Secretome profiling of primary human skeletal muscle cells

Sonja Hartwig[1,11], Silja Raschke[2,11], Birgit Knebel[1,11], Mika Scheler[3,11], Martin Irmler[3,11], Waltraud Passlack[1,11], Stefan Muller[4], Franz-Georg Hanisch[4], Thomas Franz[5], Xinping Li[5], Hans-Dieter Dicken[6], Kristin Eckardt[2,11], Johannes Beckers[3,7,11], Martin Hrabe de Angelis[3,7,11], Cora Weigert[8,9,11], Hans-Ulrich Häring[8,9,11], Hadi Al-Hasani[1,11], D. Margriet Ouwens[1,10,11], Jürgen Eckel[2,11], Jorg Kotzka[1,11], Stefan Lehr[§1,11]

[1]Institute of Clinical Biochemistry and Pathobiochemistry, German Diabetes Center, Duesseldorf, Germany; [2]Paul-Langerhans-Group for Integrative Physiology, German Diabetes Center, Duesseldorf, Germany; [3]Institute of Experimental Genetics, Helmholtz Zentrum Muenchen, German Research Center for Environmental Health (GmbH), Neuherberg, Germany; [4]Center for Molecular Medicine Cologne, Cologne, Germany; [5]Max Planck Institute for Biology of Ageing, Cologne, Germany; [6]Multimedia Center, Heinrich-Heine-University Duesseldorf, Duesseldorf, Germany; [7]Experimental Genetics, Center of Life and Food Sciences Weihenstephan, Technische Universität München, Freising-Weihenstephan, Germany; [8]Division of Endocrinology, Diabetology, Angiology, Nephrology, Pathobiochemistry and Clinical Chemistry, Department of Internal Medicine, University Tuebingen, Germany
[9]Institute for Diabetes Research and Metabolic Diseases of the Helmholtz Zentrum Muenchen at the University of Tuebingen, Tuebingen, Germany
[10]Department of Endocrinology, Ghent University Hospital, Ghent, Belgium
[11]German Center for Diabetes Research (DZD)

Address for correspondence:

[§]Dr. Stefan Lehr
Institute of Clinical Biochemistry and Pathobiochemistry
German Diabetes Center
Auf'm Hennekamp 65
40225 Duesseldorf, Germany
Tel.: +49 211 3382 531
e-mail: stefan.lehr@ddz.uni-duesseldorf.de

Abstract

The skeletal muscle is a metabolically active tissue that secretes various proteins. These so-called myokines have been proposed to affect muscle physiology and to exert systemic effects on other tissues and organs. Yet, changes in the secretory profile may participate in the pathophysiology of metabolic diseases.

The present study aimed at characterizing the secretome of differentiated primary human skeletal muscle cells (hSkMC) derived from healthy, adult donors combining three different mass spectrometry based non-targeted approaches as well as one antibody based method. This led to the identification of 548 non-redundant proteins in conditioned media from hSkmc. For 501 proteins, significant mRNA expression could be demonstrated. Applying stringent consecutive filtering using SignalP, SecretomeP and ER_retention signals databases, 305 proteins were assigned as potential myokines of which 12 proteins containing a secretory signal peptide were not previously described.

This comprehensive profiling study of the human skeletal muscle secretome expands our knowledge of the composition of the human myokinome and may contribute to our understanding of the role of myokines in multiple biological processes.

Keywords

Myokines, Combined Proteomic Profiling, Two-Dimensional Gel Electrophoresis, Mass Spectrometry

1 Introduction

The search for an “exercise factor” that communicates the energy demand of the working muscle to other organs added the skeletal muscle to the list of secretory active tissues [1]. These muscle derived secretory proteins, termed myokines, can exert auto-, para- or endocrine effects, founding a new paradigm for understanding how muscles communicate with other organs, such as adipose tissue, liver, or pancreas [2]. The discovery of the skeletal muscle as the origin of cytokines released during physical activity focuses the research on these myokines to understand the beneficial effects of exercise on metabolic disturbances like insulin resistance or type-2-diabetes. One prominent example is interleukin 6 (IL6) which is released into circulation from contracting muscles influencing glucose as well as lipid metabolism [3-5]. While exercise is still considered as one of the major contributors to the release of proteins from the muscle, other physiological and pathophysiological conditions that induce the release of myokines have been identified [6-9]. Recent studies of the past few years indicate that the skeletal muscle secretome comprises several peptides and proteins including many yet unidentified biological active factors, being involved in various biological processes [2]. Thus, the identification of the complex nature of the human myokine pattern may contribute to the understanding of various physiological crosstalks between the muscle and other organs as well as disease development and its prevention.

Complementary mass spectrometry based proteomic profiling technologies have contributed to the identification of hundreds of proteins found in conditioned media (CM) from multiple tissue derived cell lines, including the rodent skeletal muscle cell lines L6 and C_2C_{12} [9-12]. Nevertheless, little is known about the secretome of adult

human skeletal muscle [13-16]. Especially the knowledge about the secretome of mature, thus fully differentiated hSkMC obtained from healthy adults is still sparse.

In the current study we have utilized an integrated genomic and proteomic approach to define the secretome of differentiated hSkMC derived from adult donors. We applied a combination of three different mass spectrometry methods (*2D-PAGE MALDI-MS, SDS-PAGE LC-ESI-MS/MS and LC/MS^{E}*) as non-targeted approaches and multiplex immunoassay (MIA) as a targeted approach to enable a comprehensive analysis of the entire human muscle secretome and the expression of identified proteins was validated by transcriptomics. Our study expands our knowledge of the human skeletal muscle secretome by the identification of 12 potential myokines previously not described as muscle-derived factors.

2. Experimental Procedures

2.1. *Culture and preparation of conditioned media (CM) from primary human skeletal muscle cells –* *Cohort 1*. Primary hSkMC from five healthy caucasian donors (three males and two females; 31 ± 6 yrs) were used for the preparation of CM. The cells were supplied as proliferating myoblasts (PromoCell, Heidelberg, Germany) and cultured as described previously [17]. Briefly, myoblasts were seeded in six-well plates ($1x10^5$ cells/well), and were cultured in α-modified Eagle's (αMEM)/ Ham's F-12 medium containing skeletal muscle cell growth medium supplement pack (PromoCell, Heidelberg, Germany) up to near-confluence. The cells were then differentiated to myotubes in αMEM containing 2% (v/v) horse serum (Gibco, Berlin, Germany) for 5 days. On day 6 of differentiation, cells were washed twice with PBS, serum-free medium was added and CM for proteomic profiling was collected after 24h. In total more than 1l of CM was centrifuged at 85,000 xg and concentrated to mg/ml range using Amicon^{TM} Ultra 15 centrifugal filter devices (Millipore, Billerica, USA) with a cut-off mass of 3 kDa. Protein concentrations were determined using Advanced Protein Assay (Tebu-bio, Offenbach, Germany) according to manufacturers instructions and aliquots of pooled CM were stored at -80°C for non-targeted proteomic analysis.

Cohort 2. For targeted proteomic profiling (multiplex analysis (MIA)) and expression profiling, hSkMC obtained from percutaneous needle biopsies performed on the lateral portion of quadriceps femoris (vastus lateralis) muscle of twelve healthy subjects (three females, nine males; 25.6 ± 4.4 yrs) were used. They gave informed written consent to the study and the protocol was approved by the Ethics Committee of the University of Tuebingen (Number: 179/97). Cell culture and collection of CM was performed as described above.

2.2. *Non-targeted proteomic profiling*

For non-targeted proteomic profiling three independent methods were used, i.e. 1D-PAGE/LC/ESI-MS/MS, data independent LC-MS/MS (LC/MS^E) and 2D-PAGE/MALDI-MS.

2.2.1. *Validation of CM collection procedure* - To validate collection of CM, murine C_2C_{12} skeletal muscle cell model system was used. Proliferating C_2C_{12} [18] myoblasts were seeded in six-well culture dishes at a density of 1×10^5 cells/well and cultured to near-confluence in Dulbecco's modified Eagle's medium (DMEM) containing 10% (v/v) fetal calf serum (FCS) and 100 µg/ ml streptomycin (Invitrogen, Paisley, UK). The cells were then differentiated in DMEM containing 2% (v/ v) horse serum for 5 days. On day 6 cells were washed twice with PBS and three times with fresh serum-free DMEM, serum-free medium was added and CM for proteomic profiling was collected after 24h. CM was centrifuged at 85,000 xg and concentrated to mg/ml range using Amicon™ Ultra 15 centrifugal filter devices (Millipore, Billerica, USA) with a cut-off mass of 3 kDa. Aliquots of concentrated CM were subjected to Two-Dimensional Difference Gel Electrophoresis 2D-DIGE. 2D-DIGE was performed as described previously [19]. Briefly, labeled samples (50 µg each) were separated in the first dimension by isoelectric focusing (IEF) on a MultiPhor II electrophoresis unit (Amersham Biosciences) using IPG strips (24 cm, pH 4−7 linear), followed by SDS-PAGE on 12.5% polyacrylamide gels (24 cm × 18 cm) using an EttanDalt 12 system (Amersham Biosciences). Subsequently, images of protein pattern were acquired using a Typhoon 9400 (Amersham Biosciences) laser scanner according to the manufacturer's recommendations (resolution of 100 µm, photomultiplier tube of 550 V).

2.2.2. *1D-PAGE and protein identification by liquid chromatography (LC)-MS/MS* – For LC-MS analysis aliquots of concentrated CM (20 µg) were first separated by one-dimensional SDS-PAGE (5% to 15% gradient gel). Extracted peptides derived from 24

equally sized gel slices were subsequently analysed as described previously [20]. Liquid chromatography (LC)-MS data were acquired on a HCT ETD II ion trap mass spectrometer (BrukerDaltoniks, Bremen, Germany). Raw data were processed using Data Analysis 4.0 (BrukerDaltoniks, Bremen, Germany) and xml formatted peak lists were transferred to Proteinscape 3.0 (BrukerDaltoniks, Bremen, Germany). MASCOT 2.2 (Matrix Science Ltd, London, UK) was used to search a composite decoy database which was built from SwissProt_2011 (532146 sequences; 188719038 residues; 20249 human sequences). The composite database was generated with the Perl script *makeDecoyDB* (BrukerDaltoniks, Bremen, Germany) which added a randomized sequence and a tagged accession number for each entry. The tagged accessions were used for the calculation of false positive rate in Proteinscape 3.0. Searches were submitted via Proteinscape3.0 and the following parameter settings: enzyme “trypsin”, species “human”, fixed modifications “carbamidomethyl”, optional modifications “Methionine oxidation” and missed cleavages “2”. The mass tolerance was set to 0.4 Da for peptide and fragment spectra. Protein lists were compiled in Proteinscape3.0. Peptide hits were accepted when the ion score exceeded a value of 20. Protein hits required at least one peptide hit exceeding a peptide score of 40. In addition, the hits to decoy entries were used to calculate a minimal protein score which is required to keep the false positive rate below 2% on the protein level.

2.2.3 ***In-solution digestion of total protein sample and protein identification by data independent LC-MS/MS –*** 20µg of each sample was dissolved in 20 µl of 25 mM ammonium hydrogen carbonate. 2 µl of 1% RapiGest (Waters Corporation, Milford, USA) was added for denaturing the proteins. The protein solution was incubated at 80°C for 10 min on a Thermo-mixer. After the addition of 1 µl aliquot of 50 mM DTT the solution was heated at 60°C for 15 min. The protein solution was then cooled down to room temperature and centrifuged. After the addition of 1 µl aliquot of 150 mM

iodoacetamide in 25 mM ammonium, the solution was stored in the dark at room temperature for 30 min. The tryptic digestion was performed by adding Trypsin Gold mass spectrometry grade (Promega, Madison, MI, USA) at a 1:50 (w/w) ratio and incubating at 37°C overnight. 1 µl of 37% HCl was added to adjust the pH below 2. After being vortexed and centrifuged at 13,000 xg for 30 min, the supernatant was collected and transferred to a clean microcentrifuge tube. The tryptic digest was desalted with PepClean™ C-18 Spin Column (Thermo Scientific, Waltham, MA, USA) according to the manufacturer's instruction, and eluted with 40 µl of 50% acetonitrile (ACN)/ 0.1% formic acid. After drying in a SpeedVac, the digest was re-suspended in 100 µl of 0.1% formic acid.

Protein identification was performed with a Xevo Q-Tof (Waters Corporation, Milford, USA) coupled with a nanoACQUITY UPLC™ (Waters Corporation, Milford USA). 3-4 µl of tryptic digest was directly loaded into an analytical column of 75 µm x 150 mm C18 BEH 1.7 µm (Waters Corporation, Milford USA) with 3% formic acid for 25 min. The loading flow rate was 400 nl/min. The peptides were eluted with a gradient of 3% to 55% ACN in 0.1% formic acid over 180 min at a flow rate of 400 nl/ min. The Xevo Q-Tof was operated in LC/ MS^E mode over the m/z range of 50 to 1800 in nano electrospray mode. The capillary, sample cone, extraction cone and collision energy were 3.3 kV, 25.0 V, 2.0 V, and 6.0 V respectively. During elevated energy scan, the collision energy was ramped from 15 V to 35 V. Glu-fibrinopeptide B of m/z 785.84 was used as Lock Mass for mass correction. At least 3 replicates were analyzed. Data was collected using MassLynx™ 4.1 and processed and searched using ProteinLynx™ Global Server 2.5.2. (Waters Corporation, Milford USA). The following parameters were used for database search: enzyme "trypsin", minimal fragments ion per peptide matched ''3'', minimal fragments ion per protein matched ''7'', missed cleavages ''1'', fixed modification ''carbamidomethyl cystein'', variable modifications ''acetyl N-term,

oxidation Methionine'', peptide tolerance ''automatic'', fragment '' automatic'', false positive rate ''4%''. The human database was from Uniprot release knowledgebase_2012_08.

2.2.4. Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) and protein identification by MALDI-MS – For 2D-PAGE analysis aliquots (150 µg) of CM derived from primary human skeletal muscle were diluted 1:3 in a buffer containing 25 mM Tris, 4% CHAPS (w/ v), 7 M urea and 2 M thio-urea and subjected to 2D-PAGE (pH 4-7 and pH 6-9) with subsequent protein identification by MALDI-MS as described previously [20]. MS peptide mass fingerprint and fragment spectra from each individual spot were combined and used to search a human sub-set of Swiss-Prot (Sprot_2011; 532,146 sequences; 188,719,038 residues; 20,249 human sequences) non-redundant database using Mascot search engine (Version 2.2, Matrix Science Ltd, London, UK) in consideration of the following settings: enzyme "trypsin", species "human", fixed modifications "carbamidomethyl", optional modifications "Methionine oxidation" and missed cleavages "1". Mass tolerance was set to 50 ppm for peptide and 0.7 Da for fragment spectra. Using these settings, a combined mascot score of greater than 70 was taken as significant ($p<0.01$). Calculated pI and molecular mass data were obtained by Mascot. For peptides matching to different isoforms or multiple members of a protein family, we used the following reporting criteria: The experimental pI and molecular mass taken from the 2D-gels were compared with the theoretical data of the different isoforms/ protein members. If no conflicts in molecular mass or pI were found, the isoform/ protein member with the highest mascot score was reported. For verifying the results each protein spot was picked and identified from at least three independent 2D-gels.

2.3. Targeted proteomic profiling

2.3.1. Multiplex immunoassays (MIA)- Commercially available human Bio-Plex Pro™ multiplex bead-based immunoassay panels (Biorad, Hercules, CA, USA) were used to detect cytokines and chemokines in CM. Analysis was performed using a Bioplex 200 suspension array system (Biorad, Hercules, CA, USA) according to manufacturer's instructions. Protein concentrations were calculated from the appropriate optimized standard curves using Bio-Plex Manager software version 6.0 (Biorad, Hercules, CA).

2.4. *Microarray analysis* – For RNA extraction, cell lysates were disrupted using a QIAshredder, followed by purification of total RNA using a RNeasy mini kit (Qiagen, Hilden, Germany) including DNAse digestion according to the manufacturer´s instructions. RNA quality was evaluated using an Agilent 2100 Bioanalyzer and only high quality RNA (RIN>8) was used for microarray analysis. For this total RNA (150 ng) was amplified using the Ambion WT Expression Kit and the WT Terminal Labeling Kit (Affymetrix, Freiburg, Germany). Amplified cDNA was hybridized on Affymetrix Human Gene ST 1.0 arrays containing about 28,000 probe sets. Staining (Fluidics script FS450_0007) and scanning was done according to the Affymetrix expression protocol. Expression console (Affymetrix, Freiburg, Germany) was used for quality control and to obtain annotated normalized RMA gene-level data (standard settings including sketch-quantile normalisation, annotation file mogene-1_0-st-v1.na32.mm9). Background cut off was determined based on the 15000 most abundant targets. Statistical analyses were performed by utilizing the statistical programming environment R (R Development Core Team, [21]) implemented in CARMAweb [22]. GO term and pathway enrichment analyses ($p<0.01$; adj $p<0.05$) were done with Ingenuity software ($p<0.05$). Array data was submitted to GEO (accession number GSE45473) and a link for referees was created

(http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=xxmlrcugwmiumly&acc=GSE45473).

2.5. *Prediction and annotation of secretory proteins* – Secretory protein prediction and functional annotation was done by using different, independent ways. First protein information of all identified proteins was extracted from the Swiss-Prot/TrEMBL database (http://www.expasy.ch/sprot/). Gene names were used for comparison with gene expression data derived from micro array analysis, and to screen the secreted protein database (SPD; http://spd.cbi.pku.edu.cn) and gene ontology (GO; http://geneontology.org). Subsequently, proteins were analyzed by SignalP 3.0 (http://www.cbs.dtu.dk/services/SignalP/), SecretomeP 2.0. (http://www.cbs.dtu.dk/services/SecretomeP/), Prosite (http://prosite.expasy.org/), and literature for validation as potentially secreted proteins. To assign proteins as putative secretory protein, and thus potential myokine, the passing prediction of thresholds for SignalP 3.0 (*D*score cut-off: 0.43) predicting a signal peptide or SecretomeP 2.0 (NNscore cut-off: 0.5) predicting non-classical secretory proteins without signal peptide were set as mandatory.

3. Results

3.1. *Proteomic profiling of differentiated human skeletal muscle secretome* -One crucial point in analyzing cellular secretomes collected as CM is cross contamination by cell culture media components, especially serum components. Therefore, we have validated our washing and CM collection procedure by comparative 2D-DIGE (pH 4-7) analyzes of CM obtained from differentiated C_2C_{12} mouse myotubes (Suppl. Fig. 1A) and DMEM supplemented with 2% horse serum (Suppl. Fig 1B). These analyses revealed specific protein signatures in both conditions and showed no similarities between CM and medium supplemented with serum.

The secretome of differentiated hSkMC of healthy donors was analyzed by a combination of three complementary non-targeted profiling approaches, i.e. 2D-PAGE MALDI-MS and two different tandem mass spectrometry approaches (SDS-PAGE LC-ESI-MS/MS and LC/MS^E) and one targeted approach, i.e. multiplex immunoassay (Fig. 1). 2D-PAGE (pH 4-7 and pH 6-9, Suppl. Fig. 2) resolved more than 1,200 protein spots and resulted in the identification of 570 protein spots. These could be assigned to 171 non-redundant proteins (Tab. 1, Suppl. Tab.1). The LC-ESI-MS/MS from analysed replicate gels (24 gel slices/LC fractions) identified 405 non-redundant proteins (Tab. 1, Suppl. Tab. 1). Data-independent LC-MS^E from total protein lysate of CM revealed 372 proteins (Tab. 1, Suppl. Tab. 1). Collectively, we identified 530 non-redundant proteins by combining these non-targeted approaches (Tab. 2). Screening the CM for chemokines and cytokines using a multiplex analysis system added another 18 proteins to the list (Tab. 2, Suppl. Tab. 1). Thus in total we identified 548 unique proteins in CM collected from differentiated hSkMC.

3.2. *Computational filtering of identified proteins to reveal secretory proteins* - To assess whether the 548 non-redundant proteins are potentially secretory proteins

consecutive stringent filter methods were applied. To validate, if the 548 potential myokines *per se* could originate from skeletal muscle cells, we first analysed wether they are expressed in hSkMC. Table 2 shows that the expression of 501 out of the 548 proteins was confirmed by significant RNA expression levels in hSkMC. The expression level of 11 of the cytokines/chemokines detected by the MIA was below the set background cut-off of the microarray expression analysis, which may be due to the low basal expression level of these proteins. These 501 proteins were further filtered by web-based bioinformatic tools (SignalP3.0, SecretomeP2.0, ExPaSy Prosite) to predict proteins that could potentially be secreted. Using these filters, 169 proteins were found to contain secretory signal peptides (SP+), 136 follow non-classical alternative signal peptide-independent mechanisms (SP-), while 196 identified proteins did not comply with our computational filtering (NP) (Tab. 2). The complete list of identified proteins including essential information is available in our tissue specific secretome database that can be viewed at www.diabesityprot.org.

3.3. ***Literature comparison of the identified myokines*** **-** To rank our results according to recent knowledge of the skeletal muscle secretome, we first compared protein entries of the secreted protein database (SPD) with the gene expression profile derived from adult skeletal muscle (13,996 non-redundant transcripts). 1,161 transcripts are coding for proteins annotated as potentially secretory, corresponding to a sub-set of 106 proteins of our 298 predicted putative myokines resulting from the non-targeted mass spectrometry based analysis. Aligning our data with Bortoluzzi et al. [13], who reported an *in silico* skeletal muscle secretome comprising of 319 potentially secreted proteins, indicated that 61 of these proteins were confirmed by our combined proteomics and transcriptomics analyses.

Comparing our data from differentiated hSkMC from adult donors with a similar integrated genomic and proteomic approach recently published by LeBihan et al. [16], who investigated the secretome of differentiated hSkMC from a 5-day-old infant, indicated a consistency of 8,058 transcripts, whereas 1,132 transcripts are solely found in neonatal and 5,938 in adult skeletal muscle cells. On the protein level 355 proteins which were confirmed on RNA expression level, were identified in both studies, 215 out of the 305 proposed myokines are shared by both studies. In a very recent publication [23] a targeted approach using cytokine antibody arrays was applied to identify myokines released by resting and contracting hSkMC. In this study 116 myokines were reported of which 48 myokines were shown to be regulated by contraction that was induced by electrical pulse stimulation. A comparison with our study reveals an overlap of 30 factors that were detected in both reports.

Further literature search for previous global proteomic muscle secretome profiling studies, including rodent [9-12] and human [13-16] models, revealed that in total 425 of the previously described proteins were common with our study of which were 260 annotated as SP+ or SP- protein (Tab. 3, Suppl. Tab. 1). Thus, to the best of our knowledge, 12 proteins containing a secretory signal peptide (SP+) in our study have not been reported in previous approaches, and can therefore be considered as potential novel myokines (Tab. 4).

Annotation of 305 identified SP+ or SP- proteins by GO data base searches confirmed the secretion potential of these proteins (EASE score 8.2E-37) and their role in signaling processes (EASE score 4.4E-38). Many proteins also contain EGF-like conserved protein regions with high significance (EASE score 1.9E-12). Furthermore, GO data base analyses of proteins without SP+ or SP- sites (NP) identified a total of 135 proteins with MyoD and 95 with MyoGNF1 promoter consensus sequence.

4. Discussion

Detailed characterization of the human skeletal muscle secretome is essential to understand the role of secreted myokines for communication with other organs and their proposed relevance related to disease pathophysiology or disease prevention. In this context recently several attempts have been made to elucidate the complex nature of the muscle secretome [9-16, 23]. To analyze the composition of the secretory output of differentiated hSkMC we have applied an integrated genomic and proteomic approach, resulting in the identification of 305 myokines including 12 novel ones.

Composition of secretomes highly depend on the selected cell system (e.g. cell-lines, primary cells, tissue explants) as well as on the used culturing and collection conditions. In the present study, we have utilized differentiated, primary hSkMC's derived from different donors. Although, this strategy might not entirely reflect the in vivo situation, using primary human skeletal muscle cells may still be considered as the most optimal method to unravel the physiological skeletal muscle secretome. Thereby, one major challenge is to prevent or discriminate contaminating proteins in order to identify genuine secreted proteins. The secretome of skeletal muscle tissue is highly complex while concentration of myokines is potentially very low (pg to ng/ml). Thus, for example serum-derived proteins might interfere with myokine identification or could produce false-positive results by assigning contaminating proteins as myokines. Therefore, we have utilized 2D-DIGE analysis (Suppl. Fig. 1) to demonstrated that the secretome samples used for proteomic profiling did not contain significant amounts of contaminating proteins introduced due to the operational procedure. To enable a more general analysis of the mature human skeletal muscle secretome, CM from differentiated hSkMC obtained from five different donors, accounting for the biological variability, was pooled.

Up to now no biochemical technique exists that can alone efficiently separate and consistently detect the total protein composition of tissue specific secretomes. An optimized solution is to join different technical approaches [24] combining complementary proteomic profiling approaches to resolve the limitations given by each single method. Especially, utilization of orthogonal different sample preparation, *i.e.* gel-based and gel-free approaches has the potential to improve coverage of the complex nature of the analysed secretome. Together with the huge sample amount applied in our study, this strategy provides a more comprehensive approach to catalogue the complexity of the human skeletal muscle secretome. Hence, we utilized three different non-targeted proteomic approaches (2D-PAGE MALDI-MS, SDS-PAGE LC-ESI-MS/MS, LC/MS^{E}) supplemented by one targeted multiplex immunoassay to allow investigation of as many different classes of secreted proteins as possible. Comparing the three MS approaches revealed that only 19% (102) of proteins were found by all three methods. 35 proteins were detected only by 2D-PAGE/MALDI-MS, 99 only by 1-D/LC-ESI-MS/MS and 80 exclusively by LC/MS^{E} (Tab.1). In line with our previously published adipokine study [22], these results impressively illustrate the importance to combine different techniques to facilitate a comprehensive analysis. Additionally, antibody based targeted analysis exclusively detected further 18 proteins, which belong to the cytokine and chemokine families including IL6, leukemia inhibitory factor and IL8. Due to their low expression and concentrations and physically features usually detection of many cytokines by broad MS- based screening approaches are limited. Furthermore, the detection of well known myokines by the immunoassays demonstrate, that our model system is appropriate for secretome analysis.

Although we have demonstrated that the secretome samples did not contain significant amounts of contaminating proteins, we applied consecutive stringent computational filtering assigning 305 as potential myokines. Even if this strategy utilizes mainly

theoretical analysis tools and therefore includes the risk to lose interesting target proteins without further validation, it is indispensable to identify proteins most likely to be secreted. Not to lose potential myokines released by unconventional mechanisms a priori, we disclaimed to utilize common experimental strategies monitoring secretion processes, like Brefeldin A treatment blocking major ER/Golgi dependent as well as independent secretion pathways [25].

In order to put our results into the context of recent knowledge of skeletal muscle secretomes we have compared the data with available publications in this field. One of the most comprehensive studies is reported by LeBihan et *al.* [16] describing the secretome of differentiated neonatal skeletal muscle cells by an integrated genomic and proteomic approach. In contrast to this report we have analysed the secretome that was derived from differentiated hSkMC obtained from adult donors. Global gene expression profiling indicates a significant difference between this and our approach, displaying an overlap of approximately 50%. This overlap demonstrates the impact of the selected model system for the outcome of corresponding secretome analysis and underlines the importance to use different models and methodologies in order to achieve a comprehensive proteomic profiling. In order to address the secretory potential of our skeletal muscle cell model, we have first used a knowledge based approach by correlating our transcriptome data with annotated secretory proteins included in the SPD database. This comparison proposed 1,161 proteins as potentially secreted. Matching this theoretical secretome with the proteins identified by us, shows an overlap of only 106 proteins, whereas 199 predicted myokines were not found in this database. Alignment of our identified putative myokines with the *in silico* secretome (319 proteins) reported by Bortoluzzi et *al.* [13] shows an overlap of 33%, which strongly illustrates the limitations of studies using exclusively theoretical examinations. Nevertheless in our study, we confirmed the mRNA origin of secreted proteins from

412 hSkMC and confirmed the secretory potential according to the presence of classical
413 and non classical signal peptides. Of further interest is the observation that proteins of
414 our study that do not contain potential signaling peptides bare a transcription factor
415 MyoD or MyoGNF1 promoter consensus sequence and thus are likely to be involved
416 in skeletal muscle development, maturation, and functional maintenance and plasticity
417 [26].
418 Comparison of our data in the context of other previously published proteomic studies,
419 show that 85% (i.e. 265 proteins) of identified proteins were already described (Suppl.
420 Table 1). However, our approach has identified 12 additional proteins containing a
421 secretory signal peptide, which we consider as potential novel myokines secreted by
422 human skeletal myotubes.
423 Many of these potential novel myokines have been associated to various biological
424 processes already. Vitronectin for example is a cell adhesion and spreading factor,
425 which is known to interact with glycosaminoglycans and proteoglycans. It acts as a
426 potent matricellular factor in tissue remodeling or in tumors, coordinating cell migration
427 with pericellular proteolysis and growth factor signaling [27]. Ephrin type-A receptor 4
428 plays an important role in the development of the nervous system controlling different
429 steps of axonal guidance. It belongs to the receptor tyrosine kinase family and
430 modulates cell morphology and integrin-dependent cell adhesion through regulation of
431 the Rac, Rap and Rho GTPases activity [28]. Granulocyte colony stimulating factor
432 (GCSF/CSF3) is a well known cytokine that act in hematopoiesis by controlling the
433 production, differentiation, and function of granulocytes and monocyte-macrophages
434 [29]. Recombinant human GCSF is used with certain cancer patients to accelerate
435 recovery from neutropenia after chemotherapy. Other potential myokines are
436 described in the context of tumor differentiation and progression. For example the
437 tumor suppressor Retinoic acid receptor responder protein 1 (RARRES1) was shown

to play a role in tumor differentiation and staging in colorectal adenocarcinoma, in which down-regulation of RARRES1 is related to disease progression [30]. Collagen triple helix repeat containing 1 (CTHRC1) affects vascular remodeling, bone formation and developmental morphogenesis. Recent studies indicate a pivotal role of CTHRC1 in pancreatic cancer progression and metastasis [31]. Plasminogen activator urokinase (PAU) play an important role in the plasminogen-plasmin system and are described to be involved in a variety of cardiovascular diseases as well as in cell migration and tumor development [32]. Cartilage oligomeric matrix protein (COMP) plays a role in the structural integrity of cartilage by interacting with other extracellular matrix proteins such as the collagens and fibronectin. Via integrin receptors, it mediates interaction of chondrocytes with the cartilage extracellular matrix and is suggested to play a pivotal role in the pathogenesis of osteoarthritis [33]. Another novel myokine is Secreted frizzled-related protein 4 (SFRP4), which is involved in regulating of cell growth and differentiation. Very recently it has been shown, that SFRP4 represent a potential link between islet inflammation and impaired insulin secretion. SRFP4 expression correlate with inflammatory markers and serum levels are increased in type 2 diabetes patients several years before the diagnosis. This may suggest that SFRP4 could be a potential biomarker for islet dysfunction in type 2 diabetes [34]. Beta-mannosidase (MANBA) is an exoglycosidase, that cleaves single beta-linked mannose residues from N-linked glycoprotein oligosaccharides [35]. Latent transforming growth factor beta (TGF-beta) binding protein 2, belonging to the latent TGF-beta binding protein (LTBP) family, are major regulators of TGF-β bioavailability and action, which play an integral structural role in architectural organization and/or assembly the extracellular matrix [36]. Dysfunction of LTBP's is described in relation to a wide range of diseases including suppression of esophageal tumor formation. For Integrin beta-like protein 1 and Sushi repeat containing protein (SRPX) no functional data are available. Nevertheless, for all

these potential myokines additional validation studies reman required to their function in the context of the skeletal muscle secretome.

5. Conclusion

Taken together, the current knowledge comprises several hundred putative myokines, however their function and regulation in the context of muscle physiology is mainly unknown. Secretory proteins are part of a complex physiological network exerting different effects under various environmental conditions. Our extensive profiling led to the identification of 305 proteins released from human skeletal muscle cells including 12 novel myokines and thereby contribute to our understanding of the complex endocrine capacity of human skeletal muscle. Further studies have to clarify their regulation and their roles in distinct signaling pathways in order to understand their biological function for muscle plasticity and inter-organ crosstalk.

Acknowledgement

This work was partly funded by the Helmholtz Alliance ICEMED - Imaging and Curing Environmental Metabolic Diseases, through the Initiative and Network Fund of the Helmholtz Association. The study was funded in part by a grant from the Leibniz Gemeinschaft (SAW-FBN-2013-3), the German Federal Ministry of Education and Research (BMBF) to the German Center for Diabetes Research (DZD), the Commission of the European Communities (Collaborative Project ADAPT, contract number HEALTH-F2-2008-201100), and the Deutsche Forschungsgemeinschaft (EC 440/1-1).

[1] B.K. Pedersen, T.C. Akerstrom, A.R. Nielsen, C.P. Fischer, Role of myokines in exercise and metabolism, Journal of Applied Physiology 103 (2007) 1093–8.

[2] B.K. Pedersen, M.A. Febbraio, Muscles, exercise and obesity: skeletal muscle as a secretory organ, Nat. Rev. Endocrinol. 8 (2012) 457-65.

[3] C. Weigert, A.M. Hennige, K. Brodbeck, H.U. Häring, E.D. Schleicher, Interleukin-6 acts as insulin sensitizer on glycogen synthesis in human skeletal muscle cells by phosphorylation of Ser473 of Akt, Am. J. Physiol. Endocrinol. Metab. 289 (2005) 251-7.

[4] S. Glund, A. Deshmukh, Y.C. Long, T. Moller, H.A. Koistinen, K. Caidahl, J.R. Zierath, A. Krook, Interleukin-6 directly increases glucose metabolism in resting human skeletal muscle, Diabetes 56 (2007) 1630-7.

[5] A.L. Carey, G.R. Steinberg, S.L. Macaulay, W.G. Thomas, A.G. Holmes, G. Ramm, O. Prelovsek, C. Hohnen-Behrens, M.J. Watt, D.E. James, B.E. Kemp, B.K. Pedersen, M.A. Febbraio, Interleukin-6 increases insulin-stimulated glucose disposal in humans and glucose uptake and fatty acid oxidation in vitro via AMP-activated protein kinase, Diabetes 55 (2006) 2688–97.

[6] L. Zeng, Y. Akasaki, K. Sato, N. Ouchi, Y. Izumiya, K. Walsh, Insulin-like 6 is induced by muscle injury and functions as a regenerative factor, J Biol Chem 285 (2010) 36060–36069.

[7] M.M. Seldin, J.M. Peterson, M.S. Byerly, Z. Wei, G.W. Wong, Myonectin (CTRP15), a novel myokine that links skeletal muscle to systemic lipid homeostasis, J. Biol. Chem. 287 (2012) 11968-80.

[8] P. Hojman, M. Pedersen, A.R. Nielsen, R. Krogh-Madsen, C. Yfanti, T. Akerstrom, S. Nielsen, B.K. Pedersen, Fibroblast growth factor-21 is induced in human skeletal muscles by hyperinsulinemia, Diabetes 58 (2009) 2797-801.

514 [9] J. Henningsen, K.T. Rigbolt, B. Blagoev, B.K. Pedersen, I. Kratchmarova, Dynamics
515 of the skeletal muscle secretome during myoblast differentiation, Molecular & Cellular
516 Proteomics 9 (2010) 2482–96.
517 [10] C.Y. Chan, O. Masui, O. Krakovska, V.E. Belozerov, S. Voisin, S. Ghanny, J.
518 Chen, D. Moyez, P. Zhu, K.R. Evans, J.C. McDermott, K.W. Siu , Identification of
519 differentially regulated secretome components during skeletal myogenesis, Molecular
520 & Cellular Proteomics10 2011 (M110):004804.
521 [11] X.C. Chan, J.C. McDermott, K.W. Siu, Identification of secreted proteins during
522 skeletal muscle development, Journal of Proteome Research 6 (2007) 698–710.
523 [12] J.H. Yoon, K. Yea, J. Kim, Y.S. Choi, S. Park, H. Lee, C.S. Lee, P.G. Suh, S. Ryu,
524 H Comparative proteomic analysis of the insulin-induced L6 myotube secretome,
525 Proteomics 9 (2009) 51–60.
526 [13] S. Bortoluzzi, P. Scannapieco, A. Cestaro, G.A. Danieli, S. Schiaffino,
527 Computational reconstruction of the human skeletal muscle secretome, Proteins. 62
528 (2006) 776-92
529 [14] D.S. Hittel, J.R. Berggren, J. Shearer, K. Boyle, J.A. Houmard, Increased secretion
530 and expression of myostatin in skeletal muscle from extremely obese women, Diabetes
531 . 58 (2009) 30-8
532 [15] F. Norheim, T. Raastad, B. Thiede, A.C. Rustan, C.A. Drevon, F. Haugen,
533 Proteomic identification of secreted proteins from human skeletal muscle cells and
534 expression in response to strength training, Am J Physiol Endocrinol Metab, 301 (2011)
535 E1013-21
536 [16] M.C. Le Bihan, A. Bigot, S.S. Jensen, J.L. Dennis, A. Rogowska-Wrzesinska, J.
537 Lainé, V. Gache, D. Furling, O.N. Jensen, T. Voit, V. Mouly, G.R. Coulton, G. Butler-
538 Browne, In-depth analysis of the secretome identifies three major independent

539 secretory pathways in differentiating human myoblasts, J. Proteomics 77 (2012) 344-
540 56.
541 [17] D. Dietze, M. Koenen, K. Rohrig, H. Horikoshi, H. Hauner, J. Eckel, Impairment of
542 insulin signaling in human skeletal muscle cells by co-culture with human adipocytes,
543 Diabetes 51 (2002) 2369-2376
544 [18] D. Yaffe, O. Saxel, Serial passaging and differentiation of myogenic cells isolated
545 from dystrophic mouse muscle, Nature 270 (1977) 725-7.
546 [19] S. Lehr, J. Kotzka, H. Avci, B. Knebel, S. Muller, F.G. Hanisch, S. Jacob, C. Haak,
547 F. Susanto, D. Muller-Wieland, Effect of sterol regulatory element binding protein-1a
548 on the mitochondrial protein pattern in human liver cells detected by 2D-DIGE,
549 Biochemistry 44 (2005) 5117-28.
550 [20] S. Lehr, S. Hartwig, D. Lamers, S. Famulla, S. Müller, F.G. Hanisch, C. Cuvelier,
551 J. Ruige, K. Eckardt, D.M. Ouwens, H. Sell, J. Eckel, Identification and validation of
552 novel adipokines released from primary human adipocytes, Mol. Cell. Proteomics
553 (2012) M111.010504. doi: 10.1074/mcp.M111.010504.
554 [21] R Development Core Team, A language and environment for statistical computing,
555 reference index version 2.2.1. Vienna, Austria: R Foundation for Statistical Computing,
556 (2005)
557 [22] J. Rainer, F. Sanchez-Cabo, G. Stocker, A. Sturn, Z. Trajanoski, CARMAweb:
558 comprehensive R- and bioconductor-based web service for microarray data analysis,
559 Nucleic Acids Res 34 (2006) W498-W503.
560 [23] S. Raschke, K. Eckardt, K.B. Holven, J. Jensen, J. Eckel, Identification and
561 validation of novel contraction-regulated myokines released from primary human
562 skeletal muscle cells, PLOS ONE (2013) in press.
563 [24] J.H. Yoon, J. Kim, P. Song, T.G. Lee, P.G. Suh, S.H. Ryu, Secretomics for skeletal
564 muscle cells: a discovery of novel regulators?, Adv Biol Regul. (2012) 340-50.

565 [25] L. Orci, M. Tagaya, M. Amherdt, A. Perrelet, J. G. Donaldson, J. Lippincott-
566 Schwartz, R. D. Klausner, J. E. Rothman, Brefeldin A, a drug that blocks secretion,
567 prevents the assembly of non-clathrin-coated buds on Golgi cisternae, Cell 64 (1991)
568 1183-1195.
569 [26] K. Gundersen, Excitation-transcription coupling in skeletal muscle: the molecular
570 pathways of exercise, Biol. Rev. Camb. Philos. Soc. 86 (2011) 564-600.
571 [27] K.T. Preissner, U. Reuning, Vitronectin in vascular context: facets of a
572 multitalented matricellular protein, Semin Thromb Hemost. 37 (2011) 408-24.
573 [28] W.Y. Fu, Y. Chen, M. Sahin, X.S. Zhao, L. Shi, J.B. Bikoff, K.O. Lai, W.H. Yung,
574 A.K. Fu, M.E. Greenberg, N.Y. Ip, Cdk5 regulates EphA4-mediated dendritic spine
575 retraction through an ephexin1-dependent mechanism, Nat Neurosci. 10 (2007) 67-
576 76.
577 [29] Y.K. Nguyen, Granulocyte colony stimulating factor, J Fla Med Assoc. 81 (1994)
578 467-9.
579 [30] C.C. Wu, R.Y. Shyu, J.M. Chou, S.W. Jao, P.C. Chao, J.C. Kang, S.T. Wu, S.L.
580 Huang, S.Y. Jiang, RARRES1 expression is significantly related to tumour
581 differentiation and staging in colorectal adenocarcinoma, Eur J Cancer. 42 (2006) 557-
582 65.
583 [31] E.H. Park, S. Kim, J.Y. Jo, S.J. Kim, Y. Hwang, J.M. Kim, S.Y. Song, D.K. Lee,
584 S.S. Koh, Collagen triple helix repeat containing-1 promotes pancreatic cancer
585 progression by regulating migration and adhesion of tumor cells, Carcinogenesis 34
586 (2013) 694-702.
587 [32] H.C. Kwaan, B. McMahon, The role of plasminogen-plasmin system in cancer.
588 Cancer Treat Res.148 (2009) 43-66.
589 [33] A. Mobasheri, Osteoarthritis year 2012 in review: biomarkers, Osteoarthritis
590 Cartilage 20 (2012) 1451-64.

591 [34] T. Mahdi, S. Hänzelmann, A. Salehi, S.J. Muhammed, T.M. Reinbothe, Y. Tang,
592 A.S. Axelsson, Y. Zhou, X. Jing, P. Almgren, U. Krus, J. Taneera, A.M. Blom, V.
593 Lyssenko, J.L. Esguerra, O. Hansson, L. Eliasson, J. Derry, E. Zhang, C.B. Wollheim,
594 L. Groop, E. Renström, A.H. Rosengren, Secreted frizzled-related protein 4 reduces
595 insulin secretion and is overexpressed in type 2 diabetes, Cell Metab. 16 (2012) 625-
596 33.
597 [35] L.R. Moreira, E.X. Filho, An overview of mannan structure and mannan-degrading
598 enzyme systems, Appl Microbiol Biotechnol. 79 (2008) 165-78.
599 [36] V. Todorovic, D.B. Rifkin, LTBPs, more than just an escort service, J Cell Biochem.
600 113 (2012) 410-8.
601

Table legends

Table 1 Identified proteins using non-targeted proteomic profiling.

Methods	Identified proteins	Unique with applied method	RNA filtered
LC/ESI-MS/MS	405	99	386
2DE/MALDI-MS	171	35	166
LC/MS^{E}	372	80	342

Numbers of proteins (total and unique) identified by the different mass spectrometry based profiling approaches and matching with expression data are displayed.

Table 2 Identified myokines using combined proteomic profiling.

Methods	Identified proteins	Non-redundant proteins	RNA filtered	SP+	SP-	NP
MS	948	530	494	162	136	196
MIA	18	18	7	7		
all	966	548	501	169	136	196

Numbers of proteins (total and non redundant) identified by combined mass spectrometry (MS) based and targeted Multiplex immunoassay (MIA) profiling approaches. Matching with expression data and assignment of identified proteins to their secretory properties are displayed. SP+, putative secretory proteins with signal peptide (signalP positive); SP-, putative secretory proteins without signal peptide (secretomeP score above 0.5); NP, non-putative secretory proteins.

Table 3 Literature comparison

Study	Source	Identified proteins	Common with our study	RNA filtered	SP+	SP-	NP
Bortoluzzi et al.	in silico	319	62	61	60	0	1
LeBihan et al.	primary human cells	955	361	355	133	82	140
Norheim et al.	primary human cells	236	149	144	43	38	63
Hittel et al.	primary human cells	50	21	20	7	7	6
Yoon et al.	rat L6 cells	253	164	158	41	50	67
Henningsen et al.	murine C_2C_{12} cells	635	150	147	118	12	17
Chan et al.	murine C_2C_{12} cells	214	114	109	27	35	47
non-redundant proteins		**1655**	**445**	**425**	**155**	**106**	**164**

Comparison of profiling studies conducted with human and murine material. Numbers of identified proteins, overlap with our study, matching with expression data and assignment of identified proteins to their secretory properties are displayed. SP+, putative secretory proteins with signal peptide (signalP positive); SP-, putative secretory proteins without signal peptide (secretomeP score above 0.5); NP, non-putative secretory proteins.

Table 4 Novel Myokines

Protein Name	UniProtKB Accession	Gene Name	Molecular weight [Da]	pI	LC-MS/MS	2D-MALDI-MS	LC-MS E	MIA	secretion type
Vitronectin	P04004	VTN	52278	5.47	√				SP+
Ephrin type-A receptor 4	P54764	EPHA4	107788	6.32	√				SP+
Granulocyte colony-stimulating factor	P09919	CSF3	22293	5.61				√	SP+
Retinoic acid receptor responder protein 1	P49788	RARRES1	33285	8.74	√		√		SP+
Beta-mannosidase	O00462	MANBA	99117	5.26	√				SP+
Collagen triple helix repeat-containing protein 1	Q96CG8	CTHRC1	26224	9.50	√		√		SP+
Cartilage oligomeric matrix protein	P49747	COMP	80962	4.37	√		√		SP+
Plasminogen activator urokinase	E7ESM2	PLAU	44552	8.57			√		SP+
Secreted frizzled related protein 4	Q6FHJ7	SFRP4	39800	8.87			√		SP+
Integrin beta-like protein 1	O95965	ITGBL1	51460	5.27		√			SP+
Latent transforming growth factor beta binding protein 2	G3V511	LTBP2	189116	5.09			√		SP+
Sushi repeat containing protein SRPX	P78539	SRPX	41722	8.67			√		SP+

Potential novel myokines with signal peptide (signalP SP+) identified by the different mass spectrometry (MS) based and targeted Multiplex immunoassay (MIA) profiling approaches are listed.

Figure legends

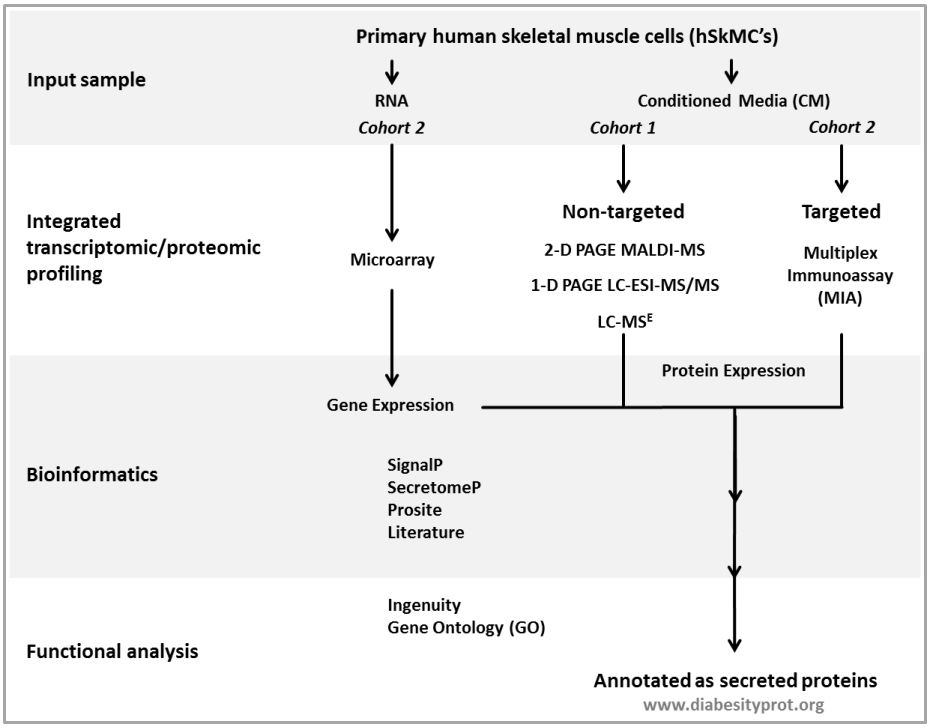

Figure 1 Flow chart – work scheme

For comprehensive secretome profiling, CM derived from primary human skeletal muscle cells (hSkMC) were analyzed using an integrated transcriptomic and proteomic approach. Candidate proteins identified by combining one targeted (Multiplex Immunoassay (MIA)) with three different mass spectrometry based non-targeted approaches (2D-PAGE MALDI-MS, 1D-PAGE LC-ESI-MS/MS and LC/MS^E) were validated by gene expression analysis and consecutive filtering using diverse bioinformatics prediction tools to assign secretory function. Cohort 1 and Cohort 2 are specified in Experimental Procedures.