Proteomics Characterization of Exosome Cargo

Kevin L. Schey¹, J. Matthew Luther^{2,3}, and Kristie L. Rose¹

¹Department of Biochemistry

²Department of Medicine

³Department of Pharmacology

Abstract

Characterization of exosomal cargo is of significant interest because this cargo can provide clues to exosome biogenesis, targeting, and cellular effects and may be a source of biomarkers for disease diagnosis, prognosis and response to treatment. With recent improvements in proteomics technologies, both qualitative and quantitative characterization of exosomal proteins is possible. Here we provide a brief review of exosome proteomics studies and provide detailed protocols for global qualitative, global quantitative, and targeted quantitative analysis of exosomal proteins. In addition, we provide an example application of a standard global quantitative analysis followed by validation via a targeted quantitative analysis of urine exosome samples from human patients. Advantages and limitations of each method are discussed as well as future directions for exosome proteomics analysis.

1. Introduction

Since the discovery of microvesicles¹ [1] and exosomes¹ [2], interest in characterizing their cargo has been intense for several reasons including: identifying clues to their biogenesis, targeting, and cellular effects; discovering their cells of origin; and generating putative biomarkers for disease diagnosis, prognosis and response to treatment. Both protein and RNA products have been used for these purposes; however, the protein cargo and methods for its characterization is the focus of this review. Protein cargoes have been characterized from many exosome sources and their identities utilized to propose mechanisms of exosome biogenesis [3]; identify the cells of origin in body fluid samples, for example from tumor cells [4, 5]; and to develop diagnostic biomarker panels for human disease [6–9]. The interest in exosomes in general is clearly evident in the rapidly growing literature on the topic (Figure 1) and this general trend holds for exosome proteomics publications as well as exosome RNA publications. The challenge of protein identification from exosome

Corresponding Author: Kevin L. Schey, 465 21st Ave So., Suite 9160, MRB III, Nashville, TN 37232, Phone: 615-936-6861, k.schey@vanderbilt.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

¹We use a size-based definition of microvesicles (diameter >100 nm) and exosomes (diameter <100 nm) for the purposes of this presentation, but acknowledge that this definition is evolving and not without controversy [48].

preparations compared to RNA identification is the lack of amplification procedures for protein cargo. Therefore, analytical sensitivity is frequently tested with exosome samples unless large amounts of exosomes can be isolated (sample requirements discussed below). This issue is reflected in the overall lower number of exosome proteomics papers published compared to exosome RNA papers. However, with rapid improvements in proteomic technologies, exosomal protein characterization is not only feasible, but continuing to expand as exemplified in numerous recent reviews [4, 7, 8, 10–12]. This purpose of this review is to present key issues in exosome proteomics, provide general protocols for both comprehensive and targeted proteomics analyses of exosome samples, and to offer example data acquired with these protocols as applied to exosome samples.

2. General Considerations

2.1 Sample Preparation

Proper sample preparation is critically important for a successful experimental outcome of any proteomics experiment. This is especially critical when working near analytical detection limits where contaminants can obscure detection of the lowest abundance components of a sample. Sample preparation for characterization of exosomes will not be discussed in great detail since it is discussed elsewhere in this issue; however, key aspects that affect proteomics analysis are discussed here.

Key issues for mass spectrometric based proteomics analysis include the sample preparation method used and potential contaminants that the method may introduce. Specific contaminants that could “damage” experimental results include detergents, lipids, and polymeric materials often found in commercially available kits and used for precipitation of exosomes. An additional source of undesired material is sucrose or other materials used to produce the density gradient in density gradient centrifugation to prepare exosomes. When these materials are used in an exosome preparation, specific clean-up steps must be used to insure optimum analytical sensitivity. Typical clean-up procedures include protein precipitation or solid phase (C18) extraction. Standard protein precipitation methods include cold acetone, trichloroacetic acid, and methanol/chloroform precipitation methods. An alternative approach is to load solubilized exosome samples onto an SDS-PAGE gel followed by in-gel digestion/extraction [13]. An alternative version of this approach is to run “short-stack” gels, i.e. run proteins about 1cm into the gel without allowing for protein separation [14].

An additional point of concern when beginning an exosome proteomics experiments is the source of exosomes and the endogenous components that may interfere with the analysis. For example, cell culture media often contain large amounts of albumin from serum, exosomes from blood may contain large amounts of albumin and other abundant serum proteins, and urine exosome samples often contain large amounts of Tamm-Horsfall (uromodulin) protein. Therefore, specific procedures should be included in the sample preparation protocol to reduce these endogenous non-exosomal proteins that will interfere with the analysis [15].

2.2 Basic Proteomics Approaches

Virtually every standard proteomics approach has been employed to examine exosomal protein cargo ranging from two-dimensional gel electrophoresis to more modern quantitative LCMS/MS approaches. For the purposes of this review, several key proteomics strategies for exosomal protein characterization will be presented along with their specific advantages for this purpose. The selection of a particular proteomics approach depends on the question being asked. The comprehensive identification of exosomal protein cargo can be best accomplished with a “Global Discovery” approach using multidimensional LC-MS/MS strategies. Quantitative analysis of exosomal protein cargo can be accomplished on a global scale, i.e. all detected proteins are quantified, or on a targeted scale, i.e. only specified proteins are detected and quantified. Each of these strategies is discussed in detail below.

2.2.1 Global Discovery for Protein and Post-translational Modification

Identification—Comprehensive identification of exosome protein cargo is best carried out by multiple dimensions of fractionation followed by sensitive tandem mass spectrometric detection. Many different modes of protein/peptide separation have been developed, but perhaps the most commonly used strategy in current use is Multidimensional Protein Identification Technology (MudPIT) where a combination of strong cation exchange chromatography (SCX) and reverse phase (RP) chromatography is used to fractionate complex mixtures of peptides [16]. Variants of this approach include reverse phase/reverse phase separation where the first dimension is run under basic pH conditions and the second dimension is run under acidic conditions [17], and 2) isoelectric focusing done prior to reverse phase chromatography [18]. GeLC-MS approaches utilize SDS-PAGE gel separation of intact proteins prior to digestion and reverse phase separation of extracted peptides [13, 19].

Figure 2 shows a schematic diagram of a MudPIT workflow where multiple peptide fractions are eluted from the SCX column via a series of salt pulses. Each salt eluted fraction is then separated by reverse phase chromatography and peptides analyzed by tandem mass spectrometry to generate molecular weight and sequence-specific fragmentation for each peptide in the mixture. Typically 8–11 salt steps are used to examine a complex mixture of proteins. During peptide elution from the reverse phase column, mass spectra and tandem mass spectra are acquired, typically in a data dependent fashion, i.e. the top 5, 10, or 20 ions in a mass spectrum are sequentially selected for fragmentation in the MS/MS scan. Tandem mass spectra are subsequently searched against protein sequence databases using software tools such as Sequest [20], Mascot [21], X!Tandem [22], and Andromeda/MaxQuant [23] to identify the eluted peptides. Peptide sequence redundancy among identified proteins is then reduced by parsimony algorithms, such as IDPicker [24] and Scaffold (Proteome Software), resulting in a final protein list. Newer methods of data independent acquisition, such as SWATH [25] and MS^E [26], attempt to acquire MS/MS data on all ions in each mass spectrum by scanning windows of selected precursor ions and fragmenting multiple ions simultaneously. Peptide fragmentation libraries are then utilized to interpret the mixed tandem mass spectra.

The end result of such a global discovery experiment is a list of proteins identified in the sample provided. Typically for exosome samples, the number of proteins identified by LC-MS/MS varies between several hundreds to several thousands depending on the amount of starting material. Standard exosome markers are typically identified including: TSG101, annexins, CD9, CD63, CD81, as well as ESCRT proteins involved in exosome sorting and trafficking. These types of lists can be used to identify the cellular source of exosomes for example podocin, NKCC2, NCC, and AQP2, in urine exosomes corresponding to podocyte, loop of Henle, distal tubule, and collecting duct cells of the nephron, respectively. Several groups have assembled results of exosome protein characterization into databases, e.g. Exocarta [27] (<http://www.exocarta.org>) and urinary exosome protein database (<http://dir.nhlbi.nih.gov/papers/lkem/exosome/>) containing 4563 exosome proteins and 1160 urine exosome proteins, respectively, as well as vesiclepedia (<http://www.microvesicles.org>) [28] containing 43,731 protein entries and EVpedia (<http://evpedia.info>) [10] containing 78,971 proteins from 176 proteome data sets from all extracellular vesicle types.

This type of analysis can also yield information on post-translational modifications (PTMs) present in exosomal proteins, e.g. phosphorylation, oxidation, acetylation. Post-translational modifications to exosome proteins likely play a role in their sorting and trafficking among other functions [29]. For low stoichiometry modifications, enrichment steps can be used to enrich for specific modifications, e.g. immobilized metal affinity chromatography (IMAC) or TiO₂ enrichment for phosphopeptides. Knepper's group reported phosphoproteins enriched from urine exosomes using such an approach [30]. Antibodies to phosphotyrosine [31], acetyllysine [32], and ubiquitin (Gly-Gly) [33] have also been used to enrich for these PTMs; however, they have not been used for exosome analysis.

2.2.2 Quantification of Exosome Proteins—The most useful types of proteomics information are typically derived from quantitative data on protein expression level. This type of information can inform, for example, on exosomal protein differences as diagnostic indicators/biomarkers of disease [7]. Quantitative proteomics data can be acquired on a global proteome scale, i.e. all detected proteins quantified, or in a targeted fashion, i.e. selected protein targets quantified. Furthermore, both isotope labeling methods and label free methods are available for quantitation.

2.2.2.1 Global Quantitation: Global quantitative methods attempt to quantify all detected proteins in a sample. The simplest approach is to use label free methods where simple LC-MS/MS data are used to produce relative quantitative information. This can be done in two general ways: 1) peak area determination, and 2) spectral count measurements. Peak area measurements for each precursor ion detected can be used to quantitatively assess the peptide, and inferred protein, levels in the sample. High mass resolution measurement of peptide molecular ions is preferred to limit interferences from isobaric, co-eluting peptides. Spectral count measurements record the number of times a specific peptide is selected in a data dependent LC-MS/MS experiment and is roughly reflective of the peptide, inferred protein, level in the sample.

Isotope labeling methods require specific incorporation of stable heavy isotope labeled amino acids into peptides to be measured. This can be done in cell culture by using heavy

isotope labeled amino acids in the culture media in the SILAC (stable isotope labeling of amino acids in cell culture) experiment [34], or by post digestion chemical reaction in iTRAQ (isotope tags for relative and absolute quantitation) [35] or TMT (tandem mass tags) [36] methods. Once labeling is accomplished, different samples can be mixed and analyzed in a single analytical run thereby reducing the variability of the results. In contrast, label free methods require each sample to be run individually and typically require some normalization method to be established. iTRAQ and TMT methods can be multiplexed to allow 8–10 different samples, respectively, to be analyzed in a single experiment. Note that database searching algorithms must be able to search with specific labels included as peptide modifications.

2.2.2.2 Targeted Peptide Quantitation: In contrast to a discovery-based, data-dependent analysis method, targeted mass spectrometry methods are used to analyze a predefined set of peptides. The most common method for targeted proteomics applications involves multiple-reaction monitoring (MRM) using a triple quadrupole mass spectrometer [37]. In an MRM experiment (depicted schematically in Figure 3), a specific peptide precursor ion is selected in the first quadrupole (Q1) analyzer, is fragmented by CID in the second quadrupole (Q2), and the intensities of peptide specific fragment ions are specifically measured by the second quadrupole mass analyzer (Q3). Detection of precursor/fragment ion pairs (also called transitions) over time provides chromatographic peaks of these monitored transitions if the specific peptide is present. A targeted method has a lower limit of detection and greater dynamic range compared to discovery methods, and while the targeted method requires *a priori* selection of peptides of interest, it is the method of choice for identifying and quantifying preselected peptides in a complex mixture with high reproducibility. For selecting proteins of interest to be quantified across exosome samples, considerations may include proteins that are involved in specific cellular pathways, proteins that are predicted to be affected under a certain condition or treatment prior to exosome collection, or proteins detected as altered in abundance in global proteomics analysis. By scheduling different peptide/fragment transitions as a function of HPLC retention time, this strategy can provide highly sensitive and quantitative information on hundreds of peptides of interest in a single experiment. Moreover, stable isotope labeled standard peptides can be spiked into each sample to produce accurate absolute quantitative information. Typically, 3–6 unique fragment ions are chosen for a given peptide and at least two unique peptides per proteins are monitored. A more recent strategy is to monitor all fragment ions for selected target peptides in a parallel reaction monitoring (PRM) experiment [38]. By monitoring all fragment ions simultaneously, higher specificity and more accurate quantitation can be achieved.

3. Specific Exosome Proteomics Protocols

3.1 Experimental Procedures for Discovery Proteomics Analysis of Exosomes

For analysis by mass spectrometry, exosomal proteins can be separated by SDS-PAGE, in-gel digested, and peptides extracted for LC-MS/MS experiments. However, a solution-phase digestion of exosome proteins often results in similar or higher yield [14], and the procedures described here involve in-solution exosomal protein digestion and analysis.

3.1.1 Sample Preparation

1. First, the protein concentration of the exosome preparation must be measured. As described in Wang *et al* [14] for urinary exosomes, the isolated exosomes once pelleted can be suspended in either HPLC-grade water or other buffer, e.g. 100mM Tris, that is amenable to protein digestion and lacks components that will interfere with LC-MS analysis. Any of numerous protein assays can be used, providing buffer components do not interfere with the assay. As a consideration, preparation of urinary exosomes involves use of a high concentration of DTT to help in reducing the major urinary Tamm-Horsfall protein [15]. The amount of remaining DTT in the exosomal pellet can interfere with protein assays such as the Lowry and BCA reagent methods. However, the Bradford reagent is a good choice for assaying protein from urinary exosomes, since it is compatible with reducing agents. A preferred protein yield of approximately 50µg from an exosome preparation is optimal for the multidimensional, comprehensive mass spectrometric analysis described in the following steps.
2. After assaying for protein content, reduce the volume of suspended exosome by speed vac concentration and resuspend the exosome proteins in a small volume (i.e. 20µL) of 50% trifluoroethanol in 100mM Tris HCl, pH 8.
3. Reduce proteins by addition of 1µL of 0.5 M Tris(2-carboxyethyl) phosphine hydrochloride (TCEP) and incubate for 45min at room temperature.
4. Alkylate available cysteine residues by addition of 1.5–2µL of 1M iodoacetamide. Incubate at room temperature for 1 hour in the dark.
5. After reduction and alkylation, dilute the exosome sample 10-fold with 100mM Tris HCl, pH 8, add sequencing-grade trypsin (Promega) at an enzyme to substrate ratio of 1:50, and digest at 37°C overnight up to 18 h.

3.1.2 Analysis by LC-coupled tandem mass spectrometry—For comprehensive characterization of digested exosome proteins we recommend using a MudPIT strategy. In addition, we promote using nanoflow separations and self-packed capillary columns, which are optimal for highly sensitive analyses of complex exosome preparations where typical yields are in the range of 50µg of protein. The experimental details below describe how to pack capillary columns needed for the described nanoLC MudPIT experiments, how to load the sample on column, and how to complete the LC-MS instrumental analysis. The following protocol involves online MudPIT separations, which we feel is optimal for analysis of the typical microgram yields from exosome preparations. However, an offline strategy where, for example, SCX fractions or basic reverse phase fractions are collected offline prior to online RP LC-MS analysis, can also be employed for analysis of complex exosome samples.

1. First, prepare the fused silica to be used for the capillary analytical column. Cut a 25cm length of 360 × 100 µm i.d. polyimide-coated fused silica (Polymicro Technologies). Remove the polyimide approximately 3 cm from the end of the fused silica using a methanol flame (1–2 seconds) and wipe the burnt coating away with a kimwipe and methanol. Using a laser puller (model P-2000, Sutter

Instrument Co), equip the end of the fused silica with a laser-pulled <3 μm orifice (emitter tip).

2. Make a slurry (15–20mg in 500 μL) of 3 μm 300Å C18 reverse phase material (Phenomenex) in HPLC-grade methanol. Using a Helium-pressurized cell (pressure bomb) which can be purchased from New Objective, pack 20 cm of the C18 material into the 100 μm i.d. analytical column. As another option, SilicaTip Emitters (New Objective) can also be purchased either pre-packed or for self-pack applications.
3. Cut a 15cm length of 360 $\times$ 150 μm i.d. polyimide-coated fused silica and frit the end of the capillary using a filter-end fitting (IDEX Health and Science) to be used for making a SCX/RP biphasic column.
4. Using a Helium-pressurized cell, prepare a slurry (~5mg in 500 μL of HPLC-grade methanol) pack 6 cm of 5 μm 100Å SCX material into the fritted 150 μm i.d. capillary, followed by 4 cm of 5 μm C18 300Å reverse phase material, making a biphasic column. Also, biphasic columns are available for purchase (New Objective) in different configurations of column length and packing materials.
5. After equilibrating the biphasic column, acidify the exosome tryptic digest by adding formic acid to a final of < 0.5% and load the digest onto the capillary column using the pressure bomb. We elect to bomb-load samples, since for the typical MudPIT analysis, the microgram amount of protein and volume of solution is too large to be loaded using a nanoLC autosampler.
6. After sample loading, connect the biphasic capillary column onto the nanoLC instrument. This is often accomplished by using a microtee (e.g. 360 μm PEEK microtee, IDEX Health and Science) to connect the output from the LC to the back-end of the MudPIT column. Voltage for nanoESI can be applied by liquid junction at the microtee, where the electrical contact is completed at the T-junction using the high voltage cable from the mass spectrometer, which is connected to a gold wire and sealed using a standard PEEK fitting and ferrule. Next connect the fritted end of the biphasic column to the back end of the analytical column using an M-520 union (IDEX Health and Science).
7. Using a nanoLC and autosampler, MudPIT analysis should be performed with a multi-step salt pulse gradient (e.g. 25 mM up to 1M ammonium acetate). For typical MudPIT experiments, we use an 11-step pulse gradient with 25, 50, 75, 100, 150, 200, 250, 500, 750, and 1M ammonium acetate salt steps. Vials of each salt concentration should be placed in the autosampler for pick-up and delivery using an automated method, and 5 μL of salt solution is sufficient for each step.
8. Following each salt pulse, peptides should be gradient-eluted from the reverse analytical column at a flow rate of 500nL/min, or lower flow rates may also be considered. For solvent A and B, use 0.1% formic acid/99.9% water and 0.1% formic acid/99.9% acetonitrile, respectively. For the first ten SCX fractions, elute the peptides using a reverse phase gradient consisting of 2–45 %B in approximately 90 min, followed by a 15 min equilibration at 2 %B prior to the next salt pulse.

Using this method, each step will take approximately 2 hours and includes autosampler delivery of the desired salt pulse, a 90-min increasing organic gradient, and column equilibration. For the last SCX-eluted peptide fraction, elute the peptides using a gradient of 2–98 %B in 100 min, followed by a 10 min equilibration at 2 %B. Upon gradient elution, peptides should be introduced via nano-electrospray into a tandem mass spectrometer such as a LTQ Velos ion trap or Velos Orbitrap instrument (Thermo Scientific), equipped with a nanospray ionization source. Using this particular MudPIT method, a single exosome sample is analyzed over a total analysis time of ~24 h/sample using a completely automated method.

9. Operate the mass spectrometer using a data-dependent method. Full scan spectra should be acquired over m/z 350–2000 or similar m/z range. If a hybrid instrument such as a Velos Orbitrap is being used, we recommend collecting the full scan data with the high mass resolution Orbitrap as the mass analyzer. The method should include data-dependent selection and fragmentation with collision-induced dissociation (CID) of the top 10–20 most abundant ions in each MS scan and tandem mass spectra should be acquired in the LTQ ion trap. Additional parameters for the instrument method include: an isolation width of 2 m/z , activation time of 10 ms, 35% normalized collision energy for MS/MS, an MS^2 AGC target value of 1×10^4 , and enabling dynamic exclusion.
10. Search data against an appropriate database using a database searching algorithm, such as Sequest [20], Mascot [21], Maxquant [23] etc. As an example, if exosome samples originated from cultured human cells, search the data against a *Homo sapiens* subset database downloaded from the uniprot KB protein database (www.uniprot.org). The database should include both forward and reversed protein sequences, so that the search results can be filtered to an acceptable false-discovery rate. Results can be filtered such that identified proteins have a certain number of distinct peptides, and the protein-level FDR should be low (e.g. <1–2%). Multiple programs can be used to accomplish this, including IDPicker [24] or Scaffold (Proteome Software).
11. Statistical analyses could be performed following databases searching if, for example, multiple exosome samples, such as exosomes collected before and after a specific treatment or from normal or diseased patients, have been analyzed by MudPIT. The number of MS/MS spectra, i.e. spectral count, from each protein in a given MudPIT analysis can be used to provide a rough estimate of relative abundance of that protein across the samples. As described in Lundgren *et al*, a normalization procedure can be applied to these data, and comparisons can be made across different groups of samples to determine those proteins that are most changed [39]. These proteins of interest can be targeted in downstream quantitative analyses, similar to those described in the following section. In addition, gene ontology enrichment analysis or pathway analysis can be performed using such programs as WebGestalt [40], DAVID [41] or Ingenuity Pathway Analysis (QIAGEN).

3.2 Experimental Procedures for Targeted Quantitative Analysis of Exosome Proteins

3.2.1 Targeted Assay Considerations—For a typical MRM experiment, we recommend analyzing 1–2 µg per analysis. The number of proteins that can be quantified in a single LC-MS analysis depends on the number of peptides analyzed per protein and the number of transitions monitored per peptide. While instrument parameters and chromatography characteristics (such as those described in the section below) will affect the total number of transitions that can be monitored in a single LC-MS experiment, it is often possible to monitor around 100 transitions per experiment. If for each protein there are 2 peptides analyzed with 3 transitions monitored per peptide, then approximately 15 proteins can be quantified in a single experiment. Depending on the specificity and sensitivity needed for a given sample, a higher number of transitions may be required per peptide, which can reduce the number of proteins quantified per analysis. If more than 15 proteins are selected for MRM analysis, then multiple MRM methods can be developed with each one requiring a separate sample injection. As an alternative, if retention time information is known or can be predicted for peptide targets, then a scheduled MRM strategy (see step 7 below) can be used. Published methods have reported targeted assays with over 1000 transitions per analysis using a scheduled MRM strategy [42, 43]. In the method described here, peptides from specific proteins observed in global exosome MudPIT analysis are selected as targets for quantitation, MRM experiments are developed and implemented, and finally these selected peptides, and by inference proteins, are quantified across many samples in a targeted analysis.

1. For specific proteins of interest, select peptide targets based on their frequency of identification (i.e. highest spectral count) in the discovery-based MudPIT experiments. We find that a given peptide is more readily quantified by MRM if it is frequently identified in data-dependent analyses. For reliable quantitation, at least two peptides should be monitored for each target protein.
2. Chemical modifications that occur during sample processing can lead to errors in quantitation; therefore avoid selection of peptides that are prone to artifactual modifications, such as peptides containing methionine residues which are prone to oxidation or peptides with N-terminal glutamine residues which will readily convert to pyro-glutamate under acidic conditions.
3. While peptides with missed cleavages or non-tryptic cleavages may be observed in MudPIT analyses, these peptides should not be targeted for quantitation since these types of peptides may be generated at variable amounts in different samples.
4. The peptides selected for quantitation should be unique to the protein of interest. Of note, it may be important to select peptides that can distinguish among different protein isoforms or variants.
5. Once peptides are selected from the MudPIT experimental results, select the observed and/or predominant charge states of the peptide targets. To develop a sensitive MRM assay, it is important to monitor transitions that are specific for the most intense fragments observed in the tandem mass spectrum of a given peptide precursor. A single precursor-product ion transition may arise from multiple

different peptide precursors; however, the probability of detecting multiple co-eluting peptides with identical panels of precursor ion-fragment ion transitions is low. The co-occurrence of several specific transitions often provides high specificity in detection of a given peptide and, therefore, we recommend monitoring 4–6 transitions per peptide. When designing MRM assays, one should be aware that relative fragment ion intensities are dependent on the type of instrument used, and this aspect should be considered when ion trap-derived data are used to select transitions for a triple quadrupole MRM analysis. For example, higher m/z b ions and doubly-protonated fragment ions are often less prominent or absent in triple quadrupole-derived MS/MS spectra compared to ion trap spectra. Also note, doubly-charged precursors yield predominantly singly-charged fragments. Singly-charged y ions are often the predominant type of ion generated in the collision cells of triple quadrupole instruments.

6. To make an MRM assay quantitative, standard peptides must be synthesized for spiking as internal standards into the exosome samples to be quantified. A standard peptide can be an unlabeled, exogenous peptide or a labeled reference peptide unrelated to the peptides being quantified, and these two options can be used for relative quantitation [44]. As another strategy, isotopically-labeled heavy peptides can be used for absolute quantitation [45]. While often utilized for these types of experiments, relative quantitation with unlabeled or labeled reference peptides can present some challenges due to variation in signal intensities across different LC-MS analyses. However, these approaches can be successful if samples are similar in background and composition [37]. For a more accurate method, labeled peptide standards identical in sequence to the endogenous peptides being quantified can be synthesized with stable heavy isotopes and, after spiking known quantities of these labeled and purified peptides into an exosome sample, quantification is based on the relative intensities of the endogenous peptide signals compared with those derived from the isotopically-labeled peptides. For quantitation of tryptic peptides from exosome samples, peptides should be synthesized using heavy Arg or Lys at their C-termini. Of note, it is recommended to use isotopes that provide a sufficiently large mass difference relative to the endogenous analyte, such as with $^{13}\text{C}_6$ $^{15}\text{N}_4$ Arg ($m=10$ Da) and $^{13}\text{C}_6$, $^{15}\text{N}_2$ Lys ($m=8$ Da). In contrast, shifts in chromatographic retention occur with deuterium-labeled standards and are therefore not advised for these types of quantitative experiments. Labeled peptides can be purchased from one of several companies, and quantitative MRM experiments can be completed once peptides are received. In the protocol provided here, a stable isotope-based method is described.
7. For targeted analyses, multiple programs such as Skyline [46] and Pinpoint (Thermo Scientific) can be used for generating MRM methods. Import protein sequences into Skyline, for example, where *in silico* digestion can be performed, transitions can be easily selected for both light (endogenous) and heavy-labeled peptides, and MRM transitions can be exported directly for use in instrumental methods. In order to maximize the number of monitored transitions in a single LC-MS, a scheduled MRM method can be considered. In a scheduled method, MRM

transitions are performed over specific elution time windows, and using this strategy, hundreds of peptides can be quantified in a single experiment. For this type of analysis, a representative sample must first be analyzed, or “scouted”, in order to determine retention times of each targeted peptide, or there are also approaches that involve retention time predictions such as that described by Escher *et al* [47] for use in informing scheduled MRM experiments. Using programs such as those listed above (i.e. Skyline), an automated method can be easily exported that includes all transition and retention time information.

3.2.2 Targeted Analysis Using Multiple-Reaction Monitoring

1. Prepare exosome samples in a similar manner to those described in section 3.1, and spike in internal peptide standards after exosome proteins have been digested. We find that low fmol/uL concentrations of internal standards work well. Of note, more accurate quantitation is achieved when peptides standards are spiked at a concentration that is within an order magnitude relative to the corresponding analyte concentration.
2. Next, prepare columns to be used for 1D reverse phase LC-MS. Prepare and pack a C18 analytical column as described in 3.2 steps 1–2.
3. Prepare a fritted capillary column as described in 3.2 step 3, and pack 5cm of 5 μ m 300Å C18 material into the column (C18 trap column).
4. Using a nanoLC and autosampler, load the spiked exosome sample onto the C18 trap column, and use mobile phase solvents and a reverse phase gradient similar to that described in 3.2 step 8, i.e. 2–45% in 90 min, to resolve the peptides over the C18 analytical column.
5. Upon gradient elution, introduce the exosome peptides via nano-electrospray into a triple quadrupole mass spectrometer. Instrument methods exported from Skyline, or a similar software, include all m/z values used for multiple-reaction monitoring. In addition, Skyline can be used to calculate and export appropriate collision energies to be used for each peptide target depending on the corresponding m/z value of the peptide and the particular instrument being used for analysis. The collision energy for each peptide can also be determined or optimized empirically by analyzing a representative exosome sample. Operate the instrument using a dwell time (i.e. 20 ms) for each transition that will allow sufficient signal to be observed. However, too few data points acquired during peptide elution will result in quantitative errors, and so the dwell time and number of monitored transitions must be appropriately adjusted depending on the chromatographic peak width of the LC-MS experiment.
6. Finally, import MRM data into Skyline, where transition peaks are grouped by peptide precursor, and transition peak areas can be summed for light or heavy peptide precursors. Ratios for light and heavy peptides pairs can be calculated, and used to derive quantitative measurements for peptides targeted in each exosome sample. Statistical methods (e.g. t-tests or paired t-tests depending on the origin of exosome samples) can be applied to determine those proteins that are most

significantly changed between individual exosome samples analyzed in replicates or between groups of exosome samples.

4. Example Exosome Proteomics Results

Using the experimental MudPIT protocols described above, exosomes were analyzed from human urine, and currently this published dataset represents the largest number of urinary exosomal proteins identified to date [14]. For this particular study, nine urine exosome samples were characterized and approximately 2500 proteins were identified per sample. A total of 3280 proteins were identified among all the samples, greatly expanding the known protein components of urinary exosomes. This extensive protein dataset was then examined to assess gene ontology and functionally categorize protein cargo. As a result, this study has provided a valuable resource considering biological fluids such as human urine have much potential in disease-related biomarker research.

More recently, this method was applied to a clinical study focused on the effects of aldosterone, a hormone that regulates sodium uptake in the kidney and that plays a role in hypertension. In the initial phase, a MudPIT strategy was used for characterizing exosome protein cargo from urine collected from patients that were administered either aldosterone or vehicle (control) only. The discovery proteomics analysis was conducted on exosome samples from ten patients and a spectral counting approach was employed to quantitate proteins that changed with aldosterone (aldo) treatment. Included in the altered proteins identified was Rac3 that was consistently found in all exosome samples was identified with decreased spectral counts following aldo treatment. To achieve a more quantitative analysis of patient samples collected in this clinical study, a targeted MRM assay was designed using isotopically-labeled peptide synthesized to match the sequences of targeted peptides. Target peptides and their labeled standards were analyzed on a Thermo TSQ Vantage triple quadrupole instrument. Among the peptides targeted a Rac3 peptide with sequence YLECSALTQR was quantified in urinary exosomes isolated from five patients with and without aldo treatment. Four transitions were monitored (y_4 – y_8) from the doubly-protonated precursor ion of Rac3 peptide (m/z 620.80) as well as the heavy isotopically-labeled ($^{13}\text{C}_6$ $^{15}\text{N}_4$ Arg) peptide standard (625.81), and the results of these data are shown in Figure 4. Three replicates were performed per patient sample and endogenous peptide intensities were normalized to the labeled peptide. Although the data show patient differences in absolute Rac3 peptide levels, they are consistent among individual patients and indicate down-regulation of Rac3 in all patients after aldosterone administration compared to vehicle.

5. Current Trends and Future Directions

While new discoveries are being published on a daily basis, there are several exciting areas that will likely improve our ability to analyze exosomal protein cargo and thereby enhance our ability to make new discoveries. Methods for improved sensitivity are continuously being developed. New mass spectrometry instrumentation is introduced on a yearly basis while improved chromatographic methods are also introduced annually. As mentioned above, improved methods for sample preparation of exosomes will only augment the analytical efficiency.

Areas of application of exosome proteomics are likely to remain similar to the current broad questions described above; however, applications to new disease states and new molecular mechanisms will surely appear. Specific examples of potential applications of improved analytical sensitivity are the exploration of the exosome surface proteome, analysis of sorted exosome populations, and new targeted assays for biomarker detection. The exosomal surface proteome likely holds clues to mechanisms of exosome formation, secretion, targeting, protein-protein interactions, and host cell capture. By examining specific, sorted populations of exosomes, increase specificity of exosome cargo in relation to the cell of origin can be achieved. Currently, exosome samples from biofluids represent contributions from entire populations of cells. By achieving cell-specific exosome analysis, higher specificity in biomarker detection can be achieved. Lastly, new biomarkers of disease are being identified on a frequent basis. With improved analytical sensitivity, new markers will be detected and new sensitive and specific assays will be developed for early disease detection, prognosis, and effect of treatment determination.

Acknowledgments

This work was supported by NIH grants HL100016, DK081662, UL1RR024975, and UL1TR000445 from NCATS/NIH (Vanderbilt Institute for Clinical and Translational Research), and the Vanderbilt Mass Spectrometry Research Center.

References

1. Wolf P. The nature and significance of platelet products in human plasma. *Br J Haematol.* 1967; 13:269–288. [PubMed: 6025241]
2. Johnstone RM, Adam M, Hammond JR, Orr L, Turbide C. Vesicle formation during reticulocyte maturation. Association of plasma membrane activities with released vesicles (exosomes). *J Biol Chem.* 1987; 262:9412–9420. [PubMed: 3597417]
3. Thery C, Boussac M, Veron P, Ricciardi-Castagnoli P, Raposo G, Garin J, Amigorena S. Proteomic analysis of dendritic cell-derived exosomes: a secreted subcellular compartment distinct from apoptotic vesicles. *J Immunol.* 2001; 166:7309–7318. [PubMed: 11390481]
4. Simpson RJ, Jensen SS, Lim JW. Proteomic profiling of exosomes: current perspectives. *Proteomics.* 2008; 8:4083–4099. [PubMed: 18780348]
5. Yang J, Wei F, Schafer C, Wong DT. Detection of tumor cell-specific mRNA and protein in exosome-like microvesicles from blood and saliva. *PLoS One.* 2014; 9:e110641. [PubMed: 25397880]
6. Hoorn EJ, Pisitkun T, Zietse R, Gross P, Frokiaer J, Wang NS, Gonzales PA, Star RA, Knepper MA. Prospects for urinary proteomics: exosomes as a source of urinary biomarkers. *Nephrology (Carlton).* 2005; 10:283–290. [PubMed: 15958043]
7. Simpson RJ, Lim JW, Moritz RL, Mathivanan S. Exosomes: proteomic insights and diagnostic potential. *Expert Rev Proteomics.* 2009; 6:267–283. [PubMed: 19489699]
8. Raimondo F, Morosi L, Chinello C, Magni F, Pitto M. Advances in membranous vesicle and exosome proteomics improving biological understanding and biomarker discovery. *Proteomics.* 2011; 11:709–720. [PubMed: 21241021]
9. Dear JW, Street JM, Bailey MA. Urinary exosomes: a reservoir for biomarker discovery and potential mediators of intrarenal signalling. *Proteomics.* 2013; 13:1572–1580. [PubMed: 23129434]
10. Choi DS, Kim DK, Kim YK, Gho YS. Proteomics, transcriptomics and lipidomics of exosomes and ectosomes. *Proteomics.* 2013; 13:1554–1571. [PubMed: 23401200]
11. Brown KJ, Formolo CA, Seol H, Marathi RL, Duguez S, An E, Pillai D, Nazarian J, Rood BR, Hathout Y. Advances in the proteomic investigation of the cell secretome. *Expert Rev Proteomics.* 2012; 9:337–345. [PubMed: 22809211]

12. Fontana S, Saieva L, Taverna S, Alessandro R. Contribution of proteomics to understanding the role of tumor-derived exosomes in cancer progression: state of the art and new perspectives. *Proteomics*. 2013; 13:1581–1594. [PubMed: 23401131]
13. Ji H, Greening DW, Barnes TW, Lim JW, Tauro BJ, Rai A, Xu R, Adda C, Mathivanan S, Zhao W, Xue Y, Xu T, Zhu HJ, Simpson RJ. Proteome profiling of exosomes derived from human primary and metastatic colorectal cancer cells reveal differential expression of key metastatic factors and signal transduction components. *Proteomics*. 2013; 13:1672–1686. [PubMed: 23585443]
14. Wang Z, Hill S, Luther JM, Hachey DL, Schey KL. Proteomic analysis of urine exosomes by multidimensional protein identification technology (MudPIT). *Proteomics*. 2012; 12:329–338. [PubMed: 22106071]
15. Fernandez-Llama P, Khositseth S, Gonzales PA, Star RA, Pisitkun T, Knepper MA. Tamm-Horsfall protein and urinary exosome isolation. *Kidney Int*. 2010; 77:736–742. [PubMed: 20130532]
16. Washburn MP, Wolters D, Yates JR 3rd. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol*. 2001; 19:242–247. [PubMed: 11231557]
17. Dwivedi RC, Spicer V, Harder M, Antonovici M, Ens W, Standing KG, Wilkins JA, Krokhin OV. Practical implementation of 2D HPLC scheme with accurate peptide retention prediction in both dimensions for high-throughput bottom-up proteomics. *Anal Chem*. 2008; 80:7036–7042. [PubMed: 18686972]
18. Slebos RJ, Brock JW, Winters NF, Stuart SR, Martinez MA, Li M, Chambers MC, Zimmerman LJ, Ham AJ, Tabb DL, Liebler DC. Evaluation of strong cation exchange versus isoelectric focusing of peptides for multidimensional liquid chromatography-tandem mass spectrometry. *J Proteome Res*. 2008; 7:5286–5294. [PubMed: 18939861]
19. Pisitkun T, Shen RF, Knepper MA. Identification and proteomic profiling of exosomes in human urine. *Proc Natl Acad Sci U S A*. 2004; 101:13368–13373. [PubMed: 15326289]
20. Eng JK, McCormack AL, Yates JR. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectrom*. 1994; 5:976–989. [PubMed: 24226387]
21. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis*. 1999; 20:3551–3567. [PubMed: 10612281]
22. Fenyo D, Beavis RC. A method for assessing the statistical significance of mass spectrometry-based protein identifications using general scoring schemes. *Anal Chem*. 2003; 75:768–774. [PubMed: 12622365]
23. Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M. Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res*. 2011; 10:1794–1805. [PubMed: 21254760]
24. Ma ZQ, Dasari S, Chambers MC, Litton MD, Sobecki SM, Zimmerman LJ, Halvey PJ, Schilling B, Drake PM, Gibson BW, Tabb DL. IDPicker 2.0: Improved protein assembly with high discrimination peptide identification filtering. *J Proteome Res*. 2009; 8:3872–3881. [PubMed: 19522537]
25. Gillet LC, Navarro P, Tate S, Rost H, Selevsek N, Reiter L, Bonner R, Aebersold R. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol Cell Proteomics*. 2012; 11:O111 016717. [PubMed: 22261725]
26. Chakraborty AB, Berger SJ, Gebler JC. Use of an integrated MS--multiplexed MS/MS data acquisition strategy for high-coverage peptide mapping studies. *Rapid Commun Mass Spectrom*. 2007; 21:730–744. [PubMed: 17279597]
27. Mathivanan S, Fahner CJ, Reid GE, Simpson RJ. ExoCarta 2012: database of exosomal proteins, RNA and lipids. *Nucleic Acids Res*. 2012; 40:D1241–1244. [PubMed: 21989406]
28. Kalra H, Simpson RJ, Ji H, Aikawa E, Altevogt P, Askenase P, Bond VC, Borrás FE, Breakefield X, Budnik V, Buzas E, Camussi G, Clayton A, Cocucci E, Falcon-Perez JM, Gabriellsson S, Gho

- YS, Gupta D, Harsha HC, Hendrix A, Hill AF, Inal JM, Jenster G, Kramer-Albers EM, Lim SK, Llorente A, Lotvall J, Marcilla A, Mincheva-Nilsson L, Nazarenko I, Nieuwland R, Nolte-'t Hoen EN, Pandey A, Patel T, Piper MG, Pluchino S, Prasad TS, Rajendran L, Raposo G, Record M, Reid GE, Sanchez-Madrid F, Schiffelers RM, Siljander P, Stensballe A, Stoorvogel W, Taylor D, Thery C, Valadi H, van Balkom BW, Vazquez J, Vidal M, Wauben MH, Yanez-Mo M, Zoeller M, Mathivanan S. Vesiclepedia: a compendium for extracellular vesicles with continuous community annotation. *PLoS Biol.* 2012; 10:e1001450. [PubMed: 23271954]
29. Moreno-Gonzalo O, Villarroja-Beltri C, Sanchez-Madrid F. Post-translational modifications of exosomal proteins. *Front Immunol.* 2014; 5:383. [PubMed: 25157254]
30. Gonzales PA, Pisitkun T, Hoffert JD, Tchapyjnikov D, Star RA, Kleta R, Wang NS, Knepper MA. Large-scale proteomics and phosphoproteomics of urinary exosomes. *J Am Soc Nephrol.* 2009; 20:363–379. [PubMed: 19056867]
31. Soskic V, Gorkach M, Poznanovic S, Boehmer FD, Godovac-Zimmermann J. Functional proteomics analysis of signal transduction pathways of the platelet-derived growth factor beta receptor. *Biochemistry.* 1999; 38:1757–1764. [PubMed: 10026255]
32. Iwabata H, Yoshida M, Komatsu Y. Proteomic analysis of organ-specific post-translational lysine-acetylation and -methylation in mice by use of anti-acetyllysine and -methyllysine mouse monoclonal antibodies. *Proteomics.* 2005; 5:4653–4664. [PubMed: 16247734]
33. Kim W, Bennett EJ, Huttlin EL, Guo A, Li J, Possemato A, Sowa ME, Rad R, Rush J, Comb MJ, Harper JW, Gygi SP. Systematic and quantitative assessment of the ubiquitin-modified proteome. *Mol Cell.* 2011; 44:325–340. [PubMed: 21906983]
34. Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics.* 2002; 1:376–386. [PubMed: 12118079]
35. Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S, Purkayastha S, Juhasz P, Martin S, Bartlett-Jones M, He F, Jacobson A, Pappin DJ. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics.* 2004; 3:1154–1169. [PubMed: 15385600]
36. Thompson A, Schafer J, Kuhn K, Kienle S, Schwarz J, Schmidt G, Neumann T, Johnstone R, Mohammed AK, Hamon C. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal Chem.* 2003; 75:1895–1904. [PubMed: 12713048]
37. Lange V, Picotti P, Domon B, Aebersold R. Selected reaction monitoring for quantitative proteomics: a tutorial. *Mol Syst Biol.* 2008; 4:222. [PubMed: 18854821]
38. Peterson AC, Russell JD, Bailey DJ, Westphall MS, Coon JJ. Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics. *Mol Cell Proteomics.* 2012; 11:1475–1488. [PubMed: 22865924]
39. Lundgren DH, Hwang SI, Wu L, Han DK. Role of spectral counting in quantitative proteomics. *Expert Rev Proteomics.* 2010; 7:39–53. [PubMed: 20121475]
40. Zhang B, Kirov S, Snoddy J. WebGestalt: an integrated system for exploring gene sets in various biological contexts. *Nucleic Acids Res.* 2005; 33:W741–748. [PubMed: 15980575]
41. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 2009; 37:1–13. [PubMed: 19033363]
42. Colangelo CM, Ivosev G, Chung L, Abbott T, Shifman M, Sakaue F, Cox D, Kitchen RR, Burton L, Tate SA, Gulcicek E, Bonner R, Rinehart J, Nairn AC, Williams KR. Development of a highly automated and multiplexed targeted proteome pipeline and assay for 112 rat brain synaptic proteins. *Proteomics.* 2014
43. Burgess MW, Keshishian H, Mani DR, Gillette MA, Carr SA. Simplified and efficient quantification of low-abundance proteins at very high multiplex via targeted mass spectrometry. *Mol Cell Proteomics.* 2014; 13:1137–1149. [PubMed: 24522978]
44. Liebler DC, Zimmerman LJ. Targeted quantitation of proteins by mass spectrometry. *Biochemistry.* 2013; 52:3797–3806. [PubMed: 23517332]

45. Kirkpatrick DS, Gerber SA, Gygi SP. The absolute quantification strategy: a general procedure for the quantification of proteins and post-translational modifications. *Methods*. 2005; 35:265–273. [PubMed: 15722223]
46. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, Kern R, Tabb DL, Liebler DC, MacCoss MJ. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics*. 2010; 26:966–968. [PubMed: 20147306]
47. Escher C, Reiter L, MacLean B, Ossola R, Herzog F, Chilton J, MacCoss MJ, Rinner O. Using iRT, a normalized retention time for more targeted measurement of peptides. *Proteomics*. 2012; 12:1111–1121. [PubMed: 22577012]
48. Lotvall J, Hill AF, Hochberg F, Buzas EI, Di Vizio D, Gardiner C, Gho YS, Kurochkin IV, Mathivanan S, Quesenberry P, Sahoo S, Tahara H, Wauben MH, Witwer KW, Thery C. Minimal experimental requirements for definition of extracellular vesicles and their functions: a position statement from the International Society for Extracellular Vesicles. *J Extracell Vesicles*. 2014; 3:26913. [PubMed: 25536934]

Highlights

- Rationale for defining exosome protein cargo is presented
- Methods for global exosome protein identification are described
- Quantitative proteomics methods, both global and targeted, are discussed
- Specific proteomics protocols are provided

Figure 1. Publications in exosome research annually since 2001 retrieved from a PubMed search of the terms “exosome” (blue), “exosome + RNA” (red), and “exosome + proteomics” (green).

Figure 2.

Workflow for discovery-based protein identification by MudPIT. Proteins are prepared, digested into peptides, separated by 2D chromatography using strong cation exchange (SCX) and reverse phase (RP C18) chromatography, and peptides are analyzed by tandem mass spectrometry. As peptides are eluted, peptide ions are selected based on ion intensity in an MS spectrum, fragmented, and tandem mass spectra (MS/MS) are searched using a database search engine to generate peptide and protein identifications.

Figure 3.

Targeted analysis using multiple-reaction monitoring (MRM). MRM is most commonly performed using a triple quadrupole mass spectrometer. A peptide precursor ion of interest is selected in the first quadrupole (Q1) mass analyzer, where a specific m/z selection step filters out co-eluting ions. The selected peptide ion is then fragmented in the second, collision cell, quadrupole, and peptide-specific fragment ions are analyzed in the third quadrupole (Q3).

Figure 4. Quantitative MRM results for Rac3 peptide YLECSALTQR in urinary exosomes. Absolute quantitation was performed using an isotopically-labeled peptide standard that elutes at the same chromatographic retention time as the endogenous peptide (A). Exosome samples from five subjects were analyzed via MRM to examine changes in Rac3 following aldosterone treatment. Three replicates were performed per sample and the summed fragment ion intensities normalized to the spiked labeled peptide shown in panel B indicate down-regulation of Rac3 in all patients after aldosterone administration compared to vehicle (B).