Published in final edited form as:

Int J Mass Spectrom. 2015 February 15; 378: 264–269. doi:10.1016/j.ijms.2014.08.035.

Top-down analysis of low mass proteins in exosomes shed by murine myeloid-derived suppressor cells

Lucía Geis-Asteggianti^{1,§}, Avantika Dhabaria^{1,§}, Nathan Edwards², Suzanne Ostrand-Rosenberg³, and Catherine Fenselau^{1,*}

¹University of Maryland, College Park, MD. 20742. United States

²Georgetown University Medical Center, Washington, D.C. 20057. United States

³University of Maryland Baltimore County, Baltimore, MD. 21250. United States

Abstract

Top-down analysis is reported for a portion of the protein cargo of exosomes shed by myeloid-derived suppressor cells that participate in intracellular signaling in the tumor microenvironment. Instrument mass resolution limited the study to proteins of molecular masses below 30 kDa. A two-step fractionation strategy was used, including open tubular gel electrophoresis and C3 reverse phase high performance liquid chromatography. Twenty-one unique proteins were identified among more than 200 proteoforms, and comprising primarily two functionally important protein families: the S100 proinflammatory mediators and an abundance of histones. Fifty-six percent of the total protein in these exosomes was determined to comprise histones, of which H2B variants contribute 42 %.

Keywords

exosomes; top-down proteomics; post-translational modifications; cytokines; histones

1. Introduction

Myeloid-derived suppressor cells (MDSCs) accumulate in the tumor microenvironment, where they promote tumor growth by inhibiting innate and adaptive anti-tumor immunity [1, 2]. We have previously shown that at least some of their immune suppressive mechanisms are mediated by exosomes [3]. Exosomes [4, 5] are small membrane-bound extracellular vesicles known to participate in intercellular communication [6, 7]. In 2013, Burke et al isolated MDSC-derived exosomes of 30 nm diameter and analyzed their protein cargo using a bottom-up proteomic strategy, identifying a total of 412 proteins [3].

© 2014 Elsevier B.V. All rights reserved.

*Corresponding author at: Department of Chemistry & Biochemistry, University of Maryland, College Park, MD, 20742, United States. fenselau@umd.edu.

§These authors contributed equally to this research.

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.

Bottom-up analysis requires the use of a digestion step, generally enzymatic (e.g. trypsin), which breaks the proteins into smaller peptides. Peptide analysis by LC-MS/MS is very efficient as it provides high throughput and does not require high-end instrumentation [8]. However, identification from the bottom-up generally presents low sequence coverage, loss of labile post-translational modifications (PTMs) and it is often unable to differentiate between protein variants [9]. Over the years, development and improvements in mass resolution have allowed for top-down analysis to become an advantageous and complementary approach to bottom-up. Top-down analysis avoids protein digestion, providing 100% sequence coverage and allowing for identification of sequence variants due to mutations and gene splicing. Moreover, this technique preserves PTMs, making possible the full characterization of proteoforms and can also define the temporal and spatial relationships of multiple modifications that are potentially important in fundamental biochemistry and in the development of therapeutics. Major limitations of top-down strategies are the need for extensive sample fractionation, high-end instrumentation and efficient fragmentation methods [8, 10].

The aim of this study is to interrogate proteins in the mass range below 25 kDa in the cargo of exosomes shed by myeloid derived suppressor cells. We combine gel-eluted liquid fraction entrapment electrophoresis (GELFrEE) and reversed-phase liquid chromatography (RP-LC) for separation, an LTQ orbitrap XL mass spectrometer for top-down analysis and the software ProSightPC 3.0 for data analysis. Full characterization of proteoforms is accomplished whenever sufficient fragmentation information was available. Special focus is given to nucleic acid binding proteins, as some exosomes are known to carry microRNA and mRNA [11], which these proteins might help stabilize.

2. Materials and methods

2.1. Myeloid-derived suppressor cells and exosomes

Following previous literature [3, 12] MDSCs were harvested from the blood of BALB/c mice carrying 4T1/IL-1 β mammary carcinoma and maintained overnight at 37°C under 5% CO₂ in serum-free media to shed exosomes. Exosomes were then isolated by ultracentrifugation and characterized using a sucrose density gradient from 0.25 to 2M and electron microscopy.

2.2. Protein analysis

Intact proteins from exosomes were obtained by lysis in 8M urea and protease inhibitor cocktail (Sigma Aldrich, St. Louis, MO). Phosphatase and histone deacetylase inhibitors were not added. Histone extraction was carried out using an EpiQuik Total Histone Extraction Kit (Epigentek, Farmingdale, NY) according to the manufacturer's specifications, except the last step where the dithiothreitol (DTT) buffer was replaced by a 50 mM ammonium bicarbonate buffer. Total protein and histone contents were determined using the Pierce BCA Protein Assay Kit (Thermo Scientific, Rockford, IL) according to the manufacturer's specifications.

Three hundred micrograms of intact proteins was fractionated by tubular electrophoresis GELFrEE. Briefly, the intact proteins were initially precipitated in a mixture of (4:1:3)

CHCl₃ – MeOH – H₂O [13] and resuspended in the SDS buffer provided by the manufacturer. Disulfide bonds were reduced with 53 mM DTT for 10 min at 50°C. The sample was electrophoresed for 2 hours using a 12% Tris-acetate polyacrylamide cartridge (Expedeon, San Diego, CA) and increasing the voltage from 50 to 85 V. In order to evaluate this fractionation, 10 µL of each fraction was electrofocused on an 8-16% Tris-HCL polyacrylamide gel (BioRad, Hercules, CA). Silver staining (Thermo Scientific, Rockford, IL) was used for visualization. Ultimately, the 12 fractions collected were each re-precipitated, as described above, and re-suspended in (97.5:2.5) H₂OMeCN in 0.1% formic acid for LC-MS/MS analysis.

2.3. Liquid Chromatography Mass Spectrometry

Intact protein analysis was achieved using a Shimadzu Prominence LC system (Columbia, MD) coupled with a Thermo Fisher LTQ-Orbitrap XL (San Jose, CA) controlled by Xcalibur 2.0 software, respectively. The LC separation used an Agilent (Wilmington, DE) C3 reversed-phase capillary column of dimensions 0.1 mm × 150 mm and 5µm particle size and a linear gradient consisting of 0.1% formic acid in both (97.5:2.4) H₂O:MeCN (solvent A) and (2.4:97.5) H₂O:MeCN (solvent B), where B was increased from 0 to 80% over 200 min with a flow rate of 300 nL/min. Prior to LC separation the sample was desalted on an Agilent Zorbax trap column of dimensions 0.3 mm × 5 mm and 5µm particle size by loading the sample and rinsing with solvent A for 15 min at a flow rate of 10 mL/min. Parameters selected for both precursor and product ions included mass resolution of 60,000 (at m/z 400) and the averaging of 5 microscans. The most abundant precursor ions were isolated in a 10 Da window, using the data dependent mode, where only ions with charges greater than +4 or unassigned were selected for further fragmentation. Fragmentation was obtained using collision induced dissociation (CID) with a nominal activation energy of 25. Automatic gain control (AGC) targets were set to 10⁶ for precursor ions and 10⁵ for product ions. A 300s exclusion period was set by dynamic exclusion with a repeat count of 2 with a duration of 240 s.

2.4. Bioinformatics

The software ProSightPC 3.0 was used for spectra deconvolution and database searching against the UniProtKB mouse reference proteome database (accessed in Jan and May 2014). Deconvolution was done using the THRASH algorithm. Database search conditions were set to provide a broad precursor mass tolerance of 2500 Da, instead a product ion mass tolerance lower than or equal to 15 ppm was set. PTMs were proposed manually using the Sequence Gazer tool in the M mode. A threshold E-value of 10⁻⁵ was required to consider the identification successful. Proteins were considered fully characterized when the theoretical and experimental mass difference and E-value were lower than or equal to 15 ppm and 10⁻⁵, respectively.

3. Results and discussion

The workflow used in this study was developed through the last five years to provide top-down analysis in the chromatographic time frame. Tandem fractionation was provided by open tubular gel electrophoresis and reversed-phase LC, prior to analysis by high resolution

tandem mass spectrometry using CID for fragmentation. From a total of 600 spectra obtained, 21 distinct proteins were identified and more than 200 proteoforms were observed (Supplemental Table 1). The mass range of proteins identified was between 5 kDa and 16 kDa. Although the number of proteins identified was significantly lower than previously reported by bottom-up [3], it should be noticed that our instrument mass resolution limited the study to proteins of molecular masses below 30 kDa. Moreover, even though GELFrEE fractionation resulted into 12 fractions, which were individually analyzed by LC-MS/MS, the separation efficiency of the GELFrEE method was low and many proteins appear in more than one fraction. The authors also analyzed the exosome lysate without GELFrEE fractionation prior to LC-MS/MS and observed similar qualitative results. Overall, the identified proteins were mainly proteins from the pro-inflammatory S100 family, and nucleic acid binding proteins including histones and a non-histone chromosomal protein HMG-17. For a detailed list of identified proteins and putative modifications, refer to the Supplementary Table.

3.1. Proteins from the S100 family

One hundred seventy-four spectra were assigned to proteins in the S100 family, which included 3 unique proteins and 44 proteoforms. Modifications were observed on 43 of the 44 proteins, consistent with the increased complexity of eukaryotic cells. A proteoform of S100 A6 with removal of M1 and acetylation of A2 was identified. In the case of S100 A9 one proteoform was identified whose structure comprises removal of M1, acetylation of A2 and methylation assigned to H107 as previously reported [14]. Figure 1 shows the sequences including fragment ions observed for these proteins. The single protein identified without modification was protein S100 A8 (Table 1).

Protein S100 A8 has been reported to be present in high abundance in both the parent myeloid-derived suppressor cells (MDSC) [15] and the shed exosomes and has been shown to be chemotactic and a cytokine [3, 15-17]. Our analysis identified a number of proteoforms of S100 A8, which are listed in Table 1. All were identified in fractions 2-4 provided by open tubular gel electrophoresis. The charge states of the molecular ions selected for analysis ranged from +10 to +13. As the Table indicates, all E-values are strong, based for the most part on series of b- and y-ions formed by cleavages of amide bonds between residues 40 and 70. Occasional fragmentation in the first 30 peptide bonds has helped to localize most modifications to the N-terminus of the protein. It should be noted that large proteolytic fragments of S100 A8 were not observed, reflecting our use of protease inhibitors during cell lysis. Many of these proteoforms appear to be formed in part by removal of N-terminal Met. Removal of N-terminal Met (10,157.03 Da proteoform) as the sole modification is supported by a reduction in protein mass of 131.07 Da, and localized by observation of one b ion due to fragmentation between D14 and V15. This is the predominant proteoform, estimated by extracted ion chromatograms to contribute ~70% of the S100 A8 proteins in our lysate. Removal of M1 and acetylation of S2 (10,200.06 Da) is localized by cleavage of the S2-G3 bond. Removal of M1 and (potentially artifactual) oxidation elsewhere (10,173.02 Da) is observed. Fragment ions localize both modifications to the first 44 residues, which include M37 and C41, but not M74. Acetylation of the protein without removal of M1 is observed (10,330.11 Da), and localized at M1 by fragmentation

between M1 and S2. Removal of M1 and (artificial) dehydration generates proteins of mass 10,139.12 Da. Both of these changes are localized to the first 41 residues, which contain three serine and two threonine residues. Besides those listed in Table 1, an isoform was partially characterized, whose mass is increased by 283 Da. The unknown modification is localized on N-terminal Met by fragmentation of the M1-P2 bond and the mass differences observed for the b-ions (see Supplementary Figure 1).

Although the modest modifications proposed at the N-terminus are not expected to alter biological activity significantly, nonetheless, the assays used to assign biological activities to protein S100 A8 and its important heterodimer with S100 A9 have not recognized or discriminated between the native proteoforms identified in this study.

3.2. Histone variants and interesting PTMs

Sixteen of the proteins identified were core histones (H2A, H2B, H3 and H4). Numerous histone variants were also identified, and as expected, most variants contained several PTMs as shown by the deviation between their observed and theoretical masses (Supplementary Table 1). Moreover, histone H3.1, H3.3, H3.C and H4 were identified without modifications. A colorimetric assay of total protein abundance and of histone abundance indicates that histones compose 56% of the protein cargo in these exosomes. Furthermore, the relative abundance of each histone type was estimated by counting the spectra for each histone type, and dividing it by the total spectra counted for all identified histones across all fractions. The most abundant histones were the H2B variants (~ 42%), followed by variants of H4 (~ 23%), H3 (~ 20%), and, lastly, H2A (~ 15%). PTMs, such as loss of N-terminal methionine, acetylation, phosphorylation and methylation, were identified and localized when possible. The loss of initial Met combined with acetylation of the new N-terminus were the most common modifications observed. Figure 2 shows the modifications observed for histones H2A.1 and H2B.1C. In the case of H2A.1 the mass difference of -88.97 Da was interpreted as the loss of M1 and acetylation of S2. For H2B.1C the mass difference of -117.03 Da was located in the first 3 amino acids (MPE) and was explained by the loss of M1 and methylation of P2. Methylation in N-terminus Pro has been observed previously for H2B of *Drosophila melanogaster* Kc cells [18].

Of greater interest, proteolytic cleavages from the N-terminal were observed. The cleavage of 22 amino acid residues was confirmed for histone H3.2 and the losses of 21 amino acid residues were observed for histone H3.3 and H3.C. Examples of these proteolytic cleavages are shown in Figure 2. Duncan et al. (2008) have previously reported regioselective proteolytic cleavage of H3.2, primarily before residue 22, by cathepsin L in mouse embryonic stem cells [19]. Moreover, several studies have suggested that the localization of cathepsin proteases in the nucleus could be an indicator of cancer, though the relation of cancer and H3 proteolysis has not been proven [20, 21]. In our study the proteolytic cleavage of 23 amino acid residues of histone H4 N-terminus was also observed (Figure 3). Full characterization was demonstrated by the 22 b-ions and 27 y-ions assigned with mass errors lower than 15ppm and a strong E-value of 2.71E-54. A top-down analysis defines such processing clearly, however, its biological significance is not yet clear.

4. Conclusions

The top-down workflow utilized in this study allowed the identification of multiple proteoforms of the abundant histone and S100 protein families in exosomes. In addition to commonly reported PTMs, a noteworthy finding was the identification and characterization of proteolytically cleaved histones H3 and H4 as part of the exosome cargo. These observations on the S100 proinflammatory mediators and histones open questions such as: which of the isoforms are active? What are their functions as part of the exosome cargo? Could the histones stabilize and support transfer of microRNA in exosomes? Given the critical role of MDSC in obstructing anti-tumor immunity and promoting tumor progression, addressing these questions may identify potential drug targets for neutralizing these immune suppressive cells.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by NIH Grant GM021248.

References

1. Gabrilovich DI, Nagaraj S. Myeloid-derived suppressor cells as regulators of the immune system. *Nat Rev Immunol.* 2009; 9:162–174. [PubMed: 19197294]
2. Marx J. Cancer immunology. Cancer's bulwark against immune attack: MDS cells. *Science.* 2008; 319:154–156. [PubMed: 18187637]
3. Burke M, Choksawangkarn W, Edwards N, Ostrand-Rosenberg S, Fenselau C. Exosomes from Myeloid-Derived Suppressor Cells Carry Biologically Active Proteins. *J Proteome Res.* 2014; 13(2):836–843. [PubMed: 24295599]
4. Gould SJ, Raposo G. As we wait: coping with an imperfect nomenclature for extracellular vesicles. *J Extracell Vesicles.* 2013; 2:2892.
5. Raposo G, Stoorvogel W. Extracellular vesicles: exosomes, microvesicles, and friends. *J Cell Biol.* 2013; 200:373–83. [PubMed: 23420871]
6. Simons M, Raposo G. Exosomes--vesicular carriers for intercellular communication. *Curr Opin Cell Biol.* 2009; 21:575–81. [PubMed: 19442504]
7. Johnstone RM. Exosomes biological significance: A concise review. *Blood Cells Mol Dis.* 2006; 36:315–321. [PubMed: 16487731]
8. Garcia BA. What does the future hold for top down mass spectrometry? *J Am Soc Mass Spectrom.* 2010; 21:193–202. [PubMed: 19942451]
9. Zhang H, Ge Y. Comprehensive Analysis of Protein Modifications by Top-Down Mass Spectrometry. *Circ Cardiovasc Genet.* 2011; 4:711–711. [PubMed: 22187450]
10. Yates JR, Kelleher NL. Top down proteomics. *Anal Chem.* 2013; 85:6151. [PubMed: 23705843]
11. Valadi H, Ekström K, Bossios A, Sjöstrand M, Lee JJ, Lötvall JO. Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nat Cell Biol.* 2007; 9:654–9. [PubMed: 17486113]
12. Chornoguz O, Grmai L, Sinha P, Artemenko KA, Zubarev RA, Ostrand-Rosenberg S. Proteomic pathway analysis reveals inflammation increases myeloid-derived suppressor cell resistance to apoptosis. *Mol Cell Proteomics.* 2011; 10(3):1–9.
13. Wessel D, Flugge UI. A method for the quantitative recovery of protein in dilute solution in the presence of detergents lipids. *Anal Biochem.* 1984; 138:141–143. [PubMed: 6731838]

14. Raftery MJ, Harrison CA, Alewood P, Jones A, Geczy CL. Isolation of the murine S100 protein MRP14 (14 kDa migration-inhibitory-factor-related protein) from activated spleen cells: characterization of post-translational modifications and zinc binding. *Biochem J.* 1996; 316(Pt 1): 285–93. [PubMed: 8645219]
15. Sinha P, Okoro C, Foell D, Freeze HH. Proinflammatory S100 proteins regulate the accumulation of myeloid-derived suppressor cells. *J Immunol.* 2008; 181:4666–4675. [PubMed: 18802069]
16. Sinha, P.; Clements, V.; Burke, M.; Fenselau, C.; Ostrand-Rosenberg, S. Tumour Immunology and Immunotherapy. Rees, R., editor. Oxford University Press; Oxford U.K: 2014. p. 95-112.
17. Cheng P, Corzo CA, Luetsteke N, Yu B, Nagaraj S, et al. Inhibition of dendritic cell differentiation and accumulation of myeloid-derived suppressor cells in cancer is regulated by S100A9 protein. *J Exp Med.* 2008; 205:2235–2249. [PubMed: 18809714]
18. Zlatanova, J.; Leuba, SH. Chromatin Structure and Dynamics: State-of-the-art. Elsevier; Netherlands: 2004. p. 217
19. Duncan EM, Muratore-Schroeder TL, Cook RG, Garcia BA, Shabanowitz J, Hunt DF, Allis CD. Cathepsin L Proteolytically Processes Histone H3 During Mouse Embryonic Stem Cell Differentiation. *Cell.* 2008; 135:284. [PubMed: 18957203]
20. Denhardt DT, Greenberg AH, Egan SE, Hamilton RT, Wright JA. Cysteine proteinase cathepsin L expression correlates closely with the metastatic potential of H ras-transformed murine fibroblasts. *Oncogene.* 1987; 2:55–59. [PubMed: 3438085]
21. Lah TT, Kos J. Cysteine proteinases in cancer progression and their clinical relevance for prognosis. *Biol Chem.* 1998; 379:125–130. [PubMed: 9524063]

Highlights

- ▶ A workflow has been developed for successful top-down analysis of low mass proteins using HPLC-MS/MS
- ▶ Multiple proteoforms are identified of the pro-inflammatory mediators S100 A8 and A9 in exosomes shed by MDSC cells that suppress the immune response to cancer.
- ▶ Histone proteoforms are found to constitute 56 % of the protein cargo carried by these exosomes.
- ▶ Four histone variants are shown to carry a rare processing modification.

Figure 1.

Examples of proteins characterized from the S100 family. Sequences are shown including the observed product ions, which are represented in blue (b-ions) and red (y-ions), and their color-coded putative PTMs.

Figure 2.

Examples of fully characterized histones found in the exosomes. Sequences are shown including the observed product ions, which are represented in blue (b-ions) and red (y-ions), and their color-coded putative PTMs.

Figure 3.

MS and MS/MS spectra of proteolytically cleaved histone H4. The sequence, including observed product ions, is shown with b-ions represented in blue and y-ions in red.

Table 1

Proteoforms identified for protein S100 A8 and their putative PTMs.

Theoretical Mass (Da)	Experimental Mass (Da)	Mass Difference (Da)	# b-ions	# y-ions	E-value	Post-translational modification
10288.10	10139.12	-148.98	9	12	1.74E-13	Modification localized to first 3 amino acids
10288.10	10157.03	-131.07	28	20	8.20E-45	Removal of initial Met
10288.10	10173.02	-115.08	14	16	1.24E-23	Removal of initial Met, oxidation of Met
10288.10	10200.06	-89.04	8	17	5.77E-14	Truncation of initial Met and acetylation of Ser
10288.10	10288.10	0	15	19	1.27E-20	None
10288.10	10330.11	+42.01	1	13	7.00E-11	Acetylation of N-terminal Met