

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

## 4,800

Open access books available

## 122,000

International authors and editors

## 135M

Downloads

Our authors are among the

## 154

Countries delivered to

## TOP 1%

most cited scientists

## 12.2%

Contributors from top 500 universities

**WEB OF SCIENCE™**Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.

For more information visit [www.intechopen.com](http://www.intechopen.com)

# Quantitative Proteomics for Investigation of Secreted Factors: Focus on Muscle Secretome

Jeanette Henningsen, Blagoy Blagoev and Irina Kratchmarova  
*Department of Biochemistry and Molecular Biology,  
University of Southern Denmark, Odense M  
Denmark*

## 1. Introduction

The response of cells to even slight changes in the cellular microenvironment determines the reaction of the whole organism and its ability to adapt to macroenvironmental alterations. Generally, it is well recognized that the communication between cells, tissues, and organs is critical for the maintenance of the entire body homeostasis. The different cell types that build the various organs and tissues have an enormous potential to produce proteins that once secreted in the extracellular space exert their action in an auto-, para- and/or endocrine manner. It is estimated that out of the total 20,500 protein-coding genes in human, approximately 10% encode secreted proteins (Clamp et al., 2007; Skalnikova et al., 2011). The separate and combinatorial action of these ~2200 secreted proteins can influence the biology not only at adjacent sites but also have a clear effect on the whole organism (Lin et al., 2008). The secreted factors, which can range from large proteins to short peptides, are divided into different groups or classes according to their structural properties and function. The prototypical secreted proteins are represented by the group of proteins found in the blood stream and other body fluids, the components of the extracellular matrix (ECM) and enzymes released in the intestine and stomach. An intriguing group of secreted factors comprise cell surface receptor ligands, such as hormones, growth factors, and cytokines. These proteins can exert their actions either on a limited number of responsive tissues or can act on virtually all cell types dependent on the expression of their specific receptors. It is essential to decipher in depth the signaling events that are triggered by the various hormones and growth factors to understand the general mechanisms of the biological processes that occur in a strictly controlled fashion in both space and time. The processes that secreted factors influence and directly regulate range from cellular differentiation, growth and survival to apoptosis, autophagy, and ageing. In addition, a growth factor can often exert a divergent and even opposite effect depending on the cell type and cellular state. Taken in consideration the role of secreted factors in directing biological processes, malfunction of the signaling cascades orchestrated by secreted factors can have severe consequences and lead to development of a series of complicated diseases and disorders (Flier, 2001; Pedersen, 2009; Walsh, 2009). Therefore, a comprehensive characterization of secreted molecules by different cellular subtypes, tissues, and organs can contribute to the elucidation of the physiological state of a given organism and to the determination of the

malfunction in diseased stages. Analyzing on a large scale and in an unbiased manner the secretome of any given cell type or tissue, which comprise a unique combination of growth factors, hormones, cytokines, inhibitory factors, and components of the extracellular environment, has become a whole distinct research field. Although still challenging, this endeavor may ultimately prove beneficial for improving human health as it can accelerate the bridging of basic research and applied medicine.

## 2. Proteomics

Proteomics has many sides and it is often difficult to combine the different aspects that can define or characterize this broad topic. The term proteomics was introduced in 1995, describing the entire set of proteins expressed by a given cell, tissue, or organism (Wasinger et al., 1995). At present, proteomics is defined as large scale studies of the proteomes that encompass protein expression, folding, and localization. It also includes functional analyses of large complexes within a cell, tissue, or organism as well as comparison of different proteomes. Some of the different aspects of proteomics include analysis of body fluids, defining proteomes of pathogens, investigation of tissue proteomes, characterization of signaling pathways and the effects of inhibitors and drugs. The term systems biology was also introduced to describe the incorporation of genomics, metabolomics, and proteomics data for creation of dynamic networks of interacting molecules at a system level. Typically, such studies involve following the changes in protein profiles in response to changes in the environment and determination of combined action of diverse signaling networks that lead to a differential outcome for the living organism. Obtaining and combining information for such networks is of particular importance when investigating the role of secreted factors in the regulation of major signaling events in any given cell or tissue. Functional quantitative mass spectrometry-based proteomics (QMSP) is a powerful approach for creation of maps that describe the differential expression and dynamic changes of secretomes. Correlation of these results with clinics can help resolve some of the still missing links in the development of different syndromes. In this review chapter, we focus on the latest advances in QMSP for the investigation of secreted factors and we discuss some of the issues and challenges that remain to be unveiled.

### 2.1 Quantitative mass spectrometry-based proteomics

The fast development of QMSP techniques added yet another dimension to the proteomic research, namely the ability to follow differences and changes of the proteomes in space and time (Aebersold and Mann, 2003; Cox and Mann, 2007; Dengjel et al., 2009). QMSP permits observation and investigation of a combination of events and interplay of pathways involving hundreds of molecules that lead to a defined outcome for the cell. It facilitates determination of even slight changes in protein expression or post-translational modifications as a result of a drug treatment, changes in the cellular environment or alterations in the total body homeostasis. Up to date, QMSP is the only available approach that can, with high confidence and in a high throughput manner, generate and combine data for the spatial and temporal order of events that take place in a cell directly at protein level in order to decipher dynamic complex processes (Dengjel et al., 2009; Rigbolt and Blagoev, 2010; Walther and Mann, 2010). There are two main QMSP strategies for relative quantitation based either on the use of stable isotopes or the label-free approach for

quantitation of changes in protein abundance (Ong and Mann, 2005; Schulze and Usadel, 2010; Walther and Mann, 2010).

### 2.1.1 Quantitation without stable isotopes

Quantitation without stable isotopes generally encompasses a gel electrophoresis approach or a chromatography-based approach. In the gel-based approach one-dimensional or two-dimensional gel electrophoresis is used as a mean of resolving the proteins from complex mixtures. This is followed by visualization of the protein bands or spots using different types of stains or fluorescent dyes. Typically, the protein samples originating from different cellular stages are separated on a gel and then the bands or spots that show distinct changes are excised, digested with proteases and identified by mass spectrometry. A major disadvantage of two dimensional gel electrophoresis is the relatively low dynamic range and inefficient access into the gel of high or very low molecular weight proteins. This results in the identification of mainly high abundant molecules, such as cytoskeletal proteins and highly expressed metabolic enzymes (Gygi et al., 2000). Reducing the complexity of the sample can at least partially overcome such limitation. The introduction of the difference in gel electrophoresis (DIGE) approach, which allows proteins from two different samples to be separated on the same gel, led to improved quantitative accuracy of this gel-based approach (Unlu et al., 1997).

The chromatography approach can be divided into two groups, namely peptide-based methods and protein-based methods. The peptide-based strategy relies on comparing the signal intensity of a peptide originating from one sample to the signal intensity of the same peptide originating from a different sample. The extracted ion chromatogram (XIC) for every peptide can be derived from the liquid chromatography profile of the two individual samples during the analysis by the mass spectrometer and the samples can thereby be compared quantitatively. Furthermore, a method called protein correlation profiling was established, where the total ion chromatograms of different samples are aligned and quantitative comparison of samples is then based on both retention time and accurate mass of the peptides. The relative protein quantitation is based on the fact that the peak areas obtained from liquid chromatography mass spectrometry correlate to the relative concentration of the protein in the sample (Andersen et al., 2003; Ong and Mann, 2005). It has been used to obtain semi quantitative data in complex mixtures such as human sera (Chelius and Bondarenko, 2002). A disadvantage is that it is only partially quantitative and requires highly reliable and reproducible analysis of the samples.

Another label-free mass spectrometry-based approach used to retrieve quantitative measurements is based on spectral count. The "spectral counting" method uses the numbers of peptide identification spectra obtained for each protein as representation of the protein abundance in a mixture (Liu et al., 2004). One disadvantage of the spectral count method is that it is biased toward high abundant proteins since they can mask or suppress the low-abundance ones in the sample, which is a key issue when analyzing e. g. plasma samples. The two label-free methods for quantitation, using either peptide ion intensities or spectral counts, are becoming increasingly popular, since they are simpler than the isotope-based strategy, despite being less accurate. In addition, both methods require very good reproducibility between the different liquid chromatography tandem mass spectrometry (LC-MS/MS) runs, high accuracy measurements and higher number of replica analyses. In general, the label-free approaches are widely applicable but the methods using stable

isotope labels result in better accuracy of quantitation (Lundgren et al., 2010; Schulze and Usadel, 2010).

### 2.1.2 Quantitation using stable isotopes

The quantitative mass spectrometry-based methods utilizing stable isotopes can be achieved either by *in vivo* metabolic labeling or *in vitro* biochemical methods. The principle of the two labeling strategies is the generation of peptides labeled with stable isotopes that differ in mass from the unlabeled peptides making it possible to distinguish them within the same spectrum.

#### 2.1.2.1 Chemical labeling strategies

The prototype of the chemical modification-based methodology for quantitation of protein is the isotope coded affinity tag (ICAT) that binds to cysteine residues (Gygi et al., 1999). It employs usage of two isotopically labeled tags - one light and one heavy, which contains eight deuterium atoms, to distinctly label the peptides originating from two separate samples. The peptides originating from one sample can thereby be distinguished from the second sample, since the heavier tag will result in a mass shift readily observable in the mass spectrum. One of the advantages of ICAT is the presence of a biotin group in the light and heavy tags allowing selective enrichment of the labeled peptides using avidin affinity chromatography, thus reducing greatly the complexity of the mixture.

ICAT has been applied to a variety of cell culture and tissue samples and has been demonstrated as a reliable and relatively easy applicable method for performing QMSP analysis. Among other applications, ICAT has been used to investigate differential expression profiles of microsomal proteins from *naive* and *in vitro*-differentiated human myeloid leukemia cells, secreted proteins during osteoclast differentiation, the dynamic changes of transcription factors during erythroid differentiation as well as comparison of livers of mice treated with different peroxisome proliferator-activated receptor agonists (Brand et al., 2004; Han et al., 2001; Kubota et al., 2003; Tian et al., 2004). Disadvantages of the ICAT strategy are that it targets only the cysteine containing peptides and the retention times of the light and heavy form during chromatographic separation are altered due to the presence of the deuterium atoms. To overcome some of those problems, a cleavable <sup>12</sup>C- and <sup>13</sup>C-based reagent (cICAT) has been developed, which has an improved peptide co-elution profile during the liquid chromatography separation and increased recovery after enrichment of the labeled peptides (Yi et al., 2005).

Several other chemical labeling strategies have been developed over the recent years. Probably the most popular of those being the isobaric tags for relative and absolute quantitation (iTRAQ) where the isobaric chemical groups are attached to the primary amine groups of the peptides. With iTRAQ, up to eight different conditions can be compared simultaneously since eight distinct isobaric tags for labeling are currently available. The quantitation is based on the intensities of the isotopically distinct fragments derived from the corresponding isobaric tags obtained in the peptide fragmentation spectrum. This is the main advantage of the method but it can also be a disadvantage since often a single fragment spectrum per peptide is available, thereby compromising the accuracy of quantitation (Ross et al., 2004).

#### 2.1.2.2 Metabolic labeling

The metabolic labeling strategies rely on the incorporation of a stable isotope in proteins, while they are being *de novo* synthesized in the cell. In contrast to the standard radioactivity-

based assays, the stable isotope is fully incorporated thereby encoding the whole proteome. There are two means of introducing the stable isotope using either media containing  $^{15}\text{N}$  labeled ammonium sulfate or media with the addition of a stable isotope labeled amino acid. The  $^{15}\text{N}$  labeling strategy has been used for quantitative analysis of protein phosphorylation in bacteria and a mouse melanoma cell line (Conrads et al., 2001; Oda et al., 1999). Additionally, entire organisms have been metabolically labeled using the  $^{15}\text{N}$  strategy, including bacteria (*E. coli* and *Deinococcus*), *C. elegans*, *D. melanogaster*, and rat (Conrads et al., 2001; Krijgsveld et al., 2003; Wu et al., 2004).

Stable Isotope Labeling by Amino Acids in Cell Culture (SILAC) is an accurate and resourceful quantitative proteomics platform, that in combination with high speed and accuracy mass spectrometry allows detailed characterization of complex biological systems (Ong et al., 2002; Ong et al., 2003). It involves usage of heavy non-radioactive stable isotope-labeled amino acids, which are incorporated directly into the newly synthesized proteins of the cell. After SILAC labeling, the entire proteome of a given cell population becomes encoded either with a light or heavier version of the same amino acid, thereby enabling direct comparison and quantitation using mass spectrometry. With SILAC, the “light” and “heavy” samples can be mixed in equal ratios at the initial stages of the workflow, which can include subsequent protein purification, interaction assay or other manipulation of the mixed sample. Combining samples prior to any further sample preparation represents a tremendous advantage, since it results in reduced quantitation errors introduced by differences in individual sample handling. Major strength of the SILAC method is the ability to discriminate true interaction partners from background, when investigating functional protein-protein interactions (Blagoev et al., 2003; Dengjel et al., 2010). Therefore, it facilitates investigation of cellular signaling cascades and creation of reliable protein interaction networks, which represents one of the biggest challenges in the field of system biology (Blagoev et al., 2004; Dengjel et al., 2009; Kratchmarova et al., 2005; Olsen et al., 2006; Osinalde et al., 2011). In addition, SILAC is invaluable for the investigation of secreted factors since it allows the distinction of specific proteins released by the cells to the extracellular environment from contaminating proteins like keratins and serum derived factors that originate from cell culture media supplements (Henningsen et al., 2010). One potential disadvantage with the SILAC protocol arises from cultures of primary cells, which usually require specific growth media with a defined formulation. Furthermore, such cells have limited division capacity in culture, whereas at least 5 population doublings are required for complete SILAC encoding of the entire proteome. Nevertheless, SILAC-based analyses have been successfully extended to include microorganisms, entire mice, and quantitation of proteins in tumor biopsies (Geiger et al., 2010; Kruger et al., 2008; Soufi et al., 2010). It was also utilized for the quantitative analyses of proteins released by omental adipose tissue explants (Alvarez-Llamas et al., 2007).

### 3. Application of QMSP for investigation of secreted proteins

Analysis of secreted proteins using QMSP allows in depth characterization of different cellular systems that secrete auto-, para-, and endocrine factors, which can influence the entire body homeostasis. Investigation of cellular models such as adult and embryonic stem cells, cells originating from a diseased state, immortalized cells representing various models for functional abnormalities, extends the knowledge of how changes in secretomes contribute to various types of human disorders. It also enables determination and discovery

of new roads of tissue cross talk and interaction. QMSP has been applied to study secretomes of a variety of cell types and tissues including adipose cells and tissues, mouse embryonic fibroblasts, astrocytes, mesenchymal stem cells, neuronal progenitor cells, kidney, and endothelial cells (Skalnikova et al., 2011). Although, there have been several proteomics reports describing the secretory function of cells from mesenchymal origin, the role of the muscle secretome has remained elusive. A limited number of studies so far have employed mass spectrometry to elucidate the secretory function of the skeletal muscle. In a study presented by Chan and coworkers condition media (CM) was collected from differentiated C2C12 myotubes at day 5 of differentiation and analyzed by 1D-gel electrophoresis combined with matrix-assisted laser desorption/ ionization tandem mass spectrometry (MALDI-MS/MS) (Chan et al., 2007). This work led to the identification of 80 proteins released from skeletal muscle of which 27 were classified as secreted proteins based on literature searches. In another study isolated primary human skeletal muscle cells were SILAC-labeled ( $^{13}\text{C}_6\text{-Lys}$ ) to make a quantitative evaluation of muscle secreted proteins between extreme obese and lean women (Hittel et al., 2009). Assessment of the identified proteins based on published literature and the Swiss-Prot database revealed 28 secreted proteins from 42 identified skeletal muscle proteins. Interestingly, the secretion of myostatin, a negative regulator of skeletal muscle growth and development but also implicated in metabolic homeostasis, was found to be markedly upregulated in extreme obesity cases. Subsequently, Yoon and colleagues presented a study investigating the effects of insulin on the secretory profile of differentiated myotubes (Yoon et al., 2009). The authors combined off-line reverse-phased HPLC fractionation with LC-MS/MS and identified 153 secreted proteins from rat L6 myotubes. Based on spectral count quantitation, 33 of these proteins were classified as differentially regulated in response to insulin. The list of secreted proteins was extracted from a total list of 254 identified proteins using three different prediction tools, Gene Ontology, SignalP, and SecretomeP. In two more recent studies, a total of 108 secreted proteins by skeletal muscle cells were identified (Chan et al., 2011; Norheim et al., 2011).

We have developed a general quantitative proteomics approach for investigation of secreted factors released by skeletal muscle cells during the course of muscle differentiation. The method utilizes a combination of SILAC labeling and advanced mass spectrometry (Fig. 1) (Henningsson et al., 2010).

Triple encoding SILAC, (Blagoev et al., 2004) was applied to investigate protein secretion at three different time points during the course of C2C12 differentiation. Initial evaluation of the differentiation protocol with SILAC-labeled cells demonstrated the formation of a high number of multinucleated myotubes and increased expression of different muscle-specific proteins. The use of three different versions of each labeled amino acid enabled the comparison of the secretome at three different time points (day 0, day 2, and day 5) during skeletal muscle differentiation. Furthermore, cells were cultured using both labeled arginine and lysine, since trypsin, which cleaves solely C-terminal to arginine and lysine, was used for in-gel digestion (Olsen et al., 2004). This “double-triple” labeling with isotopic forms of both arginine and lysine ensures that every tryptic peptide, except the C-terminal peptides of the proteins, contains at least one labeled residue and can therefore be used for quantitation (Blagoev and Mann, 2006). This increases the probability of positive protein identification and accuracy of quantitation due to the increased number of labeled peptides. It is noteworthy that under normal culture conditions cells are grown in the presence of fetal

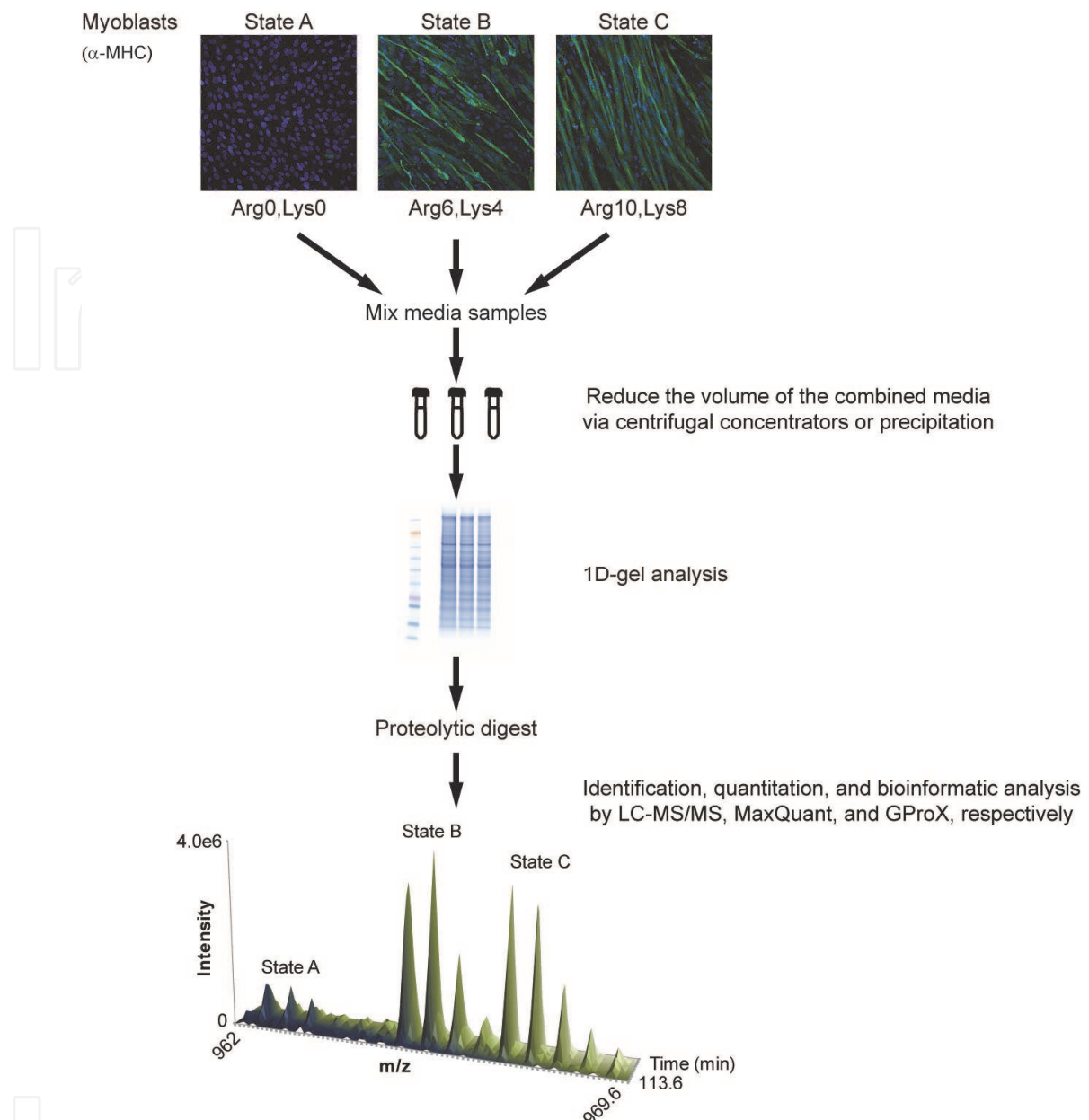


Fig. 1. General outline of QMSP for the identification and quantitation of secreted proteins.

bovine serum (FBS), but the SILAC protocol requires the use of dialyzed sera to prevent the presence of unlabeled sera-derived amino acids, which would ultimately result in inaccurate quantitation. Commercially available dFBS is dialyzed utilizing 10 kDa molecular weight cut off (MWCO) filters to remove any amino acids. Unfortunately, this also leads to the reduction of low-molecular weight proteins (<10 kDa) including certain growth factors, hormones, and cytokines, that may be needed for the growth and maintenance of certain cells. Therefore, dFBS is not compatible with all cell types and slower growth rate is observed in some cases. Ultimately, dialysis with MWCO 1,000 Da could be sufficient to remove amino acids, but it is more costly.

The myoblasts were cultured in SILAC media for at least 5 passages to ensure complete incorporation of labeled amino acids into the proteome. Before the collection of conditioned medium (CM), cells were washed and starved for 12 hours in sera-free



medium to minimize the presence of sera proteins, that would interfere with the subsequent mass spectrometry (MS)-analysis. CM was collected from myoblasts on day 0 and during conversion of myoblasts into myotubes at day 2 and day 5, followed by filtration using 0.2  $\mu\text{m}$  filters to remove any floating cells or cell debris, thereby reducing the risk of contaminating samples with intracellular proteins. The CM, collected from the three time points of differentiation was combined in a 1:1:1 ratio according to measured protein concentration. Subsequently, the pool of CM was concentrated by ultrafiltration using Vivaspin columns, MWCO 3,000 Da to ensure that proteins were retained in the concentrate. To reduce sample complexity, thereby effectively increasing the dynamic range of the MS-analysis, concentrated muscle-derived proteins were separated by size using 1D- gel electrophoresis. The excised gel bands were subjected to in-gel digestion and analyzed via LC-MS/MS using an linear ion trap (LTQ)-Orbitrap mass spectrometer followed by processing of the obtained data with the MaxQuant software (Box 2) (Cox and Mann, 2008; Cox et al., 2009). The described strategy resulted in the identification of 635 putatively secreted proteins by skeletal myoblasts based on the GO term “extracellular” and signal peptide prediction inbuilt in the MaxQuant and ProteinCenter. The commercially available database, ProteinCenter, ([www. Proxeon.com](http://www.Proxeon.com)) utilizes annotation from all major protein sequence databases including Swiss-Prot, NCBI, and Ensembl. It allows analysis of large scale proteomic studies to isolate putatively secreted factors from the total list of identified proteins. The obtained identification list of IPI numbers is filtered and extracted according to the category “extracellular” within GO term cellular component. Then, the remaining proteins are filtered using a signal peptide predictor incorporated into the ProteinCenter platform, the PrediSi algorithm. Using the SILAC strategy, 624 secreted proteins were quantitatively evaluated during the course of skeletal muscle differentiation. Proteins already known to be secreted by skeletal muscle were identified, in addition to many novel proteins not previously shown to be secreted by skeletal myoblasts. Characterization of identified secreted proteins according to GO-annotations demonstrated proteins involved in many different cellular processes including proliferation, differentiation, ECM reorganization, metabolic processes, and angiogenesis. According to the statistical analyses provided by MaxQuant, 188 secreted proteins were found to be dynamically regulated during skeletal myogenesis suggesting their regulatory involvement in skeletal muscle development, which could occur both in autocrine and paracrine manner. In a follow up study, focused on comprehensive characterization of the low abundant low molecular weight fraction of proteins secreted by muscle cells, application of triple encoding SILAC resulted in the generation of quantitative profiles of 59 growth factors and cytokines, including nine classical chemokines (Henningsen et al., 2011).

The depicted triple encoding SILAC strategy led to the characterization of the muscle secretome and creation of dynamic secretion profiles during the process of muscle differentiation. Among the identified secreted factors, we have found components of the extracellular matrix, such as collagen, fibronectin and SPARC (secreted protein acidic and rich in cysteine), growth factors, including members of the transforming growth factor and insulin-like growth factor families, members of the serpin and matrix metalloproteases classes, chemokines, and modulators. In addition, proteins such as angiopoietin-1, VEGF (Vascular endothelial growth factor), PDGF (Platelet-derived growth factor), and FGF21 (Fibroblast growth factor 21) were identified and quantitated.

Combining these results indicate that muscle is a prominent secretory organ participating actively in the general regulation of body homeostasis. The muscle specific secreted factors exert their effects in local and/or systemic manner. In Henningsen et al., 2010 we have identified and characterized the semaphorins as a new family of muscle secreted proteins. Semaphorins constitute a large family of secreted, GPI-anchored, and transmembrane proteins defined by a conserved semaphorin (sema) domain in their amino terminus (Gherardi et al., 2004; Neufeld and Kessler, 2008; Roth et al., 2009; Serini et al., 2009; Suzuki et al., 2008; Zhou et al., 2008). Initially, semaphorins were described as important regulators of axon guidance during neuronal development. However, an increasing number of studies have recognized the semaphorins as pleiotropic signaling molecules influencing a wide array of biological processes, such as angiogenesis, immune responses, and organ morphogenesis. In addition, semaphorins have also been linked to various pathologies including cancer and different diseases of the nervous system (Neufeld and Kessler, 2008; Roth et al., 2009). Currently, the mammalian semaphorin gene family consists of 20 members and although expression of individual semaphorins has been best described in the nervous system, semaphorins appear to be expressed by most if not all tissues (Yazdani and Terman, 2006). We have identified several members of the semaphorin family belonging to different subfamilies to be secreted from skeletal myoblasts including the soluble Sema3A, Sema3B, Sema3D, Sema3E, the transmembrane Sema4B, Sema4C, and Sema6A, and finally the GPI-linked Sema7A. Examination of the dynamic secretion profiles of the identified semaphorins demonstrated differential secretion of Sema3A, Sema3D, Sema3E, Sema6A, and Sema7A during the course of C2C12 myoblast differentiation. Interestingly, secretion of Sema3A, Sema3E, Sema3D, and Sema6A was markedly enhanced at the early stage of the differentiation, indicating that they may serve a role during the initial phase of the conversion process. In contrast, a gradually increased secretion of Sema7A was observed during differentiation, suggesting that Sema7A plays a role both during early and terminal differentiation. Identification of both the transmembrane and GPI-anchor semaphorins in the media would suggest that they are released from the plasma membrane in a soluble form either by proteolytic shedding, in the case of Sema4 and Sema6, or proteolytic cleavage catalyzed by a phospholipase, in the case of Sema7A. Earlier studies have shown that the enzymatic activity of metalloproteases can generate and modulate the activity of a soluble form of Sema4D (Basile et al., 2007; Elhabazi et al., 2001). Indeed, we did observe an increased secretion of various proteases including MMP-2. Western blot analysis of sema6A in conditioned media collected from C2C12 myoblasts during differentiation supported the idea of Sema6A shedding, as the secreted protein migrated at an apparent molecular weight corresponding to the size of the extracellular domain (approx. 71 kDa) and not to the size of the full-length Sema6A (approx. 113 kDa). Different members of the semaphorin family have been shown to orchestrate the development of different organs including bone, lung, kidney, and the cardiovascular system (Roth et al., 2009; Tamagnone and Giordano, 2006). The number of studies investigating the function of semaphorins in skeletal muscle development and regeneration are more limited. So far, studies have demonstrated an enhanced expression of Sema4C but no alterations of Sema4B expression during C2C12 myogenesis were detected (Ko et al., 2005; Wu et al., 2007). In addition, enhanced expression of Sema4C was also observed *in vivo* in injury-induced skeletal muscle regeneration. Targeted knockdown of Sema4C expression by

siRNA caused inhibition of C2C12 myotube formation, demonstrating that semaphorins could exert an active autocrine/paracrine function in myogenesis. Interestingly, animal models have suggested that semaphorins could be important paracrine factors regulating neurogenesis during skeletal muscle growth, development, and regeneration. A delayed transient increase of Sema3A expression was observed in response to muscle-induced injury (Tatsumi et al., 2009). In addition, a similar delay of Sema3A expression and secretion was seen in isolated skeletal muscle cells in response to HGF, which is an essential factor in muscle growth and regeneration. In our study, we have identified both Sema3A and Sema4C to be released by C2C12 myoblasts during differentiation.

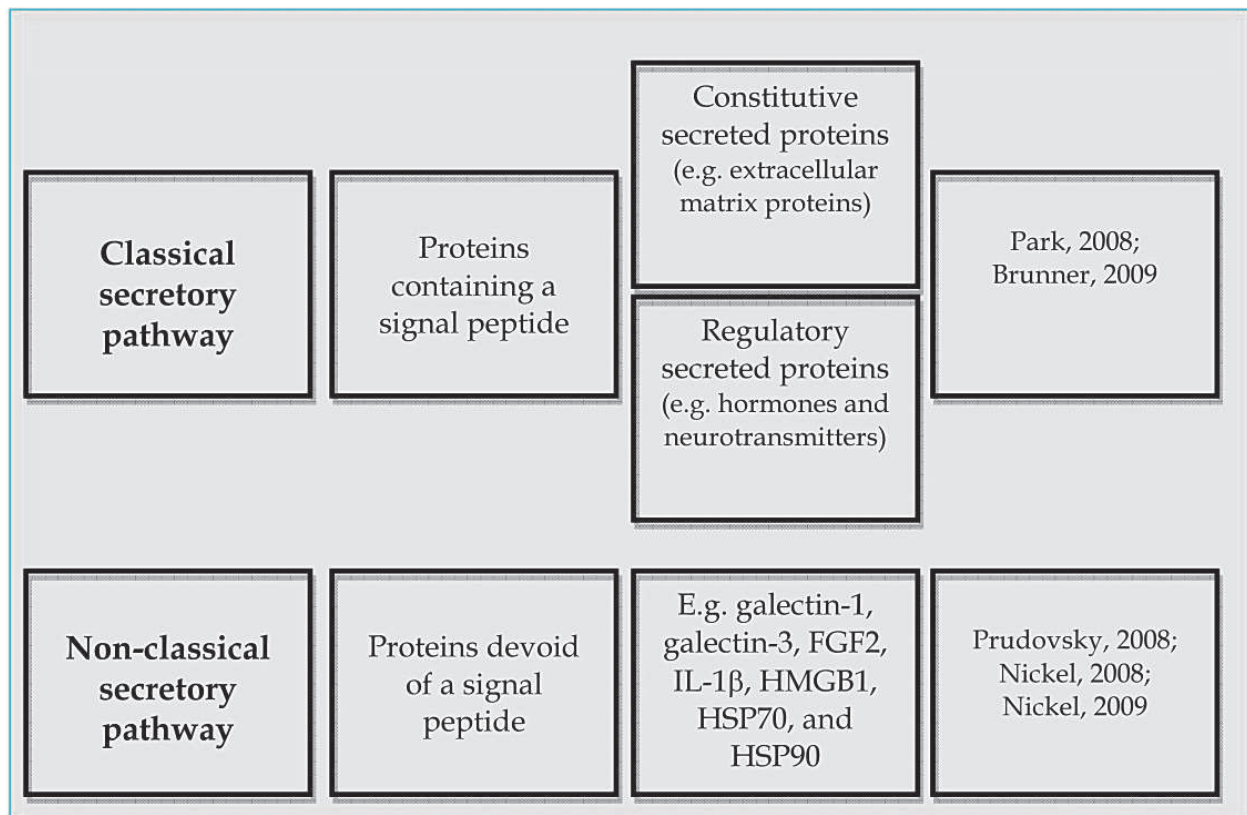
We have analyzed the mRNA and protein expression of selected regulated members of the semaphorin family (Sema3A, Sema3E, Sema6A, and Sema7A) to investigate if their dynamic secretion pattern was regulated by post-transcriptional and post-translational mechanisms. Only minor changes were observed in the mRNA expression of Sema3A, Sema3E, Sema6A, and Sema7A. The mRNA expression of Sema3A and Sema7A remained constant during differentiation, whereas there was a slight decrease and increase in the level of Sema3E and Sema6A, respectively. We found that the high levels of secreted Sema3A and Sema3E at early stage of myotube formation did not reflect the intracellular protein levels of these semaphorins. Expression of Sema3A protein remained constant, whereas a slight decrease was observed of Sema3E protein expression in accordance with the corresponding RNA profile of Sema3A and Sema3E. Moreover, although the intracellular level of Sema7A protein was increased at day 5 of differentiation, it did not correlate with the gradually enhanced level of secreted protein. These findings show that the level of secreted semaphorin proteins can be regulated both by post-transcriptional and post-translational mechanisms. This is in agreement with previous findings in which imperfect correlation between RNA and protein expression was observed (Bonaldi et al., 2008; de Godoy et al., 2008; Kratchmarova et al., 2002). It also emphasizes the necessity to quantitatively investigate protein abundance to understand the functional role exhibited by individual genes and their corresponding proteins. These results clearly illustrate, that when studying the complex nature of the secreted factors it is important to observe both the intracellular level of proteins and their secretion profiles since they might differ due to post-translational modification or modulation of their release via the secretory pathway, turnover rate, and/or processing.

#### **4. Pitfalls of the studies on secreted proteins**

One of the major challenges in secretome studies is the identification and classification of secreted proteins from the total number of identified proteins from the proteomics experiment. Secreted proteins are released in the extracellular space via two routes: the classical and non-classical secretory pathways (Box 1).

##### **4.1 Secretory pathway, classical**

Majority of eukaryotic proteins are secreted by the classical endoplasmic reticulum (ER)-Golgi secretory pathway consisting of a number of distinct membrane-bound compartments interconnected by vesicular traffic (Baines and Zhang, 2007; De Matteis and Luini; Nickel and Rabouille, 2009; Nickel and Wieland, 1998; Park and Loh, 2008; Pelham, 1996; Strating and Martens, 2009). Many basic cellular functions take place in the ER including folding of newly synthesized transmembrane and secretory proteins, lipid synthesis, and the storage



Box 1. Classification of secretory pathways

of high concentrations of calcium ions (Marie et al., 2008; Mayor and Riezman, 2004; Salles et al., 2009; Strating and Martens, 2009). In addition, post-translational modifications of soluble and membrane proteins occur in the ER lumen including oxidation of proline, N-linked glycosylation, proteolytic processing, formation of disulfide bonds, oligomerization, and attachment of a GPI-anchor. Whereas the main functions of the Golgi apparatus include carbohydrate synthesis, O-linked glycosylation, processing, post-translational modification, and sorting both proteins and lipids (De Matteis and Luini, 2008; Marie et al., 2008; Marsh and Howell, 2002). Regardless of their subsequent fate, most proteins containing a N-terminal or internal signal sequence peptide can be targeted to the ER membrane. These include transmembrane proteins destined to reside in the ER, plasma membrane or other organellar membranes as well as soluble proteins destined to the lumen of an organelle or for secretion. With the exception of mitochondria, nuclei, and peroxisomes, all other organelles receive their proteins via the ER. Signal peptides show extreme variations in their length and amino acid composition, but do contain three distinct domains: a positively charged N-terminal region, a hydrophobic core region, typically consisting of at least 6 hydrophobic residues, and C-terminal region of polar uncharged residues (Hiller et al., 2004). Soluble proteins are transported from the Golgi to the cell exterior via the constitutive secretory pathway transporting proteins directly to the cell surface or the regulated secretory pathway in which soluble proteins and other substances are initially stored in secretory vesicles, which release proteins to the extracellular space upon extracellular signals (Brunner et al., 2009; De Matteis and Luini, 2008; Strating and Martens, 2009). The latter pathway only exists in specialized secretory cells including pancreatic  $\beta$ -cell releasing insulin from secretory vesicles, nerve cells, and endocrine cells. The secretory vesicles of the

constitutive or regulated pathways fuse with the plasma membrane and release their contents by exocytosis.

#### **4.2 Secretory pathway, non-classical**

Although most identified extracellular proteins are secreted through the classical secretory pathway, emerging evidence has shown that several soluble proteins are released to the cell exterior via non-classical mechanisms (Nickel and Rabouille, 2009; Nickel and Seedorf, 2008; Prudovsky et al., 2008). For example FGF2 and IL-1 $\beta$ , well known extracellular proteins but lacking a signal peptide, are being secreted by non-classical routes either directly across the membrane or via vesicle intermediates. More specifically, studies investigating IL-1 $\beta$  secretion have demonstrated three alternative routes of extracellular translocation involving activation of caspase 1 and proteolytic processing of IL-1 $\beta$ . IL-1 $\beta$  can be released through (i) microvesicle shedding from the cell surface, (ii) translocation to secretory lysosomes, which upon fusion with the PM releases IL-1 $\beta$  to the cell exterior, and (iii) the caspase 1-IL-1 $\beta$  complex can be captured by endosomal vesicles creating multivesicular bodies that release internal vesicles as exosomes. At present, more than 20 proteins, belonging to different functional groups, have been described to be released to the cell exterior by non-classical pathways, including proteins that mainly function in the extracellular space as well as proteins that serve a role both intracellular and extracellular (Nickel and Seedorf, 2008; Prudovsky et al., 2008). Some of these proteins are constitutively secreted, whereas others are first released upon specific stimulation. Future studies are warranted to understand the biological function and regulation of the many different secretory pathways as well as the number and function of proteins devoid of a signal peptide but released to the extracellular space. Secretion of proteins by alternative pathways, which require interaction with other proteins and/or proteolytic activation, could impose additional levels of regulation to protein secretion (Nickel and Rabouille, 2009; Prudovsky et al., 2008). In addition, alternative secretion of signal peptide containing proteins that bypass the Golgi apparatus, could cause alterations in the structures of post-transcriptional modifications, such as glycosylation, or prevent proper proteolytic processing. This could be a way to modulate the biological activity of secreted proteins under certain physiological conditions.

In addition to the two general pathways of secretion, proteins are also being released to the cell exterior due to apoptosis or cell leakage, thereby contaminating the pool of true secreted proteins. In this regard, the increased performance of MS-instrumentation not only improves the dynamic range for the identification of secreted proteins but also increases the number of identified proteins originating from the intracellular space. One example was presented in the study by Henningsen et al., 2011 focusing on the low molecular weight proteins. The quantitative mass spectrometry analysis resulted in the identification of more than 2000 proteins however, less than 25% of these proteins were predicted to be secreted according to conventional database analyses based on the GO term extracellular and signal peptide prediction. Among the predicted secreted proteins, there were also tubulins, a number of ribosomal proteins, and membrane proteins that are not encountered as being truly secreted. Nevertheless, some of the cytoskeletal and ribosomal proteins have been demonstrated to be part of the exosomes and as such are being released in the extracellular environment. Major part of the exosomes consists of tubulins and Tsg101, which is a well-known exosome marker, was also identified as a secreted protein (Henningsen et al., 2011; They et al., 2002). Different tools are being used to classify the extracellular compartment in various secretome

studies, most commonly being assessment based on literature searches, GO-annotations and/or algorithms predicting secretion by classical (SignalP) or non-classical (SecretomeP) mechanisms (Box 2). Extraction of secreted proteins based on previous reported studies is extremely time-consuming considering the large number of identified proteins by today's advanced MS. In addition, this will only result in the identification of proteins already shown by experimental data to be secreted. Isolation of secreted proteins from a large list of identified proteins can be done combining GO classification as extracellular and/or prediction of a signal peptide. However, all these tools do come with certain restrictions that could lead to either false positive or false negative identifications of secretion status. The presence of a signal peptide is not restricted to extracellular proteins. Proteins destined for other intracellular compartments, such as the ER or Golgi, also contains a signal peptide. In addition, GO terms are assigned according to different parameters, including computational analyses of sequences in addition to experimental data. Again, predictions based on sequence information could result in false positive identification of secreted proteins as well. Prediction tools always have their own limitations and therefore bona fide secreted proteins could also be lost by these tools. Most affected in this regard are the proteins released from cells by non-conventional mechanisms whose number is still low but steadily increasing (Nickel and Rabouille, 2009; Prudovsky et al., 2008).

#### **Open source software and databases**

MaxQuant (<http://maxquant.org>): Advanced software program used as a tool for both protein identification and quantitation (Cox and Mann, 2008; Cox et al., 2009).

SignalP (<http://www.cbs.dtu.dk/services/SignalP>): Predicting the presence of a signal peptide, suggesting proteins could be secreted through the classical pathway (Bendtsen et al., 2004b; Emanuelsson et al., 2007).

SecretomeP (<http://www.cbs.dtu.dk/services/SecretomeP>): Prediction of proteins to be secreted by non-classical mechanisms (Bendtsen et al., 2004a).

QuickGO provided by the Gene Ontology Annotation (GOA) group (<http://www.ebi.ac.uk/GOA>): Bioinformatics resource integrating various databases to assign subcellular localization and functional annotation according to GO terms (Barrell et al., 2009; Binns et al., 2009).

GProX (<http://gprox.sourceforge.net>): Bioinformatics platform, which support the analysis and visualization of large-scale proteomics data (Rigbolt et al., 2011).

#### **Commercially available database**

ProteinCenter (<http://www.proxeon.com>): Software tool combining several data bases to analyze the biological context of complex proteomics experiments.

Box 2. Software and databases commonly used in quantitative mass spectrometry-based proteomics research of secreted proteins

The high number of identified proteins in the QMSP experiments, which are not classified as secreted could be present in the extracellular space (conditioned medium) due to cell leakage or release of intracellular proteins from necrotic or apoptotic cells. The apoptotic process is a normal process that all cells grown in culture undergo at a given time point. However, the number of dead cells is limited since the protocol for collection of media is optimized such as to reduce the number of dying cells. In addition, the collected CM is typically filtrated using a 0.2  $\mu\text{m}$  filter to ensure removal of any dead cells and thereby to reduce contamination from intracellular proteins. The presence of intracellular proteins might be explained by other structures present in the extracellular space such as exosomes and their cargo. Another point is that prediction of a signal peptide by itself does not exclude the possibility that these proteins are in fact located in other intracellular compartments of the cells, such as the ER and Golgi. On the other hand, an increasing number of proteins are being recognized as extracellular despite lacking a signal peptide and thought to be released through non-classical pathways (Nickel and Rabouille, 2009; Prudovsky et al., 2008). In the literature, more than 20 proteins devoid of any signal peptide have been shown to reside in the extracellular space and being released by non-conventional mechanisms. SecretomeP (Box 2) has been designed to predict non-classical secreted proteins (Bendtsen et al., 2004a). For that purpose 13 known human non-classical secreted proteins were analyzed, but no specific sequence motif was identified to characterize non-classical secretion. Instead, the non-classical software for prediction was developed using the multiple sequence features of the 13 non-classical secreted proteins combined with sequence information obtained from more than 3,000 classical secreted proteins. Due to the limited number of identified non-classical proteins, the value of this prediction approach is difficult to assess. Submitting either the murine or human sequence of galectin-1 to SecretomeP resulted in probability score of  $< 0.5$ , thereby exemplifying a false negative identification. Galectin-1 is a well-known extracellular protein lacking a signal peptide but released by non-conventional ways (Hughes, 1999; Sango et al., 2004). On the other hand HMGB1, also not containing a signal peptide and mainly known for its role as a chromatin modifying protein, is also serving an extracellular function suggesting that proteins with typically intracellular functions could also be released to the cell exterior (Bonaldi et al., 2003 2002; Gardella et al., 2002). Future studies will help to elucidate how many proteins deficient in a signal peptide are being released to the extracellular space. Classifying proteins as extracellular based on GO annotation can also lead to false negative or false positive classifications. GO annotation are based both on experimental data but also on computational analysis of sequence information. With the increasing number of biomarker directed studies analyzing biofluids by mass spectrometry, the number of classified secreted proteins is steadily increasing. The increased number of identified secreted proteins could be due to improvements of mass spectrometry technology, which increased the overall sensitivity of protein identification, but could also be artifacts derived from dead cells floating around in the circulation.

In summary, combination of different tools and manually curated data might be beneficial when validating the large lists obtained from QMSP experiments focusing on secreted factors. Nevertheless, release of cellular components can occur via microvesicles and/or exosomes adding to the complexity of secretome studies, thus some of the factors commonly counted as contaminants might be truly secreted ones.

### 4.3 Serum contaminants

Mammalian cell culture models are broadly used in proteomics experiments and often contamination with bovine serum proteins, originating from the serum supplement used for the culturing of cells, is observed in the results from the mass spectrometric analyses. Naturally, when studying the proteins released by specific types of cells, one of the biggest challenges remains the presence of serum proteins that could interfere with the identification of proteins secreted by the cells. Presence of serum proteins in the sample can disrupt the concentration of the CM as well as interfere with the MS analysis, masking the presence of other proteins. It has been estimated that a 10% FBS serum complement, which is commonly added to the culture media, adds 5-6 mg/ml protein to the media and even extensive washing of the cells might not be sufficient to remove the bovine proteins to levels below the detection limit of the mass spectrometer (Bunkenborg et al., 2010). The sera-derived proteins could be falsely identified as proteins being secreted by the cells due to sequence homology between species. One suggested solution to exclude bovine contaminants was based on expanding the database to include both the human and the bovine proteome. Another alternative is to extend the database to include known bovine contaminants as a common contaminant list, which is already incorporated in the data analysis by programs such as MaxQuant (Bunkenborg et al., 2010; Henningsen et al., 2011; Henningsen et al., 2010). In this way, it is possible to exclude the proteins recognized as contaminants from the initial list of identifications and thereby to minimize the number of identifications originating from sera proteins. Nevertheless, the SILAC strategy is so far the best known applicable method to investigate secreted factors by a given type of cells since the metabolic labeling makes it possible to distinguish cell-derived secreted proteins, as these are SILAC labeled, from residual sera proteins. Only proteins that are synthesized in the cells in the presence of the heavy SILAC amino acid will be labeled, thereby these are easily distinguished from sera contaminants, which remain unlabeled. In addition, the accuracy of protein quantitation could also be compromised by the presence of sera proteins in the samples. It is therefore advisable to perform a replica experiment with reverse SILAC labeling strategy (Schulze and Mann, 2004), which is an easy solution to overcome this possible drawback as well as to ensure high quality quantitation of bona fide secreted proteins.

### 4.4 Post-translational modifications of secreted proteins

Glycosylation of secreted proteins is one of the most abundant post-translational modifications (PTM), which affects the proteins folding, stability, and activity. The oligosaccharides are linked to the proteins via asparagine (N-linked) or serine/threonine (O-linked) residues. Enrichment of secreted proteins through their glycan structures is an alternative experimental strategy for the identification of secreted proteins. Various types of enrichment methods have been utilized to capture glycosylated proteins, one of the common approaches being lectin affinity chromatography. In an elegant study by Zielinska et al., the combinatorial use of optimized lectin-based enrichment step, subcellular fractionation, deglycosylation assays, SILAC labeling, advanced mass spectrometry followed by integrative bioinformatic analyses resulted in the identification of 6367 N-glycosylation sites on 2352 proteins in four mouse tissues and blood plasma. N-glycosylation was found to occur exclusively on secreted proteins, on the extracellular face of membrane proteins, and on the luminal side of ER, Golgi apparatus, and lysosomes



(Zielinska et al., 2010). In a complementary study using formalin-fixed paraffin-embedded tissue samples, 1500 N-glycosylation sites were found underlying the increased sensitivity and accuracy of the mass spectrometry-based proteomics for identification of post-translationally modified proteins, even in fixed samples. The comparison of fresh tissue using SILAC-labeled mouse (Kruger et al., 2008) with the paraffin embedded tissue showed no significant qualitative or quantitative differences between these samples, either at protein or peptide level, thereby permitting the use of this methodology in clinical studies (Ostasiewicz et al., 2010).

Hydroxyproline is another type of PTM that is identified predominantly on components of the extracellular matrix and especially on different types of collagens. We have recently reported the identification of 299 unique high-confidence hydroxyproline sites from 48 distinct secreted proteins in muscle cells, representing the largest data set so far on this proline modification. 231 of the modified prolines were located on various collagen types with a large variation of the number of modified sites per individual protein. The number ranged from 1 site on collagen alpha-3 type VI to more than 70 on collagen alpha-1 type I, highlighting the importance of the proline modification in maintaining the structure of the collagens. Motif sequence analysis revealed the canonical motif previously reported for collagen proteins as well as a novel hydroxyproline motif. Modified peptides containing hydroxyproline sites extend over more than 40 proteins including fatty acid binding protein (FABP), several components of the ECM such as SPARC, fibronectin, Lama2, perlecan, and different inhibitors of proteolytic enzymes such as serine protease inhibitors (serpinf1 and serpinh1) and the metalloproteinase inhibitor 2 (Timp2). These results indicate that hydroxyproline could serve as an important secondary modification to confer protein stability and interaction with other secreted proteins.

## 5. Conclusion

The rapid development of MS-instrumentation combined with the advances of quantitative proteomics strategies, such as SILAC, has had a tremendous impact in the analysis of complex biological systems. QMSP is increasingly becoming an essential approach, especially for the characterization of entire secretomes and generation of dynamic quantitative profiles of secreted factors during the course of cellular differentiation or in response to drugs, inhibitors, and modulators. Although there have been a marked improvement of the proteomics strategies to characterize secretomes, many challenges still remain: minimizing the suppressive effect of growth supplements present in the sample during MS analyses, identification of whole tissue secretomes, collection of samples under specific conditions to avoid induction of cell death, identification of low abundant low molecular weight proteins. However, the most critical point presented by secretome studies today is to isolate the bona fide secreted proteins and to validate the obtained results. Integrative approaches that combine highly advanced proteomics methodology followed by biological functional analyses can lead to the creation of secretome maps that underline tissue crosstalk and communication.

## 6. Acknowledgment

This work was supported by a grant from the Novo Nordisk Foundation, the Lundbeck Foundation and the Augustinus Foundation. IK is supported by grants from the Danish

Natural Science Research Council and the Danish Medical Research Council. BB is supported by grant from the Lundbeck Foundation and the Danish Natural Science Research Council.

## 7. References

- Aebersold, R., and M. Mann. 2003. Mass spectrometry-based proteomics. *Nature*. 422:198-207.
- Alvarez-Llamas, G., E. Szalowska, M.P. de Vries, D. Weening, K. Landman, A. Hoek, B.H. Wolffenbuttel, H. Roelofsen, and R.J. Vonk. 2007. Characterization of the human visceral adipose tissue secretome. *Mol. Cell. Proteomics*. 6:589-600.
- Andersen, J.S., C.J. Wilkinson, T. Mayor, P. Mortensen, E.A. Nigg, and M. Mann. 2003. Proteomic characterization of the human centrosome by protein correlation profiling. *Nature*. 426:570-574.
- Baines, A.C., and B. Zhang. 2007. Receptor-mediated protein transport in the early secretory pathway. *Trends Biochem. Sci.* 32:381-388.
- Barrell, D., E. Dimmer, R.P. Huntley, D. Binns, C. O'Donovan, and R. Apweiler. 2009. The GOA database in 2009--an integrated Gene Ontology Annotation resource. *Nucleic acids research*. 37:D396-403.
- Basile, J.R., K. Holmbeck, T.H. Bugge, and J.S. Gutkind. 2007. MT1-MMP controls tumor-induced angiogenesis through the release of semaphorin 4D. *J. Biol. Chem.* 282:6899-6905.
- Bendtsen, J.D., L.J. Jensen, N. Blom, G. Von Heijne, and S. Brunak. 2004a. Feature-based prediction of non-classical and leaderless protein secretion. *Protein Eng. Des. Sel.* 17:349-356.
- Bendtsen, J.D., H. Nielsen, G. von Heijne, and S. Brunak. 2004b. Improved prediction of signal peptides: SignalP 3.0. *Journal of molecular biology*. 340:783-795.
- Binns, D., E. Dimmer, R. Huntley, D. Barrell, C. O'Donovan, and R. Apweiler. 2009. QuickGO: a web-based tool for Gene Ontology searching. *Bioinformatics*. 25:3045-3046.
- Blagoev, B., I. Kratchmarova, S.E. Ong, M. Nielsen, L.J. Foster, and M. Mann. 2003. A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. *Nat. Biotechnol.* 21:315-318.
- Blagoev, B., and M. Mann. 2006. Quantitative proteomics to study mitogen-activated protein kinases. *Methods*. 40:243-250.
- Blagoev, B., S.E. Ong, I. Kratchmarova, and M. Mann. 2004. Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. *Nat. Biotechnol.* 22:1139-1145.
- Bonaldi, T., T. Straub, J. Cox, C. Kumar, P.B. Becker, and M. Mann. 2008. Combined use of RNAi and quantitative proteomics to study gene function in *Drosophila*. *Mol. Cell.* 31:762-772.
- Bonaldi, T., F. Talamo, P. Scaffidi, D. Ferrera, A. Porto, A. Bachi, A. Rubartelli, A. Agresti, and M.E. Bianchi. 2003. Monocytic cells hyperacetylate chromatin protein HMGB1 to redirect it towards secretion. *EMBO J.* 22:5551-5560.
- Brand, M., J.A. Ranish, N.T. Kummer, J. Hamilton, K. Igarashi, C. Francastel, T.H. Chi, G.R. Crabtree, R. Aebersold, and M. Groudine. 2004. Dynamic changes in transcription

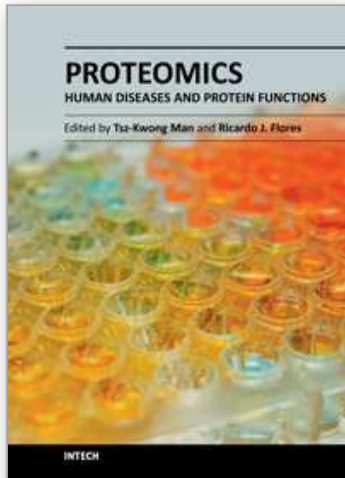
- factor complexes during erythroid differentiation revealed by quantitative proteomics. *Nat Struct Mol Biol.* 11:73-80.
- Brunner, Y., D. Schwartz, Y. Coute, and J.C. Sanchez. 2009. Proteomics of regulated secretory organelles. *Mass Spectrom. Rev.* 28:844-867.
- Bunkenborg, J., G.E. Garcia, M.I. Paz, J.S. Andersen, and H. Molina. 2010. The minotaur proteome: avoiding cross-species identifications deriving from bovine serum in cell culture models. *Proteomics.* 10:3040-3044.
- Chan, C.Y., O. Masui, O. Krakovska, V.E. Belozarov, S. Voisin, S. Ghanny, J. Chen, D. Moyez, P. Zhu, K.R. Evans, J.C. McDermott, and K.W. Siu. 2011. Identification of differentially regulated secretome components during skeletal myogenesis. *Molecular & cellular proteomics : MCP.* 10:M110 004804.
- Chan, X.C., J.C. McDermott, and K.W. Siu. 2007. Identification of secreted proteins during skeletal muscle development. *J. Proteome Res.* 6:698-710.
- Chelius, D., and P.V. Bondarenko. 2002. Quantitative profiling of proteins in complex mixtures using liquid chromatography and mass spectrometry. *J Proteome Res.* 1:317-323.
- Clamp, M., B. Fry, M. Kamal, X. Xie, J. Cuff, M.F. Lin, M. Kellis, K. Lindblad-Toh, and E.S. Lander. 2007. Distinguishing protein-coding and noncoding genes in the human genome. *Proceedings of the National Academy of Sciences of the United States of America.* 104:19428-19433.
- Conrads, T.P., K. Alving, T.D. Veenstra, M.E. Belov, G.A. Anderson, D.J. Anderson, M.S. Lipton, L. Pasa-Tolic, H.R. Udseth, W.B. Chrisler, B.D. Thrall, and R.D. Smith. 2001. Quantitative analysis of bacterial and mammalian proteomes using a combination of cysteine affinity tags and <sup>15</sup>N-metabolic labeling. *Anal Chem.* 73:2132-2139.
- Cox, J., and M. Mann. 2007. Is proteomics the new genomics? *Cell.* 130:395-398.
- Cox, J., and M. Mann. 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 26:1367-1372.
- Cox, J., I. Matic, M. Hilger, N. Nagaraj, M. Selbach, J.V. Olsen, and M. Mann. 2009. A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. *Nat. Protoc.* 4:698-705.
- de Godoy, L.M., J.V. Olsen, J. Cox, M.L. Nielsen, N.C. Hubner, F. Frohlich, T.C. Walther, and M. Mann. 2008. Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature.* 455:1251-1254.
- De Matteis, M.A., and A. Luini. 2008. Exiting the Golgi complex. *Nat. Rev. Mol. Cell Biol.* 9:273-284.
- Dengjel, J., I. Kratchmarova, and B. Blagoev. 2009. Receptor tyrosine kinase signaling: a view from quantitative proteomics. *Mol. Biosyst.* 5:1112-1121.
- Dengjel, J., I. Kratchmarova, and B. Blagoev. 2010. Mapping protein-protein interactions by quantitative proteomics. *Methods Mol Biol.* 658:267-278.
- Elhabazi, A., S. Delaire, A. Bensussan, L. Boumsell, and G. Bismuth. 2001. Biological activity of soluble CD100. I. The extracellular region of CD100 is released from the surface of T lymphocytes by regulated proteolysis. *J. Immunol.* 166:4341-4347.
- Emanuelsson, O., S. Brunak, G. von Heijne, and H. Nielsen. 2007. Locating proteins in the cell using TargetP, SignalP and related tools. *Nature protocols.* 2:953-971.
- Flier, J.S. 2001. Diabetes. The missing link with obesity? *Nature.* 409:292-293.

- Gardella, S., C. Andrei, D. Ferrera, L.V. Lotti, M.R. Torrisi, M.E. Bianchi, and A. Rubartelli. 2002. The nuclear protein HMGB1 is secreted by monocytes via a non-classical, vesicle-mediated secretory pathway. *EMBO Rep.* 3:995-1001.
- Geiger, T., J. Cox, P. Ostasiewicz, J.R. Wisniewski, and M. Mann. 2010. Super-SILAC mix for quantitative proteomics of human tumor tissue. *Nature methods.* 7:383-385.
- Gherardi, E., C.A. Love, R.M. Esnouf, and E.Y. Jones. 2004. The sema domain. *Curr. Opin. Struct. Biol.* 14:669-678.
- Gygi, S.P., G.L. Corthals, Y. Zhang, Y. Rochon, and R. Aebersold. 2000. Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. *Proc Natl Acad Sci U S A.* 97:9390-9395.
- Gygi, S.P., B. Rist, S.A. Gerber, F. Turecek, M.H. Gelb, and R. Aebersold. 1999. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* 17:994-999.
- Han, D.K., J. Eng, H. Zhou, and R. Aebersold. 2001. Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. *Nat Biotechnol.* 19:946-951.
- Henningsen, J., B.K. Pedersen, and I. Kratchmarova. 2011. Quantitative analysis of the secretion of the MCP family of chemokines by muscle cells. *Mol Biosyst.* 7:311-321.
- Henningsen, J., K.T. Rigbolt, B. Blagoev, B.K. Pedersen, and I. Kratchmarova. 2010. Dynamics of the Skeletal Muscle Secretome during Myoblast Differentiation. *Mol. Cell. Proteomics.* 9:2482-2496.
- Hiller, K., A. Grote, M. Scheer, R. Munch, and D. Jahn. 2004. PrediSi: prediction of signal peptides and their cleavage positions. *Nucleic Acids Res.* 32:W375-379.
- Hittel, D.S., J.R. Berggren, J. Shearer, K. Boyle, and J.A. Houmard. 2009. Increased secretion and expression of myostatin in skeletal muscle from extremely obese women. *Diabetes.* 58:30-38.
- Hughes, R.C. 1999. Secretion of the galectin family of mammalian carbohydrate-binding proteins. *Biochim. Biophys. Acta.* 1473:172-185.
- Ko, J.A., T. Gondo, S. Inagaki, and M. Inui. 2005. Requirement of the transmembrane semaphorin Sema4C for myogenic differentiation. *FEBS Lett.* 579:2236-2242.
- Kratchmarova, I., B. Blagoev, M. Haack-Sorensen, M. Kassem, and M. Mann. 2005. Mechanism of divergent growth factor effects in mesenchymal stem cell differentiation. *Science.* 308:1472-1477.
- Kratchmarova, I., D.E. Kalume, B. Blagoev, P.E. Scherer, A.V. Podtelejnikov, H. Molina, P.E. Bickel, J.S. Andersen, M.M. Fernandez, J. Bunkenborg, P. Roepstorff, K. Kristiansen, H.F. Lodish, M. Mann, and A. Pandey. 2002. A proteomic approach for identification of secreted proteins during the differentiation of 3T3-L1 preadipocytes to adipocytes. *Mol. Cell. Proteomics.* 1:213-222.
- Krijgsveld, J., R.F. Ketting, T. Mahmoudi, J. Johansen, M. Artal-Sanz, C.P. Verrijzer, R.H. Plasterk, and A.J. Heck. 2003. Metabolic labeling of *C. elegans* and *D. melanogaster* for quantitative proteomics. *Nat Biotechnol.* 21:927-931.
- Kruger, M., M. Moser, S. Ussar, I. Thievessen, C.A. Lubber, F. Forner, S. Schmidt, S. Zanivan, R. Fassler, and M. Mann. 2008. SILAC mouse for quantitative proteomics uncovers kindlin-3 as an essential factor for red blood cell function. *Cell.* 134:353-364.
- Kubota, K., K. Wakabayashi, and T. Matsuoka. 2003. Proteome analysis of secreted proteins during osteoclast differentiation using two different methods: two-dimensional

- electrophoresis and isotope-coded affinity tags analysis with two-dimensional chromatography. *Proteomics*. 3:616-626.
- Lin, H., E. Lee, K. Hestir, C. Leo, M. Huang, E. Bosch, R. Halenbeck, G. Wu, A. Zhou, D. Behrens, D. Hollenbaugh, T. Linnemann, M. Qin, J. Wong, K. Chu, S.K. Doberstein, and L.T. Williams. 2008. Discovery of a cytokine and its receptor by functional screening of the extracellular proteome. *Science*. 320:807-811.
- Lundgren, D.H., S.I. Hwang, L. Wu, and D.K. Han. 2010. Role of spectral counting in quantitative proteomics. *Expert review of proteomics*. 7:39-53.
- Marie, M., R. Sannerud, H. Avsnes Dale, and J. Saraste. 2008. Take the 'A' train: on fast tracks to the cell surface. *Cell. Mol. Life Sci*. 65:2859-2874.
- Marsh, B.J., and K.E. Howell. 2002. The mammalian Golgi-complex debates. *Nat. Rev. Mol. Cell Biol*. 3:789-795.
- Mayor, S., and H. Riezman. 2004. Sorting GPI-anchored proteins. *Nat. Rev. Mol. Cell Biol*. 5:110-120.
- Neufeld, G., and O. Kessler. 2008. The semaphorins: versatile regulators of tumour progression and tumour angiogenesis. *Nat. Rev. Cancer*. 8:632-645.
- Nickel, W., and C. Rabouille. 2009. Mechanisms of regulated unconventional protein secretion. *Nat. Rev. Mol. Cell Biol*. 10:148-155.
- Nickel, W., and M. Sedorf. 2008. Unconventional mechanisms of protein transport to the cell surface of eukaryotic cells. *Annu. Rev. Cell Dev. Biol*. 24:287-308.
- Nickel, W., and F.T. Wieland. 1998. Biosynthetic protein transport through the early secretory pathway. *Histochem. Cell Biol*. 109:477-486.
- Norheim, F., T. Raastad, B. Thiede, A.C. Rustan, C.A. Drevon, and F. Haugen. 2011. Proteomic identification of secreted proteins from human skeletal muscle cells and expression in response to strength training. *American journal of physiology. Endocrinology and metabolism*.
- Oda, Y., K. Huang, F.R. Cross, D. Cowburn, and B.T. Chait. 1999. Accurate quantitation of protein expression and site-specific phosphorylation. *Proc. Natl. Acad. Sci. U. S. A*. 96:6591-6596.
- Olsen, J.V., B. Blagoev, F. Gnäd, B. Macek, C. Kumar, P. Mortensen, and M. Mann. 2006. Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell*. 127:635-648.
- Olsen, J.V., S.E. Ong, and M. Mann. 2004. Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Mol. Cell. Proteomics*. 3:608-614.
- Ong, S.E., B. Blagoev, I. Kratchmarova, D.B. Kristensen, H. Steen, A. Pandey, and M. Mann. 2002. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics*. 1:376-386.
- Ong, S.E., I. Kratchmarova, and M. Mann. 2003. Properties of <sup>13</sup>C-substituted arginine in stable isotope labeling by amino acids in cell culture (SILAC). *J. Proteome Res*. 2:173-181.
- Ong, S.E., and M. Mann. 2005. Mass spectrometry-based proteomics turns quantitative. *Nat. Chem. Biol*. 1:252-262.
- Osinalde, N., H. Moss, O. Arrizabalaga, M.J. Omaetxebarria, B. Blagoev, A.M. Zubiaga, A. Fullaondo, J.M. Arizmendi, and I. Kratchmarova. 2011. Interleukin-2 signaling pathway analysis by quantitative phosphoproteomics. *Journal of proteomics*.

- Ostasiewicz, P., D.F. Zielinska, M. Mann, and J.R. Wisniewski. 2010. Proteome, phosphoproteome, and N-glycoproteome are quantitatively preserved in formalin-fixed paraffin-embedded tissue and analyzable by high-resolution mass spectrometry. *Journal of proteome research*. 9:3688-3700.
- Park, J.J., and Y.P. Loh. 2008. How peptide hormone vesicles are transported to the secretion site for exocytosis. *Mol. Endocrinol.* 22:2583-2595.
- Pedersen, B.K. 2009. The disease of physical inactivity--and the role of myokines in muscle--fat cross talk. *J. Physiol.* 587:5559-5568.
- Pelham, H.R. 1996. The dynamic organisation of the secretory pathway. *Cell Struct. Funct.* 21:413-419.
- Prudovsky, I., F. Tarantini, M. Landriscina, D. Neivandt, R. Soldi, A. Kirov, D. Small, K.M. Kathir, D. Rajalingam, and T.K. Kumar. 2008. Secretion without Golgi. *J. Cell. Biochem.* 103:1327-1343.
- Rigbolt, K.T., and B. Blagoev. 2010. Proteome-wide quantitation by SILAC. *Methods Mol Biol.* 658:187-204.
- Rigbolt, K.T., J.T. Vanselow, and B. Blagoev. 2011. GProX, a User-Friendly Platform for Bioinformatics Analysis and Visualization of Quantitative Proteomics Data. *Molecular & cellular proteomics : MCP.* 10:O110 007450.
- Ross, P.L., Y.N. Huang, J.N. Marchese, B. Williamson, K. Parker, S. Hattan, N. Khainovski, S. Pillai, S. Dey, S. Daniels, S. Purkayastha, P. Juhasz, S. Martin, M. Bartlet-Jones, F. He, A. Jacobson, and D.J. Pappin. 2004. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Molecular & cellular proteomics : MCP.* 3:1154-1169.
- Roth, L., E. Koncina, S. Satkauskas, G. Cremel, D. Aunis, and D. Bagnard. 2009. The many faces of semaphorins: from development to pathology. *Cell. Mol. Life Sci.* 66:649-666.
- Sallese, M., M. Giannotta, and A. Luini. 2009. Coordination of the secretory compartments via inter-organelle signalling. *Semin. Cell Dev. Biol.* 20:801-809.
- Sango, K., A. Tokashiki, K. Ajiki, M. Horie, H. Kawano, K. Watabe, H. Horie, and T. Kadoya. 2004. Synthesis, localization and externalization of galectin-1 in mature dorsal root ganglion neurons and Schwann cells. *Eur. J. Neurosci.* 19:55-64.
- Schulze, W.X., and M. Mann. 2004. A novel proteomic screen for peptide-protein interactions. *The Journal of biological chemistry.* 279:10756-10764.
- Schulze, W.X., and B. Usadel. 2010. Quantitation in mass-spectrometry-based proteomics. *Annual review of plant biology.* 61:491-516.
- Serini, G., F. Maione, E. Giraudo, and F. Bussolino. 2009. Semaphorins and tumor angiogenesis. *Angiogenesis.* 12:187-193.
- Skalnikova, H., J. Motlik, S.J. Gadher, and H. Kovarova. 2011. Mapping of the secretome of primary isolates of mammalian cells, stem cells and derived cell lines. *Proteomics.* 11:691-708.
- Soufi, B., C. Kumar, F. Gnad, M. Mann, I. Mijakovic, and B. Macek. 2010. Stable isotope labeling by amino acids in cell culture (SILAC) applied to quantitative proteomics of *Bacillus subtilis*. *Journal of proteome research.* 9:3638-3646.
- Strating, J.R., and G.J. Martens. 2009. The p24 family and selective transport processes at the ER-Golgi interface. *Biol. Cell.* 101:495-509.
- Suzuki, K., A. Kumanogoh, and H. Kikutani. 2008. Semaphorins and their receptors in immune cell interactions. *Nat. Immunol.* 9:17-23.

- Tamagnone, L., and S. Giordano. 2006. Semaphorin pathways orchestrate osteogenesis. *Nat. Cell Biol.* 8:545-547.
- Tatsumi, R., Y. Sankoda, J.E. Anderson, Y. Sato, W. Mizunoya, N. Shimizu, T. Suzuki, M. Yamada, R.P. Rhoads, Jr., Y. Ikeuchi, and R.E. Allen. 2009. Possible implication of satellite cells in regenerative motoneuritogenesis: HGF upregulates neural chemorepellent Sema3A during myogenic differentiation. *Am J Physiol Cell Physiol.* 297:C238-252.
- Thery, C., L. Zitvogel, and S. Amigorena. 2002. Exosomes: composition, biogenesis and function. *Nature reviews. Immunology.* 2:569-579.
- Tian, Q., S.B. Stepaniants, M. Mao, L. Weng, M.C. Feetham, M.J. Doyle, E.C. Yi, H. Dai, V. Thorsson, J. Eng, D. Goodlett, J.P. Berger, B. Gunter, P.S. Linsley, R.B. Stoughton, R. Aebersold, S.J. Collins, W.A. Hanlon, and L.E. Hood. 2004. Integrated genomic and proteomic analyses of gene expression in Mammalian cells. *Mol Cell Proteomics.* 3:960-969.
- Unlu, M., M.E. Morgan, and J.S. Minden. 1997. Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis.* 18:2071-2077.
- Walsh, K. 2009. Adipokines, myokines and cardiovascular disease. *Circ. J.* 73:13-18.
- Walther, T.C., and M. Mann. 2010. Mass spectrometry-based proteomics in cell biology. *The Journal of cell biology.* 190:491-500.
- Wasinger, V.C., S.J. Cordwell, A. Cerpa-Poljak, J.X. Yan, A.A. Gooley, M.R. Wilkins, M.W. Duncan, R. Harris, K.L. Williams, and I. Humphery-Smith. 1995. Progress with gene-product mapping of the Mollicutes: *Mycoplasma genitalium*. *Electrophoresis.* 16:1090-1094.
- Wu, C.C., M.J. MacCoss, K.E. Howell, D.E. Matthews, and J.R. Yates, 3rd. 2004. Metabolic labeling of mammalian organisms with stable isotopes for quantitative proteomic analysis. *Anal Chem.* 76:4951-4959.
- Wu, H., X. Wang, S. Liu, Y. Wu, T. Zhao, X. Chen, L. Zhu, Y. Wu, X. Ding, X. Peng, J. Yuan, X. Wang, W. Fan, and M. Fan. 2007. Sema4C participates in myogenic differentiation in vivo and in vitro through the p38 MAPK pathway. *Eur. J. Cell Biol.* 86:331-344.
- Yazdani, U., and J.R. Terman. 2006. The semaphorins. *Genome Biol.* 7:211.
- Yi, E.C., X.J. Li, K. Cooke, H. Lee, B. Raught, A. Page, V. Aneliunas, P. Hieter, D.R. Goodlett, and R. Aebersold. 2005. Increased quantitative proteome coverage with (13)C/(12)C-based, acid-cleavable isotope-coded affinity tag reagent and modified data acquisition scheme. *Proteomics.* 5:380-387.
- Yoon, J.H., K. Yea, J. Kim, Y.S. Choi, S. Park, H. Lee, C.S. Lee, P.G. Suh, and S.H. Ryu. 2009. Comparative proteomic analysis of the insulin-induced L6 myotube secretome. *Proteomics.* 9:51-60.
- Zhou, Y., R.A. Gunput, and R.J. Pasterkamp. 2008. Semaphorin signaling: progress made and promises ahead. *Trends Biochem. Sci.* 33:161-170.
- Zielinska, D.F., F. Gnad, J.R. Wisniewski, and M. Mann. 2010. Precision mapping of an in vivo N-glycoproteome reveals rigid topological and sequence constraints. *Cell.* 141:897-907.



## **Proteomics - Human Diseases and Protein Functions**

Edited by Prof. Tsz Kwong Man

ISBN 978-953-307-832-8

Hard cover, 438 pages

**Publisher** InTech

**Published online** 10, February, 2012

**Published in print edition** February, 2012

Biomedical research has entered a new era of characterizing a disease or a protein on a global scale. In the post-genomic era, Proteomics now plays an increasingly important role in dissecting molecular functions of proteins and discovering biomarkers in human diseases. Mass spectrometry, two-dimensional gel electrophoresis, and high-density antibody and protein arrays are some of the most commonly used methods in the Proteomics field. This book covers four important and diverse areas of current proteomic research: Proteomic Discovery of Disease Biomarkers, Proteomic Analysis of Protein Functions, Proteomic Approaches to Dissecting Disease Processes, and Organelles and Secretome Proteomics. We believe that clinicians, students and laboratory researchers who are interested in Proteomics and its applications in the biomedical field will find this book useful and enlightening. The use of proteomic methods in studying proteins in various human diseases has become an essential part of biomedical research.

### **How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Jeanette Henningsen, Blagoy Blagoev and Irina Kratchmarova (2012). Quantitative Proteomics for Investigation of Secreted Factors: Focus on Muscle Secretome, *Proteomics - Human Diseases and Protein Functions*, Prof. Tsz Kwong Man (Ed.), ISBN: 978-953-307-832-8, InTech, Available from: <http://www.intechopen.com/books/proteomics-human-diseases-and-protein-functions/quantitative-proteomics-for-investigation-of-secreted-factors-focus-on-muscle-secretome>

**INTECH**  
open science | open minds

### **InTech Europe**

University Campus STeP Ri  
Slavka Krautzeka 83/A  
51000 Rijeka, Croatia  
Phone: +385 (51) 770 447  
Fax: +385 (51) 686 166  
[www.intechopen.com](http://www.intechopen.com)

### **InTech China**

Unit 405, Office Block, Hotel Equatorial Shanghai  
No.65, Yan An Road (West), Shanghai, 200040, China  
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元  
Phone: +86-21-62489820  
Fax: +86-21-62489821



© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the [Creative Commons Attribution 3.0 License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

IntechOpen

IntechOpen