Infrared and Raman spectroscopy for purity assessment of extracellular vesicles

Zini, Jacopo

2022-05-01

Zini , J , Saari , H , Ciana , P , Viitala , T , Lohmus , A , Saarinen , J & Yliperttula , M 2022 , '
Infrared and Raman spectroscopy for purity assessment of extracellular vesicles ' ,
European Journal of Pharmaceutical Sciences , vol. 172 , 106135 . https://doi.org/10.1016/j.ejps.2022.106135

http://hdl.handle.net/10138/346701
https://doi.org/10.1016/j.ejps.2022.106135

European Journal of Pharmaceutical Sciences 172 (2022) 106135

Contents lists available at ScienceDirect

European Journal of Pharmaceutical Sciences

journal homepage: www.elsevier.com/locate/ejps

Infrared and Raman spectroscopy for purity assessment of extracellular vesicles

Jacopo Zini [a,*], Heikki Saari [a,b], Paolo Ciana [c], Tapani Viitala [a,d], Andres Lõhmus [a,#], Jukka Saarinen [d], Marjo Yliperttula [a,*]

[a] Division of Pharmaceutical Biosciences and Drug Research Program, Faculty of Pharmacy, University of Helsinki, Helsinki, Finland
[b] Finnish Red Cross Blood Service, Kivihaantie 7, Helsinki 00310, Finland
[c] Center of Excellence on Neurodegenerative Diseases and Department of Pharmacological and Biomolecular Sciences, University of Milan, Milan, MI, Italy
[d] Division of Pharmaceutical Chemistry and Technology and Drug Research Program, Faculty of Pharmacy, University of Helsinki, Helsinki, Finland

ARTICLE INFO

Keywords:
Extracellular vesicles
Purification protocols
Raman spectroscopy
IR spectroscopy
Ultra-centrifugation
Size exclusion chromatography
Differential centrifugation
Density gradient
Nanoparticle tracking analysis

ABSTRACT

Extracellular vesicles (EVs) are a complex and heterogeneous population of nanoparticles involved in cell-to-cell communication. Recently, numerous studies have indicated the potential of EVs as therapeutic agents, drug carriers and diagnostic tools. However, the results of these studies are often difficult to evaluate, since different characterization methods are used to assess the purity, physical and biochemical characteristics of the EV samples. In this study, we compared four methods for the EV sample characterization and purity assessment: i) the particle-to-protein ratio based on particle analyses with nanoparticle tracking and protein concentration by bicinchoninic acid assay, ii) Western Blot analysis for specific EV biomarkers, iii) two spectroscopic lipid-to-protein ratios by either the attenuated total reflection Fourier transform infrared (ATR-FTIR) or Raman spectroscopy. The results confirm the value of Raman and ATR-FTIR spectroscopy as robust, fast and operator independent tools that require only a few microliters of EV sample. We propose that the spectroscopic lipid-to-protein (Li/Pr) ratios are reliable parameters for the purity assessment of EV preparations. Moreover, apart from determining protein concentrations, we show that ATR-FTIR spectroscopy can also be used for indirect measurements of EV concentrations. Nevertheless, the Li/Pr ratios do not represent full characterization of the EV preparations. For a complete characterization of selected EV preparations, we recommend also additional use of particle size distribution and EV biomarker analysis.

1. Introduction

Extracellular vesicles (EVs) are lipid membrane containing particles that are released by cells and suspended in many body fluids, such as plasma, cerebrospinal fluid, semen, milk and urine. The EVs play a fundamental role in cell-to-cell communications by transferring DNA, RNA, proteins and lipids. (György et al., 2011, Buzas et al., 2014, Raposo and Stoorvogel, 2013, Colombo et al., 2014) Being natural cargo carriers, EVs with their payloads, are potentially useful as biomarkers and delivery vehicles for drugs, viruses and other biomolecules. (Saari et al., 2015, Garofalo et al., 2019) Several purification methods have been proposed to assess the purify of EVs and a great deal of effort has been done to validate and cross-analyze these methods. (Lobb et al., 2015, Coumans Frank et al., 2017, Van Deun et al., 2014) Likewise, a range of different methods to characterize EV suspensions have been proposed. However, the cross referencing of the purification methods is still missing. Therefore, the comparison of the results from different EV studies is difficult. This is particularly important, because the biological results from EV studies are not only influenced by the source of EVs, but also by the purification methods (Wiklander et al., 2015) and the storage

Abbreviations: ATR-FTIR, Attenuated total reflection- Fourier-transform infrared spectroscopy; BCA, Bicinchoninic acid assay; BPSM, Biotinylated proteins standard mix; CCM, Clarified cell conditioned media; DC, Differential centrifugation; DC+G, DC+ Density gradient centrifugation; DC+SEC, DC+ Size exclusion chromatography; EV, Extracellular vesicles; IR, Infra-red; Li/Pr, Lipid-to-protein ratio; Pa/Pr, Particle-to-protein ratio; PSM, Protein standard mix; UF, Ultrafiltration; UF+G, UF+ Density gradient centrifugation; UF+SEC, UF+ Size exclusion chromatography.

* Corresponding authors.
E-mail addresses: jacopo.zini@helsinki.fi (J. Zini), marjo.yliperttula@helsinki.fi (M. Yliperttula).
Present address: HansaBioMed Life Sciences, Mäealuse 2/1, 12618, Tallinn, Estonia.

https://doi.org/10.1016/j.ejps.2022.106135
Received 23 July 2021; Received in revised form 28 January 2022; Accepted 30 January 2022
Available online 1 February 2022

of EVs. Thus, reliable comparisons can be done only between well characterized EV samples.

Ideally, an EV preparation should be characterized in terms of the EV concentration, particle size distribution, morphology, enrichment of specific biomarkers and purity. For purity, particle-to-protein (Pa/Pr) (Webber and Clayton, 2013) and lipid-to-protein (Li/Pr) (Mihály et al., 2017, A. Gualerzi et al., 2019) ratios have been used. The characterization methods of EV purity should be fast, robust and operator independent. The Pa/Pr ratio and immunoblotting methods are the most popular characterization methods of EVs. Among numerous techniques for assessment of EV biomarker proteins, Western Blot is still the most commonly used method for protein analyses in this field. (Théry et al., 2018)

ATR-FTIR and Raman spectroscopy techniques are based on the interaction of electromagnetic radiation with matter. These spectroscopic techniques provide spectral information about the chemical composition and physical status of the analyte. Both methods have potential in discriminating between EV subpopulations and between EVs from different cells. (Paolini et al., 2020, A. Gualerzi et al., 2019, Smith et al., 2015)

In this study, the EVs extracted from two different cell sources were purified with six different purification protocols that involved ultrafiltration, density gradient centrifugation, size exclusion chromatography and differential centrifugation. The EV samples were characterized with four different methods: i) determination of the particle-to-protein ratio (Webber and Clayton, 2013) (using Nanoparticle Tracking Analysis and bicinchoninic acid assay (BCA), ii) identification of EV biomarkers with Western Blot, iii) determination of lipid-to-protein ratio by Attenuated Total Reflection Fourier-Transform Infrared (ATR-FTIR) and iv) by Raman spectroscopy. Additionally, a spiking method was used to evaluate the capability of each method to remove non-EV-associated proteins from the samples. To our knowledge, this is the first study to systematically compare all the aforementioned EV-characterization methods to each other, evaluating their strengths and weaknesses with samples from the most established EV purification methods. We found minor differences in the ranking of different purification methods depending on the characterization method used and consider vibrational spectroscopy methods as especially valuable tools for purity assessments. Overall, our results still highlight the need for multivariate analysis in evaluating the purity and quality of EV-preparations due to the multidimensional and heterogenous properties of EVs to be measured. Additionally, we demonstrate how the ATR-FTIR technique can further be utilized for an indirect estimation of EV concentrations, which further extends the applicability of ATR-FTIR for EV characterization.

2. Materials and methods

2.1. Workflow of the study

EVs derived from PNT2 and PC-3 cell cultures were collected and processed by centrifugation to remove large apoptotic bodies and cell debris to obtain clarified cell conditioned media (CCM). The medium was divided in aliquots as presented in Fig. 1 and processed as follows: 40 mL of medium was recovered after 2500 g centrifugation. Subsequently, the sample was split to different processes (see Fig. 1).

2.2. PNT2 and PC-3 cell cultures

PNT2 and PC-3 cell lines were obtained from the American Type Culture Collection (ATCC, USA). Both cell types were grown in CELLine AD 1000 bioreactors (Sigma-Aldrich, Missouri, US) at 37 °C and 5% of

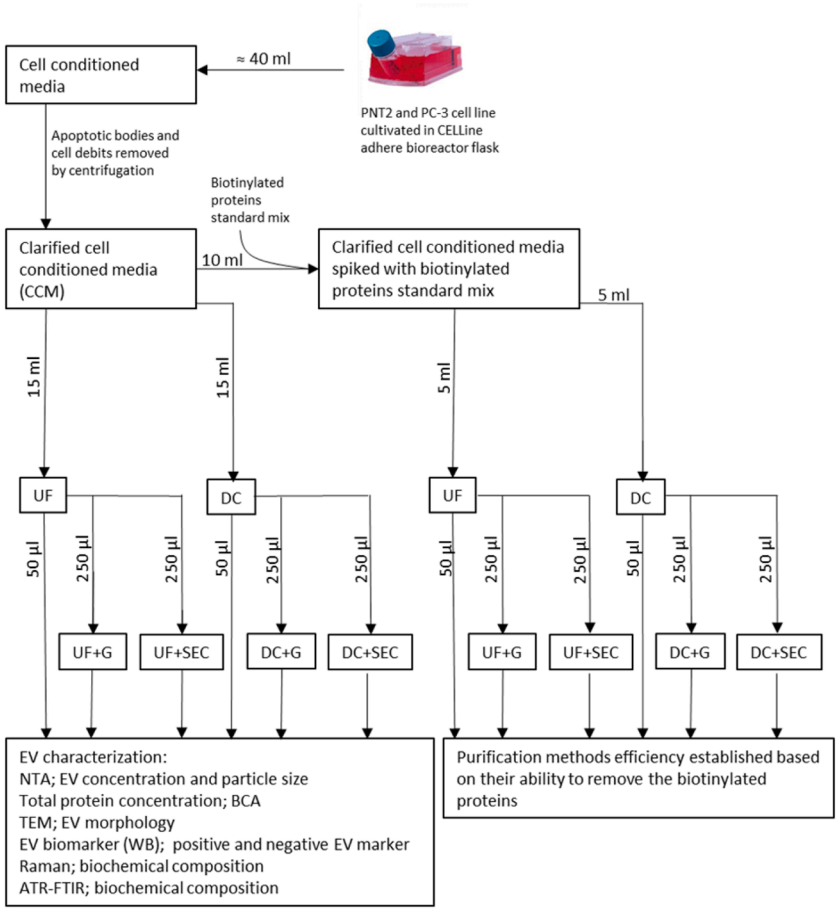

Fig. 1. The workflow of the studies is presented in the figure with the purification methods and characterization techniques used. UF = ultra-filtration, DC = differential centrifugation, G = density gradient centrifugation, SEC = size exclusion chromatography, L.C.= loading control, EV = extracellular vesicles, NTA = nanoparticle tracking analysis, BCA = bicinchoninic acid assay, TEM = transmission electron microscopy, WB = western blotting, ATR-FTIR = attenuated total reflection Fourier transform infrared.

CO_2.

The CELLine AD 1000 bioreactors have two compartments divided by a 10 kDa Nominal Molecular Weight Limit (NMWL) regenerated cellulose membrane: i) cell compartment with a 3D matrix for cell growth, and ii) Cell culture medium compartment for cell feeding and dilution of secreted factors from the cells. The cellulose membrane prevents diffusion of the EVs and macromolecules between the two compartments. Thus, macromolecules and nanoparticles in the cell culture medium do not contaminate the EVs that are released by the cells.

The two bioreactor compartments have different media. In the case of the PNT2 cell culture, the cells were grown in the cell culture compartment in 15 mL Advanced RPMI 1640 medium supplemented with glucose (4.5 g/mL) and L-glutamine (2 mM), and the outer compartment of the bioreactor contained 750 mL Advanced RPMI 1640 medium supplemented with 10% FBS and glucose (4.5 g/mL). For PC-3 cell culture, the cells were grown in the cell culture compartment in 15 mL Advanced DMEM/F-12 glucose (4.5 g/mL) and L-glutamine (2 mM), and the outer compartment of the bioreactor contained 750 mL Ham's F-12k medium with 10% FBS and glucose (4.5 g/mL). Cell culture media and FBS were purchased from Thermo Fisher Scientific (MA, USA) and glucose from Sigma-Aldrich (MO, USA). EVs were collected twice per week by aliquoting the cell culture media and washing the cell compartment with 10 mL of DPBS. The cell conditioned media was centrifuged at 2500 g (Eppendorf Centrifuge 5810 R with fixed-angle rotor Rotor FA–45–6–30, Hamburg, Germany) to remove cell debris and apoptotic bodies to obtain the clarified cell conditioned media (Fig. 1).

2.3. EV purification

2.3.1. Differential centrifugation (DC)

Clarified cell conditioned medium was loaded in 38.5 mL polyallomer tubes (Beckman Coulter, California, US) and centrifuged (Optima L-80 XP ultracentrifuge, rotor SW Ti 32) at 4 °C with 20,000 g for 1 h, k-factor 1284.0. The pellet was re-suspended in DPBS, and the supernatant was subsequently centrifuged at 4 °C with 110,000 g for 2 h, k-factor 233.4. The resulting pellet was re-suspended in DPBS and combined with the 20,000 g pellet to a total volume of 0.4 mL and stored at -80 °C. The two pellets were combined to minimize aggregation and to produce a sample containing all the EVs, as other methods do not discriminate them either.

2.3.2. Ultra-Filtration (UF)

Clarified cell conditioned medium was filtered using Amicon Ultra-15 Centrifugal Filter units with a cut-off of 10 kDa (Millipore, Massachusetts, US) at 5000 g until reaching void volume at 4 °C using the Eppendorf Centrifuge 5810 R with fixed-angle rotor (Rotor FA–45–6–30 Hamburg, Germany). After filtration, the EVs were re-suspended with 0.8–1.0 mL of Dulbecco'c Phosphate Buffer Solution (DPBS).

2.3.3. Density gradient centrifugation (G)

The EVs that were concentrated with differential centrifugation or ultrafiltration were purified by density gradient centrifugation (Fig. 1). The density gradient centrifugation was prepared by diluting a stock solution of OptiPrep (60 % w/v) (Sigma-Aldrich, MO, USA) in phosphate buffered saline (PBS), pH 7.4, to generate a 30 % w/v iodixanol solution. The density gradient centrifugation was generated by layering 6 mL of 30% (w/v) iodixanol solution below 6 mL of PBS in Ultra-ClearTM 14 × 89 mm Beckman Coulter (California, US) centrifuge tubes. The tubes were placed horizontally for 1 h at room temperature to generate a continuous density gradient through diffusion. After the incubation, 0.5 mL of samples were added on top of the tube and centrifuged (Optima L-80 XP ultracentrifuge, SWTi 41 rotor, k factor 128.3, Beckman Coulter, California, US) for 3 h with 200,000 g at 4 °C. Subsequently the density gradient was divided to 11 fractions of 1 mL each. The fractions from 5 to 7 contained most of the particles and they were collected. Iodixanol was removed through ultrafiltration with Amicon Ultra 15 mL 10 kDa Centrifugal Filters (Millipore, Massachusetts, US) at 5000 g at 4 °C. This was repeated six times by replacing the filtrate with fresh PBS.

2.3.4. Size-exclusion chromatography (SEC)

was performed with NGC Quest™ 10 Plus Chromatography System (Bio-Rad) equipped with pre-packed Superdex 200 Increase 10/300 GL column (Sigma-Aldrich, MO, USA) with exclusion limit of 1.3×10^6 Da. 250 μL of differential centrifugation or ultrafiltration samples were diluted with 250 μL of DPBS to reduce their viscosity and filtered with 0.22 μm filter. The samples were injected into the system and the eluate was collected into 1 mL fractions. The mobile phase was PBS with a flow rate of 1 mL/min. The fractions 8 and 9 with the highest EV concentrations were pooled together, and concentrated with Amicon Ultra 0.5 mL 10 kDa Centrifugal Filters (Millipore, Massachusetts, US) at 5 000 g at 4 °C.

2.4. EV characterization

2.4.1. Nanoparticle tracking analysis

Nanoparticle Tracking Analysis (NTA) was performed with a NanoSight LM10 instrument and NTA 3.1 software (Malvern Panalytical). The LM14C viewing unit was equipped with the 405 nm blue laser. To obtain the suitable concentrations for the analysis, the samples were diluted in DPBS to about 20–100 particles per frame and three 60 s videos were recorded from each of the samples using the camera level of 15. The data was analyzed using NTA software 3.1 with a detection threshold of 5 and screen gain of 10.

2.4.2. Western blot analysis

The total protein contents of EVs and cell lysates were analyzed in triplicates by the BCA protein assay kit (Thermo Scientific). Absorbance was measured with the Varioskan LUX multi-reader (v.2.4.3) (Thermo Scientific) at the wavelength of 562 nm. The cell lysate (CL) was prepared from PNT2 and PC-3 cells by culturing them on 100 mm polystyrene tissue culture plates. First, the cell monolayer was washed with 10 mL ice cold PBS. After adding 2 mL of ice cold RIPA Lysis Buffer cells were detached from the plate using a cell scraper and transferred to a 15 mL Falcon tube. The cell suspension was incubated on ice for 15 min to completely lyse the cells and, thereafter, centrifuged at 5000 g for 5 min at 4 °C. The supernatant was collected and stored at -20 °C. Equal amounts of 20 μg of EV proteins and CL were incubated in Laemmle sample buffer under reducing or non-reducing conditions at 95 °C for 10 min. The electrophoresis of the samples was done on 4–20% Mini-PROTEAN TGX™ gels (Bio-Rad). After the electrophoresis, the proteins were transferred onto the 0.2 μm nitrocellulose membrane with Trans-Blot® Turbo™ (Bio-Rad). The nitrocellulose membranes were blocked with 3% (w/v) BSA in Tris-buffered saline-0.1% (v/v) Tween 20 (TBS-T) for 1 h at room temperature. The nitrocellulose membranes were then cut into strips corresponding to the molecular weight of each protein of interest and incubated with corresponding primary antibodies in 3% BSA in TBS-T over night at room temperature. Subsequently, the strips were washed 4 times with TBS-T and incubated for 1 h at room temperature with secondary antibodies. Antibodies and their target proteins are presented in Table 1. The protein bands were visualized using ECL clarity substrate (Bio-Rad) and Chemidoc MP (Bio-Rad) imaging system. Samples for the Western blot (WB) of the TSG101, Hsp70, α-Tubulin and GM130 bands were prepared with reducing conditions, while CD9 was prepared in non-reducing conditions. The intensity of the bands was evaluated with Fiji ImageJ 1.49 software.

2.4.3. EV preparation and analysis by transmission electron microscopy (TEM)

EVs were prepared for TEM as described in Puhka et al. (2017) (Puhka et al., 2017) by loading EVs to the carbon coated and glow

Table 1
Primary and secondary antibodies and their specifications.

Primary antibody					
	Target	Condition	Dilution	Note	Manufacturer
Rabbit	GM130	Reduced	1:250	35/GM130 Isotype Mouse IgG1 Cat Code–610823	BD Transduction Laboratories, US
Mouse	Hsp70	Reduced	1:1000	Clone–7/Hsp70, Isotype Mouse IgG1 Cat Code–610607 BD	BD Transduction Laboratories, US
Mouse	α-Tubulin	Reduced	1:1500	clone 6A204	BD Transduction Laboratories, US
Mouse	TSG101	Reduced	1:250	51/TSG101, Isotype Mouse IgG1 Cat Code–612697 BD	BD Transduction Laboratories, US
Mouse	CD9	Native	1:1000	Clone–ALB 6 Isotype Mouse IgG1 Cat Code–HBM-CD9	HansaBioMed, Estonia
Secondary antibody					
goat anti-mouse IgG-HRP		//	1:5000	Polyclonal Isotype–Goat IgG Cat Code–PA1–74421	ThermoFisher Scientific, US
goat anti-rabbit IgG-HRP		//	1:5000	G-21234	ThermoFisher Scientific, US

discharged 200 mesh copper grids with piliform support membranes. EVs were fixed with 2.0% PFA in $NaPO_4$ buffer, stained with 2% neutral uranyl acetate, further stained, and embedded in uranyl acetate and methyl cellulose mixture (1.8/0.4%). EVs were viewed with TEM using Jeol JEM-1400 (Jeol Ltd., Tokyo, Japan) operating at 80 kV. Images were taken with a Gatan Orius SC 1000B CCD-camera (Gatan Inc., USA) with 4008 × 2672 px image size and no binning.

2.4.4. Attenuated total reflection Fourier transform infra-red (ATR-FTIR) spectroscopy

The ATR-FTIR data of the EVs were obtained with a PerkinElmer IR-spectrophotometer (Spectrum One spectrophotometer, Perkin Elmer Inc., Massachusetts, USA) equipped with the universal ATR sampling accessory (ZnSe Crystal). The measurements were performed at room temperature using 32 scans with a nominal resolution setting of 4 cm^{-1}. 8 μL of the sample solution with the EV concentrations of about 5×10^{11} - 2×10^{12} was placed over the ATR crystal and dried with air flow to obtain a thin layer of EVs, which covered the entire crystal surface.

The post processing of the ATR-FTIR data for lipid-to-protein (Li/Pr) ratio was calculated by the total integrated intensity of CH_2/CH_3 stretching bands from 2700 to 3000 cm^{-1} as presented in the Supplementary Fig. S1, in which the yellow area in Fig. S1C is divided by the integrated area of the deconvoluted Amide I peak at 1648 ± 5 cm^{-1} the cyan area in Fig. S1C, as suggested by Mihály J. *et al.* (Mihály et al., 2017)

ATR-FTIR was used for the concentration calibration of the EV particles. Three individual stock solutions of PNT2 derived EV suspensions (after purification by differential centrifugation) were diluted in DPBS to generate 7 different EV concentrations (21 samples in total). The EV samples were measured by NTA (see Section 2.3.1 above) and by ATR-FTIR. Each ATR-FTIR spectrum was post processed as follows: i) the baseline correction was determined at 902, 1780, 2172, 2613, 3002, 3707 and 4000 cm^{-1}, ii) the resulting spectrum was smoothed by Savitzky-Golay method, 3rd grade polynomial, with 20 smoothing points; iii) the amide I peak at 1653 ± 5 cm^{-1} was fitted with Lorentzian function. The calibration curves for each three parallel EV concentration series were obtained by fitting the amide I area under the curve (AUC) versus the EV particle concentration from NTA. We tested the calibration curve with t three new PNT2 derived EV samples purified by DC. The same data post processing as used for the construction of the calibration curve was used to estimate by interpolation the EV particle concentration of these unknown EV samples.

2.4.5. Raman spectroscopy

A confocal Raman microscope (NT-MDT Ntegra, Russia) equipped with a 532 nm laser, output power ~ 20 mW, and a 100 × objective (Mitutoyo, Japan) was used to measure the Raman spectra of EVs using a back scattering geometry. Raman peak of silicon at 520.7 cm^{-1} was used to calibrate the system daily. By measuring the full width at half maximum (FWHM) of the silicon Raman peak at 520.7 cm^{-1} a spectral resolution of ~4.4 cm^{-1} was confirmed with a 1800/500 grating. Two μL samples with EV concentrations of about 5×10^{11} - 2×10^{12} were placed on a CaF_2 substrate (LaserOptex Inc., China) and air dried for the Raman measurements. Exposure time of 10 s with the accumulation of 2 was used for the acquisition of the single spectrum. Raman spectra were measured at the edge of the dried sample. The evaporation of water from the drop induces a capillary flow which carries the dispersed material towards the edge. Therefore, the edge of the dry drop has a higher concentration of material than the central part as presented in the Supplementary Fig. S2. Each sample was measured 10 times and three parallel samples was used. For the Raman mapping (Fig. S2) an exposure time of 2 s and accumulation of 2 was used for the scanning of the area of 25 μm × 25 μm by using a mirror with 1 μm steps. Nova_Px (NT-MDT, Russia) software was used to create Raman maps and Origin2018 (OriginLab, Northampton, MA, USA) to plot spectra.

For the post processing of the Raman spectra as presented in the Supplementary Fig. S3, the baselines were drawn with Origin baseline mode least squares smoothing by the asymmetric factor of 0.001, threshold of 0.05, smoothing factor of 5, and number of interactions of 10 (Fig. S3B). The lipid component was extrapolated by the total integrated intensity of the CH_2/CH_3 scissoring peak at 1445 ± 5 cm^{-1}, and the protein component by the total integrated intensity of the Amide I peak at 1665 ± 5 cm^{-1} (Fig. S3C).

2.4.6. Spiking protein assay

The standard mix of proteins contained bovine thyroglobulin (MW ~ 670,000 D), γ-globulins from bovine blood (MW ~ 150,000 D), chicken egg albumin grade VI (MW ~ 44,300 D), and ribonuclease A type I-A from bovine pancreas (MW ~ 13,700 D) (Sigma-Aldrich, Missouri, US). The protein mix was dissolved in DPBS to yield final total protein concentration of 3.5 mg/mL. The protein solution was then biotinylated with 4.665 × 10^{-4} mmol of N-hydroxysuccinimide - biotin reagent (NHS-Biotin) dissolved in DMSO immediately before use according to the manufacturer's protocol (ThermoFisher Scientific, Massachusetts, US). Unreacted NHS-Biotin and DMSO was removed from the biotinylated proteins mix by filtration with a membrane cut-off of 3 kDa (Millipore, MA, USA) at 5000 g at 4 °C (Eppendorf Centrifuge 5810 R with fixed-angle rotor Rotor FA–45–6–30 Hamburg, Germany).

This biotinylated proteins standard mix (BPSM) was then washed twice with 1 mL of DPBS and the total protein concentration was determined with the BCA assay. After that, 1 mg of BPSM was added in 10 mL of cell conditioned media (CCM), which contained PNT2 cell derived EVs purified with the methods as previously described.

After the purification of the spiked sample, a volume containing 5 μg of the total protein was prepared and processed according to the same Western Blot (WB) protocol as described in Section 2.3.2 above. Subsequently, the samples were incubated in Laemmle buffer under non-reducing conditions at 95 °C for 10 min. The electrophoresis of the samples was done on 4–20% Mini-PROTEAN TGX™ gels (Bio-Rad). After the electrophoresis, the proteins were transferred onto 0.2 μm nitrocellulose membranes with Trans-Blot® Turbo™ (Bio-Rad). The nitrocellulose membranes were blocked with 3% (w/v) BSA in Tris-buffered saline-0.1% (v/v) Tween 20 (TBS-T) for 1 h at room temperature.

For the detection, the blocked nitrocellulose membranes were incubated with Streptavidin-HRP 1:500 in TBS-T for 2 h at room temperature, and the obtained protein bands were visualized using ECL clarity substrate (Bio-Rad Laboratories, California, US) and a Chemidoc MP (Bio-Rad Laboratories, California, US) imaging system. Signal intensities were used to estimate the residual amounts of spiking proteins after purification. The intensity of BPSM bands were then normalized based on the number of EVs measured by NTA. The results reported were then normalized against the intensity value of UF, which represents 100 % retention, as an arbitrary unit value of 1.

2.4.7. Statistics

The data were analyzed by descriptive analysis with Origin 2018 (OriginLab, Northampton, MA, US). Eventual outliers were identified by Grubbs' test ($\alpha = 0.05$) and removed. Statistical differences in the figures are displayed by compact letter display where the bars indicated by the same letter do not have any statistical differences according to the one-way ANOVA with Tukey post-hoc test.

3. Results

3.1. EV characterization

To assess the quality of the characterization methods, we first ensured that each purification method can collect EVs from the cell culture media. For minimal characterization, MISEV2018 guidelines suggest to evaluate the particle size of EVs, presence and enrichment of specific biomarkers and the EV morphology. (Théry et al., 2018) This characterization is shown in the Supplementary Figs. S4 and S5 for PNT2 and PC-3 samples, respectively. Particle sizes were measured by NTA, the presence of specific biomarkers by Western Blots and morphology by TEM. Regarding the particle sizes, neither the PNT2 (Fig. S4A) nor PC-3 EVs (Fig. S5A) show remarkable differences between different purification methods. The mean diameters were 120–150 nm for PNT2 (Fig. S4A), and 110–140 nm for PC-3 (Fig. S5A).

The PC-3 and PNT2 derived EV samples purified by ultra-filtration plus size exclusion chromatography (UF+SEC) and differential centrifugation plus size exclusion chromatography (DC+SEC) protocols displayed smaller particle sizes compared to UF and DC purified EV samples, respectively. Moreover, the SEC purified EV samples show a smaller difference between mean and mode values (Figs. S4A and S5A). The particle size distributions of PNT2 and PC-3 derived EVs are presented in the Supplementary Figs. S6 and S7, respectively. The presence of EV aggregates in the UF and DC purified samples are shown in the TEM Figs. S4C and S5C.

The EV biomarker analyzes were performed according to the MISEV2018 guidelines. (Théry et al., 2018) Two EV markers were studied: transmembrane (CD9) and cytoplasmic protein (Hsp70). The expression of the EV biomarkers differs between differently purified samples (Fig. S4B and S5B). The CD9 biomarker was present in all samples. TSG101 was particularly enriched in DC based purification protocols for both PNT2 and PC-3 derived EVs. Hsp70 was present mainly in the EV samples purified with differential centrifugation plus density gradient centrifugation (DC+G) and differential centrifugation plus size exclusion chromatography (DC+SEC) protocols used for PNT2 and in the DC+SEC purification protocols used for PC-3. MISEV2018 guidelines also recommend the blotting of a negative control, such as a protein not associated with EV biogenesis. GM130, a Golgi marker, was used as a negative control. Its expression was limited to the cell lysate and in a smaller extent in the DC purified PNT2 and PC-3 EV samples. α-Tubulin was used as representative housekeeping protein and it was present in all samples (Fig. S4B and S5B).

These results collectively show that our purification protocols can concentrate EVs from the cell culture media. The purification protocols of DC+G and DC+SEC appeared to be the most efficient, as they display high expression of EV biomarkers.

3.2. Purification protocol performance

To assess the efficacy of the purification protocols, a biotinylated proteins standard mix (BPSM) with four proteins of different molecular weights, i.e., 670, 150, 44 and 13 kDa, was added to the clarified cell conditioned media (CCM) prior to the purification. The CCM and BPSM mixture was then purified with the different purification protocols used with EV samples, namely: 1) ultra-filtration (UF), 2) UF plus gradient centrifugation (UF+G), 3) UF plus size exclusion chromatography (UF+SEC), 4) differential centrifugation (DC), 5) DC+G, 6) DC+SEC as presented in Fig. 2.

The levels of the retained BPSM were determined with WB (Fig. 2B). The intensity of the biotinylated proteins was compared to the intensity of the starting material, i.e., biotinylated proteins in cell conditioned media. Therefore, methods able to remove the biotinylated spiked proteins display a low intensity in the WB band, while methods unable to remove the biotinylated spiked proteins display high intensity bands. The cumulative bars indicate the sum of all four spiked protein bands, while 670 kDa and 150 kDa bars indicate the intensity of the 670 kDa and 150 kDa bands, respectively (Fig. 2A). The 44 kDA and 14 kDa band intensities were not plotted, because they are weak and visible only in the UF samples (Fig. 2B).

Results suggest that the purification protocol combining differential centrifugation plus size exclusion chromatography (DC+SEC) is more efficient compared to the DC plus density gradient centrifugation (DC+G) in removing the overall BPSM (cumulative) and the 150 kDa proteins (Fig. 2A). Moreover, the spiking protein method highlights that the DC based purification protocols outperforms the UF methods in all three cases. Curiously, the results point out that there is not a clear difference in removing the 670 kDa protein with the additional SEC or G purification steps within the UF and DC purification protocols.

However, we would like to underline that the results do not prove that the SEC purification protocol is superior when compared to the density gradient centrifugation (G) purification protocol, but within the purification protocols used, the DC+SEC purification protocol was the most effective in removing the spiked proteins (Fig. 2A).

3.3. Evaluation of purification methods by WB and particle-to-protein (Pa/Pr) ratio

Next, we compared particle-to-protein (Pa/Pr) ratio and EV biomarker expression as tools to evaluate the purification methods. The Pa/Pr ratio assumes that most impurities are represented by non-EV-associated proteins, that could be detected by a colorimetric protein assay (BCA) and the amount of particles measured by NTA. The comparison of the Pa/Pr ratio and EV biomarker enrichment is shown in Fig. 3. Surprisingly, the Pa/Pr ratio and the EV biomarker expressions rank the purification protocols differently for both PNT2 and PC-3 cell derived EV samples. The Pa/Pr ratios of PNT2 and PC-3 suggest that all three differential centrifugation (DC) based methods are superior to the ultra-filtration (UF) based methods, while the UF plus size exclusion chromatography (UF+SEC) is the best of the three UF based methods.

The evaluation of the purification methods with WB can be performed by comparing the expression of the EV biomarker (e.g. CD9) to the expression of the housekeeping protein, cell lysate or cell conditioned media (CCM). In our case, the comparison of the EV biomarkers between the purified samples and cell lysate was not possible for CD9, because its expression was below the detection limit. Thus, we normalized all EV biomarker band intensities obtained from WB to the ultra-filtration (UF) purified band intensities, which retains practically all molecules with molecular weight above the cut-off of 10 kDa. This method not only concentrates the EVs, as it also includes the free CCM proteins bigger than 10 kDa. Therefore, it does not qualify as a purification method *per se*. The EV biomarker enrichments relative to the UF sample are presented in Fig. S4B and S5B, and relative to the α-Tubulin in the Supplementary Fig. S8.

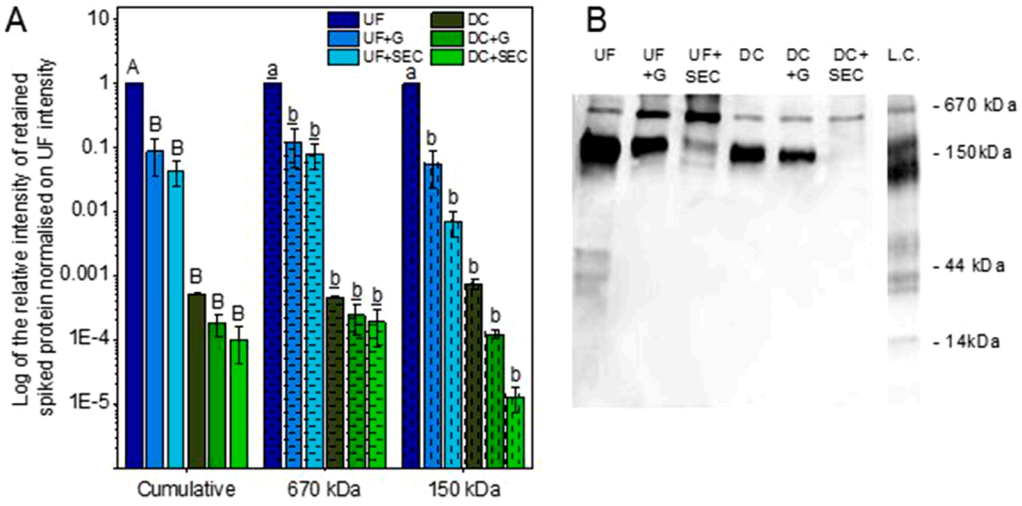

Fig. 2. Evaluation of the EV purification protocols with the spiking protein-method. (A) Relative retention of non-EV-associated proteins of the purified EV samples from the PNT2 and PC-3 cell culture media (CM) spiked with 0.1 mg/mL of four different proteins with molecular weights of 670, 150, 44 and 13 kDa. Cumulative, 670 kDa and 150 kDa bars represents the sum of the intensities of all the spiked proteins, the relative intensity of the 670 kDa and 150 kDa spiked proteins, respectively. 44 and 13 kDa proteins are not presented separately since they were detected only in the UF purified samples. (B) Representative Western Blot (WB) of the spiked proteins after each purification protocol and loading control. Intensity of the EV samples were measured with the same exposure time. Sample L.C. was separately exposed with a shorter time to avoid the saturation of the bands for sample L.C. Bars in the panels represent the mean ±1 standard error. Statistic is shown by compact letter display–Capital, lowercase and underline-lowercase letters are used for cumulative, 670 kDa and 150 kDa groups, respectively, to indicate statistically significant differences ($p < 0.05$) according to the one-way ANOVA with Tukey post-hoc test. UF = ultra-filtration, DC = differential centrifugation, G = density gradient centrifugation, SEC = size exclusion chromatography, L.C.= loading control.

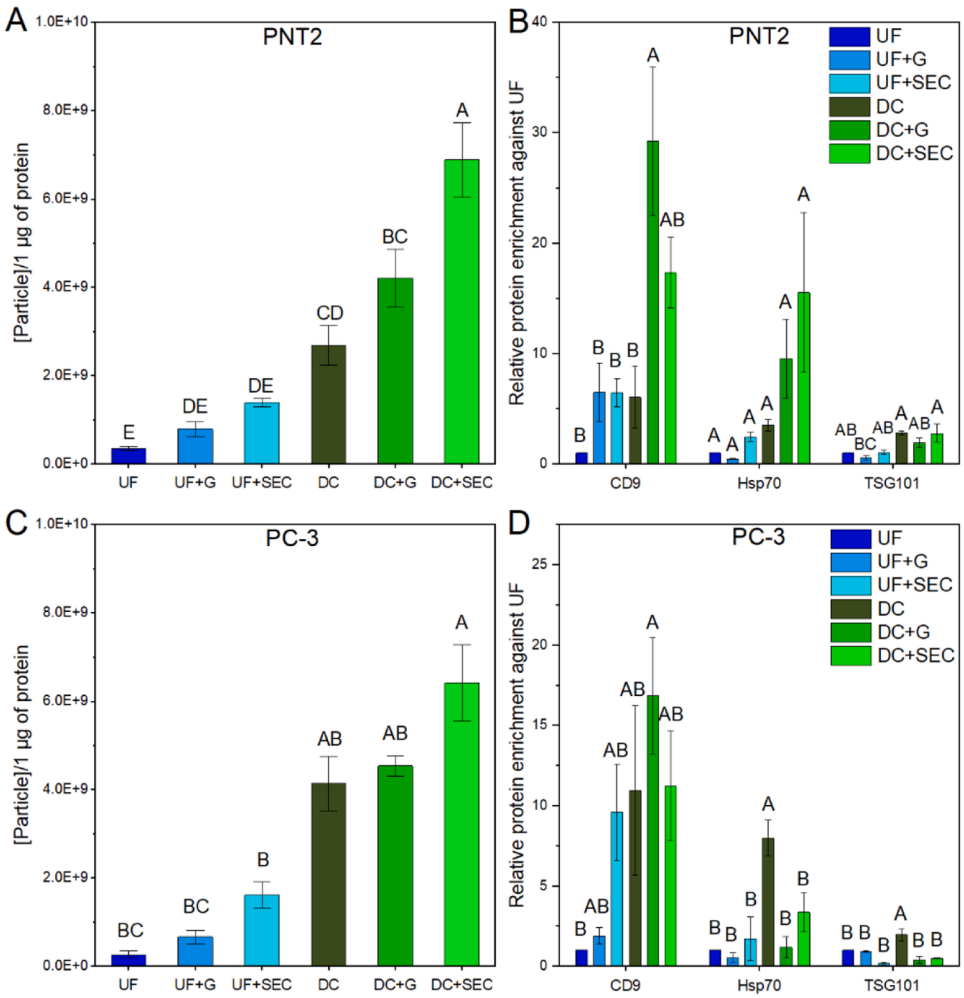

Fig. 3. Purity of the EVs evaluated by Western Blot (WB) and particle-to-protein (Pa/Pr) ratio. The data of the PNT2 samples are presented in Figs. A and B, PC-3 samples in C and D. Pa/Pr ratios of both cell line derived EVs are presented in A and C, and the EV-marker enrichment in B and D. The intensities of the enrichment of the EV markers HSP70, TSG101 and CD9 were normalized to the UF purified EVs. Bars in the panels represent the mean ±1 standard error. Statistically different groups of Pa/Pr ratios are represented by compact letter display. The statistical significance has been assessed for each individual protein. Capital letters represent the groups of Pa/Pr ratio with a statistically significant difference of $p < 0.05$ according to the one-way ANOVA with Tukey post-hoc test. UF = ultrafiltration, DC = differential centrifugation, G = density gradient centrifugation, SEC = size exclusion chromatography.

The DC+SEC and DC+G are the most efficient purification protocols in enriching EV biomarkers (Fig. S4B and S5B). On the other hand, the EV biomarker enrichment results are inconsistent, because they differ both between different biomarkers and cell lines.

In PNT2 EV samples, CD9 and Hsp70 enrichment suggest that the DC+G and DC+SEC purification protocols are the most efficient for EV biomarker enrichment (Fig. 3B). CD9 enrichment does not show any significant difference between the UF+G, UF+SEC and DC protocols, while UF is clearly the least efficient. On the other hand, the Hsp70 data ranks the purification protocols in the same order as the spiked proteins assay and Pa/Pr ratio, i.e., DC+SEC followed by DC+G, DC and UF based methods. Expression of TSG101 in PNT2 derived EVs do not provide any conclusive information other than that the DC based purification protocol seems to be more efficient than the UF (Fig. 3B and Fig. S4B).

PC-3 cell line derived EVs show the highest enrichment of CD9 in the DC+G purified samples followed by DC+SEC, DC and UF+SEC (Fig. 3D). Hsp70 and TSG101 data, however, suggest that the DC purification protocol is the most efficient method (Fig. 3B). Altogether the data obtained by calculating the Pa/Pr ratio agrees with the spiked proteins results, while the WB data presents inconsistent results for the EV populations from two different cell lines and between different biomarkers.

3.4. ATR-FTIR and Raman spectra of different EV preparations

The ATR-FTIR and Raman spectra of EVs from different protocols show remarkable similarity in peak positions as presented in the Fig. 4 and in the Supplementary Figs. S9 and S10. The peak assignments are presented in Tables 2 and 3. However, the differences are evident in the fingerprint region (800–1400 cm^{-1}) of the ATR-FTIR spectra and in the peak intensities in Raman and ATR-FTIR spectra. IR and Raman spectra are normalized based on peaks associated to lipids: CH_2 and CH_3 stretching vibration (2924 cm^{-1}) for ATR-FTIR and the CH_2 deformation vibration (1442 cm^{-1}) for Raman. Amide I and II peaks at 1536 cm^{-1} for ATR-FTIR and 1648 cm^{-1} for Raman decrease with more extensive purification protocols used, i.e., for DC to DC+SEC, in both PNT2 and PC3 derived EVs (Figs. S9 and S10) indicating decrease of protein content in relation to the lipids. A similar trend is also seen in the UF based

Table 2
Assignments of the IR vibrations (Krimm and Bandekar, 1986; Mendelsohn and Flach, 2002) for PNT2 and PC3 derived EVs purified with different purification protocols. Strong (s), medium (m), weak (w), very weak (vw), shoulder (sh).

Wave number (cm^{-1})	Functional Group	Assignments
985, sh		Phospholipids, triglycerides, and cholesterol ester
1080, sh	PO_2 symmetric stretching	Phospholipids
1239, sh	PO_2 asymmetric stretching	Phospholipids
1310, v	N-H in-plane bend	Amide III
1400, v	Symmetric stretching vibration of COO	Fatty acids and amino acids
1460, v	CH_3 bending vibration	Lipids and proteins
1545, s	C-N str and CHN bending	Amide II
1650, s	C=O stretching of amide	Amide I
1739 vw	C=O stretching	Lipids
2800 - 3000, m	CH_2 and CH_3 symmetric and asymmetric stretching	Lipids and proteins
3000 - 3600, m	N-H, O-H stretching	Water, Amides

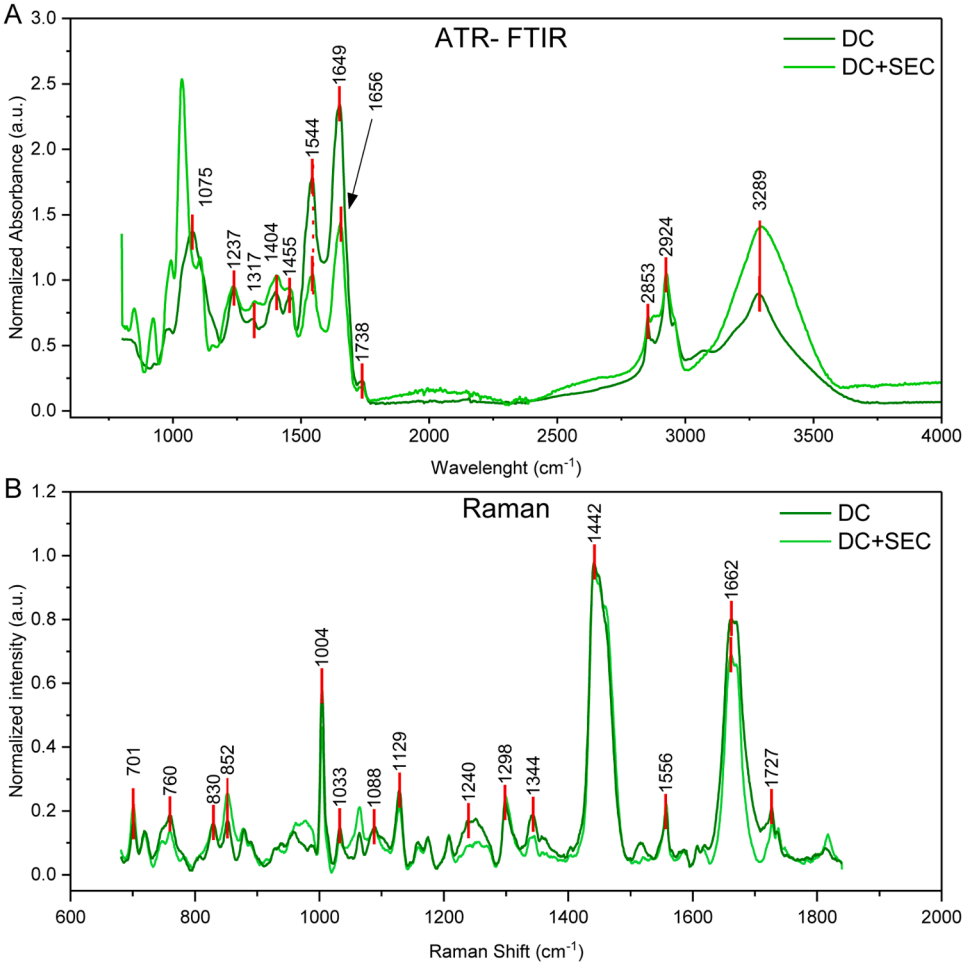

Fig. 4. ATR-FTIR and Raman spectra of PNT2 derived EVs purified by DC and DC+SEC. (A) ATR-FTIR spectra of EV preparations purified by DC (dark green) and DC+SEC (light green), normalized based on the CH_2 and CH_3 stretching vibration (2924 cm^{-1}). (B) Raman spectra of EV preparations purified by DC (dark green) and DC+SEC (light green), normalized based on the CH_2 deformation vibration (1442 cm^{-1}). Relevant peaks are highlighted by red vertical lines. UF = ultrafiltration, DC = differential centrifugation, G = density gradient centrifugation, SEC = size exclusion chromatography.

Table 3
Assignments of Raman vibrations (Czamara, 2015; Rygula, 2013), for PNT2 and PC3 derived EVs purified with different protocols. Strong (s), medium (m), weak (w) or very weak (vw).

Raman shift (cm^{-1})	Functional group	Assignments
701 vw	Cholesterol ester	Cholesterol ring deformation
760 m	Aromatic residues	Trp-NH, side chain of proteins
830 w	Aromatic residues	Trp, side chain of proteins
850 m	C–C	Protein backbone
1006 s	Aromatic residues	Phenol, side chain of proteins
1130 m	Aromatic residues	Phenol, side chain of proteins
1210 w	C-C twisting	Lipids, triacylglycerols
1298 m	CH_2 deformation	Lipids, triacylglycerols and cholesterol
1340 w	δ C–H (CH_2)	Glycosaminoglycans
1441 s	CH_2/CH_3 scissoring	Lipids and proteins
1556 w	Indole ring	Trpythophan
1661 s	Amide I	Proteins
1720 w	Ester bond	Lipid

purification (Figs. S9 and S10). The iodixanol IR spectrum as an impurity was subtracted from the IR spectra of the density gradient centrifugation purified EV samples as explained in the Supplementary material (Fig. S11) and, therefore, iodixanol has no effect on the IR spectra of the purified EVs. No differences in Raman and IR spectra were seen in the different samples, except the relative intensities of the Amide I peak at 1661 cm^{-1} and CH_2/CH_3 scissoring peak at 1441 cm^{-1} (Fig. 4). Possible interference of Iodixanol in Raman measurements should be reflected through its peculiar peak at 1513 cm^{-1}, which, however, is not present in any Raman spectra measured for the density gradient centrifugation (G) purified samples (Fig. S12). Thus, the concentration of iodixanol may be too low to be detected with this setup. Gualerzi et al. proposed to estimate the Li/Pr ratio by the integration of the peak area at 2750–3040 cm^{-1} (CH_2, CH_3 symmetric stretching) divided by the Amide I peak area at 1600–1690 cm^{-1}. In our case the Amide I peak intensity was weak compared to the CH_2 and CH_3 symmetric stretching peak intensities (Fig. S10). Hence, the Li/Pr ratio might lead to inaccurate results, and therefore we used the alternative C-H deformation peak at 1445 cm^{-1} as the lipid component for determining the Li/Pr ratios.

In both ATR-FTIR and Raman spectra, proteins affect both Amide I and C-H deformation peak intensities. However, the ratio of the integrated intensity of these two components compensates for this interference. The Li/Pr ratio can be seen as the C-H deformation intensity divided by the Amide I intensity (Li/Pr), where the intensities are proportional to the concentration of the functional groups in the analyzed area.

In a theoretical pure EV sample without free protein, the C-H deformation peak intensity should depend on the EV associated proteins and lipids, while the Amide I peak depends on the EV proteins and the resulting spectroscopic Li/Pr ratio of pure EVs. The EV samples in this study are derived from cell culture without direct contact with FBS. Thus, they do not contain HDL or other lipoproteins that interfere with both lipid and protein signals. The intensity of the Amide I depend on both the co-purified free proteins and EV-associated proteins, while the intensity of the peak of the C-H deformation depends on the EVs associated lipids and protein, and in addition of the co-purified free proteins. Thus, the Li/Pr ratio approaches lower values when the number of Amide I groups increases. Hence, by assuming that the Li/Pr ratio of pure EVs is constant, the Li/Pr ratio of the samples shift to lower values in proportion with the amount of impurities.

With the current experimental set up, the Raman spectra measured with a confocal system may be affected by the measurement position of the dried drop of the EV suspension. This could affect overall intensity and relative peak intensities. If the spectrum is measured from the edge of the dried sample, this may result in higher intensities compared to the central sample. (Deegan, 1997; Krafft, 2017; Park, 2017), Therefore, all the Raman spectra presented have been measured from the edge of the dried EV drops, because the capillary flow during the solvent evaporation concentrates the dispersed material in this region as shown in Fig. S2.

3.5. The quantification of EV concentration by ATR-FTIR

We also explored the feasibility of the ATR-FTIR for determining EV concentrations. After calibration with EV standards, we were able to estimate the EV concentrations of three PNT2 derived EV samples purified by differential centrifugation (DC) with unknown EV concentrations (Fig. 5, Table 4). The particle concentration estimates from ATR-FTIR were systematically 6 - 10% lower than those obtained by NTA measurements. Moreover, the standard deviation of the ATR-FTIR data is larger than that of NTA data. The quantification of the particle concentrations is based on the Amide I peak (1648 cm^{-1}) intensity, which can be determined by curve fitting with the Lorentz function, and reflects the total protein concentration in the sample. This process allows to suppress the interference of the other components, such as protein aggregates and free amino acids (1600 cm^{-1}). However, it can be argued that co-purified impurities, for instance non-EV proteins, can contribute to this peak. Therefore, we suggest that each of the EV purification protocols needs its own EV calibration curve. Other peaks, such as the CH_2 and CH_3 stretching (2700 - 3000 cm^{-1}) vibrations related to the lipids, or other data post processing methods, as partial least squares regression, might also give more accurate results.

3.6. Purity assessment by Raman and ATR-FTIR spectra

High Li/Pr ratios based on ATR-FTIR and Raman spectra are associated with a high purity of EV samples. The Li/Pr ratios determined from Raman spectra in Fig. 6 replicate nearly exactly with the results of the spiked proteins results presented previously in Fig. 2A. Fig. 2A shows that the DC+SEC purification protocol is the most efficient method for both PNT2 and PC-3 derived EVs followed by DC+G, DC, UF+SEC, UF+G and UF. However, the purification protocol of DC+SEC is significantly better compared to the other purification protocols only in the case of PNT2 derived EVs.

On the other hand, the Li/Pr ratios determined by ATR-FTIR indicate that the DC+SEC purification protocol is superior in the case of PC-3 derived EVs, while in the case of PNT2 derived EVs it is only slightly better than the other protocols (Fig. 6). It is interesting to notice that according to the Li/Pr ratios determined by IR spectra for PNT2 derived EVs, the UF+SEC is a more efficient purification protocol than the DC, which is slightly better than the UF and UF+G. On the other hand, for PC-3 derived EVs, the UF+SEC purification protocol appears to be more efficient compared to DC+G. Taken together, the Li/Pr ratios determined with Raman spectra for the different purification protocols follow accurately the spiked proteins results as presented in Fig. 2A. As presented in the Fig. 6, this suggests that the purification efficiency of PNT2 and PC-3 derived EVs improves from UF to DC+SEC purification protocols.

4. Discussion

The complexity of EV preparations makes their purity assessment a demanding task. Fast and robust methods for EV characterization are needed. Here, we compared four methods to assess the EV purity using samples from two different cell lines purified with different methods. Six common purification protocols were used to purify EV suspensions from PNT2 and PC-3 cell cultures. The purification protocols used were: 1) ultrafiltration (UF), 2) UF + density gradient centrifugation (UF+G), 3) UF + size exclusion chromatography (UF+SEC), 4) differential centrifugation (DC), 5) DC + density gradient (DC+G) and 6) DC + SEC. The two, UF and DC, methods to start with were chosen because they are suited for concentrating large sample volumes. The following polishing steps, SEC and G, are well established for further purification of EVs. To

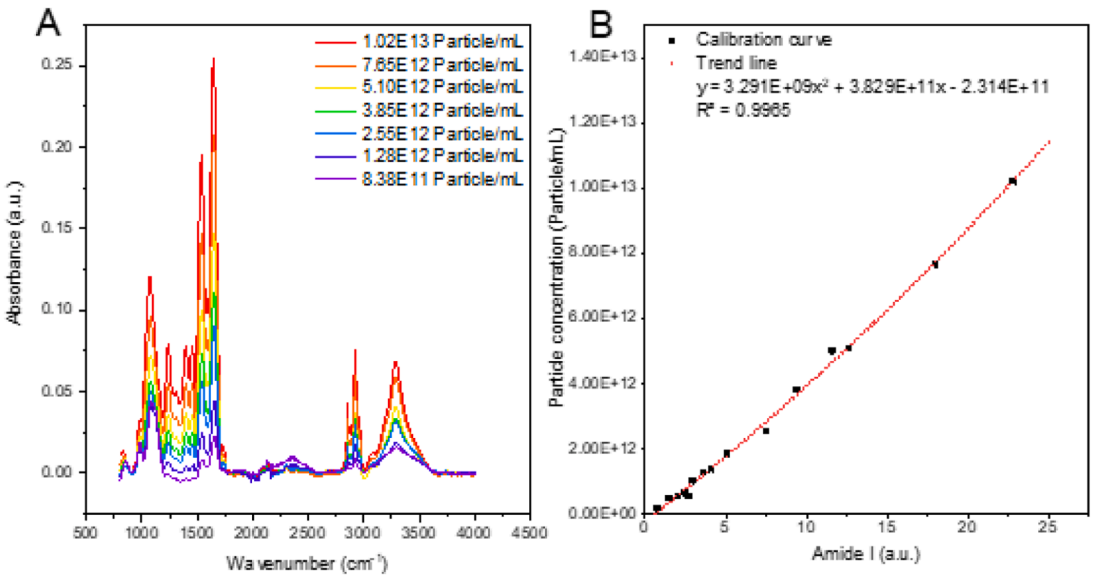

Fig. 5. Determination of the EV particle concentration by ATR-FTIR. (A) ATR-FTIR spectra of the PTN2 derived EV samples with 7 different concentrations purified with DC. (B) Calibration curve where the x-axis represents the AUC of the Amide I peak ± its standard deviation (SD) and the y-axis represents the EV particle concentrations measured by NTA ± its standard deviation (SD). Black squares indicate the EV samples, and the red dotted line represents the calculated trend line. For the calibration curve, 3 parallel PNT2 derived EV suspensions with 7 different EV concentrations diluted in DPBS were used. The particle number of each EV suspension was measured by NTA. The ATR-FTIR spectra and the respective particle numbers were used for the calibration curve.

Table 4
The AUC of the amide I peak of three different samples with unknown EV particle concentrations. The EV particle concertation was determined by interpolation with the trend line presented in the Fig. 4 and with the NTA.

Unknown EV sample	AUC	Particle concentration (number/mL) ATR-FTIR ±SD	Particle concentration (number/mL) NTA ±SD	Difference between ATR-FTIR and NTA (%)
A	2.96 ± 0.11	$9.32 \times 10^{11} \pm 3.41 \times 10^{10}$	$9.98 \times 10^{11} \pm 7.74 \times 10^{9}$	6.61
B	8.38 ± 0.35	$3.21 \times 10^{12} \pm 1.33 \times 10^{11}$	$3.44 \times 10^{12} \pm 4.80 \times 10^{10}$	6.68
C	1.85 ± 0.15	$4.90 \times 10^{11} \pm 3.90 \times 10^{10}$	$5.46 \times 10^{11} \pm 3.51 \times 10^{10}$	10.26

our knowledge, there are not yet any studies that has systematically studied and assessed this many purification protocols for EVs with a combination of up to four EV sample characterization and purity assessment methods. Commonly similar studies so far have reported the use of two to three purification protocols for EVs combined with a couple of methods for characterizing and assess the purity of EVs. Also, the use and comparison of both IR and Raman techniques to characterize and assess purity of EVs has so far not been reported in the literature.

Table 5 summarizes the information that can be obtained by the characterization methods used for EV purity assessment in this study. Also, a comparison of the time needed for the measurements and analysis, operator dependency and amounts of samples needed per measurement is included in the table.

4.1. Spiking protein assay

As there is not clear consensus about the most efficient purification protocols to remove the non-EV material, we started by ranking the protocols according to their ability to remove four spiked proteins with molecular weights of 670, 150, 44.3 and 13.7 kDa. In cell culture derived EV samples, the impurities include cell debris, organelles, non-EV proteins, aggregated and soluble proteins. Large impurities, like cell debris, are easily removed by mild centrifugation, but other impurities are removed with varying efficiency depending on the chosen protocols.

The DC+SEC purification protocol was the most efficient in removing the spiked proteins, followed by the DC+G and DC methods highlighting the power of DC based methods compared to the UF methods in EV purification (Fig. 2). However, it is worth pointing out that these data only reflects the removal of unbound proteins and not about the efficacy of removing EV-like particles (cell debris, organelles). Nevertheless, in our opinion, the protein spiking assay is a useful tool to predict the quality of the purification, but it still requires some optimization.

4.2. Pa/Pr and EV biomarker enrichments (NTA, BSA and WB)

The Pa/Pr ratio assesses the purification level of the EVs by combining information obtained from particle size distribution and particle number measurements, and concentrations of proteins. However, the Pa/Pr ratio does not provide information about morphology or biochemical composition.

Both NTA data (particle size distribution and particle number) and the BCA assay (for protein concentration) bear intrinsic errors and challenges. NTA results are dependent on the operator and instrument setting. (Gross et al., 2016; Hole et al., 2013) Moreover, the NTA software cannot discriminate between EVs and other kind of nanoparticles and the NTA measurements can also be disturbed by large particles or protein aggregates. (Filipe et al., 2010) Thus, NTA does not provide an exact absolute particle concentration. In the case of the BCA protein assay, a complex EV composition can interfere with the colorimetric assay. Namely, because membrane phospholipids, in the presence of the bicinchoninic acid, have an absorbance peak similar to that of proteins (Kessler and Fanestil, 1986; Szentirmai, 2020) which can result in false protein concentrations.

Despite these drawbacks, the Pa/Pr ratios based on NTA and BCA data are robust and reproducible in assessment of EV purity. This was seen in both PNT2 and PC-3 cell culture derived EVs, for which the Pa/Pr ratio could be used to rank the purification methods on their ability to remove the non-EV proteins (Fig. 3).

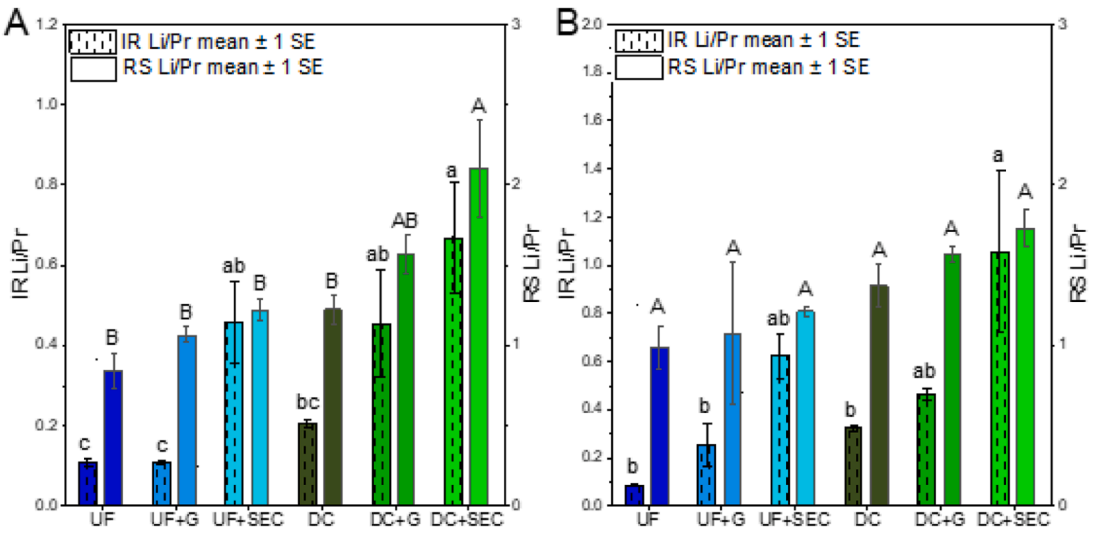

Fig. 6. Lipid-to-protein (Li/Pr) ratios obtained from ATR-FTIR and Raman spectra for PNT2 (A) and PC-3 (B) derived EVs. Bars represent the average Li/Pr ratios ± 1 standard error (SE). Statistical difference is presented as compact letter display with lower case letters for the Li/Pr ratios determined by ATR-IR and capital letters for Li/Pr ratios determined by Raman spectra. Samples with a statistically significant difference of $p < 0.05$ according to the one-way ANOVA with Tukey post-hoc test belong to different groups. In the figures, the purification protocols used were UF, UF+G, UF+SEC, DC, DC+G and DC+SEC, where UF = ultra-filtration, G = density gradient centrifugation, DC = differential centrifugation and SEC = size exclusion chromatography.

Table 5
Characterization methods for EV purity.

Method	EV concentration	Particle size distribution	Protein concertation	Biochemical composition	3D structure	Identification of EV biomarker	Time consuming	Operator dependent	Amounts of sample¤
Pa/Pr ratio	yes	yes	yes	no	no	no	IV	yes	25- 85 μL
WB	no	no	no	no	no	yes	V	yes	1–45 μL + BSA
IR spectroscopy	yes*	no	yes*	yes	yesˆ	no	II	no	10 μL
Raman spectroscopy	no*	no	no*	yes (+)	yesˆ	no	I	no	2 μL
TEM	no	no	no	no	yes	no	III	yes	10–20 μL

*= calibration is needed. IR calibration for EV protein is presented by Szentirmai *et al.* (Szentirmai et al., 2020) IR calibration for EV concentration is presented in Fig. 7.
(+) Raman spectra has more details compared with IR
ˆ= Raman and IR spectra can reveal the information of the secondary structure of the protein
¤= EV concertation ~ 5×10^{11} - 2×10^{12}

The enrichment of specific EV biomarkers is another approach to assess EV purity. Nevertheless, our data suggests that the assessment of EV purity via biomarker enrichment is not straightforward. None of the chosen EV biomarkers ranks the purification protocols according to the spiked proteins results (Fig. 3). The normalization and comparison of the different WB gels might be difficult due to the lack of reliable loading control. The MISEV 2018 position paper suggests to co-blot the EV samples with the source of the origin of the EVs.(Théry et al., 2018) However, in some cases this might be difficult if EVs are derived from different body fluids and tissues. Additionally, the analyzed biomarker might have a low expression in the cells and cannot be detected. Therefore, we suggest to co-blot the EV samples with the EV source (i.e., cell lysate) and concentrated cells conditioned media. Despite of being unreliable in purity assessment of EV samples, unlike other tested methods, the value of WB is that it can be used to show the presence of specific proteins in EV samples, including a negative marker for cell organelle derived impurities. Also, individual biomarkers may represent different EV populations that are enriched during purification. In our studies, CD9 enrichment suggested that the DC+G purification protocol is the best method for the EVs used in this study, while Hsp70 or Tsg101 did not provide any reliable information, as their expression was reduced in some purification protocols (Fig. 3). One reason for these results could be that CD9 is a membrane protein, and it is, therefore, less likely to be released as a free protein and may better represent EV enrichment.

4.3. Pa/Pr IR and Raman

IR and Raman spectroscopy methods has recently been proposed for purity assessment of air dried EV preparations. (Mihály et al., 2017, A. Gualerzi et al., 2019) However, despite of the seminal nature of these studies, only UF or UF and SEC were used to concentrate/purify EVs, and apart from other commonly used characterization methods (e.g., NTA, BCA, WB, TEM) for EVs, these studies used only either IR or Raman spectroscopy for further characterization of the EVs. Here, we have used the commonly used characterization methods for EVs (i.e., NTA, WB, BCA) and purity assesment common methods (Pa/Pr and EV marker-enrichment), as well as challenged and compared both Raman and IR techniques in assessing six different purification protocols of EVs from two different cell lines.

Our results shows that the Li/Pr ratio by Raman spectra ranks the purification protocols in the same order as the Pa/Pr ratio and spiked proteins retention assay (Fig. 6). The IR derived Li/Pr ratios display some differences. However, the UF+SEC purification protocol was as

efficient as the DC+G protocol in the case of the PNT2 derived EVs and more efficient than the DC+G protocol for PC-3 derived EVs. This partially contradicts the ranking obtained by the Pa/Pr ratio and spiked proteins retention assay. These differences in ranking the purification protocols based on the spectroscopic Li/Pr ratio might be due to the fact that different vibrations in the Raman and IR spectra were used to calculate the lipid and protein components. More specifically, in Raman, the CH_2/CH_3 scissoring peak at 1441 cm^{-1} was for lipids and the Amide I peak at 1661 cm^{-1} for protein, while for IR, the CH_2/CH_3 symmetric and asymmetric stretching at 2800–3000 cm^{-1} was used for lipids and the C=O stretching of Amide II for protein. IR and Raman spectroscopy are both based on molecular vibrations where the probability of the phenomena depends on the dipole moment for IR and polarizability of chemical bonds for Raman. Vibrations that are Raman active are not IR active and vice versa, which makes these two techniques complementary and validated against each other. Thus, it is difficult to use the same vibration of a given functional group to estimate the spectroscopic Li/Pr ratios.

IR and Raman spectroscopy have clear advantages in the characterization and purity assessment of EV preparations. They are fast, label-free, operator independent techniques and require only a small amount sample. Additionally, Li/Pr ratios based on IR and Raman spectra are self-normalizing, as both lipid and protein peaks are measured in the same measurement and their ratio compensates for the absolute quantity of these biomolecules and for the sample concentration. This enables derivation of the Li/Pr ratios, which can be compared between researchers in different laboratories. The ATR-FTIR and Raman technologies display unique advantages over the other tested characterization methods (WB, NTA/BCA), since ATR-FTIR and Raman spectra provide a fingerprint, which can be used to discriminate the EVs of different origins or different populations. (Paolini et al., 2020, Smith et al., 2015, Lee et al., 2018) Arguably, WB is able to discriminate between the EVs of different sources, but it needs pre-determined biomarkers, whereas Raman and ATR-FTIR reveal differences between unknown samples. Furthermore, IR and Raman spectroscopy provide information on the secondary protein structure that is linked to their functionality. Therefore, these spectroscopic tools could be used for the quality control of EVs in process analytical technology.

ATR-FTIR has previously been shown to be able to quantify the amounts of proteins in EVs. (Szentirmai et al., 2020) To further underline the ductility of the vibrational spectroscopic techniques, we have shown in this study, as a proof-of-concept, how to utilize ATR-FTIR to estimate EV concentrations (Fig. 5). The EV concentrations were estimated by fitting the Amide I peak intensity with a previously obtained calibration curve. This is a fast, reagent free and easy method, but it may be subjected to different imprecisions compared to NTA measurements, which are susceptible to errors as previously explained. Moreover, the Amide I intensity in ATR-FTIR seems to be related to the level of purity of the EV sample. Therefore, the calibration curve based on this specific peak might not be valid for EV samples purified with different purification protocols or methods. However, as shown in Fig. 5 and Table 4, the ATR-FTIR spectroscopy approach gives a very good estimation of the EV concentration when compared to NTA measurements. However, further optimization is still needed for ATR-FTIR to become a reliable quantification tool for EVs.

Raman spectroscopy, on the other hand, has great potential to become a quality control tool in the EV field and for EV characterization. It has been shown that the Raman spectrum of a single EV can be measured by a laser tweezer technique, (Smith et al., 2015; Lee et al., 2018) and this single EV spectrum can be used as a golden standard for EV preparations of interest. Thus, Raman spectroscopy could be used to assess the purity of plasma derived EVs with HDL and lipoprotein impurities. New innovations are continuously developed in the Raman spectroscopy field, including stimulated Raman scattering, surface/tip-enhanced Raman spectroscopy, (Park et al., 2017) coherent Raman scattering as stimulated Raman scattering (SRS) and coherent anti-Stokes Raman scattering (CARS). (Mizuguchi et al., 2020) Their developments are directly applicable to the EV field and represents potential new future characterization methods. IR spectroscopy is also continuously developed to increase its sensitivity, e.g., through Surface Enhanced Infrared Absorption Spectroscopy. However, IR spectroscopy has a clear disadvantage compared to Raman spectroscopy, i.e., water is present in all biological samples, and water has a very strong absorption in the IR region, which consequently often disturbs the spectra in the IR region. On the other hand, in Raman, water has only low intensity Raman scattering, and water containing samples can easily be analyzed with Raman spectroscopy.

Based on this study, none of the characterization methods alone can characterize the EV samples, but a combination of methods is required. Immunoblotting is valuable for identifying EV-associated and non-associated biomarkers to confirm enrichment of EVs and absence of contaminants, but it does not provide solid information about the sample purity. Biomarkers must be chosen carefully and validated, as the WB assay only gives information about the chosen biomarkers and acts as a semi-quantitative purity assay. The Pa/Pr ratio obtained by NTA and BCA is straightforward, but it is user and instrument dependent and does not discriminate between EVs and other particles. Vibrational spectroscopy-based methods are self-standardizing, operator independent and provide information about the chemical composition of the sample. They are reliable for qualitative and quantitative assays, but except tertiary protein structures, unable to give information on individual biomarkers, particle size or morphology. Therefore, we recommend the use of Pa/Pr ratio or spectroscopic Li/Pr ratio over the EV biomarker enrichment as a purity assessment method. However, neither Raman nor IR spectroscopy or Pa/Pr ratio can fully characterize EV preparations. Due to the heterogenous properties of EVs a combination of spectroscopic and protein assays together with NTA and TEM measurements are needed to fully characterize the purity, morphology, and biomarkers of EV preparations.

5. Conclusion

Our results show that Raman and ATR-FTIR spectroscopy are robust, fast and operator independent methods for assessing the quality, composition, and purity of EVs. Particles to proteins (Pa/Pr) ratios determined by NTA and BCA assay, and the spectroscopic lipids to proteins (Li/Pr) ratio are reliable parameters to assess the EV purity. Among the characterization techniques, the ATR-FTIR spectroscopy presents some advantages over Raman spectroscopy, since ATR-FTIR spectra can be calibrated to determine the amount of proteins and the concentration of EVs within one measurement. Even though ATR-FTIR and Raman spectroscopy have clear advantages, neither of these techniques provide information concerning the particle size distribution, the presence of specific biomarkers or the EV morphology. Therefore, we suggest to use at least one vibrational spectroscopy technique along with particle size measurements and a biomarker assay for EV characterization. This approach would facilitate data comparisons between different laboratories in the future.

CRediT authorship contribution statement

Jacopo Zini: Conceptualization, Methodology, Validation, Investigation, Writing – original draft, Writing – review & editing. **Heikki Saari:** Conceptualization, Methodology, Validation, Investigation, Writing – review & editing, Supervision. **Paolo Ciana:** Validation, Writing – review & editing. **Tapani Viitala:** Validation, Resources, Writing – original draft, Writing – review & editing, Supervision, Funding acquisition. **Andres Lõhmus:** Methodology, Validation, Writing – review & editing. **Jukka Saarinen:** Methodology, Validation, Writing – review & editing. **Marjo Yliperttula:** Conceptualization, Validation, Investigation, Resources, Writing – original draft, Writing – review & editing, Supervision, Funding acquisition.

Declaration of Interest Statement

The authors have no conflict of interest.

Acknowledgments

Researchers from the University of Helsinki, Faculty of Natural Sciences at Kumpula campus for allowing the use of the Raman device, Leena Pietilä for the help in cell culturing and BSc Olli-Petteri Nivaro for assisting with EV purifications. We are also thankful to prof Arto Urtti for his valuable help and advice with the revisions of the manuscript. We are thankful for the funding from Academy of Finland projects no 314406 (MY) and no 315406 (MY); Business Finland EVE ecosystem 1842/31/2019 (MY); Aaltonen Foundation (HS).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ejps.2022.106135.

References

Buzas, E.I., György, B., Nagy, G., Falus, A., Gay, S., 2014. Emerging role of extracellular vesicles in inflammatory diseases. Nat. Rev. Rheumatol. 10, 356–364.
Colombo, M., Raposo, G., Théry, C., 2014. Biogenesis, secretion, and intercellular interactions of exosomes and other extracellular vesicles. Annu. Rev. Cell Dev. Biol. 30, 255–289.
Coumans Frank, A.W, et al., 2017. Methodological guidelines to study extracellular vesicles. Circ. Res. 120, 1632–1648.
Czamara, K., et al., 2015. Raman spectroscopy of lipids–A review–Raman spectroscopy of lipids. J. Raman Spectrosc. 46, 4–20.
Deegan, R.D., et al., 1997. Capillary flow as the cause of ring stains from dried liquid drops. Nature 389, 827–829.
Filipe, V., Hawe, A., Jiskoot, W., 2010. Critical evaluation of nanoparticle tracking analysis (NTA) by nanosight for the measurement of nanoparticles and protein aggregates. Pharm. Res. 27, 796–810.
Garofalo, M., et al., 2019. Extracellular vesicles enhance the targeted delivery of immunogenic oncolytic adenovirus and paclitaxel in immunocompetent mice. J. Control Release 294, 165–175.
Gross, J., Sayle, S., Karow, A.R., Bakowsky, U., Garidel, P., 2016. Nanoparticle tracking analysis of particle size and concentration detection in suspensions of polymer and protein samples–Influence of experimental and data evaluation parameters. Eur. J. Pharm. Biopharm. 104, 30–41.
Gualerzi, A., et al., 2019a. Raman spectroscopy as a quick tool to assess purity of extracellular vesicle preparations and predict their functionality. J. Extracell. Vesicles 8, 1568780.
Gualerzi, A., et al., 2019b. Raman profiling of circulating extracellular vesicles for the stratification of Parkinson's patients. Nanomed. Nanotechnol. Biol. Med. 22, 102097.
György, B., et al., 2011. Membrane vesicles, current state-of-the-art–Emerging role of extracellular vesicles. Cell. Mol. Life Sci. 68, 2667–2688.
Hole, P., et al., 2013. Interlaboratory comparison of size measurements on nanoparticles using nanoparticle tracking analysis (NTA). J. Nanopart. Res. 15, 2101.
Kessler, R.J., Fanestil, D.D., 1986. Interference by lipids in the determination of protein using bicinchoninic acid. Anal. Biochem. 159, 138–142.
Krafft, C., et al., 2017. A specific spectral signature of serum and plasma-derived extracellular vesicles for cancer screening. Nanomed. Nanotechnol. Biol. Med. 13, 835–841.
Krimm, S. & Bandekar, J. Vibrational spectroscopy and conformation of peptides, polypeptides, and proteins. in Advances in Protein Chemistry (eds. Anfinsen, C. B., Edsall, J. T. & Richards, F. M.) vol. 38 181–364 (Academic Press, 1986).
Lee, W., et al., 2018. Label-free prostate cancer detection by characterization of extracellular vesicles using Raman spectroscopy. Anal. Chem. 90, 11290–11296.
Lobb, R.J., et al., 2015. Optimized exosome isolation protocol for cell culture supernatant and human plasma. J. Extracell. Vesicles 4, 27031.
Mendelsohn, R. & Flach, C.R. Infrared reflection-absorption spectroscopy of lipids, peptides, and proteins in aqueous monolayers. in current topics in membranes vol. 52 57–88 (Academic Press, 2002).
Mihály, J., et al., 2017. Characterization of extracellular vesicles by IR spectroscopy–Fast and simple classification based on amide and CH stretching vibrations. Biochim. Biophys. Acta 1859, 459–466.
Mizuguchi, T., et al., 2020. Multimodal multiphoton imaging of the lipid bilayer by dye-based sum-frequency generation and coherent anti-stokes Raman scattering. Anal. Chem. 92, 5656–5660.
Paolini, L., et al., 2020. Fourier-transform infrared (FT-IR) spectroscopy fingerprints subpopulations of extracellular vesicles of different sizes and cellular origin. J. Extracell. Vesicles 9, 1741174.
Park, J., et al., 2017. Exosome classification by pattern analysis of surface-enhanced raman spectroscopy data for lung cancer diagnosis. Anal. Chem. 89, 6695–6701.
Puhka, M., et al., 2017. KeepEX, a simple dilution protocol for improving extracellular vesicle yields from urine. Eur. J. Pharm. Sci. 98, 30–39.
Raposo, G., Stoorvogel, W., 2013. Extracellular vesicles–Exosomes, microvesicles, and friends. J. Cell Biol. 200, 373–383.
Rygula, A., et al., 2013. Raman spectroscopy of proteins–A review–Raman spectroscopy of proteins. J. Raman Spectrosc. 44, 1061–1076.
Saari, H., et al., 2015. Microvesicle- and exosome-mediated drug delivery enhances the cytotoxicity of Paclitaxel in autologous prostate cancer cells. J. Control Release 220, 727–737.
Smith, Z.J., et al., 2015. Single exosome study reveals subpopulations distributed among cell lines with variability related to membrane content. J. Extracell. Vesicles 4, 28533.
Szentirmai, V., et al., 2020. Reagent-free total protein quantification of intact extracellular vesicles by attenuated total reflection Fourier transform infrared (ATR-FTIR) spectroscopy. Anal. Bioanal. Chem. 412, 4619–4628.
Théry, C., et al., 2018. Minimal information for studies of extracellular vesicles 2018 (MISEV2018)–A position statement of the International Society for Extracellular Vesicles and update of the MISEV2014 guidelines. J. Extracell. Vesicles 7, 1535750.
Van Deun, J., et al., 2014. The impact of disparate isolation methods for extracellular vesicles on downstream RNA profiling. J. Extracell. Vesicles 3.
Webber, J., Clayton, A., 2013. How pure are your vesicles? J. Extracell. Vesicles 2, 19861.
Wiklander, O.P.B., et al., 2015. Extracellular vesicle in vivo biodistribution is determined by cell source, route of administration and targeting. J. Extracell. Vesicles 4, 26316.