Dynamics of the skeletal muscle secretome during myoblast differentiation

Jeanette Henningsen\textsuperscript{1,2}, Kristoffer T. G. Rigbolt\textsuperscript{1}, Blagoy Blagoev\textsuperscript{1}, Bente Klarlund Pedersen\textsuperscript{2}, and Irina Kratchmarova\textsuperscript{1}\#

\textsuperscript{1}Center for Experimental BioInformatics (CEBI), Department of Biochemistry and Molecular Biology, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark

\textsuperscript{2}The Centre of Inflammation and Metabolism (CIM), Department of Infectious Diseases and Copenhagen Muscle Research Centre, Rigshospitalet, Faculty of Health Sciences, University of Copenhagen, Blegdamsvej 9, DK-2100 Copenhagen, Denmark

\textbf{Running Title:} Identification of secreted proteins during skeletal myogenesis

\textbf{Keywords:} Quantitative proteomics, mass spectrometry, SILAC, skeletal muscle, secretome, and differentiation

\#To whom correspondence should be addressed:

Irina Kratchmarova

Tel.: 45-6550-2470

Fax: 45-6593-3018

E-mail: ihk@bmb.sdu.dk
The abbreviations used are: CM, conditioned media; dFBS, dialyzed fetal bovine serum; DMEM, Dulbecco's modified Eagle's medium; ECM, extracellular matrix; FGF-21, fibroblast growth factor-21; Fstl1, follistatin-like 1; GO, Gene Ontology; GPI, glycosylphosphatidylinositol; HIF, hypoxia-inducible factor; IGF, insulin-like growth factor; IGFBP, IGF-binding proteins; IGF-1R, IGF-1 receptor; IPI, International Protein Index; iTRAQ, isobaric tags for relative and absolute quantification; LC, liquid chromatography; LTBP, latent TGF-β binding proteins; LTQ, linear trap quadrupole; MCSF, macrophage colony-stimulating factor; MEF2, myocyte enhancer factor 2; MHC, myosin heavy chain; MMP, matrix metalloproteinases; MRF, myogenic regulatory factors; PDGF, platelet-derived growth factor; Q-PCR, quantitative-PCR; SILAC, stable isotope labeling by amino acids in cell culture; Arg6, L-Arginine-13C6,14N4; Arg10, L-Arginine-13C6,15N4; Lys4, L-Lysine-2H4; Lys8, L-Lysine-13C6,15N2; TGF-β, Transforming growth factor-β.
SUMMARY

During recent years, increased efforts have focused on elucidating the secretory function of skeletal muscle. Through secreted molecules, skeletal muscle affects local muscle biology in an auto/paracrine manner as well as having systemic effects on other tissues. Here we used a quantitative proteomics platform to investigate the factors secreted during the differentiation of murine C2C12 skeletal muscle cells. Using triple encoding Stable Isotope Labeling by Amino acids in Cell culture (SILAC), we compared the secretomes at three different time points of muscle differentiation and followed the dynamics of protein secretion. We have identified and quantitatively analyzed 635 secreted proteins including 35 growth factors, 40 cytokines and 36 metallopeptidases. The extensive presence of these proteins that can act as potent signaling mediators to other cells and tissues strongly highlights the important role of the skeletal muscle as a prominent secretory organ. In addition to previously reported molecules, we have identified many secreted proteins that have not previously been shown to be released from skeletal muscle cells, nor shown to be differentially released during the process of myogenesis. We found 188 of these secreted proteins to be significantly regulated during the process of myogenesis. Comparative analyses of selected secreted proteins revealed little correlation between their mRNA and protein levels indicating pronounced regulation by posttranscriptional mechanisms. Further, analyses of the intracellular levels of members of the semaphorin family and their corresponding secretion dynamics demonstrated that the release of secreted proteins is tightly regulated by the secretory pathway, the protein’s stability, and/or the processing of secreted proteins. Finally, we provide 299 unique hydroxyproline sites mapping to 48 distinct secreted proteins and have discovered a novel hydroxyproline motif.
INTRODUCTION

The skeletal muscle is a highly dynamic organ responsible for locomotion and generation of body heat and plays an essential role in the maintenance of metabolic homeostasis. Furthermore, it is the major target for insulin-induced glucose uptake and has a high capacity to metabolize fatty acids. Impaired glucose and lipid metabolism in the skeletal muscle are characteristic hallmarks of insulin resistance associated with different diseases such as obesity, type 2 diabetes and metabolic syndrome (1). During the last decade, data have emerged demonstrating that the skeletal muscle also plays an active role as an endocrine organ. Skeletal muscles have been suggested to be a source of secreted proteins, conceptualized as myokines that can influence metabolism and other biological processes in a systemic manner at different tissue targets. Regular exercise has many beneficial effects on whole-body well-being: it improves insulin sensitivity, metabolic processes, and blood pressure and reduces inflammation mediated through alterations of gene and protein expression in different cell types. Several studies described the association of physical inactivity with the development of different pathologies, including mental diseases, as well as obesity, type 2 diabetes, metabolic syndrome and cardiovascular diseases (1). It has been suggested that the skeletal muscle, through secreted molecules, could be transmitting many of the beneficial effects of physical activity by affecting whole body homeostasis in an autocrine, paracrine, and/or endocrine fashion.

The initial recognition of skeletal muscle as an endocrine organ originates from studies that identify interleukin (IL)-6 being expressed and secreted by contracting skeletal muscle. The plasma level of muscle-derived IL-6 increases in response to physical activity and the secreted IL-6 can affect metabolic and inflammatory processes (2, 3). In addition to IL-6, other potential myokines, with either systemic effects or local effects in an autocrine and paracrine manner, have been identified, including IL-8 and IL-15 (3). Furthermore, a mouse model of inducible muscle hypertrophy was used as a tool for the identification of new myokines, resulting in the identification of fibroblast growth factor (FGF)-21 and
follistatin-like (Fstl) as secreted skeletal muscle proteins involved in metabolic processes and vascularization (4, 5).

Despite recent discoveries, the secretome of skeletal muscle has not been fully characterized and the number of identified secreted proteins is still limited. The majority of studies investigating the secretory profile of the skeletal muscle have used traditional biochemical and molecular biology strategies to identify muscle secreted proteins; although powerful, these approaches are restricted to few proteins. Existing mass spectrometry-driven attempts to target the muscle-specific secreted proteome resulted in a limited number of identified secreted proteins (6-8). Comprehensive quantitative analyses of the muscle secretome could provide greater insight into muscle biology and the muscle-dependent cross talk with other tissues. The fast development of quantitative mass spectrometry-based proteomics has allowed the qualitative as well as the quantitative evaluation of complex biological processes in a large scale-manner (9). The SILAC methodology represents a powerful quantitative tool to study the dynamics of different biological processes including in-depth characterization of the signaling cascades involved in various types of cellular differentiation (10-12).

Here we used triple encoding SILAC to identify and quantitatively evaluate the dynamics of secreted proteins during the differentiation process of the murine C2C12 skeletal muscle cell line. We identified 635 proteins that are secreted from skeletal myoblasts; to our knowledge, this is the largest dataset covering the muscle secretome. Furthermore, we quantitatively profiled the secretion of 624 of these proteins during the course of skeletal muscle differentiation. Proteins previously known to be secreted from skeletal muscle were identified, as well as many proteins not previously shown to be secreted by skeletal myoblasts. Based on Gene Ontology (GO) annotations, the identified secreted proteins are involved in various biological processes and molecular function categories, highlighting the diversity of the muscle secretome. Moreover, 188 secreted proteins were significantly and dynamically regulated during skeletal myogenesis, suggesting involvement in skeletal muscle development in either an autocrine or paracrine manner. Among the regulated secreted factors we have identified a family of proteins, the semaphorins, that are involved in muscle development. We investigated more closely the expression
profiles of semaphorins at the mRNA and protein level during the myoblast conversion. Finally, our approach resulted in identification of 299 posttranslationally modified sites on 48 secreted proteins containing hydroxyproline residues and revealed a new motif distinct from the canonical hydroxyproline motif. To our knowledge, this also represents the largest dataset of proline hydroxylation on secreted proteins.

EXPERIMENTAL PROCEDURES

Cell culture

Murine C2C12 cells (American Type Culture Collection, ATCC) were cultured in custom prepared DMEM 4.5 g/l glucose deficient in Arginine and Lysine (Gibco, Invitrogen) supplemented with 10% dialyzed Fetal Bovine Serum (dFBS) (Gibco, Invitrogen), 1000 UI/ml penicillin, 1000 UI/ml streptomycin, and 2 mM L-Glutamine at 37°C in humidified air containing 10% CO₂. Regular L-Arg (Arg0) and L-Lys (Lys0) was added to the cell population to be harvested on day 0, whereas the cell populations harvested on day 2 and day 5 of differentiation received different isotopic forms of L-Arginine and L-Lysine (Cambridge Isotope Laboratories and Sigma-Aldrich). Day 2 cells: L-Arginine-^{13}C_{6},^{14}N_{4} (Arg6) and L-Lysine-^{2}H_{4} (Lys4) and Day 5 cells: L-Arginine-^{13}C_{6},^{15}N_{4} (Arg10) and L-Lysine-^{13}C_{6},^{15}N_{2} (Lys8). Cells were passaged at least 5 times to ensure complete incorporation of the labeled amino acids before differentiation (13). Differentiation was induced upon confluence by reducing the amount of dFBS to 2%, with media changes every 2 days.

Collection and preparation of conditioned media (CM) for analysis by mass spectrometry

Cells were passaged and differentiated as described above. On day 0, 2, and 5 of differentiation, cells were washed six times in serum-free media, to reduce the amount of contaminating serum proteins, and incubated for 12h in corresponding serum-free labeling media. At end of starvation, the CM was
collected, centrifuged and filtered using a 0.2 μm filter to ensure removal of any dead cells. The protein content of the individual media was measured using the Coomassie Plus reagent from Pierce. CM obtained from the three different time points was mixed in a 1:1:1 ratio according to the measured protein concentrations and the combined media was concentrated using spin columns with a cutoff of 3000 Da (Vivaspin, Satorius). The concentrated media was analyzed by a gradient gel Novex 4-12% Bis-Tris (Invitrogen) and proteins visualized using Colloidal Blue (Invitrogen). Whole gel lanes of CM were cut into slices and subjected to in-gel digestion according to (14) with minor modifications. In-gel reduction was done for 45 min at 56°C using 10 mM DTT in 50 mM ammonium bicarbonate followed by in-gel alkylation using 55 mM iodoacetamide in 50 mM ammonium bicarbonate for 30 min in the dark. Digestion was performed overnight with 12.5 ng/μl Sequencing Grade Modified Trypsin (Promega) at 37°C. Extracted peptide mixtures were desalted and concentrated via StageTips essentially as described (15).

Mass spectrometry and data analysis
Acidified peptide mixtures were separated by online C$_{18}$-reverse-phase nanoscale liquid chromatography and analyzed by tandem mass spectrometry (LC-MS/MS). Reverse phase nanoscale -LC was carried out by an Agilent 1200 nanoflow system (Agilent Technologies) connected to an LTQ-Orbitrap XL mass spectrometer (ThermoFisher) equipped with a nanoelectrospray ion source (Proxeon Biosystems, Odense, DK). Full scan MS spectra (m/z 300 – m/z 2000) were acquired in the Orbitrap with a resolution of 60,000 at m/z 400. The six most abundant ions in the linear ion-trap were sequentially selected for sequencing by collision-induced dissociation, using a collision energy of 35%. Ions already selected for fragmentation was dynamically excluded for 45 s. The lock mass calibration feature was enabled to improve mass accuracy (16). Additional mass spectrometric parameters included a spray voltage of 2.3 kV, no sheath and auxiliary gas flow, and the temperature of the heated capillary was 170°C.
Acquired raw mass spectra data were processed and quantitated using MaxQuant v1.0.12.5 essentially as described (17). Briefly, a peak list was generated using the Quant element of MaxQuant using the following parameters: Orbitrap/FT Ultra; SILAC triplets with medium labels Arg6 + Lys4 and heavy labels Arg10 + Lys8; maximum of 3 labeled AAs; maximum mass deviation allowed for precursor ions was 5 ppm and 0.5 Da for MS/MS events; and 6 most intense MS/MS peaks per 100 Da. A maximum of 3 missed cleavages was allowed and enzyme specificity was set to trypsin, allowing for cleavage N-terminal to proline and between aspartic acid and proline. In addition, carbamidomethyl (C) was chosen as fixed modification and variable modifications included Oxidation (M), Acetyl (Protein N-term), Deamidated (NQ), Gln->pyro-Glu (N-term Q), and Oxidation (P). The MaxQuant generated peak list was searched via the Mascot search engine v2.2 (MatrixScience, London, UK) against the mouse International Protein Index (IPI) database v3.44 containing 110461 protein sequences including a list of common contaminants such as keratins, serum albumin, trypsin and concatenated with reverse copies of all sequences. Subsequently, the acquired Mascot dat-files in association with raw files were processed and quantitated applying the Identify element of MaxQuant using the following parameters: peptide and protein FDR, 0.01; max. peptide PEP, 1; min. peptide length, 6; min unique peptides and peptides, 1. Razor and unique peptides, including unmodified and all modified peptides except oxidation (P) identified peptides, were subjected to quantitation with a min. ratio count of 1. Only proteins identified with at least 2 peptides and 1 unique peptide were included in the final list of identified proteins. We used ProteinCenter v2.0 (Proxeon Biosystems, Odense, DK), a proteomics data mining and management software to filter our large data set of proteins to isolate secreted proteins. Proteins were classified as secreted based on the predicted presence of a signal peptide in the identified protein combined with extracellular filtering based on cellular component information from GO annotations (18).

**Bioinformatic analyses**

To extract secreted proteins, protein identifications were filtered using ProteinCenter (Proxeon, Odense, Denmark). The entire IPI mouse database (version 3.52) was subjected to the same filtering process that
was used to identify secreted proteins in the mass spectrometry data, to generate a list of all proteins that are potentially secreted from any cell in the mouse. Fisher’s exact test was used to identify over-represented Gene Ontology terms in the identified proteins compared to this subset of the database. Molecular Function terms occurring in the identified proteins at least ten times and Biological Process terms occurring at least 20 were tested at a significance level 0.05 after correction for multiple testing using the Benjamini & Hochberg method. In order to account for biases that may be introduced through the MS measurements we performed identical bioinformatic analyses testing for over-represented GO terms within the cellular (non-secreted) proteins identified in our screen. These proteins were tested using the same criteria and background database. Five categories, namely nucleotide binding, protein binding, ATP binding, cellular metabolic process and primary metabolic process, were found significantly over-represented in both analyses and were therefore subtracted from the lists of enriched GO terms within the secreted proteins dataset.

Unsupervised clustering was performed using the fuzzy c-means algorithm as implemented in the Mfuzz package (19) for R (20). Only regulated protein identifications, as defined by a MaxQuant Significance value <0.05 in at least one time point, were subjected to clustering into 5 clusters with a fuzzification parameter of 3. For characterization of proteins in each cluster, ‘biological process’ and ‘molecular function’ GO terms were counted and terms occurring only once in a cluster were omitted from the analysis.

From all identified hydroxyproline sites high-confidence sites were extracted by requiring a Mascot score >30 and a MaxQuant localization probability >0.75 (21). High-confidence sites were partitioned into sites matching the proteotypic collagen hydroxyproline motif (X-P-G or G-P), sites on fibrinogen-1 and a final group containing the remaining sites. WebLogo (22) was used to illustrate the amino acid distribution in the proximity of the hydroxyprolines in the three groups. SILAC ratios for hydroxyproline containing peptides were normalized by the SILAC-ratio for the parent proteins.

**Immunostaining**
Cells were seeded on cover slips in 6-well plates in DMEM deficient in Arg and Lys supplemented with normal Arg (Arg0), Lys (lys0), 10% dFBS, L-glutamine, and antibiotics. Differentiation was induced when reaching total confluence using 2% dFBS (23). On day 0, 2, and 5 of differentiation cells were washed 6 times in serum free DMEM and subsequently starved for 12h in accordance with the treatment of cells for mass spectrometry analysis. At the end of starvation cells were fixed for 10 min using 3.7% paraformaldehyde. Permeabilization and blocking was achieved using 1% Triton X-100 in PBS and 1% Horse Sera, 0.1% Tween-20 in PBS, respectively. The nucleic acid stain Sytox (Molecular Probes, Invitrogen) was used for nuclear visualization. Antibodies including dilutions were MHC (H-300) (Santa Cruz sc-20641, 1:50), anti-Troponin T (Sigma-Aldrich T6277, 1:50), and Alexa488-conjugated secondary antibodies (Molecular Probes, Invitrogen 1:200). Images were obtained using a Zeiss LSM 510 Meta confocal laser scanning microscope.

RNA purification, cDNA synthesis, and real-time quantitative-PCR (Q-PCR)

Total cellular RNA was purified using Trizol (Invitrogen) according to manufacturer’s instructions. RNA was extracted from dishes of cells from which CM was collected for mass spectrometry analysis. 1 μg of total RNA was reversed transcribed using 200 units of M-MLV Reverse Transcriptase (Invitrogen) according to protocol given by supplier with 250 ng of Random Hexamer primers (GE Healthcare) but without any RNase inhibitor. Quantitative PCR analysis was performed using the SYBR Green JumpStart Taq ReadyMix (Sigma-Aldrich) and run on the Mx3000P QPCR machine from Stratagene. The specificity and quality of primers were evaluated including a dissociation curve analysis at the end of each PCR run. In addition, to ensure that primers only gave rise to a single product amplicons were analyzed by ethidium bromide stained 1.5% agarose gels. For normalization the expression of the reference gene TATA box binding protein (TBP) was used.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward primer</th>
<th>Reverse primer</th>
</tr>
</thead>
</table>
Whole cell extracts and Western Blotting

Whole cell extracts were prepared using RIPA buffer (50 mM Tris-HCl, 150 mM NaCl, 1% NP-40, 1 mM EDTA, 0.25% Na Deoxycholate, 0.1% SDS, 1mM Na$_3$VO$_4$, 5 mM β-glycerophosphate, 5 mM NaF) containing protease inhibitors (Roche Applied Science). Whole cell extracts were prepared from the same dishes of cells used for mass spectrometry studies. Approx. 25 µg of protein was separated using gradient gels Novex 4-12% Bis-Tris (Invitrogen). Antibodies and dilutions used were: anti-troponin T (Sigma-Aldrich T6277, 1:1000), MHC (H-300) (Santa Cruz sc-20641, 1:1000), and myogenin (F5D) (Santa Cruz sc-12732, 1:500), rabbit anti-Semaphorin 3A (Millipore, AB9604, 1:1000), anti-human Semaphorin 3E (R&D AF3239, 1:1000), Rabbit polyclonal to semaphorin 7a (Abcam ab23578, 1:2000), β-actin (C4) (Santa Cruz sc-47778, 1:5000), and HRP-conjugated secondary antibodies from GE Healthcare.

Conditioned medium for Western Blotting and Zymography

Cells were cultured as described above except that all cell populations (day 0, 2, and 5 cells) received normal Arg and Lys. Cells were washed six times and incubated in serum free DMEM supplemented with pen/strep and L-Glutamine for 12h on day 0, day 2, and day 5 during myogenesis. Differentiation was
induced at confluence using 2% dFBS as for the proteomics experiments. CM was harvested centrifuged and filtrated (0.2 µm filter) to remove any dead cells and cell debris. Enrichment of secreted proteins was done via ultrafiltration using 3000 Da molecular weight cut off (Vivaspin columns, Satorius). After concentration the protein content of the individual media were estimated via Bradford assay (Coomassie Plus reagent, Pierce).

Western Blot analysis of selected proteins was done separating 25 µg of proteins via Novex 4-12% Bis-Tris gradient gels (Invitrogen). Equal load of concentrated CM from day 0, day 2, and day 5 was verified by Colloidal Blue staining of a corresponding 4-12% Bis-Tris gradient gel. Antibodies and dilutions were: rabbit anti-Semaphorin 3A (Millipore, AB9604, 1:2000), anti-human Semaphorin 3E (R&D AF3239, 1:1000), Sema6A (QQ06) (Santa Cruz sc-74274, 1:100), Anti-human/mouse Semaphorin 7A (CD108) (R&D MAB2068, 1:1000), and HRP-conjugated secondary antibodies from GE healthcare.

In zymography studies, 5 µg of proteins were separated via 10% tris-glycine gels containing 0.1% gelatin as substrate. CM was analyzed by 1D-electrophoresis without prior reduction of samples. Following gel electrophoresis, gels were washed using Zymogram Renaturing Buffer (Invitrogen) to remove SDS. Subsequently, gels were incubated in Zymogram Developing Buffer (Invitrogen) at 37°C for approx. 12h of digestion. Colloidal Blue staining (Invitrogen) was used to visualize the activity of proteases.

RESULTS AND DISCUSSION

**Efficient differentiation of C2C12 cells in labeling media**

Development of the skeletal muscle is a complex multistep process that is under the control of numerous extracellular stimuli (24, 25). Upon induction of differentiation, myoblasts withdraw from the cell-cycle, differentiate, elongate, and fuse into multinucleated myotubes. Myogenesis requires the coordinated regulated expression of numerous genes and proteins. The myogenic program is orchestrated by a family
of transcription factors known as myogenic regulatory factors (MRFs) including MyoD, Myf-5, myogenin, and MRF4/Myf-6 (26, 27). In association with other transcription factors such as the myocyte enhancer factor 2 (MEF2) family, MRFs modulate the transcriptional activity of many muscle-specific genes. In addition to MRFs, muscle-secreted proteins also play a pivotal role, affecting local muscle biology in an autocrine and paracrine manner (24, 28).

In this study we have utilized the triple encoding SILAC approach during muscle differentiation of murine C2C12 cells induced by using 2% dialyzed sera (Fig. 1A) (see Experimental Procedures). The efficiency of differentiation was verified by analyzing the expression of well-known myogenic markers either via confocal laser scanning microscopy (Fig. 1B) or via Western blotting (Fig. 1C). Fusion of myoblasts into small myotubes was visible by day 2 of differentiation, and by day 5 of differentiation the myoblasts were engaged in formation of long multinucleated myotubes (Fig. 1B). Additionally, expression of troponin T and myosin heavy chain (MHC), two proteins involved in muscle physiology, and the muscle-regulatory transcription factor myogenin were enhanced on day 2 and throughout differentiation (Fig. 1C).

**Identification of the murine skeletal muscle secretome**

We applied triple encoding SILAC to murine C2C12 cells during the course of myogenesis to investigate which proteins are being secreted and regulated during skeletal muscle differentiation (Fig. 1A). One of the major challenges in secretome studies is to distinguish cell-derived proteins from residual sera proteins. Applying the SILAC approach offers an advantage compared to other labeling strategies since metabolic labeling of cells facilitates the discrimination between secreted proteins versus contaminating sera proteins due to the incorporation of the isotope-labeled amino acids into the proteins during their synthesis inside the cell. To quantitatively analyze protein secretion during myogenesis, C2C12 cells were cultured in Arg- and Lys-deficient medium supplemented with either normal Arg0 and Lys0 (conditioned media (CM) collected at day 0), or isotope-labeled analogues Arg6 and Lys4 (CM collected at day 2 of differentiation) and Arg10 and Lys8 (CM collected at day 5 of differentiation). Using high resolution
mass spectrometry, MaxQuant software suit for data processing and quantitation, and ProteinCenter bioinformatics program for protein annotations we have identified 635 putative secreted proteins (Supplemental Table 1). The list of identified secreted proteins from skeletal myoblasts in this study represents the largest number of muscle secreted proteins reported so far. In addition, we were able to quantitatively profile 624 of these proteins, thereby revealing the dynamics of their secretion during the course of myogenesis (Supplemental Table 1).

Functional annotation analyses of identified secreted proteins

We next characterized the distribution of the skeletal muscle secretome according to the GO Biological Process and Molecular Function annotations and performed statistical enrichment analyses to identify overrepresented categories in our dataset. We first extracted all potential secreted proteins from the entire IPI mouse database applying the same criteria for selecting proteins as being secreted as for the proteins identified in this work (see Experimental Procedures) and used these as a reference dataset.

The identified secreted proteins were found to be functionally diverse and to be involved in various biological processes. The significant enrichment of GO categories like regulation of signal transduction (n=24, p=0.001), inflammatory response (n=29, p=0.04), skeletal system development (n=36, p=3.2e-7) and neurogenesis (n=41, p=2.8e-8) highlights distinct roles of the secreted muscle proteins (Fig. 2A and Supplemental Table 2). Moreover, several molecular function categories were found to be significantly enriched compared to the IPI database, further demonstrating an enrichment of distinct proteins with a defined function in our dataset (Fig. 2B and Supplemental Table 2). Not surprisingly, among the most over-represented terms were the metal ion binding (n=212, p=9.5e-10), the extracellular matrix structural constituent (n=41, p=1.1e-12), and the enzyme regulator activity (n=54, p=1.4e-7). Key categories including enzyme inhibitor activity (n=40, p=6.4e-6), enzyme activator activity (n=12, p=0.01), integrin binding (n=14, p=0.01), growth factor activity (n=35, p=0.002), growth factor binding (n=25, p=1.4e-6), cytokine activity (n=40, p=0.01), and metalloprotease activity (n=36, p=0.003) were also found significantly overrepresented. These findings clearly characterize the skeletal muscle as an important
secretory organ in the body. For example, we found several pivotal proteins with growth factor activity such as activin, transforming growth factor (TGF)-β, FGF-21, macrophage colony-stimulating factor (MCSF) and platelet-derived growth factor (PDGF).

Characterization of differentially regulated secreted proteins during myogenesis

Statistical analysis revealed 188 of the 624 secreted proteins to be differentially regulated between proliferating and differentiating myoblasts (Fig. 3 and Supplemental Table 3). Regulated proteins were subjected to unsupervised clustering according to their relative secretion profile resulting in 5 distinct dynamic profiles. In general, a higher proportion of the identified muscle-released proteins exhibited an increased level of secretion compared to the proteins with a decreased secretion profile during the course of C2C12 differentiation (Fig. 3). More than 60% of the proteins were assigned to two clusters: cluster 2 demonstrates a marked increase in secretion during early differentiation, and cluster 4 represents secreted proteins with a pronounced upregulation in secretion levels at terminal differentiation. This observation indicates that an overall increased secretory activity is accompanying skeletal muscle cells maturation.

To examine if proteins with similar biological implications also demonstrated similar regulation, we looked at the distribution of the regulated secreted proteins into the GO annotations Biological Processes and Molecular Function (Fig. 3 and Supplemental Table 3). Upon induction of differentiation, a majority of the C2C12 myoblasts withdraw from the cell cycle and fuse to form long multinucleated myotubes. In agreement with this, we have identified several secreted proteins classified as negative regulators of cell proliferation to be upregulated at terminal differentiation. In contrast, we found that secreted proteins involved in the positive regulation of proliferation exhibited a more diverse profile, allocating to three different clusters demonstrating both a reduced secretion as well as an increased secretion during the course of differentiation, illustrating the complexity of the myogenic program. The formation of long multinucleated myotubes during muscle differentiation is a critical process that involves multiple steps such as cell migration, alignment, recognition, adhesion, cell fusion, and reorganization of the ECM. We observed several proteins involved in cell-cell adhesion and extracellular matrix organization exhibiting
differential secretion profiles during the course of differentiation. In addition, we also found differential secretion of several proteins annotated to play a role in different metabolic processes, organ morphogenesis, cell development, cell morphogenesis, blood vessel development and nervous system development - all biological processes connected with cell differentiation and tissue development (Fig. 3).

*Differential secretion of myogenic regulatory growth factors during myoblast differentiation*

Skeletal myogenesis is modulated by a number of growth factors, which may act in an autocrine and/or paracrine manner to regulate proliferation and differentiation. TGF-β and insulin-like growth factor (IGF) are two groups of growth factors demonstrated to play critical regulatory roles in the control of muscle development either inhibiting or promoting differentiation, respectively.

*Dynamic secretion profile of TGF-β1, -β2, -β3 and associated regulatory proteins*

TGF-βs are small 25 kDa pleiotropic proteins involved in a wide array of physiological processes including cell proliferation, differentiation, angiogenesis, migration, and apoptosis (29, 30). In addition, TGF-β also plays an important role in ECM morphogenesis promoting the synthesis of ECM components. TGF-β can function in an autocrine and paracrine way as well as targeting more distant cells.

We identified all three isoforms of TGF-β (TGF-β1, -β2, and -β3) as being secreted and differentially regulated during myogenic conversion (Fig. 4A and Supplemental Table 1). Little is known concerning the role of the individual isoforms of TGF-β in myogenesis, although all three isoforms of TGF-β have been shown to function as inhibitors of skeletal muscle differentiation in vitro (31-33). Only under certain specialized conditions including serum-rich media or low cell density was TGF-β found to enhance differentiation (34, 35). Recently, in vivo studies also supported TGF-β as a negative regulator of muscle development during muscle regeneration (36).

Interestingly, we observed a pronounced increase in the secreted level of both TGF-β1, -β2, and -β3 during myogenesis (Fig. 4A). This apparent paradox clearly illustrates the complexity of TGF-β signaling.
in the myogenic program. The secretion of TGF-β2 and -β3 increased gradually during the course of differentiation, whereas TGF-β1 secretion was similar on day 0 and day 2 of differentiation but increased 10 fold on day 5. Other studies have also demonstrated an increased expression of TGF-β in skeletal muscle during myogenesis and regeneration (31, 37-39). It should be emphasized, however, that an increased level of secreted TGF-β does not necessarily reflect an increased level of bioactive TGF-β. The level of TGF-β signaling is regulated in many different ways, one of them being sequestration of TGF-β to the ECM via latent TGF-β binding proteins (LTBP). We found all four LTBP members to be secreted from the C2C12 cells. While the secretion of LTBP-1, LTBP-2, and LTBP-4 was not significantly regulated during myogenesis, the release of LTBP-3 was modestly enhanced on day 2 but markedly increased on day 5 (Fig. 4A and Supplemental Table 1). The function of the individual LTBPs in skeletal myogenesis still needs to be elucidated and it is tempting to speculate that they have a major role in stabilizing the muscle-released TGF-β for endocrine purposes towards more distant target tissues.

In addition to the family of LTBP proteins, we have also identified other proteins known to function as modulators of TGF-β activity exhibiting a differentially secretion profile during differentiation. Decorin and biglycan are two small proteoglycans residing in the ECM and reduce the bioavailability of TGF-β during myogenesis (40, 41). We found that the level of secreted decorin was markedly increased during the early stages of differentiation, whereas an increased secretion of biglycan was observed throughout differentiation (Supplemental Table 1).

Dynamic secretion profile of IGF-1 and IGF-2

In contrast to the TGF-βs, IGFs known to promote differentiation demonstrated an increased level of secretion in agreement with the functionality of IGF in myogenesis. The growth factors IGF-1 and IGF-2 are known to be secreted from muscle cells and are both involved in numerous processes in skeletal muscle biology (42, 43). In addition, recent papers have suggested that IGF-1 may play a more significant role in metabolic homeostasis, including skeletal muscle metabolism, than previously anticipated (44, 45).
Here we show the dynamic profile of IGF secretion during a 5-day differentiation study of C2C12 myoblasts (Fig. 4A and Supplemental Table 1). Secretion of IGF-2 increased dramatically during the course of differentiation supporting previous findings that IGF-2 plays an essential regulatory role throughout differentiation. In contrast, the highest level of secreted IGF-1 was observed on day 2 of differentiation indicating that IGF-1 is a more important modulator of cellular processes in the initial phases of differentiation.

In the extracellular environment, the activity of IGFs is regulated by IGF-binding proteins (IGFBPs), which modulate the bioactivity of IGFs (46, 47). The IGFBPs exert a dual function by either inhibiting or enhancing the action of IGFs depending on cell type and context. Interestingly, several studies have demonstrated that IGFBPs also exert IGF-independent cellular actions modulating gene expression and activating intracellular signaling. We have identified five IGFBPs (IGFBP-2, IGFBP-4, IGFBP-5, IGFBP-6, and IGFBP-7) as proteins secreted from C2C12 cells; the secretion of IGFBP-2 and IGFBP-4 was found to be differentially regulated during the course of differentiation (Fig. 4A and Supplemental Table 1). Increased amounts of secreted IGFBP-4 was observed on day 2 of differentiation followed by a reduction on day 5, but still 6 fold higher compared to day 0 (Fig. 4A). IGFBP-2 has in many studies been shown to serve a growth promoting function (48). In support of this, we found that the level of secreted IGFBP2 indeed decreased during the course of differentiation, in which cells withdraw from the proliferative phase (Fig. 4A).

The differential secretion pattern of the IGFBPs suggests that they serve distinct functions in muscle development. Expression of IGFBP-5 increases dramatically during differentiation and IGFBP-5 has attracted a lot of attention in a number of studies investigating muscle physiology, where it was shown to be either a negative or positive regulator of differentiation, depending on the experimental conditions (49-52). However, we did not observe any significant regulation of the released IGFBP-5 during differentiation (Supplemental Table 1). In addition, we identified IGFBP-6 and IGFBP-7 to be secreted from skeletal myoblasts, but we did not observe any significant regulation of their secretion during the process of differentiation (Supplemental Table 1).
Secretion of IGFs and TGFs is regulated by posttranscriptional mechanisms

Several studies have shown that the dynamic profile of protein expression often cannot be correlated to dynamics at RNA level (53, 54). Therefore, we decided to investigate if the increased release of IGF and TGF-β was regulated at the mRNA level (Fig. 4B). Examining the expression of TGF-β mRNA using quantitative-PCR (qPCR) revealed a very modest up-regulation of TGF-β1 expression on day 5 and on both day 2 and day 5 of TGF-β2. TGF-β3 demonstrated a more pronounced increase in expression although not to the same extent as the level of secreted TGF-β3. In agreement with previous reports, we found a dramatic upregulation of IGF-2 mRNA during myogenesis (52, 55). Whereas the mRNA expression of IGF-2 was similar on both day 2 and day 5 of differentiation, we observed a higher level of secreted IGF-2 on day 5 compared to day 2. Moreover, secretion of IGF-2 was markedly higher compared to RNA expression at terminal differentiation. Although the dynamics of IGF-1 secretion resembles its mRNA levels, the secreted IGF-1 is further increased several fold on day 2 compared to the corresponding increase of mRNA expression (Fig. 4). Collectively, these data suggests that the differential release of IGFs and TGF-βs is predominantly regulated by posttranscriptional mechanisms.

Secretion of semaphorins during muscle development

The mammalian family of semaphorins is a large group of proteins consisting of 5 subclasses including secreted (Sema3A-G), transmembrane (Sema4A-G, Sema5A-B, and Sema6A-D), and glycosylphosphatidylinositol (GPI)-anchored (Sema7A) proteins. The semaphorins were initially characterized as constituents of the regulatory system responsible for axon guidance during the development of the nervous system (56). However, accumulating data have now recognized semaphorins as important modulators of angiogenesis, organogenesis, tumor progression, and immune responses (57). Moreover, recent studies have suggested that members of the semaphorin family play an important role in regulating skeletal muscle development (58, 59).
We have identified four members of the Sema3 subfamily (Sema3A, Sema3B, Sema3D, and Sema3E), two members of the Sema4 subfamily (Sema4B and Sema4C), and one member of each of the Sema6 and Sema7 subfamilies (Sema6A and Sema7A) to be secreted from both myoblast and myotubes (Fig. 5A and Supplemental Table 1). Moreover, the secretion of the soluble Sema3A, Sema3D, and Sema3E and membrane associated Sema6A and Sema7A were found to be differentially regulated during the course of differentiation. The secretion of Sema3A, Sema3E, Sema3D, and Sema6A were markedly elevated on day 2 of differentiation, whereas the secretion of Sema7A increased throughout differentiation (Fig. 5A). Western blot analyses examining the level of secreted Sema-3A, -3E, -6A, and -7A demonstrated the same trend of dynamics as the dynamic profiles obtained by mass spectrometry (Fig. 5B). Interestingly, analyzing the intracellular protein expression of Sema-3A, -3E, and -7A demonstrated a somewhat different profile suggesting that the level of secreted semaphorin is regulated posttranslationally (Fig. 5C). While protein expression of Sema3A remained more or less constant during myogenesis, the expression of Sema3E was decreased both at day 2 and day 5 of differentiation. The expression of Sema7A displayed a similar profile to secreted Sema7A although the increase in secreted Sema7A was more pronounced compared to the increase in protein expression.

Investigating the mRNA expression of Sema3A, Sema3E, and Sema7A by qPCR indicated that the level of secreted semaphorins is regulated both by posttranscriptional as well as posttranslational mechanisms (Fig. 5D). The expression profiles of Sema3A and Sema3E comparing RNA and protein were similar, implying that the high level of secreted Sema3A and Sema3E on day 2 of differentiation occurs via posttranslational regulatory mechanisms (Fig. 5A-D). The expression of Sema7A mRNA remained constant during differentiation, whereas the expression of Sema7A protein was increased at day 5 of differentiation. In addition, the level of secreted Sema7A was increased at day 2 and further increased at day 5 indicating that secreted Sema7A is regulated both at the posttranscriptional and posttranslational levels (Fig. 5A-D). Collectively these data demonstrate that care should be taken trying to extrapolate the level of secreted proteins from intracellular protein expression. The differences observed between intracellular protein expression and the secreted proteins indicate that the level of individual secreted
proteins can be regulated either via the secretory pathway or by stability and/or processing of released proteins.

The high level of secreted Sema3A, Sema3E, and Sema6A observed on day 2 suggests that they may function in an autocrine fashion regulating early events during differentiation, whereas the increased secretion of Sema7A throughout differentiation indicates that Sema7A may play a role during early and late events of myogenesis. Recent studies have shown that Sema7A is involved in bone morphogenesis promoting migration of osteoblasts and enhancing fusion of chondrocytes during differentiation, two cellular processes that are required for skeletal muscle differentiation (60). Interestingly, the increased release of semaphorins could also serve an important paracrine function (59, 61). Several studies have shown that semaphorins are important regulators of cell migration, neurogenesis and angiogenesis, processes essential for skeletal muscle development and regeneration (57, 62). In a recent work by Tatsumi and colleagues it was speculated that the increased expression of Sema3A observed in satellite cells in response to injury-induced muscle regeneration could serve a paracrine function regulating muscle innervation (59). Moreover, Sema6C was found to localize primarily at neuromuscular junctions in vivo and the expression of Sema6C was reduced in response to denervation of skeletal muscle, indicating that other classes of semaphorins could be involved in neurogenesis of skeletal muscle (63).

Class 3 semaphorins are involved in activation of MMP-2 activity in skeletal muscle

Studies have shown that proteolytic processing of class 3 semaphorins modulates their activity (64, 65). In turn, class 3 semaphorins themselves transmit some of their effects by activating a specific group of proteases, the matrix metalloproteinases (MMP) (66, 67). The growth promoting effect of Sema3A in cortical dendrites relates to activation of MMP2. The family of MMPs including MMP-2 is involved in ECM morphogenesis, release of bioactive signaling molecules from the ECM, shedding of membrane-associated proteins, and cell motility, all indispensable processes for the integrity of muscle development and regeneration (68, 69). Secretion of MMP-2 was increased during myogenesis, demonstrating the highest level during early differentiation (Fig. 5E). In order to determine whether the increased secretion
of MPP-2 correlated with increased activity of MMP-2, we carried out gelatin zymography analyzing conditioned media collected at day 0, day 2, and day 5 of differentiation (Fig. 5F). The zymograms revealed that the increased level of released MMP-2 was associated with increased protease activity of MMP-2, which reached its maximum at day 2 of differentiation. It has been demonstrated that Sema3A enhances growth of cortical dendrites via the induction of both protein expression and activity of MMP-2 (67). In contrast, Sema3A decreased both the expression and activity of MMP-3, whereas Sema3C promoted axonal growth by activating MMP-3 (66, 67). Comparing the dynamic profile of Sema3A release with the profiles of MMP-2 secretion and activity suggests that the activation of MMP-2 by Sema3A signaling might occur in other cell types than dendrites including skeletal muscle cells. Moreover, other members of the identified class 3 semaphorins could be involved in modulating the activity and response to metalloproteases. Additionally, MMPs can also function as regulators of semaphorin signaling cascades by cleavage of semaphorin proteins. Soluble Sema4D promotes angiogenesis after being shed from the membrane, a process involving MMP-14 activity (70, 71), while Sema7A was identified as a substrate of MMP-2 (72).

*Identification of hydroxyproline sites on proteins secreted by myoblasts and myotubes.*

Regulation of protein activity by posttranslational modifications such as phosphorylation and ubiquitination is well known. However, recent findings indicate that proline hydroxylation also serve an important regulatory role modulating protein activity exemplified by the transcription factor hypoxia-inducible factor (HIF) (73). The stability of the oxygen-regulated subunit HIFα is modulated by hydroxylation of a conserved proline residue in response to oxygen-level within the cells.

We have identified a total of 299 unique high-confidence hydroxyproline sites from 48 distinct secreted proteins (Supplemental Table 4). Notably 231 of these sites were derived from various collagen types with a large variation of the number of modified sites per individual protein, ranging from 1 on collagen alpha-3 type VI to over 70 on collagen alpha-1 type I (Fig. 6A). Collagens are the most abundant constituents of the ECM and the collagen triple helix structure consists of three polyproline chains.
comprising the repetitive sequence motif X-Y-Gly, where the X and Y positions are frequently occupied by proline and its posttranslational modification 4-hydroxyproline (4-Hyp), respectively. The presence of hydroxyproline residues in collagens has been known for decades and both proline and hydroxyproline are essential for formation and stabilization of the collagen triple helix structure (74). Interestingly, we found an overall down-regulation in oxidation of proline residues on day 2 of C2C12 differentiation compared to day 0 and day 5 ($p = 3\times10^{-14}$) (Fig. 6B). This could reflect an increased level of newly synthesized proteins, which have not yet been posttranslationally modified, and the formation of less rigid ECM structures to allow for the early events during fusion and formation of shorter myotubes. At terminal differentiation, in which majority of myoblasts are engaged in the formation of long multinucleated myotubes, the reorganization and stabilization of the ECM is more defined. Accordingly we observed an overall increase of the hydroxyproline residues on collagen molecules on day 5 of differentiation.

We next performed motif sequence analysis in order to identify specific linear hydroxyproline sequence motifs. Alignment of the sequences surrounding the identified 299 hydroxyproline sites exposed a motif corresponding to the canonical motif previously reported for collagen proteins (Fig. 6C, top panel). The presence of the tripeptide sequence X-P-G ($n=223$) was identified exclusively on collagens (Fig. 6C). In addition, we identified a novel hydroxyproline motif containing a tryptophan residue on the third position preceding the modified proline (Fig. 6C, middle panel). The tryptophan at -3 position was found only on peptides from fibronectin, another major compound of the ECM. Analysis of the remaining 70 hydroxyproline sites did not reveal additional sequence motifs (Fig. 6C, lower panel). These sites spread 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 serphinh1) and the metalloproteinase inhibitor 2 (Timp2). The two identified hydroxyproline motifs were found only on collagens and fibronectin, respectively, both components of the extracellular matrix. This could indicate no enzyme recognition motif requirements as it may arise merely from the need for repetitive sequences for proper function of the ECM components. In
addition, the majority of the identified hydroxyproline sites are not known as evident from the UniProt database and although beyond the scope of this study, it will be interesting to assess the role of these PTMs for the function of the corresponding secreted proteins.

CONCLUSIONS

Proteins secreted from skeletal muscle influence local muscle biology as well as have a systemic effect on other organs including brain, adipose tissue, and liver. The cross talk between tissues occurs via signaling molecules that coordinate the major cellular processes such as cell growth, differentiation, and survival. The dynamic orchestra of events underlying the proper communication between cells and tissues depends entirely on the release of secreted factors, which consequently trigger appropriate cellular responses in the target cells. The fat-muscle-liver-brain tissue cross talk is of particular interest as it can shed light on the missing links in the development of number of metabolic disorders and complications. Since the discovery that the adipose tissue is an endocrine organ, tremendous focus has been drawn to the secretory function of adipocytes and the role of adipokines in obesity, diabetes and inflammation (75, 76). Although the muscle represents the largest organ in the body, its role as an active endocrine organ that secretes factors with an auto-, para- and/or endocrine function has been recently recognized with the discovery of IL-6 (77). We applied the quantitative proteomics platform SILAC to analyze the muscle secretome resulting in identification of 635 secreted proteins from muscle cells. Not surprisingly, the diversity of identified proteins reflects well the plethora of effects they exert and the various processes they control. The identified factors have distinct molecular functions including growth factor activity, cytokine activity, enzyme activity and components of the ECM and are involved in the regulation of different cellular processes ranging from angiogenesis to neurogenesis and the inflammatory response. The overrepresentation of specific molecular functions and biological processes in our dataset underlines the significance of the skeletal muscle as a prominent secretory organ.
The dynamic secretion profile shown for 624 of the secreted factors provides valuable information for the changes of the protein expression and secretion during myoblast conversion. We compared the levels of the secreted factors at initial, medium and terminal stages of myogenesis and found 188 proteins to be significantly regulated during the course of differentiation. The regulated proteins clustered into five specific groups representing the dynamic changes occurring in the secretome. The mass spectrometry-driven platform allowed the characterization of entire families of proteins as well as their regulatory partners. The identification of all three isoforms of TGF-β1, -β2, and -β3 emphasizes another major advantage of our mass spectrometry-based proteomics analyses: the possibility to identify individual isoforms of the same protein based on identification of unique peptides found in each isoform. The number of commercially available antibodies keeps rising; however, there is no guarantee for successful functionality of these antibodies. In addition, generating antibodies recognizing specific isoforms can prove to be quite difficult and labor intensive, especially if large homology exists between protein isoforms. To verify if an antibody targets a specific isoform or recognizes all is difficult to distinguish by Western Blotting if the molecular weight of protein isoforms only differs slightly. Therefore the mass spectrometry-based proteomics is the method of choice when investigating the secreted proteins content.

We followed the dynamics of the TGF-beta superfamily members and their modulators such as LTBP s, decorin and biglycan, as well as IGF family members and their regulators, the IGFBPs. Investigation of the expression profiles of specific secreted proteins demonstrated that the level of secreted proteins is regulated posttranscriptionally at several levels. We showed several examples of mRNA expression patterns that did not correlate to their corresponding secretion profile indicating regulation by posttranscriptional mechanisms. Furthermore, analyzing the intracellular levels of semaphorins protein expression and comparing it to the dynamics of their secretion clearly illustrates that the release of secreted proteins is subjected to complex modular regulation.

One of the major strengths of mass spectrometry-based proteomics is the possibility to identify various posttranslational modifications. We identified 299 hydroxyproline sites; the majority of those were found on different collagen subtypes. However, we also identified hydroxyproline sites on many other proteins.
including several constituents and regulatory components of the ECM, which could serve as important modifications to confer protein stability and interaction with other proteins. The identification of hydroxyproline on such a large number of peptides could be an indication for a more important regulatory role of this type of modification than previously anticipated.

The large number of identified secreted proteins in this study provides a foundation for future research to understand the regulation of the myogenic program. Understanding the role played by individual secreted proteins in skeletal muscle development could offer valuable clues in the search for new target proteins with therapeutic potential for treatment of various muscular disorders, including muscular dystrophies characterized by muscle atrophy and impaired regeneration. In addition, the secreted skeletal muscle proteins could also provide valuable candidate in the search for novel proteins exhibiting a systemic effect, modulating the activity of other organs. It is anticipated that the identification of secreted proteins from skeletal muscle cells could allow greater insight into development of therapeutic strategies for patients with metabolic- and muscle-related pathologies.

ACKNOWLEDGEMENTS

We thank all members from CIM and CEBI for useful discussions. We are grateful to Dr. Carmen de Hoog (University of British Columbia, Vancouver, British Columbia, Canada), Dr. Jesper Olsen (The Novo Nordisk Foundation Center, Copenhagen, Denmark) and Dr. Joern Dengjel (Freiburg Institute for Advanced Studies, Germany) for the critical reading of the manuscript. This work was supported by a grant from the Lundbeck Foundation, the Novo Nordisk Foundation, the Danish Natural Science Research Council and the Danish Medical Research Council. The Centre of Inflammation and Metabolism is supported by a grant from the Danish National Research Foundation.


secreted proteins during the differentiation of 3T3-L1 preadipocytes to adipocytes. *Mol Cell Proteomics* **1**, 213-222


FIGURE LEGENDS

**Figure 1. Myogenic commitment of C2C12 myoblasts using dialyzed FBS (dFBS).** (A) Experimental overview. C2C12 cells were grown to confluence (day0) and myoblast differentiation was induced reducing the amount of dFBS to 2%. Conditioned media (CM) were collected on day 0, day 2, and day 5 of differentiation of C2C12 cells. “Day 0 cells” were cultured using regular Arg0 and Lys0, “day 2 cells” were cultured using Arg6 and Lys4 and “day 5 cells” were cultured using Arg10 and Lys8. CM were concentrated by ultrafiltration and proteins separated by 1D-gel electrophoresis. Excised gel bands from the entire lane were subjected to in gel-digestion and subsequently analyzed on an LTQ-Orbitrap-XL mass spectrometer. (B and C) Efficiency of differentiation was verified by examining the expression of muscle markers either by CLSM (B) or Western blotting (C). (B) Immunostaining of C2C12 cells at day 0, 2, and 5 of differentiation with an antibody against Troponin T and visualized by green fluorescence tagged secondary antibody. Nuclei were counterstained with Sytox. Scale bar represents 95 µm. (C) Western blot analysis of whole cell extracts prepared of C2C12 myoblasts at day 0, 2, and 5 of myogenesis using indicated antibodies.

**Figure 2. Functional analyses of identified secreted proteins from C2C12 skeletal myoblasts.** (A and B) Identified secreted proteins from C2C12 myoblasts were compared to a reference list of secreted proteins obtained from the entire list of IPI entries. For selected Biological Process (A) and Molecular Function (B) the frequency of the term in the data and the IPI database are illustrated. Only significantly overrepresented categories (p<0.05) are shown.
Figure 3. Dynamic and functional characterization of differentially regulated secreted proteins. Cluster analysis of regulated proteins according to their secretion profile. The heat map summarizes the occurrences of the indicated Biological Process and Molecular Function GO terms within each cluster. Only GO terms with five or more occurrences in at least one cluster are shown.

Figure 4. Secretion and expression of TGF-β1, -β2, and -β3 and IGF-1 and -2 are differentially regulated during the course of skeletal muscle differentiation. (A) Dynamic secretion profiles including ratios obtained by mass spectrometry analyses for the following proteins TGF-β1, -β2, -β3, LTBP3, IGF-1, IGF-2, IGFBP-2, and IGFBP-4 during the course of C2C12 differentiation. (B) mRNA expression of TGF-β and IGF isoforms at day 0, day 2, and day 5 of C2C12 differentiation analyzed by qPCR. The level of mRNA expression of selected genes was normalized to the mRNA level of TBP. Gene expression data are given relative to day 0 and data are shown as mean ± S.D. (n = 3).

Figure 5. Differential secretion of different subclasses of semaphorin proteins during the process of myogenesis. (A) Quantitative profiles of semaphorin secretion during the course of C2C12 differentiation including their ratios applying quantitative proteomics. (B and C) Western blot analysis of selected semaphorin proteins in conditioned media (B) or whole cell extracts (C) at day 0, day 2, and day 5 of C2C12 differentiation. (D) mRNA expression for Sema3A, Sema3E, Sema6A, and Sema7A by real time Q-PCR. Gene expression data was normalized to the expression of TBP and presented as mean ± S.D. (n = 3) relative to day 0. (E) Dynamic profile of MMP-2 secretion as observed by MS (F) Gelatin zymography assessing MMP-2 activity in conditioned medium from C2C12 myoblasts at day 0, day 2, and day 5 of differentiation (proMMP-2 72 kDa and active MMP-2 62 kDa). * Represents the glycosylated form of Sema7A.
Figure 6. Analyses of hydroxyproline sites identified from proteins secreted by C2C12 myoblasts and myotubes. (A) Distribution of hydroxyproline sites on the different collagen subtypes. (B) Distribution of the regulation (SILAC ratios) of hydroxyproline sites at day 2 (mean -1.49) and 5 (mean -0.32) of differentiation. (C) Hydroxyproline motif analysis, sites matching the prototypical collagen hydroxyproline motif (top panel), sites with a tryptophan in the -3 position, all on fibronectin (middle panel) and the remaining sites not matching significant shared motif.

Supplemental data:
Supplemental Tables 1-4, Supplemental annotated hydroxyproline MS/MS spectra
Henningsen et al., Figure 1

**A**

Mix conditioned media 1:1:1 according to protein concentration

Reduce the volume of the combined media via ultrafiltration

1D-gel electrophoresis

Digestion with trypsin

Protein identification and quantitation by LC-MS/MS

---

**B**

**α-TroponinT**

Day 0

Day 2

Day 5

---

**C**

<table>
<thead>
<tr>
<th>Protein</th>
<th>Day:</th>
<th>0</th>
<th>2</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>MHC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TroponinT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myogenin</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β-actin</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Henningsen et al., Figure 3
Henningsen et al., Figure 4

A

B

Protein Names | Ratio Day 2 | Ratio Day 5 | Cluster Annotation
---|---|---|---
TGF-β1 | 2.4 | 10.2 | 4
TGF-β2 | 7.4 | 15.8 | 4
TGF-β3 | 5.0 | 15.3 | 4
LTBP-3 | 3.0 | 11.9 | 4
IGF-1 | 1.4 | 13.1 | 2
IGF-2 | 24.9 | 6.9 | 2
IGFBP-2 | 0.8 | 63.6 | 3
IGFBP-4 | 0.3 | 6.1 | 1
Henningsen et al., Figure 5

A

B

C

D

E

F

Day 0 Day 2 Day 5

Relative Protein Level

Day 0 Day 2 Day 5

Relative Protein Level

Day 0 Day 2 Day 5

Relative Protein Level

Day 0 Day 2 Day 5

Relative Protein Level

Protein Ratio Cluster

Names Day 2 Day 5 Annotation

Sema3A 5.0 1.5 2
Sema3E 5.8 1.1 2
Sema6A 5.5 3.7 2
Sema7A 2.8 8.3 4

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A

Sema3A

Sema3E

Sema6A

Sema7A
Henningsen et al., Figure 6

**A**

Bar chart showing hydroxyproline sites for different alpha chains.

**B**

Graph showing log2(SILAC-ratio) for Day 2 and Day 5.

**C**

Images depicting protein sequences with hydroxyproline sites.