Exosomes isolated from two different cell lines using three different isolation techniques show variation in physical and molecular characteristics

Manish Dash[1], Kanagaraj Palaniyandi[3], Satish Ramalingam[4], S. Sahabudeen[5], N S Raja[2*]

[1,2,4]Department of Genetic Engineering, SRM Institute of Science and Technology, Kattankulathur, Chennai, 603203, Tamil Nadu, India

[3,5] Department of Biotechnology, SRM Institute of Science and Technology, Kattankulathur, Chennai, 603203, Tamil Nadu, India

*Corresponding Author.

Email address: rajan3@srmist.edu.in

Abstract

Exosomes are the nanoscopic lipid bi-layered extracellular vesicles that can deliver molecular medicine in the form of targeted therapeutics with great fidelity demonstrating enhanced permeability and high retention effect. In our investigation, we focused on the comparison of three major exosome isolation techniques based on the biophysical and physicochemical characteristics of exosomes isolated from COLO 205 and MCF-7 cancer cell's conditioned media. Commercially available Total Exosome Isolation reagent (TEI), Protein organic solvent precipitation (PROSPR) and differential ultracentrifugation are the three methods used for isolation with an aim to select a suitable method for clinical translation.

HRTEM images of exosomes are subjected to 3D image analysis and particle size distribution of exosomes depicted the morphological differences. Molecular and analytical characterization of exosomes using western blotting, Raman and ATR-FTIR spectroscopy and the multivariate analysis on the spectral data obtained, assessed for better molecular specifications and purity of particle. TEI method isolated exosomes with higher exosomal yield, purity, and recovery directly translatable into drug delivery and targeted therapeutics whereas ultracentrifuge had good recovery of particle morphology but showed particle aggregation and yielded exosomes with smaller mean size. PROSPR technique isolated a mixture of EVs and showed the least recovery of exosomes in particle size distribution. This comparative study should help in choosing a specific exosome isolation technique required for the objective of downstream applications.

Keywords: *exosomes, molecular medicine, extracellular vesicles, Total Exosome Isolation reagent, PROSPR, RAMAN spectroscopy, ATR-FTIR spectroscopy, multivariate analysis, targeted therapeutics, particle size distribution*

Introduction

Exosomes are the minuscule ultrafine bioparticles, the lipid vesicles with a potential to carry cellular cargo from one cell to another[1,2]. They are involved in cell–cell communication playing a major role in growth and development, immune response and even in regulating the tumor microenvironment and disease progression. Exosomes are efficient and robust in their function and have versatile cargo composed of uniquely sorted material carrying it from the cytosol of parent cell to the target cell throughout the body[3]. They are highly efficient due to their non – immunogenic delivery, their ability to easily fuse with the cell membrane, avoid phagocytosis and circumventing the lysosomes via bypassing the engulfment[4,5].

Exosomes are the ideal drug delivery system with the applications of targeted therapeutics, prognosis and diagnostics of the diseases[6]. This is because the synthetic lipid nanoparticles like lipoplexes can only be used *in vitro* and not *in vivo* as a delivery system for gene therapy and transfection[7–9]. Also, in animal models, these lipoplexes elucidate a strong immune response with increased toxicity levels and accumulation in the liver, hindering the drug activity in the target[10].

The efficiency, biocompatibility, low accumulation of exosomes in organs and tissues with low toxicity levels are been harnessed for multiple novel therapeutic strategies[11–13]. During the development of exosome-based therapeutics and diagnostics, it's important to keep in mind about the biophysical and molecular characteristics of these particles. The attributes of these particles like their morphology and topology, the size of the particle and its uniformity, optimization of the storage conditions and the mode of the delivery of particles into the target cells with improvement in the medicament potential of these particles play important role in engineering and targeting the exosomes for the treatment and therapy[14,15]. Isolation method of exosomes that can give a higher yield with better purity and recovery of the exosomes play a crucial role in the collection of high-quality exosomes for scaling up the operations in the industry[16–18].

In our present study, we compared the differential physical and molecular characteristics of exosomes isolated using three different isolation protocols namely, the isolation via Total exosome isolation reagent from Invitrogen (TEI), via Protein Organic Solvent Precipitation method (PROSPR) and via Differential Ultracentrifugation (UC). The exosomes were isolated from two metastatic cell lines, COLO 205 a colon adenocarcinoma and MCF-7 a human breast cancer cell line. The exosomes from both the cell lines were isolated from their serum-free conditioned media and were quantified for their yield and proteome. The exosomes were characterized by HRTEM and the images were analysed for the recovery of

morphology and distribution of particles, SDS PAGE analysis was done to observe the difference in their proteomic profiles and Western Blotting was done for molecular characterization using CD63 marker. Raman Spectroscopy and ATR-FTIR were performed for fingerprinting of the molecular components of exosomes respectively. Statistical analysis and multivariate analysis of the Raman and ATR-FTIR spectra were done to check for the variation in components of spectra of the exosomes isolated from two different cell line sources using three different methods. Our investigation provided the basis to select the suitable protocol for isolating exosomes, having higher yield and better recovery of exosomes with ideal morphology that could be applied to drug delivery and targeted therapeutics.

Materials and Methods

Cell culture and isolation of exosomes

COLO 205 colon adenocarcinoma cells and MCF-7 human breast cancer cells were culutured in high glucose DMEM (HiMedia®, India), supplemented with 10 % FBS (US Heat Inactivated HiMedia®, India) and 2% penicillin/streptomycin (antibiotic antimycotic solution HiMedia®, India). The cells were incubated at 37°C with 5% CO_2 supply and were allowed to grow until 70% confluency. The cells were replaced in FBS depleted media and allowed to grow up to 85-90% confluency in FBS free media from which exosomes were isolated.

The isolation of the exosomes was performed via three different methods namely, Total Exosome Isolation reagent (TEI) (Invitrogen, USA), Protein Organic Solvent Precipitation i.e. acetone precipitation mediated (PROSPR) and Differential Ultracentrifugation (UC). The protocols used for the isolation by precipitation based TEI reagent was taken from the TEI Invitrogen manufacturer's manual, the PROSPR method taken was developed by Xavier

Gallart-palau et al[19] and UC was taken from the manuscript of Clotilde Théry et al[20].The schematics for step by step protocol is given in figure1 below.

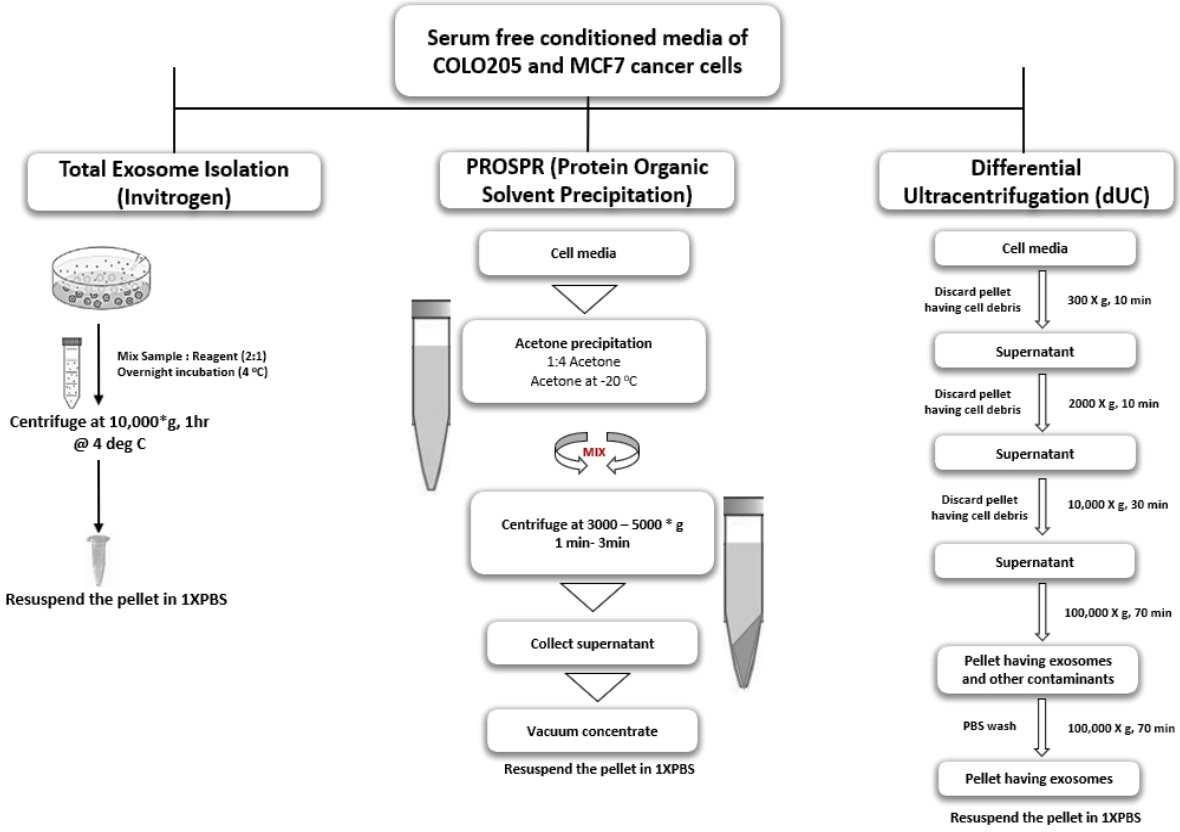


Figure 1. The three isolation techniques used in isolation of exosomes from serum-free conditioned media of COLO 205 and MCF-7 cancer cells namely, Total Exosome Isolation reagent (TEI) (Invitrogen, California, USA), the method where the media to be processed was taken and mixed with the reagent in 2:1 ratio, vortexed properly and incubated overnight at 4^{o}C. The exosomes were pelleted down at 10,000g for 60 min at 4^{o}C according to manufacturer's manual and were resuspended in 100µl of 1 X PBS for further studies. For PROSPR, the media from which exosomes had to be isolated was mixed with pre-chilled ice-cold acetone in a 1:4 ratio and vortexed. This solution was centrifuged at 3000g for 2 min and the supernatant was collected and concentrated in a vacuum concentrator in vacuum-alcoholic (V-AL) mode (Eppendorf plus, Eppendorf, Hamburg, Germany). The concentrated crystals were resuspended in 100µl of 1 X PBS and Nuclease Free Water (NFW). The

ultracentrifugation of the samples was done with a TLA-055 rotor of Beckman Coulter Optima MAX-XP ultracentrifuge (California, USA). The isolation by differential ultra-centrifugation begins after the pre-processing of the conditioned media obtained from the two cell lines COLO 205 and MCF-7 with two pre spins at 3000g for 5 min at 4°C where the pellet was discarded and the supernatant was taken and spun down at 10000g for 30 min in 4°C. This was done to remove any unnecessary cell debris. The pre-processing of the samples obtained was carried out by spinning them in lower centrifugal speed at 300g and gradually increasing the speed to 2000g and to 10,000g consecutively removing the cell debris in the form of a pellet. The supernatant collected from this stage was spun at 1,00,000g twice using ultracentrifuge at for 70 min each to isolate exosomes in the pellet at the final step. The first spin at ultra-high speed was done to remove the bigger vesicles. The supernatant was discarded and the pellet was washed with 1 X PBS in the second spin. The pellet (almost invisibly small in size) were resuspended in 100 µl of 1 X PBS and NFW.

One set of the samples of COLO 205 and MCF-7 conditioned media, were treated with 1X Proteinase K and incubated at 37°C for 30 min. The proteinase K activity was inactivated by incubation at 60°C for 10 min. This was done to check for the reduction of extracellular protein debris. The conditioned media were then processed further using the above three techniques for the isolation of exosomes.

Quantification of exosome yield and exosome proteome

The quantification of the yield of exosomes by Bradford's reagent was done by taking 5µl of exosome resuspended sample and mixing it with 500µl of Bradford's reagent, HiMedia® India. The quantification of the total protein content was done by digesting the exosomes with RIPA buffer described by Alcaraz C et al[21]. For doing so 5µl of sample from the stock suspension was mixed with 5µl 1X lysis buffer and vortexed vigorously. The mixture was

heated at 70℃ for 10 min for membrane lysis and was mixed with 500µl of Bradford's reagent and incubated in dark for 10 min. The absorbance was read at 595nm in UV-Vis spectrophotometer (Jasco V-730 UV-Vis spectrophotometer, USA).

The quantification of Proteinase K treated exosome samples and their proteome content was done via the same procedure as above stated. All the experiments were done in triplicates and the results were analysed using OriginLab Pro 2020 ver. 9.4.

High-Resolution Transmission Electron Microscopy

Jeol TEM – 2100 plus (Tokyo, Japan) was used for imaging of the exosomes. The imaging was done at high-resolution TEM with a magnification scale ranging from 20nm to 1µm with an applied voltage of 220 kV. For imaging, the samples were diluted 1:50 times and 2µl from the diluted working stock was made to set on a carbon-coated copper grid. No fixation or negative staining steps were performed for sample preparation. The grid was incubated overnight at 37℃ with a desiccant kept in a closed chamber and was taken to the facility the next day for imaging.

Image Analysis of the HRTEM images obtained for Exosomes

The 2D contour maps and 3D surface simulation of exosome images obtained from HRTEM were used to create particle size distribution plots for understanding the morphology, size variation, and distribution of exosomes. The images were prepared in Fiji ImageJ software for the particle size distribution analysis by calibrating their scale, converting the images into an 8-bit grayscale type and adjusting the threshold values for binarization and particle tagging. 2D contour maps for calculation of particle parameters were generated in OriginLab software where the parameters were extracted from the contour line data matrix of the bitmap image created. The particles were then analysed for their area by calculating their Feret or Caliper diameter and Sauter mean diameter. The parameters were normalized using a bin range of 20

for the index range used for normal distribution curves, particle frequency, particle frequency distribution curves and verification of particle size distribution. The 3D surface plots were modelled via Interactive 3D surface plots plugin and analysis manager in Fiji ImageJ software by optimizing the ratio of f(X, Y) to Z with 100% surface maxima and 0% surface minima. The size of the data grid was adjusted to the grid size of 512 nm sq. area with a perspective angle of 0.2° with optimum smoothing and z-scale enough to distinguish between individual exosomes and noise. The mode of simulation was set such that all the pixels of the image were connected without leaving holes. All the plots were computed by optimizing the parameters and plugin calling from the macro in consol. The graphing and analysis was done using OriginLab Pro 2020 ver. 9.4.9.

SDS - PAGE

Exosome samples were lysed with 1 X RIPA buffer at 70^{o}C for 10 min followed by protein precipitation with ice-cold acetone in -20^{o}C for 2 hours. The proteins were pelleted down in 0^{o}C for 15 min 10,000 rpm. The pellets were air-dried and then digested with 10µl of 1 X Laemmli's buffer for total protein denaturation and loading into PAGE. SDS PAGE under reducing condition was performed with 10 % separating gel having 1.5 M Tris pH 8.8 and 30% Acrylamide and 5 % stacking gel with 1 M Tris pH 6.8 and 30% Acrylamide at 100V. The gels were stained with Coomassie brilliant blue R-250 (Himedia® India) for visualization of protein profile.

Western Blotting

Molecular characterization of exosomes using their transmembrane protein marker was achieved via western blotting as per the ISEV guidelines. Post electrophoresis, proteins were transferred to PVDF membrane in the Bio-Rad blot module with the help of transfer buffer. The antigens were probed with anti-CD63in 1:1000 dilution (Abcam, UK) primary antibody

and rabbit anti-mouse secondary antibody in 1:2000 dilution (Abcam, UK) conjugated with HRP in phosphate buffered saline with added tween 20 (PBST) blocking buffer having 5% non-fat milk. Equal volumes of peroxide solution and luminol enhancer solution (Bio-Rad Clarity^{TM} western ECL substrates, USA) before the visualization. The module was run in the cold room to prevent protein degradation due to heat and the membrane then was blocked by primary antibody overnight in 4^{o}C. Secondary blocking of membrane with secondary antibody was done for 1 hour and post three washes the membrane was taken to ChemiDoc for characterization. The bands on the PVDF membrane (Bio-Rad, USA) were viewed under ChemiDoc XRS+ gel documentation system with varying exposures and further, the immunoblot analysis was done using the ImageJ software ver 1.52s.

Zeta Potential Analysis

The electrokinetic potential of exosomal suspension was measured to check for particle stability in suspension. Samples were diluted in 1:3000 ratio and were loaded into the disposable capillary cells, Malvern Panalytical (DTS1070) zeta potential analyser. The magnitudes were read in triplicates and the average spectra were considered for comparative purpose.

Raman Spectroscopy (Labram HR Evolution - Horiba)

Labram HR Evolution – Horiba Micro Raman Spectroscopy (Kyoto, Japan) was used for the spectral recording of exosomes for label-free molecular characterization using a 532nm laser at 25 percent laser intensity. The molecular fingerprint range and high-frequency range of Raman shift used for the study were 500 cm^{-1} to 1800 cm^{-1} and 2600 cm^{-1} to 3200 cm^{-1}. The sample was made to airdrop on a glass slide and air-dried overnight with a desiccant in a closed chamber. Brain sphingomyelin (SM) and Cholesterol (CHL) purchased from Avanti Polar

Lipids, (Alabaster, AL, USA) were used to acquire reference spectra. The acquisition times used for recording spectra for control and the samples were 10 and 15 respectively.

Fourier – transform infrared spectroscopy

FTIR IR Tracer 100-Shimadzu (Kyoto, Japan) that works on the principle of ATR was used for the spectral recording of exosome samples. The samples were dried to powder form for label-free molecular characterization. The ATR-FTIR spectral recordings were considered from the wavelength range of 500 cm^{-1} to 1800 cm^{-1} and 2600 cm^{-1} to 4000 cm^{-1}, the molecular fingerprint range and high-frequency range. Samples were experimented along with the reference samples of brain sphingomyelin (SM) lipid and Cholesterol (CHL) purchased from Avanti Polar Lipids, (Alabaster, AL, USA) each with acquisition parameters of 50 scans and 4 cm^{-1} resolution. Normalization and multipoint baseline correction of the spectra were done and Happ – Genzel function and Savitzky-Golay algorithm were used for apodization and smoothing of the spectra at the beginning and end of the time-based sampling using the LabSolutions IR.

Data analysis

Multipoint baseline correction, normalization, and smoothing of the Raman and ATR-FTIR spectra were done in Lab Spec 6 and LabSolutions IR respectively with the given parameters. After the data acquisition, the digital numeric values having wavelength and intensity values on X axis and Y axis respectively were plotted simultaneously for all the samples along with the reference spectra. For prediction and determination of the biological components from the spectra based on their molecular vibrations, regression analyses were performed in OriginLab Pro 2020 ver. 9.4.9. for curve fitting and convergence. Chi^2 and R^2 of the curves were optimized by performing 400 iterations of non – linear regression algorithm

for obtaining the statistically fit curves and for performing multivariate analysis further. Further protein to lipid ratios for the three techniques were determined from the integrated area of regions corresponding to protein and lipids in RAMAN and ATR-FTIR spectra. The area under curves were calculated by the integral function calculating mathematical area with limits fit to interpolate the region of curves having baseline mode Y = 0.

Statistical analyses were done for understanding the variation in spectral components of the Raman and ATR-FTIR spectra obtained for the exosomes isolated from two cell lines by three different isolation techniques. The operation was done in Origin Lab Pro 2020 ver. 9.4.9. Levene's test was performed to check for the equality of variances in the spectra of exosome samples from two different cell lines and by three different techniques. Following this One-way ANOVA and Tukey's test were performed between the spectral data of samples to check for the statistical evidence of significantly different population means and to find the variable responsible for significantly different means in population with the actual power of 0.05 significance level.

Multivariate analysis in the form of Principal component analysis (PCA) of the normalized spectra was performed to reduce the dimensions or the variability in the data and determine their principal components or major trends that vary the most from the mixed population where variables being the samples and observation label being the range of wavelength in cm^{-1}. Linear discriminant analysis (LDA) in the form of grouped PCA on principal component scores was used to establish discrimination in the variation of data and verification of the difference in their mean, irrespective of data being within-group variance.

Results

Quantification of the exosome yield and exosome Proteome

The Exosome samples isolated from COLO 205 and MCF 7 were resuspended in 50 µl of NFW and 1 X PBS and stored in -20℃ freezer for further protein analysis. For quantification, the samples were thawed on ice and were used for protein quantification.

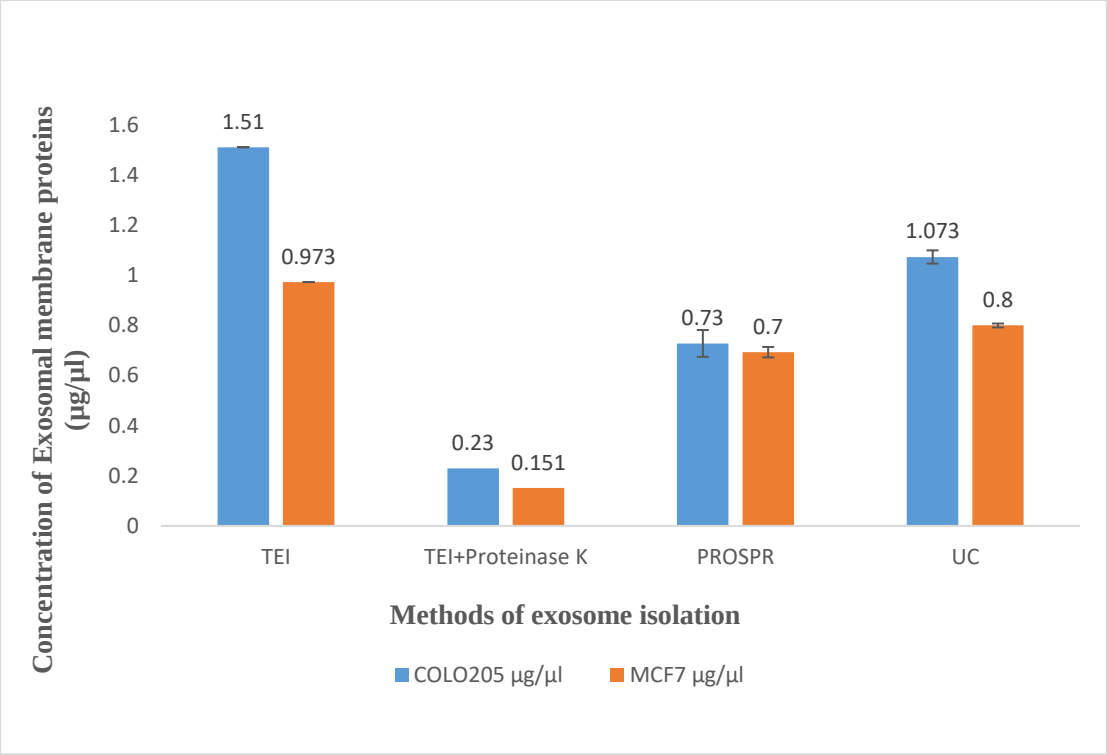


Figure 2. Bar graph representing the concentration of exosome yield per 10^6 cells obtained from COLO 205 and MCF-7 exosomes isolated by the three techniques. Blue and orange bars represent COLO 205 exosomes and MCF-7 exosomes concentration respectively. Resultant values are the average of biological triplicates. Error bars on the bar plots denote the standard error of the mean.

Quantification of exosomes using Bradford's assay is an indirect quantification approach. This protocol uses Bradford's reagent that binds to the proteins, and indirectly indicate the number of exosomes by quantifying the surface proteins and receptors[20].The

concentration of exosomes per 10^6 cells and its total protein content was determined by quantifying the samples using Bradford's assay and reading the absorbance at 595nm. The total protein concentration was determined by lysis of exosomes.

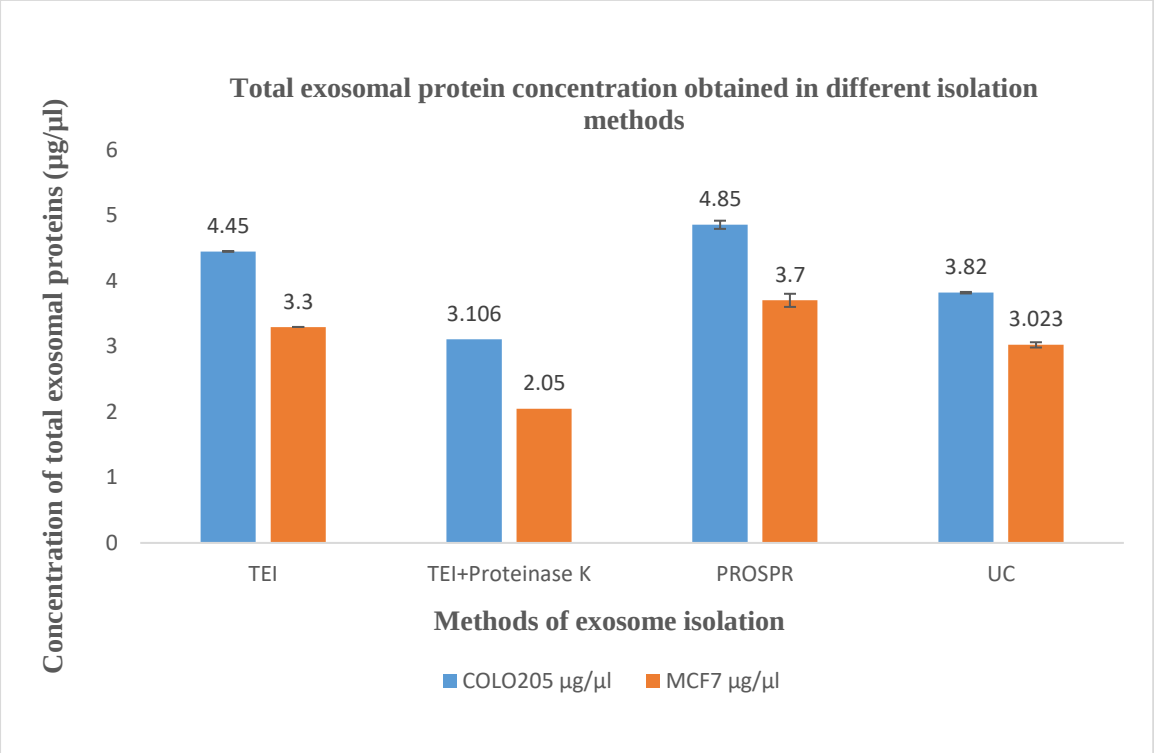


Figure 3. Bar graph representing the concentration of exosomal protein concentration per 10^6 cells obtained from COLO 205 and MCF-7 exosomes isolated by the three techniques. Blue and orange bars represent COLO 205 exosomes and MCF-7 exosome concentration respectively. Resultant values are the average of biological triplicates. Error bars on the bar plots denote standard error of the mean.

Yield from PROSPR method was less in terms of exosome quantification but high in total protein quantification when compared with TEI and UC. The proteinase K treated samples showed drastic reduction in the exosome count as the proteinase K is probably degrading the surface proteins present in the exosome membrane that were being exploited for the quantification via Bradford's assay. Proteinase K treatment will be useful for analyzing the exosome cargo but can't be used for enrichment of purity as many of the purification method rely on surface markers for exosomes. In case of exosome's hydrodynamic diameter, the

treatments like trypsin or proteinase K digestion have also been shown to wear off the electric dipole layer adhering to the surface proteins on membrane due to their digestion[22]. This will affect the morphology of exosomes isolated.

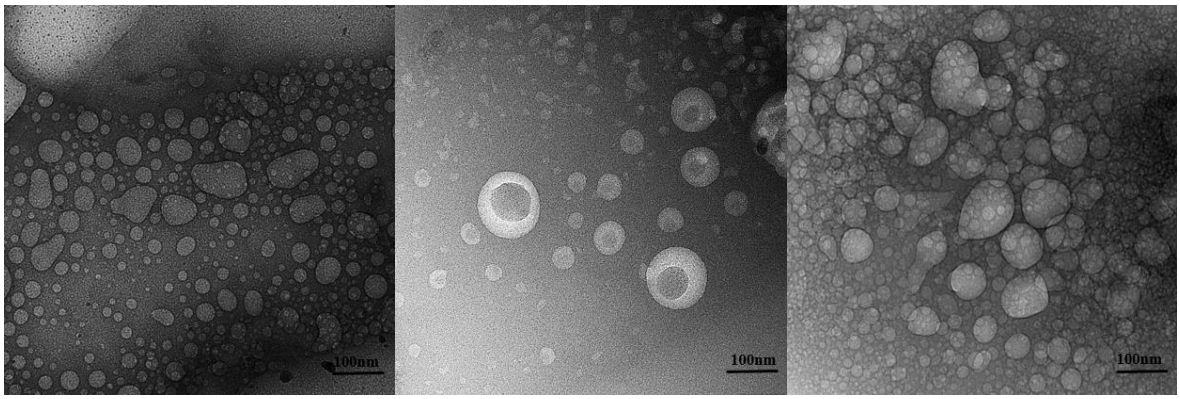
Characterization of Exosomes

High-Resolution Transmission Electron Microscopy

High (HV200kV) Resolution Transmission Electron Microscope images of COLO 205 and MCF-7 exosomes in 20000X magnification at 220 kV were captured for all the three isolation techniques, the TEI, PROSPR, and UC. The images confirmed the typical morphology of exosomes with the lipid bilayer membrane. These images were further used for image analysis to compare the particle size distribution and the morphological attributes of exosomes. For sample preparation, our protocol required no special requirements or preparations for coating and setting of exosomes on a carbon-coated copper grid. Fixation was only required if the grids had to be stored for longer period, but in general, for a quick characterization and particle size distribution and analysis, fixation of particles were not required as observed in the images.

The exosomes isolated from COLO 205 and MCF-7 conditioned media showed different morphology for the type of cells that they were extracted from as well as when different isolation techniques were employed to isolate them. The exosomes isolated from the two cell lines by the three different techniques are presented in figure 4a and 4b.

a.i) a.ii) a.iii)



a.iv) a.v) a.vi)

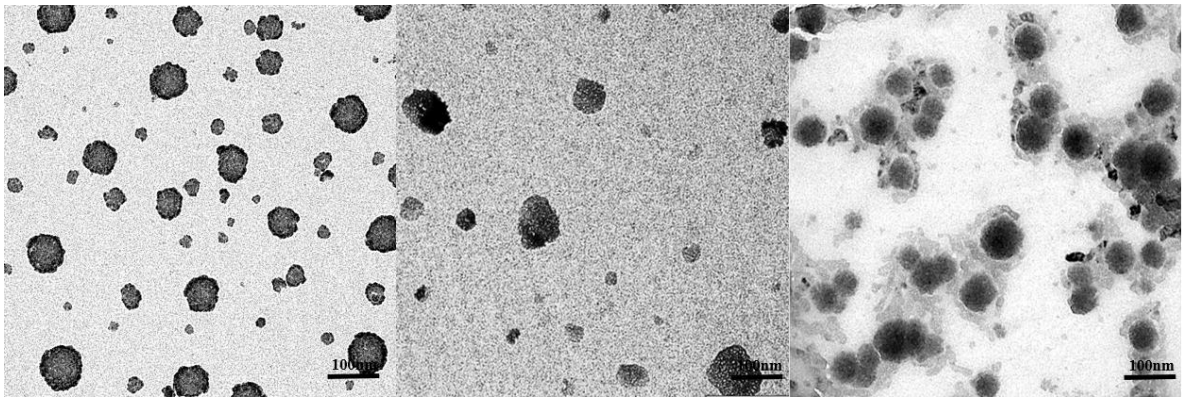
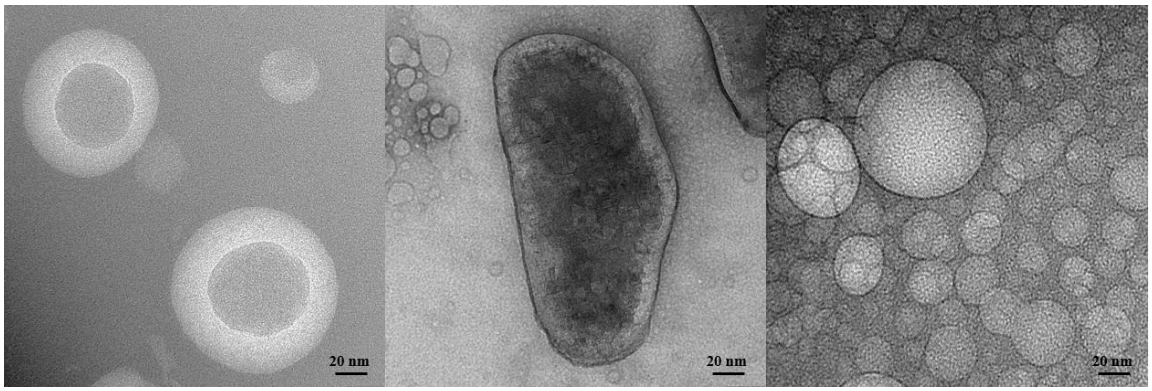


Figure 4a. High-resolution transmission electron microscopy of exosomes isolated from COLO 205 and MCF-7 media. i) COLO 205 exosomes isolated by TEI, ii) COLO 205 exosomes isolated by PROSPR, iii) COLO 205 exosomes isolated by UC, iv) MCF-7 exosomes isolated by TEI, v) MCF-7 exosomes isolated by PROSPR and vi) MCF-7 exosomes isolated by UC respectively. The images presented have a resolution of 100 nm.

b.i) b.ii) b.iii)



b.iv) b.v) b.vi)

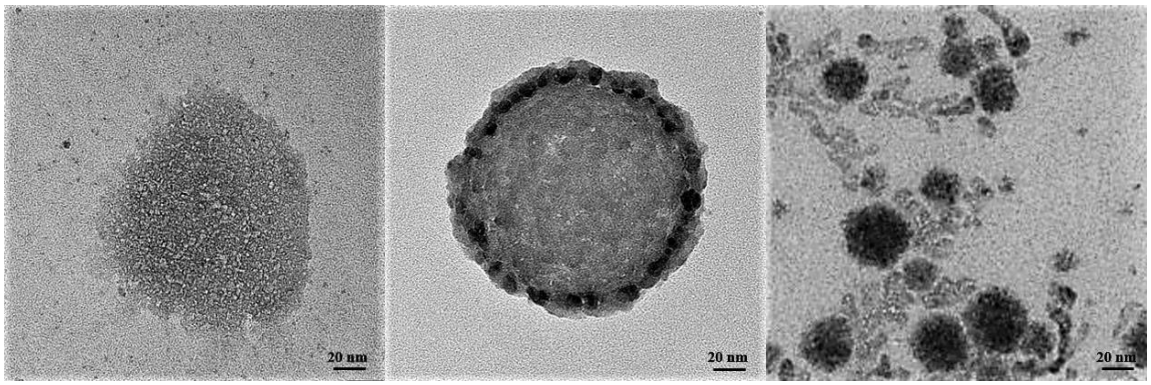


Figure 4b. High-resolution transmission electron microscopy of exosomes isolated from COLO 205 and MCF-7 media. i) COLO 205 exosomes isolated by TEI, ii) COLO 205 exosomes isolated by PROSPR, iii) COLO 205 exosomes isolated by UC, iv) MCF-7 exosomes isolated by TEI, v) MCF-7 exosomes isolated by PROSPR and vi) MCF-7 exosomes isolated by UC respectively. The images presented have a resolution of 20 nm.

The morphology of exosomes isolated from two different cell lines differed. The exosomes were of oval and cup shape and both morphologies were reported in literature[23,24]. This was consistent in exosomes isolated by TEI and UC. But in the case of PROSPR change of morphology was observed. The exosomes isolated from TEI showed better recovery of morphology and well dispersed reflecting towards better stability of particles. The size of exosomes isolated were ranging from 20 nm to around 170 nm when visualized under HRTEM.

The PROSPR technique yielded bulged and cup-shaped exosome from both COLO 205 and MCF-7. In the case of exosomes isolated by UC, the COLO 205 and MCF-7 exosomes both showed particle clustering in various regions along with background noise. This noise was supposedly of aggregated proteins and other small artifacts around these exosomes that might be responsible for the particle aggregation[25]. It was already reported that in urinary exosomes, proteins called Tamm-Horsfall induced exosome aggregation when isolated by ultracentrifugation. These proteins comes down along with exosomes down due to high centrifugal force[26,27]. The exosomes isolated by UC had better morphological recovery but had a narrow size range of 20 to 130 nm. The size range of exosomes isolated from various techniques and their distribution pattern had to be assessed by image analysis of the HRTEM images obtained.

Image Analysis of the HRTEM images obtained for Exosomes

HRTEM images provided accurate information about the size and morphology of the particles. Calculating the Sauter mean diameter or the surface - volume mean diameter and Feret's diameter or caliper diameter of exosomes based on their absolute morphology provided the accurate information of their sizes in sample holding an advantage over Dynamic Light Scattering[28,29,30]. When the dispersed exosomes get distributed in the liquid medium, the electric dipole layer adheres to its surface protein. DLS measures the size or diameter of this layer over the particle and henceforth it is measuring the hydrodynamic particle diameter which will be theoretically bigger than the exosomes imaged in HRTEM[31]. Size distribution analysis of exosomes by image analysis of HRTEM images was not biased and considered the actual shape of particles whereas in DLS the algorithm determining the exosome size calculates it through translational diffusion coefficient using the stoked Einstein equation that considers the particle to be spherical or rotationally symmetric in shape[32,33].

Plotting 2D contour map for Particle Size Distribution Analysis.

In order to calculate the size of exosomes, the Feret's diameter of each particle was calculated from the analyse particle option in ImageJ and also by extracting the contour line data from image matrix of 2D perimeter contour maps plotted for bitmap images of exosomes (Supplementary Figure 1). The data was obtained from uniformly scaled and calibrated images. The range of their surface volume mean diameter was normalized by the process of binning for the normal distribution curves in FIJI ImageJ.

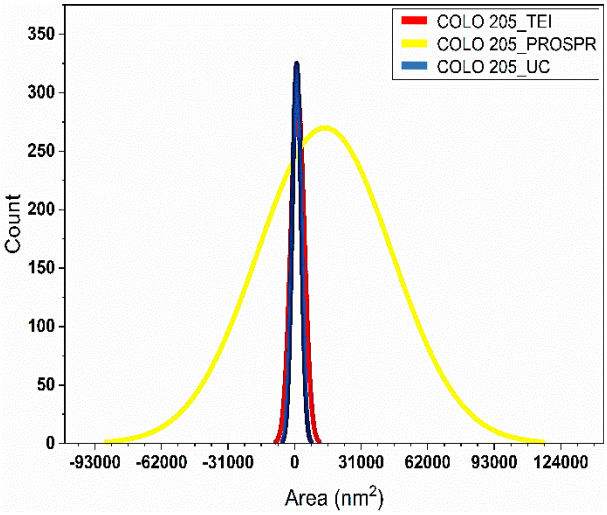
Normal Distribution curves, Particle size frequency distribution curves and comparative box plot for particle size distribution

Normal distribution curves for area and volume and size were plotted for both COLO 205 and MCF-7 exosomes based on these data (figure 5). Area distribution curves indicate the number of exosomes and the area occupied by the exosomes per nm^2. Similarly, volume distribution curves indicate number of exosomes and the volumetric dimensions of these particles per nm^3. Particle size frequency curves depict the frequency of the occurrence of these exosomes with respect to the mean size of these particles. Based on this data, particle size frequency curves (figure 6) and their frequency distribution curves (figure 7) were plotted for distribution of the exosomes based on their frequency of the occurrence with respect to their mean sizes. Comparative box (figure 8) plots were plotted based on the quartile deviations in the data containing particle count and size. The mean sizes of exosomes and the frequency of outliers in the specified window of 20 – 200 nm size range were shown. This analysis could be extrapolated towards the particle size, hydrodynamic radius of exosomes and distribution of exosomes.

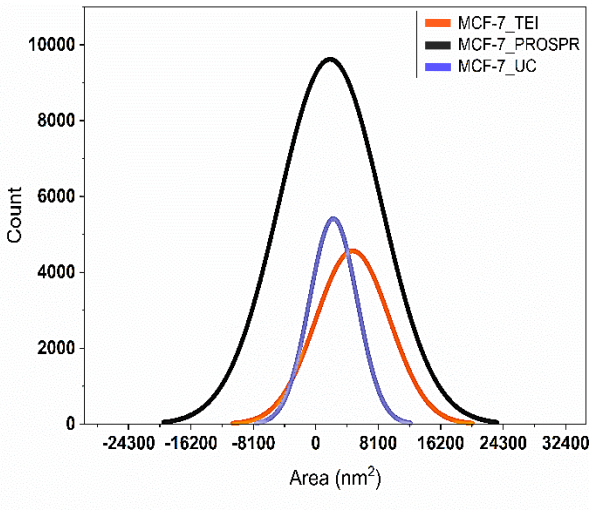
Normal distribution curves were plotted for the range of exosome sizes isolated using three different techniques. Area under these curves were correlated to the percentage of

populations within the indicated range as described in supplementary figure 2 & 3. The exosomes with size around 110 nm diameter occupy 68.26% area of the population (which is the range of data that falls within the first standard deviation of mean). The area occupying 99.74% of the population (3rd SD) had the exosomes covered with size around 150 nm as the stipulated size range. The exosomes isolated from COLO 205 and MCF-7 by TEI reagent were broadly distributed around the mean sizes in both the area and the volumetric distribution curves. These curves lie between the stipulated range of parametric area the particles occupied and the volumetric dimensions of these particles in the range of 20-200nm size. In case of PROSPR, the bell curves for exosomes isolated via this method were broadest and the distribution of their parametric area and volumetric dimensions spread outside the stipulated range. Exosomes isolated by UC had the narrowest of the bell curves among the three isolation methods with restricted distribution of the parameters to a smaller mean size. The surface and volumetric density curves had the highest number of exosomes of 50-80nm size range for all the three techniques. But the bell curves were broadest in TEI and steepest in case of UC showing that the extracted exosomes were of broader size range in TEI and narrower in the case of UC. The PROSPR method had highest outliers whereas UC had distribution with narrow upper-class limit.

a)



d)



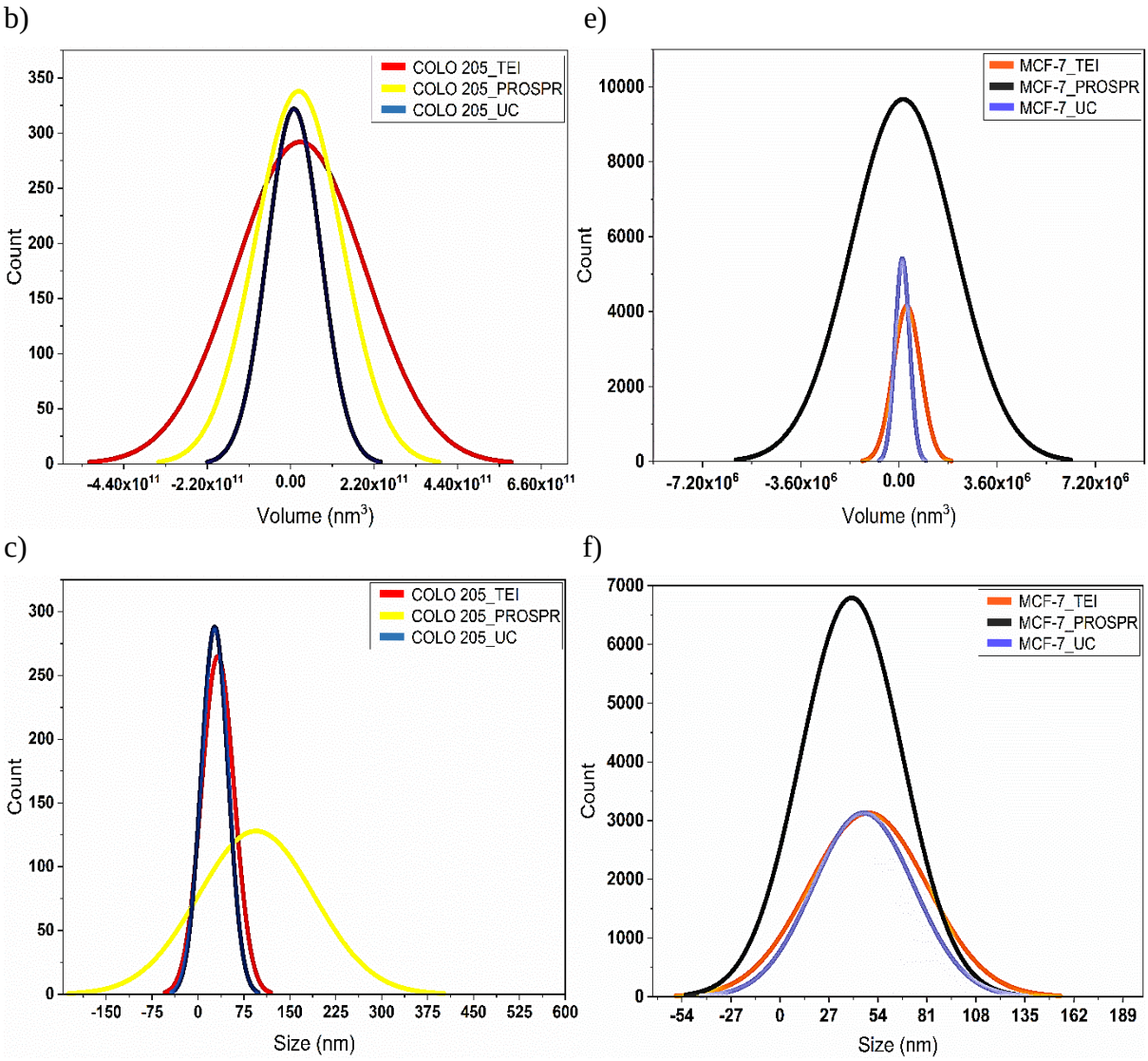


Figure 5. Normal Distribution curves of exosomes based on their area, volume and size as captured in HRTEM images. a-Area, b-Volumetric and c-Size distribution curves of COLO 205 exosomes. d-Area, e-Volumetric and f-Size distribution curves of MCF-7 exosomes.

Particle size frequency depicted the frequency of occurrences of the exosomes with respect to their mean size. The exosomes with mean size of 30 nm to 80 nm observed in high frequency in both COLO 205 and MCF-7 exosome samples with a mean size of 50 nm. The trend was common in all the three isolation techniques. Particle size frequency curves followed the similar trend as that of the normal distribution curves. With respect to the frequency of their

occurrence based on their mean size, the samples of TEI had their highest occurring frequency in the range of 40-60nm. UC had the steepest curve with exosomes highly frequent around 30nm, whereas PROSPR had their highest particle occurrence in the range of 50-70nm. PROSPR also had outliers in the range of 300-600nm.

a) b)

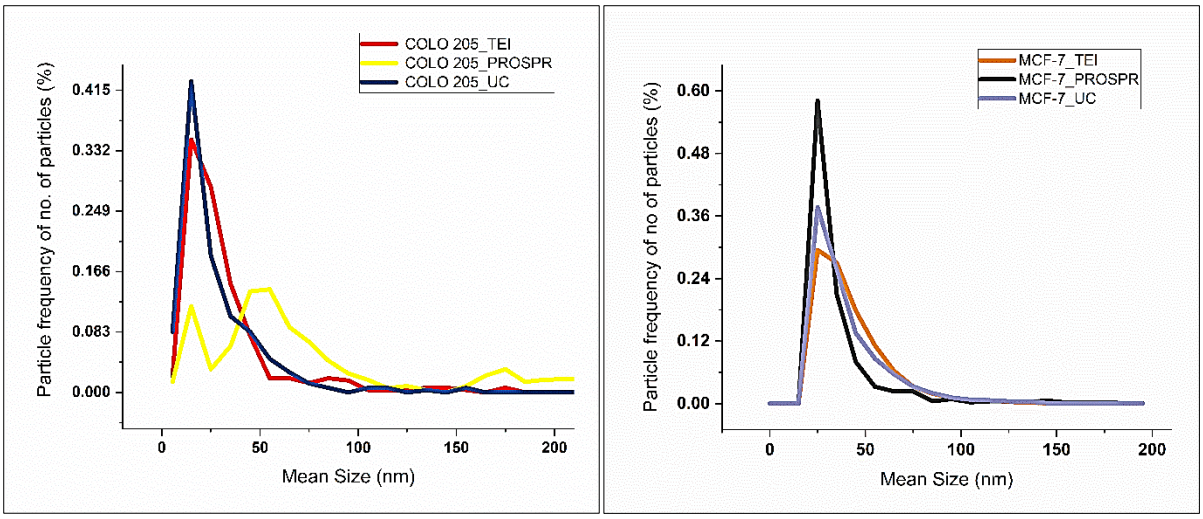


Figure 6. Frequency distribution curves of exosomes based on their size. a) Particle frequency curves of COLO 205 exosomes isolated by TEI, PROSPR, and UC for no. of particles based on their mean size. b) Particle frequency curves of MCF-7 exosomes isolated by TEI, PROSPR, and UC for no. of particles based on their mean size.

Larger size exosomes above 150 nm were lesser in count compared to the smaller ones in case of TEI and UC but since these bigger particles occupied more volumetric space in their 3D environment the particle frequency distribution of these particles were high compared to the smaller ones. In case of COLO 205 exosomes isolated by TEI had these bigger particles being distributed around the mean size of 160-180nm but in MCF-7 the distribution was balanced for all the size range of exosomes. In case of exosomes isolated from PROSPR the bigger particles occupied the space from the range of 140 nm onwards having a gradual increment in both the exosome types. The UC had bigger exosomes distributed along the range

of 130-160 nm but due to larger volume occupied by smaller particles the curve is balanced throughout more in case of MCF-7 than COLO 205.

a) b)

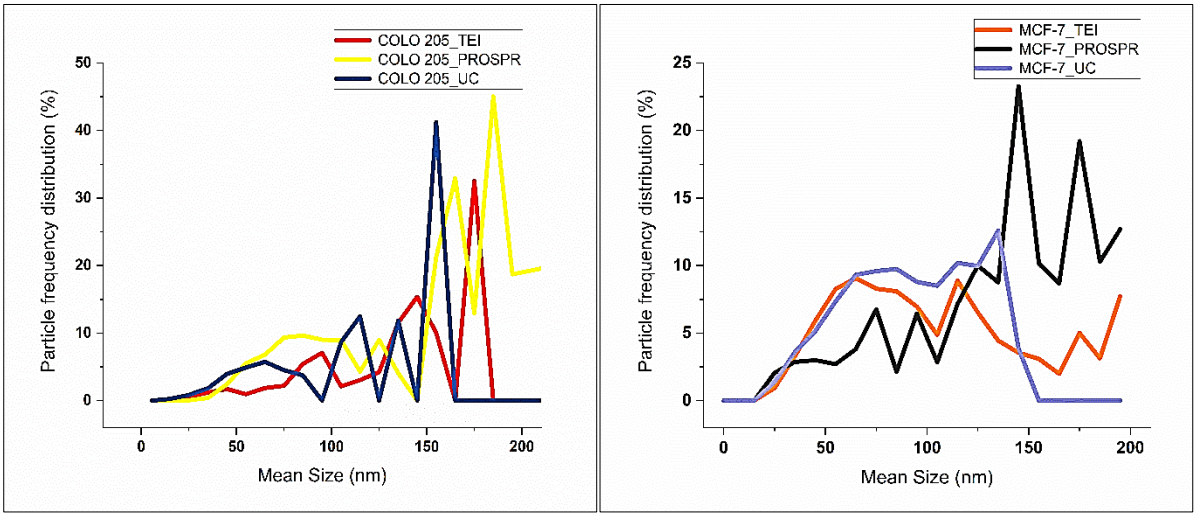
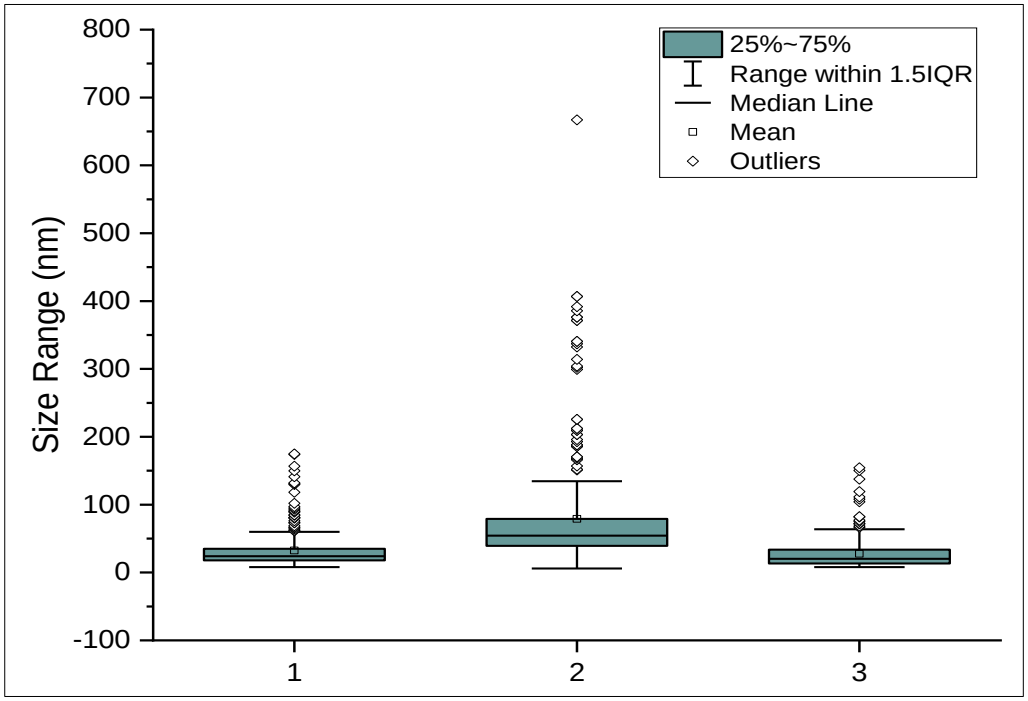


Figure 7. Frequency distribution curves of exosomes based on their size. a) Particle frequency distribution curves of COLO 205 exosomes isolated by TEI, PROSPR and UC for no. of particles based on their mean size. b) Particle frequency distribution curves of MCF-7 exosomes isolated by TEI, PROSPR and UC for no. of particles based on their mean size.

The Comparative box plots were plotted based on the quartile deviations in the size distribution data obtained from the HRTEM images (quartile parameters information available in supplementary tables S1 & S2). These box plots inferred about the particle size distribution based on their mean size, lower class limit, upper class limit and outliers with respect to the specific dataset. The class limits were defined by first quartile and third quartile in the data sets having size limit of exosomes with highest occurrence. The other size variables with low occurrence were defined by the error bars and the data points with sizes that do not lie within the given parameters lies outside the range as outliers. The exosome isolated by UC had the smallest mean size of 30nm among the three isolation techniques in both the exosome types with their outliers reaching up to exosomes with 170nm. The PROSPPR samples had the average mean size in the range of 40-60nm in both COLO 205 and MCF-7 exosome types with

outliers reaching up to 600nm and above. In case of TEI the exosomes had the mean size of 40-50nm in both COLO 205 and MCF-7 exosome samples with outliers ranging up to 190nm.

a)



b)

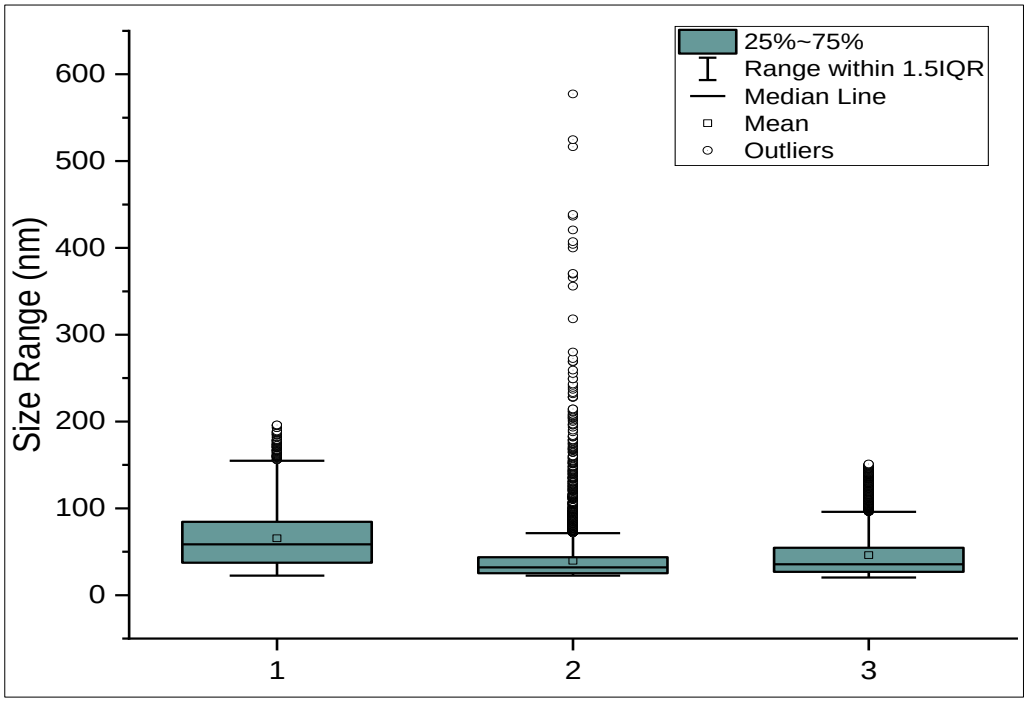


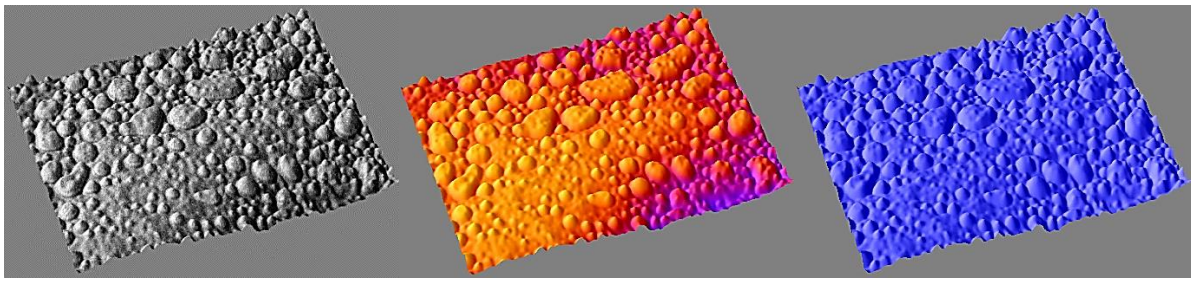
Figure 8. Comparative box plots of particle size distribution for exosomes isolated from COLO 205 and MCF-7 by TEI, PROSPR and UC techniques. The box plots were plotted based on the quartile deviation in data. The box plots comprised of the mean size and 25-75% data within 1.5 interquartile range. Outliers in the data lied outside the 1.5 IQR that helped determine the size of exosomes frequently occurring and the ones significantly differing from other data points. a) Comparative box plots of COLO 205 exosomes and b) comparative box plots of MCF-7 exosomes. The numbers 1,2 and 3 labelled on x – axis in figure 11. a) and b) represent TEI, PROSPR and UC respectively.

The normal distribution curves, particle size frequency distribution curves and comparative box plots were very useful in assessing the particle size and their distribution based on the HRTEM images obtained for these exosomes isolated from COLO 205 and MCF-7 conditioned media by the three isolation techniques the TEI, the PROSPR and the UC. It helped us assess that the exosomes isolated from the TEI had better size distribution range with the stipulated range ranging from 30nm to 190nm compared to the PROSPR and the UC techniques. Exosomes isolation by the PROSPR yielded large number of particles but had broader distribution of their size range exceeding the class limits.

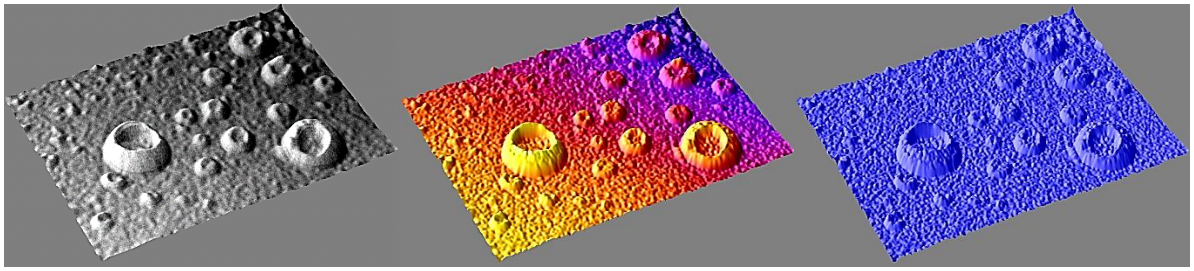
Interactive surface 3D plots for Morphology Study

The optimised parameters in the interactive environment of analysis manager of FIJI ImageJ were used to plot high resolution morphology surface plots, thermal surface plots and distribution surface plots (figure 9a and b). The images and analysis of the morphology interpret differential morphology of exosomes when isolated via the three different techniques and two different cell sources.

a.i)



a.ii)



a.iii)

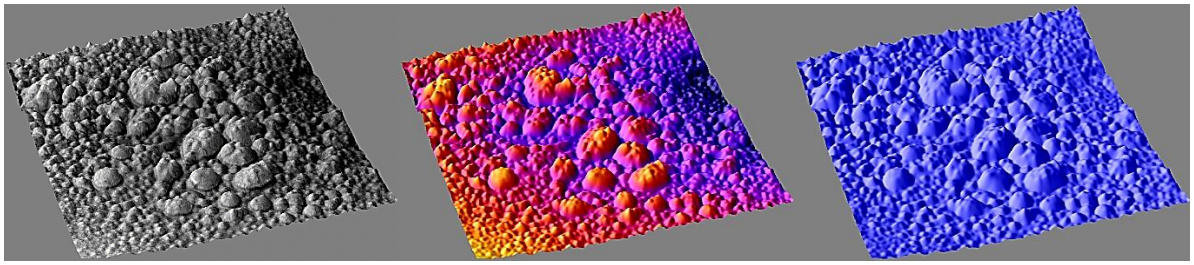
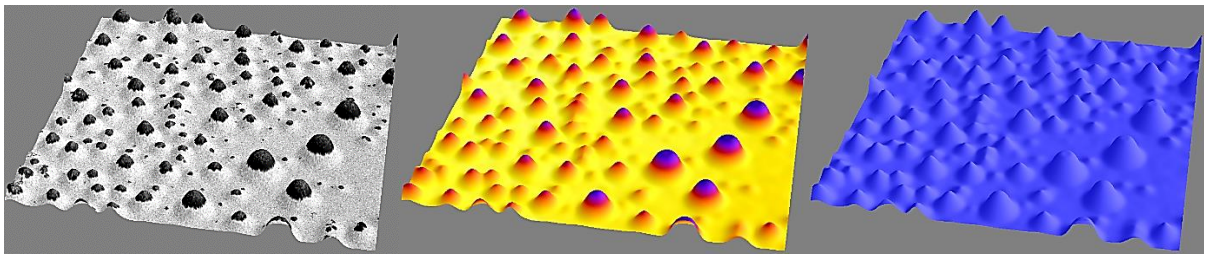
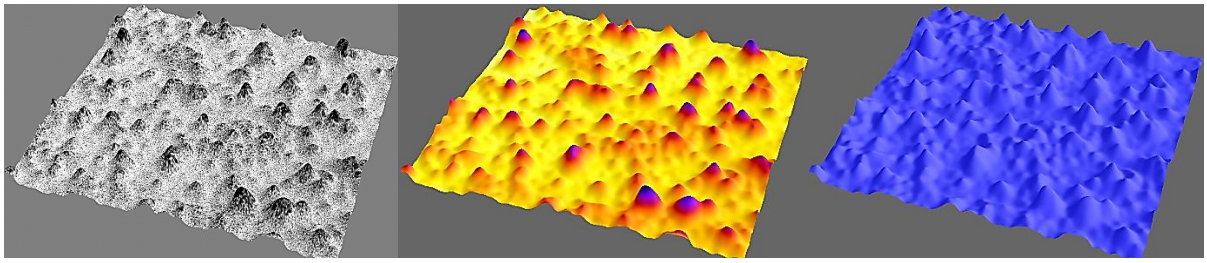


Figure 9a. 3D high resolution interactive surface plots plotted for morphology study of COLO 205 exosomes. i) Topology map plotted for the study of exosome morphology. ii) Thermal map plotted for distinguishing between exosomes and noise. iii) Distribution map plotted for studying the dispersion and distribution pattern of exosomes.

b.i)



b.ii)



b.iii)

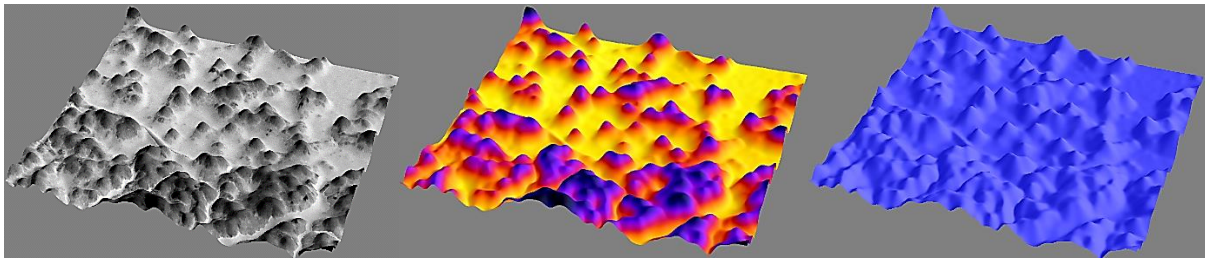


Figure 9b. 3D high resolution interactive surface plots plotted for morphology study of MCF-7 exosomes. i) Topology map plotted for the study of exosome morphology. ii) Thermal map plotted for distinguishing between exosomes and noise. iii) Distribution map plotted for studying the dispersion and distribution pattern of exosomes.

Morphology surface plots or topology plots presented the 3D morphology of exosomes and their physical attributes. When isolated by TEI reagent, exosomes had better recovery of morphology as compared to both PROSPR and UC techniques. The precipitation-based isolation method incorporating the reagent sustained the integrity of the globular membrane resulting in both oval or globular exosomes depicting that exosomes in proper hydrated state

are more of oval-shaped than cup-shaped vesicles[33,23]. Exosomes isolated via PROSPR had a differential morphology than TEI and UC based method. In case of isolation by PROSPR method, COLO 205 yielded cup-shaped exosomes with swollen or bulged up membranes whereas MCF-7 yielded mixture of oval and cup shaped grainy and swollen and not much globular shaped exosomes and other bigger microvesicles. The exosomes isolated by UC from both the conditioned media had better recovery of morphology showing the globular characteristics of exosomes but were not uniformly dispersed.

Thermal plots depicted the difference between true exosome vesicles and background noise for size distribution analysis. Thermographs with thermal signature of exosomes rising to peaks were considered true exosomes and were used for the calculation and statistics. True exosomes were readily differentiated from the background noise from the images of exosomes in case of when isolated by TEI followed by PROSPR and UC.

Distribution plots depicted the pattern of dispersion and distribution of exosomes in suspension. This map depicted whether the particles were segregated and dispersed properly or were clustered together in the form of agglomeration. It also gave insight into the pattern of distribution of small and big exosomes. The distribution map for TEI showed proper segregation between particles and uniformly spread throughout. The PROSPR samples had a mixture of bigger and smaller particles that were sparely distributed and spread randomly whereas the UC samples had better distribution compared to PROSPR but the particles were closely clustered and aggregated into groups co-localized into various region. This suggested that the exosomes tend to aggregate when isolated by UC due to very high centrifugal force along with small aggregating proteins and artefacts. Two different cellular sources, on the other hand, showed differential morphology as expected, suggesting that even exosomes have varied morphology based on the origin of cell or tissue.

Zeta Potential analysis

The comparative bar graph of Zeta potential (figure 10) depicted the electrokinetic potential of exosomes isolated by TEI, PROSPR, and UC. The analysis was done to check for the stability behaviour of exosomes in suspension. The stability behaviour range of resultant zeta potential values (mV) ranges from incipient stability to good stability. Exosomes isolated by TEI had zeta potential values above -50 mV resulting in good stability of particles with very less agglomeration. The PROSPR and UC processed samples had zeta potentials in the range -10 to -40 mV resulting in moderate and incipient stability of particles. The UC samples had the lowest zeta potential values than others that suggested exosomes in suspension might be facing aggregation or coagulation. The individual zeta potential graphs can be found in supplementary data (Supplementary Figure S4).

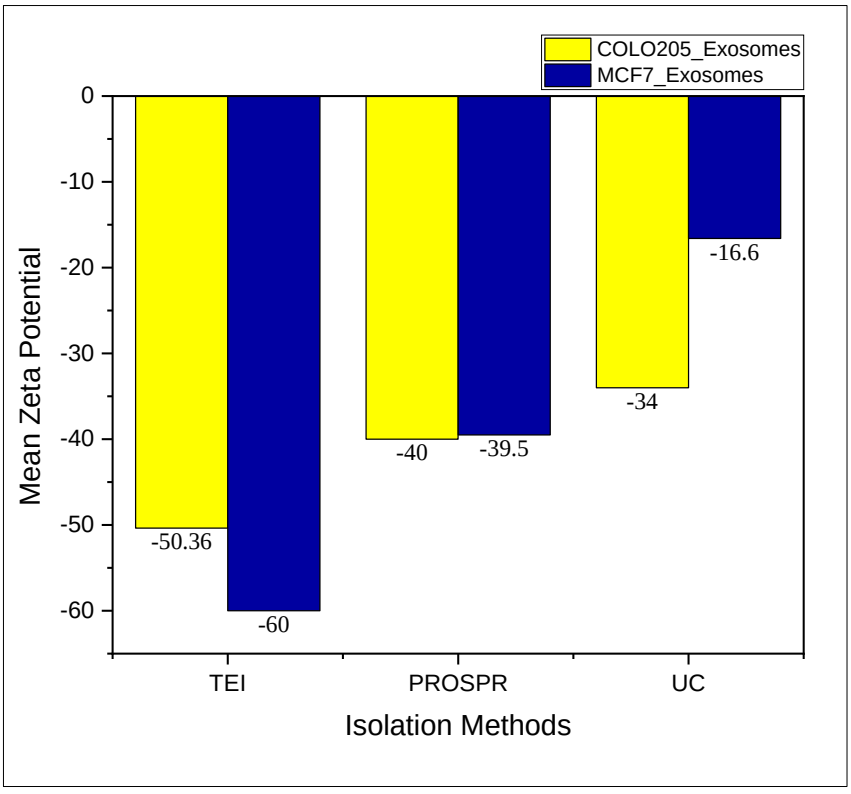


Figure 10. Comparative zeta potential analysis: The stability of exosomes in suspension was analysed by their electrokinetic potential. The yellow bars represent COLO 205 exosomes and blue bars represent MCF-7 exosomes.

SDS PAGE of total protein in exosomes and western blotting for immuno-characterization.

Exosome samples having protein concentration of ~ 150 µg to 200 µg in 100 µl of suspension were lysed with RIPA buffer and then precipitated with acetone by incubating in -20ºC for 2 hours. Following the protein precipitation at 10,000 rpm for 15 min at 0ºC, the protein pellet was resuspended in precisely 10-15µl of 1X SDS PAGE (figure 11). Many peptides that were visible in TEI and UC based isolation were absent in PROSPR based isolation method. TEI and UC based isolation methods yielded peptides in the range of 10 kDa to 250 kDa where TEI method had better protein recovery than others as evident from the intense bands observed in gel. In case of PROSPR, only 55kDa to 70kDa peptides were predominantly visible which makes it the method with least protein recovery. The results stay the same for both the cell line sources.

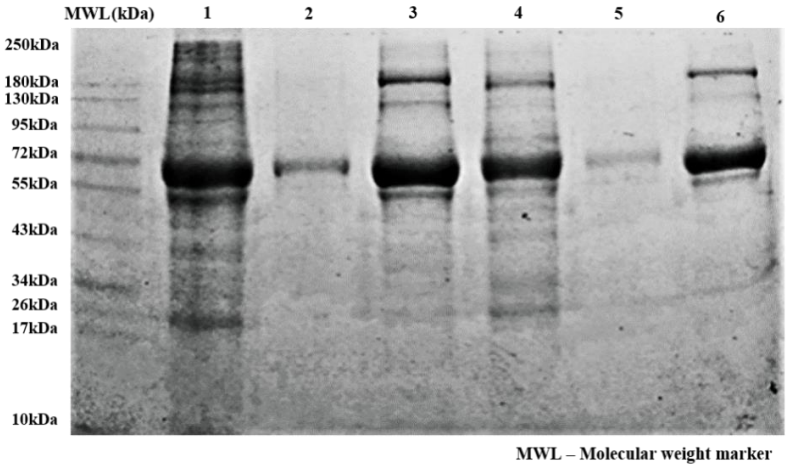


Figure 11. SDS PAGE for the exosome proteome profile. The 10% separating gel was used to run the sample in gel based on their charge and mass. 1st lane was the molecular marker with

the range of 10-250 kDa, 2nd lane onwards samples loaded were COLO 205 TEI, PROSPR, UC and MCF-7 TEI, PROSPR, UC samples consecutively.

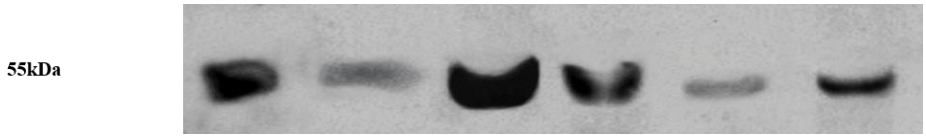


Figure 12. Western Blotting of exosomes for molecular characterization with CD63 marker. The samples blotted were COLO 205 UC, COLO 205 PROSPR, COLO 205 TEI and MCF-7 UC, MCF-7 PROSPR, MCF-7 TEI consecutively from the first lane.

Western blotting experiment for immuno-characterization of exosomes was performed comparing the CD63 recovery and expression in exosomes isolated by the three methods (figure 12). The intensities of the blots were in decreasing order w.r.t to the exosomes isolated from the three techniques. The order of intensities was TEI > UC > PROSPR showing highest recovery of CD63 in TEI versus UC and least in PROSPR.

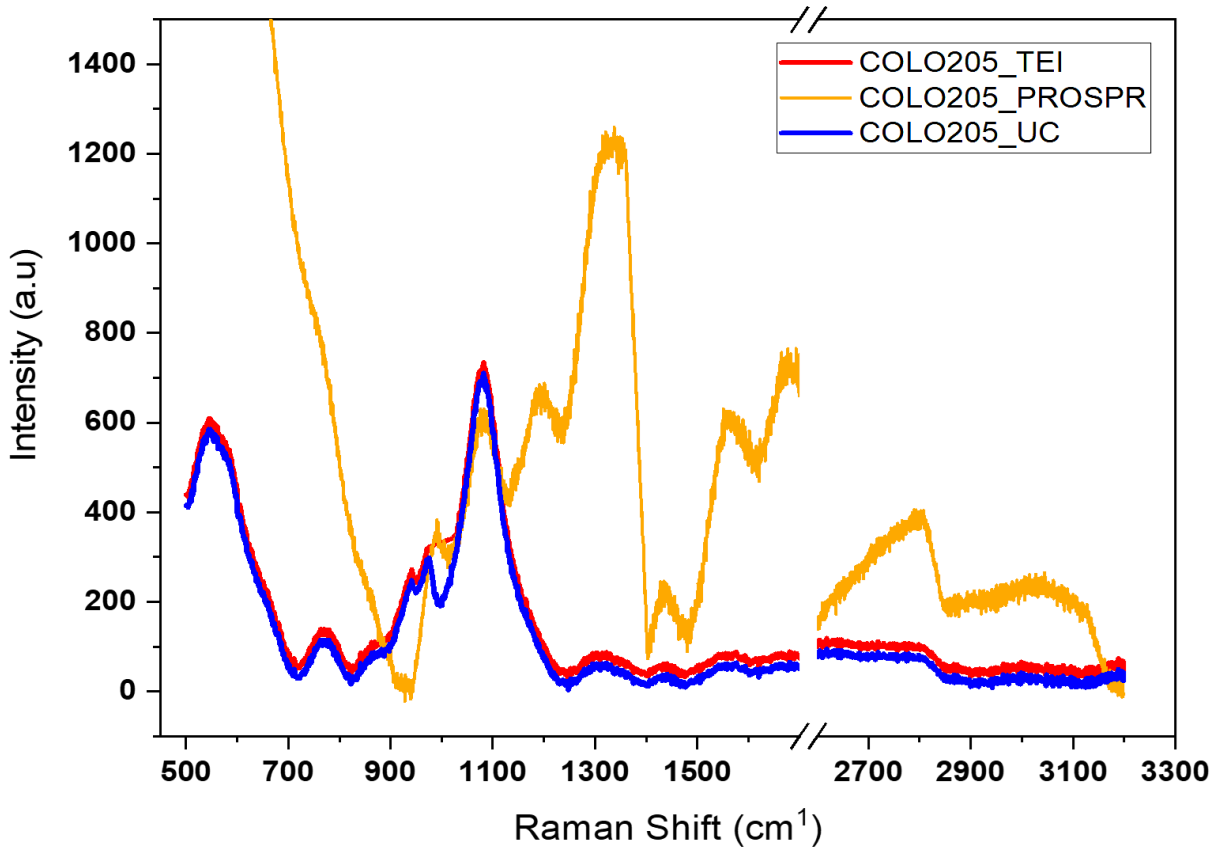
Molecular and analytical characterization of exosomes by label free fingerprinting with micro Raman Spectroscopy and ATR-FTIR

RAMAN and infrared spectroscopic approaches provide the advantage of studying various molecular constituents of exosomes through the vibrational conformation of their atoms and molecules present in structural backbone or functional groups[34,35]. These approaches were used to study the molecular dynamic characteristics of exosomes, that required very less sample (<1mM) for the analysis. ATR-FTIR spectroscopy provided much accurate spectra but couldn't provide information on macromolecular components that RAMAN spectroscopy provided. ATR-FTIR provided spectra without the problem of light scattering and background noise or fluorescence. In case of ATR-FTIR, the absorption of water molecules was subtracted from the background[36,37].

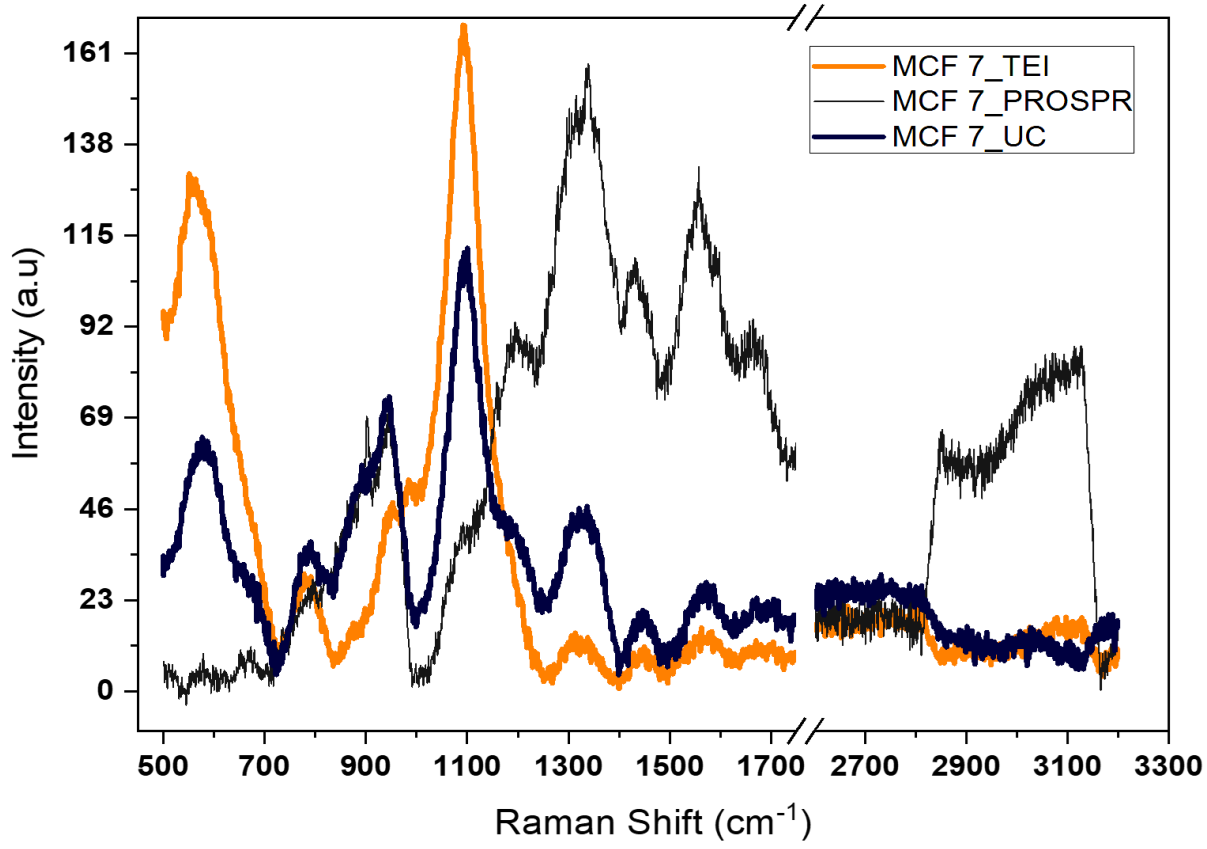
After obtaining the Raman and ATR-FTIR spectra of exosomes isolated from COLO 205 and MCF-7 conditioned media isolated by TEI, PROSPR, and UC methods, the spectra were subjected to analysis and statistics for validation. Analyzing the data and determining the biomolecules required spectral processing that consisted of 400 iterations of nonlinear regression algorithm which were performed using the Levenberg Marquardt algorithm with Lorentz function for nonlinear multiple peak fit (Supplementary Figure 6). Regression analysis was performed for the better fit and convergence of curves. Lorentz function was selected over Gaussian because of its ability to generate bell-shaped wider tails suitable for predicting the molecular components of exosomes through their atomic vibrations using the spectral data. The COD or R^2 was optimized to 0.9-1.0 and chi sq. was reduced as much as possible for statistical significance.

TEI spectra and their corresponding peaks were most accurate compared to PROSPR and UC in determining the biomolecules. The RAMAN and FTIR spectra for exosomes isolated from COLO 205 and MCF-7 by TEI, PROSPR, and UC are presented in figure 13.

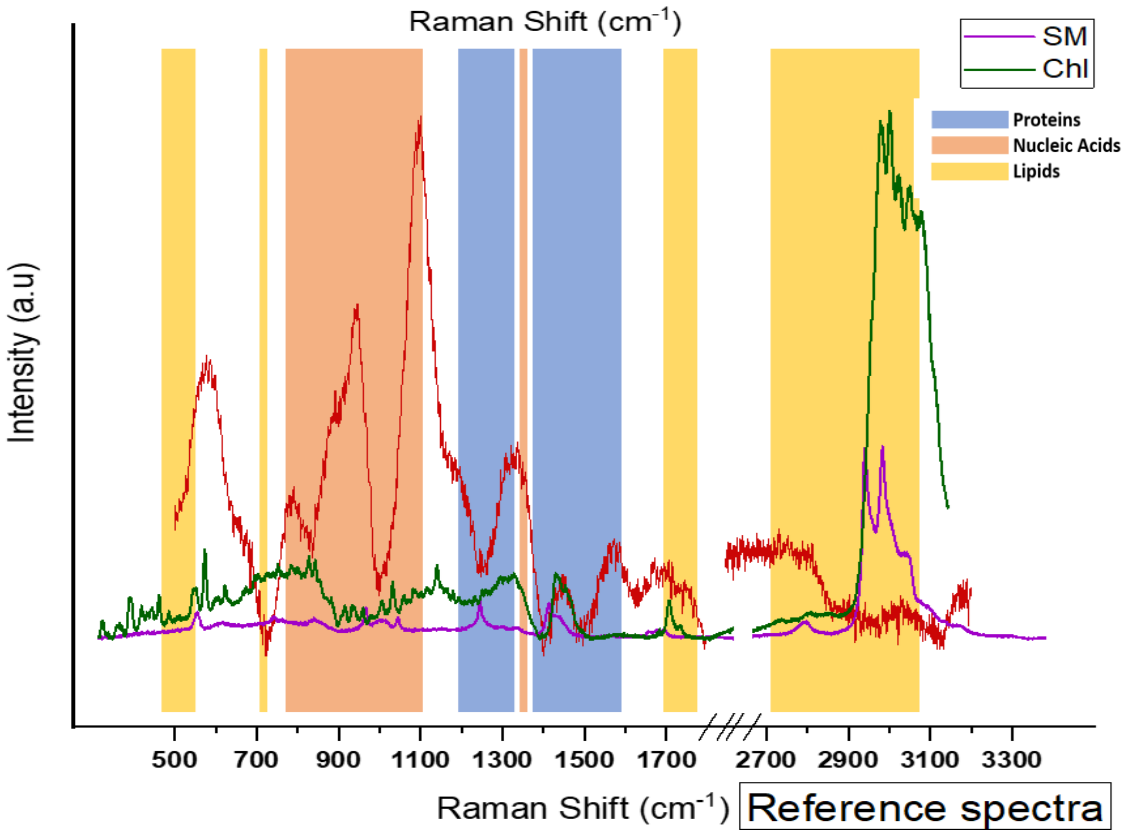
a)



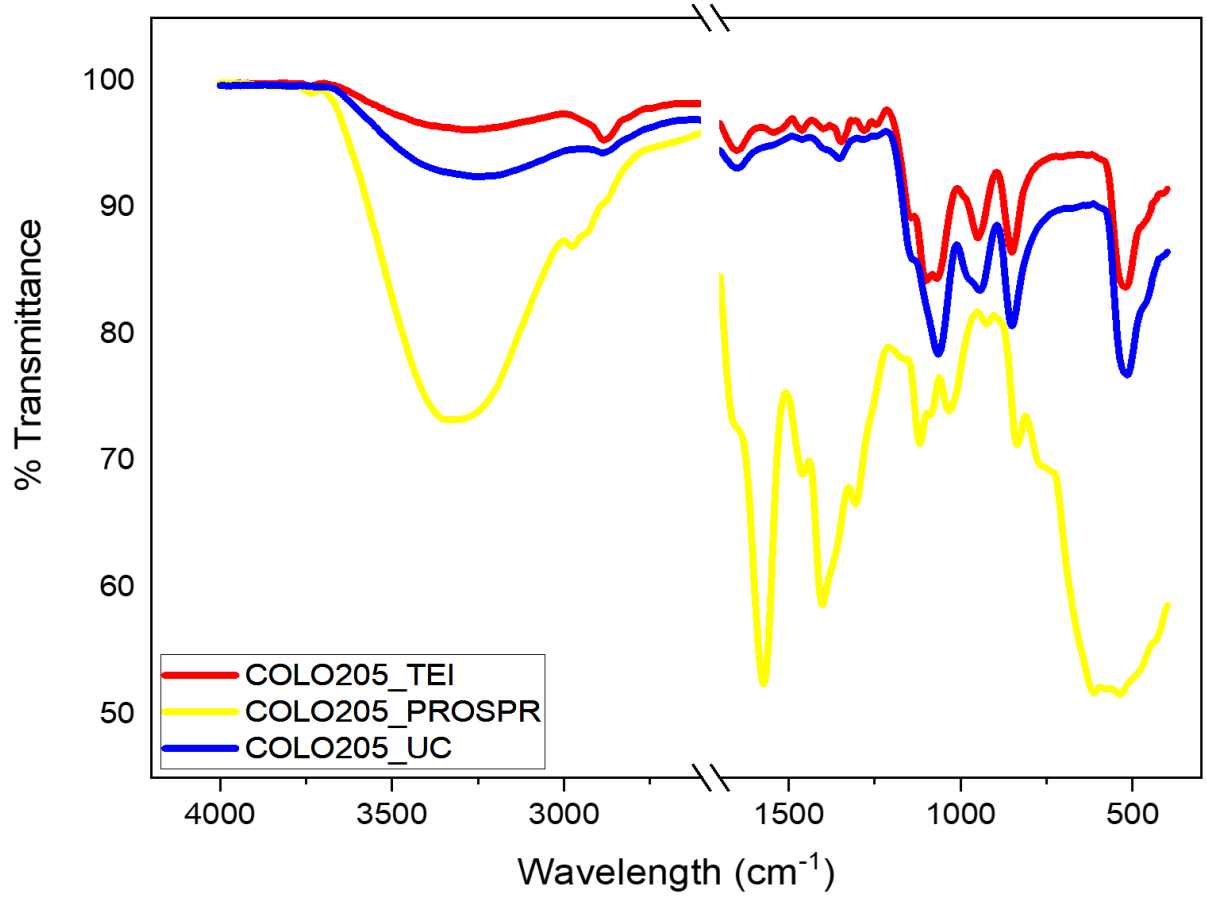
b)



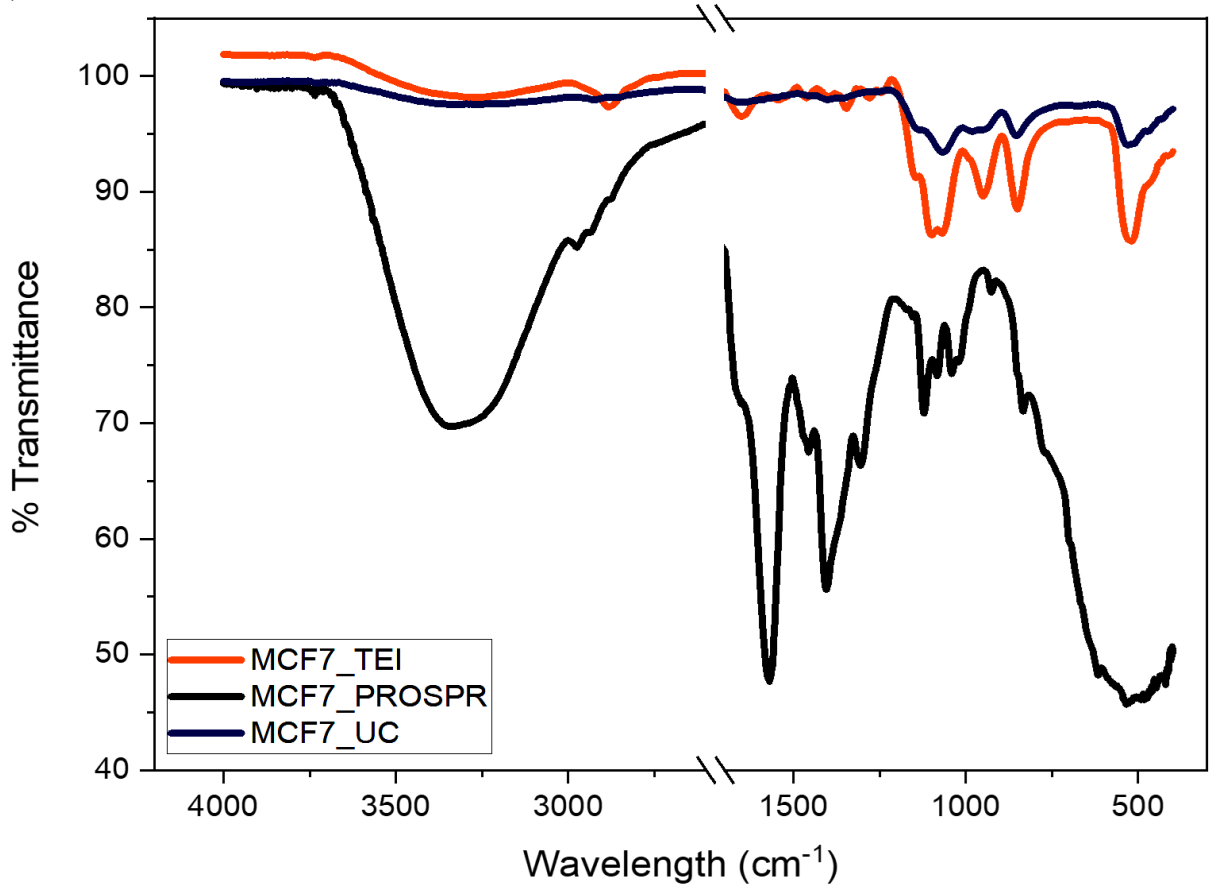
c)



d)



e)



f)

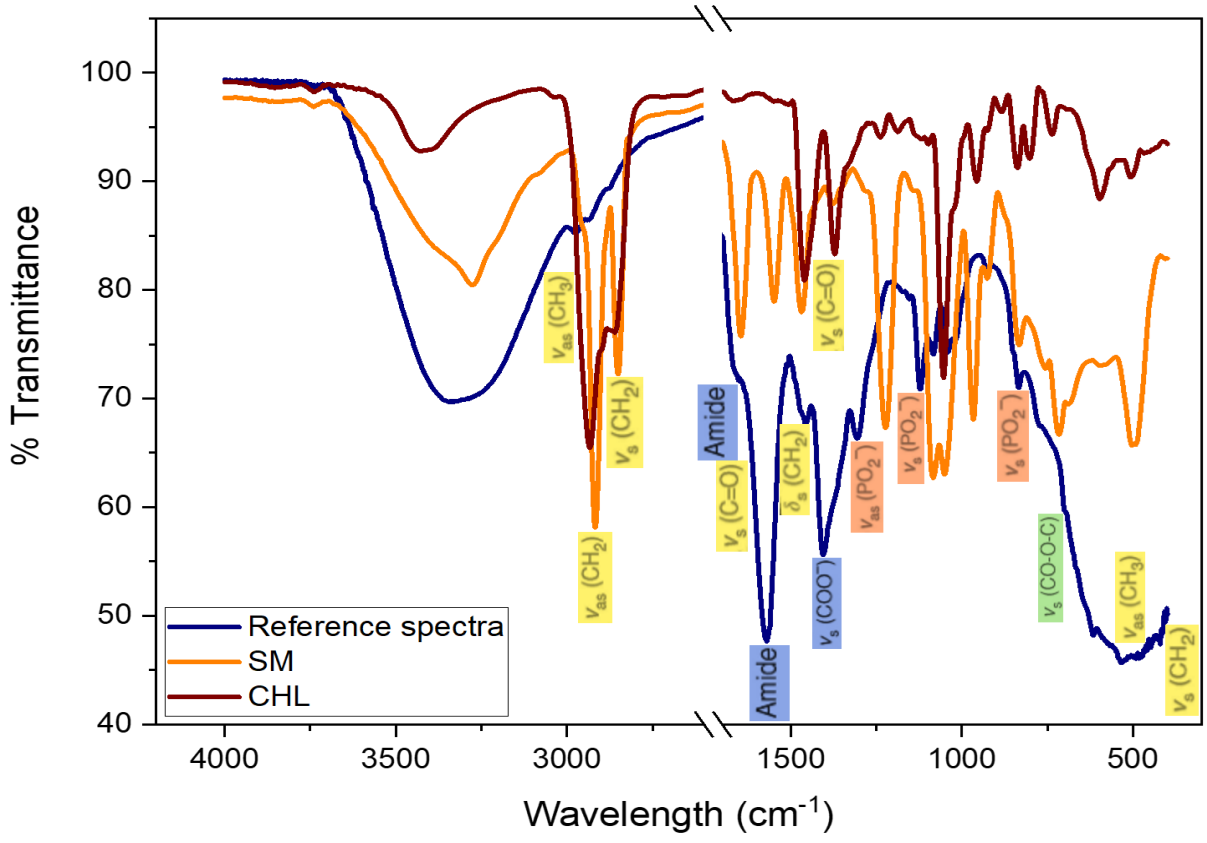


Figure 13. a) RAMAN spectral presentation of COLO 205 exosomes isolated by the three techniques, TEI (green), PROSPR (violet), and UC (orange). b) RAMAN spectral presentation of MCF-7 exosomes isolated by the three techniques, TEI (black), PROSPR (red) and UC (blue). c) Reference spectra for RAMAN spectroscopy depicting the potential regions of biomolecules representing the cargo and structural compositions of exosomes. labelling of the regions with blue represent protein species, orange regions represent the nucleic acids species and yellow regions represent the lipids species in exosomes. d) ATR-FTIR spectral presentation of COLO 205 exosomes isolated by the three techniques, TEI (green), PROSPR (violet), and UC (orange). e) ATR-FTIR spectral presentation of MCF-7 exosomes isolated by the three techniques, TEI (black), PROSPR (red) and UC (blue). f) Reference spectra for ATR-FTIR spectroscopy depicting the potential regions of biomolecules representing the cargo and structural compositions of exosomes. Molecular representation in yellow labels represent species of lipids with symmetric and asymmetric vibrations of CH_2 and CH_3, green labels represent species of carbohydrates with vibrations of carboxyl group, orange labels represent species of nucleic acids with vibrations of phosphate group and blue labels represent species of proteins with vibration of amides in exosomes.

Prediction and determination of various biomolecules and molecular components based on laser excitation using 532nm laser in RAMAN spectroscopy

Major Raman shifts for the vibration of molecules and their atoms representing the fingerprinting region range from 500 cm^{-1} to 1800 cm^{-1}.The specific molecular signature of exosome samples in this region recorded were at 520cm^{-1} for Phosphatidyinositol, 524cm^{-1} for Phosphatidylserine, 537 $cm^{-1}$for Cholesterol ester, 540 $cm^{-1}$for Glucose-saccharide band, 596 $cm^{-1}$for Phosphatidylinositol, 701 cm^{-1} –703 $cm^{-1}$for Cholesterol ester, 720 cm^{-1} –820 $cm^{-1}$for Nucleic acids, 752 cm^{-1} –760 cm^{-1} for Tryptophan, 995 $cm^{-1}$for C-O band of ribose, 1003 cm^{-1}

for Phenylalanine, 1048 $cm^{-1}$for Glycogen, 1054 $cm^{-1}$for C-O and C-N stretching of proteins,1060cm^{-1} –1095 $cm^{-1}$for C-C vibrations of lipids and carbohydrates, 1120 $cm^{-1}$for C-O band of ribose, 1127 $cm^{-1}$for C-N stretching of proteins; ceramides, 1200 cm^{-1} –1300 cm^{-1} for Amides, 1230 cm^{-1} – 1240 $cm^{-1}$for for ß-sheets, 1260 cm^{-1} –1300 cm^{-1} for α-helix, 1298 $cm^{-1}$for Fatty acids, 1336 cm^{-1} for CH_3CH_2 wagging or oscillating mode of polynucleotide chain specially purine bases, 1337 $cm^{-1}$for Tryptophan, also reported for Glycine skeleton and Proline side chain, major peaks of Guanine at 1357 cm^{-1} and 1361 cm^{-1}, 1360 $cm^{-1}$for Tryptophan, 1420 cm^{-1} –1480 $cm^{-1}$for the vibrations of CH functional groups of nucleic acids, proteins and lipids, 1555 cm^{-1} –1558 $cm^{-1}$for Tryptophan, and 1716–1740 $cm^{-1}$for C=O group stretching.

Some biological molecules like lipids have their CH_2 and CH_3 symmetric and asymmetric vibrations and stretching at high frequency region consisting of wavelength or Raman shift range from 2600 cm^{-1} to 3200 cm^{-1}. These correspond to 2853–2881 cm^{-1} for CH_2 symmetric and asymmetric stretches of lipids, 2910 $cm^{-1}$for CH_3 stretching vibrations, 2940 $cm^{-1}$for CH and CH_2 stretching vibrations in lipids and proteins. The important aspect of using Raman spectroscopy was to establish a label-free characterization method for exosomes. By using this technique major biological molecules were determined in one go also giving specific molecular signature or fingerprint of each component based on the Raman shift or the wavelength at which these individual molecules scatter light[38,39].

Determination of biomolecules present in exosomes by using the signature of their major functional groups in ATR-FTIR spectra

In ATR-FTIR spectra the wavenumbers ranging from 400 cm^{-1} to 1800 cm^{-1} constituting the fingerprinting region and 2600 cm^{-1} to 4000 cm^{-1} constituting the high frequency region were analysed and assigned to specific functional groups that were responsible for infrared

absorption corresponding to biomolecules present in exosomes. The spectral readings get recorded for vibrations of symmetric stretching (Vs), asymmetric stretching (Vas) and with change of angle as in the case of bending (δs) of the bonds present in functional groups. Vs (CO-O-C) at 988 cm^{-1} represented the carbohydrate molecules for example glycans in glycoprotein cargo in exosomes. Vs and Vas of PO_2^- in 800 cm^{-1} to 1300 cm^{-1} region corresponded to the phosphate group present in phosphodiester bonds building up the chain of nucleic acids, as functional groups in modified triglycerides like phospholipids. The amide groups corresponding to amide III (Vs COO^-) at 1314cm^{-1}, amide II at 1544 cm^{-1} and amide I at 1656 cm^{-1} represented the amide linkages or peptide bonds in proteins. Secondary structure of proteins corresponding to α-helix and β-sheets gave rise to peaks at 1651 – 1655 cm^{-1}, 1620 – 1640 cm^{-1} and 1671 – 1695 cm^{-1} respectively. Vs and Vas of CH_2 and CH_3 functional groups corresponding to lipids and proteins both gave peaks at 1394cm^{-1}. Regions from 950cm^{-1} to 1210cm^{-1} gave rise to peaks that represented the vibrations of C-O-C ester groups that are present in phospholipids, cholesterol esters, glycerides and even in the phosphodiester bonds forming nucleic acids. The region from 1800 – 2600 cm^{-1} was omitted due to non-specific noise and finally, the high frequency region gave intensities at 2700 cm^{-1} to 3010 cm^{-1} for the Vs and Vas of CH_2 and CH_3 groups generally present in lipids. The sample spectra were subtracted from the background solvent spectra for obtaining accurate signatures. ATR-FTIR proved to be a powerful and easy to use technique characterizing exosomes and analysing variation in data from the two exosome types isolated by three different methods. By using this technique major biomolecules present in exosomes were determined and characterized with the help of the functional groups of these biomolecules and their percent transmittance at specific spectral region corresponding to wavenumber[39,40,41]. The important factor was the intensity of spectral components in PROSPR when compared to TEI and UC in the region analysed between 1200 - 1700 cm^{-1}.

Based on the regions observed in previous literatures[42] related to the vibration of amides, amino acid side chains and secondary structure in proteins PROSPR differed a lot from the other two techniques. After the deconvolution of curves by multiple peak fit the 1599 cm^{-1} that corresponds to additional band components of amino acids, 1622 cm^{-1} that corresponds for aggregated proteins or apolipoproteins and the 1600 – 1700 cm^{-1} that corresponds to the region of secondary structures in proteins showed interesting patterns. Due to mixture of components there couldn't be any specific inference on secondary structure of proteins but at 1622 cm^{-1} deconvolution of the region resulted in wider peaks for UC and PROSPR but a flattened or levelled curve for TEI when observed in the x-axis scaled from 1400 – 1800 cm^{-1}.

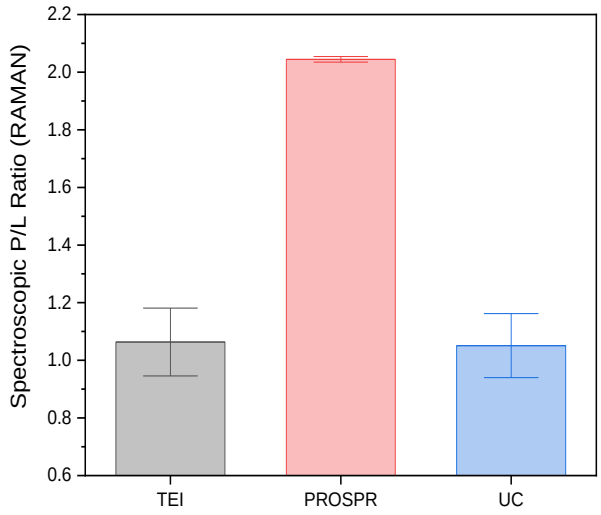
Spectroscopic protein to lipid ratio

To calculate the spectroscopic protein to lipid ratio, the regions of interest (ROI) selected were 1400 – 1800 cm^{-1} (prominently constituting the regions of amide I and II along with the information on secondary structures of proteins and their conformations)[43] and 2700 – 3040 cm^{-1} (constituting the regions of lipid-related CH_2/CH_3 stretching vibrations)[41] for both RAMAN and ATR-FTIR spectra. For proteins, integrated areas of 800 – 1300 cm^{-1} were neglected as this region with COO^- and PO_2^- and other amide stretching vibrations not just contributed to proteins but also to glycoproteins, phospholipids and nucleic acids. The integrated areas were calculated following which the ratios or relative intensities were determined by the division of these integrated areas of proteins and lipids of individual spectral recordings (supplementary table S3).

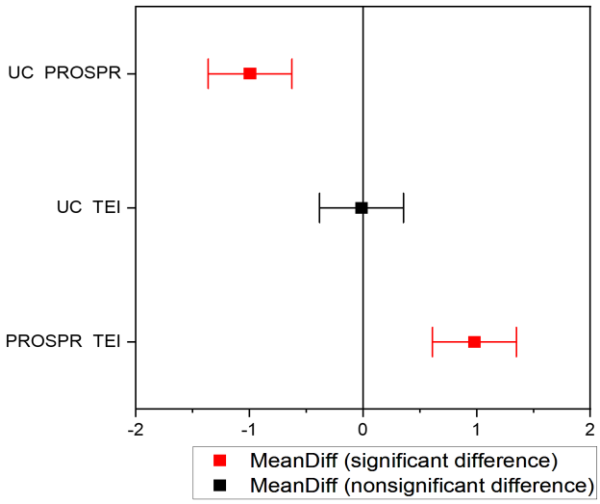
The relative intensities were plotted on to bar plots for both COLO 205 and MCF-7 exosomes isolated by TEI, PROSPR and UC (figure 14) where the order of the ratios was PROSPR > TEI > UC. Spectroscopic protein to lipid ratio was relatively close for TEI and UC (also reflecting upon the exosomal recovery, the relative intensity of TEI was higher than UC

which correlated with the total protein estimation and exosomal yield) but the ratio had a high increment in PROSPR that directly correlated the higher relative amount of protein to lipid and total protein estimation with the presence of other extracellular vesicle or microvesicles in sample. The observation was validated with one-way ANOVA (having significantly different population means at 0.05 levels) and comparison of means with Tukey's test (significance equalled 0 for the combination of TEI – UC indicating that the difference of means were not significant at 0.05 level but for the combinations of UC – PROSPR and TEI – PROSPR the significance equalled 1 indicating that the difference of means was significantly different at 0.05 level.

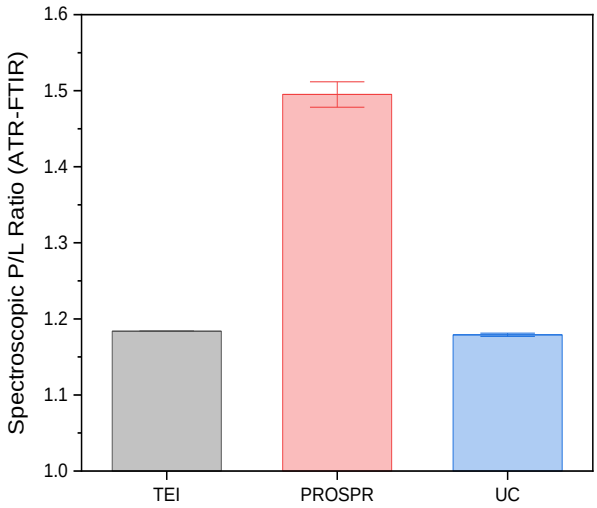
a.i)



a.ii)



b.i)



b.ii)

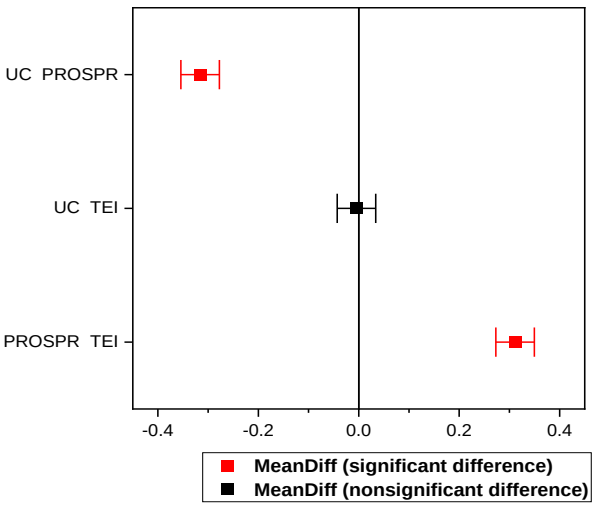


Figure 14. Representation of relative intensities for spectroscopic protein to lipid ratio determined from a.i) RAMAN spectra and b.i) ATR-FTIR spectra by division of integral areas (IA) calculated for proteins and lipids (IA (1400 – 1800 cm^{-1}) / IA (2700 – 3040 cm^{-1})) of COLO 205 and MCF-7 exosomes isolated by TEI, PROSPR, and UC. The data was statistically analysed and validated with one-way ANOVA and by comparison of means of the relative intensities of a.ii) spectroscopic P/L ratio of RAMAN spectra and b.ii) spectroscopic P/L ratio of ATR-FTIR spectra

Validation of data with statistical analysis

For statistical validation and discrimination between the exosomes isolated from two cell sources and isolation techniques, the spectral data were subjected to statistical analysis. Descriptive statistics for mean and standard deviation was calculated along with standard error of mean. Homogeneity of variance test (Levene's test) was performed to check for the presence of deviation in the equality of variances for the variables calculated for the spectral data of exosomes isolated from two cellular sources by three different isolation techniques. One-way ANOVA was then performed to check for the presence of significant difference in between means of the population. Null Hypothesis for the analysis considered that the means of all levels were equal and alternative hypothesis considered the means of one or more levels of spectral data were different. Tukey's test was performed upon the validated parameters of ANOVA to check for the significance of difference in means (figure 15). The comparison of means in data was done in combinatorial of TEI - UC, UC - PROSPR and TEI – PROSPR technique of exosome isolation in both the cell sources (Statistical analysis tabular columns available in supplementary tables S4, S5, S6 and S7).

Principal component analysis (PCA) is a statistical tool that was very helpful in analyzing spectroscopic data from RAMAN having large number of variables. The method

reduced the variability in data to new set of variables called principal components accounting for the majority of variability in data. Grouped PCAs were performed for differentiating the exosome from two cell lines (PCA loadings) and variance (PCA scores) in isolation techniques based on 95% confidence ellipses that were grouped as defined by user into data representing each type of isolation technique calculating best discriminating components defining the exosome types and isolation techniques (figure 16). It was used to analyse the variance in isolation techniques, also discriminating their vectors in the form of loadings from each other thereby making each technique unique. PCA was performed individually for both the exosome types and their isolation techniques along with a comparative 3D PCA to analyse the variance in data from both the sample types and all three techniques (figure 17). Scree plots of component selection from eigenvalues and their variance available in supplementary figures S7-12.

Statistical analysis for spectral data obtained from both the analytical techniques showed similar trend in results for COLO 205 and MCF-7 exosomes isolated by three different isolation techniques. The results of descriptive statistics reflected towards the presence of large amount of variation in the spectral data. Samples isolated from TEI and UC had similar but smaller standard deviations from mean as compared to the PROSPR samples, this meant that the data set of TEI and UC were closer to the mean than PROSPR data sets that were farther away. Levene's test proved that the population variances were significantly different at 0.05 significance level having the power as 1 rejecting the null hypothesis. The one-way ANOVA performed showed population means to be significantly different which validated the acceptance for alternate hypothesis at 0.05 levels. Tukey's test was performed to check for the component responsible for highest variance in data by comparing the means of techniques in combinatorial within upper and lower control limits. The spectral data of both COLO 205 and MCF-7 obtained from RAMAN and ATR-FTIR spectroscopy when subjected to Tukey's test

resulted in TEI and UC methods in combination to have difference of means non-significant at 0.05 levels but the other combinations of TEI – PROSPR and UC – PROSPR had difference of means significant at 0.05 levels.

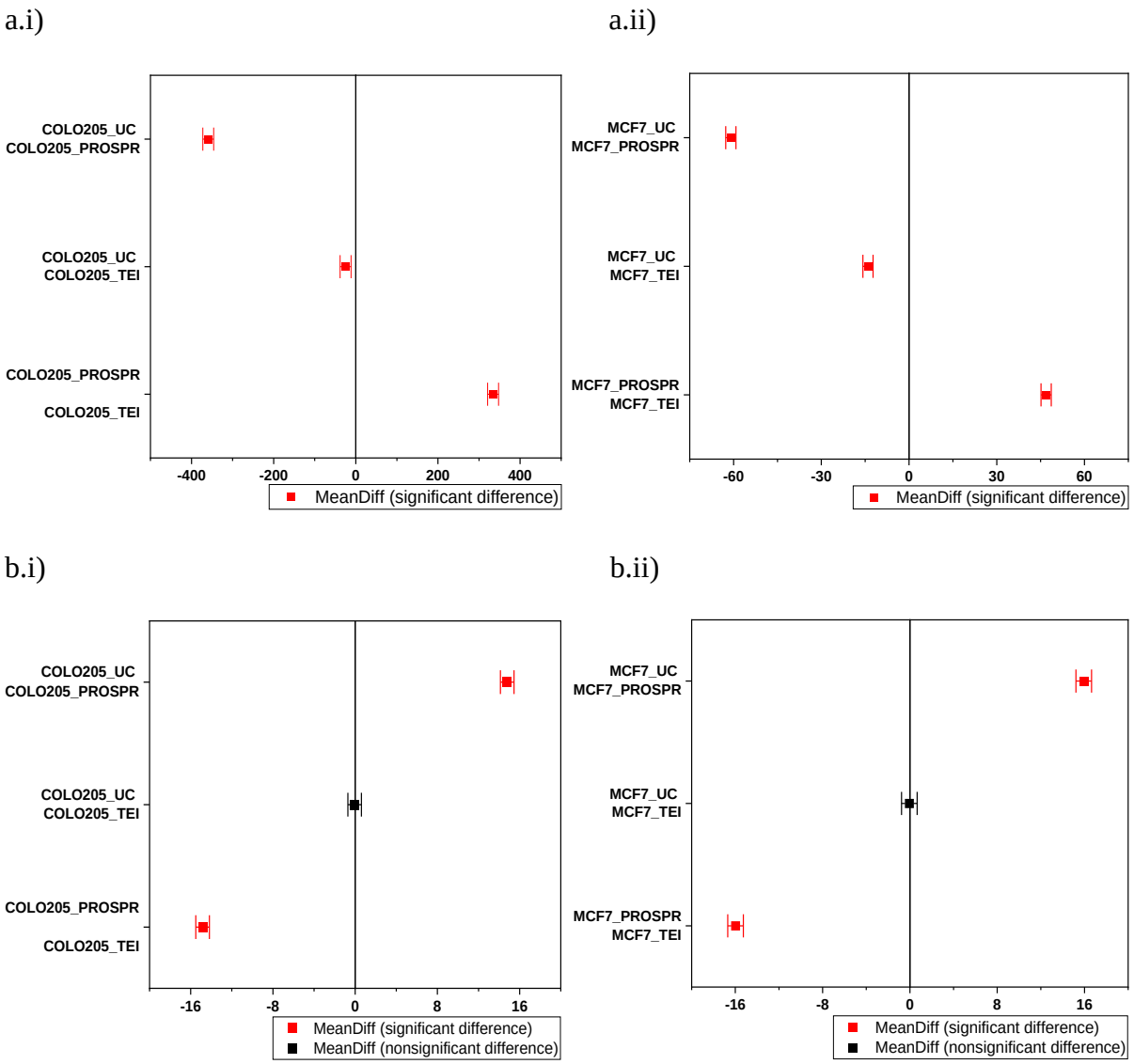


Figure 15. Mean comparison plot for Tukey's test performed on the spectral data of a) RAMAN spectroscopy i) COLO 205 exosomes ii) MCF-7 exosomes and b) ATR-FTIR done for i) COLO 205 exosomes ii) MCF-7 exosomes. Mean comparison by Tukey's test when had significance equals 1, it indicated that the difference of the means was significant at 0.05 level tested. If the

significance equalled 0, it indicated that the difference of the means was not significant at 0.05 level tested.

Multivariate analysis using Principal component analysis was performed on the spectral data obtained from RAMAN and ATR-FTIR spectroscopy for COLO 205 and MCF-7 exosomes isolated by TEI, PROSPR and UC. Grouped PCAs were plotted (figure 16) with two principal components PC1 and PC2 having 57.5% and 42.4% variance in RAMAN spectral data and 99.84% and 0.15% variance in ATR-FTIR spectral data. The plots discriminated the TEI, UC, and PROSPR samples based on component scores and COLO 205 and MCF-7 exosomes based on the vectors of loadings. The vectors were not completely out of phase by 90^{o} but were separated with a greater phase angle that indeed majorly discriminated the exosomes isolated from two cell types statistically while at the same time showed some correlation due to the same factor between the exosome types pointing out towards the species origin and structural morphology including other attributes. The component scores discriminated between the TEI, UC, and PROSPR variables that showed tight clustering and correlation between TEI and UC with overlapped confidence ellipses but in case of PROSPR the confidence ellipse occupied more space incorporating more variables out of which only a small fraction of scores clustered together with TEI and UC but the majority of them were distributed farther from the rest two, statistically differentiating between the techniques and sample types all together at one go.

a) b)

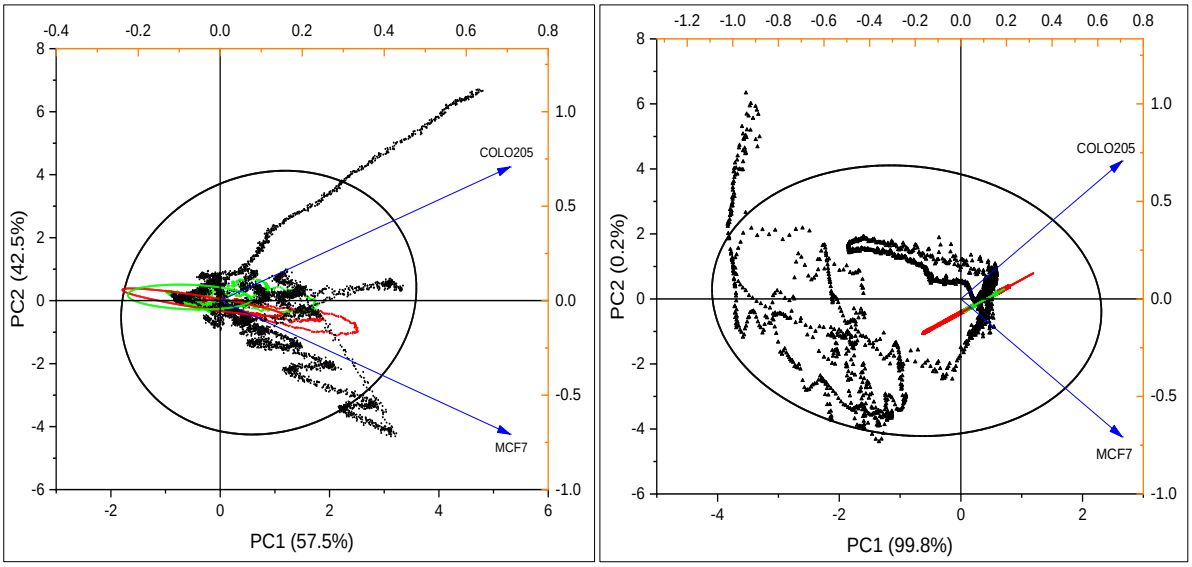


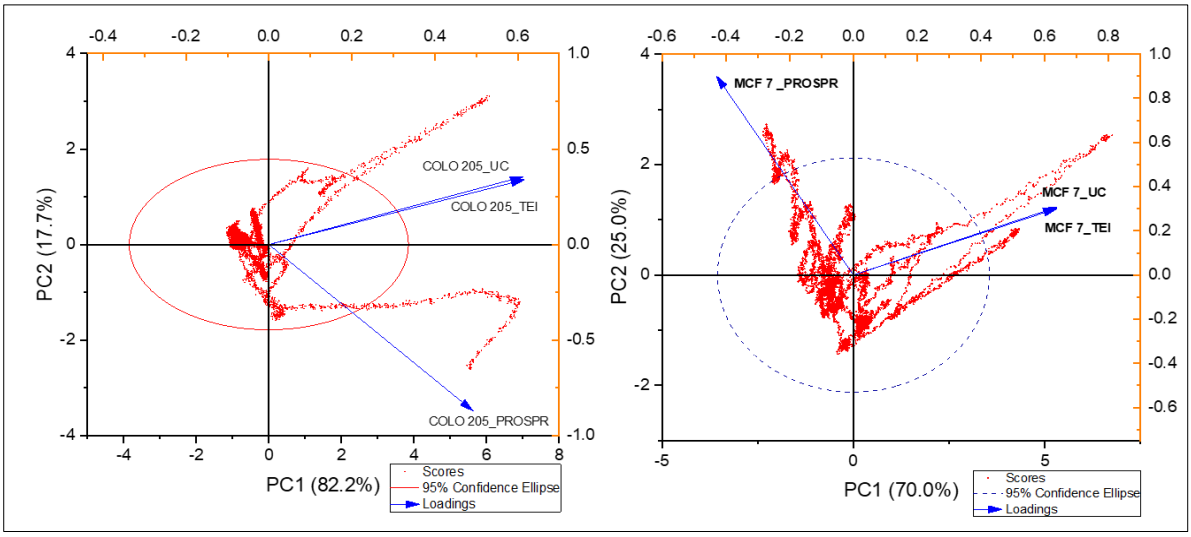
Figure 16. Grouped PCA plot for a) RAMAN spectra and b) ATR-FTIR spectra recorded for COLO 205 and MCF-7 exosomes isolated by TEI, PROSPR, and UC. PCA loadings in blue represent the exosome types and scores represent the grouped spectral data for TEI (red), PROSPR (black) and UC (green). Confidence ellipses in red represent TEI, in black represent PROSPR and in green represent UC.

In case of COLO 205 and MCF-7 exosome types individually, two principal components were selected from the scree plot with PC1 and PC2 having 82.2 % and 17.7% variance in COLO 205 exosomes and 70% and 25% in MCF-7 exosomes in RAMAN spectroscopic data. For ATR-FTIR data PC1 and PC2 had 83.82 % and 15.2% variance in COLO 205 exosomes and 83% and 17 % variance in MCF-7 exosomes. PC1 and PC2 were used to plot a PCA biplot revealing the trend in datasets. The isolation techniques were represented by loadings in biplots, where the loading vector for PROSPR in all the cases was farther away from the central axis having very small correlation with TEI and UC tending to out of phase by 90°. Many component scores corresponding to PROSPR were outliers ranging outside the 95% confidence ellipse. On the other hand, the TEI and UC vectors were very close

to each other, which meant they were highly correlated in the biplots. TEI vectors in both RAMAN and ATR-FTIR spectral analysis had the smallest angle for the variance in projected data compared to UC and PROSPR vectors making it the technique with least variance. For comparative 3D PCA plots spectral data for both the COLO 205 and MCF-7 exosomes were considered in RAMAN and ATR-FTIR data analysis where three out of six principal components were selected for the plot that corresponded to 95% of variables in data in cumulative and 99.9% of variables in cumulative in case of RAMAN and ATR-FTIR respectively. The COLO 205 and MCF-7 TEI and UC vectors or loadings were highly correlated and hence were grouped together. The COLO 205 and MCF-7 PROSPR vectors separately branched out with bigger angles accounting for the high accumulation of variance in projected data. Grouped COLO 205 and MCF-7 exosome datasets in comparative 3D PCA with PROSPR vectors showed the similar trends of non-correlation from TEI and UC datasets as it showed in PCA biplots. Grouped COLO 205 and MCF-7 exosome datasets with TEI and UC vectors were clustered together with strong correlation among the variables. Vectors for COLO 205 and MCF-7 exosomes isolated by TEI got grouped together and exosomes isolated from both the cell lines by UC got grouped together separately. Vectors for PROSPR samples were in non-correlation from each other as well as from other techniques. The trend was the same for RAMAN and ATR-FTIR spectra but with more variability in RAMAN spectra when compared to ATR-FTIR due to large number of spectral recordings and complexity in RAMAN spectroscopy.

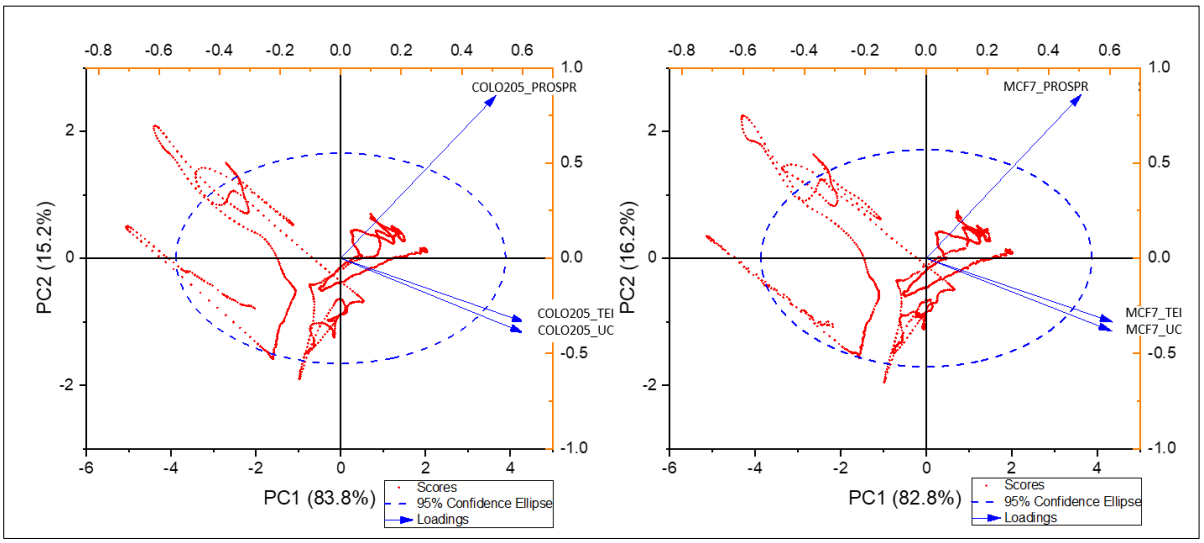
a)

b)



c)

d)



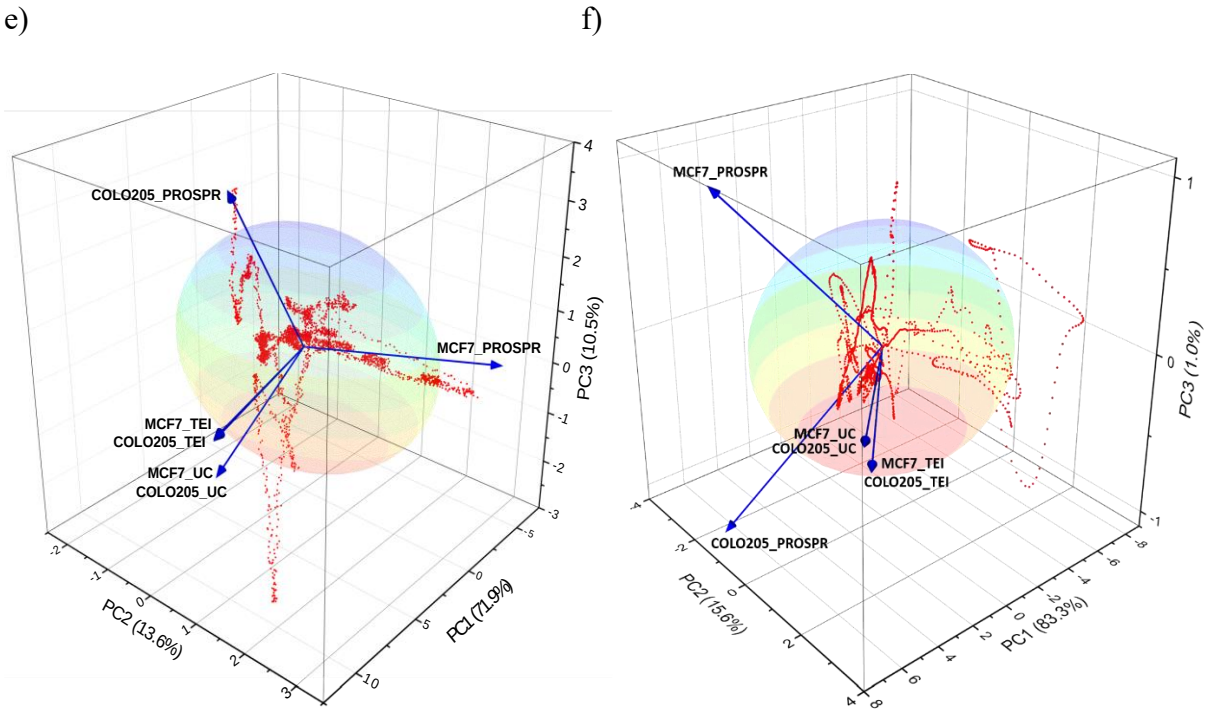


Figure 17. Multivariate analysis of RAMAN and ATR-FTIR spectra for differentiating the exosome isolation techniques and exosome types. PCA biplot for a) COLO 205 exosomes and b) MCF-7 exosomes analysis of RAMAN spectra differentiating between the techniques in the form of loading (in blue). The spectral data were represented as scores (in red). PCA biplot for c) COLO 205 exosomes and d) MCF-7 exosomes analysis of ATR-FTIR spectra differentiating between the techniques in the form of loading (in blue). The spectral data were represented as scores (in red). Comparative 3D PCA plots for e) RAMAN spectra and f) ATR-FTIR spectra of COLO 205 and MCF-7 exosomes differentiating between the techniques.

CONCLUSIONS

In our study, we compared the commonly used exosome isolation techniques TEI, PROSPR, and UC respectively. We analysed the exosomes isolated based on their morphological, biophysical and physiochemical characteristics by various methods to come to a conclusion about the usability of techniques based on the researcher's application. Exosomes isolated by TEI had the highest yield compared to UC and PROSPR with no further

requirement of post-processing for the enrichment of exosomes. PROSPR technique, on the other hand, had yielded exosomes and also some amount of microvesicles along with it that makes the technique require further post-processing for pure exosome enrichment. UC had enriched exosomes when compared to PROSPR but the presence of aggregating materials requires post-processing step a necessary one for this technique. Comparing these techniques based on the recovery and stability factors, TEI yielded good quality and well-dispersed exosomes with high stability compared to PROSPR which did not have expected morphology but had intermediate stability. UC had good recovery of morphology but the presence of aggregating particles resulted in low stability of exosomes in suspension making the particles agglomerate among themselves. PAGE analysis showed the TEI had better protein recovery compared to PROSPR and UC and the western blotting for the marker CD63 supported this observation. The determination of molecular composition through fingerprinting analysis was accurate and easy in the case of TEI followed by UC but was difficult in PROSPR spectra due to high randomness and merged spectral intensities. RAMAN and ATR-FTIR spectral analysis had similar trends for TEI and UC w.r.t statistical and multivariate analysis but differed in case of PROSPR. This may be due to the presence of non exosomal components like extracellular vesicles and apoptotic bodies contaminating the isolated exosomes. This finding was also supported with the determination of relative intensities of spectroscopic protein to lipid ratios.

In summary, TEI based isolation technique is the better technique to isolate exosomes that can be employed directly into clinical translation or other downstream applications. PROSPR can be advantageous to isolate EVs in general due to its low cost and ease of use but it could not be used for exosome with high purity and clinical applications. UC has been the conventional technique and is used frequently in research due to low cost of isolation but due to low particle stability and resultant smaller size because of high centrifugal force it’s not a suitable method for drug delivery or targeted therapeutics kind of applications.

References

1. Colombo, M., Raposo, G. & Théry, C. Biogenesis, Secretion, and Intercellular Interactions of Exosomes and Other Extracellular Vesicles. *Annu. Rev. Cell Dev. Biol.* (2014) doi:10.1146/annurev-cellbio-101512-122326.

2. Li, S. P., Lin, Z. X., Jiang, X. Y. & Yu, X. Y. Exosomal cargo-loading and synthetic exosome-mimics as potential therapeutic tools. *Acta Pharmacologica Sinica* (2018) doi:10.1038/aps.2017.178.

3. Mashouri, L. *et al.* Exosomes: Composition, biogenesis, and mechanisms in cancer metastasis and drug resistance. *Molecular Cancer* (2019) doi:10.1186/s12943-019-0991-5.

4. Théry, C., Zitvogel, L. & Amigorena, S. Exosomes: Composition, biogenesis and function. *Nature Reviews Immunology* (2002) doi:10.1038/nri855.

5. Bunggulawa, E. J. *et al.* Recent advancements in the use of exosomes as drug delivery systems. *J. Nanobiotechnology* (2018) doi:10.1186/s12951-018-0403-9.

6. Zhang, M. *et al.* Exosome-based nanocarriers as bio-inspired and versatile vehicles for drug delivery: recent advances and challenges. *Journal of Materials Chemistry B* (2019) doi:10.1039/C9TB00170K.

7. Balazs, D. A. & Godbey, W. Liposomes for Use in Gene Delivery. *J. Drug Deliv.* (2011) doi:10.1155/2011/326497.

8. Ma, B., Zhang, S., Jiang, H., Zhao, B. & Lv, H. Lipoplex morphologies and their influences on transfection efficiency in gene delivery. *Journal of Controlled Release* (2007) doi:10.1016/j.jconrel.2007.08.022.

9. Zhang, X. X., McIntosh, T. J. & Grinstaff, M. W. Functional lipids and lipoplexes for

improved gene delivery. *Biochimie* (2012) doi:10.1016/j.biochi.2011.05.005.

10. Al-Dosari, M. S. & Gao, X. Nonviral gene delivery: Principle, limitations, and recent Progress. *AAPS Journal* (2009) doi:10.1208/s12248-009-9143-y.

11. Liu, C. & Su, C. Design strategies and application progress of therapeutic exosomes. *Theranostics* (2019) doi:10.7150/thno.30853.

12. Antimisiaris, S. G., Mourtas, S. & Marazioti, A. Exosomes and exosome-inspired vesicles for targeted drug delivery. *Pharmaceutics* (2018) doi:10.3390/pharmaceutics10040218.

13. Lu, M. *et al.* Exosome-based small RNA delivery: Progress and prospects. *Asian Journal of Pharmaceutical Sciences* (2018) doi:10.1016/j.ajps.2017.07.008.

14. Lobb, R. J. *et al.* Optimized exosome isolation protocol for cell culture supernatant and human plasma. *J. Extracell. Vesicles* **4**, 1–11 (2015).

15. Rosenblum, D., Joshi, N., Tao, W., Karp, J. M. & Peer, D. Progress and challenges towards targeted delivery of cancer therapeutics. *Nature Communications* (2018) doi:10.1038/s41467-018-03705-y.

16. Patel, G. K. *et al.* Comparative analysis of exosome isolation methods using culture supernatant for optimum yield, purity and downstream applications. *Sci. Rep.* (2019) doi:10.1038/s41598-019-41800-2.

17. Yu, L. L. *et al.* A Comparison of Traditional and Novel Methods for the Separation of Exosomes from Human Samples. *BioMed Research International* (2018) doi:10.1155/2018/3634563.

18. Tang, Y. T. *et al.* Comparison of isolation methods of exosomes and exosomal RNA from cell culture medium and serum. *Int. J. Mol. Med.* (2017)

doi:10.3892/ijmm.2017.3080.

19. Gallart-Palau, X. *et al.* Extracellular vesicles are rapidly purified from human plasma by PRotein Organic Solvent PRecipitation (PROSPR). *Sci. Rep.* (2015) doi:10.1038/srep14664.

20. Théry, C., Amigorena, S., Raposo, G. & Clayton, A. Isolation and Characterization of Exosomes from Cell Culture Supernatants and Biological Fluids. *Curr. Protoc. Cell Biol.* (2006) doi:10.1002/0471143030.cb0322s30.

21. Alcaraz, C., de Diego, M., Pastor, M. J. & Escribano, J. M. Comparison of a Radioimmunoprecipitation Assay to Immunoblotting and ELISA for Detection of Antibody to African Swine Fever Virus. *J. Vet. Diagnostic Investig.* (1990) doi:10.1177/104063879000200307.

22. Skliar, M. *et al.* Membrane proteins significantly restrict exosome mobility. *Biochem. Biophys. Res. Commun.* **501**, 1055–1059 (2018).

23. Chernyshev, V. S. *et al.* Size and shape characterization of hydrated and desiccated exosomes. *Anal. Bioanal. Chem.* (2015) doi:10.1007/s00216-015-8535-3.

24. Jung, M. K. & Mun, J. Y. Sample Preparation and Imaging of Exosomes by Transmission Electron Microscopy. *J. Vis. Exp.* 5–9 (2018) doi:10.3791/56482.

25. Linares, R., Tan, S., Gounou, C., Arraud, N. & Brisson, A. R. High-speed centrifugation induces aggregation of extracellular vesicles. *Journal of Extracellular Vesicles* (2015) doi:10.3402/jev.v4.29509.

26. Fernández-Llama, P. *et al.* Tamm-Horsfall protein and urinary exosome isolation. *Kidney Int.* (2010) doi:10.1038/ki.2009.550.

27. Wachalska, M. *et al.* Protein Complexes in Urine Interfere with Extracellular Vesicle

Biomarker Studies. *J. Circ. Biomarkers* (2016) doi:10.5772/62579.

28. de la Torre Gomez, C., Goreham, R. V., Bech Serra, J. J., Nann, T. & Kussmann, M. 'Exosomics'-A review of biophysics, biology and biochemistry of exosomes with a focus on human breast milk. *Frontiers in Genetics* (2018) doi:10.3389/fgene.2018.00092.

29. Gordiychuk, A., Svanera, M., Benini, S. & Poesio, P. Size distribution and Sauter mean diameter of micro bubbles for a Venturi type bubble generator. *Experimental Thermal and Fluid Science* (2016) doi:10.1016/j.expthermflusci.2015.08.014.

30. Yap, F. Y. *et al.* Quantitative morphometric analysis of hepatocellular carcinoma: Development of a programmed algorithm and preliminary application. *Diagnostic Interv. Radiol.* (2013) doi:10.4261/1305-3825.DIR.5973-12.1.

31. Kesimer, M. & Gupta, R. Physical characterization and profiling of airway epithelial derived exosomes using light scattering. *Methods* (2015) doi:10.1016/j.ymeth.2015.03.013.

32. Davis, C. N. *et al.* The importance of extracellular vesicle purification for downstream analysis: A comparison of differential centrifugation and size exclusion chromatography for helminth pathogens. *PLoS Negl. Trop. Dis.* (2019) doi:10.1371/journal.pntd.0007191.

33. Kotrbová, A. *et al.* TEM ExosomeAnalyzer: a computer-assisted software tool for quantitative evaluation of extracellular vesicles in transmission electron microscopy images. *J. Extracell. Vesicles* (2019) doi:10.1080/20013078.2018.1560808.

34. Baker, M. J. *et al.* Using Fourier transform IR spectroscopy to analyze biological materials. *Nat. Protoc.* (2014) doi:10.1038/nprot.2014.110.

35. Smith, Z. J. *et al.* Single exosome study reveals subpopulations distributed among cell lines with variability related to membrane content. *J. Extracell. Vesicles* **4**, 1–15 (2015).

36. Ali, S. M. *et al.* A comparison of Raman, FTIR and ATR-FTIR micro spectroscopy for imaging human skin tissue sections. *Anal. Methods* (2013) doi:10.1039/c3ay40185e.

37. Blat, A. *et al.* FTIR, Raman and AFM characterization of the clinically valid biochemical parameters of the thrombi in acute ischemic stroke. *Sci. Rep.* (2019) doi:10.1038/s41598-019-51932-0.

38. De Gelder, J., De Gussem, K., Vandenabeele, P. & Moens, L. Reference database of Raman spectra of biological molecules. *J. Raman Spectrosc.* (2007) doi:10.1002/jrs.1734.

39. Griffiths, P. R. The Handbook of Infrared and Raman Characteristic Frequencies of Organic Molecules. *Vib. Spectrosc.* (1992) doi:10.1016/0924-2031(92)87021-7.

40. Zlotogorski-Hurvitz, A., Dekel, B. Z., Malonek, D., Yahalom, R. & Vered, M. FTIR-based spectrum of salivary exosomes coupled with computational-aided discriminating analysis in the diagnosis of oral cancer. *J. Cancer Res. Clin. Oncol.* (2019) doi:10.1007/s00432-018-02827-6.

41. Mihály, J. *et al.* Characterization of extracellular vesicles by IR spectroscopy: Fast and simple classification based on amide and C[sbnd]H stretching vibrations. *Biochim. Biophys. Acta - Biomembr.* (2017) doi:10.1016/j.bbamem.2016.12.005.

42. Susi, H. & Byler, D. M. [13] Resolution-Enhanced Fourier Transform Infrared Spectroscopy of Enzymes. *Methods Enzymol.* (1986) doi:10.1016/0076-6879(86)30015-6.

43. Venyaminov, S. Y. & Kalnin, N. N. Quantitative IR spectrophotometry of peptide compounds in water (H2O) solutions. I. Spectral parameters of amino acid residue absorption bands. *Biopolymers* (1990) doi:10.1002/bip.360301309.

Supplementary information for:

Exosomes isolated from two different cell lines using three different isolation techniques show variation in physical and molecular characteristics

Manish Dash[1], Kanagaraj Palaniyandi[3], Satish Ramalingam[4], S. Sahabudeen[5], N S Raja[2*]

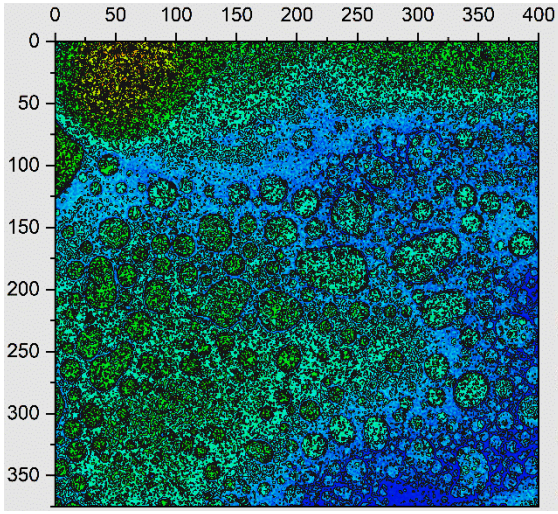
[1,2,4] Department of Genetic Engineering, SRM Institute of Science and Technology, Kattankulathur, Chennai, 603203, Tamil Nadu, India

[3,5] Department of Biotechnology, SRM Institute of Science and Technology, Kattankulathur, Chennai, 603203, Tamil Nadu, India

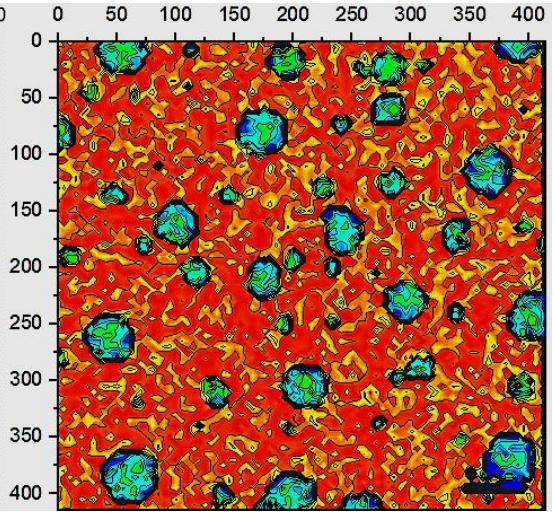
*Corresponding Author.

Email address: rajan3@srmist.edu.in

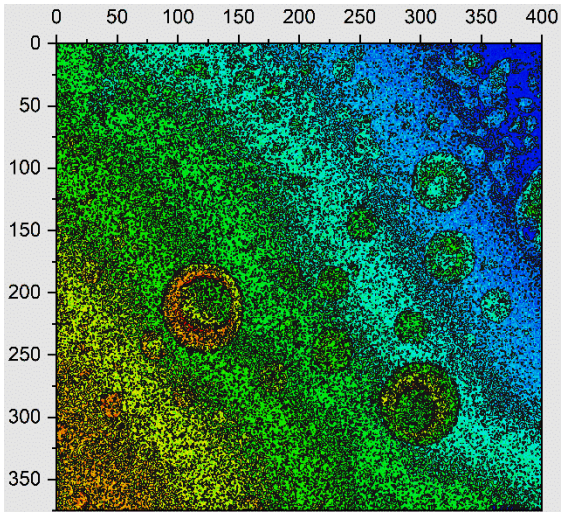
a.i)



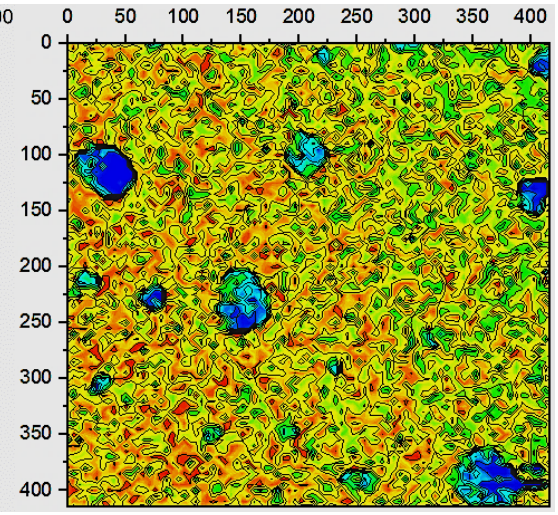
b.i)



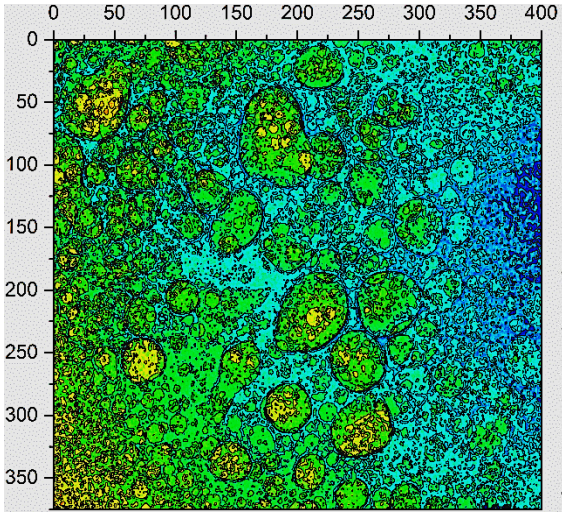
a.ii)



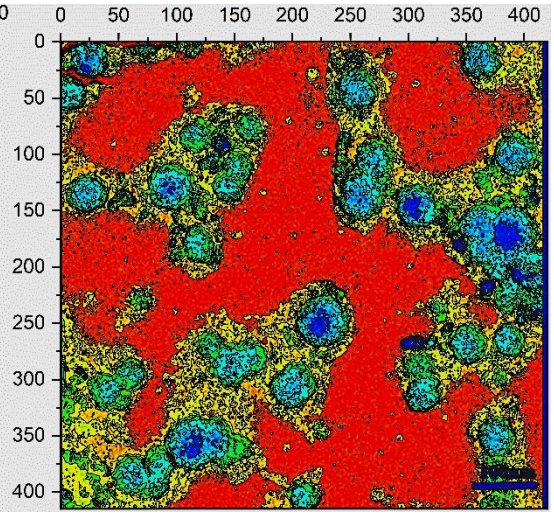
b.ii)



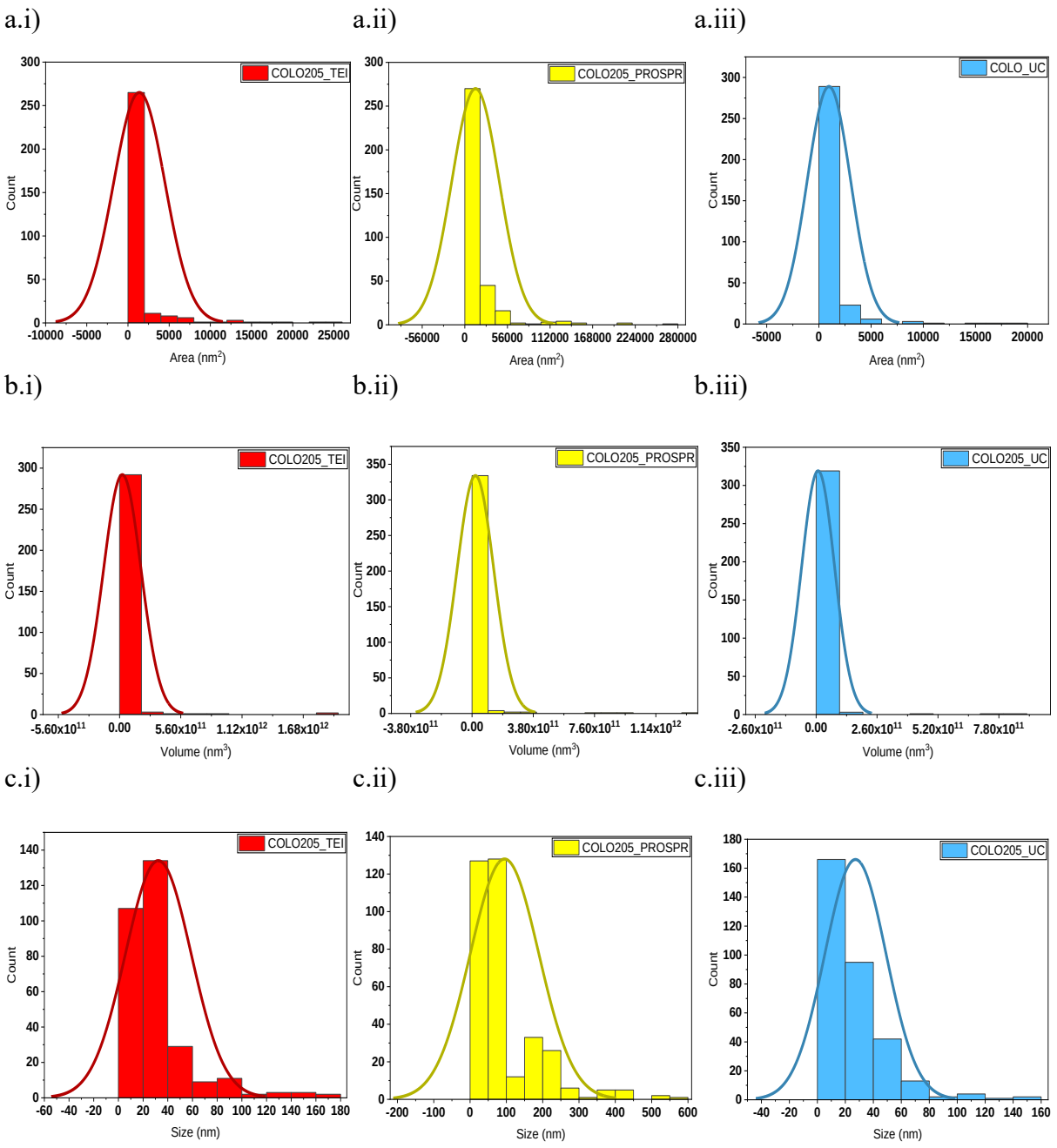
a.iii)



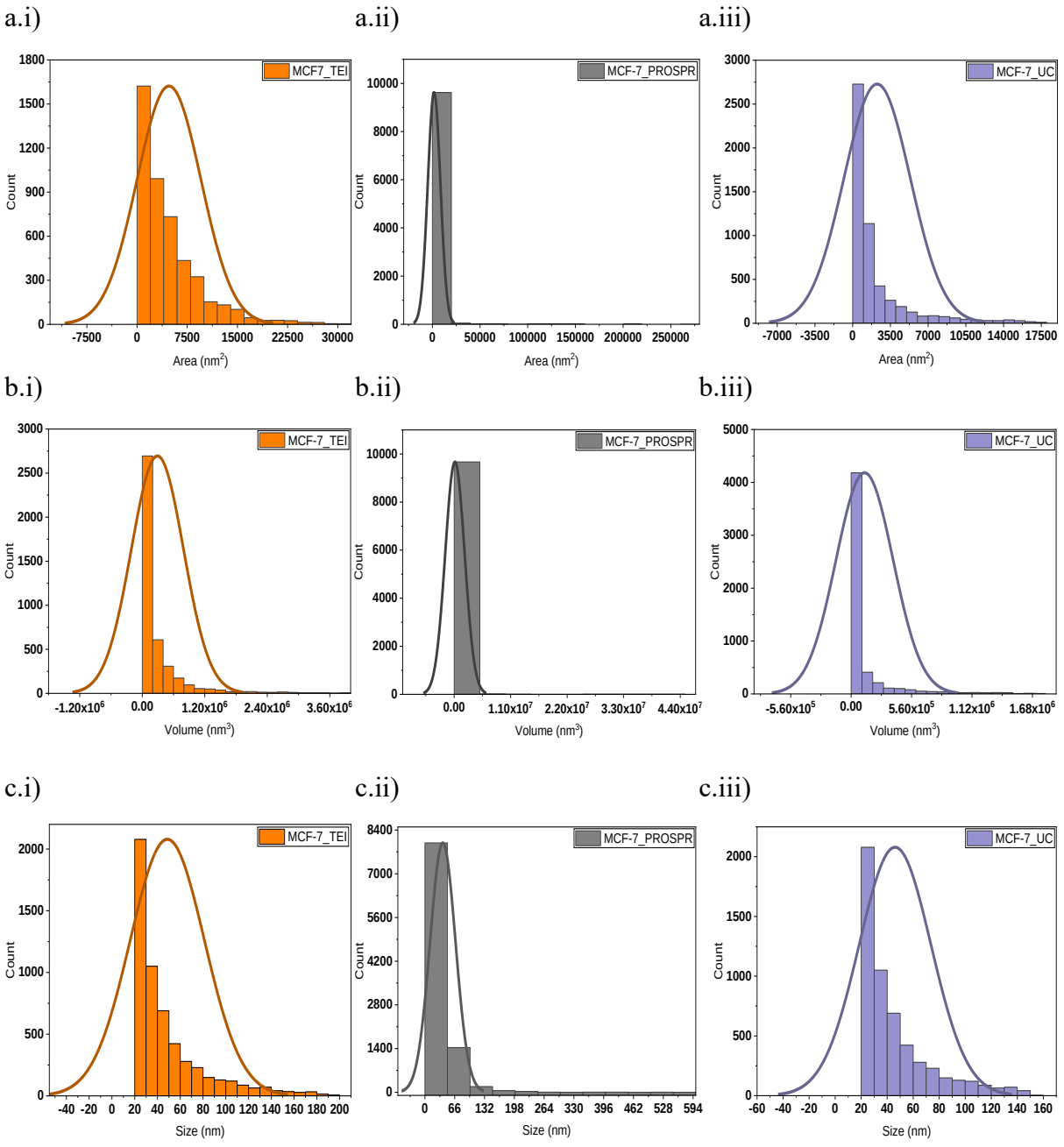
b.iii)



Supplementary Figure S1. Representation of contour maps plotted for the particle size distribution analysis of COLO 205 and MCF-7 exosome images obtained from HRTEM. The analysis was done by calculating the exosome area, volume and size form contour line data that was extracted from the matrix of bitmap images created. a) COLO 205 exosomes isolated by i) TEI, ii) PROSPR and iii) UC. b) MCF-7 exosomes isolated by i) TEI, ii) PROSPR and iii) UC. The x and y axes were scaled with arbitrary units.

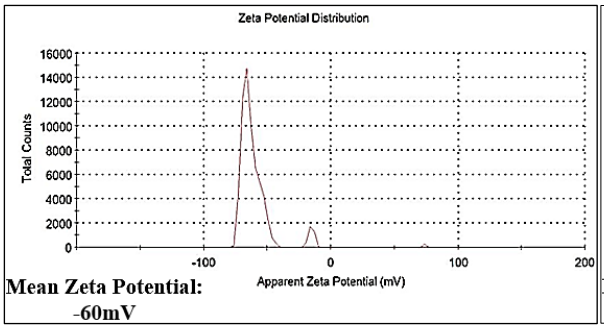


Supplementary Figure S2. Normal distribution curves of COLO 205 exosomes a) Area distribution curves of COLO 205 exosomes isolated by i) TEI, ii) PROSPR and iii) UC. b) Volumetric distribution curves of COLO 205 exosomes isolated by i) TEI, ii) PROSPR and iii) UC. c) Size distribution curves of COLO 205 exosomes isolated by i) TEI, ii) PROSPR and iii) UC.

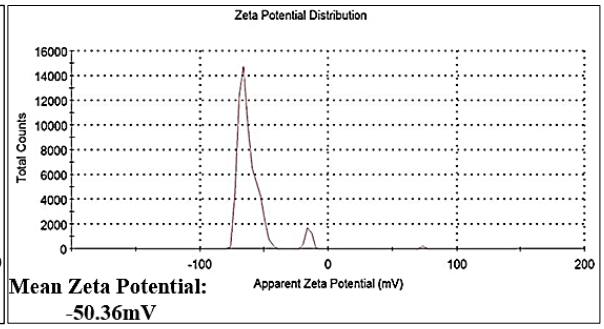


Supplementary Figure S3. Normal distribution curves of MCF-7 exosomes a) Area distribution curves of MCF-7 exosomes isolated by i) TEI, ii) PROSPR and iii) UC. b) Volumetric distribution curves of MCF-7 exosomes isolated by i) TEI, ii) PROSPR and iii) UC. c) Size distribution curves of MCF-7 exosomes isolated by i) TEI, ii) PROSPR and iii) UC.

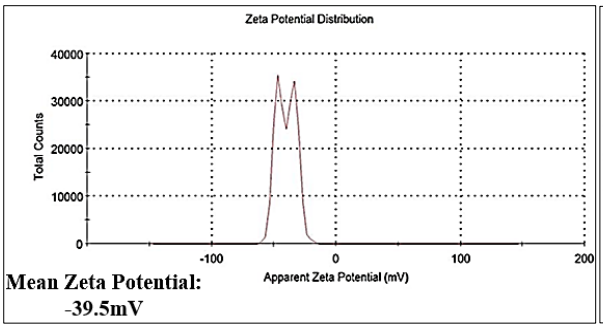
a.i) COLO 205_TEI



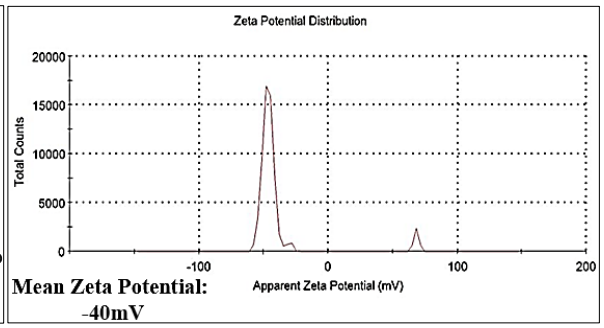
b.i) MCF-7_TEI



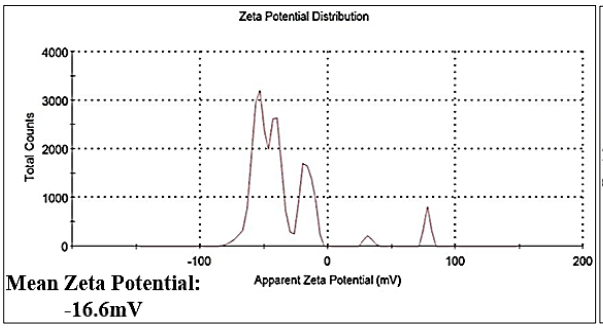
a.ii) COLO 205_PROSPR



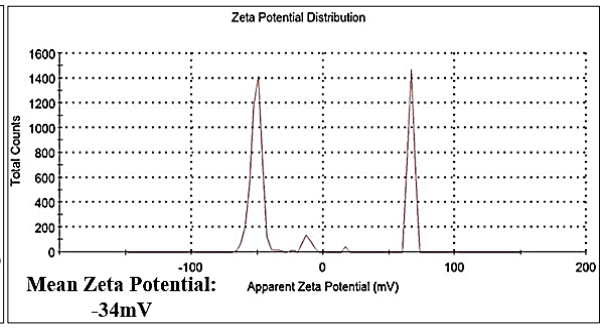
b.ii) MCF-7_PROSPR



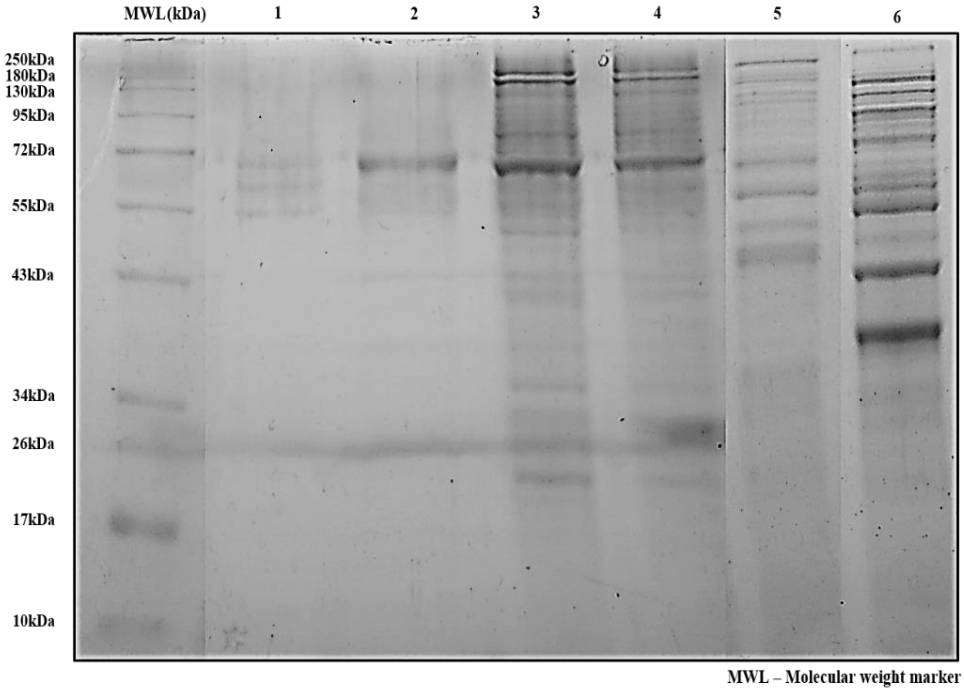
a.iii) COLO 205_UC



b.iii) MCF-7_UC



Supplementary Figure S4. System generated graphs for zeta potential analysis of a) COLO 205 exosomes isolated by i) TEI ii) PROSPR and iii) UC. b) MCF-7 exosomes isolated by i) TEI ii) PROSPR and iii) UC. The representative graphs generated were average of three instrumental triplicates. The mean zeta potential was calculated by average of three biological triplicates.



Supplementary Figure S5. SDS-PAGE for comparative exosomal proteome profile. Sample loaded in lane 1 was exosomes isolated with TEI and treated with proteinase K. In lane 2 PROSPR sample, lane 3 COLO 205 exosomes isolate by TEI, lane 4 MCF-7 exosomes isolated by TEI, lane 5 UC samples and lane 6 total extracellular proteins isolated by acetone precipitation were loaded. The samples were directly digested and denatured with SDS sample loading buffer.

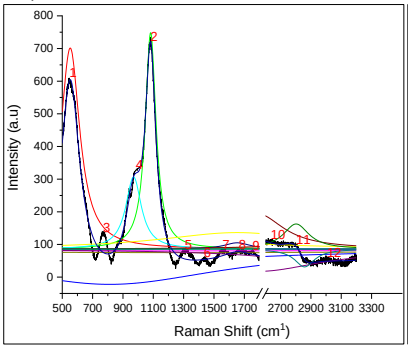
Parameters	**TEI (nm)**	**PROSPR (nm)**	**UC (nm)**
Min	8.03	5.98	8.1
First Quartile	17.82	39.301	13.41
Median Value	24.14	54.42	20.73
Third Quartile	34.93	79.01	33.59
Max	174.82	667.2	154.51

Supplementary Table S1. Quartile parameters of comparative box plots for COLO 205 exosomes.

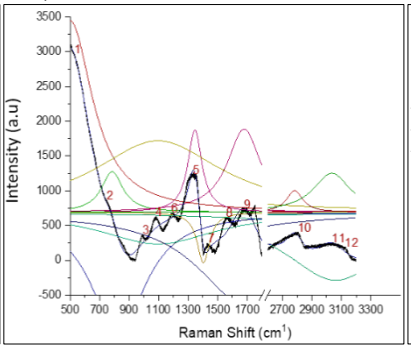
Parameters	**TEI (nm)**	**PROSPR (nm)**	**UC (nm)**
Min	22.56	22.56	20.28
First Quartile	37.42	25.23	26.92
Median Value	58.63	31.91	35.50
Third Quartile	84.44	43.70	54.31
Max	196.09	577.24	151.01

Supplementary Table S2. Quartile parameters of comparative box plots for MCF-7 exosomes.

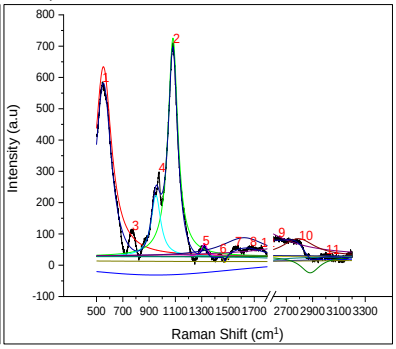
a.i)



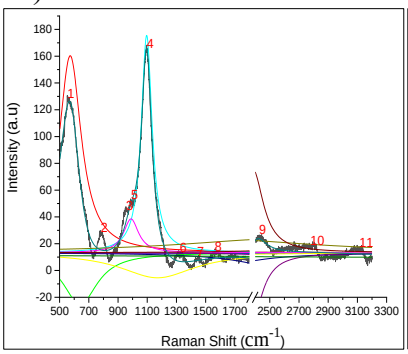
a.ii)



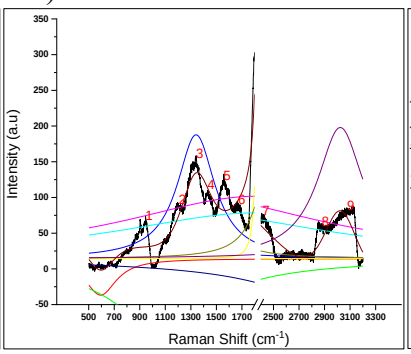
a.iii)



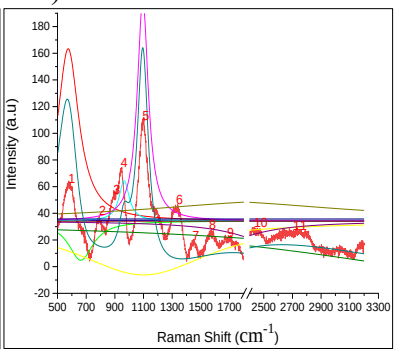
b.i)



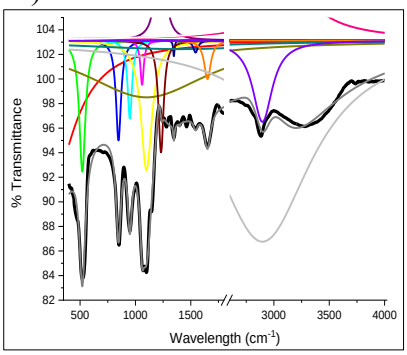
b.ii)



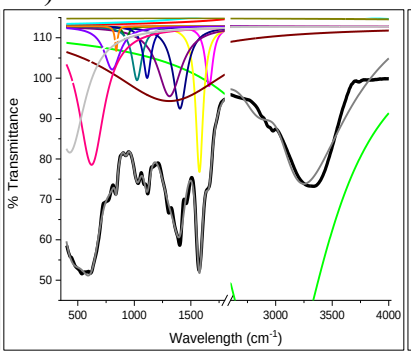
b.iii)



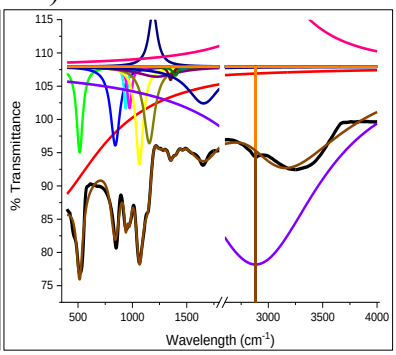
c.i)



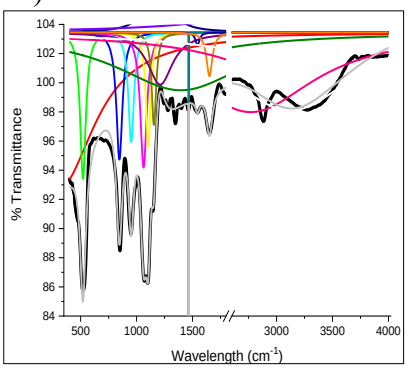
c.ii)



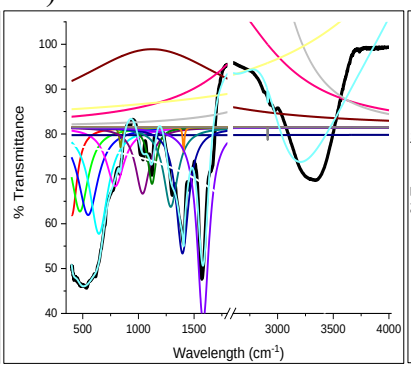
c.iii)



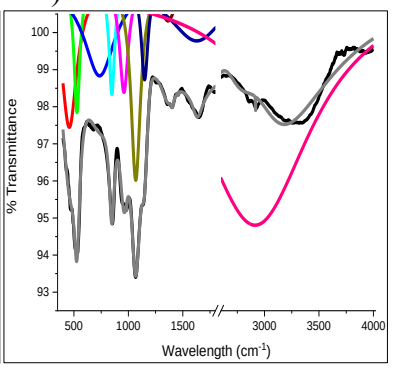
d.i)



d.ii)



d.iii)



Supplementary Figure S6. Representation of non-linear multiple curve fitting and regression analyses performed for 400 iterations with Levenberg Marquardt algorithm (Function: Lorentz) for determining the molecular components in RAMAN spectra of a) COLO 205 exosomes isolated by i) TEI with reduced chi sq. of 192 and R^2 of 0.98, ii) PROSPR with reduced chi sq. of 739 and R^2 of 0.97 and iii) UC with reduced chi sq. of 167 and R^2 of 0.99. b) MCF-7 exosomes isolated by i) TEI with reduced chi sq. of 12 and R^2 of 0.99, ii) PROSPR with reduced chi sq. of 39 and R^2 of 0.97 and iii) UC with reduced chi sq. of 14 and R^2 of 0.98. Non-linear multiple curve fitting and regression analyses were also performed for ATR-FTIR spectra for determining the molecular components. c) COLO 205 exosomes isolated by i) TEI with reduced chi sq. of 0.19 and R^2 of 0.98, ii) PROSPR with reduced chi sq. of 2.68 and R^2 of 0.98 and iii) UC with reduced chi sq. of 0.46 and R^2 of 0.98. d) MCF-7 exosomes isolated by i) TEI with reduced chi sq. of 0.24 and R^2 of 0.98, ii) PROSPR with reduced chi sq. of 7.68 and R^2 of 0.97 and iii) UC with reduced chi sq. of 0.35 and R^2 of 0.97. The RAMAN and ATR-FTIR spectra were deconvoluted for finding and retrieving the peaks withing mixture of spectrum containing the molecular components of exosomes.

a)

Isolation Techniques	**Spectroscopic protein to lipid ratio calculated from ATR-FTIR spectra**		**Mean ± SD**
	COLO 205 exosomes	MCF-7 exosomes	
TEI	1.184	1.184	1.18 ± 0.0001
PROSPR	1.484	1.51	1.49 ± 0.015
UC	1.18	1.176	1.18 ± 0.004

b)

Isolation Techniques	**Spectroscopic protein to lipid ratio calculated from RAMAN spectra**		**Mean ± SD**
	COLO 205 exosomes	MCF-7 exosomes	
TEI	1.142	0.985	1.06 ± 0.11
PROSPR	2.051	2.038	2.04 ± 0.01
UC	1.125	0.977	1.05 ± 0.1

Supplementary Table S3. Spectroscopic protein to lipid ratio calculated for a) RAMAN spectra and b) ATR-FTIR spectra of COLO 205 and MCF-7 exosomes isolated by TEI, PROSPR and UC.

a) Descriptive Statistics

	N Analysis	N Missing	Mean	Standard Deviation	SE of Mean
COLO205_TEI	6674	0	116.4	129.21	1.58
COLO205_PROSPR	6674	0	450.7	535.64	6.55
COLO205_UC	6674	0	91.48	127.5	1.56

b) Homogeneity of Variance Test (Levene's Test)

	DF	Sum of squares	Mean Square	F Value	Prob>F
Model	2	301974000	150987000	2348.61	0
Error	20019	128698000	64287.73		

c) One-way ANOVA (overall ANOVA)

	DF	Sum of squares	Mean Square	F Value	Prob>F
Model	2	537056000	268528000	2518.55	0
Error	20019	2134420000	106619.9		
Total	20021	2671480000			

d) Means Comparisons (Tukey's Test)

	MeanDiff	SEM	q Value	Prob	Alpha	Sig	LCL	UCL
COLO205_PROSPR COLO205_TEI	334.3	5.65	89.9	$3.3 * 10^{-16}$	0.05	1	345.97	372.46
COLO205_UC COLO205_TEI	-24.93	5.65	6.23	$3.1 * 10^{-5}$	0.05	1	11.67	38.17
COLO205_UC COLO205_PROSPR	-359.21	5.65	83.63	$3.3 * 10^{-16}$	0.05	1	-348	-321

Supplementary Table S4. a) Descriptive statistics b) Levene's test (The Homogeneity of Variance tested at 0.05 level resulted in the population variances to be significantly different) c) One-way ANOVA (The null hypothesis assumed that the means of all levels were equal and the alternate hypothesis assumed that the means of one or more levels or groups were different. At 0.05 level, the population means were found to be significantly different) d) Tukey's test (Significance equaled to 1 indicated that the difference of the means were significantly different at the 0.05 level but if the significance equaled to 0 that would have had indicated that the difference of the means were not significant at 0.05 level) for RAMAN spectra of COLO 205 exosomes. The statistical power was 1 for alpha 0.05 having sample size of 20022.

a) Descriptive Statistics

	N Analysis	N Missing	Mean	Standard Deviation	SE of Mean
MCF-7_TEI	6674	0	29.93	28.1	0.34
MCF-7_PROSPR	6674	0	76.82	66.07	0.8
MCF-7_UC	6674	0	15.88	19.21	0.23

b) Homogeneity of Variance Test (Levene's Test)

	DF	Sum of squares	Mean Square	F Value	Prob>F
Model	2	5327621.56	2663810.78	3076.32	0
Error	20019	17334600	865.9		

c) One-way ANOVA (overall ANOVA)

	DF	Sum of squares	Mean Square	F Value	Prob>F
Model	2	13591300	6795628.35	3690.24	0
Error	20019	36865200	1841.5		
Total	20021	5.0456400			

d) Means Comparisons (Tukey's Test)

	MeanDiff	SEM	q Value	Prob	Alpha	Sig	LCL	UCL
MCF-7_PROSPR MCF-7_TEI	46.89	0.74	89.26	$3.3 * 10^{-16}$	0.05	1	45.14	48.62
MCF-7_UC MCF-7_TEI	-14.05	0.74	26.74	$3.3 * 10^{-16}$	0.05	1	-15.8	-12.3
MCF-7_UC MCF-7_PROSPR	-60.93	0.74	116	$3.3 * 10^{-16}$	0.05	1	-62.6	-59.19

Supplementary Table S5. a) Descriptive statistics b) Levene's test (The Homogeneity of Variance tested at 0.05 level resulted in the population variances to be significantly different) c) One-way ANOVA (The null hypothesis assumed that the means of all levels were equal and the alternate hypothesis assumed that the means of one or more levels or groups were different. At 0.05 level, the population means were found to be significantly different) d) Tukey's test (Significance equaled to 1 indicated that the difference of the means were significantly different at the 0.05 level but if the significance equaled to 0 that would have had indicated that the difference of the means were not significant at 0.05 level) for RAMAN spectra of MCF-7 exosomes. The statistical power was 1 for alpha 0.05 having sample size of 20022.

a) Descriptive Statistics

	N Analysis	N Missing	Mean	Standard Deviation	SE of Mean
COLO205_TEI	1868	0	98	3.53	0.08
COLO205_PROSPR	1868	0	83.15	14.41	0.33
COLO205_UC	1868	0	97.94	1.27	0.03

b) Homogeneity of Variance Test (Levene's Test)

	DF	Sum of squares	Mean Square	F Value	Prob>F
Model	2	148015.43	74007.71	3892.16	0
Error	5601	106500.56	19.01		

c) One-way ANOVA (overall ANOVA)

	DF	Sum of squares	Mean Square	F Value	Prob>F
Model	2	273355.35	136677.67	1848.104	0
Error	5601	414225.26	73.95		
Total	5603	687580.61			

d) Means Comparisons (Tukey's Test)

	MeanDiff	SEM	q Value	Prob	Alpha	Sig	LCL	UCL
COLO205_PROSPR COLO205_TEI	-14.84	0.28	74.6	$3.3 * 10^{-16}$	0.05	1	-15.5	-14.18
COLO205_UC COLO205_TEI	-0.05	0.28	0.25	0.98	0.05	0	-0.7	0.6
COLO205_UC COLO205_PROSPR	14.8	0.28	74.33	$3.3 * 10^{-16}$	0.05	1	14.13	15.45

Supplementary Table S6. a) Descriptive statistics b) Levene's test (The Homogeneity of Variance tested at 0.05 level resulted in the population variances to be significantly different) c) One-way ANOVA (The null hypothesis assumed that the means of all levels were equal and the alternate hypothesis assumed that the means of one or more levels or groups were different. At 0.05 level, the population means were found to be significantly different) d) Tukey's test (Significance equaled to 1 indicated that the difference of the means were significantly different at the 0.05 level and significance equaled to 0 indicated that the difference of the means were not significant at 0.05 level) for ATR-FTIR spectra of COLO 205 exosomes. The statistical power was 1 for alpha 0.05 having sample size of 20022.

a) Descriptive Statistics

	N Analysis	N Missing	Mean	Standard Deviation	SE of Mean
MCF-7_TEI	1868	0	97.99	3.5	0.08
MCF-7_PROSPR	1868	0	82.01	15.72	0.36
MCF-7_UC	1868	0	97.94	1.27	0.03

b) Homogeneity of Variance Test (Levene's Test)

	DF	Sum of squares	Mean Square	F Value	Prob>F
Model	2	171238.54	85619.27	3415.9	0
Error	5601	140388.46	25.06		

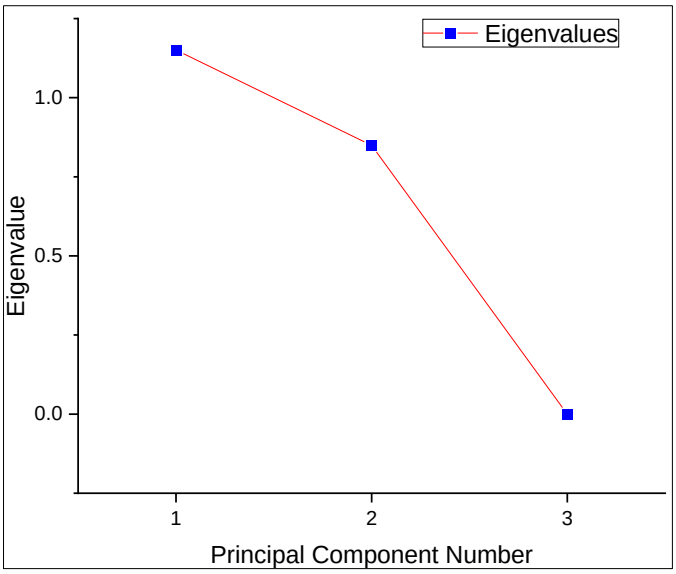
c) One-way ANOVA (overall ANOVA)

	DF	Sum of squares	Mean Square	F Value	Prob>F
Model	2	317082.46	158541.23	1818.41	0
Error	5601	488330.63	87.18		
Total	5603	805413.09			

d) Means Comparisons (Tukey's Test)

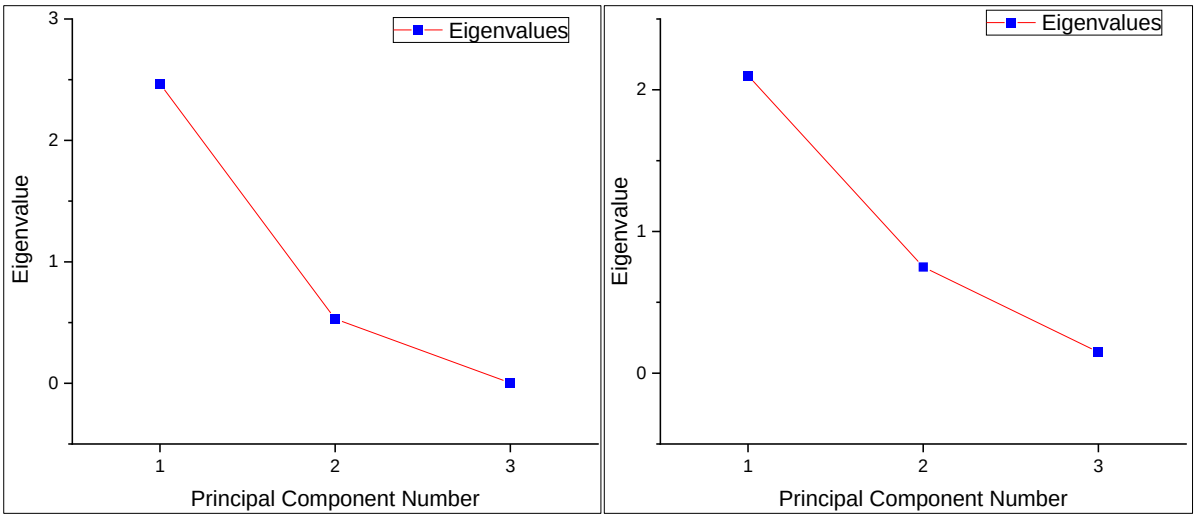
	MeanDiff	SEM	q Value	Prob	Alpha	Sig	LCL	UCL
MCF-7_PROSPR MCF-7_TEI	-15.98	0.3	73.97	$3.3 * 10^{-16}$	0.05	1	-16.69	-15.26
MCF-7_UC MCF-7_TEI	-0.04	0.3	0.23	0.98	0.05	0	-0.7	0.66
MCF-7_UC MCF-7_PROSPR	15.93	0.3	73.74	$3.3 * 10^{-16}$	0.05	1	15.21	16.64

Supplementary Table S7. a) Descriptive statistics b) Levene's test (The Homogeneity of Variance tested at 0.05 level resulted in the population variances to be significantly different) c) One-way ANOVA (The null hypothesis assumed that the means of all levels were equal and the alternate hypothesis assumed that the means of one or more levels or groups were different. At 0.05 level, the population means were found to be significantly different) d) Tukey's test (Significance equaled to 1 indicated that the difference of the means were significantly different at the 0.05 level but if the significance equaled to 0 that would have had indicated that the difference of the means were not significant at 0.05 level) for ATR-FTIR spectra of MCF-7 exosomes. The statistical power was 1 for alpha 0.05 having sample size of 20022.

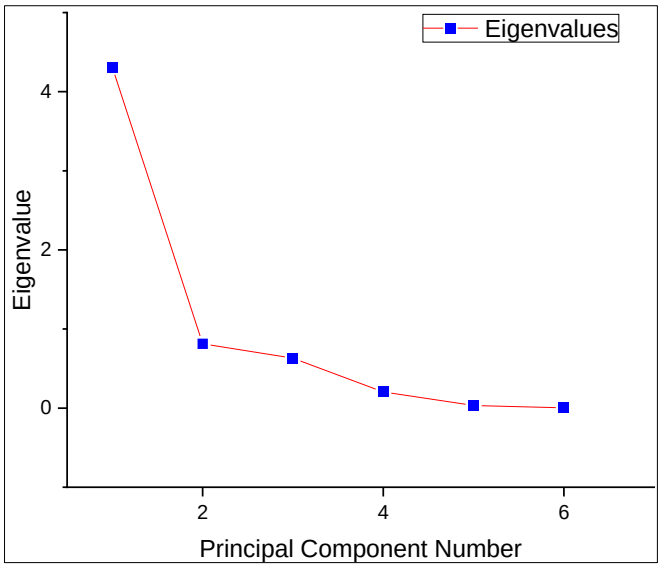


Supplementary Figure S7. Scree plot for grouped PCA of RAMAN spectra for COLO 205 and MCF-7 exosomes.

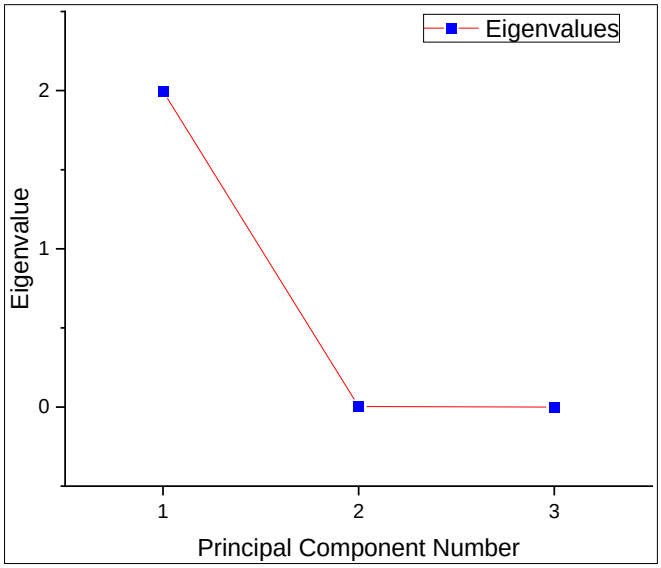
a) b)



Supplementary Figure S8. Scree plots for eigen values of PCA biplots plot for RAMAN spectra of a) COLO 205 exosomes and b) MCF-7 exosomes.

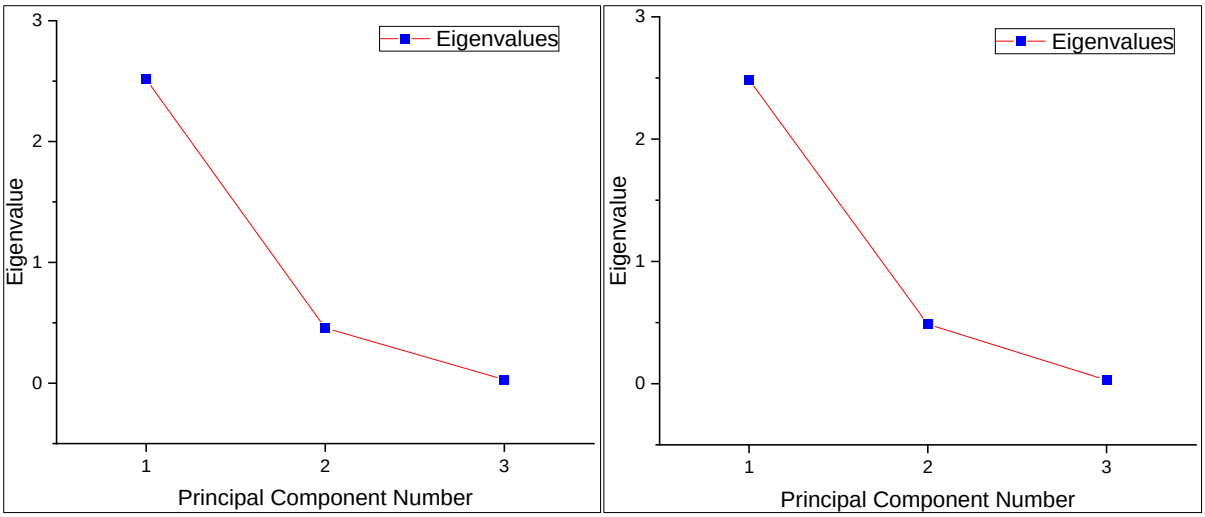


Supplementary Figure S9. Scree plots for eigen values of 3D comparative PCA for RAMAN spectra of COLO 205 and MCF-7 exosomes.

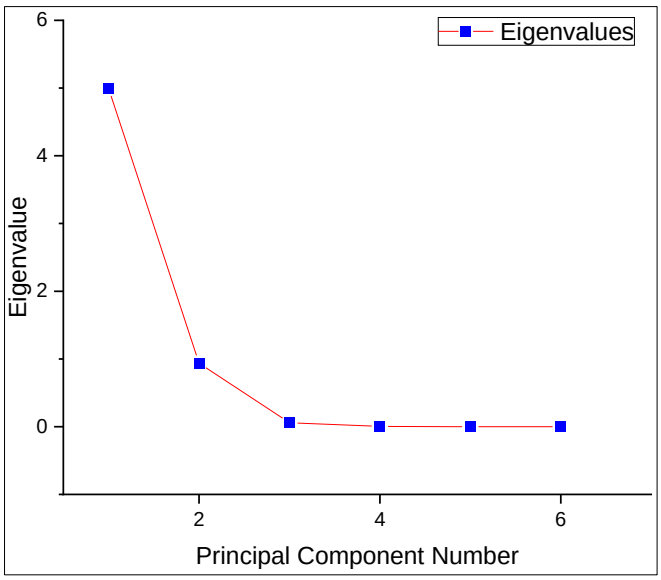


Supplementary Figure S10. Scree plot for grouped PCA of ATR-FTIR spectra for COLO 205 and MCF-7 exosomes.

a)



Supplementary Figure S11. Scree plots for eigen values of PCA biplots plot for ATR-FTIR spectra of a) COLO 205 exosomes and b) MCF-7 exosomes.



Supplementary Figure S12. Scree plots for eigenvalues of 3D comparative PCA for ATR-FTIR spectra of COLO 205 and MCF-7 exosomes.