This methodology resulted in a strong increase in the selectivity of purification of phosphorylated peptides from complex mixtures of non-phosphorylated and phosphorylated peptides. In conclusion, the mass spectrometric identification and characterisation of PTMs is still an extremely difficult task since 100 % protein sequence coverage is still not possible.

9 General Summary

Top-down and bottom-up approaches are the general methods used to analyse proteomic samples today, however, the bottom-up approach has been dominant in the last decade. Establishing a bottom-up method involves not only the choice of adequate instruments and the optimisation of the experimental parameters, but also choosing the right experimental conditions and sample preparation steps. LC-ESI MS/MS has widely been used in this field due to its advanced automation. The primary objective of the present study was to establish a sensitive high-throughput nLC-MALDI MS/MS method for the identification and characterisation of proteins in biological samples. The method establishment included optimisation and validation of parameters such as the capillaries in the HPLC systems, gradient slopes, column temperature, spotting frequencies or the MS and MS/MS acquisition methods. The optimisation was performed using two HPLC-systems (Agilent 1100 series and Proxeon Easy nLC system), three spotters and the 4800 MALDI-TOF/TOF analyzer. Furthermore, samples preparation protocols were modified to fit to the established nLC-MALDI-TOF/TOF-platform.

The potentials of this method was demonstrated by the successful analysis of complex protein samples isolated from lipid particles, pre-adipocytes/adipocytes tissues, membrane proteins and proteins pulled-down from protein-proteins interaction studies. Despite the small amount of proteins in the lipid particles or oil bodies, and the challenges encountered in studying such proteins, 41(6 novel + 14 mammal specific + 21 visceral specific) proteins were added to the already existing proteins of the secretome of human subcutaneous (pre)adipocytes and 6 novel proteins localised in the yeast lipid particles. Protein-protein interaction studies present another area of application. Here the analytical challenges are mostly due to the loss of binding partner upon sample clean-up and to differentiate from non-specific background. Novel interaction partners for AF4•MLL and AF4 protein complex were identified.

Furthermore, a novel sample protocol for the analysis of membrane proteins, based on the less specific protease, elastase, was established. Compared to trypsin, a higher sequence coverage and higher coverage of the transmembrane domains were achieved. The use of this enzyme in proteomics has been limited because of its non specific cleavage. However, from the results obtained in these studies, elastase was found to cleave preferentially at the C-terminal site of the amino acids AVLIST. The advantage of the established protocol over conventional protocols is that the same enzyme can be used for shaving of the soluble dormains of intact proteins in membranes and

the digestion of the hydrophobic domain after solubilisation. Furthermore, the solvents used are compatible with the nLC-MALDI method setup. In addition, it was also shown that for less specific enzymes, a higher mass accuracy is required to reduce the rate of false positive identifications, since current search engines are not perfectly adapted for these types of enzymes.

A brief statistical analysis of the MS/MS data obtained from the LC-MALDI-TOF/TOF system showed that for less specific enzymes, under high-energy collision conditions, approximately 43 % of the fragment ions could not be matched to the known y- b- type ions and their resultant internal fragments. This limitation greatly influenced the search results. However, this limitation can be overcome by modifying the N-terminal amino acids with basic moieties such as TMT.

The use of elastase as a digestion enzyme in proteomic workflow further increased the complexity of the sample. Therefore, orthogonal multidimensional separation was necessary. Offgel-IEF was used as the separation technique for the first dimension. Here peptides are separated according to the p*I*. However, the acquired samples could not be loaded to the nLC due to the high viscosity of the concentrated samples when using the standard protocol. In order to achieve compatibility of the Offgel-IEF to the nLC-MALDI-TOF/TOF-platform, the separation protocol of the Offgel-IEF was modified by omitting the glycerol, which was the cause of the viscous solution. The novel glycerol free protocol is advantageous over the conventional method because the samples could directly be picked-up and loaded onto the pre-column without resulting in an increase in back pressure or a subsequent pre-column clogging. The glycerol free protocol was then assessed using purple membrane and membrane fraction of *C. glutamicum*. The results obtained were comparable to those applied in published reports. Therefore, the absence of glycerol did not affect the separation efficiency of the Offgel-IEF.

In addition the applicability of elastase and the glycerol free Offgel-IEF for quantitation of membrane proteins was assessed. Most of the unique peptides identified were in the acidic region and 85 % were focused only into one fraction and approximately 95 % in only two fractions. These results are in accordance with previously published results (Lengqvist *et al*., 2007). When compared with theoretical digests of the proteins identified in this study, it can be concluded that basic moiety (TMT) on the peptide backbone, did not affect the separation efficiency of the Offgel-IEF. In an applied study, changes in the protein content of yeast strain grown in two different media were relatively quantified. For example, prominent proteins, such as the hexose tranporter proteins responsible for transporting glucose accross the membrane, were successfully quantified.

Last but not least, the nLC-MALDI-TOF/TOF platform also served as a basis for the development of a high-throughput method for the identification of protein phosphorylation. The establishment of such a method using MALDI has been challenging due to the lack of sensitive matrices, such as CHCA for non-modified peptides, which exhibit a homogenous crystallisation and thus yield stable signal intensity over a long period of time in an automated setup. The first step of this method was the establishment of a matrix/matrix mixture with better crystal morphology and higher analyte signal intensity than the matrix of choice, i.e. DHB. From MS and MS/MS measurements of standard phosphopeptides, a combination of FCCA and CHAC in a 3:1 ratio and 3 mM $NH_4H_2PO_4$ facilitated high analyte signal intensities and good fragmentation behaviour. Combining a custom-packed biphasic column for the enrichment of phosphopeptides, the applicability of the matrix mixture was assessed in anautomated phosphopeptide analysis using standard phosphopeptides spiked to a 20-fold excess BSA digest. These analyses showed that this method is reproducibile and both flow throughs can be analysed. Applying the method to the analysis of 2 standard phosphoproteins, α/β-casein, and a leukemia related protein, ENL, 13 phosphopeptides from both α/β-Casein and 13 phosphopeptides with 6 phosphorylation sites from the ENL were identified.

As a general conclusion, it can be stated that the nLC-MALDI-TOF/TOF method established here in various modifications for different analytical purposes is a robust platform for proteomic analyses.

9.1 Zusammenfassung

Der Begriff *Proteom* wurde 1994 vom Australier Marc R. Wilkins in Anlehnung an die Begriffe Genom und Transkriptom beschrieben. Später wurde das Gebiet der *Proteomics* durch Patterson & Aebersold als *"systematische Erforschung der zahlreichen und unterschiedlichsten Eigenschaften von Proteinen in paralleler Weise mit der Zielsetzung einer detaillierten Beschreibung der Struktur, Funktion und Steuerung biologischer Systeme im gesunden und krankhaften Zustand*" definiert (Patterson *et al*., 2003). Somit nimmt dieses Themengebiet eine wichtige Position in der Biologie ein und übt einen erheblichen Einfluss auf verwandte Bereiche der Biowissenschaften aus. Mit Hilfe der Massenspektrometrie und der weichen Ionisierungstechniken Matrix-unterstützte Laser-Desorption/Ionisation (*engl Matrix assisted laser desorption/ ionization*,MALDI) und Elektrospray ionization (ESI) war die Ausbreitung dieses Forschungsgebiets in vielen Laboratorien nicht mehr aufzuhalten.

Für die massenspektrometrische Bestimmung von Proteinen werden meist zwei Identifizierungswege – *Top-down* und *Bottom-up* – verwendet. Beim *Top-down*-Verfahren werden die getrennten Proteine direkt gemessen, während beim *Bottom-up*-Verfahren die Proteine zuvor entweder enzymatisch oder chemisch gespalten und die so entstandenen Peptide analysiert werden. Letzteres wird häufiger für *Proteomics*-Experimente angewandt, da mehrere Proteine nebeneinander bestimmt werden können. Trotz dessen ist die *"Large-Scale"*-Proteomanalyse von komplexen biologischen Proben in vielerlei Hinsicht erschwert, da eine hohe Massengenauigkeit, Trennleistung bei den LC-Läufen und hochauflösende Massenspektrometrie erforderlich ist.

Diese Arbeit befasst sich mit der Etablierung einer nLC-MALDI-TOF/TOF Methode, basierend auf *reversed phase*-Chromatographie, zur Identifizierung von Proteinen aus komplexen biologischen Proben. Dabei wurde die Probe an zwei nanoLC-Anlagen, die jeweils mit einem automatischen Spotter versehen waren, getrennt. Die anschließende massenspektrometrische Messung fand an einem MALDI-TOF/TOF Gerät statt.

Zuerst wurde die Methode mit Hilfe von Standardpeptiden und Standardproteinen optimiert. Dabei wurde u.a. untersucht, welchen Einfluss verschiedene Parameter, wie der Fließmittelgradient, die Säulentemperatur oder die verwendete Matrixkonzentration auf die Trennung und anschließende Identifizierung haben. Aufgrund der großen Anzahl von Spektren, die bei *Bottom-up-Proteomic* generiert

werden, ist eine manuelle Interpretation nicht möglich. Aus diesem Grund ist eine hohe Qualität der Spektren notwendig, um die richtige Peptidsequenz zu identifizieren. Daher wurde neben der chromatographischen Trennung auch die massenspektrometrische Methode optimiert. Es zeigte sich, dass die besten MS/MS-Ergebnisse mit 1 kV Fragmentierungsmethode erzeugt wurden.

Während der Optimierung wurde außerdem festgestellt, dass Matrixkonzentrationen (α-cyano-4-hydroxyzimtsäure) über 3,5 mg/ml zu einer Peak-Verbreiterung führten. Vermutlich kommt es aufgrund der Kristallbildung an der Nadelspitze des Spotter zu Verschleppungen des Analyten und dadurch zu der Peak-Verbreiterung. Um die optimale Säulentemperatur zu bestimmen, wurde die Trennung bei 25, 30, 40 und 60 °C durchgeführt. Die erhöhte Säulentemperatur führte zu einer Druckverminderung während der LC-Trennung. Durch die Verringerung des Drucks ist der Einsatz von längeren Trennsäulen oder höheren Flussraten möglich. Ferner zeigte sich, dass bei 40°C Säulentemperatur die beste Auftrennung und damit verbunden deutlich mehr Peptide identifiziert werden konnten. Bei 60 °C nahm die Menge der identifizierten Peptide als Folge des Abbaus von Säulenmaterial ab.

Nach der Optimierung wurde die Robustheit der Methode und nLC-Anlage durch Analyse verschiedener biologischer Proben überprüft. Dabei wurden Proteine aus *Saccharomyces*-Lipidpartikeln und aus humanen Adipocyten untersucht. Die Herausforderung für die Analytik dieser Proben war zum einen die niedrige Proteinkonzentration und zum anderen der hohe Lipidanteil der Partikel (> 90 %), der nach dem Verdau anfangs zu einem Verstopfen der Vorsäule führte. Um eine Analyse der Proben mittels nLC zu ermöglichen, mussten daher die Lipide zuvor mit Diethylether extrahiert werden. Dank der verbesserten Methode konnten neben bereits bekannten jeweils sechs neue Proteine identifiziert werden. Insgesamt wurden 41 Proteine aus humanen subkutanen Steatoblasten, bzw. Adipozyten identifiziert, davon neben den sechs neuen Proteinen, 14 säugetierspezifische und 21 viszeral-spezifische Proteine.

Weiterhin wurde die Methode zur Analytik von Proteinkomplexen verwendet. Die Untersuchung von Protein-Protein-Wechselwirkung gibt Aufschlüsse über Zellmechanismen und über das Entstehen bestimmter Krankheiten, wie etwa Leukämien, woraus sich Therapieansätze und -maßnahmen ableiten lassen. Das *mixed-lineage leukemia* (*MLL*-Gen*)* auf Chromosom 11 Bande q23 ist in zahlreiche, reziproke chromosomale Translokationen verwickelt. Bei der akuten Leukämie t(4;11), häufig anzutreffen bei Säuglingen und Kleinkindern, sind das *MLL*-Gen auf Chromosom 11 und das *AF4*-Gen auf Chromosom 4 beteiligt. Durch die Translokation entstehen zwei neue Derivatchromosomen – (*MLL•AF4*) und (*AF4•MLL*). Chromosomale Translokationen des *MLL*-Gens werden auf Grund ihrer

sehr schlechten Prognose und Therapierbarkeit als Hochrisiko-Leukämien eingestuft. Bis heute konnten 64 Translokations-Partnergene identifiziert werden, wobei das *AF4*-Gen mit 42 % bei allen untersuchten Leukämien und mit ca. 66 % bei der akuten lymphatischen Leukämie (ALL) den größten Prozentsatz ausmacht.

Dabei war das Ziel dieser Anwendung die Identifizierung neuer Interaktionspartner mit den Fusionsproteinen MLL•AF4 und AF4•MLL. Hierfür wurden die Strep-Elutionen von AF4, AF4•MLL, ENL und die Negativkontrolle in Lösung verdaut, getrennt und analysiert. Insgesamt wurden 17 bekannte Interaktionspartner, deren Funktion und Rolle in der Leukämogenese eindeutig in zahlreichen Studien nachgewiesen ist, aus AF4- und AF4•MLL-Komplexen identifiziert; daneben auch neue Fusionsproteine des AF4-Proteins (DDX6, NPM1, NFkB, BRD4, CARM1, NSD1) und des AF4•MLL-Fusionsprotein-Komplexes (DDX6, NPM1, NFkB, HEXIM1, CARM1, DOT1L). Leider konnten bei der ENL keine Fusionsproteine von bekannten Interaktionspartnern nachgewiesen werden.

Trotz der Identifizierung bekannter und neuer Interaktionspartner aus den AF4- und AF4•MLL- Proteinkomplexen gibt es weitere Herausforderungen, wie bei der Zellzucht oder der Probenvorbereitung zu bewältigen. Viele Interaktionen von Proteinen sind leider nicht von Dauer, oder die Interaktionspartner sind in zu geringen Mengen vorhanden, sodass die Analyse schwierig ist.

Als weiterer Anwendungsbereich wurde die nLC-MALDI-TOF/TOF-Methode zur Identifizierung von Membranproteinen verwendet. Membranproteine sind aufgrund ihrer hydrophoben Eigenschaften in wässrigen Puffersystemen schlecht löslich und somit im Vergleich zu hydrophilen Proteinen analytisch schwer zugänglich.

Weiterhin besitzen sie weniger tryptische Schnittstellen, weshalb die Verwendung von Trypsin zur Hydrolyse der Proteine nur zu mäßigen Ergebnissen führt. Weniger spezifische Enzyme bilden überlappende Fragmentcluster, da sie Proteine an mehr Aminosäuren als Trypsin spalten. Wegen seiner unspezifischen Spaltung ist die Verwendung von Elastase im Bereich Proteomik bisher eingeschränkt. Es stellte sich allerdings heraus, dass nach dem Verdau von Membranproteinen aus *Halobacterium salinarium* und *Corynebacterium glutamicum* eine, im Vergleich zu Trypsin, höhere Sequenzabdeckung des gesamten Proteins sowie der Transmembrandomänen erreicht wurde. Insgesamt konnten bei der Analyse der Membranfraktion 379 Proteine identifiziert werden. Davon waren 44 % integrale Membranproteine. Zum Vergleich wurde mit Trypsinverdau lediglich weniger als 44 % integrale Membranproteine identifiziert Diese Ergebnisse zeigen, dass der Einsatz von weniger spezifischen Enzymen bei Membranproteinen zu einer höheren Sequenzabdeckung

führen kann. Ein Problem, das sich nach dem Verdau mit Elastase stellte, war die Auswertung der erzeugten MS/MS-Daten. Etwa 43 % der Fragmentionen gehörten nicht zu den bekannten y- oder b-Ionen und konnten daher nicht den internen Fragmenten zugeordnet werden.

Ein Grund dafür ist vermutlich die fehlende Lokalisierung der positiven Ladung an den C- und N- Termini der entstandenen Peptide. Eine Erhöhung der internen Fragmentionenbildung führte zu Schwierigkeiten bei der Interpretation der Spektren. Die Fragmentierung konnte wiederum durch Derivatisierung mit einem basischen Rest (z.B. TMT) verbessert werden.

Durch Verwendung von TMT konnten mehr als 50 % der Fragmentionen wieder zugeordnet werden (Bäumlisberger *et al*., 2010).

Selbst die verbesserte Trennleistung der entwickelten nLC-Methode war für sehr komplexe biologische Proben, wie z.B. Zelllysate, nicht ausreichend. Diese Problematik wurde durch Kombination mit der Offgel-IEF-Trennung verbessert. Diese Technik trennt die Proteine/Peptide der komplexen Probe anhand des isoelektrischen Punkt in verschiedene Fraktionen auf. Die erhaltenen Fraktionen wurden anschließend einzeln auf der nLC getrennt. Nach Optimierung des Methodenprotokolls war es möglich die vorfraktionierten Proben mittels nLC weiter aufzutrennen und anschließend mittels MALDI-TOF/TOF zu identifizieren. Es zeigte sich, dass durch die optimierte Offgel-IEF Methode ca. 90-95 % aller Peptide in einer oder zwei Fraktionen vorlagen.

Ferner wurde überprüft, ob die optimierte Methode für eine relative Quantifizierungstechnik geeignet ist. Hierfür wurden mit TMT markierte Peptide mittels IEF und anschließend nLC getrennt. Bei der Überprüfung der Trennleistung wurde kein negativer Einfluss des basischen TMT auf die Trennung festgestellt. Auch mit TMT lagen 95 % der identifizierten Peptide in nur einer oder zwei Fraktionen vor. Generell induziert TMT eine geringe Verschiebung des isoelektrischen Punktes, der bei Verwendung von IPG-Streifen mit großen pH-Bereich (pH 3-10) aber vernachlässigbar ist. Trotz der effizienten Trennung, ist die Fokussierung im Bereich von neutralen und alkalischen p*I*´s bei Weitem noch nicht optimal. Dies ist durch die allgemein schlechtere Fokussierung von neutralen Peptiden zu erklären, da aliphatische Peptide ohne polare Seitenketten im pH-Intervall zwischen 2,4 und 9,6 keine Nettoladung aufweisen. Vorgeschlagene Strategien zur Verbesserung der Trennleistung sind durch die verwendeten Zusätze wie Sorbit oder Methylcellulose leider nicht mit der nLC kompatibel.

Der letzte Teil dieser Arbeit befasst sich mit der Etablierung einer automatisierten Methode zur Identifizierung und Charakterisierung von Phosphopeptiden. Die Analyse von phosphorylierten Aminosäuren gestaltet sich trotz ständiger instrumenteller Verbesserungen immer noch als schwierig. Dies liegt vor allem daran, dass in der Regel nur ein geringer Prozentsatz eines potentiellen Phosphoproteins tatsächlich phosphoryliert vorliegt und zudem noch reversibel modifiziert ist.

Die Identifizierung und Charakterisierung von Phosphopeptiden wird meistens mittels ESI-MS durchgeführt. Der Grund dafür sind unter anderem die fortgeschrittenen Fragmentierungstechniken, wie z.B. der Einsatz der Elektronentransferdissoziation (ETD). Des Weiteren ermöglicht die direkte Kopplung mit der HPLC eine schnelle Identifizierung von komplexen Proteingemischen. MALDI hingegen wird meistens nur für manuelle Messungen von nicht komplexen Proben verwendet. Die üblicherweise für Phosphopeptidanalysen verwendete MALDI-Matrix DHB weist eine geringe Analytensensivität und eine inhomogene Kristallisation auf. Dies macht eine automatische Messung nahezu unmöglich.

Zuerst musste eine für die Phosphopeptidanalysen besser geeignete Matrix gefunden werden. Dazu wurden vier verschiedene Matrizes (4-Chlor-α-cyano-zimtsäure [ClCCA], 4-Hydroxy-α-cyano-zimtsäure [CHCA], 2,5-Dihydroxybenzoesäure [DHB] und Di-2,4-Difluor-α-cyano-zimtsäure [diFCCA]), sowie Mischungen daraus getestet. Die besten Ergebnisse wurden durch Verwendung eines 3:1-Gemisches aus diFCCA und CHCA mit 3 mM $NH_4H_2PO_4$-Lösung erzielt. Neben der Optimierung der Matrix und Probenpräparation wurde eine Anreicherung und Trennung der Phosphopeptide mittels einer 2-phasigen Vorsäule (TiO2/C18) erzielt. Bei pH < 4 bleiben die Phosphopeptide auf dem TiO2 Material und die nicht phosphorylierten Peptide auf dem C18 Material haften. Danach konnten die nicht phosphorylierten Peptide mit einem steigenden ACN/0,1 % TFA-Gradient getrennt werden, während die phosphorylierten Peptide auf die TiO_2-material haften blieben. Durch Änderung des pH-Wertes mittels Ammoniumhydrogencarbonat (pH 10) wurden die Phosphospeptide in den C_{18}-Teil der Vorsäule eluiert und konnten dann mit einem weiteren LC-Lauf aufgetrennt werden. Mit dieser Methode können Phosphopeptide einer Probe separat von den nicht phosphorylierten Peptiden analysiert werden. Dies bewirkt bei der anschließenden massenspektrometrischen Detektion eine deutlich einfachere Identifizierung der Peptide. Der Vorteil ist, dass die Reproduzierbarkeit durch die Automatisierung steigt und die Vorsäule mehrmals verwendet werden kann. Die Methode wurde erfolgreich sowohl mit Standardphosphoproteinen, als auch mit ENL-Proteinkomplexen validiert.

Abschließend lässt sich sagen, dass die massenspektrometrische Identifikation und Charakterisierung von PTMs immer noch eine außerordentlich schwierige Aufgabe darstellt, da das standardmäßige Erreichen von 100 % Sequenzabdeckung bis jetzt noch nicht möglich ist.

Diese Arbeit zeigt, dass durch Verwendung der nLC auch in Kombination mit anderen Trennverfahren, eine erheblichen Verbesserung der Identifizierung der Identifizierung von Proteine mittels MALDI-TOF/TOF erreicht werden konnte..

10 References

Aggarwal Kunal, Choe H. Leila and Lee H Kelvin.Briefings in functional genomics and proteomics.5.(2). 112-120

Akao, Y., and Isobe, M. 2000. Cancer 27, 412-417.

Alpert, Andrew J. 1990, Journal of Chromatography 499: 177–196.

Alvarez-Llamas G, Szalowska E, de Vries MP, Weening D, Landman K, Hoek A, et al., 2007, Mol Cell Proteomics. 6(4):589-600.

Anne Benedikt 2009, Der AF4- und AF4•MLL-Multiproteinkomplex und deren potentielle Rollen in transkriptionellen und epigenetischen Prozessen

Aoki N, Jin-no S, Nakagawa Y, Asai N, Arakawa E, Tamura N, et al. 2007 Endocrinology,148(8):3850-62.

Arnold RJ, Jayasankar N, Aggarwal D, Tang H, Radivojac P 2006, Pac Symp Biocomput. 219-230.

Asara, J. M., and Allison, J. 1999, J. Am. Soc. Mass Spectrom. 10, 35– 44.

Athenstaedt K. and G. Daum 2000, J. Biol. Chem. 275, 235–240

Athenstaedt Karin, Zweytick Dagmar, Jandrositz Anita, Kohlwein Dieter Sepp, and Daum Günther 1999, Journal of Bacteriology 181, 20, 6441-6448

Athenstaedt, K., and G. Daum 1997. J. Bacteriol. 179:7611-7616,

Athenstaedt, K., Daum,G. 2006, Cell Mol.Life Sci. 63, 1355-1369.

Baldus, C.D., Bullinger, L. 2008, Oncol. 35, 356-364.

Baltruschat Sabrina 2009, Untersuchung hochmolekularer Proteinkomplexe in menschlichen Leukämien mittels Proteomics Werkzeugen.

Barr, J. R., Maggio, V. L., Patterson, D. G., Jr., Cooper, G. R., et al., 1996 Clin. Chem. 42, 1676–1682

Baxter RC, Twigg SM. 2009, Trends Endocrinol Metab., 20(10):499-505.

Beavis RC, Chait BT. 1989, Rapid Commun. Mass Spectrom. 3 (12): 432–5.

Beavis RC, Chait BT. 1989, Rapid Commun. Mass Spectrom. 3 (12): 436–9.

Beavis, R. C. 1992, Org. Mass Spectrom. 27: 156–8.

Beller M, Sztalryd C, Southall N, Bell M, Jäckle H, Auld DS, Oliver B. 2008, PLoS Biol. 6(11):e292.

Ben Herbert, Righetti Pier Giorgio 2000, Electrophoresis 21 (17), 3639 – 3648

Bendz M, Möller MC, Arrigoni G, Wåhlander A, Stella R, Cappadona S, Levander F, Hederstedt L, James P. 2009, J.Proteome Res. 8(12):5666-73.

Benedikt Anne 2009, Der AF4- und AF4•MLL-Multiproteinkomplex und deren potentielle Rollen in transkriptionellen und epigenetischen Prozessen.

Beynon, R. J.; Doherty, M. K.; Pratt, J. M.; Gaskell, S. J. 2005, Nat. Methods 2, 587–589.

Bickel, P.E., Tansey, J.T., Welte, M.A. 2009, Biochim.Biophys.Acta 1791, 419-440.

Bieth, J.G., 2004, Pancreatic Elastase. Handbook of Proteolytic Enzymes 2nd Edn (Elsevier, Amsterdam, Netherlands)

Billeci T. M. and Stults J. T. 1993, Anal .Chem. 65, 1709

Bitoun, E., Oliver, P.L., Davies, K.E., 2007, Hum. Mol. Genet. 16, 92-106.

Blackburn K, Goshe MB. 2009, Brief Funct Genomic Proteomic. 8(2):90-103.

Blackburn Kevin and Goshe B. Michael.2009, Briefings in Functional Genomics 8(2):90-103.

Blacken GR, Volný M, Vaisar T, Sadílek M, Turecek F. 2007, Anal Chem., 79(14):5449-56.

Blonder J, Conrads T P, Veenstra TD, 2004. Expert Rev. Proteomics. 153-63

Blonder J, Conrads TP, Yu LR, Terunuma A, Janini GM, Issaq HJ, Vogel JC, Veenstra TD. 2004, Proteomic. 4 (1):31-45.

Bodenmiller B, Mueller LN, Mueller M, Domon B, Aebersold R. 2007, Nat Methods, 4(3):231-7.

Bodnar W.M, Blackburn R.K., Krise J.M. and Moseley M.A 2003, J. Am. Soc. Mass Spectrom. 14,.971–979.

Boles, E., and C. P. Hollenberg.1997, Microbiol.Rev. 21:85–111.

Borner GH, Sherrier DJ, Weimar T, Michaelson LV, Hawkins ND, Macaskill A, Napier JA, Beale MH, Lilley KS, Dupree P. 2005, Plant Physiol. 137(1):104-16.

Bowie JU. 1997, J. Mol. Biol. 272(5):780-9.

Brasaemle DL, Dolios G, Shapiro L, Wang R. 2004, Journal of Biological Chemistry, 279, 46835-46842.

Brill L.M. and Brock A. 2003, Anal. Chem. 75,. 2309–2315.

Bogan M.J. and Agnes G.R. 2004, Rapid Commun.Mass Spectrom. 18,. 2673–2681.

Bursen, A., Moritz, S., Gaussmann, A., Moritz, S., Dingermann, T., Marschalek, R. 2004, Oncogene 23, 6237-6249

Carboni L, Piubelli C, Righetti PG, Jansson B, Domenici E. 2002, Electrophoresis. (24):4132-41.

Cargile BJ, Bundy JL, Freeman TW, et al. 2004, J Proteome Res., 3:112–119

Cargile BJ, Bundy JL, Stephenson JL Jr. 2004, J. Proteome Res., 3, 112-119.

Cargile BJ, Sevinsky JR, Essader AS, Stephenson JL Jr, Bundy JL.2000, J. Biomol. Tech., 16, 181-190.

Cargile BJ, Stephenson JL, Jr. 2004, Anal Chem., 76: 267–275.

Cargile BJ, Talley DL, Stephenson JL, Jr. 2004, Electrophoresis, 25:936–945

Cargile J. Benjamin, Joel R. Sevinsky, Amal S. Essader, James L. Stephenson, Jr., and Jonathan L. Bundy 2005, J Biomol Tech. 16(3): 181–189

Cazzaniga, G., Dell'Oro, M.G., Mecucci, C., Giarin, E., Masetti, R., et al. 2005, Blood 106, 1419-1422.

Cech N.B. and Enke C.G. 2000, Relating Electrospray Ionization Response to Nonpolar Character of Small Peptides.Anal. Chem. 72 (13), 2717–2723.

Chan SS, Schedlich LJ, Twigg SM, Baxter RC. 2009, Am J Physiol Endocrinol Metab. 296(4):E654-63.

Chan SS, Twigg SM, Firth SM, Baxter RC 2005, J Clin Endocrinol Metab. 90(12):6588-95.

Chen CP, Rost B. 2002, Appl Bioinformatics.1(1):21-35.

Chen X, Cushman SW, Pannell LK, Hess S. 2005, J Proteome Res., 4(2):570-7.

Chiellini C, Cochet O, Negroni L, Samson M, Poggi M, Ailhaud G, et al. 2008;. BMC Mol Biol. 9:26. 1471-2199

Cho, Y. W., Hong, T., Hong, S., Guo,et al., (2007)., J Biol Chem 282, 20395-20406.

Choi J.Y and C.E. Martin 1999, J. Biol. Chem. 274 (), pp. 4671–4683

Chu, C. Y., and Rana, T. M. 2006, PLoS Biol 4, e210.

Cotter R. J. 1984 Anal. Chem 56(3), 485A-504A

Czabany T. Wagner, A. Zweytick et al. 2008, J. of biological chemistry 283, 17065-17074

Dahlman I, Eriksson P, Kaaman M, Jiao H, Lindgren CM, Kere J, et al. 2004, Diabetologia. 47(11):1974-9.

Dayon Loïc, Hainard Alexandre, Licker Virginie, et al. 2008, Anal. Chem., 80 (8), 2921–2931

Decitre M, Gleyzal C, Raccurt M, Peyrol S, Aubert-Foucher E, Csiszar K, et al., 1998, Lab Invest. 78(2):143-51.

DeGnore JP, Qin J.1998, J Am Soc Mass Spectrom., 9(11):1175-88.

Delom F, Szponarski W, Sommerer N, Boyer JC, Bruneau JM, Rossignol M, Gibrat R. 2006, Proteomics. 6(10):3029-39.

Dobson, K. D., and McQuillan, A. J. 2000, Acta Part A Mol. Biomol. Spectrosc.56, 557– 565.

Domer PH, Fakharzadeh SS, Chen CS, Jockel J, Johansen L, Silverman GA, Kersey JH, Korsmeyer SJ. 1993, Proc Natl Acad Sci U S A., 90(16):7884–7888.

Doroshenko V.M. and. Cotter R.J. 1995, Anal. Chem. 67 (13), 2180–2187

Dunn J. D, Igrisan EA, Palumbo AM, Reid GE, Bruening ML. 2008, Anal Chem. 80(15):5727-35.

Dunn JD, Watson JT, Bruening ML. 2006 Anal Chem., 78(5):1574-80.

Edman Pehr. 1950, ACTA Chemica Scandinavica 4, 283-293

Egen NB, Bliss M, Mayersohn M 1988. Anal Biochem; 172:488–494.

Eguchi M, Eguchi-Ishimae M, Greaves M. 2003, Int J Hematol.;78:390–40.

Eichacker LA, Granvogl B, Mirus O, Müller BC, Miess C, Schleiff E. 2004,. J Biol. Chem., 279(49):50915-22.

Ericson C., Phung Q.T., Horn D.M., Peters E.C., Fitchett J.R., Ficarro S.B., Salomon A.R.,

Ejaz A, Wu D, Kwan P, Meydani M. 2009, J Nutr. 139(5):919-25.

Ekstrom S., Nisson J., Helldin G., Laurell T. and Marko-Varga G. 2001, Electrophoresis 22, 3987–3992.

Eng K. Jimmy, McCormack L. Ashley and Yates III R. John.1994, Am. Soc. Mass Spectrom. 5, 976–989

Eric m. Phizicky1* and Stanley fields 1995, American Society for Microbiology p. 94–123

Essader AS, Cargile BJ, Bundy JL, et al. 2005, Proteomics 5(1):24–34.

F. Hillenkamp, M. Karas, R.C. Beavis and B.T. Chait. 1991, Anal.Chem., 63. A1193–A1203

Farese,R.V., Jr., Walther,T.C. 2009, Cell 139, 855-860.

Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F, Whitehouse, C.M. 1989, Science. 246, 64-71

Ficarro, S. B., McCleland, M. L., Stukenberg, P. T., Burke, D. J., Ross, M. M., Shabanowitz, J., Hunt, D. F., and White, F. M. 2002, Nat. Biotechnol.20, 301– 305.

Fink SL, Cookson BT. 2005, 73(4):1907-16.

Firth SM, Baxter RC. 2002, Endocr Rev. 23(6):824-54.

Foster, L. J.; Zeemann, P. A.; Li, C.; Mann, M.; Jensen, O. N.; Kassem, M. Stem Cells 2005, 23, 1367.

Fountain S.T., Lee H. and Lubman D.M.1994, Rapid Commun. Mass Spectrom. 8 (5), 407–416

Fraterman, S.; Zeiger, U.; Khurana, T. S.; Rubinstein, N. A.; Wilm, M. 2007.. Proteomics, 7, 3404-3416.

Fujimoto Y, Itabe H, Sakai J, Makita M, Noda J, Mori M, Higashi Y, Kojima S, Takano T. 2004. Biochim.Biophys. Acta 1644, 47–59

Garzon, R., Garofalo, M., Martelli, M.P., et al. 2008. Proc. Natl. Acad. Sci. USA 105, 3945-3950.

Gerber A Scott, Rush John, Stemman Olaf, Kirschner W Marc, and Gygi P.Steven 2003, PNAS. 100 (12) 6940-6945

Gettins PG. 2002, Chem Rev. 102(12):4751-804.

Glückmann, M.; Karas, M. J. 1999, Mass Spectrom. 34, 467

Goo YA, Yi EC, Baliga NS, Tao WA, Pan M, Aebersold R, Goodlett DR, Hood L, Ng WV. 2003, Mol Cell Proteomics, 2(8):506-24.

Goodman, J. M. 2009, J. Lipid Res. 50: 2148-2156

Goodman,J.M. 2008,. J.Biol.Chem. 283, 28005-28009.

Görg A, Postel W, Günther S. 1988 , Electrophoresis 9: 531-546

Görg Angelika , Weiss Walter, Dunn J. Michael, 2004, Proteomics. 4 (12),3665 – 3685.

Görg, A., Weiss, W., Dunn, M.J. 2004, Proteomics. 4, 3665-3685

Goshe, M.B., Blonder, B. & Smith, R.D. 2003. J. Proteome Res. 2, 153–161

Griffin T.J., Gygi S.P., Rist B., Aebersold,R. Loboda A., Jilkine A., Ens W. and Standing K.G. 2001, Anal. Chem. 73, 978–986

Grillitsch Karlheinz, Connerth Melanie, Köfeler Harald, Arrey N. Tabiwang, Rietschel Benjamin, Karas Michael and Dauma Günther (Submitted).

Grummitt, C.G., Townsley, F.M., Johnson, C.M., Warren, A.J., Bycroft, M. 2008, J. Biol. Chem. 283, 23326-23332.

Guenther G. Matthew, Jenner G. Richard, Brett Chevalier, Tatsuya Nakamura, Carlo M. Croce, Eli Canaani, and. Richard A. Young. 2005, PNAS. 102. (24) 8603-8608. ¶

Guenther, M. G., Jenner, R. G., Chevalier, B., Nakamura, T., Croce, C. M., Canaani, E. & Young, R. A. 2005, Proc. Natl. Acad. Sci. USA 102, 8603-8608

Gurumurthy, M., Tan, C.H., Ng, R., Zeiger, L., et al. 2008, J. Mol. Biol. 378, 302-317.

Gygi S.P., Corthals G. Zhang L., Y., Rochon Y. and Aebersold R., 2000.Proc. Natl. Acad. Sci. U.S.A. 97 (17)9390–9395.

Gygi, S. P.; Rist, B.; Gerber, S. A.; Turecek, F.; Gelb, M. H.; Aebersold, R., 1999, Nat. Biotechnol. 17 (10), 994−999.

Gygi, S. P.; Rist, B.; Gerber, S. A.; Turecek, F.; Gelb, M. H.; Aebersold, R. 1999, Nat. Biotechnol. 17 (10), 994−999.

Haas W, Faherty BK, Gerber SA, Elias JE, Beausoleil SA, Bakalarski CE, Li X, Villén J, Gygi SP 2006, Mol Cell Proteomics. (7):1326-37

Han Xuemei, Aslanian Aaron and Yates III R John 2008, Current Opinion in Chemical Biology 12, (5) 483-490

Hardman M, Makarov A. 2003; Anal. Chem. 75: 1699.

Harper P. David, and Peter D. Aplan1 2008, Cancer 68; 10024.

Harvey Lodish, Arnold Berk, S Lawrence Zipursky, Paul Matsudaira, David Baltimore, and James Darnell 2000, Molecular Cell Biology 4th edition. New York: W. H. Freeman ISBN: 0-7167-3136-3

Hausman GJ, Poulos SP, Richardson RL, Barb CR, Andacht T, Kirk HC, et al. 2006, J Anim Sci. 84(7):1666-81.

Haystead TAJ, Garrison JC: Study of protein phosphorylation inintact cells. In Protein Phosphorylation: a Practical Approach, edn 2. Edited by Hardie DG. Oxford: Oxford University Press; 1999:1-31.

Heller M, Michel PE, Morier P, Crettaz D, Wenz C, Tissot JD, Reymond F, Rossier JS:. 2005, Electrophoresis, 26:1174-1188

Heller M, Ye M, Michel PE, Morier P, Stalder D, Junger MA, Aebersold R, Reymond F, Rossier JS. 2005 J Proteome Res. 4:2273–2282

Herbert B, Righetti PG. 2000,. Electrophoresis, 21(17):3639-48.

Hess JL. MLL: 2004, Trends Mol Med.; 10:500–507.

Hess, J.L., Roeder, R.G. 2005, Cell 121, 873-885.

Hillenkamp F. and Peter-Katalinić J. 2007, MALDI MS: a practical guide to instrumentation, methods and applications.

Hoang T, Roth U, Kowalewski K, Belisle C, Steinert K, Karas M. 2010 Anal Chem., 82(1):219-28.

Hoang-Trang Lam 2008 Of Gels, Offgel and Off with the Gel: Development of Isoelectric Focusing Tools for Proteomics.

Hochuli E, Döbeli H, Schacher A. 1987 J Chromatogr. 411:177-84.

Hörth, P.; Miller, C. A.; Preckel, T.; Wenz, C. 2006, Mol. Cell Proteomics, 5, 1968-1974.

Hossner KL, Yemm R, Vierck J, Dodson MV. 1997, In Vitro Cell Dev Biol Anim. 33(10):791-5.

Hu Qizhi, Noll t J. Rober , Li Hongyan, Makarov Alexander, Hardman Mark, Cooks R. Graham 2005, 40,(4) 430-443

Hubner, N. C.; Ren, S.; Mann, M. 2008, Proteomics 8, 4862-4872.

Hunt F D , Yates3rd J R , Shabanowitz J , S Winston, and Hauer C R. 1986, Proc. Nat. Acad. Sci. USA 83, 6233–6237

Hunter and Sefton, 1980 T. Hunter and B.M. Sefton 1980, Proc. Natl. Acad. Sci. USA 77, 1311–1315.

Hurwitz N, Pellegrini-Calace M, Jones DT 2006, Philos Trans R Soc Lond B Biol Sci. 361(1467):465-75.

Ikeguchi, Y., and Nakamura, H. 1997, Anal. Sci. 13, 479– 483.

James P., Quadroni M., Carafoli E. and Gonnet G. 1993, Biochem. Biophys. Res. Commun. 195 (1) 58–64

Jaskolla TW, Papasotiriou DG, Karas M. 2009, J Proteome Res. 8(7):3588-97.

Jaskolla, T.W., Lehmann, W.D., Karas , M. 2008, Proc. Natl. Acad. Sci. U S A. 105, 12200-12205

Jérôme Chenau, Sylvie Michelland, Jonathan Sidibe, and Michel Seve 2008 , Proteome Sci.6: (9), 477-5956.

Ji Chengjie and Li Liang 2005, J. Proteome Res. 4, 734-742.

Jin Z-G, Lungu AO, Xie L, Wang M, Wong C, Berk BC. 2004,. Arterioscler Thromb Vasc Biol. 24(7):1186-91.

Jin ZG, Melaragno MG, Liao DF, Yan C, Haendeler J, Suh YA, et al. 2000, Circ Res. 27;87(9):789-96.

Jürgen Cox. and Mann mathias. 2007, Cell, 130, 395-398.

Kallio P, Tolppanen AM, Kolehmainen M, Poutanen K, Lindstrom J, Tuomilehto J, et al. 2009, Int J Obes (Lond). 33(1):80-8.

Karas M. and Hillenkamp F. 1988, Anal. Chem. 60 20), 2299–2301

Karas M., Bahr U., lngedoh A. and Hillenkamp F. 1989, Angew.Chem.lnt. Ed. Engl. 28, 760.

Karas, M.; Bachman, D.; Bahr, U.; Hillenkamp, F. 1987, Int J Mass Spectrom. Ion Proc 78: 53–68.

Kaufmann R., Spengler B. and Lutzenkirchen F. 1993, Rapid Commun. Mass Spectrom. 7 (10), 902–910

Kellermann J. 2008, Methods Mol Biol. 424:113-23.

Kinseth MA, Anjard C, Fuller D, Guizzunti G, Loomis WF, Malhotra V. 2007, Cell, 10;130(3):524-34.

Kjellstrom, S., and Jensen, O. N. 2004 Anal. Chem.,76, 5109– 5117.

Knoll L.J. et al. J. Biol. Chem. 269 1994, 16348–16356).

Kostner K 2002 Journal für Kardiologie - Austrian Journal of Cardiology, 9 (7-8): 328-331

Kratchmarova I, Kalume DE, Blagoev B, Scherer PE, Podtelejnikov AV, Molina H, et al.2002, Mol Cell Proteomics. 1(3):213-22.

Kratzer R., Eckerskorn C. Karas, M. and Lottspeich F. 1998, Electrophoresis 19 (11), 1910–1919.

Krogh, A., Larsson, B., von Heijne, G., Sonnhammer, E.L. 2001, J Mol Biol 305, 567-580.

Kyte, J., Doolittle, R.F. 1982, J. Mol. Biol. 157, 105-132

Lam, H. T.; Josserand, J.; Lion, N. Girault, H. H. 2007, J. Proteome Res., 6, 1666-1676

Larsen MR, Thingholm TE, Jensen ON, Roepstorff P, Jørgensen TJ. 2005, Mol Cell Proteomics 4(7):873-86.

Laurell T., Nisson J. and Marko-Varga G. 2001, J. Chromatogr.B 752, 217–232.

Lavebratt C, Dungner E, Hoffstedt J. 2005, J Lipid Res.46(10):2278-81.

Le Bihan T, Goh T, Stewart II, Salter AM, Bukhman YV, Dharsee M, Ewing R, Wiśniewski JR. 2006. J. Proteome Res., 5(10):2701-10.

Le Maire M, Deschamps S, Møller JV, Le Caer JP, Rossier J. 1993, Anal Biochem. 214(1):50-7.

Leber R, Landl K, Zinser E, Ahorn H, Spök A, Kohlwein SD, Turnowsky F, Daum G.1998, Mol. Biol. Cell 9:375-386

Lengqvist J, Uhlen K, Lehtio J: 2007, Proteomics., 7:1746–1752.

Li Y. and McIver Jr R.T.1994, Rapid Commun Mass Spectrom.8(9):743-9.

Liang, H., and R. F. Gaber. 1996,. Mol. Biol. Cell 7:1953–1966.

Liebler C. Dainiel, Introduction to proteomics: Tool for the new biology.

Lim JM, Sherling D, Teo CF, Hausman DB, Lin D, Wells L. 2008, J Proteome Res. Mar;7(3):1251-63.

Lin, C.Y., Liang, Y.C., Yung, B.Y., 2006, Cell Signal 18, 2041-2048.

Link A.J., Eng J., Schieltz D.M., Carmack E., Mize G.J., Morris D.R., Garvik B.M. and J.R., Yates I.I.I 1999, Nat. Biotechnol. 17 (7),676–682.

Liu Tao, Belov E. Mikhail, Jaitly Navdeep, Qian Wei-Jun, and Smith D. Richard,2007, Chem. Rev., 107 (8), 3621–3653

Liu, P., Ying, Y., Zhao, Y., Mundy, D. I., Zhu, M., Anderson, R. G. W. 2004, J. Biol. Chem. 279, 3787–3792

Loboda A.V., Krutchinsky A.N., Bromirski M., Ens W.and Standing K.G. 2000, Rapid Commun. Mass Spectrom. 14 (12),1047–1057

Lodish H., Berk H., Zipursky A., Matsudaira, P. Baltimore D. and Darnell J., Molecular Cell Biology, W. H. Freeman, New York,4th edn, 1999.

Loïc Dayon, Alexandre Hainard, Virginie Licker, Natacha Turck, Karsten Kuhn, Denis F. Hochstrasser, Pierre R. Burkhard, and Jean-Charles Sanchez. 2008, Anal. Chem., 80 (8), 2921–2931.

MacCoss, M.J., McDonald, W.H., et al. 2002. Proc. Natl. Acad. Sci. U S A. 99, 7900-7905.

Makarov A. 2000; Anal. Chem 72: 1156

Makarov A. Mass spectrometer.U.S. Patent 5,886,346 1999.

Mann M. and Wilm M. 1994, Anal. Chem. 66, 4390–4399

Mann Matthias and Jensen N Ole 2003, Biotechnology 21, 255 – 261.

Mann Matthias and Ole N. Jensen 2003, Nature Biotechnology 21, 255 – 261

Marschalek R, Nilson I, Löchner K, Greim R, Siegler G, Greil J, Beck JD, Fey GH., 1997 Leuk Lymphoma. 27(5-6):417-28.

Mariman EC, Wang P. 2010, Cell Mol Life Sci. 67(8):1277-92.

Mastro R, Hall M. 1999, Anal Biochem. 10;273(2):313-5.

McLachlin, D. T., and Chait, B. T. 2003.Anal.Chem. 75, 6826– 6836.

Medzihradszky K.F., Campbell J.M., Baldwin M.A., Falick A.M., Juhasz P., Vestal M.L. and Burlingame A.L. 2000, Anal.Chem. 72 (3),. 552–558

Meyer C, Kowarz E, Hofmann J, Renneville A, et al. 2009, Leukemia. 23(8):1490-9.

Meyer, E., Cole, G., Radhakrishnan, R., Epp, O. 1988, Acta Crytallogr. B. 44, 28-38

Miliotis T., Kjellstrom S., Nilsson J., Laurell T., Edholm L. and Marko-Varga G. 2000, J. Mass Spectrom. 35, pp. 369–377.

Mirgorodskaya E., . Braeuer C, Fucini P., Lehrach H. and Gobom J. 2005, Proteomics 5, 399–408.

Milne, T. A., Briggs, S. D., Brock, H. W., Martin, M. E., Gibbs, D., Allis, C. D. & Hess, J. L., 2002, Mol. Cell 10, 1107-1117.

Miyamoto, Y., Imamoto, N., Sekimoto, T., Tachibana, T., et al 1997. J. Biol. Chem. 272: 26375-26381

Molina H, Yang Y, Ruch T, Kim JW, Mortensen P, Otto T, et al. 2009, J Proteome Res., 8(1):48-58.

Morandell S, Stasyk T, Grosstessner-Hain K, Roitinger E, Mechtler K, Bonn GK, Huber LA., 2006, Proteomics. 6(14):4047-56.

Morandell Sandra, Stasyk Taras, Grosstessner-Hain Karin, Roitinger Elisabeth, Mechtler Karl, Bonn K. Guenther and Huber A. Lukas. 2006, Proteomics 6, 4047–4056.

Moser K, White FM. 2006 J Proteome Res. 5(1):98-104.

Nabetani T, Miyazaki K, Tabuse Y, Tsugita A. 2006, Proteomics, 6(16):4456-65.

Nagasawa A, Kubota R, Imamura Y, Nagamine K, Wang Y, Asakawa S, et al. 1997, Genomics. 15;44(3):273-9.

Nagiec M. M., Wells G. B., Lester R. L., Dickson R. C.1993, J. Biol. Chem. 268:22156–22163.

Nakamura T, Alder H, Gu Y, Prasad R, Canaani O, Kamada N, Gale RP, Lange B, Crist WM, Nowell PC, et al. 1993, Proc Natl Acad Sci U S A.;90(10):4631–4635

Navarro R.E., Shim, E., Kohara Y. Y., Singson A. and Blackwell T.K. 2001, Development 127, 3221–3232.

Ndassa, Y. M., Orsi, C., Marto, J. A., Chen, S., and Ross, M. M. 2006, J. Proteome Res. 5, 2789– 2799.

Nielsen PA, Olsen JV, Podtelejnikov AV, Andersen JR, Mann M, Wisniewski JR. 2005, Mol Cell Proteomics. 4(4):402-8.

Nielsen Peter Aa. , Olsen V. Jesper , Alexandre V. Podtelejnikov, Jens R. Andersen, Matthias Mann and Jacek R. Wiśniewski 2005, Molecular & Cellular Proteomics, 4, 402-408.

Nuhse, T. S., Stensballe, A., Jensen, O. N., and Peck, S. C. 2003, Mol. Cell. Proteomics 2, 1234– 1243.

O'Farrell, P.H., 1975, J. Biol. Chem. 250, 4007-4021

Ohsaki,Y., Cheng,J., Suzuki,M., Shinohara,Y., Fujita,A., Fujimoto,T. 2009,. Biochim.Biophys.Acta 1791, 399-407.

Olofsson,S.O., Bostrom,P., Andersson,L., Rutberg,M., Levin,M., Perman,J., Boren,J. 2008, Curr.Opin.Lipidol. 19, 441-447.

Olsen JV, Andersen JR, Nielsen PA, Nielsen ML, Figeys D, Mann M, Wisniewski JR. 2004, Mol Cell Proteomics. 3(1):82-92.

Ong Shao-En, Blagoev Blagoy, Kratchmarova Irina, Kristensen Dan Bach, Steen Hanno ¶, Pandey Akhilesh and Mann Matthias 2002, Molecular & Cellular Proteomics, 1, 376-386.

Opiteck, G. J.; Lewis, K.C.; Jorgenson, J. W.; Anderegg, R. J. 1997, Anal. Chem., 69, 1518–1524.

Ott CM, Lingappa VR. 2003 Cell Sci., 15;115(Pt 10):2003-9

Özcan Sabire and Johnston Mark, Microbiology and Molecular Biology Reviews, 1999, 554–569.

Pappin.C D.J, Hojrup P. and Bleasby.J A. 1993, Current Biology, 3, 6, 327-332.

Patel L, Buckels AC, Kinghorn IJ, et al. 2003, Biochem Biophys Res Commun., 300(2):472-6

Patterson D. Scott &Aebersold H. Ruedi. 2003, Nature Genetics 33, 311 - 323

Pedersen K. Susanne, Harry L. Jenny., Lucille Sebastian, Baker Jasmine, et al., 2003, Journal of proteomics 2 (3), 303–311.

Pedersen SK, Harry JL, Sebastian L, Baker J, et al. 2003, J Proteome Res. 2(3):303-11

Peng, J.; Elias, J. E.; Thoreen, C. C.; Licklider, L. J.; Gygi, S. P. 2003, J. Proteome Res., 2, 43–50.

Perkins, D.N., Pappin, D.J., Creasy, D.M., Cottrell, J.S. 1999, Electrophoresis. 20, 3551-3567

Pieles, U., Zürcher, W., Schär, M., and Moser, H.E. 1993. Nucleic Acids Res. 21, 3191-3196

Pinkse, M. W., Uitto, P. M., Hilhorst, M. J., Ooms, B., and Heck, A. J. 2004, Anal.Chem. 76, 3935– 3943.

Posewitz, M. C., and Tempst, P. 1999. Anal.Chem.71, 2883– 2892.

Prasad, R. Yano T, Sorio C, Nakamura T, Rallapalli R, Gu Y, Leshkowitz D, Croce C M, and Canaani E 1995. Proc. Natl. Acad. Sci. USA 92, 12160−12164.

Pui Ching-Hon, Frankel Lawrence S., et al. 1991, Blood 77: 440–447.

R. Keller, F. Lottspeich, H. E. Meyer (Eds.), VCH, Weinheim 1994, 268, ISBN 3-527-30048-1

Ramus C, Gonzalez de Peredo A, Dahout C, Gallagher M, Garin J 2006.Mol Cell Proteomics. 5(1):68-78

Ramus C, Gonzalez de Peredo A, Dahout C, Gallagher M, Garin J. 2006, Mol Cell Proteomics. 5(1):68-78.

Rand JH, Wu XX, Quinn AS, Taatjes DJ. 2008, Lupus. 17(10):922-30.

Reid E. G. and McLuckey A. Scott.2002, J. Mass Spectrom. 37, 663–675

Reifenberger, E., E. Boles, and M. Ciriacy 1997,. Eur. J. Biochem. 245:324–333.

Righetti P. 1990, J Chromatogr. 7; 516(1):3-22.

Righetti PG, Castagna A, Herbert B, Reymond F, Rossier JS. 2003 Proteomics.3 (8):1397-407.

Righetti PG, Gianazza E. 1987, Methods Biochem Analysis;32:215–278.

Righetti P. G., Castagna, A., Antonioli, P. and Boschetti, E 2005. Electrophoresis 26, 297-319

Righetti P. G.; Campostrini, N.; Pascali, J.; Hamdan, M.; Astner, H. Eur. J. 2004, Mass Spectrom. 10 (3), 335−348.

Robenek H., Buers I., Hofnagel,O., Robenek,M.J., Troyer,D., Severs,N.J. 2009, Biochim.Biophys.Acta 1791, 408-418.

Robinette Martin, S. Cotter, Susan; Van de Water (2001). Quick Look Series in Veterinary Medicine: Hematology. Teton NewMedia.p. 105.ISBN 1-893441-36-9.

Rodríguez-Ortega MJ, Norais N, Bensi Get. al., 2006, Nat Biotechnol. 24(2):191-7.

Roelofsen H, Dijkstra M, Weening D, de Vries MP, Hoek A, Vonk RJ. 2009, Mol Cell Proteomics. 8(2):316-24.

Roepstorff, P. & Fohlman 1984, Biomed. Mass Spectrom.11, 601.

Ros Alexandra , Faupel Michel, Mees Hervé , Oostrum Jan van , Ferrigno Rosaria , Reymond Frédéric , Michel Philippe, Rossier S. Joël, Girault H. Hubert 2002. Proteomics, 2(2):151-6.

Rosenberger S., Connerth,M., Zellnig,G., Daum,G. 2009, Biochim.Biophys.Acta 1791, 379-387.

Rosenow Anja, Arrey N.Tabiwang, Bouwman Freek et al. im Press

Ross P. L. Huang, Y. L. N.; Marchese et al. 2004, Mol. Cell.Proteomics 3 (12), 1154−1169.

Ross P. L.; Huang, Y. L. N.; Marchese, J. N. et al. 2004, Mol. Cell. Proteomics, 3 (12), 1154−1169.

Sabrina Baltruschat 2009, Untersuchung hochmolekularer Proteinkomplexe in menschlichen Leukämien mittels Proteomics Werkzeugen.

Santoni, V., Molloy, M. & Rabilloud, T. 2000, Electrophoresis 21, 1054–1070.

Scheurer, S. B.; Rybak, J. N.; Roesli, C.; Brunisholz, R. A.; Potthast, F.; Schlapbach, R.; Neri, D.; Elia, G. Proteomics 2005, 5, 2718.

Schluesener D, Fischer F, Kruip J, Rögner M, Poetsch A. 2005, Proteomics, 5(5):1317-30.

Seeley EH, Riggs LD, Regnier F. E. 2005, J Chromatogr B Analyt Technol Biomed Life Sci.817(1):81-8.

Sell H, Dietze-Schroeder D, Eckardt K, Eckel J. 2006, Biochem Biophys Res Commun. 12;343(3):700-6.

Serrano, R. 1988, Methods Enzymol.157:533–544

Seward, R. J., Perlman, D. H., Berg, E. A., Hu, J., and Costello, C. E. 2004, in 52nd ASMS Conference on Mass Spectrometry and Allied Topics, Nashville, May 23–27, 2004, Abstr. A042056, American Society for Mass Spectrometry, Santa Fe, NM

Shama P. Mirza, Brian D. Halligan , Andrew S. Greene and Michael Olivier 2007, Physiological Genomics 30:89-94

Shao-En Ong, Blagoy Blagoev, Irina Kratchmarova, Dan Bach Kristensen, Hanno Steen¶, Akhilesh Pandey and Matthias Mann. 2002, Molecular & Cellular Proteomics, 1, 376-386.

Sherry B, Yarlett N, Strupp A, Cerami A., 1992, Proc Natl Acad Sci U S A. 15;89(8):3511-5.

Shevchenko A, Loboda A, Shevchenko A, Ens W, Standing KG.2000 Anal Chem., 72(9):2132-41.

Shukla A.K. and Futrell J.H. 2000, J Mass Spectrom 35, pp. 1069–1090

Sickmann A, Mreyen M, Meyer H.E., 2003, Adv Biochem Eng Biotechnol. 83:141-76.

Sickmann A, Reinders J, Wagner Y, Joppich C, et al.. 2003, Proc Natl Acad Sci U S A. 100(23):13207-12.

Silverman GA, Bird PI, Carrell RW, Church FC, Coughlin PB, Gettins PG, et al. 2001. J Biol Chem. 7;276(36):33293-6.

Simpson, R. J.; Connolly, L. M.; Eddes, J. S.; Pereira, J. J.; Moritz, R. L.; Reid, G. E. 2000, Electrophoresis, 21, 1707-1732.

Small Hamish 1989.Ion chromatography. New York: Plenum Press. ISBN 0-306-43290-0.

Smid A. et al. J. Biol. Chem. 270 1995,. 13534–13540

Snyder R. Lloyd and Dolan W. John. 2006, High-Performance Gradient Elution: The Practical Application of the Linear-Solvent-Strength Model. Wiley Interscience.

Søren S. Jensen, Martin R. Larsen. 2007, Rapid Communications in Mass Spectrometry 21 (22), 3635-3645.

Speers E Anna, Blackler R. Adele, and Wu C Christine 2007, Anal. Chem., 79 (12), pp 4613–4620

Speers E Anna.and Wu C .Christine. 2007, Chem. Rev, 107 (8), 3687–3714

Spyropoulos, I.C., Liakopoulos, T.D., Bagos, P.G., Hamodrakas, S.J. 2004, Bioinformatics. 20, 3258-3260.

Stalder D, Haeberli A, Heller M. 2008, Proteomics.8(3):414-24.

Stass Sanford A. Schumacher, Harold R.; Rock, William R. (2000).Handbook of hematologic pathology. New York, N.Y: Marcel Dekker. pp. 193–194. ISBN 0-8247-0170-4.

Steen H. and Mann M. 2004, Nat Rev Mol Cell Biol. 5: 699-711.

Stewart I. Thomson T.and Figeys, D. 2001, Rapid Commun. Mass Spectrom.15, 2456– 246.

Stoop MP, Lamers RJ, Burgers PC, Sillevis Smitt PA, Hintzen RQ, Luider TM. 2008, J Proteome Res. 7(11):4841-7.

Strupat K, Karas M, Hillenkamp F. 1991, Int. J. Mass Spectrom. Ion Processes 72 (111): 89–102.

Sugiyama N, Masuda T, Shinoda K, Nakamura A, Tomita M, Ishihama Y. 2007, Mol Cell Proteomics. 6(6):1103-9.

Suzuki J, Jin Z-G, Meoli DF, Matoba T, Berk BC., 2006, Circ Res. 31; 98(6):811-7.

Syka E. P. John. , Coon J. Joshua , Schroeder J. Melanie , Shabanowitz Jeffrey , and Hunt F. Donald. 2004, Proceedings of the National Academy of Sciences of the United States of America, 101, 26, 9528-9533

Tanaka K, Arao T, Maegawa M, Matsumoto K, Kaneda H, Kudo K, et al. 2009 Int J Cancer. 1;124(5):1072-80.

Tanaka. K.; I&, Y.; Akita, S.; Yoshida, Y.; Yoshida, T. Presented at the Second Japan-China Joint Symposium on Mass Spectrometry (abstract),Takarazuka Hotel, Osaka, Japan; Sept 15-18, 1987

Tang K, Taranenko NI, Allman SL, Cháng LY, Chen CH. 1994, Rapid Commun. Mass Spectrom. 8 (9): 727–30.

Tauchi-Sato K, Ozeki S, Houjou T, Taguchi R, Fujimoto T. 2002, J Biol Chem., 277(46):44507-12.

Tegeler T.J., Mechref Y.,. Boraas K, Reilly J.P. and Novotny M.V.2004, *Anal.Chem.***76**, 6698–6706

Teuber K, Schiller J, Fuchs B, Karas M, Jaskolla TW. 2010, Chem Phys Lipids. 163(6):552-60.

Thiele,C., Spandl,J. 2008, Curr.Opin.Cell Biol. 20, 378-385.

Thirman MJ, Gill HJ, Burnett RC, et al. 1993 N Engl J Med , 329:909–14

Thompson, A.; Schafer, J.; Kuhn, K.; Kienle, S.; Schwarz, J.; Schmidt, G.; Neumann, T.; Hamon, C., 2003, Anal. Chem. 75 (8), 1895−1904.

Tonge, R.; Shaw, J.; Middleton, B.; Rowlinson, R.; Rayner, S.; Young, J.; Pognan, F.; Hawkins, E.; Currie, I.; Davison 2001, M. Proteomics, 1 (3), 377−396.

Torres J, Stevens TJ, Samsó M. 2003, Biochem. Sci., 28, 137-44

Torta F, Fusi M, Casari CS, Bottani CE, Bachi A. 2009 Proteome Res., 8(4):1932-42.

Umlauf E, Csaszar E, Moertelmaier M, Schuetz GJ, Parton RG, Prohaska R. 2004, J. Biol. Chem. 279, 23699–23709

Urso B, Niesler CU, O'Rahilly S, Siddle K. 2001, Cell Signal. 13(4):279-85.

Uwaje, N. C.; Mueller, N. S.; Maccarrone, G.; Turck, C. W. 2007, Electrophoresis, 28, 1867-1874.

Van der Geer P, Luo K, Sefton BM, Hunter T. In Protein Phosphorylation: a Practical Approach, edn 2. Edited by Hardie D.G. Oxford: Oxford University Press 1999:97-126.

Vanhoenacker Gerd & Sandra Pat 2008, Anal Bioanal Chem., 390:245–248.

Vazquez-Vela ME, Torres N, Tovar AR. 2008, Arch Med Res. 39(8):715-28.

Voet and Voet, 1995, Biochemistry John Wiley and Sons;

Von Heijne, G.1996, Prog Biophys Mol Biol 66 (2) : 113-139.

Wabitsch M, Brenner RE, Melzner I, Braun M, Moller P, Heinze E, et al. 2001, Int J Obes Relat Metab Disord. 25(1):8-15.

Wabitsch M, Heinze E, Debatin KM, Blum WF. 2000, Horm Metab Res. 32(11-12):555-9.

Wall D.B., Berger S.J., Finch J.W., Cohen S.A., Richardson K., Chapman R., Drabble D., Brown J. and Gostick D. 2002, Electrophoresis 23 (18), 3193–3204.

Wallin E, von Heijne G. 1998., Protein Sci. 7(4):1029-38

Wang P, Bouwman FG, Mariman EC. 2009, Proteomics. 9(11):2955-66.

Wang P, Mariman E, Keijer J, Bouwman F, Noben JP, Robben J, et al. 2004, Cell Mol Life Sci., 61(18):2405-17.

Wang, W.; Guo, T.; Song, T.; Lee, C. S.; Balgley, B. M. 2007, Proteomics 7, 1178-87.

Washburn M.P., Wolters D. and Yates J.R., I.I.I 2001 Nat. Biotechnol. 19 (3), 242–247.

Watkins P.A. et al. J. Biol. Chem. 273 1998,. 18210–18219,

Wehr Tim. 2006, LCGC North Amarica, 24 (9)

Wei C, Yang J, Zhu J, Zhang X, Leng W, Wang J, Xue Y, Sun L, Li W, Wang J, Jin Q. 2006, J. Proteome Res. 5(8):1860-5.

Weiss Tatjana, Weiss Joachim 2005, Handbook of Ion Chromatography.Weinheim: Wiley-VCH. ISBN 3-527-28701-9.

Wessel D, Flügge UI. 1984, Anal Biochem. 138(1):141-3.

Whitehouse M. Craig, Dreyer N. Robert, Yamashita Masamichi, Fenn B. John. 1985, Anal.Chem. 57 3 pp. 675–679.

Wiernik, Peter H. (2001). Adult leukemias. New York: B. C. Decker. pp. 3–15. ISBN 1-55009-111-5.

Wilkins R M., Sanchez J.C., Gooley A.A., Appel R.D., I. Humphery-Smith, D.F. Hochstrasser and Williams K.L. 1995, Biotechnol.Genet. Eng. Rev. 13, 19–50.

Wolters D.A., Washburn M.P. and Yates J.R., I.I.I. 2001, Anal. Chem. 73 (23), 5683–5690.

Wu C Christine., MacCoss J Michael., Howell E. Kathryn & John R. Yates, III. 2003, Nature Biotechnology 21, 532 - 538

Wu CC, Howell KE, Neville MC, Yates JR 3rd, McManaman JL.,2000, Electrophoresis 21, 3470–3482

Wu CC, MacCoss MJ, Howell KE, Yates JR 3rd. 2003, Nat Biotechnol. 21(5):532-8.

Wu M, Gan K, Huang C, Tang Y, Chen Q, Tang K, et al. 2006, J Neurooncol. 80(2):133-42.

Wu, W. W.; Wang, G. H.; Baek, S. J.; Shen, R. F. 2006, J. Proteome Res. 5 (3), 651−658.

Wu.J K., Stedding A. and Becker C.H. 1993, Rapid Commun. Mass Spectrom 7, 142–146.

Xu Nafei Zhen Yuejun, Richardson Bonnie, Becklin Robert, Savage R. Justin, Blake Kelly and Peltier M. John 2004, J Am Soc Mass Spectrom. (6):803-22

Yamaguchi,T., Osumi,T., 2009, Biochim.Biophys.Acta 1791, 519-523.

Yang, X.; Wu, H.; Kobayashi, T.; Solaro, R. J.; van Breemen, R. B. Anal. Chem. 2004, 76, 1532−1536

Yao Xudong, Freas Amy, Ramirez Javier, Demirev A. Plamen, and Fenselau Catherine.2001. Anal. Chem., 73 (13), 2836–2842

Yates J.R.I.I.I, Speicher S. Griffin P.R and Hunkapiller T. 1993, Anal Biochem. 214 2, 397–408.

Yermanos D.M. 1975, J. Am. Oil Chem. Soc,

Zeisig T Deniz, Bittner B Claudia, Zeisig B Bernd, García-Cuéllar Maria-Paz, Hess Jay L and Slan Robert K 2005, Oncogene 24, 5525–5532.

Zhang B., McDonald C. and Li L. 2004, Anal.Chem. 76, 992–1001.

Zhou C, Zheng Y, Zhou Y. 2004, Genomics Proteomics Bioinformatics. 2(1):1-5.

Zhou XR, Sun CH, Liu JR, Zhao D. 2008, Growth Horm IGF Res. 18(5):361-8.

Zubarev A R., Kelleher L. Neil, and McLafferty W. Fred. 1998, J. Am. Chem. Soc. 120 (13), 3265–3266

Zubarev A. Roman. 2003, Current Opinion in Biotechnology, 15, 1, 12-16

Zvonic S, Lefevre M, Kilroy G, Floyd ZE, DeLany JP, Kheterpal I, et al. 2007,Mol Cell Proteomics. Jan;6(1):18-28.

Zweytick D, Hrastnik C, Kohlwein SD, Daum G. 2000, Febs letters 470(1):83-7

Zweytick D, Leitner E, Kohlwein SD, Yu C, Rothblatt J, Daum G. 2000, Eur J Biochem. 267(4):1075-82.

Zweytick, D et al. 2000, Biochim. Biophys. Acta 1469, 101–120

11 Publications

11.1 Papers in jorunals

Gerold Jerza, Tabiwang N. Arrey, Victor Wrayb, Qizhen Duc and Peter Winterhalter (2007) **Structural characterization of 132-hydroxy-(132-S)-phaeophytin-a from leaves and stems of Amaranthus tricolor isolated by high-speed countercurrent chromatography**, Innovative Food Science & Emerging Technologies 8(3), 413-418

Rietschel B., Arrey T.N., Meyer B,, Bornemann S., Schuerken M., Karas M., Poetsch A.

(2009) **Elastase digests: new ammunition for shotgun membrane proteomics**, Mol Cell

Proteomics, 8(5), 1029-1043

Rietschel B., Baeumlisberger D., Arrey T,N., Bornemann S., Rohmer M., Schuerken M.,

Karas M., Meyer B. (2009) **The benefit of combining nLC-MALDI-Orbitrap MS data**

with nLC-MALDI-TOF/TOF data for proteomic analyses employing elastase, J

Proteome Res, 8(11), 5317-5324

Rietschel B., Bornemann S., Arrey T.N., Baeumlisberger D., Karas M., Meyer B. (2009)

Membrane protein analysis using an improved peptid in-solution digestion protocol,

Proteomics, 9(24), 5553-5557

Arrey T.N., Rietschel B., Papasotiriou D.G., Bornemann S., Baeumlisberger D., Karas M.,

Meyer B. (2010) **Approaching the Complexity of Elastase-Digested Membrane**

Proteomes using Off-Gel IEF/nLC-MALDI-MS/MS, Anal Chem 82:2145-21499.3.2

Anja Rosenow, Tabiwang Arrey, Freek Bouwman, Jean-Paul Noben, Martin Wabitsch, Edwin C.M. Mariman, Micheal Karas and Johan Renes (2010) **Identification of novel human adipocyte secreted proteins by using SGBS cells** J. Proteome Res., 2010, 9 (10), pp 5389–5401

Dominic Baeumlisberger†, Tabiwang N. Arrey†, Benjamin Rietschel, Marion Rohmer, Dimitrios G. Papasotiriou, Benjamin Mueller, Tobias Beckhaus and Michael Karas (2010) **Labelling elastase digests with TMT – informational gain by identification of poorl detectable peptides with MALDI TOF/TOF mass spectrometry**. PROTEOMICS (pmic.201000288.R1) im press.

Anne Benedikt, Sabrina Baltruschat, Bastian Scholz, Adelheid Bursen, Tabiwang N.Arrey, Björn Meyer, Linda Varagnolo, Albrecht M. Müller, Michael Karas, TheoDingermann and Rolf Marschalek (2010) **Purification of the AF4 and AF4-MLL protein complexes reveals a central role of activated PTEFb kinase and epigenetic dysregulation in t(4;11) leukemia** (submitted)

11.2 Posterpresentationen

43nd Annual Meeting of Germany mass spectrometric Society in Halle

42nd Annual Meeting of Germany mass spectrometric Society in Konstanz

57th ASMS Conference of mass spectrometry in Philadelphia (Poster)

11.3 Eidesstattliche Versicherung

Ich erkläre hiermit an Eides Statt, daß ich die vorgelegte Dissertation über „Method establishment for the high-throughput analysis of complex biological samples using nLMALDI MS/MS “ selbstständig angefertigt und mich anderer Hilfsmittel als der in ihr angegebenen nicht bedient habe, insbesondere, daß aus Schriften Entlehnungen, soweit sie in der Dissertation nicht ausdrücklich als solche mit Angabe der betreffenden Schriften bezeichnet sind, nicht stattgefunden haben.

Frankfurt am Main, den

Supplementary table 1A (S1A)

Gradient 1 Proteins identified with their Protein accession number and number of peptides using gradient 1.

Protein hit	Protein Protein accession	protein description	peptides
1	P00924	Enolase 1	36
2	P00925	Enolase 2	24
3	P01317	Insulin	23
4	P00698	Lysozyme C	16
5	P02666	Beta-casein	14
6	P02662	Alpha-S1-casein	14
7	P00711	Alpha-lactalbumin 2	12
8	P02789	Ovotransferrin	15
9	P00692	Alpha-amylase	16
10	P68082	Myoglobin	8
11	P02768	Serum albumin	9
12	P02769	Serum albumin	11
13	P01012	Ovalbumin	9
14	P68871	Hemoglobin subunit beta	5
15	P00921	Carbonic anhydrase 2	5
16	P00004	Cytochrome c	3
17	P61823	Ribonuclease pancreatic	4
18	P02754	Beta-lactoglobulin	2
19	P02787	Serotransferrin	5

Supplementary table 1B (S1B)

Gradient 2 Proteins identified with their Protein accession number and number of peptides using gradient 2.

Protein hit number	Protein accession number	protein description	peptides
1	P00924	Enolase 1	76
2	P00925	Enolase	52
3	P01317	Insulin	29
4	P00711	Alpha-lactalbumin	21
5	P02666	Beta-casein	33
6	P02789	Ovotransferrin	37
7	P02662	Alpha-S1-casein	23
8	P02769	Serum albumin	28
9	P00692	Alpha-amylase 1	34
10	P00698	Lysozyme C	22
11	P68082	Myoglobin	14
12	P02768	Serum albumin	24
13	P00921	Carbonic anhydrase 2	20
14	P02787	Serotransferrin	23
15	P02754	Beta-lactoglobulin	14
16	P01012	Ovalbumin	16
17	P61823	Ribonuclease pancreatic	8
18	P68871	Hemoglobin subunit beta	11
19	P00004	Cytochrome c	8

Supplementary table 1C (S1C)

Gradient 3 Proteins identified with their Protein accession number and number of peptides using gradient 3.

protein hit number	Protein accession	protein description	peptides
1	P00924	Enolase 1	88
2	P00925	Enolase 2	62
3	P02789	Ovotransferrin	63
4	P01317	Insulin	41
5	P00711	Alpha-lactalbumin	25
6	P00692	Alpha-amylase	53
7	P02666	Beta-casein	36
8	P02769	Serum albumin	42
9	P02662	Alpha-S1-casein	36
10	P02787	Serotransferrin	40
11	P00698	Lysozyme C	36
12	P01012	Ovalbumin	36
13	P00921	Carbonic anhydrase 2	32
14	P02768	Serum albumin	41
15	P68082	Myoglobin	20
16	P02754	Beta-lactoglobulin	17
17	P61823	Ribonuclease pancreatic SV=1	10
18	P00004	Cytochrome c	11
19	P68871	Hemoglobin subunit beta	13
20	P02588	Troponin C	5
21	P22088	Lipase	1

Supplementary table 2 (S2) Summary, showing the identified proteins, protein hit number, Protein accession number r and number of peptides assigned to each protein from YPD (A) and YPO (B) with their respective Protein accession number number.

Supplementary table 2 A (S2 A) YPD

Protein hit number	Protein accession number	protein description	peptides
1	YML008C	reverse complement, Verified ORF, ~~Delta(24)-sterol C-methyltransferase, converts zymosterol to fecosterol in the ergosterol biosynthetic pathway by methylating position C-24; localized to both lip	37
2	YIL124W	Verified ORF, ~~NADPH-dependent 1-acyl dihydroxyacetone phosphate reductase found in lipid particles, ER, and mitochondrial outer membrane; involved in phosphatidic acid biosynthesis; required for spo	15
3	YOR317W	Verified ORF, ~~Long chain fatty acyl-CoA synthetase with a preference for C12:0-C16:0 fatty acids; involved in the activation of imported fatty acids; localized to both lipid particles and mitochondr	20
4	YBR041W	Verified ORF, ~~Fatty acid transporter and very long-chain fatty acyl-CoA synthetase, may form a complex with Faa1p or Faa4p that imports and activates exogenous fatty acids~~	13
5	YKR046C	complement, Verified ORF, ~~Protein of unknown function that co-purifies with lipid particles; expression pattern suggests a role in respiratory growth; computational analysis of large-scale	12
6	YMR246W	Verified ORF, ~~Long chain fatty acyl-CoA synthetase, regulates protein modification during growth in the presence of ethanol, functions to incorporate palmitic acid into phospholipids and neutral l	13
7	YKL140W	Verified ORF, ~~Steryl ester hydrolase, one of three gene products (Yeh1p, Yeh2p, Tgl1p) responsible for steryl ester hydrolase activity and involved in sterol homeostasis; localized to lipid particle	12

8	YBR177C	reverse complement, Verified ORF, ~~Acyl-coenzymeA:ethanol O-acyltransferase that plays a minor role in medium-chain fatty acid ethyl ester biosynthesis; possesses short-chain esterase activity; local	19
9	YPL206C	reverse complement, Verified ORF, ~~Phosphatidyl Glycerol phospholipase C; regulates the phosphatidylglycerol (PG) content via a phospholipase C-type degradation mechanism; contains glycerophosphodie	9
10	YJL034W	Verified ORF, ~~ATPase involved in protein import into the ER, also acts as a chaperone to mediate protein folding in the ER and may play a role in ER export of soluble proteins; regulates the unfolded	9
11	YLR100W	Verified ORF, ~~3-keto sterol reductase, catalyzes the last of three steps required to remove two C-4 methyl groups from an intermediate in ergosterol biosynthesis; mutants are sterol auxotrophs~~	9
12	YCL043C	, reverse complement, Verified ORF, ~~Protein disulfide isomerase, multifunctional protein resident in the endoplasmic reticulum lumen, essential for the formation of disulfide bonds in secretory and cel	6
13	YLR378C	reverse complement, Verified ORF, ~~Essential subunit of Sec61 complex (Sec61p, Sbh1p, and Sss1p); forms a channel for SRP-dependent protein import and retrograde transport of misfolded proteins out	5
14	YMR110C	reverse complement, Verified ORF, ~~Putative fatty aldehyde dehydrogenase, located in the mitochondrial outer membrane and also in lipid particles; has similarity to human fatty aldehyde dehydrogena	7
15	YKR067W	Verified ORF, ~~Glycerol-3-phosphate/dihydroxyacetone phosphate dual substrate-specific sn-1 acyltransferase located in lipid particles and the ER; involved in the stepwise acylation of glycerol-3-pho	12
16	YDR304C	reverse complement, Verified ORF, ~~Peptidyl-prolyl cis-trans isomerase (cyclophilin) of the endoplasmic reticulum, catalyzes the cis-trans isomerization of peptide bonds N-terminal to proline resid	7
17	YDL193W	Verified ORF, ~~Putative prenyltransferase, required for cell viability; proposed to be involved in protein trafficking because tet-repressible mutant shows accumulation of hypoglycosylated forms of C	9
18	YML001W	Verified ORF, ~~GTPase; GTP-binding protein of the rab family; required for homotypic fusion event in vacuole inheritance, for endosome-endosome fusion, similar to mammalian Rab7~~	4

19	YMR148W	Uncharacterized ORF, ~~Putative protein of unknown function; predicted to contain a transmembrane domain; YMR148W is not an essential gene~~	2
20	YPR165W	Verified ORF, ~~GTP-binding protein of the rho subfamily of Ras-like proteins, involved in establishment of cell polarity; regulates protein kinase C (Pkc1p) and the cell wall synthesizing enzyme 1,3	4
21	YOR081C	Treverse complement, Verified ORF, ~~Triacylglycerol lipase involved in TAG mobilization; localizes to lipid particles; potential Cdc28p substrate~~	5
22	YPR139C	reverse complement, Verified ORF, ~~Cytoplasmic protein of unknown function involved in vacuolar protein sorting.~~	6
23	YOR254C	reverse complement, Verified ORF, ~~Essential subunit of Sec63 complex (Sec63p, Sec62p, Sec66p and Sec72p); with Sec61 complex, Kar2p/BiP and Lhs1p forms a channel competent for SRP-dependent and pos	4
24	YHR072W	Verified ORF, ~~Lanosterol synthase, an essential enzyme that catalyzes the cyclization of squalene 2,3-epoxide, a step in ergosterol biosynthesis~~	7
25	YML013W	Verified ORF, ~~Protein involved in ER-associated protein degradation; proposed to coordinate the assembly of proteins involved in ERAD; contains a UBX (ubiquitin regulatory X) domain and a ubiquiti	5
26	YMR313C	reverse complement, Verified ORF, ~~Triacylglycerol lipase of the lipid particle, responsible for all the TAG lipase activity of the lipid particle; contains the consensus sequence motif GXSXG, whic	4
27	YIL009W	Verified ORF, ~~Long chain fatty acyl-CoA synthetase, has a preference for C16 and C18 fatty acids; green fluorescent protein (GFP)-fusion protein localizes to the cell periphery~~	3
28	YAL023C	reverse complement, Verified ORF, ~~Protein O-mannosyltransferase, transfers mannose residues from dolichyl phosphate-D-mannose to protein serine/threonine residues; acts in a complex with Pmt1p, can i	3
29	YKR089C	reverse complement, Verified ORF, ~~Triacylglycerol lipase involved in triacylglycerol mobilization and degradation; found in lipid particles; phosphorylated and activated by Cdc28p; required with Tgl	3
30	YDL052C	reverse complement, Verified ORF, ~~1-acyl-sn-gylcerol-3-phosphate acyltransferase, catalyzes the acylation of lysophosphatidic acid to form phosphatidic acid, a key intermediate in lipid metabolism;	2

31	YPR183W	Verified ORF, ~~Dolichol phosphate mannose (Dol-P-Man) synthase of the ER membrane, catalyzes the formation of Dol-P-Man from Dol-P and GDP-Man; required for glycosyl phosphatidylinositol membrane an	2
32	YKL094W	Verified ORF, ~~Serine hydrolase with sequence similarity to monoglyceride lipase (MGL), localizes to lipid particles~~	3
33	YML128C	reverse complement, Verified ORF, ~~Protein of unknown function; mutant is defective in directing meiotic recombination events to homologous chromatids; the authentic, non-tagged protein is detected i	2
34	YIR038C	reverse complement, Verified ORF, ~~ER associated glutathione S-transferase capable of homodimerization; expression induced during the diauxic shift and throughout stationary phase; functional overlap	2
35	YDL204W	Verified ORF, ~~Protein of unknown function; has similarity to mammalian reticulon proteins; member of the RTNLA (reticulon-like A) subfamily~~	3
36	YGR175C	reverse complement, Verified ORF, ~~Squalene epoxidase, catalyzes the epoxidation of squalene to 2,3-oxidosqualene; plays an essential role in the ergosterol-biosynthesis pathway and is the specific	5
37	YKL212W	Verified ORF, ~~Phosphatidylinositol phosphate (PtdInsP) phosphatase involved in hydrolysis of PtdIns[4]P; transmembrane protein localizes to ER and Golgi; involved in protein trafficking and processing	1
38	YCR017C	reverse complement, Verified ORF, ~~Putative sensor/transporter protein involved in cell wall biogenesis; contains 14-16 transmembrane segments and several putative glycosylation and phosphorylation	2
39	YBR196C-B	Yreverse complement, Uncharacterized ORF, ~~Putative protein of unknown function; identified by expression profiling and mass spectrometry~~	1
40	YDL095W	Verified ORF, ~~Protein O-mannosyltransferase, transfers mannose residues from dolichyl phosphate-D-mannose to protein serine/threonine residues; acts in a complex with Pmt2p, can instead interact wit	2
41	YNL231C	reverse complement, Verified ORF, ~~Phosphatidylinositol transfer protein (PITP) controlled by the multiple drug resistance regulator Pdr1p, localizes to lipid particles and microsomes, controls lev	5
42	YGL228W	Verified ORF, ~~Putative glycosylphosphatidylinositol (GPI)-anchored	4

		protein of unknown function; overexpression causes growth arrest~~	
43	YNL219C	reverse complement, Verified ORF, ~~Mannosyltransferase, involved in N-linked glycosylation; catalyzes the transfer of mannose from Dol-P-Man to lipid-linked oligosaccharides; mutation of the human o	2
44	YOR085W	Verified ORF, ~~Gamma subunit of the oligosaccharyltransferase complex of the ER lumen, which catalyzes asparagine-linked glycosylation of newly synthesized proteins; Ost3p is important for N-glycosyl	1
45	YFL038C	reverse complement, Verified ORF, ~~Ras-like small GTPase, involved in the ER-to-Golgi step of the secretory pathway; complex formation with the Rab escort protein Mrs6p is required for prenylation of Y	1
46	YFL005W	Verified ORF, ~Secretory vesicle-associated Rab GTPase essential for exocytosis; associates with the exocyst component Sec15p and may regulate polarized delivery of transport vesicles to the exocyst at the	1
47	YDR411C	reverse complement, Verified ORF, ~~ER localized derlin-like family member involved in ER stress and homeostasis; not involved in ERAD or substrate retrotranslocation; interacts with CDC48; contains	2
48	YOL048C	reverse complement, Uncharacterized ORF, ~~Putative protein of unknown function~~	4
49	YDR086C	reverse complement, Verified ORF, ~~Subunit of the Sec61p translocation complex (Sec61p-Sss1p-Sbh1p) that forms a channel for passage of secretory proteins through the endoplasmic reticulum membrane,	1
50	YKR094C	reverse complement, Verified ORF, ~~Fusion protein, identical to Rpl40Ap, that is cleaved to yield ubiquitin and a ribosomal protein of the large (60S) ribosomal subunit with similarit	1
51	YDL212W	Verified ORF, ~~Endoplasmic reticulum packaging chaperone, required for incorporation of amino acid permeases into COPII coated vesicles for transport to the cell surface~~	1
52	YER031C	reverse complement, Verified ORF, ~~GTPase of the Ypt/Rab family, very similar to Ypt32p; involved in the exocytic pathway; mediates intra-Golgi traffic or the budding of post-Golgi vesicles from the	1
53	YDR196C	reverse complement, Uncharacterized ORF, ~~Putative dephospho-CoA kinase (DPCK) that catalyzes the final step in Coenzyme A biosynthesis; essential for viability; the authentic, non-tagged protein	1

54	YOR175C	reverse complement, Verified ORF, ~~Lysophospholipid acyltransferase, partially redundant with Slc1p; part of MBOAT family of membrane-bound O-acyltransferases; key component of Lands cycle; may have	1
55	YOR246C	reverse complement, Uncharacterized ORF, ~~Protein with similarity to oxidoreductases, found in lipid particles; required for replication of Brome mosaic virus in S. cerevisiae, which is a model sy	1
56	YNL302C	reverse complement, Verified ORF, ~~Protein component of the small (40S) ribosomal subunit, required for assembly and maturation of pre-40 S particles; mutations in human RPS19 are associ	2
57	YGR038C-B	reverse complement, transposable_element_gene, ~~Retrotransposon TYA Gag and TYB Pol genes; transcribed/translated as one unit; polyprotein is processed to make a nucleocapsid-like	2
58	YPR114W	Uncharacterized ORF, ~~Putative protein of unknown function~~	1
59	YBR265W	Verified ORF, ~~3-ketosphinganine reductase, catalyzes the second step in phytosphingosine synthesis, essential for growth in the absence of exogenous dihydrosphingosine or phytosphingosine, member o	2
60	YJL002C	reverse complement, Verified ORF, ~~Alpha subunit of the oligosaccharyltransferase complex of the ER lumen, which catalyzes asparagine-linked glycosylation of newly synthesized proteins~~	1
61	YDR233C	reverse complement, Verified ORF, ~~ER membrane protein that interacts with exocyst subunit Sec6p and with Yip3p; also interacts with Sbh1p; null mutant has an altered (mostly cisternal) ER morphology	1
62	YDL015C	reverse complement, Verified ORF, ~~Enoyl reductase that catalyzes the last step in each cycle of very long chain fatty acid elongation, localizes to the ER, highly enriched in a structure marking nu	1
63	YLR292C	reverse complement, Verified ORF, ~~Non-essential subunit of Sec63 complex (Sec63p, Sec62p, Sec66p and Sec72p); with Sec61 complex, Kar2p/BiP and Lhs1p forms a channel competent for SRP-dependent an	1
64	YAL047C	reverse complement, Verified ORF, ~Component of the cytoplasmic Tub4p (gamma-tubulin) complex, binds spindle pole bodies and links them to microtubules; has roles in astral microtubule formation and stabiliz	1
65	YGR027C	reverse complement, Verified ORF, ~~Protein component of the small (40S) ribosomal subunit; nearly identical to Rps25Bp and has similarity to	1

		rat S25 ribosomal protein~~	
66	YNL059C	reverse complement, Verified ORF, ~~Nuclear actin-related protein involved in chromatin remodeling, component of chromatin-remodeling enzyme complexes~~	1
67	YEL002C	WBP1 SGDID:S000000728, Chr V from 150013-148721, reverse complement, Verified ORF, ~~Beta subunit of the oligosaccharyl transferase (OST) glycoprotein complex; required for N-linked glycosylation of proteins in the endoplasmic reticulum~~	3
68	YDR294C	reverse complement, Verified ORF, ~~Dihydrosphingosine phosphate lyase, regulates intracellular levels of sphingolipid long-chain base phosphates (LCBPs), degrades phosphorylated long chain bases, p	2
69	YDL075W	Verified ORF, ~~Protein component of the large (60S) ribosomal subunit, nearly identical to Rpl31Bp and has similarity to rat L31 ribosomal protein; associates with the karyopherin Sxm	1
70	YPR147C	reverse complement, Uncharacterized ORF, ~~Putative protein of unknown function; green fluorescent protein (GFP)-fusion protein localizes to the cytoplasm and is induced in response to the DNA-dam	1
71	YDR519W	Verified ORF, ~~Membrane-bound peptidyl-prolyl cis-trans isomerase (PPIase), binds to the drugs FK506 and rapamycin; expression pattern suggests possible involvement in ER protein trafficking~~	1
72	YBR283C	reverse complement, Verified ORF, ~~Subunit of the Ssh1 translocon complex; Sec61p homolog involved in co-translational pathway of protein translocation; not essential~~	1
73	YML048W	Verified ORF, ~~ER localized integral membrane protein that may promote secretion of certain hexose transporters, including Gal2p; involved in glucose-dependent repression~~	1
74	YBR204C	reverse complement, Uncharacterized ORF, ~~Serine hydrolase; YBR204C is not an essential gene~~	1
75	YOR270C	reverse complement, Verified ORF, ~~Subunit a of vacuolar-ATPase V0 domain, one of two isoforms (Vph1p and Stv1p); Vph1p is located in V-ATPase complexes of the vacuole while Stv1p is located in V-ATP	1
76	YGL103W	RPL28 SGDID:S000003071, Chr VII from 310970-311018,311530-311930, Verified ORF, ~~Ribosomal protein of the large (60S) ribosomal	1

		subunit, has similarity to E. coli L15 and rat L27a ribosomal proteins; may have peptidyl transferase activity; can mutat	
77	YML029W	Verified ORF, ~~Protein involved in ER-associated protein degradation (ERAD); component of the Hrd1p complex; interacts with the U1 snRNP-specific protein, Snp1p~~	1
78	YPL095C	reverse complement, Verified ORF, ~~Acyl-coenzymeA:ethanol O-acyltransferase responsible for the major part of medium-chain fatty acid ethyl ester biosynthesis during fermentation; possesses short-ch	1
79	YIL043C	reverse complement, Verified ORF, ~~Microsomal cytochrome b reductase, not essential for viability; also detected in mitochondria; mutation in conserved NADH binding domain of the human ortholog resul	1
80	YHR190W	Verified ORF, ~~Farnesyl-diphosphate farnesyl transferase (squalene synthase), joins two farnesyl pyrophosphate moieties to form squalene in the sterol biosynthesis pathway~~	1
81	YLR065C	reverse complement, Uncharacterized ORF, ~~Putative protein of unknown function; YLR065C is not an essential gene~~	1
82	YDR307W	Uncharacterized ORF, ~~Putative protein of unknown function~~	1
83	YDR518W	Verified ORF, ~~Protein disulfide isomerase of the endoplasmic reticulum lumen, function overlaps with that of Pdi1p; may interact with nascent polypeptides in the ER~~	1
84	YBR106W	Verified ORF, ~~Probable membrane protein, involved in phosphate transport; pho88 pho86 double null mutant exhibits enhanced synthesis of repressible acid phosphatase at high inorganic phosphate conc	1
85	YMR152W	Verified ORF, ~~Protein of unknown function; null mutant displays sensitivity to DNA damaging agents; the authentic, non-tagged protein is detected in highly purified mitochondria in high-throughput	1
86	YML101C	reverse complement, Verified ORF, ~~Protein of unknown function; has a CUE domain that binds ubiquitin, which may facilitate intramolecular monoubiquitination~~	1
87	YFR037C	, reverse complement, Verified ORF, ~Component of the RSC chromatin remodeling complex; essential for viability and mitotic growth; homolog of SWI/SNF subunit Swi3p, but unlike Swi3p, does not activate trans	1
88	YHR098C	reverse complement, Verified ORF, ~Member of the Sec24p family; forms	1

		a complex, with Sec23p, that is involved in sorting of Pma1p into COPII vesicles; peripheral ER membrane protein; potential Cdc28p su	
89	YPR063C	reverse complement, Uncharacterized ORF, ~~ER-localized protein of unknown function~~	1
90	YGR130C	reverse complement, Uncharacterized ORF, ~~Putative protein of unknown function; green fluorescent protein (GFP)-fusion protein localizes to the cytoplasm; specifically phosphorylated in vitro by	1
91	YHR042W	Verified ORF, ~~NADP-cytochrome P450 reductase; involved in ergosterol biosynthesis; associated and coordinately regulated with Erg11p~~	1
92	YGL200C	reverse complement, Verified ORF, ~~Integral membrane component of endoplasmic reticulum-derived COPII-coated vesicles, which function in ER to Golgi transport~~	1
93	YOR089C	reverse complement, Verified ORF, ~~GTPase required for transport during endocytosis and for correct sorting of vacuolar hydrolases; localized in endocytic intermediates; detected in mitochondria; ge	1
94	YOR216C	reverse complement, Verified ORF, ~Golgi matrix protein involved in the structural organization of the cis-Golgi; interacts genetically with COG3 and USO1~	1
95	YDL102W	PVerified ORF, ~Catalytic subunit of DNA polymerase delta; required for chromosomal DNA replication during mitosis and meiosis, intragenic recombination, repair of double strand DNA breaks, and DNA replicat	1
96	YPL094C	reverse complement, Verified ORF, ~~Essential subunit of Sec63 complex (Sec63p, Sec62p, Sec66p and Sec72p); with Sec61 complex, Kar2p/BiP and Lhs1p forms a channel competent for SRP-dependent and po	1
97	YFR031C	reverse complement, Verified ORF, ~~Subunit of the condensin complex; essential SMC chromosomal ATPase family member that forms a complex with Smc4p to form the active ATPase; Smc2p/Smc4p complex bind	1
98	YJL117W	Verified ORF, ~~Endoplasmic reticulum (ER) resident protein required for ER exit of the high-affinity phosphate transporter Pho84p, specifically required for packaging of Pho84p into COPII vesicles~~	1
99	YBR171W	Verified ORF, ~~Non-essential subunit of Sec63 complex (Sec63p, Sec62p, Sec66p and Sec72p); with Sec61 complex, Kar2p/BiP and Lhs1p	1

		forms a channel competent for SRP-dependent and post-translational	
100	YLR020C	reverse complement, Verified ORF, ~~Steryl ester hydrolase, catalyzes steryl ester hydrolysis at the plasma membrane; involved in sterol metabolism~~	1
101	YOR165W	Verified ORF, ~~Protein of unknown function, contains two predicted GTP-binding motifs GXXXXGKS and DXXG near the N-terminus, homolog of the Arabidopsis gene RHD3 (Root Hair Defective)~~	1
102	YGR270W	Verified ORF, ~Protein that localizes to chromatin and has a role in regulation of histone gene expression; has a bromodomain-like region that interacts with the N-terminal tail of histone H3, and an AT	1
103	YAL004W	Dubious ORF, ~~Dubious open reading frame unlikely to encode a protein, based on available experimental and comparative sequence data; completely overlaps verified gene SSA1/YAL005C~~	1
104	YIL046W	Verified ORF, ~F-box protein containing five copies of the WD40 motif, controls cell cycle function, sulfur metabolism, and methionine biosynthesis as part of the ubiquitin ligase complex; interacts with	1
105	YDR537C	reverse complement, Dubious ORF, ~Dubious open reading frame unlikely to encode a protein, almost completely overlaps verified ORF PAD1/YDR538W~	1
106	YIL004C	reverse complement, Verified ORF, ~~Type II membrane protein required for vesicular transport between the endoplasmic reticulum and Golgi complex; v-SNARE with similarity to synaptobrevi	1
107	YHR079C	reverse complement, Verified ORF, ~Serine-threonine kinase and endoribonuclease; transmembrane protein that mediates the unfolded protein response (UPR) by regulating Hac1p synthesis through HAC1 mRNA sp	1
108	YNL139C	reverse complement, Verified ORF, ~Subunit of the THO complex, which is required for efficient transcription elongation and involved in transcriptional elongation-associated recombination; required for La	1
109	YMR243C	reverse complement, Verified ORF, ~~Vacuolar membrane zinc transporter, transports zinc from the cytosol into the vacuole for storage; also has a role in resistance to zinc shock resulting from a su	1
110	YPR080W	Verified ORF, ~~Translational elongation factor EF-1 alpha; also encoded	1

		by TEF2; functions in the binding reaction of aminoacyl-tRNA (AA-tRNA) to ribosomes~~	
111	Q0250	Verified ORF, ~Subunit II of cytochrome c oxidase, which is the terminal member of the mitochondrial inner membrane electron transport chain; one of three mitochondrially-encoded subunits~	1
112	YPR010C	reverse complement, Verified ORF, ~~RNA polymerase I subunit A135~~	1
113	YGL116W	Verified ORF, ~Cell-cycle regulated activator of anaphase-promoting complex/cyclosome (APC/C), which is required for metaphase/anaphase transition; directs ubiquitination of mitotic cyclins, Pds1p, and o	1
114	YIL126W	Verified ORF, ~ATPase component of the RSC chromatin remodeling complex; required for expression of early meiotic genes; essential helicase-related protein homologous to Snf2p~	1
115	YLR206W	Verified ORF, ~Epsin-like protein required for endocytosis and actin patch assembly and functionally redundant with Ent1p; contains clathrin-binding motif at C-terminus~	1
116	YDL161W	Verified ORF, ~Epsin-like protein involved in endocytosis and actin patch assembly and functionally redundant with Ent2p; binds clathrin via a clathrin-binding domain motif at C-terminus~	1
117	YIR002C	reverse complement, Verified ORF, ~Member of the DEAH family of helicases, functions in an error-free DNA damage bypass pathway that involves homologous recombination, binds to flap DNA and stimulates acti	1
118	YIL117C	reverse complement, Verified ORF, ~Pheromone-regulated protein, predicted to have 1 transmembrane segment; induced during cell integrity signaling~	1
119	YHL037C	reverse complement, Dubious ORF, ~~Dubious open reading frame unlikely to encode a functional protein, based on available experimental and comparative sequence data~~	1
121	YDR481C	reverse complement, Verified ORF, ~~Repressible alkaline phosphatase, a glycoprotein localized to the vacuole; regulated by levels of inorganic phosphate and by a system consisting of Pho4p, Pho9p,	1
122	YBR079C	reverse complement, Verified ORF, ~Subunit of the core complex of translation initiation factor 3(eIF3), essential for translation; part of a subcomplex (Prt1p-Rpg1p-Nip1p) that stimulates binding of mRNA	1

123	YGL022W	Verified ORF, ~~Subunit of the oligosaccharyltransferase complex of the ER lumen, which catalyzes asparagine-linked glycosylation of newly synthesized proteins; forms a subcomplex with Ost3p and Ost4	1

Supplementary table 2 B (S2 B) YPD

Protein hit number	Protein accession number	Protein description	peptides
1	YML008C	reverse complement, Verified ORF, ~~Delta(24)-sterol C-methyltransferase, converts zymosterol to fecosterol in the ergosterol biosynthetic pathway by methylating position C-24; localized to both lip	30
2	YKR046C	reverse complement, Verified ORF, ~~Protein of unknown function that co-purifies with lipid particles; expression pattern suggests a role in respiratory growth; computational analysis of large-scale	24
3	YMR110C	reverse complement, Verified ORF, ~~Putative fatty aldehyde dehydrogenase, located in the mitochondrial outer membrane and also in lipid particles; has similarity to human fatty aldehyde dehydrogena	18
4	YKL140W	Verified ORF, ~~Steryl ester hydrolase, one of three gene products (Yeh1p, Yeh2p, Tgl1p) responsible for steryl ester hydrolase activity and involved in sterol homeostasis; localized to lipid particle	14
5	YBR177C	reverse complement, Verified ORF, ~~Acyl-coenzymeA:ethanol O-acyltransferase that plays a minor role in medium-chain fatty acid ethyl ester biosynthesis; possesses short-chain esterase activity; local	18
6	YOR317W	Verified ORF, ~~Long chain fatty acyl-CoA synthetase with a preference for C12:0-C16:0 fatty acids; involved in the activation of imported fatty acids; localized to both lipid particles and mitochondr	15
7	YIL124W	Verified ORF, ~~NADPH-dependent 1-acyl dihydroxyacetone phosphate reductase found in lipid particles, ER, and mitochondrial outer membrane; involved in phosphatidic acid biosynthesis; required for spo	12
8	YGR175C	reverse complement, Verified ORF, ~~Squalene epoxidase, catalyzes the epoxidation of squalene to 2,3-oxidosqualene; plays an essential role in the ergosterol-biosynthesis pathway and is the specific	15
9	YNL208W	Verified ORF, ~~Protein of unknown function; may interact with ribosomes, based on co-purification experiments; authentic, non-tagged protein is detected in purified mitochondria in high-throughpu	3

10	YJR121W	Verified ORF, ~~Beta subunit of the F1 sector of mitochondrial F1F0 ATP synthase, which is a large, evolutionarily conserved enzyme complex required for ATP synthesis; phosphorylated~~	9
11	YNL231C	reverse complement, Verified ORF, ~~Phosphatidylinositol transfer protein (PITP) controlled by the multiple drug resistance regulator Pdr1p, localizes to lipid particles and microsomes, controls lev	9
12	YIR038C	reverse complement, Verified ORF, ~~ER associated glutathione S-transferase capable of homodimerization; expression induced during the diauxic shift and throughout stationary phase; functional overlap	8
13	YFL014W	Verified ORF, ~~Plasma membrane localized protein that protects membranes from desiccation; induced by heat shock, oxidative stress, osmostress, stationary phase entry, glucose depletion, oleate and	6
14	YOL109W	Verified ORF, ~~Peripheral membrane protein of the plasma membrane that interacts with Mid2p; regulates the cell integrity pathway mediated by Pkc1p and Slt2p; the authentic protein is detected in a p	9
15	YMR246W	Verified ORF, ~~Long chain fatty acyl-CoA synthetase, regulates protein modification during growth in the presence of ethanol, functions to incorporate palmitic acid into phospholipids and neutral l	6
16	YLR100W	Verified ORF, ~~3-keto sterol reductase, catalyzes the last of three steps required to remove two C-4 methyl groups from an intermediate in ergosterol biosynthesis; mutants are sterol auxotrophs~~	4
17	YCL043C	reverse complement, Verified ORF, ~~Protein disulfide isomerase, multifunctional protein resident in the endoplasmic reticulum lumen, essential for the formation of disulfide bonds in secretory and cel	7
18	YPR080W	Verified ORF, ~~Translational elongation factor EF-1 alpha; also encoded by TEF2; functions in the binding reaction of aminoacyl-tRNA (AA-tRNA) to ribosomes~~	5
19	YHL001W	Verified ORF, ~~Protein component of the large (60S) ribosomal subunit, nearly identical to Rpl14Ap and has similarity to rat L14 ribosomal protein~~	1
20	YER004W	Verified ORF, ~~Protein of unknown function, localized to the mitochondrial outer membrane; induced by treatment with 8-methoxypsoralen and UVA irradiation~~	4
21	YBR041W	Verified ORF, ~~Fatty acid transporter and very long-chain fatty acyl-CoA synthetase, may form a complex with Faa1p or Faa4p that imports and activates	9

		exogenous fatty acids~~	
22	YBR204C	reverse complement, Uncharacterized ORF, ~~Serine hydrolase; YBR204C is not an essential gene~~	3
23	YNL178W	Verified ORF, ~~Protein component of the small (40S) ribosomal subunit, has apurinic/apyrimidinic (AP) endonuclease activity; essential for viability; has similarity to E. coli S3 and rat S3 ribosoma	5
24	YLR167W	Verified ORF, ~~Fusion protein that is cleaved to yield a ribosomal protein of the small (40S) subunit and ubiquitin; ubiquitin may facilitate assembly of the ribosomal protein into ribosomes; inter	4
25	YDL193W	Verified ORF, ~~Putative prenyltransferase, required for cell viability; proposed to be involved in protein trafficking because tet-repressible mutant shows accumulation of hypoglycosylated forms of C	4
26	YGR192C	reverse complement, Verified ORF, ~~Glyceraldehyde-3-phosphate dehydrogenase, isozyme 3, involved in glycolysis and gluconeogenesis; tetramer that catalyzes the reaction of glyceraldehyde-3-phosphate	5
27	YMR148W	Uncharacterized ORF, ~~Putative protein of unknown function; predicted to contain a transmembrane domain; YMR148W is not an essential gene~~	1
28	YPL145C	reverse complement, Verified ORF, ~~Member of the oxysterol binding protein family, which includes seven yeast homologs; involved in negative regulation of Sec14p-dependent Golgi complex secretory fu	5
29	YDL052C	reverse complement, Verified ORF, ~~1-acyl-sn-gylcerol-3-phosphate acyltransferase, catalyzes the acylation of lysophosphatidic acid to form phosphatidic acid, a key intermediate in lipid metabolism;	3
30	YDR304C	reverse complement, Verified ORF, ~~Peptidyl-prolyl cis-trans isomerase (cyclophilin) of the endoplasmic reticulum, catalyzes the cis-trans isomerization of peptide bonds N-terminal to proline resid	3
31	YMR313C	reverse complement, Verified ORF, ~~Triacylglycerol lipase of the lipid particle, responsible for all the TAG lipase activity of the lipid particle; contains the consensus sequence motif GXSXG, whic	5
32	YKR089C	reverse complement, Verified ORF, ~~Triacylglycerol lipase involved in triacylglycerol mobilization and degradation; found in lipid particles; phosphorylated and activated by Cdc28p; required with Tgl	4
33	YJL052W	Verified ORF, ~~Glyceraldehyde-3-phosphate dehydrogenase, isozyme 1, involved	4

		in glycolysis and gluconeogenesis; tetramer that catalyzes the reaction of glyceraldehyde-3-phosphate to 1,3 bis-phosphogly	
34	YMR152W	Verified ORF, ~~Protein of unknown function; null mutant displays sensitivity to DNA damaging agents; the authentic, non-tagged protein is detected in highly purified mitochondria in high-throughput	4
35	YML073C	reverse complement, Verified ORF, ~~N-terminally acetylated protein component of the large (60S) ribosomal subunit, has similarity to Rpl6Bp and to rat L6 ribosomal protein; binds to	1
36	YML013W	Verified ORF, ~~Protein involved in ER-associated protein degradation; proposed to coordinate the assembly of proteins involved in ERAD; contains a UBX (ubiquitin regulatory X) domain and a ubiquiti	3
37	YLR378C	reverse complement, Verified ORF, ~~Essential subunit of Sec61 complex (Sec61p, Sbh1p, and Sss1p); forms a channel for SRP-dependent protein import and retrograde transport of misfolded proteins out	1
38	YBL002W	Verified ORF, ~~Histone H2B, core histone protein required for chromatin assembly and chromosome function; nearly identical to HTB1; Rad6p-Bre1p-Lge1p mediated ubiquitination regulates transcriptional	1
39	YGL008C	reverse complement, Verified ORF, ~~Plasma membrane H+-ATPase, pumps protons out of the cell; major regulator of cytoplasmic pH and plasma membrane potential; part of the P2 subgroup of cation-transp	3
40	YER065C	reverse complement, Verified ORF, ~~Isocitrate lyase, catalyzes the formation of succinate and glyoxylate from isocitrate, a key reaction of the glyoxylate cycle; expression of ICL1 is induced by growt	1
41	YPR139C	reverse complement, Verified ORF, ~~Cytoplasmic protein of unknown function involved in vacuolar protein sorting.~~	4
42	YOL086C	reverse complement, Verified ORF, ~~Alcohol dehydrogenase, fermentative isozyme active as homo- or heterotetramers; required for the reduction of acetaldehyde to ethanol, the last step in the glycolyt	2
43	YKL094W	Verified ORF, ~~Serine hydrolase with sequence similarity to monoglyceride lipase (MGL), localizes to lipid particles~~	2
44	YOR245C	reverse complement, Verified ORF, ~~Diacylglycerol acyltransferase, catalyzes the terminal step of triacylglycerol (TAG) formation, acylates diacylglycerol using acyl-CoA as an acyl donor, localized t	3

45	YMR173W	Verified ORF, ~~DNA damage-responsive protein, expression is increased in response to heat-shock stress or treatments that produce DNA lesions; contains multiple repeats of the amino acid sequence	1
46	YAL005C	reverse complement, Verified ORF, ~~ATPase involved in protein folding and nuclear localization signal (NLS)-directed nuclear transport; member of heat shock protein 70 (HSP70) family; forms a chaperon	2
47	YGL205W	Verified ORF, ~~Fatty-acyl coenzyme A oxidase, involved in the fatty acid beta-oxidation pathway; localized to the peroxisomal matrix~~	3
48	YEL060C	reverse complement, Verified ORF, ~~Vacuolar proteinase B (yscB), a serine protease of the subtilisin family; involved in protein degradation in the vacuole and required for full protein degradation	1
49	YHR174W	Verified ORF, ~~Enolase II, a phosphopyruvate hydratase that catalyzes the conversion of 2-phosphoglycerate to phosphoenolpyruvate during glycolysis and the reverse reaction during gluconeogenesis;	2
50	YNL202W	Verified ORF, ~~Peroxisomal 2,4-dienoyl-CoA reductase, auxiliary enzyme of fatty acid beta-oxidation; homodimeric enzyme required for growth and sporulation on petroselineate medium; expression indu	1
51	YPL206C	reverse complement, Verified ORF, ~~Phosphatidyl Glycerol phospholipase C; regulates the phosphatidylglycerol (PG) content via a phospholipase C-type degradation mechanism; contains glycerophosphodie	1
52	YIL041W	Verified ORF, ~~BAR domain-containing protein that localizes to both early and late Golgi vesicles; required for adaptation to varying nutrient concentrations, fluid-phase endocytosis, polarization	3
53	YML063W	Verified ORF, ~~Ribosomal protein 10 (rp10) of the small (40S) subunit; nearly identical to Rps1Ap and has similarity to rat S3a ribosomal protein~~	2
54	YLR075W	Verified ORF, ~~Protein component of the large (60S) ribosomal subunit, responsible for joining the 40S and 60S subunits; regulates translation initiation; has similarity to rat L10 ribosomal protein	2
55	YNL055C	reverse complement, Verified ORF, ~~Mitochondrial porin (voltage-dependent anion channel), outer membrane protein required for the maintenance of mitochondrial osmotic stability and mitochondrial membrane	2
56	YDR537C	reverse complement, Dubious ORF, ~~Dubious open reading frame unlikely to encode a protein, almost completely overlaps verified ORF PAD1/YDR538W~~	1

57	YER031C	reverse complement, Verified ORF, ~~GTPase of the Ypt/Rab family, very similar to Ypt32p; involved in the exocytic pathway; mediates intra-Golgi traffic or the budding of post-Golgi vesicles	1
58	YPL095C	reverse complement, Verified ORF, ~Acyl-coenzymeA:ethanol O-acyltransferase responsible for the major part of medium-chain fatty acid ethyl ester biosynthesis during fermentation; possesses short-chain	1
59	YPR183W	Verified ORF, ~~Dolichol phosphate mannose (Dol-P-Man) synthase of the ER membrane, catalyzes the formation of Dol-P-Man from Dol-P and GDP-Man; required for glycosyl phosphatidylinositol membrane	1
60	YLR044C	reverse complement, Verified ORF, ~~Major of three pyruvate decarboxylase isozymes, key enzyme in alcoholic fermentation, decarboxylates pyruvate to acetaldehyde; subject to glucose-, ethanol	1
61	YBL015W	Verified ORF, ~~Acetyl-coA hydrolase, primarily localized to mitochondria; phosphorylated; required for acetate utilization and for diploid pseudohyphal growth~~	2
62	YKL103C	reverse complement, Verified ORF, ~~Vacuolar aminopeptidase yscI; zinc metalloproteinase that belongs to the peptidase family M18; often used as a marker protein in studies of autophagy and cytosol to	3
63	YDL082W	Verified ORF, ~~Protein component of the large (60S) ribosomal subunit, nearly identical to Rpl13Bp; not essential for viability; has similarity to rat L13 ribosomal protein~~	1
64	YPL131W	Verified ORF, ~~Protein component of the large (60S) ribosomal subunit with similarity to E. coli L18 and rat L5 ribosomal proteins; binds 5S rRNA and is required for 60S subunit assembly~~	3
65	YML128C	reverse complement, Verified ORF, ~~Protein of unknown function; mutant is defective in directing meiotic recombination events to homologous chromatids; the authentic, non-tagged protein is detected i	1
66	YFL037W	Verified ORF, ~~Beta-tubulin; associates with alpha-tubulin (Tub1p and Tub3p) to form tubulin dimer, which polymerizes to form microtubules~~	2
67	YDR519W	Verified ORF, ~~Membrane-bound peptidyl-prolyl cis-trans isomerase (PPIase), binds to the drugs FK506 and rapamycin; expression pattern suggests possible involvement in ER protein trafficking~~	1
68	YFL038C	reverse complement, Verified ORF, ~~Ras-like small GTPase, involved in the ER-	1

		to-Golgi step of the secretory pathway; complex formation with the Rab escort protein Mrs6p is required for prenylation of Y	
69	YFL005W	Verified ORF, ~Secretory vesicle-associated Rab GTPase essential for exocytosis; associates with the exocyst component Sec15p and may regulate polarized delivery of transport vesicles to the exocyst at the	1
70	YBL027W	Verified ORF, ~~Protein component of the large (60S) ribosomal subunit, nearly identical to Rpl19Ap and has similarity to rat L19 ribosomal protein; rpl19a and rpl19b single null mutat	1
71	YJR019C	reverse complement, Verified ORF, ~~Peroxisomal acyl-CoA thioesterase likely to be involved in fatty acid oxidation rather than fatty acid synthesis; conserved protein also found in human peroxisomes;	1
72	YNL059C	reverse complement, Verified ORF, ~~Nuclear actin-related protein involved in chromatin remodeling, component of chromatin-remodeling enzyme complexes~~	1
73	YOR081C	reverse complement, Verified ORF, ~~Triacylglycerol lipase involved in TAG mobilization; localizes to lipid particles; potential Cdc28p substrate~~	3
74	YEL039C	reverse complement, Verified ORF, ~~Cytochrome c isoform 2, expressed under hypoxic conditions; electron carrier of the mitochondrial intermembrane space that transfers electrons from ubiquinone-cytochro	1
75	YNL134C	reverse complement, Uncharacterized ORF, ~~Putative protein of unknown function with similarity to dehydrogenases from other model organisms; green fluorescent protein (GFP)-fusion protein localiz	2
76	YDR196C	reverse complement, Uncharacterized ORF, ~~Putative dephospho-CoA kinase (DPCK) that catalyzes the final step in Coenzyme A biosynthesis; essential for viability; the authentic, non-tagged protein	1
77	YFR014C	reverse complement, Verified ORF, ~Calmodulin-dependent protein kinase; may play a role in stress response, many CA++/calmodulan dependent phosphorylation substrates demonstrated in vitro, amino acid seque	1
78	YBR196C-B	reverse complement, Uncharacterized ORF, ~~Putative protein of unknown function; identified by expression profiling and mass spectrometry~~	1
79	YKL085W	Verified ORF, ~~Mitochondrial malate dehydrogenase, catalyzes interconversion of malate and oxaloacetate; involved in the tricarboxylic acid (TCA) cycle; phosphorylated~~	1
80	YJL082W	Verified ORF, ~~Protein of unknown function; the authentic, non-tagged protein is	1

		detected in highly purified mitochondria in high-throughput studies~~	
81	YBL099W	Verified ORF, ~~Alpha subunit of the F1 sector of mitochondrial F1F0 ATP synthase, which is a large, evolutionarily conserved enzyme complex required for ATP synthesis; phosphorylated~~	1
82	YHR072W	Verified ORF, ~~Lanosterol synthase, an essential enzyme that catalyzes the cyclization of squalene 2,3-epoxide, a step in ergosterol biosynthesis~~	2
83	YML001W	Verified ORF, ~~GTPase; GTP-binding protein of the rab family; required for homotypic fusion event in vacuole inheritance, for endosome-endosome fusion, similar to mammalian Rab7~~	2
84	Q0050	Verified ORF, ~Reverse transcriptase required for splicing of the COX1 pre-mRNA, encoded by a mobile group II intron within the mitochondrial COX1 gene~	1
85	YJL034W	Verified ORF, ~~ATPase involved in protein import into the ER, also acts as a chaperone to mediate protein folding in the ER and may play a role in ER export of soluble proteins; regulates the unfolded	1
86	YGR130C	reverse complement, Uncharacterized ORF, ~Putative protein of unknown function; green fluorescent protein (GFP)-fusion protein localizes to the cytoplasm; specifically phosphorylated in vitro by mammal	1
87	YGR270W	Verified ORF, ~Protein that localizes to chromatin and has a role in regulation of histone gene expression; has a bromodomain-like region that interacts with the N-terminal tail of histone H3, and an AT	1
88	YJR104C	reverse complement, Verified ORF, ~~Cytosolic superoxide dismutase; some mutations are analogous to those that cause ALS (amyotrophic lateral sclerosis) in humans~~	1
89	YOL048C	reverse complement, Uncharacterized ORF, ~~Putative protein of unknown function~~	2
90	YIL126W	Verified ORF, ~ATPase component of the RSC chromatin remodeling complex; required for expression of early meiotic genes; essential helicase-related protein homologous to Snf2p~	1
91	YLR340W	Verified ORF, ~~Conserved ribosomal protein P0 similar to rat P0, human P0, and E. coli L10e; shown to be phosphorylated on serine 302~~	1
92	Q0250	Verified ORF, ~Subunit II of cytochrome c oxidase, which is the terminal member of the mitochondrial inner membrane electron transport chain; one of three mitochondrially-encoded subunits~	1

93	YLR206W	Verified ORF, ~Epsin-like protein required for endocytosis and actin patch assembly and functionally redundant with Ent1p; contains clathrin-binding motif at C-terminus~	1
94	YKR067W	Verified ORF, ~~Glycerol-3-phosphate/dihydroxyacetone phosphate dual substrate-specific sn-1 acyltransferase located in lipid particles and the ER; involved in the stepwise acylation of glycerol-3-pho	1
95	YFL010C	reverse complement, Verified ORF, ~~WW domain containing protein of unknown function; binds to Mca1p, a caspase-related protease that regulates H2O2-induced apoptosis; overexpression causes Gi phase g	1
96	YAL004W	Dubious ORF, ~Dubious open reading frame unlikely to encode a protein, based on available experimental and comparative sequence data; completely overlaps verified gene SSA1/YAL005C~	1
97	YGR027C	reverse complement, Verified ORF, ~~Protein component of the small (40S) ribosomal subunit; nearly identical to Rps25Bp and has similarity to rat S25 ribosomal protein~~	1
98	YPL154C	reverse complement, Verified ORF, ~~Vacuolar aspartyl protease (proteinase A), required for the posttranslational precursor maturation of vacuolar proteinases; important for protein turnover after ox	1
99	YLR029C	reverse complement, Verified ORF, ~~Protein component of the large (60S) ribosomal subunit, nearly identical to Rpl15Bp and has similarity to rat L15 ribosomal protein; binds to 5.8 S rRNA~~	1
100	YDR171W	Verified ORF, ~~Small heat shock protein (sHSP) with chaperone activity; forms barrel-shaped oligomers that suppress unfolded protein aggregation; involved in cytoskeleton reorganization after heat s	1
101	YER002W	Verified ORF, ~Constituent of 66S pre-ribosomal particles, involved in 60S ribosomal subunit biogenesis~	1
102	YOR207C	reverse complement, Verified ORF, ~~Second-largest subunit of RNA polymerase III, which is responsible for the transcription of tRNA and 5S RNA genes, and other low molecular weight RNAs~~	1
103	YGR263C	reverse complement, Verified ORF, ~~Sterol deacetylase; component of the sterol acetylation/deacetylation cycle along with Atf2p; integral membrane protein with active site in the ER lumen; green f	1
104	YDL091C	reverse complement, Verified ORF, ~~UBX (ubiquitin regulatory X) domain-	1

		containing protein that interacts with Cdc48p, green fluorescent protein (GFP)-fusion protein localizes to the cytoplasm in a pu	
105	YGL121C	reverse complement, Verified ORF, ~~Proposed gamma subunit of the heterotrimeric G protein that interacts with the receptor Gpr1p; involved in regulation of pseudohyphal growth; requires Gpb1p or Gpb	1
106	YGL116W	Verified ORF, ~~Cell-cycle regulated activator of anaphase-promoting complex/cyclosome (APC/C), which is required for metaphase/anaphase transition; directs ubiquitination of mitotic cyclins, Pds1p,	1
107	YLR295C	reverse complement, Verified ORF, ~~Subunit h of the F0 sector of mitochondrial F1F0 ATP synthase, which is a large, evolutionarily conserved enzyme complex required for ATP synthesis~~	1
108	YOR216C	reverse complement, Verified ORF, ~~Golgi matrix protein involved in the structural organization of the cis-Golgi; interacts genetically with COG3 and USO1~~	1
109	YOR246C	reverse complement, Uncharacterized ORF, ~~Protein with similarity to oxidoreductases, found in lipid particles; required for replication of Brome mosaic virus in S. cerevisiae, which is a model sy	1
110	YOL127W	Verified ORF, ~~Primary rRNA-binding ribosomal protein component of the large (60S) ribosomal subunit, has similarity to E. coli L23 and rat L23a ribosomal proteins; binds to 26S rRNA via a	1
111	YBL003C	reverse complement, Verified ORF, ~~Histone H2A, core histone protein required for chromatin assembly and chromosome function; one of two nearly identical (see also HTA1) subtypes; DNA damage-dependen	1
112	YAR035W	Verified ORF, ~Outer mitochondrial carnitine acetyltransferase, minor ethanol-inducible enzyme involved in transport of activated acyl groups from the cytoplasm into the mitochondrial matrix; phosphorylated	1
113	YNL007C	reverse complement, Verified ORF, ~~Type II HSP40 co-chaperone that interacts with the HSP70 protein Ssa1p; not functionally redundant with Ydj1p due to due to substrate specificity; shares similarit	1
114	YBR265W	Verified ORF, ~~3-ketosphinganine reductase, catalyzes the second step in phytosphingosine synthesis, essential for growth in the absence of exogenous dihydrosphingosine or phytosphingosine, member o	1
115	YDL015C	reverse complement, Verified ORF, ~~Enoyl reductase that catalyzes the last step	1

		in each cycle of very long chain fatty acid elongation, localizes to the ER, highly enriched in a structure marking nu	
116	YGL082W	Uncharacterized ORF, ~~Putative protein of unknown function; predicted prenylation/proteolysis target of Afc1p and Rce1p; green fluorescent protein (GFP)-fusion protein localizes to the cytoplasm	1
117	YML029W	Verified ORF, ~~Protein involved in ER-associated protein degradation (ERAD); component of the Hrd1p complex; interacts with the U1 snRNP-specific protein, Snp1p~~	1
118	YGL195W	Verified ORF, ~~Positive regulator of the Gcn2p kinase activity, forms a complex with Gcn20p; proposed to stimulate Gcn2p activation by an uncharged tRNA~~	1
119	YMR032W	Verified ORF, ~Bud neck-localized, SH3 domain-containing protein required for cytokinesis; regulates actomyosin ring dynamics and septin localization; interacts with the formins, Bni1p and Bnr1p, and wit	1
120	YDL231C	reverse complement, Verified ORF, ~Zinc finger protein containing five transmembrane domains; null mutant exhibits strongly fragmented vacuoles and sensitivity to brefeldin A, a drug which is known to affect	1
121	YOL117W	Verified ORF, ~~Subunit of the COP9 signalosome (CSN) complex that cleaves the ubiquitin-like protein Nedd8 from SCF ubiquitin ligases; plays a role in the mating pheromone response~~	1
122	YJL200C	reverse complement, Verified ORF, ~~Putative mitochondrial aconitase isozyme; similarity to Aco1p, an aconitase required for the TCA cycle; expression induced during growth on glucose, by amino acid star	1
123	YLR430W	Verified ORF, ~~Presumed helicase required for RNA polymerase II transcription termination and processing of RNAs; homolog of Senataxin which causes Ataxia-Oculomotor Apraxia 2 and a dominant form o	1
124	YML028W	Verified ORF, ~~Thioredoxin peroxidase, acts as both a ribosome-associated and free cytoplasmic antioxidant; self-associates to form a high-molecular weight chaperone complex under oxidative stress;	1
125	YER131W	Verified ORF, ~~Protein component of the small (40S) ribosomal subunit; nearly identical to Rps26Ap and has similarity to rat S26 ribosomal protein~~	1
126	YHL037C	reverse complement, Dubious ORF, ~~Dubious open reading frame unlikely to encode a functional protein, based on available experimental and comparative sequence data~~	1

127	YOR059C	reverse complement, Uncharacterized ORF, ~~Hypothetical protein~~	1
128	YGL076C	reverse complement, Verified ORF, ~~Protein component of the large (60S) ribosomal subunit, nearly identical to Rpl7Bp and has similarity to E. coli L30 and rat L7 riboso	1
130	YDR050C	reverse complement, Verified ORF, ~~Triose phosphate isomerase, abundant glycolytic enzyme; mRNA half-life is regulated by iron availability; transcription is controlled by activators Reb1p, Gcr1p, an	1
132	YLR192C	reverse complement, Verified ORF, ~Dual function protein involved in translation initiation as a substoichiometric component of eukaryotic translation initiation factor 3 (eIF3) and required for processin	1

Supplementary table 2 C (S2 C) Proteome of yeast LP. [§]novel LP protein, found on glucose as well on oleate grown cells, [¥] numbers indicate fragments used for identification of proteins from cells grown on glucose (D) or oleate (O). C..cytosol;M..mitochondria; PM..plasma membrane; ER..endoplasmic reticulum; LP..lipid particle; End..endosomes; G..golgi; Mic..microsomes; V..vacuole; Px..peroxisome; N..nucleus; nEnv..nuclear envelope; ext..extrinsic to membrane; mem..integral to membrane; bud..cellular bud; rib..ribosomal subunit; CW..cell wall; R..ribosome; Retro..retrotransposon; mTub..microtubule

Gene name	Systematic name	SGD GFP YPL MIPs	This study	D[¥]	O[¥]	Localisation (SGD)	Protein description
ACH1	YBL015W		✓		2	C / M	CoA transferase activity
ADH1	YOL086C		✓		2	C / PM	Alcohol dehydrogenase
ALG9	YNL219C		✓	2		ER	Mannosyltransferase
ATF1	YOR377W	✓				LP / End	Alcohol acyltransferase
ATP2	YJR121W				9	M	Subunit of mitochondria ATP synthase
AYR1	YIL124W	✓ ✓	✓	15	12	C / ER / LP / M	NADPH-dependent 1-acyl dihydroxyacetone phosphate reductase
BSC2	YDR275W	✓ ✓ ✓				LP	Unknown
COY1	YKL179C	✓ ✓				G	Golgi membrane protein
CPR5[§]	YDR304C		✓	7	3	C / ER	Peptidyl-prolyl cis-trans isomerase
CSR1	YLR380W	✓				LP / C / M / Mic	Phosphatidylinositol transfer protein
CST26	YBR042C	✓ ✓ ✓ ✓				LP	Required for incorporation of stearic acid into phosphatidylinostiol
CWH43	YCR017C		✓	2		PM	Putative sensor/transporter protein
DFM1	YDR411C		✓	2		ER	ER localized derlin-like family member
DGA1	YOR245C	✓	✓		3	LP	Diacylglycerol acyltransferase
DPL1	YDR294C		✓	2		ER	Dihydrosphingosine phosphate lyase
DPM1	YPR183W		✓	2		ER / M	Dolichol phosphate mannose synthase
EHT1	YBR177C	✓ ✓ ✓	✓	19	18	LP / M	Acyl-coenzymeA:ethanol O-acyltransferase
ENO2	YHR174W		✓		2	V / PM / M	Enolase II

ERG1	YGR175C	✓		✓	✓	✓	5	15	ER / LP	Squalene epoxidase
ERG6	YML008C	✓	✓	✓	✓	✓	37	30	ER / LP / M	Delta(24)-sterol C-methyltransferase
ERG7	YHR072W	✓	✓	✓	✓	✓	7	2	ER /LP / PM	Lanosterol synthase
ERG27	YLR100W		✓	✓	✓	✓	9	4	ER / M	3-Keto sterol reductase
FAA1	YOR317W	✓				✓	20	15	LP / PM / M	Long chain fatty acyl-CoA synthetase
FAA3	YIL009W					✓	3		unknown	Long chain fatty acyl-CoA synthetase
FAA4	YMR246W	✓	✓	✓	✓	✓	13	6	LP / C	Long chain fatty acyl-CoA synthetase
FAT1	YBR041W	✓	✓	✓	✓	✓	13	9	PM / LP / Mic / PX	Fatty acid transporter
FMP52	YER004W					✓		4	ER / M	Unknown
GPT2	YKR067W					✓	12		C / ER	sn-1 Acyltransferase
GTT1[§]	YIR038C					✓	2	8	ER / M / PM	glutathione S-transferase
GVP36	YIL041W					✓		3	C / G	BAR domain-containing protein
HFD1	YMR110C	✓	✓		✓	✓	7	18	M / LP / End	Putative fatty aldehyde dehydrogenase
HSP12	YFL014W					✓		6	C / PM / N	Heat shock protein
KAR2	YJL034W					✓	9		ER	ATPase
LAP4	YKL103C					✓		3	V	Vacuolar aminopeptidase
LDB16	YCL005W	✓	✓		✓				LP / M	Unknown
MSC1	YML128C					✓	2		M / ER / PM	Unknown
NUS1	YDL193W	✓		✓		✓	9	4	ER / LP /nEnv	Putative prenyltransferase
OSH4[§]	YPL145C						3	5	C / G / ext	Oxysterol binding protein
OSW5	YMR148W		✓		✓	✓	2		mem	Unknown
PDI1	YCL043C					✓	6	7	ER	Disulfide isomerase
PDR16	YNL231C	✓	✓		✓	✓	5	9	LP / Mic / PM / C	Phosphatidylinositol transfer protein
PET10	YKR046C	✓	✓	✓	✓	✓	12	24	LP	Unknown
PGC1	YPL206C	✓				✓			.P / M	Phosphatidylglycerol phospholipase C
PIL1	YGR086C	✓							C / M / PM	Primary component of eisosomes
PMA1	YGL008C					✓		3	PM / M / mem	Plasma membrane H+-ATPase
PMT1	YDL095W					✓	2		ER	Protein O-mannosyltransferase
PMT2	YAL023C					✓	3		ER	Protein O-mannosyltransferase
POR1	YNL055C					✓		2	M	Mitochondrial porin

POX1	YGL205W		✓		3	PX	Fatty-acyl coenzyme A oxidase
RHO1	YPR165W		✓	4		mem / PX / PM / M / bud	GTP-binding protein
RPL5	YPL131W		✓		3	rib	Protein component of the large (60S) ribosomal subunit
RPL10	YLR075W		✓		2	rib	Protein component of the large (60S) ribosomal subunit
RPS1B	YML063W		✓		2	rib	Ribosomal protein 10 (rp10) of the small (40S) subunit
RPS3	YNL178W		✓		5	rib	Protein component of the small (40S) ribosomal subunit
RPS19B	YNL302C		✓	2		rib	Protein component of the small (40S) ribosomal subunit
RPS31	YLR167W		✓		4	rib / C	Fusion protein that is cleaved to yield a ribosomal protein of the small (40S) subunit and ubiquitin
RRT8	YOL048C	✓ ✓ ✓	✓	4	2	LP	Unknown
RTN2	YDL204W		✓	3		ER / nEnv	Unknown
SEC61	YLR378C		✓	5		ER	Essential subunit of Sec61 complex
SEC63	YOR254C		✓	4		ER / M	Essential subunit of Sec63 complex
SHE10	YGL228W		✓	4		unknown	Putative glycosylphosphatidylinositol (GPI)-anchored protein of unknown function
SLC1	YDL052C	✓ ✓ ✓	✓	2	3	LP	1-Acyl-sn-glycerol-3-phosphate acyltransferase
SNA2	YDR525W-a	✓ ✓ ✓				Mem / C	Unknown
SNX41	YDR425W	✓ ✓				End	Sorting nexin
SRT1	YMR101C	✓				LP	Cis-prenyltransferase
SSA1	YAL005C		✓		2	C / PM / N	ATPase
SSO1	YPL232W	✓ ✓				PM	Plasma membrane t-snare
TDH1	YJL052W	✓	✓		4	C / LP / M / PM / CW	Glyceraldehyde-3-phosphate dehydrogenase, isozyme 1
TDH2	YJR009C	✓				C / LP / M / PM / CW	Glyceraldehyde-3-phosphate dehydrogenase, isozyme 2

TDH3	YGR192C	✓				✓		5	C / LP / M / PM / CW	Glyceraldehyde-3-phosphate dehydrogenase, isozyme 3
TEF1	YPR080W							5	M / R	Translational elongation factor EF-1 alpha
TGL1	YKL140W	✓			✓	✓	12	14	LP / mem	Steryl ester hydrolase
TGL3	YMR313C	✓	✓		✓	✓	4	5	LP	Triacylglycerol lipase
TGL4	YKR089C	✓	✓		✓	✓	3	4	LP	Triacylglycerol lipase
TGL5	YOR081C	✓	✓		✓	✓	5	3	LP	Triacylglycerol lipase
USE1	YGL098W		✓		✓				ER	SNARE protein
TSC10	YBR265W					✓	2		C / ER / M	3-ketosphinganine reductase
TUB2	YFL037W					✓		2	mTub	Beta-tubulin
UBX2[§]	YML013W					✓	5	3	ER / M	Protein involved in ER-associated protein degradation
VPS66[§]	YPR139C					✓	6	4	C	Cytoplasmic protein of unknown function involved in vacuolar protein sorting
WBP1	YEL002C					✓	3		ER / nEnv	Beta subunit of the oligosaccharyl transferase (OST) glycoprotein complex
YBR204C	YBR204C					✓		3	unknown	Serine hydrolase
YDR018C	YDR018C			✓					unknown	Similarity to acyltransferase
YEH1	YLL012W	✓	✓		✓				LP / mem	Steryl ester hydrolase
YIM1	YMR152W	✓		✓		✓		4	LP / C / M	Unknown
YGR038C-B	YGR038C-B					✓	2		Retro	Retrotransposon TYA Gag and TYB Pol genes
YJU3	YKL094W	✓		✓	✓	✓	3	2	LP / C / M / PM	Serine hydrolase
YNL134C	YNL134C					✓		2	C / N	Unknown
YNL208W	YNL208W					✓		3	M / R	Unknown
YOR059C	YOR059C	✓							LP	Unknown
YOR246C	YOR246C	✓		✓	✓				LP	Similarity to oxidoreductase
YPT7[§]	YML001W					✓	4	2	V / M	GTPase
ZEO1	YOL109W					✓		9	M / PM / ext	Peripheral membrane protein

Supplementary table 2 D (S2 D) Summary, showing the identified proteins and number of peptides assigned to each protein from Adipocytes

Protein hit number	Protein accession number	Protein descrition	peptides
1	CO1A2	Collagen alpha-2(I) chain	31
2	CO1A1	Collagen alpha-1(I) chain	20
3	CO3A1	Collagen alpha-1(III) chain	23
4	FINC	Fibronectin	26
5	SPRC	SPARC	9
6	CO6A1	Collagen alpha-1(VI) chain	11
7	VIME	Vimentin	8
8	ACTB	Actin, cytoplasmic 1	6
9	PRDX6	Peroxiredoxin-6	3
10	CO4A2	Collagen alpha-2(IV) chain	6
11	FABP4	Fatty acid-binding protein, adipocyte	3
12	ACTC	Actin, alpha cardiac muscle 1	4
13	FSTL1	Follistatin-related protein 1	4
14	ALDOA	Fructose-bisphosphate aldolase A	4
15	PEDF	Pigment epithelium-derived factor	4
16	HSPB6	Heat shock protein beta-6	1
17	ACBP	Acyl-CoA-binding protein	3
18	CO5A2	Collagen alpha-2(V) chain	3
19	FAS	Fatty acid synthase	5
20	GRP78	78 kDa glucose-regulated protein	4
21	CO4A1	Collagen alpha-1(IV) chain	3
22	LEG1	Galectin-1	3
23	CFAD	Complement factor D	3
24	FABP5	Fatty acid-binding protein, epidermal	2

25	IBP5	Insulin-like growth factor-binding protein 5	2
26	MMP2	72 kDa type IV collagenase	2
27	ALDOC	Fructose-bisphosphate aldolase C	1
28	SERPH	Serpin H1	1
29	TIMP1	Metalloproteinase inhibitor 1	3
30	TBA1B	Tubulin alpha-1B chain	2
31	PGS2	Decorin	1
32	HSP7C	Heat shock cognate 71 kDa protein	3
33	G3P	Glyceraldehyde-3-phosphate dehydrogenase	1
34	IBP4	Insulin-like growth factor-binding protein 4	3
35	TKT	Transketolase	1
36	RTN4	Reticulon-4	1
37	HSPB1	Heat shock protein beta-1	1
38	GPDA	Glycerol-3-phosphate dehydrogenase [NAD+], cytoplasmic	1
39	SAP	Proactivator polypeptide	1
40	PPIA	Peptidyl-prolyl cis-trans isomerase A	1
41	ENOB	Beta-enolase	1
42	ENOA	Alpha-enolase	1
43	CALU	Calumenin	1
44	NID1	Nidogen-1	3
45	CO6A3	Collagen alpha-3(VI) chain	1
46	AKA12	A-kinase anchor protein 12	2
47	PDIA1	Protein disulfide-isomerase	1
48	FLNA	Filamin-A	3
49	CO5A3	Collagen alpha-3(V) chain	3
50	TRFE	Serotransferrin	2
51	ACTN1	Alpha-actinin-1	2
52	ACTN4	Alpha-actinin-4	2
53	TSP1	Thrombospondin-1	2

54	SODM	Superoxide dismutase [Mn], mitochondrial	1
55	SH3L3	SH3 domain-binding glutamic acid-rich-like protein 3	2
56	LDHA	L-lactate dehydrogenase A chain	1
57	CH10	10 kDa heat shock protein, mitochondrial	1
58	PGK1	Phosphoglycerate kinase 1	2
59	CYB5	Cytochrome b5	1
60	LEG3	Galectin-3	1
61	INS	Insulin	1
62	ADIPO	Adiponectin	1
63	6PGD	6-phosphogluconate dehydrogenase, decarboxylating	1
64	TPIS	Triosephosphate isomerase	1
65	GSTP1	Glutathione S-transferase P	2
66	RCN3	Reticulocalbin-3	1
67	CO5A1	Collagen alpha-1(V) chain	1
68	PPIB	Peptidyl-prolyl cis-trans isomerase B	1
69	TCPE	T-complex protein 1 subunit epsilon	1
70	BASP	Brain acid soluble protein 1	1
71	MIF	Macrophage migration inhibitory factor	1
72	IBP3	Insulin-like growth factor-binding protein 3	1
73	CO6A2	Collagen alpha-2(VI) chain	1
74	VINC	Vinculin	1
75	LG3BP	Galectin-3-binding protein	2
76	EF1A1	Elongation factor 1-alpha 1	2
77	MRLC2	Myosin regulatory light chain MRLC2	1
78	MOES	Moesin	1
79	CO2A1	Collagen alpha-1(II) chain	1
80	AK1C1	Aldo-keto reductase family 1 member C1	1
81	IC1	Plasma protease C1 inhibitor	2
82	YBOX1	Nuclease-sensitive element-binding protein 1	1
83	FKB1A	Peptidyl-prolyl cis-trans isomerase FKBP1A	1

84	UCHL1	Ubiquitin carboxyl-terminal hydrolase isozyme L1	1
85	K1881	Protein KIAA1881	3
86	CTGF	Connective tissue growth factor	1
87	PGAM1	Phosphoglycerate mutase 1	1
88	RD23B	UV excision repair protein RAD23 homolog B	1
89	CYTB	Cystatin-B	1
90	PDIA3	Protein disulfide-isomerase A3	1
91	CD44	CD44 antigen	1
92	ECHM	Enoyl-CoA hydratase, mitochondrial	1
93	NEDD8	NEDD8	1
94	TIMP2	Metalloproteinase inhibitor 2	1
95	ZSC29	Zinc finger and SCAN domain-containing protein 29	1
98	COF1	Cofilin-1	1
101	CALR	Calreticulin	1

Supplementary table 2 E (S2 E) Summary, showing the identified proteins and number of peptides assigned to each protein from Pre-adipocytes with their respective Protein accession number, protein hit number and the number of peptides matched to each protein.

Protein hit number	Protein accession number	Protein descrition	peptides
1	FINC	Fibronectin	54
2	CO1A2	Collagen alpha-2(I) chain	39
3	CO1A1	Collagen alpha-1(I) chain	36
4	BGH3	Transforming growth factor-beta-induced protein ig-h3	14
5	TSP1	Thrombospondin-1	17
6	TIMP1	Metalloproteinase inhibitor 1	7
7	PAI1	Plasminogen activator inhibitor 1	14
8	MMP2	72 kDa type IV collagenase	7
9	CO3A1	Collagen alpha-1(III) chain	11
10	SPRC	SPARC	6
11	CO6A1	Collagen alpha-1(VI) chain	10
12	PTX3	Pentraxin-related protein PTX3	7
13	ACTB	Actin, cytoplasmic 1	8
14	VIME	Vimentin	3
15	LOXL2	Lysyl oxidase homolog 2	4
16	SERPH	Serpin H1	5
17	POSTN	Periostin	3
18	POTEE	POTE ankyrin domain family member E	4
19	GDN	Glia-derived nexin	3
20	ACTC	Actin, alpha cardiac muscle 1	4
21	QSOX1	Sulfhydryl oxidase 1	5
22	CO4A2	Collagen alpha-2(IV) chain	7

23	IBP6	Insulin-like growth factor-binding protein 6	2
24	LEG1	Galectin-1	2
25	FBLN3	EGF-containing fibulin-like extracellular matrix protein 1	2
26	LUM	Lumican	4
27	ALBU	Serum albumin	2
28	CALU	Calumenin	3
29	IBP3	Insulin-like growth factor-binding protein 3	3
30	ACTN1	Alpha-actinin-1	2
31	FBN1	Fibrillin-1	3
32	PGS2	Decorin	2
33	MOES	Moesin	2
34	IBP7	Insulin-like growth factor-binding protein 7	4
35	ENOB	Beta-enolase	1
36	CO5A2	Collagen alpha-2(V) chain	3
37	PGBM	Basement membrane-specific heparan sulfate proteoglycan core protein	3
38	TBA1B	Tubulin alpha-1B chain	2
39	LYOX	Protein-lysine 6-oxidase	3
40	TRFL	Lactotransferrin	1
41	1433F	14-3-3 protein eta	1
42	LDHA	L-lactate dehydrogenase A chain	2
43	HSPB1	Heat shock protein beta-1	1
44	TBB2A	Tubulin beta-2A chain	1
45	GREM1	Gremlin-1	1
46	LG3BP	Galectin-3-binding protein	2
47	FABP4	Fatty acid-binding protein, adipocyte	1
48	TSP2	Thrombospondin-2	3
49	PCOC1	Procollagen C-endopeptidase enhancer 1	1
50	COCA1	Collagen alpha-1(XII) chain	2

51	ACTN2	Alpha-actinin-2	1
52	ARP3B	Actin-related protein 3B	1
53	GRP78	78 kDa glucose-regulated protein	2
54	YBOX1	Nuclease-sensitive element-binding protein	1
55	FSTL1	Follistatin-related protein 1	2
56	PROF1	Profilin-1	1
57	ISLR	Immunoglobulin superfamily containing leucine-rich repeat protein	1
58	TPIS	Triosephosphate isomerase	1
59	CO2A1	Collagen alpha-1(II) chain	1
60	INS	Insulin	1
61	ALDOA	Fructose-bisphosphate aldolase A	2
62	CO4A4	Collagen alpha-4(IV) chain	1
63	C2D1A	Coiled-coil and C2 domain-containing protein 1A	1
64	IBP4	Insulin-like growth factor-binding protein 4	2
65	MIF	Macrophage migration inhibitory factor	1
66	CO6A3	Collagen alpha-3(VI) chain	1
67	CO5A1	Collagen alpha-1(V) chain	1
68	PDIA3	Protein disulfide-isomerase A3	1
69	GDIA	Rab GDP dissociation inhibitor alpha	1
70	SH3L3	SH3 domain-binding glutamic acid-rich-like protein 3	1
71	LIPA3	Liprin-alpha-3	1
72	LDHC	L-lactate dehydrogenase C chain	1
73	VINC	Vinculin	1
74	FBLN1	Fibulin-1	2
75	PXDN	Peroxidasin homolog	1
76	FLNA	Filamin-A	1
77	NUCB1	Nucleobindin-1	1
78	FETUA	Alpha-2-HS-glycoprotein	1
79	PPIA	Peptidyl-prolyl cis-trans isomerase A	1

80	PPIB	Peptidyl-prolyl cis-trans isomerase B	1
81	CENPJ	Centromere protein J	1
82	FSCN2	Fascin-2	1
83	NEBU	Nebulin	1
84	CO6A2	Collagen alpha-2(VI) chain	1
85	HRX	Histone-lysine N-methyltransferase HRX	1
86	HSP72	Heat shock-related 70 kDa protein 2	1
87	HSP7C	Heat shock cognate 71 kDa protein SV=1	1
88	PGK1	Phosphoglycerate kinase 1	1
89	RNF31	RING finger protein 31	1
90	K1671	Uncharacterized protein KIAA1671	1
91	DJC21	DnaJ homolog subfamily C member 21	1
92	ITSN1	Intersectin-1	1

Supplementary table 2 F (S2 F) Protein comparison between adipocytes and pre-adipocytas.

overlap	Adipocytes	Pre-adipocytes
Collagen alpha-2(I) chain	Peroxiredoxin-6	Transforming growth factor-beta-induced protein ig-h3
Collagen alpha-1(I) chain	Pigment epithelium-derived factor	Plasminogen activator inhibitor 1
Collagen alpha-1(III) chain	Heat shock protein beta-6	Pentraxin-related protein PTX3
Fibronectin	Acyl-CoA-binding protein	Lysyl oxidase homolog 2
SPARC	Fatty acid synthase	Periostin
Collagen alpha-1(VI) chain	Collagen alpha-1(IV) chain	POTE ankyrin domain family member E
Vimentin	Complement factor D	Glia-derived nexin
Actin, cytoplasmic 1	Fatty acid-binding protein, epidermal	Sulfhydryl oxidase 1
Collagen alpha-2(IV) chain	Insulin-like growth factor-binding protein 5	Insulin-like growth factor-binding protein 6
Fatty acid-binding protein, adipocyte	Fructose-bisphosphate aldolase C	EGF-containing fibulin-like extracellular matrix protein 1
Actin, alpha cardiac muscle 1	Glyceraldehyde-3-phosphate dehydrogenase	Lumican
Follistatin-related protein 1	Transketolase	Serum albumin
Fructose-bisphosphate aldolase A	Reticulon-4	Fibrillin-1
Collagen alpha-2(V) chain	Glycerol-3-phosphate dehydrogenase [NAD+], cytoplasmic	Insulin-like growth factor-binding protein 7
78 kDa glucose-regulated protein	Proactivator polypeptide	Basement membrane-specific heparan sulfate proteoglycan core protein
Galectin-1	Alpha-enolase	Protein-lysine 6-oxidase
72 kDa type IV collagenase	Nidogen-1	Lactotransferrin
Serpin H1	A-kinase anchor protein 12	14-3-3 protein eta OS=Homo sapiens GN=YWHAH PE=1 SV=4

Metalloproteinase inhibitor 1	Protein disulfide-isomerase	Tubulin beta-2A chain
Tubulin alpha-1B chain	Collagen alpha-3(V) chain	Gremlin-1
Decorin =1	Serotransferrin	Thrombospondin-2
Heat shock cognate 71 kDa protein	Alpha-actinin-4	Procollagen C-endopeptidase enhancer 1
Insulin-like growth factor-binding protein 4	Superoxide dismutase [Mn], mitochondrial	Collagen alpha-1(XII) chain sapiens
Heat shock protein beta-1	10 kDa heat shock protein, mitochondrial	Alpha-actinin-2
Peptidyl-prolyl cis-trans isomerase A	Cytochrome b5	Actin-related protein 3B
Beta-enolase	Galectin-3	Profilin-1
Calumenin	Adiponectin	Immunoglobulin superfamily containing leucine-rich repeat protein
Collagen alpha-3(VI) chain	6-phosphogluconate dehydrogenase, decarboxylating	Collagen alpha-4(IV) chain
Filamin-A	Glutathione S-transferase P	Coiled-coil and C2 domain-containing protein 1A
Alpha-actinin-1	Reticulocalbin-3	Rab GDP dissociation inhibitor alpha
Thrombospondin-1	T-complex protein 1 subunit epsilon	Liprin-alpha-3
SH3 domain-binding glutamic acid-rich-like protein 3	Brain acid soluble protein 1	L-lactate dehydrogenase C chain
L-lactate dehydrogenase A chain	Elongation factor 1-alpha 1	Fibulin-1
Phosphoglycerate kinase 1	Myosin regulatory light chain MRLC2	Peroxidasin homolog
Insulin	Aldo-keto reductase family 1 member C1	Nucleobindin-1
Triosephosphate isomerase	Plasma protease C1 inhibitor	Alpha-2-HS-glycoprotein
Collagen alpha-1(V) chain	Peptidyl-prolyl cis-trans isomerase FKBP1A	Centromere protein J
Peptidyl-prolyl cis-trans isomerase B	Ubiquitin carboxyl-terminal hydrolase isozyme L1	Fascin-2
Macrophage migration inhibitory factor	Protein KIAA1881	Nebulin

Insulin-like growth factor-binding protein 3	Connective tissue growth factor	Histone-lysine N-methyltransferase HRX
Collagen alpha-2(VI) chain	Phosphoglycerate mutase 1	Heat shock-related 70 kDa protein 2
Vinculin	UV excision repair protein RAD23 homolog B	RING finger protein 31
Galectin-3-binding protein	Cystatin-B	Uncharacterized protein KIAA1671
Moesin	CD44 antigen	DnaJ homolog subfamily C member 21
Collagen alpha-1(II) chain	Enoyl-CoA hydratase, mitochondrial	Intersectin-1
Nuclease-sensitive element-binding protein 1	NEDD8	
Protein disulfide-isomerase A3	Metalloproteinase inhibitor	
	Zinc finger and SCAN domain-containing protein 29	
	Cofilin-1	
	Calreticulin	

Supplementary table 2 G (S2 G) All identified proteins compared to previous adipokine profiling literature. * indicates all secreted proteins which are not identified by previous proteomic adipocyte studies. # indicates all secreted proteins which are only identified by rodent proteomic adipocyte studies. 1= have been identified before and 0= not been identified before

protein name	known in literature		
	human visceral	human subcutaneous	rodent
extracellular matrix			
Annexin A2	1	1	1
Basement membrane-specific heparan sulfate proteoglycan core protein (HSPG) (perlecan)	1	1	1
Collagen alpha-1(I) chain	1	1	1
Collagen alpha-1(II) chain #	0	0	1
Collagen alpha-1(III) chain	1	1 (not pre)	1
Collagen alpha-1(IV) chain #	0	0	1
Collagen alpha-1(V) chain	1	0	1
Collagen alpha-1(VI) chain	1	1	1
Collagen alpha-1(XII) chain	1	0	1
Collagen alpha-2(I) chain	1	1	1
Collagen alpha-2(IV) chain	1	0	[1
Collagen alpha-2(V) chain #	0	0	1
Collagen alpha-2(VI) chain	1	1	1
Collagen alpha-3(V) chain #	0	0	1
Collagen alpha-3(VI) chain	1	1	1
Connective tissue growth factor (Hypertrophic chondrocyte-specific protein 24) #	0	0	1
Decorin (Bone proteoglycan II) (PG-S2) (PG40)	1	1	1

Dermatopontin (tyrosine-rich acidic matrix protein) (tramp)	1	1	0
EGF-containing fibulin-like extracellular matrix protein 1 (Fibulin-3)	1	0	1
Fibrillin-1 (FBN1) #	0	0	1
Fibronectin (FN) (Cold-insoluble globulin)	1	1	1
Fibulin-1 (FBLN1)	1	0	1
Galectin-1 (Lectin galactoside-binding soluble 1) #	0	0	1
Galectin-3-binding protein (Lectin galactoside-binding soluble 3-binding protein)	1	1	1
Laminin subunit beta-1 (laminin B1 chain)	1	0	1
Laminin subunit gamma-1 (laminin B2 chain)	1	1	1
Lumican (Keratan sulfate proteoglycan lumican)	1	1	1
Nidogen-1 (entactin)	1	1	1
Periostin (PN) (Osteoblast-specific factor 2)	1	1 (not adi)	1
SPARC (Osteonectin)	1	1	1
Thrombospondin-1	1	0	0
Thrombospondin-2	1	0	1
Transforming growth factor-beta-induced protein ig-h3 (Beta ig-h3)	1	1	1
turnover			
72 kDa type IV collagenase (Matrix metalloproteinase-2) (MMP-2)	1	0	1
Lysyl oxidase homolog 2 (Lysyl oxidase-like protein 2) *	0	0	0
Metalloproteinase inhibitor 1 (Tissue inhibitor of metalloproteinases) (TIMP-1)	1	1	1
Metalloproteinase inhibitor 2 (Tissue inhibitor of metalloproteinases 2) (TIMP-2)	1	1	1
Procollagen C-endopeptidase enhancer 1 (PCPE-1)	1	1	1
Protein-lysine 6-oxidase (Lysyl oxidase) #	0	0	1
regulation/signaling			
Acyl-CoA-binding protein (ACBP) (Diazepam-binding inhibitor)	1	0	0
Adiponectin	1	1	1

Alpha-2-HS-glycoprotein (Ba-alpha-2-glycoprotein) (Alpha-2-Z-globulin) (Fetuin-A) *	0	0	0
Apolipoprotein E (Apo-E)	1	0	1
Calreticulin (CRP55) (calregulin)	1	1	1
Calumenin (Crocalbin)	0	1	1
Cystatin-C (Neuroendocrine basic polypeptide)	1	0	1
Follistatin-related protein 1 (Follistatin-like 1)	1	1	1
Gelsolin isoform 1 (Actin-depolymerizing factor)	1	0	1
Glia-derived nexin (GDN) (Protease nexin I) (Protease inhibitor 7) *	0	0	0
Insulin-like growth factor-binding protein 3 (IBP-3) *	0	0	0
Insulin-like growth factor-binding protein 4 (IBP-4)	1	0	1
Insulin-like growth factor-binding protein 5 (IBP-5) *	0	0	0
Insulin-like growth factor-binding protein 6 (IBP-6)	1	0	0
Insulin-like growth factor-binding protein 7 (IBP-7)	1	1 (osteo)	0
Tetranectin (TN) (Plasminogen kringle 4-binding protein) #	0	0	1
Latent-transforming growth factor beta-binding protein 2 (LTBP-2)	1	0	1
Peroxidasin homolog (Vascular peroxidase 1) #	0	0	1
Pigment epithelium-derived factor (PEDF) (Serpin-F1) (EPC-1)	1	[1	1
Plasminogen activator inhibitor 1 (PAI-1)	1	1	1
Prostaglandin-H2 D-isomerase (PGDS2)	1	0	0
Retinol-binding protein 4 (PRBP) (RBP)	1	1	1
Sulfhydryl oxidase 1 (hQSOX)	1	1 (osteo)	1
Sushi repeat-containing protein SRPX *	0	0	0
Vitamin D-binding protein (DBP) (VDB) #	0	0	1
immune regulation			
Clusterin (Complement cytolysis inhibitor)(CLI)	1	0	1
Complement C1r subcomponent	1	1	1
Complement C1s subcomponent	1	1	1

Complement C3	1	0	1
Complement factor D (C3 convertase activator) (Properdin factor D) (Adipsin)	1	0	1
Immunoglobulin superfamily containing leucine-rich repeat protein *	0	0	0
Interleukin-25 (IL-25) #	0	0	1
Macrophage migration inhibitory factor (MIF) #	0	0	1
Pentraxin-related protein PTX3 (Tumor necrosis factor-inducible gene 14 protein)	1	1	0
Plasma protease C1 inhibitor (C1 Inh)	1	1	0
other			
Coiled-coil domain-containing protein 96 *	0	0	0
Transgelin-2 (SM22-alpha homolog)	1	0	1
not known as secreted proteins			
Annexin A5	1	1	1
Peptidyl-prolyl cis-trans isomerase A (PPIase A) (Cyclophilin A)	1	1	1

Table S3A Complete list of proteins identified from both *Halobacterium salinarium* purple membrane (PM) and *Corynebacterium glutamicum* membranes (CB) using elastase/trypsin as proteolytic enzymes.

Supplementary table 3A (S3 A) PM_Elastase

Protein hit number	Protein accession number number	protein description	peptides
1	P02945	Bacteriorhodopsin - Halobacterium salinarium	145
2	P00772	Elastase-1 precursor - Sus scrofa	73
3	Q9HMW9	Dipeptide ABC transporter dipeptide-binding - Halobacterium salinarium	72
4	Q9HR99	Halocyanin-like - Halobacterium salinarium	23
5	Q9HN93	Halocyanin-like - Halobacterium salinarium	28
6	Q9HMG3	Putative uncharacterized protein - Halobacterium salinarium	17
7	Q9HMI3	Dipeptide ABC transporter ATP-binding - Halobacterium salinarium	25
8	Q9HHN1	Na+/H+ antiporter - Halobacterium salinarium	27
9	Q48302	Precursor proteolipid - Halobacterium salinarium	27
10	Q9HSA8	Bifunctional short chain isoprenyl diphosphate synthase - Halobacterium salinarium	17
11	Q9HRQ9	Putative uncharacterized protein - Halobacterium salinarium	15
12	Q9HND9	H+-transporting ATP synthase subunit K - Halobacterium salinarium	17
13	Q9HH34	Rhodopsin - Halobacterium salinarium	16
14	Q9HN95	Cytochrome c oxidase subunit I - Halobacterium salinarium	17
15	Q9HHP3	Vng6297c - Halobacterium salinarium	8
16	Q9HRS0	Putative uncharacterized protein - Halobacterium salinarium	13
17	Q9HN94	Cytochrome c oxidase subunit II - Halobacterium salinarium	7
18	Q9HM69	Cell surface glycoprotein - Halobacterium salinarium	18
19	Q9HNQ9	Protein-export membrane protein - Halobacterium salinarium	7
20	Q9HRR1	Putative uncharacterized protein - Halobacterium salinarium	6

21	Q9HMK0	Putative uncharacterized protein - Halobacterium salinarium	7
22	Q9HRL1	NADH dehydrogenase/oxidoreductase - Halobacterium salinarium	5
23	Q9HRR3	Putative uncharacterized protein - Halobacterium salinarium	10
24	Q9HMW4	Oligopeptide binding protein - Halobacterium salinarium	13
25	Q9HRL3	F420H2:quinone oxidoreductase chain L - Halobacterium salinarium	9
26	Q9HSS1	Phosphoglycerate dehydrogenase - Halobacterium salinarium	12
27	Q9HRR0	Putative uncharacterized protein - Halobacterium salinarium	6
28	Q9HQA6	Putative uncharacterized protein - Halobacterium salinarium	3
29	Q9HMX1	Dipeptide ABC transporter permease - Halobacterium salinarium	9
30	Q9HNB9	Putative uncharacterized protein - Halobacterium salinarium	3
31	Q9HSQ2	Probable protease htpX homolog - Halobacterium salinarium	4
32	Q9HS72	Preprotein translocase subunit secE - Halobacterium salinarium	3
33	Q9HND8	V-type ATP synthase subunit I - Halobacterium salinarium	8
34	Q9HQF4	Membrane protein - Halobacterium salinarium	4
35	Q9HSQ7	ABC transport protein - Halobacterium salinarium	7
36	Q9HRL4	NADH dehydrogenase/oxidoreductase-like protein - Halobacterium salinarium	3
37	Q9HPB1	Preprotein translocase subunit secY - Halobacterium salinarium	5
38	Q9HRL8	NADH dehydrogenase/oxidoreductase - Halobacterium salinarium	5
39	Q9HMT7	Probable thiosulfate sulfurtransferase - Halobacterium salinarium	6
40	Q9HMF1	Aconitase - Halobacterium salinarium	6
41	Q9HRJ0	Putative uncharacterized protein - Halobacterium salinarium	2
42	Q9HN92	Putative uncharacterized protein - Halobacterium salinarium	3
43	Q9HNL7	Putative uncharacterized protein - Halobacterium salinarium	3
44	Q9HS26	Na+/H+ antiporter - Halobacterium salinarium	3
45	Q9HMU9	Nitrite/nitrate reduction protein - Halobacterium salinarium	6
46	Q9HMH4	Putative uncharacterized protein - Halobacterium salinarium	2
47	Q9HRU5	Immunogenic protein - Halobacterium salinarium	3

48	Q9HPQ2	Daunorubicin resistance ABC transporter ATP-binding protein - Halobacterium salinarium	1
49	Q9HMZ3	Heterodisulfide reductase - Halobacterium salinarium	4
50	Q9HMB7	Probable carboxypeptidase - Halobacterium salinarium	3
51	Q9HP76	Putative uncharacterized protein - Halobacterium salinarium	7
52	Q9HNN6	Putative uncharacterized protein - Halobacterium salinarium	3
53	Q9HRF2	Putative uncharacterized protein - Halobacterium salinarium	2
54	Q9HRL2	F420H2:quinone oxidoreductase chain M - Halobacterium salinarium	2
55	Q9HS51	Putative uncharacterized protein - Halobacterium salinarium	2
56	Q9HMB6	Alcohol dehydrogenase - Halobacterium salinarium	2
57	Q9HSC3	Putative protease La homolog type - Halobacterium salinarium	3
58	Q9HR04	Iron-binding protein - Halobacterium salinarium	3
59	Q9HR22	Ribose ABC transporter permease - Halobacterium salinarium	1
60	P33518	Cytochrome c oxidase polypeptide 1 - Halobacterium salinarium	2
61	Q9HPD4	50S ribosomal protein L3P - Halobacterium salinarium	1
62	Q9HQP2	Transmembrane oligosaccharyl transferase - Halobacterium salinarium	5
63	Q9HNE6	Brp-like homolog - Halobacterium salinarium	2
64	Q9HRS4	Putative uncharacterized protein - Halobacterium salinarium	1
65	Q9HSM4	NADP-specific glutamate dehydrogenase B - Halobacterium salinarium	2
66	Q9HQV6	Putative uncharacterized protein - Halobacterium salinarium	2
67	Q9HNS6	Putative uncharacterized protein - Halobacterium salinarium	1
68	Q9HQ64	Membrane anchor - Halobacterium salinarium	2
69	Q9HRR2	Cytochrome b6 - Halobacterium salinarium	2
70	Q9HM89	Elongation factor 1-alpha - Halobacterium salinarium	3
71	Q9HP88	Phytoene dehydrogenase - Halobacterium salinarium	2
72	Q9HHS5	Vng6251h - Halobacterium salinarium	1
73	Q9HS12	Phosphate ABC transporter permease - Halobacterium salinarium	5
74	Q9HRZ2	Probable oxidoreductase - Halobacterium salinarium	2

75	Q9HMU8	Copper transport ATP-binding protein - Halobacterium salinarium	1
76	P25964	Sensory rhodopsin-1 - Halobacterium salinarium	4
77	Q9HRJ9	Cytochrome c oxidase subunit 2 - Halobacterium salinarium	1
78	Q9HRL9	NADH dehydrogenase/oxidoreductase - Halobacterium salinarium	2
79	Q9HRM1	NADH dehydrogenase/oxidoreductase-like protein - Halobacterium salinarium	2
80	Q9HR74	Dimethylsulfoxide reductase - Halobacterium salinarium	1
81	P57697	Proteasome subunit alpha - Halobacterium salinarium	1
82	Q9HQ04	Putative uncharacterized protein - Halobacterium salinarium	3
83	Q9HQU7	Probable oxidoreductase - Halobacterium salinarium	2
84	Q9HR98	Cystathionine gamma-synthase - Halobacterium salinarium	1
85	Q9HN07	Acylaminoacyl-peptidase - Halobacterium salinarium	1
86	Q9HMD1	Preprotein translocase subunit secG - Halobacterium salinarium	1
87	Q9HMX0	Dipeptide ABC transporter permease - Halobacterium salinarium	4
88	Q9HN25	Putative uncharacterized protein - Halobacterium salinarium	1
89	Q9HN22	Putative uncharacterized protein - Halobacterium salinarium	2
90	Q9HQV3	Putative uncharacterized protein - Halobacterium salinarium	2
91	Q9HSC9	N-methyltransferase homolog - Halobacterium salinarium	2
92	Q9HPG8	Ribonucleoside-diphosphate reductase - Halobacterium salinarium	1
93	Q9HP89	Photolyase/cryptochrome - Halobacterium salinarium	2
94	Q9HR16	Proline permease - Halobacterium salinarium	1
95	Q9HHR4	Vng6268c - Halobacterium salinarium	3
96	Q9HRD0	Anion permease - Halobacterium salinarium	1
97	Q9HPA6	Htr-like protein - Halobacterium salinarium	2
98	Q9HN36	Protein export - Halobacterium salinarium	1
99	Q9HQ52	Thymidylate synthase thyX - Halobacterium salinarium	2
100	Q9HMB4	Repair helicase - Halobacterium salinarium	1
101	Q9HQQ9	UDP-glucose dehydrogenase - Halobacterium salinarium	1

102	Q9HP48	NADH oxidase - Halobacterium salinarium	2
103	Q9HRN3	Proteinase IV homolog - Halobacterium salinarium	1
104	Q9HRW7	Putative uncharacterized protein - Halobacterium salinarium	1
105	Q9HNW7	Phosphoribosyl transferase - Halobacterium salinarium	1
106	Q9HNR0	Protein-export membrane protein - Halobacterium salinarium	3
107	Q9HSA0	Putative uncharacterized protein - Halobacterium salinarium	1
108	Q9HT00	Glucosamine--fructose-6-phosphate aminotransferase [isomerizing] - Halobacterium salinarium	2
109	Q9HRT0	Putative uncharacterized protein - Halobacterium salinarium	3
110	Q9HPT9	Carboxylesterase - Halobacterium salinarium	1
111	Q9HRS9	Putative uncharacterized protein - Halobacterium salinarium	1
112	Q9HM85	Elongation factor 2 - Halobacterium salinarium	2
113	Q9HT01	Putative uncharacterized protein - Halobacterium salinarium	1
114	Q9HHK5	Transcription initiation factor IIB 3 - Halobacterium salinarium	1
115	Q9HSF7	Transcription initiation factor IIB 7 - Halobacterium salinarium	1
116	Q9HQU8	Acetyl-CoA synthetase - Halobacterium salinarium	1
117	Q9HNI7	Possible phosphate binding protein - Halobacterium salinarium	1
118	Q9HMQ2	Argininosuccinate synthase - Halobacterium salinarium	2
119	Q9HMK7	Glutaryl-CoA dehydrogenase - Halobacterium salinarium	1
120	Q9HP01	Spermidine/putrescine ABC transporter permease - Halobacterium salinarium	1
121	Q9HQU3	Putative uncharacterized protein - Halobacterium salinarium	1
122	Q9HRC9	Putative uncharacterized protein - Halobacterium salinarium	1
123	Q9HQ43	Putative uncharacterized protein - Halobacterium salinarium	1
124	Q9HN02	Hemolysin protein - Halobacterium salinarium	1
125	Q9HN85	Phosphomannomutase - Halobacterium salinarium	1
126	Q9HPY6	Putative uncharacterized protein - Halobacterium salinarium	1
127	Q9HRH6	Putative uncharacterized protein - Halobacterium salinarium	1
128	Q9HRY9	Putative uncharacterized protein - Halobacterium salinarium	1
129	O59634	Methyl-accepting aerotaxis transducer protein - Halobacterium	2

		salinarium	
130	Q9HPQ5	Htr8 transducer - Halobacterium salinarium	2
131	Q9HNS4	Glycerol-3-phosphate dehydrogenase chain A - Halobacterium salinarium	1
132	Q9HSF1	Putative uncharacterized protein - Halobacterium salinarium	1
133	Q9HQ63	Succinate dehydrogenase hydrophobic membrane anchor subunit - Halobacterium salinarium	2
134	O51960	Putative uncharacterized protein - Halobacterium salinarium	1
135	Q9HSZ8	Glucose-1-phosphate thymidylyltransferase - Halobacterium salinarium	1
136	Q9HMS1	Putative uncharacterized protein - Halobacterium salinarium	1
137	Q9HQH9	Small heat shock protein - Halobacterium salinarium	2
138	Q9HRK1	Cytochrome c oxidase subunit III - Halobacterium salinarium	1
139	Q9HMU6	Putative uncharacterized protein - Halobacterium salinarium	1
140	Q9HNP7	Putative uncharacterized protein - Halobacterium salinarium	1
141	Q48294	Arginine deiminase - Halobacterium salinarium	1
142	Q9HNT9	Putative uncharacterized protein - Halobacterium salinarium	1
143	Q9HT02	Putative uncharacterized protein - Halobacterium salinarium	1
144	Q9HNU0	Formyltetrahydrofolate deformylase - Halobacterium salinarium	1
145	Q9HNY9	Quinolinate phosphoribosyltransferase - Halobacterium salinarium	2
146	Q9HN43	Phosphoenolpyruvate carboxylase - Halobacterium salinarium	1
147	Q9HN54	Putative uncharacterized protein - Halobacterium salinarium	1
148	Q9HP72	Glutamyl-tRNA reductase - Halobacterium salinarium	1
149	Q9HT03	Amino acid ABC transporter, ATP-binding protein - Halobacterium salinarium	1
150	Q9HRR5	Putative uncharacterized protein - Halobacterium salinarium	1
151	Q9HR92	Halobacterial transducer protein 6 - Halobacterium salinarium	2
152	P71415	Transducer HtD protein - Halobacterium salinarium	2
153	Q9HME2	Putative uncharacterized protein - Halobacterium salinarium	1
154	Q9HR32	Type II DNA topoisomerase VI subunit A - Halobacterium	1

		salinarium	
155	Q9HSL5	Cationic amino acid transporter - Halobacterium salinarium	1
156	Q9HN08	Thioredoxin reductase - Halobacterium salinarium	1
157	Q9HR62	Putative uncharacterized protein - Halobacterium salinarium	1
158	Q9HP98	Htr-like protein - Halobacterium salinarium	1
159	Q9HS86	Htr14 transducer - Halobacterium salinarium	1
160	Q9HNB2	DNA repair protein - Halobacterium salinarium	2
161	Q9HPU7	Bacteriorhodopsin related protein - Halobacterium salinarium	1
162	Q9HRJ5	Putative uncharacterized protein - Halobacterium salinarium	1
163	Q9HSJ5	Putative uncharacterized protein - Halobacterium salinarium	1
164	Q9HS40	Putative uncharacterized protein - Halobacterium salinarium	1
165	Q9HHJ8	DNA polymerase B2 - Halobacterium salinarium	1
166	Q9HPX0	Aspartate-semialdehyde dehydrogenase - Halobacterium salinarium	1
167	Q9HQU6	Putative uncharacterized protein - Halobacterium salinarium	1
168	Q9HNI3	Putative uncharacterized protein - Halobacterium salinarium	1
169	Q9HQE2	Putative uncharacterized protein - Halobacterium salinarium	1
170	Q9HN76	Pyruvate dehydrogenase beta subunit - Halobacterium salinarium	1
171	Q48290	Cell division protein ftsZ homolog - Halobacterium salinarium	1
172	Q9HNZ0	L-aspartate oxidase - Halobacterium salinarium	1
173	Q9HQU4	Inosine monophosphate dehydrogenase - Halobacterium salinarium	1
174	Q9HPK2	UPF0100 protein VNG1595C - Halobacterium salinarium	1
175	Q9HNH3	Potassium channel homolog - Halobacterium salinarium	1
176	Q9HNE5	Geranylgeranyl-diphosphate geranylgeranyltransferase - Halobacterium salinarium	1
177	Q9HQ46	Deoxyribodipyrimidine photo-lyase - Halobacterium salinarium	1
178	Q9HP32	CTP synthase - Halobacterium salinarium	1
179	P16102	Halorhodopsin - Halobacterium salinarium	1
180	Q9HRM0	NADH dehydrogenase/oxidoreductase - Halobacterium salinarium	1

181	Q9HSL6	DNA mismatch repair protein mutS 2 - Halobacterium salinarium	1
182	Q9HSW6	Cell division control protein 6 homolog 1 - Halobacterium salinarium	1
183	Q9HQ03	Putative uncharacterized protein - Halobacterium salinarium	1
184	Q9HMJ2	Putative uncharacterized protein - Halobacterium salinarium	1
185	Q9HPW3	Putative uncharacterized protein - Halobacterium salinarium	1
186	Q9HN41	Putative uncharacterized protein - Halobacterium salinarium	1

Supplementary table 3 B (S3 B) PM_Trypsin

Protein hit number	Protein accession number number	Protein protein description	peptides
1	Q9HRQ9	Putative uncharacterized protein - Halobacterium salinarium	3
2	Q9HHP3	Vng6297c - Halobacterium salinarium	3
3	Q9HSA8	Bifunctional short chain isoprenyl diphosphate synthase - Halobacterium salinarium	5
4	Q9HMW9	Dipeptide ABC transporter dipeptide-binding - Halobacterium salinarium	8
5	Q9HS72	Preprotein translocase subunit secE - Halobacterium salinarium	3
6	Q9HHN1	Na+/H+ antiporter - Halobacterium salinarium	5
7	Q9HNQ9	Protein-export membrane protein - Halobacterium salinarium	4
8	Q9HN95	Cytochrome c oxidase subunit I - Halobacterium salinarium	5
9	Q9HMI3	Dipeptide ABC transporter ATP-binding - Halobacterium salinarium	4
10	Q9HMW4	Oligopeptide binding protein - Halobacterium salinarium	6
11	Q9HND9	H+-transporting ATP synthase subunit K - Halobacterium salinarium	1
12	Q9HRR3	Putative uncharacterized protein - Halobacterium salinarium	4
13	Q9HRL4	NADH dehydrogenase/oxidoreductase-like protein - Halobacterium salinarium	2
14	Q9HR99	Halocyanin-like - Halobacterium salinarium	2
15	Q9HRJ0	Putative uncharacterized protein - Halobacterium salinarium	2
16	Q9HS26	Na+/H+ antiporter - Halobacterium salinarium	3
17	Q9HPB1	Preprotein translocase subunit secY - Halobacterium salinarium	4
18	Q9HRL3	F420H2:quinone oxidoreductase chain L - Halobacterium salinarium	2
19	P02945	Bacteriorhodopsin - Halobacterium salinarium	4
20	Q9HRF2	Putative uncharacterized protein - Halobacterium salinarium	3

21	Q9HRR1	Putative uncharacterized protein - Halobacterium salinarium	2
22	Q9HMZ3	Heterodisulfide reductase - Halobacterium salinarium	6
23	Q9HN93	Halocyanin-like - Halobacterium salinarium	2
24	Q9HNE6	Brp-like homolog - Halobacterium salinarium	2
25	Q9HP76	Putative uncharacterized protein - Halobacterium salinarium	5
26	Q9HMG3	Putative uncharacterized protein - Halobacterium salinarium	1
27	Q9HN94	Cytochrome c oxidase subunit II - Halobacterium salinarium	2
28	Q9HSQ2	Probable protease htpX homolog - Halobacterium salinarium	2
29	Q9HPD4	50S ribosomal protein L3P - Halobacterium salinarium	1
30	Q9HR22	Ribose ABC transporter permease - Halobacterium salinarium	1
31	P33518	Cytochrome c oxidase polypeptide 1 - Halobacterium salinarium	2
32	Q9HMX1	Dipeptide ABC transporter permease - Halobacterium salinarium	4
33	Q9HRS0	Putative uncharacterized protein - Halobacterium salinarium	2
34	Q9HRL8	NADH dehydrogenase/oxidoreductase - Halobacterium salinarium	2
35	Q9HSC3	Putative protease La homolog type - Halobacterium salinarium	2
36	Q9HQP2	Transmembrane oligosaccharyl transferase - Halobacterium salinarium	3
37	Q9HHS5	Vng6251h - Halobacterium salinarium	1
38	Q9HRL1	NADH dehydrogenase/oxidoreductase - Halobacterium salinarium	1
39	Q9HRR2	Cytochrome b6 - Halobacterium salinarium	1
40	Q9HSC9	N-methyltransferase homolog - Halobacterium salinarium	2
41	Q9HQ64	Membrane anchor - Halobacterium salinarium	1
42	Q9HHR4	Vng6268c - Halobacterium salinarium	3
43	Q9HMU9	Nitrite/nitrate reduction protein - Halobacterium salinarium	3
44	Q9HMD1	Preprotein translocase subunit secG - Halobacterium	1

		salinarium	
45	P25964	Sensory rhodopsin-1 - Halobacterium salinarium	2
46	Q9HQU7	Probable oxidoreductase - Halobacterium salinarium	1
47	Q9HN07	Acylaminoacyl-peptidase - Halobacterium salinarium	1
48	Q9HR16	Proline permease - Halobacterium salinarium	2
49	Q9HP89	Photolyase/cryptochrome - Halobacterium salinarium	3
50	Q9HN25	Putative uncharacterized protein - Halobacterium salinarium	1
51	Q9HQ04	Putative uncharacterized protein - Halobacterium salinarium	2
52	Q9HRM1	NADH dehydrogenase/oxidoreductase-like protein - Halobacterium salinarium	2
53	Q9HSQ7	ABC transport protein - Halobacterium salinarium	3
54	Q9HS12	Phosphate ABC transporter permease - Halobacterium salinarium	2
55	Q9HRZ2	Probable oxidoreductase - Halobacterium salinarium	1
56	Q9HMX0	Dipeptide ABC transporter permease - Halobacterium salinarium	2
57	Q9HT01	Putative uncharacterized protein - Halobacterium salinarium	1
58	Q9HND8	V-type ATP synthase subunit I - Halobacterium salinarium	1
59	Q9HRC9	Putative uncharacterized protein - Halobacterium salinarium	1
60	Q9HMK7	Glutaryl-CoA dehydrogenase - Halobacterium salinarium	2
61	Q9HRT0	Putative uncharacterized protein - Halobacterium salinarium	2
62	Q9HRR0	Putative uncharacterized protein - Halobacterium salinarium	1
63	Q9HNN6	Putative uncharacterized protein - Halobacterium salinarium	1
64	Q9HRN3	Proteinase IV homolog - Halobacterium salinarium	1
65	Q9HRL2	F420H2:quinone oxidoreductase chain M - Halobacterium salinarium	1
66	Q9HSF7	Transcription initiation factor IIB 7 - Halobacterium salinarium	1
67	Q9HPU7	Bacteriorhodopsin related protein - Halobacterium salinarium	2
68	Q9HRL9	NADH dehydrogenase/oxidoreductase - Halobacterium salinarium	1

69	Q9HPA6	Htr-like protein - Halobacterium salinarium	1
70	Q9HRQ7	Putative uncharacterized protein - Halobacterium salinarium	1
71	Q9HQ63	Succinate dehydrogenase hydrophobic membrane anchor subunit - Halobacterium salinarium	1
72	O51960	Putative uncharacterized protein - Halobacterium salinarium	1
73	Q9HR80	Putative uncharacterized protein - Halobacterium salinarium	2
74	Q9HT02	Putative uncharacterized protein - Halobacterium salinarium	1
75	Q9HQ52	Thymidylate synthase thyX - Halobacterium salinarium	1
76	Q9HR62	Putative uncharacterized protein - Halobacterium salinarium	1
77	P33741	Sensory rhodopsin I transducer - Halobacterium salinarium	2
78	Q9HPQ5	Htr8 transducer - Halobacterium salinarium	1
79	Q9HSL5	Cationic amino acid transporter - Halobacterium salinarium	1
80	Q9HQU6	Putative uncharacterized protein - Halobacterium salinarium	1
81	Q9HQP7	Putative uncharacterized protein - Halobacterium salinarium	1
82	Q9HN76	Pyruvate dehydrogenase beta subunit - Halobacterium salinarium	1
83	Q9HPT0	Putative uncharacterized protein - Halobacterium salinarium	1
84	Q48290	Cell division protein ftsZ homolog - Halobacterium salinarium	1
85	Q9HN02	Hemolysin protein - Halobacterium salinarium	1
86	Q9HRM0	NADH dehydrogenase/oxidoreductase - Halobacterium salinarium	1
87	Q9HNK1	Putative uncharacterized protein - Halobacterium salinarium	1
88	Q9HQ81	TRK potassium uptake system protein - Halobacterium salinarium	1
89	Q9HN58	Putative uncharacterized protein - Halobacterium salinarium	1
90	Q9HQV1	Putative uncharacterized protein - Halobacterium salinarium	1
91	Q9HQQ6	Glycosyl transferase-like - Halobacterium salinarium	1
92	Q9HSA3	L-tyrosine decarboxylase - Halobacterium salinarium	1
93	Q9HP84	Halobacterial transducer protein 4 - Halobacterium salinarium	1

94	Q9HNI9	Transport protein - Halobacterium salinarium	1
95	O52005	Putative uncharacterized protein - Halobacterium salinarium	1
96	Q9HNG0	Proline dehydrogenase - Halobacterium salinarium	1
97	Q9HPX5	Htr-like protein - Halobacterium salinarium	1
98	Q9HR54	Putative uncharacterized protein - Halobacterium salinarium	1
99	Q9HMH6	Putative uncharacterized protein - Halobacterium salinarium	1
100	P57699	Potassium-transporting ATPase B chain - Halobacterium salinarium	1
101	Q9HRC5	Putative uncharacterized protein - Halobacterium salinarium	1
102	Q9HRS2	Putative uncharacterized protein - Halobacterium salinarium	1
103	Q9HR04	Iron-binding protein - Halobacterium salinarium	1
104	Q9HNS2	Glycerol-3-phosphate dehydrogenase chain C - Halobacterium salinarium	1
105	P16102	Halorhodopsin - Halobacterium salinarium	1
106	Q9HMW2	Oligopeptide transport permease protein - Halobacterium salinarium	1
107	Q9HRK2	Putative uncharacterized protein - Halobacterium salinarium	1
108	Q9HRY6	Probable acetyltransferase - Halobacterium salinarium	1
109	Q9HP19	Putative uncharacterized protein - Halobacterium salinarium	1
110	Q9HNB2	DNA repair protein - Halobacterium salinarium	1
111	Q9HHI4	Vng6379c - Halobacterium salinarium	1
112	Q9HSJ6	Putative uncharacterized protein - Halobacterium salinarium	1
113	Q9HQS2	CDP-diacylglycerol-glycerol-3-phosphate 3-phosphatidyltransferase - Halobacterium salinarium	1
114	Q9HQS6	Putative uncharacterized protein - Halobacterium salinarium	1
115	Q9HHW0	Vng6203h - Halobacterium salinarium	1
116	Q9HQ56	Putative uncharacterized protein - Halobacterium salinarium	1

Supplementary table 3 C (S3 C) CB_Elastase

Protein hit number	Protein accession number number	Protein descrition	peptides
1	P00772	Elastase-1	152
2	Cg2705	MALTOSE-BINDING PROTEIN PRECURSOR ABC-type sugar transport system, periplasmic component	37
3	Cg2409	CYTOCHROME C OXIDASE CHAIN II	35
4	Cg1537	GLUCOSE-SPECIFIC ENZYME II BC COMPONENT OF PTS	41
5	Cg0446	succinate dehydrogenase A	46
6	Cg3138	Membrane protease subunit, stomatin/prohibitin homolog	18
7	Cg0587	ELONGATION FACTOR TU	22
8	Cg2911	ABC-type Mn/Zn transport system, secreted Mn/Zn-binding (lipo)protein (surface adhesin)	10
9	Cg3009	hypothetical protein predicted by Glimmer/Critica	14
10	Cg2875	hypothetical protein predicted by Glimmer	14
11	Cg0447	succinate dehydrogenase B	16
12	Cg1656	NADH DEHYDROGENASE	13
13	Cg2195	putative secreted or membrane protein	4
14	Cg1368	ATP SYNTHASE ALPHA SUBUNIT	12
15	Cg2780	PROBABLE CYTOCHROME C OXIDASE POLYPEPTIDE SUBUNIT	12
16	Cg1429	putative membrane protein	6
17	Cg2196	putative secreted or membrane protein	9
18	Cg0924	ABC-type cobalamin/Fe3+-siderophores transport sys	5
19	Cg2403	CYTOCHROME B, MEMBRANE PROTEIN	11
20	Cg0445	succinate dehydrogenase CD	6
21	Cg2404	RIESKE IRON-SULFUR PROTEIN	13

22	Cg1685	Sec-independent protein secretion pathway component	8
23	Cg2444	hypothetical protein predicted by Glimmer/Critica	7
24	Cg2708	ABC-type sugar transport system, ATPase component	9
25	Cg2843	ABC-type phosphate transport system, ATPase component	9
26	Cg1290	Homocysteine methyltransferase	6
27	Cg1081	ABC-type multidrug transport system, ATPase component	5
28	Cg0957	FATTY ACID SYNTHASE	10
29	Cg1791	GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE	6
30	Cg2181	ABC-type peptide transport system, secreted component	6
31	Cg3186	Trehalose corynomycolyl transferase	6
32	Cg1363	ATP SYNTHASE C CHAIN	3
33	Cg0737	ABC-type transport system, secreted lipoprotein component	5
34	Cg1880	THREONYL-TRNA SYNTHETASE	7
35	Cg1121	Permease of the major facilitator superfamily	7
36	Cg0791	PYRUVATE CARBOXYLASE	6
37	Cg0766	ISOCITRATE DEHYDROGENASE	5
38	Cg1362	ATP SYNTHASE F0 SUBUNIT 6	3
39	Cg2291	PYRUVATE KINASE	1
40	Cg2137	GLUTAMATE SECRETED BINDING PROTEIN	3
41	Cg1787	PHOSPHOENOLPYRUVATE CARBOXYLASE	3
42	Cg2157	TELLURIUM RESISTANCE MEMBRANE PROTEIN	3
43	Cg2405	CYTOCHROME C1	5
44	Cg1001	LARGE CONDUCTANCE MECHANOSENSITIVE CHANNEL	4
45	Cg1169	Na+-dependent transporters of the SNF family	4
46	Cg1364	ATP synthase B chain	2
47	Cg2211	putative membrane protein	6
48	Cg1556	conserved hypothetical protein	2
49	Cg2499	GLYCYL-TRNA SYNTHETASE (GLYCINE--TRNA LIGASE)	4
50	Cg1790	PHOSPHOGLYCERATE KINASE	6
51	Cg0834	Bacterial extracellular solute-binding protein, fa	3

52	Cg2496	putative secreted protein	3
53	Cg0656	50S RIBOSOMAL PROTEIN L17	1
54	Cg1813	PUTATIVE CARBAMOYL-PHOSPHATE SYNTHASE SUBUNIT	3
55	Cg0583	ELONGATION FACTOR G	3
56	Cg1111	ENOLASE	4
57	Cg1579	putative secreted protein	1
58	Cg3227	PUTATIVE L-LACTATE DEHYDROGENASE	5
59	Cg0647	preprotein translocase subunit SecY	5
60	Cg3219	L-LACTATE DEHYDROGENASE	3
61	Cg1737	ACONITASE	3
62	Cg0935	conserved hypothetical protein	2
63	Cg0756	PUTATIVE CARBON STARVATION PROTEIN A	6
64	Cg1365	H+-ATPASE DELTA SUBUNIT	2
65	Cg3195	Flavin-containing monooxygenase (FMO)	2
66	Cg1437	KETOL-ACID REDUCTOISOMERASE	3
67	Cg0528	putative secreted protein	1
68	Cg1796	putative membrane protein - C. ammoniagenes RibX homolog	1
69	Cg2613	MALATE DEHYDROGENASE OXIDOREDUCTASE PROTEIN	2
70	Cg1366	PROBABLE ATP SYNTHASE ALPHA CHAIN PROTEIN	5
71	Cg2262	Signal recognition particle GTPase	1
72	Cg0047	conserved hypothetical protein	2
73	Cg4005	putative secreted protein |	2
74	Cg2695	ABC-type transport system, ATPase component	2
75	Cg3008	PORIN	2
76	Cg0508	iron/thiamine transport system, secreted component	1
77	Cg2192	MALATE:QUINONE OXIDOREDUCTASE OXIDOREDUCTASE	2
78	Cg3365	putative ribitol transport membrane protein	2
79	Cg1280	2-OXOGLUTARATE DEHYDROGENASE	1
80	Cg2408	putative membrane protein	1

81	Cg2292	PUTATIVE PROLIPOPROTEIN DIACYLGLYCEROL TRANSFERASE	1
82	Cg1586	ARGININOSUCCINATE SYNTHASE	2
83	Cg0594	50S RIBOSOMAL PROTEIN L3	1
84	Cg3244	conserved hypothetical protein	1
85	Cg3100	Heat shock protein hsp70	1
86	Cg2325	hypothetical protein predicted by Glimmer/Critica	3
87	Cg3149	Aminotransferases class-I	3
88	Cg2466	PYRUVATE DEHYDROGENASE E1 COMPONENT	3
89	Cg0928	ABC-type cobalamin/Fe3+-siderophores transport system, ATPase component	1
90	Cg2470	SECRETED ABC TRANSPORTER SUBSTRATE-BINDING PROTEIN	2
91	Cg0303	2-ISOPROPYLMALATE SYNTHASE	1
92	Cg2558	related to aldose 1-epimerase	3
93	Cg2862	PHOSPHORIBOSYLFORMYLGLYCINAMIDINE SYNTHASE	1
94	Cg2523	4-ALPHA-GLUCANOTRANSFERASE	2
95	Cg0953	Na+/proline, Na+/panthothenate symporter or related permease	2
96	Cg2120	SUGAR SPECIFIC PTS SYSTEM, FRUCTOSE/MANNITOL-SPECIFIC TRANSPORT PROTEIN	2
97	Cg1238	putative membrane protein	3
98	Cg1911	putative secreted protein	2
99	Cg0952	PUTATIVE INTEGRAL MEMBRANE PROTEIN	3
100	Cg1227	putative membrane protein	2
101	Cg3429	Preprotein translocase subunit YidC	1
102	Cg2845	ABC-type phosphate transport system, permease component	3
103	Cg0414	cell surface polysaccharide biosynthesis / Chain length determinant protein	2
104	Cg0523	membrane protein required for cytochrome c biosynthesis	1
105	Cg2166	PUTATIVE POLYRIBONUCLEOTIDE PHOSPHORYLASE / GUANOSINE PENTAPHOSPHATESYNTHETASE	1

106	Cg0545	PUTATIVE LOW-AFFINITY PHOSPHATE TRANSPORT PROTEIN	2
107	Cg2437	THREONINE SYNTHASE	1
108	Cg0914	CELL DIVISION ATP-BINDING PROTEIN	1
109	Cg2154	CDP-DIACYLGLYCEROL--GLYCEROL-3-PHOSPHATE 3-PHOSPHATIDYLTRANSFERASE	1
110	Cg1128	similar to ribosomal protein S2	1
111	Cg0949	CITRATE SYNTHASE	2
112	Cg3182	Trehalose corynomycolyl transferase	3
113	Cg1418	ABC-type cobalamin/Fe3+-siderophores transport system secreted component	1
114	Cg1314	PROLINE TRANSPORT SYSTEM	1
115	Cg0007	DNA GYRASE SUBUNIT B	1
116	Cg1203	Mg-chelatase subunit ChlI	1
117	Cg1789	TRIOSEPHOSPHATE ISOMERASE	1
118	Cg2561	secreted protein potentially involved into thiamin biosynthesis	1
119	Cg2833	O-Acetylserine (Thiol)-Lyase	2
120	Cg2467	ABC TRANSPORTER ATP-BINDING PROTEIN	1
121	Cg1229	ABC-type cobalt transport system, permease component CbiQ	2
122	Cg0896	membrane protein	3
123	Cg1432	DIHYDROXY-ACID DEHYDRATASE	1
124	Cg3257	conserved hypothetical protein	2
125	Cg0161	putative secreted or membrane protein	1
126	Cg2644	ATP-DEPENDENT CLP PROTEASE PROTEOLYTIC SUBUNIT CLPP2	1
127	Cg2799	putative secreted protein	2
128	Cg2361	Cell division initiation protein - Antigen 84 homolog	1
129	Cg2300	IMIDAZOLEGLYCEROL-PHOSPHATE SYNTHASE, AMIDOTRANSFERASE	1
130	Cg2406	CYTOCHROME C OXIDASE SUBUNIT 3	1
131	Cg3068	fructose-bisphosphate aldolase	1

132	Cg1635	putative membrane protein	1
133	Cg1811	putative integration host factor cIHF	1
134	Cg1082	putative membrane protein	1
135	Cg2704	ABC-type sugar transport system, permease component	2
136	Cg0489	hypothetical membrane protein	1
137	Cg0040	PUTATIVE SECRETED PROTEIN	1
138	Cg1072	Ribosomal protein L25 (general stress protein Ctc)	1
139	Cg2460	[putative membrane protein	1
140	Cg1247	putative secreted protein	1
141	Cg0966	PROBABLE THYMIDYLATE SYNTHASE	1
142	Cg3079	PROBABLE ATP-DEPENDENT PROTEASE (HEAT SHOCK PROTEIN)	2
143	Cg1341	RESPIRATORY NITRATE REDUCTASE 2 GAMMA CHAIN	1
144	Cg2678	ABC-type dipeptide/oligopeptide/nickel transport systems, secreted component	1
145	Cg1525	DNA POLYMERASE I	1
146	Cg3096	ALDEHYDE DEHYDROGENASE	1
147	Cg3323	Myo-inositol-1-phosphate synthase	1
148	Cg2958	L-2,3-butanediol dehydrogenase/acetoin reductase	2
149	Cg1629	Similar to preprotein translocase subunit SecA	1
150	Cg1643	6-PHOSPHOGLUCONATE DEHYDROGENASE	2
151	Cg1228	ABC-type cobalt transport system, ATPase component	1
152	Cg1345	PUTATIVE NITRATE/NITRITE TRANSPORTER	1
153	Cg2924	CYSTEINE TRNA SYNTHETASE	1
154	Cg2527	PROBABLE PEPTIDYL-DIPEPTIDASE A PROTEIN	1
155	Cg3139	conserved hypothetical protein	1
156	Cg0238	L-GULONOLACTONE OXIDASE	1
157	Cg0683	permease	1
158	Cg0062	PROTEIN PHOSPHATASE	1
159	Cg1502	ABC-type polar amino acid transport system, ATPase component	1

160	Cg0674	30S RIBOSOMAL PROTEIN S9	1
161	Cg3343	putative secreted membrane protein	2
162	Cg1505	putative secreted protein	1
163	Cg0326	NADH-QUINONE OXIDOREDUCTASE CHAIN 5	2
164	Cg1836	secreted solute-binding protein, aminodeoxychorismate lyase-like	1
165	Cg1463	PUTATIVE GLUTAMYL-TRNA SYNTHETASE	1
166	Cg1794	Uncharacterised P-loop ATPase protein	1
167	Cg0291	3,4-dioxygenase beta subunit	1
168	Cg0652	(Q9RA65) RIBOSOMAL PROTEIN S13	1
169	Cg1603	CONSERVED MEMBRANE PROTEIN	1
170	Cg1195	Sulfate permease or related transporter (MFS superfamily)	1
171	Cg0915	PUTATIVE CELL DIVISION PROTEIN	1
172	Cg2963	]: PROBABLE ATP-DEPENDENT PROTEASE (HEAT SHOCK PROTEIN)	1
173	Cg2174	Exopolyphosphatase-related protein	1
174	Cg1672	polyprenol-phosphate-mannose synthase domain 1	1
175	Cg1861	PPGPP SYNTHETASE, PPGPP PYROPHOSPHORYLASE	1
176	Cg1694	RecB family exonuclease	1
178	Cg0779	TRYPTOPHAN TRNA SYNTHETASE	1
179	Cg2811	ABC-type transport system, involved in lipoprotein release, permease component	1
181	Cg0835	ABC-type sugar transport systems, ATPase component	1
182	Cg1268	GLYCOSYL TRANSFERASE	1
183	Cg2703	Sugar permease	1
184	Cg2052	putative secreted protein	2
185	Cg1259	similar to tetrahydrodipicolinate N-succinyltransferase	1
186	Cg3237	MANGANESE SUPEROXIDE DISMUTASE	1
187	Cg0610	50S RIBOSOMAL PROTEIN L5	1
188	Cg1829	PUTATIVE CHORISMATE SYNTHASE	1
189	Cg2841	Predicted TIM-barrel enzyme, possibly dehydrogenase, nifR3 family	1

190	Cg0933	DNA or RNA helicase of superfamily II	1
191	Cg2033	putative secreted protein	1
192	Cg2362	putative membrane protein	1
193	Cg1479	PUTATIVE GLYCOGEN PHOSPHORYLASE	1
194	Cg2840	BUTYRYL-COA:ACETATE COENZYME A TRANSFERASE	1
195	Cg0592	PUTATIVE BUTYRYL-COA:ACETATE COENZYME A TRANSFERAS	1
197	Cg0048	PEPTIDYL-PROLYL CIS-TRANS ISOMERASE B	1

Supplementary table 3 D (S3 D) CB_Trypsin

Protein hit number	Protein accession number number	Protein protein description	peptides
1	gi|19553194	ABC-type transporter, periplasmic component [Corynebacterium glutamicum ATCC 13032]	12
2	gi|19551617	succinate dehydrogenase flavoprotein subunit [Corynebacterium glutamicum ATCC 13032]	19
3	gi|19554025	putative membrane protease subunit [Corynebacterium glutamicum ATCC 13032]	9
4	gi|19553141	phosphotransferase system, fructose-specific IIC component [Corynebacterium glutamicum ATCC 13032]	7
5	gi|19551738	elongation factor G [Corynebacterium glutamicum ATCC 13032]	9
6	gi|19551746	50S ribosomal protein L3 [Corynebacterium glutamicum ATCC 13032]	5
7	gi|19552590	hypothetical protein NCgl1320 [Corynebacterium glutamicum ATCC 13032]	9
8	gi|19553391	cytochrome b subunit of the bc complex [Corynebacterium glutamicum ATCC 13032]	7
9	gi|19553660	ABC-type transporter, ATPase component [Corynebacterium glutamicum ATCC 13032]	5

10	gi|19553827	ABC-type transporter, periplasmic component [Corynebacterium glutamicum ATCC 13032]	4
11	gi|19552680	NADH dehydrogenase, FAD-containing subunit [Corynebacterium glutamicum ATCC 13032]	6
12	gi|19553397	cytochrome C oxidase subunit II [Corynebacterium glutamicum ATCC 13032]	6
13	gi|19551739	elongation factor Tu [Corynebacterium glutamicum ATCC 13032]	5
14	gi|19553427	hypothetical protein NCgl2145 [Corynebacterium glutamicum ATCC 13032]	5
15	gi|19553197	ABC-type transporter, duplicated ATPase component [Corynebacterium glutamicum ATCC 13032]	8
16	gi|21325295	Heme/copper-type cytochrome/quinol oxidases, subunit 1 [Corynebacterium glutamicum ATCC 13032]	7
17	gi|19552365	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase [Corynebacterium glutamicum ATCC 13032]	5
18	gi|21324774	Hypothetical membrane protein [Corynebacterium glutamicum ATCC 13032]	4
19	gi|19554075	K+ transport flavoprotein [Corynebacterium glutamicum ATCC 13032]	4
20	gi|19551592	esterase [Corynebacterium glutamicum ATCC 13032]	5
21	gi|19552574	phosphotransferase system IIC component, glucose/maltose/N-acetylglucosamine-specific [Corynebacterium glutamicum ATCC 13032]	3
22	gi|19551747	50S ribosomal protein L4 [Corynebacterium glutamicum ATCC 13032]	4
23	gi|19551691	hypothetical protein NCgl0431 [Corynebacterium glutamicum ATCC 13032]	4
24	gi|19552045	hypothetical protein NCgl0784 [Corynebacterium glutamicum ATCC 13032]	2
25	gi|41326199	putative membrane protein [Corynebacterium glutamicum ATCC 13032]	1
26	gi|19551749	50S ribosomal protein L2 [Corynebacterium glutamicum ATCC 13032]	8
27	gi|19552104	large-conductance mechanosensitive channel [Corynebacterium glutamicum ATCC 13032]	2
28	gi|19551824	hypothetical protein NCgl0565 [Corynebacterium glutamicum ATCC 13032]	2

29	gi|19552797	phosphoenolpyruvate carboxylase [Corynebacterium glutamicum ATCC 13032]	6
30	gi|19551869	ABC-type transporter, periplasmic component [Corynebacterium glutamicum ATCC 13032]	5
31	gi|19552799	phosphoglycerate kinase [Corynebacterium glutamicum ATCC 13032]	5
32	gi|19553220	hypothetical protein NCgl1941 [Corynebacterium glutamicum ATCC 13032]	2
33	gi|19553155	glutamate ABC-type transporter, ATPase component [Corynebacterium glutamicum ATCC 13032]	1
34	gi|19551717	preprotein translocase subunit SecE [Corynebacterium glutamicum ATCC 13032]	2
35	gi|19552314	hypothetical protein NCgl1043 [Corynebacterium glutamicum ATCC 13032]	2
36	gi|19551548	hypothetical protein NCgl0293 [Corynebacterium glutamicum ATCC 13032]	1
37	gi|19554067	putative esterase [Corynebacterium glutamicum ATCC 13032]	5
38	gi|19552487	hypothetical protein NCgl1218 [Corynebacterium glutamicum ATCC 13032]	2
39	gi|19553166	phage shock protein A (IM30) [Corynebacterium glutamicum ATCC 13032]	3
40	gi|19552493	ketol-acid reductoisomerase [Corynebacterium glutamicum ATCC 13032]	4
41	gi|19552215	phospho-2-dehydro-3-deoxyheptonate aldolase [Corynebacterium glutamicum ATCC 13032]	3
42	gi|19551798	30S ribosomal protein S4 [Corynebacterium glutamicum ATCC 13032]	7
43	gi|19553196	ABC-type transporter, permease component [Corynebacterium glutamicum ATCC 13032]	2
44	gi|21323302	Ribosomal protein S8 [Corynebacterium glutamicum ATCC 13032]	1
45	gi|21325344	ABC-type transporter, ATPase component [Corynebacterium glutamicum ATCC 13032]	2
46	gi|19553765	acetyl-CoA hydrolase [Corynebacterium glutamicum ATCC 13032]	1
47	gi|19551731	DNA-directed RNA polymerase subunit beta [Corynebacterium glutamicum ATCC 13032]	4

48	gi|19551797	30S ribosomal protein S11 [Corynebacterium glutamicum ATCC 13032]	4
49	gi|19553032	hypothetical protein NCgl1756 [Corynebacterium glutamicum ATCC 13032]	2
50	gi|19554080	1-acyl-sn-glycerol-3-phosphate acyltransferase [Corynebacterium glutamicum ATCC 13032]	1
51	gi|19553392	Rieske Fe-S protein [Corynebacterium glutamicum ATCC 13032]	4
52	gi|21324403	Hypothetical protein [Corynebacterium glutamicum ATCC 13032]	2
53	gi|21323137	Succinate dehydrogenase/fumarate reductase Fe-S protein [Corynebacterium glutamicum ATCC 13032]	3
54	gi|19553205	malate:quinone oxidoreductase [Corynebacterium glutamicum ATCC 13032]	3
55	gi|19554065	putative esterase [Corynebacterium glutamicum ATCC 13032]	5
56	gi|19553828	ABC-type transporter, ATPase component [Corynebacterium glutamicum ATCC 13032]	4
57	gi|19551529	membrane carboxypeptidase [Corynebacterium glutamicum ATCC 13032]	1
58	gi|19554105	L-lactate dehydrogenase [Corynebacterium glutamicum ATCC 13032]	4
59	gi|19551956	ABC-type transporter, periplasmic component [Corynebacterium glutamicum ATCC 13032]	2
60	gi|21322997	Uncharacterized ACR [Corynebacterium glutamicum ATCC 13032]	1
61	gi|19552403	arginyl-tRNA synthetase [Corynebacterium glutamicum ATCC 13032]	2
62	gi|19552219	serine hydroxymethyltransferase [Corynebacterium glutamicum ATCC 13032]	2
63	gi|19552305	ABC-type transporter, duplicated ATPase component [Corynebacterium glutamicum ATCC 13032]	4
64	gi|19551937	detergent sensitivity rescuer dtsR2 [Corynebacterium glutamicum ATCC 13032]	2
65	gi|19551984	ribosome-associated protein Y [Corynebacterium glutamicum ATCC 13032]	3
66	gi|19554117	hypothetical protein NCgl2831 [Corynebacterium glutamicum ATCC 13032]	1
67	gi|19551441	FAD/FMN-containing dehydrogenase [Corynebacterium glutamicum ATCC 13032]	2

68	gi|21323598	Uncharacterized membrane protein [Corynebacterium glutamicum ATCC 13032]	3
69	gi|19554227	ascorbate-specific PTS system enzyme IIC [Corynebacterium glutamicum ATCC 13032]	3
70	gi|19553840	phosphotransferase system IIC component [Corynebacterium glutamicum ATCC 13032]	2
71	gi|19551616	hypothetical protein NCgl0359 [Corynebacterium glutamicum ATCC 13032]	3
72	gi|19551276	hypothetical protein NCgl0025 [Corynebacterium glutamicum ATCC 13032]	3
73	gi|19552361	hypothetical protein NCgl1090 [Corynebacterium glutamicum ATCC 13032]	1
74	gi|19553738	hypothetical protein NCgl2452 [Corynebacterium glutamicum ATCC 13032]	1
75	gi|19553114	transcriptional regulator [Corynebacterium glutamicum ATCC 13032]	1
76	gi|19552644	GTP-binding protein EngA [Corynebacterium glutamicum ATCC 13032]	1
77	gi|19553350	cell division initiation protein [Corynebacterium glutamicum ATCC 13032]	1
78	gi|41325061	FATTY ACID SYNTHASE [Corynebacterium glutamicum ATCC 13032]	3
79	gi|19552869	preprotein translocase subunit YajC [Corynebacterium glutamicum ATCC 13032]	3
80	gi|19552098	50S ribosomal protein L31 type B [Corynebacterium glutamicum ATCC 13032]	1
81	gi|19552504	D-3-phosphoglycerate dehydrogenase [Corynebacterium glutamicum ATCC 13032]	2
82	gi|19552746	membrane protease subunit [Corynebacterium glutamicum ATCC 13032]	4
83	gi|19553156	glutamate ABC-type transporter, periplasmic component [Corynebacterium glutamicum ATCC 13032]	2
84	gi|41326948	conserved hypothetical protein [Corynebacterium glutamicum ATCC 13032]	1
85	gi|19551619	hypothetical protein NCgl0362 [Corynebacterium glutamicum ATCC 13032]	1

86	gi|21323688	Hypothetical protein [Corynebacterium glutamicum ATCC 13032]	2
87	gi|21324968	Hypothetical membrane protein [Corynebacterium glutamicum ATCC 13032]	1
88	gi|19551799	DNA-directed RNA polymerase subunit alpha [Corynebacterium glutamicum ATCC 13032]	2
89	gi|19551732	DNA-directed RNA polymerase subunit beta~ [Corynebacterium glutamicum ATCC 13032]	3
90	gi|21324681	DNA-directed RNA polymerase sigma subunits (sigma70/sigma32) (SigA) [Corynebacterium glutamicum ATCC 13032]	1
91	gi|19552573	30S ribosomal protein S1 [Corynebacterium glutamicum ATCC 13032]	2
92	gi|21324098	ABC-type transporter, ATPase component [Corynebacterium glutamicum ATCC 13032]	1
93	gi|21325656	Hypothetical membrane protein [Corynebacterium glutamicum ATCC 13032]	3
94	gi|21322889	Hypothetical membrane protein [Corynebacterium glutamicum ATCC 13032]	2
95	gi|19553476	glycyl-tRNA synthetase [Corynebacterium glutamicum ATCC 13032]	1
96	gi|21325487	porin [Corynebacterium glutamicum ATCC 13032]	1
97	gi|19552775	ABC-type transporter, ATPase component [Corynebacterium glutamicum ATCC 13032]	2
98	gi|19553694	3-oxoacyl-(acyl-carrier-protein) synthase [Corynebacterium glutamicum ATCC 13032]	1
99	gi|19553123	RNA polymerase sigma factor SigB [Corynebacterium glutamicum ATCC 13032]	1
100	gi|19552143	hypothetical protein NCgl0880 [Corynebacterium glutamicum ATCC 13032]	3
101	gi|19551752	30S ribosomal protein S3 [Corynebacterium glutamicum ATCC 13032]	2
102	gi|19551290	putative septation inhibitor protein [Corynebacterium glutamicum ATCC 13032]	1
103	gi|19553618	hypothetical protein NCgl2336 [Corynebacterium glutamicum ATCC 13032]	2
104	gi|19552407	homoserine dehydrogenase [Corynebacterium glutamicum ATCC 13032]	3

105	gi|19552617	argininosuccinate synthase [Corynebacterium glutamicum ATCC 13032]	2
106	gi|19552088	bifunctional phosphoribosylaminoimidazolecarboxamide formyltransferase/IMP cyclohydrolase [Corynebacterium glutamicum ATCC 13032]	1
107	gi|19553751	ABC-type transporter, ATPase component [Corynebacterium glutamicum ATCC 13032]	2
108	gi|19551541	cAMP-binding domain-containing protein [Corynebacterium glutamicum ATCC 13032]	1
109	gi|19551759	50S ribosomal protein L24 [Corynebacterium glutamicum ATCC 13032]	2
110	gi|19552252	hypothetical protein NCgl0987 [Corynebacterium glutamicum ATCC 13032]	1
111	gi|19553333	dehydrogenase [Corynebacterium glutamicum ATCC 13032]	1
112	gi|19552805	hypothetical protein NCgl1531 [Corynebacterium glutamicum ATCC 13032]	1
113	gi|19553189	translation initiation factor IF-2 [Corynebacterium glutamicum ATCC 13032]	2
114	gi|41325436	ATP SYNTHASE ALPHA SUBUNIT [Corynebacterium glutamicum ATCC 13032]	2
115	gi|19554063	hypothetical protein NCgl2775 [Corynebacterium glutamicum ATCC 13032]	1
116	gi|19552423	transcription termination factor Rho [Corynebacterium glutamicum ATCC 13032]	1
117	gi|19551816	30S ribosomal protein S9 [Corynebacterium glutamicum ATCC 13032]	2
118	gi|19551753	50S ribosomal protein L16 [Corynebacterium glutamicum ATCC 13032]	2
119	gi|19552490	small-conductance mechanosensitive channel [Corynebacterium glutamicum ATCC 13032]	1
120	gi|19554256	putative proline-betaine transporter [Corynebacterium glutamicum ATCC 13032]	1
121	gi|19552092	30S ribosomal protein S18 [Corynebacterium glutamicum ATCC 13032]	2
122	gi|19553658	ABC-type transporter, periplasmic component [Corynebacterium glutamicum ATCC 13032]	2
123	gi|19551720	50S ribosomal protein L1 [Corynebacterium glutamicum ATCC 13032]	1

124	gi|19551777	30S ribosomal protein S5 [Corynebacterium glutamicum ATCC 13032]	1
125	gi|19552432	F0F1 ATP synthase subunit delta [Corynebacterium glutamicum ATCC 13032]	1
126	gi|21324061	Glutamyl- and glutaminyl-tRNA synthetases [Corynebacterium glutamicum ATCC 13032]	1
127	gi|19552608	hypothetical protein NCgl1337 [Corynebacterium glutamicum ATCC 13032]	2
128	gi|19553872	ATPase with chaperone activity, ATP-binding subunit [Corynebacterium glutamicum ATCC 13032]	1
129	gi|19553650	putative ABC transporter ATP-binding protein [Corynebacterium glutamicum ATCC 13032]	2
130	gi|19552060	Na+/proline, Na+/panthothenate symporter or related permease [Corynebacterium glutamicum ATCC 13032]	1
131	gi|19552085	metalloendopeptidase-like membrane protein [Corynebacterium glutamicum ATCC 13032]	1
132	gi|19554026	hypothetical protein NCgl2738 [Corynebacterium glutamicum ATCC 13032]	1
133	gi|19553758	cysteine synthase [Corynebacterium glutamicum ATCC 13032]	1
134	gi|21324976	Leucyl aminopeptidase [Corynebacterium glutamicum ATCC 13032]	1
135	gi|19551579	hypothetical protein NCgl0323 [Corynebacterium glutamicum ATCC 13032]	1
136	gi|19552323	hypothetical protein NCgl1052 [Corynebacterium glutamicum ATCC 13032]	1
137	gi|19554184	hypothetical protein NCgl2888 [Corynebacterium glutamicum ATCC 13032]	1
138	gi|19552825	pyrimidine regulatory protein PyrR [Corynebacterium glutamicum ATCC 13032]	1
139	gi|19552029	cell-division ATP-binding protein [Corynebacterium glutamicum ATCC 13032]	1
140	gi|19551957	ABC-type transporter, ATPase component [Corynebacterium glutamicum ATCC 13032]	1
141	gi|19552093	30S ribosomal protein S14 [Corynebacterium glutamicum ATCC 13032]	1

142	gi|19551647	phosphoglyceromutase [Corynebacterium glutamicum ATCC 13032]	1
143	gi|19551919	pyruvate carboxylase [Corynebacterium glutamicum ATCC 13032]	1
144	gi|19551440	short chain dehydrogenase [Corynebacterium glutamicum ATCC 13032]	1
145	gi|19552095	50S ribosomal protein L28 [Corynebacterium glutamicum ATCC 13032]	1
146	gi|41326869	putative membrane protein [Corynebacterium glutamicum ATCC 13032]	1
147	gi|19554069	phosphoribose diphosphate:decaprenyl-phosphate phosphoribosyltransferase [Corynebacterium glutamicum ATCC 13032]	1
148	gi|19552821	carbamoyl phosphate synthase large subunit [Corynebacterium glutamicum ATCC 13032]	1
149	gi|19553913	hypothetical protein NCgl2624 [Corynebacterium glutamicum ATCC 13032]	1
150	gi|21324427	Preprotein translocase subunit SecD [Corynebacterium glutamicum ATCC 13032]	1
151	gi|21323174	Hypothetical membrane protein [Corynebacterium glutamicum ATCC 13032]	1
152	gi|41325890	THREONYL-TRNA SYNTHETASE [Corynebacterium glutamicum ATCC 13032]	1
153	gi|19551295	protein serine/threonine phosphatase [Corynebacterium glutamicum ATCC 13032]	1
154	gi|21324859	Pyruvate kinase [Corynebacterium glutamicum ATCC 13032]	1
155	gi|21323977	F0F1-type ATP synthase alpha subunit [Corynebacterium glutamicum ATCC 13032]	1
156	gi|19552251	SNF family Na(+)-dependent transporter [Corynebacterium glutamicum ATCC 13032]	1
157	gi|19554257	hypothetical protein NCgl2962 [Corynebacterium glutamicum ATCC 13032]	2
158	gi|19551908	hypothetical protein NCgl0648 [Corynebacterium glutamicum ATCC 13032]	1
159	gi|19552803	hypothetical protein NCgl1529 [Corynebacterium glutamicum ATCC 13032]	1
160	gi|19551297	hypothetical protein NCgl0046 [Corynebacterium glutamicum ATCC 13032]	1

161	gi|19553838	rRNA methylase [Corynebacterium glutamicum ATCC 13032]	1
162	gi|21325721	Hypothetical protein [Corynebacterium glutamicum ATCC 13032]	1
163	gi|19553542	30S ribosomal protein S20 [Corynebacterium glutamicum ATCC 13032]	2
164	gi|19554029	hypothetical protein NCgl2741 [Corynebacterium glutamicum ATCC 13032]	1
165	gi|19553473	chromosome segregation ATPase [Corynebacterium glutamicum ATCC 13032]	1
166	gi|19551452	ABC-type transporter, permease component [Corynebacterium glutamicum ATCC 13032]	1
167	gi|19551736	30S ribosomal protein S12 [Corynebacterium glutamicum ATCC 13032]	1
168	gi|19552431	F0F1 ATP synthase subunit B [Corynebacterium glutamicum ATCC 13032]	1
169	gi|19553687	hypothetical protein NCgl2402 [Corynebacterium glutamicum ATCC 13032]	1
170	gi|19552819	hypothetical protein NCgl1545 [Corynebacterium glutamicum ATCC 13032]	1
171	gi|19551930	acyl-CoA carboxylase [Corynebacterium glutamicum ATCC 13032]	1
172	gi|21325042	Glucosamine 6-phosphate synthetase, contains amidotransferase and phosphosugar isomerase domains [Corynebacterium glutamicum ATCC 13032]	1
173	gi|19551283	hypothetical protein NCgl0032 [Corynebacterium glutamicum ATCC 13032]	1
174	gi|19552037	ABC-type cobalamin/Fe3+-siderophore transport system, periplasmic component [Corynebacterium glutamicum ATCC 13032]	1
175	gi|19553179	polynucleotide phosphorylase/polyadenylase [Corynebacterium glutamicum ATCC 13032]	1
176	gi|19553829	ABC-type transporter, permease component [Corynebacterium glutamicum ATCC 13032]	1
177	gi|19551877	ABC-type Fe3+-siderophores transport system, periplasmic component [Corynebacterium glutamicum ATCC 13032]	1
178	gi|19552631	hypothetical protein NCgl1360 [Corynebacterium glutamicum ATCC 13032]	1

179	gi|161486709	alpha-ketoglutarate decarboxylase [Corynebacterium glutamicum ATCC 13032]	1
180	gi|19552172	ABC-type transporter, ATPase component [Corynebacterium glutamicum ATCC 13032]	1
181	gi|19554061	putative polyketide synthase [Corynebacterium glutamicum ATCC 13032]	1
182	gi|41326753	ABC-type phosphate transport system, permease component [Corynebacterium glutamicum ATCC 13032]	1
183	gi|19553415	glutamine synthase [Corynebacterium glutamicum ATCC 13032]	1
184	gi|19551985	preprotein translocase subunit SecA [Corynebacterium glutamicum ATCC 13032]	1
185	gi|19552595	translation initiation factor IF-3 [Corynebacterium glutamicum ATCC 13032]	1
186	gi|19554128	hypothetical protein NCgl2842 [Corynebacterium glutamicum ATCC 13032]	1
187	gi|19552127	D-lactate dehydrogenase [Corynebacterium glutamicum ATCC 13032]	1
188	gi|19552705	twin argininte translocase protein A [Corynebacterium glutamicum ATCC 13032]	1
189	gi|19553229	30S ribosomal protein S2 [Corynebacterium glutamicum ATCC 13032]	1
190	gi|19554208	hypothetical protein NCgl2912 [Corynebacterium glutamicum ATCC 13032]	1
191	gi|19552345	hypothetical protein NCgl1074 [Corynebacterium glutamicum ATCC 13032]	1
192	gi|19553083	integrase [Corynebacterium glutamicum ATCC 13032]	1
193	gi|19551706	naphthoate synthase [Corynebacterium glutamicum ATCC 13032]	1
194	gi|19552350	hypothetical protein NCgl1079 [Corynebacterium glutamicum ATCC 13032]	1
195	gi|19554180	penicillin-binding protein [Corynebacterium glutamicum ATCC 13032]	1
196	gi|19552199	phosphopyruvate hydratase [Corynebacterium glutamicum ATCC 13032]	1
197	gi|19551758	50S ribosomal protein L14 [Corynebacterium glutamicum ATCC 13032]	1
198	gi|19553103	hypothetical protein NCgl1824 [Corynebacterium glutamicum ATCC 13032]	1
200	gi|19552283	Mg-chelatase subunit ChlI [Corynebacterium glutamicum ATCC 13032]	1

Supplementry table 4 complete list of the proteins identified from atryptic digestion of AF4, AF4•MLL and ENL

Supplementry table 4 A (S4A) Summary of all the proteins identified from AF4

Protein hit number	Protein accession number	prot_protein description	peptides
1	AFF1	AF4/FMR2 family member 1 (Protein AF-4) (Proto-oncogene AF4) (Protein FEL) -	106
2	HSP71	Heat shock 70 kDa protein 1 (HSP70.1) (HSP70-1/HSP70-2) -	49
3	HSP7C	Heat shock cognate 71 kDa protein (Heat shock 70 kDa protein 8) -	36
4	TBB5	Tubulin beta chain (Tubulin beta-5 chain) -	25
5	HSP76	Heat shock 70 kDa protein 6 (Heat shock 70 kDa protein B~) -	32
6	GRP78	78 kDa glucose-regulated protein precursor (GRP 78) (Heat shock 70 kDa protein 5) (Immunoglobulin heavy chain-binding protein) (BiP) (Endoplasmic reticulum lumenal Ca(2+)-binding protein grp78) -	31
7	HS70L	Heat shock 70 kDa protein 1L (Heat shock 70 kDa protein 1-like) (Heat shock 70 kDa protein 1-Hom) (HSP70-Hom) -	20
8	TBAK	Tubulin alpha-ubiquitous chain (Alpha-tubulin ubiquitous) (Tubulin K-alpha-1) -	21
9	TBA6	Tubulin alpha-6 chain (Alpha-tubulin 6) -	19
10	TBB2C	Tubulin beta-2C chain (Tubulin beta-2 chain) -	23
11	TBB2A	Tubulin beta-2A chain -	20
12	TBA3	Tubulin alpha-3 chain (Alpha-tubulin 3) (Tubulin B-alpha-1) -	20
13	TBB4	Tubulin beta-4 chain (Tubulin 5 beta) -	18
14	TBA2	Tubulin alpha-2 chain (Alpha-tubulin 2) -	15
15	GRP75	Stress-70 protein, mitochondrial precursor (75 kDa glucose-regulated protein) (GRP 75) (Peptide-binding protein 74) (PBP74) (Mortalin) (MOT) -	14
16	K2C1	Keratin, type II cytoskeletal 1 (Cytokeratin-1) (CK-1) (Keratin-1) (K1) (67 kDa cytokeratin) (Hair alpha protein) -	11
17	HS90A	Heat shock protein HSP 90-alpha (HSP 86) (Renal carcinoma antigen NY-REN-38) -	22
18	HS90B	Heat shock protein HSP 90-beta (HSP 84) (HSP 90) -	26

19	ACTB	Actin, cytoplasmic 1 (Beta-actin) -	13
20	ENL	Protein ENL (YEATS domain-containing protein 1) -	22
21	EF2	Elongation factor 2 (EF-2) -	19
22	ACTC	Actin, alpha cardiac muscle 1 (Alpha-cardiac actin) -	8
23	ATX2L	Ataxin-2-like protein (Ataxin-2 domain protein) (Ataxin-2-related protein) -	14
24	HS105	Heat-shock protein 105 kDa (Heat shock 110 kDa protein) (Antigen NY-CO-25) -	14
25	PGRC1	Membrane-associated progesterone receptor component 1 (mPR) -	9
26	ENOA	Alpha-enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (Non-neural enolase) (NNE) (Enolase 1) (Phosphopyruvate hydratase) (C-myc promoter-binding protein) (MBP-1) (MPB-1) (Plasminogen-binding protein) -	10
27	FAS	Fatty acid synthase (EC 2.3.1.85) [Includes: [Acyl-carrier-protein] S-acetyltransferase (EC 2.3.1.38); [Acyl-carrier-protein] S-malonyltransferase (EC 2.3.1.39); 3-oxoacyl-[acyl-carrier-protein] synthase (EC 2.3.1.41); 3-oxoacyl-[acyl-carrier-protei	14
28	TCPB	T-complex protein 1 subunit beta (TCP-1-beta) (CCT-beta) -	7
29	CH60	60 kDa heat shock protein, mitochondrial precursor (Hsp60) (60 kDa chaperonin) (CPN60) (Heat shock protein 60) (HSP-60) (Mitochondrial matrix protein P1) (P60 lymphocyte protein) (HuCHA60) -	8
30	G3P	Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH) -	8
31	ATPB	ATP synthase subunit beta, mitochondrial precursor (EC 3.6.3.14) -	10
32	CALX	Calnexin precursor (Major histocompatibility complex class I antigen-binding protein p88) (p90) (IP90) -	8
33	K1C9	Keratin, type I cytoskeletal 9 (Cytokeratin-9) (CK-9) (Keratin-9) (K9) -	6
34	IMA2	Importin alpha-2 subunit (Karyopherin alpha-2 subunit) (SRP1-alpha) (RAG cohort protein 1) -	7
35	TPIS	Triosephosphate isomerase (EC 5.3.1.1) (TIM) (Triose-phosphate isomerase) -	8
36	EF1G	Elongation factor 1-gamma (EF-1-gamma) (eEF-1B gamma) -	10
37	EF1A1	Elongation factor 1-alpha 1 (EF-1-alpha-1) (Elongation factor 1 A-1) (eEF1A-1) (Elongation factor Tu) (EF-Tu) (Leukocyte receptor cluster member 7) -	8
38	PYC	Pyruvate carboxylase, mitochondrial precursor (EC 6.4.1.1) (Pyruvic carboxylase) (PCB) -	11
39	CDK9	Cell division protein kinase 9 (EC 2.7.11.22) (EC 2.7.11.23) (Cyclin-dependent kinase 9) (Serine/threonine-protein kinase PITALRE) (C-2K) (Cell division cycle 2-like	12

		protein kinase 4) -	
40	ROA1	Heterogeneous nuclear ribonucleoprotein A1 (Helix-destabilizing protein) (Single-strand RNA-binding protein) (hnRNP core protein A1) -	5
41	DDX6	Probable ATP-dependent RNA helicase DDX6 (EC 3.6.1.-) (DEAD box protein 6) (ATP-dependent RNA helicase p54) (Oncogene RCK) -	9
42	4F2	4F2 cell-surface antigen heavy chain (4F2hc) (Lymphocyte activation antigen 4F2 large subunit) (4F2 heavy chain antigen) (CD98 antigen) -	7
43	RBM25	Probable RNA-binding protein 25 (RNA-binding motif protein 25) (RNA-binding region-containing protein 7) (Protein S164) -	5
44	PPIA	Peptidyl-prolyl cis-trans isomerase A (EC 5.2.1.8) (PPIase A) (Rotamase A) (Cyclophilin A) (Cyclosporin A-binding protein) -	5
45	NPM	Nucleophosmin (NPM) (Nucleolar phosphoprotein B23) (Numatrin) (Nucleolar protein NO38) -	6
46	TCPE	T-complex protein 1 subunit epsilon (TCP-1-epsilon) (CCT-epsilon) -	5
47	TCPH	T-complex protein 1 subunit eta (TCP-1-eta) (CCT-eta) (HIV-1 Nef-interacting protein) -	6
48	ATPA	ATP synthase subunit alpha, mitochondrial precursor (EC 3.6.3.14) -	4
49	PRDX6	Peroxiredoxin-6 (EC 1.11.1.15) (Antioxidant protein 2) (1-Cys peroxiredoxin) (1-Cys PRX) (Acidic calcium-independent phospholipase A2) (EC 3.1.1.-) (aiPLA2) (Non-selenium glutathione peroxidase) (EC 1.11.1.7) (NSGPx) (24 kDa protein) (Liver 2D page	2
50	TCPA	T-complex protein 1 subunit alpha (TCP-1-alpha) (CCT-alpha) -	7
51	SF3A1	Splicing factor 3 subunit 1 (Spliceosome-associated protein 114) (SAP 114) (SF3a120) -	4
52	IMB1	Importin beta-1 subunit (Karyopherin beta-1 subunit) (Nuclear factor P97) (Importin 90) -	3
53	KPYM	Pyruvate kinase isozymes M1/M2 (EC 2.7.1.40) (Pyruvate kinase muscle isozyme) (Pyruvate kinase 2/3) (Cytosolic thyroid hormone-binding protein) (CTHBP) (THBP1) -	5
54	NDKB	Nucleoside diphosphate kinase B (EC 2.7.4.6) (NDK B) (NDP kinase B) (nm23-H2) (C-myc purine-binding transcription factor PUF) -	5
55	TCPD	T-complex protein 1 subunit delta (TCP-1-delta) (CCT-delta) (Stimulator of TAR RNA-binding) -	3

56	TCPZ	T-complex protein 1 subunit zeta (TCP-1-zeta) (CCT-zeta) (CCT-zeta-1) (Tcp20) (HTR3) (Acute morphine dependence-related protein 2) -	2
57	KV206	Ig kappa chain V-II region RPMI 6410 precursor -	1
58	HNRH1	Heterogeneous nuclear ribonucleoprotein H (hnRNP H) -	4
59	CAH2	Carbonic anhydrase 2 (EC 4.2.1.1) (Carbonic anhydrase II) (Carbonate dehydratase II) (CA-II) (Carbonic anhydrase C) -	3
60	GDIB	Rab GDP dissociation inhibitor beta (Rab GDI beta) (Guanosine diphosphate dissociation inhibitor 2) (GDI-2) -	4
61	ACLY	ATP-citrate synthase (EC 2.3.3.8) (ATP-citrate (pro-S-)-lyase) (Citrate cleavage enzyme) -	8
62	GDIA	Rab GDP dissociation inhibitor alpha (Rab GDI alpha) (Guanosine diphosphate dissociation inhibitor 1) (GDI-1) (XAP-4) (Oligophrenin-2) -	3
63	CCNT1	Cyclin-T1 (CycT1) (Cyclin-T) -	11
64	IMA4	Importin alpha-4 subunit (Karyopherin alpha-4 subunit) (Qip1 protein) -	9
65	YBOX1	Nuclease sensitive element-binding protein 1 (Y-box-binding protein 1) (Y-box transcription factor) (YB-1) (CCAAT-binding transcription factor I subunit A) (CBF-A) (Enhancer factor I subunit A) (EFI-A) (DNA-binding protein B) (DBPB) - (	4
66	CNBP	Cellular nucleic acid-binding protein (CNBP) (Zinc finger protein 9) -	2
67	HSP74	Heat shock 70 kDa protein 4 (Heat shock 70-related protein APG-2) (HSP70RY) -	5
68	CH10	10 kDa heat shock protein, mitochondrial (Hsp10) (10 kDa chaperonin) (CPN10) (Early-pregnancy factor) (EPF) -	3
69	GSTP1	Glutathione S-transferase P (EC 2.5.1.18) (GST class-pi) (GSTP1-1) -	2
70	DNJA1	DnaJ homolog subfamily A member 1 (Heat shock 40 kDa protein 4) (DnaJ protein homolog 2) (HSJ-2) (HSDJ) -	2
71	TERA	Transitional endoplasmic reticulum ATPase (TER ATPase) (15S Mg(2+)-ATPase p97 subunit) (Valosin-containing protein) (VCP) -	4
72	ROA2	Heterogeneous nuclear ribonucleoproteins A2/B1 (hnRNP A2 / hnRNP B1) -	4
73	KCRB	Creatine kinase B-type (EC 2.7.3.2) (Creatine kinase B chain) (B-CK) -	2
74	HNRPF	Heterogeneous nuclear ribonucleoprotein F (hnRNP F) (Nucleolin-like protein mcs94-1) -	3
75	RL15	60S ribosomal protein L15 -	5
76	ALDOA	Fructose-bisphosphate aldolase A (EC 4.1.2.13) (Muscle-type aldolase) (Lung cancer	4

		antigen NY-LU-1) -	
77	PRDX2	Peroxiredoxin-2 (EC 1.11.1.15) (Thioredoxin peroxidase 1) (Thioredoxin-dependent peroxide reductase 1) (Thiol-specific antioxidant protein) (TSA) (PRP) (Natural killer cell-enhancing factor B) (NKEF-B) -	3
78	PRP40	Pre-mRNA-processing factor 40 homolog A (Formin-binding protein 3) (Formin-binding protein 11) (Huntingtin-interacting protein HYPA) (Huntingtin yeast partner A) (Fas ligand-associated factor 1) (NY-REN-6 antigen) -	5
79	PRDX1	Peroxiredoxin-1 (EC 1.11.1.15) (Thioredoxin peroxidase 2) (Thioredoxin-dependent peroxide reductase 2) (Proliferation-associated gene protein) (PAG) (Natural killer cell-enhancing factor A) (NKEF-A) -	5
80	IF5A1	Eukaryotic translation initiation factor 5A-1 (eIF-5A-1) (eIF-5A1) (Eukaryotic initiation factor 5A isoform 1) (eIF-5A) (eIF-4D) (Rev-binding factor) -	2
81	PGAM1	Phosphoglycerate mutase 1 (EC 5.4.2.1) (EC 5.4.2.4) (EC 3.1.3.13) (Phosphoglycerate mutase isozyme B) (PGAM-B) (BPG-dependent PGAM 1) -	5
82	RAB10	Ras-related protein Rab-10 -	2
83	PP2CG	Protein phosphatase 2C isoform gamma (EC 3.1.3.16) (PP2C-gamma) (Protein phosphatase magnesium-dependent 1 gamma) (Protein phosphatase 1C) -	2
84	ADT2	ADP/ATP translocase 2 (Adenine nucleotide translocator 2) (ANT 2) (ADP,ATP carrier protein 2) (Solute carrier family 25 member 5) (ADP,ATP carrier protein, fibroblast isoform) -	4
85	NDKA	Nucleoside diphosphate kinase A (EC 2.7.4.6) (NDK A) (NDP kinase A) (Tumor metastatic process-associated protein) (Metastasis inhibition factor nm23) (nm23-H1) (Granzyme A-activated DNase) (GAAD) -	3
86	COF1	Cofilin-1 (Cofilin, non-muscle isoform) (18 kDa phosphoprotein) (p18) -	5
87	AFF4	AF4/FMR2 family member 4 (ALL1-fused gene from chromosome 5q31) (Major CDK9 elongation factor-associated protein) -	7
88	COA1	Acetyl-CoA carboxylase 1 (EC 6.4.1.2) (ACC-alpha) [Includes: Biotin carboxylase (EC 6.3.4.14)] -	9
89	ROA3	Heterogeneous nuclear ribonucleoprotein A3 (hnRNP A3) -	1
90	RL6	60S ribosomal protein L6 (TAX-responsive enhancer element-binding protein 107) (TAXREB107) (Neoplasm-related protein C140) -	2
91	TCPG	T-complex protein 1 subunit gamma (TCP-1-gamma) (CCT-gamma) (hTRiC5) -	6

92	TCPQ	T-complex protein 1 subunit theta (TCP-1-theta) (CCT-theta) (Renal carcinoma antigen NY-REN-15) -	1
93	RAP1A	Ras-related protein Rap-1A precursor (GTP-binding protein smg-p21A) (Ras-related protein Krev-1) (C21KG) (G-22K) -	2
94	VIME	Vimentin -	10
95	IMA3	Importin alpha-3 subunit (Karyopherin alpha-3 subunit) (SRP1-gamma) -	7
96	CLH1	Clathrin heavy chain 1 (CLH-17) -	5
97	SERA	D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95) (3-PGDH) -	1
98	PIMT	Protein-L-isoaspartate(D-aspartate) O-methyltransferase (EC 2.1.1.77) (Protein-beta-aspartate methyltransferase) (PIMT) (Protein L-isoaspartyl/D-aspartyl methyltransferase) (L-isoaspartyl protein carboxyl methyltransferase) -	5
99	SYEP	Bifunctional aminoacyl-tRNA synthetase [Includes: Glutamyl-tRNA synthetase (EC 6.1.1.17) (Glutamate--tRNA ligase); Prolyl-tRNA synthetase (EC 6.1.1.15) (Proline--tRNA ligase)] -	5
100	VINC	Vinculin (Metavinculin) -	1
101	NEDD8	NEDD8 precursor (Ubiquitin-like protein Nedd8) (Neddylin) -	1
102	SF3A3	Splicing factor 3A subunit 3 (Spliceosome-associated protein 61) (SAP 61) (SF3a60) -	2
103	UBE1	Ubiquitin-activating enzyme E1 (A1S9 protein) -	3
104	CALM	Calmodulin (CaM) -	2
105	ANXA5	Annexin A5 (Annexin V) (Lipocortin V) (Endonexin II) (Calphobindin I) (CBP-I) (Placental anticoagulant protein I) (PAP-I) (PP4) (Thromboplastin inhibitor) (Vascular anticoagulant-alpha) (VAC-alpha) (Anchorin CII) -	1
106	HNRPK	Heterogeneous nuclear ribonucleoprotein K (hnRNP K) (Transformation up-regulated nuclear protein) (TUNP) -	4
107	PSMD1	26S proteasome non-ATPase regulatory subunit 1 (26S proteasome regulatory subunit RPN2) (26S proteasome regulatory subunit S1) (26S proteasome subunit p112) -	1
108	EF1D	Elongation factor 1-delta (EF-1-delta) (Antigen NY-CO-4) -	5
109	EFTU	Elongation factor Tu, mitochondrial precursor (EF-Tu) (P43) -	3
110	IF3A	Eukaryotic translation initiation factor 3 subunit 10 (eIF-3 theta) (eIF3 p167) (eIF3 p180) (eIF3 p185) (eIF3a) -	4
111	KU70	ATP-dependent DNA helicase 2 subunit 1 (ATP-dependent DNA helicase II 70 kDa subunit) (Lupus Ku autoantigen protein p70) (Ku70) (70 kDa subunit of Ku antigen)	2

		(Thyroid-lupus autoantigen) (TLAA) (CTC box-binding factor 75 kDa subunit) (CTCBF) (CTC75)	
112	K1C10	Keratin, type I cytoskeletal 10 (Cytokeratin-10) (CK-10) (Keratin-10) (K10) -	6
113	RLA1	60S acidic ribosomal protein P1 -	1
114	RIB1	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 67 kDa subunit precursor (EC 2.4.1.119) (Ribophorin I) (RPN-I) -	2
115	TOM70	Mitochondrial precursor proteins import receptor (Translocase of outer membrane TOM70) -	2
116	SFPQ	Splicing factor, proline- and glutamine-rich (Polypyrimidine tract-binding protein-associated-splicing factor) (PTB-associated-splicing factor) (PSF) (DNA-binding p52/p100 complex, 100 kDa subunit) (100 kDa DNA-pairing protein) (hPOMp100) - Homo sap	2
117	RAN	GTP-binding nuclear protein Ran (GTPase Ran) (Ras-like protein TC4) (Androgen receptor-associated protein 24) -	2
118	KINH	Kinesin heavy chain (Ubiquitous kinesin heavy chain) (UKHC) -	2
119	SSB	Single-stranded DNA-binding protein, mitochondrial precursor (Mt-SSB) (MtSSB) (PWP1-interacting protein 17) -	1
120	RS18	40S ribosomal protein S18 (Ke-3) (Ke3) -	2
121	TRAP1	Heat shock protein 75 kDa, mitochondrial precursor (HSP 75) (Tumor necrosis factor type 1 receptor-associated protein) (TRAP-1) (TNFR-associated protein 1) -	2
122	HMGB1	High mobility group protein B1 (High mobility group protein 1) (HMG-1) -	2
123	LDHB	L-lactate dehydrogenase B chain (EC 1.1.1.27) (LDH-B) (LDH heart subunit) (LDH-H) (Renal carcinoma antigen NY-REN-46) -	2
124	CALR	Calreticulin precursor (CRP55) (Calregulin) (HACBP) (ERp60) (grp60) -	3
125	MOES	Moesin (Membrane-organizing extension spike protein) -	3
126	MPCP	Phosphate carrier protein, mitochondrial precursor (PTP) (Solute carrier family 25 member 3) -	2
127	C1QBP	Complement component 1 Q subcomponent-binding protein, mitochondrial precursor (Glycoprotein gC1qBP) (C1qBP) (GC1q-R protein) (Hyaluronan-binding protein 1) (Mitochondrial matrix protein p32) (p33) -	2
128	IPO4	Importin-4 (Importin 4b) (Imp4b) (Ran-binding protein 4) (RanBP4) -	1
129	DDB1	DNA damage-binding protein 1 (Damage-specific DNA-binding protein 1) (UV-	1

		damaged DNA-binding factor) (DDB p127 subunit) (DDBa) (UV-damaged DNA-binding protein 1) (UV-DDB 1) (Xeroderma pigmentosum group E-complementing protein) (XPCe) (XPE-binding fa	
130	IF3I	Eukaryotic translation initiation factor 3 subunit 6-interacting protein -	2
131	PSMD2	26S proteasome non-ATPase regulatory subunit 2 (26S proteasome regulatory subunit RPN1) (26S proteasome regulatory subunit S2) (26S proteasome subunit p97) (Tumor necrosis factor type 1 receptor-associated protein 2) (55.11 protein) - (	1
132	AAAT	Neutral amino acid transporter B(0) (ATB(0)) (Sodium-dependent neutral amino acid transporter type 2) (RD114/simian type D retrovirus receptor) (Baboon M7 virus receptor) -	1
133	G6PI	Glucose-6-phosphate isomerase (EC 5.3.1.9) (GPI) (Phosphoglucose isomerase) (PGI) (Phosphohexose isomerase) (PHI) (Neuroleukin) (NLK) (Sperm antigen 36) (SA-36) -	1
134	ARP3	Actin-like protein 3 (Actin-related protein 3) -	1
135	IPYR	Inorganic pyrophosphatase (EC 3.6.1.1) (Pyrophosphate phospho-hydrolase) (PPase) -	1
136	U5S1	116 kDa U5 small nuclear ribonucleoprotein component (U5 snRNP-specific protein, 116 kDa) (U5-116 kDa) (Elongation factor Tu GTP-binding domain protein 2) (hSNU114) -	3
137	NOLC1	Nucleolar phosphoprotein p130 (Nucleolar 130 kDa protein) (140 kDa nucleolar phosphoprotein) (Nopp140) (Nucleolar and coiled-body phosphoprotein 1) -	1
138	HMGB2	High mobility group protein B2 (High mobility group protein 2) (HMG-2) -	1
139	K22E	Keratin, type II cytoskeletal 2 epidermal (Cytokeratin-2e) (K2e) (CK 2e) -	5
140	NP1L1	Nucleosome assembly protein 1-like 1 (NAP-1-related protein) (hNRP) -	2
141	MIF	Macrophage migration inhibitory factor (MIF) (Phenylpyruvate tautomerase) (EC 5.3.2.1) (Glycosylation-inhibiting factor) (GIF) -	2
142	K2C6A	Keratin, type II cytoskeletal 6A (Cytokeratin-6A) (CK 6A) (K6a keratin) -	3
143	PYRG1	CTP synthase 1 (EC 6.3.4.2) (UTP--ammonia ligase 1) (CTP synthetase 1) -	1
144	MDHM	Malate dehydrogenase, mitochondrial precursor (EC 1.1.1.37) -	4
145	IF4A1	Eukaryotic initiation factor 4A-I (EC 3.6.1.-) (ATP-dependent RNA helicase eIF4A-1) (eIF4A-I) (eIF-4A-I) -	2
146	K2C5	Keratin, type II cytoskeletal 5 (Cytokeratin-5) (CK-5) (Keratin-5) (K5) (58 kDa cytokeratin) -	4
147	PHB2	Prohibitin-2 (B-cell receptor-associated protein BAP37) (Repressor of estrogen	2

		receptor activity) (D-prohibitin) -	
148	PGK1	Phosphoglycerate kinase 1 (EC 2.7.2.3) (Primer recognition protein 2) (PRP 2) -	3
149	SERPH	Serpin H1 precursor (Collagen-binding protein) (Colligin) (47 kDa heat shock protein) (Rheumatoid arthritis-related antigen RA-A47) (Arsenic-transactivated protein 3) (AsTP3) (Proliferation-inducing gene 14 protein) -	2
150	PSME2	Proteasome activator complex subunit 2 (Proteasome activator 28-subunit beta) (PA28beta) (PA28b) (Activator of multicatalytic protease subunit 2) (11S regulator complex subunit beta) (REG-beta) -	1
151	PGRC2	Membrane-associated progesterone receptor component 2 (Progesterone membrane-binding protein) (Steroid receptor protein DG6) -	1
152	NP1L4	Nucleosome assembly protein 1-like 4 (Nucleosome assembly protein 2) (NAP2) -	1
153	RS24	40S ribosomal protein S24 -	1
154	MLRM	Myosin regulatory light chain 2, nonsarcomeric (Myosin RLC) -	2
155	FUS	RNA-binding protein FUS (Oncogene FUS) (Oncogene TLS) (Translocated in liposarcoma protein) (POMp75) (75 kDa DNA-pairing protein) -	1
156	SYRC	Arginyl-tRNA synthetase, cytoplasmic (EC 6.1.1.19) (Arginine--tRNA ligase) (ArgRS) -	2
157	HORN	Hornerin -	4
158	PDIA1	Protein disulfide-isomerase precursor (EC 5.3.4.1) (PDI) (Prolyl 4-hydroxylase subunit beta) (Cellular thyroid hormone-binding protein) (p55) -	2
159	HMOX1	Heme oxygenase 1 (EC 1.14.99.3) (HO-1) -	2
160	LA	Lupus La protein (Sjoegren syndrome type B antigen) (SS-B) (La ribonucleoprotein) (La autoantigen) -	1
161	TBCA	Tubulin-specific chaperone A (Tubulin-folding cofactor A) (CFA) (TCP1-chaperonin cofactor A) -	1
162	IF35	Eukaryotic translation initiation factor 3 subunit 5 (eIF-3 epsilon) (eIF3 p47 subunit) (eIF3f) -	1
163	PDIA3	Protein disulfide-isomerase A3 precursor (EC 5.3.4.1) (Disulfide isomerase ER-60) (ERp60) (58 kDa microsomal protein) (p58) (ERp57) (58 kDa glucose-regulated protein) -	3
164	RAB7	Ras-related protein Rab-7 -	2
165	AT1A1	Sodium/potassium-transporting ATPase alpha-1 chain precursor (EC 3.6.3.9) (Sodium	4

		pump 1) (Na(+)/K(+) ATPase 1) -	
166	RS19	40S ribosomal protein S19 -	3
167	TEBP	Prostaglandin E synthase 3 (EC 5.3.99.3) (Cytosolic prostaglandin E2 synthase) (cPGES) (Telomerase-binding protein p23) (Hsp90 co-chaperone) (Progesterone receptor complex p23) -	1
168	API5	Apoptosis inhibitor 5 (API-5) (Fibroblast growth factor 2-interacting factor) (FIF) (Protein XAGL) (Antiapoptosis clone 11 protein) (AAC-11) -	1
169	AN32A	Acidic leucine-rich nuclear phosphoprotein 32 family member A (Potent heat-stable protein phosphatase 2A inhibitor I1PP2A) (Acidic nuclear phosphoprotein pp32) (Leucine-rich acidic nuclear protein) (Lanp) (Putative HLA-DR-associated protein I) (PHAP	1
170	HMGB3	High mobility group protein B3 (High mobility group protein 4) (HMG-4) (High mobility group protein 2a) (HMG-2a) -	1
171	PUR6	Multifunctional protein ADE2 [Includes: Phosphoribosylaminoimidazole-succinocarboxamide synthase (EC 6.3.2.6) (SAICAR synthetase); Phosphoribosylaminoimidazole carboxylase (EC 4.1.1.21) (AIR carboxylase) (AIRC)] -	3
172	AKA12	A-kinase anchor protein 12 (A-kinase anchor protein 250 kDa) (AKAP 250) (Myasthenia gravis autoantigen gravin) -	1
173	K2C4	Keratin, type II cytoskeletal 4 (Cytokeratin-4) (CK-4) (Keratin-4) (K4) -	2
174	RS4X	40S ribosomal protein S4, X isoform (Single copy abundant mRNA protein) (SCR10) -	4
175	IMA1	Importin alpha-1 subunit (Karyopherin alpha-1 subunit) (SRP1-beta) (RAG cohort protein 2) (Nucleoprotein interactor 1) (NPI-1) -	4
176	IF36	Eukaryotic translation initiation factor 3 subunit 6 (eIF-3 p48) (eIF3e) (Viral integration site protein INT-6 homolog) -	1
177	IPO7	Importin-7 (Imp7) (Ran-binding protein 7) (RanBP7) -	1
178	RS29	40S ribosomal protein S29 -	1
179	BASI	Basigin precursor (Leukocyte activation antigen M6) (Collagenase stimulatory factor) (Extracellular matrix metalloproteinase inducer) (EMMPRIN) (5F7) (Tumor cell-derived collagenase stimulatory factor) (TCSF) (OK blood group antigen) (CD147 antigen)	1
180	SEPT2	Septin-2 (Protein NEDD5) -	1

181	DAZP1	DAZ-associated protein 1 (Deleted in azoospermia-associated protein 1) -	1
182	SSRD	Translocon-associated protein subunit delta precursor (TRAP-delta) (Signal sequence receptor subunit delta) (SSR-delta) -	1
183	SYAC	Alanyl-tRNA synthetase, cytoplasmic (EC 6.1.1.7) (Alanine--tRNA ligase) (AlaRS) (Renal carcinoma antigen NY-REN-42) -	1
184	GCN1L	GCN1-like protein 1 (HsGCN1) -	1
185	SF3A2	Splicing factor 3A subunit 2 (Spliceosome-associated protein 62) (SAP 62) (SF3a66) -	3
186	THIO	Thioredoxin (Trx) (ATL-derived factor) (ADF) (Surface-associated sulphydryl protein) (SASP) -	1
187	DHSA	Succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial precursor (EC 1.3.5.1) (Fp) (Flavoprotein subunit of complex II) -	1
188	IMA7	Importin alpha-7 subunit (Karyopherin alpha-6) -	4
189	TKT	Transketolase (EC 2.2.1.1) (TK) -	2
190	RAB6B	Ras-related protein Rab-6B -	1
191	DDX21	Nucleolar RNA helicase 2 (EC 3.6.1.-) (Nucleolar RNA helicase II) (Nucleolar RNA helicase Gu) (RH II/Gu) (Gu-alpha) (DEAD box protein 21) -	1
192	IF39	Eukaryotic translation initiation factor 3 subunit 9 (eIF-3 eta) (eIF3 p116) (eIF3 p110) (eIF3b) (Prt1 homolog) (hPrt1) -	5
193	RBBP4	Histone-binding protein RBBP4 (Retinoblastoma-binding protein 4) (RBBP-4) (Retinoblastoma-binding protein p48) (Chromatin assembly factor 1 subunit C) (CAF-1 subunit C) (Chromatin assembly factor I p48 subunit) (CAF-I 48 kDa subunit) (CAF-I p48) (Nu	2
194	LDHA	L-lactate dehydrogenase A chain (EC 1.1.1.27) (LDH-A) (LDH muscle subunit) (LDH-M) (Proliferation-inducing gene 19 protein) (Renal carcinoma antigen NY-REN-59) -	2
195	GPSN2	Synaptic glycoprotein SC2 -	1
196	MYH9	Myosin-9 (Myosin heavy chain 9) (Myosin heavy chain, nonmuscle IIa) (Nonmuscle myosin heavy chain IIa) (NMMHC II-a) (NMMHC-IIA) (Cellular myosin heavy chain, type A) (Nonmuscle myosin heavy chain-A) (NMMHC-A) -	3
197	ENPL	Endoplasmin precursor (Heat shock protein 90 kDa beta member 1) (94 kDa glucose-regulated protein) (GRP94) (gp96 homolog) (Tumor rejection antigen 1) -	2
198	APT	Adenine phosphoribosyltransferase (EC 2.4.2.7) (APRT) -	2
199	RBBP7	Histone-binding protein RBBP7 (Retinoblastoma-binding protein 7) (RBBP-7)	3

		(Retinoblastoma-binding protein p46) (Histone acetyltransferase type B subunit 2) (Nucleosome remodeling factor subunit RBAP46) -	
200	EWS	RNA-binding protein EWS (EWS oncogene) (Ewing sarcoma breakpoint region 1 protein) -	1
201	KAD2	Adenylate kinase isoenzyme 2, mitochondrial (EC 2.7.4.3) (ATP-AMP transphosphorylase) -	3
202	RS5	40S ribosomal protein S5 -	1
203	ACTN1	Alpha-actinin-1 (Alpha-actinin cytoskeletal isoform) (Non-muscle alpha-actinin-1) (F-actin cross-linking protein) -	2
204	SURF4	Surfeit locus protein 4 -	1
205	SODC	Superoxide dismutase [Cu-Zn] (EC 1.15.1.1) -	2
206	AT1A3	Sodium/potassium-transporting ATPase alpha-3 chain (EC 3.6.3.9) (Sodium pump 3) (Na(+)/K(+) ATPase 3) (Alpha(III)) -	2
207	SMC4	Structural maintenance of chromosomes protein 4 (Chromosome-associated polypeptide C) (hCAP-C) (XCAP-C homolog) -	3
208	PYR1	CAD protein [Includes: Glutamine-dependent carbamoyl-phosphate synthase (EC 6.3.5.5); Aspartate carbamoyltransferase (EC 2.1.3.2); Dihydroorotase (EC 3.5.2.3)] -	2
209	HS74L	Heat shock 70 kDa protein 4L (Osmotic stress protein 94) (Heat shock 70-related protein APG-1) -	3
210	ECHM	Enoyl-CoA hydratase, mitochondrial precursor (EC 4.2.1.17) (Short chain enoyl-CoA hydratase) (SCEH) (Enoyl-CoA hydratase 1) -	1
211	RS28	40S ribosomal protein S28 -	1
212	PRS6B	26S protease regulatory subunit 6B (Proteasome 26S subunit ATPase 4) (MIP224) (MB67-interacting protein) (TAT-binding protein 7) (TBP-7) -	1
213	SAP	Proactivator polypeptide precursor [Contains: Saposin A (Protein A); Saposin B-Val; Saposin B (Sphingolipid activator protein 1) (SAP-1) (Cerebroside sulfate activator) (CSAct) (Dispersin) (Sulfatide/GM1 activator); Saposin C (Co-beta-glucosidase) (	1
214	P5CS	Delta 1-pyrroline-5-carboxylate synthetase (P5CS) (Aldehyde dehydrogenase 18 family member A1) [Includes: Glutamate 5-kinase (EC 2.7.2.11) (Gamma-glutamyl kinase) (GK); Gamma-glutamyl phosphate reductase (GPR) (EC 1.2.1.41) (Glutamate-5-semialdehyde	1
215	K2C7	Keratin, type II cytoskeletal 7 (Cytokeratin-7) (CK-7) (Keratin-7) (K7) (Sarcolectin) -	1

216	K0310	Uncharacterized protein KIAA0310 -	1
217	ERP29	Endoplasmic reticulum protein ERp29 precursor (ERp31) (ERp28) -	1
218	PCCB	Propionyl-CoA carboxylase beta chain, mitochondrial precursor (EC 6.4.1.3) (PCCase subunit beta) (Propanoyl-CoA:carbon dioxide ligase subunit beta) -	2
219	TXLNA	Alpha-taxilin -	1
220	PAIRB	Plasminogen activator inhibitor 1 RNA-binding protein (PAI1 RNA-binding protein 1) (PAI-RBP1) (SERPINE1 mRNA-binding protein 1) -	2
221	DDX3X	ATP-dependent RNA helicase DDX3X (EC 3.6.1.-) (DEAD box protein 3, X-chromosomal) (Helicase-like protein 2) (HLP2) (DEAD box, X isoform) -	1
222	PYGL	Glycogen phosphorylase, liver form (EC 2.4.1.1) -	1
223	RANG	Ran-specific GTPase-activating protein (Ran-binding protein 1) (RanBP1) -	1
224	RENT2	Regulator of nonsense transcripts 2 (Nonsense mRNA reducing factor 2) (Up-frameshift suppressor 2 homolog) (hUpf2) -	1
225	RLA2	60S acidic ribosomal protein P2 (Renal carcinoma antigen NY-REN-44) -	1
226	ACTZ	Alpha-centractin (Centractin) (Centrosome-associated actin homolog) (Actin-RPV) (ARP1) -	2
227	PSD7	26S proteasome non-ATPase regulatory subunit 7 (26S proteasome regulatory subunit rpn8) (26S proteasome regulatory subunit S12) (Proteasome subunit p40) (Mov34 protein homolog) -	1
228	IF4G1	Eukaryotic translation initiation factor 4 gamma 1 (eIF-4-gamma 1) (eIF-4G1) (eIF-4G 1) (p220) -	1
229	HNRPU	Heterogeneous nuclear ribonucleoprotein U (hnRNP U) (Scaffold attachment factor A) (SAF-A) (p120) (pp120) -	3
230	PSME3	Proteasome activator complex subunit 3 (Proteasome activator 28-gamma subunit) (PA28gamma) (PA28g) (Activator of multicatalytic protease subunit 3) (11S regulator complex subunit gamma) (REG-gamma) (Ki nuclear autoantigen) -	1
231	ATPO	ATP synthase O subunit, mitochondrial precursor (EC 3.6.3.14) (Oligomycin sensitivity conferral protein) (OSCP) -	1
232	DYHC	Dynein heavy chain, cytosolic (DYHC) (Cytoplasmic dynein heavy chain 1) (DHC1) (Dynein heavy chain 1, cytoplasmic 1) -	2
233	TADBP	TAR DNA-binding protein 43 (TDP-43) -	1
234	MYL6	Myosin light polypeptide 6 (Smooth muscle and nonmuscle myosin light chain alkali	2

		6) (Myosin light chain alkali 3) (Myosin light chain 3) (MLC-3) (LC17) -	
235	ASNS	Asparagine synthetase [glutamine-hydrolyzing] (EC 6.3.5.4) (Glutamine-dependent asparagine synthetase) (Cell cycle control protein TS11) -	1
236	ATP5J	ATP synthase coupling factor 6, mitochondrial precursor (EC 3.6.3.14) (ATPase subunit F6) -	1
237	RIB2	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 63 kDa subunit precursor (EC 2.4.1.119) (Ribophorin II) (RPN-II) (RIBIIR) -	1
238	UBIQ	Ubiquitin -	2
239	1433Z	14-3-3 protein zeta/delta (Protein kinase C inhibitor protein 1) (KCIP-1) -	1
240	PRPS1	Ribose-phosphate pyrophosphokinase I (EC 2.7.6.1) (Phosphoribosyl pyrophosphate synthetase I) (PRS-I) (PPRibP) -	2
241	SET	Protein SET (Phosphatase 2A inhibitor I2PP2A) (I-2PP2A) (Template-activating factor I) (TAF-I) (HLA-DR-associated protein II) (PHAPII) (Inhibitor of granzyme A-activated DNase) (IGAAD) -	1
242	RSSA	40S ribosomal protein SA (p40) (34/67 kDa laminin receptor) (Colon carcinoma laminin-binding protein) (NEM/1CHD4) (Multidrug resistance-associated protein MGr1-Ag) -	2
243	IMB3	Importin beta-3 (Karyopherin beta-3) (Ran-binding protein 5) (RanBP5) -	1
244	1433E	14-3-3 protein epsilon (14-3-3E) -	1
245	ROAA	Heterogeneous nuclear ribonucleoprotein A/B (hnRNP A/B) (APOBEC-1-binding protein 1) (ABBP-1) -	2
246	NDUS8	NADH dehydrogenase [ubiquinone] iron-sulfur protein 8, mitochondrial precursor (EC 1.6.5.3) (EC 1.6.99.3) (NADH-ubiquinone oxidoreductase 23 kDa subunit) (Complex I-23kD) (CI-23kD) (TYKY subunit) -	1
247	RS10	40S ribosomal protein S10 -	2
248	CND3	Condensin complex subunit 3 (Non-SMC condensin I complex subunit G) (Chromosome-associated protein G) (Condensin subunit CAP-G) (hCAP-G) (XCAP-G homolog) (Antigen NY-MEL-3) -	1
249	DHCA	Carbonyl reductase [NADPH] 1 (EC 1.1.1.184) (NADPH-dependent carbonyl reductase 1) (Prostaglandin-E(2) 9-reductase) (EC 1.1.1.189) (Prostaglandin 9-ketoreductase) (15-hydroxyprostaglandin dehydrogenase [NADP+]) (EC 1.1.1.197) -	1
250	DYL1	Dynein light chain 1, cytoplasmic (Dynein light chain LC8-type 1) (8 kDa dynein light	1

		chain) (DLC8) (Protein inhibitor of neuronal nitric oxide synthase) (PIN) -	
251	RL24	60S ribosomal protein L24 (Ribosomal protein L30) -	1
252	STMN1	Stathmin (Phosphoprotein p19) (pp19) (Oncoprotein 18) (Op18) (Leukemia-associated phosphoprotein p18) (pp17) (Prosolin) (Metablastin) (Protein Pr22) -	1
253	RUVB1	RuvB-like 1 (EC 3.6.1.-) (49 kDa TATA box-binding protein-interacting protein) (49 kDa TBP-interacting protein) (TIP49a) (Pontin 52) (Nuclear matrix protein 238) (NMP 238) (54 kDa erythrocyte cytosolic protein) (ECP-54) (TIP60-associated protein 54-	1
254	ATPK	ATP synthase f chain, mitochondrial (EC 3.6.3.14) -	1
255	HRX	Zinc finger protein HRX (ALL-1) (Trithorax-like protein) -	1
256	PDIA6	Protein disulfide-isomerase A6 precursor (EC 5.3.4.1) (Protein disulfide isomerase P5) (Thioredoxin domain-containing protein 7) -	1
257	SYV	Valyl-tRNA synthetase (EC 6.1.1.9) (Valine--tRNA ligase) (ValRS) (Protein G7a) -	1
258	RCC2	Protein RCC2 (Telophase disk protein of 60 kDa) (RCC1-like protein TD-60) -	1
259	RM10	39S ribosomal protein L10, mitochondrial precursor (L10mt) (MRP-L10) -	1
260	PCBP1	Poly(rC)-binding protein 1 (Alpha-CP1) (hnRNP-E1) (Nucleic acid-binding protein SUB2.3) -	2
261	SPEE	Spermidine synthase (EC 2.5.1.16) (Putrescine aminopropyltransferase) (SPDSY) -	1
262	SYIC	Isoleucyl-tRNA synthetase, cytoplasmic (EC 6.1.1.5) (Isoleucine--tRNA ligase) (IleRS) (IRS) -	3
263	MTA2	Metastasis-associated protein MTA2 (Metastasis-associated 1-like 1) (MTA1-L1 protein) (p53 target protein in deacetylase complex) -	1
264	MYH10	Myosin-10 (Myosin heavy chain 10) (Myosin heavy chain, nonmuscle IIb) (Nonmuscle myosin heavy chain IIb) (NMMHC II-b) (NMMHC-IIB) (Cellular myosin heavy chain, type B) (Nonmuscle myosin heavy chain-B) (NMMHC-B) -	2
265	PRS4	26S protease regulatory subunit 4 (P26s4) (Proteasome 26S subunit ATPase 1) -	1
266	SYTC	Threonyl-tRNA synthetase, cytoplasmic (EC 6.1.1.3) (Threonine--tRNA ligase) (ThrRS) -	1
267	EF1B	Elongation factor 1-beta (EF-1-beta) -	2
268	RL38	60S ribosomal protein L38 -	1
269	PSB2	Proteasome subunit beta type 2 (EC 3.4.25.1) (Proteasome component C7-I) (Macropain subunit C7-I) (Multicatalytic endopeptidase complex subunit C7-I) -	1
270	OST48	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 48 kDa subunit	2

		precursor (EC 2.4.1.119) (Oligosaccharyl transferase 48 kDa subunit) (DDOST 48 kDa subunit) -	
271	KNG1	Kininogen-1 precursor (Alpha-2-thiol proteinase inhibitor) [Contains: Kininogen-1 heavy chain; Bradykinin (Kallidin I); Lysyl-bradykinin (Kallidin II); Kininogen-1 light chain; Low molecular weight growth-promoting factor] -	1
272	APEX1	DNA-(apurinic or apyrimidinic site) lyase (EC 4.2.99.18) (AP endonuclease 1) (APEX nuclease) (APEN) (REF-1 protein) -	2
273	TFR1	Transferrin receptor protein 1 (TfR1) (TR) (TfR) (Trfr) (CD71 antigen) (T9) (p90) -	1
274	RL10A	60S ribosomal protein L10a (CSA-19) -	1
275	H2A1A	Histone H2A type 1-A -	1
276	DPY30	Dpy-30-like protein -	1
277	PRDX3	Thioredoxin-dependent peroxide reductase, mitochondrial precursor (EC 1.11.1.15) (Peroxiredoxin-3) (PRX III) (Antioxidant protein 1) (AOP-1) (Protein MER5 homolog) (HBC189) -	1
278	RL12	60S ribosomal protein L12 -	1
279	KLDC2	Kelch domain-containing protein 2 (Hepatocellular carcinoma-associated antigen 33) (Host cell factor homolog LCP) -	1
280	DDX46	Probable ATP-dependent RNA helicase DDX46 (EC 3.6.1.-) (DEAD box protein 46) (PRP5 homolog) -	1
281	RCN2	Reticulocalbin-2 precursor (Calcium-binding protein ERC-55) (E6-binding protein) (E6BP) -	2
282	SFRS3	Splicing factor, arginine/serine-rich 3 (Pre-mRNA-splicing factor SRP20) -	1
283	CPNE1	Copine-1 (Copine I) -	1
284	RS27L	40S ribosomal protein S27-like protein -	1
285	UBP7	Ubiquitin carboxyl-terminal hydrolase 7 (EC 3.1.2.15) (Ubiquitin thioesterase 7) (Ubiquitin-specific-processing protease 7) (Deubiquitinating enzyme 7) (Herpesvirus-associated ubiquitin-specific protease) -	1
286	C1TC	C-1-tetrahydrofolate synthase, cytoplasmic (C1-THF synthase) [Includes: Methylenetetrahydrofolate dehydrogenase (EC 1.5.1.5); Methenyltetrahydrofolate cyclohydrolase (EC 3.5.4.9); Formyltetrahydrofolate synthetase (EC 6.3.4.3)] - (Human	1
287	PSMD3	26S proteasome non-ATPase regulatory subunit 3 (26S proteasome regulatory subunit	1

		S3) (Proteasome subunit p58) -	
288	RCC1	Regulator of chromosome condensation (Chromosome condensation protein 1) (Cell cycle regulatory protein) -	1
289	HDGF	Hepatoma-derived growth factor (HDGF) (High-mobility group protein 1-like 2) (HMG-1L2) -	1
290	RD23B	UV excision repair protein RAD23 homolog B (hHR23B) (XP-C repair-complementing complex 58 kDa protein) (p58) -	1
291	NUP50	Nucleoporin 50 kDa (Nuclear pore-associated protein 60 kDa-like) -	2
292	FLNA	Filamin-A (Alpha-filamin) (Filamin-1) (Endothelial actin-binding protein) (Actin-binding protein 280) (ABP-280) (Nonmuscle filamin) -	2
293	SYQ	Glutaminyl-tRNA synthetase (EC 6.1.1.18) (Glutamine--tRNA ligase) (GlnRS) -	1
294	RS15A	40S ribosomal protein S15a -	1
295	HXK1	Hexokinase-1 (EC 2.7.1.1) (Hexokinase type I) (HK I) (Brain form hexokinase) -	1
296	FXR1	Fragile X mental retardation syndrome-related protein 1 (hFXR1p) -	1
297	HINT1	Histidine triad nucleotide-binding protein 1 (Adenosine 5~-monophosphoramidase) (Protein kinase C inhibitor 1) (Protein kinase C-interacting protein 1) (PKCI-1) -	1
298	JIP4	C-jun-amino-terminal kinase-interacting protein 4 (JNK-interacting protein 4) (JIP-4) (JNK-associated leucine-zipper protein) (JLP) (Sperm-associated antigen 9) (Mitogen-activated protein kinase 8-interacting protein 4) (Human lung cancer protein 6)	1
299	RS23	40S ribosomal protein S23 -	1
300	G3BP2	Ras GTPase-activating protein-binding protein 2 (G3BP-2) (GAP SH3 domain-binding protein 2) -	1
301	SRP09	Signal recognition particle 9 kDa protein (SRP9) -	1
302	TIF1B	Transcription intermediary factor 1-beta (TIF1-beta) (Tripartite motif-containing protein 28) (Nuclear corepressor KAP-1) (KRAB-associated protein 1) (KAP-1) (KRAB-interacting protein 1) (KRIP-1) (RING finger protein 96) -	1
303	PERQ2	PERQ amino acid-rich with GYF domain-containing protein 2 (Grb10-interacting GYF protein 2) (Trinucleotide repeat-containing protein 15) -	1
304	GCSH	Glycine cleavage system H protein, mitochondrial precursor -	2
305	PSMD6	26S proteasome non-ATPase regulatory subunit 6 (26S proteasome regulatory subunit S10) (p42A) (Proteasome regulatory particle subunit p44S10) (Phosphonoformate immuno-associated protein 4) (Breast cancer-associated protein SGA-113M) - (	1

306	PSA	Puromycin-sensitive aminopeptidase (EC 3.4.11.-) (PSA) -	1
307	HCDH	Hydroxyacyl-coenzyme A dehydrogenase, mitochondrial precursor (EC 1.1.1.35) (Short chain 3-hydroxyacyl-CoA dehydrogenase) (HCDH) (Medium and short chain L-3-hydroxyacyl-coenzyme A dehydrogenase) -	1
308	U520	U5 small nuclear ribonucleoprotein 200 kDa helicase (EC 3.6.1.-) (U5 snRNP-specific 200 kDa protein) (U5-200KD) (Activating signal cointegrator 1 complex subunit 3-like 1) (BRR2 homolog) -	1
309	RL7A	60S ribosomal protein L7a (Surfeit locus protein 3) (PLA-X polypeptide) -	1
310	IF34	Eukaryotic translation initiation factor 3 subunit 4 (eIF-3 delta) (eIF3 p44) (eIF-3 RNA-binding subunit) (eIF3 p42) (eIF3g) -	2
311	GBF1	Golgi-specific brefeldin A-resistance guanine nucleotide exchange factor 1 (BFA-resistant GEF 1) -	1
312	AMPL	Cytosol aminopeptidase (EC 3.4.11.1) (Leucine aminopeptidase) (LAP) (Leucyl aminopeptidase) (Leucine aminopeptidase 3) (Proline aminopeptidase) (EC 3.4.11.5) (Prolyl aminopeptidase) (Peptidase S) -	1
313	KAP2	cAMP-dependent protein kinase type II-alpha regulatory subunit -	1
314	SRP68	Signal recognition particle 68 kDa protein (SRP68) -	2
315	RL23	60S ribosomal protein L23 (Ribosomal protein L17) -	1
316	IGHG2	Ig gamma-2 chain C region -	1
317	HSPB1	Heat-shock protein beta-1 (HspB1) (Heat shock 27 kDa protein) (HSP 27) (Stress-responsive protein 27) (SRP27) (Estrogen-regulated 24 kDa protein) (28 kDa heat shock protein) -	1
318	TXND5	Thioredoxin domain-containing protein 5 precursor (Thioredoxin-like protein p46) (Endoplasmic reticulum protein ERp46) -	2
319	RL18	60S ribosomal protein L18 -	1
320	RGS3	Regulator of G-protein signaling 3 (RGS3) (RGP3) -	1
321	LAMP1	Lysosome-associated membrane glycoprotein 1 precursor (LAMP-1) (CD107a antigen) -	1
322	PHB	Prohibitin -	1
323	RS20	40S ribosomal protein S20 -	2
324	COTL1	Coactosin-like protein -	1
325	RL30	60S ribosomal protein L30 -	1

326	TRY1	Trypsin-1 precursor (EC 3.4.21.4) (Trypsin I) (Cationic trypsinogen) (Serine protease 1) -	1
327	RS7	40S ribosomal protein S7 -	1
328	RL17	60S ribosomal protein L17 (L23) -	1
329	SC23A	Protein transport protein Sec23A (SEC23-related protein A) -	1
330	LRC59	Leucine-rich repeat-containing protein 59 -	1
331	SAHH	Adenosylhomocysteinase (EC 3.3.1.1) (S-adenosyl-L-homocysteine hydrolase) (AdoHcyase) -	1
332	RS9	40S ribosomal protein S9 -	1
333	SGTA	Small glutamine-rich tetratricopeptide repeat-containing protein A (Vpu-binding protein) (UBP) -	1
334	LAP2A	Lamina-associated polypeptide 2 isoform alpha (Thymopoietin isoform alpha) (TP alpha) (Thymopoietin-related peptide isoform alpha) (TPRP isoform alpha) [Contains: Thymopoietin (TP) (Splenin); Thymopentin (TP5)] -	1
335	GANAB	Neutral alpha-glucosidase AB precursor (EC 3.2.1.84) (Glucosidase II subunit alpha) -	1
336	OXRP	150 kDa oxygen-regulated protein precursor (Orp150) (Hypoxia up-regulated 1) -	1
337	RL8	60S ribosomal protein L8 -	1
338	ENAM	Enamelin precursor -	1
339	MBB1A	Myb-binding protein 1A -	1
340	YES	Proto-oncogene tyrosine-protein kinase Yes (EC 2.7.10.2) (p61-Yes) (c-Yes) -	1
341	CD2L7	Cell division cycle 2-related protein kinase 7 (EC 2.7.11.22) (CDC2-related protein kinase 7) (Cdc2-related kinase, arginine/serine-rich) (CrkRS) -	1
342	RS14	40S ribosomal protein S14 -	1
343	SF3B3	Splicing factor 3B subunit 3 (Spliceosome-associated protein 130) (SAP 130) (SF3b130) (Pre-mRNA-splicing factor SF3b 130 kDa subunit) (STAF130) -	1
345	STT3A	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit STT3A (EC 2.4.1.119) (Oligosaccharyl transferase subunit STT3A) (STT3-A) (B5) (Integral membrane protein 1) (TMC) -	1
346	GP162	Probable G-protein coupled receptor 162 (Gene-rich cluster gene A protein) -	1
347	AT2A2	Sarcoplasmic/endoplasmic reticulum calcium ATPase 2 (EC 3.6.3.8) (Calcium pump 2) (SERCA2) (SR Ca(2+)-ATPase 2) (Calcium-transporting ATPase sarcoplasmic reticulum type, slow twitch skeletal muscle isoform) (Endoplasmic reticulum class 1/2	1

		Ca(2+) AT	
348	PCCA	Propionyl-CoA carboxylase alpha chain, mitochondrial precursor (EC 6.4.1.3) (PCCase subunit alpha) (Propanoyl-CoA:carbon dioxide ligase subunit alpha) -	1
349	LPPRC	130 kDa leucine-rich protein (LRP 130) (GP130) (Leucine-rich PPR motif-containing protein) -	1
350	ANM1	Protein arginine N-methyltransferase 1 (EC 2.1.1.-) (Interferon receptor 1-bound protein 4) -	1
352	CF060	Uncharacterized protein C6orf60 -	1
353	LAT1	Large neutral amino acids transporter small subunit 1 (L-type amino acid transporter 1) (4F2 light chain) (4F2 LC) (4F2LC) (CD98 light chain) (Integral membrane protein E16) (hLAT1) -	1
354	LAP2B	Lamina-associated polypeptide 2, isoforms beta/gamma (Thymopoietin, isoforms beta/gamma) (TP beta/gamma) (Thymopoietin-related peptide isoforms beta/gamma) (TPRP isoforms beta/gamma) [Contains: Thymopoietin (TP) (Splenin); Thymopentin (TP5)] - Homo	1
355	SND1	Staphylococcal nuclease domain-containing protein 1 (p100 co-activator) (100 kDa coactivator) (EBNA2 coactivator p100) -	1
356	K0146	Uncharacterized protein KIAA0146 -	1
357	BTNL3	Butyrophilin-like protein 3 precursor (Butyrophilin-like receptor) -	1
358	LAMP2	Lysosome-associated membrane glycoprotein 2 precursor (LAMP-2) (CD107b antigen) -	1
359	PROF1	Profilin-1 (Profilin I) -	1
360	CORT	Cortistatin precursor [Contains: Cortistatin-29; Cortistatin-17] -	1
361	PUR2	Trifunctional purine biosynthetic protein adenosine-3 [Includes: Phosphoribosylamine--glycine ligase (EC 6.3.4.13) (GARS) (Glycinamide ribonucleotide synthetase) (Phosphoribosylglycinamide synthetase); Phosphoribosylformylglycinamidine cyclo-ligase	1
362	PTBP1	Polypyrimidine tract-binding protein 1 (PTB) (Heterogeneous nuclear ribonucleoprotein I) (hnRNP I) (57 kDa RNA-binding protein PPTB-1) -	1
363	XPO5	Exportin-5 (Exp5) (Ran-binding protein 21) -	1
364	PR285	Peroxisomal proliferator-activated receptor A-interacting complex 285 kDa protein (EC 3.6.1.-) (ATP-dependent helicase PRIC285) (PPAR-alpha-interacting complex	1

		protein 285) (PPAR-gamma DBD-interacting protein 1) (PDIP1) -	
365	PFD2	Prefoldin subunit 2 -	1
366	PIGR	Polymeric-immunoglobulin receptor precursor (Poly-Ig receptor) (PIGR) (Hepatocellular carcinoma-associated protein TB6) [Contains: Secretory component] -	1
367	MMP15	Matrix metalloproteinase-15 precursor (EC 3.4.24.-) (MMP-15) (Membrane-type matrix metalloproteinase 2) (MT-MMP 2) (MTMMP2) (Membrane-type-2 matrix metalloproteinase) (MT2-MMP) (MT2MMP) (SMCP-2) -	1
368	PUR9	Bifunctional purine biosynthesis protein PURH [Includes: Phosphoribosylaminoimidazolecarboxamide formyltransferase (EC 2.1.2.3) (5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase) (AICAR transformylase); IMP cyclohydrolase (EC 3.5.4.10	1
369	ROA0	Heterogeneous nuclear ribonucleoprotein A0 (hnRNP A0) -	1
370	IF4H	Eukaryotic translation initiation factor 4H (eIF-4H) (Williams-Beuren syndrome chromosome region 1 protein) -	1
371	ENSA	Alpha-endosulfine (ARPP-19e) -	1
372	H1X	Histone H1x -	1
373	UAP56	Spliceosome RNA helicase BAT1 (EC 3.6.1.-) (DEAD box protein UAP56) (56 kDa U2AF65-associated protein) (ATP-dependent RNA helicase p47) (HLA-B-associated transcript-1) -	1
374	MCA3	Eukaryotic translation elongation factor 1 epsilon-1 (Multisynthetase complex auxiliary component p18) (Elongation factor p18) -	1
375	PARP1	Poly [ADP-ribose] polymerase 1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+) ADP-ribosyltransferase 1) (Poly[ADP-ribose] synthetase 1) -	1
377	RL36	60S ribosomal protein L36 -	1
378	ANM5	Protein arginine N-methyltransferase 5 (EC 2.1.1.125) (EC 2.1.1.-) (Shk1 kinase-binding protein 1 homolog) (SKB1Hs) (Jak-binding protein 1) (72 kDa ICln-binding protein) -	1
379	FLNB	Filamin-B (FLN-B) (Beta-filamin) (Actin-binding-like protein) (Thyroid autoantigen) (Truncated actin-binding protein) (Truncated ABP) (ABP-280 homolog) (ABP-278) (Filamin 3) (Filamin homolog 1) (Fh1) -	1
380	FAK1	Focal adhesion kinase 1 (EC 2.7.10.2) (FADK 1) (pp125FAK) (Protein-tyrosine kinase 2) -	1

381	CENPU	Centromere protein U (CENP-U) (CENP-U(50)) (CENP-50) (Interphase centromere complex protein 24) (MLF1-interacting protein) (KSHV latent nuclear antigen-interacting protein 1) -	1
382	ZN516	Zinc finger protein 516 -	1
383	GP175	Integral membrane protein GPR175 -	1
384	PRKDC	DNA-dependent protein kinase catalytic subunit (EC 2.7.11.1) (DNA-PK catalytic subunit) (DNA-PKcs) (DNPK1) (p460) -	1
385	IF38	Eukaryotic translation initiation factor 3 subunit 8 (eIF3 p110) (eIF3c) -	1
386	AF9	Protein AF-9 (ALL1 fused gene from chromosome 9 protein) (Myeloid/lymphoid or mixed-lineage leukemia translocated to chromosome 3 protein) (YEATS domain-containing protein 3) -	1
387	GRN	Granulins precursor (Proepithelin) (PEPI) [Contains: Acrogranin; Paragranulin; Granulin-1 (Granulin G); Granulin-2 (Granulin F); Granulin-3 (Granulin B); Granulin-4 (Granulin A); Granulin-5 (Granulin C); Granulin-6 (Granulin D); Granulin-7 (Granulin	1
388	CSK2B	Casein kinase II subunit beta (CK II beta) (Phosvitin) (G5a) -	2
389	LRMP	Lymphoid-restricted membrane protein (Protein Jaw1) -	1
390	SEPT7	Septin-7 (CDC10 protein homolog) -	1
391	TM165	Transmembrane protein 165 (Transmembrane protein TPARL) (Transmembrane protein PT27) -	1
392	SUHW3	Suppressor of hairy wing homolog 3 -	1
393	IF37	Eukaryotic translation initiation factor 3 subunit 7 (eIF-3 zeta) (eIF3 p66) (eIF3d) -	1
394	SYDC	Aspartyl-tRNA synthetase, cytoplasmic (EC 6.1.1.12) (Aspartate--tRNA ligase) (AspRS) (Cell proliferation-inducing gene 40 protein) -	1
395	RM27	Mitochondrial 39S ribosomal protein L27 (L27mt) (MRP-L27) -	1
397	NHERF	Ezrin-radixin-moesin-binding phosphoprotein 50 (EBP50) (Na(+)/H(+) exchange regulatory cofactor NHE-RF) (NHERF-1) (Regulatory cofactor of Na(+)/H(+) exchanger) (Sodium-hydrogen exchanger regulatory factor 1) (Solute carrier family 9 isoform 3 regula	1
398	BOLA2	BolA-like protein 2 -	1
399	CCNT2	Cyclin-T2 (CycT2) -	1
400	RL3L	60S ribosomal protein L3-like -	1

401	IGF1B	Insulin-like growth factor IB precursor (IGF-IB) (Somatomedin C) (Mechano growth factor) (MGF) -	1
402	TCOF	Treacle protein (Treacher Collins syndrome protein) -	1
403	NPA1P	Nucleolar preribosomal-associated protein 1 (Fragment) -	1
404	RAB2A	Ras-related protein Rab-2A -	1
405	UN45B	UNC45 homolog B (UNC-45B) -	1
406	LN28A	Lin-28 homolog A (Zinc finger CCHC domain-containing protein 1) -	1
407	2AAA	Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A alpha isoform (PP2A, subunit A, PR65-alpha isoform) (PP2A, subunit A, R1-alpha isoform) (Medium tumor antigen-associated 61 kDa protein) -	1
408	GROA	Growth-regulated protein alpha precursor (CXCL1) (Melanoma growth stimulatory activity) (MGSA) (Neutrophil-activating protein 3) (NAP-3) (GRO-alpha(1-73)) [Contains: GRO-alpha(4-73); GRO-alpha(5-73); GRO-alpha(6-73)] -	1
410	RUXE	Small nuclear ribonucleoprotein E (snRNP-E) (Sm protein E) (Sm-E) (SmE) -	1
411	CG029	Uncharacterized protein C7orf29 -	1
412	MSH2	DNA mismatch repair protein Msh2 (MutS protein homolog 2) -	1
414	INOC1	Putative DNA helicase INO80 complex homolog 1 (EC 3.6.1.-) (hINO80) -	1
416	CKAP4	Cytoskeleton-associated protein 4 (63 kDa membrane protein) (p63) -	1
417	SFRS2	Splicing factor, arginine/serine-rich 2 (Splicing factor SC35) (SC-35) (Splicing component, 35 kDa) (Protein PR264) -	1
419	CENA1	Centaurin-alpha 1 (Putative MAPK-activating protein PM25) -	1

Supplementry table 4 B (S4B) AF4•MLL

Protein hit number	Protein accession number number	Protein protein description	peptides
1	HRX	Zinc finger protein HRX (ALL-1) (Trithorax-like protein)	131
2	HSP71	Heat shock 70 kDa protein 1 (HSP70.1) (HSP70-1/HSP70-2)	38
3	AFF1	AF4/FMR2 family member 1 (Protein AF-4) (Proto-oncogene AF4) (Protein FEL)	44
4	HSP7C	Heat shock cognate 71 kDa protein (Heat shock 70 kDa protein 8)	23
5	GRP78	78 kDa glucose-regulated protein precursor (GRP 78) (Heat shock 70 kDa protein 5) (Immunoglobulin heavy chain-binding protein) (BiP) (Endoplasmic reticulum lumenal Ca(2+)-binding protein grp78)	22
6	HSP76	Heat shock 70 kDa protein 6 (Heat shock 70 kDa protein B~)	13
7	GRP75	Stress-70 protein, mitochondrial precursor (75 kDa glucose-regulated protein) (GRP 75) (Peptide-binding protein 74) (PBP74) (Mortalin) (MOT)	10
8	TBA6	Tubulin alpha-6 chain (Alpha-tubulin 6)	7
9	TBB5	Tubulin beta chain (Tubulin beta-5 chain)	11
10	TBB2C	Tubulin beta-2C chain (Tubulin beta-2 chain)	9
11	ACTB	Actin, cytoplasmic 1 (Beta-actin)	6
12	HCFC1	Host cell factor (HCF) (HCF-1) (C1 factor) (VP16 accessory protein) (VCAF) (CFF) [Contains: HCF N-terminal chain 1; HCF N-terminal chain 2; HCF N-terminal chain 3; HCF N-terminal chain 4; HCF N-terminal chain 5; HCF N-terminal chain 6; HCF C-termina	14
13	PIMT	Protein-L-isoaspartate(D-aspartate) O-methyltransferase (EC 2.1.1.77) (Protein-beta-aspartate methyltransferase) (PIMT) (Protein L-isoaspartyl/D-aspartyl methyltransferase) (L-isoaspartyl protein carboxyl methyltransferase)	7
14	RBBP5	Retinoblastoma-binding protein 5 (RBBP-5) (Retinoblastoma-binding protein RBQ-3)	4
15	CSK2B	Casein kinase II subunit beta (CK II beta) (Phosvitin) (G5a)	2
16	ATX2L	Ataxin-2-like protein (Ataxin-2 domain protein) (Ataxin-2-related protein)	9
17	CH60	60 kDa heat shock protein, mitochondrial precursor (Hsp60) (60 kDa chaperonin) (CPN60) (Heat shock protein 60) (HSP-60) (Mitochondrial matrix protein P1) (P60 lymphocyte protein) (HuCHA60)	4

18	HS105	Heat-shock protein 105 kDa (Heat shock 110 kDa protein) (Antigen NY-CO-25)	6
19	EF1G	Elongation factor 1-gamma (EF-1-gamma) (eEF-1B gamma)	5
20	WDR5	WD repeat protein 5 (BMP2-induced 3-kb gene protein)	4
21	MIF	Macrophage migration inhibitory factor (MIF) (Phenylpyruvate tautomerase) (EC 5.3.2.1) (Glycosylation-inhibiting factor) (GIF)	2
22	DDX6	Probable ATP-dependent RNA helicase DDX6 (EC 3.6.1.-) (DEAD box protein 6) (ATP-dependent RNA helicase p54) (Oncogene RCK)	4
23	HSP74	Heat shock 70 kDa protein 4 (Heat shock 70-related protein APG-2) (HSP70RY)	2
24	CSK21	Casein kinase II subunit alpha (EC 2.7.11.1) (CK II)	3
25	ASH2L	Set1/Ash2 histone methyltransferase complex subunit ASH2 (ASH2-like protein)	4
26	ADT2	ADP/ATP translocase 2 (Adenine nucleotide translocator 2) (ANT 2) (ADP,ATP carrier protein 2) (Solute carrier family 25 member 5) (ADP,ATP carrier protein, fibroblast isoform)	2
27	HS90B	Heat shock protein HSP 90-beta (HSP 84) (HSP 90)	2
28	CDK9	Cell division protein kinase 9 (EC 2.7.11.22) (EC 2.7.11.23) (Cyclin-dependent kinase 9) (Serine/threonine-protein kinase PITALRE) (C-2K) (Cell division cycle 2-like protein kinase 4)	2
29	DPY30	Dpy-30-like protein	2
30	HS90A	Heat shock protein HSP 90-alpha (HSP 86) (Renal carcinoma antigen NY-REN-38)	3
31	EF1A1	Elongation factor 1-alpha 1 (EF-1-alpha-1) (Elongation factor 1 A-1) (eEF1A-1) (Elongation factor Tu) (EF-Tu) (Leukocyte receptor cluster member 7)	3
32	EF2	Elongation factor 2 (EF-2)	1
33	COA1	Acetyl-CoA carboxylase 1 (EC 6.4.1.2) (ACC-alpha) [Includes: Biotin carboxylase (EC 6.3.4.14)]	1
34	CCNT1	Cyclin-T1 (CycT1) (Cyclin-T)	3
35	PRDX1	Peroxiredoxin-1 (EC 1.11.1.15) (Thioredoxin peroxidase 2) (Thioredoxin-dependent peroxide reductase 2) (Proliferation-associated gene protein) (PAG) (Natural killer cell-enhancing factor A) (NKEF-A)	2
36	RS29	40S ribosomal protein S29	1
37	SET	Protein SET (Phosphatase 2A inhibitor I2PP2A) (I-2PP2A) (Template-activating factor I) (TAF-I) (HLA-DR-associated protein II) (PHAPII) (Inhibitor of granzyme A-activated DNase) (IGAAD)	1

38	PPIA	Peptidyl-prolyl cis-trans isomerase A (EC 5.2.1.8) (PPIase A) (Rotamase A) (Cyclophilin A) (Cyclosporin A-binding protein)	2
39	UTS2	Urotensin-2 precursor (Urotensin-II) (U-II) (UII)	1
40	GCSH	Glycine cleavage system H protein, mitochondrial precursor	1
41	ACTN1	Alpha-actinin-1 (Alpha-actinin cytoskeletal isoform) (Non-muscle alpha-actinin-1) (F-actin cross-linking protein)	1
42	PRDX4	Peroxiredoxin-4 (EC 1.11.1.15) (Prx-IV) (Thioredoxin peroxidase AO372) (Thioredoxin-dependent peroxide reductase A0372) (Antioxidant enzyme AOE372) (AOE37-2)	2
43	MCCC2	Methylcrotonoyl-CoA carboxylase beta chain, mitochondrial precursor (EC 6.4.1.4) (3-Methylcrotonyl-CoA carboxylase 2) (MCCase subunit beta) (3-methylcrotonyl-CoA:carbon dioxide ligase subunit beta) (3-Methylcrotonyl-CoA carboxylase non-biotin-contai	2
44	GDE	Glycogen debranching enzyme (Glycogen debrancher) [Includes: 4-alpha-glucanotransferase (EC 2.4.1.25) (Oligo-1,4-1,4-glucantransferase); Amylo-alpha-1,6-glucosidase (EC 3.2.1.33) (Amylo-1,6-glucosidase) (Dextrin 6-alpha-D-glucosidase)] - Homo sapien	2
45	TCPE	T-complex protein 1 subunit epsilon (TCP-1-epsilon) (CCT-epsilon)	1
46	PSME3	Proteasome activator complex subunit 3 (Proteasome activator 28-gamma subunit) (PA28gamma) (PA28g) (Activator of multicatalytic protease subunit 3) (11S regulator complex subunit gamma) (REG-gamma) (Ki nuclear autoantigen)	3
47	RS9	40S ribosomal protein S9	1
48	TCPB	T-complex protein 1 subunit beta (TCP-1-beta) (CCT-beta)	2
49	FAS	Fatty acid synthase (EC 2.3.1.85) [Includes: [Acyl-carrier-protein] S-acetyltransferase (EC 2.3.1.38); [Acyl-carrier-protein] S-malonyltransferase (EC 2.3.1.39); 3-oxoacyl-[acyl-carrier-protein] synthase (EC 2.3.1.41); 3-oxoacyl-[acyl-carrier-protei	1
50	HNRPF	Heterogeneous nuclear ribonucleoprotein F (hnRNP F) (Nucleolin-like protein mcs94-1)	1
51	SERA	D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95) (3-PGDH)	1
52	ROA1	Heterogeneous nuclear ribonucleoprotein A1 (Helix-destabilizing protein) (Single-strand RNA-binding protein) (hnRNP core protein A1)	1
53	SFPQ	Splicing factor, proline- and glutamine-rich (Polypyrimidine tract-binding protein-	1

		associated-splicing factor) (PTB-associated-splicing factor) (PSF) (DNA-binding p52/p100 complex, 100 kDa subunit) (100 kDa DNA-pairing protein) (hPOMp100) - Homo sap	
54	HCFC2	Host cell factor 2 (HCF-2) (C2 factor)	2
55	PP2AA	Serine/threonine-protein phosphatase 2A catalytic subunit alpha isoform (EC 3.1.3.16) (PP2A-alpha) (Replication protein C) (RP-C)	1
56	SF3A1	Splicing factor 3 subunit 1 (Spliceosome-associated protein 114) (SAP 114) (SF3a120)	1
57	TCPG	T-complex protein 1 subunit gamma (TCP-1-gamma) (CCT-gamma) (hTRiC5)	1
58	CLH1	Clathrin heavy chain 1 (CLH-17)	1
59	SYT4	Synaptotagmin-4 (Synaptotagmin IV) (SytIV)	1
60	PGAM1	Phosphoglycerate mutase 1 (EC 5.4.2.1) (EC 5.4.2.4) (EC 3.1.3.13) (Phosphoglycerate mutase isozyme B) (PGAM-B) (BPG-dependent PGAM 1)	1
61	CH10	10 kDa heat shock protein, mitochondrial (Hsp10) (10 kDa chaperonin) (CPN10) (Early-pregnancy factor) (EPF)	1
62	ASPP1	Apoptosis-stimulating of p53 protein 1 (Protein phosphatase 1 regulatory subunit 13B)	1
63	NPM	Nucleophosmin (NPM) (Nucleolar phosphoprotein B23) (Numatrin) (Nucleolar protein NO38)	1
64	PGRC1	Membrane-associated progesterone receptor component 1 (mPR)	1
65	G3P	Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH)	1
66	ANK2	Ankyrin-2 (Brain ankyrin) (Ankyrin-B) (Ankyrin, nonerythroid)	1
67	HS74L	Heat shock 70 kDa protein 4L (Osmotic stress protein 94) (Heat shock 70-related protein APG-1)	1
68	SUCA	Succinyl-CoA ligase [GDP-forming] subunit alpha, mitochondrial precursor (EC 6.2.1.4) (Succinyl-CoA synthetase subunit alpha) (SCS-alpha)	1
69	COG1	Conserved oligomeric Golgi complex component 1	1
70	IMB1	Importin beta-1 subunit (Karyopherin beta-1 subunit) (Nuclear factor P97) (Importin 90)	1
71	KCRB	Creatine kinase B-type (EC 2.7.3.2) (Creatine kinase B chain) (B-CK)	1
72	K2C1	Keratin, type II cytoskeletal 1 (Cytokeratin-1) (CK-1) (Keratin-1) (K1) (67 kDa cytokeratin) (Hair alpha protein)	2
73	L1CAM	Neural cell adhesion molecule L1 precursor (N-CAM L1) (CD171 antigen)	1
74	WASF2	Wiskott-Aldrich syndrome protein family member 2 (WASP-family protein member 2)	1

		(Protein WAVE-2) (Verprolin homology domain-containing protein 2)	
75	K22E	Keratin, type II cytoskeletal 2 epidermal (Cytokeratin-2e) (K2e) (CK 2e)	1
76	HTRA3	Probable serine protease HTRA3 precursor (EC 3.4.21.-) (High-temperature requirement factor A3) (Pregnancy-related serine protease)	1
78	STMN1	Stathmin (Phosphoprotein p19) (pp19) (Oncoprotein 18) (Op18) (Leukemia-associated phosphoprotein p18) (pp17) (Prosolin) (Metablastin) (Protein Pr22)	1
81	RAN	GTP-binding nuclear protein Ran (GTPase Ran) (Ras-like protein TC4) (Androgen receptor-associated protein 24)	1
83	XYLT1	Xylosyltransferase 1 (EC 2.4.2.26) (Xylosyltransferase I) (XylT-I) (XT-I) (Peptide O-xylosyltransferase 1)	1
84	CSK22	Casein kinase II subunit alpha~ (EC 2.7.11.1) (CK II)	1
85	MAGI1	Membrane-associated guanylate kinase, WW and PDZ domain-containing protein 1 (BAI1-associated protein 1) (BAP-1) (Membrane-associated guanylate kinase inverted 1) (MAGI-1) (Atrophin-1-interacting protein 3) (AIP3) (WW domain-containing protein 3)	1
87	SNAT	Serotonin N-acetyltransferase (EC 2.3.1.87) (Aralkylamine N-acetyltransferase) (AA-NAT) (Serotonin acetylase)	1

Supplementry table 4 C (S4C) ENL

Protein hit number	Protein accession number	Protein description	peptides
1	K2C1	Keratin, type II cytoskeletal 1 (Cytokeratin-1) (CK-1) (Keratin-1) (K1) (67 kDa cytokeratin) (Hair alpha protein)	11
2	HSP71	Heat shock 70 kDa protein 1 (HSP70.1) (HSP70-1/HSP70-2)	14
3	K1C9	Keratin, type I cytoskeletal 9 (Cytokeratin-9) (CK-9) (Keratin-9) (K9)	6
4	AFF1	AF4/FMR2 family member 1 (Protein AF-4) (Proto-oncogene AF4) (Protein FEL)	21
5	KV206	Ig kappa chain V-II region RPMI 6410 precursor	1
6	HSP76	Heat shock 70 kDa protein 6 (Heat shock 70 kDa protein B~)	3
7	TBAK	Tubulin alpha-ubiquitous chain (Alpha-tubulin ubiquitous) (Tubulin K-alpha-1)	6
8	TBB5	Tubulin beta chain (Tubulin beta-5 chain)	6
9	HSP7C	Heat shock cognate 71 kDa protein (Heat shock 70 kDa protein 8)	5
10	ACTB	Actin, cytoplasmic 1 (Beta-actin)	5
11	K1C10	Keratin, type I cytoskeletal 10 (Cytokeratin-10) (CK-10) (Keratin-10) (K10)	5
12	GRP78	78 kDa glucose-regulated protein precursor (GRP 78) (Heat shock 70 kDa protein 5) (Immunoglobulin heavy chain-binding protein) (BiP) (Endoplasmic reticulum lumenal Ca(2+)-binding protein grp78)	3
13	ALBU	Serum albumin precursor	4
14	K2C4	Keratin, type II cytoskeletal 4 (Cytokeratin-4) (CK-4) (Keratin-4) (K4)	2
15	K2C6A	Keratin, type II cytoskeletal 6A (Cytokeratin-6A) (CK 6A) (K6a keratin)	3
16	K2C5	Keratin, type II cytoskeletal 5 (Cytokeratin-5) (CK-5) (Keratin-5) (K5) (58 kDa cytokeratin)	2
17	EF1A1	Elongation factor 1-alpha 1 (EF-1-alpha-1) (Elongation factor 1 A-1) (eEF1A-1) (Elongation factor Tu) (EF-Tu) (Leukocyte receptor cluster member 7)	3
18	PGRC1	Membrane-associated progesterone receptor component 1 (mPR)	2
19	K22E	Keratin, type II cytoskeletal 2 epidermal (Cytokeratin-2e) (K2e) (CK 2e)	1
20	ENL	Protein ENL (YEATS domain-containing protein 1)	3
21	DCD	Dermcidin precursor (Preproteolysin) [Contains: Survival-promoting peptide; DCD-1]	1

22	YBOX1	Nuclease sensitive element-binding protein 1 (Y-box-binding protein 1) (Y-box transcription factor) (YB-1) (CCAAT-binding transcription factor I subunit A) (CBF-A) (Enhancer factor I subunit A) (EFI-A) (DNA-binding protein B) (DBPB) (	2
23	HORN	Hornerin	4
24	KV401	Ig kappa chain V-IV region precursor (Fragment)	1
25	CH60	60 kDa heat shock protein, mitochondrial precursor (Hsp60) (60 kDa chaperonin) (CPN60) (Heat shock protein 60) (HSP-60) (Mitochondrial matrix protein P1) (P60 lymphocyte protein) (HuCHA60)	2
26	IGHG2	Ig gamma-2 chain C region	1
27	TCPD	T-complex protein 1 subunit delta (TCP-1-delta) (CCT-delta) (Stimulator of TAR RNA-binding)	2
28	CALX	Calnexin precursor (Major histocompatibility complex class I antigen-binding protein p88) (p90) (IP90)	2
29	RS28	40S ribosomal protein S28	1
30	EFTU	Elongation factor Tu, mitochondrial precursor (EF-Tu) (P43)	1
31	TCPA	T-complex protein 1 subunit alpha (TCP-1-alpha) (CCT-alpha)	2
32	TCPG	T-complex protein 1 subunit gamma (TCP-1-gamma) (CCT-gamma) (hTRiC5)	2
33	MLRM	Myosin regulatory light chain 2, nonsarcomeric (Myosin RLC)	1
34	CDK9	Cell division protein kinase 9 (EC 2.7.11.22) (EC 2.7.11.23) (Cyclin-dependent kinase 9) (Serine/threonine-protein kinase PITALRE) (C-2K) (Cell division cycle 2-like protein kinase 4)	1
35	CND3	Condensin complex subunit 3 (Non-SMC condensin I complex subunit G) (Chromosome-associated protein G) (Condensin subunit CAP-G) (hCAP-G) (XCAP-G homolog) (Antigen NY-MEL-3)	1
36	IMB1	Importin beta-1 subunit (Karyopherin beta-1 subunit) (Nuclear factor P97) (Importin 90)	1
37	RL10A	60S ribosomal protein L10a (CSA-19)	1
38	PYC	Pyruvate carboxylase, mitochondrial precursor (EC 6.4.1.1) (Pyruvic carboxylase) (PCB)	2
39	HS90A	Heat shock protein HSP 90-alpha (HSP 86) (Renal carcinoma antigen NY-REN-38)	1

40	UBIQ	Ubiquitin	1
41	PSMD3	26S proteasome non-ATPase regulatory subunit 3 (26S proteasome regulatory subunit S3) (Proteasome subunit p58)	1
42	RCN2	Reticulocalbin-2 precursor (Calcium-binding protein ERC-55) (E6-binding protein) (E6BP)	1
43	PERQ2	PERQ amino acid-rich with GYF domain-containing protein 2 (Grb10-interacting GYF protein 2) (Trinucleotide repeat-containing protein 15)	1
44	TCPH	T-complex protein 1 subunit eta (TCP-1-eta) (CCT-eta) (HIV-1 Nef-interacting protein)	1
45	RS10	40S ribosomal protein S10	1
46	MYH10	Myosin-10 (Myosin heavy chain 10) (Myosin heavy chain, nonmuscle IIb) (Nonmuscle myosin heavy chain IIb) (NMMHC II-b) (NMMHC-IIB) (Cellular myosin heavy chain, type B) (Nonmuscle myosin heavy chain-B) (NMMHC-B)	1
47	SRP68	Signal recognition particle 68 kDa protein (SRP68)	2
48	WDR7	WD repeat protein 7 (TGF-beta resistance-associated protein TRAG) (Rabconnectin-3 beta)	1
49	RL30	60S ribosomal protein L30	1
50	ENAM	Enamelin precursor	1
51	SMC4	Structural maintenance of chromosomes protein 4 (Chromosome-associated polypeptide C) (hCAP-C) (XCAP-C homolog)	1
52	ANM1	Protein arginine N-methyltransferase 1 (EC 2.1.1.-) (Interferon receptor 1-bound protein 4)	1
53	LAT1	Large neutral amino acids transporter small subunit 1 (L-type amino acid transporter 1) (4F2 light chain) (4F2 LC) (4F2LC) (CD98 light chain) (Integral membrane protein E16) (hLAT1)	1
54	RS4X	40S ribosomal protein S4, X isoform (Single copy abundant mRNA protein) (SCR10)	1
55	LAMP2	Lysosome-associated membrane glycoprotein 2 precursor (LAMP-2) (CD107b antigen)	1
56	H1X	Histone H1x	1
57	GRP75	Stress-70 protein, mitochondrial precursor (75 kDa glucose-regulated protein) (GRP 75) (Peptide-binding protein 74) (PBP74) (Mortalin) (MOT)	1

58	RS20	40S ribosomal protein S20	1
59	ANM5	Protein arginine N-methyltransferase 5 (EC 2.1.1.125) (EC 2.1.1.-) (Shk1 kinase-binding protein 1 homolog) (SKB1Hs) (Jak-binding protein 1) (72 kDa ICln-binding protein)	1
60	FLNB	Filamin-B (FLN-B) (Beta-filamin) (Actin-binding-like protein) (Thyroid autoantigen) (Truncated actin-binding protein) (Truncated ABP) (ABP-280 homolog) (ABP-278) (Filamin 3) (Filamin homolog 1) (Fh1)	1
61	NPM	Nucleophosmin (NPM) (Nucleolar phosphoprotein B23) (Numatrin) (Nucleolar protein NO38)	1
62	ZN516	Zinc finger protein 516	1
63	GP175	Integral membrane protein GPR175	1
64	RBBP4	Histone-binding protein RBBP4 (Retinoblastoma-binding protein 4) (RBBP-4) (Retinoblastoma-binding protein p48) (Chromatin assembly factor 1 subunit C) (CAF-1 subunit C) (Chromatin assembly factor I p48 subunit) (CAF-I 48 kDa subunit) (CAF-I p48) (Nu	1
65	SUHW3	Suppressor of hairy wing homolog 3	1
66	SYDC	Aspartyl-tRNA synthetase, cytoplasmic (EC 6.1.1.12) (Aspartate--tRNA ligase) (AspRS) (Cell proliferation-inducing gene 40 protein)	1
67	CARD9	Caspase recruitment domain-containing protein 9 (hCARD9)	1
68	IQEC2	IQ motif and Sec7 domain-containing protein 2	1
69	ZEP2	Human immunodeficiency virus type I enhancer-binding protein 2 (HIV-EP2) (MHC-binding protein 2) (MBP-2)	1
71	ESPL1	Separin (EC 3.4.22.49) (Separase) (Caspase-like protein ESPL1) (Extra spindle poles-like 1 protein)	1
72	INOC1	Putative DNA helicase INO80 complex homolog 1 (EC 3.6.1.-) (hINO80)	1
73	SO1B3	Solute carrier organic anion transporter family member 1B3 (Solute carrier family 21 member 8) (Organic anion transporter 8) (Organic anion-transporting polypeptide 8) (OATP8) (Liver-specific organic anion transporter 2) (LST-2) (Huma	1
74	HNRH1	Heterogeneous nuclear ribonucleoprotein H (hnRNP H)	1
75	TMM33	Transmembrane protein 33 (DB83 protein)	1

Supplementry table 4 D (S4D) **Negative control**

Protein hit number	Protein accession number	Protein description	peptides
1	PCCA	Propionyl-CoA carboxylase alpha chain, mitochondrial precursor (EC 6.4.1.3) (PCCase subunit alpha) (Propanoyl-CoA:carbon dioxide ligase subunit alpha)	24
2	PYC	Pyruvate carboxylase, mitochondrial precursor (EC 6.4.1.1) (Pyruvic carboxylase) (PCB)	29
3	PCCB	Propionyl-CoA carboxylase beta chain, mitochondrial precursor (EC 6.4.1.3) (PCCase subunit beta) (Propanoyl-CoA:carbon dioxide ligase subunit beta)	20
4	MCCC2	Methylcrotonoyl-CoA carboxylase beta chain, mitochondrial precursor (EC 6.4.1.4) (3-Methylcrotonyl-CoA carboxylase 2) (MCCase subunit beta) (3-methylcrotonyl-CoA:carbon dioxide ligase subunit beta) (3-Methylcrotonyl-CoA carboxylase non-biotin-contai	11
5	K2C1	Keratin, type II cytoskeletal 1 (Cytokeratin-1) (CK-1) (Keratin-1) (K1) (67 kDa cytokeratin) (Hair alpha protein)	8
6	HSP71	Heat shock 70 kDa protein 1 (HSP70.1) (HSP70-1/HSP70-2)	10
7	GRP78	78 kDa glucose-regulated protein precursor (GRP 78) (Heat shock 70 kDa protein 5) (Immunoglobulin heavy chain-binding protein) (BiP) (Endoplasmic reticulum lumenal Ca(2+)-binding protein grp78)	7
8	ATX2L	Ataxin-2-like protein (Ataxin-2 domain protein) (Ataxin-2-related protein)	4
9	HSP7C	Heat shock cognate 71 kDa protein (Heat shock 70 kDa protein 8)	3
10	ACTB	Actin, cytoplasmic 1 (Beta-actin)	5
11	GRP75	Stress-70 protein, mitochondrial precursor (75 kDa glucose-regulated protein) (GRP 75) (Peptide-binding protein 74) (PBP74) (Mortalin) (MOT)	5
12	K1C9	Keratin, type I cytoskeletal 9 (Cytokeratin-9) (CK-9) (Keratin-9) (K9)	2
13	MCCA	Methylcrotonoyl-CoA carboxylase subunit alpha, mitochondrial precursor (EC 6.4.1.4) (3-methylcrotonyl-CoA carboxylase 1) (MCCase subunit alpha) (3-methylcrotonyl-CoA:carbon dioxide ligase subunit alpha) (3-methylcrotonyl-CoA carboxylase biotin-conta	4
14	GCSH	Glycine cleavage system H protein, mitochondrial precursor	1
15	EF1G	Elongation factor 1-gamma (EF-1-gamma) (eEF-1B gamma)	2

16	PIMT	Protein-L-isoaspartate(D-aspartate) O-methyltransferase (EC 2.1.1.77) (Protein-beta-aspartate methyltransferase) (PIMT) (Protein L-isoaspartyl/D-aspartyl methyltransferase) (L-isoaspartyl protein carboxyl methyltransferase)	2
17	KNG1	Kininogen-1 precursor (Alpha-2-thiol proteinase inhibitor) [Contains: Kininogen-1 heavy chain; Bradykinin (Kallidin I); Lysyl-bradykinin (Kallidin II); Kininogen-1 light chain; Low molecular weight growth-promoting factor]	1
18	UBP35	Ubiquitin carboxyl-terminal hydrolase 35 (EC 3.1.2.15) (Ubiquitin thioesterase 35) (Ubiquitin-specific-processing protease 35) (Deubiquitinating enzyme 35)	1
19	CORT	Cortistatin precursor [Contains: Cortistatin-29; Cortistatin-17]	1
20	K2C6A	Keratin, type II cytoskeletal 6A (Cytokeratin-6A) (CK 6A) (K6a keratin)	1

Suplementry table S 5 (S5A) shows the identified proteins and number of peptides assigned to each protein from PM with their respective Protein accession number. Membrane proteins are indicated by bold typeface.

Protein accession number	Protein identifeind	No. of peptide
gi|10581760	dipeptide ABC transporter dipeptide-binding	101
gi|50979297	pancreatic elastase I [Sus scrofa]	64
gi|10580369	halocyanin precursor-like	14
gi|10580964	**bacteriorhodopsin**	81
gi|10581613	**halocyanin precursor-like**	4
gi|10584350	**Na+/H+ antiporter**	10
gi|10581921	dipeptide ABC transporter ATP-binding	75
gi|10579949	**bifunctional short chain isoprenyl diphosphate synthase**	4
gi|10581945	**conserved hypothetical protein**	11
gi|16554513	**cytochrome c oxidase subunit I**	14
gi|10580181	**conserved hypothetical protein**	4
gi|10580237	**NADH dehydrogenase/oxidoreductase**	6
gi|10580777	**hypothetical protein**	15
gi|10581558	**H+-transporting ATP synthase subunit K**	5
gi|10580180	**hypothetical protein**	6
gi|10581468	hypothetical protein	1
gi|10584336	Vng6297c	12
gi|10580235	**F420H2:quinone oxidoreductase chain L**	2
gi|10580236	**F420H2:quinone oxidoreductase chain M**	5
gi|10580721	membrane protein	3
gi|10581015	**daunorubicin resistance ABC transporter ATP-binding protein**	3
gi|10581732	**heterodisulfide reductase**	3
gi|10582055	**cell surface glycoprotein**	16

gi|10580483	iron-binding protein	12
gi|10581758	**dipeptide ABC transporter permease**	3
gi|10579770	**ABC transport protein**	1
gi|10580230	**NADH dehydrogenase/oxidoreductase**	4
gi|10581182	**protein translocase**	4
gi|10580113	**hypothetical protein**	1
gi|10580234	NADH dehydrogenase/oxidoreductase-like protein	1
gi|10581612	cytochrome c oxidase subunit II	6
gi|10580168	conserved hypothetical protein	40
gi|10584307	adhesion protein	43
gi|10580139	immunogenic protein	44
gi|10579932	**lipoate protein ligase**	10
gi|16554486	metallo-beta-lactamase superfamily hydrolase	5
gi|10581419	**protein-export membrane protein**	7
gi|10581255	hypothetical protein	5
gi|10580177	**conserved hypothetical protein**	5
gi|10580062	phosphate ABC transporter periplasmic phosphate-binding	6
gi|10580210	hypothetical protein	4
gi|10581692	conserved hypothetical protein	3
gi|10581371	**hypothetical protein**	4
gi|10581882	phosphate ABC transporter binding	6
gi|10581159	50S ribosomal protein L13P	6
gi|10581787	**hypothetical protein**	5
gi|10580017	hypothetical protein	3
gi|10580676	50S ribosomal protein L13P	4
gi|10584333	Vng6296c	4
gi|10581759	**dipeptide ABC transporter permease**	3
gi|10581802	conserved hypothetical protein	3

gi|10581968	**hypothetical protein**	3
gi|10580962	**bacteriorhodopsin related protein**	3
gi|10581997	probable carboxypeptidase	4
gi|10580022	conserved hypothetical protein	2
gi|10580178	**cytochrome b6**	3
gi|10580246	**cytochrome c oxidase subunit I**	3
gi|10581140	30S ribosomal protein S8E	3
gi|10581420	**protein-export membrane protein**	2
gi|10580517	**flagellin B1 precursor**	3
gi|10584332	peroxidase / catalase	2
gi|10580211	**proteinase IV homolog**	2
gi|16554512	hypothetical protein VNG2166Cm	1
gi|10581401	**hypothetical protein**	1
gi|16554520	nicotinate phosphoribosyltransferase	2
gi|2822317	**CydA**	2
gi|10580518	flagellin B2 precursor	2
gi|10580564	**flagellin A1 precursor**	2
gi|10580175	**hypothetical protein**	2
gi|10582039	30S ribosomal protein S7P	2
gi|10579826	50S ribosomal protein L15E	2
gi|10581923	**dipeptide ABC transporter permease**	2
gi|10579829	**halorhodopsin**	3
gi|10579959	**hypothetical protein**	2
gi|10581502	possible phosphate binding protein	3
gi|10580541	**hypothetical protein**	1
gi|10581366	pyruvate dehydrogenase alpha subunit	1
gi|10581215	**Htr5 transducer**	2
gi|10580278	**conserved hypothetical protein**	1
gi|10581131	**Htr1 transducer**	2
gi|161349984	**putative monovalent cation/H+ antiporter**	1

	subunit D	
gi|10581818	**hypothetical protein**	2
gi|10580251	**cytochrome c oxidase subunit II**	3
gi|10580565	flagellin A2 precursor	2
gi|10580828	**succinate dehydrogenase subunit C**	3
gi|10581411	hypothetical protein	1
gi|10581208	photolyase/cryptochrome	1
gi|10579815	proteasome, subunit beta	1
gi|10580617	**transmembrane oligosaccharyl transferase**	1
gi|10581350	**hypothetical protein**	2
gi|10580166	conserved hypothetical protein	2
gi|10584311	**Vng6268c**	3
gi|10580667	putative 2-ketoglutarate ferredoxin oxidoreductase (beta)	2
gi|10579748	50S ribosomal protein L10E	3
gi|10580916	hypothetical protein	1
gi|10581815	GTP-binding protein homolog	1
gi|10580374	**conserved hypothetical protein**	1
gi|10581307	**cation efflux system protein (zinc/cadmium)**	1
gi|10581785	copper transport ATP-binding protein	2
gi|10581174	30S ribosomal protein S8P	1
gi|10580182	**hypothetical protein**	1
gi|10581298	**hypothetical protein**	1
gi|10580755	hypothetical protein	1
gi|10581693	conserved hypothetical protein	1
gi|10580262	conserved hypothetical protein	1
gi|10581447	**conserved hypothetical protein**	1
gi|10579799	regulatory protein	1
gi|10580172	**hypothetical protein**	1
gi|10581680	**protein export**	1

gi|10580137	**conserved hypothetical protein**	1
gi|10584298	**Vng6251h**	1
gi|10580337	**hypothetical protein**	1
gi|10580060	**phosphate ABC transporter permease**	2
gi|10580277	**conserved hypothetical protein**	1
gi|10580156	**conserved hypothetical protein**	1
gi|10581800	probable thiosulfate sulfurtransferase	1
gi|10580519	flagellin B3 precursor	1
gi|10579654	glutamine-fructose-6-phosphate transaminase	1
gi|10580668	putative 2-ketoglutarate ferredoxin oxidoreductase (alpha)	1
gi|10581224	translation initiation factor eIF-5A	1
gi|10580553	probable oxidoreductase	1
gi|10581433	oligopeptidase	1
gi|10580424	**conserved hypothetical protein**	2
gi|10580248	**hypothetical protein**	1
gi|10579777	**heat shock protein**	1
gi|10580874	**conserved hypothetical protein**	1
gi|10582040	30S ribosomal protein S12P	1
gi|10581312	**spermidine/putrescine ABC transporter permease**	1
gi|10581108	halocyanin precursor-like	1
gi|10584354	transcription regulator	1
gi|10581486	**conserved hypothetical protein**	1
gi|10580917	hypothetical protein	1
gi|10581181	50S ribosomal protein L15P	1
gi|10581199	Htr-like protein	1
gi|10580365	**Htr6 transducer**	1
gi|10580990	50S ribosomal protein L37E	1
gi|10580748	glucose-6-phosphate isomerase	1

gi|10580872	**hypothetical protein**	1
gi|10581417	**conserved hypothetical protein**	1
gi|10584305	**ABC transporter, permease protein**	1
gi|10580459	**ribose ABC transporter permease**	1
gi|10579905	conserved hypothetical protein	1
gi|10580439	**possible signaling protein**	1
gi|10581507	**hypothetical protein**	1
gi|10580225	**NADH dehydrogenase/oxidoreductase-like protein**	1
gi|10581107	**cobalamin biosynthesis protein**	1
gi|10580692	small heat shock protein	1

Suplementry table S 5B (S5B) shows the identified proteins and number of peptides assigned to each protein from CGM with their respective Protein accession number Membrane proteins are indicated by bold typeface.

Protein hit number	Protein accession number	Protein description	No. of peptides
1	Cg0446	succinate dehydrogenase A	38
2	Cg2181	ABC-type peptide transport system, secreted component	21
3	Cg2911	ABC-type Mn/Zn transport system, secreted Mn/Zn-binding (lipo)protein	20
4	P00772	Elastase-1	14
5	**Cg2404**	**RIESKE IRON-SULFUR PROTEIN**	**18**
6	Cg2705	MALTOSE-BINDING PROTEIN PRECURSOR	**17**
7	Cg2137	GLUTAMATE SECRETED BINDING PROTEIN component/domain	**16**
8	**Cg2195**	**putative secreted or membrane protein**	**10**
9	Cg0737	ABC-type transport system, secreted lipoprotein component	12
10	**Cg1656**	**NADH DEHYDROGENASE**	**11**
11	**Cg1537**	**GLUCOSE-SPECIFIC ENZYME II BC COMPONENT OF PTS**	**7**
12	Cg0935	conserved hypothetical protein	8
13	Cg0040	PUTATIVE SECRETED PROTEIN	**7**
14	Cg3186	Trehalose corynomycolyl transferase	**8**
15	Cg0447	succinate dehydrogenase B	9
16	**Cg2409**	**CYTOCHROME C OXIDASE CHAIN II**	**11**
17	**Cg2403**	**CYTOCHROME B, MEMBRANE PROTEIN**	**7**
18	Cg0413	Trehalose corynomycolyl transferase	9
19	**Cg2211**	**putative membrane protein**	**4**
20	Cg2708	ABC-type sugar transport system, ATPase component	**9**
21	Cg0834	Bacterial extracellular solute-binding protein, fa	5
22	**Cg2780**	**PROBABLE CYTOCHROME C OXIDASE POLYPEPTIDE SUBUNIT**	**8**

23	Cg1787	PHOSPHOENOLPYRUVATE CARBOXYLASE	7
24	**Cg2657**	**putative membrane protein - fragment**	3
25	Cg3182	Trehalose corynomycolyl transferase	3
26	Cg2151	Similar to phage shock protein A	6
27	Cg1085	hypothetical protein predicted by Glimmer/Critica	5
28	**Cg0953**	**Na+/proline, Na+/panthothenate symporter or related permease**	**4**
29	Cg2429	GLUTAMINE SYNTHETASE I	5
30	**Cg2925**	**ENZYME II SUCROSE PROTEIN**	**5**
31	Cg1081	ABC-type multidrug transport system, ATPase component	6
32	Cg2843	ABC-type phosphate transport system, ATPase component	2
33	Cg1790	PHOSPHOGLYCERATE KINASE	2
34	Cg2949	putative secreted protein	2
35	Cg1332	PUTATIVE SECRETED HYDROLASE	2
36	**Cg1001**	**LARGE CONDUCTANCE MECHANOSENSITIVE CHANNEL**	**2**
37	**Cg3138**	**Membrane protease subunit, stomatin/prohibitin homolog**	**4**
38	Cg0753	secreted protein	**3**
39	Cg2782	Ferritin-like protein	3
40	Cg3257	conserved hypothetical protein	3
41	Cg1290	Homocysteine methyltransferase	3
42	Cg0654	RIBOSOMAL PROTEIN S4	2
43	Cg2499	GLYCYL-TRNA SYNTHETASE (GLYCINE--TRNA LIGASE)	2
44	Cg2419	LEUCINE AMINOPEPTIDASE	2
45	Cg3192	putative secreted or membrane protein	1
46	Cg0655	DNA-DIRECTED RNA POLYMERASE ALPHA SUBUNIT	2
47	Cg3195	Flavin-containing monooxygenase (FMO)	2
48	Cg1111	ENOLASE	2
49	**Cg1363**	**ATP SYNTHASE C CHAIN**	**2**
50	Cg2184	ATPase component of peptide ABC-type transport system, contains duplicated ATPase domains	1
51	Cg2912	ABC-type cobalamin/Fe3+-siderophores transport system, ATPase	1

		component	
52	**Cg3395**	**PROLINE/ECTOINE CARRIER**	**4**
53	Cg1109	Anion-specific porin precursor	2
54	Cg2325	hypothetical protein predicted by Glimmer/Critica	2
55	**Cg1796**	**putative membrane protein - C. ammoniagenes RibX homolog**	**2**
56	Cg0699	INOSITOL-MONOPHOSPHATE DEHYDROGENASE	1
57	Cg3100	Heat shock protein hsp70	1
58	Cg0350	TRANSCRIPTIONAL REGULATOR, CRP/FNR FAMILY	1
59	Cg1836	secreted solute-binding protein, aminodeoxychorismate lyase-like	2
60	Cg0577	DNA-DIRECTED RNA POLYMERASE BETA~ CHAIN	2
61	Cg2052	putative secreted protein	1
62	**Cg1227**	**putative membrane protein**	**1**
63	Cg1604	putative secreted protein	1
64	Cg1794	Uncharacterised P-loop ATPase protein	1
65	Cg1713	DIHYDROOROTATE DEHYDROGENASE	1
66	**Cg2604**	**putative secreted or membrane protein**	**1**
67	Cg2492	PROBABLE GLUCOSAMINE--FRUCTOSE-6-PHOSPHATE AMINOTRANSFERASE	1
68	Cg1108	putative secreted protein	1
69	Cg1052	corynomycolyl transferase	1
70	Cg1366	PROBABLE ATP SYNTHASE ALPHA CHAIN PROTEIN	2
71	**Cg2704**	**ABC-type sugar transport system, permease component**	**1**
72	Cg0938	Cold shock protein	1
73	**Cg2405**	**CYTOCHROME C1**	**1**
74	**Cg3343**	**putative secreted membrane protein**	**2**
75	**Cg3008**	**PORIN**	**1**
76	Cg0441	DIHYDROLIPOAMIDE DEHYDROGENASE	1
77	Cg1654	PHOSPHOMETHYLPYRIMIDINE KINASE / HYDROXYMETHYLPYRI	2
78	Cg3149	Aminotransferases class-I	2

79	**Cg2183**	**ABC-type peptide transport system, permease component**	**1**
80	**Cg2460**	**putative membrane protein**	**1**
81	Cg2275	putative F0F1-type ATP synthase b subunit	1
82	**Cg0811**	**ACETYL/PROPIONYL COA CARBOXYLASE, BETA SUBUNIT**	**1**
83	Cg2136	GLUTAMATE UPTAKE SYSTEM ATP-BINDING PROTEIN	1
84	Cg0336	PENICILLIN-BINDING PROTEIN 1B	1
85	**Cg3252**	**putative preprotein translocase subunit YidC, SpoIIIJ homolog**	**1**
86	Cg1286	conserved hypothetical protein	1
87	Cg2466	PYRUVATE DEHYDROGENASE E1 COMPONENT	1
88	**Cg0464**	**COPPER-TRANSPORTING ATPASE**	**1**
89	Cg2963	PROBABLE ATP-DEPENDENT PROTEASE (HEAT SHOCK PROTEIN)	1
90	Cg0625	secreted protein	1
91	Cg1368	ATP SYNTHASE ALPHA SUBUNIT	2
92	**Cg0448**	**conserved hypothetical membrane protein**	**1**
93	**Cg0896**	**membrane protein**	**1**
94	Cg0835	ABC-type sugar transport systems, ATPase component	2
95	**Cg0254**	**AMINO ACID CARRIER PROTEIN (SODIUM/ALANINE SYMPORTER)**	**1**
96	Cg1367	ATP SYNTHASE GAMMA SUBUNIT	1
97	Cg1730	secreted protease subunit, stomatin/prohibitin homolog	2
98	Cg1479	PUTATIVE GLYCOGEN PHOSPHORYLASE	2
99	Cg0107	secreted protein	1
100	**Cg0451**	**PUTATIVE MEMBRANE PROTEIN**	**1**
101	**Cg1868**	**Preprotein translocase subunit YajC homolog**	**1**
102	Cg0587	ELONGATION FACTOR TU	1
103	Cg0237	PUTATIVE OXIDOREDUCTASE	1
104	**Cg0161**	**putative secreted or membrane protein**	**1**
105	Cg1046	RNA POLYMERASE SIGMA FACTOR, PUTATIVE	1
106	Cg0410	PUTATIVE PROLYL ENDOPEPTIDASE	1

107	Cg1504	ABC-type amino acid transport system, secreted component	1
108	**Cg3356**	**Na+/H+-dicarboxylate symporter**	**1**
109	**Cg0756**	**PUTATIVE CARBON STARVATION PROTEIN A**	**1**
110	Cg2747	SECRETED PEPTIDASE, M23/M37 FAMILY	1
111	Cg1813	PUTATIVE CARBAMOYL-PHOSPHATE SYNTHASE SUBUNIT	1
112	**Cg1229**	**ABC-type cobalt transport system, permease component CbiQ**	**1**
113	Cg2450	putative pyridoxine biosynthesis enzyme	1
114	Cg0564	50S RIBOSOMAL PROTEIN L1	1
115	Cg0603	50S RIBOSOMAL PROTEIN L29	1
116	**Cg3216**	**GLUCONATE PERMEASE**	**1**
117	Cg1590	secreted Mg-chelatase subunit	1
118	Cg0980	secreted protein related to metalloendopeptidases	1
119	**Cg2120**	**SUGAR SPECIFIC PTS SYSTEM, FRUCTOSE/MANNITOL-SPECIFIC TRANSPORT PROTEIN**	**1**
120	Cg2444	hypothetical protein predicted by Glimmer/Critica	1
121	Cg4005	putative secreted protein	1
122	**Cg0359**	**PUTATIVE MEMBRANE PROTEIN**	**1**
123	Cg2291	PYRUVATE KINASE	1
124	Cg2467	ABC TRANSPORTER ATP-BINDING PROTEIN	1
125	Cg1778	GLUCOSE-6-PHOSPHATE 1-DEHYDROGENASE	1
126	Cg0576	DNA-DIRECTED RNA POLYMERASE BETA CHAIN	1
127	Cg1133	Serine Hydroxymethyltransferase	1
128	Cg2160	HYDROLASE OF METALLO-BETA-LACTAMASE SUPERFAMILY	1
129	**Cg3174**	**exporter of the MMPL family**	**1**
130	Cg0601	30S RIBOSOMAL PROTEIN S3	1
131	Cg2235	50S RIBOSOMAL PROTEIN L19	1
132	Cg2773	Uncharacterized protein with SCP/PR1 domain	1
133	Cg0791	PYRUVATE CARBOXYLASE	1
134	**Cg1169**	**Na+-dependent transporters of the SNF family**	**1**
135	Cg1791	GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE	1

136	Cg0648	ADENYLATE KINASE	1
137	Cg2132	conserved hypothetical protein	1
138	Cg3244	conserved hypothetical protein	1
139	Cg0610	50S RIBOSOMAL PROTEIN L5	1
140	Cg0508	iron/thiamine transport system, secreted component	1
141	**Cg0524**	**Cytochrome c assembly membrane protein**	**1**
143	**Cg1238**	**putative membrane protein**	**1**
144	Cg0445	succinate dehydrogenase CD	1
145	Cg0518	GLUTAMATE-1-SEMIALDEHYDE 2,1-AMINOMUTASE	1
146	Cg1577	putative secreted hydrolase	1
147	Cg1672	polyprenol-phosphate-mannose synthase domain 1	1
148	Cg0783	conserved hypothetical protein	1
149	Cg2213	ABC-type multidrug transport system, ATPase component	1
150	**Cg1277**	**conserved hypothetical membrane protein**	**1**

Supplementry table S5 C (S5C) summary of the fractions, peptides sequence, calculated mass (pep_calc_mr), mass acuuracy (pep_delta), peptide score (pep_score), homology (pep_homol), identity threshold (pep_ident), the rank number (pep_Rank), hydrophobicity score (GRAVY_score) and isoelectric point (pI_Value) for PM and CGM.

Fraction	Peptide sequence	pep_calc_mr	pep_delta	pep_score	pep_homol	pep_ident	pep_rank	GRAVY_Score	pI_Value
1	HSIAAQGGVNSA	1110.5418	0.0024	71.29	33	35	1	0.125	7.84
1	SVDGWFTLPFTIPNYLGPLLGSERLSEDAPEAQ	3618.7882	-0.0069	94.94	20	39	1	-0.1727	3.5
1	SFEDELSGRSINQPSAV	1834.8697	-0.0186	45.19	28	35	1	-0.5941	3.82
1	SDSLQGADVLTYDPERA	1835.8537	-0.0054	57.85	31	35	1	-0.7176	3.5
1	FSSFEDELSGRSINQPS	1898.8646	-0.0003	55.22	30	34	1	-0.8294	3.82
1	FSSFEDELSGRSINQPSAV	2068.9702	-0.0162	104.39		35	1	-0.4263	3.82
1	DDTTFTVELTQPESDFPLRL	2323.122	-0.0092	55.62	23	37	1	-0.555	3.42
1	NGTEPQNPLVPGNTNEVGGGRIVDSI	2633.3045	-0.0537	41.05	18	37	1	-0.6192	3.82
1	SDMVESTPEGYRATTL	1755.7985	-0.0108	41.07	28	34	1	-0.6437	3.82
1	VAPDVDIEAIITGGDIDPH	1945.9633	-0.016	113.08		37	1	0.2421	3.36
1	DHDVDNEHVWYSTEYV	2006.8283	0.0031	32.16	19	28	1	-1.2375	3.84
1	NTENFLDAFTKAVDDLTAAT	2156.0273	0.0054	119.41		37	1	-0.155	3.5
1	NTENFLDAFTKAVDDLTAATN	2270.0703	0.004	39.45	29	36	1	-0.3143	3.5
1	GGGGYDSWLYGTLEDDDRIIH	2338.0502	-0.0167	109.18	26	33	1	-0.7428	3.76
1	GGGGYDSWLYGTLEDDDRIIHAL	2522.1714	-0.0002	118.47	32	36	1	-0.4347	3.76
1	VVHPYWNTDDVAAGYDIA	2004.9218	-0.0273	75.38	26	33	1	-0.05	3.6
1	KQLTDDIPPLTTEDGEVRGVPF	2426.2329	0.0071	65.18	24	38	1	-0.6272	3.84
1	TDDIPPLTTEDGEVRGVPFAVEGF	2560.2334	-0.0226	41.47	24	36	1	-0.2208	3.38
1	NAPIWQEYQDKGVEDTNEIEFS	2611.1714	0.0144	48.86	24	34	1	-1.2181	3.44
1	KQLTDDIPPLTTEDGEVRGVPFAVEGF	2929.4709	-0.0573	104.04	33	37	1	-0.3296	3.76
1	YEGDFRVVEMEKDGEPF	2045.9041	-0.0104	41.56	24	32	1	-1.0117	3.82
1	EYVVNSIADDKGWDHPTIEWRESPS	2929.3519	-0.0127	79.9	25	35	1	-1.068	4.06
1	EYVVNSIADDKGWDHPTIEWRESPSAQ	3128.4475	-0.0083	99.29	31	35	1	-1.0518	4.06
1	TGEDIFGEEHERDNENTPA	2158.9039	0.0034	33.32	19	29	1	-1.7842	3.74
1	GEDWDADQPVTGEDIFGEEH	2244.9084	0.0202	55.64	20	27	1	-1.29	3.24

1	EDIFGEEHERDNENTPAWA	2257.9512	0.0071	33.16	23	29	1	-1.6789	3.74
1	TGEDIFGEEHERDNENTPAWA	2416.0203	-0.0061	79.32	23	28	1	-1.5714	3.74
1	DQPVTGEDIFGEEHERDNENTPA	2598.1106	0.0193	38.86	17	31	1	-1.6655	3.62
1	GEDWDADQPVTGEDIFGEEHERDNENTPA	3271.345	0.012	153.74		26	1	-1.6655	3.42
1	VGEDWDADQPVTGEDIFGEEHERDNENTPA	3370.4134	-0.0081	141.52	23	26	1	-1.47	3.42
1	AVGEDWDADQPVTGEDIFGEEHERDNENTPA	3441.4505	-0.0017	114.39	20	27	1	-1.3645	3.42
1	AAVGEDWDADQPVTGEDIFGEEHERDNENTPA	3512.4876	0.0007	93.92	25	27	1	-1.2656	3.42
1	DGIDGESVAIPNDPSNQGRAI	2124.0083	0.0117	48.8	33	37	1	-0.6523	3.5
1	DRPGADLQEILDRLEADQENA	2367.1302	-0.0223	39.35	22	36	1	-1.1761	3.62
1	SVPVDRPGADLQEILDRLEADQENA	2749.3518	-0.01	60.47	28	38	1	-0.748	3.62
1	RFQVKDQSIVDQQEIDSDPS	2333.1135	0.0132	84.12	35	37	1	-1.185	3.76
1	QAPDSYGRIGNQSGSDVSPV	2032.945	0	39.45	24	35	1	-0.81	3.88
1	NTYPVIDAWTEPEVWQEARPT	2501.1863	-0.0186	153.72		36	1	-0.9047	3.68
1	DFLRMGWPDGITPE	1632.7606	-0.0266	40.82	28	34	1	-0.5428	3.7
1	DVSYLEFNRIETLGTTDEIPV	2410.1904	-0.0295	41.67	28	37	1	-0.1857	3.5
1	VEPDWTYDQATLTGTWGNLTDSQLRSV	3052.4414	-0.0385	144.76		35	1	-0.6666	3.5
1	GFRPEALEIIPEGESTDLSIPI	2382.2318	-0.0133	70.22		38	1	0.0136	3.58
1	SESYEDVLTELFERAQVTA	2186.0379	-0.0264	37.96	31	35	1	-0.3894	3.58
1	TAEQSVEDTKQLHDQALQQADQA	2553.1943	0.0247	55.34	32	37	1	-1.2521	3.84
1	EFRAPDPSGNPYLGFAA	1807.8529	-0.0122	40.18	29	35	1	-0.4647	4.08
1	FENPDQKETEDYISGRFG	2130.9494	-0.0097	103.44	27	32	1	-1.5555	3.92
1	DATVSAPVLESGVENREVGGDTA	2272.0819	-0.0077	78.39	31	36	1	-0.2304	3.5
1	DATVSAPVLESGVENREVGGDTAEA	2472.1616	-0.0024	77.28	28	36	1	-0.28	3.44
1	SFSVEDFEEFLHQQ	1740.7631	-0.0172	31.93	23	31	1	-0.6714	3.68
1	ALGFLDYLDEEQRA	1638.7889	-0.0163	38.14	27	36	1	-0.4214	3.58
1	FFDQETNERWIPFVIEPA	2237.0793	-0.0275	53.54	33	36	1	-0.3833	3.68
1	EDFGWRIPFLT	1379.6874	-0.0021	49.73	33	36	1	-0.1090	4.08
1	GVLEEGQYNRELAEAI	1789.8846	-0.034	55.41		36	1	-0.4812	3.8
1	AGVLEEGQYNRELAEAI	1860.9217	-0.0331	64.33	31	36	1	-0.3470	3.8
1	RLELPDFEFVPM	1491.7432	-0.0119	55.51	27	37	1	0.0916	3.82

1	FAEFERTTAEDVANGVPEAI	2165.0277	-0.0112	76.2	25	36	1	-0.16	3.58
1	DDSLSQEVFGVTFPRPL	1905.9473	-0.0302	78.43		37	1	-0.1647	3.7
1	AELFDASGQIPSSQEFFRV	2127.0273	-0.0246	56.88	24	36	1	-0.1	3.82
1	SASDDSDALDWVRELTSYLPS	2326.0601	-0.0039	56.27		34	1	-0.5380	3.36
1	LRDLLEGTDERIQDLEFFQDAV	2621.2973	-0.0053	41.65	26	38	1	-0.4681	3.62
1	SLDAVRQEVEEIIGQGSQPT	2155.0757	-0.0322	58.01		37	1	-0.535	3.68
1	AFQDALGVRDADFIAQLV	1948.0054	-0.0295	56.89	33	37	1	0.6055	3.6
1	LFDQPEDDQPNDGFDGRT	2064.8661	0.0233	50.21	27	31	1	-1.7388	3.26
1	INDEVINEIRDGISSF	1819.8952	-0.006	37.43	31	37	1	-0.1562	3.58
1	DWADDDSMNSYFLANPRMQDYI	2666.1053	-0.0154	37.26		25	1	-0.9409	3.3
2	KYPYGIDET	1084.5077	0.0031	50.97	33	36	1	-1.3	4.08
2	GYGPDNKYELN	1268.5673	0.0013	43.39	30	36	1	-1.7363	4.08
2	SLLETTKQGEHEAF	1588.7733	0.0102	41.29	24	39	1	-0.821	4.54
2	MAYVINREQFVSDV	1669.8134	-0.0003	42.55	27	38	1	0.2142	4.08
2	RKDAATNMEEGIWES	1735.7835	0.0025	127.83	36	37	1	-1.2466	4.16
2	SLLETTKQGEHEAFTL	1802.905	0.0019	101.43	35	39	1	-0.525	4.54
2	RKDAATNMEEGIWESA	1806.8206	0.0022	124.52		37	1	-1.0562	4.16
2	RKDAATNMEEGIWESA	1822.8155	-0.0024	41.51	22	36	1	-1.05	4.16
2	FHLQPPQLFSDGQEGY	1861.8635	-0.0406	102.03	29	36	1	-0.85	4.06
2	KVMEEAGYGPDNKYELN	1955.8934	0.0059	56.03	26	37	1	-1.3	4.16
2	RLEQAHIEMTINDANFS	1987.9421	-0.009	72.76	30	38	1	-0.5352	4.42
2	QTRKDAATNMEEGIWESA	2035.9269	0.0057	48.63	18	38	1	-1.1722	4.16
2	NSDPGVREFMTEQYERLV	2169.0161	-0.016	40.69	20	38	1	-0.9944	4.16
2	RKDAATNMEEGIWESAALIPV	2300.147	-0.0261	163.74	34	39	1	-0.2	4.16
2	EARKVMEEAGYGPDNKYELN	2312.0742	-0.0026	129.93	32	38	1	-1.415	4.36
2	KVMDLFDAPTNYENGRTEIT	2313.0947	-0.0104	118.67	30	38	1	-0.79	4.06
2	HLQPPQLFSDGQEGYNNHYQGE	2557.1258	0.0109	49.07	16	36	1	-1.6045	4.4
2	RLYWKEDANSDPGVREFMTEQ	2570.186	-0.0057	41.02	29	38	1	-1.3619	4.28
2	KVMDLFDAPTNYENGRTEITNLLIT	2867.4375	-0.0605	39.62	21	39	1	-0.316	4.06

2	GYGPDNKYELNWLQYQSPTWKEMANTI	3245.5128	-0.0188	134.24	31	38	1	-1.1814	4.44
2	STLGIEHEVDENGAVKRATGVEAVDDYTVRL	3342.6692	0.0308	56.64	19	40	1	-0.3548	4.22
2	ESMRPEVA	917.4276	0.0063	38.22	31	38	1	-0.75	4.26
2	RAIFGEAEAPEPS	1372.6623	0.0045	58.85	32	38	1	-0.5153	3.96
2	RYADWLFTTPL	1381.703	-0.0071	37.85	27	37	1	-0.0909	6.7
2	RYADWLFTTPLLL	1607.8712	-0.0049	78.48	26	36	1	0.5076	6.7
2	RYADWLFTTPLLLL	1720.9552	-0.0039	91.58	28	35	1	0.7428	6.7
2	PFGGEQNPIYWARYA	1767.8369	-0.0039	52.41	23	38	1	-0.7733	6.82
2	YWARYADWLFTTPL	1801.8828	-0.0097	64.41	30	39	1	-0.1	6.64
2	RYADWLFTTPLLLLD	1835.9822	-0.0076	73.55	24	37	1	0.46	3.88
2	RYADWLFTTPLLLLDLA	2020.1033	-0.0051	123.79		36	1	0.7352	3.88
2	AVEGVSQAQITGRPEWIWL	2139.1113	-0.0153	67.79	29	38	1	0.0315	4.26
2	AVEGVSQAQITGRPEWIWLA	2210.1484	-0.0317	70.37	30	39	1	0.12	4.26
2	GGEQNPIYWARYADWLFTTPL	2497.2066	-0.0169	39.11	22	39	1	-0.4666	4.08
2	GGPFLDRVELA	1172.619	0.0037	69.23		38	1	0.2272	4.08
2	AILSRDQIESL	1243.6772	-0.0031	57.32		37	1	0.16363	4.08
2	AAILSRDQIESL	1314.7143	-0.0005	61.86		37	1	0.3	4.08
2	SVGGPFLDRVELA	1358.7194	0.0029	103.78		38	1	0.4538	4.08
2	YTPTSVGGPFLDRVELA	1820.9309	-0.0118	55.89	34	38	1	0.0941	4.08
2	RAQGLINGDAHITDNLDAQ	2020.9926	-0.0125	47.66	22	39	1	-0.6157	4.16
2	AHDQAPWVYLDYADLIR	2045.0007	-0.0315	78.01	22	39	1	-0.3	4.16
2	TNQIFDTLIQFKPGTSGEL	2108.079	-0.0236	87.01	26	39	1	-0.2368	4.08
2	GPFEFDQLDNETQRVRLT	2164.0549	-0.0041	52.78	25	39	1	-1.0833	4.06
2	AHDQAPWVYLDYADLIRGV	2201.0905	-0.0131	83.58	24	39	1	-0.0684	4.16
2	TNQIFDTLIQFKPGTSGELT	2209.1267	-0.0407	126.1	35	39	1	-0.26	4.08
2	TNQIFDTLIQFKPGTSGELTS	2296.1587	-0.0412	133.61	30	39	1	-0.2857	4.08
2	GTGPFEFDQLDNETQRVRLT	2322.1241	0.0013	89.78	28	39	1	-1.03	4.06
2	TNQIFDTLIQFKPGTSGELTSGL	2466.2642	-0.01	129.24		39	1	-0.1130	4.08
2	AHDQAPWVYLDYADLIRGVNEA	2515.2132	-0.0229	136.45	31	39	1	-0.2954	3.96
2	SQIAHDQAPWVYLDYADLIRGV	2529.2652	-0.0147	99.59	22	39	1	-0.05	4.16

2	SQIAHDQAPWVYLDYADLIRGVNEA	2843.3878	-0.0203	111.99	26	40	1	-0.252	3.96
2	TSGEVSKVTNQIFDTLIQFKPGTSGELT	2996.5343	0.0078	155.26	30	39	1	-0.246428571	4.44
2	TSGEVSKVTNQIFDTLIQFKPGTSGELTS	3083.5663	0.0021	149.95	28	40	1	-0.265517241	4.44
2	TSGEVSKVTNQIFDTLIQFKPGTSGELTSGL	3253.6718	0.0079	145.54	26	39	1	-0.138709677	4.44
2	SAWVDTDYMEFVDKGQRTYDDAARQQYYHQA	3756.6539	-0.0389	49.82		33	1	-1.274193548	4.24
2	VHPYWNTDDV	1244.5462	0.0046	46.2	26	36	1	-0.98	3.88
2	VVHPYWNTDDV	1343.6146	0.0022	75.9	34	37	1	-0.509090909	3.88
2	VVHPYWNTDDVA	1414.6517	0.0012	81.71		37	1	-0.316666667	3.88
2	VVHPYWNTDDVAA	1485.6889	0.0028	93.03	31	37	1	-0.153846154	3.88
2	VVHPYWNTDDVAAG	1542.7103	0.0014	70.29	29	37	1	-0.171428571	3.88
2	KIVVHPYWNTDDVAAGYDIA	2246.1008	0.0047	56.89	20	39	1	-0.015	4.16
2	VGEHNLNQNDGTEQYVGVQKI	2341.1299	-0.0009	118.84	27	39	1	-0.995238095	4.42
2	QKIVVHPYWNTDDVAAGYDIA	2357.1328	0.0655	70.12	22	39	1	-0.180952381	4.16
2	QKIVVHPYWNTDDVAAGYDIA	2374.1594	0.0021	146.56	29	39	1	-0.180952381	4.16
2	GGLGDGSFNDQAKQGLERA	1918.9133	-0.0048	50.06	24	39	1	-0.957894737	4.3

2	GTGGLGDGSFNDQAKQGLERA	2076.9825	-0.0065	131.48	32	39	1	-0.919047619	4.3
2	YGDQLGAEIPQSVKDAVSKS	2091.0484	-0.0022	72.49	27	39	1	-0.555	4.3
2	VYGDQLGAEIPQSVKDAVSKS	2190.1168	0.0024	89.95	30	39	1	-0.328571429	4.3
2	YGTGGLGDGSFNDQAKQGLERA	2240.0458	-0.0045	134.55	29	38	1	-0.936363636	4.3
2	VYGTGGLGDGSFNDQAKQGLERA	2339.1142	0.0033	88.9	28	39	1	-0.713043478	4.3
2	KDAVSKSRTELIAGDIDVPTEP	2340.2173	-0.0067	91.47	31	39	1	-0.527272727	4.2
2	GTVYGTGGLGDGSFNDQAKQGLER	2426.1463	-0.0068	164.99	33	39	1	-0.804166667	4.3
2	TVYGTGGLGDGSFNDQAKQGLERA	2440.1619	0.0007	90.19	23	39	1	-0.7125	4.3
2	GTVYGTGGLGDGSFNDQAKQGLERA	2497.1834	0.0045	194.94	38	39	1	-0.7	4.3
2	AYDIEMRPLVV	1304.6798	0.003	65.89	28	37	1	0.545454545	4.08
2	EWNATWRNFE	1351.5945	0.002	58.02		36	1	-1.64	4.26
2	IFDRADRDVVQL	1445.7627	-0.003	56.86		38	1	-0.141666667	4.18
2	IFDRADRDVVQLA	1516.7998	-0.0005	65.28	37	38	1	0.007692308	4.18
2	GGAEWNATWRNFE	1536.6746	-0.002	78.09		36	1	-1.184615385	4.26
2	HVGPDFQPWVDRAIESL	1964.9745	-0.0183	71.43	27	39	1	-0.335294118	4.3
2	IHVGPDFQPWVDRAIESL	2078.0585	-0.0111	102.38	31	39	1	-0.066666667	4.3

2	KQSIDNLTEGFVQLAPSHENTL	2440.2234	-0.0376	61.92		39	1	-0.540909091	4.42
2	ITKEILDADALIHVGPDFQPWVDRA	2818.4654	-0.0684	125	28	40	1	-0.004	4.12
2	KNGTGIDAFQDRAVDNLAGVA	2131.0658	-0.0094	93.6	31	39	1	-0.219047619	4.18
2	KGNADFALIQNDIAYFAKNGTGIDAF	2773.3711	0.0152	79.58	25	40	1	-0.030769231	4.18
2	GASVENVGSLAKGNADFALIQNDIAYFAKNGTGI	3424.7263	-0.0272	119.15	23	40	1	0.132352941	4.3
2	TTGASVENVGSLAKGNADFALIQNDIAYFAKNGTGI	3626.8216	0.0246	46.95		40	1	0.086111111	4.3
2	NVSAWPDMHDKISIPDGA	1951.9098	0.0036	98.65	29	38	1	-0.472222222	4.16
2	KVGPNNQNVFDPAEV	1626.8002	0.0039	70.73	28	38	1	-0.753333333	4.08
2	VKVGPNNQNVFDPAEV	1725.8686	0.0013	58.05		38	1	-0.44375	4.08
2	FGEYRGLLEPG	1236.6139	0.0032	65.88	32	38	1	-0.472727273	4.26
2	FGEYRGLLEPGIN	1463.7409	-0.0015	67.81	26	38	1	-0.323076923	4.26
2	QGERQSAIENAQGDKQSNII	2185.0723	0.0015	43.21	24	39	1	-1.26	4.44
2	FGEYRGLLEPGINVIPPFVS	2203.1677	-0.0178	95.13	29	38	1	0.375	4.26
2	NLAQTTLRAVLGDMELDDTLNKRQEI	2956.5287	0.0133	45.97	22	39	1	-0.526923077	4.2
2	AQFPLHTWLPDAMEGPTPV	2106.0245	-0.0402	87.13	21	38	1	-0.142105263	4.06
2	KGVTLPLARGADTFDQGVIDGIVNAV	2625.4126	-0.0158	158.69	27	38	1	0.426923077	4.18
2	QFRLVDGLTTIEA	1461.7827	-0.0374	40.46	27	38	1	0.31538461	4.08

								5	
2	TNRKADDNTITVELL	1701.8897	-0.0041	80.63	33	38	1	-0.66	4.3
2	GTTNRKADDNTITVELL	1859.9589	-0.006	83.51		39	1	-0.647058824	4.3
2	VMGTTNRKADDNTITVELL	2090.0678	-0.0066	67.91	25	39	1	-0.257894737	4.3
2	NGQVADDDNQIDVEGTAPGKDNVAAIIIGSRGKVKFQ	3838.945	-0.0279	124.14	27	40	1	-0.5	4.24
2	YWQPGEGMLVEKPFGI	1849.9073	-0.0123	59.78	29	39	1	-0.2375	4.26
2	AYWQPGEGMLVEKPFGI	1920.9444	-0.0201	110.62	36	39	1	-0.117647059	4.26
2	VEFEPRSY	1025.4818	-0.0001	43.18	29	36	1	-1.025	4.26
2	TFIDKFPTEA	1167.5812	0.001	77.9		37	1	-0.2	4.08
2	EKDGNLPHFAPDAIRELILEA	2347.2172	-0.0325	43.04	21	39	1	-0.361904762	4.28
2	SEAFRGGDTGEAHPDVNFNPKST	2432.0993	-0.0067	47.13	23	37	1	-1.156521739	4.54
2	TMSEAFRGGDTGEAHPDVNFNPK	2476.1077	-0.0147	44.49	16	36	1	-1.039130435	4.54
2	MSEAFRGGDTGEAHPDVNFNPKST	2563.1398	-0.0069	59.68	17	36	1	-1.029166667	4.54
2	TMSEAFRGGDTGEAHPDVNFNPKST	2664.1875	-0.0116	63.73	20	36	1	-1.016	4.54
2	HFHGSQTAWEDDGVPT	1782.7598	0.0006	61.81	23	34	1	-1.06875	4.28
2	AGGSGSFSAGELSVPGHDSPNIA	2112.9712	0.0051	51.76	21	38	1	-0.095652174	4.06
2	SMFENTVEQFKT	1459.6653	0.0001	86.33	30	37	1	-0.7	4.26
2	SMFENTVEQFKT	1475.6602	0.0007	51.22	23	37	1	-0.7	4.26
2	SMFENTVEQFKTA	1530.7024	0.0021	54.28	24	37	1	-0.507692308	4.26

								8	
2	NFLLSEFGQEKFV	1556.7875	-0.0098	38.48	32	38	1	0.084615385	4.26
2	ADYFASEYGYEAGGERSSPPFRV	2554.1401	-0.0042	95.46		36	1	-0.804347826	4.16
2	NFLLSEFGQEKFVTSNDYFALPESRL	3050.5025	-0.0364	41.95	17	40	1	-0.265384615	4.16
2	NFLLSEFGQEKFVTSNDYFALPESRLS	3137.5346	-0.0259	45.18		40	1	-0.285185185	4.16
2	GENLGAEEYPLSRDLHAYTWQDTSRKEAAFI	3566.7066	-0.0124	43.57	18	40	1	-0.874193548	4.42
2	ASVPFHFWAPEAYEGAPAPV	2142.0211	-0.0432	115.88	29	37	1	0.15	4.24
2	RVGGGTLADRLEFAADLIEDGGVHAF	2685.3511	-0.0117	48.85	17	40	1	0.180769231	4.2
2	VWIIDNLVTMVEHFTTA	1988.0078	0.007	54.39	21	39	1	0.894117647	4.06
2	VWIIDNLVTMVEHFTTA	2004.0027	-0.0097	70.47	29	39	1	0.894117647	4.06
2	VWIIDNLVTMVEHFTTAA	2075.0398	-0.0056	42.93	23	39	1	0.944444444	4.06
2	SNRSQFVPSWLVPEAAGDLPLTV	2482.2856	-0.0331	50.87	21	39	1	0.117391304	4.08
2	FWEAFLPEMKT	1397.6689	-0.0232	53.32	32	37	1	-0.090909091	4.26
2	GGFWESVDRA	1122.5094	0.0036	55.72		37	1	-0.52	4.08
2	NIESAITGGTGDAVAKELR	1900.9854	-0.0017	98.93		38	1	-0.178947368	4.44
2	VRDAFDIVHERTDENPI	2024.9916	-0.0102	51.85	36	39	1	-0.794117647	4.2
2	GTAEIFDTLRVLTVV	1632.9087	-0.0044	40.31	27	35	1	1.02	4.08
2	SAGTAEIFDTLRVLTVV	1790.9778	-0.0082	68.45	29	36	1	0.95882352	4.08

								9	
2	SSAGTAEIFDTLRVLTVV	1878.0099	-0.0095	38.5	23	37	1	0.861111111	4.08
2	SLEQFQNTAVEQLQSRVA	2047.0334	-0.0143	60.77	33	39	1	-0.494444444	4.26
2	RSSGEAIGEFQKGREEVEQEL	2377.1509	-0.0044	69.16	23	39	1	-1.266666667	4.26
2	DGFEQWEKRYLPA	1637.7838	-0.0017	64.3		38	1	-1.4	4.44
2	DFFEFEARKV	1286.6295	0.0033	53.91	32	38	1	-0.45	4.44
3	SLLETTKQGEHEAF	1588.7733	0.0017	94.93	34	39	1	-0.821428571	4.54
3	NFSSLLETTKQGEHEA	1789.8482	-0.0072	101.49		38	1	-0.9875	4.54
3	SLLETTKQGEHEAFTL	1802.905	-0.0003	62.55	35	39	1	-0.525	4.54
3	STLGIEHEVDENGAVKR	1852.9279	-0.007	80.39	29	39	1	-0.758823529	4.62
3	RKVMEEAGYGPDNKYE	1884.8676	-0.0038	77.11	25	37	1	-1.68125	4.64
3	STLGIEHEVDENGAVKRA	1923.965	-0.0037	150.85	34	39	1	-0.616666667	4.62
3	NFSSLLETTKQGEHEAF	1936.9167	-0.0024	55.2	24	38	1	-0.764705882	4.54
3	RKVMEEAGYGPDNKYEL	1997.9516	-0.0097	47.36	27	38	1	-1.358823529	4.64
3	RLYWKEDANSDPGVREF	2080.9966	-0.0073	64.45	28	39	1	-1.341176471	4.56
3	RKVMEEAGYGPDNKYELN	2111.9946	0.0004	139.76	30	38	1	-1.477777778	4.64
3	RKVMEEAGYGPDNKYELN	2127.9895	0.005	109.2	27	38	1	-	4.64

								1.477777778	
3	NFSSLLETTKQGEHEAFTL	2151.0484	-0.011	103.17	32	39	1	-0.521052632	4.54
3	RKVMEEAGYGPDNKYELNWL	2411.1579	-0.0002	81.06	24	39	1	-1.185	4.64
3	KVAEARKVMEEAGYGPDNKYELN	2610.2747	-0.0049	61.57	25	39	1	-1.139130435	4.8
3	FHLQPPQLFSDGQEGYNNHYQGE	2704.1942	-0.0155	108.55	20	35	1	-1.413043478	4.4
3	RYADWLFTTPL	1381.703	-0.0014	69.44	30	37	1	-0.090909091	6.7
3	RYADWLFTTPLLL	1607.8712	-0.0085	77.71	28	36	1	0.507692308	6.7
3	RSRAIFGEAEAPEPS	1615.7954	0.0013	123.56	33	39	1	-0.8	4.56
3	RYADWLFTTPLLLL	1720.9552	-0.0007	90.62	25	35	1	0.742857143	6.7
3	PFGGEQNPIYWARYA	1767.8369	-0.0261	69.24	22	37	1	-0.773333333	6.82
3	YWARYADWLFTTPL	1801.8828	-0.0041	46.43	29	39	1	-0.1	6.64
3	LLRSRAIFGEAEAPEPS	1841.9635	-0.004	113.22	36	38	1	-0.258823529	4.56
3	ILLRSRAIFGEAEAPEPS	1955.0476	-0.0068	82.15	28	38	1	0.005555556	4.56
3	RYADWLFTTPLLLLDLA	2020.1033	-0.0149	50.09	25	37	1	0.735294118	3.88
3	GFGLILLRSRAIFGEAEAPEPS	2329.243	0.0007	124.85	32	38	1	0.268181818	4.56
3	VVHPYWNTDDV	1343.6146	-0.004	51.98	25	37	1	-0.509090909	3.88
3	VVGGTEAQRNSWPSQI	1727.8591	-0.0039	50.89		39	1	-0.5875	6.98
3	SYRFREEAGSYLVGEL	1874.9162	-0.0054	44.94	26	39	1	-0.50625	4.56

3	HVGPDFQPWVDRA	1522.7317	0.0003	61.17	29	38	1	-0.746153846	5.1
3	ASLDRDEEGIGQNRGKDPHFWLDPNRA	3092.4812	-0.0097	54.71	19	39	1	-1.47037037	4.7
3	SAPHEYSHTFEGPTGEYSYV	2256.96	-0.0078	36.24		33	1	-0.99	4.52
3	YVSPGDTVKWVWEGSTGHNVHATSVPDEA	3124.4527	-0.0096	74.83	19	38	1	-0.524137931	4.52
3	TSLVGRYGKHLAGSDVGDGGSDLE	2389.151	-0.0074	40.9	19	39	1	-0.45	4.44
3	SLVGRYGKHLAGSDVGDGGSDLESM	2506.1758	-0.0166	60.11	20	38	1	-0.36	4.44
3	TSLVGRYGKHLAGSDVGDGGSDLESM	2607.2235	-0.0128	96.59	31	38	1	-0.373076923	4.44
3	HILSSGRDGKFGEDTANSI	2002.9708	-0.0059	60.55	24	39	1	-0.657894737	5.28
3	RGDELVEVDQPKRDQNGRPFVPV	2649.3623	-0.0251	46.56	22	39	1	-1.2	4.64
3	VLRGDELVEVDQPKRDQNGRPFVPV	2861.5148	-0.0326	40.53	23	39	1	-0.784	4.64
3	RGDELVEVDQPKRDQNGRPFVPVQL	2890.505	-0.0019	39.2	19	39	1	-1.092	4.64
3	VSPDDETPTWRERKERTT	2202.0665	-0.0069	42.15	17	39	1	-2.133333333	4.72
3	VSPDDETPTWRERKERTTGLS	2459.2041	0.0001	55.35	22	39	1	-1.704761905	4.72
3	SGKVVVIDEAHGNGVGEF	1812.9007	0.0081	100.97	26	38	1	0.133333333	4.42
3	DHHVIAPDLPGFG	1373.6728	0.0032	61.28	27	39	1	-0.023076923	4.94
3	RLVVLEESDVQPH	1519.7995	0.0047	41.06	24	37	1	-0.3	4.42
3	NALQHLFSGIDPFAGVPV	1880.9785	0.0099	64.68	23	38	1	0.622222222	4.94
3	SKPPAEDIAQDVLEHRA	1874.9486	-0.0084	52.9	25	39	1	-0.894117647	4.54

3	ARSFYSHIKEAWEDPDDGKLA	2434.1553	-0.0183	82.01	25	38	1	-1.061904762	4.62
3	HALAGDDWLFPTYRS	1747.8318	-0.0282	56.34	26	38	1	-0.426666667	5.1
3	NQFLFGNLSTV	1238.6295	0.0022	39.09	23	37	1	0.454545455	6.02
4	FHLQPPQLFSDGQEGYNNH	2227.0083	0.0017	126.25	26	37	1	-1.252631579	5.04
4	YHNLAQSFWYDRLDYNPPGA	2426.108	-0.0214	39.28	17	36	1	-1.02	5.1
4	GFGLILL	731.4582	-0.0019	38.5		28	2	2.557142857	6.02
4	RYADWLFT	1070.5185	0.0042	49.98	29	37	1	-0.3125	6.7
4	TGRPEWIWLA	1227.64	0.0016	49.85		37	1	-0.24	6.98
4	RYADWLFTTPL	1381.703	0.0396	77.38	30	36	1	-0.090909091	6.7
4	YWARYADWLFT	1490.6983	-0.0016	39.03	29	38	1	-0.263636364	6.64
4	RYADWLFTTPLL	1494.7871	-0.0034	62.87	27	37	1	0.233333333	6.7
4	TGRPEWIWLALGT	1498.7932	-0.0051	78.74		38	1	0.023076923	6.98
4	GGEQNPIYWARYA	1523.7157	-0.0021	72.77	30	38	1	-0.984615385	6.82
4	DGIMIGTGLVGALTKV	1543.8644	-0.0013	57.9		35	1	1.14375	6.66
4	RYADWLFTTPLLL	1607.8712	-0.0023	84.54	31	36	1	0.507692308	6.7
4	RSRAIFGEAEAPEPS	1615.7954	-0.0044	43.18	24	39	1	-0.8	4.56
4	GADGIMIGTGLVGALTKV	1671.923	-0.0019	73.16	30	36	1	1.094444444	6.66
4	RYADWLFTTPLLLL	1720.9552	-0.0035	100.7	25	35	1	0.742857143	6.7
4	PFGGEQNPIYWARYA	1767.8369	-0.0041	108.32	30	38	1	-	6.82

								0.77333333	
4	YWARYADWLFTTPL	1801.8828	-0.0042	78.53	35	39	1	-0.1	6.64
4	GADGIMIGTGLVGALTKVY	1834.9863	-0.0312	38.14	24	38	1	0.968421053	6.6
4	GFGLILLRSRAIFGEAEAPEPS	2329.243	-0.0135	69.4	26	38	1	0.268181818	4.56
4	AHDQAPWVY	1085.493	0.0048	52.05	32	36	1	-0.688888889	4.94
4	NQPLPPDVLGH	1185.6142	-0.001	86.06		38	1	-0.645454545	4.94
4	AHDQAPWVYL	1198.5771	-0.0026	82.78	34	37	1	-0.24	4.94
4	SANQPLPPDVLGH	1343.6834	-0.0005	122.41		39	1	-0.469230769	4.94
4	SQIAHDQAPWVYL	1526.7518	-0.005	93.41	28	38	1	-0.169230769	4.94
4	DLTQQYAPFLRNL	1577.8202	-0.0047	77.55		38	1	-0.469230769	6.7
4	YDDAARQQYYHQA	1627.7015	-0.006	60.76	25	34	1	-1.823076923	5.1
4	DLTQQYAPFLRNLA	1648.8573	-0.0058	78.03	35	38	1	-0.307142857	6.7
4	NQIFDTLIQFKPGTS	1707.8832	-0.0062	70.08	24	38	1	-0.246666667	6.66
4	DLTQQYAPFLRNLAI	1761.9413	-0.003	73.98	28	37	1	0.013333333	6.7
4	TNQIFDTLIQFKPGTS	1808.9309	-0.0059	70.64	28	38	1	-0.275	6.66
4	YDDAARQQYYHQASQI	1955.8762	-0.0033	119.53	32	36	1	-1.46875	5.1
4	DLTQQYAPFLRNLAIFAA	2051.084	-0.012	43.69	25	38	1	0.366666667	6.7

4	DSGGPLHCLV	996.4699	0.0058	49.07	30	36	1	0.44	4.94
4	SWINNVIASN	1116.5564	0.0048	51.54	35	39	1	0.2	6.02
4	SAYISWINNVIA	1349.6979	0.0042	87.64	34	37	1	0.875	5.92
4	SAYISWINNVIASN	1550.7729	0.0042	45.67	23	38	1	0.442857143	5.92
4	QKIVVHPYWNTDDV	1712.8522	0.0006	59.05		38	1	-0.607142857	5.1
4	VVGGTEAQRNSWPSQI	1727.8591	-0.0069	86.56		39	1	-0.5875	6.98
4	QKIVVHPYWNTDDVA	1783.8893	-0.0001	67.1	22	38	1	-0.446666667	5.1
4	QKIVVHPYWNTDDVAA	1837.8999	0.0498	81.35	25	38	1	-0.30625	5.1
4	QKIVVHPYWNTDDVAA	1854.9265	-0.0008	107.42	34	39	1	-0.30625	5.1
4	QKIVVHPYWNTDDVAAG	1911.9479	-0.0042	91.25	24	38	1	-0.311764706	5.1
4	QKIVVHPYWNTDDVAAGY	2075.0112	-0.0012	123.26	38	39	1	-0.366666667	5.1
4	MFNSGADIIYHASAA	1566.7137	-0.0341	62.85	27	35	1	0.493333333	4.94
4	IHVGPDFQPW	1194.5822	0.0019	40.55	28	37	1	-0.32	4.94
4	LIHVGPDFQPW	1307.6663	-0.0009	53.39	26	37	1	0.054545455	4.94
4	AHNAFQYLGTAYDI	1582.7416	0.0017	126.64		38	1	-0.064285714	4.94
4	IHVGPDFQPWVDRA	1635.8158	-0.006	93.69	32	38	1	-0.371428571	5.1
4	GLHGHGWEPNASITKEILDA	2144.065	-0.0035	104.74	26	39	1	-0.465	5.18
4	AHNAFQYLGTAYDIEMRPL	2209.0626	-0.0096	151.89	34	39	1	-0.252631579	5.22
4	AHNAFQYLGTAYDIEMRPLV	2308.131	-0.0121	61.03	21	39	1	-0.03	5.22

4	AHNAFQYLGTAYDIEMRPLVV	2407.1994	-0.005	162.4	32	39	1	0.171428571	5.22
4	AHNAFQYLGTAYDIEMRPLVV	2423.1944	-0.0092	96.95	35	39	1	0.171428571	5.22
4	FVVGGWPVGA	987.5178	0.0022	55.25		38	1	1.35	6.02
4	NDIAYFAKNGTGI	1382.683	-0.0284	38.68	22	37	1	-0.138461538	6.6
4	LIQNDIAYFAKNGTGI	1736.9097	0.0013	38.46	23	37	1	0.1875	6.6
4	NVSAWPDMHDKI	1411.6554	0.0033	62.88	31	37	1	-0.708333333	5.1
4	SHNVSAWPDMHDKISIPDGA	2176.0007	-0.0346	145.71	22	36	1	-0.625	5.02
4	SHNVSAWPDMHDKISIPDGA	2191.9957	0.0003	106.1	25	37	1	-0.625	5.02
4	GGGGPVERSPHEMGVPIQA	1873.9105	-0.0064	70.13	26	39	1	-0.484210526	5.3
4	WLLYSGPAPT	1103.5651	0.0044	53.68		38	1	0.21	5.92
4	GYHAADIAPFVPA	1327.6561	-0.0069	58.76		38	1	0.546153846	4.94
4	AQFPLHTWLPDAM	1525.7388	-0.0027	78.45	29	38	1	0.069230769	4.94
4	FAGYHAADIAPFVPA	1545.7616	-0.0006	41.03	37	39	1	0.78	4.94
4	GLHIEFLHTWLAGPEGT	1876.9472	-0.0098	92.95	31	39	1	0.117647059	5.12
4	TGLHIEFLHTWLAGPEGT	1977.9949	-0.0108	111.52	31	39	1	0.072222222	5.12
4	QQFRLVDGLTTI	1389.7616	0.0019	49.76	31	36	1	0.191666667	6.78
4	HILSSGRDGKFGEDTA	1688.8118	-0.0061	41.7	20	39	1	-0.79375	5.28
4	AHILSSGRDGKFGEDTANSI	2074.0079	-0.0089	63.47	25	39	1	-0.535	5.28
4	DLWDHVRG	996.4777	0.0032	54.2	33	37	1	-1	5.1
4	MLGDVRHDPFQSGGMATPS	2001.9037	-0.004	56.4	23	37	1	-0.447368421	5.1
4	SFEWVSNTHNIL	1445.6939	0.0012	41.08	24	38	1	-0.13333333	5.12

								3	
4	GFTHTHTFDTEGV	1447.6369	0.0003	50.86	25	36	1	-0.553846154	5.04
4	GFTHTHTFDTEGVYT	1711.7479	0	119.47	33	35	1	-0.613333333	5.04
4	DWGGHDPIENTGFTHTH	1919.8187	-0.0032	171.33	29	34	1	-1.3	5.02
4	QRFTGDMTQYGFTSQA	1836.8101	-0.0104	43.75	23	35	1	-0.8875	6.7
4	EMHPMMREMFGGG	1508.6033	0.0008	32.54	16	31	1	-0.546153846	5.3
4	FHVTSGDVTHGFELV	1643.7944	-0.0015	53.92	24	38	1	0.373333333	5.04
4	SFRHWVGDDVPPAAG	1609.7638	-0.0018	67.4	28	38	1	-0.373333333	5.1
4	FFAPHDRDQLPGGPPAAL	1904.9533	-0.0061	86.28	27	39	1	-0.377777778	5.1
4	FFAPHDRDQLPGGPPAALNQ	2147.0548	-0.0108	126.05	33	39	1	-0.69	5.1
4	NYGDIHRYEPA	1333.6051	0.0006	67.98	25	37	1	-1.5	5.22
4	ATHAADVLFAG	1071.5349	0.0063	49.93	34	38	1	0.927272727	4.94
4	GLNPHDWTVY	1200.5564	-0.0044	69.93	31	37	1	-0.71	4.94
4	TFTQDQTTVSRGGVVPV	1790.9163	-0.0045	40.66	22	38	1	-0.082352941	6.78
4	TFTQDQTTVSRGGVVPVTL	2005.0481	-0.0088	52.51	21	38	1	0.089473684	6.78
4	DHHVIAPDLPGFGRSDRPPIEYA	2558.2666	-0.0195	93.68	24	40	1	-0.665217391	5.16
4	DHHVIAPDLPGFGRSDRPPIEYAA	2629.3037	-0.018	41.69	20	40	1	-0.5625	5.16
4	SNALQHLFSGIDPF	1544.7623	0.0049	40.85	29	39	1	0.157142857	4.94
4	SNALQHLFSGIDPFAGVPV	1968.0105	-0.0018	128.71	30	38	1	0.54736842	4.94

								1	
4	LHIDPAGRALTYDTGYI	1874.9527	-0.0047	104.2	35	39	1	-0.052941176	5.1
4	IYNDPDDLPHHRVGFL	1906.9326	0.002	51.97	35	39	1	-0.66875	5.02
4	HLGRYNGEFEAPREVV	1871.9278	-0.0075	61.86	25	39	1	-0.81875	5.4
4	SGALLGFYYSPS	1260.6026	0.0023	57.4	36	38	1	0.4	5.86
4	NRVGPWGLGTIVVDSV	1667.8995	-0.0058	58.08	30	37	1	0.525	6.78
4	WLPDAHVQAPTPA	1401.7041	-0.0089	63.84		38	1	-0.246153846	4.94
4	ALIGGFDGFSHV	1218.6033	0.0037	75.66	31	38	1	0.933333333	4.94
4	ALIGGFDGFSHVA	1289.6405	-0.0006	63.9	31	38	1	1	4.94
4	FDRLDHPAWTIV	1468.7463	-0.0025	42.02	30	37	1	-0.066666667	5.1
4	SNLGTVGDRLAPGWAGSGGATQ	2071.0083	-0.0036	77.04	28	39	1	-0.259090909	6.78
4	FEHGLPSGPHLADAVPG	1699.8318	-0.0028	64.03	28	39	1	-0.117647059	5.04
4	SLPHVPQSQLPEQV	1557.8151	0.0013	45.49	29	38	1	-0.542857143	5.12
4	SIEPKVAADHQAVLQ	1604.8522	-0.0027	57.88		37	1	-0.093333333	5.22
4	FGGLLLI	731.4582	-0.0019	42.75		28	1	2.557142857	6.02
4	ALMPLHSWLPDAM	1480.7207	-0.0004	70.04	31	38	1	0.553846154	4.94
4	FHAPELGLKMDALPG	1594.8177	-0.0048	55.22	30	38	1	0.106666667	5.22
4	FHAPELGLKMDALPGQQ	1850.9349	-0.0058	45.81	25	38	1	-0.317647059	5.22

4	PTFLTGAF	852.4382	0.0077	43.49	31	35	1	0.975	6.02
4	SVPTFLTGAF	1038.5386	-0.0001	45.84	29	35	1	1.12	6.02
4	FGVALLI	731.4582	-0.0019	36.68		28	3	2.928571429	6.02
4	NSYFDSYGFHTI	1449.6201	0.0011	41.38	28	35	1	-0.45	4.94
4	SDGMRQAFGVPVGT	1420.6769	-0.0002	38.5	37	38	1	-0.064285714	6.78
4	AHYVGPEIDVSLH	1435.7096	0.0064	59.85	24	38	1	0.076923077	5.04
4	NILLDFVRPT	1186.671	0.0032	53.07		35	1	0.53	6.78
4	NILLDFVRPTTG	1344.7402	-0.0003	40.65	21	36	1	0.35	6.78
4	SFTLHPLGDSGIIA	1426.7456	-0.0008	64.25	27	38	1	0.7	4.94
4	EVTGPWPTYQHDFAR	1802.8376	-0.0057	43.31	24	38	1	-1.106666667	5.22
4	APDLPGFGHTDRPSIA	1649.8162	-0.0001	59.99	27	38	1	-0.44375	5.1
4	VLRERYAAGDIDDEEFTARRR	2537.2735	0.0018	50.66	24	39	1	-1.2	4.78
4	TYEWSHTAPVAV	1359.6459	0.0013	50.07		38	1	-0.058333333	5.12
4	AVGGHDLPYAFQPA	1441.699	0.0031	42.09	22	38	1	0.05	4.94
4	FLPFLPQSVVPS	1329.7333	0.001	41.15	27	36	1	0.975	6.02
4	ALSATPHTDWVTVT	1497.7464	0.0072	40.21	29	38	1	0.214285714	4.94
4	GETPAQRGVNFAR	1401.7113	-0.0161	38.69		38	2	-0.923076923	10.88
5	QSFWYDRL	1113.5243	0.0036	46.32	34	37	1	-0.9875	6.7
5	RLEQAHIEMTI	1339.6918	-0.0001	46.99	25	37	1	-0.218181818	5.3
5	QYQSPTWKEMA	1367.618	0.0085	44.62	23	37	1	-1.454545455	6.8

5	MAYVINREQFV	1368.686	0.004	39.81	32	38	1	0.281818182	6.88
5	MAYVINREQFVS	1455.718	-0.0007	38.21	25	38	1	0.191666667	6.88
5	KQGEHEAFTLGWVA	1571.7732	-0.0024	90.74	37	39	1	-0.4	5.3
5	QYQSPTWKEMANTI	1695.7926	-0.0359	104.48	30	37	1	-1.121428571	6.8
5	DYPRPRNFMQLVQPDNT	2090.0004	-0.0481	62.79	20	37	1	-1.417647059	6.84
5	FWYDRLDYNPPGAMDSSKQKT	2518.1587	-0.0107	103.62	20	38	1	-1.376190476	6.74
5	NNAVAGIDGKGRLSGISSTFNSLDPV	2588.3194	-0.0242	67.15	25	39	1	-0.061538462	6.8
5	DGKGRLSGISSTFNSLDPVASGNTASGA	2665.2944	-0.0221	57.81	23	39	1	-0.275	6.8
5	FWYDRLDYNPPGAMDSSKQKTNNA	2817.2816	-0.0228	75.23	19	36	1	-1.420833333	6.74
5	SLLETTKQGEHEAFTLGWVADYPRPR	3000.5094	-0.0151	112.87	27	40	1	-0.796153846	5.44
5	NLAQSFWYDRLDYNPPGAMDSSKQKT	3031.4134	-0.0206	103.27	24	38	1	-1.196153846	6.74
5	SLLETTKQGEHEAFTLGWVADYPRPRNF	3261.6207	-0.0176	46.74	21	40	1	-0.764285714	5.44
5	GFGLILL	731.4582	-0.0068	46.47		29	3	2.557142857	6.02
5	RYADWLFT	1070.5185	-0.0059	44.53	27	36	1	-0.3125	6.7
5	TGRPEWIWL	1156.6029	-0.0239	42.73		37	1	-0.466666667	6.98
5	RYADWLFTT	1171.5662	0.0024	57.89	34	38	1	-0.355555556	6.7

5	SDPDAKKFYAI	1253.6292	0.0064	52.12		37	1	-0.690909091	6.68
5	RYADWLFTTP	1268.619	0.0025	58.93	31	37	1	-0.48	6.7
5	SDPDAKKFYAIT	1354.6769	-0.0011	66.4	37	38	1	-0.691666667	6.68
5	RYADWLFTTPL	1381.703	0.0272	72.72	29	37	1	-0.090909091	6.7
5	SDPDAKKFYAITT	1455.7245	0.0012	44.4	28	38	1	-0.692307692	6.68
5	RYADWLFTTPLL	1494.7871	-0.0094	56.2	26	37	1	0.233333333	6.7
5	TGRPEWIWLALGT	1498.7932	-0.0249	85.84		38	1	0.023076923	6.98
5	GGEQNPIYWARYA	1523.7157	-0.0032	89.94	32	38	1	-0.984615385	6.82
5	MGVSDPDAKKFYAI	1540.7596	-0.0006	73.05		38	1	-0.135714286	6.68
5	RYADWLFTTPLLL	1607.8712	-0.0065	84.31	30	36	1	0.507692308	6.7
5	MGVSDPDAKKFYAIT	1641.8072	0.0041	100.82	37	38	1	-0.173333333	6.68
5	RYADWLFTTPLLLL	1720.9552	-0.0038	96.63	25	35	1	0.742857143	6.7
5	SQAQITGRPEWIWLA	1754.9104	-0.0057	40.58	34	38	1	-0.26	6.98
5	PFGGEQNPIYWARYA	1767.8369	-0.0061	110.1	31	38	1	-0.773333333	6.82
5	YWARYADWLFTTPL	1801.8828	0.0072	82.02		39	1	-0.1	6.64
5	SDPDAKKFYAITTLVPA	1835.9669	-0.0028	89.47	32	37	1	-0.047058824	6.68

5	SDPDAKKFYAITTLVPAIA	2020.0881	-0.0046	97.05	32	37	1	0.289473684	6.68
5	MGVSDPDAKKFYAITTLVPA	2123.0973	-0.0087	42.71	29	39	1	0.245	6.68
5	ANQPLPPDVLGH	1256.6513	0.0022	77.01	30	38	1	-0.441666667	4.94
5	TLTLREDATFH	1302.6568	0.0033	64.21	30	38	1	-0.418181818	5.22
5	SANQPLPPDVLGH	1343.6834	0.0018	54.99	22	39	1	-0.469230769	4.94
5	DLTQQYAPFLR	1350.6932	-0.0082	44.53	32	38	1	-0.581818182	6.7
5	DLTQQYAPFLRNL	1577.8202	-0.0019	89.79		38	1	-0.469230769	6.7
5	TNQIFDTLIQFKPG	1620.8512	-0.0045	100.81	34	38	1	-0.207142857	6.66
5	DLTQQYAPFLRNLA	1648.8573	-0.006	80.81	33	38	1	-0.307142857	6.7
5	IDLTQQYAPFLRNLA	1761.9413	-0.0035	44	27	37	1	0.013333333	6.7
5	TNQIFDTLIQFKPGTS	1808.9309	-0.0044	113	33	38	1	-0.275	6.66
5	DLTQQYAPFLRNLAIF	1909.0098	-0.0177	43.47	22	38	1	0.1875	6.7
5	DLTQQYAPFLRNLAIFA	1980.0469	-0.0085	92.87	31	38	1	0.282352941	6.7
5	SFELATFSNPRGYNPNPI	2022.9799	-0.0136	52.25	27	39	1	-0.644444444	6.88
5	DLTQQYAPFLRNLAIFAA	2051.084	-0.0093	89.38		38	1	0.366666667	6.7
5	SFELATFSNPRGYNPNPIQTA	2323.1233	-0.0171	60.53	18	39	1	-0.666666667	6.88

5	IFKTTGSNQTRAQGLINGDAHITDNLDAQ	3098.5381	-0.0007	50.35	22	40	1	-0.64137931	5.2
5	AVIFKTTGSNQTRAQGLINGDAHITDNLDAQ	3268.6436	-0.0049	64.25	18	40	1	-0.40645161 3	5.2
5	QRNSWPSQI	1097.5254	0.0034	41.23		37	1	-1.62222222 2	11.04
5	AYISWINNVIA	1262.6659	0.0028	38.61	29	37	1	1.02727272 7	5.92
5	TEAQRNSWPSQI	1415.6793	0.0022	47.38		38	1	-1.41666666 7	6.98
5	AYISWINNVIASN	1463.7408	0.0039	54.01	31	38	1	0.53846153 8	5.92
5	GGTEAQRNSWPSQI	1529.7223	0.0016	53.03		38	1	-1.27142857 1	6.98
5	VGGTEAQRNSWPSQI	1628.7907	0.0002	41.86	29	38	1	-0.90666666 7	6.98
5	QKIVVHPYWNTDDV	1712.8522	-0.0051	98.87		38	1	-0.60714285 7	5.1
5	VVGGTEAQRNSWPSQI	1727.8591	0.0012	102.85		39	1	-0.5875	6.98
5	QKIVVHPYWNTDDVA	1783.8893	0.0007	45.61	19	38	1	-0.44666666 7	5.1
5	GVQKIVVHPYWNTDDV	1868.9421	-0.0074	111.15	31	39	1	-0.29375	5.1
5	VGVQKIVVHPYWNTDDV	1968.0105	-0.0033	77.63	25	38	1	-0.02941176 5	5.1
5	AQDIGRFAIGV	1145.6193	0.0063	48.32	33	36	1	0.66363636 4	6.78
5	RSMFNSGADII	1209.5812	-0.0156	82.84	30	37	1	0.18181818 2	6.78
5	TGDGVFRAAQDIGRF	1608.8009	-0.0055	64.61	30	38	1	-0.23333333 3	6.92

5	TGDGVFRAAQDIGRFA	1679.838	-0.0077	86.51	37	38	1	-0.10625	6.92
5	TGDGVFRAAQDIGRFAIGV	1949.0119	-0.0004	89.82	29	38	1	0.34736842 1	6.92
5	GLHGHGWEPNASITKEILDA	2144.065	-0.0044	133.6	32	39	1	-0.465	5.18
5	AAHNAFQYLGTAYDIEMRPL	2280.0997	-0.0084	75.05	24	39	1	-0.15	5.22
5	RNINMPTFKLLLGNAAPDDVGGA	2383.2318	-0.0128	46.65	22	39	1	-0.03478260 9	6.8
5	KQSIDNLTEGFVQLAPSHENTLRT	2697.3722	-0.0316	148.39	26	40	1	-0.7125	5.34
5	TLYPETITLITRA	1490.8344	-0.0019	44.59		34	1	0.36153846 2	6.88
5	QNDIAYFAKNGTGI	1510.7416	-0.0015	97.63	35	38	1	-0.37857142 9	6.6
5	LIQNDIAYFAKNGTGI	1736.9097	-0.0039	114.3		38	1	0.1875	6.6
5	GGGGPVERSPHEMGVPI	1674.8148	-0.0028	92.43	34	39	1	-0.44117647 1	5.3
5	GGGGPVERSPHEMGVPIQA	1873.9105	-0.0065	68.12	24	39	1	-0.48421052 6	5.3
5	YVSPGDTVKWVWEGSTGH	2003.9378	-0.0065	111.45		38	1	-0.57777777 8	5.22
5	YVSPGDTVKWVWEGSTGHNV	2217.0491	-0.0168	116.4	32	38	1	-0.485	5.22
5	YTFDMRTQTI	1274.5965	0.0018	45.15	31	38	1	-0.57	6.7
5	RGLLEPGINVIPPFVS	1706.9719	-0.005	68.65	25	33	1	0.61875	6.98
5	HAFPHMAEQA	1137.5025	0.0009	61.99	33	36	1	-0.49	6.02
5	RLVDGLTTI	986.5761	0.0053	64.42		33	1	0.72222222 2	6.78
5	QQFRLVDGL	1074.5822	0.0022	48.54	30	36	1	-0.08888888 9	6.78
5	QFRLVDGLTTI	1261.703	0.0008	57.75	28	35	1	0.52727272 7	6.78
5	QQFRLVDGLTTI	1389.7616	0.0012	77.72	35	36	1	0.19166666 7	6.78

5	SGSPTGDQIRDRILSNTV	1914.9759	-0.0049	52.93	23	39	1	-0.67777778	6.92
5	SLTVMGTTNRKADDNTITV	2036.0208	-0.0094	50.49	33	39	1	-0.35263159	6.8
5	SDVQPTVKLLKDVGILS	1811.0404	-0.0072	94.12		34	1	0.34705884	6.72
5	SLSDVQPTVKLLKDVGILS	2011.1565	-0.0157	81.09		34	1	0.46842105	6.72
5	DRFQGTLGWAPTYGAFQGFVTA	2389.1492	-0.0099	50.58	20	39	1	-0.03181818	6.7
5	LSSGTRDPVLSEQFTTEHWVGLAGRS	2829.4046	-0.0247	68.15	23	40	1	-0.41538461	5.34
5	SILIQGRGFAEAWQI	1687.9046	-0.0071	41.86	25	37	1	0.41333333	6.98
5	HAGAMLGDVRHDPFQSGGMATPS	2338.0583	-0.0016	92.43	36	37	1	-0.36956521	6.02
5	NKGVLFIDEINTLDVRSQQKL	2429.3278	-0.0197	95.24	25	37	1	-0.33809523	6.9
5	NKGVLFIDEINTLDVRSQQKLMTA	2732.4531	-0.0195	59.21	20	38	1	-0.17083333	6.9
5	RLLPKESLQDVLVYHNPDDSNEPKVRTVPA	3428.8052	-0.0098	55.79	19	39	1	-0.80666667	5.42
5	MSRLLPKESLQDVLVYHNPDDSNEPKVRTVPA	3646.8777	-0.0823	42.69	18	40	1	-0.721875	5.42
5	YGDVLRRLLTTDT	1521.8151	-0.0021	56.68	34	37	1	-0.32307692	6.84
5	AQRFTGDMTQYGFTSQA	1907.8472	-0.007	72.39	24	36	1	-0.72941176	6.7
5	MERYIPQVT	1135.5696	0.0044	37.57	24	37	1	-0.5	6.88
5	GNDFFGQFRSLLAGGQGT	1870.8962	-0.0107	118.54	29	38	1	-	6.78

								0.23333333	
5	FRGGDTGEAHPDVNFNPKST	2144.9876	-0.0063	72.03	20	37	1	-1.205	5.28
5	INTLTENGHEVRFY	1691.8267	-0.0009	99.49	30	39	1	-0.678571429	5.3
5	SGKVVVIDEAHGNGVGEFQLRPL	2420.2812	-0.0209	39.82	23	38	1	0.017391304	5.34
5	SNYGDIHRYEPA	1420.6371	0.0001	58.28	21	36	1	-1.441666667	5.22
5	ISPDQRYGQNQQLQQAIGQ	2171.0719	-0.013	51.18	22	39	1	-1.363157895	6.7
5	ASVPFHFWAPEAY	1520.7088	-0.0031	63.76	31	38	1	0.176923077	5.12
5	TFTQDQTTVSRGGVVPV	1790.9163	-0.0061	76.21	27	39	1	-0.082352941	6.78
5	SNRSQFVPSWLVPEAA	1786.9002	-0.0011	74.51	29	39	1	-0.18125	6.98
5	RLEDWRSYTPL	1434.7255	-0.023	47.28	33	38	1	-1.245454545	6.96
5	TIQWIGPDRA	1155.6037	0.0048	55.48		37	1	-0.43	6.78
5	TYWAQVPESLKDKNAVTL	2062.0735	-0.0237	62.7	25	39	1	-0.455555556	6.8
5	TFRQSFLDVV	1210.6346	0.0053	54.99	30	36	1	0.48	6.78
5	TYWANVPESLKDKNAV	1833.9261	-0.0026	68.07	24	39	1	-0.70625	6.8
5	TYWANVPESLKDKNAVTL	2048.0578	-0.0226	68.57	36	39	1	-0.455555556	6.8
5	TYWANVPESLKDKNAV	1833.9261	-0.0026	68.07	24	39	1	-0.70625	6.8
5	TYWANVPESLKDKNAVTL	2048.0578	-0.0226	68.57	36	39	1	-0.455555556	6.8
5	SNLGTVGDRLAPGWA	1512.7685	0.0003	56.83	31	38	1	-0.08666666	6.78

								7	
5	GAFVSLPHVPQSQLPEQV	1932.0105	-0.0211	51.86	23	39	1	0.04444444	5.12
5	FGGLLLI	731.4582	-0.0068	56.27		29	1	2.557142857	6.02
5	SSRRDEIGQLF	1306.6629	0.002	48.19	28	38	1	-0.94545454 5	7.04
5	SYTGLRWLAGGAET	1480.731	-0.0004	93.64	37	38	1	-0.171428571	6.88
5	TIQWIGPDKA	1127.5975	0.006	47.44		37	1	-0.37	6.66
5	TYWANVPESLKDKNAV	1833.9261	-0.0026	68.07	24	39	1	-0.70625	6.8
5	FGVALLI	731.4582	-0.0068	49.98		29	2	2.928571429	6.02
5	FSAGQTRGFAIGGDFTPA	1798.8639	-0.0311	38.63	23	38	1	0.077777778	6.78
5	FNLLLL	731.4581	-0.0067	39.89		29	4	2.416666667	6.02
5	FRALLL	731.4694	-0.018	32.38		29	7	1.916666667	11.04
5	TAYDQRLPFLA	1293.6717	-0.0013	38.72	24	37	1	-0.1	6.7
5	TIQWIGPDKA	1127.5975	0.006	47.44		37	1	-0.37	6.66
5	FWRTLLEPV	1159.639	-0.0005	38.95	32	36	1	0.377777778	6.98
5	NGVSLIELVRGNAEKPV	1793.9999	-0.0038	42.79	23	35	1	0.052941176	7
5	FARLLL	731.4694	-0.018	34.6		29	6	1.916666667	11.04
5	SRPQNFEIFLQ	1377.7041	0.0015	39.65	29	38	1	-0.636363636	6.98
5	NRFSEWTVPV	1233.6142	0.0045	38.71	26	38	1	-0.43	6.98
5	FRALLL	731.4694	-0.018	32.38		29	7	1.916666667	11.04
6	YVINREQFV	1166.6084	0.0006	45.73		36	1	-	6.88

								0.066666667	
6	RYNYFLNENA	1302.5993	0.0024	45.35	27	37	1	-1.27	6.82
6	MAYVINREQFV	1368.686	0.0019	70.17	34	37	1	0.281818182	6.88
6	KYPYGIDETKVA	1382.7082	-0.0017	42.78	30	37	1	-0.8	6.78
6	KYPYGIDETKVAEAR	1738.889	-0.002	111.72	30	38	1	-1.053333333	6.88
6	RKVMEEAGYGPDNKY	1755.825	0.0012	105.86	26	38	1	-1.56	6.88
6	YVINREQFVSDVFKG	1799.9206	-0.0091	76.01	24	39	1	-0.146666667	6.86
6	DYPRPRNFMQLVQPD	1874.9097	-0.0055	55.8	28	39	1	-1.326666667	6.84
6	DGKGRLSGISSTFNSLDPV	1948.9854	-0.005	121.38		39	1	-0.342105263	6.8
6	WLQYQSPTWKEMANTI	1994.956	-0.0016	48.74	22	38	1	-0.8	6.8
6	MAYVINREQFVSDVFKG	2001.9982	-0.0016	121.1	32	39	1	0.088235294	6.86
6	DYPRPRNFMQLVQPDNT	2090.0004	-0.002	79.45	26	39	1	-1.417647059	6.84
6	GIDGKGRLSGISSTFNSLDPV	2119.091	-0.0049	95.83	32	39	1	-0.114285714	6.8
6	VAGIDGKGRLSGISSTFNSLDPV	2289.1965	0.0005	93.16	29	38	1	0.156521739	6.8
6	FWYDRLDYNPPGAMDSSKQKT	2518.1587	-0.001	40.2	19	38	1	-1.376190476	6.74
6	FWYDRLDYNPPGAMDSSKQKT	2534.1536	0.0022	40.29	17	38	1	-1.376190476	6.74
6	SSGETLPGGRKLGTYGPMENDKTVNSS	2781.3239	-0.0131	68.21	20	39	1	-1.033333333	6.92

								3	
6	FWYDRLDYNPPGAMDSSKQKTNNA	2817.2816	-0.0094	88.95	22	37	1	-1.420833333	6.74
6	GFGLILL	731.4582	-0.0008	53.07		28	2	2.557142857	6.02
6	IAFTMYL	857.4357	-0.0019	39.95	33	37	1	1.828571429	5.92
6	GLGTLYFLV	981.5535	-0.0009	41.73		33	1	1.733333333	5.92
6	RYADWLFT	1070.5185	0.0039	46.14	29	37	1	-0.3125	6.7
6	TGRPEWIWL	1156.6029	-0.0045	37.05	34	37	1	-0.466666667	6.98
6	TGRPEWIWLA	1227.64	0.0088	61.11		37	1	-0.24	6.98
6	GGEQNPIYWAR	1289.6153	0.0054	47.04	23	38	1	-1.209090909	6.88
6	ALMGLGTLYFLV	1296.7152	-0.0008	40.94		35	1	1.925	5.92
6	DVSAKVGFGLILL	1330.786	-0.0034	68.32		31	1	1.530769231	6.66
6	ITGRPEWIWLA	1340.7241	0.0018	48.1	33	36	1	0.190909091	6.98
6	SDPDAKKFYAIT	1354.6769	0	71.48		38	1	-0.691666667	6.68
6	GRPEWIWLALGT	1397.7456	-0.0018	68.79		37	1	0.083333333	6.98
6	ARYADWLFTTPL	1452.7401	-0.0061	45.58	22	38	1	0.066666667	6.7
6	QITGRPEWIWLA	1468.7827	0.0003	51.92		37	1	-0.116666667	6.98
6	RYADWLFTTPLL	1494.7871	0.0028	41.6	26	37	1	0.233333333	6.7
6	TGRPEWIWLALGT	1498.7932	-0.0005	80.58		38	1	0.023076923	6.98
6	GVSDPDAKKFYAIT	1510.7668	0.002	39.81	33	38	1	-	6.68

								0.321428571	
6	GGEQNPIYWARYA	1523.7157	-0.0024	56.08	25	38	1	-0.984615385	6.82
6	PFGGEQNPIYWAR	1533.7365	0.0064	51.14	26	38	1	-0.930769231	6.88
6	DGIMIGTGLVGALTKV	1543.8644	-0.0029	60.35	28	35	1	1.14375	6.66
6	MGVSDPDAKKFYAIT	1641.8072	-0.0036	88.36	35	38	1	-0.173333333	6.68
6	GADGIMIGTGLVGALTKV	1671.923	-0.0026	175.97		36	1	1.094444444	6.66
6	RYADWLFTTPLLLL	1720.9552	-0.001	43.64	26	35	1	0.742857143	6.7
6	QITGRPEWIWLALGT	1739.9359	-0.004	43.67	30	37	1	0.086666667	6.98
6	PFGGEQNPIYWARYA	1767.8369	-0.0018	100.64	34	38	1	-0.773333333	6.82
6	YWARYADWLFTTPL	1801.8828	0.0041	83.9	37	39	1	-0.1	6.64
6	GADGIMIGTGLVGALTKVY	1834.9863	-0.0073	150.03		37	1	0.968421053	6.6
6	SDPDAKKFYAITTLVPA	1835.9669	-0.0035	72.81	28	38	1	-0.047058824	6.68
6	LVGADGIMIGTGLVGALTKV	1884.0754	-0.0042	141.11		33	1	1.385	6.66
6	GVSQAQITGRPEWIWLA	1911.0003	-0.0076	67.86	30	38	1	-0.005882353	6.98
6	ALVGADGIMIGTGLVGALTKV	1955.1125	-0.0065	104.57		34	1	1.404761905	6.66
6	SDPDAKKFYAITTLVPAIA	2020.0881	-0.0107	77.3	26	37	1	0.289473684	6.68
6	LVGADGIMIGTGLVGALTKVY	2047.1388	-0.0042	90.73	32	35	1	1.257142857	6.6
6	DLTQQYAPFLRNL	1577.8202	-0.0006	55.89	34	38	1	-0.46923076	6.7

								9	
6	DLTQQYAPFLRNLA	1648.8573	-0.0029	68.15	31	38	1	-0.307142857	6.7
6	IDLTQQYAPFLRNL	1690.9042	-0.004	72.62		37	1	-0.114285714	6.7
6	DLTQQYAPFLRNLAI	1761.9413	-0.0034	89.38	35	37	1	0.013333333	6.7
6	IDLTQQYAPFLRNLAI	1875.0254	-0.0091	63.48	29	36	1	0.29375	6.7
6	DLTQQYAPFLRNLAIFA	1980.0469	-0.0059	77.39	34	37	1	0.282352941	6.7
6	DLTQQYAPFLRNLAIFAA	2051.084	-0.0054	85.29	31	38	1	0.366666667	6.7
6	KVTNQIFDTLIQFKPGTSGELT	2436.2901	-0.0077	168.02	31	38	1	-0.222727273	6.84
6	AVIFKTTGSNQTRAQGLINGDAHITDNLDAQ	3268.6436	-0.0037	41.49	17	40	1	-0.406451613	5.2
6	AYISWINNVIA	1262.6659	-0.0036	85.49		37	1	1.027272727	5.92
6	VVGGTEAQRNSWPS	1486.7165	-0.001	43.31		38	1	-0.742857143	6.98
6	GGTEAQRNSWPSQI	1529.7223	0.0008	76.96		38	1	-1.271428571	6.98
6	VVGGTEAQRNSWPSQI	1727.8591	0.0196	83.84		39	1	-0.5875	6.98
6	GLLTQREFSA	1120.5877	0.0039	41.19	26	37	1	-0.12	6.98
6	AQDIGRFAIGV	1145.6193	0.0024	66.73		37	1	0.663636364	6.78
6	RSMFNSGADII	1209.5812	0.0026	83.64	29	37	1	0.181818182	6.78
6	GDGVFRAAQDIGRFAIGV	1847.9642	-0.004	38.27	30	38	1	0.405555556	6.92
6	GFVGGIETPLIKKFEAGFTA	2081.1197	-0.0127	60.36	19	37	1	0.48	6.92
6	AQDIGRFAIGVDRDQSVTKS	2162.108	-0.0057	58.71	28	39	1	-0.485	6.9

6	TVGFVGGIETPLIKKFEAGFT	2210.1987	-0.0183	68.42	26	37	1	0.538095238	6.92
6	GFVGGIETPLIKKFEAGFTAGV	2237.2096	-0.0111	130.62		37	1	0.609090909	6.92
6	GTVYGTGGLGDGSFNDQAKQGLERAK	2625.2783	-0.0136	79.31	19	39	1	-0.823076923	6.86
6	KVGTVYGTGGLGDGSFNDQAKQGLER	2653.3096	-0.0181	170.96	32	39	1	-0.730769231	6.86
6	VFETRRPAEQLL	1457.799	-0.0014	58.33	29	37	1	-0.45	7.12
6	AVFETRRPAEQLL	1528.8362	0.0024	55.37	35	36	1	-0.276923077	7.12
6	RNINMPTFKLLLGNAAPDDV	2198.1517	-0.0086	147.14	29	38	1	-0.09	6.8
6	RNINMPTFKLLLGNAAPDDVGGA	2383.2318	-0.0047	104.66	30	39	1	-0.034782609	6.8
6	ARNINMPTFKLLLGNAAPDDVGGA	2454.2689	-0.0047	117.2	28	39	1	0.041666667	6.8
6	LIQNDIAYFAKNGTGI	1736.9097	0.0033	102.23		37	1	0.1875	6.6
6	AGGQSGTYYPLSNEFKT	1818.8424	0.0004	57.86	21	38	1	-0.847058824	6.76
6	KGNADFALIQNDIAYFAKNGTGI	2440.2386	-0.011	187.3		39	1	-0.082608696	6.68
6	GSLAKGNADFALIQNDIAYFAKNGTGI	2768.4133	-0.0143	65.5	17	40	1	0.092592593	6.68
6	ALAAGYAERGIGS	1234.6306	0.0048	63.68	36	37	1	0.323076923	6.88
6	ALAAGYAERGIGSA	1305.6677	0.0024	91.94		38	1	0.428571429	6.88
6	AGYAERGIGSAAVGA	1348.6735	0.005	39.26	31	38	1	0.4	6.88
6	ALAAGYAERGIGSAAV	1475.7732	0.0006	93.55		38	1	0.75	6.88
6	ALAAGYAERGIGSAAVG	1532.7947	-0.0002	94.01	32	38	1	0.682352941	6.88
6	ALAAGYAERGIGSAAVGA	1603.8318	-0.0012	144.17		38	1	0.74444444	6.88

								4	
6	AVGLAALAAGYAERGIGS	1645.8787	-0.0036	37.6	33	37	1	0.85555555 6	6.88
6	AALAAGYAERGIGSAAVGA	1674.8689	-0.0002	79.74	31	38	1	0.8	6.88
6	AALAVGLAALAAGYAERGI	1756.9835	-0.0013	106.39		35	1	1.26315789 5	6.88
6	LAVGLAALAAGYAERGIGS	1758.9628	-0.0021	117.45		36	1	1.01052631 6	6.88
6	LAVGLAALAAGYAERGIGSA	1829.9999	-0.0052	106.28		37	1	1.05	6.88
6	AALAVGLAALAAGYAERGIGS	1901.037	-0.0032	154.57		37	1	1.08571428 6	6.88
6	AALAVGLAALAAGYAERGIGSA	1972.0741	-0.0064	155.29		37	1	1.11818181 8	6.88
6	AALAVGLAALAAGYAERGIGSAAV	2142.1797	-0.0128	50.99	29	37	1	1.275	6.88
6	AALAVGLAALAAGYAERGIGSAAVGA	2270.2382	-0.008	78.78		37	1	1.23076923 1	6.88
6	KTVKVGPNNQNVFDPAEV	1955.0112	-0.006	130.83		39	1	-0.65	6.84
6	DAVVYIRVRDA	1275.6935	0.0033	46.38	31	36	1	0.30909090 9	6.84
6	LLEPGINVIPPFVSRT	1750.9982	-0.0026	109.53	28	34	1	0.6	6.98
6	GLLEPGINVIPPFVSRT	1808.0196	-0.0053	39.93	20	34	1	0.54117647 1	6.98
6	FGEYRGLLEPGINVIPPFVSRT	2460.3165	-0.0125	42.9	17	38	1	0.10454545 5	7.04
6	ALGYEHYHHLLESFA	1785.8474	-0.0026	68.31	24	38	1	-0.17333333 3	5.74
6	AFSALGYEHYHHLLES	1872.8795	-0.0025	129.72	32	38	1	-0.2125	5.74
6	AFSALGYEHYHHLLESFA	2090.985	0.0004	60.3	19	38	1	0.06666666 7	5.74
6	RLVDGLTTI	986.5761	0.0041	47.6		34	1	0.72222222 2	6.78
6	IYLGGGEGGRYV	1239.6248	0.0006	63.5	32	37	1	-0.00833333 3	6.82
6	IYLGGGEGGRYVAL	1423.746	-0.0004	38.14	28	37	1	0.39285714 3	6.82

6	IYLGGGEGGRYVALI	1536.83	-0.0003	65.56	32	36	1	0.666666667	6.82
6	MRQLRGDEATREWLT	1860.9264	-0.0047	57.25	26	39	1	-1.26	7.16
6	GAILSGVYMGDKQSPL	1634.8338	-0.0044	115.26	35	38	1	0.2125	6.6
6	KPLFRGEGAAPFDPV	1599.8409	-0.0005	55.66	28	38	1	-0.253333333	6.92
6	AKPLFRGEGAAPFDPV	1670.878	-0.0037	59.47	26	38	1	-0.125	6.92
6	KPLFRGEGAAPFDPVVG	1755.9308	-0.0016	87.88	34	38	1	-8.48994E-17	6.92
6	AKPLFRGEGAAPFDPVVG	1826.9679	-0.0017	70.43	27	38	1	0.1	6.92
6	AKPLFRGEGAAPFDPVVGTI	2041.0997	-0.0069	46.58	22	37	1	0.28	6.92
6	DNYIQRRKDYELG	1668.8219	-0.0007	45.3	22	38	1	-1.930769231	6.88
6	HAGAMLGDVRHDPFQS	1752.8002	-0.0043	46.32	20	37	1	-0.51875	6.02
6	AQRFTGDMTQYGFTSQA	1907.8472	-0.0062	93.18	24	36	1	-0.729411765	6.7
6	AFSQPAIQQPSAQPQRPAFEQSPM	2640.2755	-0.0207	81.37	22	39	1	-0.841666667	6.98
6	AFSQPAIQQPSAQPQRPAFEQSPMSQSQPA	3238.5465	-0.0211	67.15	26	39	1	-0.953333333	6.98
6	AFSQPAIQQPSAQPQRPAFEQSPMSQSQPAQQAPA	3733.7907	-0.0182	81.51	23	40	1	-0.96	6.98
6	RPVEQPFWA	1128.5716	0.0028	54.8	34	38	1	-0.755555556	6.98
6	HLPLNPHLSFDLA	1472.7776	0.0012	106.2	29	38	1	0.184615385	6
6	GNDFFGQFRS	1173.5203	0.0037	36.48	20	36	1	-0.82	6.78
6	GNDFFGQFRSLLAGGQGT	1870.8962	-0.0053	138.07	36	38	1	-0.233333333	6.78
6	RQFAFSPGTSEPI	1435.7096	0.0022	106.06	34	38	1	-	6.98

								0.423076923	
6	SFEQTQRGGSTSQIGSAFETRPY	2533.1833	-0.0099	114.49		38	1	-1.043478261	7.04
6	ASFEQTQRGGSTSQIGSAFETRPY	2604.2205	-0.0182	41.39		38	1	-0.925	7.04
6	SFEQTQRGGSTSQIGSAFETRPYTT	2735.2787	-0.0116	55.06	18	38	1	-1.016	7.04
6	ISPDQRYGQNQQLQQAIGQ	2171.0719	-0.0106	57.65		39	1	-1.363157895	6.7
6	RVGGGTLADRLEFA	1460.7736	-0.0001	51.17	22	37	1	0.021428571	7.04
6	RVGGGTLADRLEFAA	1531.8107	0.0017	38.84	27	37	1	0.14	7.04
6	RLYEDTPYGKKPLTEV	1907.9993	-0.0022	70.37	29	38	1	-1.1625	6.88
6	RILGTGFPSFETV	1422.7507	0.0009	114.45		37	1	0.423076923	6.98
6	AGRPLDPRDVILDLKSY	1927.0527	-0.0049	44.76	18	37	1	-0.423529412	6.84
6	SNRSQFVPSWLVPE	1644.826	-0.004	89.57	32	39	1	-0.464285714	6.98
6	HWSQGSIGYFPT	1378.6306	0.0027	68.45	25	37	1	-0.525	7.76
6	LLETLYYRV	1168.6492	0.0013	41.83		35	1	0.477777778	6.82
6	HLGLVWMQGIAEPH	1586.8028	-0.0014	117.69	30	39	1	0.235714286	6.02
6	RLEDWRSYTPL	1434.7255	0.0007	40.61	26	38	1	-1.245454545	6.96
6	WFRAKEETTVLP	1475.7773	-0.0004	76.39	34	38	1	-0.558333333	7
6	TVGEPRFRTVDVRGDAEKV	2130.1182	-0.0003	72.12	31	38	1	-0.726315789	7.08
6	TIEDVSANDANPNYARRNIITK	2474.2513	-0.0213	57.9	18	39	1	-0.88181818	6.92

								2	
6	RLLSKELVIPEGV	1451.8711	0.0006	51.28		30	1	0.46923076 9	7
6	KGSSADVLVDQSDRIKVI	1929.0531	-0.0022	74.81	36	37	1	- 0.10555555 6	6.82
6	GGRQRLPPVDSWPDNV	1791.9016	-0.0033	100.44	35	39	1	-1.13125	6.92
6	DGRGGAAGGRQRLPPVDSWPDNV	2376.1683	-0.0003	56.36	24	39	1	- 1.03043478 3	7.02
6	HLALGFEPHGGISPG	1487.7521	-0.0016	106.95	31	39	1	0.08	6.02
6	HTWLPDAHVQAPTPA	1639.8107	-0.0055	92.03		38	1	- 0.47333333 3	6
6	GWYVTEIGRQPWVI	1702.8831	-0.006	77.43	30	39	1	- 0.02142857 1	6.88
6	DTAPQRRVDQALKFLADGA	2071.081	0.0003	84.72	27	38	1	- 0.59473684 2	6.9
6	FGGLLLI	731.4582	-0.0008	56.31		28	1	2.55714285 7	6.02
6	ASVPTFLTGAF	1109.5757	-0.0015	52	28	36	1	1.18181818 2	6.02
6	RLRSVYKDQIEGLPDLS	1988.0691	-0.0021	81.77	28	37	1	- 0.68823529 4	6.92
6	FGVALLI	731.4582	-0.0008	50.01		28	3	2.92857142 9	6.02
6	GLRDVVGWSVLAGVA	1497.8304	0.0013	76.65		36	1	1.14	6.78
6	VFNNEVFGLTKGQTSPT	1837.9211	-0.0008	48.82	36	38	1	- 0.31764705 9	6.84
6	GADRVSDGMRQAFGVPV	1760.8628	0.0051	42.13	27	39	1	- 0.12941176 5	6.92
6	GADRVSDGMRQAFGVPVGT	1918.932	-0.0045	47.15	20	39	1	- 0.17368421	6.92

								1	
6	SRTKTKLEDLGLVDTEKVPI	2241.258	-0.0092	74.65	30	35	1	-0.54	6.94
6	GLFADLRDSVSRV	1433.7627	-0.0001	69.51		37	1	0.2	6.92
6	DAGYDVDFLPGRRRIGSA	1963.9864	-0.0037	60.97	28	39	1	-0.555555556	6.92
6	RLPFERPETAT	1315.6884	0.0013	45.2	30	38	1	-1.109090909	7.12
6	FNLLLL	731.4581	-0.0008	39.9		28	4	2.416666667	6.02
6	FRALLL	731.4694	-0.012	32.39		28	7	1.916666667	11.04
6	LLVTQLVGLPGLDIGRAA	1805.0775	-0.0073	49.16		29	1	1.138888889	6.78
6	HELGHSLHSQYTSEHQPYVYS	2498.1251	-0.0006	45.3	25	37	1	-1.180952381	6.02
6	AFLYETQRL	1139.5975	0.0032	45.08	31	36	1	-0.144444444	6.88
6	LDLTKATIHTYMATL	1690.8964	0.0038	41.41		37	2	0.446666667	7.66
6	FARLLL	731.4694	-0.012	34.61		28	6	1.916666667	11.04
6	NYEDVGKSGFKRPEKTDRDVAVL	2622.3402	-0.0163	40.52	20	39	1	-1.126086957	6.94
6	VVANVFLAATGVDLGRA	1671.9308	-0.0105	39.86		36	2	1.258823529	6.78
6	GTWLAADLIGQLIRNT	1740.9523	-0.0065	39.67	32	37	1	0.36875	6.78
6	AGVVGVIAAITEVGLFAIAGGA	1955.1092	-0.0032	38.11		34	2	1.886363636	3.3
6	VLIFPWTVIYRDA	1591.8763	-0.0017	36.52	31	35	1	1.023076923	6.7
6	GVVDSIVRFA	1061.587	0.0035	38.73	32	35	1	1.25	6.78
6	FRALLL	731.4694	-0.012	32.39		28	7	1.916666667	11.04

7	FHLQPPQLF	1125.5971	-0.0044	47.7		36	1	-0.022222222	7.84
7	ANGARLYWKEDA	1392.6786	0.0001	72.24	29	38	1	-1.025	6.86
7	DYPRPRNFMQLVQPD	1874.9097	-0.0014	55.37	33	39	1	-1.326666667	6.84
7	GIDGKGRLSGISSTFNSLDPV	2119.091	0.0024	59.61	31	39	1	-0.114285714	6.8
7	GRPEWIWLA	1126.5923	-0.0033	46.02		36	1	-0.188888889	6.98
7	TGRPEWIWLA	1227.64	-0.002	58.1		37	1	-0.24	6.98
7	DVSAKVGFGLILL	1330.786	-0.0044	63.26		31	1	1.530769231	6.66
7	RYADWLFTTPL	1381.703	-0.0009	76.21	32	37	1	-0.090909091	6.7
7	GVSDPDAKKFYAI	1409.7191	0.0023	62.74	35	37	1	-0.292307692	6.68
7	TGRPEWIWLALGT	1498.7932	-0.0013	50.44	31	38	1	0.023076923	6.98
7	GVSDPDAKKFYAIT	1510.7668	0.0007	71.91		38	1	-0.321428571	6.68
7	PFGGEQNPIYWAR	1533.7365	-0.0002	46.21	31	38	1	-0.930769231	6.88
7	GADGIMIGTGLVGALTKV	1671.923	-0.0021	68.91	29	36	1	1.094444444	6.66
7	YWARYADWLFTTPL	1801.8828	0.0004	73.05	30	39	1	-0.1	6.64
7	GADGIMIGTGLVGALTKVY	1834.9863	-0.0043	76.35	28	37	1	0.968421053	6.6
7	LVGADGIMIGTGLVGALTKV	1884.0754	0.0002	50.42	28	33	1	1.385	6.66
7	LVGADGIMIGTGLVGALTKVY	2047.1388	0.0068	34.18	16	34	1	1.25714285	6.6

								7	
7	GGTEAQRNSWPSQI	1529.7223	0.0012	66.09		38	1	-1.27142857	6.98
7	VVGGTEAQRNSWPSQI	1727.8591	0.0009	90.02		39	1	-0.5875	6.98
7	GFVGGIETPLIKKFEA	1704.9451	0.0025	49.93	22	35	1	0.38125	6.92
7	GFVGGIETPLIKKFEAGFT	2010.0826	0.0024	91.75	25	37	1	0.41052636	6.92
7	GFVGGIETPLIKKFEAGFTA	2081.1197	0.0065	84.66	22	37	1	0.48	6.92
7	GFVGGIETPLIKKFEAGFTAGV	2237.2096	0.0038	86.16	27	37	1	0.60909090	6.92
7	GTVYGTGGLGDGSFNDQAKQGLERAK	2625.2783	0.0108	46.13	16	39	1	-0.82307692	6.86
7	GLHGHGWEPNA	1173.5316	-0.0073	38.2	26	36	1	-1.04545454	6.02
7	GLHGHGWEPNAS	1260.5636	-0.0036	75.93	35	36	1	-1.025	6.02
7	GLHGHGWEPNASI	1373.6476	-0.0004	40.48		38	1	-0.6	6.02
7	GLHGHGWEPNASIT	1474.6953	-0.0022	95.88	33	38	1	-0.60714285	6.02
7	AVFETRRPAEQLL	1528.8362	0.0001	73.19	29	37	1	-0.27692307	7.12
7	GLHGHGWEPNASITKEI	1844.9169	0.004	109.1	36	39	1	-0.67058823	6.02
7	ALAAGYAERGIGS	1234.6306	-0.004	43.55	35	38	1	0.32307692	6.88
7	ALAAGYAERGIGSA	1305.6677	0.0008	50.32	33	38	1	0.42857142	6.88
7	ALAAGYAERGIGSAAVGA	1603.8318	0	119.28		38	1	0.74444444	6.88
7	LAVGLAALAAGYAERGIGS	1758.9628	0.007	39.41	33	36	1	1.01052631	6.88
7	AALAVGLAALAAGYAERGIGS	1901.037	0.0017	107.75		36	1	1.08571428	6.88

7	AALAVGLAALAAGYAERGIGSA	1972.0741	0.0072	113.8		36	1	1.11818181 8	6.88
7	SHNVSAWPDMH	1279.5404	-0.0046	69.8	26	34	1	-0.87272727 3	6
7	SHNVSAWPDMHDKI	1635.7464	0.0002	94.94	25	37	1	-0.89285714 3	6.02
7	SHNVSAWPDMHDKI	1651.7413	0.0082	56.1	22	37	1	-0.89285714 3	6.02
7	YVSPGDTVKWVWEGSTGHNVHAT	2526.1928	0.006	151.14	35	39	1	-0.51304347 8	6.02
7	AKPLFRGEGAAPFDPV	1670.878	0.0008	70.53	28	38	1	-0.125	6.92
7	AKPLFRGEGAAPFDPVVG	1826.9679	0.0014	55.74	29	38	1	0.1	6.92
7	GTYDSFVGEGSAGRWLKA	1899.9115	0.0052	78.43	30	39	1	-0.39444444 4	6.86
7	VHGDAYGRHVEDTQPQKFA	2154.0243	0.0016	52.23	26	39	1	-1.15263157 9	6.02
7	FEHGLPSGPHLA	1260.6251	-0.0034	69.96	31	38	1	-0.20833333 3	6.02
7	PLVDNFLQNYHIGHVL	1877.9788	-0.017	42.22	31	38	1	0.2125	6
7	PLVDNFLQNYHIGHVLA	1949.0159	0	62.52	22	38	1	0.30588235 3	6
7	IGTGEPLARALEKI	1466.8456	-0.0041	43.57	27	33	1	0.12142857 1	7
8	FHLQPPQLF	1125.5971	-0.0066	63.39		36	1	-0.02222222 2	7.84
8	YHNLAQSFWYDRL	1711.8107	0.0012	83.27	31	38	1	-0.79230769 2	7.7
8	SKYDPSLHSSGETLPGGRKLGTYGPMENDKT	3321.5936	0.0164	45.71	17	40	1	-	7.54

								1.235483871	
8	SKYDPSLHSSGETLPGGRKLGTYGPMENDKTVNSS	3708.769	0.0302	75.18	19	40	1	-1.12	7.54
8	TGRPEWIWL	1156.6029	-0.0055	40.28		37	1	-0.466666667	6.98
8	TGRPEWIWLA	1227.64	-0.0037	51.45		38	1	-0.24	6.98
8	RYADWLFTTPL	1381.703	-0.003	73.67	32	37	1	-0.090909091	6.7
8	TGRPEWIWLALGT	1498.7932	-0.002	52.05	25	38	1	0.023076923	6.98
8	YWARYADWLFTTPL	1801.8828	0.0022	68.64	29	39	1	-0.1	6.64
8	AHCVDRELT	1042.4866	-0.0017	51.11	25	36	1	-0.344444444	5.22
8	TLNSYVQLGVLPRAG	1586.878	-0.0031	38.1	20	36	1	0.326666667	9.84
8	VVGGTEAQRNSWPSQI	1727.8591	-0.0008	82.57		39	1	-0.5875	6.98
8	AHNAFQYLGT	1120.5302	-0.0064	59.76	25	38	1	-0.24	7.76
8	WHAGGQSGTYYPL	1435.6521	-0.0043	108.86	31	37	1	-0.684615385	7.7
8	SWHAGGQSGTYYPL	1522.6841	-0.0055	82.69	27	36	1	-0.692857143	7.7
8	WHAGGQSGTYYPLSNEFK	2040.933	-0.0006	57.91	18	37	1	-0.988888889	7.62
8	WHAGGQSGTYYPLSNEFKT	2141.9807	-0.0018	149.89	30	37	1	-0.973684211	7.62
8	NVLYTFYPPLQAHPA	1729.8828	-0.0157	42.9		39	1	-0.006666667	7.7
8	NFPDDLGMFGHRKS	1619.7515	-0.0118	70.31	24	37	1	-1	7.72
8	AHRYDAPIVGSPFTI	1642.8467	-0.0024	70.35	26	38	1	0.13333333	7.76

								3	
8	NFPDDLGMFGHRKSVDRA	2060.985	-0.0001	61.93	30	39	1	-0.888888889	7.72
8	LFLGATLPPHLGQPA	1530.8558	-0.0027	45.19	22	35	1	0.573333333	7.84
8	HVPVEKHVFDDEGHGFSKRA	2290.1243	0.0019	41.48	22	39	1	-0.955	6.32
8	SAHVPVEKHVFDDEGHGFSKRA	2448.1935	-0.0034	114.22	25	39	1	-0.822727273	6.32
8	NAIGFGPIVGPHIA	1361.7456	-0.0027	106.93	33	36	1	0.928571429	7.84
8	MAGWYRTPEFGHGQLI	1861.8934	-0.0019	63.46	21	39	1	-0.35	7.76
8	VLHPRLPIDPA	1226.7135	-0.003	40.05	24	34	1	0.190909091	7.84
9	GAFHLQPPQLF	1253.6557	-0.0064	54.39	27	37	1	0.109090909	7.84
9	TLTFKEGVQFH	1305.6718	-0.0049	78.27		38	1	-0.209090909	7.72
9	ALIPVYHNLAQS	1324.7139	-0.0054	54.6	24	36	1	0.5	7.76
9	GRPEWIWLA	1126.5923	-0.0011	41.74		36	1	-0.188888889	6.98
9	TGRPEWIWLA	1227.64	-0.0023	54.23		37	1	-0.24	6.98
9	STFKVLRNVTVV	1361.8031	-0.0023	30.14	22	30	1	0.775	11.24
9	RYADWLFTTPL	1381.703	-0.0008	75.28	32	37	1	-0.090909091	6.7
9	TGRPEWIWLALGT	1498.7932	-0.0022	76.64		38	1	0.023076923	6.98
9	YWARYADWLFTTPL	1801.8828	-0.0007	68.1	26	39	1	-0.1	6.64
9	GCNVTRKPTVFT	1321.6813	-0.0046	53.65	32	38	1	-0.191666667	9.78
9	TLNSYVQLGVLPRAG	1586.878	-0.0016	98.63		36	1	0.32666666	9.84

								7	
9	GQNRGKDPHFWLDPN	1779.8441	0.005	79.39	30	38	1	-1.826666667	7.72
9	KWVWEGSTGH	1185.5567	-0.0061	59.67	35	36	1	-1.05	7.72
9	KWVWEGSTGHNV	1398.6681	-0.0023	105.22		37	1	-0.816666667	7.72
9	KWVWEGSTGHNVH	1535.727	-0.001	68.39	31	38	1	-1	7.88
9	KWVWEGSTGHNVHAT	1707.8118	0.0003	112.34	27	38	1	-0.793333333	7.88
9	GLTNHYYPMLV	1306.638	-0.0058	67.17	33	38	1	0.154545455	7.7
9	GLTNHYYPMLVF	1453.7064	-0.0029	82.86	26	38	1	0.375	7.7
9	LYTFYPPLQAHPA	1516.7714	-0.0018	70.95	23	39	1	-0.061538462	7.7
9	ANVLYTFYPPLQAHPA	1800.9199	0.0499	72.84		37	1	0.10625	7.7
9	LFLGATLPPHLGQPA	1530.8558	0.001	75.3	29	35	1	0.573333333	7.84
9	AFLRFWLEHAT	1389.7193	0	73	36	38	1	0.363636364	7.84
9	VHDVHPLDSMGGPVKLPRVW	2238.1732	0.0104	43.3	18	38	1	-0.14	7.88
9	VHDVHPLDSMGGPVKLPRVWA	2309.2103	0.0009	60.8	19	38	1	-0.047619048	7.88
9	GDSPPRWHAPLLGGLV	1670.8893	0.0003	49.86	28	38	1	-0.09375	7.84
9	DSPPRWHAPLLGGLVA	1684.9049	0.0006	68.34	25	37	1	0.04375	7.84
9	GDSPPRWHAPLLGGLVA	1741.9264	0.0021	50.99	25	37	1	0.017647059	7.84
9	DSPPRWHAPLLGGLVAT	1785.9526	0.0025	49.02	24	37	1	-7.83687E-17	7.84
9	GDSPPRWHAPLLGGLVAT	1842.9741	0.0035	74.5	22	38	1	-0.022222222	7.84
9	GYTHEEKPGGIIRWFTTV	2090.0585	0.0035	43.64	19	39	1	-	7.68

								0.53888889	
9	GYSVDSDILLNNHVLRRS	2057.0654	0.0064	101.2	21	38	1	-0.33333333	7.76
9	GWYRTPEFGHGQLI	1659.8158	0.0011	67.21	28	39	1	-0.66428571	7.76
9	GFALPPLPHLLA	1244.7281	-0.0048	62.68	30	33	1	1.1	7.84
9	AFSWFHGNLSGQA	1420.6524	-0.0003	60.77	30	37	1	-0.03846153	7.84
10	SAKVGFGLILL	1116.6907	-0.0045	39.11		29	1	1.74545454	9.74
10	GRPEWIWLA	1126.5923	-0.0037	47.53		36	1	-0.18888889	6.98
10	TGRPEWIWLA	1227.64	-0.0023	56.01		37	1	-0.24	6.98
10	STFKVLRNVTVV	1361.8031	0	60.39	29	30	1	0.775	11.24
10	RYADWLFTTPL	1381.703	-0.0006	77.23	32	37	1	-0.09090909	6.7
10	YWARYADWLFTTPL	1801.8828	0.0011	59	28	39	1	-0.1	6.64
10	SYVQLGVLPRAG	1258.7034	-0.0014	68.53	30	36	1	0.44166667	9.84
10	VVHPYWNTDDV	1343.6146	-0.0069	45.26	24	36	1	-0.50909090	3.88
10	NSYVQLGVLPRAG	1372.7463	0.0007	37.8	24	37	1	0.13846153	9.84
10	SYVQLGVLPRAGTI	1472.8351	0.0022	56.02		34	1	0.65	9.84
10	TLNSYVQLGVLPRAG	1586.878	-0.003	101.22		36	1	0.32666667	9.84
10	TLNSYVQLGVLPRAGTI	1801.0098	0.0029	108.86	33	34	1	0.51176470	9.84
10	TLNSYVQLGVLPRAGTILA	1985.131	0.0073	53.99	20	33	1	0.75263157	9.84

10	VVHPYWNTDDVAAGYDIA	2004.9218	-0.0007	126.4	36	38	1	-0.05	3.6
10	QKIVVHPYWNTDDVAAGYDIA	2374.1594	0.0099	42.1	21	39	1	-0.180952381	4.16
11	NREQFVSDVFKGRGV	1736.8958	0.001	97.11	34	38	1	-0.686666667	9.74
11	NREQFVSDVFKGRGVG	1793.9173	0.0033	59.79	25	38	1	-0.66875	9.74
11	DTYYVGFNMTKVPKPV	1857.9335	0.0011	66.06	25	39	1	-0.31875	9.24
11	NREQFVSDVFKGRGVGA	1864.9544	0.0036	88.64	28	38	1	-0.523529412	9.74
11	YVINREQFVSDVFKGRGV	2112.1116	0.0049	80.68	27	38	1	-0.161111111	9.52
11	SKYDPSLHSSGETLPGGRKL	2128.0913	0.0079	101.26	27	38	1	-1.035	9.34
11	STIDTYYVGFNMTKVPKPV	2159.0973	0.005	88.97	20	39	1	-0.110526316	9.24
11	KVGFGLILL	958.6215	-0.0016	28.26		21	1	2.022222222	9.74
11	SAKVGFGLILL	1116.6907	-0.0067	61.5		29	1	1.745454545	9.74
11	GRPEWIWLA	1126.5923	-0.0012	38.97		36	1	-0.188888889	6.98
11	TGRPEWIWLA	1227.64	0	40.42		37	1	-0.24	6.98
11	STFKVLRNVTVV	1361.8031	-0.0005	53.35	24	30	1	0.775	11.24
11	RYADWLFTTPL	1381.703	-0.0011	74.73	32	37	1	-0.090909091	6.7
11	TGRPEWIWLALGT	1498.7932	-0.0022	67.44	34	38	1	0.023076923	6.98
11	KGMGVSDPDAKKFYAI	1725.876	0.0057	47.08	30	38	1	-0.3875	9.2
11	YWARYADWLFTTPL	1801.8828	0.0036	57.73	26	39	1	-0.1	6.64

11	KGMGVSDPDAKKFYAIT	1826.9237	-0.0012	58.06	23	39	1	-0.405882353	9.2
11	QQYAPFLRNLA	1319.6986	-0.0035	43.09	29	37	1	-0.354545455	9.84
11	TFSNPRGYNPNPI	1475.7157	-0.0027	68.84	27	38	1	-1.207692308	9.84
11	RQNWVMTA	1004.4862	-0.0028	47.15	29	38	1	-0.65	11.04
11	QRNSWPSQI	1114.5519	-0.0039	48.65	32	38	1	-1.622222222	11.04
11	YVQLGVLPRAG	1171.6713	-0.0037	56.06		35	1	0.554545455	9.84
11	SLQYRSGSSWA	1240.5836	-0.0046	54.35	30	37	1	-0.745454545	9.84
11	SYVQLGVLPRAG	1258.7034	-0.0012	97.55		36	1	0.441666667	9.84
11	NSYVQLGVLPRAG	1372.7463	-0.0013	88.15	33	37	1	0.138461538	9.84
11	YVQLGVLPRAGTI	1385.8031	-0.0011	44.2		33	1	0.761538462	9.84
11	TLNSYVQLGVLPR	1458.8195	0.0001	78.84	26	35	1	0.269230769	9.84
11	SYVQLGVLPRAGTI	1472.8351	-0.0004	83.76		34	1	0.65	9.84
11	TLNSYVQLGVLPRAG	1586.878	-0.002	109.75		36	1	0.326666667	9.84
11	TLNSYVQLGVLPRAGTI	1801.0098	0.0018	103.95	30	34	1	0.511764706	9.84
11	QSVTLNSYVQLGVLPRAG	1901.0371	0.004	92.05	32	36	1	0.266666667	9.84
11	VVHPYWNTDDVAAGYDIA	2004.9218	0.0003	54.96	21	38	1	-0.05	3.6
11	SLFGVPYNFERPKLS	1752.9199	0.0032	86.26	35	37	1	-0.3	9.52
11	SGSLFGVPYNFERPKLS	1896.9734	0.0043	85.32		38	1	-0.335294118	9.52

11	FKRVMNEGKENFLPI	1820.9607	0.0001	62.96	28	37	1	-0.553333333	9.46
11	FKRVMNEGKENFLPIVT	2021.0768	0.0028	53.02	16	37	1	-0.282352941	9.46
11	NLGSDRGPYYPKRPDGI	1903.954	0.0068	70.08	31	39	1	-1.441176471	9.38
11	LFRAGGLLL	958.5964	0.0236	23.61		21	2	1.611111111	11.04
12	GWVADYPRPR	1215.6149	-0.0017	47.15		38	1	-1.23	9.84
12	GWVADYPRPRN	1329.6578	-0.0019	43.57		39	1	-1.436363636	9.84
12	DYPRPRNFMQL	1435.703	-0.0036	46.97	29	38	1	-1.409090909	9.84
12	LHSSGETLPGGRKL	1450.7892	0.0004	52.83	26	35	1	-0.628571429	9.74
12	GWVADYPRPRNF	1476.7262	-0.0006	41.69	32	39	1	-1.083333333	9.84
12	NREQFVSDVFKGRGV	1736.8958	0.0017	103.39	34	38	1	-0.686666667	9.74
12	NREQFVSDVFKGRGVG	1793.9173	0.0041	84.21	28	38	1	-0.66875	9.74
12	DTYYVGFNMTKVPKPV	1857.9335	-0.0021	109.96	30	39	1	-0.31875	9.24
12	NREQFVSDVFKGRGVGA	1864.9544	0.0017	93.7	30	38	1	-0.523529412	9.74
12	DTYYVGFNMTKVPKPV	1873.9284	0.0049	70.28	25	39	1	-0.31875	9.24
12	VINREQFVSDVFKGRGV	1949.0483	-0.0009	74.89	33	37	1	-0.094117647	9.74
12	KYDPSLHSSGETLPGGRKL	2041.0592	0.005	117.52	34	38	1	-	9.34

								1.047368421	
12	YVINREQFVSDVFKGRGV	2112.1116	0.0081	95.89	29	38	1	-0.161111111	9.52
12	SKYDPSLHSSGETLPGGRKL	2128.0913	0.0055	121.56	33	39	1	-1.035	9.34
12	STIDTYYVGFNMTKVPKPV	2159.0973	0.0048	100.01	30	39	1	-0.110526316	9.24
12	YVINREQFVSDVFKGRGVG	2169.1331	0.0027	58.76	23	38	1	-0.173684211	9.52
12	STIDTYYVGFNMTKVPKPV	2175.0922	-0.0006	54.82	24	39	1	-0.110526316	9.24
12	MAYVINREQFVSDVFKGRGV	2314.1892	0.0056	48.35	19	39	1	0.04	9.52
12	GFGLILLR	887.5593	-0.0011	60.17		28	1	1.675	11.04
12	KVGFGLILL	958.6215	-0.0028	62.67		21	1	2.022222222	9.74
12	RSRAIFGEA	1005.5356	-0.005	48.41		36	1	-0.311111111	10.88
12	AKVGFGLILL	1029.6586	-0.0062	59.13		25	1	2	9.74
12	SAKVGFGLILL	1116.6907	-0.0046	79.9		29	1	1.745454545	9.74
12	GTGLVGALTKVY	1177.6707	-0.0075	52.3		33	1	0.833333333	9.52
12	TGRPEWIWLA	1227.64	-0.0014	47.53		37	1	-0.24	6.98
12	STFKVLRNVTVV	1361.8031	-0.0006	56.55	24	30	1	0.775	11.24
12	RYADWLFTTPL	1381.703	-0.0023	75.57	32	37	1	-0.090909091	6.7
12	MIGTGLVGALTKVY	1421.7952	-0.0044	39.57	21	34	1	1.171428571	9.52
12	TGRPEWIWLALGT	1498.7932	-0.0026	46.54	36	38	1	0.023076923	6.98
12	KGMGVSDPDAKKFYAI	1725.876	0.0059	71.61	32	38	1	-0.3875	9.2

12	YWARYADWLFTTPL	1801.8828	0.0024	59.75	25	39	1	-0.1	6.64
12	KGMGVSDPDAKKFYAIT	1826.9237	0.0014	73.56	26	38	1	-0.405882353	9.2
12	KGMGVSDPDAKKFYAIT	1842.9186	0.0051	42.62	19	39	1	-0.405882353	9.2
12	GYMAFNHDRKPA	1405.6561	-0.0007	71.79	30	38	1	-1.133333333	9.52
12	TFSNPRGYNPNPI	1475.7157	-0.0003	69.32	27	38	1	-1.207692308	9.84
12	TGWGLTRT	890.461	-0.0019	42.78		37	1	-0.5625	11.04
12	RQNWVMT	933.4491	-0.0031	43.54		38	1	-1	11.04
12	RQNWVMTA	1004.4862	-0.0047	59.75		38	1	-0.65	11.04
12	YVQLGVLPR	1043.6128	-0.0062	38.31		34	1	0.522222222	9.84
12	QRNSWPSQI	1114.5519	-0.0018	47.58	36	38	1	-1.622222222	11.04
12	QLGVLPRAGTI	1123.6713	-0.0049	31.98		31	1	0.636363636	11.04
12	SYVQLGVLPR	1130.6448	-0.0009	42.4	25	34	1	0.39	9.84
12	YVQLGVLPRAG	1171.6713	-0.0032	74.78		35	1	0.554545455	9.84
12	SLQYRSGSSWA	1240.5836	-0.0036	62.5	28	37	1	-0.745454545	9.84
12	SYVQLGVLPRAG	1258.7034	0.0011	93.25		36	1	0.441666667	9.84
12	QKIVVHPYWN	1282.6822	-0.0037	64.04	25	37	1	-0.5	9.52
12	QLGVLPRAGTILA	1307.7925	-0.0052	36.65		29	1	0.969230769	11.04
12	NSYVQLGVLPRAG	1372.7463	-0.0026	84.63	31	37	1	0.138461538	9.84
12	YVQLGVLPRAGTI	1385.8031	-0.0024	87.41		33	1	0.761538462	9.84

12	TLNSYVQLGVLPR	1458.8195	-0.0047	106.54	31	35	1	0.269230769	9.84
12	SYVQLGVLPRAGTI	1472.8351	-0.0023	90.68		34	1	0.65	9.84
12	TLNSYVQLGVLPRAG	1586.878	0.0083	118.94		35	1	0.326666667	9.84
12	TLNSYVQLGVLPRAGTI	1801.0098	0.0012	135.77	33	34	1	0.511764706	9.84
12	QSVTLNSYVQLGVLPRAG	1901.0371	0	109.99	32	36	1	0.266666667	9.84
12	GFVGGIETPLIKKF	1504.8654	-0.0021	75.39	28	32	1	0.557142857	9.46
12	GQNRGKDPHFWLDPNRA	2006.9823	0.0037	44.48	24	39	1	-1.770588235	9.74
12	LFGVPYNFERPKLS	1665.8879	0.0014	76	30	37	1	-0.264285714	9.52
12	SLFGVPYNFERPKLS	1752.9199	0.0035	91.39	31	37	1	-0.3	9.52
12	SGSLFGVPYNFERPKLS	1896.9734	0.0027	97.93	37	38	1	-0.335294118	9.52
12	VEKPTSGSLFGVPYNFERPKLS	2451.2798	0.002	90.13	22	38	1	-0.509090909	9.34
12	GHLRSLLPI	1004.6131	-0.0018	30.59		28	1	0.6	11.04
12	WYTILPKLA	1103.6379	-0.013	40.26	23	34	1	0.611111111	9.52
12	FKRVMNEGKENFLPI	1820.9607	-0.0006	88.01	30	37	1	-0.553333333	9.46
12	FKRVMNEGKENFLPIVT	2021.0768	-0.004	68.94	22	37	1	-0.282352941	9.46
12	DQPKRDQNGRPFVPV	1751.9067	0.0032	80.6	29	38	1	-1.626666667	9.74
12	QQSIAQQQPMARPQPV	1805.9206	-0.0003	76.01	26	38	1	-1.05625	11.04
12	GIRVGPGEKHTYTIPA	1694.9104	0.0021	98.18	29	37	1	-0.45	9.52

12	GLKYFLI	852.5109	-0.0061	30.56	27	30	1	1.328571429	9.52
12	AVRLYEDTPYGKKPL	1748.9461	0.0031	45.32	21	36	1	-0.84	9.24
12	AAVRLYEDTPYGKKPL	1819.9832	0.003	55.41	20	36	1	-0.675	9.24
12	LGSDRGPYYPKRPDGI	1789.9111	0.0016	55.48	32	39	1	-1.3125	9.38
12	NLGSDRGPYYPKRPDGI	1903.954	0.0032	61.9	30	39	1	-1.441176471	9.38
12	ANLGSDRGPYYPKRPDGI	1974.9912	0.0033	43.14	15	39	1	-1.261111111	9.38
12	QSRLWENHIGRSPA	1649.8386	0.0024	40.26	21	38	1	-1.221428571	10.88
12	NIWQFPFKT	1179.6077	-0.0072	43.21	28	37	1	-0.444444444	9.74
12	GLGNIWQFPFKT	1406.7347	-0.0012	74.28	26	37	1	-0.083333333	9.74
12	ELQWQRKQDWRKEGAIERV	2454.288	0.0096	44.84	19	38	1	-1.773684211	9.46
12	DLVQGAPRLKYL	1371.7874	0.0012	40.77	26	34	1	-0.108333333	9.52
12	DLVQGAPRLKYLLG	1541.8929	0.0039	79.65		33	1	0.15	9.52
12	YVFDDTVLSQVGRPKRA	1950.0323	0.0026	81.19	28	37	1	-0.423529412	9.52
12	KVVLHPRLPIDPA	1453.8769	-0.0021	51.94	23	30	1	0.184615385	9.74
12	FLGGLLRA	845.5123	0.0003	35.23		30	1	1.3375	11.04
12	LIGVGYIIGPRIA	1340.818	-0.0011	60.83		30	1	1.476923077	9.84
12	GNGLFRYGVFTL	1342.7034	-0.0012	50.94	32	38	1	0.516666667	9.84
12	RIEAEGVFDGKKRSLSQPV	2115.1436	0.009	37.61	20	36	1	-	9.46

								0.710526316	
12	LFRAGGLLL	958.5964	0.0224	32.03		21	2	1.611111111	11.04
12	RLPIVLVVPT	1105.7223	-0.0032	43.43		19	1	1.63	11.04
12	VTSLLLSVG	887.5328	0.0254	42.1		28	2	1.9	6.02
12	FMSAVGPRGIIPA	1314.7118	-0.0025	41.38	27	37	1	0.938461538	11.04
12	SAKRRDYEWQGKTGDN	1909.9031	0.0036	40.21	25	38	1	-2.3125	9.34
	CGM								
1	HSIAAQGGVNSA	1110.5418	0.0024	71.29	33	35	1	0.125	7.84
1	SVDGWFTLPFTIPNYLGPLLGSERLSEDAPEAQ	3618.7882	-0.0069	94.94	20	39	1	-0.172727273	3.5
1	SFEDELSGRSINQPSAV	1834.8697	-0.0186	45.19	28	35	1	-0.594117647	3.82
1	SDSLQGADVLTYDPERA	1835.8537	-0.0054	57.85	31	35	1	-0.717647059	3.5
1	FSSFEDELSGRSINQPS	1898.8646	-0.0003	55.22	30	34	1	-0.829411765	3.82
1	FSSFEDELSGRSINQPSAV	2068.9702	-0.0162	104.39		35	1	-0.426315789	3.82
1	DDTTFTVELTQPESDFPLRL	2323.122	-0.0092	55.62	23	37	1	-0.555	3.42
1	NGTEPQNPLVPGNTNEVGGGRIVDSI	2633.3045	-0.0537	41.05	18	37	1	-0.619230769	3.82
1	SDMVESTPEGYRATTL	1755.7985	-0.0108	41.07	28	34	1	-0.64375	3.82
1	VAPDVDIEAIITGGDIDPH	1945.9633	-0.016	113.08		37	1	0.242105263	3.36
1	DHDVDNEHVWYSTEYV	2006.8283	0.0031	32.16	19	28	1	-1.2375	3.84

1	NTENFLDAFTKAVDDLTAAT	2156.0273	0.0054	119.41		37	1	-0.155	3.5
1	NTENFLDAFTKAVDDLTAATN	2270.0703	0.004	39.45	29	36	1	-0.314285714	3.5
1	GGGGYDSWLYGTLEDDDRIIH	2338.0502	-0.0167	109.18	26	33	1	-0.742857143	3.76
1	GGGGYDSWLYGTLEDDDRIIHAL	2522.1714	-0.0002	118.47	32	36	1	-0.434782609	3.76
1	VVHPYWNTDDVAAGYDIA	2004.9218	-0.0273	75.38	26	33	1	-0.05	3.6
1	KQLTDDIPPLTTEDGEVRGVPF	2426.2329	0.0071	65.18	24	38	1	-0.627272727	3.84
1	TDDIPPLTTEDGEVRGVPFAVEGF	2560.2334	-0.0226	41.47	24	36	1	-0.220833333	3.38
1	NAPIWQEYQDKGVEDTNEIEFS	2611.1714	0.0144	48.86	24	34	1	-1.218181818	3.44
1	KQLTDDIPPLTTEDGEVRGVPFAVEGF	2929.4709	-0.0573	104.04	33	37	1	-0.32962963	3.76
1	YEGDFRVVEMEKDGEPF	2045.9041	-0.0104	41.56	24	32	1	-1.011764706	3.82
1	EYVVNSIADDKGWDHPTIEWRESPS	2929.3519	-0.0127	79.9	25	35	1	-1.068	4.06
1	EYVVNSIADDKGWDHPTIEWRESPSAQ	3128.4475	-0.0083	99.29	31	35	1	-1.051851852	4.06
1	TGEDIFGEEHERDNENTPA	2158.9039	0.0034	33.32	19	29	1	-1.784210526	3.74
1	GEDWDADQPVTGEDIFGEEH	2244.9084	0.0202	55.64	20	27	1	-1.29	3.24
1	EDIFGEEHERDNENTPAWA	2257.9512	0.0071	33.16	23	29	1	-1.678947368	3.74
1	TGEDIFGEEHERDNENTPAWA	2416.0203	-0.0061	79.32	23	28	1	-1.57142857	3.74

								1	
1	DQPVTGEDIFGEEHERDNENTPA	2598.1106	0.0193	38.86	17	31	1	-1.665217391	3.62
1	GEDWDADQPVTGEDIFGEEHERDNENTPA	3271.345	0.012	153.74		26	1	-1.665517241	3.42
1	VGEDWDADQPVTGEDIFGEEHERDNENTPA	3370.4134	-0.0081	141.52	23	26	1	-1.47	3.42
1	AVGEDWDADQPVTGEDIFGEEHERDNENTPA	3441.4505	-0.0017	114.39	20	27	1	-1.364516129	3.42
1	AAVGEDWDADQPVTGEDIFGEEHERDNENTPA	3512.4876	0.0007	93.92	25	27	1	-1.265625	3.42
1	DGIDGESVAIPNDPSNQGRAI	2124.0083	0.0117	48.8	33	37	1	-0.652380952	3.5
1	DRPGADLQEILDRLEADQENA	2367.1302	-0.0223	39.35	22	36	1	-1.176190476	3.62
1	SVPVDRPGADLQEILDRLEADQENA	2749.3518	-0.01	60.47	28	38	1	-0.748	3.62
1	RFQVKDQSIVDQQEIDSDPS	2333.1135	0.0132	84.12	35	37	1	-1.185	3.76
1	QAPDSYGRIGNQSGSDVSPV	2032.945	0	39.45	24	35	1	-0.81	3.88
1	NTYPVIDAWTEPEVWQEARPT	2501.1863	-0.0186	153.72		36	1	-0.904761905	3.68
1	DFLRMGWPDGITPE	1632.7606	-0.0266	40.82	28	34	1	-0.542857143	3.7
1	DVSYLEFNRIETLGTTDEIPV	2410.1904	-0.0295	41.67	28	37	1	-0.185714286	3.5
1	VEPDWTYDQATLTGTWGNLTDSQLRSV	3052.4414	-0.0385	144.76		35	1	-0.666666667	3.5
1	GFRPEALEIIPEGESTDLSIPI	2382.2318	-0.0133	70.22		38	1	0.013636364	3.58
1	SESYEDVLTELFERAQVTA	2186.0379	-0.0264	37.96	31	35	1	-0.389473684	3.58

1	TAEQSVEDTKQLHDQALQQADQA	2553.1943	0.0247	55.34	32	37	1	-1.252173913	3.84
1	EFRAPDPSGNPYLGFAA	1807.8529	-0.0122	40.18	29	35	1	-0.464705882	4.08
1	FENPDQKETEDYISGRFG	2130.9494	-0.0097	103.44	27	32	1	-1.555555556	3.92
1	DATVSAPVLESGVENREVGGDTA	2272.0819	-0.0077	78.39	31	36	1	-0.230434783	3.5
1	DATVSAPVLESGVENREVGGDTAEA	2472.1616	-0.0024	77.28	28	36	1	-0.28	3.44
1	SFSVEDFEEFLHQQ	1740.7631	-0.0172	31.93	23	31	1	-0.671428571	3.68
1	ALGFLDYLDEEQRA	1638.7889	-0.0163	38.14	27	36	1	-0.421428571	3.58
1	FFDQETNERWIPFVIEPA	2237.0793	-0.0275	53.54	33	36	1	-0.383333333	3.68
1	EDFGWRIPFLT	1379.6874	-0.0021	49.73	33	36	1	-0.109090909	4.08
1	GVLEEGQYNRELAEAI	1789.8846	-0.034	55.41		36	1	-0.48125	3.8
1	AGVLEEGQYNRELAEAI	1860.9217	-0.0331	64.33	31	36	1	-0.347058824	3.8
1	RLELPDFEFVPM	1491.7432	-0.0119	55.51	27	37	1	0.091666667	3.82
1	FAEFERTTAEDVANGVPEAI	2165.0277	-0.0112	76.2	25	36	1	-0.16	3.58
1	DDSLSQEVFGVTFPRPL	1905.9473	-0.0302	78.43		37	1	-0.164705882	3.7
1	AELFDASGQIPSSQEFFRV	2127.0273	-0.0246	56.88	24	36	1	-0.1	3.82
1	SASDDSDALDWVRELTSYLPS	2326.0601	-0.0039	56.27		34	1	-0.53809523	3.36

								8	
1	LRDLLEGTDERIQDLEFFQDAV	2621.2973	-0.0053	41.65	26	38	1	-0.468181818	3.62
1	SLDAVRQEVEEIIGQGSQPT	2155.0757	-0.0322	58.01		37	1	-0.535	3.68
1	AFQDALGVRDADFIAQLV	1948.0054	-0.0295	56.89	33	37	1	0.605555556	3.6
1	LFDQPEDDQPNDGFDGRT	2064.8661	0.0233	50.21	27	31	1	-1.738888889	3.26
1	INDEVINEIRDGISSF	1819.8952	-0.006	37.43	31	37	1	-0.15625	3.58
1	DWADDDSMNSYFLANPRMQDYI	2666.1053	-0.0154	37.26		25	1	-0.940909091	3.3
2	HAEPLFFESVPL	1384.7027	0.0092	62.75	31	36	1	0.416666667	4.24
2	ETTRPEFIHPVSVLPEVS	2036.0579	-0.0489	38.22	27	37	1	-0.127777778	4.54
2	SVDGWFTLPFTIPNYLGPLLGSERL	2791.4585	-0.0036	88.21	28	38	1	0.284	4.08
2	TQPESDFPLRL	1301.6616	0.0025	39.47		37	1	-0.845454545	4.08
2	FEGTRTPATDFTSPVIDGH	2046.9647	-0.0219	88.26	26	35	1	-0.510526316	4.3
2	FSSFEDELSGRSINQPSAV	2068.9702	-0.0572	39.89	23	33	1	-0.426315789	3.82
2	IFEGTRTPATDFTSPVIDGH	2160.0488	-0.009	64.93	26	37	1	-0.26	4.3
2	NTNEVGGGRIVDSIFSGLVYY	2259.1172	-0.0673	46.76	30	35	1	0.152380952	4.08
2	NGTEPQNPLVPGNTNEVGGGRIVDSI	2633.3045	-0.0782	54.18	23	36	1	-0.619230769	3.82
2	ELEGDKTYRITIKDGQTFTDGTPV	2683.3341	-0.0118	48.28	27	38	1	-0.87916666	4.2

								7	
2	SYNADGGHQAWVDATANSIRNTLGIDAI	2929.3954	-0.045	44.99	18	35	1	-0.289285714	4.16
2	VVHPYWNTDDV	1343.6146	-0.012	55.07	31	33	1	-0.509090909	3.88
2	VVHPYWNTDDVAAGY	1705.7737	-0.0038	44.85	19	34	1	-0.246666667	3.88
2	FGPAARALPQLPITVDEEGYLIA	2440.3002	-0.0409	80.48	35	38	1	0.369565217	3.82
2	DAEKVIEREGQESFHYGDYYAYSKI	2996.3828	-0.072	44.97	22	31	1	-1.12	4.42
2	ADDKGWDHPTIEWRESPS	2124.9501	-0.0179	60.15	23	32	1	-1.694444444	4.2
2	ADDKGWDHPTIEWRESPSAQ	2324.0458	-0.0187	100.05	27	33	1	-1.61	4.2
2	SIADDKGWDHPTIEWRESPS	2325.0662	-0.0171	43.46	22	34	1	-1.34	4.2
2	NSIADDKGWDHPTIEWRESPS	2439.1091	-0.0058	50.88	24	34	1	-1.442857143	4.2
2	SIADDKGWDHPTIEWRESPSAQ	2524.1619	-0.0036	71.11	33	35	1	-1.295454545	4.2
2	NSIADDKGWDHPTIEWRESPSAQ	2638.2048	-0.0041	69.9	21	34	1	-1.391304348	4.2
2	TKYDQPGLGLRNPDNSMSGLDVDVA	2661.2704	-0.0687	55.58	21	34	1	-0.696	3.88
2	NNSFLDRAGIDPN	1431.6742	-0.011	40.18	32	35	1	-0.915384615	3.88
2	NNSFLDRAGIDPNLA	1615.7954	-0.039	37.39	29	35	1	-0.42	3.88
2	SVIPVDAAQAPTAYQEGRPA	2040.0276	-0.012	57	24	37	1	-0.23	4.08
2	SVIPVDAAQAPTAYQEGRPAII	2266.1957	-0.0536	78.95	34	37	1	0.2	4.08
2	RIGTTDAAKEAWTVFEDKAAEEGITLDIVPF	3392.714	-0.0082	60.36	25	39	1	-0.083870968	3.96

2	SLGEFTRVFEYDSLVV	1859.9305	-0.02	87.27	30	37	1	0.425	3.82
2	AGASSPWGDDKEFPVSAEETGYVHPYTRI	3165.468	-0.0819	54.68	16	31	1	-0.73448275	4.28
2	VLGDYKVDKDADLERPERV	2216.1437	-0.0004	71.82	36	38	1	-1.005263158	4.3
2	SGLWSPQDDGVRV	1414.6841	-0.0109	74.37		35	1	-0.59230769	3.88
2	ALDRVMEQGGYVTI	1550.7763	-0.0135	62.5		37	1	0.185714286	4.08
2	SILELLDHVNNKFIEEGKEPFAFA	2759.417	-0.0156	72.28	24	38	1	-0.07083333	4.36
2	IDRWNEAHPDEQVTL	1821.8646	-0.0241	50.34	31	35	1	-1.173333333	4.06
2	NGQLLFRNTEIIPEAPA	1881.9948	-0.0415	85.79	35	37	1	-0.194117647	4.26
2	NTNGQLLFRNTEIIPEAPA	2097.0854	-0.0483	49.59	27	37	1	-0.394736842	4.26
2	SVEDLRAIPWVL	1396.7714	-0.0003	52.95		35	1	0.625	4.08
2	SSVEDLRAIPWVL	1483.8035	-0.0053	47.02	33	36	1	0.515384615	4.08
2	ATDETPNYYPGGNVAGRPVPA	2145.0127	-0.0044	122.33	30	36	1	-0.776190476	4.08
2	TASPVATDETPNYYPGGNVAGRPVPA	2600.2507	-0.0034	88.43	28	37	1	-0.515384615	4.08
2	SNSVEGRMAEVEQAGVQ	1789.8265	-0.0069	47.24	30	35	1	-0.576470588	3.96
2	LFDSKIEENADPKVQIQQAIEDAQRQH	3149.5741	-0.0926	44.12	25	36	1	-1.118518519	4.36

2	AIMNAGEIAVEGTLDELVAREKSII	2641.3996	-0.0336	48.51	31	38	1	0.456	4
2	SIEAELLTEIRDLLQEQKRLQ	2524.386	-0.0159	56.72	29	36	1	-0.6	4.36
2	YLQLSYVLDDLGLTGMRDWMKA	2587.2814	0.0055	56.86	27	38	1	0.027272727	4.18
2	QISRTLFDLLFPSNTDAA	2008.0266	-0.0242	57.86	27	37	1	0.111111111	3.88
2	AQISRTLFDLLFPSNTDAA	2079.0637	-0.0226	49.45	27	37	1	0.2	3.88
2	FLGPVDDEEGLHYLRPE	1984.953	-0.045	68.91	24	34	1	-0.711764706	3.92
2	LVELFGLAREL	1258.7285	0.007	60.35		32	1	1.1	4.26
2	SYNEDTKQWDVTVNRDGAESTLHPT	2862.3056	-0.0019	50.43	21	34	1	-1.368	4.2
2	AMEDFGWRIPFLT	1581.765	-0.0095	46.86	32	36	1	0.192307692	4.08
2	VATFRDLENKYMPED	1826.8509	-0.0198	53.85	24	34	1	-1	4.06
2	SIIDHIVEDVVAGTLTDRPVLV	2360.2952	-0.015	73.88	25	36	1	0.818181818	3.96
2	VFTQEQIDQLKELKSRDQEAA	2475.2605	-0.0222	72.44	36	38	1	-1.076190476	4.28
2	QAPTVLERQPVEEPL	1704.9046	-0.0227	48.03	29	37	1	-0.646666667	3.96
2	GFLTPEEVESFLAGRV	1749.8938	-0.0133	57.77	33	37	1	0.28125	3.96
2	FAFDSIDKAPEEKERGITI	2165.1004	-0.0457	45.63	36	37	1	-0.557894737	4.28
2	MWDERYPGGAPV	1376.6183	-0.0123	46.45	24	33	1	-0.816666667	4.08
2	DLGDGLAQLLKLPERFVIGA	2124.1943	-0.0136	45.71	32	35	1	0.445	4.3
2	LVDGPPGSDFGVLPEFKRT	2030.0473	-0.0528	41.24	24	37	1	-0.252631579	4.3
2	GFVRDVEERLNELEAA	1845.9221	-0.0174	40.06	33	37	1	-0.5	4
2	TLVIIPTYNELENLPLIVDRV	2423.3675	-0.0284	37.29	28	35	1	0.638095238	3.82

3	SAWEPGENGTFVRHAEPL	1995.9439	-0.0175	43.01	20	35	1	-0.761111111	4.54
3	SFTIPESLEHFSGEEGVLRRQ	2417.1975	-0.0107	61.84	27	37	1	-0.63333333	4.7
3	FQSFTIPESLEHFSGEEGVLRRQ	2692.3245	-0.0769	82.64	29	36	1	-0.608695652	4.7
3	FTIPESLEHFSGEEGVLRRQAISLA	2785.4399	-0.0533	44.6	30	38	1	-0.056	4.7
3	SFTIPESLEHFSGEEGVLRRQAISLA	2872.4719	-0.0571	84.98	33	38	1	-0.084615385	4.7
3	FQSFTIPESLEHFSGEEGVLRRQAISLA	3147.5989	-0.0713	78.94	26	38	1	-0.103571429	4.7
3	IEREGQESFHYGDYYAYSKI	2454.1128	-0.0158	51.75	23	33	1	-1.155	4.62
3	EMVNDGAEYDPAKDVYEHQMHSVHGPRNA	3295.4411	-0.019	66.2	20	30	1	-1.210344828	4.72
3	KWEDVVKHFNDNWAAEKES	2331.092	-0.0085	82.37	25	35	1	-1.431578947	4.7
3	KWEDVVKHFNDNWAAEKESNWG	2688.2357	0.0151	37.9	22	36	1	-1.454545455	4.7
3	TSYQKLKEVAEDMQAKKDELGIEGA	2780.3902	-0.0161	44.97	22	38	1	-0.944	4.44
3	ATKDFVDWLFTSEAGKEHVVKDLGFIAPF	3266.6652	-0.035	46.99	17	38	1	0.117241379	4.62
3	RLVELWHDPEVL	1504.8038	-0.0252	56.96	32	37	1	-0.075	4.42
3	NYYQEILKLDGMKHFADGEAT	2442.1525	-0.004	53.59	22	36	1	-0.747619048	4.54
3	HHGVAVVLENGDLLHTLGDEDSRNGAVV	2922.4584	-0.0209	57.98	18	38	1	-0.067857143	4.58
3	NNKFIEEGKEPFA	1521.7463	0.0015	36.72	32	36	1	-	4.56

								1.184615385	
3	AYPVIKDMVVDRSALDRVMEQGGYVTI	3024.5413	-0.0251	46.38	27	38	1	0.159259259	4.46
3	AAYPVIKDMVVDRSALDRVMEQGGYVTI	3095.5784	-0.0163	41.02	24	38	1	0.217857143	4.46
3	GGFFGPDPEDIRAKAKEIEHA	2283.1284	-0.0169	44.49	23	37	1	-0.828571429	4.7
3	NGQLLFRNTEIIPEAPA	1881.9948	-0.0139	58.61	32	37	1	-0.194117647	4.26
3	NTNGQLLFRNTEIIPEAPA	2097.0854	-0.0165	46.58	30	37	1	-0.394736842	4.26
3	TYNGTLYALPQNTNGQLLFRNTEIIPEAPA	3318.6884	-0.0993	75.13	25	37	1	-0.37	4.26
3	GHDIYNAVIRNGQSTEAEQV	2200.0509	-0.0063	64.5	30	36	1	-0.74	4.42
3	VGHDIYNAVIRNGQSTEAEQV	2299.1193	-0.0184	62.68	23	37	1	-0.504761905	4.42
3	AVGHDIYNAVIRNGQSTEAEQV	2370.1564	-0.0116	54.82	21	37	1	-0.4	4.42
3	SAAVGHDIYNAVIRNGQSTEAEQV	2528.2255	-0.0231	68.55	23	37	1	-0.325	4.42
3	LRPTPQEFELY	1391.7085	-0.0007	40.55	32	37	1	-0.890909091	4.26
3	SFVLRGGQVELPVLSGAEII	2083.1677	-0.0302	46.48	33	36	1	0.91	4.26
3	SIEAELLTEIRDLLQEQKRLQ	2524.386	-0.0127	78.86	34	36	1	-0.6	4.36
3	SFSVEDFEEFLHQQAESALRHV	2604.2245	0.0005	45.75	16	36	1	-0.445454545	4.34
3	ALRVLESLFNEHAYELV	2002.0523	-0.024	88.71	27	37	1	0.364705882	4.54
3	TALRVLESLFNEHAYELV	2103.1	-0.0165	41.08	26	37	1	0.305555556	4.54
3	NHDLTQEEIAQLVGASRETVNKA	2522.2725	-0.0192	72.9	23	38	1	-0.708695652	4.62

3	SLIEREAPAGDFDKVARVI	2085.1218	-0.0134	52.04	30	37	1	-0.015789474	4.56
3	ASLIEREAPAGDFDKVARVI	2156.1589	-0.0029	53.9	27	37	1	0.075	4.56
3	GAQVAEILRVVAKELES	1811.0152	-0.0119	70.17	33	35	1	0.382352941	4.56
3	SFGYTEEQAKEKWPDREIKVA	2510.2441	-0.0119	65.56	34	38	1	-1.323809524	4.8
3	MYKVFEALDDLVQAVQRA	2095.0772	-0.0154	61	33	37	1	0.172222222	4.3
3	RVIGAVVDVEFPRGELPA	1923.0578	-0.0179	37.53	26	36	1	0.444444444	4.44
3	AAAPYAEEIQRVLERLASAS	2144.1225	-0.028	52.57	33	37	1	-0.02	4.56
3	ANKQDFAEWILERQ	1746.8689	-0.0047	45.29	31	37	1	-1.114285714	4.44
3	WGPPAGRLLEQLLAEGKAEPV	2230.211	-0.03	50.26	26	37	1	-0.214285714	4.56
3	GGLGVLLLKEFLEGGRDFQPV	2243.2314	-0.0309	48.07	33	37	1	0.333333333	4.44
3	AIQLDRLGILEKYGVELI	2042.1775	-0.0288	38.49	29	34	1	0.566666667	4.44
3	NEAKEELFNLRFQ	1636.8209	-0.0072	45.22	26	37	1	-1.107692308	4.56
3	SHQSLNLRSIYGDEWVEFA	2250.0705	-0.0673	43.3	23	33	1	-0.489473684	4.42
3	HLLRDDNWWAPG	1478.7055	0.0003	42.06	24	36	1	-1.05	5.1
3	NDLTDNKTLSTLLKFDSIMGRLGQEV	2907.5012	-0.0689	40.99	23	38	1	-0.342307692	4.46
3	GVFADSEKKEAAERYI	1811.9053	-0.009	38.68	28	37	1	-0.74375	4.64
3	QVDVYGEFLREIGSPILPH	2168.1266	-0.0297	37.88	22	37	1	-2.33731E-17	4.42
3	SGFNVWEEPNDAKRS	1734.7961	-0.0029	36.81	25	35	1	-1.44	4.44

4	DHMNAIGAPFAREY	1590.7249	-0.0026	37.7	22	34	1	-0.492857143	5.22
4	RHAEPLFFESVPLQT	1769.9101	-0.0101	59.74	27	37	1	-0.246666667	5.3
4	DHMNAIGAPFAREYGGA	1775.8049	-0.0046	58.07	25	34	1	-0.347058824	5.22
4	DHMNAIGAPFAREYGGALA	1959.9261	-0.0148	43.62	22	35	1	-0.015789474	5.22
4	DHMNAIGAPFAREYGGALAT	2060.9738	-0.015	60.19	24	35	1	-0.05	5.22
4	ADGGHQAWVDATANSIRNTLGI	2266.1091	-0.023	70.26	29	37	1	-0.240909091	5.1
4	SYNADGGHQAWVDATANSIRNTLGI	2630.2473	-0.0214	57.42	33	36	1	-0.436	5.1
4	FQSFTIPESLEHFSGEEGVLRRQAISLA	3147.5989	-0.0588	42.99	23	38	1	-0.103571429	4.7
4	NDGAEYDPAKDVYEHQMHSVHGPRNA	2936.2896	-0.0105	47.65	16	31	1	-1.45	5.16
4	SGEDWRWQTHLA	1484.6797	-0.0029	52.79	29	34	1	-1.358333333	5.22
4	SEAGKEHVVKDLGFIAPF	1943.0153	-0.0124	73.42	29	37	1	0.066666667	5.34
4	LAQDQLDVNLFQHLKFL	2041.0997	-0.008	64.68		37	1	0.147058824	5.1
4	DVDVTLIDRTNHHLFQPL	2132.1015	-0.0196	53.6	22	37	1	-0.238888889	5.02
4	DVDVTLIDRTNHHLFQPLLYQVA	2706.413	-0.0342	53.42	23	38	1	0.030434783	5.02
4	EETGYVHPYTRI	1463.7045	-0.001	49.41	22	36	1	-1	5.3
4	STGGGFQLFDTGAWTEPHGDHTHSYT	2805.2056	-0.0317	32.79		27	1	-0.896153846	5.02

4	DHRLTGSVLEFVAMT	1674.84	-0.0077	57.91	31	37	1	0.34666666 7	5.22
4	QALEHIGQAPYAT	1397.6939	0.0035	49.9	24	36	1	-0.30769230 8	5.12
4	NAQALEHIGQAPY	1410.6891	-0.0023	38.96	26	36	1	-0.52307692 3	5.12
4	NAQALEHIGQAPYAT	1582.7739	-0.0004	59.32	23	36	1	-0.38	5.12
4	NHIEEANTLRAL	1379.7157	0.0033	75.23		36	1	-0.55833333 3	5.3
4	FIEVHQPLGPVDDHGHPI	2006.001	-0.0129	54.23	24	37	1	-0.28888888 9	5.02
4	FIEVHQPLGPVDDHGHPIPLPY	2476.254	-0.0246	89.29	25	38	1	-0.26818181 8	5.02
4	FIEVHQPLGPVDDHGHPIPLPYAGA	2675.3496	-0.0263	55.51	23	38	1	-0.108	5.02
4	SLLFPQHFPGFYDAA	1708.825	-0.0024	76.36	30	36	1	0.24666666 7	4.94
4	HVNAADERVKEMWAYSPS	2088.9687	-0.0114	39.61	25	35	1	-0.85	5.34
4	GTHSWGWWQDDLRGSWTT	2174.9559	-0.0237	36.31		30	1	-1.27222222 2	5.1
4	GTHSWGWWQDDLRGSWTTFA	2393.0614	-0.0213	50.29	21	31	1	-0.915	5.1
4	TLTGTWGNLTDSQLRSV	1847.9378	-0.0074	52.74	31	37	1	-0.32352941 2	6.78
4	NYRELSEAEKLEVLLKELR	2331.2797	-0.0133	62.99	29	36	1	-0.78421052 6	4.86
4	NHWFEWSTNNWVNMT	1964.8264	-0.0053	63.48	25	29	1	-1.11333333 3	5.12
4	NHWFEWSTNNWVNMT	1980.8213	0.0089	38.36	15	29	1	-1.11333333 3	5.12

4	NHWFEWSTNNWVNMTGLPRET	2618.1761	-0.0215	72.25	24	32	1	-1.123809524	5.3
4	NHWFEWSTNNWVNMTGLPRET	2634.171	-0.0081	45.51	16	32	1	-1.123809524	5.3
4	NGPGHGQSIGGTNQAINDIVVDYLRT	2695.3314	-0.023	104.18	33	37	1	-0.453846154	5.1
4	WYDHLPYIPFPDHWPV	2080.9836	-0.0004	65.85		36	1	-0.55625	4.94
4	THDVNPLMGLVDRI	1578.8188	-0.0032	73.6	28	36	1	0.107142857	5.1
4	DFFRPMGI	981.4742	0.0117	36.18	30	34	1	0.25	6.78
4	DFLHVWGLTVA	1256.6554	0.0025	80.88	32	36	1	1.081818182	4.94
4	NLRNSAAVGQLDFPLNIA	1912.0166	-0.0094	71.87		37	1	0.194444444	6.78
4	NDIIAGAFGFRPV	1375.7248	0.0007	50.01	27	36	1	0.653846154	6.78
4	NALAMFDSRGPQIQAA	1688.8304	-0.0073	51.15	28	37	1	-0.06875	6.78
4	AAFQHLIDTGDYQRI	1746.8689	-0.0072	47.14	27	37	1	-0.326666667	5.1
4	GHEDHIGAIPWLL	1456.7463	0.0015	42.23	28	37	1	0.130769231	5.04
4									
5	ETTRPEFIHPV	1324.6776	-0.0216	55.55	33	35	1	-0.709090909	5.3
5	SETTRPEFIHPV	1411.7096	-0.025	41.13	24	36	1	-0.716666667	5.3
5	RHAEPLFFESVPL	1540.8038	-0.0058	63.32		36	1	0.038461538	5.3
5	RHAEPLFFESVPLQT	1769.9101	-0.0113	49.6	22	37	1	-0.246666667	5.3
5	VIDHMNAIGAPFAREYGGA	1987.9574	-0.0287	58.95	30	36	1	0.14736842	5.22

								1	
5	KDMWEYQKDHMNLVSPL	2133.0023	-0.0164	50.44	23	35	1	-1.029411765	5.28
5	VIDHMNAIGAPFAREYGGALAT	2273.1263	-0.0224	40.45	25	37	1	0.35	5.22
5	STHSETTRPEFIHPVSVLPEVS	2448.2285	-0.0261	39.43	19	38	1	-0.354545455	5.22
5	ADYTGGRQAQNDGVKFI	1838.8911	-0.042	67.85	30	35	1	-0.823529412	6.76
5	AQTEPIADHILSH	1430.7154	-0.0103	41.92	30	36	1	-0.276923077	5.04
5	VVGGTEAQRNSWPSQI	1727.8591	-0.0369	65.1		36	1	-0.5875	6.98
5	NFIEPLGPAFWERKS	1789.9151	-0.0269	69.67	34	37	1	-0.566666667	7
5	MVNDGAEYDPAKDVYEHQMHSVHGPRNA	3166.3985	-0.0491	63.4	18	28	1	-1.128571429	5.16
5	SLTSGEDWRWQTHLA	1785.8434	-0.0499	45.7	26	33	1	-0.933333333	5.22
5	TSLTSGEDWRWQTHLA	1886.8911	-0.0543	47.55	24	33	1	-0.91875	5.22
5	KHFNDNWAAEKESNWG	1931.855	-0.0354	94	27	30	1	-1.7875	5.34
5	STSLTSGEDWRWQTHLA	1973.9232	-0.0574	60.35	30	32	1	-0.911764706	5.22
5	GKEHVVKDLGFIAPFESY	2035.0415	-0.0247	53.57	20	37	1	-0.105555556	5.34
5	FASTSLTSGEDWRWQTHLA	2192.0287	-0.0265	58.35	25	35	1	-0.573684211	5.22
5	QAPTAYQEGRPAII	1513.7889	-0.0138	53.39	28	37	1	-0.442857143	6.88

5	SAEETGYVHPYTRI	1621.7736	-0.0062	74.86	28	36	1	-0.785714286	5.3
5	LFINKDRAYVSDPSNNELHVV	2429.2339	-0.0364	76.87	31	38	1	-0.371428571	5.28
5	QGRVLDVMNAW	1287.6394	0.006	47.69	34	37	1	-0.036363636	6.78
5	YLEQNFGVARNNNSIGGL	1964.9704	-0.0288	68.39	25	37	1	-0.4388888 9	6.88
5	ADYPMYGIHGWAQFNSQLERT	2483.1328	-0.0339	100.52	28	33	1	-0.761904762	5.22
5	FWVPEFLFKRDA	1553.8031	-0.0045	47.56	31	36	1	0.025	6.92
5	FWVPEFLFKRDAYAHPEAN	2336.1378	-0.0274	49.12	27	36	1	-0.484210526	5.34
5	SLLPWEYLKL	1260.7118	-0.0001	36.95		34	1	0.32	6.8
5	SLLPWEYLKLTLDRG	1802.993	-0.0095	39.12	25	36	1	-0.14	6.86
5	GTHSWGWWQDDLRGSWT	2073.9082	-0.0271	64.41	26	30	1	-1.305882353	5.1
5	GTHSWGWWQDDLRGSWTT	2174.9559	-0.0306	72.98	26	30	1	-1.272222222	5.1
5	TLTGTWGNLTDSQLRSV	1847.9378	-0.0158	77.79		37	1	-0.323529412	6.78
5	AGSSHTIEPEIYRGVSTL	1915.9639	-0.0252	47.03	29	37	1	-0.1888888 9	5.3
5	IFIGDKDVTHVAPRDRDIA	2137.128	-0.0313	66.52	26	37	1	-0.236842105	5.26
5	ALEGLHYMNAAREIMGM	1905.8899	-0.0185	53.29	24	35	1	0.170588235	5.3

5	SIVNHWFEWSTNNWVNMTGLPRET	2917.3606	-0.0332	38.47	20	34	1	-0.654166667	5.3
5	SFVLRGGQVELPV	1399.7824	-0.0011	60.63	30	35	1	0.638461538	6.98
5	QSKEELEHAQKFAQHLLDRDYTPQ	2910.426	-0.052	63.79	25	37	1	-1.533333333	5.28
5	QISRTLFDLLFPSNT	1750.9254	-0.0056	53.11	27	36	1	0.126666667	6.78
5	QWSVDSFRLA	1207.5986	0.0082	38.42	32	36	1	-0.14	6.78
5	SGFELGVGPGVWDIHSPRVPSAQ	2391.1972	-0.0307	57.41	35	37	1	-0.104347826	5.22
5	SLLLPFVGMWSDRV	1618.8541	-0.0042	49.14	36	37	1	0.857142857	6.78
5	MSLLLPFVGMWSDRV	1749.8946	-0.0026	38.54	31	37	1	0.926666667	6.78
5	FLPNPFWVPELRPF	1757.9294	-0.0058	54.1	34	37	1	0.1	6.98
5	KWAGAVADLIGLLG	1382.7922	-0.0012	62.39		34	1	1.142857143	6.66
5	WRGVFPDTV	1075.5451	0.0001	39.51		37	1	-0.044444444	6.78
5	SGGLTLLVPFVDRV	1471.8399	-0.0021	54.24		34	1	1.064285714	6.78
5	LYPWLLKDHQSVTEI	1840.9723	-0.022	45.43	31	37	1	-0.186666667	5.22
5	GWFAEHPDALTDTYRRPTTA	2304.0923	-0.0327	44.22	16	35	1	-0.965	5.28
5	ADPFRVIPA	984.5393	-0.0291	39.18		35	1	0.433333333	6.78
5	GWSAYPRHLDFEAFQ	1822.8427	-0.0076	42.43		35	1	-0.68	5.22
5	SFGIGVERTFPV	1307.6874	-0.0029	41.3	27	36	1	0.55	6.98
5	NYERNLKGLVSISPELELH	2210.1695	-0.0281	41.23	25	37	1	-0.531578947	5.4

6	DRLMGNRPEWV	1371.6717	-0.0357	48.39	31	34	1	-1.136363636	7.04
6	SAWEPGENGTFVRH	1585.7274	-0.0008	53.34	27	34	1	-1.014285714	5.3
6	HGPEYYHRQLGDILYF	2006.9639	-0.0198	64.19	22	36	1	-0.8	6.02
6	ADYTGGRQAQNDGVKFI	1838.8911	-0.0525	85.28	33	34	1	-0.823529412	6.76
6	VVGGTEAQRNSWPSQI	1727.8591	-0.0461	80.32		35	1	-0.5875	6.98
6	GNFIEPLGPAFWERKS	1846.9366	-0.0153	55.49	32	37	1	-0.55625	7
6	AGNFIEPLGPAFWERKS	1917.9737	-0.0166	91.6	32	37	1	-0.417647059	7
6	EYDPAKDVYEHQMHSVHGPRNA	2579.1611	-0.0288	49.89		32	1	-1.459090909	5.76
6	QAPTAYQEGRPAII	1513.7889	-0.0358	55.44	28	36	1	-0.442857143	6.88
6	SSDLPIELRDNQRFSGK	1960.9966	-0.0569	64.86	23	36	1	-1.170588235	6.98
6	KVQAPDSYGRIGNQSGSDVSPV	2260.1084	-0.0241	58.08	23	37	1	-0.722727273	6.76
6	YLEQNFGVARNNNSIGGL	1964.9704	-0.0545	38.82	23	36	1	-0.438888889	6.88
6	FWVPEFLFKRDA	1553.8031	-0.0036	44.34	33	36	1	0.025	6.92
6	HSFWVPEFLFKRDAYAHPEAN	2560.2288	-0.017	68.89	29	36	1	-0.628571429	6.04
6	SLLPWEYLKLTLDRG	1802.993	-0.0107	40.81	27	36	1	-0.14	6.86
6	SVLDNMAQVMSKAELRLA	1975.023	-0.022	80.5		37	1	0.277777778	6.92
6	NYRELSEAEKLEVLLKELRSPRPL	2881.6024	-0.0372	62.94	26	36	1	-	7.08

								0.816666667	
6	GGQTIDFQNGTIRQV	1632.822	-0.0421	71.97	35	36	1	-0.573333333	6.78
6	SNYKKTGPVNTYGLGEIEAGANLL	2508.286	-0.0152	96.58	26	38	1	-0.4	6.84
6	ILRPELGIMLQ	1281.7478	-0.0078	35.21	29	33	1	0.8	6.98
6	KTIFWNGPMGVFEFA	1742.8491	-0.005	56.39	26	36	1	0.393333333	6.84
6	YQNWSNKIVAET	1451.7045	-0.034	44.45	22	35	1	-0.925	6.8
6	SIGSTEVEGLGFHIRAGNAA	1984.9966	-0.0549	46.04	28	36	1	0.155	5.3
6	LQGVEKNLEQVKQTTERNAPA	2352.2397	-0.0656	69.29	36	38	1	-1.147619048	7
6	TIGPGLGIGILVGKALEGMARQPEMA	2578.3975	-0.0293	61.22	28	37	1	0.503846154	7
6	GLVPQDPMTNLNPVWRI	1949.0193	-0.0152	52.56	34	37	1	-0.170588235	6.78
6	GALDRFAAGREV	1260.6575	-0.02	53.11	33	37	1	-0.05	7.04
6	FLPNPFWVPELRPF	1757.9294	-0.0043	50.61	34	37	1	0.1	6.98
6	VYALGGGDGGQGGQNWVTRT	1991.945	-0.0586	72.8	29	33	1	-0.545	6.7
6	ALNIINTLFNRMNVEGV	1917.0142	-0.0116	47.22	31	37	1	0.494117647	6.98
6	QLENAIERLGNFLSTYKQ	2123.1011	-0.0132	50.81	35	37	1	-0.672222222	6.94
6	FNGLGGLDDVEAAYKKLTGK	2095.095	-0.0026	43.64	22	37	1	-0.365	6.8
6	LAIPELMRLL	1167.7049	0.0053	36.8		30	1	1.38	6.98
6	SFDAKKLAENYGALLDEIIRIKPS	2690.4642	-0.0296	45.32	22	37	1	-0.229166667	6.9
6	NFSHGDHPDHEQNYKWVREA	2465.0897	-0.022	43.16	18	31	1	-1.935	5.76
6	KLTVRDLLQEFA	1431.8085	0.0011	43.13	29	34	1	0.05	6.92
6	ELPMIWRIA	1127.6161	0.0078	37.93		35	1	0.666666667	6.98

7	IDRLMGNRPEWV	1484.7558	-0.0049	39.56	30	36	1	-0.666666667	7.04
7	STHSETTRPEFIHPV	1736.8482	-0.0059	52.95	27	36	1	-0.886666667	6.02
7	ADYTGGRQAQNDGVKFI	1838.8911	-0.0112	66.33		36	1	-0.823529412	6.76
7	KMDELHNQIHDLPAVR	1914.9734	-0.0124	45.46	25	37	1	-0.86875	6.02
7	KMDELHNQIHDLPAVRI	2028.0574	-0.0177	72.34	28	37	1	-0.552941176	6.02
7	KMDELHNQIHDLPAVRIA	2099.0946	-0.0225	84.76	28	38	1	-0.422222222	6.02
7	KMDELHNQIHDLPAVRIA	2115.0895	-0.0323	37.63	24	37	1	-0.422222222	6.02
7	TKMDELHNQIHDLPAVRI	2129.1051	-0.025	51.15	25	37	1	-0.561111111	6.02
7	TTKMDELHNQIHDLPAVRI	2230.1528	-0.0629	79.23	26	37	1	-0.568421053	6.02
7	TTKMDELHNQIHDLPAVRIA	2301.1899	-0.0251	85.03	28	38	1	-0.45	6.02
7	GNFIEPLGPAFWERKS	1846.9366	-0.0196	73.72	30	37	1	-0.55625	7
7	GGMIKNPWNPKEGPMDV	1868.8913	-0.0104	41.14	20	36	1	-0.952941176	6.84
7	AGNFIEPLGPAFWERKS	1917.9737	-0.0204	84.79	29	37	1	-0.417647059	7
7	AAGNFIEPLGPAFWERKS	1989.0108	-0.0242	63.3	30	37	1	-0.294444444	7
7	GLTIIAPMGGMIKNPWNPKEGPMDV	2665.343	-0.0158	65.77	29	38	1	-0.096	6.84

7	NFGGPYLLTHQALLV	1641.8879	-0.0108	86	30	36	1	0.626666667	7.76
7	NAGRSESVNFGGPYLLTHQALLV	2442.2655	-0.034	50.69	23	38	1	0.082608696	7.76
7	ARIIGAFERAEI	1344.7513	-0.0019	38.74	33	35	1	0.441666667	7.12
7	ANVEKKHFVDPAWPEHNPA	2185.0705	-0.0172	63.66	22	37	1	-1.094736842	6.04
7	ANVEKKHFVDPAWPEHNPADGH	2494.1778	-0.0225	78.52	27	35	1	-1.268181818	5.76
7	ANVEKKHFVDPAWPEHNPADGHVVTELISKV	3462.7684	-0.0269	83.63	25	39	1	-0.512903226	5.78
7	ANVEKKHFVDPAWPEHNPADGHVVTELISKVAGA	3661.8641	-0.0279	57.66	27	39	1	-0.373529412	5.78
7	GRGGNAGLYLLDGMRATEYS	2100.0058	-0.0251	59.24	28	36	1	-0.435	6.9
7	HLSLLPLGKEERGL	1560.8987	-0.0413	60.95	26	35	1	-0.2	7.72
7	EKRGEGEFPKQL	1416.7361	0.0037	40.22	31	37	1	-1.841666667	7
7	HSFWVPEFLFKRDA	1777.894	-0.0099	44.42	23	37	1	-0.264285714	7.72
7	FIEVHQPLGPVDDHGHPIPLPYAGAAVPKQM	3329.702	-0.0549	74.73	22	38	1	-0.122580645	5.74
7	GTWGNLTDSQLRSV	1532.7583	-0.0082	76.38	34	37	1	-0.564285714	6.78
7	FELHYPHMI	1185.5641	-0.0162	40.54	28	34	1	0.022222222	6.02
7	FELHYPHMIER	1470.7078	-0.0001	59.84	31	36	1	-0.709090909	6.02
7	FELHYPHMIERM	1601.7483	-0.0425	56.71	25	33	1	-0.49166666	6.02

								7	
7	GGQTIDFQNGTIRQV	1632.822	-0.0047	62.84	34	37	1	-0.573333333	6.78
7	NLNGTHFNPLKKQGDEV	1909.9646	-0.0143	51.12	28	37	1	-1.205882353	7.64
7	ALRHVATQHPYDSPVDG	1861.9071	-0.0127	53.5	27	37	1	-0.705882353	6.02
7	HVDIAGPAYNTAGEFGYTPKRA	2334.1393	-0.0211	59.25	25	37	1	-0.536363636	7.62
7	LQGVEKNLEQVKQTTERNAPA	2352.2397	-0.0221	58.44	32	38	1	-1.147619048	7
7	AGKTLVELFGLAREL	1615.9297	-0.0021	61.67		33	1	0.593333333	7
7	ATIGPGLGIGILVGKALEGMARQPEMA	2649.4346	-0.0327	51.64	22	37	1	0.551851852	7
7	GLVPQDPMTNLNPVWRI	1949.0193	-0.0171	55.2	31	37	1	-0.170588235	6.78
7	GALDRFAAGREV	1260.6575	0.0057	54.5		36	1	-0.05	7.04
7	ALNIINTLFNRMNVEGV	1917.0142	-0.0072	48.5	27	37	1	0.494117647	6.98
7	GPQNMPKFSDRQLSADEKKDIIAFI	2847.4589	-0.0563	57.47	22	38	1	-0.7	6.88
7	FLSASVHEREFFDQATRHGWTMVG	2807.3238	-0.0408	46.32	23	35	1	-0.333333333	6.04
7	NVISHLLSTIPYEKI	1725.9665	-0.013	38.44	26	35	1	0.4	7.66
8	TTRPEFIHPV	1195.635	0.0068	51.36	31	34	1	-0.43	7.84
8	FIQFHPTGLPV	1254.6761	0.0048	69.67		35	1	0.645454545	7.84
8	FIQFHPTGLPVNSTWQS	1957.9687	-0.0124	54.79	29	37	1	-0.182352941	7.84

8	RVIDHMNAIGAPFAREYGGA	2144.0585	-0.0128	43.3	24	37	1	-0.085	7.76
8	ADYTGGRQAQNDGVKFI	1838.8911	-0.0155	66.76	30	36	1	-0.823529412	6.76
8	AVTTKMDELHNQIHDLPAVR	2287.1743	-0.0366	59.62	26	38	1	-0.465	6.02
8	AVTTKMDELHNQIHDLPAVRI	2400.2583	-0.0168	90.71	29	38	1	-0.228571429	6.02
8	AVTTKMDELHNQIHDLPAVRI	2416.2533	-0.042	56.47	20	38	1	-0.228571429	6.02
8	AVTTKMDELHNQIHDLPAVRIA	2471.2955	-0.0351	89.53	36	38	1	-0.136363636	6.02
8	AVTTKMDELHNQIHDLPAVRIA	2487.2904	-0.0623	49.02	25	38	1	-0.136363636	6.02
8	VVGGTEAQRNSWPSQI	1727.8591	-0.0102	44.42	31	36	1	-0.5875	6.98
8	FHYGDYYAYSKI	1525.6878	-0.0025	48.05	30	34	1	-0.658333333	7.56
8	QRRHDGPSEEVDRRTIVA	2120.0835	-0.0066	53.58	27	37	1	-1.527777778	7.86
8	GKEHVVKDLGFIAPF	1655.9035	-0.0116	77.16	31	36	1	0.246666667	7.64
8	NAGRSESVNFGGPYLLTH	1917.9333	-0.0181	52.89	28	36	1	-0.455555556	7.76
8	NAGRSESVNFGGPYLLTHQ	2045.9919	-0.0272	60.8	25	36	1	-0.615789474	7.76
8	TYSINAGRSESVNFGGPYLLTHQ	2510.219	-0.0256	39.67	24	37	1	-0.434782609	7.7
8	DRTNHHLFQPLLYQVA	1951.0064	-0.0176	43.23	23	37	1	-0.51875	7.9
8	NYYQEILKLDGMKHF	1897.9396	-0.0112	60.73	25	37	1	-0.746666666	7.58

								7	
8	NYYQEILKLDGMKHFA	1968.9767	-0.0138	44.72	27	37	1	-0.5875	7.58
8	VHLSLLPLGKEERGL	1659.9672	-0.0103	74.67		32	1	0.09333333	7.72
8	FWVPEFLFKRDAYAHPEANKS	2551.2648	-0.0287	61.71	23	37	1	-0.66190476	7.6
8	SLLFPQHFPG	1141.592	0.0014	38.93	28	35	1	0.21	7.84
8	SLLFPQHFPGFY	1451.7238	-0.0025	39.03	23	36	1	0.3	7.76
8	FQNYALYPHMTV	1482.6966	-0.0023	45.15	30	35	1	-0.05	7.7
8	SVFHARIVEGGQHNFS	1783.8754	-0.0097	51.5	26	36	1	-0.21875	8
8	SVFHARIVEGGQHNFSA	1854.9125	-0.0154	39.33	22	36	1	-0.1	8
8	MNLGTFSVKDGDATSGHARI	2076.0058	-0.0219	68.52	27	36	1	-0.275	7.72
8	LKESLENAAPRPVSPFYPAIS	2285.2055	-0.0321	58.06	32	38	1	-0.21428571	6.94
8	FELHYPHMIERMRAEAH	2166.0251	-0.0158	46.65	22	36	1	-0.79411764	6.32
8	NYRELSEAEKLEVLLKELRSPRPLIPH	3228.7982	-0.0301	40.43	15	36	1	-0.73703703	7.62
8	NIATHHPDMFKFVGSFS	1933.9145	-0.0116	66.24	22	36	1	-0.04705882	7.88
8	GHSRLEQIRAEMAGGSLT	1911.9585	-0.0104	60.72	22	37	1	-0.47777778	7.86
8	GHSRLEQIRAEMAGGSLTA	1982.9956	-0.0129	63.22	27	37	1	-0.35789473	7.86
8	KFFVHDPFTREA	1492.7463	-0.0045	58.78	31	37	1	-0.54166667	7.72
8	LNFHVQRGEVFGLLGT	1785.9526	-0.0082	59.13	35	37	1	0.33125	7.84
8	LNFHVQRGEVFGLLGTNGA	2028.0541	-0.0166	60.49		37	1	0.16842105	7.84

8	SLAHEQKISGLIREL	1692.9522	-0.0126	39.98	25	34	1	-0.126666667	7.72
8	WLQVVPENLRILVH	1714.9883	-0.0049	35.16	30	33	1	0.557142857	7.84
8	WLQVVPENLRILVHQLPERA	2409.3645	-0.0187	59.96	29	35	1	0.015	7.86
8	SAGLKESHPHHIQQTVEAPNYH	2479.1992	-0.02	82.13	25	37	1	-1.15	6.48
8	FHQGGVGGDITGGLPRVQ	1793.9173	-0.0092	42.69	26	37	1	-0.188888889	7.84
8	HFFPQGRGEALVRPGQEHFIA	2392.2189	-0.0192	50.3	24	38	1	-0.442857143	8
8	SYLNPYTNHDAVFERRNTVLGA	2536.2459	-0.0126	46.93	22	37	1	-0.640909091	7.7
8	FQNYALYPHMTV	1482.6966	-0.0023	45.15	30	35	1	-0.05	7.7
8	IIFVHPLFQYV	1374.77	0.0003	48.53	27	34	1	1.563636364	7.76
8	QHSDVPVLNGFRSLPQSFQPRVDVA	2792.4358	-0.0261	48.11	23	38	1	-0.36	7.84
8	SHVNDSGFWLVGRLM	1716.8406	0.0096	45.13	24	37	1	0.18	7.84
8	SYTSGFGARWGTNHNGVDIA	2108.9664	-0.0045	35.19	19	34	1	-0.495	7.76
9	SALQRQIHLGSVEIFTH	1935.0326	-0.0143	78.23	28	37	1	0.076470588	8
9	GELGYDVKAFTYHDAPRRA	2165.0654	-0.0075	44.64	26	37	1	-0.847368421	7.62
9	LGELGYDVKAFTYHDAPRRA	2278.1494	-0.0146	78.38	27	37	1	-0.615	7.62
9	ALGELGYDVKAFTYHDAPRRA	2349.1866	0.0018	87.53	31	38	1	-0.5	7.62
9	AALGELGYDVKAFTYHDAPRRA	2420.2237	-0.0195	76.56	28	38	1	-0.395454545	7.62
9	ADYTGGRQAQNDGVKFI	1838.8911	-0.0109	60.04	32	36	1	-0.823529412	6.76

9	KMDELHNQIHDLPAVRI	2028.0574	-0.0168	43.74	26	37	1	-0.552941176	6.02
9	KDVYEHQMHSVHGPRNA	2003.9384	-0.0133	68.34	23	35	1	-1.411764706	7.9
9	GKEHVVKDLGFIAPF	1655.9035	-0.0092	59.05	31	36	1	0.246666667	7.64
9	AATYSINAGRSESVNFGGPYLLTHQ	2652.2932	-0.0229	69.56	18	37	1	-0.256	7.7
9	DRTNHHLFQPL	1376.6949	0.0042	57.53	35	37	1	-1.209090909	8
9	LIDRTNHHLFQPL	1602.8631	-0.0041	51.46	26	36	1	-0.384615385	8
9	LIDRTNHHLFQPLLYQVA	2177.1745	-0.0187	56.04	25	37	1	-3.70074E-17	7.9
9	FQNYALYPH	1151.54	-0.0019	39.68	30	35	1	-0.666666667	7.7
9	LYDRKLGGHLYDPA	1616.8311	-0.0079	48.73	27	37	1	-0.742857143	7.62
9	FELHYPHMIERMR	1757.8494	-0.0055	45.14	23	36	1	-0.8	7.92
9	AVAPRVDGHVAPQRPEPT	1895.9966	-0.0082	43.62	23	37	1	-0.677777778	7.84
9	ALGVLYDRKLGGHLYDPA	1957.0421	-0.0129	39.01	20	37	1	-0.055555556	7.62
9	MMMAGLDGIKNRIEPH	1811.8844	-0.0096	56.13	33	36	1	-0.2625	7.72
9	ALRHVATQHPYDSPVDGRV	2117.0766	-0.0121	51.11	27	37	1	-0.647368421	7.92
9	WLQVVPENLRILVH	1714.9883	-0.0029	66.24		33	1	0.557142857	7.84
9	AADSAGLPLFRYIGGPNAHVLPVPM	2562.3417	-0.0253	44.65	20	38	1	0.472	7.76
9	LEWSTRWLNEALRHA	1880.9645	-0.0074	42.59	25	37	1	-	7.86

								0.73333333	
9	GFNWPHHDHFRVV	1646.7855	-0.0062	47.2	32	36	1	-0.769230769	8.08
9	GFNHFFRGKDHPGGGDQI	1984.9293	-0.01	45.62	26	35	1	-1.083333333	7.88
9	FQNYALYPH	1151.54	-0.0019	39.68	30	35	1	-0.666666667	7.7
9	GFGPRWGTFHNGIDIA	1743.8482	-0.0068	49.07	25	36	1	-0.19375	7.84
9	LEREENFHNLRKA	1654.8539	-0.0067	44.77	30	36	1	-1.646153846	7.72
9	ARDVIQNHLIQLL	1531.8834	0.0005	43.05		32	1	0.361538462	7.84
9	ITSDWKSHWYADKS	1722.8002	-0.0102	39.39	23	35	1	-1.278571429	7.6
9	LRREFDLPVHV	1379.7674	0.0014	41.03	26	33	1	-0.181818182	7.84
9	AILSEKLGIPHISTGDLFRA	2137.1895	-0.0193	40.84	24	35	1	0.4	7.72
9	LQSFWFPIHV	1272.6655	0.0026	38.45	24	37	1	0.81	7.84
9	DIFRFAPFFH	1295.6451	0.0037	37.5	31	35	1	0.47	7.84
9	AEIVGGPWHPSVKG	1432.7463	-0.0042	37.36	29	37	1	-0.142857143	7.72
10	NLVPRDVASRAISQQI	1765.9799	-0.0043	50.41	33	35	1	-0.0875	10.88
10	ALGELGYDVKAFTYHDAPRRA	2349.1866	-0.0129	62.15	27	38	1	-0.5	7.62
10	NSTWQSKTILMSESLRNDGRI	2435.2227	-0.0197	45.73	28	37	1	-0.847619048	9.74
10	GNGPYKLQEWNH	1441.6738	-0.0045	51.23	34	35	1	-1.825	7.66
10	SYVQLGVLPRAG	1258.7034	0.0056	35.88	23	35	1	0.44166666	9.84

								7	
10	NSYVQLGVLPRAG	1372.7463	0.0021	74.56		36	1	0.138461538	9.84
10	TLNSYVQLGVLPRAG	1586.878	-0.0037	87.05		35	1	0.326666667	9.84
10	NSYVQLGVLPRAGTILA	1770.9992	-0.0031	46.47	31	35	1	0.658823529	9.84
10	TLNSYVQLGVLPRAGTI	1801.0098	-0.0081	48.98	25	35	1	0.511764706	9.84
10	NVLVQAGLVTLKTPGLVTPAPV	2186.3039	-0.0145	124.39		30	1	0.963636364	9.74
10	NTGAFGLVGRI	1103.6087	0.0051	44.79		35	1	0.654545455	11.04
10	AGDNGSGMHAHQSLWKDGKPLFH	2489.1659	-0.0161	67.49	23	35	1	-0.92173913	7.88
10	NTIGMVFQKANPFPT	1663.8392	-0.0073	63.13	35	37	1	-0.093333333	9.74
10	NARPVLVGPLTFL	1395.8238	0.0048	47.15	20	31	1	0.930769231	11.04
10	AADSAGLPLFRYIGGPNAH	1925.9748	-0.0064	68.46	30	37	1	0.047368421	7.76
10	NKLLGSFELGGIAPAPR	1738.973	-0.009	80.45	30	35	1	0.1	9.74
10	NFIIRIPQAPT	1268.7241	0.0043	64.16	31	34	1	0.245454545	11.04
10	GVQKYFGDFH	1196.5615	-0.0008	60.54	30	34	1	-0.64	7.66
10	NQLKLGDVFITRPGEKI	1927.0891	-0.0069	58.36	27	35	1	-0.341176471	9.46
10	NQLKNLVPGQASEKLEKA	1966.0847	-0.0119	43.28	27	36	1	-0.905555556	9.3
10	NFKLREGMPI	1203.6434	0.004	38.93	33	36	1	-0.44	9.74
11	SWTYHGANRL	1203.5785	-0.0044	48.55		36	1	-0.97	9.84
11	YYTRGQTGQQLQLS	1641.8111	-0.0003	82.73		36	1	-1.178571429	9.58

11	TYYTRGQTGQQLQLS	1742.8588	0.0003	60		37	1	-1.146666667	9.58
11	SWTYHGANRLGANSLL	1758.8801	0.0004	45.57	25	37	1	-0.3125	9.84
11	NLVPRDVASRAISQQI	1765.9799	-0.0063	45.71	26	35	1	-0.0875	10.88
11	SYVQLGVLPRAG	1258.7034	-0.0014	65.75		35	1	0.441666667	9.84
11	NSYVQLGVLPRAG	1372.7463	-0.0051	64.07	32	36	1	0.138461538	9.84
11	TLNSYVQLGVLPR	1458.8195	-0.0103	73.12		35	1	0.269230769	9.84
11	SYVQLGVLPRAGTI	1472.8351	-0.0074	43.47	31	35	1	0.65	9.84
11	TLNSYVQLGVLPRAG	1586.878	-0.0007	98.11		35	1	0.326666667	9.84
11	SYVQLGVLPRAGTILA	1656.9563	-0.0039	39.89	27	34	1	0.91875	9.84
11	TLNSYVQLGVLPRAGTI	1801.0098	0	68.8	28	35	1	0.511764706	9.84
11	QSVTLNSYVQLGVLPRAG	1901.0371	0.0538	38.27		33	2	0.266666667	9.84
11	TLNSYVQLGVLPRAGTILA	1985.131	-0.002	59.76		34	1	0.752631579	9.84
11	FGPAARALPQLPI	1349.7819	-0.0027	50.96	27	32	1	0.546153846	11.04
11	FGPAARALPQLPITV	1549.898	-0.0032	38.2	30	32	1	0.706666667	11.04
11	YNYLVYTGVIKG	1388.734	-0.0072	62.22	28	37	1	0.325	9.28
11	NVLVQAGLVTLKTPGLVTPAPV	2186.3039	0.0034	65.66		29	1	0.963636364	9.74
11	HRTLAGEYKNFNTNSA	1821.8758	0.0003	90.43	25	36	1	-1.20625	9.52
11	FAAGKLGPGIAIQPT	1439.8136	-0.0063	106.27		34	1	0.566666667	9.74
11	TAARKLGANEWIGA	1456.7786	-0.0049	38.36	31	36	1	-0.164285714	9.74
11	FAAGKLGPGIAIQPTGNTVVAPA	2149.1895	0.0039	79.08	27	35	1	0.62173913	9.74
11	MGRNIPVQIQPA	1322.7129	-0.0041	52.74	31	35	1	-0.14166666	11.04

								7	
11	HLSLLPLGKEERGLRA	1788.037	-0.0062	55.68	22	32	1	-0.34375	9.74
11	KLVHLSLLPLGKEERGL	1901.1462	-0.0554	74.53		33	1	0.07647058	9.46
11	HHNLLQRPRDVPV	1579.8695	-0.004	37.43	20	35	1	-1.00769230	10.88
11	TLGLTEFLERKPKAL	1714.9981	-0.0042	76.46		32	1	-0.19333333	9.46
11	NTGAFGLVGRI	1103.6087	-0.0053	44.4	30	35	1	0.65454545	11.04
11	STPNLDQVREKIERRYA	2074.0919	0.0016	49.44	22	37	1	-1.47058823	9.52
11	NWFSRAV	878.4399	-0.0001	35.45		35	1	-0.12857142	11.04
11	NNANWFSRAV	1177.5628	-0.0043	43.78	27	36	1	-0.61	11.04
11	TKRFGGNEFLGA	1295.6622	-0.0026	37.86	27	35	1	-0.50833333	9.74
11	HLGRPKGEVNEKYSLAPV	1993.0745	0.0005	79.9	25	36	1	-0.79444444	9.34
11	NARPVLVGPLTFL	1395.8238	-0.0003	44.84	25	32	1	0.93076923	11.04
11	YNTAGEFGYTPKRA	1573.7525	-0.0042	48.91	25	36	1	-1.1	9.38
11	THIWYFKGVPS	1333.6819	-0.0098	62.39	34	37	1	-0.11818181	9.52
11	NFIIRIPQAPT	1268.7241	-0.0079	40.73		34	1	0.24545454	11.04
11	FLGRVINPLGQPI	1422.8347	-0.0018	41.16		31	1	0.62307692	11.04
11	QAGIKLEIPRY	1286.7346	-0.0034	44.78		34	1	-0.37272727	9.52

11	HYIKGYVPV	1074.5862	-0.0098	55.73		36	1	0.133333333	9.38
11	QTKSVVDIPGAPQRY	1657.8788	-0.0073	42.46	30	37	1	-0.706666667	9.52
11	SLEKKRPEGWPV	1424.7776	-0.0061	36.21		35	1	-1.383333333	9.46
12	SYVQLGVLPRAG	1258.7034	0.0016	49.08	33	35	1	0.441666667	9.84
12	TLNSYVQLGVLPR	1458.8195	-0.003	42	24	35	1	0.269230769	9.84
12	TLNSYVQLGVLPRAG	1586.878	-0.0035	95.17		35	1	0.326666667	9.84
12	TLNSYVQLGVLPRAGTI	1801.0098	-0.0046	44.24	26	35	1	0.511764706	9.84
12	FGPAARALPQLPI	1349.7819	-0.0009	42.89	29	32	1	0.546153846	11.04
12	HRTLAGEYKNFNTN	1663.8066	-0.0025	49.56	24	36	1	-1.45	9.52
12	HRTLAGEYKNFNTNSA	1821.8758	-0.0026	54.72	24	36	1	-1.20625	9.52
12	FAAGKLGPGIAIQPT	1439.8136	-0.0086	97.34	34	35	1	0.566666667	9.74
12	GYVHPYTRI	1104.5716	0.0011	36.04	24	36	1	-0.477777778	9.58
12	ANRDWLRA	1000.5202	-0.0013	41.49		36	1	-1.1875	10.88
12	MGRNIPVQIQPA	1322.7129	-0.001	39.62	33	35	1	-0.141666667	11.04
12	HLSLLPLGKEERGLRA	1788.037	-0.0045	40.53	21	32	1	-0.34375	9.74
12	GLTEFLERKPKAL	1500.8664	-0.0032	41.26	27	33	1	-0.461538462	9.46
12	GIKLEIPRY	1087.6389	0.0044	57.28	31	32	1	-0.266666667	9.52
12	YQYLEKLPKIA	1364.7703	-0.0077	35.91	32	33	1	-	9.24

								0.463636364	
12	AFKGAQDYFFDKRNA	1776.8584	0.0062	38.63	32	37	1	-0.946666667	9.34

Supplementry table S5D Total number of proteins identified from elastade diegest of plama membrane from YEAST cells grown in glucose and oleic acid.

Protein hit number	Protein accession number	protein description	No. of peptides
1	P05030	Plasma membrane ATPase 1	145
2	P19657	Plasma membrane ATPase 2	91
3	P02994	Elongation factor 1-alpha	77
6	P32324	Elongation factor 2	67
4	P39004	High-affinity hexose transporter HXT6	55
5	P39003	High-affinity hexose transporter HXT6	54
7	P38695	Probable glucose transporter HXT5	41
8	P32466	Low-affinity glucose transporter HXT3	40
9	P32467	Low-affinity glucose transporter HXT4	38
10	P10592	Heat shock protein SSA2	35
11	P10591	Heat shock protein SSA1	35
12	P32465	Low-affinity glucose transporter HXT1	28
13	P22146	1,3-beta-glucanosyltransferase GAS1	26
14	Q12117	Protein MRH1	23
15	Q08193	1,3-beta-glucanosyltransferase GAS5	23
22	P33302	Pleiotropic ABC efflux transporter of multiple drugs	21
17	P53252	Sphingolipid long chain base-responsive protein PIL1	20
23	P38079	Protein YRO2	20
16	P23585	High-affinity glucose transporter HXT2	19
18	P54862	Hexose transporter HXT11	19
19	P40885	Hexose transporter HXT9	19
21	P40441	Putative transporter-like protein YIL170W	17
20	P09435	Heat shock protein SSA3	16
24	P40886	Hexose transporter HXT8	16
25	P22202	Heat shock protein SSA4	15

33	Q04182	ATP-dependent permease	13
26	Q12230	Sphingolipid long chain base-responsive protein LSP1	12
27	P11484	Heat shock protein SSB1	12
34	P16521	Elongation factor 3A	12
29	P40150	Heat shock protein SSB2	11
43	Q07824	Polyamine transporter 1	11
48	P39105	Lysophospholipase 1	11
49	P00890	Citrate synthase, mitochondrial	11
31	A6ZL22	Cell wall protein ECM33 (strain YJM789)	10
32	P38248	Cell wall protein ECM33	10
35	P36008	Elongation factor 1-gamma 2	10
36	P00560	Phosphoglycerate kinase	9
42	P38968	Protein transport protein SEC31	9
30	P12398	Heat shock protein SSC1, mitochondrial	8
39	P38216	Uncharacterized protein YBR016W	8
40	P02365	40S ribosomal protein S6	8
41	Q03655	Probable 1,3-beta-glucanosyltransferase GAS3	8
45	P39926	Protein SSO2	8
46	P36035	Carboxylic acid transporter protein homolog	8
58	P25349	Flavoprotein-like protein YCP4	8
62	P47068	Myosin tail region-interacting protein MTI1	8
28	P25619	30 kDa heat shock protein	7
44	P39730	Eukaryotic translation initiation factor 5B	7
47	P15891	Actin-binding protein	7
60	B3RHV0	40S ribosomal protein S1-A (strain RM11-1a)	7
63	B3LLJ2	40S ribosomal protein S1-B (strain RM11-1a)	7
38	Q03482	Uncharacterized protein YDR210W	6
52	P39987	Heat shock protein SSC3, mitochondrial	6
53	P60010	Actin	6
54	P23291	Casein kinase I homolog 1	6

56	P29547	Elongation factor 1-gamma 1	6
59	Q08972	[NU+] prion formation protein 1	6
61	P61864	Ubiquitin GN=UBI1 PE=1 SV=1	6
65	P13181	Galactose transporter	6
81	P23292	Casein kinase I homolog 2	6
37	P46992	Uncharacterized protein YJL171C	5
51	P2529	Protein SIS1	5
68	P32568	Protein SNQ2	5
69	P05453	Eukaryotic peptide chain release factor GTP-binding subunit	5
71	Q08745	40S ribosomal protein S10-A	5
83	P51533	ATP-dependent permease PDR10	5
87	P38427	Trehalose synthase complex regulatory subunit TSL1	5
113	P33892	Translational activator GCN1	5
121	P52910	Acetyl-coenzyme A synthetase 2	5
50	Q05931	Heat shock protein SSQ1, mitochondrial	4
55	P87284	Plasma membrane proteolipid 3	4
57	Q12256	Polyamine transporter 4	4
66	P34167	Eukaryotic translation initiation factor 4B	4
70	P00549	Pyruvate kinase 1	4
72	P16474	78 kDa glucose-regulated protein homolog	4
73	P32457	Cell division control protein 3	4
74	Q07478	ATP-dependent RNA helicase SUB2	4
76	P46784	40S ribosomal protein S10-B	4
77	P54003	Protein SUR7	4
78	P38219	Uncharacterized GTP-binding protein OLA1	4
92	Q3E795	Uncharacterized protein YLR361C-A GN=YLR361C-A PE=1 SV=1	4
94	P25491	Mitochondrial protein import protein MAS5 GN=YDJ1 PE=1 SV=1	4
97	Q3E792	40S ribosomal protein S25-A GN=RPS25A PE=1 SV=1	4
98	P35271	40S ribosomal protein S18 GN=RPS18A PE=1 SV=4	4
99	Q02785	ATP-dependent permease PDR12 GN=PDR12 PE=1 SV=1	4

129	P24276	Protein SSD1 GN=SSD1 PE=1 SV=1	4
131	P00830	ATP synthase subunit beta, mitochondrial	4
135	P15019	Transaldolase	4
67	P32867	Protein SSO1	3
75	P39935	Eukaryotic initiation factor 4F subunit p150	3
79	P40159	Uncharacterized protein YNL208W	3
80	P00925	Enolase 2	3
82	P32329	Aspartic proteinase 3	3
84	P38788	Ribosome-associated complex subunit SSZ1	3
88	P01120	Ras-like protein 2	3
90	P53894	Serine/threonine-protein kinase	3
91	P33417	Intrastrand cross-link recognition protein	3
95	Q12489	Uncharacterized protein YDL012C	3
96	Q06451	Polyamine transporter 3	3
100	P00330	Alcohol dehydrogenase 1	3
101	P32356	Neutral trehalase	3
109	P13587	Sodium transport ATPase 1	3
110	P07246	Alcohol dehydrogenase 3, mitochondrial	3
112	P53049	Oligomycin resistance ATP-dependent permease YOR1	3
114	Q12460	Nucleolar protein 56	3
115	Q04491	Protein transport protein SEC13	3
117	P43581	Hexose transporter HXT10	3
119	P38701	40S ribosomal protein S20	3
124	P04806	Hexokinase-1	3
125	P04807	Hexokinase-2	3
126	Q01896	Sodium transport ATPase 2	3
132	A6ZP43	Metacaspase-1 (strain YJM789)	3
134	P04451	60S ribosomal protein L23	3
137	P49573	Copper transport protein CTR1	3
138	P40485	Phosphatidylinositol 4,5-bisphosphate-binding protein SLM1	3

147	P08679	Citrate synthase, peroxisomal	3
153	P25613	Accumulation of dyads protein 2	3
168	A6ZPE5	Nucleolar protein 58 (strain YJM789)	3
174	A6ZMG6	Myosin-5 (strain YJM789)	3
175	Q12122	Homocitrate synthase, mitochondrial	3
176	P48570	Homocitrate synthase, cytosolic isozyme	3
180	P00360	Glyceraldehyde-3-phosphate dehydrogenase 1	3
64	P38249	Eukaryotic translation initiation factor 3 subunit A	2
85	P35732	Uncharacterized protein YKL054C	2
86	Q12207	Non-classical export protein 2	2
89	P06780	GTP-binding protein RHO1	2
93	P27796	3-ketoacyl-CoA thiolase, peroxisomal	2
102	P53165	SAGA-associated factor 73	2
103	P38873	Target of rapamycin complex 1 subunit KOG1	2
104	Q08887	Nuclear division defective protein 1	2
105	Q12518	Epsin-1	2
106	Q00772	Mitogen-activated protein kinase SLT2/MPK1	2
107	P53968	Transcriptional regulator CRZ1	2
108	P11746	Pheromone receptor transcription factor	2
111	A6ZPJ1	Eukaryotic translation initiation factor 3 subunit B (strain YJM789)	2
116	P52911	Glucan 1,3-beta-glucosidase 2	2
118	P35997	40S ribosomal protein S27-A	2
120	P10823	Guanine nucleotide-binding protein alpha-2 subunit	2
122	P40482	Protein transport protein SEC24	2
123	Q03516	Uncharacterized protein RSN1	2
128	Q99271	Na(+)/H(+) antiporter	2
130	Q05050	Uncharacterized protein YMR031C	2
133	A6ZQJ1	ATP-dependent RNA helicase eIF4A (strain YJM789)	2
136	P14126	60S ribosomal protein L3	2
139	P32527	Zuotin	2

140	Q12213	60S ribosomal protein L7-B	2
141	P05737	60S ribosomal protein L7-A	2
143	P38720	6-phosphogluconate dehydrogenase, decarboxylating 1	2
144	P06367	40S ribosomal protein S14-A	2
145	P40213	40S ribosomal protein S16	2
146	P32497	Eukaryotic translation initiation factor 3 subunit C	2
148	P15108	ATP-dependent molecular chaperone HSC82	2
149	P01119	Ras-like protein 1	2
150	P16140	V-type proton ATPase subunit B	2
151	P05753	40S ribosomal protein S4	2
152	P38631	1,3-beta-glucan synthase component FKS1	2
154	P40989	1,3-beta-glucan synthase component GSC2	2
155	P07278	cAMP-dependent protein kinase regulatory subunit	2
156	Q07651	SUR7 family protein FMP45	2
157	P18480	SWI/SNF chromatin-remodeling complex subunit SNF5	2
158	P46974	Zinc finger protein ZMS1	2
159	P38856	Clathrin coat assembly protein AP180A	2
160	P43572	Enhancer of polycomb-like protein 1	2
161	P48562	Serine/threonine-protein kinase CLA4	2
162	Q08831	Protein VTS1	2
163	P39008	Poly(A) ribonuclease POP2	2
164	P39081	Protein PCF11	2
165	Q99395	Uncharacterized protein YPL229W	2
166	P08539	Guanine nucleotide-binding protein alpha-1 subunit	2
169	P39940	E3 ubiquitin-protein ligase RSP5	2
170	P46943	GTP-binding protein GUF1	2
171	P38993	Iron transport multicopper oxidase FET3	2
172	P32623	Probable glycosidase CRH2	2
173	P32473	Pyruvate dehydrogenase E1 component subunit beta, mitochondrial	2
177	P25087	Sterol 24-C-methyltransferase	2

182	P32599	Fimbrin	2
183	P02405	60S ribosomal protein L42	2
185	P05317	60S acidic ribosomal protein P0	2
193	P52917	Vacuolar protein sorting-associated protein 4	2
196	Q12359	Ammonia transport outward protein 3	2
199	P26782	40S ribosomal protein S24	2
201	Q12691	Sodium transport ATPase 5	2
202	P05755	40S ribosomal protein S9-B	2
203	O13516	40S ribosomal protein S9-A	2
205	P32769	Elongation factor 1 alpha-like protein	2
206	P15303	Protein transport protein SEC23	2
208	P40440	Putative transporter-like protein YIL171W	2
209	P38011	Guanine nucleotide-binding protein subunit beta-like protein	2
211	Q12335	Protoplast secreted protein 2	2
213	P07342	Acetolactate synthase catalytic subunit, mitochondrial	2
218	Q08108	Lysophospholipase 3	2
219	P33322	H/ACA ribonucleoprotein complex subunit 4	2
220	P02293	Histone H2B.1	2
223	P05736	60S ribosomal protein L2	2
224	Q08954	Smr domain-containing protein YPL199C	2
233	P00359	Glyceraldehyde-3-phosphate dehydrogenase 3	2
234	Q12449	Hsp90 co-chaperone AHA1	2
243	P18852	Guanine nucleotide-binding protein subunit gamma	2
248	P38266	Uncharacterized protein YBR108W	2
249	P40088	Plasma membrane iron permease	2
127	P22147	5~-3~ exoribonuclease 1	1
142	P32471	Elongation factor 1-beta	1
167	O13577	Uncharacterized protein YLR437C	1
178	P12385	Eukaryotic peptide chain release factor subunit 1	1
179	P32907	Ammonia transport outward protein 2	1

181	Q04067	Eukaryotic translation initiation factor 3 subunit G	1
184	P00924	Enolase 1	1
186	P36091	Mannan endo-1,6-alpha-mannosidase DCW1	1
187	P38353	Sec sixty-one protein homolog	1
188	P26786	40S ribosomal protein S7-A	1
189	P40053	Uncharacterized protein YER080W	1
190	P33332	Exocyst complex component SEC3	1
191	P27692	Transcription elongation factor SPT5	1
192	P40169	Uncharacterized plasma membrane protein YNL194C	1
194	P00331	Alcohol dehydrogenase 2	1
195	P04840	Mitochondrial outer membrane protein porin 1	1
197	P53879	GTP-binding protein RHO5	1
198	P0C2H6	60S ribosomal protein L27-A	1
200	P17079	60S ribosomal protein L12	1
204	P09547	SWI/SNF chromatin-remodeling complex subunit SWI1	1
207	P05750	40S ribosomal protein S3	1
210	P02829	ATP-dependent molecular chaperone HSP82	1
214	A7A261	LAS seventeen-binding protein 3 (strain YJM789)	1
215	P00927	Threonine dehydratase, mitochondrial	1
216	P31539	Heat shock protein 104	1
217	Q08913	UPF0695 membrane protein YOR390W	1
221	A6ZP47	ATP-dependent RNA helicase DED1 (strain YJM789)	1
222	Q12001	Dolichyl pyrophosphate Man9GlcNAc2 alpha-1,3-glucosyltransferase	1
225	Q3E756	Uncharacterized protein YBL029C-A	1
226	P04147	Polyadenylate-binding protein, cytoplasmic and nuclear	1
227	P53893	Ribosome assembly protein 1	1
228	Q06991	Cell membrane protein YLR414C	1
229	P40075	Vesicle-associated membrane protein-associated protein SCS2	1
230	P19073	Cell division control protein 42	1
231	P38085	Valine/tyrosine/tryptophan amino-acid permease	1

232	Q03690	Protein TIF31	1
235	P40494	Actin-regulating kinase PRK1	1
236	Q12511	Protein phosphatase 2C homolog 5	1
237	P07267	Saccharopepsin	1
238	P31383	Protein phosphatase PP2A regulatory subunit A	1
239	P34231	Uncharacterized protein YKL187C	1
240	P39961	Uncharacterized transcriptional regulatory protein YER184C	1
241	P35691	Translationally-controlled tumor protein homolog	1
242	P17967	Protein disulfide-isomerase	1
244	P40552	Cell wall protein TIR3	1
245	P23254	Transketolase 1	1
246	A6ZMK5	Eukaryotic translation initiation factor 3 subunit I (strain YJM789)	1
247	P0C2I0	60S ribosomal protein L20	1
250	P32589	Heat shock protein homolog SSE1	1
251	P28274	CTP synthase 1	1
252	P38627	CTP synthase 2	1
253	A6ZQ59	CTP synthase 2 (strain YJM789)	1
254	Q06205	FK506-binding protein 4	1
255	P38911	FK506-binding nuclear protein	1
256	P32458	Cell division control protein 11	1
257	P28321	Serine hydrolase YJU3	1
258	P53030	60S ribosomal protein L1	1
259	P42833	Hexose transporter HXT14	1
260	P53978	Elongation factor 3B	1
261	P34216	EH domain-containing and endocytosis protein 1	1
262	Q03718	Non-structural maintenance of chromosome element 5	1
263	P21373	NAD(+) kinase	1
264	P53075	Uncharacterized protein YGL228W	1
265	P40513|	Mitochondrial acidic protein MAM33	1
266	Q99325	Uncharacterized membrane protein YOR152C	1

267	P32074	Coatomer subunit gamma	1
268	P12945	N-terminal acetyltransferase A complex subunit NAT1	1
269	P32901	Peptide transporter PTR2	1
270	P38682	ADP-ribosylation factor GTPase-activating protein GLO3	1
271	P47116	Serine/threonine-protein kinase PTK2/STK2	1
272	Q08299	Siderophore iron transporter ENB1	1
273	P15992	Heat shock protein 26	1
274	Q12344	GTPase-activating protein GYP5	1
275	P23900	Glycerol uptake/efflux facilitator protein	1
276	Q12252	Phosphate metabolism protein 7	1
278	Q03390	Vacuolar protein-sorting-associated protein 60	1
279	Q07950	Sterol esterase 2	1
280	P40474	Quinidine resistance protein 2	1
281	P39932	Sugar transporter STL1	1
282	Q12329	Heat shock protein 42	1
283	P53121	Putative flavin carrier protein 3	1
284	Q03361	Uncharacterized protein JIP4	1
285	P42845	Protein STB1	1
286	P39078	T-complex protein 1 subunit delta	1
287	P35192	Metal-binding activator 1	1
288	A6ZXG9	ATP-dependent RNA helicase DHH1 (strain YJM789)	1
289	P32912	Vacuolar morphogenesis protein 7	1
290	P39962	Casein kinase I homolog 3	1
291	Q12259	BTB/POZ domain-containing protein YLR108C	1
292	P35127	Ubiquitin carboxyl-terminal hydrolase YUH1	1
294	P38903	Serine/threonine-protein phosphatase 2A 56 kDa regulatory subunit delta isoform	1
295	P31412	V-type proton ATPase subunit C	1
296	Q12355	Cell wall mannoprotein PST1	1
297	P40576	Uncharacterized protein YIR024C	1

299	P53319	6-phosphogluconate dehydrogenase, decarboxylating 2	1
300	P07347	N-terminal acetyltransferase A complex catalytic subunit ARD1	1
301	P07991	Ornithine aminotransferase	1
302	P28777|	Chorismate synthase	1
303	P53200	N(6)-adenine-specific DNA methyltransferase-like 1	1
304	P40204	Small nuclear ribonucleoprotein G	1
305	P38853	Kelch repeat-containing protein 1	1
306	P53756	Uncharacterized ABC transporter ATP-binding protein/permease YNR070W	1
307	Q12087	40S ribosomal protein S30	1
308	P53145	Large subunit GTPase 1	1
309	P46971	Dolichyl-phosphate-mannose--protein mannosyltransferase 4	1
310	P38137	Peroxisomal-coenzyme A synthetase	1
311	P10363	DNA primase small subunit	1
312	Q07732	Accumulates dyads protein 3	1
314	Q00684	Tyrosine-protein phosphatase CDC14	1
315	A6ZZH2	NADH-cytochrome b5 reductase 2 (strain YJM789)	1
318	P36006	Myosin-3	1
319	A6ZZJ1	Myosin-3 (strain YJM789)	1
320	Q04781	RING finger protein YMR247C	1
321	P41835	Thiamine biosynthetic bifunctional enzyme	1
322	Q12059	NEDD8-activating enzyme E1 regulatory subunit	1
323	P32660	Probable phospholipid-transporting ATPase DNF1	1
327	P25582	AdoMet-dependent rRNA methyltransferase SPB1	1
328	P53008	Mannosyl-oligosaccharide glucosidase	1
329	P43588	26S proteasome regulatory subunit RPN11	1
330	P36516	54S ribosomal protein L3, mitochondrial	1
331	P36059	Uncharacterized protein YKL151C	1
333	P32837	GABA-specific permease	1
334	P16861	6-phosphofructokinase subunit alpha	1

Supplementry table S5 E (S5E) Protein quantification using TMT6plex. Yeast cell grown in oleic acid were labeled with the tags of m/z 126.1, 127.1 and 128.1 and those grown in glucose with 129.1, 130.1 and 131.1.

ACCESION	Protein	131/129	Stdev	126/129	Stdev	127/129	Stdev	0.99	Stdev	No. of Peptides
P05030|PMA1	Plasma membrane ATPase 1	1.02	0.25	0.61	0.29	0.61	0.28	0.58	0.24	145
P19657|PMA2	Plasma membrane ATPase 2	1.00	0.24	0.62	0.29	0.60	0.27	0.60	0.23	91
P02994|EF1A	Elongation factor 1-alpha	1.01	0.26	0.60	0.29	0.56	0.20	0.60	0.27	77
P32324|EF2	Elongation factor 2	0.98	0.22	0.33	0.12	0.32	0.10	0.34	0.12	67
P39004|HXT7	High-affinity hexose transporter HXT6	1.01	0.29	1.72	0.97	1.57	0.72	1.58	0.81	55
P39003|HXT6	High-affinity hexose transporter HXT6	1.02	0.29	1.73	0.98	1.59	0.73	1.59	0.82	54
P38695|HXT5	Probable glucose transporter HXT5	0.99	0.25	1.65	1.08	1.46	0.65	1.55	0.88	40
P32466|HXT3	Low-affinity glucose transporter HXT3	0.99	0.23	1.10	0.69	1.07	0.63	1.06	0.65	40
P32467|HXT4	Low-affinity glucose transporter HXT4	0.98	0.23	1.16	0.66	1.13	0.59	1.13	0.61	38
P10592|HSP72	Heat shock protein SSA2	0.95	0.19	0.45	0.18	0.47	0.19	0.47	0.16	35
P10591|HSP71	Heat shock protein SSA1	0.93	0.21	0.51	0.21	0.51	0.20	0.50	0.17	35
P32465|HXT1	Low-affinity glucose transporter HXT1	0.96	0.25	1.42	1.01	1.35	0.72	1.32	0.88	28
P22146|GAS1	1,3-beta-glucanosyltransferase GAS1	0.98	0.25	0.95	0.28	0.85	0.23	0.94	0.31	26
Q12117|MRH1	Protein MRH1	0.99	0.26	0.67	0.39	0.71	0.44	0.76	0.57	23
Q08193|GAS5	1,3-beta-glucanosyltransferase GAS5	1.07	0.29	0.69	0.22	0.69	0.21	0.67	0.17	23
P33302|PDR5	Pleiotropic ABC efflux transporter of multiple drugs	1.01	0.25	0.46	0.25	0.44	0.30	0.45	0.27	21
P53252|PIL1	Sphingolipid long chain base-responsive protein PIL1	0.95	0.24	0.28	0.16	0.30	0.17	0.29	0.13	20
P38079|YRO2	Protein YRO2	0.98	0.21	0.68	0.26	0.65	0.19	0.67	0.20	20
P23585|HXT2	High-affinity glucose transporter HXT2	0.98	0.23	1.21	0.71	1.06	0.43	1.18	0.63	19
P54862|HXT11	Hexose transporter HXT11	0.91	0.20	1.22	0.49	1.32	0.57	1.17	0.50	19
P40885|HXT9	Hexose transporter HXT9	0.91	0.20	1.22	0.49	1.32	0.57	1.17	0.50	19
P40441|YIR0	Putative transporter-like protein YIL170W	0.93	0.20	1.27	0.49	1.36	0.59	1.22	0.50	17
P09435|HSP73	Heat shock protein SSA3	0.90	0.17	0.50	0.20	0.47	0.14	0.48	0.17	16
P40886|HXT8	Hexose transporter HXT8	0.93	0.21	1.31	0.48	1.41	0.57	1.26	0.48	16

P22202|HSP74	Heat shock protein SSA4	0.97	0.21	0.55	0.20	0.52	0.20	0.53	0.18	15
Q04182|PDR15	ATP-dependent permease PDR15	1.19	0.35	1.71	1.54	1.78	1.93	1.65	1.53	9
Q12230|LSP1	Sphingolipid long chain base-responsive protein LSP1	1.02	0.23	0.26	0.14	0.25	0.10	0.37	0.14	12
P11484|HSP75	Heat shock protein SSB1	0.97	0.23	0.44	0.16	0.46	0.15	0.46	0.18	12
P16521|EF3A	Elongation factor 3A	0.87	0.11	0.24	0.12	0.26	0.08	0.27	0.13	12
P40150|HSP76	Heat shock protein SSB2	0.97	0.22	0.43	0.16	0.45	0.14	0.45	0.20	11
Q07824|TPO1	Polyamine transporter 1	1.06	0.23	1.23	0.51	1.08	0.46	1.17	0.49	11
P39105|PLB1	Lysophospholipase 1	0.98	0.18	0.57	0.18	0.63	0.12	0.61	0.15	11
P00890|CISY1	Citrate synthase, mitochondrial	0.83	0.09	2.88	1.50	3.22	2.04	2.85	1.26	9
A6ZL22|ECM33	Cell wall protein ECM33 (strain YJM789)	0.96	0.16	0.79	0.27	0.71	0.10	0.69	0.23	10
P38248|ECM33	Cell wall protein ECM33	0.96	0.16	0.79	0.27	0.71	0.10	0.69	0.23	10
P36008|EF1G2	Elongation factor 1-gamma 2	1.03	0.17	0.44	0.11	0.44	0.18	0.42	0.21	10
P00560|PGK	Phosphoglycerate kinase	0.96	0.21	2.06	0.79	1.68	0.24	1.86	0.68	9
P38968|SEC31	Protein transport protein SEC31	1.11	0.27	0.47	0.14	0.42	0.09	0.51	0.12	9
P12398|HSP77	Heat shock protein SSC1, mitochondrial	0.87	0.17	0.41	0.09	0.37	0.08	0.39	0.11	8
P38216|YBM6	Uncharacterized protein YBR016W	0.99	0.31	0.55	0.27	0.59	0.28	0.51	0.35	8
P02365|RS6	40S ribosomal protein S6	1.04	0.27	0.61	0.24	0.54	0.12	0.55	0.22	8
Q03655|GAS3	Probable 1,3-beta-glucanosyltransferase GAS3	0.98	0.24	0.68	0.23	0.64	0.24	0.59	0.16	8
P39926|SSO2	Protein SSO2	0.98	0.27	0.61	0.26	0.56	0.11	0.61	0.17	8
P36035|JEN1	Carboxylic acid transporter protein homolog	1.08	0.21	2.16	0.78	2.28	0.87	2.21	0.38	5
P25349|YCP4	Flavoprotein-like protein YCP4	1.07	0.30	0.76	0.26	0.73	0.23	0.77	0.20	8
P47068|BBC1	Myosin tail region-interacting protein MTI1	1.08	0.20	0.42	0.11	0.38	0.06	0.41	0.08	8
P25619|HSP30	30 kDa heat shock protein	1.05	0.19	0.80	0.34	0.89	0.23	0.98	0.44	7
P39730|IF2P	Eukaryotic translation initiation factor 5B	0.99	0.19	0.34	0.16	0.32	0.22	0.29	0.22	7
P15891|ABP1	Actin-binding protein	1.01	0.20	0.34	0.14	0.32	0.05	0.33	0.11	7
B3RHV0|RS3A1	40S ribosomal protein S1-A (strain RM11-1a)	1.08	0.34	0.52	0.19	0.51	0.20	0.51	0.21	7
B3LLJ2|RS3A2	40S ribosomal protein S1-B (strain RM11-1a)	1.06	0.42	0.48	0.19	0.44	0.09	0.45	0.21	7
Q03482|YD210	Uncharacterized protein YDR210W	1.13	0.30	0.42	0.27	0.50	0.33	0.56	0.20	6
P39987|HSP7E	Heat shock protein SSC3, mitochondrial	0.91	0.18	0.38	0.09	0.36	0.18	0.36	0.10	6

P60010|ACT	Actin	0.93	0.16	0.52	0.08	0.47	0.11	0.51	0.17	6
P23291|KC11	Casein kinase I homolog 1	1.01	0.19	0.49	0.28	0.66	0.16	0.54	0.29	6
P29547|EF1G1	Elongation factor 1-gamma 1	1.16	0.22	0.52	0.12	0.59	0.15	0.58	0.20	6
Q08972|NEW1	[NU+] prion formation protein 1	0.87	0.20	0.41	0.14	0.38	0.10	0.35	0.07	6
P61864|UBIQ	Ubiquitin	1.14	0.25	1.07	0.30	1.03	0.12	1.02	0.15	6
P13181|GAL2	Galactose transporter	0.83	0.10	0.95	0.43	0.87	0.22	0.95	0.34	6
P23292|KC12	Casein kinase I homolog 2	1.06	0.31	0.60	0.10	0.71	0.09	0.76	0.21	6
P46992|YJR1	Uncharacterized protein YJL171C	1.06	0.47	0.45	0.17	0.42	0.04	0.41	0.12	5
P25294|SIS1	Protein SIS1	1.08	0.20	0.45	0.08	0.37	0.16	0.41	0.12	5
P32568|SNQ2	Protein SNQ2	0.96	0.19	0.35	0.13	0.24	0.06	0.35	0.11	5
P05453|ERF3	Eukaryotic peptide chain release factor GTP-binding subunit	0.78	0.08	0.37	0.11	0.42	0.03	0.49	0.14	5
Q08745|RS10A	40S ribosomal protein S10-A	1.18	0.34	0.35	0.17	0.31	0.21	0.38	0.19	5
P51533|PDR10	ATP-dependent permease PDR10	1.15	0.19	1.58	2.12	1.96	3.07	1.67	2.22	4
P38427|TSL1	Trehalose synthase complex regulatory subunit TSL1	0.89	0.11	0.39	0.25	0.26	0.25	0.34	0.19	5
P33892|GCN1	Translational activator GCN1	0.96	0.15	0.40	0.06	0.46	0.09	0.38	0.10	5
P52910|ACS2	Acetyl-coenzyme A synthetase 2	1.06	0.43	0.41	0.42	0.40	0.37	0.54	0.35	5
Q05931|HSP7Q	Heat shock protein SSQ1, mitochondrial	0.84	0.14	0.43	0.12	0.41	0.05	0.41	0.15	4
P87284|PMP3	Plasma membrane proteolipid 3	1.31	0.62	0.83	0.35	0.80	0.24	0.75	0.38	4
Q12256|TPO4	Polyamine transporter 4	1.18	0.54	0.86	0.22	1.02	0.31	0.88	0.23	4
P34167|IF4B	Eukaryotic translation initiation factor 4B	0.84	0.16	0.33	0.10	0.45	0.12	0.53	0.26	4
P00549|KPYK1	Pyruvate kinase 1	0.92	0.19	0.22	0.06	0.23	0.12	0.26	0.13	4
P16474|GRP78	78 kDa glucose-regulated protein homolog	0.86	0.10	0.37	0.10	0.46	0.20	0.38	0.13	4
P32457|CDC3	Cell division control protein 3	1.31	0.16	0.49	0.58	0.99	0.47	0.86	0.60	4
Q07478|SUB2	ATP-dependent RNA helicase SUB2	0.87	0.12	0.63	0.34	0.49	0.13	0.45	0.23	4
P46784|RS10B	40S ribosomal protein S10-B	1.15	0.38	0.33	0.19	0.32	0.24	0.37	0.22	4
P54003|SUR7	Protein SUR7	0.90	0.07	0.69	0.14	0.58	0.10	0.76	0.25	4
P38219|OLA1	Uncharacterized GTP-binding protein OLA1	1.00	0.16	0.49	0.20	0.35	0.14	0.42	0.14	4
Q3E795|YL361	Uncharacterized protein YLR361C-A	1.09	0.43	0.90	0.24	0.87	0.25	0.98	0.28	4
P25491|MAS5	Mitochondrial protein import protein MAS5	0.89	0.06	0.49	0.14	0.51	0.12	0.38	0.05	4

Q3E792|RS25A	40S ribosomal protein S25-A	0.97	0.30	0.46	0.17	0.47	0.09	0.43	0.13	4
P35271|RS18	40S ribosomal protein S18	1.15	0.24	0.49	0.05	0.48	0.11	0.52	0.16	4
Q02785|PDR12	ATP-dependent permease PDR12	0.83	0.09	0.21	0.11	0.19	0.09	0.25	0.11	4
P24276|SSD1	Protein SSD1	0.79	0.17	0.53	0.25	0.46	0.14	0.59	0.30	4
P00830|ATPB	ATP synthase subunit beta, mitochondrial	1.11	0.28	0.73	0.16	0.62	0.20	0.64	0.23	4
P15019|TAL1	Transaldolase	0.96	0.22	0.50	0.15	0.44	0.10	0.46	0.10	4
P32867|SSO1	Protein SSO1	0.85	0.09	0.55	0.16	0.46	0.10	0.53	0.24	3
P39935|IF4F1	Eukaryotic initiation factor 4F subunit p150	1.04	0.15	0.40	0.07	0.39	0.04	0.40	0.12	3
P40159|YNU8	Uncharacterized protein YNL208W	0.75	0.08	0.61	0.35	0.44	0.10	0.45	0.23	3
P00925|ENO2	Enolase 2	1.37	0.07	0.40	0.10	0.36	0.15	0.39	0.11	3
P32329|YPS1	Aspartic proteinase 3	0.86	0.12	0.50	0.12	0.47	0.08	0.58	0.20	3
P38788|SSZ1	Ribosome-associated complex subunit SSZ1	0.90	0.08	0.47	0.20	0.41	0.13	0.38	0.14	3
P01120|RAS2	Ras-like protein 2	1.16	0.30	1.58	1.17	1.33	0.72	1.43	0.93	3
P53894|CBK1	Serine/threonine-protein kinase CBK1	0.90	0.11	0.37	0.36	0.63	0.20	0.30	0.27	3
P33417|IXR1	Intrastrand cross-link recognition protein	0.92	0.07	0.40	0.37	0.64	0.19	0.34	0.30	3
Q12489|YD012	Uncharacterized protein YDL012C	0.75	0.10	0.37	0.15	0.34	0.08	0.34	0.04	3
Q06451|TPO3	Polyamine transporter 3	1.35	0.37	0.76	0.19	0.71	0.13	0.89	0.14	3
P00330|ADH1	Alcohol dehydrogenase 1	1.04	0.51	0.80	0.15	0.72	0.35	0.84	0.25	3
P32356|TREA	Neutral trehalase	0.90	0.14	0.33	0.06	0.41	0.08	0.41	0.03	3
P13587|ATN1	Sodium transport ATPase 1	1.19	0.06	0.14	0.07	0.12	0.05	0.15	0.04	3
P07246|ADH3	Alcohol dehydrogenase 3, mitochondrial	1.04	0.51	0.80	0.15	0.72	0.35	0.84	0.25	3
P53049|YOR1	Oligomycin resistance ATP-dependent permease YOR1	0.73	0.13	0.28	0.06	0.34	0.16	0.35	0.21	3
Q12460|NOP56	Nucleolar protein 56	1.14	0.24	0.43	0.09	0.50	0.16	0.49	0.07	3
Q04491|SEC13	Protein transport protein SEC13	1.09	0.23	0.60	0.21	0.62	0.20	0.64	0.07	3
P43581|HXT10	Hexose transporter HXT10	0.76	0.06	0.86	0.13	0.95	0.06	0.91	0.29	3
P38701|RS20	40S ribosomal protein S20	1.07	0.45	0.65	0.20	0.58	0.15	0.60	0.20	3
P04806|HXKA	Hexokinase-1	1.22	0.22	0.82	0.18	0.75	0.13	0.87	0.18	3
P04807|HXKB	Hexokinase-2	0.76	0.71	0.39	0.41	0.40	0.40	0.44	0.40	3
Q01896|ATN2	Sodium transport ATPase 2	1.19	0.06	0.14	0.07	0.12	0.05	0.15	0.04	3

A6ZP43|MCA1	Metacaspase-1 (strain YJM789)	1.06	0.21	0.64	0.07	0.52	0.06	0.61	0.09	3
P04451|RL23	60S ribosomal protein L23	0.90	0.13	0.40	0.16	0.41	0.16	0.38	0.19	3
P49573|CTR1	Copper transport protein CTR1	1.02	0.15	2.62	1.09	2.68	0.13	2.49	1.11	3
P40485|SLM1	Phosphatidylinositol 4,5-bisphosphate-binding protein SLM1	0.92	0.07	0.44	0.38	0.63	0.20	0.27	0.26	3
P08679|CISY2	Citrate synthase, peroxisomal	0.74	0.09	2.25	1.20	1.92	1.46	2.07	1.01	3
P25613|ADY2	Accumulation of dyads protein 2	1.32	0.72	1.84	0.25	1.97	0.53	1.81	0.02	2
A6ZPE5|NOP58	Nucleolar protein 58 (strain YJM789)	1.11	0.21	0.44	0.10	0.58	0.22	0.50	0.07	3
A6ZMG6|MYO5	Myosin-5 (strain YJM789)	1.00	0.15	0.73	0.13	0.59	0.17	0.75	0.11	3
Q12122|HOSM	Homocitrate synthase, mitochondrial	1.20	0.32	0.56	0.17	0.61	0.19	0.58	0.18	3
P48570|HOSC	Homocitrate synthase, cytosolic isozyme	1.20	0.32	0.56	0.17	0.61	0.19	0.58	0.18	3
P00360|G3P1	Glyceraldehyde-3-phosphate dehydrogenase 1	1.19	0.34	0.58	0.38	0.52	0.30	0.47	0.12	3
P38249|EIF3A	Eukaryotic translation initiation factor 3 subunit A	0.89	0.03	0.21	0.02	0.26	0.12	0.29	0.02	2
P35732|YKF4	Uncharacterized protein YKL054C	0.96	0.01	0.36	0.51	0.74	0.11	0.26	0.36	2
Q12207|NCE2	Non-classical export protein 2	0.79	0.15	0.50	0.21	0.81	0.56	0.53	0.34	2
P06780|RHO1	GTP-binding protein RHO1	0.99	0.24	0.90	0.07	0.78	0.11	0.92	0.12	2
P53165|SGF73	SAGA-associated factor 73	0.96	0.01	0.36	0.51	0.74	0.11	0.26	0.36	2
P38873|KOG1	Target of rapamycin complex 1 subunit KOG1	0.96	0.01	0.36	0.51	0.74	0.11	0.26	0.36	2
Q08887|NDD1	Nuclear division defective protein 1	0.96	0.01	0.36	0.51	0.74	0.11	0.26	0.36	2
Q12518|ENT1	Epsin-1	0.96	0.01	0.36	0.51	0.74	0.11	0.26	0.36	2
Q00772|SLT2	Mitogen-activated protein kinase SLT2/MPK1	0.96	0.01	0.36	0.51	0.74	0.11	0.26	0.36	2
P53968|CRZ1	Transcriptional regulator CRZ1	0.96	0.01	0.36	0.51	0.74	0.11	0.26	0.36	2
P11746|MCM1	Pheromone receptor transcription factor	0.96	0.01	0.36	0.51	0.74	0.11	0.26	0.36	2
A6ZPJ1|EIF3B	Eukaryotic translation initiation factor 3 subunit B (strain YJM789)	1.00	0.23	0.46	0.28	0.56	0.28	0.42	0.18	2
P52911|EXG2	Glucan 1,3-beta-glucosidase 2	1.17	0.57	0.55	0.26	0.47	0.17	0.77	0.46	2
P35997|RS27A	40S ribosomal protein S27-A	1.10	0.41	0.50	0.12	0.66	0.05	0.67	0.06	2
P10823|GPA2	Guanine nucleotide-binding protein alpha-2 subunit	0.85	0.12	1.30	0.99	1.50	1.18	1.24	1.01	2
P40482|SEC24	Protein transport protein SEC24	1.05	0.26	0.57	0.02	0.57	0.10	0.55	0.01	2
Q03516|RSN1	Uncharacterized protein RSN1	1.10	0.49	0.51	0.03	0.35	0.01	0.45	0.10	2
Q99271|NAH1	Na(+)/H(+) antiporter	0.87	0.08	0.38	0.03	0.45	0.06	0.39	0.01	2

Q05050|YMS1	Uncharacterized protein YMR031C	1.20	0.04	0.58	0.32	0.46	0.16	0.44	0.01	2
A6ZQJ1|IF4A	ATP-dependent RNA helicase eIF4A (strain YJM789)	1.22	0.09	0.38	0.09	0.39	0.14	0.48	0.01	2
P14126|RL3	60S ribosomal protein L3	1.10	0.42	0.63	0.23	0.86	0.51	0.55	0.19	2
P32527|ZUO1	Zuotin	1.30	0.40	0.47	0.18	0.38	0.06	0.51	0.23	2
Q12213|RL7B	60S ribosomal protein L7-B	1.00	0.63	0.84	0.25	0.96	0.53	1.02	0.63	2
P05737|RL7A	60S ribosomal protein L7-A	1.00	0.63	0.84	0.25	0.96	0.53	1.02	0.63	2
P38720|6PGD1	6-phosphogluconate dehydrogenase, decarboxylating 1	1.11	0.43	0.88	0.36	0.81	0.47	0.82	0.37	2
P06367|RS14A	40S ribosomal protein S14-A	1.27	0.09	0.68	0.07	0.59	0.29	0.56	0.01	2
P40213|RS16	40S ribosomal protein S16	0.86	0.10	0.30	0.04	0.40	0.09	0.39	0.05	2
P32497|EIF3C	Eukaryotic translation initiation factor 3 subunit C	1.01	0.37	0.45	0.28	0.61	0.29	0.44	0.24	2
P15108|HSC82	ATP-dependent molecular chaperone HSC82	0.98	0.43	0.40	0.15	0.33	0.01	0.37	0.09	2
P01119|RAS1	Ras-like protein 1	1.01	0.47	0.71	0.31	0.59	0.26	0.73	0.28	2
P16140|VATB	V-type proton ATPase subunit B	1.05	0.15	0.18	0.26	0.22	0.31	0.21	0.29	2
P05753|RS4	40S ribosomal protein S4	0.92	0.12	0.60	0.14	0.49	0.05	0.59	0.02	2
P38631|FKS1	1,3-beta-glucan synthase component FKS1	0.93	0.23	0.41	0.15	0.55	0.10	0.63	0.12	2
P40989|FKS2	1,3-beta-glucan synthase component GSC2	0.94	0.23	0.47	0.23	0.61	0.18	0.72	0.25	2
P07278|KAPR	cAMP-dependent protein kinase regulatory subunit	0.82	0.07	0.47	0.04	0.55	0.04	0.54	0.21	2
Q07651|FMP45	SUR7 family protein FMP45	0.90	0.16	1.51	0.33	1.69	0.60	1.42	0.36	2
P18480|SNF5	SWI/SNF chromatin-remodeling complex subunit SNF5	0.96	0.01	0.36	0.51	0.74	0.11	0.26	0.36	2
P46974|ZMS1	Zinc finger protein ZMS1	0.96	0.01	0.36	0.51	0.74	0.11	0.26	0.36	2
P38856|AP18A	Clathrin coat assembly protein AP180A	0.96	0.01	0.36	0.51	0.74	0.11	0.26	0.36	2
P43572|EPL1	Enhancer of polycomb-like protein 1	0.96	0.01	0.36	0.51	0.74	0.11	0.26	0.36	2
P48562|CLA4	Serine/threonine-protein kinase CLA4	0.96	0.01	0.36	0.51	0.74	0.11	0.26	0.36	2
Q08831|VTS1	Protein VTS1	0.96	0.01	0.36	0.51	0.74	0.11	0.26	0.36	2
P39008|POP2	Poly(A) ribonuclease POP2	0.96	0.01	0.36	0.51	0.74	0.11	0.26	0.36	2
P39081|PCF11	Protein PCF11	0.96	0.01	0.36	0.51	0.74	0.11	0.26	0.36	2
Q99395|YP229	Uncharacterized protein YPL229W	0.96	0.01	0.36	0.51	0.74	0.11	0.26	0.36	2
P08539|GPA1	Guanine nucleotide-binding protein alpha-1 subunit	0.74	0.02	0.25	0.05	0.31	0.09	0.28	0.01	2

P39940|RSP5	E3 ubiquitin-protein ligase RSP5	0.83	0.04	0.38	0.08	0.50	0.00	0.30	0.01	2
P46943|GUF1	GTP-binding protein GUF1	1.41	0.07	0.39	0.06	0.48	0.07	0.43	0.02	2
P38993|FET3	Iron transport multicopper oxidase FET3	1.09	0.54	0.60	0.22	0.60	0.19	0.55	0.23	2
P32623|CRH2	Probable glycosidase CRH2	0.80	0.15	0.40	0.00	0.25	0.04	0.29	0.06	2
P32473|ODPB	Pyruvate dehydrogenase E1 component subunit beta, mitochondrial	1.14	0.38	0.82	0.18	0.57	0.07	0.60	0.02	2
P25087|ERG6	Sterol 24-C-methyltransferase	1.11	0.31	0.51	0.12	0.61	0.13	0.78	0.12	2
P32599|FIMB	Fimbrin	1.08	0.42	0.40	0.08	0.52	0.34	0.42	0.15	2
P02405|RL44	60S ribosomal protein L42	1.21	0.40	0.62	0.07	0.59	0.10	0.79	0.04	2
P05317|RLA0	60S acidic ribosomal protein P0	0.81	0.03	0.29	0.08	0.24	0.06	0.24	0.05	2
P52917|VPS4	Vacuolar protein sorting-associated protein 4	1.25	0.53	0.25	0.03	0.33	0.06	0.25	0.09	2
Q12359|ATO3	Ammonia transport outward protein 3	1.09	0.09	2.13	0.41	2.23	0.92	1.66	0.58	2
P26782|RS24	40S ribosomal protein S24	1.24	0.33	0.40	0.00	0.34	0.06	0.40	0.01	2
Q12691|ATN5	Sodium transport ATPase 5	1.19	0.08	0.18	0.03	0.14	0.05	0.17	0.01	2
P05755|RS9B	40S ribosomal protein S9-B	0.87	0.07	0.53	0.07	0.58	0.04	0.50	0.07	2
O13516|RS9A	40S ribosomal protein S9-A	0.87	0.07	0.53	0.07	0.58	0.04	0.50	0.07	2
P32769|HBS1	Elongation factor 1 alpha-like protein	0.73	0.07	0.56	0.08	0.47	0.06	0.69	0.12	2
P15303|SEC23	Protein transport protein SEC23	0.78	0.12	0.34	0.00	0.41	0.04	0.37	0.01	2
P40440|YIR1	Putative transporter-like protein YIL171W	0.79	0.06	0.79	0.05	0.93	0.07	0.74	0.02	2
P38011|GBLP	Guanine nucleotide-binding protein subunit beta-like protein	0.75	0.01	0.34	0.03	0.55	0.25	0.49	0.18	2
Q12335|PST2	Protoplast secreted protein 2	1.05	0.29	1.04	0.44	0.89	0.23	0.98	0.37	2
P07342|ILVB	Acetolactate synthase catalytic subunit, mitochondrial	0.86	0.09	0.26	0.11	0.43	0.23	0.27	0.13	2
Q08108|PLB3	Lysophospholipase 3	1.03	0.07	0.51	0.03	0.69	0.15	0.70	0.13	2
P33322|CBF5	H/ACA ribonucleoprotein complex subunit 4	0.88	0.06	0.36	0.04	0.40	0.04	0.53	0.00	2
P02293|H2B1	Histone H2B.1	1.26	0.06	0.75	0.06	0.72	0.33	0.73	0.10	2
P05736|RL2	60S ribosomal protein L2	1.21	0.31	0.97	0.27	0.92	0.02	0.90	0.11	2
Q08954|YP199	Smr domain-containing protein YPL199C	0.88	0.01	0.22	0.31	0.31	0.04	0.43	0.11	2
P00359|G3P3	Glyceraldehyde-3-phosphate dehydrogenase 3	1.26	0.44	0.68	0.47	0.66	0.26	0.54	0.00	2
Q12449|AHA1	Hsp90 co-chaperone AHA1	1.21	0.02	0.77	0.02	0.54	0.09	0.81	0.22	2
P18852|GBG	Guanine nucleotide-binding protein subunit gamma	0.99	0.59	0.36	0.05	0.41	0.07	0.37	0.11	2

P38266|YBV8	Uncharacterized protein YBR108W	0.96	0.01	0.36	0.51	0.74	0.11	0.26	0.36	2
P40088|FTR1	Plasma membrane iron permease	1.34	0.02	0.95	0.01	1.09	0.42	0.97	0.02	2

Supplementry table S5 F (S5F) summary of the fractions, peptides sequence, calculated mass (pep_calc_mr), mass acuuracy (pep_delta), peptide score (pep_score), homology (pep_homol), identity threshold (pep_ident), the rank number (pep_Rank), hydrophobicity score (GRAVY_score) and isoelectric point (pI_Value) for yeast cells grown in glucose and oleic acid.

Fraction	pep_seq	pep_calc_mr	pep_delta	pep_score	pep_homol	pep_ident	GRAVY_Score	pI_Value
1	AIIDAL	843,5268	-0,0186	42,28		39	2,15	3,1
1	EGATDAA	862,4235	-0,0059	48,9		40	-0,385714286	3
1	TATGDNT	907,445	-0,0027	46,71		40	-1,1	3,1
1	GFVQEF	954,5014	0,0155	46,4		39	0,4	3,3
1	VEGATDAA	961,4919	0,0127	79,77		42	0,1875	3
1	SLDIDLI	1016,5957	0,0416	48,1		41	1,257142857	2,92
1	GLGGGGDMPG	1045,5066	-0,0132	57,62		38	-0,18	3,1
1	FYYEMS	1067,4837	0,0236	40,53		39	-0,366666667	3,3
1	ENWTDIV	1104,5654	-0,0222	52,21		40	-0,485714286	3
1	ATGDNTFVG	1109,5556	-0,0122	54,84	36	40	-0,044444444	3,1
1	IAYDNAPY	1154,5811	-0,0128	47,97	35	40	-0,3875	3,1
1	LSDWVDFG	1166,5811	0,0408	59,96	38	41	0,2125	2,92
1	NEVVPGDIL	1183,6651	-0,0152	53,52	38	41	0,466666667	3
1	SENWTDIV	1191,5975	-0,0171	54,19		40	-0,525	3
1	EEDHPIPE	1193,5767	0,0421	53,1	33	41	-1,9875	3,68
1	TATGDNTFVG	1210,6033	-0,0164	77,53	39	40	-0,11	3,1
1	GLSDWVDFG	1223,6026	0,0482	67,3	39	41	0,144444444	2,92
1	DGFAEVFPQ	1237,6182	0,0019	63,69	40	41	-0,1	3
1	SDWVDFGVI	1265,6495	-0,0014	50,35		41	0,733333333	2,92
1	LFGWWSEN	1266,6236	-0,0151	49,72		40	-0,425	3,3

1	MTGDGVNDAPS	1291,5917	0,0564	96,08		41	-0,590909091	2,92
1	SLTENWLIF	1350,7386	-0,0173	50,34	38	42	0,611111111	3,3
1	GLGGGGDMPGSEL	1374,6652	0,0154	106,24		40	-0,176923077	3
1	MTGDGVNDAPSL	1404,6758	-0,0215	81,93	34	38	-0,225	2,92
1	EEDHPIPEDV	1407,6721	-0,0284	58,25	28	38	-1,52	3,5
1	GLGGGGDMPGSELA	1445,7023	0,0014	107,6	38	40	-0,035714286	3
1	GLGGGGDMPGSELA	1461,6973	-0,0023	66,45	27	39	-0,035714286	3
1	ILDNSLDIDLI	1471,8336	-0,0348	94,71		41	0,918181818	2,82
1	AILDNSLDIDLI	1542,8708	-0,0351	113,86		42	0,991666667	2,82
1	EEDHPIPEDVH	1544,731	0,0485	61,01	25	41	-1,672727273	3,92
1	WWSENWTDIV	1563,7561	-0,0251	84,07		39	-0,6	3
1	GWWSENWTDIV	1620,7776	-0,0365	97,61	36	38	-0,581818182	3
1	VEEDHPIPEDVH	1643,7994	-0,0193	68,3	26	39	-1,183333333	3,92
1	FGWWSENWTDIV	1767,846	0,0446	95,48	35	41	-0,3	3
1	EEDHPIPEDVHEN	1787,8165	0,0712	88,16	30	41	-1,953846154	3,82
1	GEARPVPEEYLQTD	1831,9155	-0,0337	57,84	28	40	-1,271428571	3,68
1	LFGWWSENWTDIV	1880,93	-0,0223	50,49	32	40	0,015384615	3
1	VEEDHPIPEDVHEN	1886,8849	-0,0076	80,97	29	39	-1,514285714	3,82
1	RDGQLVEIPANEVVPG	1921,0472	0,0578	77,09	32	42	-0,20625	3,82
1	EEDHPIPEDVHENY	1950,8798	-0,0254	107,56	26	36	-1,907142857	3,82
1	EARPVPEEYLQTDPS	1958,9788	-0,0399	49,08	27	40	-1,32	3,68
1	GEARPVPEEYLQTDPS	2016,0003	-0,0252	72,97	35	40	-1,2625	3,68
1	VEEDHPIPEDVHENY	2049,9482	-0,0545	45,1	16	36	-1,5	3,82
1	EEDHPIPEDVHENYE	2079,9224	-0,0062	65,26	26	36	-2,013333333	3,74
1	EEDHPIPEDVHENYEN	2193,9653	0,01	90,43	28	37	-2,10625	3,74
1	VEEDHPIPEDVHENYEN	2293,0338	-0,0353	37,81	21	36	-1,735294118	3,74
1	AIIDAL	843,5268	-0,0186	42,28		39	2,15	3,1
1	EGATDAA	862,4235	-0,0059	48,9		40	-0,385714286	3
1	TATGDNT	907,445	-0,0027	46,71		40	-1,1	3,1
1	VEGATDAA	961,4919	0,0127	79,77		42	0,1875	3

1	GLGGGGDMPG	1045,5066	-0,0132	57,62		38	-0,18	3,1
1	ENWTDIV	1104,5654	-0,0222	52,21		40	-0,485714286	3
1	ATGDNTFVG	1109,5556	-0,0122	54,84	36	40	-0,044444444	3,1
1	IAYDNAPY	1154,5811	-0,0128	47,97	35	40	-0,3875	3,1
1	SENWTDIV	1191,5975	-0,0171	54,19		40	-0,525	3
1	EEDHPIPE	1193,5767	0,0421	53,1	33	41	-1,9875	3,68
1	TATGDNTFVG	1210,6033	-0,0164	77,53	39	40	-0,11	3,1
1	DGFAEVFPQ	1237,6182	0,0019	63,69	40	41	-0,1	3
1	LFGWWSEN	1266,6236	-0,0151	49,72		40	-0,425	3,3
1	MTGDGVNDAPS	1291,5917	0,0564	96,08		41	-0,590909091	2,92
1	SLTENWLIF	1350,7386	-0,0173	50,34	38	42	0,611111111	3,3
1	GLGGGGDMPGSEL	1374,6652	0,0154	106,24		40	-0,176923077	3
1	MTGDGVNDAPSL	1404,6758	-0,0215	81,93	34	38	-0,225	2,92
1	EEDHPIPEDV	1407,6721	-0,0284	58,25	28	38	-1,52	3,5
1	GLGGGGDMPGSELA	1445,7023	0,0014	107,6	38	40	-0,035714286	3
1	GLGGGGDMPGSELA	1461,6973	-0,0023	66,45	27	39	-0,035714286	3
1	EEDHPIPEDVH	1544,731	0,0485	61,01	25	41	-1,672727273	3,92
1	WWSENWTDIV	1563,7561	-0,0251	84,07		39	-0,6	3
1	GWWSENWTDIV	1620,7776	-0,0365	97,61	36	38	-0,581818182	3
1	VEEDHPIPEDVH	1643,7994	-0,0193	68,3	26	39	-1,183333333	3,92
1	FGWWSENWTDIV	1767,846	0,0446	95,48	35	41	-0,3	3
1	EEDHPIPEDVHEN	1787,8165	0,0712	88,16	30	41	-1,953846154	3,82
1	LFGWWSENWTDIV	1880,93	-0,0223	50,49	32	40	0,015384615	3
1	VEEDHPIPEDVHEN	1886,8849	-0,0076	80,97	29	39	-1,514285714	3,82
1	EEDHPIPEDVHENY	1950,8798	-0,0254	107,56	26	36	-1,907142857	3,82
1	VEEDHPIPEDVHENY	2049,9482	-0,0545	45,1	16	36	-1,5	3,82
1	EEDHPIPEDVHENYE	2079,9224	-0,0062	65,26	26	36	-2,013333333	3,74
1	EEDHPIPEDVHENYEN	2193,9653	0,01	90,43	28	37	-2,10625	3,74
1	VEEDHPIPEDVHENYEN	2293,0338	-0,0353	37,81	21	36	-1,735294118	3,74
1	DAIEQPS	987,5076	0,0228	44,45		41	-0,942857143	3

1	GGVGEFEA	993,497	-0,0192	61,19	38	40	0,075	3,12
1	GGVGEFEAG	1050,5185	-0,0181	52,66	30	39	0,022222222	3,12
1	AGGVGEFEA	1064,5341	-0,0044	72,36		41	0,266666667	3,12
1	AGGVGEFEAG	1121,5556	-0,0236	74,72		39	0,2	3,12
1	EQGVPGDNV	1142,5771	0,042	44,03	35	41	-0,888888889	3
1	GGVGEFEAGI	1163,6026	0,0267	63,62	41	42	0,47	3,12
1	EQLEQGVPG	1184,624	0,0022	69,05	35	40	-0,933333333	3,12
1	AGGVGEFEAGI	1234,6397	0,0187	95,19	41	42	0,590909091	3,12
1	GGVGEFEAGIS	1250,6346	0,0195	93,3		41	0,354545455	3,12
1	LEQGVPGDNV	1255,6611	-0,0307	60,27		40	-0,42	3
1	AGGVGEFEAGIS	1321,6717	0,0579	104,9	40	42	0,475	3,12
1	EQGVPGDNVGF	1346,667	-0,0125	55,18	29	40	-0,509090909	3
1	SGWNGDNMIEA	1421,6448	0,0175	87,67	34	38	-0,754545455	3
1	GWNGDNMIEAT	1435,6605	0,0171	71,54		39	-0,745454545	3
1	LEQGVPGDNVGF	1459,751	0,0298	79,99	36	42	-0,15	3
1	EQLEQGVPGDNV	1512,7623	-0,0316	96,77	35	40	-0,933333333	2,94
1	SGWNGDNMIEAT	1522,6925	-0,0011	42,09	23	38	-0,75	3
1	HEQLEQGVPGDNVGF	1853,9111	0,0185	108,38	32	41	-0,8	3,82
1	VDDPSVL	972,5331	-0,0074	49,88	36	40	0,4	2,92
1	GDNYFF	990,465	-0,0168	36,33	26	36	-0,516666667	3,1
1	QQLTGDN	1003,5137	0,0058	42,74		41	-1,614285714	3,1
1	VFGWDTG	1009,5072	-0,0214	41,07	31	40	0,157142857	3,1
1	AVDDPSVL	1043,5702	-0,016	49,21	38	40	0,575	2,92
1	TGDNYFF	1091,5127	-0,0197	47,27	34	37	-0,542857143	3,1
1	NYDAEEM	1099,4695	0,0193	45,15	35	36	-1,657142857	2,94
1	NYDAEEM	1115,4644	0,0044	39,83		33	-1,657142857	2,94
1	FGWDTGTI	1124,5705	-0,011	43,24	29	41	0,0875	3,1
1	SWGELFSS	1140,5654	-0,0023	48,9		40	-0,075	3,3
1	GDNYFFY	1153,5283	-0,0172	46,94	32	36	-0,628571429	3,1
1	FVFGWDTG	1156,5756	0,0259	56,35		40	0,4875	3,1

1	QQLTGDNY	1166,5771	0,0117	68,02		41	-1,575	3,1
1	GEGEEHEPV	1210,5669	0,007	62,32	33	39	-1,711111111	3,78
1	FGWDTGTIS	1211,6026	-0,0144	55,83	38	40	-0,011111111	3,1
1	GFVFGWDTG	1213,5971	0,0554	65,06	38	41	0,388888889	3,1
1	TGDNYFFY	1254,576	-0,0213	54,53	29	36	-0,6375	3,1
1	WPNGQDQPS	1256,5989	-0,0081	50,07	27	38	-2,144444444	3,1
1	GGFVFGWDTG	1270,6186	0,0468	65,53	34	41	0,31	3,1
1	GFINQTDFI	1282,6761	-0,0063	71,56	30	41	0,333333333	3,1
1	QQLTGDNYF	1313,6455	0,0324	60,57	33	41	-1,088888889	3,1
1	GDNYFFYY	1316,5916	0,0003	55,44	23	37	-0,7125	3,1
1	YGEGEEHEPV	1373,6302	0,0587	63,01	35	40	-1,67	3,78
1	MWEEGVLPW	1374,6845	-0,0127	54,24	30	40	-0,1	3,12
1	TGDNYFFYY	1417,6393	-0,0056	53,33		36	-0,711111111	3,1
1	FGGFVFGWDTG	1417,687	-0,0243	67,53	28	38	0,536363636	3,1
1	GFVFGWDTGTI	1427,7289	-0,021	75,9	37	40	0,663636364	3,1
1	AYGEGEEHEPV	1444,6673	0,0184	67,28	29	39	-1,354545455	3,78
1	FVFGWDTGTIS	1457,7394	-0,0017	71,44	32	41	0,627272727	3,1
1	GDNYFFYYGT	1474,6608	-0,0311	72,74	28	34	-0,68	3,1
1	GFVFGWDTGTIS	1514,7609	0,0271	89,13	36	41	0,541666667	3,1
1	TGDNYFFYYGT	1575,7085	-0,0307	62,71	32	34	-0,681818182	3,1
1	NTMWEEGVLPW	1589,7751	-0,0109	56,81	31	40	-0,463636364	3,12
1	QQLTGDNYFFY	1623,7772	-0,0282	78,85	30	38	-0,754545455	3,1
1	FGGFVFGWDTGTI	1631,8187	0,0545	92,84	29	42	0,746153846	3,1
1	YGEGEEHEPVVEI	1714,8253	0,0685	95,16		42	-0,884615385	3,66
1	FGGFVFGWDTGTIS	1718,8508	-0,0156	131,21	37	40	0,635714286	3,1
1	QQLTGDNYFFYY	1786,8405	0,0037	89,81	31	39	-0,8	3,1
1	QQLTGDNYFFYYGT	1944,9097	-0,0342	126,24	31	37	-0,764285714	3,1
1	VDDPSVL	972,5331	-0,0074	49,88	36	40	0,4	2,92
1	GDNYFF	990,465	-0,0168	36,33	26	36	-0,516666667	3,1
1	QQLTGDN	1003,5137	0,0058	42,74		41	-1,614285714	3,1

1	VFGWDTG	1009,5072	-0,0214	41,07	31	40	0,157142857	3,1
1	AVDDPSVL	1043,5702	-0,016	49,21	38	40	0,575	2,92
1	TGDNYFF	1091,5127	-0,0197	47,27	34	37	-0,542857143	3,1
1	NYDAEEM	1099,4695	0,0193	45,15	35	36	-1,657142857	2,94
1	NYDAEEM	1115,4644	0,0044	39,83		33	-1,657142857	2,94
1	FGWDTGTI	1124,5705	-0,011	43,24	29	41	0,0875	3,1
1	SWGELFSS	1140,5654	-0,0023	48,9		40	-0,075	3,3
1	GDNYFFY	1153,5283	-0,0172	46,94	32	36	-0,628571429	3,1
1	FVFGWDTG	1156,5756	0,0259	56,35		40	0,4875	3,1
1	QQLTGDNY	1166,5771	0,0117	68,02		41	-1,575	3,1
1	GEGEEHEPV	1210,5669	0,007	62,32	33	39	-1,711111111	3,78
1	FGWDTGTIS	1211,6026	-0,0144	55,83	38	40	-0,011111111	3,1
1	GFVFGWDTG	1213,5971	0,0554	65,06	38	41	0,388888889	3,1
1	TGDNYFFY	1254,576	-0,0213	54,53	29	36	-0,6375	3,1
1	WPNGQDQPS	1256,5989	-0,0081	50,07	27	38	-2,144444444	3,1
1	GGFVFGWDTG	1270,6186	0,0468	65,53	34	41	0,31	3,1
1	GFINQTDFI	1282,6761	-0,0063	71,56	30	41	0,333333333	3,1
1	QQLTGDNYF	1313,6455	0,0324	60,57	33	41	-1,088888889	3,1
1	GDNYFFYY	1316,5916	0,0003	55,44	23	37	-0,7125	3,1
1	YGEGEEHEPV	1373,6302	0,0587	63,01	35	40	-1,67	3,78
1	MWEEGVLPW	1374,6845	-0,0127	54,24	30	40	-0,1	3,12
1	TGDNYFFYY	1417,6393	-0,0056	53,33		36	-0,711111111	3,1
1	FGGFVFGWDTG	1417,687	-0,0243	67,53	28	38	0,536363636	3,1
1	GFVFGWDTGTI	1427,7289	-0,021	75,9	37	40	0,663636364	3,1
1	AYGEGEEHEPV	1444,6673	0,0184	67,28	29	39	-1,354545455	3,78
1	FVFGWDTGTIS	1457,7394	-0,0017	71,44	32	41	0,627272727	3,1
1	GDNYFFYYGT	1474,6608	-0,0311	72,74	28	34	-0,68	3,1
1	GFVFGWDTGTIS	1514,7609	0,0271	89,13	36	41	0,541666667	3,1
1	TGDNYFFYYGT	1575,7085	-0,0307	62,71	32	34	-0,681818182	3,1
1	NTMWEEGVLPW	1589,7751	-0,0109	56,81	31	40	-0,463636364	3,12

1	QQLTGDNYFFY	1623,7772	-0,0282	78,85	30	38	-0,754545455	3,1
1	FGGFVFGWDTGTI	1631,8187	0,0545	92,84	29	42	0,746153846	3,1
1	YGEGEEHEPVVEI	1714,8253	0,0685	95,16		42	-0,884615385	3,66
1	FGGFVFGWDTGTIS	1718,8508	-0,0156	131,21	37	40	0,635714286	3,1
1	QQLTGDNYFFYY	1786,8405	0,0037	89,81	31	39	-0,8	3,1
1	QQLTGDNYFFYYGT	1944,9097	-0,0342	126,24	31	37	-0,764285714	3,1
1	EGPADDA	902,4184	-0,0016	35,72		35	-1,271428571	2,88
1	VESESSQ	993,4818	-0,0117	48,08	38	39	-1,242857143	3,12
1	DGNSFLI	993,5334	-0,0121	47,03	36	42	0,414285714	3,1
1	DFSSEVT	1012,4916	-0,0106	52,4	37	38	-0,328571429	3
1	GIDQFLL	1033,6011	0,015	48,58	39	42	1,071428571	3,1
1	VESESSQT	1094,5294	-0,0134	59,57		39	-1,175	3,12
1	EPIDEEVS	1145,5655	-0,0051	47,56		39	-0,9625	2,88
1	YEGPADDAN	1179,5247	0,0107	46,38	25	37	-1,522222222	2,88
1	LGDVQVYPA	1189,6546	0,0393	66,36	34	42	0,411111111	3,1
1	FGPDGNGPNLV	1314,6771	0,0207	64,78		41	-0,372727273	3,1
1	NESFGFTGEL	1328,6451	-0,0241	85,33		38	-0,34	3,12
1	FNMFILDPI	1337,7256	0,044	59,07	40	42	1,3	3,1
1	EQLYEGPADDA	1435,667	-0,0289	70,04	31	36	-1,218181818	2,84
1	RIMADDYGWDVT	1669,7973	-0,0211	56,82	31	39	-0,491666667	3,6
1	SEMSDEDVKEI	1738,8701	-0,0203	62,86	36	41	-1,127272727	3,5
1	GDNYFF	990,465	-0,0168	36,33	26	36	-0,516666667	3,1
1	GYNDNLA	994,4922	0,0126	52,7	38	41	-0,942857143	3,1
1	QQLTGDN	1003,5137	0,0058	42,74		41	-1,614285714	3,1
1	VFGWDTG	1009,5072	-0,0214	41,07	31	40	0,157142857	3,1
1	TGDNYFF	1091,5127	-0,0197	47,27	34	37	-0,542857143	3,1
1	FGWDTGTI	1124,5705	-0,011	43,24	29	41	0,0875	3,1
1	GDNYFFY	1153,5283	-0,0172	46,94	32	36	-0,628571429	3,1
1	NEMYEEN	1156,4909	0,0296	46,47	30	36	-2,414285714	3,02
1	FVFGWDTG	1156,5756	0,0259	56,35		40	0,4875	3,1

1	QQLTGDNY	1166,5771	0,0117	68,02		41	-1,575	3,1
1	FGWDTGTIS	1211,6026	-0,0144	55,83	38	40	-0,011111111	3,1
1	GFVFGWDTG	1213,5971	0,0554	65,06	38	41	0,388888889	3,1
1	TGDNYFFY	1254,576	-0,0213	54,53	29	36	-0,6375	3,1
1	WPNGQDQPS	1256,5989	-0,0081	50,07	27	38	-2,144444444	3,1
1	GGFVFGWDTG	1270,6186	0,0468	65,53	34	41	0,31	3,1
1	QQLTGDNYF	1313,6455	0,0324	60,57	33	41	-1,088888889	3,1
1	GDNYFFYY	1316,5916	0,0003	55,44	23	37	-0,7125	3,1
1	TGDNYFFYY	1417,6393	-0,0056	53,33		36	-0,711111111	3,1
1	FGGFVFGWDTG	1417,687	-0,0243	67,53	28	38	0,536363636	3,1
1	GFVFGWDTGTI	1427,7289	-0,021	75,9	37	40	0,663636364	3,1
1	FVFGWDTGTIS	1457,7394	-0,0017	71,44	32	41	0,627272727	3,1
1	GDNYFFYYGT	1474,6608	-0,0311	72,74	28	34	-0,68	3,1
1	GFVFGWDTGTIS	1514,7609	0,0271	89,13	36	41	0,541666667	3,1
1	TGDNYFFYYGT	1575,7085	-0,0307	62,71	32	34	-0,681818182	3,1
1	QQLTGDNYFFY	1623,7772	-0,0282	78,85	30	38	-0,754545455	3,1
1	FGGFVFGWDTGTI	1631,8187	0,0545	92,84	29	42	0,746153846	3,1
1	FGGFVFGWDTGTIS	1718,8508	-0,0156	131,21	37	40	0,635714286	3,1
1	QQLTGDNYFFYY	1786,8405	0,0037	89,81	31	39	-0,8	3,1
1	QQLTGDNYFFYYGT	1944,9097	-0,0342	126,24	31	37	-0,764285714	3,1
1	GDNYFF	990,465	-0,0168	36,33	26	36	-0,516666667	3,1
1	QQLTGDN	1003,5137	0,0058	42,74		41	-1,614285714	3,1
1	VFGWDTG	1009,5072	-0,0214	41,07	31	40	0,157142857	3,1
1	TGDNYFF	1091,5127	-0,0197	47,27	34	37	-0,542857143	3,1
1	FGWDTGTI	1124,5705	-0,011	43,24	29	41	0,0875	3,1
1	GDNYFFY	1153,5283	-0,0172	46,94	32	36	-0,628571429	3,1
1	FVFGWDTG	1156,5756	0,0259	56,35		40	0,4875	3,1
1	QQLTGDNY	1166,5771	0,0117	68,02		41	-1,575	3,1
1	FGWDTGTIS	1211,6026	-0,0144	55,83	38	40	-0,011111111	3,1
1	GFVFGWDTG	1213,5971	0,0554	65,06	38	41	0,388888889	3,1

1	TGDNYFFY	1254,576	-0,0213	54,53	29	36	-0,6375	3,1
1	GGFVFGWDTG	1270,6186	0,0468	65,53	34	41	0,31	3,1
1	QQLTGDNYF	1313,6455	0,0324	60,57	33	41	-1,088888889	3,1
1	GDNYFFYY	1316,5916	0,0003	55,44	23	37	-0,7125	3,1
1	TGDNYFFYY	1417,6393	-0,0056	53,33		36	-0,711111111	3,1
1	FGGFVFGWDTG	1417,687	-0,0243	67,53	28	38	0,536363636	3,1
1	GFVFGWDTGTI	1427,7289	-0,021	75,9	37	40	0,663636364	3,1
1	FVFGWDTGTIS	1457,7394	-0,0017	71,44	32	41	0,627272727	3,1
1	GDNYFFYYGT	1474,6608	-0,0311	72,74	28	34	-0,68	3,1
1	GFVFGWDTGTIS	1514,7609	0,0271	89,13	36	41	0,541666667	3,1
1	NDMYAEGVLPW	1522,7329	0,0348	64,77	26	41	-0,272727273	3
1	TGDNYFFYYGT	1575,7085	-0,0307	62,71	32	34	-0,681818182	3,1
1	QQLTGDNYFFY	1623,7772	-0,0282	78,85	30	38	-0,754545455	3,1
1	FGGFVFGWDTGTI	1631,8187	0,0545	92,84	29	42	0,746153846	3,1
1	FGGFVFGWDTGTIS	1718,8508	-0,0156	131,21	37	40	0,635714286	3,1
1	QQLTGDNYFFYY	1786,8405	0,0037	89,81	31	39	-0,8	3,1
1	QQLTGDNYFFYYGT	1944,9097	-0,0342	126,24	31	37	-0,764285714	3,1
1	NMPEEKGVQDDFQAEADQV	2607,2528	0,0543	48,34	41	42	-1,273684211	3,38
1	MNMPEEKGVQDDFQAEADQV	2738,2933	0,0265	67,7	29	41	-1,115	3,38
1	GDNYFF	990,465	-0,0168	36,33	26	36	-0,516666667	3,1
1	QQLTGDN	1003,5137	0,0058	42,74		41	-1,614285714	3,1
1	VFGWDTG	1009,5072	-0,0214	41,07	31	40	0,157142857	3,1
1	AYQEDTA	1025,4868	0,0027	58,23		39	-1,271428571	3
1	TGDNYFF	1091,5127	-0,0197	47,27	34	37	-0,542857143	3,1
1	FGWDTGTI	1124,5705	-0,011	43,24	29	41	0,0875	3,1
1	GDNYFFY	1153,5283	-0,0172	46,94	32	36	-0,628571429	3,1
1	FVFGWDTG	1156,5756	0,0259	56,35		40	0,4875	3,1
1	QQLTGDNY	1166,5771	0,0117	68,02		41	-1,575	3,1
1	FGWDTGTIS	1211,6026	-0,0144	55,83	38	40	-0,011111111	3,1
1	GFVFGWDTG	1213,5971	0,0554	65,06	38	41	0,388888889	3,1

1	TGDNYFFY	1254,576	-0,0213	54,53	29	36	-0,6375	3,1
1	GGFVFGWDTG	1270,6186	0,0468	65,53	34	41	0,31	3,1
1	QQLTGDNYF	1313,6455	0,0324	60,57	33	41	-1,088888889	3,1
1	GDNYFFYY	1316,5916	0,0003	55,44	23	37	-0,7125	3,1
1	LWEEGVLPW	1356,7281	-0,0261	41,02	32	41	0,111111111	3,12
1	TGDNYFFYY	1417,6393	-0,0056	53,33		36	-0,711111111	3,1
1	FGGFVFGWDTG	1417,687	-0,0243	67,53	28	38	0,536363636	3,1
1	GFVFGWDTGTI	1427,7289	-0,021	75,9	37	40	0,663636364	3,1
1	FVFGWDTGTIS	1457,7394	-0,0017	71,44	32	41	0,627272727	3,1
1	GDNYFFYYGT	1474,6608	-0,0311	72,74	28	34	-0,68	3,1
1	GFVFGWDTGTIS	1514,7609	0,0271	89,13	36	41	0,541666667	3,1
1	TGDNYFFYYGT	1575,7085	-0,0307	62,71	32	34	-0,681818182	3,1
1	QQLTGDNYFFY	1623,7772	-0,0282	78,85	30	38	-0,754545455	3,1
1	FGGFVFGWDTGTI	1631,8187	0,0545	92,84	29	42	0,746153846	3,1
1	FGGFVFGWDTGTIS	1718,8508	-0,0156	131,21	37	40	0,635714286	3,1
1	QQLTGDNYFFYY	1786,8405	0,0037	89,81	31	39	-0,8	3,1
1	QQLTGDNYFFYYGT	1944,9097	-0,0342	126,24	31	37	-0,764285714	3,1
1	GAPEGAAPG	954,4973	-0,0124	76,26	36	39	-0,277777778	3,3
1	YADNQPG	992,4766	-0,0075	55,32	30	39	-1,714285714	3,1
1	TYADNQPG	1093,5243	0,0022	43,64	33	40	-1,5875	3,1
1	YADNQPGVL	1204,6291	-0,0249	62,97	32	40	-0,444444444	3,1
1	FDLGGGTFDV	1255,6288	0,0098	50,52	36	40	0,47	2,92
1	FDLGGGTFDVS	1342,6608	-0,0257	83,54		39	0,354545455	2,92
1	LGGEDFDNRLV	1462,7619	-0,0014	63,66		42	-0,427272727	3,7
1	GAPPAPEAEGPTVEEVD	1892,9206	-0,0367	98,21	31	39	-0,682352941	2,84
1	GFPGGAPPAPEAEGPTVEEVD	2251,0848	-0,0135	76,99	26	40	-0,533333333	2,84
1	YADNQPG	992,4766	-0,0075	55,32	30	39	-1,714285714	3,1
1	TYADNQPG	1093,5243	0,0022	43,64	33	40	-1,5875	3,1
1	YADNQPGVL	1204,6291	-0,0249	62,97	32	40	-0,444444444	3,1
1	FDLGGGTFDV	1255,6288	0,0098	50,52	36	40	0,47	2,92

1	FDLGGGTFDVS	1342,6608	-0,0257	83,54		39	0,354545455	2,92
1	LGGEDFDNRLV	1462,7619	-0,0014	63,66		42	-0,427272727	3,7
1	GAPPAPEAEGPTVEEVD	1892,9206	-0,0367	98,21	31	39	-0,682352941	2,84
1	GFPGGAPPAPEAEGPTVEEVD	2251,0848	-0,0135	76,99	26	40	-0,533333333	2,84
1	GDNYFF	990,465	-0,0168	36,33	26	36	-0,516666667	3,1
1	QQLTGDN	1003,5137	0,0058	42,74		41	-1,614285714	3,1
1	TGDNYFF	1091,5127	-0,0197	47,27	34	37	-0,542857143	3,1
1	FGWDTGTI	1124,5705	-0,011	43,24	29	41	0,0875	3,1
1	GDNYFFY	1153,5283	-0,0172	46,94	32	36	-0,628571429	3,1
1	QQLTGDNY	1166,5771	0,0117	68,02		41	-1,575	3,1
1	FGWDTGTIS	1211,6026	-0,0144	55,83	38	40	-0,011111111	3,1
1	TGDNYFFY	1254,576	-0,0213	54,53	29	36	-0,6375	3,1
1	WPNGQDQPS	1256,5989	-0,0081	50,07	27	38	-2,144444444	3,1
1	QQLTGDNYF	1313,6455	0,0324	60,57	33	41	-1,088888889	3,1
1	GDNYFFYY	1316,5916	0,0003	55,44	23	37	-0,7125	3,1
1	TGDNYFFYY	1417,6393	-0,0056	53,33		36	-0,711111111	3,1
1	GDNYFFYYGT	1474,6608	-0,0311	72,74	28	34	-0,68	3,1
1	NDMYAEGVLPW	1522,7329	0,0348	64,77	26	41	-0,272727273	3
1	TGDNYFFYYGT	1575,7085	-0,0307	62,71	32	34	-0,681818182	3,1
1	QQLTGDNYFFY	1623,7772	-0,0282	78,85	30	38	-0,754545455	3,1
1	QQLTGDNYFFYY	1786,8405	0,0037	89,81	31	39	-0,8	3,1
1	QQLTGDNYFFYYGT	1944,9097	-0,0342	126,24	31	37	-0,764285714	3,1
1	GYSSND	870,3922	-0,0036	34,67		34	-1,716666667	3,1
1	LFTEVEA	1036,5644	-0,0237	48,32	39	41	0,7	3,12
1	LNDADIY	1051,5389	0,0168	49,36		41	-0,242857143	2,92
1	DVWSGGIV	1060,5756	-0,0224	48,38		41	0,8625	3,1
1	DDEDTRV	1077,5141	0,0052	40,13		40	-2,142857143	3,5
1	LFTEVEAL	1149,6484	-0,0211	55,24	41	42	1,0875	3,12
1	NDADIYVI	1150,6073	0	56,13		42	0,4	2,92
1	YMYFEET	1210,5419	0,0192	50,53		38	-0,8	3,12

1	LNDADIYVI	1263,6913	-0,0077	72		42	0,777777778	2,92
1	GYSSNDDEDTRV	1585,7059	0,077	69,87	30	40	-1,816666667	3,5
1	FFSEYGCNEVTPR	1776,8344	0,0158	85,79	26	40	-0,576923077	4,26
1	FFSEYGCNEVTPRLF	2036,9869	0,0007	103,26	30	41	-0,06	4,26
1	NVLQPDSA	1071,5763	-0,0168	45,23	37	40	-0,3875	3,1
1	FYGIIDLI	1181,6899	0,0372	64,79		39	1,8625	3,1
1	GIFYGIIDL	1238,7114	0,052	53,87	34	39	1,611111111	3,1
1	GIFYGIIDLI	1351,7954	-0,0146	68,99	35	42	1,9	3,1
1	SNPEGNGGY	1122,5144	-0,007	50,76	30	37	-1,711111111	3,3
1	YQDYSVPV	1198,6073	0,0025	48,23	34	40	-0,45	3,1
1	YTVDNSQDH	1306,5993	0,0048	65,58	26	38	-1,755555556	3,88
1	SNPEGNGGYST	1310,5942	-0,0171	76,58		36	-1,536363636	3,3
1	YTVDNSQDHS	1393,6313	0,0011	79,54	31	38	-1,66	3,88
1	SNPEGNGGYSTS	1397,6262	-0,0228	77,34		35	-1,475	3,3
1	VFGWDTG	1009,5072	-0,0214	41,07	31	40	0,157142857	3,1
1	FGWDTGTI	1124,5705	-0,011	43,24	29	41	0,0875	3,1
1	FVFGWDTG	1156,5756	0,0259	56,35		40	0,4875	3,1
1	FGWDTGTIS	1211,6026	-0,0144	55,83	38	40	-0,011111111	3,1
1	GFVFGWDTG	1213,5971	0,0554	65,06	38	41	0,388888889	3,1
1	GGFVFGWDTG	1270,6186	0,0468	65,53	34	41	0,31	3,1
1	FGGFVFGWDTG	1417,687	-0,0243	67,53	28	38	0,536363636	3,1
1	GFVFGWDTGTI	1427,7289	-0,021	75,9	37	40	0,663636364	3,1
1	FVFGWDTGTIS	1457,7394	-0,0017	71,44	32	41	0,627272727	3,1
1	GFVFGWDTGTIS	1514,7609	0,0271	89,13	36	41	0,541666667	3,1
1	FGGFVFGWDTGTI	1631,8187	0,0545	92,84	29	42	0,746153846	3,1
1	FGGFVFGWDTGTIS	1718,8508	-0,0156	131,21	37	40	0,635714286	3,1
1	YDGYEA	945,4283	0,0154	41,21		36	-1,366666667	3
1	LEQELV	958,5538	0,029	45,32		40	0,216666667	3,12
1	LNEWTLDSA	1276,6502	-0,0288	56,73		40	-0,388888889	3
1	ELDDQFIDRY	1541,7565	0,0124	54,68	30	41	-1,22	3,5

1	DDSPVTPGETRPA	1569,7838	0,0753	59,73	32	42	-1,261538462	3,7
1	SELDDQFIDRY	1628,7885	0,0286	79,26	33	41	-1,181818182	3,5
1	LDDSPVTPGETRPA	1682,8678	-0,0291	88,47		41	-0,9	3,7
1	ENDDDVSDITDKL	1935,9679	0,019	87,41		42	-1,338461538	3,26
1	IEEPEQEEEGAVEEH	1981,8955	-0,0105	110	30	37	-1,746666667	3,44
1	YEVSELDDQFIDRY	2019,9628	-0,0318	62,72	24	38	-0,971428571	3,42
1	SIWGLENDDDVSDITDKL	2492,2688	0,0591	140,43		43	-0,622222222	3,26
1	LSFKQDYEDFEPEEGEEEEEEDGQ	3364,4672	-0,0396	56,62		34	-2,145833333	3,12
1	GDNYFF	990,465	-0,0168	36,33	26	36	-0,516666667	3,1
1	QQLTGDN	1003,5137	0,0058	42,74		41	-1,614285714	3,1
1	TGDNYFF	1091,5127	-0,0197	47,27	34	37	-0,542857143	3,1
1	FGWDTGTI	1124,5705	-0,011	43,24	29	41	0,0875	3,1
1	GDNYFFY	1153,5283	-0,0172	46,94	32	36	-0,628571429	3,1
1	QQLTGDNY	1166,5771	0,0117	68,02		41	-1,575	3,1
1	FGWDTGTIS	1211,6026	-0,0144	55,83	38	40	-0,011111111	3,1
1	TGDNYFFY	1254,576	-0,0213	54,53	29	36	-0,6375	3,1
1	QQLTGDNYF	1313,6455	0,0324	60,57	33	41	-1,088888889	3,1
1	GDNYFFYY	1316,5916	0,0003	55,44	23	37	-0,7125	3,1
1	TGDNYFFYY	1417,6393	-0,0056	53,33		36	-0,711111111	3,1
1	GDNYFFYYGT	1474,6608	-0,0311	72,74	28	34	-0,68	3,1
1	TGDNYFFYYGT	1575,7085	-0,0307	62,71	32	34	-0,681818182	3,1
1	QQLTGDNYFFY	1623,7772	-0,0282	78,85	30	38	-0,754545455	3,1
1	QQLTGDNYFFYY	1786,8405	0,0037	89,81	31	39	-0,8	3,1
1	QQLTGDNYFFYYGT	1944,9097	-0,0342	126,24	31	37	-0,764285714	3,1
1	GDNYFF	990,465	-0,0168	36,33	26	36	-0,516666667	3,1
1	QQLTGDN	1003,5137	0,0058	42,74		41	-1,614285714	3,1
1	TGDNYFF	1091,5127	-0,0197	47,27	34	37	-0,542857143	3,1
1	FGWDTGTI	1124,5705	-0,011	43,24	29	41	0,0875	3,1
1	GDNYFFY	1153,5283	-0,0172	46,94	32	36	-0,628571429	3,1
1	QQLTGDNY	1166,5771	0,0117	68,02		41	-1,575	3,1

1	FGWDTGTIS	1211,6026	-0,0144	55,83	38	40	-0,011111111	3,1
1	TGDNYFFY	1254,576	-0,0213	54,53	29	36	-0,6375	3,1
1	QQLTGDNYF	1313,6455	0,0324	60,57	33	41	-1,088888889	3,1
1	GDNYFFYY	1316,5916	0,0003	55,44	23	37	-0,7125	3,1
1	TGDNYFFYY	1417,6393	-0,0056	53,33		36	-0,711111111	3,1
1	GDNYFFYYGT	1474,6608	-0,0311	72,74	28	34	-0,68	3,1
1	TGDNYFFYYGT	1575,7085	-0,0307	62,71	32	34	-0,681818182	3,1
1	QQLTGDNYFFY	1623,7772	-0,0282	78,85	30	38	-0,754545455	3,1
1	QQLTGDNYFFYY	1786,8405	0,0037	89,81	31	39	-0,8	3,1
1	QQLTGDNYFFYYGT	1944,9097	-0,0342	126,24	31	37	-0,764285714	3,1
1	YADNQPG	992,4766	-0,0075	55,32	30	39	-1,714285714	3,1
1	TYADNQPG	1093,5243	0,0022	43,64	33	40	-1,5875	3,1
1	YADNQPGVL	1204,6291	-0,0249	62,97	32	40	-0,444444444	3,1
1	FDLGGGTFDV	1255,6288	0,0098	50,52	36	40	0,47	2,92
1	FDLGGGTFDVS	1342,6608	-0,0257	83,54		39	0,354545455	2,92
1	LGGEDFDNRLV	1462,7619	-0,0014	63,66		42	-0,427272727	3,7
1	GDNYFF	990,465	-0,0168	36,33	26	36	-0,516666667	3,1
1	QQLTGDN	1003,5137	0,0058	42,74		41	-1,614285714	3,1
1	TGDNYFF	1091,5127	-0,0197	47,27	34	37	-0,542857143	3,1
1	GDNYFFY	1153,5283	-0,0172	46,94	32	36	-0,628571429	3,1
1	QQLTGDNY	1166,5771	0,0117	68,02		41	-1,575	3,1
1	TGDNYFFY	1254,576	-0,0213	54,53	29	36	-0,6375	3,1
1	QQLTGDNYF	1313,6455	0,0324	60,57	33	41	-1,088888889	3,1
1	GDNYFFYY	1316,5916	0,0003	55,44	23	37	-0,7125	3,1
1	TGDNYFFYY	1417,6393	-0,0056	53,33		36	-0,711111111	3,1
1	GDNYFFYYGT	1474,6608	-0,0311	72,74	28	34	-0,68	3,1
1	TGDNYFFYYGT	1575,7085	-0,0307	62,71	32	34	-0,681818182	3,1
1	QQLTGDNYFFY	1623,7772	-0,0282	78,85	30	38	-0,754545455	3,1
1	QQLTGDNYFFYY	1786,8405	0,0037	89,81	31	39	-0,8	3,1
1	QQLTGDNYFFYYGT	1944,9097	-0,0342	126,24	31	37	-0,764285714	3,1

1	NDDEASREA	1234,5628	-0,0078	44,08	35	37	-2,133333333	3,58
1	QDYVLGDDFIRG	1625,8252	-0,0242	63,82	27	41	-0,441666667	3,6
1	VPGQDYVLGDDFIRG	1878,9679	-0,0311	71,89		41	-0,206666667	3,6
1	FSDPEAPGYDPKLDPN	2219,1152	0,055	73,09	33	43	-1,36875	3,5
1	NVLQPDSA	1071,5763	-0,0168	45,23	37	40	-0,3875	3,1
1	FYGIIDLL	1181,6899	0,0372	64,79		39	1,775	3,1
1	GLIFDEEPA	1218,6335	-0,0097	46,55	37	41	0,044444444	2,94
1	GLIFDEEPAEH	1484,735	0,0188	74,49	29	41	-0,572727273	3,68
1	KKAQEEEED	1791,9823	0,0357	51,17	32	42	-3	4
1	GDNYFF	990,465	-0,0168	36,33	26	36	-0,516666667	3,1
1	QQLTGDN	1003,5137	0,0058	42,74		41	-1,614285714	3,1
1	TGDNYFF	1091,5127	-0,0197	47,27	34	37	-0,542857143	3,1
1	GDNYFFY	1153,5283	-0,0172	46,94	32	36	-0,628571429	3,1
1	QQLTGDNY	1166,5771	0,0117	68,02		41	-1,575	3,1
1	TGDNYFFY	1254,576	-0,0213	54,53	29	36	-0,6375	3,1
1	QQLTGDNYF	1313,6455	0,0324	60,57	33	41	-1,088888889	3,1
1	GDNYFFYY	1316,5916	0,0003	55,44	23	37	-0,7125	3,1
1	TGDNYFFYY	1417,6393	-0,0056	53,33		36	-0,711111111	3,1
1	GDNYFFYYGT	1474,6608	-0,0311	72,74	28	34	-0,68	3,1
1	TGDNYFFYYGT	1575,7085	-0,0307	62,71	32	34	-0,681818182	3,1
1	QQLTGDNYFFY	1623,7772	-0,0282	78,85	30	38	-0,754545455	3,1
1	QQLTGDNYFFYY	1786,8405	0,0037	89,81	31	39	-0,8	3,1
1	QQLTGDNYFFYYGT	1944,9097	-0,0342	126,24	31	37	-0,764285714	3,1
1	YADNQPG	992,4766	-0,0075	55,32	30	39	-1,714285714	3,1
1	TYADNQPG	1093,5243	0,0022	43,64	33	40	-1,5875	3,1
1	YADNQPGVL	1204,6291	-0,0249	62,97	32	40	-0,444444444	3,1
1	FDLGGGTFDV	1255,6288	0,0098	50,52	36	40	0,47	2,92
1	FDLGGGTFDVS	1342,6608	-0,0257	83,54		39	0,354545455	2,92
1	YDGYEA	945,4283	0,0154	41,21		36	-1,366666667	3
1	LEQELV	958,5538	0,029	45,32		40	0,216666667	3,12

1	TEVPVDDEAH	1339,6459	0,0563	71,76	37	41	-0,93	3,58
1	ELQDQFIDKY	1755,9449	-0,0089	53,04	39	42	-1,16	3,7
1	YELGELQDQFIDK	2055,093	-0,0228	78,32	41	42	-0,9	3,58
1	YELGELQDQFIDKY	2218,1563	0,0462	99,18	31	43	-0,928571429	3,58
1	LWGADNDDDVSDVTDKL	2335,1586	-0,0187	109,28	33	41	-0,788235294	3,22
1	SLWGADNDDDVSDVTDKL	2422,1906	-0,0395	131,78	38	40	-0,788888889	3,22
1	FHQTVDDVYEDEDGEEEEEPEIQ	3009,2726	0,0285	107,88	30	35	-1,760869565	3,1
1	TLEPVEQV	1142,6386	-0,0243	40,64	37	40	-0,075	3,12
1	FDLGGGTFDV	1255,6288	0,0098	50,52	36	40	0,47	2,92
1	FDLGGGTFDVS	1342,6608	-0,0257	83,54		39	0,354545455	2,92
1	ATDDVEDAAPETKEA	2019,005	-0,0097	69,74	29	41	-1,1	3,38
1	HATDDVEDAAPETKEA	2156,0639	0,0195	123,13	30	42	-1,23125	3,74
1	ATDDVEDAAPETKEAVPESPRAS	2842,4238	-0,069	67,83	27	40	-1,013043478	3,68
1	TLEPVEQV	1142,6386	-0,0243	40,64	37	40	-0,075	3,12
1	FDLGGGTFDV	1255,6288	0,0098	50,52	36	40	0,47	2,92
1	FDLGGGTFDVS	1342,6608	-0,0257	83,54		39	0,354545455	2,92
1	FDNLVWA	1092,5807	0,0446	53,84		42	0,671428571	3,1
1	SLQEVDSI	1118,6022	0,0121	51,09	41	42	0,05	3
1	FDNLVWAN	1206,6236	-0,0143	63,38	37	41	0,15	3,1
1	SFGSLQEVDSI	1409,7241	-0,0204	87,79		41	0,181818182	3
1	FDNLVWA	1092,5807	0,0446	53,84		42	0,671428571	3,1
1	SLQEVDSI	1118,6022	0,0121	51,09	41	42	0,05	3
1	FDNLVWAN	1206,6236	-0,0143	63,38	37	41	0,15	3,1
1	SFGSLQEVDSI	1409,7241	-0,0204	87,79		41	0,181818182	3
1	SQDAYDLFDKV	1757,9242	-0,0128	65,37	32	42	-0,672727273	3,6
1	DLIDERIIDQQ	1585,8514	0,0354	48,35	33	42	-0,745454545	3,5
1	IGADLIDERIIDQQ	1826,9941	-0,0363	65,85	35	42	-0,164285714	3,5
1	LENTEIGDSIFDKA	2009,0723	-0,0209	88,32	34	42	-0,421428571	3,58
1	GVVFDSSGNVV	1307,6925	0,0394	88,38	35	42	0,927272727	3,1
1	FQGIEDLGTGI	1377,7343	-0,0132	74,79		42	0,290909091	3

1	IGQEVDGEAVGDEFKG	2107,0839	0,0809	112,23	33	43	-0,5625	3,5
1	YEYTEEA	1132,5127	0,008	42,63	24	37	-1,714285714	3,02
1	SATNSESDNEVFFV	1773,826	-0,0332	115,92	27	37	-0,342857143	2,94
1	GFFSGNEVINDQSDYA	1990,9111	-0,01	104,56	26	38	-0,5375	2,88
1	SIEQNPSQGA	1258,6356	0,0068	81,36		40	-1,13	3,3
1	DEQETETKQQ	1692,8572	-0,0283	48,57	31	41	-2,98	3,68
1	NTLDEQETETKQ	1892,9733	0,075	100,06		42	-2,225	3,68
1	NTLDEQETETKQQ	2021,0319	0,0892	98,75	34	43	-2,323076923	3,68
1	EGYGFTENGELPY	1703,7882	0,0224	59,82	28	39	-1,038461538	3,02
1	TFDGPNDPLHPFNWPM	2112,9931	-0,02	43,43	27	38	-0,89375	3,88
1	ATPTPNPESSAGADDTSREA	2202,0239	0,0446	80,75	33	41	-1,225	3,58
1	SHENEDQNQGEEEEEGEEEEEEEEERA	3476,3569	0,0222	126,37		29	-3	3,36
1	SHENEDQNQGEEEEEGEEEEEEEEERAH	3613,4158	-0,0012	35,2	26	30	-3,007142857	3,56
1	QDVEQG	903,4501	-0,0111	44,75	34	39	-1,7	3
1	NQIDAQ	916,4817	-0,0138	40,86		39	-1,283333333	3,1
1	QLFNDMEEL	1366,6641	0,0204	69,64		41	-0,577777778	2,94
1	QLFNDMEELVIEQ	1835,9178	-0,0348	128,13		40	-0,269230769	2,88
1	SGEQQQPA	1072,5352	0,0009	49,46	25	39	-1,875	3,3
1	NQVEEYEADGL	1494,7041	-0,0134	79,91	29	38	-1,172727273	2,88
1	STFAEPPLVDGEGNPI	1870,9516	-0,026	82,44	31	41	-0,25	2,94
1	YSQDQGVEYEEDEEDKPNLS	2831,3027	-0,0727	65,87	19	35	-2,03	3,28
1	ARDEDDLDENELLM	1905,8828	0,0137	78,99	37	39	-1,242857143	3,28
1	QKEEPEQEEIAPSLPS	2268,1891	-0,0378	106,44	33	42	-1,54375	3,68
1	EFVDDDWWLGELEKDGSKG	2911,4848	-0,0488	81,79	34	42	-1,131578947	3,62
1	EAPKPEVPEDEPEGEPDVKDL	3005,5689	0,0889	55,54	26	43	-1,604761905	3,58
1	LFEFNPF	1141,6011	-0,0036	42,91	38	40	0,514285714	3,3
1	LFEFNPFEMG	1458,7056	0,0243	46,3	26	40	0,16	3,12
1	TDFLEDLSDNSDDIAIYAPNPFKEA	3257,6021	-0,08	74,96	37	40	-0,584	3,22
1	NQEVLEWL	1258,676	0,0324	57,94		41	-0,3875	3,12
1	QEQQLQEQ	1258,6356	-0,0187	63,15		39	-2,5875	3,12

1	SGTILDIFDVDQV	1649,8715	-0,031	101,88	40	42	0,623076923	2,82
1	SYGGGQES	1012,4664	-0,0071	62,11		37	-1,3875	3,3
1	LLEYDAF	1098,58	0,0373	45,36	36	40	0,557142857	3
1	SYGGGQESTV	1212,5826	0,0353	74,78	39	40	-0,76	3,3
1	DIYRVEETLPDEV	1805,925	-0,0266	72,85	34	41	-0,684615385	3,5
1	QDKEGIPPDQQ	1711,9147	0,0719	62,32	29	41	-2,181818182	3,7
1	SKEVSEEENTEPEVQ	2191,0898	0,0552	99,46	34	42	-1,826666667	3,58
1	FGWDTGTI	1124,5705	-0,011	43,24	29	41	0,0875	3,1
1	FGWDTGTIS	1211,6026	-0,0144	55,83	38	40	-0,011111111	3,1
1	LWEEGVLPW	1356,7281	-0,0261	41,02	32	41	0,111111111	3,12
1	RTEDDDEDEEAEKQ	2165,9966	-0,0248	85,2		37	-3,021428571	3,52
1	QAWVEDGLVKPEAVEEV	2355,2728	-0,0435	80,16	40	43	-0,211764706	3,58
1	RTEDDDEDEEAEKQNGDAKEN	3123,4684	0,0128	47,62	25	40	-2,8	3,64
1	QLFNDMEEL	1366,6641	0,0204	69,64		41	-0,577777778	2,94
1	QLFNDMEELVIEQ	1835,9178	-0,0348	128,13		40	-0,269230769	2,88
1	DSSENPA	947,4399	-0,0031	54,13	29	37	-1,7	3
1	NDQEEEEGQGEGH	1685,6968	-0,0089	87,85	27	31	-2,761538462	3,5
1	ANDQEEEEGQGEGH	1756,7339	-0,0077	45,78	22	32	-2,435714286	3,5
1	NDQEEEEGQGEGHENQ	2056,8409	-0,0192	57,31	23	30	-2,9	3,44
1	NEADQY	967,445	0,0024	40,67		38	-2,25	3
1	GDDQQESQQQRGY	1766,8023	-0,018	49,22	29	36	-2,723076923	3,7
1	GDDQQESQQQRGYT	1867,8499	-0,0267	97,87		36	-2,578571429	3,7
1	GIPNAGDET	1101,5505	0,041	43,78	29	41	-0,811111111	3
1	ATGPEDMQQ	1204,5597	0,0141	58,96	32	39	-1,444444444	3
1	KDYVVEDGDII	1722,9446	-0,0194	73,35	32	42	-0,2	3,5
1	EFGGPGGQGFGGPNPQ	1730,8216	-0,0469	54,45	18	37	-1	3,3
1	ANDEQNA	989,4617	0,0064	42,62		38	-1,985714286	3
1	YDEEREAS	1226,5618	-0,0096	39,74	28	37	-2,35	3,68
1	AYGIFTDSLYGVFA	1751,8973	0,0016	118,8	33	42	0,95	3,1
1	QQQQQDPT	1200,5938	-0,0006	48,52		39	-2,9125	3,1

1	QQNPSDQEQQPPASPT	1979,9388	-0,0349	80,15	34	38	-2,18125	3
1	NIDNSTGQAGQA	1403,6844	-0,0053	100,84		40	-0,975	3,1
1	RFFPTDELENVPDSPA	2062,021	-0,0393	43,07	23	40	-0,80625	3,58
1	SDPEKRDIYDQFGEDGL	2441,2116	-0,0521	40,92	25	40	-1,547058824	3,68
1	FKRDGDDLVYEAEIDLL	2468,3204	0,0336	65,65	27	43	-0,376470588	3,68
1	GELTQEELERIV	1643,8933	-0,0276	73,14	31	42	-0,566666667	3,8
1	EAEYGNQGA	1166,5406	-0,0003	76,7	36	38	-1,388888889	3,12
1	NSDDFDTEFREPDMPT	2143,9207	0,0908	73,63	28	39	-1,68125	3,32
1	NSEKEDEEQEAV	1863,9103	0,0844	64,23	32	43	-2,225	3,5
1	SQNSEKEDEEQEAV	2079,0009	0,0856	84,01	36	42	-2,214285714	3,5
1	SSRNEFSEEDGDHNGEFS	2270,9515	-0,0145	32,77	19	31	-1,894444444	3,82
1	ALLDQLDKDINTF	1963,1032	-0,0071	78,76	33	43	-0,123076923	3,6
1	FGWDTGTI	1124,5705	-0,011	43,24	29	41	0,0875	3,1
1	FGWDTGTIS	1211,6026	-0,0144	55,83	38	40	-0,011111111	3,1
1	DQSDPNDPKS	1559,7833	-0,0224	74,47		40	-2,62	3,6
1	NMMDIADLEEPFLPRPN	2230,0965	-0,0148	74,71	32	41	-0,576470588	3,58
1	GLVDPDERVNDFDIS	1918,9475	0,0105	43,37	32	41	-0,586666667	3,36
1	RIEDDPFENLEDTDDIFQKDFG	3115,5028	-0,0799	41,57	23	38	-1,290909091	3,44
1	RIEEDPFENLEDTDDLFQNEFGI	3013,4032	-0,0679	42,19	29	37	-1,052173913	3,2
1	EYFRDEEGQDVL	1727,8205	0,0352	50,27	25	41	-1,366666667	3,5
1	QGTDQPY	1036,5028	0,0116	48,65	37	41	-2,071428571	3,1
1	GQQYEQQ	1108,5352	-0,0184	56,3	31	39	-2,742857143	3,3
1	YSDLPQQERDTI	1692,8522	-0,0263	74,63	31	40	-1,508333333	3,7
1	YFEDNA	986,4548	-0,0159	35,64		35	-1,2	3
1	YFEDNAGVI	1255,6288	0,0016	62,64		40	0,122222222	3
1	YFEDNAGVIA	1326,6659	0,0177	80,15	39	41	0,29	3
1	KGDDEANGA	1333,6879	-0,006	53,9	33	41	-1,677777778	3,7
1	DYFGDADKA	1458,7397	-0,016	41,64	28	40	-1,077777778	3,6
1	SIDDLIHEII	1395,7812	-0,0207	70,11	33	42	0,73	3,7
1	SIDDLIHEII	1395,7812	-0,0207	70,11	33	42	0,73	3,7

1	SLDDLQQSIEEDEDHVQST	2416,1081	-0,014	83,83	31	38	-1,315789474	3,28
1	DFIEDEPKEDSDGV	2051,9941	0,0228	56,45	39	41	-1,407142857	3,28
1	NQEVLEWL	1258,676	0,0324	57,94		41	-0,3875	3,12
1	RDDADEDEEDPDTRSS	2079,8668	-0,0434	50,51	23	29	-2,6625	3,44
1	YLDEEPPLTEGEEPRIY	2278,1208	-0,0276	58,04	26	40	-1,288235294	3,5
1	YTGGDNNEY	1260,5461	0,0579	41,08	32	39	-2,011111111	3
1	YLDEEPPLNEGEEPRIY	2291,116	-0,0223	63,2	36	40	-1,452941176	3,5
1	REGDQGENFY	1442,6629	-0,0124	37,24	26	37	-2,13	3,82
1	QYESPSVDTG	1310,6193	0,0623	41,11	30	41	-1,19	3
1	KDLNQEGDDQMFV	1995,9977	0,0577	67,5	35	42	-1,238461538	3,5
1	EILPEEIEERV	1583,8609	-0,0303	50,08	28	42	-0,6	3,68
1	FSADDERFGEVV	1598,778	-0,0067	46,41	31	40	-0,325	3,58
1	NEVDENEERQV	1588,7532	0,0382	42,7	36	41	-2,190909091	3,58
1	LNEVDENEERQV	1701,8372	-0,0266	76,09	36	39	-1,691666667	3,58
1	DSTLREGEQFA	1480,7361	-0,026	50,62	35	40	-1,090909091	3,82
1	RALDDFGVDYI	1511,7823	0,0572	53,2	30	42	0,036363636	3,6
1	DSTLREGEQFA	1480,7361	-0,026	50,62	35	40	-1,090909091	3,82
1	RALDDFGVDYI	1511,7823	0,0572	53,2	30	42	0,036363636	3,6
1	LFDDLKDGLI	1605,9384	-0,018	65,12	40	42	0,39	3,6
1	QVYDNG	923,4552	-0,0016	41,5		40	-1,333333333	3,1
1	TAENGGSSVGDGA	1349,6262	-0,0097	71,86		38	-0,507692308	3
1	KVGEDDQTEAQ	1676,8623	0,0161	75,04	32	42	-1,818181818	3,58
1	LDRGDGEEGN	1289,6051	-0,0023	42,28	28	38	-1,94	3,58
1	GYQDDPQYA	1284,5825	0,0541	53,02	32	40	-1,866666667	2,92
1	FGWDTGTI	1124,5705	-0,011	43,24	29	41	0,0875	3,1
1	FGWDTGTIS	1211,6026	-0,0144	55,83	38	40	-0,011111111	3,1
1	YLVDDLRPEFA	1565,8292	0,0397	42,01	37	42	-0,136363636	3,7
1	KENAGDVDTQ	1533,804	0,01	62,94		42	-1,65	3,7
1	EKQDEVSGQ	1476,7826	-0,008	59,72	40	42	-2,044444444	3,82
1	DRFGDHNLGDDDDADFEKQV	2765,2935	-0,0455	52,36	23	38	-1,6	3,68

1	LQQPQEQQQQQQ	1738,8801	-0,0068	52,49	39	41	-2,733333333	3,3
1	EVGQYEIV	1164,6229	0,0406	47,08	34	42	0,0875	3,12
1	LDNGWDIF	1207,6076	-0,0105	42,53	35	41	-0,0875	2,92
1	IFGDAFDVDRA	1453,7404	0,0515	52,32	36	42	0,227272727	3,6
1	LTRELV	958,6014	-0,0186	50,48		40	0,516666667	6,98
1	ITRELV	958,6014	-0,0186	50,48		40	0,633333333	6,98
1	ITRELV	958,6014	-0,0186	50,48		40	0,633333333	6,98
1	LTRELV	958,6014	-0,0186	50,48		40	0,516666667	6,98
1	QDMQDEMLDLIEQ	1835,8484	0,0346	47,86		40	-0,930769231	2,76
1	AKNDDDSGNNLDNN	1962,9285	-0,0214	47,81	38	39	-2,214285714	3,42
1	FYGILNAR	1181,676	0,0512	43,46		39	0,4	9,84
1	LEQELV	958,5538	0,029	45,32		40	0,216666667	3,12
1	IEQELV	958,5538	0,029	45,32		40	0,333333333	3,12
1	KMNDDPDQQ	1547,7655	0,0326	44,66	30	42	-2,733333333	3,6
1	INDWFII	1148,6433	0,0488	43,92		41	1,2	3,1
1	ALLDAL	843,5268	-0,0186	42,28		39	1,916666667	3,1
1	AILDAL	843,5268	-0,0186	42,28		39	2,033333333	3,1
1	ALLDAL	843,5268	-0,0186	42,28		39	1,916666667	3,1
1	EQEEKGESS	1479,7458	-0,0194	41,67	33	40	-2,6	3,8
1	VMGGGLGGGP	1045,543	-0,0496	41,47		38	0,59	6,02
2	AITGESL	918,5225	-0,0125	47,97		43	0,671428571	3,3
2	LAGVEIL	942,5953	-0,001	45,54		39	2,028571429	3,3
2	GFVQEF	954,5014	0,0467	46,29		39	0,4	3,3
2	AITGESLA	989,5596	0,0007	66,69		42	0,8125	3,3
2	QFVMEAA	1023,5262	-0,0135	53,56		41	0,785714286	3,3
2	GFVQEFQ	1082,56	0,0078	53,99	37	41	-0,157142857	3,3
2	SLTENWL	1090,5862	-0,0137	44,31		42	-0,257142857	3,3
2	LTENWLI	1116,6382	0,0461	51,27		41	0,5	3,3
2	GEGHWEIL	1168,608	-0,006	55,09	33	41	-0,45	4,24

2	SLTENWLI	1203,6702	0,0332	64,62		42	0,3375	3,3
2	EFHPFDPV	1215,6127	0,035	53,63	37	41	-0,45	4,06
2	ESPEGERIV	1243,6611	0,0454	46,93	34	41	-1,011111111	3,96
2	LFGWWSEN	1266,6236	0,0119	47,03	37	41	-0,425	3,3
2	RDGQLVEIPA	1325,7506	-0,0024	58,39	36	41	-0,27	4,08
2	LEFHPFDPV	1328,6968	0,0006	50,73	32	41	0,022222222	4,06
2	VESPEGERIV	1342,7295	0,0103	44,81	34	41	-0,49	3,96
2	GEARPVPEEY	1374,6982	0,0374	47,24	30	42	-1,39	3,96
2	AVVESPEGERI	1413,7666	0,0215	68,25	40	41	-0,281818182	3,96
2	ENYENKVA	1423,7713	0,0167	46,23	33	42	-1,65	4,26
2	IRDGQLVEIPA	1438,8347	0,0278	41,65	33	40	0,163636364	4,08
2	ADEKESLVV	1446,8335	-0,0201	65,2		42	-0,133333333	3,82
2	GEARPVPEEYL	1487,7823	-0,0211	48,38	30	41	-0,918181818	3,96
2	AVVESPEGERIV	1512,8351	-0,0183	70,52	38	42	0,091666667	3,96
2	FLAPGLSAIIDAL	1528,9068	-0,0217	41,11	30	41	1,715384615	3,1
2	GDGVNDAPSLK	1529,8455	0,0458	59,76	37	41	-0,709090909	3,88
2	VIRDGQLVEIPA	1537,9031	0,0457	53,5	30	39	0,5	4,08
2	GEARPVPEEYLQ	1615,8409	0,0005	61,38		42	-1,133333333	3,96
2	VVIRDGQLVEIPA	1636,9715	0,0343	63,33	28	40	0,784615385	4,08
2	EEDHPIPEDVHE	1673,7736	-0,0014	52,93	23	38	-1,825	3,82
2	GEARPVPEEYLQT	1716,8885	-0,0169	56,37	26	41	-1,1	3,96
2	MTGDGVNDAPSLK	1761,9337	-0,0245	69,6	32	42	-0,507692308	3,88
2	EEDHPIPEDVHEN	1787,8165	-0,0179	64,93	27	37	-1,953846154	3,82
2	RDGQLVEIPANEVVPG	1921,0472	0,0069	84,95	31	43	-0,20625	3,82
2	KTVEEDHPIPEDVH	2102,105	0,0614	64,11	29	43	-1,342857143	4,28
2	KTVEEDHPIPEDVHEN	2345,1905	0,0656	45,97	25	43	-1,6125	4,12
2	KTVEEDHPIPEDVHENY	2508,2538	0,0622	105,31	27	43	-1,594117647	4,12
2	EEDHPIPEDVHENYENK	2551,2232	0,0653	100,11	34	42	-2,211764706	4,02
2	EEDHPIPEDVHENYENKV	2650,2916	0,0185	75,34	31	42	-1,855555556	4,02
2	EEDHPIPEDVHENYENKVA	2721,3288	0,0605	96,94	29	43	-1,663157895	4,02

2	VEEDHPIPEDVHENYENKV	2749,3601	-0,0233	84,36	26	41	-1,536842105	4,02
2	KTVEEDHPIPEDVHENYEN	2751,3393	0,1033	57,58	32	43	-1,794736842	4,02
2	VEEDHPIPEDVHENYENKVA	2820,3972	-0,0136	106,11	36	42	-1,37	4,02
2	KTVEEDHPIPEDVHENYENKV	3207,6656	-0,0168	57,58	33	43	-1,60952381	4,26
2	KTVEEDHPIPEDVHENYENKVA	3278,7027	-0,0004	65,78	26	43	-1,454545455	4,26
2	AITGESL	918,5225	-0,0125	47,97		43	0,671428571	3,3
2	LAGVEIL	942,5953	-0,001	45,54		39	2,028571429	3,3
2	AITGESLA	989,5596	0,0007	66,69		42	0,8125	3,3
2	QFVMEAA	1023,5262	-0,0135	53,56		41	0,785714286	3,3
2	SLTENWL	1090,5862	-0,0137	44,31		42	-0,257142857	3,3
2	LTENWLI	1116,6382	0,0461	51,27		41	0,5	3,3
2	GEGHWEIL	1168,608	-0,006	55,09	33	41	-0,45	4,24
2	SLTENWLI	1203,6702	0,0332	64,62		42	0,3375	3,3
2	EFHPFDPV	1215,6127	0,035	53,63	37	41	-0,45	4,06
2	ESPEGERIV	1243,6611	0,0454	46,93	34	41	-1,011111111	3,96
2	LFGWWSEN	1266,6236	0,0119	47,03	37	41	-0,425	3,3
2	LEFHPFDPV	1328,6968	0,0006	50,73	32	41	0,022222222	4,06
2	VESPEGERIV	1342,7295	0,0103	44,81	34	41	-0,49	3,96
2	AVVESPEGERI	1413,7666	0,0215	68,25	40	41	-0,281818182	3,96
2	ENYENKVA	1423,7713	0,0167	46,23	33	42	-1,65	4,26
2	AVVESPEGERIV	1512,8351	-0,0183	70,52	38	42	0,091666667	3,96
2	FLAPGLSAIIDAL	1528,9068	-0,0217	41,11	30	41	1,715384615	3,1
2	GDGVNDAPSLK	1529,8455	0,0458	59,76	37	41	-0,709090909	3,88
2	EEDHPIPEDVHE	1673,7736	-0,0014	52,93	23	38	-1,825	3,82
2	MTGDGVNDAPSLK	1761,9337	-0,0245	69,6	32	42	-0,507692308	3,88
2	EEDHPIPEDVHEN	1787,8165	-0,0179	64,93	27	37	-1,953846154	3,82
2	KTVEEDHPIPEDVH	2102,105	0,0614	64,11	29	43	-1,342857143	4,28
2	KTVEEDHPIPEDVHEN	2345,1905	0,0656	45,97	25	43	-1,6125	4,12
2	KTVEEDHPIPEDVHENY	2508,2538	0,0622	105,31	27	43	-1,594117647	4,12
2	EEDHPIPEDVHENYENK	2551,2232	0,0653	100,11	34	42	-2,211764706	4,02

2	EEDHPIPEDVHENYENKV	2650,2916	0,0185	75,34	31	42	-1,855555556	4,02
2	EEDHPIPEDVHENYENKVA	2721,3288	0,0605	96,94	29	43	-1,663157895	4,02
2	VEEDHPIPEDVHENYENKV	2749,3601	-0,0233	84,36	26	41	-1,536842105	4,02
2	KTVEEDHPIPEDVHENYEN	2751,3393	0,1033	57,58	32	43	-1,794736842	4,02
2	VEEDHPIPEDVHENYENKVA	2820,3972	-0,0136	106,11	36	42	-1,37	4,02
2	KTVEEDHPIPEDVHENYENKV	3207,6656	-0,0168	57,58	33	43	-1,60952381	4,26
2	KTVEEDHPIPEDVHENYENKVA	3278,7027	-0,0004	65,78	26	43	-1,454545455	4,26
2	HEQLEQ	1011,5188	-0,0123	41,12	29	40	-2,233333333	4,24
2	AGGVGEFEA	1064,5341	0,0025	67,31	40	41	0,266666667	3,12
2	HEQLEQGVPG	1321,6829	-0,0077	67,13	32	41	-1,16	4,24
2	AGGVGEFEAGIS	1321,6717	0,028	95,76	41	42	0,475	3,12
2	EMHHEQLEQ	1408,6608	-0,0266	69,78	32	37	-2,022222222	4,52
2	IDAIEQPSRPT	1454,7932	0,0193	57,37	34	41	-0,809090909	4,08
2	KTLLEAIDAI	1543,9591	-0,0222	75,37		40	0,86	4,08
2	HEQLEQGVPGDNV	1649,8212	0,037	101,04	33	42	-1,107692308	3,82
2	EAFSEYPPLGRFA	1711,8772	-0,025	65,11	28	41	-0,323076923	4,26
2	EMHHEQLEQGVPG	1718,8249	0,0064	109,24		40	-1,261538462	4,52
2	CRFDELLEKN	1723,9333	0,0601	43,51	26	42	-0,95	4,44
2	RFDELLEKND	1735,951	0,0615	51,59	30	42	-1,55	4,06
2	HEQLEQGVPGDNVGF	1853,9111	-0,0313	95,89	28	40	-0,8	3,82
2	DAIEQPSRPTDKPL	2024,1308	0,0029	61,13		42	-1,328571429	4,3
2	EMHHEQLEQGVPGDNV	2046,9632	0,0795	135,54	36	42	-1,2	4,14
2	EMHHEQLEQGVPGDNVG	2103,9847	0,047	73,51	26	41	-1,152941176	4,14
2	IDAIEQPSRPTDKPL	2137,2148	0,0103	53,89	33	42	-0,94	4,3
2	EMHHEQLEQGVPGDNVGF	2251,0531	-0,0292	146,19	33	38	-0,933333333	4,14
2	EAIDAIEQPSRPTDKPL	2337,2945	0,0121	83,42	30	43	-0,929411765	4,06
2	NYDAEEM	1099,4695	-0,0074	44,92	27	33	-1,657142857	2,94
2	GFINQTDFI	1282,6761	-0,0232	74,63	33	41	0,333333333	3,1
2	YGEGEEHEPV	1373,6302	-0,0015	79,81		37	-1,67	3,78
2	FGGFVFGWDTG	1417,687	-0,0158	68,88	29	39	0,536363636	3,1

2	NYDAEEM	1099,4695	-0,0074	44,92	27	33	-1,657142857	2,94
2	GFINQTDFI	1282,6761	-0,0232	74,63	33	41	0,333333333	3,1
2	YGEGEEHEPV	1373,6302	-0,0015	79,81		37	-1,67	3,78
2	FGGFVFGWDTG	1417,687	-0,0158	68,88	29	39	0,536363636	3,1
2	QDLEHDHA	1192,5676	-0,0164	53,94	32	38	-1,85	4,28
2	LQDLEHDHA	1305,6516	-0,0085	49,96	33	40	-1,222222222	4,28
2	YRETVESES	1327,6458	-0,0201	39,41	36	39	-1,6	3,96
2	FNMFILDPI	1337,7256	0,0127	55,31		42	1,3	3,1
2	SEEQRPGTPL	1341,7091	0,0246	47,71	26	41	-1,63	4,26
2	FNMFILDPI	1353,7206	0,0168	47,55	28	42	1,3	3,1
2	AYRETVESES	1398,683	-0,0206	39,78	33	39	-1,26	3,96
2	YRETVESESS	1414,6779	-0,0179	38,71		38	-1,52	3,96
2	SKEDLYQT	1440,7866	0,0555	48,38	34	41	-1,675	4,08
2	ADPKIQEPV	1453,8546	0,0068	49,75	36	40	-0,788888889	4,08
2	WQEYYDKL	1601,8495	-0,0216	49,53	39	41	-1,7625	4,08
2	YRETVESESSQT	1643,7841	-0,0293	55,51	25	38	-1,616666667	3,96
2	QVSKEDLYQT	1667,9136	0,0081	77,12	32	42	-1,27	4,08
2	LQDLEHDHAGVPL	1671,8783	-0,0094	81,57	36	42	-0,384615385	4,28
2	NESFGFTGELRQA	1683,8419	0,0145	90,34	34	41	-0,738461538	4,26
2	TKEGPIFGEEM	1694,8955	0,041	70,75	38	42	-0,754545455	3,96
2	AYRETVESESSQT	1714,8212	-0,0277	64,31	33	38	-1,353846154	3,96
2	FLLADPKIQEPV	1827,0912	0,0009	82,2	30	40	0,275	4,08
2	GFLLADPKIQEPV	1884,1126	-0,0202	84,35	27	41	0,223076923	4,08
2	FGGFVFGWDTG	1417,687	-0,0158	68,88	29	39	0,536363636	3,1
2	NKTTEDSPLVT	1661,9242	0,0301	81,86	28	42	-0,990909091	4,08
2	FGGFVFGWDTG	1417,687	-0,0158	68,88	29	39	0,536363636	3,1
2	FGGFVFGWDTG	1417,687	-0,0158	68,88	29	39	0,536363636	3,1
2	LWEEGVLPWKSPS	1985,1028	0,0364	42,62		42	-0,469230769	4,26
2	IQVFEGERA	1276,6978	-0,0067	45,19	40	42	-0,233333333	4,26
2	LGGEDFDNRLV	1462,7619	0,0509	75,42		42	-0,427272727	3,7

2	IGRNFNDPEVQ	1516,7837	0,0145	45,04	34	42	-1,136363636	4,08
2	DVDGKPQIQV	1555,8976	-0,0324	53,11	36	42	-0,7	3,88
2	HLGGEDFDNRLV	1599,8208	0,0138	49,71	35	41	-0,658333333	4,3
2	IDVDGKPQIQV	1668,9816	0,0379	58,16	33	39	-0,227272727	3,88
2	GDTHLGGEDFDNRLV	1872,9169	-0,0094	74,82	30	40	-0,833333333	3,96
2	AGDTHLGGEDFDNRLV	1943,954	-0,0134	85,14	31	40	-0,66875	3,96
2	KEEFDDQLKEL	2080,166	0,0555	42,52	31	42	-1,672727273	3,92
2	ATKEEFDDQLKEL	2252,2508	0,0131	60,2	32	42	-1,330769231	3,92
2	AFTDTERL	1180,6291	-0,0153	43,28	31	41	-0,5625	4,08
2	IQVFEGERA	1276,6978	-0,0067	45,19	40	42	-0,233333333	4,26
2	LGGEDFDNRLV	1462,7619	0,0509	75,42		42	-0,427272727	3,7
2	IGRNFNDPEVQ	1516,7837	0,0145	45,04	34	42	-1,136363636	4,08
2	DVDGKPQIQV	1555,8976	-0,0324	53,11	36	42	-0,7	3,88
2	IGRNFNDPEVQA	1587,8208	-0,0078	43,93	31	41	-0,891666667	4,08
2	HLGGEDFDNRLV	1599,8208	0,0138	49,71	35	41	-0,658333333	4,3
2	IDVDGKPQIQV	1668,9816	0,0379	58,16	33	39	-0,227272727	3,88
2	GDTHLGGEDFDNRLV	1872,9169	-0,0094	74,82	30	40	-0,833333333	3,96
2	AGDTHLGGEDFDNRLV	1943,954	-0,0134	85,14	31	40	-0,66875	3,96
2	GFFAGNEV	1068,5443	0,0345	48,48	34	40	0,475	3,3
2	DDEDTRVK	1434,772	-0,0237	57,99	35	42	-2,3625	3,96
2	TSINRDDPTWT	1533,7626	0,0215	62,4	35	41	-1,445454545	3,88
2	TSINRDDPTWTV	1632,8311	-0,0071	47,39	30	41	-0,975	3,88
2	SNDDEDTRVK	1635,847	-0,0286	51,78	31	41	-2,32	3,96
2	SSNDDEDTRVK	1722,879	-0,0226	79,6		41	-2,181818182	3,96
2	YFEETNKYGLV	1819,9762	-0,0114	78,37	28	42	-0,663636364	4,26
2	SNDDEDTRVKMT	1867,9351	0,0154	46,68		42	-1,833333333	3,96
2	GYSSNDDEDTRVK	1942,9638	-0,0257	95,47	30	40	-1,976923077	3,96
2	SSNDDEDTRVKMT	1954,9672	-0,0063	80,06		41	-1,753846154	3,96
2	GYSSNDDEDTRVKMT	2175,052	0,0366	85,2	29	41	-1,633333333	3,96
2	FYGIIDL	1068,6058	-0,0023	48,58	36	40	1,485714286	3,1

2	FYGIIDLI	1181,6899	0,0355	64,83		39	1,8625	3,1
2	GIFYGIIDL	1238,7114	0,0019	54,65	34	41	1,611111111	3,1
2	IFYGIIDLI	1294,774	-0,0183	70,72		41	2,155555556	3,1
2	GIFYGIIDLI	1351,7954	-0,0079	74,18	39	41	1,9	3,1
2	YHFGPSDAEAV	1420,6826	-0,0231	68,95	24	39	-0,336363636	4,06
2	GYHFGPSDAEAV	1477,7041	-0,0261	78,26	29	38	-0,341666667	4,06
2	GFFAGNEV	1068,5443	0,0345	48,48	34	40	0,475	3,3
2	FFAGNEVI	1124,6069	-0,02	56,14		41	1,0875	3,3
2	NSESGERFYI	1429,704	0,0006	47,77	36	40	-1,1	4,26
2	FFNSESGERF	1447,6935	0,049	65,58	30	41	-0,86	4,26
2	AFFNSESGERF	1518,7306	0,0605	71,25	29	42	-0,618181818	4,26
2	FFNSESGERFY	1610,7568	0,0017	69,66	32	39	-0,9	4,26
2	DFENLKNEYS	1715,8772	-0,0139	61,88	39	41	-1,69	3,82
2	FFNSESGERFYI	1723,8409	0,0116	72,95	31	41	-0,45	4,26
2	AFFNSESGERFYI	1794,878	0,0044	45,48	23	41	-0,276923077	4,26
2	TDFENLKNEYS	1816,9249	0,004	60,01	30	42	-1,6	3,82
2	FGGFVFGWDTG	1417,687	-0,0158	68,88	29	39	0,536363636	3,1
2	FDSIIEH	1088,5705	0,0081	45,11		41	0,114285714	4,06
2	RDIEGSVQPS	1315,6935	-0,0147	66,1	35	41	-0,99	4,08
2	IRDIEGSVQPS	1428,7775	0,0135	64,56	36	41	-0,490909091	4,08
2	FNYQFDSIIEH	1640,8038	-0,0243	83,37	29	39	-0,427272727	4,06
2	AFTDTERL	1180,6291	-0,0153	43,28	31	41	-0,5625	4,08
2	LGGEDFDNRLV	1462,7619	0,0509	75,42		42	-0,427272727	3,7
2	HLGGEDFDNRLV	1599,8208	0,0138	49,71	35	41	-0,658333333	4,3
2	GDTHLGGEDFDNRLV	1872,9169	-0,0094	74,82	30	40	-0,833333333	3,96
2	AGDTHLGGEDFDNRLV	1943,954	-0,0134	85,14	31	40	-0,66875	3,96
2	AYLFESL	1070,5851	0,0164	53,47		40	0,942857143	3,3
2	NYNGFDEH	1223,541	-0,0023	51,01	29	36	-2,0125	4,06
2	GEGLNVEQR	1229,6567	-0,0097	44,55	36	41	-1,255555556	4,26
2	LMQEFDRL	1279,6797	0,0139	55	27	42	-0,3375	4,08

2	AGEGLNVEQR	1300,6938	-0,0029	70,18		41	-0,95	4,26
2	FSSGVEGVNPI	1333,7081	0,0177	64,34	35	42	0,427272727	3,3
2	ANDPENVGERS	1415,6844	-0,0206	41,28	28	38	-1,709090909	3,82
2	SQDAYDLFNKV	1756,9401	-0,0317	60,85	38	41	-0,672727273	3,88
2	SIEEKNRYVEEVI	2065,1461	0,0538	42,02	33	42	-0,815384615	4,26
2	FYGIIDL	1068,6058	-0,0023	48,58	36	40	1,485714286	3,1
2	FYGIIDLL	1181,6899	0,0355	64,83		39	1,775	3,1
2	IFYGIIDLL	1294,774	-0,0183	70,72		41	2,077777778	3,1
2	TIFYGIIDLL	1395,8216	-0,0233	78,77		41	1,8	3,1
2	KKAQEEEEDVATDSE	2394,237	0,0183	45,26	28	43	-1,966666667	3,74
2	AFTDTERL	1180,6291	-0,0153	43,28	31	41	-0,5625	4,08
2	GNFGPELA	1032,5443	-0,0154	60,35	35	41	-0,125	3,3
2	LLGEFDLK	1391,843	0,0175	43,35	34	40	0,3625	4,08
2	HEPEPEAEQA	1364,6411	-0,0223	45,81	25	37	-2,03	3,78
2	IHEPEPEAEQA	1477,7251	-0,0058	75,85	27	40	-1,436363636	3,78
2	HATDDVEDAAPETKEA	2156,0639	0,0095	110,34	29	41	-1,23125	3,74
2	HATDDVEDAAPETKEAVPESPRAS	2979,4827	0,0944	90,51	27	43	-1,104166667	3,96
2	LLGEFDLK	1391,843	0,0175	43,35	34	40	0,3625	4,08
2	RVQGGEEVNAEEL	1657,8474	0,0472	66,81	32	42	-0,946153846	3,8
2	SNDELSKA	1320,7291	0,0064	42,75		42	-1,3	4,08
2	SNDELSKA	1320,7291	0,0064	42,75		42	-1,3	4,08
2	AYLFESL	1070,5851	0,0164	53,47		40	0,942857143	3,3
2	GEGLNVEQR	1229,6567	-0,0097	44,55	36	41	-1,255555556	4,26
2	AGEGLNVEQR	1300,6938	-0,0029	70,18		41	-0,95	4,26
2	SQDAYDLFDKV	1757,9242	-0,019	46,3	25	42	-0,672727273	3,6
2	SIEEKNRYVEEVI	2065,1461	0,0538	42,02	33	42	-0,815384615	4,26
2	EELDWMEKNLPGRS	2161,1243	-0,0193	55,94	28	42	-1,435714286	4,16
2	VQEELDWMEKNLPGRS	2388,2513	0,0147	78,98	34	43	-1,2125	4,16
2	DNKDIERFIPS	1790,9932	0,0352	42,78	31	42	-1,181818182	4,3
2	GYKYNDELTLT	1773,9554	-0,0093	61,67	30	42	-1,018181818	4,08

2	NGEDKEIVDGKV	1989,1351	-0,039	47,67	35	43	-1,1	4,06
2	ALPWFWEHYNPEEY	2108,9835	0,0213	107,08	33	40	-1,171428571	3,96
2	ALPWFWEHYNPEEYS	2196,0155	0,054	59,1	36	41	-1,146666667	3,96
2	QLIDNLLDKV	1627,9914	-0,0243	70,49	40	41	0,22	3,88
2	FHTFQGIEDLGTGI	1762,9093	-0,0058	108,15	31	42	0,15	4,06
2	GLSKEDDVRDFV	1836,9987	0,0491	45,75	27	42	-0,716666667	3,96
2	GLSKEDDVRDFVI	1950,0828	-0,0285	52,91	26	43	-0,315384615	3,96
2	FFDKRIGQEVDGEAVGDEFKG	3029,6067	-0,0162	58,22	28	43	-0,728571429	4,06
2	GFFSGNEVI	1197,6233	-0,0175	58,48	39	40	0,633333333	3,3
2	NYGESLNRADPS	1550,7528	-0,0037	72,13	35	40	-1,483333333	4,08
2	NSGNYGESLNRADPS	1808,8492	-0,0243	84,68	32	38	-1,5	4,08
2	SATNSESDNEVFFVKGV	2287,1738	0,0275	46,26	36	42	-0,288235294	3,82
2	KFNDLDWS	1481,792	0,0401	48,05	37	42	-1,1875	3,88
2	KFAPEAPDLFA	1662,9387	-0,0312	50,98	28	42	0,063636364	4,08
2	SKREVDDLVENLDVS	2175,1789	-0,0193	44,12	34	43	-0,72	3,84
2	KVIDDVNEEDWNLLEKL	2758,5361	0,0128	82,53	34	43	-0,729411765	3,76
2	DMFDTNVRG	1282,6179	-0,0127	49,94	26	39	-0,8	3,88
2	AFADMFDTNVRG	1571,7605	-0,021	62,67	30	39	-0,066666667	3,88
2	FDVKVDKEAEQY	2157,1926	0,0323	56,86	24	42	-1,133333333	4,06
2	DSNYKEESKEQA	2114,11	-0,0158	69,62	28	42	-2,491666667	4,16
2	QSIQDTRDDELLERV	2045,0592	-0,0016	51,53	31	42	-1,246666667	3,84
2	NYEPEVYNPDHEKL	2204,1155	-0,0317	75,34	28	41	-1,85	4,16
2	LFEFNPF	1141,6011	-0,014	45,81		40	0,514285714	3,3
2	DVFKNGEMPFPI	1851,0006	0,0642	66,79	29	42	-0,15	4,08
2	IYEVAPG	976,5432	-0,0199	48,03	28	42	0,528571429	3,3
2	GYENKDFIDLM	1801,9326	-0,0194	56,7	27	41	-0,6	3,7
2	RELPKAEGSTEPLPEAL	2294,2887	0,0362	55,86	24	42	-0,829411765	4,26
2	TIPEIQRELPKAEGSTEPLPEAL	2975,6584	-0,0174	49,06	27	43	-0,626086957	4,08
2	GFAGDDAPRAV	1303,6724	0,0396	44,99	33	42	-0,136363636	3,88
2	SKQEYDESGPS	1683,8357	-0,0341	71,04	27	40	-2,145454545	3,82

2	GMGQKDSYVGDEA	1813,8922	0,0638	68,28	30	42	-1,023076923	3,7
2	SKQEYDESGPSIV	1895,9882	-0,0308	103,32	33	42	-1,146153846	3,82
2	YYFGQEG	1091,5127	-0,0177	45,36	31	37	-1,085714286	3,3
2	NDGRGWDLN	1274,6207	0,019	43,81	38	40	-1,822222222	3,88
2	ALPWFWEHYNPEEY	2108,9835	0,0213	107,08	33	40	-1,171428571	3,96
2	ALPWFWEHYNPEEYS	2196,0155	0,054	59,1	36	41	-1,146666667	3,96
2	RDAETGELLPVVPEKF	2257,2724	-0,0321	85,05	34	43	-0,38125	4,16
2	SYGGGQES	1012,4664	0,0147	43,37	27	38	-1,3875	3,3
2	NAPQKPEDIPVAT	1837,0351	0,0015	47,15	28	42	-0,853846154	4,08
2	VFGGSDSQREFV	1555,7834	0,0356	59,6	32	42	-0,283333333	4,08
2	LEEGRDFLLDHL	1684,8987	0,0501	50,68	30	42	-0,341666667	4,06
2	DSNVNDWKRLEDFL	2208,1581	0,0512	45,63	34	43	-1,178571429	3,96
2	FYQEKERADEDEGIEIV	2527,2848	0,0201	48,33	24	42	-1,194117647	3,74
2	LFYQEKERADEDEGIEIV	2640,3688	-0,0261	79,98	25	42	-0,916666667	3,74
2	QDKEGIPPDQQRLI	2094,1839	-0,0371	56,73	29	43	-1,442857143	4,3
2	EFNAQERWWI	1606,8095	0,0231	45,87	29	42	-1,12	4,26
2	SYKPEEGIEQL	1749,9554	-0,0157	43,25	39	42	-1,245454545	3,96
2	SEFKDEVIQPI	1761,9918	0,0387	56,25	29	42	-0,390909091	3,82
2	QGSEVGKEAESSPNTG	2034,0271	-0,048	54,12		40	-1,33125	3,96
2	SREADEGTNDIEKEQ	2178,0806	-0,0344	53,44	31	40	-2,133333333	3,82
2	TAPPLPRAPPVPPATFEFDSEPT	2662,3846	-0,0077	44,27	23	43	-0,508695652	3,82
2	LFEEFDRL	1296,6917	0,0277	42,33	39	41	-0,225	3,82
2	KGVEEEDISPG	1616,8663	0,0359	82,71	40	42	-1,127272727	3,68
2	QTEEKTEEKS	1895,0456	-0,031	63,25	36	42	-2,75	4,26
2	RGDLGIEIPAPEVL	1706,977	-0,0141	87,68	37	42	0,257142857	3,82
2	KQDPEQEERQ	1743,9157	-0,033	41,67	29	41	-3,45	4,16
2	ATEGDNNNNTAAGDKKG	2363,2285	-0,0376	81,2	35	42	-1,735294118	4,3
2	SATEGDNNNNTAAGDKKG	2450,2606	-0,0394	85,22	33	42	-1,683333333	4,3
2	DKYSPGDLIIPF	1822,0282	0,0321	45,77	32	42	-0,083333333	3,88
2	MQKGKDYVVEDGDII	2396,323	0,0763	65,82	24	42	-0,54	3,96

2	QYGADNGNPNGERG	1676,7706	-0,0261	65,57	27	36	-2,014285714	4,08
2	EFGGPGGQGFGGPNPQEFGGQGR	2462,1567	-0,0297	61,14	25	39	-1,126086957	4,26
2	NLDVKDQKAVDDFLLS	2506,4251	-0,0127	73,69	30	43	-0,325	3,88
2	YYFGQEG	1091,5127	-0,0177	45,36	31	37	-1,085714286	3,3
2	HYDEEREA	1276,5887	-0,0089	37,92	29	37	-2,65	4,16
2	GEGLNVEQR	1229,6567	-0,0097	44,55	36	41	-1,255555556	4,26
2	LMQEFDRL	1279,6797	0,0139	55	27	42	-0,3375	4,08
2	EQFPFEKDVNVV	1908,0399	-0,0221	55,86	35	42	-0,4	3,82
2	AEQFPFEKDVNVV	1979,077	-0,0261	75,98	34	43	-0,230769231	3,82
2	NFIEPLAWPL	1427,8016	-0,0173	58,22	31	41	0,56	3,3
2	EQLRQEGQQPV	1539,8208	-0,0018	53,56	26	42	-1,772727273	4,26
2	VFENYVADVEVDGRRV	2095,0901	-0,0175	80,36	27	42	-0,1625	4,06
2	ARFPVEDTAGGLL	1573,8667	-0,02	49,96	28	42	0,276923077	4,08
2	QSDEGPRPNVT	1427,7208	-0,0278	47,87	22	40	-1,763636364	4,08
2	KNEGLFEDEILPI	1974,1079	0,0678	86,43	32	41	-0,307692308	3,68
2	DNNEKGKEGDSS	1966,0212	-0,0365	54,3	30	41	-2,6	4,06
2	GELTQEELERI	1544,8249	0,0104	50,16	33	42	-1	3,8
2	NDITDGKDYH	1634,8306	-0,002	45,35	37	41	-1,9	4,16
2	LFEQFDRL	1295,7077	0,0498	49,01		41	-0,225	4,08
2	NILSNEEGIERL	1614,8779	0,0552	62,19		41	-0,55	3,96
2	RQLEQGIEPGDSGI	1726,9053	-0,0242	59,57	30	41	-0,914285714	3,82
2	FGKEEETSVGDRL	1924,0307	-0,0107	44,83	26	42	-1,069230769	4,16
2	FTKDQVNDDQLKNIT	2465,3734	0,0217	47,77	33	43	-1,226666667	4,18
2	HNPENDLESNNKRDPFEA	2583,2719	-0,0401	70,58	29	40	-2,15	4,28
2	EEIVEEVLEETKEKVE	2618,451	0	56,96	28	43	-0,975	3,78
2	GIVENDFENRS	1507,747	0,02	43,52		41	-1,063636364	3,82
2	KFLVDDGGLVDREPI	2130,2091	-0,0317	53,08	27	42	-0,1	3,96
2	TIEEDGEHNGTKES	2002,9849	-0,0291	70,14	28	40	-1,9	4
2	DWTQDYRDRV	1581,7739	-0,0251	48,63	32	39	-2,17	4,18
2	KEKVEEQEQQQQ	2217,2209	-0,017	81,05	36	43	-2,925	4,26

2	ASSEKNGSTPDTQ	1778,9052	-0,0263	54,5	28	41	-1,684615385	4,08
2	VDEQSPRPG	1212,6302	-0,0177	55,75		40	-1,688888889	4,08
2	KYDITIDEESPRPG	2077,1097	-0,0361	67,93	25	42	-1,414285714	4,06
2	TKYDITIDEESPRPG	2178,1574	-0,0389	50,38	29	42	-1,366666667	4,06
2	FGKEEETSVGDRL	1924,0307	-0,0107	44,83	26	42	-1,069230769	4,16
2	FTKDQVNDDQLKNIT	2465,3734	0,0217	47,77	33	43	-1,226666667	4,18
2	HNPENDLESNNKRDPFEA	2583,2719	-0,0401	70,58	29	40	-2,15	4,28
2	SGDEGEAEEGGGRKG	1891,9277	-0,0228	68,63		40	-1,82	4
2	LSGDEGEAEEGGGRKG	2005,0118	-0,02	80,91	37	41	-1,46875	4
2	VFIQELI	1089,6637	-0,0071	48,78		42	1,828571429	3,3
2	KDAQPNDS	1331,7087	-0,0132	61,4	34	41	-2,3125	3,88
2	MNSTEPFPRVLDPV	1829,9549	-0,0283	42,29	23	41	-0,314285714	4,08
2	NIERDILENS	1430,7568	0,0091	42,95	33	42	-1	3,82
2	SLDQDGFFKII	1739,9864	0,0296	71,99	37	41	0,254545455	3,88
2	VKDREQFVDDLEQA	2149,1421	0,0438	53,02	27	43	-1,15	3,84
2	KADRDESSPY	1624,8462	-0,0087	42,62	28	42	-2,16	4,3
2	IYEVAPG	976,5432	-0,0199	48,03	28	42	0,528571429	3,3
2	RNPSDITQEEYNA	1764,8481	0,0065	40,24	27	40	-1,815384615	3,82
2	ILDEFYDRA	1369,7081	-0,016	48,49		40	-0,377777778	3,7
2	DATNENFRLV	1406,7357	-0,0148	61,1		41	-0,66	4,08
2	PWFGILEA	1160,6433	-0,0217	41,03	30	41	0,8125	3,3
2	IYNNLFDWI	1425,7495	0,0418	43,87		41	0,322222222	3,1
2	KSPNDYDDRQV	1793,9314	-0,009	46,42	27	42	-2,309090909	4,18
2	TYQERDPA	1207,6036	-0,0165	41,56	25	40	-2,1	4,08
2	ATYQERDPA	1278,6407	-0,016	43,73	27	40	-1,666666667	4,08
2	SLREELLFPFAPI	1760,0075	-0,0077	85,54		42	0,6	4,26
2	NIDVKDQKAVDDFLIS	2506,4251	-0,0127	73,69	30	43	-0,2375	3,88
2	YHYPEIEDLVDRI	1889,9726	0,0649	69,62	31	43	-0,684615385	4,06
2	DLKEDITDFDKL	2138,2079	0,0134	64,36	32	42	-0,925	3,76
2	KSEPQQPEDNAETA	2001,0056	-0,0418	74,57	27	40	-2,107142857	3,68

2	SQEEGEDNGGEDNKKLRGA	2719,3981	0,0076	51,93	33	43	-2,136842105	4,12
2	DISRNINDLLNKDFY	2297,2421	0,0517	71,73	31	43	-0,806666667	4,18
2	ETFEQPSQREEA	1678,8001	-0,0075	66,61		39	-2	3,8
2	FGKEEETSVGDRL	1924,0307	-0,0107	44,83	26	42	-1,069230769	4,16
2	FTKDQVNDDQLKNIT	2465,3734	0,0217	47,77	33	43	-1,226666667	4,18
2	TIIEPKEEEPI	1755,0071	-0,0001	65,58	40	41	-0,754545455	3,8
2	RVVPNEKADDDSVTII	2228,2418	0,0336	43,2	25	42	-0,35	3,96
2	RNPSDITQEEYNA	1764,8481	0,0065	40,24	27	40	-1,815384615	3,82
2	EDQVAEEERRA	1559,7742	-0,0282	50,6	34	39	-2,018181818	4
2	FVNDIFERI	1380,7604	0,0146	47,12	32	42	0,422222222	4,08
2	RDYNHSTDEEYQ	1784,7804	-0,0337	38,55	21	33	-2,733333333	4,06
2	NAPIEEKPLI	1580,9543	0,051	54,68	31	37	-0,3	4,26
2	GNYEGKGGDEA	1553,7727	-0,0302	53,28	26	40	-1,727272727	3,82
2	TYQERDPA	1207,6036	-0,0165	41,56	25	40	-2,1	4,08
2	ATYQERDPA	1278,6407	-0,016	43,73	27	40	-1,666666667	4,08
2	VITGESR	989,5708	-0,0106	58,88		42	-0,171428571	6,98
2	DVSEDKEAQRPA	1801,9576	-0,0321	58,67	36	42	-1,708333333	4,06
2	KEILSILDR	1543,9703	-0,0335	56,29		40	0,044444444	6,92
2	IGSFKDWEFF	1732,9231	-0,0347	55,14	28	42	-0,01	4,08
2	VLREGLEAV	1213,7233	0,0521	43,29	28	38	0,655555556	4,26
2	REELEELVKPL	1812,0762	-0,0183	57,69		41	-0,763636364	4,26
2	KEEPEEQEQLRES	2088,074	-0,0192	46,27	31	42	-2,692307692	3,96
2	VAQDVNGERQQ	1471,7582	-0,0246	50,09	32	40	-1,427272727	4,08
2	GPWYNEPL	1203,6127	0,0325	45,96	33	42	-1,125	3,3
2	RQNDYDSQ	1253,5839	0,0109	41,15	30	39	-3,0125	3,88
2	RNDNLDDKSTV	1733,9314	0,0171	44,23	32	42	-1,763636364	4,18
2	NMESNESPRNVPI	1714,8511	0,0243	48,43	25	41	-1,246153846	4,26
2	KDIQTIPEEERTEPVPEGY	2687,4059	-0,0387	48,41	34	42	-1,452631579	3,88
2	VIERKFDIDEELV	2062,1716	-0,0202	47,33	30	43	-0,146153846	3,92
2	IFFFYPETA	1362,7063	-0,0188	47,07	29	41	0,844444444	3,3

2	EDGGEDSNSRRYPS	1796,8128	-0,0247	47,02	27	36	-2,328571429	4,06
2	FIWQEFA	1168,612	0,0022	46,83	38	41	0,571428571	3,3
2	FYGILNAR	1181,676	0,0494	41,17		39	0,4	9,84
2	YYFGQEG	1091,5127	-0,0177	45,36	31	37	-1,085714286	3,3
2	GNSIISLEALDAI	1543,866	0,0708	44,14		40	0,938461538	3
2	LFEQFDRL	1295,7077	0,0498	49,01		41	-0,225	4,08
2	EYDIYSPEDNYKRV	2248,1417	-0,0091	42,21	22	42	-1,678571429	4,06
2	KREREDDDEPA	1816,9321	-0,0402	43,35	33	41	-3,063636364	4,06
2	IYNNLFDWI	1425,7495	0,0418	43,87		41	0,322222222	3,1
2	IYNNLFDWI	1425,7495	0,0418	43,87		41	0,322222222	3,1
2	IAREIL	942,6065	-0,0122	41,86		39	1,1	6,98
2	IAREII	942,6065	-0,0122	41,86		39	1,216666667	6,98
3	LAGVEIL	942,5953	-0,0091	49,59		39	2,028571429	3,3
3	GFVQEFQ	1082,56	0,0075	52,69	34	41	-0,157142857	3,3
3	SLTENWLI	1203,6702	0,018	53,7		42	0,3375	3,3
3	EFHPFDPV	1215,6127	0,001	42,04	33	40	-0,45	4,06
3	RDGQLVEIPA	1325,7506	-0,0154	65,9	29	41	-0,27	4,08
3	GEARPVPEEY	1374,6982	0,0318	45,81	27	42	-1,39	3,96
3	DVHENYENKVA	1774,9255	0,0438	73,02	36	42	-1,427272727	4,42
3	ADEKESLVVKFV	2050,2283	-0,0295	53,79	35	42	0,158333333	4,44
3	EEDHPIPEDVHENYENKVA	2721,3288	-0,0281	57,7	23	40	-1,663157895	4,02
3	LAGVEIL	942,5953	-0,0091	49,59		39	2,028571429	3,3
3	SLTENWLI	1203,6702	0,018	53,7		42	0,3375	3,3
3	EFHPFDPV	1215,6127	0,001	42,04	33	40	-0,45	4,06
3	DVHENYENKVA	1774,9255	0,0438	73,02	36	42	-1,427272727	4,42
3	EEDHPIPEDVHENYENKVA	2721,3288	-0,0281	57,7	23	40	-1,663157895	4,02
3	GGIGTVPV	927,5593	-0,0222	41,3	27	39	1,175	6,02
3	GGVGEFEA	993,497	-0,0133	54,76	33	40	0,075	3,12
3	AGGVGEFEA	1064,5341	-0,0208	69,68		40	0,266666667	3,12

3	AGGVGEFEAGIS	1321,6717	-0,0117	58,78	28	41	0,475	3,12
3	DAIEQPSRPT	1341,7091	-0,0141	51,93		41	-1,34	4,08
3	EMHHEQLEQ	1408,6608	-0,0225	61,16	28	37	-2,022222222	4,52
3	IDAIEQPSRPT	1454,7932	-0,0149	50,4	35	42	-0,809090909	4,08
3	ELLEKNDR	1473,8557	-0,0289	43,23		42	-1,85	4,44
3	RFDELLEK	1506,8811	-0,0071	42,3	37	42	-1,0625	4,44
3	KTLLEAIDAI	1543,9591	-0,0187	61,98		40	0,86	4,08
3	RFDELLEKN	1620,9241	0,0319	44,88	35	41	-1,333333333	4,44
3	RFDELLEKNDR	1892,0521	0,0119	53,36	40	42	-1,818181818	4,56
3	DAIEQPSRPTDKPL	2024,1308	-0,0488	52,2	35	43	-1,328571429	4,3
3	EMHHEQLEQGVPGDNV	2046,9632	0,0346	45,99	29	41	-1,2	4,14
3	FFYYGT	1025,5061	-0,0184	39,64	29	39	0,316666667	5,86
3	AFGGFVFG	1029,5487	0,0233	42,67	27	41	1,65	6,02
3	FVFGWDTG	1156,5756	0,0271	47,92		40	0,4875	3,1
3	GFVFGWDTG	1213,5971	0,0113	59,07	34	40	0,388888889	3,1
3	MWEEGVLPW	1374,6845	-0,0214	53,89	35	39	-0,1	3,12
3	LAGVEAEKL	1386,8488	0,0295	64,33		39	0,455555556	4,26
3	YGEGEEHEPVVEIPKRPA	2493,3269	-0,0451	115,92	32	43	-1,183333333	4,42
3	FFYYGT	1025,5061	-0,0184	39,64	29	39	0,316666667	5,86
3	AFGGFVFG	1029,5487	0,0233	42,67	27	41	1,65	6,02
3	FVFGWDTG	1156,5756	0,0271	47,92		40	0,4875	3,1
3	GFVFGWDTG	1213,5971	0,0113	59,07	34	40	0,388888889	3,1
3	MWEEGVLPW	1374,6845	-0,0214	53,89	35	39	-0,1	3,12
3	LAGVEAEKL	1386,8488	0,0295	64,33		39	0,455555556	4,26
3	YGEGEEHEPVVEIPKRPA	2493,3269	-0,0451	115,92	32	43	-1,183333333	4,42
3	TGGQAFPQ	1033,5396	0,0113	53,7		42	-0,6875	6,02
3	LLEKLEIV	1413,9212	0,0073	38,44		37	1,15	4,26
3	DTRKDEQERGI	1803,9845	0,0097	50,1	37	42	-2,454545455	4,56
3	TKEGPIFGEEMR	1850,9966	-0,0085	68,43	36	42	-1,066666667	4,56
3	NKDTDAEGKPL	1874,0718	-0,0089	58,42	31	42	-1,718181818	4,3

3	KEGPIFGEEMRSV	1936,0494	0,0405	47,26	29	42	-0,669230769	4,56
3	TKEGPIFGEEMRS	1938,0286	-0,0133	60,07	31	42	-1,046153846	4,56
3	TNKDTDAEGKPL	1975,1194	-0,0123	50,93	27	42	-1,633333333	4,3
3	TKEGPIFGEEMRSV	2037,097	0,0433	58,8	29	43	-0,671428571	4,56
3	HGMKEEVPGWQEY	2047,0239	0,0094	86,25	27	42	-1,507692308	4,54
3	KGDEKDLEGKA	2105,2503	-0,0021	46,46	29	41	-1,9	4,56
3	KDTDAEGKPLERA	2116,2096	0,0305	49,02	28	42	-1,661538462	4,56
3	LKGDEKDLEGK	2147,2973	-0,0465	48,98	30	42	-1,718181818	4,56
3	LKGDEKDLEGKA	2218,3344	-0,0437	64,79	33	42	-1,425	4,56
3	NKDTDAEGKPLERA	2230,2526	0,0859	62,79	29	41	-1,792857143	4,56
3	TNKDTDAEGKPLERA	2331,3002	0,0714	49,75	26	42	-1,72	4,56
3	HGMKEEVPGWQEYYDKL	2795,4561	-0,0356	47,97	31	43	-1,441176471	4,62
3	FFYYGT	1025,5061	-0,0184	39,64	29	39	0,316666667	5,86
3	AFGGFVFG	1029,5487	0,0233	42,67	27	41	1,65	6,02
3	FVFGWDTG	1156,5756	0,0271	47,92		40	0,4875	3,1
3	GFVFGWDTG	1213,5971	0,0113	59,07	34	40	0,388888889	3,1
3	FFYYGT	1025,5061	-0,0184	39,64	29	39	0,316666667	5,86
3	AFGGFVFG	1029,5487	0,0233	42,67	27	41	1,65	6,02
3	FVFGWDTG	1156,5756	0,0271	47,92		40	0,4875	3,1
3	GFVFGWDTG	1213,5971	0,0113	59,07	34	40	0,388888889	3,1
3	FFYYGT	1025,5061	-0,0184	39,64	29	39	0,316666667	5,86
3	AFGGFVFG	1029,5487	0,0233	42,67	27	41	1,65	6,02
3	FVFGWDTG	1156,5756	0,0271	47,92		40	0,4875	3,1
3	GFVFGWDTG	1213,5971	0,0113	59,07	34	40	0,388888889	3,1
3	IQVFEGERA	1276,6978	-0,0078	56,01	40	42	-0,233333333	4,26
3	KEEDEKESQRIA	2148,1994	0,0528	47,14	41	42	-2,316666667	4,36
3	EFKGETKNFTPEQI	2354,309	0,0353	61,69	27	42	-1,328571429	4,56
3	EFKGETKNFTPEQIS	2441,3411	-0,0077	68,2	26	43	-1,293333333	4,56
3	EKFKEEDEKESQRIA	2781,5683	-0,0577	64,34	32	43	-2,16	4,5
3	IQVFEGERA	1276,6978	-0,0078	56,01	40	42	-0,233333333	4,26

3	KEEDEKESQRIA	2148,1994	0,0528	47,14	41	42	-2,316666667	4,36
3	EFKGETKNFTPEQI	2354,309	0,0353	61,69	27	42	-1,328571429	4,56
3	EFKGETKNFTPEQIS	2441,3411	-0,0077	68,2	26	43	-1,293333333	4,56
3	EKFKEEDEKESQRIA	2781,5683	-0,0577	64,34	32	43	-2,16	4,5
3	FFYYGT	1025,5061	-0,0184	39,64	29	39	0,316666667	5,86
3	FFSEYGCNEVTPR	1776,8344	-0,044	70,06	27	37	-0,576923077	4,26
3	GYSSNDDEDTRVKMT	2175,052	0,0646	51,3	25	42	-1,633333333	3,96
3	FYGIIDLI	1181,6899	0,0234	60,05		39	1,8625	3,1
3	GIFYGIIDLI	1351,7954	-0,0259	73,69	36	42	1,9	3,1
3	FFAGNEVI	1124,6069	-0,0136	52,71		41	1,0875	3,3
3	SRPFTEIEAI	1390,7659	0,0237	52,41	34	42	-0,1	4,26
3	DFENLKNEYSKV	2172,2035	0,0615	58,1	25	42	-1,383333333	4,44
3	TDFENLKNEYSKV	2273,2512	0,044	63,31	29	42	-1,330769231	4,44
3	FFYYGT	1025,5061	-0,0184	39,64	29	39	0,316666667	5,86
3	AFGGFVFG	1029,5487	0,0233	42,67	27	41	1,65	6,02
3	FVFGWDTG	1156,5756	0,0271	47,92		40	0,4875	3,1
3	GFVFGWDTG	1213,5971	0,0113	59,07	34	40	0,388888889	3,1
3	KQIIIDAESA	1544,9179	0,0072	67,15		41	0,19	4,08
3	FFYYGT	1025,5061	-0,0184	39,64	29	39	0,316666667	5,86
3	FFYYGT	1025,5061	-0,0184	39,64	29	39	0,316666667	5,86
3	FFYYGT	1025,5061	-0,0184	39,64	29	39	0,316666667	5,86
3	ELDWMERELPKKG	2317,3072	-0,0088	50,49	29	43	-1,515384615	4,64
3	QSELDWMERELPKKG	2532,3978	-0,0143	49,01	29	43	-1,6	4,64
3	FYGIIDLL	1181,6899	0,0234	60,05		39	1,775	3,1
3	FFYYGT	1025,5061	-0,0184	39,64	29	39	0,316666667	5,86
3	KIPVLEQELV	1625,0169	-0,0277	46,22	30	39	0,45	4,26
3	LSDFFDGKQLEKS	2200,2348	-0,0488	64,05	28	43	-0,815384615	4,3
3	HHHATDDVEDAAPETKEA	2430,1817	0,013	106,46	34	41	-1,45	4,36
3	LSDFFDGKQLEKS	2200,2348	-0,0488	64,05	28	43	-0,815384615	4,3
3	GRRFEDAEVQ	1434,7418	0,0089	48,78	31	42	-1,46	4,44

3	IGRRFEDAEVQ	1547,8259	-0,0213	49,28	27	42	-0,918181818	4,44
3	SFGSLQ	866,4701	-0,0117	41,34		41	0,183333333	6,02
3	NNIIVSDTTLESVEGFSTLKKV	3080,7578	0,0763	124,06	38	41	0,081818182	4,44
3	SFGSLQ	866,4701	-0,0117	41,34		41	0,183333333	6,02
3	NNIIVSDTTLESVEGFSTLKKV	3080,7578	0,0763	124,06	38	41	0,081818182	4,44
3	KVLEELF	1334,8215	0,0286	43,41		38	0,528571429	4,26
3	LKEFEGGVI	1448,8645	-0,0335	51,85		42	0,4	4,26
3	KVLEELFQ	1462,8801	0,0209	50,14		40	0,025	4,26
3	KDILDEFR	1492,8655	-0,0008	48,63	34	42	-0,975	4,3
3	LKEFEGGVIII	1675,0326	0,0448	87,62		38	1,145454545	4,26
3	NELLKDETVAPR	1842,0616	-0,0427	81,35		42	-0,925	4,44
3	NQVADEKERA	1616,8887	-0,0185	59,61	33	42	-1,81	4,44
3	ALPWFWEHYNPE	1816,8776	0,0451	59,84	30	41	-0,966666667	4,24
3	NGEDKEIVDGKVLK	2459,4771	-0,0128	56,17	34	41	-0,95	4,56
3	RYHIEEEGS	1347,6622	-0,0028	45,14	24	40	-1,8	4,54
3	AGFLLEKEL	1476,8957	0,0037	53,21	39	41	0,522222222	4,26
3	VWNGPPGVFEFEK	1963,0609	-0,0083	82,28	32	43	-0,407692308	4,26
3	WNGPPGVFEFEKF	2011,0609	0,0424	52,4	30	42	-0,515384615	4,26
3	VWNGPPGVFEFEKF	2110,1294	-0,002	89,92	35	43	-0,178571429	4,26
3	VWNGPPGVFEFEKFA	2181,1665	-0,0131	78,17	41	43	-0,046666667	4,26
3	KKGEQELEGL	1817,0867	-0,025	47,55	37	42	-1,5	4,56
3	VKKGEQELEGL	1916,1551	0,0035	61,48	33	40	-0,981818182	4,56
3	VIVKKGEQELEGL	2128,3076	-0,0499	57,62	29	41	-0,161538462	4,56
3	ALDSNNERLKESV	1932,0682	-0,0413	46,58	32	42	-1,069230769	4,44
3	NYKEESKEQA	1912,051	0,0195	62,39	27	42	-2,56	4,56
3	SPNAKKEYEPEST	2166,1777	0,0443	59,97	24	42	-2,061538462	4,56
3	EEAEQPKTDYKKI	2494,4454	-0,0531	49,18	26	43	-2,038461538	4,64
3	LQKERELDVI	1700,0238	0,0351	43,12	32	39	-0,61	4,44
3	READVFKNGEMPFPI	2207,1814	-0,0037	46,56	30	43	-0,533333333	4,44
3	RANQEVLEWL	1485,8142	-0,0113	53,49	38	41	-0,58	4,26

3	GRRFEDAEVQ	1434,7418	0,0089	48,78	31	42	-1,46	4,44
3	IGRRFEDAEVQ	1547,8259	-0,0213	49,28	27	42	-0,918181818	4,44
3	LFLPPVAV	1083,6895	-0,0241	40,16	34	38	2,175	6,02
3	ALPWFWEHYNPE	1816,8776	0,0451	59,84	30	41	-0,966666667	4,24
3	LHGEGSGEEKG	1556,82	-0,0103	69,78		41	-1,472727273	4,54
3	KIQDKEGIPPDQQ	2182,2566	-0,0156	58,6	27	42	-1,8	4,3
3	LHGEGSGEEKG	1556,82	-0,0103	69,78		41	-1,472727273	4,54
3	LHAEQDAEVRQQ	1651,8481	-0,0185	82,95	36	41	-1,425	4,42
3	FFYYGT	1025,5061	-0,0184	39,64	29	39	0,316666667	5,86
3	LGGGSSDRREEYPVPDAPPYRA	2617,3087	-0,038	55,77	26	41	-1,259090909	4,56
3	NYKEENKEQA	1939,0619	0,0025	62,77	33	42	-2,83	4,56
3	ELPKVEDLKI	1870,1748	-0,0447	51,47	39	40	-0,36	4,44
3	AETKEPTKEPT	1917,1027	0,0055	59,23	31	42	-1,981818182	4,56
3	WVMDTNKEERNDGKTI	2622,4044	0,0421	63,96	30	43	-1,5875	4,56
3	REVLGEQGKDV	1686,967	-0,0148	57,19	28	42	-1	4,44
3	VFPFVFEKEPV	1795,0326	-0,0368	65,5	26	42	0,627272727	4,26
3	KHEEIDT	1328,7342	-0,0095	45,27	34	41	-1,971428571	4,42
3	KELNPDITDETNEGKTGPKLV	3213,8268	0,0754	45,2	21	41	-1,238095238	4,28
3	KHEEIDT	1328,7342	-0,0095	45,27	34	41	-1,971428571	4,42
3	SHMSPEDAEEELKKL	2429,308	0,0055	56,79	28	43	-1,36	4,36
3	EAPQLKDEYRTY	1970,0515	-0,0519	42,26	28	42	-1,808333333	4,44
3	HAEQFPFEKDVNVV	2116,1359	-0,0163	78,72	31	42	-0,442857143	4,42
3	KFHTENAEDQDRV	2046,0536	0,0261	45,94	27	42	-1,884615385	4,54
3	VGNKSDLENEKQV	2146,2202	-0,0374	46,68	30	43	-1,369230769	4,44
3	KEVEEKLGENPKIT	2529,5189	-0,0151	56,69		41	-1,385714286	4,72
3	KKDYSIDSIE	1884,0813	-0,0405	45,87	28	42	-1,22	4,3
3	ILDQEKYDRI	1750,0031	0,0274	43,35	27	41	-1,09	4,3
3	VILDQEKYDRI	1849,0715	0,039	42,76	27	41	-0,609090909	4,3
3	LYKYEIDIA	1584,9169	-0,0421	63,65	37	42	0,122222222	4,08
3	MEEHFNPSTGEKT	1963,9715	-0,0072	49,64	29	41	-1,584615385	4,54

3	LDNQEEGKKS	1834,0405	0,0237	44,93	29	42	-2,27	4,44
3	KWEPAGEV	1372,7756	-0,0252	46,97	33	41	-0,975	4,26
3	LLKEEKFPDVL	2017,2432	0,0112	61,98	35	39	-0,136363636	4,44
3	FFYYGT	1025,5061	-0,0184	39,64	29	39	0,316666667	5,86
3	RYVFDIHPEDVL	1730,9195	-0,0315	41,44	25	41	-0,133333333	4,3
3	KALQEDHENSSPF	1959,0103	-0,0386	93,6	28	41	-1,492307692	4,42
3	SFDVFIPEFGIEKRV	2240,2611	0,0211	42,36	37	42	0,273333333	4,44
3	LVKKDELTLEGI	2044,2752	-0,0437	49,77	31	41	0,058333333	4,44
3	KLFEQEGV	1406,8175	0,0137	49,95		42	-0,5	4,26
3	LVEFGKEIL	1504,927	0,0328	52,03		39	0,866666667	4,26
3	STGKDYKGEADPG	2011,0831	0,0124	55,26	37	42	-1,7	4,3
3	KFDDVDGKPLVEKI	2518,5182	-0,0394	44,73	24	42	-0,585714286	4,46
3	TLKDEHDLERA	1783,9834	-0,0009	47,67	25	42	-1,536363636	4,54
3	RANQEVLEWL	1485,8142	-0,0113	53,49	38	41	-0,58	4,26
3	LKDILGDQVEKV	2043,2548	-0,045	55,53	30	41	-0,141666667	4,3
3	FFDTEKKIEIA	2027,1911	-0,0393	43,95	35	42	-0,236363636	4,44
3	FFDTEKKIEIA	2027,1911	-0,0393	43,95	35	42	-0,236363636	4,44
3	DKYDENNPEHR	1873,9324	-0,0353	46,6	24	40	-3,227272727	4,54
3	VHDKILEDLVFPT	1983,1447	0,0337	75,42	30	41	0,261538462	4,3
3	NVSKDELREKLAEV	2316,3621	0,0122	55,7	31	41	-0,914285714	4,64
3	LKVEDFLER	1605,9496	-0,0343	50,18	28	42	-0,477777778	4,44
3	LKVEDFLER	1605,9496	-0,0343	50,18	28	42	-0,477777778	4,44
3	EAEKKGIEA	1660,9968	0,0119	51,01	34	41	-1,177777778	4,56
3	EAEKKGIEAA	1732,0339	0,009	46,76	34	41	-0,88	4,56
3	KKQEELDA	1646,9811	0,0483	50,45		39	-2,025	4,44
3	FAGEESGDLPFRKG	1967,0518	0,0203	59,22	30	42	-0,807142857	4,44
3	SNVEEKPGDRG	1644,8837	-0,0299	41,55	31	41	-1,945454545	4,44
3	SNVEEKPGDRGALA	1900,042	0,0436	54,06	39	42	-1	4,44
3	DLRDDKVIIEKL	2143,3185	-0,0449	47,76	32	41	-0,458333333	4,46
3	KWTGIPVGEEDRV	1943,0882	0,0266	43,6	29	42	-0,769230769	4,44

3	IQQDLLDEKRFA	1933,1038	0,0055	53,49	35	42	-0,766666667	4,3
3	KEILSILDR	1543,9703	-0,0299	41,36		40	0,044444444	6,92
3	GSRTPGHPEFELPGVEVT	2137,1007	0,043	52,8	25	43	-0,633333333	4,54
3	YFDFKYDVEKA	2111,1548	-0,0484	55,55	25	43	-0,845454545	4,3
3	GHGLKEIDVEFI	1814,0344	-0,0368	46,37	27	42	0,116666667	4,42
3	FFYYGT	1025,5061	-0,0184	39,64	29	39	0,316666667	5,86
3	FNQKSETTLDTTPEAESVPEKRA	3264,7446	-0,0537	50,4	33	43	-1,304347826	4,36
3	TLNEESKDKV	1849,0765	-0,03	43,6	30	42	-1,53	4,44
3	FYGILNAR	1181,676	0,0373	39,72		39	0,4	9,84
3	TKLPNFNEVSPEERIPL	2440,3731	-0,0499	44,97	25	43	-0,770588235	4,56
3	SFGSLQ	866,4701	-0,0117	41,34		41	0,183333333	6,02
3	KFDDVDGKPLVEKI	2518,5182	-0,0394	44,73	24	42	-0,585714286	4,46
3	ELLEKNDR	1473,8557	-0,0289	43,23		42	-1,85	4,44
4	LHLEIF	999,5956	0,0022	41,29		39	1,366666667	5,12
4	MFFVGPI	1038,5775	0,0058	42,51	26	41	2,028571429	6,02
4	NFGAMNGI	1051,5323	0,0019	51,6	37	41	0,4	6,02
4	FHPFDPV	1086,5701	0,0088	43,34	29	40	-0,014285714	4,94
4	LHLEIFLG	1169,7011	0,0365	42,76	23	38	1,45	5,12
4	QNFGAMNGI	1179,5909	-0,0125	42,57	36	40	-0,033333333	6,02
4	RDGQLVEIPA	1325,7506	-0,0123	41,31	29	41	-0,27	4,08
4	DKHYGDQT	1420,7352	-0,0157	45,36	31	41	-2,5	5,1
4	VDKHYGDQTF	1666,8721	0,0154	50,47	27	42	-1,3	5,1
4	YDNAPYSPKPV	1707,9238	0,0122	69,95	25	42	-1,190909091	6,56
4	AYDNAPYSPKPV	1778,9609	0,0012	68,44	30	42	-0,941666667	6,56
4	MTGDGVNDAPSLKKA	2190,2287	0,0052	48,25	35	43	-0,58	6,72
4	LHLEIF	999,5956	0,0022	41,29		39	1,366666667	5,12
4	MFFVGPI	1038,5775	0,0058	42,51	26	41	2,028571429	6,02
4	FHPFDPV	1086,5701	0,0088	43,34	29	40	-0,014285714	4,94
4	LHLEIFLG	1169,7011	0,0365	42,76	23	38	1,45	5,12

4	MTGDGVNDAPSLKKA	2190,2287	0,0052	48,25	35	43	-0,58	6,72
4	GHVDSG	799,4027	-0,0063	41,95	27	37	-0,683333333	4,94
4	GGIGTVPV	927,5593	-0,0226	46,17	31	40	1,175	6,02
4	DAPGHRDF	1142,5672	-0,0001	39,7	31	39	-1,5125	5,1
4	EMHHEQL	1151,5596	0,0031	47,69	29	40	-1,6	5,12
4	DAPGHRDFI	1255,6512	0,0308	45,14	33	41	-0,844444444	5,1
4	IDAPGHRDFI	1368,7353	0,0012	46,86	34	41	-0,31	5,1
4	HHEQLEQGVPG	1458,7418	-0,016	59,73	33	40	-1,345454545	5,12
4	VIDAPGHRDFI	1467,8037	0,0315	45,29		41	0,1	5,1
4	RFDELLEKN	1620,9241	0,0438	42,23	38	41	-1,333333333	4,44
4	RFDELLEKNDR	1892,0521	0,0365	46,98	34	42	-1,818181818	4,56
4	FGGFVFG	958,5116	0,0189	42,77	37	41	1,628571429	6,02
4	FGGFVFGW	1144,5909	0,0055	45,87	26	40	1,3125	6,02
4	NWLWGFL	1163,6331	0,0319	42,29	33	42	0,671428571	6,02
4	YGEGEEHEPVVEIPKRPA	2493,3269	0,0066	62,36	27	43	-1,183333333	4,42
4	FGGFVFG	958,5116	0,0189	42,77	37	41	1,628571429	6,02
4	FGGFVFGW	1144,5909	0,0055	45,87	26	40	1,3125	6,02
4	NWLWGFL	1163,6331	0,0319	42,29	33	42	0,671428571	6,02
4	YGEGEEHEPVVEIPKRPA	2493,3269	0,0066	62,36	27	43	-1,183333333	4,42
4	TGGQAFPQ	1033,5396	0	46,74		42	-0,6875	6,02
4	QGPNYVPG	1059,5552	-0,0046	50,1	37	41	-1,0125	5,92
4	HEIKDSVVA	1454,8499	0,0369	55,73	38	39	-0,022222222	5,22
4	DTRKDEQERGI	1803,9845	0,0199	42,22	33	42	-2,454545455	4,56
4	DTRKDEQERGITI	2018,1162	0,0347	47,95		42	-1,784615385	4,56
4	TKEGPIFGEEMRSV	2037,097	0,0786	65,21	30	42	-0,671428571	4,56
4	LKGDEKDLEGKA	2218,3344	0,0087	45,01	29	40	-1,425	4,56
4	NKDTDAEGKPLERA	2230,2526	0,0918	44,94	23	41	-1,792857143	4,56
4	FGGFVFG	958,5116	0,0189	42,77	37	41	1,628571429	6,02
4	HQGPLEGS	1052,5454	-0,0242	52,03	26	40	-1,2	5,12
4	FGGFVFGW	1144,5909	0,0055	45,87	26	40	1,3125	6,02

4	NWIWGFL	1163,6331	0,0319	42,29	33	42	0,771428571	6,02
4	FGGFVFG	958,5116	0,0189	42,77	37	41	1,628571429	6,02
4	FGGFVFGW	1144,5909	0,0055	45,87	26	40	1,3125	6,02
4	NWLWGFL	1163,6331	0,0319	42,29	33	42	0,671428571	6,02
4	APDHPFIQQ	1280,6716	0,0432	44,99	32	42	-0,866666667	4,94
4	FGGFVFG	958,5116	0,0189	42,77	37	41	1,628571429	6,02
4	FGGFVFGW	1144,5909	0,0055	45,87	26	40	1,3125	6,02
4	NWIWGFL	1163,6331	0,0319	42,29	33	42	0,771428571	6,02
4	YFNDSQRQ	1285,6254	0,0178	52,88	26	40	-2,225	6,7
4	AGGAPGGFPG	1015,529	0,0178	45,63	27	40	0,12	6,02
4	YFNDSQRQ	1285,6254	0,0178	52,88	26	40	-2,225	6,7
4	HFANDRVDII	1427,7724	0,0153	59,36	35	41	-0,04	5,1
4	NWIWGFL	1163,6331	0,0319	42,29	33	42	0,771428571	6,02
4	DLFNSYKTV	1543,8652	0,0663	55,48	30	40	-0,322222222	6,6
4	SNLGWIPV	1113,6385	-0,01	44,96		40	0,6625	6,02
4	GYHFGPSDA	1178,5559	0,0127	52,52	31	40	-0,733333333	4,94
4	FIPTLLVPI	1240,7998	0,0278	36,69	30	34	2,188888889	6,02
4	MLAGTPFWQ	1278,6634	0,005	72,19	35	42	0,355555556	6,02
4	GIFYGIIDLI	1351,7954	0,0074	53,81	34	41	1,9	3,1
4	LKDLGINTV	1429,891	0,0294	53,05		39	0,477777778	6,66
4	YTVDNSQDHSH	1530,6902	-0,0098	36,26	26	36	-1,8	4,94
4	KLYQDYSVPV	1668,9492	0,0495	80,22	34	40	-0,37	6,56
4	DNSQDHSHCMKLL	1984,9865	0,0185	60,73	28	42	-1,069230769	6
4	FGGFVFG	958,5116	0,0189	42,77	37	41	1,628571429	6,02
4	FGGFVFGW	1144,5909	0,0055	45,87	26	40	1,3125	6,02
4	NWIWGFL	1163,6331	0,0319	42,29	33	42	0,771428571	6,02
4	GHQQSESLPQ	1338,6731	-0,022	70,8	36	40	-1,7	5,12
4	YFNDSQRQ	1285,6254	0,0178	52,88	26	40	-2,225	6,7
4	NHLAEVA	981,5446	0,0047	54,98		40	0,2	5,12
4	SHANQDYY	1225,5566	-0,0154	45,54	28	36	-1,9125	4,94

4	GYCQQQDLHLKTA	1962,0399	0,0249	69,43	31	43	-0,892307692	7,14
4	SNLGWIPV	1113,6385	-0,01	44,96		40	0,6625	6,02
4	IAPNLFMSI	1233,6994	0,0005	48,82	32	42	1,488888889	6,02
4	AIAPNLFMSI	1304,7365	0,0203	57,49	35	42	1,52	6,02
4	YFNDSQRQ	1285,6254	0,0178	52,88	26	40	-2,225	6,7
4	DMFGIVVPR	1261,7056	-0,0105	45,11	26	42	0,844444444	6,78
4	LSDFFDGKQLEKS	2200,2348	0,0143	42,14	23	42	-0,815384615	4,3
4	DIFGIVVPR	1243,7492	-0,0027	45,84	28	41	1,133333333	6,78
4	LSDFFDGKQLEKS	2200,2348	0,0143	42,14	23	42	-0,815384615	4,3
4	YFNDSQRQ	1285,6254	0,0178	52,88	26	40	-2,225	6,7
4	SFGSLQTV	1066,5862	-0,0086	65,01	34	41	0,575	6,02
4	SFGSLQTV	1066,5862	-0,0086	65,01	34	41	0,575	6,02
4	HKCPPDANPA	1506,8019	-0,0116	44,73	19	42	-1,28	7,14
4	AHKCPPDANPA	1577,839	-0,0102	55,86	30	42	-1	7,14
4	GYCQQQDLHLKTA	1962,0399	0,0249	69,43	31	43	-0,892307692	7,14
4	NELLKDETVAPR	1842,0616	-0,0051	46,15	32	42	-0,925	4,44
4	SGPQSYQ	994,4923	-0,0157	48,42	26	40	-1,7	5,92
4	YYGYGGTT	1109,5232	0,0247	58,83	32	40	-0,8125	5,82
4	QQGPPQQG	1067,5563	0,006	42,14	34	41	-2,25	6,02
4	YQQGPPQQG	1230,6196	-0,0244	49,43	25	39	-2,144444444	5,92
4	YYQQGPPQQG	1393,6829	-0,0106	54,93	29	40	-2,06	5,86
4	GYYQQGPPQQG	1450,7044	-0,0136	42,13	27	40	-1,909090909	5,86
4	YYQQGPPQQGYPQ	1781,8576	0,0539	58,15	28	42	-2,076923077	5,82
4	YYQQGPPQQGYPQQQPV	2234,0959	0,1088	64,16	22	43	-1,847058824	5,82
4	NQQGYNQQ	1207,5784	0,0117	50,2	34	40	-2,8375	5,92
4	QQYYQQQ	1213,593	-0,02	44,59	32	39	-2,871428571	5,86
4	FFEETHHPT	1372,6615	-0,0096	47	28	39	-1,2	5,12
4	LFFEETHHPT	1485,7455	0,0417	51,73	23	41	-0,7	5,12
4	YHLFDSFT	1257,6233	-0,0064	50,34		39	-0,0125	4,94
4	DAQHKDLLTQV	1724,9827	-0,0048	93,7	36	42	-0,745454545	5,1

4	NYKEESKEQA	1912,051	0,0277	47,48	27	42	-2,56	4,56
4	HLAELAPA	1049,6072	-0,023	49,23	26	42	0,5875	5,12
4	HIHEFSWENVNPIPELR	2345,2119	0,0121	48,27	26	43	-0,794117647	5,22
4	IYAPNPF	1049,5749	-0,0021	46,08	31	42	0,157142857	5,92
4	SHSILTPDGI	1267,6975	0,0036	51,66	29	41	0,18	4,94
4	NEHGLGGLLQGA	1393,7517	0,0242	81,43	29	42	-0,175	5,12
4	YFNDSQRQ	1285,6254	0,0178	52,88	26	40	-2,225	6,7
4	GGGAGMGGMPG	1076,4946	0,008	79,29	30	38	0,109090909	6,02
4	QFFGGSSPF	1201,5971	-0,0099	43,98	31	39	0,1	6,02
4	LAWFPG	918,5166	-0,0043	42,82		42	0,916666667	6,02
4	LFLPPVAV	1083,6895	-0,0001	42,48	34	36	2,175	6,02
4	LSLFLPPVAV	1283,8056	-0,0056	58	30	38	2,04	6,02
4	AGGLILPF	1015,6269	0,0199	40,93		40	1,7875	6,02
4	YGHIDVL	1044,5807	0,0106	48	39	41	0,585714286	4,94
4	NFHGMDFT	1196,5488	-0,0019	47,82	26	38	-0,475	4,94
4	NFHGMDFT	1196,5488	-0,0019	47,82	26	38	-0,475	4,94
4	NLPEHIVPG	1203,6815	-0,0186	70,7	39	42	-0,144444444	5,12
4	NLPEHIVPG	1203,6815	-0,0186	70,7	39	42	-0,144444444	5,12
4	DNFNTHI	1088,5454	0,0045	42,32	30	40	-1,014285714	4,94
4	FDFGGFHI	1167,5916	0,0319	44,46	29	41	0,675	4,94
4	VFDFGGFHI	1266,66	0,021	64,42	41	42	1,066666667	4,94
4	HKCPPDANPA	1506,8019	-0,0116	44,73	19	42	-1,28	7,14
4	AHKCPPDANPA	1577,839	-0,0102	55,86	30	42	-1	7,14
4	GYCQQQDLHLKTA	1962,0399	0,0249	69,43	31	43	-0,892307692	7,14
4	QNIKEVEEKLGENPKIT	2884,7045	0,1331	62,6	21	36	-1,288235294	4,72
4	YVQQQPA	1061,5708	-0,027	41,37	34	41	-1,057142857	5,92
4	YYQQQQPQ	1310,6458	-0,0233	59,7		39	-2,7125	5,86
4	GYTHDGSFQ	1239,5723	0,002	67,99		38	-1,222222222	4,94
4	WHGDWPLPV	1334,6975	0,0183	47,56	26	42	-0,455555556	4,94
4	AWHGDWPLPV	1405,7346	0,0147	53,17	23	42	-0,23	4,94

4	GYTHDGSFQ	1239,5723	0,002	67,99		38	-1,222222222	4,94
4	WHGDWPLPV	1334,6975	0,0183	47,56	26	42	-0,455555556	4,94
4	AWHGDWPLPV	1405,7346	0,0147	53,17	23	42	-0,23	4,94
4	LIEHQPRVDPL	1544,8878	0,0255	45,4	25	41	-0,463636364	5,22
4	GHWNLELNPT	1408,7302	0,0272	93,78		42	-0,97	5,12
4	LVQDLLHPT	1263,739	-0,0232	46,46	31	42	0,344444444	4,94
4	VLVQDLLHPT	1362,8074	-0,0115	53,66	26	42	0,73	4,94
4	HYFTGDGAGRDHDGY	1895,839	0,0758	47,19	27	40	-1,446666667	5,02
4	HWQPGVDSA	1224,609	0,0335	42,3	30	41	-0,877777778	4,94
4	QHLGENTV	1125,5981	-0,004	42,57	35	41	-0,85	5,12
4	HIKDSLDNT	1499,8349	0,0538	43,47	28	41	-1,2	5,1
4	FGGTLNPG	990,5338	-0,0174	61,35		41	-0,05	6,02
4	APVVGQPQPT	1221,692	0,0059	56,9	26	42	-0,27	6,02
4	SHEMFNPF	1236,58	0,0232	44,03	29	40	-0,6375	5,12
4	QHSQDFSPW	1359,6411	-0,0083	41,28	22	38	-1,666666667	4,94
4	QHSQDFSPWYG	1579,7259	0,0037	66,36	29	38	-1,518181818	4,94
4	FGGTLNPG	990,5338	-0,0174	61,35		41	-0,05	6,02
4	SHDDKHIIIDGV	1806,0042	0,0012	55,09	40	42	-0,358333333	5,02
4	HFLLDVL	1084,6484	0,0316	46,1	37	38	1,671428571	4,94
4	KHFNDPDAPPILL	1934,1031	0,0054	74,95	36	42	-0,438461538	5,1
4	THFKDVSTEDDETRKL	2607,4113	0,0171	47,75	37	43	-1,56875	4,62
4	GYTHDGSFQ	1239,5723	0,002	67,99		38	-1,222222222	4,94
4	DAPGHRDF	1142,5672	-0,0001	39,7	31	39	-1,5125	5,1
4	YHGEQIA	1045,5395	-0,0036	42,65	32	41	-0,8	5,12
4	HQLNPDFI	1211,6502	0,0109	51,83	38	41	-0,525	4,94
4	NIPPFDHPYVIA	1610,866	0,006	67,12	35	42	0,125	4,94
4	KETEEEPKKA	2104,2551	0,0547	46,97	31	39	-2,62	4,72
4	GWGNSGGSNNS	1264,5635	-0,0131	44,41	26	35	-1,4	6,02
4	SWINNVAL	1144,6443	-0,0136	44,05	33	41	0,7	6,02
4	HLYDIFSPI	1332,7281	0,0251	48,36	26	42	0,577777778	4,94

4	NHFIEGL	1057,5759	0,009	40,94	33	40	0,071428571	5,12
4	HVGPEQVV	1092,6131	0,0255	53,17		41	0,05	5,12
4	HVGPEQVV	1092,6131	0,0255	53,17		41	0,05	5,12
4	HVGPEQVV	1092,6131	0,0255	53,17		41	0,05	5,12
4	HGFGEYT	1038,4974	0,0239	52,38	29	40	-0,957142857	5,12
4	FHNDVYAQ	1221,5981	0,0373	44,45	27	42	-0,775	4,94
4	FHLDFGPM	1191,595	0,0071	41,42		41	0,325	4,94
5	FLAPGLS	932,5534	0,0034	46,42	39	41	1,342857143	6,02
5	EILQNRG	1057,6083	-0,001	44,1	37	40	-1,014285714	6,98
5	SIPSWQLA	1129,6334	0,0206	52,71		41	0,3125	6,02
5	NIYNAERL	1220,6716	-0,0055	44,33	41	42	-0,775	6,88
5	RGEGHWEIL	1324,7091	0,0115	49,58	27	41	-0,9	5,3
5	IIQNFGAMNGI	1405,759	-0,0239	69,37	36	42	0,790909091	6,02
5	STQHEKET	1416,7614	-0,0004	63,21	33	42	-2,475	5,3
5	DKHYGDQT	1420,7352	-0,0147	42,84	29	41	-2,5	5,1
5	VDKHYGDQT	1519,8037	-0,0221	51,57	31	42	-1,755555556	5,1
5	HENYENKVA	1560,8302	-0,0138	57,87	30	42	-1,822222222	5,3
5	DKHYGDQTF	1567,8037	0,026	47,98	28	42	-1,911111111	5,1
5	AVDKHYGDQT	1590,8408	-0,0301	61,35	29	41	-1,4	5,1
5	LAVDKHYGDQT	1703,9248	0,0335	58,05	32	42	-0,927272727	5,1
5	YDNAPYSPKPV	1707,9238	-0,0052	55,01	26	42	-1,190909091	6,56
5	DGFAEVFPQHK	1731,935	-0,0066	76,32	37	42	-0,727272727	5,22
5	AYDNAPYSPKPV	1778,9609	-0,0041	62,01	28	42	-0,941666667	6,56
5	KVLEFHPFDPV	1785,0231	-0,0278	45,52	29	42	0,045454545	5,22
5	IAYDNAPYSPKPV	1892,0449	-0,0289	70,82	31	42	-0,523076923	6,56
5	DGFAEVFPQHKY	1894,9983	0,0281	70,95	33	42	-0,775	5,22
5	GDGVNDAPSLKKA	1958,1405	0,028	69,74	38	41	-0,761538462	6,72
5	AIAYDNAPYSPKPV	1963,082	-0,0358	43,47	25	42	-0,357142857	6,56
5	FLAPGLSAIIDALKT	1987,2123	0,0386	101,04	35	39	1,18	6,66

5	MTGDGVNDAPSLKKA	2190,2287	-0,0288	93,71	34	43	-0,58	6,72
5	MTGDGVNDAPSLKKA	2206,2236	-0,0302	54,47	27	43	-0,58	6,72
5	ENADGFAEVFPQHKYRVV	2563,3589	0,0427	57,31	32	43	-0,588888889	5,34
5	FLAPGLS	932,5534	0,0034	46,42	39	41	1,342857143	6,02
5	EILQNRG	1057,6083	-0,001	44,1	37	40	-1,014285714	6,98
5	SIPSWQLA	1129,6334	0,0206	52,71		41	0,3125	6,02
5	NIYNAERL	1220,6716	-0,0055	44,33	41	42	-0,775	6,88
5	RGEGHWEIL	1324,7091	0,0115	49,58	27	41	-0,9	5,3
5	HENYENKVA	1560,8302	-0,0138	57,87	30	42	-1,822222222	5,3
5	DGFAEVFPQHK	1731,935	-0,0066	76,32	37	42	-0,727272727	5,22
5	KVLEFHPFDPV	1785,0231	-0,0278	45,52	29	42	0,045454545	5,22
5	DGFAEVFPQHKY	1894,9983	0,0281	70,95	33	42	-0,775	5,22
5	GDGVNDAPSLKKA	1958,1405	0,028	69,74	38	41	-0,761538462	6,72
5	FLAPGLSAIIDALKT	1987,2123	0,0386	101,04	35	39	1,18	6,66
5	MTGDGVNDAPSLKKA	2190,2287	-0,0288	93,71	34	43	-0,58	6,72
5	MTGDGVNDAPSLKKA	2206,2236	-0,0302	54,47	27	43	-0,58	6,72
5	ENADGFAEVFPQHKYRVV	2563,3589	0,0427	57,31	32	43	-0,588888889	5,34
5	GGIGTVPV	927,5593	-0,02	61,5	34	39	1,175	6,02
5	DAPGHRDFI	1255,6512	0,0356	44,97	32	41	-0,844444444	5,1
5	ETGVIKPG	1257,7698	0,0108	49,83	39	40	-0,225	6,84
5	EYPPLGRFA	1277,6971	0,0082	49,17	38	42	-0,5	6,88
5	VIDAPGHRDF	1354,7197	-0,0048	59,84		41	-0,34	5,1
5	VIDAPGHRDFI	1467,8037	0,0179	64,08		42	0,1	5,1
5	FSEYPPLGRFA	1511,7975	-0,0169	51,08	31	41	-0,227272727	6,88
5	TVIDAPGHRDFI	1568,8514	-0,0166	50,18	28	42	0,033333333	5,1
5	FGGFVFG	958,5116	0,0136	42,76	35	40	1,628571429	6,02
5	FGGFVFGW	1144,5909	-0,0067	52,99	36	40	1,3125	6,02
5	NWLWGFL	1163,6331	0,0175	44,09	41	42	0,671428571	6,02
5	RLWPNGQDQPS	1525,784	0,0309	51,95	32	42	-1,818181818	6,78
5	SWGELFSSKT	1598,871	-0,0296	48,47	31	42	-0,52	6,84

5	TRLWPNGQDQPS	1626,8317	0,015	60,19	27	42	-1,725	6,78
5	FGGFVFG	958,5116	0,0136	42,76	35	40	1,628571429	6,02
5	FGGFVFGW	1144,5909	-0,0067	52,99	36	40	1,3125	6,02
5	NWLWGFL	1163,6331	0,0175	44,09	41	42	0,671428571	6,02
5	AHDDKPLY	1415,7814	0,0318	45,67	28	42	-1,425	5,1
5	RLWPNGQDQPS	1525,784	0,0309	51,95	32	42	-1,818181818	6,78
5	SWGELFSSKT	1598,871	-0,0296	48,47	31	42	-0,52	6,84
5	TRLWPNGQDQPS	1626,8317	0,015	60,19	27	42	-1,725	6,78
5	TGGQAFPQ	1033,5396	-0,0076	52,53		42	-0,6875	6,02
5	NLIDSPGHV	1179,6451	-0,0082	55,84	32	42	-0,055555556	4,94
5	HGMKEEVPG	1440,7801	0,0013	59,71	22	42	-1,155555556	5,3
5	WGDSFFNPKT	1655,8714	-0,0117	43,11	32	42	-0,97	6,66
5	QYLHEIKDSV	1688,9503	-0,0031	54,55	34	42	-0,72	5,22
5	QYLHEIKDSVV	1788,0187	0,0374	76,53	31	41	-0,272727273	5,22
5	QYLHEIKDSVVA	1859,0558	0,0235	89,95	37	42	-0,1	5,22
5	FGGFVFG	958,5116	0,0136	42,76	35	40	1,628571429	6,02
5	FGGFVFGW	1144,5909	-0,0067	52,99	36	40	1,3125	6,02
5	NWIWGFL	1163,6331	0,0175	44,09	41	42	0,771428571	6,02
5	RLWPNGQDQPS	1525,784	0,0309	51,95	32	42	-1,818181818	6,78
5	TRLWPNGQDQPS	1626,8317	0,015	60,19	27	42	-1,725	6,78
5	FFFVPETKGL	1641,9536	0,0399	44,58	26	40	0,63	6,84
5	FGGFVFG	958,5116	0,0136	42,76	35	40	1,628571429	6,02
5	FGGFVFGW	1144,5909	-0,0067	52,99	36	40	1,3125	6,02
5	NWLWGFL	1163,6331	0,0175	44,09	41	42	0,671428571	6,02
5	MHDDQPFYK	1637,8278	-0,0207	53,44	26	41	-1,755555556	5,1
5	FFFVPETKGL	1641,9536	0,0399	44,58	26	40	0,63	6,84
5	FGGFVFG	958,5116	0,0136	42,76	35	40	1,628571429	6,02
5	FGGFVFGW	1144,5909	-0,0067	52,99	36	40	1,3125	6,02
5	NWIWGFL	1163,6331	0,0175	44,09	41	42	0,771428571	6,02
5	MHDDQPFYK	1637,8278	-0,0207	53,44	26	41	-1,755555556	5,1

5	FFFVPETKGL	1641,9536	0,0399	44,58	26	40	0,63	6,84
5	NHFIQEF	1162,5974	-0,008	49,11	32	41	-0,514285714	5,12
5	YFNDSQRQ	1285,6254	-0,0103	41,27	29	39	-2,225	6,7
5	YFNDSQRQA	1356,6625	-0,0147	44,41	29	39	-1,777777778	6,7
5	YFNDSQRQAT	1457,7102	-0,0055	42,94	31	39	-1,67	6,7
5	YFNDSQRQATKDA	2001,0321	0,03	60,65	30	42	-1,715384615	6,76
5	YFNDSQRQATKDAGTIA	2343,2225	-0,0231	46,1	24	42	-1,005882353	6,76
5	IGRNFNDPEVQGDMKHFPF	2705,379	-0,0473	44,35	23	42	-0,926315789	5,28
5	NHFIQEF	1162,5974	-0,008	49,11	32	41	-0,514285714	5,12
5	YFNDSQRQ	1285,6254	-0,0103	41,27	29	39	-2,225	6,7
5	YFNDSQRQA	1356,6625	-0,0147	44,41	29	39	-1,777777778	6,7
5	YFNDSQRQAT	1457,7102	-0,0055	42,94	31	39	-1,67	6,7
5	YFNDSQRQATKDA	2001,0321	0,03	60,65	30	42	-1,715384615	6,76
5	YFNDSQRQATKDAGTIA	2343,2225	-0,0231	46,1	24	42	-1,005882353	6,76
5	NWIWGFL	1163,6331	0,0175	44,09	41	42	0,771428571	6,02
5	RLWPNGQDQPS	1525,784	0,0309	51,95	32	42	-1,818181818	6,78
5	TRLWPNGQDQPS	1626,8317	0,015	60,19	27	42	-1,725	6,78
5	MHDDQPFYK	1637,8278	-0,0207	53,44	26	41	-1,755555556	5,1
5	FFFVPETKGL	1641,9536	0,0399	44,58	26	40	0,63	6,84
5	DLFNSYKTV	1543,8652	0,0331	49,84	27	41	-0,322222222	6,6
5	NLGWIPV	1026,6065	0,011	40,32		39	0,871428571	6,02
5	FIPTLLVPI	1240,7998	0,0128	47,74	32	35	2,188888889	6,02
5	MLAGTPFWQ	1278,6634	-0,0191	75,14		41	0,355555556	6,02
5	MLAGTPFWQMA	1480,741	0,0367	47,8	32	42	0,627272727	6,02
5	LKDLGINTV	1429,891	0,0312	50,86		39	0,477777778	6,66
5	KLYQDYSVPV	1668,9492	0,0017	48,32	28	41	-0,37	6,56
5	FGGFVFG	958,5116	0,0136	42,76	35	40	1,628571429	6,02
5	FGGFVFGW	1144,5909	-0,0067	52,99	36	40	1,3125	6,02
5	NWIWGFL	1163,6331	0,0175	44,09	41	42	0,771428571	6,02
5	GMLMVPESPRFL	1604,8621	0,0015	66,91		42	0,5	6,98

5	FFFVPETKGL	1641,9536	0,0399	44,58	26	40	0,63	6,84
5	FFFVPETKGL	1641,9536	0,0399	44,58	26	40	0,63	6,84
5	YFNDSQRQ	1285,6254	-0,0103	41,27	29	39	-2,225	6,7
5	YFNDSQRQA	1356,6625	-0,0147	44,41	29	39	-1,777777778	6,7
5	YFNDSQRQAT	1457,7102	-0,0055	42,94	31	39	-1,67	6,7
5	YFNDSQRQATKDA	2001,0321	0,03	60,65	30	42	-1,715384615	6,76
5	YFNDSQRQATKDAGTIA	2343,2225	-0,0231	46,1	24	42	-1,005882353	6,76
5	FFFVPETKGL	1641,9536	0,0399	44,58	26	40	0,63	6,84
5	SHANQDYYEVWRN	1909,891	-0,0088	80,81	31	39	-1,807692308	5,22
5	NLGWIPV	1026,6065	0,011	40,32		39	0,871428571	6,02
5	AIAPNLFMSI	1304,7365	0,027	43,43	30	42	1,52	6,02
5	LLGGTPLWQI	1325,791	-0,0029	66,57	39	40	0,84	6,02
5	LLGGTPLWQIA	1396,8281	0,0191	57,48		39	0,927272727	6,02
5	FFFVPETKGL	1641,9536	0,0399	44,58	26	40	0,63	6,84
5	YFNDSQRQ	1285,6254	-0,0103	41,27	29	39	-2,225	6,7
5	YFNDSQRQA	1356,6625	-0,0147	44,41	29	39	-1,777777778	6,7
5	YFNDSQRQAT	1457,7102	-0,0055	42,94	31	39	-1,67	6,7
5	YFNDSQRQATKDA	2001,0321	0,03	60,65	30	42	-1,715384615	6,76
5	YFNDSQRQATKDAGTIA	2343,2225	-0,0231	46,1	24	42	-1,005882353	6,76
5	DMFGIVVPR	1261,7056	-0,0137	57,1	25	42	0,844444444	6,78
5	YFNDAQRQ	1269,6305	-0,0047	42,45	24	40	-1,9	6,7
5	GDMFGIVVPR	1318,7271	-0,0169	53,11	24	42	0,72	6,78
5	DIFGIVVPR	1243,7492	-0,0205	53,67	32	41	1,133333333	6,78
5	YFNDAQRQ	1269,6305	-0,0047	42,45	24	40	-1,9	6,7
5	YFNDSQRQ	1285,6254	-0,0103	41,27	29	39	-2,225	6,7
5	YFNDSQRQA	1356,6625	-0,0147	44,41	29	39	-1,777777778	6,7
5	YFNDSQRQAT	1457,7102	-0,0055	42,94	31	39	-1,67	6,7
5	KHSNGDAWVEA	1670,8782	0,0271	99,34		42	-1,081818182	5,22
5	YFNDSQRQATKDA	2001,0321	0,03	60,65	30	42	-1,715384615	6,76
5	SFGSLQTV	1066,5862	-0,0119	70,38	34	41	0,575	6,02

5	DGFNKVQTV	1464,8342	-0,0217	61,63	34	42	-0,477777778	6,66
5	SFGSLQTV	1066,5862	-0,0119	70,38	34	41	0,575	6,02
5	DGFNKVQTV	1464,8342	-0,0217	61,63	34	42	-0,477777778	6,66
5	QEKHRPGDIENN	1894,0062	-0,0343	52,24	26	42	-2,508333333	5,34
5	TADNRHEIA	1254,6519	-0,0226	44,68	27	40	-1,2	5,22
5	HLDKTPSEYI	1659,9238	-0,0321	42,47	36	42	-1,02	5,22
5	ALPWFWEH	1313,676	0,0218	49,93	35	41	-0,2125	5,12
5	YYGYGGTT	1109,5232	0,0136	57,04	30	40	-0,8125	5,82
5	YYQQGPPQQG	1393,6829	-0,0264	60,83	27	39	-2,06	5,86
5	YYQQGPPQQGYPQQQPV	2234,0959	0,0387	51,74	24	42	-1,847058824	5,82
5	SGGNDKQGFPM	1594,8179	-0,0274	61,4	34	41	-1,209090909	6,66
5	FIGFGLMGI	1182,6674	0,0027	59,79		41	2,122222222	6,02
5	APMGGDRPYPPSLPS	1769,8973	-0,0224	43,36	26	41	-0,813333333	6,7
5	LFNQFLL	1122,664	-0,0045	43,2		41	1,428571429	6,02
5	STLFNQFLL	1310,7437	0,0057	58,28		41	0,944444444	6,02
5	ARSEIPEHVI	1378,7771	-0,0008	47,3	29	42	-0,21	5,3
5	DSLPKDLHPM	1609,8903	-0,0306	44,33		42	-0,86	5,1
5	KHFPDYELF	1652,8968	-0,0283	52,12	38	42	-0,844444444	5,22
5	LKHFPDYELF	1765,9809	-0,0333	53,82	29	42	-0,38	5,22
5	QLLDSLPKDLHPM	1964,1171	-0,0395	45,61	28	43	-0,346153846	5,1
5	YFNDSQRQ	1285,6254	-0,0103	41,27	29	39	-2,225	6,7
5	YFNDSQRQA	1356,6625	-0,0147	44,41	29	39	-1,777777778	6,7
5	YFNDSQRQAT	1457,7102	-0,0055	42,94	31	39	-1,67	6,7
5	YFNDSQRQATKDA	2001,0321	0,03	60,65	30	42	-1,715384615	6,76
5	TYKNQGDYNPQ	1784,9099	-0,0314	52,43	26	41	-2,427272727	6,56
5	YFNDAQRQ	1269,6305	-0,0047	42,45	24	40	-1,9	6,7
5	KHSNGDAWVEA	1670,8782	0,0271	99,34		42	-1,081818182	5,22
5	RVAPEEHPV	1261,6982	-0,024	46,56	25	42	-0,855555556	5,3
5	LFLPPVAV	1083,6895	0,0274	44,31		35	2,175	6,02
5	SLFLPPVAV	1170,7215	0,0196	42,02	30	38	1,844444444	6,02

5	LSLFLPPVAV	1283,8056	-0,0172	60,6	30	39	2,04	6,02
5	ALPWFWEH	1313,676	0,0218	49,93	35	41	-0,2125	5,12
5	AGGLILPF	1015,6269	0,0228	48,28	39	40	1,7875	6,02
5	DLQGSEDHSFRKI	1989,0685	-0,0141	81,67	29	42	-1,269230769	5,28
5	LADLQGSEDHSFRKI	2173,1897	-0,007	83,26	29	43	-0,726666667	5,28
5	ALPKHNEVEEHV	1859,0307	0,0228	66,97	29	42	-0,991666667	5,22
5	DLQGSEDHSFRKV	1975,0529	0,0274	45,75	31	42	-1,292307692	5,28
5	LADLQGSEDHSFRKV	2159,1741	-0,0322	93,68	33	42	-0,746666667	5,28
5	FFFVPETKGL	1641,9536	0,0399	44,58	26	40	0,63	6,84
5	KEDWHEKWLN	2071,1459	-0,0228	45,07	34	42	-2,3	5,34
5	NLPEHIVPG	1203,6815	-0,0255	52,18	28	42	-0,144444444	5,12
5	REYLNLPEHIVPG	1764,9725	0,0196	57,94	22	42	-0,523076923	5,3
5	KHEEIDTKNL	1913,119	0,0133	51,3	30	41	-1,74	5,34
5	YFNDAQRQ	1269,6305	-0,0047	42,45	24	40	-1,9	6,7
5	NEADQYLHRVG	1529,7789	0,0277	54,83	33	41	-1,236363636	5,22
5	NLPEHIVPG	1203,6815	-0,0255	52,18	28	42	-0,144444444	5,12
5	REYLNLPEHIVPG	1764,9725	0,0196	57,94	22	42	-0,523076923	5,3
5	KHEEIDTKNL	1913,119	0,0133	51,3	30	41	-1,74	5,34
5	VYAGENFH	1164,5767	0,0205	51,91	37	41	-0,3875	5,12
5	FDFGGFHI	1167,5916	0,0233	51,79	29	41	0,675	4,94
5	VFDFGGFHI	1266,66	0,0263	62,37	34	42	1,066666667	4,94
5	YKYDNFPI	1516,8331	0,0002	43,06	26	42	-0,975	6,56
5	FEDHSWKFY	1715,8713	-0,0077	41,55	28	41	-1,277777778	5,22
5	HDEELKYNGRVV	1916,0521	0,0211	48,16	22	42	-1,258333333	5,34
5	GGHRPQISDEEVSK	1996,0743	-0,0228	48,4	25	42	-1,492857143	5,34
5	ALKYHPDKNPSEEA	2285,2624	0,0461	47,82	41	42	-1,635714286	5,34
5	YYQQQQPQ	1310,6458	-0,0259	50,39	38	39	-2,7125	5,86
5	QQQYQQQPQ	1403,6996	-0,0069	42,34		40	-3,044444444	5,92
5	KAPVPDVNAPQ	1592,9292	0,0166	41,45	19	41	-0,654545455	6,66
5	NTNVDGNIKIV	1643,9612	-0,0331	48,7	36	42	-0,145454545	6,66

5	LSDVFGYVKPG	1638,9387	-0,0328	47,15	28	42	0,281818182	6,6
5	WHFPELA	1127,5967	-0,0184	48,26	33	40	-0,114285714	5,12
5	KEWYGWHFPELA	2020,0612	0,0063	49,9	32	42	-0,9	5,3
5	VLVQDLLHPT	1362,8074	-0,0247	61,08	29	42	0,73	4,94
5	KHQEELWS	1513,8295	-0,008	48,72	30	42	-1,9375	5,3
5	THGNDDDVTKRL	1827,9845	0,0176	61,19	27	42	-1,616666667	5,2
5	GKVDPDQYGHY	1735,8935	0,0243	43,97	25	42	-1,672727273	5,1
5	KTDEHGNLNQF	1759,9259	0,0261	57,79	29	42	-1,736363636	5,22
5	SAVEDRDIHNIGKT	2012,1056	0,015	65,9	24	42	-0,892857143	5,28
5	KYDNIPEHA	1543,84	0,0137	44,5	29	42	-1,577777778	5,22
5	HYGEIKNDFI	1692,9241	0,0197	51,18	27	42	-0,75	5,22
5	NNFLWPF	1165,6123	-0,0102	48,76	37	41	-0,014285714	6,02
5	NNFLWPF	1165,6123	-0,0102	48,76	37	41	-0,014285714	6,02
5	SFNDTFVH	1194,5873	-0,0192	59,34	34	39	-0,2375	4,94
5	YHQKYVDEQ	1666,8721	-0,0157	63,39	29	42	-2,166666667	5,22
5	LGDGKSDNQNHA	1712,8847	-0,0422	61,4	31	41	-1,716666667	5,1
5	FGGTLNPG	990,5338	-0,0103	50,59		41	-0,05	6,02
5	LGDGKSDNQNHA	1712,8847	-0,0422	61,4	31	41	-1,716666667	5,1
5	NKLKDHDYFGEVA	2222,2304	-0,0329	62,92	26	43	-1,084615385	5,28
5	WHFPELA	1127,5967	-0,0184	48,26	33	40	-0,114285714	5,12
5	KEWYGWHFPELA	2020,0612	0,0063	49,9	32	42	-0,9	5,3
5	DKYDENNPEHRKI	2344,2744	0,0175	55,68	37	43	-2,684615385	5,38
5	FGGTLNPG	990,5338	-0,0103	50,59		41	-0,05	6,02
5	SQGFPNIQ	1118,5923	-0,0066	43,6	35	41	-0,75	6,02
5	RVDFKNPHDIIEGI	2110,1941	0,0288	67,87	23	42	-0,507142857	5,28
5	RLDSEKHIDFAPT	1986,094	-0,0467	43,29	26	42	-0,946153846	5,28
5	RLDSEKHIDFAPT	1986,094	-0,0467	43,29	26	42	-0,946153846	5,28
5	IVDAPGHRDF	1354,7197	-0,0048	44,63		41	-0,34	5,1
5	LFNQFLL	1122,664	-0,0045	43,2		41	1,428571429	6,02
5	STLFNQFLL	1310,7437	0,0057	58,28		41	0,944444444	6,02

5	SWINNVAL	1144,6443	-0,0102	44,28	34	41	0,7	6,02
5	VPHENEKQTS	1625,8779	-0,0402	60,02		42	-2	5,3
5	SLFDLKIT	1393,8586	-0,0284	45,99	32	41	0,75	6,66
5	KLDHTENDVRGV	1840,0209	0,015	46,8	24	42	-1,208333333	5,28
5	NQINEAHPT	1251,641	-0,0288	55,26	28	40	-1,466666667	5,12
5	HLDKTPSEYI	1659,9238	-0,0321	42,47	36	42	-1,02	5,22
5	AHTGGDTQKEA	1571,8309	-0,0203	44,08	29	41	-1,472727273	5,22
5	SKDFIQEHVPGPKES	2384,3308	-0,0387	43,01	25	43	-1,246666667	5,34
5	DLFRHEWPS	1414,7196	0,0052	41,23	32	41	-1,266666667	5,22
6	STQHEKET	1416,7614	0,0211	52,74	32	42	-2,475	5,3
6	EILQNRGYLVA	1503,8612	0,0023	49,77		42	0,127272727	6,88
6	YDNAPYSPKPV	1707,9238	0,0316	47,73	26	42	-1,190909091	6,56
6	AYDNAPYSPKPV	1778,9609	0,0229	54,72	24	42	-0,941666667	6,56
6	GDGVNDAPSLKKA	1958,1405	0,0435	52,69	36	40	-0,761538462	6,72
6	FLAPGLSAIIDALKT	1987,2123	-0,0058	103,8	36	40	1,18	6,66
6	MTGDGVNDAPSLKKA	2190,2287	0,0265	85,74	32	42	-0,58	6,72
6	EILQNRGYLVA	1503,8612	0,0023	49,77		42	0,127272727	6,88
6	GDGVNDAPSLKKA	1958,1405	0,0435	52,69	36	40	-0,761538462	6,72
6	FLAPGLSAIIDALKT	1987,2123	-0,0058	103,8	36	40	1,18	6,66
6	MTGDGVNDAPSLKKA	2190,2287	0,0265	85,74	32	42	-0,58	6,72
6	GGIGTVPV	927,5593	0,0079	45,9	32	39	1,175	6,02
6	ETGVIKPG	1257,7698	0,0366	39,72	37	39	-0,225	6,84
6	FSEYPPLGRFA	1511,7975	-0,0053	51,51	28	42	-0,227272727	6,88
6	RFQEIVKET	1606,9448	0,0384	59,11		40	-0,9	7
6	EQPSRPTDKPL	1724,9827	0,0227	47,62	34	41	-1,945454545	6,92
6	DAIEQPSRPTDKPLRLPL	2503,4528	-0,0275	48,6	25	43	-0,95	6,98
6	GQDQPSSKGA	1431,7724	0,0049	62,31	33	42	-1,66	6,66
6	RLWPNGQDQPS	1525,784	0,0414	50,3	26	42	-1,818181818	6,78
6	GQDQPSSKGA	1431,7724	0,0049	62,31	33	42	-1,66	6,66

6	RLWPNGQDQPS	1525,784	0,0414	50,3	26	42	-1,818181818	6,78
6	TGGQAFPQ	1033,5396	0,0155	51,42		42	-0,6875	6,02
6	GQDQPSSKGA	1431,7724	0,0049	62,31	33	42	-1,66	6,66
6	RLWPNGQDQPS	1525,784	0,0414	50,3	26	42	-1,818181818	6,78
6	SWGELFTGKPA	1649,9183	-0,0138	45,94	28	42	-0,345454545	6,84
6	YFNDSQRQ	1285,6254	0,0201	43,21	27	40	-2,225	6,7
6	YFNDSQRQA	1356,6625	0,0243	43,94	30	40	-1,777777778	6,7
6	YFNDSQRQATKDA	2001,0321	0,0216	61,05	27	42	-1,715384615	6,76
6	YFNDSQRQ	1285,6254	0,0201	43,21	27	40	-2,225	6,7
6	YFNDSQRQA	1356,6625	0,0243	43,94	30	40	-1,777777778	6,7
6	YFNDSQRQATKDA	2001,0321	0,0216	61,05	27	42	-1,715384615	6,76
6	GQDQPSSKGA	1431,7724	0,0049	62,31	33	42	-1,66	6,66
6	RLWPNGQDQPS	1525,784	0,0414	50,3	26	42	-1,818181818	6,78
6	SWGELFTGKPA	1649,9183	-0,0138	45,94	28	42	-0,345454545	6,84
6	LKDLGINTV	1429,891	-0,0078	44,43	38	40	0,477777778	6,66
6	KLYQDYSVPV	1668,9492	-0,0116	68,98	33	42	-0,37	6,56
6	QNPPPPPSTT	1263,6662	-0,0134	54,29		41	-1,72	6,02
6	YFNDSQRQ	1285,6254	0,0201	43,21	27	40	-2,225	6,7
6	YFNDSQRQA	1356,6625	0,0243	43,94	30	40	-1,777777778	6,7
6	YFNDSQRQATKDA	2001,0321	0,0216	61,05	27	42	-1,715384615	6,76
6	YFNDSQRQ	1285,6254	0,0201	43,21	27	40	-2,225	6,7
6	YFNDSQRQA	1356,6625	0,0243	43,94	30	40	-1,777777778	6,7
6	YFNDSQRQATKDA	2001,0321	0,0216	61,05	27	42	-1,715384615	6,76
6	YFNDAQRQ	1269,6305	0,0247	49,21	31	41	-1,9	6,7
6	NEQGNRVTPS	1329,684	-0,0013	45,58		41	-1,78	6,98
6	ANEQGNRVTPS	1400,7211	0,0469	55,19	38	41	-1,454545455	6,98
6	YFNDAQRQ	1269,6305	0,0247	49,21	31	41	-1,9	6,7
6	NEQGNRVTPS	1329,684	-0,0013	45,58		41	-1,78	6,98
6	ANEQGNRVTPS	1400,7211	0,0469	55,19	38	41	-1,454545455	6,98
6	YFNDSQRQ	1285,6254	0,0201	43,21	27	40	-2,225	6,7

6	YFNDSQRQA	1356,6625	0,0243	43,94	30	40	-1,777777778	6,7
6	YFNDSQRQATKDA	2001,0321	0,0216	61,05	27	42	-1,715384615	6,76
6	KKITGDLNMQELII	2302,3902	-0,0213	42,12	32	41	0,007142857	6,84
6	KKITGDLNMQELII	2302,3902	-0,0213	42,12	32	41	0,007142857	6,84
6	ALPWFWEH	1313,676	0,002	51,12	31	41	-0,2125	5,12
6	SGPQSYQ	994,4923	0,0237	53,1	27	41	-1,7	5,92
6	YQQGPPQQG	1230,6196	0,0001	45,43	27	40	-2,144444444	5,92
6	YYQQGPPQQG	1393,6829	-0,0023	67,71	39	40	-2,06	5,86
6	NQQGYNQQ	1207,5784	-0,0166	54,95	33	38	-2,8375	5,92
6	QQYYQQQ	1213,593	-0,01	50,45	37	39	-2,871428571	5,86
6	QQYYQQQQ	1341,6516	-0,0131	55,27		39	-2,95	5,86
6	NQQGYNQQGY	1427,6632	-0,0113	58,32	34	37	-2,44	5,86
6	GYNQQGYNQQ	1427,6632	-0,0063	69,45	29	38	-2,44	5,86
6	NQQGYNQQGYNQQ	1797,8233	0,0169	103,6	35	39	-2,684615385	5,86
6	FIGFGLMGI	1182,6674	0,0205	56,84	36	40	2,122222222	6,02
6	AFADMFDTNVRGK	1929,0184	-0,0076	46,07	29	42	-0,361538462	6,8
6	KQNSGNAET	1405,7567	0,0021	53,79	35	42	-2	6,84
6	YFNDSQRQ	1285,6254	0,0201	43,21	27	40	-2,225	6,7
6	YFNDSQRQA	1356,6625	0,0243	43,94	30	40	-1,777777778	6,7
6	YFNDSQRQATKDA	2001,0321	0,0216	61,05	27	42	-1,715384615	6,76
6	YFNDAQRQ	1269,6305	0,0247	49,21	31	41	-1,9	6,7
6	QLQQQQQQQ	1384,7262	-0,0207	86,35		41	-2,688888889	6,02
6	QQQQQQQQYA	1505,7426	-0,0151	86,44		40	-2,75	5,92
6	QLQQQQQQQQ	1512,7848	-0,0308	77,56		41	-2,77	6,02
6	LFLPPVAV	1083,6895	0,0119	46,72		36	2,175	6,02
6	LSLFLPPVAV	1283,8056	0,0223	45,2	21	36	2,04	6,02
6	ALPWFWEH	1313,676	0,002	51,12	31	41	-0,2125	5,12
6	KYSNEDTRPV	1665,9092	0,0108	51,05	31	42	-1,91	6,86
6	AEIQGKT	1203,7229	0,011	44,27		41	-0,814285714	6,84
6	SSSQEQLRQT	1391,7207	0,0263	72,91	38	42	-1,78	6,98

6	YFNDAQRQ	1269,6305	0,0247	49,21	31	41	-1,9	6,7
6	AATENQPKGI	1485,8557	0,0137	52,85	34	41	-0,9	6,84
6	QLQQQQQQ	1384,7262	-0,0207	86,35		41	-2,688888889	6,02
6	QLQQQQQQQ	1512,7848	-0,0308	77,56		41	-2,77	6,02
6	QIQQQQQQ	1384,7262	-0,0207	86,35		41	-2,611111111	6,02
6	QIQQQQQQQ	1512,7848	-0,0308	77,56		41	-2,7	6,02
6	LQQQQQQQ	1384,7262	-0,0207	66,37		41	-2,688888889	6,02
6	QQQQQQQQYA	1505,7426	-0,0151	86,44		40	-2,75	5,92
6	LQQQQQQQQ	1512,7848	-0,0308	61,3		41	-2,77	6,02
6	QQQYQQQPQ	1403,6996	-0,0012	63		41	-3,044444444	5,92
6	KAPVPDVNAPQ	1592,9292	0,0391	41,1	24	40	-0,654545455	6,66
6	QIQQQQQQ	1384,7262	-0,0207	86,35		41	-2,611111111	6,02
6	QIQQQQQQQ	1512,7848	-0,0308	77,56		41	-2,7	6,02
6	QLQQQQQQ	1384,7262	-0,0207	86,35		41	-2,688888889	6,02
6	QLQQQQQQQ	1512,7848	-0,0308	77,56		41	-2,77	6,02
6	QIQQQQQQ	1384,7262	-0,0207	86,35		41	-2,611111111	6,02
6	QIQQQQQQQ	1512,7848	-0,0308	77,56		41	-2,7	6,02
6	QLQQQQQQ	1384,7262	-0,0207	86,35		41	-2,688888889	6,02
6	QLQQQQQQQ	1512,7848	-0,0308	77,56		41	-2,77	6,02
6	QLQQQQQQ	1384,7262	-0,0207	86,35		41	-2,688888889	6,02
6	QLQQQQQQQ	1512,7848	-0,0308	77,56		41	-2,77	6,02
6	QLQQQQQQ	1384,7262	-0,0207	86,35		41	-2,688888889	6,02
6	QLQQQQQQQ	1512,7848	-0,0308	77,56		41	-2,77	6,02
6	QLQQQQQQ	1384,7262	-0,0207	86,35		41	-2,688888889	6,02
6	QLQQQQQQQ	1512,7848	-0,0308	77,56		41	-2,77	6,02
6	ATDSGKNQA	1348,7352	0,0378	51,84	34	42	-1,411111111	6,66
6	LQQQQQQQ	1384,7262	-0,0207	66,37		41	-2,688888889	6,02
6	LQQQQQQQQ	1512,7848	-0,0308	61,3		41	-2,77	6,02
6	SKQSAEPQ	1331,7451	-0,0021	42,91	34	42	-1,975	6,84
6	NQNNNNEGNTKYS	1953,9546	-0,024	74,43	34	40	-2,7	6,8

6	LQQQQQQQ	1384,7262	-0,0207	66,37		41	-2,688888889	6,02
6	LQQQQQQQQ	1512,7848	-0,0308	61,3		41	-2,77	6,02
6	PSGQQQQQQ	1384,6898	0,0157	66,37		41	-2,73	6,02
6	PSGQQQQQQQ	1512,7484	0,0056	61,3		41	-2,8	6,02
6	LQQQQQQQ	1384,7262	-0,0207	66,37		41	-2,688888889	6,02
6	LQQQQQQQQ	1512,7848	-0,0308	61,3		41	-2,77	6,02
6	LQQQQQQQ	1384,7262	-0,0207	66,37		41	-2,688888889	6,02
6	LQQQQQQQQ	1512,7848	-0,0308	61,3		41	-2,77	6,02
6	LQQQQQQQ	1384,7262	-0,0207	66,37		41	-2,688888889	6,02
6	LQQQQQQQQ	1512,7848	-0,0308	61,3		41	-2,77	6,02
6	LQQQQQQQ	1384,7262	-0,0207	66,37		41	-2,688888889	6,02
6	LQQQQQQQQ	1512,7848	-0,0308	61,3		41	-2,77	6,02
6	LQQQQQQQ	1384,7262	-0,0207	66,37		41	-2,688888889	6,02
6	LQQQQQQQQ	1512,7848	-0,0308	61,3		41	-2,77	6,02
6	LQQQQQQQ	1384,7262	-0,0207	66,37		41	-2,688888889	6,02
6	LQQQQQQQQ	1512,7848	-0,0308	61,3		41	-2,77	6,02
6	LQQQQQQQ	1384,7262	-0,0207	66,37		41	-2,688888889	6,02
6	LQQQQQQQQ	1512,7848	-0,0308	61,3		41	-2,77	6,02
6	QLQQQQQQ	1384,7262	-0,0207	86,35		41	-2,688888889	6,02
6	LQQQQQQQ	1384,7262	-0,0207	66,37		41	-2,688888889	6,02
6	YNNNGGQPQGQ	1404,6585	0,0264	45,84	26	40	-2,281818182	5,92
6	QLQQGVMPQQ	1384,7336	-0,0281	49,53		41	-0,96	6,02
6	QLQQGVMPQQQ	1512,7922	-0,0382	46,64		41	-1,190909091	6,02
6	ELIQNRLNME	1503,7918	0,0717	43,75		42	-0,8	4,26
6	QQGYPPQQEH	1439,6996	0,0227	46,29	21	41	-2,56	5,12
6	VYLSGDV	980,5382	0,0209	42	39	41	0,885714286	3,1
6	LADALL	843,5268	0,0269	40,71		38	1,916666667	3,1
6	LADAIL	843,5268	0,0269	40,71		38	2,033333333	3,1
7	TIIGVPV	926,6004	-0,0176	38,04		36	2,1	6,02

7	EILQNRG	1057,6083	0,0048	42,19	38	40	-1,014285714	6,98
7	NIYNAER	1107,5875	0,0002	46,78		42	-1,428571429	6,88
7	QNFGAMNGI	1179,5909	-0,0182	56,96	37	40	-0,033333333	6,02
7	EILQNRGYL	1333,7557	-0,0017	50,44		42	-0,511111111	6,88
7	QPTQEKPA	1355,7814	-0,0152	46,63	35	41	-2,0625	6,84
7	EILQNRGYLVA	1503,8612	0,0053	52,65		42	0,127272727	6,88
7	DNAPYSPKPV	1544,8604	0,0162	44,93	31	42	-1,18	6,6
7	YDNAPYSPKPV	1707,9238	-0,0063	48,26	26	42	-1,190909091	6,56
7	SIVDELKKT	1719,0751	0,0024	48,1		39	-0,422222222	6,84
7	AYDNAPYSPKPV	1778,9609	-0,0095	50,13	26	42	-0,941666667	6,56
7	IAYDNAPYSPKPV	1892,0449	-0,0304	59,25	36	43	-0,523076923	6,56
7	GDGVNDAPSLKKA	1958,1405	0,0138	51,08	34	41	-0,761538462	6,72
7	AIAYDNAPYSPKPV	1963,082	-0,0346	45,47	28	42	-0,357142857	6,56
7	FLAPGLSAIIDALKT	1987,2123	0,0272	99,97	37	39	1,18	6,66
7	MTGDGVNDAPSLKKA	2190,2287	-0,0185	92,97	34	43	-0,58	6,72
7	MTGDGVNDAPSLKKA	2206,2236	-0,0247	44,73	25	43	-0,58	6,72
7	TIIGVPV	926,6004	-0,0176	38,04		36	2,1	6,02
7	EILQNRG	1057,6083	0,0048	42,19	38	40	-1,014285714	6,98
7	NIYNAER	1107,5875	0,0002	46,78		42	-1,428571429	6,88
7	EILQNRGYL	1333,7557	-0,0017	50,44		42	-0,511111111	6,88
7	EILQNRGYLVA	1503,8612	0,0053	52,65		42	0,127272727	6,88
7	SIVDELKKT	1719,0751	0,0024	48,1		39	-0,422222222	6,84
7	GDGVNDAPSLKKA	1958,1405	0,0138	51,08	34	41	-0,761538462	6,72
7	FLAPGLSAIIDALKT	1987,2123	0,0272	99,97	37	39	1,18	6,66
7	MTGDGVNDAPSLKKA	2190,2287	-0,0185	92,97	34	43	-0,58	6,72
7	MTGDGVNDAPSLKKA	2206,2236	-0,0247	44,73	25	43	-0,58	6,72
7	ASFNAT	838,4388	0,0028	41,06	40	41	0,233333333	6,02
7	GGIGTVPV	927,5593	-0,0143	45,37	28	39	1,175	6,02
7	KTLLEAI	1244,8109	-0,0174	41,23		39	0,828571429	6,84
7	ETGVIKPG	1257,7698	0,0077	46,34		39	-0,225	6,84

7	EYPPLGRFA	1277,6971	0,0141	43,64	30	42	-0,5	6,88
7	LRLPLQDVY	1344,7968	-0,0168	43,11	32	42	0,133333333	6,7
7	SEYPPLGRFA	1364,7291	0,0042	43,44	37	42	-0,53	6,88
7	FSEYPPLGRF	1440,7604	-0,0142	46,29	36	41	-0,43	6,88
7	FSEYPPLGRFA	1511,7975	-0,0099	68,52	32	42	-0,227272727	6,88
7	RFQEIVKET	1606,9448	0,0133	57,21		41	-0,9	7
7	RFQEIVKETS	1693,9769	-0,0145	52,64	29	42	-0,89	7
7	NAPWYKGWEKET	2195,1983	0,018	52,88	28	43	-1,858333333	6,88
7	TNAPWYKGWEKET	2296,246	0,0315	63,5	33	43	-1,769230769	6,88
7	FGGFVFG	958,5116	0,0066	50,03		40	1,628571429	6,02
7	FGGFVFGW	1144,5909	-0,0002	43,12	35	40	1,3125	6,02
7	RLWPNGQDQPS	1525,784	0,0233	46,46	28	42	-1,818181818	6,78
7	FGGFVFG	958,5116	0,0066	50,03		40	1,628571429	6,02
7	FGGFVFGW	1144,5909	-0,0002	43,12	35	40	1,3125	6,02
7	RLWPNGQDQPS	1525,784	0,0233	46,46	28	42	-1,818181818	6,78
7	TGGQAFPQ	1033,5396	-0,0006	51,58		42	-0,6875	6,02
7	IENGIINPR	1253,7294	-0,0067	49,51	35	40	-0,388888889	6,98
7	RFTDTRKDEQERGI	2208,2017	0,0203	43,59	22	43	-2,1	7,08
7	FGGFVFG	958,5116	0,0066	50,03		40	1,628571429	6,02
7	FGGFVFGW	1144,5909	-0,0002	43,12	35	40	1,3125	6,02
7	RLWPNGQDQPS	1525,784	0,0233	46,46	28	42	-1,818181818	6,78
7	FGGFVFG	958,5116	0,0066	50,03		40	1,628571429	6,02
7	FGGFVFGW	1144,5909	-0,0002	43,12	35	40	1,3125	6,02
7	SWGELFTGKPA	1649,9183	-0,0253	51,99	30	42	-0,345454545	6,84
7	FGGFVFG	958,5116	0,0066	50,03		40	1,628571429	6,02
7	FGGFVFGW	1144,5909	-0,0002	43,12	35	40	1,3125	6,02
7	YFNDSQRQ	1285,6254	0,0039	45,33	25	39	-2,225	6,7
7	YFNDSQRQA	1356,6625	0,0047	45,03	27	40	-1,777777778	6,7
7	YFNDSQRQAT	1457,7102	-0,0032	59,83	32	40	-1,67	6,7
7	VTDYFNGKEPNRS	1984,042	-0,0132	47,45	36	42	-1,553846154	6,86

7	YFNDSQRQ	1285,6254	0,0039	45,33	25	39	-2,225	6,7
7	YFNDSQRQA	1356,6625	0,0047	45,03	27	40	-1,777777778	6,7
7	YFNDSQRQAT	1457,7102	-0,0032	59,83	32	40	-1,67	6,7
7	VTDYFNGKEPNRS	1984,042	-0,0132	47,45	36	42	-1,553846154	6,86
7	RLWPNGQDQPS	1525,784	0,0233	46,46	28	42	-1,818181818	6,78
7	SWGELFTGKPA	1649,9183	-0,0253	51,99	30	42	-0,345454545	6,84
7	MNLYYEKS	1504,8001	-0,0204	51,51	32	42	-1,075	6,76
7	SNLGWIPV	1113,6385	-0,0152	40,33	37	40	0,6625	6,02
7	FIPTLLVPI	1240,7998	0,0021	51,69	32	36	2,188888889	6,02
7	MLAGTPFWQ	1278,6634	-0,0053	61,73	34	41	0,355555556	6,02
7	LKDLGINTV	1429,891	0,0058	46,41		39	0,477777778	6,66
7	FGGFVFG	958,5116	0,0066	50,03		40	1,628571429	6,02
7	FGGFVFGW	1144,5909	-0,0002	43,12	35	40	1,3125	6,02
7	SWGELFSNKG	1581,8557	-0,0144	45,69	31	42	-0,76	6,84
7	QLQNPPPPPSTT	1504,8089	-0,0214	48,12	33	42	-1,408333333	6,02
7	YFNDSQRQ	1285,6254	0,0039	45,33	25	39	-2,225	6,7
7	YFNDSQRQA	1356,6625	0,0047	45,03	27	40	-1,777777778	6,7
7	YFNDSQRQAT	1457,7102	-0,0032	59,83	32	40	-1,67	6,7
7	KGVDRESYA	1481,8244	-0,0047	47,4	31	42	-1,322222222	6,86
7	AGEGLNVEQRK	1657,9517	0,0073	61,13	35	42	-1,218181818	7
7	SNLGWIPV	1113,6385	-0,0152	40,33	37	40	0,6625	6,02
7	LLGGTPLWQIA	1396,8281	-0,0032	40,14	30	40	0,927272727	6,02
7	YFNDSQRQ	1285,6254	0,0039	45,33	25	39	-2,225	6,7
7	YFNDSQRQA	1356,6625	0,0047	45,03	27	40	-1,777777778	6,7
7	YFNDSQRQAT	1457,7102	-0,0032	59,83	32	40	-1,67	6,7
7	YFNDAQRQ	1269,6305	0,0019	53,36	28	40	-1,9	6,7
7	YFNDAQRQA	1340,6676	0,0135	45,05	25	41	-1,488888889	6,7
7	ANEQGNRVTPS	1400,7211	0,002	49,83	32	41	-1,454545455	6,98
7	YFNDAQRQATKDA	1985,0372	0,028	43,63	27	42	-1,515384615	6,76
7	YFNDAQRQ	1269,6305	0,0019	53,36	28	40	-1,9	6,7

7	YFNDAQRQA	1340,6676	0,0135	45,05	25	41	-1,488888889	6,7
7	ANEQGNRVTPS	1400,7211	0,002	49,83	32	41	-1,454545455	6,98
7	YFNDAQRQATKDA	1985,0372	0,028	43,63	27	42	-1,515384615	6,76
7	YFNDSQRQ	1285,6254	0,0039	45,33	25	39	-2,225	6,7
7	YFNDSQRQA	1356,6625	0,0047	45,03	27	40	-1,777777778	6,7
7	YFNDSQRQAT	1457,7102	-0,0032	59,83	32	40	-1,67	6,7
7	SFGSLQ	866,4701	0,0108	42,6		40	0,183333333	6,02
7	KKITGDLNMQELII	2302,3902	-0,0176	59,89	31	41	0,007142857	6,84
7	SFGSLQ	866,4701	0,0108	42,6		40	0,183333333	6,02
7	KKITGDLNMQELII	2302,3902	-0,0176	59,89	31	41	0,007142857	6,84
7	AGEGLNVEQRK	1657,9517	0,0073	61,13	35	42	-1,218181818	7
7	KVLEELFQKL	1933,222	-0,0209	71,09	32	39	0,01	6,92
7	YQQGPPQQG	1230,6196	-0,0091	54,94	26	40	-2,144444444	5,92
7	YYQQGPPQQG	1393,6829	-0,0196	65,74	27	40	-2,06	5,86
7	NQQGYNQQGY	1427,6632	-0,0197	37,51	28	37	-2,44	5,86
7	QERGYPPQQQ	1458,7418	-0,0036	50,94	28	41	-2,69	6,88
7	DAFKNYPNVL	1637,9183	0,0072	52,84	31	42	-0,47	6,6
7	YFNDSQRQ	1285,6254	0,0039	45,33	25	39	-2,225	6,7
7	YFNDSQRQA	1356,6625	0,0047	45,03	27	40	-1,777777778	6,7
7	YFNDSQRQAT	1457,7102	-0,0032	59,83	32	40	-1,67	6,7
7	YKNQGDYNPQ	1683,8622	-0,0155	56,73	30	41	-2,6	6,56
7	ANEQELKKG	1703,0186	-0,0049	54,3	35	41	-1,844444444	6,92
7	TYKNQGDYNPQ	1784,9099	-0,017	71,93	28	41	-2,427272727	6,56
7	YFNDAQRQ	1269,6305	0,0019	53,36	28	40	-1,9	6,7
7	YFNDAQRQA	1340,6676	0,0135	45,05	25	41	-1,488888889	6,7
7	YFNDAQRQATKDA	1985,0372	0,028	43,63	27	42	-1,515384615	6,76
7	LFLPPVAV	1083,6895	-0,0022	36,41	31	36	2,175	6,02
7	SLFLPPVAV	1170,7215	0,0082	41,13	36	38	1,844444444	6,02
7	LSLFLPPVAV	1283,8056	-0,0152	56,16	28	39	2,04	6,02
7	KMNAPQKPEDIPVAT	2325,3335	-0,0314	44,88	32	42	-0,873333333	6,84

7	SKLIPEVI	1355,8794	-0,0242	41,88		38	0,9	6,84
7	GKQLEDGRT	1460,8353	-0,0168	46,5	34	42	-1,844444444	6,92
7	FAGKQLEDGRT	1678,9408	0,015	68,74		42	-1,090909091	6,92
7	FAGKQLEDGRTL	1792,0249	-0,0175	61,38	41	42	-0,683333333	6,92
7	SKLIPEVI	1355,8794	-0,0242	41,88		38	0,9	6,84
7	VFIWKDV	1363,827	0,0087	42,33	37	40	1,057142857	6,66
7	YFNDAQRQ	1269,6305	0,0019	53,36	28	40	-1,9	6,7
7	YFNDAQRQA	1340,6676	0,0135	45,05	25	41	-1,488888889	6,7
7	YFNDAQRQATKDA	1985,0372	0,028	43,63	27	42	-1,515384615	6,76
7	AATENQPKGI	1485,8557	0,0017	51,21	39	41	-0,9	6,84
7	VYAGENFHHGDKL	1944,0259	-0,0229	62,19	25	42	-0,792307692	6,02
7	YKYDNFPI	1516,8331	-0,0044	46,3	28	42	-0,975	6,56
7	TATDVEIKKA	1762,0809	0,0059	54,32	31	40	-0,39	6,84
7	KQATDEGVKT	1763,0397	-0,0203	48,87	33	42	-1,41	6,84
7	KQLENVS	1274,76	-0,0046	48,59		42	-1,028571429	6,84
7	DNKQNALEKFI	2006,1769	-0,0227	47,77	30	42	-1,127272727	6,84
7	FGGTLNPG	990,5338	-0,0072	42,05	39	42	-0,05	6,02
7	KFAEPEAKDKS	2165,2867	0,0172	41,28	28	41	-1,654545455	6,9
7	FGGTLNPG	990,5338	-0,0072	42,05	39	42	-0,05	6,02
7	NVSKDELREKL	2017,214	-0,0227	44,02	28	41	-1,390909091	6,96
7	VRLILPGELA	1308,8332	-0,0153	49,84	34	39	1,19	6,98
7	SWINNVAL	1144,6443	-0,0141	41,77	40	41	0,7	6,02
7	KTLLEAI	1244,8109	-0,0174	41,23		39	0,828571429	6,84
7	KTLLEAI	1244,8109	-0,0174	41,23		39	0,828571429	6,84
7	LQTQNELLT	1287,7237	0,0541	46,18		41	-0,444444444	3,3
7	SIGKIIEIL	1442,9478	0,001	44,95		36	1,466666667	6,84
7	SFGSLQ	866,4701	0,0108	42,6		40	0,183333333	6,02
7	LADALL	843,5268	0,0079	42,22		39	1,916666667	3,1
7	LADAIL	843,5268	0,0079	42,22		39	2,033333333	3,1
7	TRLILLDAI	1255,8066	-0,0065	43,58		37	1,5	6,78

7	KIPGLNPGDEPRAL	1934,1355	-0,0027	42,22	34	42	-0,757142857	6,92
8	EILQNRG	1057,6083	0,0093	44,42	38	40	-1,014285714	6,98
8	NIYNAER	1107,5875	0	51,72		42	-1,428571429	6,88
8	QNFGAMNGI	1179,5909	-0,0118	49,87	31	40	-0,033333333	6,02
8	QNFGAMNGI	1195,5858	-0,0183	45,17	37	39	-0,033333333	6,02
8	NIYNAERL	1220,6716	0,0039	42		42	-0,775	6,88
8	EILQNRGYL	1333,7557	0,0093	46,88		42	-0,511111111	6,88
8	EILQNRGYLVA	1503,8612	0,012	42,43	41	42	0,127272727	6,88
8	YDNAPYSPKPV	1707,9238	-0,0095	47,28	26	42	-1,190909091	6,56
8	SIVDELKKT	1719,0751	0,0073	48		39	-0,422222222	6,84
8	AYDNAPYSPKPV	1778,9609	-0,0131	58,81	33	42	-0,941666667	6,56
8	MTGDGVNDAPSLKKA	2190,2287	-0,0315	62,09	30	43	-0,58	6,72
8	MTGDGVNDAPSLKKA	2206,2236	-0,0362	49,92	25	43	-0,58	6,72
8	DGFAEVFPQHKYRVV	2249,2363	0,0168	59,17	26	43	-0,36	7,66
8	EILQNRG	1057,6083	0,0093	44,42	38	40	-1,014285714	6,98
8	NIYNAER	1107,5875	0	51,72		42	-1,428571429	6,88
8	NIYNAERL	1220,6716	0,0039	42		42	-0,775	6,88
8	EILQNRGYL	1333,7557	0,0093	46,88		42	-0,511111111	6,88
8	EILQNRGYLVA	1503,8612	0,012	42,43	41	42	0,127272727	6,88
8	SIVDELKKT	1719,0751	0,0073	48		39	-0,422222222	6,84
8	MTGDGVNDAPSLKKA	2190,2287	-0,0315	62,09	30	43	-0,58	6,72
8	MTGDGVNDAPSLKKA	2206,2236	-0,0362	49,92	25	43	-0,58	6,72
8	DGFAEVFPQHKYRVV	2249,2363	0,0168	59,17	26	43	-0,36	7,66
8	KTLLEAI	1244,8109	-0,0113	44,69		39	0,828571429	6,84
8	ETGVIKPG	1257,7698	0,0102	50,46	37	40	-0,225	6,84
8	EYPPLGRFA	1277,6971	0,0247	43,59	33	42	-0,5	6,88
8	FSEYPPLGRFA	1511,7975	-0,0094	58,99	30	42	-0,227272727	6,88
8	RFQEIVKET	1606,9448	-0,0038	54,11		42	-0,9	7
8	EQPSRPTDKPL	1724,9827	0,014	45,99	38	42	-1,945454545	6,92

8	NAPWYKGWEKET	2195,1983	0,011	58,88	28	43	-1,858333333	6,88
8	TNAPWYKGWEKET	2296,246	0,049	57,26	29	42	-1,769230769	6,88
8	FGGFVFG	958,5116	0,0146	40,77	31	40	1,628571429	6,02
8	FGGFVFGW	1144,5909	0,0014	46,8	38	40	1,3125	6,02
8	RLWPNGQDQPS	1525,784	0,0134	76,28	30	42	-1,818181818	6,78
8	FGGFVFG	958,5116	0,0146	40,77	31	40	1,628571429	6,02
8	FGGFVFGW	1144,5909	0,0014	46,8	38	40	1,3125	6,02
8	RLWPNGQDQPS	1525,784	0,0134	76,28	30	42	-1,818181818	6,78
8	TGGQAFPQ	1033,5396	0,0027	45,43		42	-0,6875	6,02
8	QGPNYVPG	1059,5552	-0,007	43,21	26	40	-1,0125	5,92
8	FGGFVFG	958,5116	0,0146	40,77	31	40	1,628571429	6,02
8	FGGFVFGW	1144,5909	0,0014	46,8	38	40	1,3125	6,02
8	RLWPNGQDQPS	1525,784	0,0134	76,28	30	42	-1,818181818	6,78
8	FGGFVFG	958,5116	0,0146	40,77	31	40	1,628571429	6,02
8	FGGFVFGW	1144,5909	0,0014	46,8	38	40	1,3125	6,02
8	FGGFVFG	958,5116	0,0146	40,77	31	40	1,628571429	6,02
8	FGGFVFGW	1144,5909	0,0014	46,8	38	40	1,3125	6,02
8	YFNDSQRQA	1356,6625	0,0099	51,83	35	40	-1,777777778	6,7
8	QATKDAGTIA	1432,8291	0,0122	43,98	39	42	-0,28	6,66
8	YFNDSQRQAT	1457,7102	-0,0001	65,4	33	40	-1,67	6,7
8	YFNDSQRQA	1356,6625	0,0099	51,83	35	40	-1,777777778	6,7
8	QATKDAGTIA	1432,8291	0,0122	43,98	39	42	-0,28	6,66
8	YFNDSQRQAT	1457,7102	-0,0001	65,4	33	40	-1,67	6,7
8	RLWPNGQDQPS	1525,784	0,0134	76,28	30	42	-1,818181818	6,78
8	DLFNSYKT	1444,7968	0,0089	44,39	28	42	-0,8875	6,6
8	MLAGTPFWQ	1278,6634	0,0062	55,34	38	42	0,355555556	6,02
8	FGGFVFG	958,5116	0,0146	40,77	31	40	1,628571429	6,02
8	FGGFVFGW	1144,5909	0,0014	46,8	38	40	1,3125	6,02
8	QLQNPPPPPSTT	1504,8089	-0,0211	49,47	34	42	-1,408333333	6,02
8	YFNDSQRQA	1356,6625	0,0099	51,83	35	40	-1,777777778	6,7

8	QATKDAGTIA	1432,8291	0,0122	43,98	39	42	-0,28	6,66
8	YFNDSQRQAT	1457,7102	-0,0001	65,4	33	40	-1,67	6,7
8	YFNDSQRQA	1356,6625	0,0099	51,83	35	40	-1,777777778	6,7
8	QATKDAGTIA	1432,8291	0,0122	43,98	39	42	-0,28	6,66
8	YFNDSQRQAT	1457,7102	-0,0001	65,4	33	40	-1,67	6,7
8	YFNDAQRQ	1269,6305	0,0086	40,46	22	40	-1,9	6,7
8	ANEQGNRVTPS	1400,7211	-0,0109	54,99	32	41	-1,454545455	6,98
8	YFNDAQRQ	1269,6305	0,0086	40,46	22	40	-1,9	6,7
8	ANEQGNRVTPS	1400,7211	-0,0109	54,99	32	41	-1,454545455	6,98
8	YFNDSQRQA	1356,6625	0,0099	51,83	35	40	-1,777777778	6,7
8	YFNDSQRQAT	1457,7102	-0,0001	65,4	33	40	-1,67	6,7
8	SFGSLQTV	1066,5862	-0,0049	58,69		41	0,575	6,02
8	SFGSLQTV	1066,5862	-0,0049	58,69		41	0,575	6,02
8	YYGYGGTT	1109,5232	0,0123	53,3	25	39	-0,8125	5,82
8	YYQQGPPQQG	1393,6829	-0,0208	63,29	37	40	-2,06	5,86
8	NQQGYNQQGY	1427,6632	-0,0266	70,16	33	36	-2,44	5,86
8	LFLKYGESI	1526,9114	0,0007	48,36	33	40	0,555555556	6,8
8	YFNDSQRQA	1356,6625	0,0099	51,83	35	40	-1,777777778	6,7
8	YFNDSQRQAT	1457,7102	-0,0001	65,4	33	40	-1,67	6,7
8	ANEQELKKG	1703,0186	0,0001	45,86	33	41	-1,844444444	6,92
8	YFNDAQRQ	1269,6305	0,0086	40,46	22	40	-1,9	6,7
8	LFLPPVAV	1083,6895	0,0016	38,59	35	36	2,175	6,02
8	SLFLPPVAV	1170,7215	0,009	41,37	31	38	1,844444444	6,02
8	LSLFLPPVAV	1283,8056	-0,0183	47,44	29	39	2,04	6,02
8	YFNDAQRQ	1269,6305	0,0086	40,46	22	40	-1,9	6,7
8	QATKDAGTIA	1432,8291	0,0122	43,98	39	42	-0,28	6,66
8	QQQYQQQPQ	1403,6996	0,0013	50,22		41	-3,044444444	5,92
8	KQLENVS	1274,76	-0,0058	44,56		42	-1,028571429	6,84
8	WDIQREIQKA	1744,0037	0,0071	55,3	27	42	-1,25	6,92
8	KTLLEAI	1244,8109	-0,0113	44,69		39	0,828571429	6,84

8	KTLLEAI	1244,8109	-0,0113	44,69		39	0,828571429	6,84
9	EILQNRGYLVA	1503,8612	-0,0096	44,8	38	42	0,127272727	6,88
9	DGFAEVFPQHKYRVV	2249,2363	0,0091	87,37	30	43	-0,36	7,66
9	EILQNRGYLVA	1503,8612	-0,0096	44,8	38	42	0,127272727	6,88
9	DGFAEVFPQHKYRVV	2249,2363	0,0091	87,37	30	43	-0,36	7,66
9	NHPGQIS	980,5242	-0,0034	48,72	33	41	-1,214285714	7,84
9	VLNHPGQI	1105,6447	-0,0006	47,27	33	42	0,0375	7,84
9	VLNHPGQIS	1192,6767	-0,0015	59,94	40	42	-0,055555556	7,84
9	GHVDSGKST	1344,7403	-0,0127	42,49	31	41	-1,055555556	7,72
9	RLWPNGQDQPS	1525,784	-0,0039	57,92	34	41	-1,818181818	6,78
9	RLWPNGQDQPS	1525,784	-0,0039	57,92	34	41	-1,818181818	6,78
9	FGSGLHGWA	1159,5977	0,0119	50,86	38	41	0,255555556	7,84
9	RLWPNGQDQPS	1525,784	-0,0039	57,92	34	41	-1,818181818	6,78
9	MHDDQPFYKKM	2126,1261	-0,049	47,02	22	42	-1,618181818	7,6
9	MHDDQPFYKKM	2126,1261	-0,049	47,02	22	42	-1,618181818	7,6
9	RLWPNGQDQPS	1525,784	-0,0039	57,92	34	41	-1,818181818	6,78
9	ANHFGA	844,4394	0,0007	39,1		39	-0,116666667	7,84
9	MLAGTPF	964,5255	-0,0074	44,11	38	42	1,085714286	6,02
9	ANHFGADKL	1429,8083	-0,0033	44,35	27	41	-0,477777778	7,72
9	YQQQQQHPG	1341,6629	-0,0206	45,79	25	39	-2,666666667	7,76
9	LFLPPVAV	1083,6895	0,0058	42,5		36	2,175	6,02
9	AATENQPKGI	1485,8557	-0,022	49,92	30	41	-0,9	6,84
9	HFYVAPT	1062,5701	-0,0042	45,01	31	42	0,285714286	7,76
9	HFYNSQPG	1177,5719	-0,0142	45,23	25	39	-1,4375	7,76
9	GHSGTLDPKVT	1568,8928	-0,0274	43,89	27	42	-0,654545455	7,72
9	ELIQNRLNME	1503,7918	0,0598	44,8	38	42	-0,8	4,26
9	VYLSGDV	980,5382	0,0178	41,83	39	41	0,885714286	3,1
9	VYISSIA	980,5745	-0,0186	41,83	39	41	1,728571429	5,92
9	AYISEAV	980,5381	0,0178	41,83	39	41	0,957142857	3,3

10	NHPGQIS	980,5242	-0,0017	46,29	33	41	-1,214285714	7,84
10	LNHPGQIS	1093,6083	0,0018	49,91	34	42	-0,5875	7,84
10	VLNHPGQI	1105,6447	0,0028	48,63	34	42	0,0375	7,84
10	GHVDSGKST	1344,7403	0,0005	43,09	29	41	-1,055555556	7,72
10	GISKDGQTREH	1684,9262	0,0015	54,89	22	42	-1,809090909	7,72
10	AHVDHGKS	1307,7352	-0,0022	45,6	23	42	-1,125	7,88
10	AHVDHGKST	1408,7829	-0,0004	44,38	27	42	-1,077777778	7,88
10	ANHFGADKL	1429,8083	0,0002	45,43	31	41	-0,477777778	7,72
10	KNAEQH	1183,6715	0,0126	40,83		40	-2,633333333	7,72
10	AHVDHGKS	1307,7352	-0,0022	45,6	23	42	-1,125	7,88
10	AHVDHGKST	1408,7829	-0,0004	44,38	27	42	-1,077777778	7,88
11	MFLPKGGIIQ	1560,9467	-0,0419	46,18	36	42	0,77	9,74
11	NDAPSLKKA	1630,0022	-0,0142	46,83	29	41	-1,088888889	9,46
11	KFVMFFVGPI	1641,9723	-0,0221	50,56	27	41	1,73	9,74
11	KFVMFFVGPI	1657,9672	-0,0573	53,34	26	42	1,73	9,74
11	FLPKGGIIQNF	1691,0176	-0,0029	49,32	27	41	0,463636364	9,74
11	FLPKGGIIQNFGA	1819,0762	-0,0432	54,39	32	42	0,5	9,74
11	MFLPKGGIIQNF	1822,0581	0,0029	63,94	29	42	0,583333333	9,74
11	MFLPKGGIIQNFGAM	2081,1572	-0,0377	47,49	30	43	0,686666667	9,74
11	MTGDGVNDAPSLKKA	2190,2287	-0,0302	80,8	28	43	-0,58	6,72
11	LKQYPKAKDALT	2291,4388	-0,029	41,52	26	40	-0,925	9,64
11	NDAPSLKKA	1630,0022	-0,0142	46,83	29	41	-1,088888889	9,46
11	MTGDGVNDAPSLKKA	2190,2287	-0,0302	80,8	28	43	-0,58	6,72
11	KIGGIGTVPV	1397,9012	-0,0229	62,28	34	38	1	9,74
11	SGKKLEDHPK	2054,2659	0,0123	45,86	32	39	-2,09	9,3
11	SFKYAWVLDKL	2056,233	-0,0586	49,75	33	42	0,190909091	9,34
11	GKKLEDHPKFL	2227,3864	0,0312	38,61	26	37	-1,227272727	9,3
11	SGKKLEDHPKFL	2314,4184	0,0454	45,77	32	38	-1,191666667	9,3

11	DKPLRLPLQDVYKI	2384,4764	-0,0505	44,58	26	41	-0,514285714	9,34
11	SINKWYQYFIG	1876,0289	-0,0301	80,77		42	-0,345454545	9,38
11	THDDKPLYKR	1959,151	-0,0049	58,09	26	42	-2,23	9,34
11	THDDKPLYKRM	2090,1915	-0,0473	45,06	25	43	-1,854545455	9,34
11	SWGELFSSKTKVL	2168,2814	0,0333	67,25	37	40	-0,084615385	9,46
11	SINKWYQYFIG	1876,0289	-0,0301	80,77		42	-0,345454545	9,38
11	SWGELFSSKTKVL	2168,2814	0,0333	67,25	37	40	-0,084615385	9,46
11	KISPPVV	1196,7898	0,016	40,05		32	0,714285714	9,74
11	KISPPVVA	1267,8269	-0,0111	47,34	31	35	0,85	9,74
11	AHVDHGKST	1408,7829	0,0089	46,23	28	42	-1,077777778	7,88
11	KNANDLPKLV	1798,1285	0,0242	54,48	34	37	-0,63	9,46
11	SINKWYQYFIG	1876,0289	-0,0301	80,77		42	-0,345454545	9,38
11	SINKWYQYFIG	1876,0289	-0,0301	80,77		42	-0,345454545	9,38
11	AVEKGTGKS	1562,96	0,0218	41,55	26	39	-0,844444444	9,46
11	SAVEKGTGKS	1649,9921	0,0384	52,74	33	39	-0,84	9,46
11	KTKDNNLLG	1689,0393	-0,0496	45,46	33	42	-1,311111111	9,46
11	SAVEKGTGKSN	1764,035	0,0053	53,1	30	41	-1,081818182	9,46
11	AVEKGTGKS	1562,96	0,0218	41,55	26	39	-0,844444444	9,46
11	SAVEKGTGKS	1649,9921	0,0384	52,74	33	39	-0,84	9,46
11	KTKDNNLLG	1689,0393	-0,0496	45,46	33	42	-1,311111111	9,46
11	SAVEKGTGKSN	1764,035	0,0053	53,1	30	41	-1,081818182	9,46
11	SINKWYQYFIG	1876,0289	-0,0301	80,77		42	-0,345454545	9,38
11	KYNHVQT	1346,7712	-0,0061	47,55	32	42	-1,7	9,52
11	IKINPPTGA	1367,8542	-0,0219	41,12		39	-0,1	9,74
11	KYNHVQTS	1433,8033	0,0068	56,62	25	42	-1,5875	9,52
11	STQKEHPGYRQI	1901,0525	0,0548	52,59	33	42	-1,866666667	9,52
11	SINKWYQYFIG	1876,0289	-0,0301	80,77		42	-0,345454545	9,38
11	SINKWYQYFIG	1876,0289	-0,0301	80,77		42	-0,345454545	9,38
11	SINKWYQYFIG	1876,0289	-0,0301	80,77		42	-0,345454545	9,38
11	KYNHVQT	1346,7712	-0,0061	47,55	32	42	-1,7	9,52

11	IKINPPTGA	1367,8542	-0,0219	41,12		39	-0,1	9,74
11	KYNHVQTS	1433,8033	0,0068	56,62	25	42	-1,5875	9,52
11	STQKEHPGYRQI	1901,0525	0,0548	52,59	33	42	-1,866666667	9,52
11	AVEKGTGKS	1562,96	0,0218	41,55	26	39	-0,844444444	9,46
11	SAVEKGTGKS	1649,9921	0,0384	52,74	33	39	-0,84	9,46
11	SAVEKGTGKSN	1764,035	0,0053	53,1	30	41	-1,081818182	9,46
11	AHLKYKDPQ	1786,071	-0,0132	42,82	23	41	-1,7	9,34
11	LFLQKGGQTV	1547,9441	-0,024	48,75	30	41	0,22	9,74
11	EKPKAEKS	1832,1542	0,0497	45,23	34	36	-2,4125	9,3
11	KAEKPKAEKS	2260,4492	0,0524	34,9	24	33	-2,14	9,6
11	IKDIVQKT	1631,059	-0,0081	38,96		38	-0,2875	9,46
11	QKSKGESDKPSAS	2264,3147	-0,0007	59,5	28	42	-1,969230769	9,3
11	QKEKNKELNKQ	2531,5773	0,0626	40,39	26	35	-2,981818182	9,6
11	QRDGLHDSNKQAQ	1954,0386	-0,001	43,34	25	42	-2,169230769	7,72
11	EAPMNPKSNREK	2087,1765	0,0417	49,66	32	42	-2,216666667	9,46
11	ASKKKEEAKPA	2331,4863	0,0407	37,19	28	35	-1,781818182	9,6
11	KFGSVALPI	1388,8797	-0,0244	43,49	36	41	1,155555556	9,74
11	KQPKFDVG	1604,9859	-0,0447	43,84		42	-1,225	9,46
11	LQKIKPGDILVI	2023,3378	0,008	52,74		32	0,708333333	9,46
11	KQPKFDVG	1604,9859	-0,0447	43,84		42	-1,225	9,46
11	SINKWYQYFIG	1876,0289	-0,0301	80,77		42	-0,345454545	9,38
11	RKSEELYPGRPL	1902,1092	-0,0438	44,6	27	42	-1,5	9,52
11	MLMPKEDRNKI	2061,2047	0,04	47,65	30	41	-1,118181818	9,46
11	RKQQAAPGGNTSEA	1872,0219	-0,0228	56,26	29	42	-1,492857143	9,74
11	LRHLEKEGIIKPI	2232,429	-0,0466	39,24	25	39	-0,261538462	9,46
11	ALRHLEKEGIIKPI	2303,4661	-0,0642	54,97	26	40	-0,114285714	9,46
11	VQKFFLPL	1448,9161	-0,053	43,58	30	41	1,05	9,74
11	YHQKYVDEQSKNELKKA	3252,8844	0,0776	42,92	19	40	-1,964705882	9,06
11	KWFKDTPFI	1868,1169	0,0159	41,71	30	40	-0,488888889	9,46
11	AHVDHGKST	1408,7829	0,0089	46,23	28	42	-1,077777778	7,88

11	KYQYPQTPS	1568,8604	0,02	56,1	30	42	-2,022222222	9,38
11	TQKYQYPQTPS	1797,9667	0,0238	50,07	27	42	-2,036363636	9,38
11	KHFELGGEKKQKG	2630,6246	-0,0356	47,58	27	40	-1,838461538	9,6
11	KHFELGGEKKQKGQ	2758,6832	-0,0303	56,05	27	41	-1,957142857	9,6
11	THKRNEYIYDRS	2039,0954	0,0496	42,98	26	42	-2,183333333	9,38
11	KKGQKDKIHEQS	2570,5882	0,0838	50,9	25	34	-2,458333333	9,6
11	AHVDHGKTS	1408,7829	0,0089	46,23	28	42	-1,077777778	7,88
11	KQALPGIPA	1351,8593	-0,0095	48,52	29	39	0,1	9,74
11	KQALPGIPA	1351,8593	-0,0095	48,52	29	39	0,1	9,74
12	MFLPKGGII	1533,9359	-0,0422	49,9	39	42	1,244444444	9,74
12	TMFLPKGGII	1533,9851	-0,0744	56,79	37	42	1,05	9,74
12	LKQYPKA	1679,9799	-0,0776	44,9	35	42	-1,228571429	9,82
12	TVKRGEGFMVV	1767,0119	-0,0756	44,12	30	42	0,354545455	9,74
12	STVKRGEGFMVV	1930,1194	-0,0314	44,03	27	42	0,258333333	9,74
12	AEVFPQHKYRVV	2102,371	-0,0449	46,91	25	42	-0,358333333	9,52
12	ASRKKKGLDAI	1930,1194	-0,0314	39,91	27	38	-0,818181818	10,24
12	AEVFPQHKYRVV	2102,371	-0,0449	46,91	25	42	-0,358333333	9,52
12	ASRKKKGLDAI	1569,9608	-0,0348	39,91	27	38	-0,818181818	10,24
12	GRVETGVIKPG	2098,1325	-0,0989	54,31	33	41	-0,227272727	9,74
12	RLWPNGQDQPSSKGA	2168,2814	-0,0623	42,52	24	41	-1,553333333	9,74
12	SWGELFSSKTKVL	2098,1325	-0,0989	59,77	29	42	-0,084615385	9,46
12	RLWPNGQDQPSSKGA	2168,2814	-0,0623	42,52	24	41	-1,553333333	9,74
12	SWGELFSSKTKVL	2098,1325	-0,0989	59,77	29	42	-0,084615385	9,46
12	RLWPNGQDQPSSKGA	2098,1325	-0,0989	42,52	24	41	-1,553333333	9,74
12	RLWPNGQDQPSSKGA	1750,0329	-0,0397	42,52	24	41	-1,553333333	9,74
12	MAPKAPVASPRPA	1814,0204	-0,0584	46,15	22	42	-0,176923077	11,24
12	TQKEHPGYRQI	1901,0525	-0,0553	49,33	35	42	-1,963636364	9,52
12	STQKEHPGYRQI	1814,0204	-0,0584	61,2	37	42	-1,866666667	9,52
12	TQKEHPGYRQI	1901,0525	-0,0553	49,33	35	42	-1,963636364	9,52

12	STQKEHPGYRQI	2201,1707	-0,095	61,2	37	42	-1,866666667	9,52
12	SHLGRPNGERNEKYS	1622,9146	-0,0553	79,86	37	41	-2,106666667	9,52
12	TFENRNVGKT	1622,9146	-0,0553	44,21	25	42	-1,37	9,74
12	TFENRNVGKT	2613,5363	-0,1198	44,21	25	42	-1,37	9,74
12	GFHLIKPGYENKVPRTA	2076,1718	-0,0928	44,41	25	43	-0,670588235	9,82
12	QQQQKQMQQKS	2446,2831	-0,0548	77,21		42	-2,836363636	10,06
12	NQQEKQQQRQQQPSPH	1615,9452	-0,0769	53,11	23	42	-3,1625	9,74
12	RGEIGPFASPK	2501,3756	-0,0725	49,31	41	42	-0,690909091	9,74
12	SGNNNIPTPPQNRDVPKPV	2081,1747	-0,0766	46,4	26	43	-1,389473684	9,74
12	IQEQQNLRQKQPN	1301,7984	-0,0088	49,43	27	42	-2,284615385	9,74
12	FLPILLNML	1301,7798	0,0098	58,5		41	2,144444444	6,02
12	FIPLLAEGTL	1806,9643	-0,0721	50,67		41	1,43	3,3
12	NHQNPHFRDKG	1308,7427	0,0075	49,57	39	41	-2,545454545	9,74
12	GLIRISFMQ	2061,2969	-0,0736	43,2	40	42	0,922222222	11,04
12	KSQTPKVEKT			43,82	29	41	-1,83	9,78

Supplementry table S6 B (S6B) summary of all proteins, peptides sequence, calculated mass (pep_calc_mr), mass acuuracy (pep_delta), peptide score (pep_score), homology (pep_homol), identity threshold (pep_ident), and the type of modifications on the peptides from a tryptic digestion of purified ENL fraction.

prot_desc	pep_calc_mr	pep_delta	pep_score	pep_homol	pep_ident	pep_seq	pep_var_mod
Protein ENL	723,4028	0,0232	45,47		28	LELGHR	
Protein ENL	760,3967	0,0044	52,04		34	LESTSPK	
Protein ENL	771,3875	0,012	37,75		31	NKEEPR	
Protein ENL	771,3875	0,0128	55,46		31	SAPGTSPR	
Protein ENL	789,4385	0,0224	44,43		32	AAFKEPK	
Protein ENL	840,4454	-0,0064	49,21	29	31	TKPSHGSK	
Protein ENL	851,3538	-0,0164	37,87		28	SAPGTSPR	Phospho (S)
Protein ENL	914,4821	-0,0009	60,85		33	VKAESEPR	
Protein ENL	1001,5141	-0,0442	48,55		32	ELEREQAK	
Protein ENL	1019,511	-0,0359	36,97		33	VCKEPPYK	
Protein ENL	1031,4407	-0,047	73,19		28	TSSSSSFSDK	
Protein ENL	1118,5139	-0,0242	66,75		32	DNQCTVQVR	
Protein ENL	1118,5819	-0,0049	39,31		34	ETKLESTSPK	
Protein ENL	1147,5179	0,0397	62,94		33	SPESCSKPEK	
Protein ENL	1161,6295	-0,0522	95,28		34	GGPPPPPPPPPR	
Protein ENL	1198,5482	-0,0359	46,52	31	32	ETKLESTSPK	Phospho (S)
Protein ENL	1227,4842	-0,0497	54,81	22	26	SPESCSKPEK	Phospho (S)

Protein ENL	1303,619	-0,0649	63,89		31	RSPESCSKPEK	
Protein ENL	1324,7099	-0,043	96,06		34	RPATADSPKPSAK	
Protein ENL	1338,6568	-0,0568	91,13		33	LTFNNPTTEFR	
Protein ENL	1344,6674	0,05	96,13		34	AYTDELVELHR	
Protein ENL	1383,5853	-0,0253	54,01		31	RSPESCSKPEK	Phospho (S)
Protein ENL	1404,6762	-0,0091	69,02		34	RPATADSPKPSAK	Phospho (S)
Protein ENL	1455,7205	-0,0386	87,38		34	TSSSSSFSDKKPAK	
Protein ENL	1500,7685	0,0253	97,11		34	AYTDELVELHRR	
Protein ENL	1508,7987	0,0029	81,88	31	34	EPPPPQKPPPPNSK	
Protein ENL	1585,7195	0,0489	106,15		35	GPEQCDIQHFVEK	
Protein ENL	1782,7795	0,0159	54,81	32	34	ALEVEESNSEDEASFK	
Protein ENL	1782,8084	0,0479	44,04	31	35	VSGRRSPESCSKPEK	Phospho (S)
Protein ENL	1862,7459	-0,0065	135,56		31	ALEVEESNSEDEASFK	Phospho (S)
Protein ENL	1942,7122	-0,0203	74,63		28	ALEVEESNSEDEASFK	2 Phospho (S)
Protein ENL	1983,9819	-0,0283	113,45		35	LESTSPKGGPPPPPPPPR	Phospho (S)
Protein ENL	1990,8408	0,001	88,48	28	33	KALEVEESNSEDEASFK	Phospho (S)
Protein ENL	1999,8607	-0,0505	144,82		32	LSFSDSESDNSADSSLPSR	
Protein ENL	2052,0026	0,0162	96,78	34	35	VEESGYAGFIMPIEVHFK	
Protein ENL	2079,827	0,0095	142,7		32	LSFSDSESDNSADSSLPSR	Phospho (S)
Protein ENL	2193,1178	0,0997	44,47	29	32	KGTYDKAYTDELVELHRR	
Protein ENL	2342,1671	-0,075	58,49	26	35	ETKLESTSPKGGPPPPPPPPR	Phospho (S)
Protein ENL	2517,2475	-0,0521	149,14		36	VCFTYDLFLNLEGNPPVNHLR	
Protein ENL	2630,0297	-0,0277	59,27	29	30	SMVEDLQSEESDEDDSSSGEEAAGK	
Protein ENL	2645,3424	0,0813	156,72		33	KVCFTYDLFLNLEGNPPVNHLR	
Protein ENL	2666,309	0,0122	68,66	28	36	EPPYKVEESGYAGFIMPIEVHFK	
Protein ENL	3398,7193	0,0352	73,48	30	35	NKEEPRKVCFTYDLFLNLEGNPPVNHLR	
Protein ENL	3490,6488	0,0083	136,32		36	LSFSDSESDNSADSSLPSREPPPPQKPPPPNSK	
Protein ENL	3570,6151	-0,0261	66,74	30	36	LSFSDSESDNSADSSLPSREPPPPQKPPPPNSK	Phospho (S)
Protein ENL	3598,48	0,0309	225,02		34	SESAQSSPSNSSSSSDSSSDSDFEPSQNHSQGPLR	
Heat shock 70 kDa	800,414	0,0168	36,37		33	DNNLLGR	

protein 1							
Heat shock 70 kDa protein 1	859,4512	0,0189	45,94		34	DISQNKR	
Heat shock 70 kDa protein 1	1002,5345	-0,0224	73,08		33	LSKEEIER	
Heat shock 70 kDa protein 1	1108,5665	0,0475	88,73		32	LLQDFFNGR	
Heat shock 70 kDa protein 1	1136,5462	0,0006	75,48		33	YKAEDEVQR	
Heat shock 70 kDa protein 1	1182,6397	-0,0304	44,02	29	33	FELSGIPPAPR	
Heat shock 70 kDa protein 1	1196,6877	-0,0136	113,32		31	DAGVIAGLNVLR	
Heat shock 70 kDa protein 1	1227,6207	-0,0057	65,12		34	VEIIANDQGNR	
Heat shock 70 kDa protein 1	1260,6503	0,0403	63,54		33	LVNHFVEEFK	
Heat shock 70 kDa protein 1	1314,5914	0,0118	79,46		33	FEELCSDLFR	
Heat shock 70 kDa protein 1	1416,7514	0,0578	87		32	LVNHFVEEFKR	
Heat shock 70 kDa protein 1	1464,8049	-0,0123	101,03		33	AQIHDLVLVGGSTR	
Heat shock 70 kDa protein 1	1486,694	-0,0081	76,09		34	TTPSYVAFTDTER	
Heat shock 70 kDa protein 1	1541,7296	0,0591	75,56		34	ARFEELCSDLFR	
Heat shock 70 kDa protein 1	1657,8424	-0,0097	100,74		34	NQVALNPQNTVFDAK	
Heat shock 70 kDa protein 1	1674,7234	0,0213	128,27		34	ATAGDTHLGGEDFDNR	
Heat shock 70 kDa protein 1	1679,842	0,0457	82,08	33	34	HWPFQVINDGDKPK	
Heat shock 70 kDa protein 1	1686,894	0,081	136,28		30	IINEPTAAAIAYGLDR	
Heat shock 70 kDa protein 1	1813,9435	-0,043	110,5		35	NQVALNPQNTVFDAKR	
Heat shock 70 kDa protein 1	1821,0108	-0,0053	94,69		32	LDKAQIHDLVLVGGSTR	
Heat shock 70 kDa protein 1	2264,126	0,0896	132,08		33	AAAIGIDLGTTYSCVGVFQHGK	
Heat shock 70 kDa protein 1	2785,3559	-0,1093	142,81		36	QTQIFTTYSDNQPGVLIQVYEGER	

Heat shock 70 kDa protein 1	2980,4553	-0,0895	125,71		36	TLSSSTQASLEIDSLFEGIDFYTSITR	
Heat shock 70 kDa protein 1	3000,4869	0,0422	160,55		35	EIAEAYLGYPVTNAVITVPAYFNDSQR	
Heat shock 70 kDa protein 1	3054,4869	0,0899	147,92		35	ELEQVCNPIISGLYQGAGGPGPGGFGAQGPK	
Heat shock 70 kDa protein 1	3259,6223	-0,0143	115,47		36	MKEIAEAYLGYPVTNAVITVPAYFNDSQR	
Heat shock 70 kDa protein 1	3338,683	0,0623	186,14		34	RKELEQVCNPIISGLYQGAGGPGPGGFGAQGPK	
Heat shock cognate 71 kDa protein	860,4352	0,0016	43,45		33	DISENKR	
Heat shock cognate 71 kDa protein	1080,5604	-0,0393	73,58		33	LLQDFFNGK	
Heat shock cognate 71 kDa protein	1165,5727	0,0065	73,97		34	YKAEDEKQR	
Heat shock cognate 71 kDa protein	1227,6207	-0,0057	65,12		34	VEIIANDQGNR	
Heat shock cognate 71 kDa protein	1252,6088	0,0156	70,53		34	FEELNADLFR	
Heat shock cognate 71 kDa protein	1409,6609	0,05	54,83		34	RFDDAVVQSDMK	
Heat shock cognate 71 kDa protein	1479,747	0,0493	85,25		33	ARFEELNADLFR	
Heat shock cognate 71 kDa protein	1480,7998	0,053	60,76	26	30	SQIHDIVLVGGSTR	
Heat shock cognate 71 kDa protein	1486,694	-0,0081	76,09		34	TTPSYVAFTDTER	
Heat shock cognate 71 kDa protein	1652,8246	0,0584	61,28	29	33	HWPFMVVNDAGRPK	
Heat shock cognate 71 kDa protein	1658,8879	0,0065	51,93	30	34	IINEPTAAAIAYGLDK	
Heat shock cognate 71 kDa protein	1690,7183	0,0059	117,52		33	STAGDTHLGGEDFDNR	
Heat shock cognate 71 kDa protein	1744,8016	-0,0244	89,19		34	NQTAEKEEFEHQQK	
Heat shock cognate 71 kDa protein	1786,9828	-0,0829	113,07		35	IINEPTAAAIAYGLDKK	
Heat shock cognate 71 kDa protein	1980,9905	0,058	122,45		34	TVTNAVVTVPAYFNDSQR	
Heat shock cognate 71 kDa protein	2259,1383	0,0137	91,8	34	35	SINPDEAVAYGAAVQAAILSGDK	
Heat shock cognate 71 kDa protein	2262,1104	-0,0822	116,08	33	35	GPAVGIDLGTTYSCVGVFQHGK	

Heat shock cognate 71 kDa protein	2773,3195	-0,0373	155,25		36	QTQTFTTYSDNQPGVLIQVYEGER	
Heat shock cognate 71 kDa protein	2996,4502	-0,012	128,61		36	TLSSSTQASIEIDSLYEGIDFYTSITR	
Heat shock 70 kDa protein 1L	800,414	0,0168	36,37		33	DNNLLGR	
Heat shock 70 kDa protein 1L	859,4512	0,0189	45,94		34	DISQNKR	
Heat shock 70 kDa protein 1L	1002,5345	-0,0224	73,08		33	LSKEEIER	
Heat shock 70 kDa protein 1L	1136,5462	0,0006	75,48		33	YKAEDEVQR	
Heat shock 70 kDa protein 1L	1182,6397	-0,0304	41,96	29	33	FDLTGIPPAPR	
Heat shock 70 kDa protein 1L	1196,6877	-0,0136	113,32		31	DAGVIAGLNVLR	
Heat shock 70 kDa protein 1L	1227,6207	-0,0057	65,12		34	VEIIANDQGNR	
Heat shock 70 kDa protein 1L	1464,8413	-0,0487	101,03		33	AKIHDIVLVGGSTR	
Heat shock 70 kDa protein 1L	1486,694	-0,0081	76,09		34	TTPSYVAFTDTER	
Heat shock 70 kDa protein 1L	1658,8879	0,0065	51,93	30	34	IINEPTAAAIAYGLDK	
Heat shock 70 kDa protein 1L	1674,7234	0,0213	128,27		34	ATAGDTHLGGEDFDNR	
Heat shock 70 kDa protein 1L	2785,3559	-0,1093	142,81		36	QTQIFTTYSDNQPGVLIQVYEGER	
Heat shock 70 kDa protein 6	800,414	0,0168	36,37		33	DNNLLGR	
Heat shock 70 kDa protein 6	1080,5604	-0,0393	73,58		33	LLQDFFNGK	
Heat shock 70 kDa protein 6	1182,6397	-0,0304	44,02	29	33	FELSGIPPAPR	
Heat shock 70 kDa protein 6	1227,6207	-0,0057	65,12		34	VEILANDQGNR	
Heat shock 70 kDa protein 6	1314,5914	0,0118	79,46		33	FEELCSDLFR	
Heat shock 70 kDa protein 6	1486,694	-0,0081	76,09		34	TTPSYVAFTDTER	
Heat shock 70 kDa protein 6	1541,7296	0,0591	75,56		34	ARFEELCSDLFR	
Heat shock 70 kDa protein 6	1674,7234	0,0213	128,27		34	ATAGDTHLGGEDFDNR	

Heat shock 70 kDa protein 6	1686,894	0,081	136,28		30	IINEPTAAAIAYGLDR	
Heat shock 70 kDa protein 6	1716,8584	-0,0114	62,97	31	35	HAVITVPAYFNDSQR	
Heat shock 70 kDa protein 6	2980,4553	-0,0895	40,51		36	TLSSSTQATLEIDSLFEGVDFYTSITR	
Tubulin alpha-1B chain	1022,4417	0,0103	53,1		31	EDAANNYAR	
Tubulin alpha-1B chain	1084,6128	-0,0189	57,97		33	EIIDLVLDR	
Tubulin alpha-1B chain	1409,7667	-0,005	46,11		33	QLFHPEQLITGK	
Tubulin alpha-1B chain	1486,8719	0,0056	127,37		30	LISQIVSSITASLR	
Tubulin alpha-1B chain	1700,8985	-0,0824	122,31		35	AVFVDLEPTVIDEVR	
Tubulin alpha-1B chain	1717,8747	0,0338	69,07		34	NLDIERPTYTNLNR	
Tubulin alpha-1B chain	1755,9559	-0,0277	80,63	30	34	IHFPLATYAPVISAEK	
Tubulin alpha-1B chain	1863,8971	-0,0398	91,42	34	35	AVCMLSNTTAIAEAWAR	
Tubulin alpha-1B chain	2006,8858	-0,0318	109,22		33	TIGGGDDSFNTFFSETGAGK	
Tubulin alpha-1B chain	2329,011	-0,0727	33,68	26	32	AFVHWYVGEGMEEGEFSEAR	
Tubulin alpha-1B chain	2408,2012	-0,0298	154,58		36	FDGALNVDLTEFQTNLVPYPR	
Tubulin alpha-1B chain	2414,1978	0,0363	79,65		35	QLFHPEQLITGKEDAANNYAR	
Tubulin alpha-1B chain	2749,284	0,1104	54,42	34	35	AYHEQLSVAEITNACFEPANQMVK	
Tubulin beta chain	1038,5862	0,003	49,63		31	YLTVAAVFR	
Tubulin beta chain	1076,525	-0,0283	42,35		32	IREEYPDR	
Tubulin beta chain	1129,588	-0,0212	96,81		34	FPGQLNADLR	
Tubulin beta chain	1300,6299	-0,0099	45,22	27	33	ISVYYNEATGGK	
Tubulin beta chain	1957,9745	-0,0487	175,3		35	GHYTEGAELVDSVLDVVR	
Tubulin beta chain	2086,0695	-0,0734	144,68		35	GHYTEGAELVDSVLDVVRK	
Tubulin beta chain	2707,331	0,0979	125,06	33	34	LTTPTYGDLNHLVSATMSGVTTCLR	
Tubulin beta chain	2797,3361	-0,0461	110,72		36	SGPFGQIFRPDNFVFGQSGAGNNWAK	
Tubulin beta chain	3101,4003	-0,0972	114,52		34	FWEVISDEHGIDPTGTYHGDSDLQLDR	
Tubulin alpha-1A chain	1022,4417	0,0103	53,1		31	EDAANNYAR	
Tubulin alpha-1A chain	1084,6128	-0,0189	57,97		33	EIIDLVLDR	
Tubulin alpha-1A chain	1409,7667	-0,005	46,11		33	QLFHPEQLITGK	
Tubulin alpha-1A chain	1456,8613	-0,0445	46,63		33	LIGQIVSSITASLR	

Tubulin alpha-1A chain	1700,8985	-0,0824	122,31		35	AVFVDLEPTVIDEVR	
Tubulin alpha-1A chain	1717,8747	0,0338	69,07		34	NLDIERPTYTNLNR	
Tubulin alpha-1A chain	1755,9559	-0,0277	80,63	30	34	IHFPLATYAPVISAEK	
Tubulin alpha-1A chain	1863,8971	-0,0398	91,42	34	35	AVCMLSNTTAIAEAWAR	
Tubulin alpha-1A chain	2006,8858	-0,0318	109,22		33	TIGGGDDSFNTFFSETGAGK	
Tubulin alpha-1A chain	2329,011	-0,0727	33,68	26	32	AFVHWYVGEGMEEGEFSEAR	
Tubulin alpha-1A chain	2408,2012	-0,0298	154,58		36	FDGALNVDLTEFQTNLVPYPR	
Tubulin alpha-1A chain	2414,1978	0,0363	79,65		35	QLFHPEQLITGKEDAANNYAR	
Tubulin alpha-1A chain	2749,284	0,1104	54,42	34	35	AYHEQLSVAEITNACFEPANQMVK	
Heat shock-related 70 kDa protein 2	1080,5604	-0,0393	73,58		33	LLQDFFNGK	
Heat shock-related 70 kDa protein 2	1182,6397	-0,0304	41,96	29	33	FDLTGIPPAPR	
Heat shock-related 70 kDa protein 2	1227,6207	-0,0057	65,12		34	VEIIANDQGNR	
Heat shock-related 70 kDa protein 2	1252,6088	0,0156	70,53		34	FEELNADLFR	
Heat shock-related 70 kDa protein 2	1479,747	0,0493	85,25		33	ARFEELNADLFR	
Heat shock-related 70 kDa protein 2	1486,694	-0,0081	76,09		34	TTPSYVAFTDTER	
Heat shock-related 70 kDa protein 2	1658,8879	0,0065	51,93	30	34	IINEPTAAAIAYGLDK	
Heat shock-related 70 kDa protein 2	1690,7183	0,0059	117,52		33	STAGDTHLGGEDFDNR	
Heat shock-related 70 kDa protein 2	1786,9828	-0,0829	113,07		35	IINEPTAAAIAYGLDKK	
Heat shock-related 70 kDa protein 2	3014,4539	-0,115	71,06	34	35	EIAEAYLGGKVHSAVITVPAYFNDSQR	Phospho (S)
Heat shock protein HSP 90-beta	828,5221	0,006	38,52		31	ALLFIPR	
Heat shock protein HSP 90-beta	900,5181	-0,0278	43,03		34	TKPIWTR	
Heat shock protein HSP 90-beta	950,457	-0,03	42,16		32	ADHGEPIGR	
Heat shock protein HSP 90-beta	1235,6299	0,0151	57,08		33	RAPFDLFENK	
Heat shock protein HSP 90-beta	1248,6098	-0,0366	63,04		33	EQVANSAFVER	

Heat shock protein HSP 90-beta	1310,5626	-0,0276	94,05		31	EDQTEYLEER	
Heat shock protein HSP 90-beta	1363,7248	-0,0132	39,51		34	RAPFDLFENKK	
Heat shock protein HSP 90-beta	1512,7784	-0,0072	73,2	32	34	GVVDSEDLPLNISR	
Heat shock protein HSP 90-beta	1526,7365	0,0552	47,38	32	34	SLTNDWEDHLAVK	
Heat shock protein HSP 90-beta	1781,9424	-0,0203	81,64	34	35	HLEINPDHPIVETLR	
Heat shock protein HSP 90-beta	1807,9509	0,0102	130,1		34	HSQFIGYPITLYLEK	
Heat shock protein HSP 90-beta	1846,7897	-0,0654	81,52		31	NPDDITQEEYGEFYK	
Heat shock protein HSP 90-beta	1910,0374	-0,0043	53,09	28	33	KHLEINPDHPIVETLR	
Heat shock protein HSP 90-beta	2014,0371	-0,0175	85,05		35	VILHLKEDQTEYLEER	
Heat shock protein HSP 90-beta	2175,9379	-0,0757	46,87	24	32	YHTSQSGDEMTSLSEYVSR	
Heat shock protein HSP 90-beta	2254,9516	-0,009	118,18		34	HNDDEQYAWESSAGGSFTVR	
Heat shock protein HSP 90-beta	2373,1385	-0,0052	102,9		36	VFIMDSCDELIPEYLNFIR	
Heat shock protein HSP 90-beta	2431,097	-0,0011	66,87	30	35	LVSSPCCIVTSTYGWTANMER	
Heat shock protein HSP 90-beta	3531,6376	-0,1257	112,46		35	LGLGIDEDEVAAEEPNAAVPDEIPPLEGDEDASR	
Heat shock protein HSP 90-alpha	814,5065	-0,0056	50,11		33	ALLFVPR	
Heat shock protein HSP 90-alpha	900,5181	-0,0278	43,03		34	TKPIWTR	
Heat shock protein HSP 90-alpha	1223,6186	-0,0052	34,47	29	34	HIYYITGETK	
Heat shock protein HSP 90-alpha	1263,636	0,0131	43,3	31	34	RAPFDLFENR	
Heat shock protein HSP 90-alpha	1310,5626	-0,0276	94,05		31	EDQTEYLEER	
Heat shock protein HSP 90-alpha	1512,7784	-0,0072	73,2	32	34	GVVDSEDLPLNISR	
Heat shock protein HSP 90-alpha	1526,7365	0,0552	47,38	32	34	SLTNDWEDHLAVK	
Heat shock protein HSP 90-alpha	1588,8685	0,0444	43,87	29	31	ELHINLIPNKQDR	

Heat shock protein HSP 90-alpha	1777,9403	-0,0783	120,26		35	HSQFIGYPITLFVEK	
Heat shock protein HSP 90-alpha	1785,9373	0,0232	93,5		34	HLEINPDHSIIETLR	
Heat shock protein HSP 90-alpha	1832,7741	-0,0693	80,59		31	NPDDITNEEYGEFYK	
Heat shock protein HSP 90-alpha	1914,0323	-0,0793	84,08		35	KHLEINPDHSIIETLR	
Heat shock protein HSP 90-alpha	2014,0371	-0,0175	85,05		35	VILHLKEDQTEYLEER	
Heat shock protein HSP 90-alpha	2063,084	-0,0498	56,96	27	35	HSQFIGYPITLFVEKER	
Heat shock protein HSP 90-alpha	2254,9516	-0,009	118,18		34	HNDDEQYAWESSAGGSFTVR	
Heat shock protein HSP 90-alpha	2414,165	0,0344	99,03		36	VFIMDNCEELIPEYLNFIR	
Heat shock protein HSP 90-alpha	2445,1127	0,0189	39,45	30	36	LVTSPCCIVTSTYGWTANMER	
Tubulin beta-2A chain	1076,525	-0,0283	42,35		32	IREEYPDR	
Tubulin beta-2A chain	1129,588	-0,0212	96,81		34	FPGQLNADLR	
Tubulin beta-2A chain	1957,9745	-0,0487	175,3		35	GHYTEGAELVDSVLDVVR	
Tubulin beta-2A chain	2086,0695	-0,0734	144,68		35	GHYTEGAELVDSVLDVVRK	
Tubulin beta-2A chain	2707,331	0,0979	125,06	33	34	LTTPTYGDLNHLVSATMSGVTTCLR	
Tubulin beta-2A chain	2797,3361	-0,0461	110,72		36	SGPFGQIFRPDNFVFGQSGAGNNWAK	
Tubulin beta-2A chain	3101,4003	-0,0413	59,01		35	FWEVISDEHGIDPTGSYHGDSDLQLER	
Tubulin beta-2C chain	1038,5862	0,003	49,63		31	YLTVAAVFR	
Tubulin beta-2C chain	1076,525	-0,0283	42,35		32	IREEYPDR	
Tubulin beta-2C chain	1129,588	-0,0212	96,81		34	FPGQLNADLR	
Tubulin beta-2C chain	1957,9745	-0,0487	175,3		35	GHYTEGAELVDSVLDVVR	
Tubulin beta-2C chain	2086,0695	-0,0734	144,68		35	GHYTEGAELVDSVLDVVRK	
Tubulin beta-2C chain	2707,331	0,0979	125,06	33	34	LTTPTYGDLNHLVSATMSGVTTCLR	
Tubulin beta-2C chain	2797,3361	-0,0461	110,72		36	SGPFGQIFRPDNFVFGQSGAGNNWAK	
Tubulin alpha-8 chain	1022,4417	0,0103	53,1		31	EDAANNYAR	
Tubulin alpha-8 chain	1409,7667	-0,005	46,11		33	QLFHPEQLITGK	
Tubulin alpha-8 chain	1486,8719	0,0056	127,37		30	LISQIVSSITASLR	

Tubulin alpha-8 chain	1717,8747	0,0338	69,07		34	NLDIERPTYTNLNR	
Tubulin alpha-8 chain	1863,8971	-0,0398	91,42	34	35	AVCMLSNTTAIAEAWAR	
Tubulin alpha-8 chain	2329,011	-0,0727	33,68	26	32	AFVHWYVGEGMEEGEFSEAR	
Tubulin alpha-8 chain	2408,2012	-0,0298	154,58		36	FDGALNVDLTEFQTNLVPYPR	
Tubulin alpha-8 chain	2414,1978	0,0363	79,65		35	QLFHPEQLITGKEDAANNYAR	
Tubulin alpha-3C/D chain	1022,4417	0,0103	53,1		31	EDAANNYAR	
Tubulin alpha-3C/D chain	1409,7667	-0,005	46,11		33	QLFHPEQLITGK	
Tubulin alpha-3C/D chain	1456,8613	-0,0445	46,63		33	LIGQIVSSITASLR	
Tubulin alpha-3C/D chain	1686,8829	0,0208	34,6		34	AVFVDLEPTVVDEVR	
Tubulin alpha-3C/D chain	1717,8747	0,0338	69,07		34	NLDIERPTYTNLNR	
Tubulin alpha-3C/D chain	1755,9559	-0,0277	80,63	30	34	IHFPLATYAPVISAEK	
Tubulin alpha-3C/D chain	1863,8971	-0,0398	91,42	34	35	AVCMLSNTTAIAEAWAR	
Tubulin alpha-3C/D chain	2006,8858	-0,0318	109,22		33	TIGGGDDSFNTFFSETGAGK	
Tubulin alpha-3C/D chain	2329,011	-0,0727	33,68	26	32	AFVHWYVGEGMEEGEFSEAR	
Tubulin alpha-3C/D chain	2408,2012	-0,0298	154,58		36	FDGALNVDLTEFQTNLVPYPR	
Tubulin alpha-3C/D chain	2414,1978	0,0363	79,65		35	QLFHPEQLITGKEDAANNYAR	
Tubulin alpha-3C/D chain	2749,284	0,1104	54,42	34	35	AYHEQLSVAEITNACFEPANQMVK	
Tubulin alpha-3E chain	1409,7667	-0,005	46,11		33	QLFHPEQLITGK	
Tubulin alpha-3E chain	1456,8613	-0,0445	46,63		33	LIGQIVSSITASLR	
Tubulin alpha-3E chain	1686,8829	0,0208	34,6		34	AVFVDLEPTVVDEVR	
Tubulin alpha-3E chain	1717,8747	0,0338	69,07		34	NLDIERPTYTNLNR	
Tubulin alpha-3E chain	1755,9559	-0,0277	80,63	30	34	IHFPLATYAPVISAEK	
Tubulin alpha-3E chain	1863,8971	-0,0398	91,42	34	35	AVCMLSNTTAIAEAWAR	
Tubulin alpha-3E chain	2006,8858	-0,0318	109,22		33	TIGGGDDSFNTFFSETGAGK	
Tubulin alpha-3E chain	2408,2012	-0,0298	154,58		36	FDGALNVDLTEFQTNLVPYPR	

Tubulin alpha-3E chain	2749,284	0,1104	54,42	34	35	AYHEQLSVAEITNACFEPANQMVK	
78 kDa glucose-regulated protein	996,5101	0,0069	48,01		32	ALSSQHQAR	
78 kDa glucose-regulated protein	1227,6207	-0,0057	65,12		34	VEIIANDQGNR	
78 kDa glucose-regulated protein	1429,6838	-0,0634	49,57		33	TWNDPSVQQDIK	
78 kDa glucose-regulated protein	1511,7442	-0,0037	46,83	29	34	AKFEELNMDLFR	
78 kDa glucose-regulated protein	1565,7726	-0,0396	58,3		34	ITPSYVAFTPEGER	
78 kDa glucose-regulated protein	1658,8879	0,0065	51,93	30	34	IINEPTAAAIAYGLDK	
78 kDa glucose-regulated protein	1814,989	-0,0754	97,63		35	IINEPTAAAIAYGLDKR	
78 kDa glucose-regulated protein	1886,9639	-0,0011	87,82	27	35	VTHAVVTVPAYFNDAQR	
78 kDa glucose-regulated protein	1973,9007	-0,0546	46,82	30	34	IEWLESHQDADIEDFK	
78 kDa glucose-regulated protein	2163,9848	-0,0482	114,71		34	IEIESFYEGEDFSETLTR	
78 kDa glucose-regulated protein	2174,9855	0,0091	120,11	33	35	LYGSAGPPPTGEEDTAEKDEL	
Tubulin beta-4 chain	1038,5862	0,003	49,63		31	YLTVAAVFR	
Tubulin beta-4 chain	1129,588	-0,0212	96,81		34	FPGQLNADLR	
Tubulin beta-4 chain	1941,9796	-0,0006	88,47		35	GHYTEGAELVDAVLDVVR	
Tubulin beta-4 chain	2707,331	0,0979	125,06	33	34	LTTPTYGDLNHLVSATMSGVTTCLR	
Tubulin beta-4 chain	2797,3361	-0,0461	110,72		36	SGPFGQIFRPDNFVFGQSGAGNNWAK	
Importin subunit alpha-2	824,4545	-0,0379	46,45		33	FVSFLGR	
Importin subunit alpha-2	1376,6871	-0,0552	56,37		33	NLTWTLSNLCR	
Importin subunit alpha-2	1548,9239	-0,0472	80,25		32	LLGASELPIVTPALR	
Importin subunit alpha-2	1875,8599	-0,0123	94,24		34	NVSSFPDDATSPLQENR	
Importin subunit alpha-2	1942,0887	0,0406	85,82	26	29	NPAPPIDAVEQILPTLVR	
Importin subunit alpha-2	1955,8262	-0,0021	35,43		33	NVSSFPDDATSPLQENR	Phospho (S)
Importin subunit alpha-2	2184,2266	0,043	109,38		28	NKNPAPPIDAVEQILPTLVR	
Putative heat shock protein HSP 90-beta-	1235,6299	0,0151	57,08		33	RAPFDLFENK	

Putative heat shock protein HSP 90-beta-	1248,6098	-0,0366	63,04		33	EQVANSAFVER	
Putative heat shock protein HSP 90-beta-	1310,5626	-0,0276	94,05		31	EDQTEYLEER	
Putative heat shock protein HSP 90-beta-	1363,7248	-0,0132	39,51		34	RAPFDLFENKK	
Putative heat shock protein HSP 90-beta-	1807,9509	0,0102	130,1		34	HSQFIGYPITLYLEK	
Putative heat shock protein HSP 90-beta-	2014,0371	-0,0175	85,05		35	VILHLKEDQTEYLEER	
Putative tubulin-like protein alpha-4B	1409,7667	-0,005	46,11		33	QIFHPEQLITGK	
Putative tubulin-like protein alpha-4B	1486,8719	0,0056	127,37		30	LISQIVSSITASLR	
Stress-70 protein, mitochondrial	1289,6728	-0,0232	37,63	31	34	VQQTVQDLFGR	
Stress-70 protein, mitochondrial	1360,7351	-0,0137	97,33		34	AQFEGIVTDLIR	
Stress-70 protein, mitochondrial	1693,8424	-0,0323	75,21		35	NAVITVPAYFNDSQR	
Stress-70 protein, mitochondrial	2054,9545	0,0044	102,07		35	STNGDTFLGGEDFDQALLR	
Tubulin beta-6 chain	1129,588	-0,0212	96,81		34	FPGQLNADLR	
Tubulin beta-6 chain	1941,9796	-0,0006	88,47		35	GHYTEGAELVDAVLDVVR	
DnaJ homolog subfamily A member 1	1867,751	-0,0394	90,98		30	HYNGEAYEDDEHHPR	
DnaJ homolog subfamily A member 1	2128,0913	0,0786	119,12		33	NVVHQLSVTLEDLYNGATR	
Putative heat shock protein HSP 90-beta 2	900,5181	-0,0278	43,03		34	TKPIWTR	
Putative heat shock protein HSP 90-beta 2	1807,9509	0,0102	130,1		34	HSQFLGYPITLYLEK	
Heat shock 70 kDa protein 4 OS=Homo sapiens GN=HSPA4 PE=1 SV=4	1417,6183	-0,0105	48,49	31	32	NAVEEYVYEMR	Oxidation (M)
Heat shock 70 kDa protein 4	1494,695	-0,0503	52,53		33	AGGIETIANEYSDR	
Heat shock 70 kDa protein 4	1734,9192	-0,0841	77,88	31	35	EFSITDVVPYPISLR	
Heat shock 70 kDa protein 4	2140,1164	-0,1049	62,83		35	TSTVDLPIENQLLWQIDR	

4F2 cell-surface antigen heavy chain	1370,8286	0,0074	69,38		28	LLTSFLPAQLLR	
4F2 cell-surface antigen heavy chain	1504,7522	-0,0355	53,03		34	GQSEDPGSLLSLFR	
4F2 cell-surface antigen heavy chain	2291,1487	-0,1015	94,89	30	35	SLLHGDFHAFSAGPGLFSYIR	
Reticulocalbin-2	2213,96	-0,1103	100,12	29	31	VIDFDENTALDDAEEESFR	
Reticulocalbin-2	2640,2442	-0,1032	134,9		35	LSEEEILENPDLFLTSEATDYGR	
Putative tubulin beta chain-like protein ENSP00000290377	1076,525	-0,0283	42,35		32	IREEYPDR	
Putative tubulin beta chain-like protein ENSP00000290377	1129,588	-0,0212	96,81		34	FPGQLNADLR	
Tubulin beta-1 chain	1076,525	-0,0283	42,35		32	IREEYPDR	
Tubulin beta-1 chain OS=Homo sapiens GN=TUBB1 PE=1 SV=1	1129,588	-0,0212	96,81		34	FPGQLNADLR	
Translocon-associated protein subunit delta	1404,6561	0,0074	65,73	28	34	FFDEESYSLLR	
Translocon-associated protein subunit delta	2165,1117	0,0094	104,54		35	NNEDISIIPPLFTVSVDHR	
DnaJ homolog subfamily A member 4	2128,1164	-0,0549	101,92		35	NVVHQLSVTLEDLYNGVTK	
Heat shock protein 105 kDa	1417,6514	0,0383	51,42		34	NAVEEYVYEFR	
Heat shock protein 105 kDa	1605,7358	0,044	46,17		35	FICEQDHQNFLR	
Heat shock protein 105 kDa	1781,9927	0,0524	71,67		29	IEVPLYSLLEQTHLK	
Heat shock protein 75 kDa, mitochondrial	1512,7784	-0,0072	73,2	32	34	GVVDSEDIPLNLSR	
Casein kinase II subunit alpha	1864,0418	-0,012	39,1	26	32	GGPNIITLADIVKDPVSR	
Casein kinase II subunit alpha	2306,112	-0,0542	60,84	29	36	LIDWGLAEFYHPGQEYNVR	
Transitional endoplasmic reticulum ATPase	1809,9876	0,0656	76,3	27	29	NAPAIIFIDELDAIAPK	
Casein kinase II subunit beta	1083,5938	-0,0316	61,47		33	RPANQFVPR	

Casein kinase II subunit beta	1389,5772	0,0111	57,75		32	YQQGDFGYCPR	
ATP synthase subunit alpha, mitochondrial	2337,1601	-0,0262	101,12		36	EVAAFAQFGSDLDAATQQLLSR	
Calnexin	1507,6667	-0,0164	59,27		33	AEEDEILNRSPR	Phospho (S)
Calnexin OS=Homo sapiens GN=CANX PE=1 SV=2	1862,9203	-0,0848	64,54	31	34	KIPNPDFFEDLEPFR	
D-3-phosphoglycerate dehydrogenase	985,596	0,0094	55,45		31	DLPLLLFR	
Protein transport protein Sec16A	1246,6744	0,033	44,2	22	32	KAPPPPPTSMPK	
Membrane-associated progesterone receptor component 1	1515,7147	-0,0105	83,33	30	34	FYGPEGPYGVFAGR	
Myosin-14	1957,9929	-0,0671	57,19		35	AELSSLQTARQEGEQRR	
Importin subunit alpha-3	2035,8858	0,0066	75,36	31	34	NVPQEESLEDSDVDADFK	
Squalene synthetase	2039,9324	-0,0845	74,94		33	LFSASEFEDPLVGEDTER	
Importin subunit alpha-8	1376,6871	-0,0552	56,37		33	NITWTLSNLCR	
Pyruvate carboxylase, mitochondrial	1362,6667	-0,0546	73,56		33	IAEEFEVELER	
Tubulin alpha chain-like 3	1022,4417	0,0103	53,1		31	EDAANNYAR	
Elongation factor 1-alpha 1	1024,603	-0,0107	54,15		30	IGGIGTVPVGR	
Probable Xaa-Pro aminopeptidase 3	814,5065	-0,0056	50,11		33	AILFVPR	
Protein AF-9	723,4028	0,0232	45,47		28	LELGHR	
Protein AF-9	771,3875	0,012	37,75		31	NKEEPR	
ADP/ATP translocase 3	1120,5665	-0,0482	61,47		33	EQGVLSFWR	
Phosphatidylinositol-5-phosphate 4-kinase type-2 gamma	1372,6598	0,0156	45,35		34	FKEYCPQVFR	
Phosphatidylinositol-5-phosphate 4-kinase type-2 gamma	1793,8696	0,0633	41,22		34	HGAGAEISTVHPEQYAK	
45 kDa calcium-binding protein	1319,563	-0,004	58,14		32	DLGGFDEDAEPR	
Protein spinster homolog 1	1611,7376	-0,0111	57,18	32	34	SEEPEVPDQEGLQR	

STIP1 homology and U box-containing protein 1	1284,5921	0,0513	51,72	30	33	KYPEAAACYGR	
Phosphatidylinositol-5-phosphate 4-kinase type-2 alpha	1174,6387	-0,0289	45,37	27	34	FLDFIGHILT	
Elongation factor 2	1443,7609	0,0522	45,85	28	32	EGIPALDNFLDKL	
Protein CASC5	1726,7788	-0,0365	39,93		33	ERTDRPNFELSQR	Phospho (S)
Leucine zipper protein 1	1002,5345	0,0171	44,51		33	ELKSEIER	
Proteasome subunit beta type-	1305,7445	-0,018	44,17	31	32	FILNLPTFSVR	
Regulator of nonsense transcripts 2	814,5065	-0,0295	38,7		33	ALFIVPR	
Tight junction-associated protein 1	1428,6732	-0,033	42,91	29	33	LEQEEPLTDAER	
POTE ankyrin domain family member E	1514,7419	-0,0092	42,07	33	34	IWHHTFYNELR	
Actin, cytoplasmic 1	1514,7419	-0,0092	42,07	33	34	IWHHTFYNELR	
Beta-actin-like protein 3	1514,7419	-0,0092	42,07	33	34	IWHHTFYNELR	
Band 4.1-like protein	1377,7445	0,0389	36,81		31	NLPWLFTFNVK	
Probable ATP-dependent RNA helicase DDX6	1230,6972	-0,0175	40,48	25	33	SGAYLIPLLER	
DnaJ homolog subfamily C member 1	1002,5094	0,0028	40,35		33	LGASEKNER	
Splicing factor 3B subunit 1	914,5185	-0,0372	39,58		33	VGAAEIISR	
Adipocyte enhancer-binding protein 1	789,3868	0,0363	38,09		34	KEDSSPK	
Elongation factor 1-gamma	1706,8641	-0,03	38,03	27	35	VLSAPPHFHFGQTNR	
Keratin, type II cytoskeletal 1	2382,9447	0,042	33,02	19	33	GGGGGGYGSGGSSYGSGGGSYGSGGGGGGGR	
Protein FAM184A	1002,4982	0,014	37,6		33	LSQLEEER	
Ataxin-2-like protein	2289,1713	-0,0431	36,93	31	36	TLSSPSNRPSGETSVPPPPAVGR	
DnaJ homolog subfamily B member 12	1958,1214	-0,0902	36,64		34	ALRFLEKAQRLYPTPR	
Microtubule-associated protein 1A	1118,5455	0,0315	35,48		34	DKDLEEKDK	
Serine/arginine repetitive matrix protein	1183,6462	-0,0586	35,92	26	33	RYSPSPPPKR	

1							
Poly(ADP-ribose) glycohydrolase	1366,6629	-0,0452	34,41	30	33	FQARDADIEFR	
Histone H1oo	941,5116	-0,0158	34,2		32	KMAPATAPR	
UPF0419 protein C12orf48	1248,554	0,0614	34,17	25	34	GYAPPPSDPLR	Phospho (S)