Complex Nanostructures and Bio-nanoscale Interactions：Well Defined Synthesis, Identification and Biological Effects

Wei Zhang, PhD

UCD Student Number: 16211384

This thesis is submitted to University College Dublin in fulfilment of the requirements for the degree of Doctor of Philosophy in the College of Science.

UCD School of Chemistry

Head of School: Professor Michael Casey

Principal Supervisor: Professor Kenneth A. Dawson

December 2021

cbni
Centre for BioNano Interactions

CHINA SCHOLARSHIP COUNCIL

HEA HIGHER EDUCATION AUTHORITY
AN tÚDARÁS um ARD-OIDEACHAS

Contents

Abstract

Since Richard Feynman firstly introduced the nanotechnology concept in 1959 and further defined by Norio Taniguchi in 1974 as “nanotechnology mainly consists of the processing of separation, consolidation and deformation of materials by one atom or one molecule”, nanoscience has attracted a huge amount of researchers’ interest during the past six decades and therefore has achieved accelerated development, and the applications have grown from physiochemical fields to the nanomedicine area.

Nowadays it has become inevitable for human beings to be exposed to nanostructures with distinctive complex shapes, due to the widespread development of nanotechnologies. Meanwhile bio-nanoscale interactions engage in nearly all aspects of coordinated biological actions ranging from viral infection to exosome-based cell-cell communication. To date, the shape-dependent biological effects have been acknowledged and phenomenological study has been performed. However, the challenges to investigate the shape-dependent bio-nanoscale recognitions and interactions in such an immense universe are significant. One of the obstacles is how to achieve a systematic exploration of the shape effects instead of a case-by-case study.

To untangle this challenge, we introduce a microfluidic synthesis platform, which allows reproducible and tuneable synthesis across the boundless nanoscale shape universe. After achieving high-quality synthesis of complex nano shapes, we use a quantitative computational shape species identification method to classify the shape ensembles and then combine the shape analysis with high dimensional biological responses *in vitro* and *in vivo*, to illustrate those different classes of nanoscale shape ensembles that elicit distinct biological responses.

Meanwhile, based on the advanced computational shape analysis method, two shapes were picked — “spherical” and “star-like” for a deeper investigation of shape impacts on the epigenome. To achieve robust results for the shape effects in the epigenome, the chosen two shapes are synthesized with similar size, surface area and surface charge. Different batches used during this project were cautiously synthesized to achieve high reproducibility. After high standard control of the quality of the two shapes, a basic cell viability assay was done for two cell lines to get a suitable treatment concentration without obvious cytotoxicity. Following the cell viability test, cellular uptake was detected by inductively coupled plasma mass spectrometry (ICP-MS) which showed a higher uptake of the spherical shape than the branched one. Then transcriptomics was performed to investigate the key cellular responses induced by shapes and differential expression gene analysis found that these genes are related with epigenome modifications. Based on the transcriptomics result, proteomics were

operated to reveal the global changes of histone methylation and acetylation induced by shapes and an obvious difference in the post-translational modifications on histone H3 was observed. ELISA and western blot results also validated the proteomics discoveries. Finally, chromatin immunoprecipitation sequencing (ChIP-seq) was carried out to profile H3K27me3 modifications that are tuned by different shapes. The ChIP-seq analysis demonstrated that these two distinctive shapes triggered H3K27me3 marks differently and targeted distinct pathways.

In addition to the investigation of the biological effects of nanostructures, how to artificial design these nanostructures to achieve their unique function like nano-therapy is also an important issue to be solved. Thus, in the third part of my work, endogenous cellular processes were used to re-engineer the surface of the nanoparticles, these were subsequently isolated in sufficient abundance to enable the characterization of their physiochemical properties and a study of their functional properties. These re-engineered particle complexes possess a well-defined and reproducible structure and composition that is distinctive, including lipids, proteins and nucleic acid. Conventional biotechnology modifications of the producer cells within which these particle complexes are formed, allowed us to further control the new nanoscale constructs. As a consequence, they showed striking cellular interactions with target cells, adopting different pathways to the original nanoparticles, including evasion of the lysosome, and enabling gene transfer to recipient cells.

In this thesis, a framework was proposed in chapter II, aiming to identify distinct shape populations and build a quantitative linkage of well-defined nanoscale shapes to biological impacts. This inductive nanoscale shape discovery and evaluation framework is biologically relevant, and we believe by utilizing machine learning it could benefit the field of shape dependent therapy. In chapter III, the shape dependent histone modifications were reported. As histone modifications are one of the crucial epigenetic regulators that control chromatin structure and gene transcription, shape dependent histone modifications indicate that some important cellular phenotypes differences induced by nanoscale shapes may be related to the histone modifications, which opens a new window for the investigation of nanoscale shape effects and nano therapy. In chapter IV, we proposed a method to modify the surface of the nanostructures by endogenous cellular processes and studies found that this re-engineered particle complex was able to transfer the loading genes to recipient cells, which indicates their potential to work as an efficient nucleic acid delivery machine.

Statement of Original Authorship

I hereby certify that the submitted work is my own work, was completed while registered as a candidate for the degree stated on the Title Page, and I have not obtained a degree elsewhere on the basis of the research presented in this submitted work.

All the data included in this thesis is my own, with the exception of the following collaborations.

Chapter II:

- The transcriptomics was done in cooperation with Dr. Jingji Li and Dr. Valeria A. Giannone.
- The *in vivo* part was done by André Perez-Potti, Stéphanie Devineau, Nicolò Panini, Sandra Lara Martinez, Valeria A. Giannone, Paolo Bigini, Mario Salmona, Laura Talamini, Martina B. Violatto and Giuseppe Ristagno.
- The computational shape analysis was done in cooperation with Dr. Hender Lopez and Dr. Joao M. de Araujo.

Chapter III:

- The THP-1 cell line related experiments including transcriptomics, proteomics, western blot, ELISA and TEM was done in cooperation with Dr. Jingji Li.
- Dr. Qi Cai and Dr. Camila P. Silveira supplied some of the nanoparticles used in the THP-1 cell line study.

Chapter IV:

- The magnetic silica nanoparticles synthesis was done by Dr. Laurent Adumeau, Dr. Lorenzo Cursi and Mr. Zengchun Xie. TEM and SEM images were done by Dr. Vanya Petseva and Dr. Koen Evers respectively.
- Most of the confocal images were captured by Ms. Silvia Vercellino, some were imaged by Mr. Xiaoliang Yang and the left part of the confocal images were done in cooperation with Dr. Yuchen Lin and Ms. Ying Ling Dee.
- The proteomics was done by Ms. Cara Gaffney, Dr. Mura McCafferty, Mr. Xiaoliang Yang and Dr. Linlin Song.
- The RNA extraction for RNA-seq was cooperated with Dr. Mura McCafferty.

Wei Zhang

December 2021

College of Science Guidelines – Collection of Papers for PhD Thesis

Declaration of co-authorship

Name: Wei Zhang	
Student Number: 16211384	E-mail: wei.zhang1@ucdconnect.ie
School: School of Chemistry	
Principal supervisor: Kenneth A. Dawson	E-mail: kenneth.a.dawson@cbni.ucd.ie
Research Degree for which thesis is being submitted: PhD	
Title of Research thesis: Complex Nanostructures and Bio-nanoscale Interactions: Well Defined Synthesis, Identification and Biological Effects	

This co-authorship declaration applies to the following paper: A nanoscale shape-discovery framework supporting systematic investigations of shape-dependent biological effects and immunomodulation

The student's contribution to the paper

(Outline briefly the student's intellectual and practical contribution to the research presented in the paper. A percentage figure only will not suffice). This should be separated into experimental and theoretical basis.

Experimental Part: Wei advanced the microfluidic synthesis platform and was central to the setup of the tunable shape library synthesis and variable surface coatings.
Theoretical Part: Wei and Hender drove the calculations for the paper, Wei and Zengchun performed the computational shape analysis for NPs used in the paper. Wei drew all of the figures in the paper and wrote the paper.

Other Authors contribution to the paper

(Outline briefly the contribution of other authors to the paper)

Author	Contribution
Hender Lopez	Drove the calculations for the paper with Wei and worked out the code for the computational shape analysis with Joao.
Luca Boselli, Qi Cai, Valentina Castagnola	Initiated the microfluidic setup, synthesized and characterized the nanoparticles used for biological investigations cooperation with Camila P. Silveira.
Paolo Bigini, Mario Salmona	Led the first efforts *in vivo* set up.
André Perez-Potti	Performed the ELISA with Stéphanie Devineau and B cell receptor repertoire analysis with Valeria A. Giannone.
Zengchun Xie	Carried out much of the current microfluidic synthesis, characterization and contour analysis.
Laura Talamini, Giuseppe Ristagno, Martina B. Violatto	Performed the *in vivo* immunization and tissue collection.
Nicolò Panini	Was responsible for FACS analysis for cell sorting from spleens of rats.

PhD student: Wei Zhang **Date:** 01/12/2021

Signature (dd/mm/yy)

College of Science Guidelines – Collection of Papers for PhD Thesis

Declaration of co-authorship

Name: Wei Zhang	
Student Number: 16211384	E-mail: wei.zhang1@ucdconnect.ie
School: School of Chemistry	
Principal supervisor: Kenneth A. Dawson	E-mail: kenneth.a.dawson@cbni.ucd.ie
Research Degree for which thesis is being submitted: PhD	
Title of Research thesis: Complex Nanostructures and Bio-nanoscale Interactions: Well Defined Synthesis, Identification and Biological Effects	

This co-authorship declaration applies to the following paper: Shape-Dependent Histone Modifications Induced by Gold Nanoparticles

The student's contribution to the paper

(Outline briefly the student's intellectual and practical contribution to the research presented in the paper. A percentage figure only will not suffice). This should be separated into experimental and theoretical basis.

Experimental Part: Wei did the synthesis and characterization of the gold nanoparticles and led the A549 cell line related histone modifications experiments including cytotoxicity, western blot, ChIP-seq and data analysis. Theoretical Part: Wei did the computational shape analysis of the gold nanoparticles and performed the data analysis of ChIP-seq. Wei drew all of the panels in figure 1 and figure 4 and wrote the paper.

Other Authors contribution to the paper

(Outline briefly the contribution of other authors to the paper)

Author	Contribution
Jingji Li	Drove the THP-1 cell line related experiments.
Qi Cai	Supplied some of the GNP1 used in THP-1 cell line study.
Camila P. Silveira	Supplied some of the GNP2 used in THP-1 cell line study.
Valeria A. Giannone	Assisted Jingji Li for the THP-1 cell line related experiments.

PhD student: Wei Zhang **Date:** 01/12/2021

Signature (dd/mm/yy)

College of Science Guidelines – Collection of Papers for PhD Thesis

Declaration of co-authorship

Name: Wei Zhang	
Student Number: 16211384	E-mail: wei.zhang1@ucdconnect.ie
School: School of Chemistry	
Principal supervisor: Kenneth A. Dawson	E-mail: kenneth.a.dawson@cbni.ucd.ie
Research Degree for which thesis is being submitted: PhD	
Title of Research thesis: Complex Nanostructures and Bio-nanoscale Interactions: Well Defined Synthesis, Identification and Biological Effects	

This co-authorship declaration applies to the following paper: Horizontal cellular transfer of biomolecular condensates

The student's contribution to the paper

1. Wei was central to the RNA-seq related experiment and also lead the setup of RNA-seq data analysis platform, all the RNA-seq data analysis related figures (Figure 4.2 f, g, 4.2 h provided analyzed data, 4.4 h, i left, S4.20, S4.21, S4.22 provided analyzed data, S4.37, S4.38 left, S4.40) were done by Wei.
2. Wei investigated the RNA delivery kinetics of particle complexes to different recipient cells. Wei led the validation experiment –RT-qPCR related experiment, and all the qPCR related figures were from Wei (Figure 4.4 i right, j, k, S4.38 right, S4.39).
3. Wei supported the proteomics data analysis and contributed as an independent replicate for part of the proteomics data analysis. (Figure 4.2 a, b, e, h, S4.22, S4.34, S4.35)
4. Wei contributed to the splitting events observation in confocal microscope by cooperation with Ying Ling and Yuchen. (Figure 4.3 c)
5. Wei contributed to TEM imaging as an independent replicate for the images. (Replicate of Figure 4.1 g)

Other Authors contribution to the paper

(Outline briefly the contribution of other authors to the paper)

Author	Contribution
Silvia Vercellino	Contributed to most of the confocal imaging experiment and particle complex production protocol set up.
Mura McCafferty	Drove the RNA-seq experiment together with Wei, contributed to the set up of particle complex production protocol.
Xiaoliang Yang	Contributed to the proteomics data analysis, part of the mRNA-seq data analysis and part of the confocal imaging.
Laurent Adumeau, Lorenzo Cursi, Zengchun Xie	Drove the nanoparticles synthesis and characterization.
YingLing Dee	Contributed to part of the confocal imaging.
Cara Gaffney, Mura McCafferty, Linlin Song	Contributed to the sample preparation of proteomics.
Vanya Petseva, Koen Evers	Contributed to the TEM (Vanya) and SEM, confocal imaging (Koen) of the particle complex.
LinLin Song, Yuchen Lin	Established the plasmids and stable cell line.

PhD student: Wei Zhang **Date:** 01/12/2021

Signature (dd/mm/yy)

Acknowledgement

First of all, I would like to express my deepest appreciation to my supervisor Prof. Kenneth A. Dawson and co-supervisor Prof. Yan Yan. During the past four years, Prof. Kenneth and Prof. Yan not only guided me to grow up to a real researcher in science but also cared about my life, my personal career development, which makes me feel they are not only my supervisors but someone who I can trust and ask for help. Their true love in science and precious supervision developed my passion in science. From them I was able to know what's a real scientist and how he/she performs the scientific research. Their way of dealing with difficulties, supervising students also inspired me to think about what kind of supervisor I want to become and how to grow up to the one I want to be.

Many thanks to CBNI professional admin team for the valuable help and support. I would like to extend my sincere thanks to Rhona, James, Martin and Ruth, you all played an important role for my personal growth in the past four years. Rhona always helped me with the organization of my exams, from DSP to transfer to third year talk and finally thesis submission, viva examination and your kind concern whether I was loading too much really made me feel cured. James, thanks for your open-handed sharing your experience to me which greatly helped and inspired me how to deal with the management of people and my project, and also for those overnight experiment, thanks James for those hard days! I cannot imagine how I can get my grant application done so efficiently and smoothly without your professional guidance Martin, really appreciated for your contribution for suggestions on my grant application! Ruth, thank you so much for your outstanding management for all the meetings and discussions, our RNA-seq and ChIP-seq analysis platform won't be established without your help. Appreciation also goes to our CBNI outstanding researchers, it's a great honor to be able to work with the most advanced techniques, collaborate with excellent researchers and learn the meaning of "cutting edge of science" in CBNI. For the unforgettable four-years' experience in CBNI, thanks to Dr. Qi, Dr. Valentina, Dr. Luca, Dr. Jingji and Dr. Kai who trained me in the first year of my PhD. Thanks to Ms. Emily for the biological background introduction in my second year and also many good funs you brought us; Appreciation to Dr. Mura for the third year bioanalyzer training and being my partner for those overnight experiment; Gratitude to Mr. Xiaoliang for the patient training of proteomics data analysis. Thanks to Ms. Ying Ling for lovely chatting and treasurable emotional support; To Qiwei and Wanting, thanks for all the delicious food you shared and recommended, it's been a pleasure to live with you two under the same roof.

Special thanks to Prof. Bing Yan who brought me to the bio-nano world, introduced me to CBNI and supported me for my postdoc application. Words cannot express my appreciation, as a light shining in my scientific research road, hope one day I can grow up to an independent PI, a real scientist and thank you!

I also extend my appreciation to Tiina O'Neill, Dr. Niamh Stephens, Dr. Dimitri Scholz, Dr. Alfonso, Philippa Kavanagh and Alison Murphy for the professional technical support in Conway lab and Conway core tech center.

Great thanks to Professor Matthias Wilm, Assoc Professor Gerard Cagney and Professor Epple for reviewing my thesis and I am looking forward to meeting you all in my viva.

I gratefully acknowledge the financial support from Chinese Scholarship Council (CSC) and High Education Authority (HEA).

To my forever best friends, my family members, Zengchun, Peng and Liting, how lucky I am to have you accompany for all the happy or sad moments in Ireland, especially for those dark days that makes me glimpsed the refined cruelty of the world. Thanks for your company and give me the courage to insist those truth and kindness, to look up for my moon instead of just accepting or joining the majority without thinking.

Most of all, I thank my parents, simply for everything. The thesis is dedicated to you.

List of Publications

1. Luca Boselli, Hender Lopez, Wei Zhang, Qi Cai, Valeria A. Giannone, Jingji Li, Alirio Moura, João M. de Araujo, Jennifer Cookman, Valentina Castagnola, Yan Yan & Kenneth A. Dawson. **"Classification and biological identity of complex nano shapes".** Commun Mater 1, 35 (2020).
2. Qi Cai, Valentina Castagnola, Luca Boselli, Alirio Moura, Hender Lopez, Wei Zhang, João M. de Araujo, Kenneth A. Dawson. **"A microfluidic approach for synthesis and kinetic profiling of branched gold nanostructure".** Nanoscale Horiz., 2021, DOI: 10.1039/d1nh00540e.
3. Wei Zhang, Hender Lopez, Luca Boselli, Paolo Bigini, André Perez-Potti, Zengchun Xie, Valentina Castagnola, Qi Cai, Camila P. Silveira, Joao M. de Araujo, Laura Talamini, Nicolò Panini, Giuseppe Ristagno, Martina B. Violatto, Stéphanie Devineau, Marco Monopoli, Mario Salmona, Valeria A. Giannone, Sandra Lara, Kenneth A. Dawson & Yan Yan**. "A Nanoscale Shape-Discovery Framework Supporting Systematic Investigations of Shape-Dependent Biological Effects and Immunomodulation".** ACS Nano 2022, 16, 1, 1547–1559.
4. Wei Zhang, Jingji Li, Qi Cai, Camila P. Silveira, Valeria A. Giannone, Kenneth A. Dawson & Yan Yan. **"Shape-Dependent Histone Modifications Induced by Gold Nanoparticles".** Manuscript under reviewing.
5. Wei Zhang, Laurent Adumeau, Mura McCafferty, Silvia Vercellino, Xiaoliang Yang, Cara Gaffney, Ying Ling Dee, Linlin Song, Yuchen Lin, Koen Evers, Lorenzo Cursi, Vanya Petseva, Zengchun Xie, Aisling Fleming, Emily Sheridan, Ingrid Morera, Yan Yan & Kenneth A. Dawson. **"Horizontal cellular transfer of biomolecular condensates".** Submitted to Nature.

The work presented in the thesis is based on the results included from the above scientific articles and manuscripts. Chapter II, III, IV are based on the results in manuscripts 3, 4 and 5 respectively. Paper 1 is included as a reference in Appendix III.

Abbreviation

NPs: nanoparticles

NNI: National Nanotechnology Initiative

LSPR: localized surface plasmon resonance

GNPs: gold nanoparticles

$mSiO_2$: magnetic SiO_2

SAXS: small-angle X-ray scattering spectroscopy

XANES: X-ray absorption near edge spectroscopy

BSPP: bis(p-sulfonatophenyl) phenylphosphine dihydrate dipotassium salt

AA: ascorbic acid

CTAB: cetyltrimethylammonium bromide

HQ: hydroquinone

Na_3Cit: trisodium citrate dihydrate

PTFE: polytetrafluoroethylene

DMF: digital microfluidic reactor

LPS: lipopolysaccharides

FDA: Food and Drug Administration

LAL: limulus amebocyte lysate

pNA: p-nitroaniline

DCS: differential centrifugal sedimentation

DLS: dynamic light scatterting

PDI: polydispersity index

TEM: transmission electron microscopy

NTA: nanoparticle trafficking analysis

UV-vis-NIR: ultraviolet visible near infrared spectroscopy

ICP-MS: inductively coupled plasma mass spectroscopy

PCA: principal component analysis

DFT: Discrete Fourier Transform

MR: microfluidic reactor

TR: tank reactor

BSA: bovine serum albumin

HAS: human serum albumin

OVA: albumin from chicken egg white

BCA: bicinchoninic acid

WNA: West Nile Virus

ApoB: Apolipoprotein Eb

APOE: Apolipoprotein E

µg: microgram

mg: milligram

nm: nanometre

min: minute

h: hour

µl: microliter

ml: millilitre

l: litre

EU/ml: endotoxin unit per millilitre

MEM: minimum essential medium

cMEM: minimum essential medium supplemented with 10% foetal bovine serum and 1% penicillin

FBS: fetal bovine serum

HS: human serum

PBS: phosphate-buffered saline

HEPES: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

JAWSII: mouse dendritic cells

THP-1: monocytic leukaemia cell line

A549: adenocarcinomic human alveolar basal epithelial cells

ELISA: Enzyme-linked immunoassay

BCR: B cell receptor

ChIP-seq: chromatin immunoprecipitation sequencing

RNA-seq: RNA sequencing

DEGs: differential expression genes

GO: gene ontology

DNA: deoxyribonucleic acid

RNA: ribonucleic acid

FSC: forward scatter

SSC: side scatter

H3K27me3: histone H3-lys27 trimethylation

HCA: high content analysis

SRRF: super-resolution radial fluctuations

MTS: CellTiter 96® AQueous One Solution Cell Proliferation

LDH: CytoTox 96® Non-Radioactive Cytotoxicity

TMRM: tetramethyl rhodamine methyl ester

MW: molecular weight

PMT: post-translational modification

RPMI 1640: Roswell Park Memorial Institute 1640

DMEM: Dulbecco's Modified Eagle Medium

RT: room temperature

EGFP: enhanced green fluorescent protein

EF: exosome free

DTT: dithiothreitol

CMO: cell mask orange

P: pellet

SN: supernatant

EV: extracellular vesicles

ER: endoplasmic reticulum

SDS-PAGE: Sodium dodecyl-sulfate polyacrylamide gel electrophoresis

List of Figures

List of Tables

Chapter I

Introduction and Thesis Motivation

1 Introduction

1.1 Introduction to nanoparticles (NPs)

The first nanotechnology concept was laid out by Richard P. Feynman in the famous speech "There's plenty of Room at the Bottom" in 1959. Since then, various developments have been made in the field of nanotechnology and advances in nanotechnology have made it currently a central tool in biomedical research nowadays. According to the National Nanotechnology Initiative (NNI), nanotechnology is defined as research and technology development at the atomic, molecular or macromolecular scale, leading to the controlled creation and use of structures, devices and systems with a length scale of 1-100 nanometres (nm).[1] Nanoparticles (NPs) are those materials which have at least one dimension smaller than 100 nm.

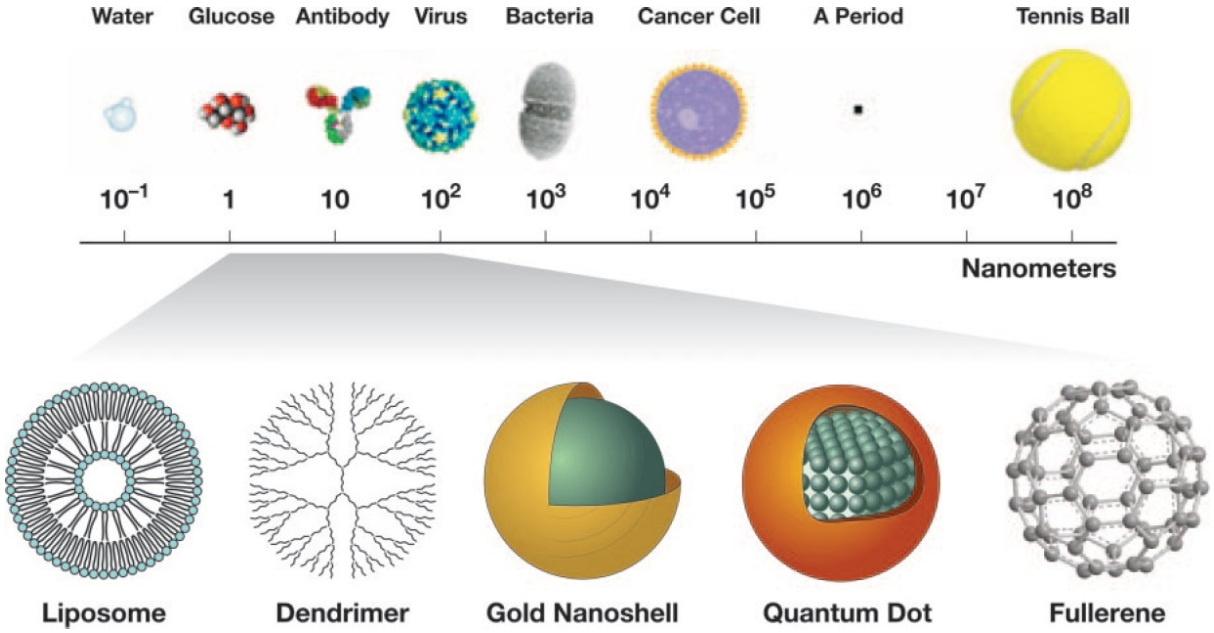


Figure 1. 1 The definition of nanoscale comparing with familiar items.[1]

1.1.1 Properties of NPs

Compared with bulk materials or atoms, NPs show unique physical and chemical properties, there are two main factors that make NPs show significant differences from other materials.

a. Surface effects: due to their small size, NPs have high surface-to-volume ratio, and therefore the number of atoms on the surface of particles are greatly increased, thus causing the surface of the NPs to be very active.

b. Quantum effects: due to the quantum confinement effects in NPs with delocalized electrons[2], they display discontinuous behaviour.

These factors onward affect the chemical reactivity of NPs, as well as their mechanical, optical, electric and magnetic properties. For example, colloidal gold NPs (GNPs) exhibit distinctive size and shape dependent colours (Figure 1.2).

Figure 1. 2 Colours dependence of GNPs on size and shape.[3]

This interaction with light is due to localized surface plasmon resonance (LSPR). Surface plasmon resonance is an optical phenomenon that causes a collective oscillation of valence electrons and subsequent absorption within the ultraviolet-visible-near infrared (UV-Vis-NIR) band, due to the interactions between the incident photons and the conductions band of a noble metal surface. In the case of noble metal NPs, this induced plasmon oscillation occurs locally to the nanostructure, which is called LSPR, as shown in Figure 1.3.[4] The properties of LSPR are highly dependent on the material used and the size and shape of the metallic NPs (more details will be discussed in 1.3 characterization section). In noble metal NPs, both the absorption and scattering efficiencies are strongly enhanced due to the LSPR. This makes these NPs very promising in a wide range of applications including sensing, imaging and photothermal therapy. For example, the application of GNPs in thermal therapy to harm bacteria,[5] viruses,[6] and cancer cells[7] based on their heating effects under laser irradiation has been demonstrated, which is due to the enhanced absorption induced by LSPR.

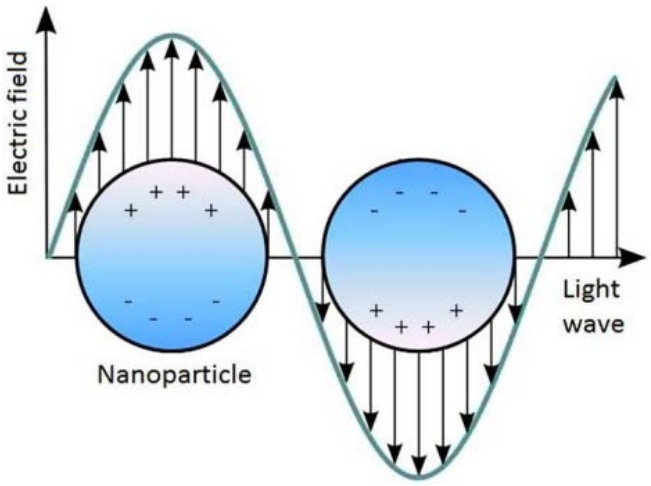


Figure 1. 3 Schematic of LSPR. [4]

1.1.2 Applications of NPs in nanomedicine area

During the past thirty years, various NPs are synthesised and designed for cellular biology and medical research and applications, especially in the rapidly developing field of nanomedicine.[8-9] (Figure 1.4) For example, various kinds of fluorescent NPs have been widely used in biological imaging, ranging from fixed cells to living cells and model organisms.[10] NPs also provide new tools for drug delivery and gene therapy due to their unique properties.[11]

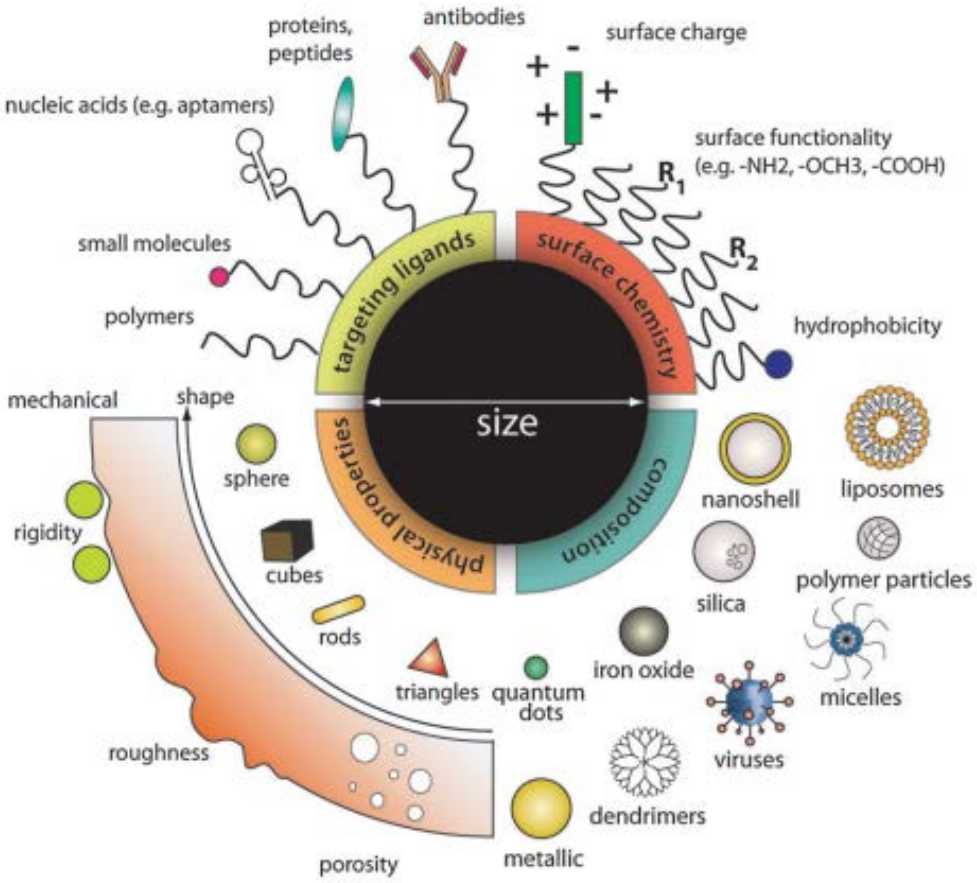


Figure 1. 4 Designing NPs for biological application.[9]

Among these NPs, GNPs with specific optical properties which can be tuned based on their size and shape[12] are one of the most prominent materials.[13-15] Magnetic iron oxide NPs with unique magnetic properties such as superparamagnetic, high magnetic susceptibility *etc.* also are widely used in biomedicine applications including magnetic bio separation, clinic diagnosis and therapy (Figure 1.5).[16-17]

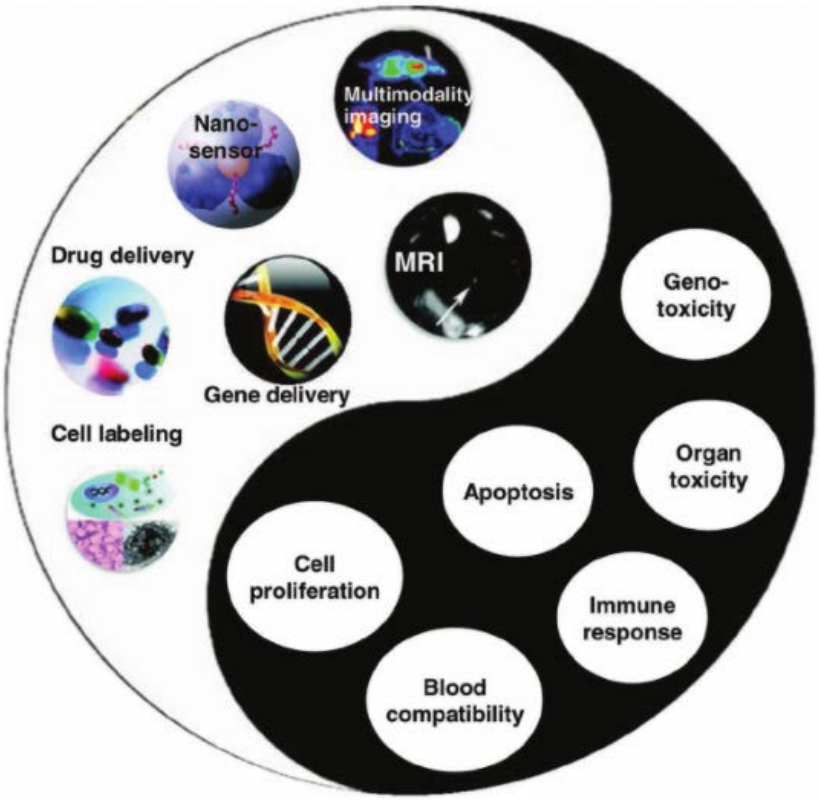


Figure 1. 5 Biomedical applications and potential toxicity of magnetic iron oxide NPs.[16]

No matter which NPs are used for the research or the nanomedicine application, the synthesis of good quality NPs is the primary step, this will be introduced in the following section.

1.2 Synthesis of NPs

1.2.1 Nucleation and growth—mechanism of the NPs formation

It has been widely accepted that nucleation and growth are the key processes in the evolution of NPs. Nucleation usually is the first step of NPs formation whereby nuclei or seeds act as template for the furthering growth of NPs.[18] There are two types of nucleation, homogenous and heterogeneous nucleation. Homogenous nucleation occurs spontaneously and randomly without the influence of foreign particles while heterogeneous nucleation requires the presence of alien particles to aid the nucleation. Classical homogenous nucleation mechanism[18-19] consider the process as a thermodynamic process determined by the total free energy (ΔG) which is defined as the sum of the surface free energy (γ) and bulk crystal free energy (ΔG_υ), as shown in equation 1.1, where r is the radius of the particle. The crystal free energy ΔG_υ is dependent on the temperature T, Boltzmann's constant k_B, the supersaturation of the solution S and its molar volume υ as shown in equation 1.2.

$$\Delta G = 4\pi r^2 \gamma + \frac{4}{3}\pi r^3 \Delta G_\upsilon \quad \text{(eq 1.1)}$$

$$\Delta G_\upsilon = \frac{-k_B T ln(S)}{\upsilon} \quad \text{(eq 1.2)}$$

According to the above equations it is possible to find a maximum ΔG where $d\Delta G/dr = 0$, and this critical radius r_{crit} corresponds to the minimum size at which a particle can survive in solution without being redissolved, Figure 1.6.

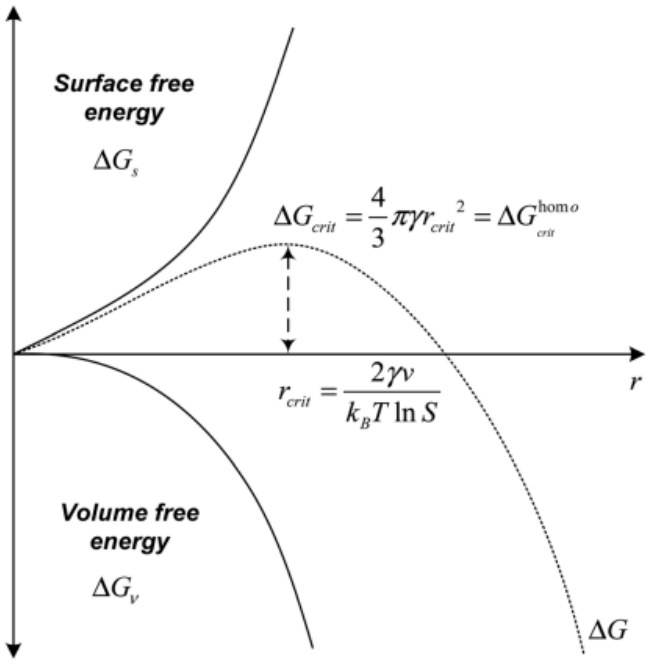


Figure 1. 6 Schematic illustration of the free energy diagram for nucleation.[20]

To better explain the correlation between the nucleation and growth, in 1950 LaMer *et al.*[21] proposed another model to explain how monodispersed particles are obtained in a homogenous solution. Briefly, the LaMer theory can be divided into three steps as shown in Figure 1.7: (I) prenucleation stage, a rapid increase in the concentration of free monomers in solution, (II) nucleation stage, the monomer undergoes “burst-nucleation” which significantly reduces the concentration of free monomers in solution. The rate of this nucleation is described as “effectively infinite” and after this point, there is almost no nucleation occurring due to the concentration of monomers is lower than the critical minimum concentration for nucleation; (III) growth stage which occurs under the control of the diffusion of the monomers through the solution.

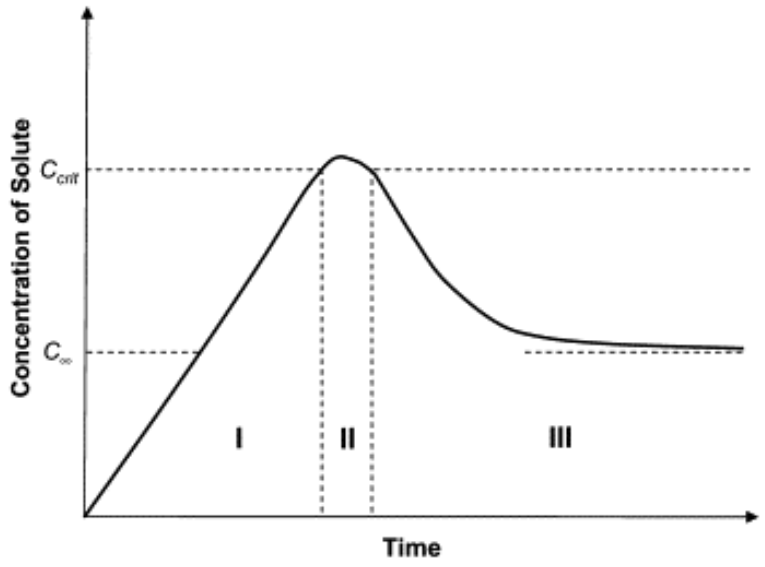


Figure 1. 7 Schematic illustration of the LaMer theory.

C_{∞} and C_{crit} are the equilibrium concentration of solute with the bulk solid and the critical concentration as the minimum for nucleation, respectively. [22]

In addition to the LaMer theory, many other mechanisms for the growth process have been put forward such as Ostwald ripening[23], orientated attachment[24-25], intraparticle growth[26-27] *etc.* In addition to the above model/theories, another point to understand about the NPs evolution process is through the thermodynamic versus kinetics control[28-30] perspective. Figure 1.8 shows a schematic illustration of the concept of thermodynamic versus kinetics control. In this case, shape II is the predominance product when the synthesis undergoes thermodynamic control because shape II is more thermodynamically stable than shape I. Nevertheless, the activation energy barrier (Ea) involved in the formation of shape I is lower than that for the formation of shape II, so shape I is expected to be in the majority when the synthesis is conducted under kinetic control. During the nanocrystals synthesis process it is possible to get a series of shapes/products sharing the same elements, for such case, Figure 1.8 (B) is more suitable to illustrate the concept of thermodynamic versus kinetics control. During this process, the shape with the global minimum Gibbs free energy will be the dominant product for thermodynamic control, while for kinetically controlled synthesis, shapes corresponding to local minimum Gibbs free energy will be trapped.

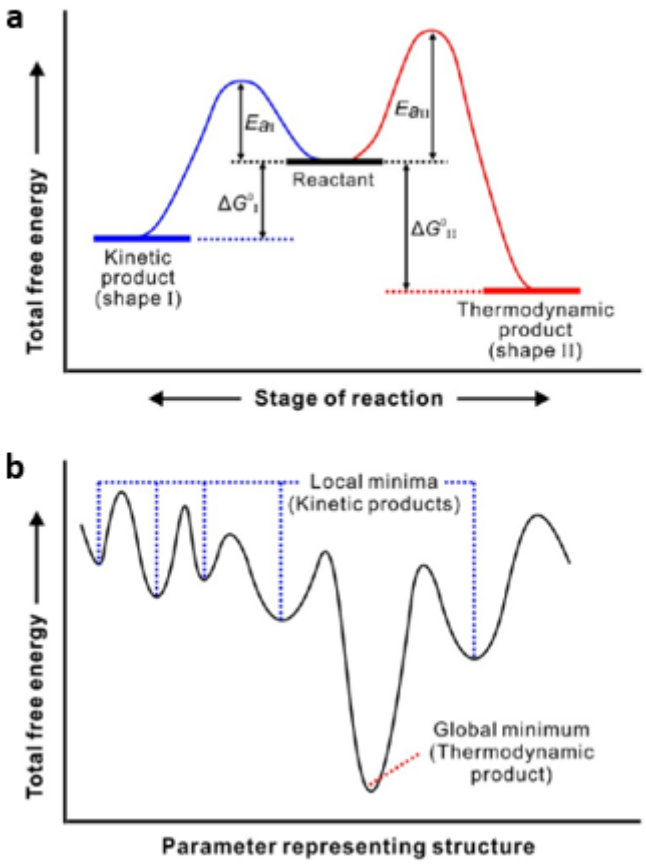


Figure 1. 8 Schematic illustration of two different scenarios (a and b) of thermodynamic versus kinetic control.[30]

After a broad view of the nucleation and growth theory development, now we move to the mechanism development specifically for GNPs formation. Since 1857 Faraday successfully synthesized gold colloids, numerous studies have been carried out to reveal the evolution mechanisms for GNPs. In 1951 Turkevich did the first trial of studying the process of nucleation and growth in gold colloids by using electron microscope[31], a half century later researchers are able to track the formation process more precisely by using UV-visible spectroscopy, X-ray spectroscopy and many other advanced technologies. Ji *et al.*[32] described a pH dependent GNP formation pathway as shown in Figure 1.9. For a low pH (3.7-6.5), the formation of final GNPs includes three steps: a LaMer nucleation[22] within 10 s, a fast random attachment, and finally the intra-particle ripening. For a higher pH (6.5-7.7), the formation starting from a reduction and then a longer nucleation (60 s) and finally a slow growth to obtain the particles. By the combination of small-angle X-ray scattering (SAXS), X-ray absorption near edge spectroscopy (XANES) and UV-visible spectroscopy analysis Polte *et al.*[33] proposed a more detailed sequential process on the formation of the GNPs *via* citrate methods (Figure 1.10). Briefly, the four-step mechanism comprising a fast reduction and nucleation, growth of nuclei *via* coalescence, diffusion- controlled growth of particles by ongoing reduction

of Au (III) and finally fast consumption of Au (III) to achieve bigger particles. Further studies carried out by Polte[34], Yao[35], Mikhlin[36] *et al.* confirmed Polte's observation and further developed the mechanism of the evolution of spherical GNPs (Figure 1.11).

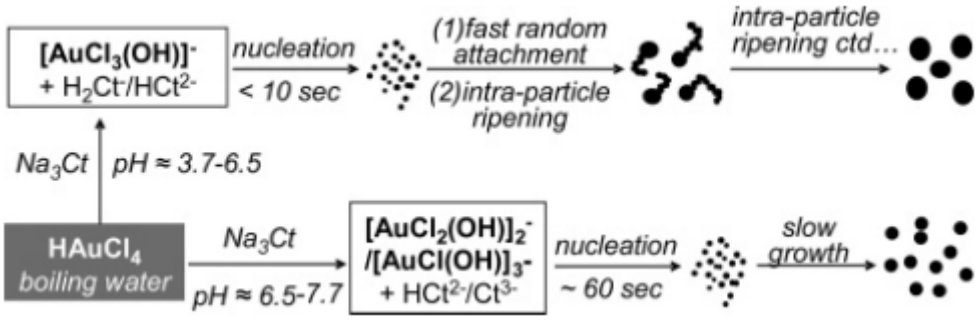


Figure 1. 9 Schematic illustration of the pH dependent GNP formation pathways. [32]

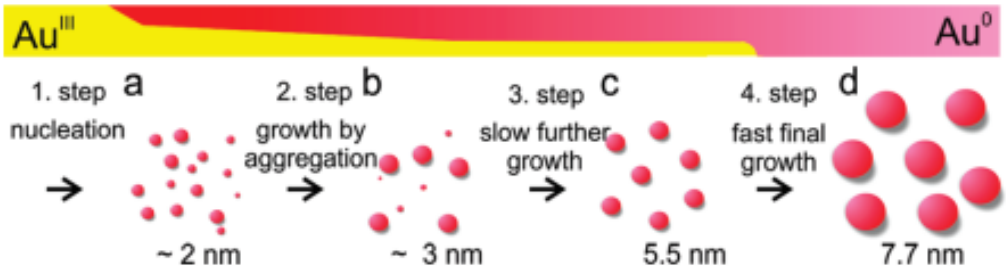


Figure 1. 10 Schematic illustration for the deduced process of GNP formation.[33]

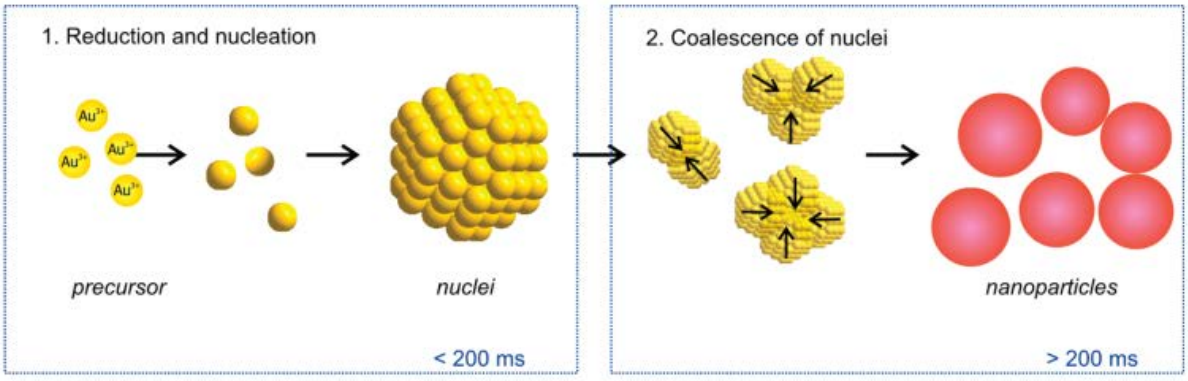


Figure 1. 11 Schematic illustration for the deduced process of GNP formation. [34]

Besides spherical GNPs, the synthesis and related evolution mechanism of GNPs with well-defined shapes is another popular subject of research from the beginning of 21st century. The same as spherical GNPs formation, shape-controlled GNPs synthesis can be carried out *via* either homogenous or heterogenous nucleation. Heterogenous nucleation is conducted by adding pre-synthesized seeds to a reactant mixture, effectively isolating GNPs nucleation and growth as separate steps.[37] This method effectively benefits the shape formation because it

allows thoughtful design of shapes through the choice of different seeds[38-41]. In addition to choosing different seeds, shape control can be achieved by employing capping agents that selectively absorb to specific crystal planes. The general strategy to generate shape anisotropy during nanocrystal growth is by stabilizing a particular facet through this molecular interaction; growth is limited on the crystal plane where binding is strong and promoted on the crystal plane where binding is weak[38, 42-44]. Figure 1.12 is an example of shape-controlled synthesis of GNPs in the presence of capping agents.

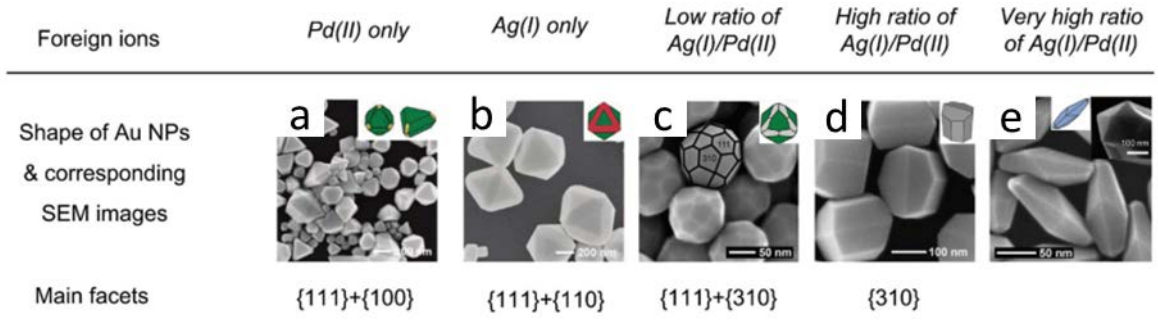


Figure 1. 12 (a)-(e) Morphology evolution of Au nanocrystals synthesized in the presence of Ag(I) and Pd(II) in a mediated polyol process.[45]

Based on the above nucleation and growth mechanisms, various methods are proposed to obtain GNPs with different sizes, shapes and coatings. Details will be discussed in the following sections.

1.2.2 Benchtop tank reactor-based synthesis—state of the art

After the first report by Faraday in 1857 on the reduction of a tetrachloroaurate solution by phosphorus in carbon disulfide (a biphasic reaction), the preparation of GNPs with varied sizes and shapes has raised increased attention. Since 20th century, many different approaches have been developed to generate GNPs. There are two strategies to obtain GNPs on the nanoscale: top down method where material is removed from the bulk material, leaving only the desired nanostructures, and bottom up method, where the atoms produced from reduction of ions are assembled to generate nanostructures.[18] In the bottom up strategy, the formation of GNPs originates from individual molecules, which involves a chemical reduction of a gold precursor (normally a coordination compound presenting the gold in its oxidation state (+3 or +1). As mentioned above, the chemical reduction method includes two steps: nucleation and successive growth. When the nucleation and successive growth are completed in the same process, it is called in situ(one-pot) synthesis; otherwise it is called seed-mediated method.[46]

I. Synthesis of spherical GNPs

Up to now various spherical GNPs synthesis methods have been proposed, for spherical GNPs with a small size, such as size below 20 nm, one-pot synthesis is widely used. For example, 10-20 nm spherical GNPs can be synthesized *via* the modified Turkevich-Frens method[47]. For spherical GNPs larger than 20 nm, the seed-mediated method is more versatile, because by separating the nucleation and growth processes, it can be easier to control the monodispersity of GNPs. Also, more moderate reagents can be used, greatly increasing the synthesis possibility of GNPs. For example, 50 nm spherical GNPs can be synthesized *via* a two-step seeded mediated method, in which 15 nm spherical GNPs are pre-synthesized as seeds, and then citrate sodium is used as reducing agent to help the GNPs grow to a larger size.[48]

II. Synthesis of branched GNPs

Compared to spherical GNPs, branched GNPs show more advanced properties and potential for nanomedicine applications. For the branched GNPs synthesis, in 2003 Chen *et al.*[49] firstly proposed a room temperature method for the synthesis of a mixture of branched GNPs by adding ascorbic acid (AA) to Au-cetyltrimethylammonium bromide (CTAB) solution, Hao *et al.*[50] reported another room temperature method by adding $HAuCl_4$ to bis-(p-sulfonatophenyl) phenylphosphine dihydrate dipotassium (BSPP), H_2O_2 and sodium citrate solution. After that Sau *et al.*[51] put forward a seeds mediated growth for branched GNPs by varying the quantities of seeds and growth solution (containing desired quantities of CTAB, $HAuCl_4$, AA and in some cases a small amount of $AgNO_3$) (Figure 1.13). Since then more protocols[43, 52-54] were presented for various branched GNPs. In 2011 Li *et al.*[55] successfully used hydroquinone (HQ) to tune the activity of $HAuCl_4$ to obtain stable “urchin-like” GNPs and in the past decade researchers moved further on the optimization of the synthesis methods to get more homogenous distributed branched GNPs.[38, 42, 56-58]

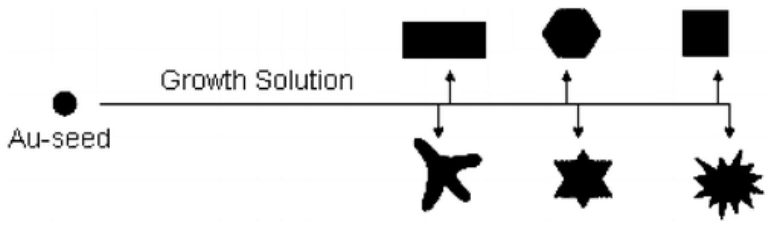


Figure 1. 13 Schematic illustration of seeds mediated synthesis of nanoshapes.[51]

Although traditional methods for the synthesized of GNPs are widely used, there are some drawbacks to the approach. For example, due to a lack of effective mixing in a bulk volume and low control over the experimental conditions, the traditional synthesis can cause a high variability in the GNPs size and shape distribution. However, GNPs with high batch-to-batch reproducibility and high mono dispersity in terms of size and shape distribution are of vital

importance for biological investigations. Therefore, well-controlled synthetic methods for homogenous distribution and high reproducibility of NPs is important to facilitate their biological investigations and applications.

1.2.3 Microfluidic reactor-based synthesis—a novel developing approach

Microfluidic refers to the process or manipulation of fluids in channels with dimensions constrained to sub-milimeters.[59-60] Since microfluidic reactors were first reported in 1992[61], microfluidics as a technology has shown compelling benefits that attracts researchers' interests from various arenas. The rapid development of microfluidics makes it widely applied in the pharmaceutical[62], biotechnology[63] and chemical industries[59, 64]. Here we mainly focused on the capability of a microfluidic reactor as an advanced approach for the formation of NPs.

I. Components of a microfluidic synthesis system

The key component of a microfluidic synthesis system is the microfluidic reactor, which can be made from polymers[65], glass[66], silicon[67], metal[68] and other materials, Figure 1.14 shows a schematic of a microfluidic reactor[69]. Although a significant proportion of microfluidic reactors research has been performed in academia, industry has also taken an inquisitive interest in this technology[70]. Commercially available microfluidic reactors make microfluidic synthesis more achievable while releasing researchers from the need to design microfluidic reactors.

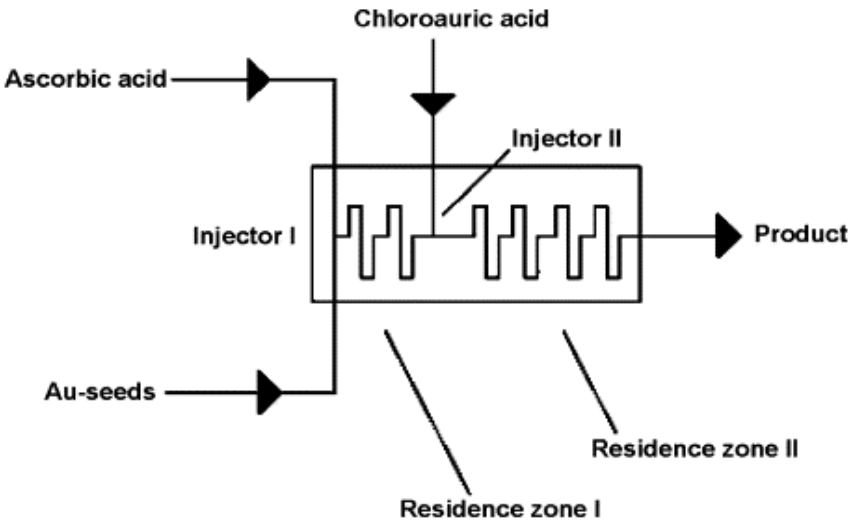


Figure 1. 14 Schematic drawing of microfluidic reactor for GNPs synthesis.[69]

In the past twenty years, a diverse range of microfluidic reactors have been designed for the synthesis of various materials. Generally microfluidic reactors can be classified into three categories: continuous-flow[71-73], droplet-based[74] and digital microfluidics[75].

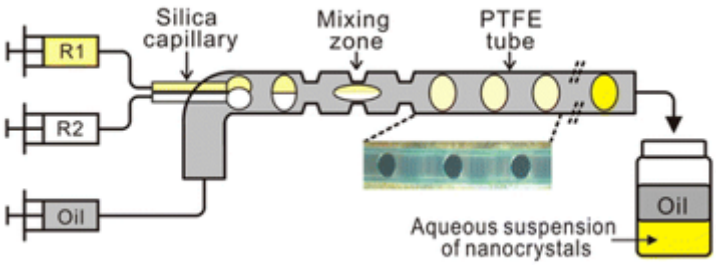
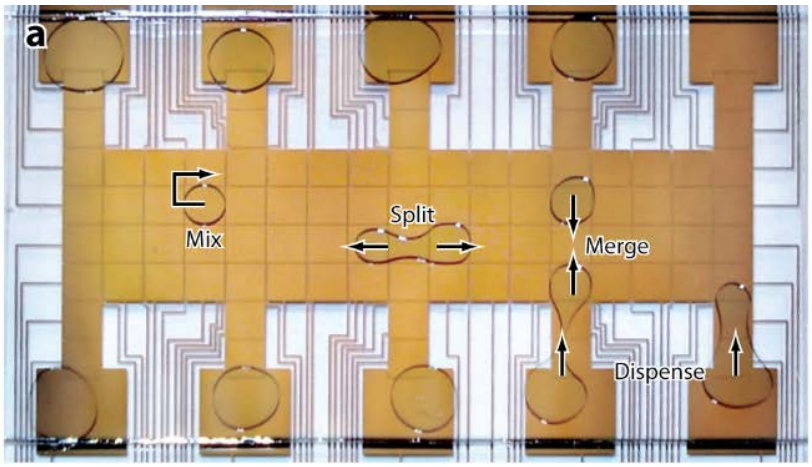


Figure 1. 15 Schematic drawing of droplet-based microfluidic reactor for the generation of aqueous droplets separated by oil, followed by the mixing and formation of NPs in the droplets. Inset: photograph of droplets containing blue dye in polytetrafluoroethylene (PTFE) tube.[74]



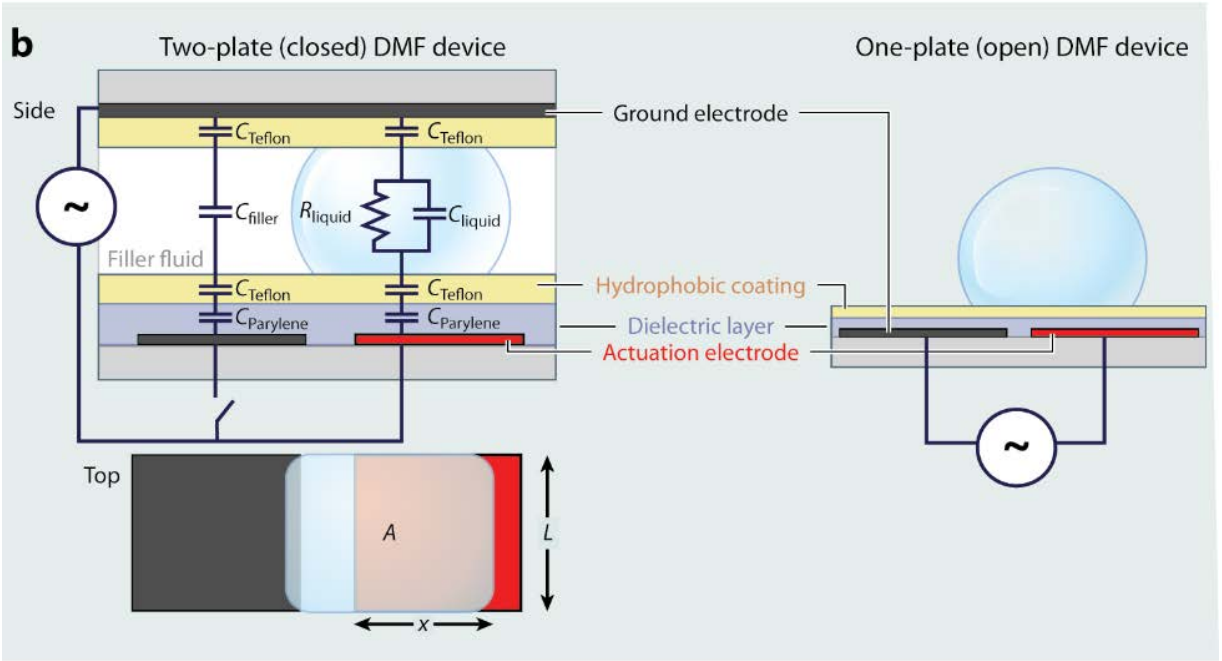


Figure 1. 16 Schematic drawing of digital microfluidic reactor (DMF). (a) Droplet operations in DMF, including mixing, splitting, merging and dispensing from reservoirs. (b) Side view

schematics of two-plate and one-plate DMF devices. A top-view schematic of the two-plate device is shown below the side view schematic.[75]

In the microfluidic reactor, chemical reagents will be mixed by diffusion under the laminar flow, in which the viscous forces are more prominent than inertial forces. Reynold number (Re) is a measure of the ratio between inertial forces to viscous forces:

$$Re = \frac{\rho u L}{\mu} = \frac{4\rho u A}{\mu P_{wet}} \text{ (eq 1.4)}$$

Where ρ is the density of the fluid, u is the velocity of the fluid with respect to the object, L is a characteristic linear dimension, μ is the dynamic viscosity of the fluid, A is the channel cross-section area and P_{wet} is the wetted perimeter. When the Re<1000, the flow falls into the laminar flow regime, in the microfluidic reactor Re can be smaller than 100. In the laminar flow there is no turbulence, the mass and heat transports depend only on diffusion, and the average mixing time can be calculated by equation 1.5, where x is the diffusion length and D is the diffusion coefficient.

$$\tau_{mix} \propto \frac{x^2}{D} \text{ (eq 1.5)}$$

Mixing times can be significantly decreased by reducing the required diffusion length in microfluidic reactor, and rapid mixing makes it possible for the separation of the nucleation and growth steps in the microfluidic reactor which benefits the homogenous size and shape distribution of NPs.[76]

In low Re microfluidic systems, microfluidic reactors can be classified in to active and passive mixing reactors based on the mixing mechanisms (Figure 1.17).[77] Active microfluidic reactors enhance the mixing performance by stirring or agitating the fluid flow using some form of external energy supply, while passive microreactors contain no moving parts and require no energy input other than the pressure head used to drive the fluid flows at a constant rate.[77]

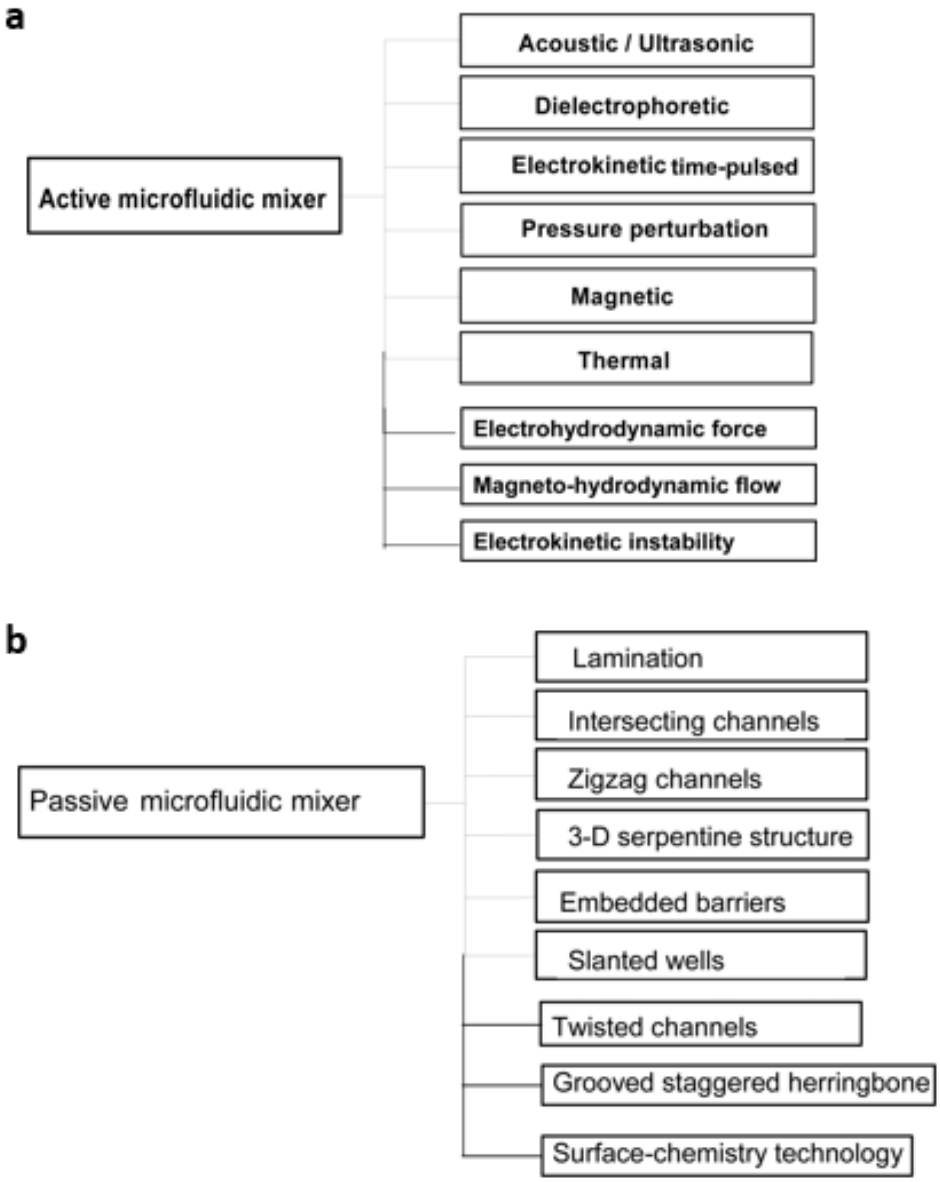


Figure 1. 17 Categories of active (a) and passive (b) microfluidic reactors.[77]

Apart from the microfluidic reactor, microfluidic systems are usually equipped with the following accessories to achieve high quality synthesis of NPs: sensors and a control centre for real time parameters (flow rate, reaction temperature *etc.*) to enable monitoring of the microfluidic system; pressure pump or other methods to introduce chemical reagents; nowadays microfluidic reactors can even be compatible with analysis equipment to realise the real time analysis of the reaction.

II. Properties of microfluidic reactors

As microfluidic reactors are reactors whose internal dimensions are less than 1 mm, the main feature of these reactors is a larger interfacial surface area-to-volume ratio than conventional tank reactors (it can be 50-500 times larger)[60]. Since the heat transfer coefficient is inversely

proportional to the channel diameter, the heat exchange coefficients of microfluidic reactors can be up to one order of magnitude higher than those in conventional tank reactors.[78] The high heat transfer capacity allows the solutions to be heated and cooled quickly so that reactions can be performed under isothermal conditions with precisely defined residence times, this is beneficial for higher selectivities, yields and product qualities.[79]

The large surface area-to-volume ratio of microfluidic reactors also contributes to fast substances transport. The diffusion and mixing times in microfluidic reactors are usually just a few milliseconds which greatly decrease the possibility of the materials transport having an influence on the reaction rate.[80-81]

A third advantage of the microfluidic reactor-based synthesis is that it is much easier to control the reactions' pressure, temperature, residence time and flow rate than traditional tank reactor-based synthesis.

Last but not least, the microfluidic synthesis set up also makes it possible for the real time in situ characterization of the synthesis when the characterization facilities are coupled to the microfluidic reactor[82-83] (Figure 1.18).

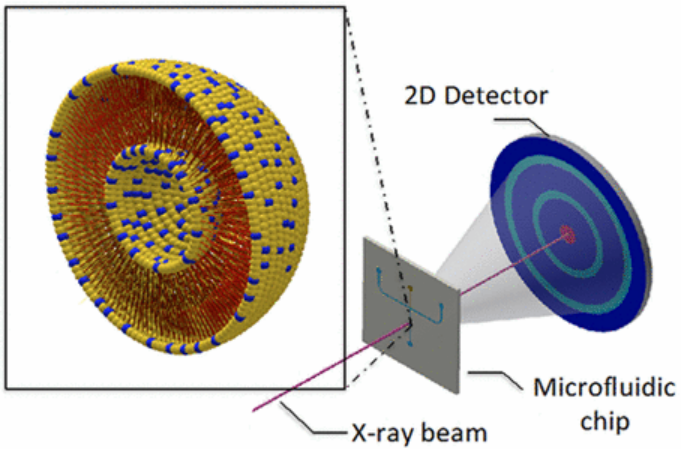


Figure 1. 18 Schematic of a microfluidic reactor coupled with SAXS. [82]

In summary, the synthesis of GNPs using microfluidic reactors[78, 84-90] has been recognized as an efficient approach as it shows advantages such as fast mixing, highly controlled reaction parameters, good particle quality and reproducibility, convenient real-time monitoring for quality control, *etc.*, compared to traditional benchtop tank reactor synthesis.

1.2.4 Endotoxin—the uninvited guest

Endotoxins, also called lipopolysaccharide (LPS), are a key component of the outer cell membrane of Gram-negative bacteria[91], and can be easily released during the activities of the bacterial membrane and cell wall. LPS are large (molecular weight: 3000 to 4000 Da), heat-

stable (up to 100°C), molecules consisting of a core polysaccharide chain (outer and inner core), O-specific polysaccharide side chains (O-antigen) and a lipid component, Lipid A (Figure 1.19)[92]. Lipid A is the most conserved part of endotoxin and is responsible for most of the endotoxicity.[93]

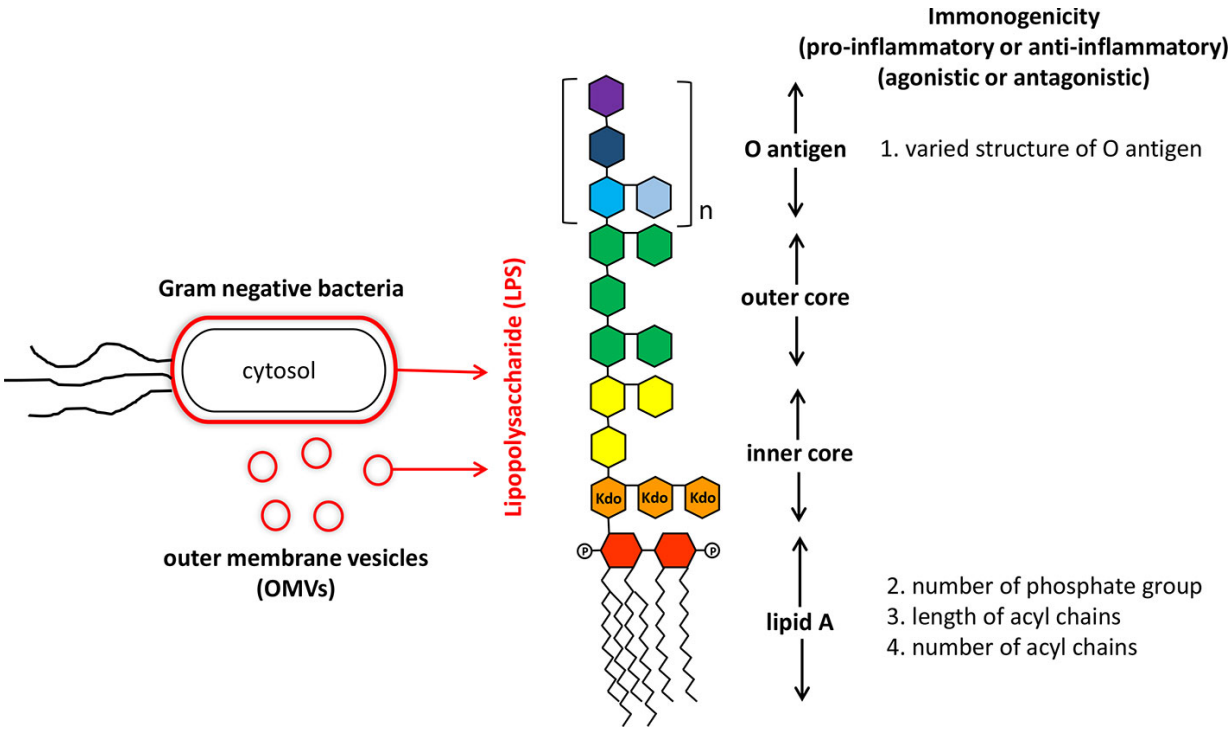


Figure 1. 19 Diagram of a lipopolysaccharide (LPS).[92]

It has been demonstrated that LPS is able to strongly adsorb on the NPs' surface and be part of the biomolecular corona, therefore potentially altering the biological fate of NPs. For example, researchers found that the absorption of LPS on the surface of engineered NPs can greatly affect the inflammatory or immune toxic effects, which might lead to erroneous interpretation of biological outcomes.[94] Moreover, LPS adsorbed onto the NP surface can be recognized by the scavenger receptors and toll-like receptor 4, which are abundant in the mononuclear phagocyte system.[95] This will modify their "natural" binding cells, and potentially influence their biodistribution.[96] Therefore, distinguishing the true NPs effects from those caused by LPS is crucial for a correct and reliable nano-safety and bio-nano interactions evaluation.

Considering LPS is prevalent in air, tap water and human fingers *etc.* endotoxin contamination is hard to avoid and must not be ignored in NPs synthesis that is performed in the open laboratory environment. Simple distillation and deionizing columns don't remove endotoxin and other more vigorous purification methods (such as strong base or acid washing) may impact the colloidal stability of the NPs. Thus, the most efficient way to protect NPs from

endotoxin contamination is by performing the NPs synthesis under an endotoxin-free synthesis protocol.[48] As shown in Figure 1.20, our lab designed a special “clean room” with air filtration system and strictly limited access for the endotoxin-free(“clean”) synthesis. Before entering the “clean room”, disposable lab coat, overshoes, face mask and hair cover should be worn in the ante room. In the endotoxin-free synthesis protocol, all of the standard operational procedures associated with cell culture are applied to the synthesis to prevent LPS contamination. Meanwhile all of the reagents and consumables used for the clean synthesis are endotoxin-free certified and are only strictly opened inside the cabinet. All the glassware used has to be cleaned with aqua regia before being brought to the cabinet and then thoroughly rinsing with endotoxin-free water. Bactericidal solution is used in the water bath and the chiller for the reflux apparatus. All the reagents stock solutions need to be filtered through 0.2 mm filters and evaluated for their LPS level before use.

Figure 1. 20 Standard operation in a clean synthesis process.

The endotoxin level of the prepared NPs needs to be evaluated to ensure safe use in the biological studies. The current US Food and Drug Administration (FDA) limits require eluates from medical devices to be less than 0.5 EU/ml, unless the device comes into contact with cerebrospinal fluid where the limit is then 0.06 EU/ml [97]. Limulus amebocyte lysate (LAL)

assay, which is based on a cell lysate of the horseshoe crab Limulus polyphemus that coagulates in the presence of even very low levels of endotoxins, is now a widely accepted method for the evaluation of the endotoxin level in NPs. Pierce™ LAL Chromogenic Endotoxin Quantitation Kit is a commercial kit based on the LAL assay. As shown in Figure 1.21, endotoxin activates factor C, B and pro-clothing enzyme sequentially in the amebocyte lysate, and the activated pro-clotting enzyme can catalyze the chromogenic substrate (Ac-lle-Glu-Ala-Arg-pNA) to release p-nitroaniline (pNA), which can be measured at 405 nm after the reaction is stopped. The absorption at 405 nm is proportional to the concentration of endotoxin in the sample and can be calculated *via* a standard curve.[98]

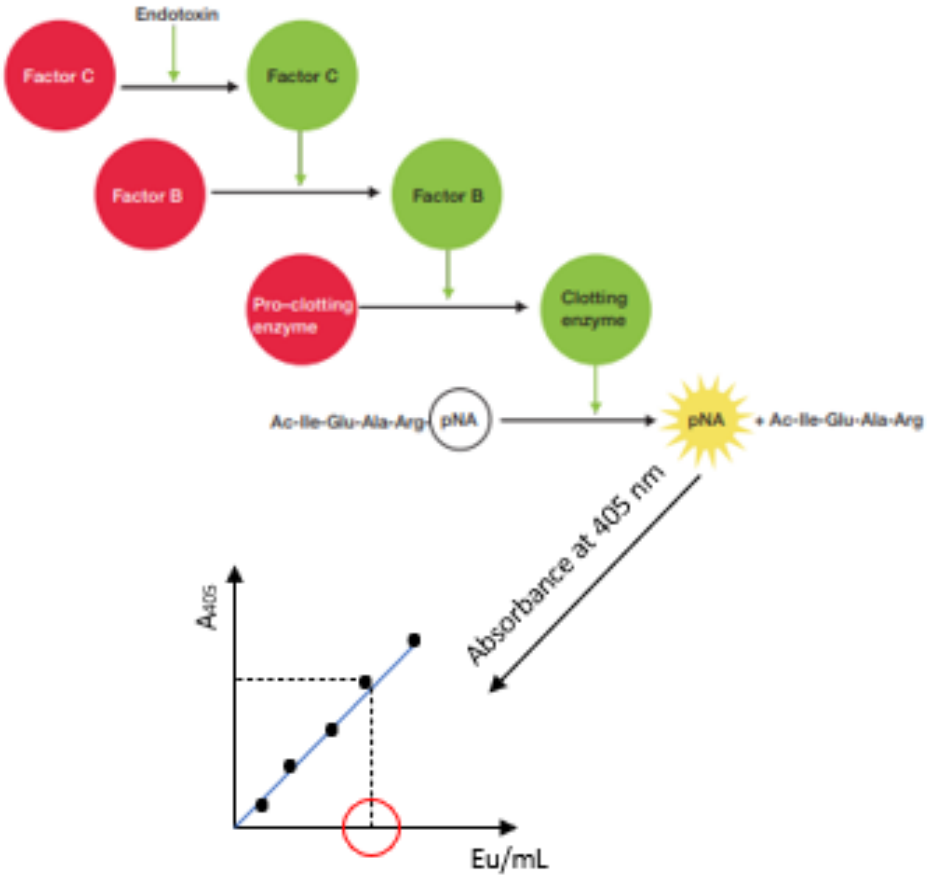


Figure 1. 21 Principle of the Pierce™ LAL Chromogenic Endotoxin Quantitation Kit. Red indicates inactive enzymes and green indicates active enzymes.[98]

1.3 Characterization of NPs

Adequate characterization of nanostructures is of vital importance for the fingerprint of nanostructures and further investigation of their effects and the development of their applications. Nowadays with the development of nanotechnologies, various methods can be

used to determine the properties of nanostructures including size, morphology, optical properties and so on.

1.3.1 Classical characterization methods

When it comes to a proper characterization of nanostructures, determination of their size, morphology and optical properties are crucial. Among these traditional characterization methods, DCS, DLS and NTA are widely used for size measurement, while TEM is one of the most powerful tools to identity the morphology of various nanostructures and UV-vis spectroscopy is one of the traditional methods that can give a hint of the nanostructures' complicated surface properties including size, shape, surface coating *etc.*

I. DCS (Differential centrifugal sedimentation)

DCS, as the name indicates, it is a particle size distribution analysis tool using centrifugal sedimentation within an optically clear spinning disc that is filled with fluid. Sedimentation is stabilized by a density gradient within the fluid and the accuracy of the measured size is insured through the use of a known size calibration standard usually run before each measurement.[99] As shown in Figure 1.22, after injected into the disc, particles are proceeded from the centre of the disc to a detector placed at the outer rim of the disc under a strong centrifugal force. The time taken for this process is recorded and for materials with homogenous density and simple shape (like spherical particles), the time recorded can be related to the size of the particles according to equation 1.6, where D corresponds to the particle diameter, η is the fluid viscosity, R_f is the final radius of rotation, R_0 is the initial radius of rotation, ρ_p is the particle density, ρ_f is the fluid density, ω is the rotational velocity, t is the time recorded to sediment from R_0 to R_f.

$$D = \sqrt{\frac{18\eta \ln\left(\frac{R_f}{R_0}\right)}{\left((\rho_p - \rho_f)\omega^2 t\right)}} \quad \text{(eq 1.6)}$$

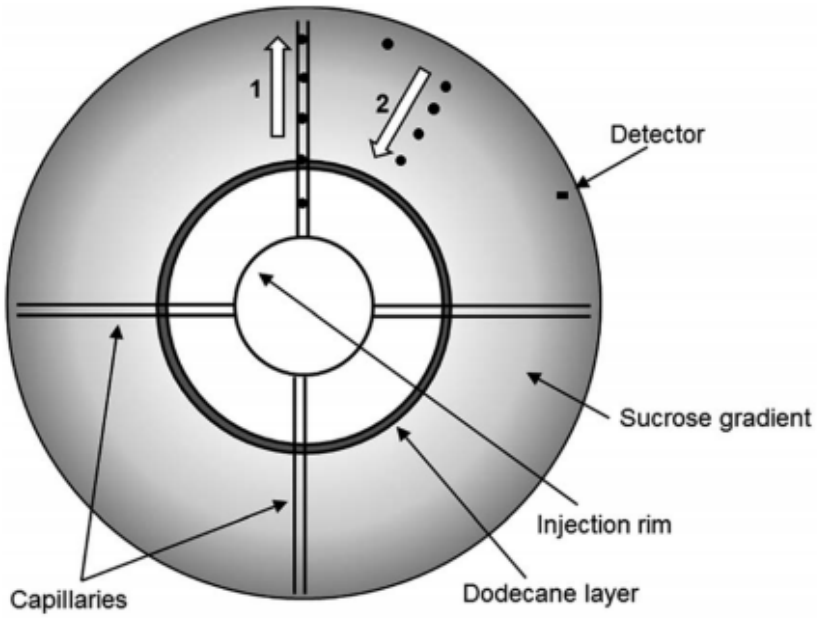


Figure 1. 22 Schematic illustration of the working mechanisms of DCS.[100]

II. DLS (Dynamic light scattering)

Compared with DCS, DLS is used to detect the hydrodynamic diameter of particles dispersed in solutions. The principle of the DLS instrument for size measurement is based on the dispersed particles' scattering of incident light and the detector in the DLS can detect and record the intensity of the scattered light and then calculate the fluctuations in scattered light intensity due to the Brownian movement of the particles (Figure 1.23). According to the Stokes-Einstein equation, the hydrodynamic radius (R_H) of spherical particles can be obtained, where K_B=Boltzmann constant, T=temperature, η=absolute viscosity and D_t=translational diffusion coefficient.

$$D_t = \frac{k_B T}{6\pi\eta R_H} \text{ (eq 1.7)}$$

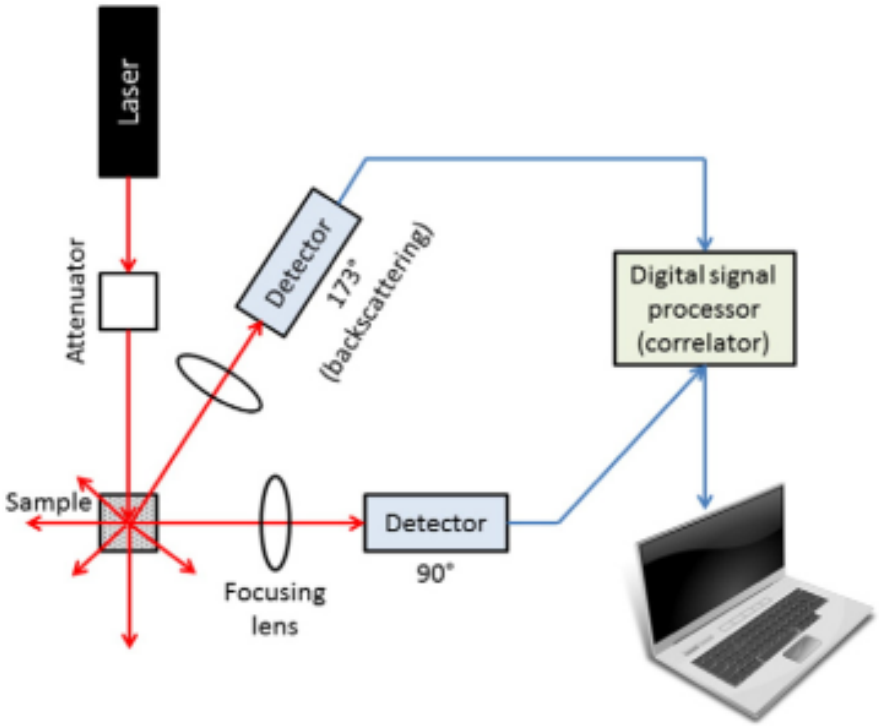


Figure 1. 23 Schematic illustration of the working mechanism of DLS.[101]

III. NTA (Nanosight trafficking analysis)

NTA is another user-friendly tool for the hydrodynamic size determination of nanostructures. The same as DLS, it is based on the dispersed particles' scattering of incident light. By combining the laser light scattering microscopy with a charge-coupled device (CCD) camera, NTA enables the visualization and recording of particles in solution and then the software is able to identify and track individual particles moving under Brownian motion and relates the movement to the hydrodynamic size according to the following formula derived from the Stokes-Einstein equation:

$$\overline{(x,y)^2} = \frac{2k_BT}{3\pi\eta R_H} \quad \text{(eq 1.8)}$$

where $\overline{(x,y)^2}$ is the mean-squared speed of a particle[102].

In addition to the size characterization, the existence of the camera enables NTA to detect the concentration of nanoparticles based on the counting of particles in a time course. As shown in Figure 1.24, the Brownian motion of NPs suspended in the liquid and numbers of NPs in a time course are captured by the microscope with a CCD camera and then transferred to the size distribution and concentration (NPs/ml) by the software.

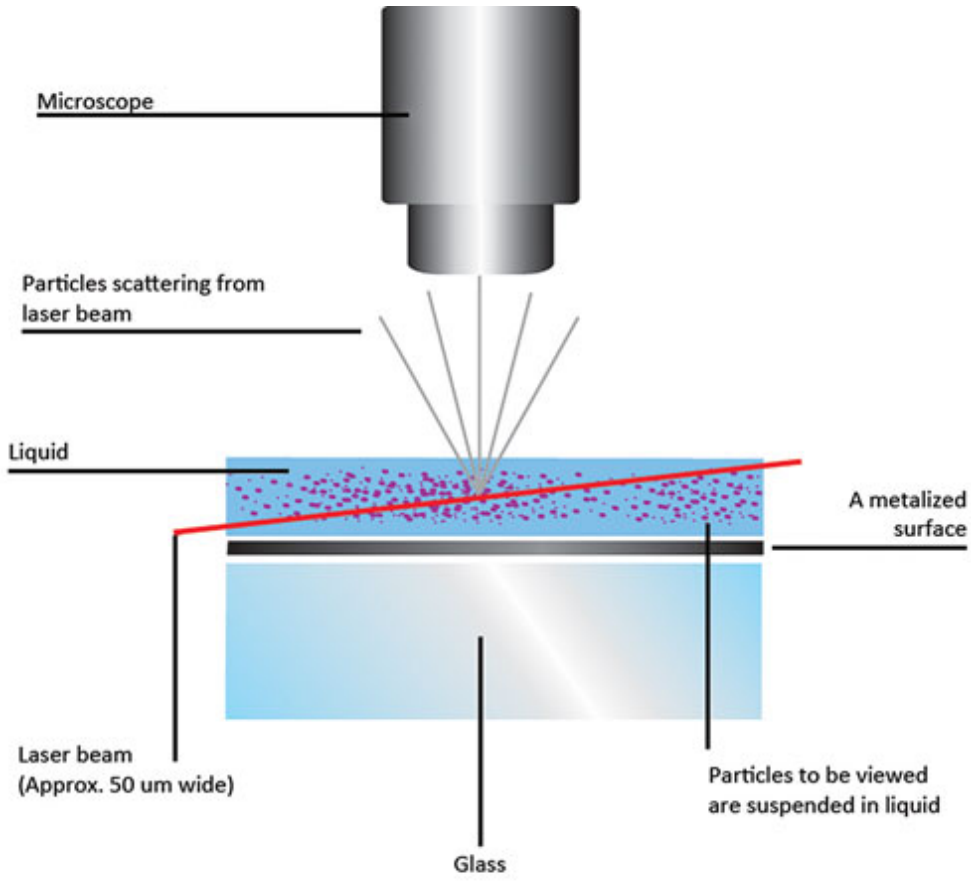


Figure 1. 24 Schematic illustration of the instrument of NTA, copied from website https://www.malvernpanalytical.com/en/products/technology/light-scattering/nanoparticle-tracking-analysis.

IV. TEM (Transmitted electron microscopy)

To determine the morphology characterization of nanostructures, microscopic techniques such as TEM, SEM and polarized optical microscopy are used. TEM uses a beam of electrons transmitted on a specimen and produces a highly magnified and detailed image of the specimen. TEM can identify the NPs' size, structure, shape and even surface biomolecular coatings combining with different labelling or mapping methods. In addition, the images from TEM can in conjunction with machine learning be used for further characterization of complex nanostructures, more details will be discussed in chapter III and appendix III.

V. UV-vis spectroscopy (Ultraviolet-visible spectroscopy)

As a traditional optical characterization method, UV-vis spectroscopy still plays an important role in the nanoparticles' characterization, here we mainly focused on the LSPR characterization which is crucial for GNPs whose LSPR is closely related with the size[103] and shape[104] properties.

Based on Mie scattering theory, Wolfgang et.al. proposed a UV-vis spectra based method for the determination of size and concentration of spherical GNPs.[103] As shown in Figure 1.25, researchers found that the ratio of the absorbance of spherical GNPs at the LSPR peak to the absorbance at 450 nm shifted from 1.3 to 2.5 with size range from 5 to 25 nm[103-105]. With increasing size the LSPR induced absorption peak also shifts from 520 nm to 590 nm[106].

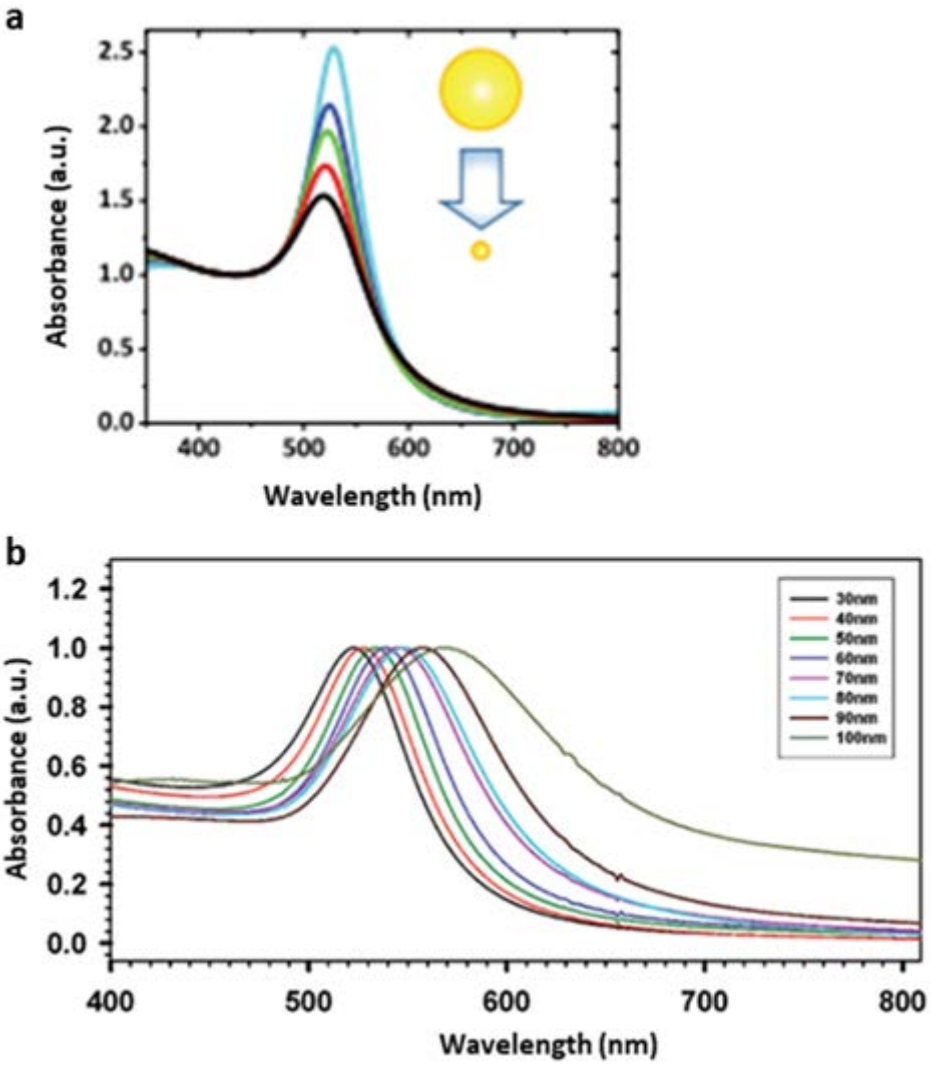


Figure 1. 25 (a) LSPR broadening and damping with spherical GNPs size shifted from 25 to 5 nm.[104] (b) Normalized UV-vis spectra for spherical GNPs with size from 30-100 nm.[106]

As for the shape effects for the LSPR, as shown in Figure 1.26, compared with spherical GNPs, nanorod, nanostar and nanocage show near-infrared shifting of the resonance peak.

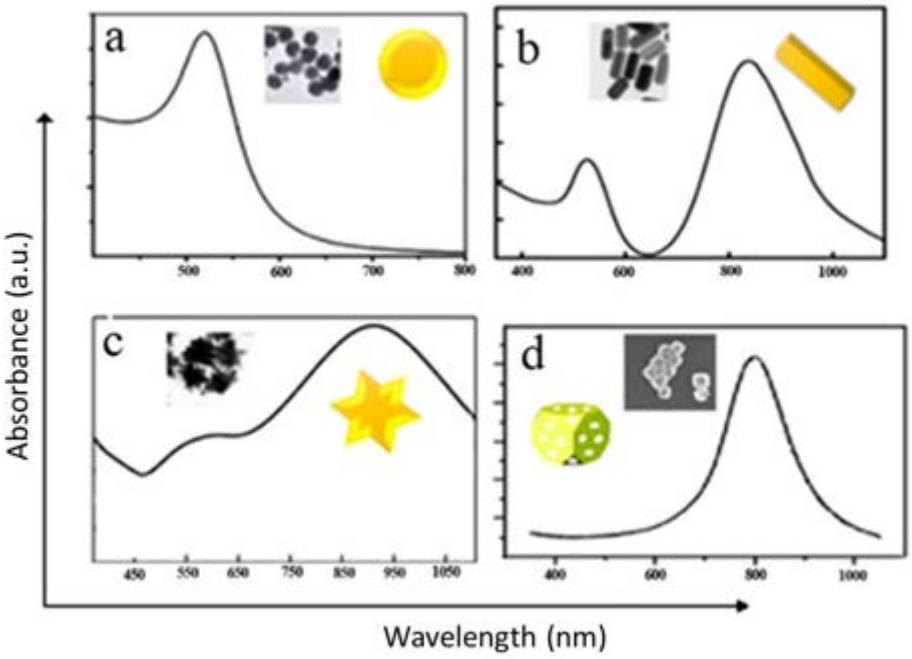


Figure 1. 26 LSPR different shapes of GNPs and their corresponding LSPR. (a) spherical GNPs. (b) gold nanorod. (c) gold nanostar. (d) gold nanocages.[107]

Although NPs' shape can influence their LSPR and makes it possible for UV-vis to detect the possible shape of NPs, however, considering LSPR can be affected by many other factors (size, material, temperature *etc.*), UV-vis is not precise enough to identify and classify different shapes. As shown in Figure 1.27, while UV-vis absorption spectra show a good overlap of the LSPR peaks, TEM micrographs show a very different nanoshapes. This is the same for the other optical properties dependent shape characterization methods such as single particle extinction and scattering (SPES) based optical characterization methods for nanoshapes identification.[108]

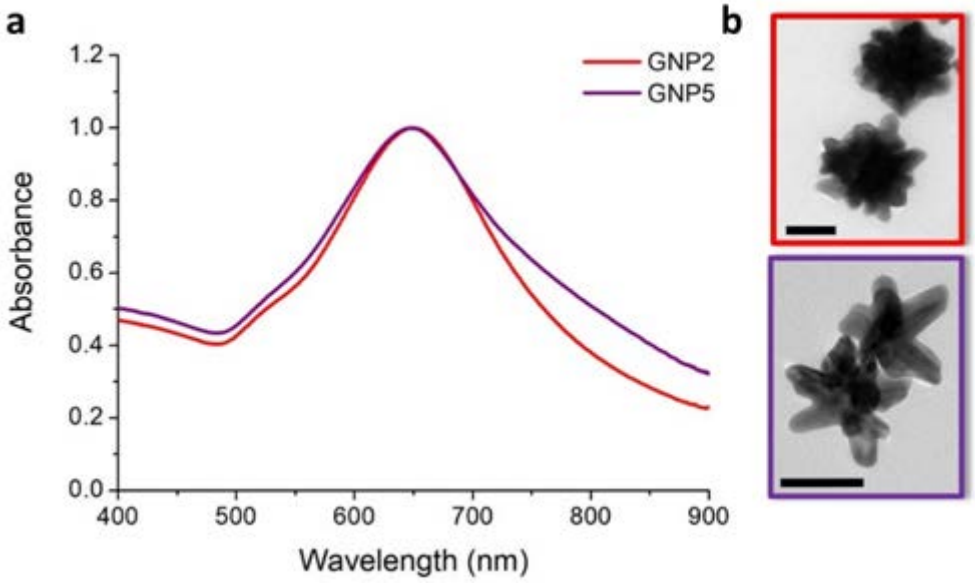


Figure 1. 27 comparison of shapes by optical properties and TEM. (a) UV-vis absorption spectra. (b) Representative TEM images (scale bar is 50 nm). [109]

1.3.2 Computational characterization of complex nanostructures

Although classical characterization methods provide a broad view of nanostructures' properties, when it comes to complex nanostructures, relying on the classical methods is not enough. For example, with the synthetic methods developed, more complicated shapes can be obtained and sometimes we even cannot give a suitable name to these complex nanoshapes, not to say compare them with other research's work. Computer science and artificial intelligence development, provides the possibility of statistical analysis of complex nanoshapes and gives a fingerprint for each shape ensembles, which traditional characterization methods cannot achieve. Briefly, well spread TEM images of each shape ensembles goes through computational analysis and transfers the shape information to digital coordinates and then the principal component analysis is used to define and visualize each shape ensembles. More details for this method will be discussed in chapter III and appendix III.

1.4 Bio-nanoscale Interactions

1.4.1 Biomolecular corona—the biological identity of NPs

The surface of NPs have higher free energy than that of the bulk material, the surfaces of NPs will progressively and selectively absorb biomolecules (such as proteins, lipids and so on) when they come into contact with complex biological fluids, unless they have been designed to do otherwise.[110] This corona of biomolecules lowers the surface energy of NPs and promotes their dispersion. Therefore, when NPs interact with biological systems it is the biomolecular corona that constitutes a major element of the biological identity of the NPs. The biomolecules adsorption on the NPs surface is mainly dominated by the biomolecule affinity interactions towards the NP surface and the affinity between biomolecules in the fluid media. As shown in Figure 1.28[110], the composition of the corona changes over time, and it can be classified based on this exchange into a "hard" corona in which the biomolecules have higher affinity for the NPs surface and therefore strongly adsorb onto the NP (slowly exchanging with the biological fluid) and on top of this "hard" corona is a "soft" corona which consists of biomolecules that are loosely associated with the NP surface and rapidly exchange with the surrounding biomolecules in the medium.[99, 111-112]

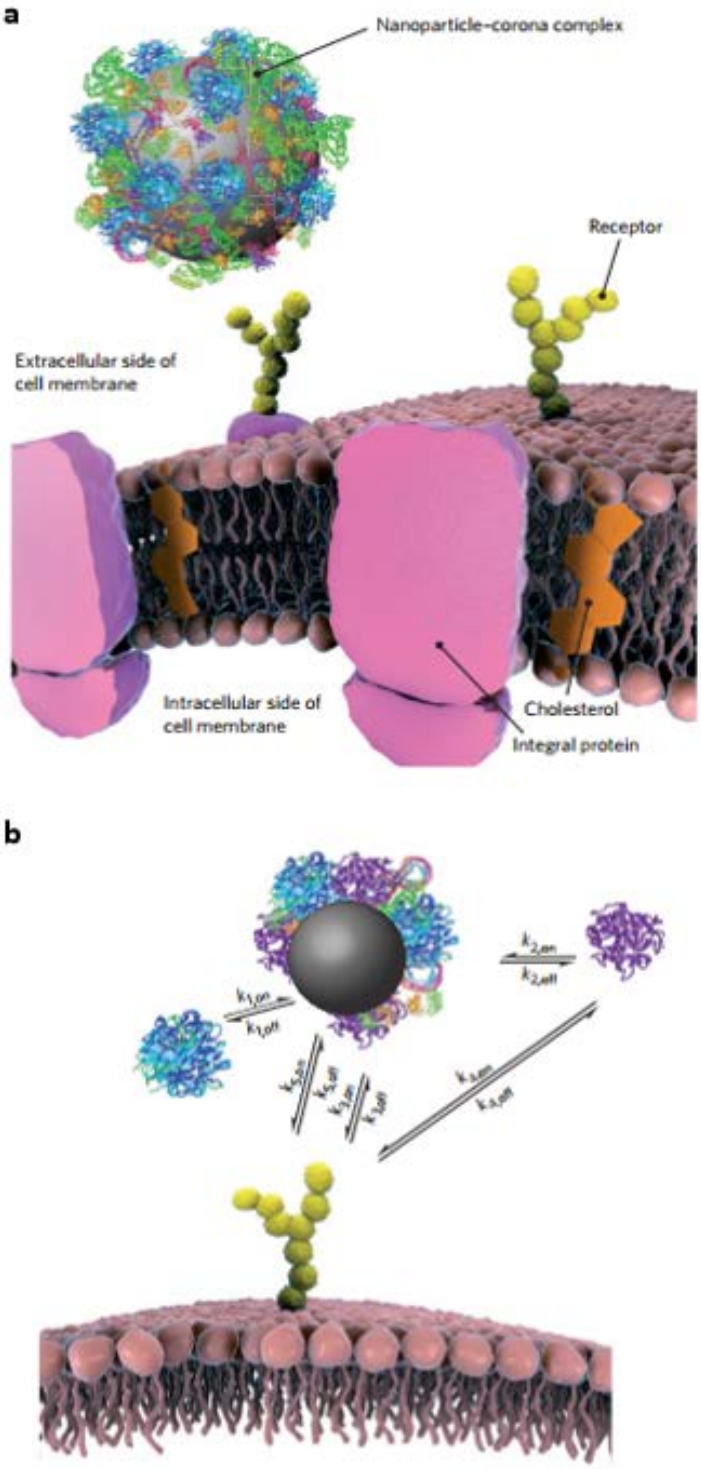


Figure 1. 28 The nanoparticle-corona complex in a biological environment.[110]

As the hard corona is distinctively stable, any subsequent exposure of the NPs to a new environment with different biomolecules may lead to only partial displacement of the original hard corona by new molecules[113-114]. Biomolecules that are not replaced would serve as a corona 'memory' of the nanoparticle's previous environment. Therefore, the corona

composition could potentially depend not only on the current environment of the NPs, but on all environments that they have gone through.

In cases where the corona biomolecules have the propensity of being recognized by the biological machinery there may be a combination of two effects:

(I) A non-specific lowering of the direct physical interactions between the bare particle and the cell surface. For example, in several cases, such as carbon nanotubes,[115] silica and graphene oxide[116], cell damage occurs in the absence of a protein corona (that is, cells exposed to particles in a medium without proteins), but is mitigated in the presence of protein corona.

(II) A specific interaction between absorbed biomolecules and the biological machinery, resulting in the activation of more specific regulated pathways. For example, it has been found that NPs induced protein unfolding leading to initiation of the nuclear factor-κB (NF-κB) pathway and inflammation.[117]

1.4.2 Biological effects of NPs—shape matters

When it comes to the biological effects of NPs, the effect of shape in the bio-nano interactions is one of the fields that attracts lots of researchers' interest. Studies show that in serum free medium (which means there is no biomolecular corona on the NPs surface during the nanoparticle-cell interactions), NPs with a sharp shape can pierce the membrane of the endosome, escape to the cytoplasm, therefore reduce the cellular exocytosis (Figure 1.29).[118]

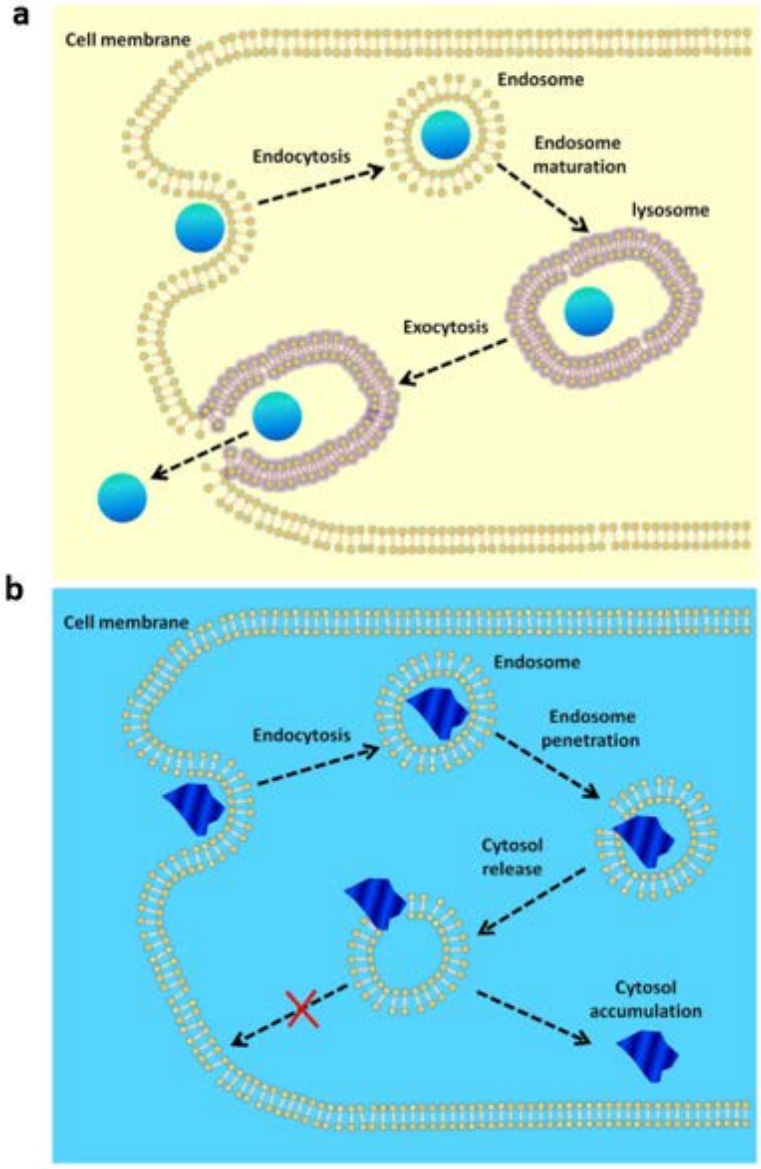


Figure 1. 29 Shape effect on cellular exocytosis.[118]

In such cases, particles in the absence of protein may instead acquire a corona from biomolecules in the cell membrane, which suggests a mechanism for the damage.[119] However, as mentioned above, due to NPs having a large surface area and high surface free energy, when they are exposed to the biological milium, their surface will be modified by the adsorption of biomolecules. The biomolecular corona gives NPs a new biological identity during their interactions with cells and can influence the nanoparticle intracellular fate. For example, studies show that the presence or absence of corona can have great effects on NPs uptake and impact on cells.[119] Hence it is more realistic meaningful to investigate the shape effect of NPs with a biomolecular corona, in order to have a better understanding of nanotoxicity and the rational design of NPs for medical use. For the GNPs shape *in vivo* study, researchers found that NPs' shape can influence their anatomical biodistribution which is an important and

fundamental investigation associated with further nanoparticle-enable drug delivery application. (Figure 1.30)[48]

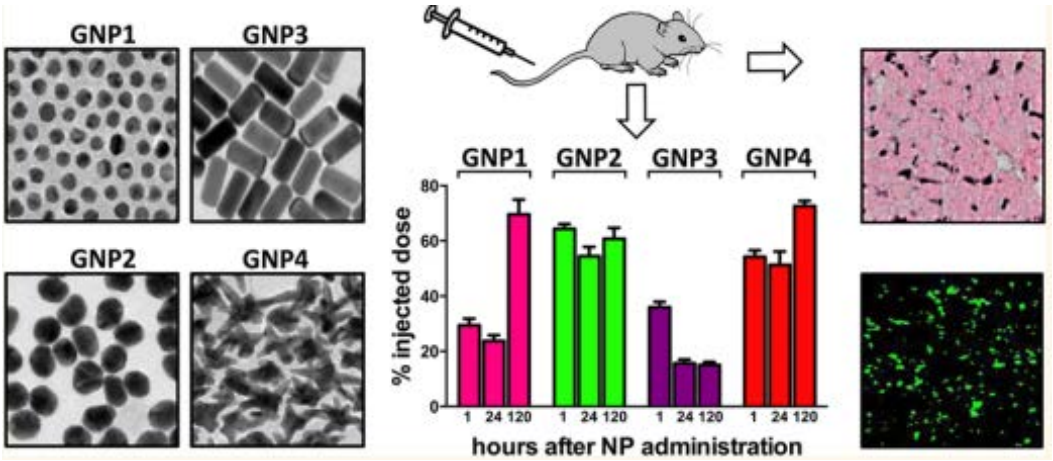


Figure 1. 30 Influence of size and **s**hape on the anatomical biodistribution of GNPs.[48]

In 2013 Niikura *et al.*[120] reported GNPs' shape can influence the immunological responses *in vitro* and *in vivo* for the production of antibodies for West Nile Virus (WNV), as shown in Figure 1.31, 40 nm spherical GNPs produced the highest level of WNV envelope specific antibodies, which makes it promising for GNPs to work as effective vaccine adjuvants.

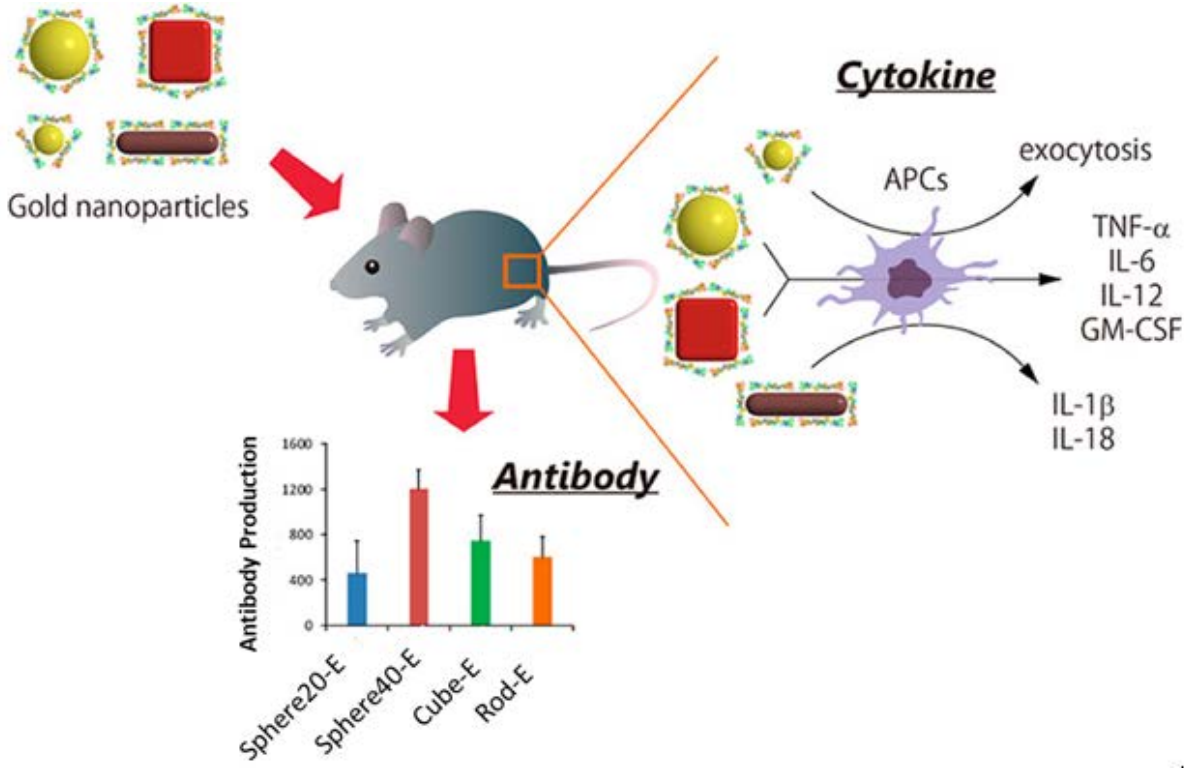


Figure 1. 31 Size and Shape effects of GNPs on the *in vitro* and *in vivo* immunological responses.[120]

1.4.3 Endogenous process—re-engineering of NPs

As mentioned above, NPs show great potential in the nanomedicine area, in such applications NPs should efficiently enter target cells and leave after fulfilling their function to avoid adverse health effects. Therefore, it is necessary to understand the mechanisms by which NPs are internalised by cells as well as their intracellular trafficking routes and final destinations (endogenous process).

The endogenous process (intracellular fate) of NPs mainly includes the NPs cellular uptake pathways, intracellular trafficking pathways and final destinations.[121] (Figure 1.32)

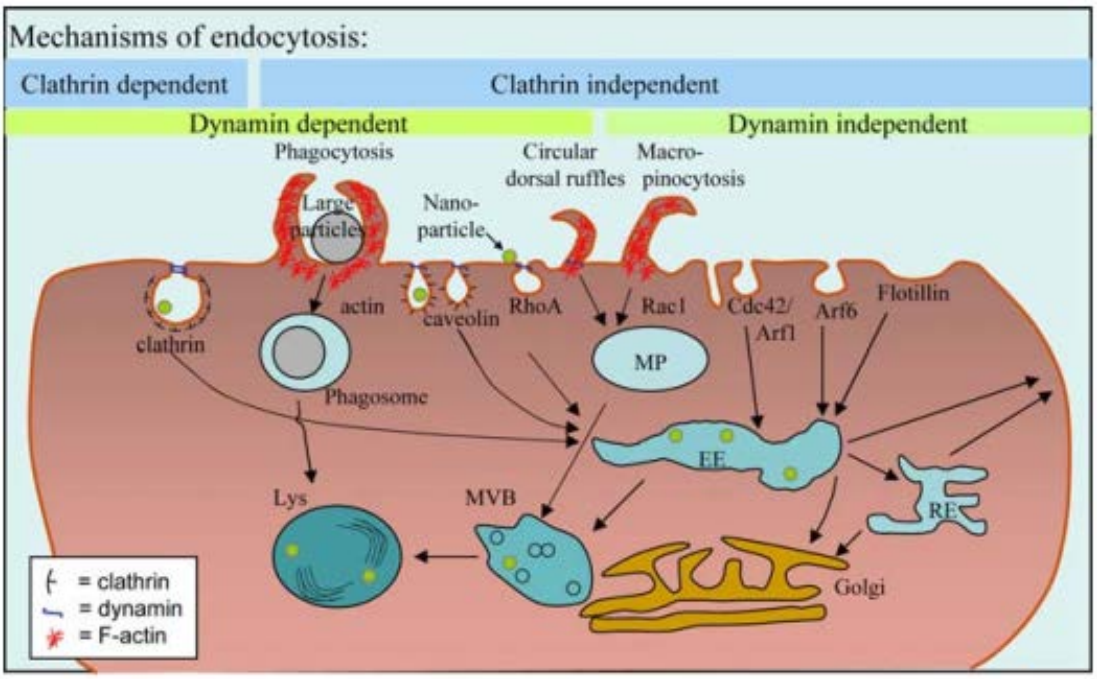


Figure 1. 32 Models of endocytosis mechanisms and intracellular transport.[121]

Cellular uptake of various types of NPs has been extensively studied, revealing active endocytosis processes and pathways involved in their internalization. Almost all cells can internalize NPs by pinocytosis.[122-123] Four basic pinocytic mechanisms are currently known, macropinocytosis, clathrin-mediated endocytosis, caveolin-mediated endocytosis, and mechanisms independent of clathrin and caveolin.[124] Endocytosis of NPs can be regulated by their physical and chemical properties, including their size and shape, surface charge, and modifications. In addition, cell type and biomolecules in the environment were also found to strongly modulate the cellular uptake efficiency.[125-126] Following endocytosis, the majority of the NPs entered the endosomes and were transported into lysosomes, small subpopulations can escape this fate and are released to the exterior of the cell (exocytosis).

In contrast to those studies on endocytosis, investigations on NPs that are secreted by cells are rare and there are many questions needing further study. For example, where do these

secreted NPs go after secretion from cells (intracellular trafficking pathways), how to link the NPs surface presentations to their intracellular trafficking pathways, why these NPs can escape from the endo-lysosome fate and so on.

To answer these questions, the investigation can be roughly divided into three steps:

I. isolation of these secreted NPs in sufficient quantity

The first challenge in the exploration is the efficient extraction of these released NPs without damaging their surface coatings, as their surface coatings are re-engineered during their intracellular trafficking process which contain their intracellular trafficking information. There are many aspects that need to pay attention in order to achieve this goal. For instance, the NPs chosen for this study, NPs with magnetite cores would facilitate their isolation and analysis by differential centrifugation and magnetic extraction, in addition, NPs with shells that can be easily bonded with fluorescence molecules would allow ease of flow cytometry sorting and fluorescent microscopy imaging.

II. characterization of the surface components of these re-engineered NPs

As mentioned above, when NPs traffic to different vesicles or organelles of the cells, some of the biomolecules of the corona will change, but some will not, and this “memory” can give us a crude view of the NPs subcellular locations through analysis of the composition of the biomolecular condensates. What’s more, a full characterization of their surface compositions, it can help us figure out the key component that enables the NPs to escape the lysosomes and finally achieve the target therapy in future.

Understanding the NPs intracellular fate can be elucidated by a comprehensive characterization of the key proteins of the corona involved in this process. Proteomics[127-128] analysis is an effective way to reveal the impact of shape on the nanoparticle intracellular fate. (Figure 1.33)

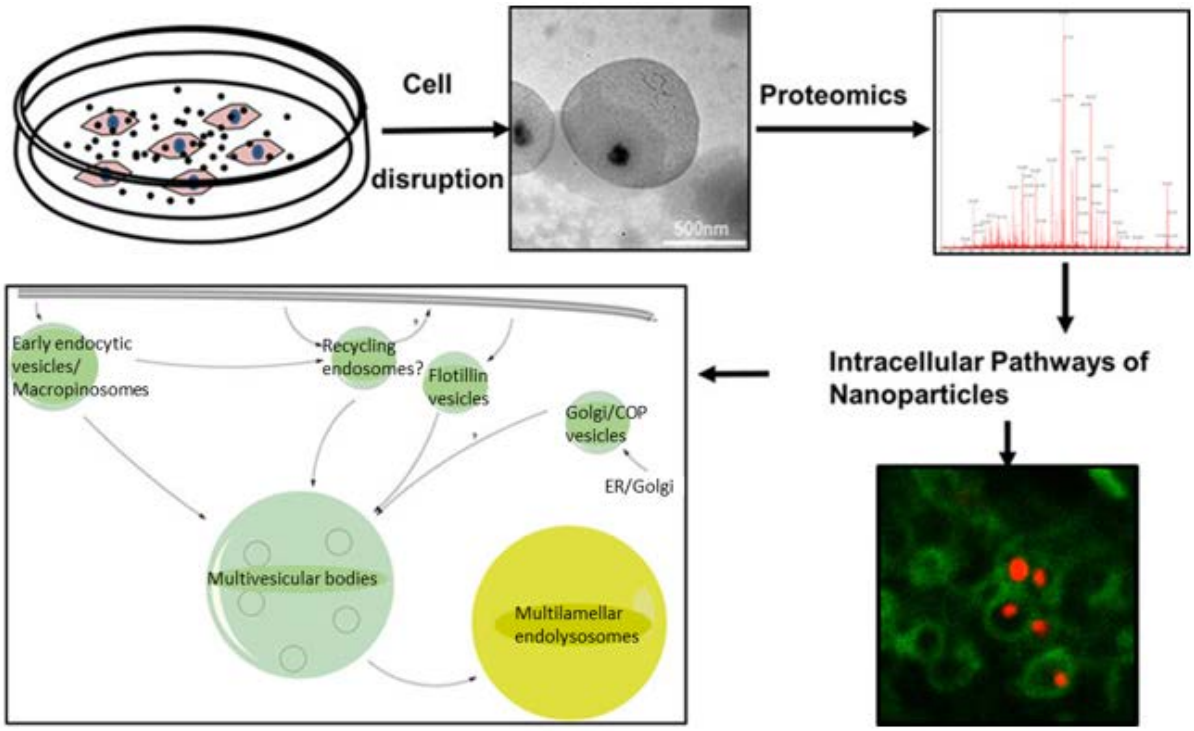


Figure 1. 33 Process of proteomics method for analysing the composition of nanoparticle corona. [127]

Epitope mapping is a novel method for protein characterization in which through suitably designed immunogold labels and subsequent image processing, the statistical distribution of exposed protein epitopes presented across the nanoparticle corona surface are determined on a particle-by-particle basis, identifying the epitopes expressed as well as their organization in relation to one another.[129-133] In this way we can figure out the compositions of corona and then link them to the nanoparticle intracellular trafficking pathways. (Figure 1.34)

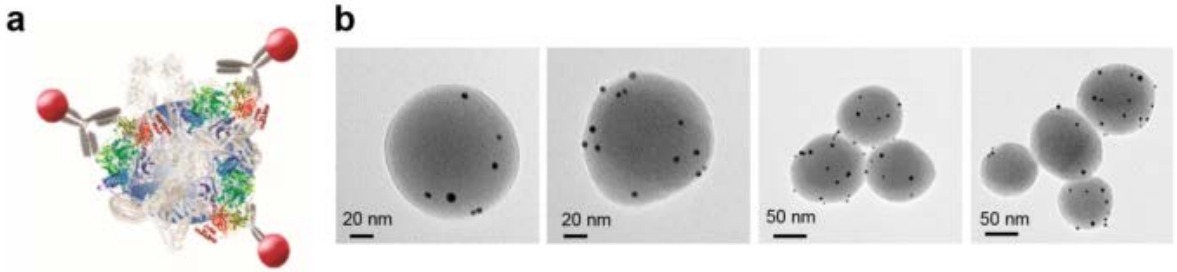


Figure 1. 34 (a) Schematic representation of epitope mapping method for analysing the composition of the nanoparticle corona. (b) TEM micrographs of Apolipoprotein Eb-100 (ApoB-100) epitopes on SiO_2 NPs with a corona formed in 50% of human serum and subsequently incubated in 50% of de-lipidized serum for 4 h.[132]

For the surface components characterization, in addition to identification of proteins, identification of RNAs is also of vital importance as they can give the information about the biogenesis of these NPs. After extraction of integrated RNA from these released NPs, different types of RNA (mRNA, small RNA and so on) can be detected by various types of RNA sequencing methods.

III. further investigation of these re-engineered NPs and their intracellular fate

This research step can help us understand the role these surface components play in the cellular activities and will guide the artificial engineering of NPs for nanomedicine application, for example the possible application of such engineered NPs as an efficient RNA delivery machine.

2 Thesis motivation

Nanoscale cellular processes participate nearly all aspects of coordinated biological action ranging from viral infection to exosome-based cell-cell communication. While it is now apparent that biology is able to recognize and differentially respond to distinct regimes of nanostructures, significant fundamental questions remain unclear especially for the shape-dependent recognition. On the other hand, nanoscale shape is inexhaustibly diverse, which increases the difficulty for their identification and investigation. Hence, the challenges to systematically investigate shape-specificity have hitherto appeared significant, perhaps insurmountable.

The aim of the thesis seeks solutions for the below three fields:

(I) Well-defined control of complex nanostructure's synthesis and classification of shape ensembles.
(II) A systematic study of shape responses *in vitro* and *in vivo* based on (I).
(III) A thorough exploration of the bio-nanoscale recognition and interaction including cellular processes to help understand those effects observed on (II).

Chapter II demonstrates a framework for identifying and generating distinct shape populations and building a quantitative linkage of well-defined nanoscale shape populations to biological effects. Combining our group previously published computational analysis method[134] to define nanoscale shape, we were able to exploit new tunable microfluidic flow chemistry synthetic capacities to (reproducibly) traverse 'shape space'. Based on *in vitro* readouts a shape regime of biological interest was identified, and using microfluidic synthesis inductively generated a

tuning trajectory within the regime. We then observed a consistent directional shape-specificity in antibody response and B-cell receptor repertoire changes using nanoparticles that reached this shape regime but were generated by scaled up synthesis, suggesting that such quantitative definition of the targeted shape distribution is transferrable.

Chapter III investigates the shape-dependent histone modifications by using two distinctive shapes of GNPs. Firstly, through human transcriptome microarray analysis, we found many of the cellular pathways triggered by different shapes of GNPs were related with histone modifications. Based on the transcriptomic results, mass-spectrometry based analysis was carried out to draw the shapes effects on the global histone modifications profile. The proteomics results were validated by antibody-based assay. Finally, chromatin immunoprecipitation sequencing (ChIP-seq) was performed to map the global binding sites differences of H3K27me3 protein induced by different shapes.

Chapter IV reports a novel method for the re-engineering of the surface of nanoparticles via an active endogenous process. After the endogenous process the nanoparticles are coated with a new biomolecular layer which contains lipids, proteins and nucleic acids. Biogenesis investigation found that these re-engineered particle complexes were interacted with mitochondrion, endoplasmic reticulum (ER) etc. before secreting to the extracellular milieu, which is consistent with the composition characterization results, showing that the identified proteins and RNAs in particle complexes are related with mitochondrion and ER. Due to their unique surface coatings, particle complexes show distinct bio-nanoscale interactions from their nanoparticle cores, including escaping of the lysosome destination and transferring of foreign genes to recipient cells. Since the surface coatings of the particle complex are producer cells derived, we are able to load specific genes to the particle complex via modifications of the producer cells' genome and hence make it possible for the particle complexes to deliver specific genes to recipient cells.

In summary, work presented in the thesis provides a strategy for the study of the complex nanostructure and bio-nanoscale interactions, starting from well-defined synthesis to quantitative identification and finally seeking for the systematic investigation of their biological effects.

3 References

1. McNeil, S. E., Nanotechnology for the biologist. *Journal of Leukocyte Biology* **2005,** *78* (3), 585-594.
2. Roduner, E., Size matters: why nanomaterials are different. *Chem Soc Rev* **2006,** *35* (7), 583-92.
3. Dreaden, E. C.; Alkilany, A. M.; Huang, X.; Murphy, C. J.; El-Sayed, M. A., The golden age: gold nanoparticles for biomedicine. *Chem Soc Rev* **2012,** *41* (7), 2740-79.
4. Hammond, J. L.; Bhalla, N.; Rafiee, S. D.; Estrela, P., Localized surface plasmon resonance as a biosensing platform for developing countries. *Biosensors (Basel)* **2014,** *4* (2), 172-88.
5. Mohamed, M. M.; Fouad, S. A.; Elshoky, H. A.; Mohammed, G. M.; Salaheldin, T. A., Antibacterial effect of gold nanoparticles against Corynebacterium pseudotuberculosis. *International journal of veterinary science and medicine* **2017,** *5* (1), 23-29.
6. Cagno, V.; Andreozzi, P.; D'Alicarnasso, M.; Jacob Silva, P.; Mueller, M.; Galloux, M.; Le Goffic, R.; Jones, S. T.; Vallino, M.; Hodek, J.; Weber, J.; Sen, S.; Janecek, E. R.; Bekdemir, A.; Sanavio, B.; Martinelli, C.; Donalisio, M.; Rameix Welti, M. A.; Eleouet, J. F.; Han, Y.; Kaiser, L.; Vukovic, L.; Tapparel, C.; Kral, P.; Krol, S.; Lembo, D.; Stellacci, F., Broad-spectrum non-toxic antiviral nanoparticles with a virucidal inhibition mechanism. *Nat Mater* **2018,** *17* (2), 195-203.
7. Han, J.; Li, J.; Jia, W.; Yao, L.; Li, X.; Jiang, L.; Tian, Y., Photothermal therapy of cancer cells using novel hollow gold nanoflowers. *Int J Nanomedicine* **2014,** *9*, 517-26.
8. Salata, O., Applications of nanoparticles in biology and medicine. *Journal of Nanobiotechnology* **2004,** *2* (1), 3.
9. Chou, L. Y.; Ming, K.; Chan, W. C., Strategies for the intracellular delivery of nanoparticles. *Chem Soc Rev* **2011,** *40* (1), 233-45.
10. Karimi, M.; Ghasemi, A.; Sahandi Zangabad, P.; Rahighi, R.; Moosavi Basri, S. M.; Mirshekari, H.; Amiri, M.; Shafaei Pishabad, Z.; Aslani, A.; Bozorgomid, M.; Ghosh, D.; Beyzavi, A.; Vaseghi, A.; Aref, A. R.; Haghani, L.; Bahrami, S.; Hamblin, M. R., Smart micro/nanoparticles in stimulus-responsive drug/gene delivery systems. *Chem Soc Rev* **2016,** *45* (5), 1457-501.
11. Huo, S.; Jin, S.; Ma, X.; Xue, X.; Yang, K.; Kumar, A.; Wang, P. C.; Zhang, J.; Hu, Z.; Liang, X.-J., Ultrasmall Gold Nanoparticles as Carriers for Nucleus-Based Gene Therapy Due to Size-Dependent Nuclear Entry. *ACS Nano* **2014,** *8* (6), 5852-5862.
12. Eustis, S.; El-Sayed, M. A., Why gold nanoparticles are more precious than pretty gold: Noble metal surface plasmon resonance and its enhancement of the radiative and nonradiative properties of nanocrystals of different shapes. *Chemical Society Reviews* **2006,** *35* (3), 209-217.
13. Sperling, R. A.; Rivera Gil, P.; Zhang, F.; Zanella, M.; Parak, W. J., Biological applications of gold nanoparticles. *Chemical Society Reviews* **2008,** *37* (9), 1896-1908.
14. Saha, K.; Agasti, S. S.; Kim, C.; Li, X.; Rotello, V. M., Gold Nanoparticles in Chemical and Biological Sensing. *Chemical Reviews* **2012,** *112* (5), 2739-2779.
15. Jain, P. K.; El-Sayed, I. H.; El-Sayed, M. A., Au nanoparticles target cancer. *Nano Today* **2007,** *2* (1), 18-29.
16. Liu, G.; Gao, J.; Ai, H.; Chen, X., Applications and potential toxicity of magnetic iron oxide nanoparticles. *Small* **2013,** *9* (9-10), 1533-45.
17. Laurent, S.; Forge, D.; Port, M.; Roch, A.; Robic, C.; Vander Elst, L.; Muller, R. N., Magnetic Iron Oxide Nanoparticles: Synthesis, Stabilization, Vectorization, Physicochemical Characterizations, and Biological Applications. *Chemical Reviews* **2008,** *108* (6), 2064-2110.
18. Thanh, N. T.; Maclean, N.; Mahiddine, S., Mechanisms of nucleation and growth of nanoparticles in solution. *Chem Rev* **2014,** *114* (15), 7610-30.
19. Crystallisation, 4th Edition By J. W. Mullin. 2001. Butterworth Heinemann: Oxford, UK. 600 pp. £75.00. ISBN 075-064-833-3. *Organic Process Research & Development* **2002,** *6* (2), 201-202.

20. Wu, Z.; Yang, S.; Wu, W., Shape control of inorganic nanoparticles from solution. *Nanoscale* **2016,** *8* (3), 1237-59.
21. LaMer, V. K.; Dinegar, R. H., Theory, Production and Mechanism of Formation of Monodispersed Hydrosols. *Journal of the American Chemical Society* **1950,** *72* (11), 4847-4854.
22. Sugimoto, T., Underlying mechanisms in size control of uniform nanoparticles. *Journal of Colloid and Interface Science* **2007,** *309* (1), 106-118.
23. Marqusee, J. A.; Ross, J., Kinetics of phase transitions: Theory of Ostwald ripening. *The Journal of Chemical Physics* **1983,** *79* (1), 373-378.
24. Niederberger, M.; Cölfen, H., Oriented attachment and mesocrystals: Non-classical crystallization mechanisms based on nanoparticle assembly. *Physical Chemistry Chemical Physics* **2006,** *8* (28), 3271-3287.
25. Li, D.; Nielsen, M. H.; Lee, J. R. I.; Frandsen, C.; Banfield, J. F.; De Yoreo, J. J., Direction-Specific Interactions Control Crystal Growth by Oriented Attachment. *Science* **2012,** *336* (6084), 1014.
26. Peng, Z. A.; Peng, X., Mechanisms of the Shape Evolution of CdSe Nanocrystals. *Journal of the American Chemical Society* **2001,** *123* (7), 1389-1395.
27. Peng, X.; Manna, L.; Yang, W.; Wickham, J.; Scher, E.; Kadavanich, A.; Alivisatos, A. P., Shape control of CdSe nanocrystals. *Nature* **2000,** *404* (6773), 59-61.
28. Wang, Y.; He, J.; Liu, C.; Chong, W. H.; Chen, H., Thermodynamics versus kinetics in nanosynthesis. *Angew Chem Int Ed Engl* **2015,** *54* (7), 2022-51.
29. Xia, Y.; Xiong, Y.; Lim, B.; Skrabalak, S. E., Shape-controlled synthesis of metal nanocrystals: simple chemistry meets complex physics? *Angew Chem Int Ed Engl* **2009,** *48* (1), 60-103.
30. Xia, Y.; Xia, X.; Peng, H. C., Shape-Controlled Synthesis of Colloidal Metal Nanocrystals: Thermodynamic versus Kinetic Products. *J Am Chem Soc* **2015,** *137* (25), 7947-66.
31. Turkevich, J.; Stevenson, P. C.; Hillier, J., A study of the nucleation and growth processes in the synthesis of colloidal gold. *Discussions of the Faraday Society* **1951,** *11*, 55-75.
32. Ji, X.; Song, X.; Li, J.; Bai, Y.; Yang, W.; Peng, X., Size Control of Gold Nanocrystals in Citrate Reduction: The Third Role of Citrate. *Journal of the American Chemical Society* **2007,** *129* (45), 13939-13948.
33. Polte, J.; Ahner, T. T.; Delissen, F.; Sokolov, S.; Emmerling, F.; Thünemann, A. F.; Kraehnert, R., Mechanism of Gold Nanoparticle Formation in the Classical Citrate Synthesis Method Derived from Coupled In Situ XANES and SAXS Evaluation. *Journal of the American Chemical Society* **2010,** *132* (4), 1296-1301.
34. Polte, J.; Erler, R.; Thünemann, A. F.; Sokolov, S.; Ahner, T. T.; Rademann, K.; Emmerling, F.; Kraehnert, R., Nucleation and Growth of Gold Nanoparticles Studied via in situ Small Angle X-ray Scattering at Millisecond Time Resolution. *ACS Nano* **2010,** *4* (2), 1076-1082.
35. Yao, T.; Sun, Z.; Li, Y.; Pan, Z.; Wei, H.; Xie, Y.; Nomura, M.; Niwa, Y.; Yan, W.; Wu, Z.; Jiang, Y.; Liu, Q.; Wei, S., Insights into Initial Kinetic Nucleation of Gold Nanocrystals. *Journal of the American Chemical Society* **2010,** *132* (22), 7696-7701.
36. Mikhlin, Y.; Karacharov, A.; Likhatski, M.; Podlipskaya, T.; Zubavichus, Y.; Veligzhanin, A.; Zaikovski, V., Submicrometer intermediates in the citrate synthesis of gold nanoparticles: New insights into the nucleation and crystal growth mechanisms. *Journal of Colloid and Interface Science* **2011,** *362* (2), 330-336.
37. Tao, A. R.; Habas, S.; Yang, P., Shape Control of Colloidal Metal Nanocrystals. *Small* **2008,** *4* (3), 310-325.
38. Atta, S.; Beetz, M.; Fabris, L., Understanding the role of AgNO3 concentration and seed morphology in the achievement of tunable shape control in gold nanostars. *Nanoscale* **2019,** *11* (6), 2946-2958.
39. Bastus, N. G.; Comenge, J.; Puntes, V., Kinetically controlled seeded growth synthesis of citrate-stabilized gold nanoparticles of up to 200 nm: size focusing versus Ostwald ripening. *Langmuir* **2011,** *27* (17), 11098-105.

40. Perezjuste, J.; Pastorizasantos, I.; Lizmarzan, L.; Mulvaney, P., Gold nanorods: Synthesis, characterization and applications. *Coordination Chemistry Reviews* **2005,** *249* (17-18), 1870-1901.
41. Murphy, C. J.; Thompson, L. B.; Chernak, D. J.; Yang, J. A.; Sivapalan, S. T.; Boulos, S. P.; Huang, J.; Alkilany, A. M.; Sisco, P. N., Gold nanorod crystal growth: From seed-mediated synthesis to nanoscale sculpting. *Current Opinion in Colloid & Interface Science* **2011,** *16* (2), 128-134.
42. De Silva Indrasekara, A. S.; Johnson, S. F.; Odion, R. A.; Vo-Dinh, T., Manipulation of the Geometry and Modulation of the Optical Response of Surfactant-Free Gold Nanostars: A Systematic Bottom-Up Synthesis. *ACS Omega* **2018,** *3* (2), 2202-2210.
43. Senthil Kumar, P.; Pastoriza-Santos, I.; Rodriguez-Gonzalez, B.; Javier Garcia de Abajo, F.; Liz-Marzan, L. M., High-yield synthesis and optical response of gold nanostars. *Nanotechnology* **2008,** *19* (1), 015606.
44. Jana, N. R., Gram-scale synthesis of soluble, near-monodisperse gold nanorods and other anisotropic nanoparticles. *Small* **2005,** *1* (8-9), 875-82.
45. Tran, T. T.; Lu, X., Synergistic Effect of Ag and Pd Ions on Shape-Selective Growth of Polyhedral Au Nanocrystals with High-Index Facets. *The Journal of Physical Chemistry C* **2011,** *115* (9), 3638-3645.
46. Zhao, P.; Li, N.; Astruc, D., State of the art in gold nanoparticle synthesis. *Coordination Chemistry Reviews* **2013,** *257* (3-4), 638-665.
47. Krpetić, Ž.; Davidson, A. M.; Volk, M.; Lévy, R.; Brust, M.; Cooper, D. L., High-Resolution Sizing of Monolayer-Protected Gold Clusters by Differential Centrifugal Sedimentation. *ACS Nano* **2013,** *7* (10), 8881-8890.
48. Talamini, L.; Violatto, M. B.; Cai, Q.; Monopoli, M. P.; Kantner, K.; Krpetic, Z.; Perez-Potti, A.; Cookman, J.; Garry, D.; C, P. S.; Boselli, L.; Pelaz, B.; Serchi, T.; Cambier, S.; Gutleb, A. C.; Feliu, N.; Yan, Y.; Salmona, M.; Parak, W. J.; Dawson, K. A.; Bigini, P., Influence of Size and Shape on the Anatomical Distribution of Endotoxin-Free Gold Nanoparticles. *ACS Nano* **2017,** *11* (6), 5519-5529.
49. Chen, S.; Wang, Z. L.; Ballato, J.; Foulger, S. H.; Carroll, D. L., Monopod, Bipod, Tripod, and Tetrapod Gold Nanocrystals. *Journal of the American Chemical Society* **2003,** *125* (52), 16186-16187.
50. Hao, E.; Bailey, R. C.; Schatz, G. C.; Hupp, J. T.; Li, S., Synthesis and Optical Properties of "Branched" Gold Nanocrystals. *Nano Letters* **2004,** *4* (2), 327-330.
51. Sau, T. K.; Murphy, C. J., Room Temperature, High-Yield Synthesis of Multiple Shapes of Gold Nanoparticles in Aqueous Solution. *Journal of the American Chemical Society* **2004,** *126* (28), 8648-8649.
52. Xie, J.; Lee, J. Y.; Wang, D. I. C., Seedless, Surfactantless, High-Yield Synthesis of Branched Gold Nanocrystals in HEPES Buffer Solution. *Chemistry of Materials* **2007,** *19* (11), 2823-2830.
53. Khoury, C. G.; Vo-Dinh, T., Gold Nanostars For Surface-Enhanced Raman Scattering: Synthesis, Characterization and Optimization. *The Journal of Physical Chemistry C* **2008,** *112* (48), 18849-18859.
54. Jeong, G. H.; Lee, Y. W.; Kim, M.; Han, S. W., High-yield synthesis of multi-branched gold nanoparticles and their surface-enhanced Raman scattering properties. *Journal of Colloid and Interface Science* **2009,** *329* (1), 97-102.
55. Li, J.; Wu, J.; Zhang, X.; Liu, Y.; Zhou, D.; Sun, H.; Zhang, H.; Yang, B., Controllable Synthesis of Stable Urchin-like Gold Nanoparticles Using Hydroquinone to Tune the Reactivity of Gold Chloride. *The Journal of Physical Chemistry C* **2011,** *115* (9), 3630-3637.
56. Wang, W.; Pang, Y.; Yan, J.; Wang, G.; Suo, H.; Zhao, C.; Xing, S., Facile synthesis of hollow urchin-like gold nanoparticles and their catalytic activity. *Gold Bulletin* **2012,** *45* (2), 91-98.
57. Li, Y.; Zhai, M.; Xu, H., Controllable synthesis of sea urchin-like gold nanoparticles and their optical characteristics. *Applied Surface Science* **2019,** *498*, 143864.
58. Yuan, H.; Khoury, C. G.; Hwang, H.; Wilson, C. M.; Grant, G. A.; Vo-Dinh, T., Gold nanostars: surfactant-free synthesis, 3D modelling, and two-photon photoluminescence imaging. *Nanotechnology* **2012,** *23* (7), 075102.

59. DeMello, A. J., Control and detection of chemical reactions in microfluidic systems. *Nature* **2006,** *442* (7101), 394-402.
60. Song, Y.; Hormes, J.; Kumar, C. S., Microfluidic synthesis of nanomaterials. *Small* **2008,** *4* (6), 698-711.
61. Manz, A.; Harrison, D. J.; Verpoorte, E. M. J.; Fettinger, J. C.; Paulus, A.; Lüdi, H.; Widmer, H. M., Planar chips technology for miniaturization and integration of separation techniques into monitoring systems: Capillary electrophoresis on a chip. *Journal of Chromatography A* **1992,** *593* (1), 253-258.
62. Yager, P.; Edwards, T.; Fu, E.; Helton, K.; Nelson, K.; Tam, M. R.; Weigl, B. H., Microfluidic diagnostic technologies for global public health. *Nature* **2006,** *442* (7101), 412-418.
63. Grodrian, A.; Metze, J.; Henkel, T.; Martin, K.; Roth, M.; Köhler, J. M., Segmented flow generation by chip reactors for highly parallelized cell cultivation. *Biosensors and Bioelectronics* **2004,** *19* (11), 1421-1428.
64. Watts, P.; Wiles, C., Micro reactors: a new tool for the synthetic chemist. *Organic & Biomolecular Chemistry* **2007,** *5* (5), 727-732.
65. Rolland, J. P.; Van Dam, R. M.; Schorzman, D. A.; Quake, S. R.; DeSimone, J. M., Solvent-Resistant Photocurable "Liquid Teflon" for Microfluidic Device Fabrication. *Journal of the American Chemical Society* **2004,** *126* (8), 2322-2323.
66. Löwe, H.; Ehrfeld, W., State-of-the-art in microreaction technology: concepts, manufacturing and applications. *Electrochimica Acta* **1999,** *44* (21), 3679-3689.
67. Jensen, K. F., Silicon-Based Microchemical Systems: Characteristics and Applications. *MRS Bulletin* **2006,** *31* (2), 101-107.
68. Snyder, D. A.; Noti, C.; Seeberger, P. H.; Schael, F.; Bieber, T.; Rimmel, G.; Ehrfeld, W., Modular Microreaction Systems for Homogeneously and Heterogeneously Catalyzed Chemical Synthesis. *Helvetica Chimica Acta* **2005,** *88* (1), 1-9.
69. Wagner, J.; Kirner, T.; Mayer, G.; Albert, J.; Köhler, J. M., Generation of metal nanoparticles in a microchannel reactor. *Chemical Engineering Journal* **2004,** *101* (1), 251-260.
70. Elvira, K. S.; Casadevall i Solvas, X.; Wootton, R. C.; deMello, A. J., The past, present and potential for microfluidic reactor technology in chemical synthesis. *Nat Chem* **2013,** *5* (11), 905-15.
71. Seo, M.; Nie, Z.; Xu, S.; Mok, M.; Lewis, P. C.; Graham, R.; Kumacheva, E., Continuous Microfluidic Reactors for Polymer Particles. *Langmuir* **2005,** *21* (25), 11614-11622.
72. Jahn, A.; Reiner, J. E.; Vreeland, W. N.; DeVoe, D. L.; Locascio, L. E.; Gaitan, M., Preparation of nanoparticles by continuous-flow microfluidics. *Journal of Nanoparticle Research* **2008,** *10* (6), 925-934.
73. Shalom, D.; Wootton, R. C. R.; Winkle, R. F.; Cottam, B. F.; Vilar, R.; deMello, A. J.; Wilde, C. P., Synthesis of thiol functionalized gold nanoparticles using a continuous flow microfluidic reactor. *Materials Letters* **2007,** *61* (4), 1146-1150.
74. Zhang, L.; Wang, Y.; Tong, L.; Xia, Y., Synthesis of Colloidal Metal Nanocrystals in Droplet Reactors: The Pros and Cons of Interfacial Adsorption. *Nano Letters* **2014,** *14* (7), 4189-4194.
75. Choi, K.; Ng, A. H. C.; Fobel, R.; Wheeler, A. R., Digital Microfluidics. *Annual Review of Analytical Chemistry* **2012,** *5* (1), 413-440.
76. Lu, M.; Ozcelik, A.; Grigsby, C. L.; Zhao, Y.; Guo, F.; Leong, K. W.; Huang, T. J., Microfluidic hydrodynamic focusing for synthesis of nanomaterials. *Nano Today* **2016,** *11* (6), 778-792.
77. Lee, C.-Y.; Chang, C.-L.; Wang, Y.-N.; Fu, L.-M., Microfluidic Mixing: A Review. *International Journal of Molecular Sciences* **2011,** *12* (5).
78. Murphy*, C. J., A Simple Millifluidic Benchtop Reactor System for the High-Throughput Synthesis and Functionalization of Gold Nanoparticles with Different Sizes and Shapes. *ACS Nano* **2013,** *VOL. 7 ' NO. 5 ' 4135–4150 ' 2013*.
79. Jähnisch, K.; Hessel, V.; Löwe, H.; Baerns, M., Chemie in Mikrostrukturreaktoren. *Angewandte Chemie* **2004,** *116* (4), 410-451.

80. Tonhauser, C.; Natalello, A.; Löwe, H.; Frey, H., Microflow Technology in Polymer Synthesis. *Macromolecules* **2012,** *45* (24), 9551-9570.
81. Mills, P. L.; Quiram, D. J.; Ryley, J. F., Microreactor technology and process miniaturization for catalytic reactions—A perspective on recent developments and emerging technologies. *Chemical Engineering Science* **2007,** *62* (24), 6992-7010.
82. Ilhan-Ayisigi, E.; Yaldiz, B.; Bor, G.; Yaghmur, A.; Yesil-Celiktas, O., Advances in microfluidic synthesis and coupling with synchrotron SAXS for continuous production and real-time structural characterization of nano-self-assemblies. *Colloids and Surfaces B: Biointerfaces* **2021,** *201*, 111633.
83. Ghazal, A.; Gontsarik, M.; Kutter, J. P.; Lafleur, J. P.; Ahmadvand, D.; Labrador, A.; Salentinig, S.; Yaghmur, A., Microfluidic Platform for the Continuous Production and Characterization of Multilamellar Vesicles: A Synchrotron Small-Angle X-ray Scattering (SAXS) Study. *The Journal of Physical Chemistry Letters* **2017,** *8* (1), 73-79.
84. Yang, S.-Y.; Cheng, F.-Y.; Yeh, C.-S.; Lee, G.-B., Size-controlled synthesis of gold nanoparticles using a micro-mixing system. *Microfluidics and Nanofluidics* **2009,** *8* (3), 303-311.
85. Jamal, F.; Jean-Sébastien, G.; Maël, P.; Edmond, P.; Christian, R., Gold nanoparticle synthesis in microfluidic systems and immobilisation in microreactors designed for the catalysis of fine organic reactions. *Microsystem Technologies* **2011,** *18* (2), 151-158.
86. Thiele, M.; Soh, J. Z. E.; Knauer, A.; Malsch, D.; Stranik, O.; Müller, R.; Csáki, A.; Henkel, T.; Köhler, J. M.; Fritzsche, W., Gold nanocubes – Direct comparison of synthesis approaches reveals the need for a microfluidic synthesis setup for a high reproducibility. *Chemical Engineering Journal* **2016,** *288*, 432-440.
87. Uson, L.; Sebastian, V.; Arruebo, M.; Santamaria, J., Continuous microfluidic synthesis and functionalization of gold nanorods. *Chemical Engineering Journal* **2016,** *285*, 286-292.
88. Rahman, M.; Rebrov, E., Microreactors for Gold Nanoparticles Synthesis: From Faraday to Flow. *Processes* **2014,** *2* (2), 466-493.
89. Boleininger, J.; Kurz, A.; Reuss, V.; Sönnichsen, C., Microfluidic continuous flow synthesis of rod-shaped gold and silver nanocrystals. *Phys. Chem. Chem. Phys.* **2006,** *8* (33), 3824-3827.
90. Duraiswamy, S.; Khan, S. A., Droplet-based microfluidic synthesis of anisotropic metal nanocrystals. *Small* **2009,** *5* (24), 2828-34.
91. Gorbet, M. B.; Sefton, M. V., Endotoxin: The uninvited guest. *Biomaterials* **2005,** *26* (34), 6811-6817.
92. Lin, T.-L.; Shu, C.-C.; Chen, Y.-M.; Lu, J.-J.; Wu, T.-S.; Lai, W.-F.; Tzeng, C.-M.; Lai, H.-C.; Lu, C.-C., Like Cures Like: Pharmacological Activity of Anti-Inflammatory Lipopolysaccharides From Gut Microbiome. *Front Pharmacol* **2020,** *11*, 554-554.
93. Raetz, C. R.; Whitfield, C., Lipopolysaccharide endotoxins. *Annu Rev Biochem* **2002,** *71*, 635-700.
94. Li, Y.; Shi, Z.; Radauer-Preiml, I.; Andosch, A.; Casals, E.; Luetz-Meindl, U.; Cobaleda, M.; Lin, Z.; Jaberi-Douraki, M.; Italiani, P.; Horejs-Hoeck, J.; Himly, M.; Monteiro-Riviere, N. A.; Duschl, A.; Puntes, V. F.; Boraschi, D., Bacterial endotoxin (lipopolysaccharide) binds to the surface of gold nanoparticles, interferes with biocorona formation and induces human monocyte inflammatory activation. *Nanotoxicology* **2017,** *11* (9-10), 1157-1175.
95. Vallhov, H.; Qin, J.; Johansson, S. M.; Ahlborg, N.; Muhammed, M. A.; Scheynius, A.; Gabrielsson, S., The Importance of an Endotoxin-Free Environment during the Production of Nanoparticles Used in Medical Applications. *Nano Letters* **2006,** *6* (8), 1682-1686.
96. Mano, S. S.; Kanehira, K.; Taniguchi, A., Comparison of cellular uptake and inflammatory response via toll-like receptor 4 to lipopolysaccharide and titanium dioxide nanoparticles. *International journal of molecular sciences* **2013,** *14* (7), 13154-13170.
97. Gorbet, M. B.; Sefton, M. V., Endotoxin: the uninvited guest. *Biomaterials* **2005,** *26* (34), 6811-7.
98. <A highly sensitive assay for endotoxin detection and quantification for a variety of samples.pdf>.

99. Walczyk, D.; Bombelli, F. B.; Monopoli, M. P.; Lynch, I.; Dawson, K. A., What the Cell "Sees" in Bionanoscience. *Journal of the American Chemical Society* **2010,** *132* (16), 5761-5768.
100. Minelli, C.; Sikora, A.; Garcia-Diez, R.; Sparnacci, K.; Gollwitzer, C.; Krumrey, M.; Shard, A. G., Measuring the size and density of nanoparticles by centrifugal sedimentation and flotation. *Analytical Methods* **2018,** *10* (15), 1725-1732.
101. Bhattacharjee, S., DLS and zeta potential - What they are and what they are not? *J Control Release* **2016,** *235*, 337-351.
102. Filipe, V.; Hawe, A.; Jiskoot, W., Critical evaluation of Nanoparticle Tracking Analysis (NTA) by NanoSight for the measurement of nanoparticles and protein aggregates. *Pharm Res* **2010,** *27* (5), 796-810.
103. Haiss, W.; Thanh, N. T. K.; Aveyard, J.; Fernig, D. G., Determination of Size and Concentration of Gold Nanoparticles from UV−Vis Spectra. *Analytical Chemistry* **2007,** *79* (11), 4215-4221.
104. Amendola, V.; Pilot, R.; Frasconi, M.; Marago, O. M.; Iati, M. A., Surface plasmon resonance in gold nanoparticles: a review. *J Phys Condens Matter* **2017,** *29* (20), 203002.
105. Bonaccorso, F.; Zerbetto, M.; Ferrari, A. C.; Amendola, V., Sorting Nanoparticles by Centrifugal Fields in Clean Media. *The Journal of Physical Chemistry C* **2013,** *117* (25), 13217-13229.
106. Njoki, P. N.; Lim, I. I. S.; Mott, D.; Park, H.-Y.; Khan, B.; Mishra, S.; Sujakumar, R.; Luo, J.; Zhong, C.-J., Size Correlation of Optical and Spectroscopic Properties for Gold Nanoparticles. *The Journal of Physical Chemistry C* **2007,** *111* (40), 14664-14669.
107. Fathi, F.; Rashidi, M.-R.; Omidi, Y., Ultra-sensitive detection by metal nanoparticles-mediated enhanced SPR biosensors. *Talanta* **2019,** *192*, 118-127.
108. Potenza, M. A.; Krpetic, Z.; Sanvito, T.; Cai, Q.; Monopoli, M.; de Araujo, J. M.; Cella, C.; Boselli, L.; Castagnola, V.; Milani, P.; Dawson, K. A., Detecting the shape of anisotropic gold nanoparticles in dispersion with single particle extinction and scattering. *Nanoscale* **2017,** *9* (8), 2778-2784.
109. Cai Q, C. V., Boselli L, Moura A, Lopez H, Zhang W, Araujo J, Dawson K., A microfluidic approach for synthesis and kinetic profiling of branched gold nanostructure. *ChemRxiv. Cambridge: Cambridge Open Engage* **2021**.
110. Monopoli, M. P.; Aberg, C.; Salvati, A.; Dawson, K. A., Biomolecular coronas provide the biological identity of nanosized materials. *Nat Nanotechnol* **2012,** *7* (12), 779-86.
111. Lundqvist, M.; Stigler, J.; Elia, G.; Lynch, I.; Cedervall, T.; Dawson, K. A., Nanoparticle size and surface properties determine the protein corona with possible implications for biological impacts. *Proc Natl Acad Sci U S A* **2008,** *105* (38), 14265-70.
112. Vilanova, O.; Mittag, J. J.; Kelly, P. M.; Milani, S.; Dawson, K. A.; Radler, J. O.; Franzese, G., Understanding the Kinetics of Protein-Nanoparticle Corona Formation. *ACS Nano* **2016,** *10* (12), 10842-10850.
113. Gasser, M.; Rothen-Rutishauser, B.; Krug, H. F.; Gehr, P.; Nelle, M.; Yan, B.; Wick, P., The adsorption of biomolecules to multi-walled carbon nanotubes is influenced by both pulmonary surfactant lipids and surface chemistry. *Journal of Nanobiotechnology* **2010,** *8* (1), 31.
114. Lundqvist, M.; Stigler, J.; Cedervall, T.; Berggård, T.; Flanagan, M. B.; Lynch, I.; Elia, G.; Dawson, K., The Evolution of the Protein Corona around Nanoparticles: A Test Study. *ACS Nano* **2011,** *5* (9), 7503-7509.
115. Ge, C.; Du, J.; Zhao, L.; Wang, L.; Liu, Y.; Li, D.; Yang, Y.; Zhou, R.; Zhao, Y.; Chai, Z.; Chen, C., Binding of blood proteins to carbon nanotubes reduces cytotoxicity. *Proceedings of the National Academy of Sciences* **2011,** *108* (41), 16968-16973.
116. Hu, W.; Peng, C.; Lv, M.; Li, X.; Zhang, Y.; Chen, N.; Fan, C.; Huang, Q., Protein Corona-Mediated Mitigation of Cytotoxicity of Graphene Oxide. *ACS Nano* **2011,** *5* (5), 3693-3700.
117. Deng, Z. J.; Liang, M.; Monteiro, M.; Toth, I.; Minchin, R. F., Nanoparticle-induced unfolding of fibrinogen promotes Mac-1 receptor activation and inflammation. *Nat Nanotechnol* **2011,** *6* (1), 39-44.

118. Chu, Z.; Zhang, S.; Zhang, B.; Zhang, C.; Fang, C. Y.; Rehor, I.; Cigler, P.; Chang, H. C.; Lin, G.; Liu, R.; Li, Q., Unambiguous observation of shape effects on cellular fate of nanoparticles. *Sci Rep* **2014,** *4*, 4495.
119. Lesniak, A.; Fenaroli, F.; Monopoli, M. P.; Åberg, C.; Dawson, K. A.; Salvati, A., Effects of the Presence or Absence of a Protein Corona on Silica Nanoparticle Uptake and Impact on Cells. *ACS Nano* **2012,** *6* (7), 5845-5857.
120. Niikura, K.; Matsunaga, T.; Suzuki, T.; Kobayashi, S.; Yamaguchi, H.; Orba, Y.; Kawaguchi, A.; Hasegawa, H.; Kajino, K.; Ninomiya, T.; Ijiro, K.; Sawa, H., Gold Nanoparticles as a Vaccine Platform: Influence of Size and Shape on Immunological Responses in Vitro and in Vivo. *ACS Nano* **2013,** *7* (5), 3926-3938.
121. Iversen, T.-G.; Skotland, T.; Sandvig, K., Endocytosis and intracellular transport of nanoparticles: Present knowledge and need for future studies. *Nano Today* **2011,** *6* (2), 176-185.
122. Linares, J.; Matesanz, M. C.; Vila, M.; Feito, M. J.; Gonçalves, G.; Vallet-Regí, M.; Marques, P. A. A. P.; Portolés, M. T., Endocytic Mechanisms of Graphene Oxide Nanosheets in Osteoblasts, Hepatocytes and Macrophages. *ACS Applied Materials & Interfaces* **2014,** *6* (16), 13697-13706.
123. Zhang, W.; Ji, Y.; Wu, X.; Xu, H., Trafficking of gold nanorods in breast cancer cells: uptake, lysosome maturation, and elimination. *ACS Appl Mater Interfaces* **2013,** *5* (19), 9856-65.
124. Kumari, S.; Mg, S.; Mayor, S., Endocytosis unplugged: multiple ways to enter the cell. *Cell Res* **2010,** *20* (3), 256-75.
125. Walkey, C. D.; Chan, W. C., Understanding and controlling the interaction of nanomaterials with proteins in a physiological environment. *Chem Soc Rev* **2012,** *41* (7), 2780-99.
126. Voigt, J.; Christensen, J.; Shastri, V. P., Differential uptake of nanoparticles by endothelial cells through polyelectrolytes with affinity for caveolae. *Proceedings of the National Academy of Sciences* **2014,** *111* (8), 2942-2947.
127. Hofmann, D.; Tenzer, S.; Bannwarth, M. B.; Messerschmidt, C.; Glaser, S.-F.; Schild, H.; Landfester, K.; Mailänder, V., Mass Spectrometry and Imaging Analysis of Nanoparticle-Containing Vesicles Provide a Mechanistic Insight into Cellular Trafficking. *ACS Nano* **2014,** *8* (10), 10077-10088.
128. Wang, L.; Yang, L.; Pan, L.; Kadasala, N. R.; Xue, L.; Schuster, R. J.; Parker, L. L.; Wei, A.; Tao, W. A., Time-Resolved Proteomic Visualization of Dendrimer Cellular Entry and Trafficking. *Journal of the American Chemical Society* **2015,** *137* (40), 12772-12775.
129. Kelly, P. M.; Aberg, C.; Polo, E.; O'Connell, A.; Cookman, J.; Fallon, J.; Krpetic, Z.; Dawson, K. A., Mapping protein binding sites on the biomolecular corona of nanoparticles. *Nat Nanotechnol* **2015,** *10* (5), 472-9.
130. O'Connell, D. J.; Bombelli, F. B.; Pitek, A. S.; Monopoli, M. P.; Cahill, D. J.; Dawson, K. A., Characterization of the bionano interface and mapping extrinsic interactions of the corona of nanomaterials. *Nanoscale* **2015,** *7* (37), 15268-76.
131. Herda, L. M.; Hristov, D. R.; Lo Giudice, M. C.; Polo, E.; Dawson, K. A., Mapping of Molecular Structure of the Nanoscale Surface in Bionanoparticles. *J Am Chem Soc* **2017,** *139* (1), 111-114.
132. Lara, S.; Alnasser, F.; Polo, E.; Garry, D.; Lo Giudice, M. C.; Hristov, D. R.; Rocks, L.; Salvati, A.; Yan, Y.; Dawson, K. A., Identification of Receptor Binding to the Biomolecular Corona of Nanoparticles. *ACS Nano* **2017,** *11* (2), 1884-1893.
133. Ye, D.; Anguissola, S.; O'Neill, T.; Dawson, K. A., Immunogold labeling reveals subcellular localisation of silica nanoparticles in a human blood-brain barrier model. *Nanoscale* **2015,** *7* (22), 10050-8.
134. Boselli, L.; Lopez, H.; Zhang, W.; Cai, Q.; Giannone, V. A.; Li, J.; Moura, A.; de Araujo, J. M.; Cookman, J.; Castagnola, V.; Yan, Y.; Dawson, K. A., Classification and biological identity of complex nano shapes. *Commun. Mater.* **2020,** 1, 1-12.

Chapter II

A Nanoscale Shape-Discovery Framework Supporting Systematic Investigations of Shape-Dependent Biological Effects and Immunomodulation

Wei Zhang[a,§], Hender Lopez[a,b,§], Luca Boselli[a,§], Paolo Bigini[c,§], André Perez-Potti[a,§], Zengchun Xie[a,§], Valentina Castagnola[a], Qi Cai[a], Camila P. Silveira[a], Joao M. de Araujo[a,d], Laura Talamini[c], Nicolò Panini[c], Giuseppe Ristagno[e], Martina B. Violatto[c], Stéphanie Devineau[a], Marco Monopoli[a], Mario Salmona[c], Valeria A. Giannone[a], Sandra Lara[a], Kenneth A. Dawson[a,f] & Yan Yan[a,g]**

a. Centre for BioNano Interactions, School of Chemistry, University College Dublin, Belfield, Dublin 4, Ireland

b. School of Physics and Optometric & Clinical Sciences, Technological University Dublin, City Campus, Kevin Street, Dublin 8, Ireland

c. Istituto di Ricerche Farmacologiche Mario Negri IRCCS, Via Mario Negri 2, Milan, Italy

d. Departamento de Física Teórica e Experimental, Universidade Federal do Rio Grande do Norte, 59078970, Natal, RN, Brazil

e. Department of Pathophysiology and Transplantation, University of Milan, Italy

f. Guangdong Provincial Education Department Key Laboratory of Nano-Immunoregulation Tumor Microenvironment, The Second Affiliated Hospital, Guangzhou Medical University, Guangzhou, 510260 Guangdong, P.R. China

g. School of Biomolecular and Biomedical Science, UCD Conway Institute of Biomolecular and Biomedical Research, University College Dublin, Belfield, Dublin 4, Ireland

*Email: Kenneth.a.dawson@cbni.ucd.ie; yan.yan@cbni.ucd.ie

Abstract

Since it is now possible to make, in a controlled fashion, an almost unlimited variety of nanostructure shapes, it is of increasing interest to understand the novel forms of biological control that nanoscale shape allows. However, *a priori* rational investigation of such a vast universe of shapes appears to present intractable fundamental and practical challenges. This has limited the useful systematic investigation of their biological interactions and development of innovative nanoscale shape-dependent therapies. Here we introduce a new concept of biologically relevant inductive nanoscale shape discovery and evaluation that is ideally suited to, and will ultimately become, a vehicle for machine learning discovery. Combining the unprecedented reproducibility and tunability of microfluidic flow nanochemistry syntheses, quantitative computational shape analysis and iterative feedback from biological responses *in vitro* and *in vivo*, we show that these challenges can be mastered, allowing shape biology to be explored within accepted scientific and biomedical research paradigms. Early applications identify significant forms of shape-induced biological and adjuvant-like immunological control.

Introduction

While the underlying principles and paradigms of nanostructure biological recognition and processing are quite different from those for biomolecules[1-5], this fundamental distinction is only beginning to be appreciated and applied in biology and medicine. This has resulted in much emphasis on the use of nanostructures as vehicles to 'carry' drugs and other cargoes, and more limited appreciation of the fundamental role that the nanostructure itself can play in biological control. Increasingly we understand that, in contrast to molecular ligand-receptor binding, numerous interactions distributed across the whole nanostructure-cell interface ('synapse') collectively induce a complex set of membrane and peri-membrane molecular events that we term 'bionanoscale recognition'. In determining the nanostructure's biological identity these processes take account of the details of the molecular presentation at the nanostructure's surface[3-6], the detailed organization of nanoscale shape features[7-10] and possibly other collective features yet to be discovered[11,12]. Membrane signaling responses are sensitive to stress relaxational phenomena on the nanoscale, and since relaxation of peri-membrane recognition processes occurs on comparable length- and timescales, we hypothesize that correlated nanoscale shape features can be detected by the cell recognition machinery as patterns of differentiated stress relaxation at the membrane[2]. However, extensive and detailed mechanistic investigations will be required to fully determine the mechanistic drivers of shape recognition on the nanoscale, greatly advanced and facilitated by recent advances in shape control and characterization[6,8].

Clearly, for the field of shape-dependent biological effects to progress we now require new concepts that will allow us to explore the science by systematic rather than (only) phenomenological trial and error investigations. Those of us seeking to develop this field of research see many competing priorities as to where to begin. In many ways we are like the early astronomers recording the apparently unlimited variety of events in the sky; somewhat awestruck by the infinitude of particle shapes and the extent of the nanoparticle shape universe, but unable as yet to make sense of the diversity and 'meaning' (degree of biological control) that can be exerted by nanoscale shape biology. Certainly, the few early snapshots of shape biology we do have suggest an extraordinary richness of responses, distinct from simpler biological interactions, and hint at practical possibilities to control immunological, metabolic and other system level responses[13-21]. Still for the field to grow we will need executable research programmes. Firstly, it is important for shape characteristic data to be reportable and transferrable between different laboratories and across different approaches to the science. Investigators will need to use different types of reactors and syntheses, but have transferrable structural inputs and biological outcomes in much the same way as we take

for granted in small or biomolecular biological investigations. Clearly, the usual physiochemical data alone (while important) do not specify shape and are insufficient for this purpose. Shape quantification is required to make meaningful, reportable and reproducible connections between shape distributions and biological outcomes. However, the major challenge is that in such a vast universe of shape, we need to know where to look for interesting biological effects without the impracticalities of guessing, or randomly searching a vast unknown shape space.

These and many other detailed technical features of nanoscale shape recognition have hitherto appeared to make systematic exploration of the relationship between nanoscale shape and biology a daunting proposition. For instance, (whether endogenous or man-made) even when the objective is to create a single shape identity, nanostructures are typically fabricated in weakly constrained assembly processes, leading to structural variations between individual particles and thereby to heterogenous distributions. That, and the difficulty in controlling the process itself, often make it difficult to reproduce shape distributions and characterize them in a meaningful way. However, this challenge has now been addressed by new computational geometry methods applied to electron microscopy that digitize, capture and analyze particle shape[22-26]. Those methods now allow us to check the reproducibility of shape and dispersion characteristics of the ensembles and to develop methods to ensure those standards are met. Such advances now allow us to carry out meaningful, reportable and systematic biological investigations of nanoscale shape, if we know which shapes are of interest.

In this paper we report on the next step, presenting a new ‘shape discovery’ approach that enables disciplined biological studies on interesting nanoscale shapes. Based on previous investigations of the mechanism behind the shape formation of branched GNPs, combining new tunable microfluidic nanostructure flow synthesis capacities with a quantitative framework that captures and quantifies nanoscale shape, we are able to vary shape in a flexible manner, reproducibly making shape ensembles[27]. Then coupling those microfluidic syntheses and digital shape characterization we use feedback from cellular (in the example discussed in this article: immune relevant) *in vitro* read-outs to inductively tune along a trajectory of different shape distributions to a regime of biological interest. Such inductively located particle ensembles are essentially ‘lead shape distributions’ for further investigations. As a proof of concept, using this discovery process, we have identified a new immunologically interesting nanoscale shape regime and confirmed the distinctive shape-dependent immunological properties by detailed analysis of antibody responses and B-cell receptor repertoire.

Results and Discussions

Definition and characterization of nanoscale shape ensemble distributions

While nanoscale shape ensembles have in the past been described using 'typical' electron microscope images and evocative names (*e.g.*, stars, flowers, and urchins) (Figure 2.1A), here we use statistical nanostructure image capture, digitization and quantitative computational analysis of shape ensembles. The key steps have been described before[7], including abstraction of minimal information capture of shape ensembles (*de facto* choosing a mathematical representation), and further condensation of that information (by principal component analysis) to allow it to be manipulated and analyzed, so here we only briefly summarize the new features. Summarily, we first capture and digitize hundreds of nanostructure electron microscopy generated surfaces (or more condensed surface-projected contour descriptions) and, in much the same way signals are Fourier analysed[28-31], transform the image contour of each particle into (typically) hundreds to thousands of discretized coefficients in a suitable representation (Figure 2.1B). At that point these coefficients are simply equivalent to the contour itself, but identification of the principal components in such a (suitable) representation produces a ranked ordering of the most relevant (principal component eigenvector) combinations to describe (and differentiate) those shapes. In this representation, particle surfaces are sampled by their projected contours and the presence of correlated typical nanoscale 'features' (*e.g.*, spikes and bumps) dominates the principal component description.

Each structure may then be represented by a single point in a low (often two or three) dimensional representation in the space of major principal component directions, representing an easily understood 'shape space'. As an example, several well-known shape types (formerly named as stars, flowers *etc.*) are easily differentiated as separated clustered 'shape space' identities (Figure 2.1C). We can also determine an 'average' shape and capture the largest particle-to-particle ensemble shape fluctuations (deviations) from that mean, thereby quantifying shape polydispersity in different directions in principal component space (Figure 2.1D), providing a rather complete description of the shape ensemble. This information can now be used to guide the optimization of flow reactor design and parameters until an appropriate level of reproducibility, 'purity' and dispersion of shape-distributions has been achieved.

Since biology is sensitive to nanoscale shape then the question of particle shape dispersity is important. Two nearby distributions with different mean shapes may still contain numbers of near identically shaped biologically active particles, making it is necessary to determine how

different two (statistically shape independent) ensembles are from each other. In Figures 2.1E-G we visually represent the issue of shape-distinctiveness between two distributions by comparing the spread in their shape distributions (projected onto the line connecting the two mean shapes) with the ‘distance’ between the mean shapes. It is also possible to numerically quantify the fraction of particles common to both, and to stipulate a threshold for independence. These points illustrate the type of shape characterization required for interoperable, reproducible and reportable nanoscale shape biology. As we show later, they also provide the basis for shape discovery.

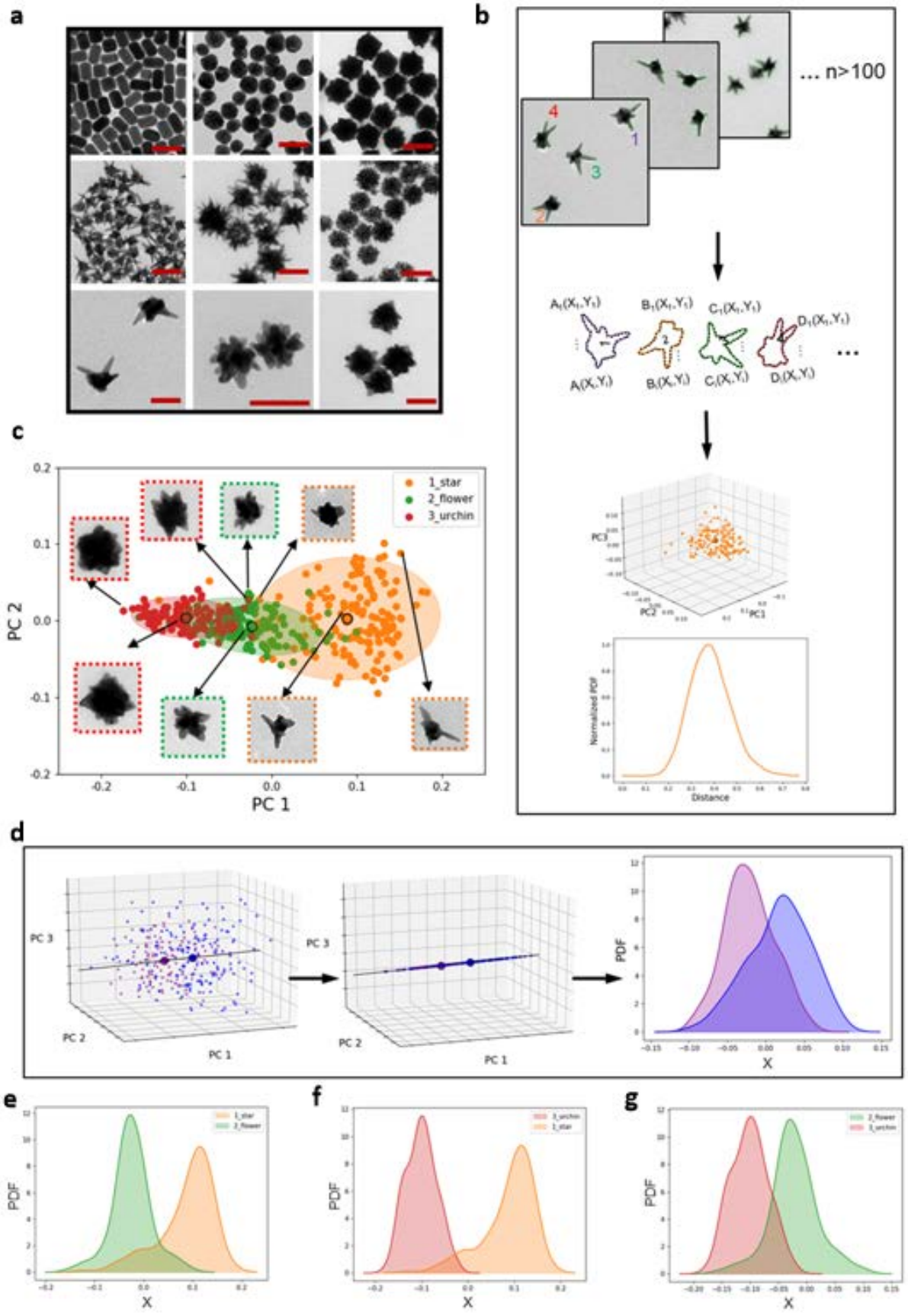


Figure 2. 1 Definition of nanoscale shape ensemble distributions.

(a) TEM micrographs showing the shape library of gold nanoparticles (GNPs), scale bar is 100 nm. (b) Schematic showing the process of nanoscale shape identification: capture and digitization of the contours of the nanostructures, shape space classification based on Fourier transform descriptor. (c) 2D PCA scatter plot and TEM micrographs for selected points (particles) of the first two principal components obtained from the analysis of the shape descriptor for three nanostructures: star, flower and urchin. The ellipses represent the regions which contain 95% of the points for each shape. (d) Schematic showing the process used to quantify the level of overlap between two nanoscale shapes. The coordinate X represents the line that joins the centre of gravity of the two shapes in the 3D scatter plot of the first three principal components. The projected points onto the line X are used to calculate the probability distribution function (PDF) for each shape which is then used to measure the level of overlap between two shape distributions. Examples of the overlap quantification between star and flower (e), urchin and star (f), and flower and urchin (g).

Synthesis of nanoscale shapes for biological application

Defined shape ensembles typically result from kinetically controlled growth around supercritical spherical-symmetry-broken nuclei (or 'seeds')[6,32-37]. Those seeds possess different crystal growth faces which can be differentially grown by control of the growth kinetics at the different interfaces[32,38,39]. Consequently, shape-ensemble-growth control features include the density and geometry of the growable (and quenchable) surfaces represented by the seeds[40-42], the nature of the reaction (composition and reactants)[24,43], the rate at which reactants can be deposited at the growing interface, and the nature and amount of the surface-active substance enhancing ('catalyzing') or inhibiting ('blocking') the growth kinetics at those surfaces[44-46]. As slow mixing heterogeneities occur on time-scales comparable (or greater) than interface growth kinetics, macroscopic reaction vessels limit our control of shape synthesis[47,48]. High reproducibility and tunability of shape ensembles require small mixing-volume flow chemistries in which suitable reactor control parameters provide fixed and reproducible constraints between mass transport, interface growth and quenching kinetics[49-52].

Here we employed a fast-mixing microfluidic reactor to tune across a large range of nanoscale shapes while achieving a high shape ensemble reproducibility. Our flow reactor consists of a Luer Lock T-junction, microfluidic sample injection shut-off valves, flow sensors, PTFE tubes (to allow the mixing and further reactions), reaction reservoirs for reagents, and a computer for real-time monitoring and control (Figure 2.2A, detailed microfluidic synthesis protocol and potential scalability are discussed in supporting method). The reactor leads to highly reproducible symmetry-broken gold seeds (Figure 2.2B-D). For instance, we observe that it is

almost impossible to resolve between three batches of independently produced 5 nm gold seeds using any macroscopic measurement (*e.g.*, UV-vis-NIR absorption spectra and differential centrifugation sedimentation (DCS) shown in Figures 2.2B and C). Benchtop tank reactor-based synthesis methods do have the advantage of being able to make much more material, but efforts to reproduce this seed type using macroscopic syntheses are challenging and typically lead to numerous failures before suitable batches are made (a typical example is shown in Figure S2.1).

Starting with high quality and reproducible seeds, we now synthesize a wide range of nanoparticle shapes using various reconfigurable microfluidic setups. As an example, a reactor, including four reservoirs: growth feed solution $HAuCl_4 \cdot 3H_2O$ in reservoir 1, seeds dispersed in trisodium citrate dihydrate (Na_3Cit) in reservoir 2, and reducing solution hydroquinone in reservoirs 3 and 4 (Figure 2.2E), was used to synthesize a family of branched gold nanoparticles. The resulting particles drain into the collection vessel containing the initial surface agent of interest (small exchangeable molecules, proteins, other biopolymers or surface-active agents) where the reaction is quenched, and nanoparticles remain dispersed. When proteins are used to quench the system, we observed that, beyond a critical protein concentration, different choices of surface quenching proteins lead to very modest changes in the final shape distribution (Figure S2.2). An unprecedented reproducibility of shape dispersions across independent microfluidic syntheses is confirmed by analyses of localized surface plasmon resonance (Figure 2.2F) and DCS size distribution (Figure 2.2G). These results are consistent with the well overlapped shape variance profiles (Figure 2.2H) and two-dimensional shape scatter plots (Figure 2.2I). Together these results illustrate the quite general capacity of such flow reactors to reproduce ‘typical shapes’ accompanied by reproducible and narrow distributions around that average (a comparison of the shape distribution between tank reactor- and flow reactor-based synthesis methods is shown in Figure S2.3). We note carefully that many flow reactors operate on complex and different principles, some of which are not fully understood as yet.[52-55]. Therefore, it is essential to tune parameters, and quantitatively compare the output shape distributions, rather than (only) replicate reactor design.

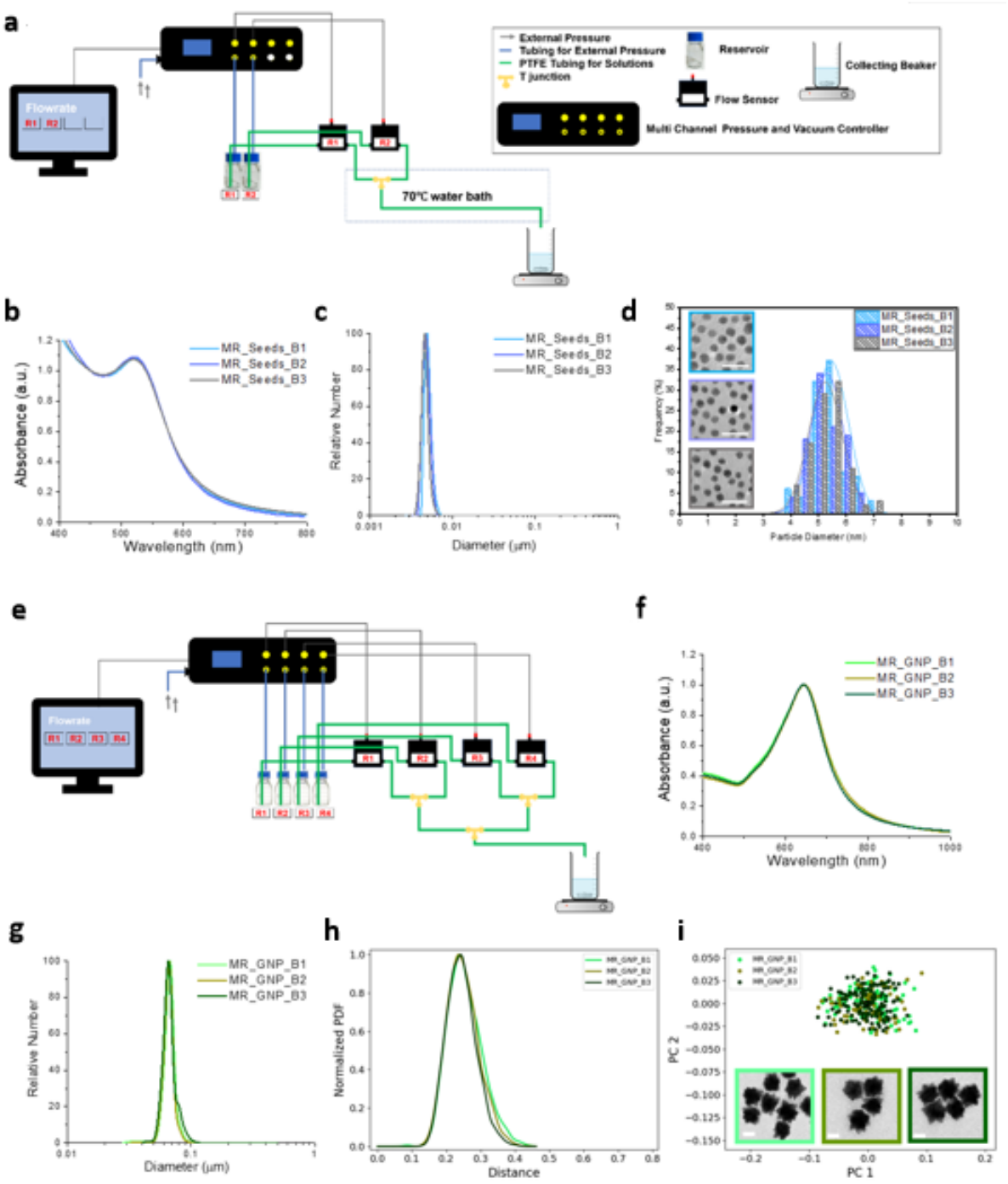


Figure 2. 2 A novel microfluidic reactor (MR) which can achieve high reproducibility and narrow shape distribution for 5 nm seeds and GNPs.

(a) Diagram of the microfluidic reactor synthesis set-up for 5 nm seeds. (b) Normalized UV-vis-NIR spectra absorption. (c) DCS analysis showing the high reproducibility of different batches of 5 nm MR_Seeds. (d) Representative TEM micrographs and TEM size distribution, scale bar is 20 nm. (e) Diagram of the microfluidic reactor synthesis set up for MR_GNPs. (f-g) Normalized UV-vis-NIR spectra absorption and DCS analysis showing the high

reproducibility of different batches of MR_GNP. (h) Shape variance expressed as a probability distribution function (PDF) over distance showing the similarity of three batches of MR_GNP. (i) 2D scatter plot of the first two principal components and representative TEM micrographs for each batch MR_GNP. The scale bar is 50 nm.

Inductive navigation along shape space trajectories of biological interest

Recognizing the unlimited numbers of potential shapes, and complete lack of *a priori* knowledge on which shapes are of biological interest, unguided combinatorial search screening would not be efficient. Therefore, to develop the selection of interesting shape-space regimes we propose an inductive approach. Using a flow reactor, we iteratively make small changes in the reaction parameters, represent the shape ensemble in principal component shape space, and then at appropriate check points, use a cellular readout to decide if we are moving towards (or away from) an interesting area of shape space. Thereby we can build a 'shape learning trajectory' by simultaneously varying (single and multiple) combinations of different flow reactor parameters, using shape computation to calibrate the scale of the shape-space increments required to reach the next shape ensemble along the trajectory. There are some novel and noteworthy features associated with searching across shape space, derived from the fact that we are comparing distributions, not simply fixed shapes. For instance, since this is essentially an experimentally based 'gradient optimization' search process, when exploring such shape learning trajectories, if we make overlarge changes in the flow reactor parameters in a single step the change in typical shape is discontinuous and it is difficult to discern which aspect ('direction') of the flow reactor leads to the new (desirable) structural features or specific biological outcomes. That information therefore may not be helpful in choosing which direction to take in the next move. However, there is also a trade-off between the need for these close (small change) points along the learning trajectory that contribute to directional guidance and the fact that nearby distributions have significant 'overlaps' corresponding to numerous similar structures being common to the two distributions.

Next, at appropriately chosen steps along this sequence of synthesis-characterization points, we use biological readouts to select useful directions ('slopes') in the shape-space. Biological readouts from such contiguous distributions (containing many common structures) do not contribute to new biological information, so computational methods of shape characterization must be used to identify when (along the trajectory) there is sufficient independence to learn anything new about the biology. This inductive shape learning process culminates in a useful 'lead location' in shape space. Evidently all of these steps can (and will be) implemented *via* automated machine learning to optimize the shape learning process. Here, as proof of concept

we seek to show how the essential elements of the discovery process may be framed. Since we apply the approach manually, on a relatively limited scale (and seek *in vivo* immunological readouts), we use a high-dimensional ('transcriptome') read-out to choose directions along the learning trajectory.

Given the fact that we are dealing with a multidimensional, highly non-linear and *a priori* unknown relationship between flow reactor parameters and output shape, increments in flow parameters along the shape learning trajectory have to be locally adapted to the regime being explored (Figures 2.3A and B). For instance, different flow parameter settings (parameter details in Table 2.1) associated with MR_GNP06 and MR_GNP07 lead to distributions that are barely different statistically (Figure 2.3C). Many of those along the rest of the trajectory are sufficiently independent to be of interest, but in some cases represent such large steps in shape that it is necessary for the shape learning process to return close to previous coordinates to recommence more smooth progressions. It is also worth noting that while highly congruent ('nearby') shape ensembles are useful to understand shape-space tuning in synthesis, meaningful structure-function relationships rely on an independent (investigator-controlled) variable involving the evolution of shape identity along which the (dependent variable) biological read-outs can be measured. For instance, the shape trajectory (Figures 2.3D and E), in distinction to a shape learning trajectory, is composed of nearly independent ensembles and biological readouts from these will reflect fundamental changes in shape and would constitute the basis of a useful 'structure-function' relationship. Along this trajectory each particle shape ensemble is labelled by a different color, large dots with black borders represent the 'typical' (mean) shape, and illustrations of the overlaps are given in various examples by confidential ellipses (Figure S2.4). The associated physiochemical properties, including surface plasmon resonance (Figure 2.3F), effective size distribution of different shapes (Figure 2.3G), hydrodynamic diameter and zeta potential (Table 2.2), are shown.

To direct the trajectory illustrated in Figure 2.3A we use principal component analysis (PCA) of the whole transcriptome data reported previously[7] to identify the key changes in biological responses. Those outcomes then help us choose the direction of the learning trajectory to achieve a target biological outcome. Mouse dendritic cells (JAWSII) were treated with the three distinct shapes (marked *in vitro*_GNPa-c, Figure 2.3H) along the shape learning trajectory (The details of synthesis and PCA of the transcriptome are described in the supporting information.). The transcriptomic changes are captured in the principal component analysis (Figure 2.3I), illustrating that the transcriptome of *in vitro*_GNPc treated cells returned to the untreated transcriptome where the adjacent shape (*in vitro*_GNPb) gave rise to a distinct transcriptome from the untreated one. It suggests a sharply shape-responsive regime of interest (*e.g.*, region between *in vitro*_GNPb and *in vitro*_GNPc). In summary, despite the fact

that the intrinsic non-linear relationships between shape and flow reactor conditions combined with the complex dependence of biological readouts on shape make it far from obvious how to *a priori* tune shape for biological outcomes, the inductive process outlined here converges relatively quickly along the directions of primary interest.

We next sought to illustrate these concepts in a scientifically interesting, challenging and practically important shape regime within an *in vivo* setting. This example also illustrates the practical role of computational shape characterization in the 'hand-over' of target shape distributions between synthetic approaches (including benchtop tank reactor-based synthesis) that may make the whole workflow, including scale-up, feasible for extended *in vivo* studies.

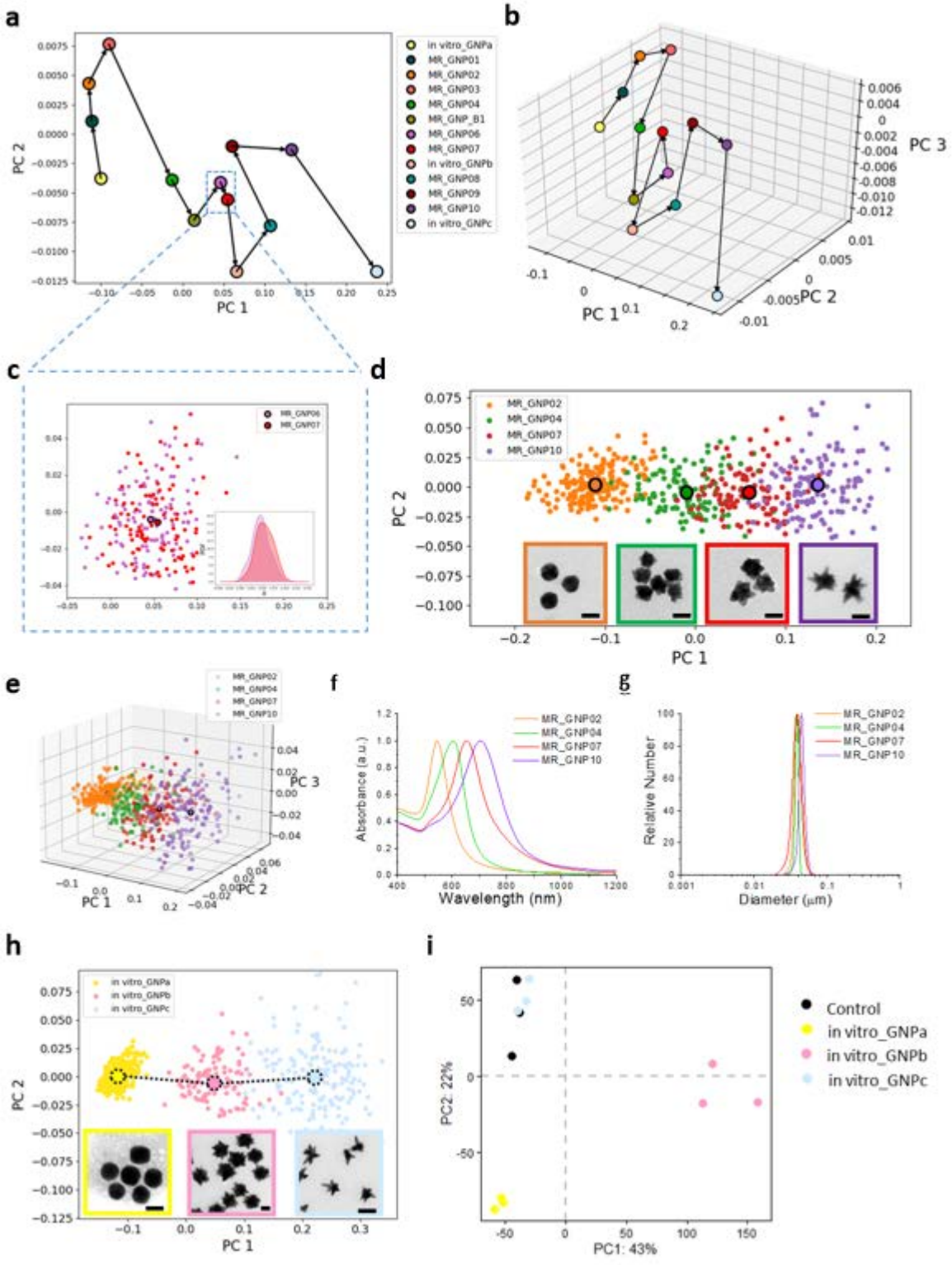


Figure 2. 3 Inductive navigation by microfluidic synthesis along shape space trajectories of biological significance.

(a and b) 2D and 3D PCA scatter plots of the first two and three principal components, showing the centre of gravity of each shape ensemble to illustrate the “shape learning trajectory”. The arrows indicate the shape tuning direction. The synthesis methodology for the shape trajectory is reported in supporting information. (c) 2D PCA scatter plot of the first two principal components for MR_GNP06 and MR_GNP07. The inset shows the overlap quantification for these two shapes. (d and e) 2D and 3D PCA scatter plots of the first two principal components for four different shapes. The larger dots with black borders represent the centre of gravity of each shape distribution. The insets show TEM micrographs of each shape. The scale bar is 50 nm. (f) Normalized UV-vis-NIR spectra absorption showing the LSPR (localized surface plasmon resonance) of different shapes. (g) DCS analysis showing a similar size distribution of different shapes. (h) 2D scatter plot of the first two principal components for three distinct shape ensembles used in the previously reported transcriptome study[7]. The larger dots with black dash borders represent the centre of gravity of each shape distribution. The insets show representative TEM micrographs for each shape, and scale bar is 50 nm. (i) PCA analysis illustrating distinctively different transcriptome profiles induced by the three shape ensembles. The percentages shown in the axis labels represent the variance explained by each PC.

Decisive biological readouts for nanoscale shape

Here we illustrate the larger ‘shape discovery’ potential of the inductive screening approach described by searching for novel specific adjuvant-like shape-controlled immune responses. We have investigated the ‘lead’ shape ensembles (using large-scale batches, defined *via* their shape geometry) to allow *in vivo* investigation of the shape region of interest. After optimization particle batches (*i.e.*, *in vivo*_GNP(B)) were prepared that occupy the target shape region of interest (*i.e.*, region between *in vitro*_GNPb and *in vitro*_GNPc) (Figure 2.4A), being of an acceptable structural quality and lipopolysaccharide (LPS) free (the full characterization is shown in Figure S2.6). The other two shape ensembles (*i.e.*, *in vivo*_GNP(A) and *in vivo*_GNP(C)) with the gravity centre of shape distribution shifted away from the target shape region were used as the control. Healthy rats were subcutaneously injected with the same number of gold nanoparticles (Figure 2.4B). The concentration of serum IgG at Day 0, 30, and 60 (determined by ELISA) was observed to have a general gradual increase of circulating IgG for all the groups over the immunization course, with the steepest rise of the *in vivo*_GNP(B)-treated group (Figure 2.4C). At Day 60, the concentration of circulating IgG was shown to be approximately two- to three-fold higher in *in vivo*_GNP(B)-treated rats than the *in vivo*_GNP(A)- and *in vivo*_GNP(C)-treated groups, respectively. In contrast, the *in vivo*_GNP(C) treated-group exhibited comparable levels of IgG to the control and *in vivo*_GNP(A)-treated groups (Figure 2.4C). To ensure the accuracy of circulating IgG

concentration, we carried out the ELISA measurements by independent operators with each assay performed in replicates (some examples were shown in Figure S2.6).

Higher IgG levels are prevalent in many autoimmune diseases, such as systemic lupus erythematosus (SLE) and rheumatoid arthritis (RA)[56], and since the shape ensembles studied here had no exogenous biological antigen, the origin of this rise in IgG levels was explored further for such effects. We assayed serum IgM from Day 0-7 and for several autoantibodies in the *in vivo*_GNP(B)-treated group. The concentration of IgM at Day 5 was found to be about 2-fold higher than Day 0 in the *in vivo*_GNP(C) group, whereas the *in vivo*_GNP(A) treatment or control did not exhibit significant change over the time (Figure 2.4D). To ensure reproducibility, we repeated the ELISA measurements with independent operators (some examples are shown in Figure S2.7). One of the autoantibodies against complement C1q (anti-C1q), consistently showed a three-fold elevation in ELISA for the *in vivo*_GNP(B)-treated group, in contrast to unchanged levels of anti-C1q antibodies observed in the control and *in vivo*_GNP(A)-treated groups (Figure 2.4E).

Changes in B cell tolerance are often associated with an increase in B cell clonal diversity, resulting in increased self-reactivity[57], and B cell clonal diversity is also an important proxy for BCR and antibody repertoires. We therefore harvested the draining lymph node B cells and analyzed the B cell receptor (BCR) repertoire using next generation sequencing. It was shown that *in vivo*_GNP(B) treatment resulted in a significantly more diverse BCR repertoire than the control and *in vivo*_GNP(A)-treated groups (Figure S2.8, bioinformatic analysis parameter details in Table 2.3), consistent with the elevation of autoantibodies. Evidently these results are quite striking, suggesting a role for nanoparticle shape in the over-riding of usual autoreactive controls. Certainly, under usual circumstances, the frequency and affinity of autoreactive B cells is highly regulated through multiple mechanisms at several sites, including central tolerance in the bone marrow and peripheral tolerance in the spleen, lymph nodes, and other tissues. While B cell tolerance is centrally enforced, leaky self-reactive mature naïve B cells are sometimes found in the periphery[58], and their fate is subject to a number of microenvironment contextual signals (such as the specifics of antigen presentation, innate signaling (*e.g.*, TLRs), dendritic cell input (*e.g.*, BAFF), and T cell collaboration (*e.g.*, CD40L)). We therefore hypothesize that exquisitely controlled nanoparticle shape ensembles are able to modulate those contextual signals and lead to a form of self-recognition. While it is not the purpose of this paper to enter into detailed mechanistic investigations of such phenomena, we consider these results constitute a striking example of a definitive readout for nanoscale shape regulation, and the potential for inductive discovery processes.

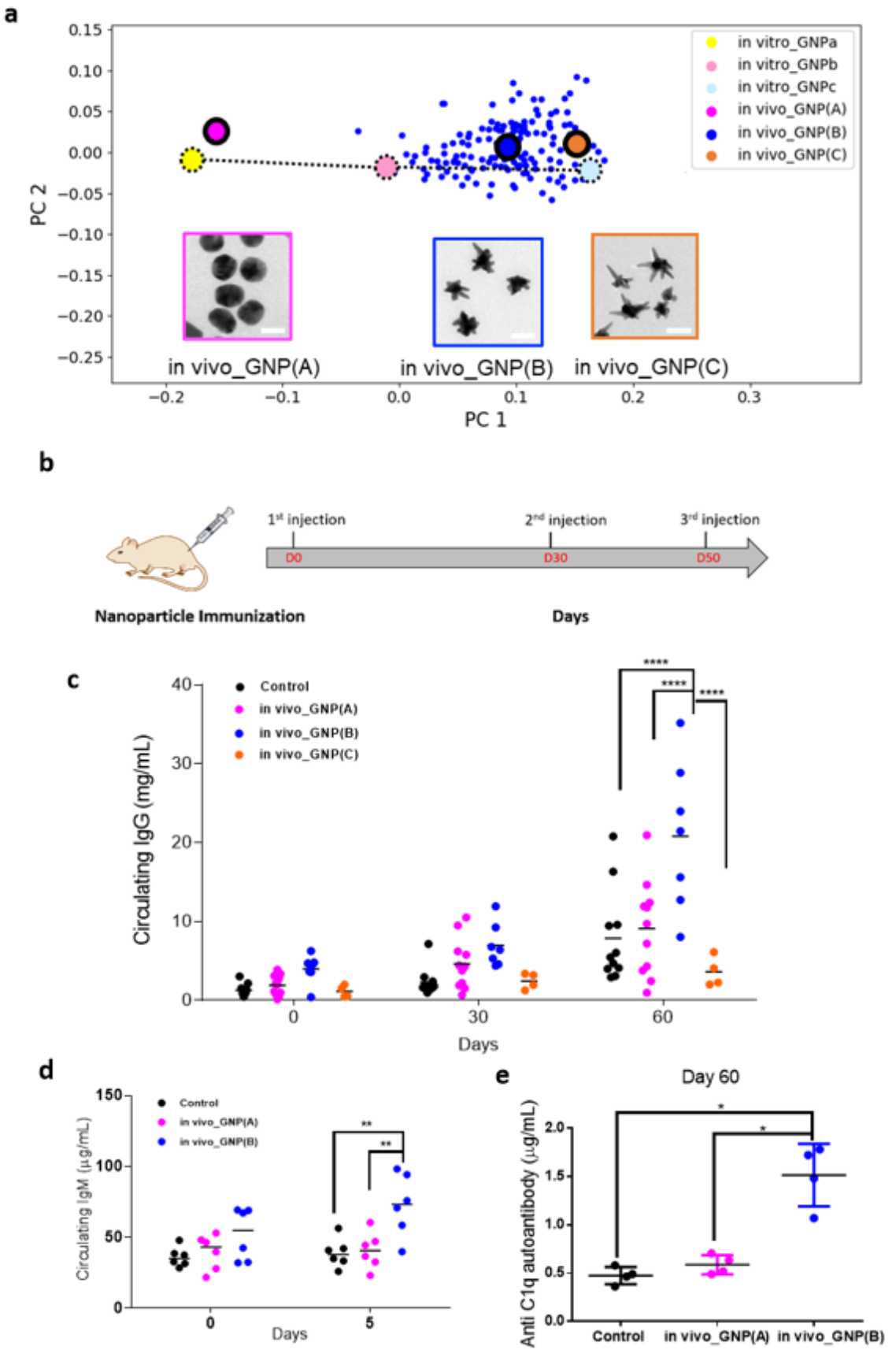


Figure 2. 4 Antibody responses to nanoscale shape ensembles.

(a) 2D PCA scatter plot showing the shape distribution of *in vivo*_GNPs shape ensembles in relation to the biological responsive shape regime identified by *in vitro*_GNPs. The larger dots with black borders represent the centre of gravity of each shape distribution. Representative TEM micrographs of each shape are shown, and the scale bar corresponds to 50 nm. (b) Subcutaneous immunization schedule in rats. (c) Levels of circulating IgG determined by ELISA. Data are presented as dot plots of individual rats, showing the mean of duplicates. Statistical significance was determined by two-way ANOVA analysis using the Tukey's test, ****p<0.0001. (d-e) Circulating IgM and anti-C1q autoantibodies induced by *in vivo*_GNP(A) and *in vivo*_GNP(B) were evaluated by ELISA. Data are presented as dot plots of individual rats, showing the mean of duplicates. Statistical significance was determined by ANOVA analysis using the Tukey's test, *p<0.05, **p<0.01.

Conclusion

In this paper we propose a generally applicable framework that will enable the discovery of important biological and medical outcomes in nanoscale shape biology, while supporting the systematic unraveling of the nanoscale shape-dependent biological mechanisms. Our purpose was to highlight and illustrate that these two mutually supportive agendas of 'discovery' and systematic mechanistic can be achieved within the same conceptual framework and *via* workflows that are now firmly within the realm of well-defined scientific investigation. Given the novelty of the field, we chose to illustrate the ideas with an important example of nanoscale shape control that could lead to key practical outcomes.

We stress that much has yet to be achieved to understand the detailed mechanisms in shape control biology. The issues are subtle, and it would be premature to make definitive statements on that topic. For instance, we have known for some time that it is the composition and collective organization of the surface biomolecular (corona) layer, rather than only individual surface molecules, that are recognized by nanoscale biological mechanisms specifically evolved for that purpose[1-3]. And it is that collective recognition that determines many early and later downstream biological outcomes. There are certainly common features between that corona paradigm (largely based on multiple and simultaneous molecular motif engagements between the nanoscale surface and cognate cell receptors and other interacting membrane proteins) and repetitive features on various particle geometries, such as those reported here. Indeed, potentially spatially correlated nano-shape features could be recognized and transduced at the cell membrane as patterns of differentiated stress relaxation but could also (at least to some degree) be coupled to collective surface molecular recognition. That issue of decoupling the surface and shape effects is subtle and will take time to clarify.

It will be necessary to acquire detailed and elaborated evidence to reliably assign mechanisms of shape regulation *in vivo*, including those shape-induced self-immune responses discussed here. Still, practically speaking our observations of *in vivo* shape regulation are potentially highly significant, suggesting the possibility to control the breaking of balances of immunity and tolerance. We note carefully that such effects may point towards a novel and unique avenue to develop adjuvants and immunotherapies possessing local regulatory functions without affecting systemic immune tolerance. The implications for significantly improved safety and efficacy in vaccine applications are clear. In a scientifically related issue, the results reported here also raise questions related to the environmental impact of processes producing rich and novel varieties of nanoscale shape fragments. To confirm such a link to human health and autoimmune disease it would be necessary to carefully and fully investigate a variety of material types, conditions and species in the ecosystem. However, given the longstanding suspicion that autoimmune diseases are linked to environmental dusts the issues involved appear significant and should be investigated[59-61]. In that context, we note carefully that the difficulty in locating these effects (indeed requiring inductive searching) may suggest such autoimmune-shape biology effects are not ubiquitous and may be confined to highly specific shapes. The good news is that their limited nature could make conceivable their isolation and elimination, with significant implications for human health. Shape searches using the full machinery of inductive shape learning (or otherwise very insightful hypotheses) to isolate and identify these effects may therefore constitute a new practical, feasible and scientifically well-founded frontier in hazard identification for a new generation of environmental health investigations.

However, these considerations all represent specific tasks that should be considered by science. The central purpose of this paper was to create conceptual and practical order out what at first sight looks like a nanoscale shape cacophony. That is to create a new systematic shape-discovery framework that will lead to interesting discoveries. We believe that automation and implementation of machine learning of the framework of inductive shape discovery presented in this article will play a central role in the search for valuable (or elimination of harmful) shape space entities. Simultaneously it will allow the shape-biology-medicine research enterprise to be moved onto well-established scientific principles, enabling widely shared reproducible and validated results from scientific research. That will frame a secure basis on which to pursue future research in the field.

Methods

Chemicals

The following chemicals were purchased from Sigma Aldrich and were of highest available purity and used as received: Hydrogen tetrachloroaurate trihydrate ($HAuCl_4 \cdot 3H_2O$, ≥99.9%), trisodium citrate dihydrate ($C_6H_9Na_3O_9$, meets USP testing specifications), potassium carbonate (K_2CO_3, ≥99%), tannic acid ($C_{76}H_{52}O_{46}$, ACS Reagent Grade), hydroquinone (HQ, $C_6H_6O_2$, ≥99%), silver nitrate ($AgNO_3$, ≥99.9%), glycerol ($C_3H_8O_3$, ≥99%), sucrose ($C_{12}H_{22}O_{11}$, ≥99.5%), dodecane ($CH_3(CH_2)_{10}CH_3$, ≥99%), Clean water (CHROMASOLV® Plus, for HPLC), bis(p-sulfonatophenyl)phenylphosphine dihydrate dipotassium salt (BSPP, $C_{18}H_{17}K_2O_8PS_2$, 97%), bovine serum albumin (BSA, lyophilized powder, ≥98%), human serum albumin (HAS, lyophilized powder, ≥98%), ovalbumin (OVA, lyophilized powder, ≥98%). Sodium hydroxide (NaOH, ACS Reagent Grade) was purchased from Fluka. Poly(vinyl chloride) (PVC) calibration standard for differential centrifugal sedimentation (DCS) measurements (263 nm) was purchased from Analytik Ltd.

Microfluidics

Oil-free air compressor (8 bar) and Luer Lock T-junctions (microfluidic manifold 3 port small kit) were purchased from Darwin microfluidics. The other microfluidic equipment and adjuncts were purchased from Elveflow including microfluidic flow controller (OB1 MK3+, channel pressure range 0-8 bar), microfluidic flow sensors (MFS, flow rate range 0-5 ml/min), PTFE tubing (1/16" OD × 1/32" ID, 50 m), microfluidic reservoir for 100 ml bottles (bottle cap with two 1/4" - 28 threaded ports) and microfluidic fittings (1/4" - 28 thread).

Computational shape analysis

To avoid aggregation due to drying effects and to obtain well-dispersed imaging of nanoparticles (NPs), sample preparation for TEM imaging used a modified method based on the previously reported protocol[7]. Grids (Agar Scientific) were pre-treated with a glow discharger and a 1 µl of 1×10^{10} - 1×10^{11} NPs/ml sample solution was deposited on the grid. Imaging was performed using a FEI Tecnai G2 20 Twin TEM at 200 kV, with magnifications no less than 19000×. TEM images containing well-spread NPs were used to extract their contours following the protocol previously reported by our group[7]. The obtained contours were then used to analyze the shapes of the different batches of NPs.

Nanoparticle immunization

All rat work was performed in accordance with institutional guidance, the NIH Guide for the Care and Use of Laboratory Animals (2011 edition) and EU directives and guidelines (EEC Council Directive 2010/63/UE). Adult CD (Charles River) male rats (approx. 300 grams in body weight) were housed paired in individually ventilated cages (Tecniplast S.p.A., Varese, Italy) and maintained under specific pathogen-free conditions in the Institute's animal care facilities.

They received food and water *ad libitum* and were regularly checked by a certified veterinarian responsible for animal welfare supervision and experimental protocol revision. The investigators were not blinded to allocation during experiments and outcome assessment. To exclude contamination, all the procedure concerning animals were performed in a Class 2 laminar flow hood following strict precautions. The reagents used for NP preparation were opened inside the laminar flow fume hood. Immediately before the treatment, NPs were dissolved in water to reach a concentration of 1.5×10^{12} NP/ml in a final volume of 500 µl and injected subcutaneously into the loose skin over the interscapular area. NPs and water were administered at day 1 (1st boost), at day 30 (2nd boost) and at day 50 (3rd boost). Rats were randomly assigned to the following treatment groups: control (n = 11 for IgG ELISA; n = 6 for IgM ELISA), *in vivo*_GNP(A) (n = 11 for IgG ELISA; n = 6 for IgM ELISA; n = 4 for anti-C1q autoantibody ELISA), *in vivo*_GNP(B) (n = 7 for IgG ELISA; n = 6 for IgM ELISA; n = 4 for anti-C1q autoantibody ELISA), *in vivo*_GNP(C) (n = 4 for IgG ELISA).

Blood collection and ELISA

Under general anesthesia (continuous flow of 5% isoflurane/oxygen mixture for induction and 2-3% for maintenance) blood was taken from the lateral tail vein at each indicated time point. Day 0 is defined as the day of injection. Blood on Day 0 was taken before the injection. Blood was collected into EDTA-tubes and centrifuged for 15 minutes at 1500g at 4°C. The supernatant (plasma) was aliquoted and immediately frozen. Animals were sacrificed on Day 60. Blood was collected through a terminal cardiac puncture and then animals were euthanized by CO_2 inhalation.

The levels of circulating IgG and IgM in rat plasma were determined by Ready-SET-Go! ® total rat IgG and IgM ELISA (Cat. No. 88-50490 and 88- 50540, eBioscience) with pre-dilution of 1:250,000 and 1:5,000 in the provided sample diluent for IgG and IgM, respectively. The concentration of autoantibody anti-C1q IgG was determined by rat anti-complement 1q antibody ELISA (Cat. No. MBS722996, MyBioSource, USA) with pre-dilution of 1:10 in provided sample diluent.

Author Information

Corresponding Authors

*Email: Kenneth.a.dawson@cbni.ucd.ie; yan.yan@cbni.ucd.ie

Author contributions

Y.Y. and K.A.D. conceived and developed the overall concept of shape specification and its relationship to biology, developed the methods and approaches, and wrote the paper. Y.Y. led the project overall, including biological development, established and developed current working methods for the materials in biology, and interpreted and supervised the experiments. W.Z. was central to the setup of the tunable shape library synthesis and variable surface coatings and wrote the paper. Q.C., V.C. and L.B. developed the original microfluidic synthesis and foundations of the work, and assisted by C.P.S., supported and carried out some of the synthetic and characterisation experiments. Q.C., V.C. and L.B. initiated the microfluidic setup and synthesis protocol, and W.Z. further advanced this to the current status. Z.X. carried out much of the current microfluidic synthesis, characterization and contour analysis. V.C, L.B, M.M, Q.C. and C.P.S all played central roles (at different periods) in the introductions of shape synthesis overall in the laboratory, from the early batch syntheses, and ensuring the establishment of 'clean syntheses'. W.Z. (and Laurent Adumeau, CBNI) subsequently brought the clean room syntheses protocols to the highest operational level. H.L. and J.M.A wrote the original computational analysis code (with support from Alirio Moura, CBNI). Adaptations were made by H.L, and he and W.Z. drove the calculations for the present paper. W.Z and Z.X. performed the computational analysis of shapes and this was checked by H.L. and J.M.A. A.P.P. and S.D. performed the ELISA. A.P.P. carried out the BCR repertoire analysis. N.P. was responsible for FACS analysis for cell sorting from spleens of rats. S.L. and V.A.G. assisted with the extraction of B cells, S.L. checked B cell data and V.A.G assisted with BCR PCR analysis. P.B and M.S. led the first efforts *in vivo* set up, and L.T., M.B.V. and G.R. performed the *in vivo* immunization and tissue collection. §. These authors contributed equally: Wei Zhang, Hender Lopez, Luca Boselli, Paolo Bigini, André Perez-Potti and Zengchun Xie.

The authors declare no competing interests.

Animal ethics: The animal work was reviewed by the IRCCS-IRFMN Animal Care and Use Committee (IACUC) and approved by the Italian "Istituto Superiore di Sanità" (code: 42/2016-PR).

Acknowledgement and Fund Sources

The authors acknowledge that this publication has emanated from research supported in part by grants from Science Foundation Ireland (15/SIRG/3423, Y.Y. and 17/NSFC/4898, K.A.D.). K.A.D. acknowledges the funding of Guangdong Provincial Education Department Key Laboratory of Nano-Immunoregulation Tumor Microenvironment (2019KSYS008). H.L., V.C., and Q.C. acknowledge financial support from the Irish Research Council (EPSPD/2015/5,

H.L.; GOIPD/2016/128, V.C.; GOIPG/2014/874, Q.C.). W.Z., Z.X. and Q.C. acknowledge the Chinese Scholarship Council (agreement no. 201706220062, 201806220054 and 201408300003). The work described in this article, prior to reaching this level, evolved over some years with contributions from authors and other previous researchers at the Centre for BioNano Interactions (CBNI), University College Dublin. They were helpful in bringing the work to the present stage of development. Inge Nelissen performed the transcriptome array measurements. Marco Monopoli was involved in early phases when syntheses and dispersions of shape particles were quite irreproducible and helped parameterize the nature of the problems involved. David Garry was also involved in the early phases during the isolation of B cells from the spleen and the creation of the cDNA library. Laurent Adumeau supported the development of the clean room syntheses protocols. Alirio Moura contributed to the writing of the original computational analysis code.

TOC

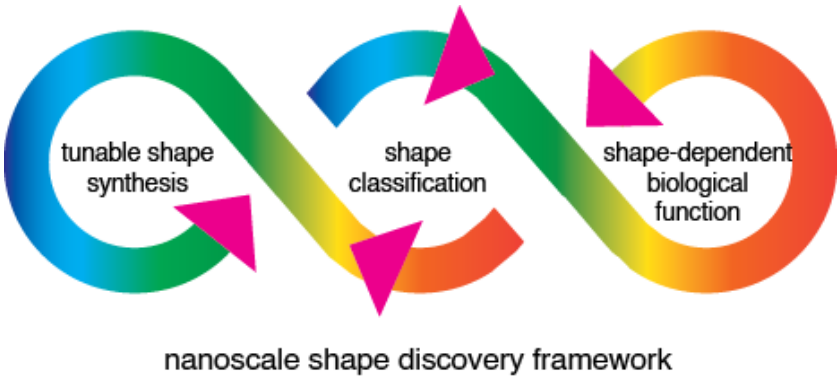


Figure 2. 5 TOC

Supporting Information

Supporting methods

Synthesis of TR_Seeds and MR_Seeds

TR_Seeds were synthesized *via* a previously reported method[18]. Briefly, TR_Seeds were prepared by adding a reducing solution (10 ml) containing trisodium citrate (0.068 mmol, 6.80 mM), tannic acid (0.003 mmol, 0.29 mM) and potassium carbonate (0.013 mmol, 1.25 mM) to 40 ml of gold precursor ($HAuCl_4 \cdot 3H_2O$ 0.013 mmol, 0.32 mM) at 60 °C under vigorous stirring. The mixture was then heated to reflux for 2-3 min and consequently cooled to room temperature. The pH was adjusted to 8.5 using NaOH (500 mM). TR_Seeds were filtered with a 0.2 µm filter before use.

For the synthesis of MR_Seeds, 50 ml of $HAuCl_4 \cdot 3H_2O$ (2.8×10^{-5} mol, 0.56 mM) was placed in reservoir 1 (R1, inlet 1), 50 ml of reducing agent containing trisodium citrate (8.2×10^{-5} mol, 0.4 mM), tannic acid (3.5×10^{-6} mol, 0.07 mM), and potassium carbonate (1.5×10^{-5} mol, 0.3 mM) was placed in reservoir 2 (R2, inlet 2). The reagents were mixed in the T-junction at 70°C and the outlet tubing was also immersed in a 70°C water bath during the synthesis. The flow rate was 2700 µl/min. The tubing length between the reservoir and flow sensor was 30 cm, the tubing length between the flow sensor and T-junctions was 10 cm and the total length of outlet tubing was 10 m. 50 ml of the resulting particles were collected in a 250 ml beaker and the particles were stirred to cool to room temperature. 500 mM NaOH solution was used to adjust the dispersion pH to 8.5 before characterization. MR_Seeds were filtered with a 0.2 µm filter before use.

Synthesis of TR_GNP and MR_GNPs

TR_GNP was synthesized by the same method as described for *in vivo*_GNP(A), details are discussed below in the synthesis of *in vivo*_GNPs section.

By using the high reproducibility and homogenous size distributed MR_Seeds, we succeeded in synthesising high reproducibility branched MR_GNPs using a similar microfluidic set up. Briefly, four reservoirs containing 50 ml of reagents were used. Four flow sensors were placed between reservoirs and T-junctions to monitor the flow rate. All of the flow rate was set as 2700 µl/min. The reaction took place in PTFE tubing and T-junctions. The tubing length between the reservoir and flow sensor was 30 cm, the tubing length between the flow sensor and T-junctions was 10 cm and the total length of outlet tubing was 15 m.

MR_GNP_B1-B3 (characterization shown in Figure 2. 2 f-i) synthesis in the microfluidic system was initiated through the mixing of three solutions: growth solution (reservoir 1), seeds solution (reservoir 2) and reducing solution (reservoir 3 and 4). 5×10^{-5} mol, 276 mM $HAuCl_4 \cdot 3H_2O$ was added in reservoir 1, 1.5×10^{-5} mol, 1% w/v sodium dihydrate citrate and 300 µl MR_Seeds was added in reservoir 2, and 2×10^{-4} mol, 200 mM hydroquinone was added to reservoirs 3 and 4, respectively. The resulting particles were collected in a beaker with BSA (1 mg/ml, 10 ml) and stirred at 37 °C for 30 mins.

For the synthesis of MR_GNP01-10, all the microfluidic set up parameters remained the same, the tuning of shapes was achieved by changing the reagents quantities as shown in Table 2.1.

UV-Vis-NIR spectroscopy

All UV-Vis-NIR spectra were recorded on an Agilent Cary 6000i UV-Vis-NIR spectrophotometer in the range 400-1200 nm. The measurements were carried out using a quartz cell with a path length of 1 cm.

Differential centrifugal sedimentation (DCS)

DCS experiments were run using a CPS instrument (DC24000). An 8-24% sucrose gradient either water-based or phosphate-buffered saline (PBS)-based (particles with protein coating) was used. Calibration was performed using polyvinyl chloride (PVC) particles (0.263 µm, Analytik Ltd.), following successful PVC standard calibration, 0.1 ml of the sample was injected for analysis.

Transmission electron microscopy (TEM)

Samples for TEM imaging were prepared by evaporating ca. 10 µl of the nanoparticle suspension onto formvar-coated copper grids (Agar Scientific). Imaging was performed by

using FEI Tecnai G2 20 Twin TEM and data analysis using ImageJ. Images for computational shape analysis were taken by a specific method mentioned in the main text computational analysis part.

Quantification of the overlap between two NP shape distributions

To quantify the overlap between two NP shapes, we first perform a principal component analysis (PCA) of the shape descriptors (Fourier coefficients) as explained in detail in our previous paper[7]. Once the PCA is calculated, the pair of NP groups for which the overlap is being studied is represented in the 3D space corresponding to the first three PCs. The line that joins the centre-of-gravity of the two groups will be considered the axis in which to calculate the shapes distribution and which we label as X. The “clouds” corresponding to the PCs of the NPs of the two groups are then projected onto X. Finally, the probability distribution function of the projected points on X are calculated and the overlap of the shape distributions can be represented.

Calculation of the principal components for the shape learning trajectory

The PCA for the shape space trajectories was performed as explained in our previous paper[2] but only the data from the initial and final shapes are used in the calculation of the eigenvectors during the PCA. In this way we obtained a projection matrix, which projects the Fourier coefficients into the PC space defined by the initial and final shapes. Then, the calculated projection matrix is used to project all the intermediate shapes into this PC space.

Transcriptome

The synthesis of *in vitro*_GNPa-c was reported previously[2]. In brief, *in vitro*_GNPa and *in vitro*_GNPb were synthesized using a microfluidic method, and *in vitro*_GNPc was synthesized by a benchtop method. After synthesis, the three types of GNPs were dispersed in 1 mg/ml OVA to coat the particle surface with OVA.

Transcriptome experiment details were reported in our previous work[2]. JAWS II cells were treated with 3×10^{10} *in vitro*_GNPs for 24 h. After the treatment, the cells were washed with PBS, and subsequently the total RNA was extracted by InviTrap® Spin Universal RNA Mini Kit. Each treatment was performed in replicates and repeated independently three times. The transcriptome was analyzed by using Whole-Mouse Genome One-Color Microarray (Agilent). After correcting the background and filtering low intensity signals, data were expressed in the mean of replicates in a log2 scale. PCA was performed by R studio (v.4.0.3) and the additional package ggplot2.

Synthesis of *in vivo*_GNPs

*In vivo*_GNPs were prepared in water *via* the seed-mediated method reported previously[1]. For *in vivo*_GNP(A), gold seeds were prepared by adding 4.5 ml of trisodium citrate (0.15 mmol, 34 mM) to 150 ml of a boiling gold precursor solution ($HAuCl_4 \cdot 3H_2O$, 0.038 mmol, 0.25 mM). The mixture was stirred under reflux for 30 min and overnight at room temperature. 12 ml of the seeds and 1.53 ml of trisodium citrate (0.052 mmol, 34 mM) were added to 300 ml of aqueous solution of $HAuCl_4 \cdot 3H_2O$ (0.089 mmol, 0.30 mM) under reflux and then stirred for 30 min. After a further addition of trisodium citrate (11.8 ml, 0.400 mmol, 34 mM) the mixture was stirred under reflux for a further 1 h, then cooled to room temperature and filtered through 0.2 µm filters. 200 mg of BSPP (0.38 mmol) were dissolved in 5 ml of water and added to 150 ml of the prepared NP suspension and stirred overnight. The sample was then washed several times using centrifugal filters (10000 MWCO, 600 rcf) using water with low BSPP concentration (4 µM) and concentrated.

For *in vivo*_GNP(B) and *in vivo*_GNP(C), 0.5 ml of the TR_Seeds described above was added to 100 ml 4 °C ultrapure water containing 0.5 ml of glycerol, 0.1 ml of $HAuCl_4 \cdot 3H_2O$ (0.022 mmol), 0.5 ml $AgNO_3$ solution (0.0005 mmol, 1 mM) and ascorbic acid (0.5 ml, 0.05 mmol, 100 mM) were subsequently added under vigorous stirring. After 30 s 1.0 ml of BSPP solution (0.311 mmol, 311 mM) was added and the mixture was gently stirred at room temperature overnight. The dispersion was finally filtered through 0.2 µm filters and washed.

In order to avoid lipopolysaccharide (LPS) contamination, all the synthetic processes for *in vitro* and *in vivo*_GNPs were performed using the standard operational procedures associated with cell culture. All the processes were carried out in a laminar flow cabinet (Class II, type B2) located in a laboratory with filtered air and limited access. All the solvents and the reagents were strictly opened inside the laminar flow cabinet after purchase and tested for LPS contaminant. All the plasticware used (*e.g.*, falcon tubes, Eppendorf tubes, pipette tips, syringes, *etc.*) were endotoxin-free certified. All the glassware used in the reaction were previously soaked in fresh aqua regia for at least 30 mins, and then carefully rinsed with endotoxin-free water. The LPS level of GNPs were detected by Limulus Amebocyte Lysate (LAL) assay before using.

Lymphatic B cell isolation

Armpit lymph nodes were extracted from treated animals and transferred to Dulbecco's Modified Eagle Medium (DMEM) Glutamax (GIBCO, 10566-016) supplemented with 10% heat inactivated fetal bovine serum (HI-FBS) on ice and processed immediately after extraction. The same organs from the same treatment group were pooled and mechanically homogenized by disrupting them through a 70 µm cell strainer (Thermo Fisher, 22363548) using a plunger of a sterile syringe until no more tissue was left. Cells were collected in 50 ml Falcon tubes

with 50 ml DMEM 10% HI-FBS and split into 2 tubes. Cells were spun down for 10 min at 1300 rpm for further erythrocyte lysis. Erythrocyte lysis was performed using the erythrocyte lysing kit (WL2000) following the manufacturer's instructions for each of the resulting tubes from each organ and treatment group. Non-erythrocytic cells were resuspended in 5 ml DMEM with 10% HI-FBS carefully and transferred to a new tube. Cells were counted in order to further perform the B cell isolation.

B cell isolation was performed by means of the MagCellect B cell isolation kit (MAGR303). Two aliquots of 2 $\times 10^8$ cells from the spleen homogenates and whole lymph node homogenates were used. Cells were spun down at 1300 rpm for 5 minutes and resuspended in 1 ml of ice-cold 1X MagCellect Buffer. 200 µl of MagCellect Rat B Cell Biotinylated Antibody cocktail was added for negative selection of B cells and incubated at 4°C for 15 minutes. 250 µl MagCellect Streptavidin Ferrofluid were added to the cell suspension and incubated at 4°C for 15 minutes. After that, 1.55 ml 1X MagCellect Buffer was added. Samples were applied to a magnetic stand for 6 minutes at room temperature and the supernatant containing B cells was extracted. This step was repeated to ensure removal of all tagged cells.

Total RNA extraction and cDNA library preparation

B cell total RNA isolation was performed using the Invitrap® Spin cell RNA mini kit (10061100300) following the manufacturer's instructions. In the case of the B cells isolated from the spleens, purification was done from 2 aliquots of 1×10^7 B cells, while for the lymph nodes, only 1 aliquot containing all the B cells was performed. RNA quality and concentration were checked immediately after the purification by Nanodrop. 1 µg and 150 ng of purified RNA from spleen and lymph nodes, respectively, of each group were used for retrotranscription RNA, at concentrations of 50 and 7.5 ng/µl respectively. High-capacity cDNA reverse transcription kit (4368814) was used for RT-PCR, and the 2X master mix was prepared as follows: 4.0 µl 10X RT buffer, 1.6 µl 25X dNTP Mix, 4.0 µl 10X RT Random primers, 2.0 µl Reverse Transcriptase, 8.4 µl Nuclease-free water. PCR was performed as follows: 25°C for 10 minutes; 37°C for 60 minutes; 37 °C for 60 minutes; 85°C for 5 minutes; 4°C. 4 µl of cDNA at 150 and 20 ng/ µl for spleen and lymph nodes were used from each sample to amplify the corresponding Ig specific variable regions. Amplification of Ig specific rat variable amplicons were amplified following a previously reported primer mix targeting the framework regions 1 and 4. A total of 50 PCR reactions were performed for each of the samples (*i.e.*, different experimental group and organ). Twenty-four reactions for heavy chain amplification (each reaction contained 1 heavy chain specific forward primer + combination of 4 heavy chain backward primers) and 26 reactions for light chain amplification (each reaction contained 1 light chain specific forward primer + combination of 5 light chain specific backward primers).

Q5® High-Fidelity DNA Polymerase kit (M0491L) was used following the suggested PCR master mix by the manufacturer.

PCR was performed as follows: 90°C for 5 minutes; 35 cycles of: 95°C for 30 seconds, 60°C for 30 seconds, 72°C for 30 seconds; 72°C for 10 minutes. Final PCR products were mixed and cleaned by a AMPure XP PCR purification Kit (A63880) using a ratio of beads of 0.8X of the total volume of the mixture and elution was done in 40 µl of 10 mM tris-acetate pH 8.0. Illumina Nextera indexes were included for Next Generation Sequencing using the Nextera XT Illumina Index Kit (FC-131-1096). For each group, 3 aliquots of 12.5 µl (containing 100 ng/µl of cleaned variable region amplicons) a short cycle PCR for attachment of indexes was performed following manufacturer's instructions. PCR was performed as follows: 72°C for 3 minutes; 92°C for 30 seconds; 12 cycles of: 95°C for 10 minutes, 55°C for 30 seconds, 72°C for 30 seconds. The clean-up procedure was performed as before, and amplicons were eluted in 40 µl of 10 mM tris-acetate pH 8.0. Quality and concentration were checked by Bioanalyzer prior to sequencing and 30 µl at 5 nM were used for sequencing.

V(D)J Repertoire analysis

IMGT/High-VQUEST alignment files for the different groups were analyzed by means of the tools provided in the Immcantation (V2.5.0) portal (https://immcantation.readthedocs.io/en/version-2.5.0/), consisting of the packages Change-O, Alakazam, SHazaM and TIgGER for BCR repertoire analysis.

Supporting Figures

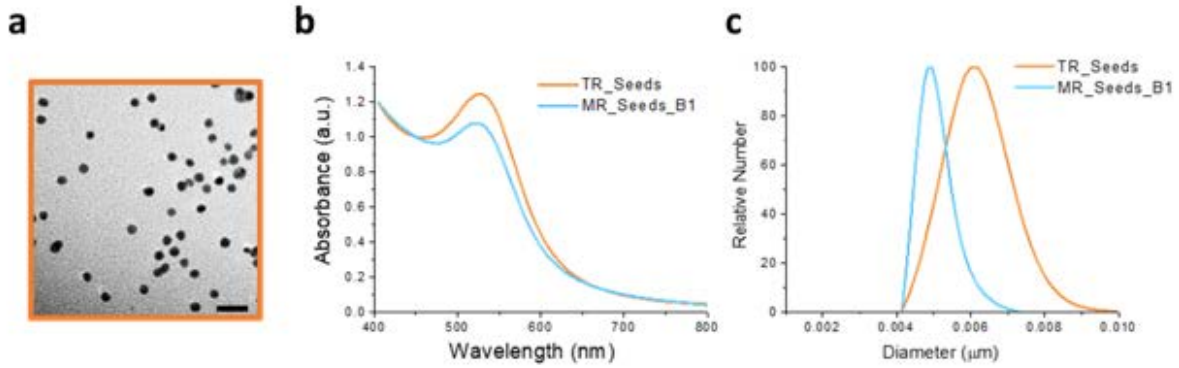


Figure S2. 1 Characterization of 5 nm seeds synthesized by benchtop tank reactor (TR_Seeds).

(a) TEM micrographs for TR_Seeds, scale bar is 20 nm. (b) Normalized UV-Vis-NIR absorption for TR_Seeds and MR_Seeds. (c) DCS analysis showing a narrower size distribution of MR_Seeds compared with TR_Seeds.

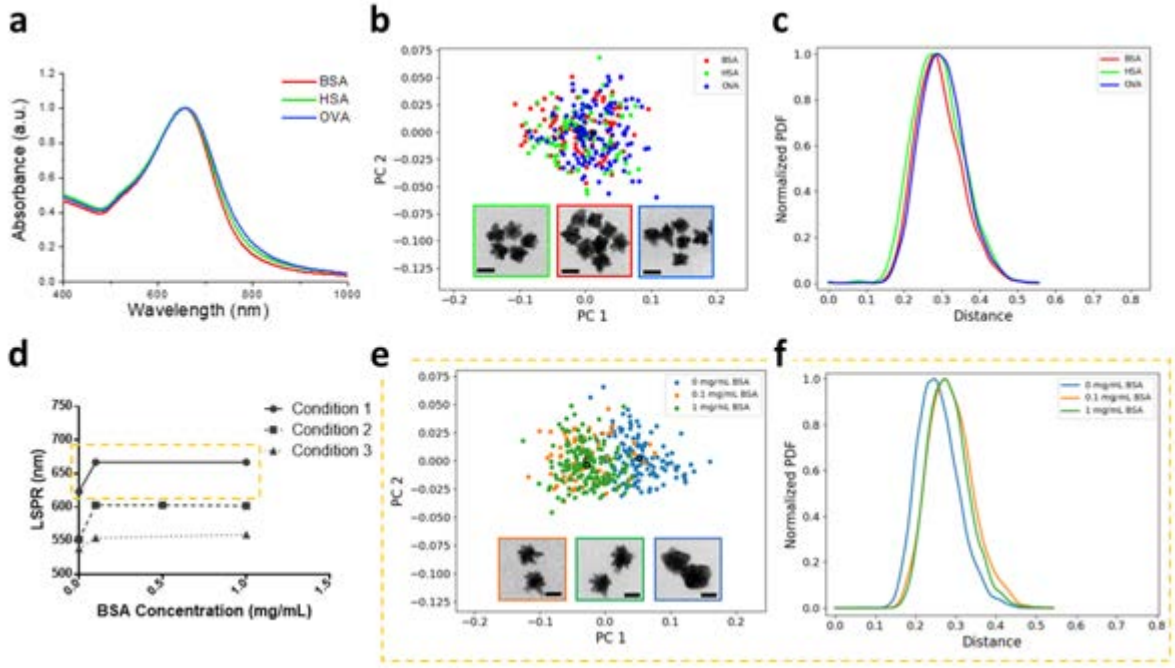


Figure S2. 2 Characterization of MR_GNPs with different protein coatings.

(a) Normalized UV-Vis-NIR absorption showing similar localization surface plasmon resonance (LSPR) of different protein associated shape synthesis, bovine serum albumin (BSA), human serum albumin (HSA) and ovalbumin (OVA) concentrations are 1 mg/ml. (b-c) 2D scatter plot, representative TEM micrographs and shape variance showing similar shape and shape distribution for different protein coating GNPs. (d) LSPR corresponding to BSA concentration plot showing the protein amount effect on shape, condition 1 is 400 µl seeds, 1 × 10^{-4} mol hydroquinone and 3 × 10^{-7} mol $AgNO_3$; condition 2 is 400 µl seeds and 4 × 10^{-4} mol hydroquinone; condition 3 is 200 µl seeds and 3 × 10^{-5} mol hydroquinone. $HAuCl_4 \cdot 3H_2O$ (5 × 10^{-5} mol) and sodium dihydrate citrate (1.5 × 10^{-5} mol) are kept the same for condition 1-3. (e-f) 2D scatter plot, representative TEM micrographs and shape variance for condition 1. Scale bar is 50 nm.

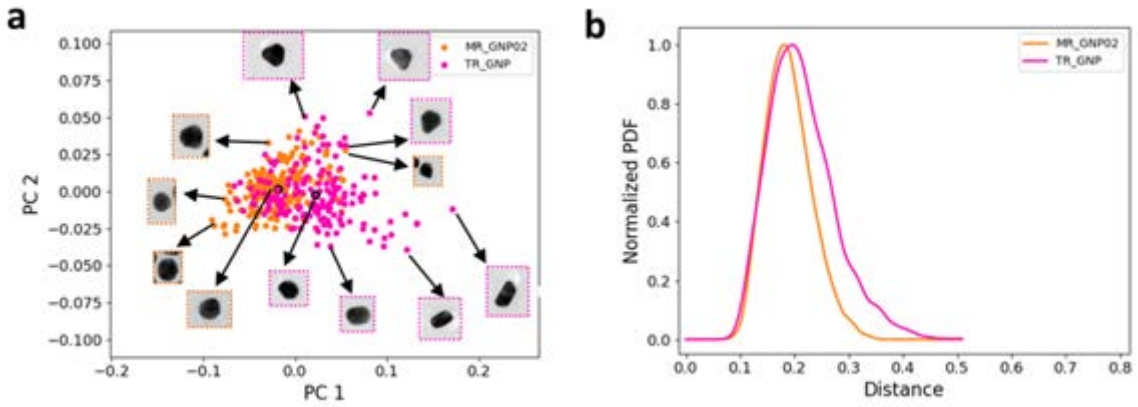


Figure S2. 3 Shape distribution comparison between flow reactor- (MR) and benchtop tank reactor (TR)-based synthesis.

(a-b) 2D scatter plot and shape variance showing more homogenous and narrower shape distribution of MR_GNP02.

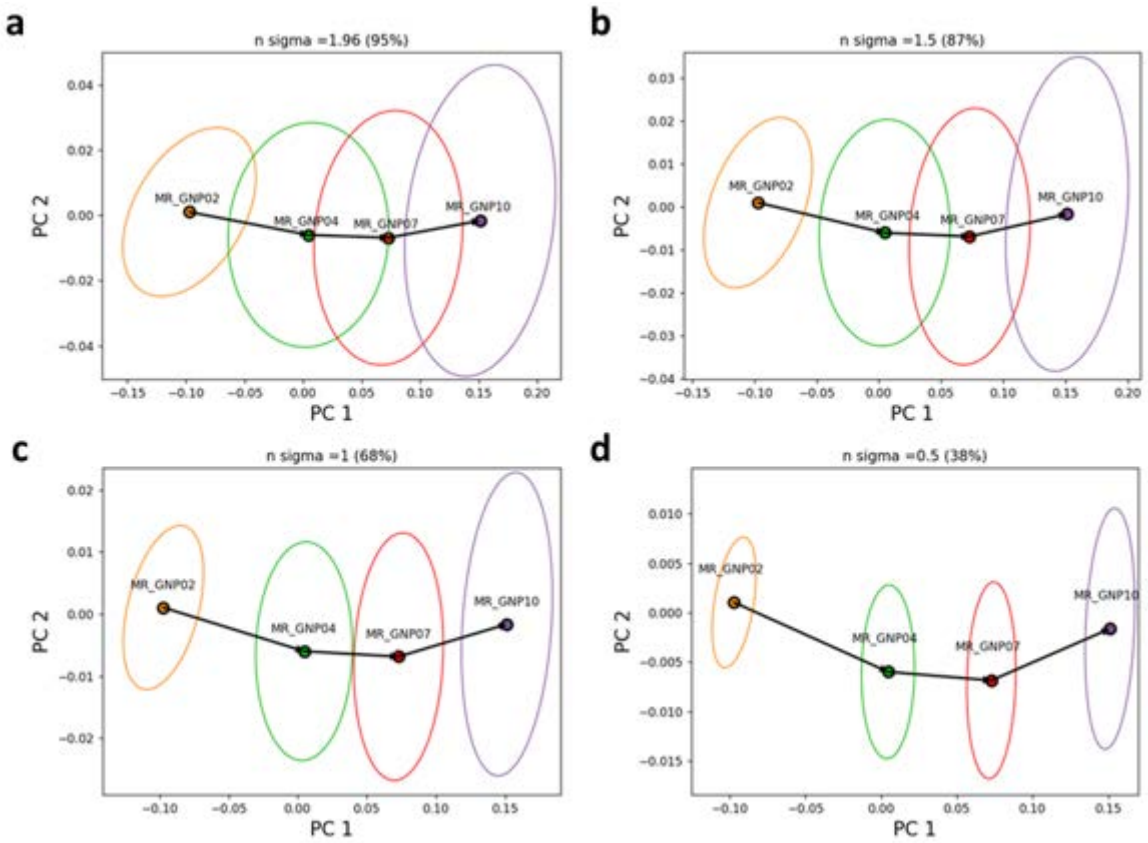


Figure S2. 4 Shape distributions of MR_GNPs illustrated by confidential ellipses.

In each panel an ellipse represents the likelihood (interval of confidence) that a certain percentage of the points are contained in that ellipse. The axis of each ellipse is obtained as follows. We assume that the distribution of points for all groups are normally distributed. Based on this assumption we then calculate the variance and orientation of the each of the gaussians for each group on the 2D PC space. Finally, an ellipse with certain number of standard deviations (sigma) of the distributions previously calculated is drawn. For example: n sigma = 1.96 (95%) means the ellipse contains 95% of the data outlines with an internal of confidence.

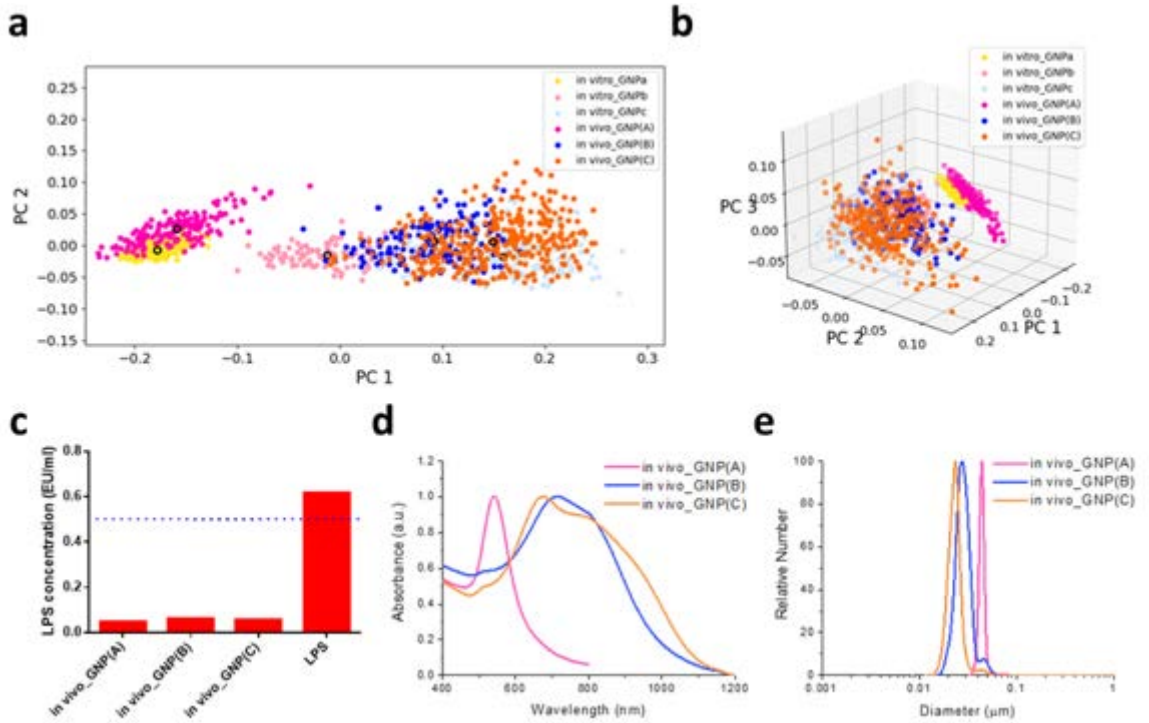


Figure S2. 5 Characterization of GNPs used for *in vivo* study.

(a-b) 2D and 3D scatter plot for *in vitro*_GNPs and *in vivo*_GNPs, (c) LPS level of *in vivo*_GNPs, (d) Normalized UV-Vis-NIR absorption of *in vivo*_GNPs, (e) DCS size distribution analyzed by relative number representing the size distribution of *in vivo*_GNPs.

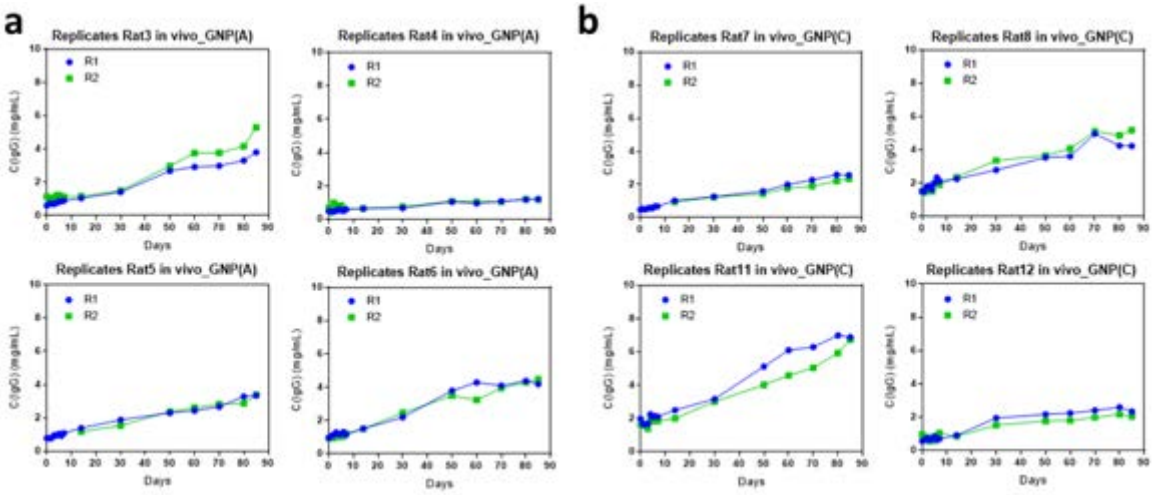


Figure S2. 6 Replicates of IgG expression for *in vivo*_GNP(A) and *in vivo*_GNP(C).

Data are the mean of duplicates. R1 and R2 present two different operators.

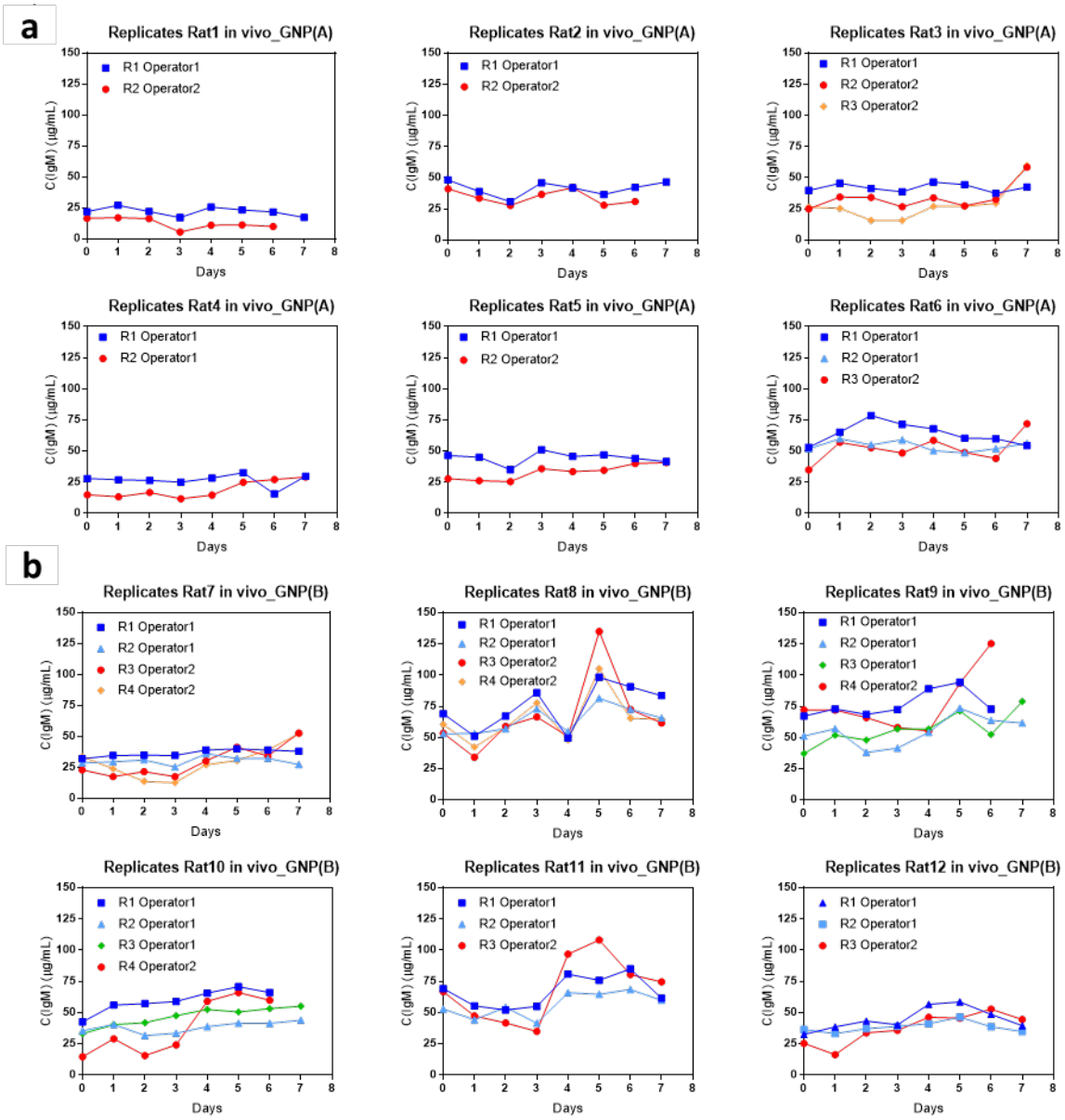


Figure S2. 7 Replicates of IgM expression performed by different operators for *in vivo*_GNP(A) and *in vivo*_GNP(B).

Data are the mean of duplicates. R1 and R2 present two different operators.

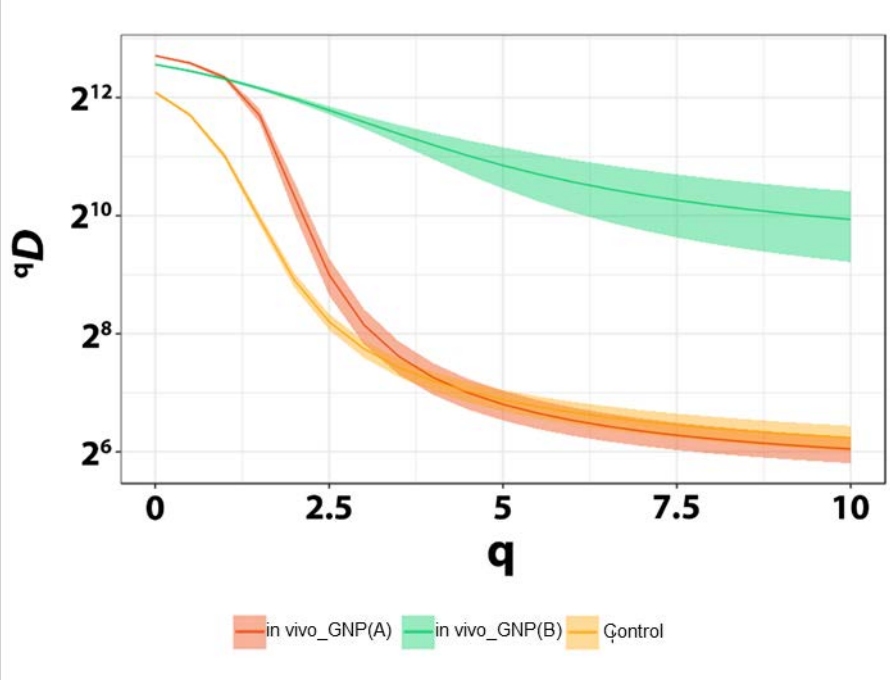


Figure S2. 8 Comparison of clonal diversity between repertoires by Hill diversity curves using Change-O.

Diversity indexes (^qD) were estimated by uniform sampling of the lymphatic B cell repertoires across different diversity orders (q).

Reservoirs / GNPs	R1			R2		R3	R4
	$HAuCl_4 \cdot 3H_2O$	$AgNO_3$	Glycerol	MR_ Seeds	Na_3Cit	HQ	HQ
MR_GNP01	5×10^{-5} mol	0	0	400 μl	1.5×10^{-5} mol	3×10^{-5} mol	3×10^{-5} mol
MR_GNP02	5×10^{-5} mol	0	0	300 μl	1.5×10^{-5} mol	3×10^{-5} mol	3×10^{-5} mol
MR_GNP03	5×10^{-5} mol	0	0	200 μl	1.5×10^{-5} mol	3×10^{-5} mol	3×10^{-5} mol
MR_GNP04	5×10^{-5} mol	0	0	300 μl	1.5×10^{-5} mol	4×10^{-4} mol	4×10^{-4} mol
MR_GNP06	5×10^{-5} mol	0	0	300 μl	1.5×10^{-5} mol	1×10^{-4} mol	1×10^{-4} mol
MR_GNP07	5×10^{-5} mol	5×10^{-8} mol	0	300 μl	1.5×10^{-5} mol	1×10^{-4} mol	1×10^{-4} mol
MR_GNP08	5×10^{-5} mol	1×10^{-7} mol	0	300 μl	1.5×10^{-5} mol	1×10^{-4} mol	1×10^{-4} mol
MR_GNP09	5×10^{-5} mol	3×10^{-7} mol	0	300 μl	1.5×10^{-5} mol	1×10^{-4} mol	1×10^{-4} mol
MR_GNP10	5×10^{-5} mol	1×10^{-7} mol	500 μL	300 μl	1.5×10^{-5} mol	1×10^{-4} mol	1×10^{-4} mol

Table 2. 1 Synthesis recipe for MR_GNP01-10. Reaction volume for each reservoir is 50 ml.

	Hydrodynamic diameter (nm)	Polydispersity index (PDI)	Zeta potential in H_2O (mV)
MR_GNP02	51 ±1	0.15 ± 0.01	-28 ± 1
MR_GNP04	50 ±1	0.12 ± 0.01	-24 ± 1
MR_GNP07	49 ±1	0.20 ± 0.01	-24 ± 1
MR_GNP10	57 ±1	0.08 ± 0.01	-26 ± 2

Table 2. 2 Hydrodynamic diameter and zeta potential of MR_GNPs.

Action	Tool/Program	Parameters
Read processing	Trimmomatic V0.32	ILLUMINACLIP:2:30:10 LEADING:3 TRAILING:3 MINLEN:36 NexteraPE-PE.fa (adapters)
Paired-End read alignment	PEAR V0.9.8	-p 0.05 -v 15 –q 30 –n 300 –m 500
V(D)J germline alignment	IMGT/High-VQUEST	Species:Rattus norvegicus locus:IG Single individual:NO
IMGT output formatting for Immcantation tools: Parsing	Change-O "MakeDb"	--partial true --scores true --regions true --junction true
IMGT output formatting for Immcantation tools: Removing non-functional	Change-O "ParseDb"	select -f FUNCTIONAL -u T
IMGT output formatting for Immcantation tools: Separating heavy and light chains	Change-O "ParseDb"	select
IMGT output formatting for Immcantation tools: Adding group name and organ	Change-O "ParseDb"	add columns with corresponding names
Polymorphism identification	TIgGER "findNovelAlleles"	germline_min = 100 min_seqs = 50 auto_mutrange = TRUE mut_range = 1:10 y_intercept = 0.125 alpha = 0.05
Inferring Genotype	TIgGER "inferGenotype"	default
Correcting allele calls	TIgGER "reassignAlleles"	method = hamming keep_gene = gene
Calculating distance to nearest neighbor	SHazaM "distToNearest"	model = ham normalize = none symmetry = avg first = FALSE
Finding distance threshold for clonal assignment	SHazaM "findThreshold"	method = density

Clonal clustering	Change-O "defineClones"	--mode allele --act set --model ham --dist (calculated before) --sym avg
Reconstruct germline alignments	Change-O "CreateGermlines"	--vf V_CALL_GENOTYPED
V(D)J family and gene usage	Alakazam "countGenes"	groups = GROUP/ORGAN mode = family/gene clone = CLONE
Clonal diversity	Alakazam "rarefyDiversity"	clone = CLONE max_n = 9000 min_q = 0 max_q = 10 step_q = 0.5 ci = 0.95 nboot = 2000 uniform = TRUE

Table 2. 3 Detailed step-by-step bioinformatic analysis with the corresponding parameters for each of the procedures.

References

1. Monopoli, M. P.; Aberg, C.; Salvati, A.; Dawson, K. A., Biomolecular coronas provide the biological identity of nanosized materials. *Nat. Nanotechnol.* **2012,** *7*, 779-786.
2. Dawson, K. A.; Yan, Y., Current understanding of biological identity at the nanoscale and future prospects. *Nat. Nanotechnol.* **2021,** *16*, 229-242.
3. Kelly, P. M.; Aberg, C.; Polo, E.; O'Connell, A.; Cookman, J.; Fallon, J.; Krpetic, Z.; Dawson, K. A., Mapping protein binding sites on the biomolecular corona of nanoparticles. *Nat. Nanotechnol.* **2015,** *10*, 472-479.
4. Lara, S.; Perez-Potti, A.; Herda, L. M.; Adumeau, L.; Dawson, K. A.; Yan, Y., Differential recognition of nanoparticle protein corona and modified low-density lipoprotein by macrophage receptor with collagenous structure. *ACS Nano* **2018,** *12*, 4930-4937.
5. Lara, S.; Alnasser, F.; Polo, E.; Garry, D.; Lo Giudice, M. C.; Hristov, D. R.; Rocks, L.; Salvati, A.; Yan, Y.; Dawson, K. A., Identification of receptor binding to the biomolecular corona of nanoparticles. *ACS Nano* **2017,** *11*, 1884-1893.
6. Theiss, S.; Voggel, M.; Kuper, H.; Hoermann, M.; Krings, U.; Baum, P.; Becker, J. A.; Wittmann, V.; Polarz, S., Ligand-programmed consecutive symmetry break(s) in nanoparticle based materials showing emergent phenomena: Transitioning from sixfold to threefold symmetry in anisotropic ZnO colloids. *Adv. Funct. Mater.* **2021,** *31*, 2009104.
7. Boselli, L.; Lopez, H.; Zhang, W.; Cai, Q.; Giannone, V. A.; Li, J.; Moura, A.; de Araujo, J. M.; Cookman, J.; Castagnola, V.; Yan, Y.; Dawson, K. A., Classification and biological identity of complex nano shapes. *Commun. Mater.* **2020,** *1, 1-12*.
8. Valov, I.; Yang, Y., The route towards nanoparticle shape metrology. *Nat. Nanotechnol.* **2020,** *15*, 510-511.
9. Wang, J.; Chen, H. J.; Hang, T.; Yu, Y.; Liu, G.; He, G.; Xiao, S.; Yang, B. R.; Yang, C.; Liu, F.; Tao, J.; Wu, M. X.; Xie, X., Physical activation of innate immunity by spiky particles. *Nat. Nanotechnol.* **2018,** *13*, 1078-1086.
10. Kumar, S.; Anselmo, A. C.; Banerjee, A.; Zakrewsky, M.; Mitragotri, S., Shape and size-dependent immune response to antigen-carrying nanoparticles. *J. Control. Release* **2015,** *220*, 141-148.
11. Faria, M.; Bjornmalm, M.; Thurecht, K. J.; Kent, S. J.; Parton, R. G.; Kavallaris, M.; Johnston, A. P. R.; Gooding, J. J.; Corrie, S. R.; Boyd, B. J.; Thordarson, P.; Whittaker, A. K.; Stevens, M. M.; Prestidge, C. A.; Porter, C. J. H.; Parak, W. J.; Davis, T. P.; Crampin, E. J.; Caruso, F., Minimum information reporting in bio-nano experimental literature. *Nat. Nanotechnol.* **2018,** *13*, 777-785.
12. Rivera-Gil, P.; Jimenez de Aberasturi, D.; Wulf, V.; Pelaz, B.; del Pino, P.; Zhao, Y.; de la Fuente, J. M.; Ruiz de Larramendi, I.; Rojo, T.; Liang, X. J.; Parak, W. J., The challenge to relate the physicochemical properties of colloidal nanoparticles to their cytotoxicity. *Acc. Chem. Res.* **2013,** *46*, 743-749.
13. Shi, J.; Kantoff, P. W.; Wooster, R.; Farokhzad, O. C., Cancer nanomedicine: progress, challenges and opportunities. *Nat. Rev. Cancer* **2017,** *17*, 20-37.
14. Gong, N.; Zhang, Y.; Teng, X.; Wang, Y.; Huo, S.; Qing, G.; Ni, Q.; Li, X.; Wang, J.; Ye, X.; Zhang, T.; Chen, S.; Wang, Y.; Yu, J.; Wang, P. C.; Gan, Y.; Zhang, J.; Mitchell, M. J.; Li, J.; Liang, X. J., Proton-driven transformable nanovaccine for cancer immunotherapy. *Nat. Nanotechnol.* **2020,** *15*, 1053-1064.
15. Kuai, R.; Ochyl, L. J.; Bahjat, K. S.; Schwendeman, A.; Moon, J. J., Designer vaccine nanodiscs for personalized cancer immunotherapy. *Nat. Mater.* **2017,** *16*, 489-496.
16. Singh, A., Eliciting B cell immunity against infectious diseases using nanovaccines. *Nat. Nanotechnol.* **2021,** *16*, 16-24.
17. Jiang, W.; Wang, Y.; Wargo, J. A.; Lang, F. F.; Kim, B. Y. S., Considerations for designing preclinical cancer immune nanomedicine studies. *Nat. Nanotechnol.* **2021,** *16*, 6-15.

18. Talamini, L., *et al*, Influence of size and shape on the anatomical distribution of endotoxin-free gold nanoparticles. *ACS Nano* **2017,** *11*, 5519-5529.
19. Kenichi Niikura, T. M., Tadaki Suzuki, Shintaro Kobayashi, Hiroki Yamaguchi,; Yasuko Orba, A. K., Hideki Hasegawa, Kiichi Kajino,z Takafumi Ninomiya, Kuniharu Ijiro,; and Hirofumi Sawa, Gold nanoparticles as a vaccine platform: influence of size and shape on immunological responses *in vitro* and *in vivo*. *ACS Nano* **2013,** *7*, 3926-3938.
20. Yang, L.; Zhou, Z.; Song, J.; Chen, X., Anisotropic nanomaterials for shape-dependent physicochemical and biomedical applications. *Chem. Soc. Rev.* **2019,** *48*, 5140-5176.
21. Albanese, A.; Tang, P. S.; Chan, W. C., The effect of nanoparticle size, shape, and surface chemistry on biological systems. *Annu. Rev. Biomed. Eng.* **2012,** *14*, 1-16.
22. Keys, A. S.; Iacovella, C. R.; Glotzer, S. C., Characterizing complex particle morphologies through shape matching: Descriptors, applications, and algorithms. *J. Comput. Phys.* **2011,** *230*, 6438-6463.
23. Rice, K. P.; Saunders, A. E.; Stoykovich, M. P., Classifying the shape of colloidal nanocrystals by complex fourier descriptor analysis. *Cryst. Growth Des.* **2012,** *12*, 825-831.
24. Bals, S.; Goris, B.; Liz-Marzan, L. M.; Van Tendeloo, G., Three-dimensional characterization of noble-metal nanoparticles and their assemblies by electron tomography. *Angew. Chem. Int. Ed.* **2014,** *53*, 10600-10610.
25. Bigdeli, A.; Hormozi-Nezhad, M. R.; Jalali-Heravi, M.; Abedini, M. R.; Sharif-Bakhtiar, F., Towards defining new nano-descriptors: extracting morphological features from transmission electron microscopy images. *RSC Adv.* **2014,** *4*, 60135-60143.
26. Wang, X.; Li, J.; Ha, H. D.; Dahl, J. C.; Ondry, J. C.; Moreno-Hernandez, I.; Head-Gordon, T.; Alivisatos, A. P., AutoDetect-mNP: An unsupervised machine learning algorithm for automated analysis of transmission electron microscope images of metal nanoparticles. *JACS Au* **2021,** *1*, 316-327.
27. Cai Q.; Castagnola V.; Boselli L.; Moura A.; Lopez H.; Zhang W.; Araujo J.; Dawson K., A microfluidic approach for synthesis and kinetic profiling of branched gold nanostructure. *ChemRxiv. Cambridge: Cambridge Open Engage* **2021**; This content is a preprint and has not been peer-reviewed.
28. Vranic, D.; Saupe, D., 3D shape descriptor based on 3D Fourier transform. *the 3rd EURASIP Conference on Digital Signal Processing for Multimedia Communications and Services* **2001**, 271-274.
29. Zhang*, D.; Lu, G., Shape-based image retrieval using generic Fourier descriptor. *Signal Process. Image Commun.* **2002,** *17*, 825-848.
30. Zhang, H., Fiume, E., Shape matching of 3D contours using normalized Fourier descriptors. Proceedings SMI. Shape Modeling International 2002, Banff, AB, Canada, pp. 261-268 (2002). doi: 10.1109/SMI.2002.1003554
31. Elisabeth T. Bowman, K. S., Tom W. Drummond, Particle shape characterisation using fourier analysis. *Geotechnique* **2000,** *51*, 545-554.
32. Grzelczak, M.; Perez-Juste, J.; Mulvaney, P.; Liz-Marzan, L. M., Shape control in gold nanoparticle synthesis. *Chem. Soc. Rev.* **2008,** *37*, 1783-1791.
33. Day, R. W.; Mankin, M. N.; Gao, R.; No, Y. S.; Kim, S. K.; Bell, D. C.; Park, H. G.; Lieber, C. M., Plateau-Rayleigh crystal growth of periodic shells on one-dimensional substrates. *Nat. Nanotechnol.* **2015,** *10*, 345-352.
34. Finney, E. E.; Finke, R. G., Nanocluster nucleation and growth kinetic and mechanistic studies: a review emphasizing transition-metal nanoclusters. *J. Colloid Interface Sci.* **2008,** *317*, 351-374.
35. Frens, G., Controlled nucleation for the regulation of the particle size in monodisperse gold suspensions. *Nature* **1973,** *241*, 20-22.
36. JOHN TURKEVICH, P. C. S., JAMES HILLIE, A study of the nucleation and growth processes in the synthesis of colloidal gold. *Discuss. of the Faraday Soc.* **1951,** *11*, 55-75.
37. Niu, W.; Chua, Y. A.; Zhang, W.; Huang, H.; Lu, X., Highly symmetric gold nanostars: crystallographic control and surface-enhanced raman scattering property. *J. Am. Chem. Soc.* **2015,** *137*, 10460-10463.

38. Thanh, N. T.; Maclean, N.; Mahiddine, S., Mechanisms of nucleation and growth of nanoparticles in solution. *Chem. Rev.* **2014,** *114*, 7610-7630.
39. Sun, Y.; Xia, Y., Shape-controlled synthesis of gold and silver nanoparticles. *Science* **2002,** *298*, 2176-2179.
40. El-Sayed*, B. N. a. M. A., Preparation and growth mechanism of gold nanorods (NRs) using seed-mediated growth method. *Chem. mater.* **2003,** *15*, 1957-1962.
41. Jana N R; Gearheart L; S, M. C. J., Seed-mediated growth approach for shape-controlled synthesis of spheroidal and rod-like gold nanoparticles using a surfactant template. *Adv. Mater.* **2001,** *13*, 1389-1393.
42. Zhao L; Ji X; Sun X, e. a., Formation and stability of gold nanoflowers by the seeding approach: the effect of intraparticle ripening. *J. Phys. Chem. C* **2009,** *113*, 16645-16651.
43. Li, J.; Wu, J.; Zhang, X.; Liu, Y.; Zhou, D.; Sun, H.; Zhang, H.; Yang, B., Controllable synthesis of stable urchin-like gold nanoparticles using hydroquinone to tune the reactivity of gold chloride. *J. Phys. Chem. C* **2011,** *115*, 3630-3637.
44. Osman M. Bakr; Benjamin H. Wunsch; Stellacci*, F., High-yield synthesis of multi-branched urchin-like gold nanoparticles. *Chem. mater.* **2006,** *18*, 3297-3301.
45. Xia, Y.; Xiong, Y.; Lim, B.; Skrabalak, S. E., Shape-controlled synthesis of metal nanocrystals: simple chemistry meets complex physics? *Angew. Chem. Int. Ed.* **2009,** *48*, 60-103.
46. Sun, Y., Controlled synthesis of colloidal silver nanoparticles in organic solutions: empirical rules for nucleation engineering. *Chem. Soc. Rev.* **2013,** *42*, 2497-2511.
47. Rahman, M.; Rebrov, E., Microreactors for gold nanoparticles synthesis: From Faraday to flow. *Processes* **2014,** *2*, 466-493.
48. Richmond, C. J., *et al*, A flow-system array for the discovery and scale up of inorganic clusters. *Nat. Chem.* **2012,** *4*, 1037-1043.
49. Zhao, C.-X.; He, L.; Qiao, S. Z.; Middelberg, A. P. J., Nanoparticle synthesis in microreactors. *Chem. Eng. Sci.* **2011,** *66*, 1463-1479.
50. Hou, X., Zhang, Y. S., Santiago, T., Alvarez, M. M., Ribas, J., Jonas, S., Weiss, P., Andrews, A.M., Aizenberg, J., Khademhosseini, A., Interplay between materials and microfluidics. *Nat. Rev. Mater.* **2017,** *2*, 1-15.
51. Rodrigues, T.; Schneider, P.; Schneider, G., Accessing new chemical entities through microfluidic systems. *Angew. Chem. Int. Ed.* **2014,** *53*, 5750-8.
52. Samuel E. Lohse, J. R. E., Sean T. Sivapalan, Michael R. Plews, and Catherine J. Murphy*, A Simple millifluidic benchtop reactor system for the high-throughput synthesis and functionalization of gold nanoparticles with different sizes and shapes. *ACS Nano* **2013,** *7*, 4135-4150.
53. Duraiswamy, S.; Khan, S. A., Droplet-based microfluidic synthesis of anisotropic metal nanocrystals. *Small* **2009,** *5*, 2828-2834.
54. Song, Y.; Hormes, J.; Kumar, C. S., Microfluidic synthesis of nanomaterials. *Small* **2008,** *4*, 698-711.
55. Lignos, I.; Maceiczyk, R.; deMello, A. J., Microfluidic technology: uncovering the mechanisms of nanocrystal nucleation and growth. *Acc. Chem. Res.* **2017,** *50*, 1248-1257.
56. Kallenberg, C. G., Anti-C1q autoantibodies. *Autoimmun Rev.* **2008,** *7*, 612-615.
57. Bashford-Rogers, R. J. M.; Smith, K. G. C.; Thomas, D. C., Antibody repertoire analysis in polygenic autoimmune diseases. *Immunology* **2018,** *155*, 3-17.
58. Nemazee, D., Mechanisms of central tolerance for B cells. *Nat. Rev. Immunol.* **2017,** *17*, 281-294.
59. Vojdani, A., A potential link between environmental triggers and autoimmunity. *Autoimmune Dis.* **2014,** *2014*, 437231.
60. Zhao, C. N.; Xu, Z.; Wu, G. C.; Mao, Y. M.; Liu, L. N.; Qian, W.; Dan, Y. L.; Tao, S. S.; Zhang, Q.; Sam, N. B.; Fan, Y. G.; Zou, Y. F.; Ye, D. Q.; Pan, H. F., Emerging role of air pollution in autoimmune diseases. *Autoimmun Rev.* **2019,** *18*, 607-614.

61. Khan, M. F.; Wang, G., Environmental agents, oxidative stress and autoimmunity. *Curr. Opin. Toxicol.* **2018,** *7*, 22-27.

Chapter III

Shape-Dependent Histone Modifications Induced by Gold Nanoparticles

Wei Zhang[a,§], Jingji Li[a,§], Qi Cai[a], Camila P. Silveira[a], Valeria A. Giannone[a], Kenneth A. Dawson[a,b] & Yan Yan[a,c]**

a. Centre for BioNano Interactions, School of Chemistry, University College Dublin, Belfield, Dublin 4, Ireland

b. Guangdong Provincial Education Department Key Laboratory of Nano-Immunoregulation Tumor Microenvironment, The Second Affiliated Hospital, Guangzhou Medical University, Guangzhou, 510260 Guangdong, P.R. China

c. School of Biomolecular and Biomedical Science, UCD Conway Institute of Biomolecular and Biomedical Research, University College Dublin, Belfield, Dublin 4, Ireland

*Email: Kenneth.a.dawson@cbni.ucd.ie; yan.yan@cbni.ucd.ie

§. These authors contributed equally: Wei Zhang and Jingji Li

Abstract

In this study, we used two gold nanoparticles with distinct shapes to investigate the shape effects on the epigenome. The gold nanoparticles used in this study were endotoxin free with fully characterization to determine their physiochemical and shape features, to assure the reproducibility of nanoparticles used in the biological investigations. Microarray based transcriptomic study revealed that the differential expression genes between different shapes treated cells are related with the histone modifications. To provide a comprehensive analysis on the histone modification profiles induced by different shapes, proteomics was employed and confirmed the shape-dependent global histone modifications, especially on the methylation and acetylation of histone H3. Further antibody-based analysis (western blot and ELISA) validated the findings in proteomics. To investigate the biological process which could be governed by the shape induced epigenetic markers, chromatin immunoprecipitation sequencing (ChIP-seq) was done and found that H3K27me3 mark was specifically changed in the genes associated with neuron and synapse related signalling pathways.

Introduction

In contrast to smaller biomolecules or larger bulk materials, when it comes to the bionanoscale recognition and interaction[1-2], it is widely admitted that the nanoscale materials are processed *via* sophisticated endogenous cellular networks, which allows them to have a specific access to and control over key biological processes[3-4] that makes them show great potential in the application of nanomedicine and nano therapy[5]. For the nanoscale shape biological processing, we assume it is the shape features that makes the cell recognise them and then coat them with different biomolecular coronas[6-7] which work as a new identity of the shapes for the cellular interaction and following downstream outcomes. The detailed mechanism how the cells recognise different shapes yet to be discovered, however the investigation of the shape induced biological effects will advance our understanding of this question.

For the shape-dependent biological effects field, we are like still in the early stage, groping our way from phenomenological based trials to a systematic study. There are many reasons accounting for the difficulty of a systematic investigation of the shape-dependent impacts, for instance, when we talk about shape effects, we are talking about a shape ensemble instead of a pure single shape which is almost impossible to achieve in current nanotechnology process. Our previous work found that slight shape changes in the shape ensemble may induce obvious different biological outcomes (chapter II), which means for a robust shape-dependent biological effects investigation, we need to execute high reproducibility synthesis

and also well-defined characterization of the shapes to record and compare the nanoscale shapes used in different biological studies. Secondly, shape information needs to be able to store and transfer so that we can communicate the shapes between different laboratories, traditional physiochemical characterizations are important but not enough. Additionally, it is worth notion that a specific biological outcome can be induced by a combination of various factors, which also increases the difficulty to figure out the mechanism of shape-dependent biological effects.

Meanwhile increasing use and synthesis ability of unlimited nanoscale shape universe entails the need for reliable risk assessment of the potential toxicity of nanoscale shapes, not only on cytotoxicity, genotoxicity but also epigenetic changes. Since Conrad H. Waddington introduced the term "epigenotype" in 1942[8], it is widely accepted that epigenetic modifications[9] including DNA methylation, histone modifications and non-coding RNAs, can influence the gene expression without changing DNA sequence[10] and could be the cause for many diseases[11] like cancers, cardiovascular disease[12] and diabetes[13]. The potential impact of nanoparticles on epigenome has just started to receive researchers' attention in the past decade[14-16]. To date, most of the nanoparticles induced epigenetic modifications studies are mainly focused on the DNA methylation[17-19] and alterations in miRNA expression[20]. However, for the shape-dependent epigenetic modifications, it is still an intriguing field. Investigation into this will be beneficial both for the safety risk assessment of the nanoparticles and also for their potential nano therapy applications.

In this paper we report the shape-dependent histone modifications in both dTHP-1 and A549 cell lines. By strictly quality control of the synthesis of gold nanoparticles with two distinct shapes, facilitated with our previously published computational shape analysis method[21], we are able to achieve the reproducible clean synthesis of gold nanoparticles, recording, visualizing and comparing the shapes for the investigation of the shape effects. After preliminary cytotoxicity study, we figured out a suitable exposure concentration of these two types of gold nanoparticles and further explored the cellular uptake differences between the two shapes, which showed that GNP1 (spherical) has a higher uptake amount than GNP2 (star-like). We believe the uptake amount difference is caused by the different cellular recognition mechanisms, which needs more time to clarify the detailed recognition mechanism. After this, microarray-based transcriptome was performed to investigate the shape effects on the whole cell profile and found that most of the differential expression genes are clustered to the histone modifications related biological process, which stimulated us to do the proteomics analysis to figure out the global changes of histone modifications after treating with GNP1/GNP2. Proteomics result revealed that GNP2 induced a global decrease of the methylation and acetylation of histone H3, which was further validated by ELISA in dTHP-1

cell line and western blot in both dTHP-1 and A549 cell lines. To specify the histone modification marks with related biological pathways, we operated ChIP-seq study and found that H3K27me3 is closely related with neuron and synapse related signalling pathways. We believe this will further guide the investigation of shape-dependent epigenetic modifications with specific phenotype.

Results and Discussions

Characterization of Gold Nanoparticles in Biological Media

The gold nanoparticles were synthesized in a clean environment to limit contamination and characterized by a range of techniques to determine their physiochemical properties. The gold nanoparticles with two distinct shapes (spherical and star-like, namely GNP1 and GNP2, respectively) were illustrated by transmission electron microscopy (TEM) images (Figure 3.1A). TEM images also allowed us to code the shape[21]. Briefly, the contours of nanoparticles were extracted from TEM images by using the Open Source Computer Vision Library (Figure S3.1). Then the coordinates of each contour were used to do the Discrete Fourier Transform (DFT) calculations and principal component analysis (PCA). After this the shape variance and scatter plots of the nanoparticles were obtained (Figure 3.1B-D). Shape variance (Figure 3.1B) showed the distribution of each shape group. The distribution peak became wider and shifted towards higher distance values when the nanoparticles shape was more irregular (from sphere to star). Scatter plots (Figure 3.1C and D) showed the geometrical fluctuations of each shape group. In the scatter plots each point represented a nanoparticle, so the more centre-concentrated of the cluster, the more similar of the nanoparticles shape and different shapes can be identified by clusters with different centres. Considering the nanoparticle shape is controlled *via* irreversible nucleation and growth processes, the synthesis procedure is prone to generate drift from one apparent recognizable shape to another. By the shape contour analysis, we have confirmed that the two distinct shapes were reproduced and consistent throughout the study (Figure S3.2 and S3.3). Based the TEM images, we further evaluated the average size of GNP1 and GNP2. With the definition of 'size' as the longest tip-to-tip distance for GNP2, the average size over 100 nanoparticles was shown to be 53 ± 5 nm and 59 ± 9 nm for GNP1 and GNP2, respectively (Figure 3.1E). The distinct two shapes of gold nanoparticles were further confirmed by the UV-vis absorbance spectra (Figure 3.1F). The absorbance spectrum of GNP2 exhibited a significant shift to the far-red region in comparison to the GNP1, which is characteristic for branched gold nanoparticles of similar diameter. The GNP1 and GNP2 dispersion were incubated with 100% human serum to form hard corona and subsequently dispersed in serum-free macrophage media. The DCS data showed that

the corona-GNP1 and corona-GNP2 exhibited similar colloidal stability compared with the pristine GNP1 and GNP2 dispersed in water (Figure 3.1G). The decrease in calculated apparent size of corona-nanoparticle complexes compared with the pristine nanoparticles is consistent with the formation of corona, which is due to a decrease in the density of complexes. As endotoxin (LPS) stimulates various innate sensors (*e.g.* toll-like receptor 4, TLR4) in macrophages[22], it is important to ensure the absence of LPS contamination. Hence, prior to cell uptake, the concentration of LPS in dispersion containing 1×10^{11} nanoparticles per mL (which is the highest dose used in this study) was measured by the Limulus Amebocyte Lysate (LAL) test. As confirmed, the levels of endotoxin in GNP1 and GNP2 dispersion were well below 0.5 EU/mL (FDA LPS limit for products that directly or indirectly contact the cardiovascular and lymphatic systems) (Figure 3.1H). Other physiochemical properties like hydrodynamic size, zeta potential and computational calculated surface area were shown in table 3.1. Taken together, the gold nanoparticles used in this study are immunologically clean and colloidally stable in relevant biological media. They represent two contrasting particle shapes with similar size and surface area, providing a good model system to uncover differential impact of particle shape on cellular responses.

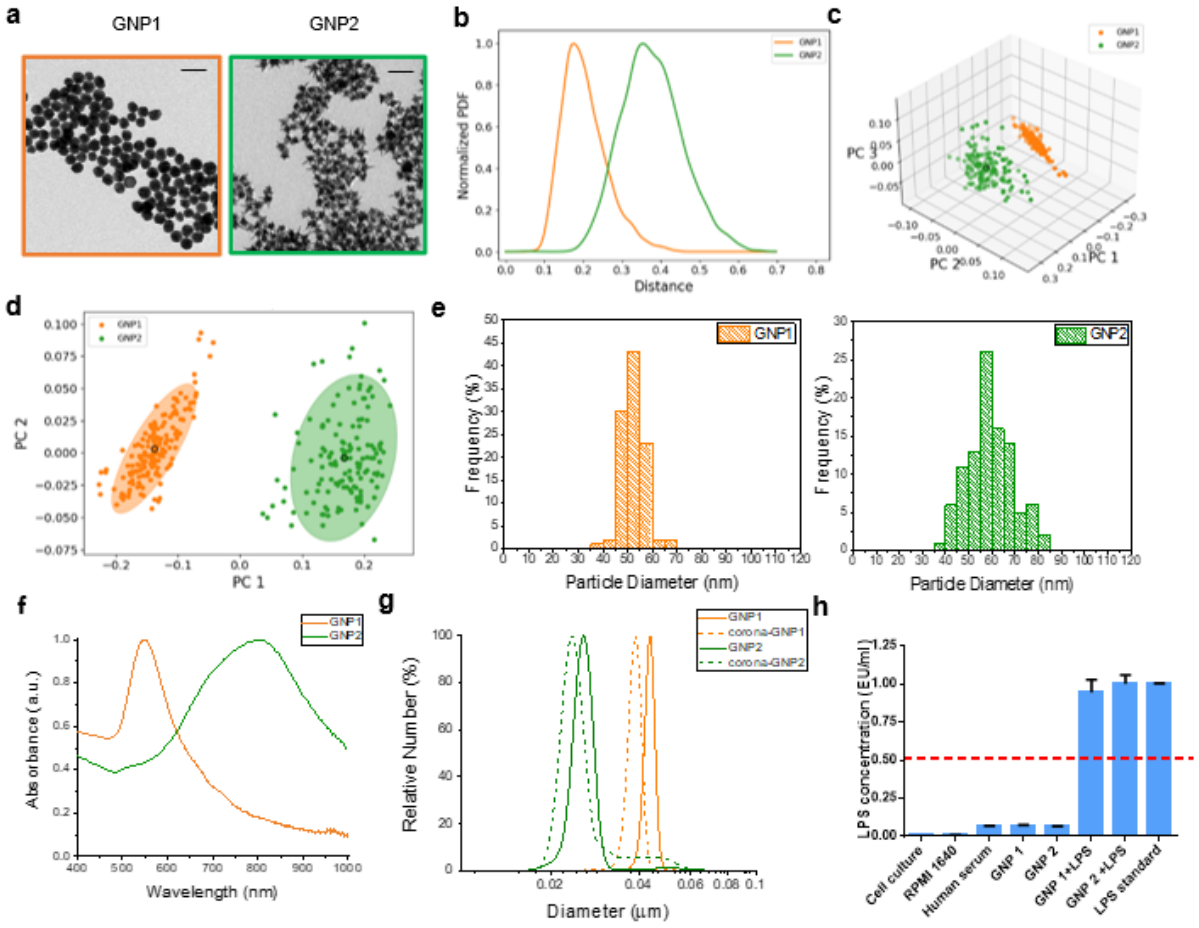


Figure 3. 1 Characterization of GNPs.

(a) Transmission electron microscopy (TEM) micrographs of GNPs, scale bar 100 nm. (b) Shape variance, (c) 3D principle component analysis (PCA) scatter plot and (d) 2D PCA scatter plot of GNPs from the shape contour analysis. (e) Histograms representing the size distribution of GNPs. The average size of GNP1 is 53 ± 5 nm, GNP2 is 59 ± 9 nm. The average size of GNP2 is the longest tip-to-tip distance. More than 100 particles were counted. (f) UV-vis absorption spectra analysis. (g) Differential centrifugal sedimentation (DCS) analysis of GNPs and corona-GNPs. In the case of GNP2, due to the DCS software incorrect estimation, it shows a representative particle distribution instead of the real particle size. (h) Limulus Amebocyte Lysate (LAL) test of GNPs.

Distinct Transcriptome Induced by GNP1 and GNP2

To investigate the biological effect at cell level, we incubated the human serum coated-GNP1 or -GNP2 with a human macrophage model (dTHP-1) in serum free macrophage cell culture media. The human macrophage-like cells were obtained by differentiation of a human leukemia monocytic cell line (THP-1) with phorbol-12-myristate-13-acetate (PMA). The macrophage-like differentiation was optimized by examining the cell morphology by optical microscopy and surface marker CD14 by flow cytometry (Figure S3.4). Firstly, we intended to identify the suitable GNP concentration for the cellular study. MTS cell assay (a test for quantification of mitochondrial activity) and LDH-release assay (a test for examining the integrity of cell membrane by measuring the LDH release in the cell media) were employed to ensure negligible cytotoxicity induced by the treatment with GNPs. It was shown that dTHP-1 cells did not exhibit significant viability change compared with untreated cells after the incubation with different concentration of GNPs (1×10^{11} GNPs/mL as the highest concentration tested) for 24 h (Figure S3.5A and B). The negligible cytotoxicity within such concentration range of GNPs was also confirmed in murine macrophage cell line, RAW 264.7 and Human lung adenocarcinoma epithelial cell line, A549 (Figure 3.5C-F).

Cellular uptake, as one of the important aspects for evaluating GNP-cell interactions, was quantified by Inductively Coupled Plasma Mass Spectrometry (ICP-MS). The macrophage-like cells, dTHP-1, were exposed to GNP1 and GNP2 at 1e11 GNPs/mL for different time intervals up to 24 h. The gold content of internalized by the cells and remained in the cell media was determined, and the percentage of GNPs that were internalized was calculated and compared between GNP1 and GNP2. It is shown that GNP1 exhibited an enhanced uptake compared with GNP2 at all the time intervals (Figure S3.6A). Consistently, the higher uptake of GNP1 was demonstrated when three different concentrations of GNPs were applied (Figure S3.6B). The internalization of GNPs was further confirmed by TEM, where the majority of GNP1 and GNP2 was shown to be confined in membrane-enclosed vesicles (Figure S3.6C).

It is also noted that the GNP-containing vesicles appeared in a diverse range of size and morphology, suggesting complex intracellular trafficking trajectories involving vesicle scission and fusion for both GNP1 and GNP2 intracellular processing.

Next, we chose to analyse the transcriptome of dTHP-1 cells after the incubation with GNP1 and GNP2 (*i.e.* 1×10^{11} GNPs/mL for 4h and 24 h) to reveal key global cellular responses. The total RNA was extracted from the dTHP-1 cells after the treatment. To identify the transcriptional changes induced by GNPs, a dual-colour Agilent whole human genome transcriptome microarray was used. Briefly, the cDNA generated from GNP-treated cells and untreated cells were labelled independently with Cy3 and Cy5 and co-hybridised on the array slides. The ratio of signal intensities after background correction and Loess normalization was used as a relative measure of gene expression against untreated cells. Volcano plot was used to highlight the Differentially Expressed Genes (DEGs) with threshold of fold-change > 1.5 and p-value < 0.05 (One sample t-test). After 4 h incubation, 88 and 41 DEGs were identified for the treatment with GNP1 and GNP2, respectively (supporting Figure 7). However, these DEGs did not result in enrichment of known networks or pathways. After 24 h incubation widespread changes in transcriptome were observed with 3013 and 3604 DEGs for GNP1 and GNP2, respectively (Figure 3.2A and B). It is also noted that most of the DEGs were found to be downregulated after the treatment of GNPs (2693 out of 3013 for GNP1, 3364 out of 3604 for GNP2). To differentiate the effect of particle shape on cellular responses, we directly compared the transcriptome between GNP1 and GNP2 by a volcano plot. 523 DEGs were identified (Figure 3.2C) and clustered for enrichment analysis by Metascape[23] (Figure 3.2D and E). We found that several overrepresented GO-Biological Processing terms (boxed in Figure 3.2D) are related epigenome modifications. These statistically enriched terms were then clusters into a hierarchical tree to visualize the similarity between the GO terms (Figure 3.2E). The five clusters formed two completely separate networks, and within a network the terms were only occasionally bridged, suggesting that the epigenome modifications involve distinct processes. It is noted that ‘histone modification’ is related with ‘negative regulation of chromosome organization’ in the enrichment network.

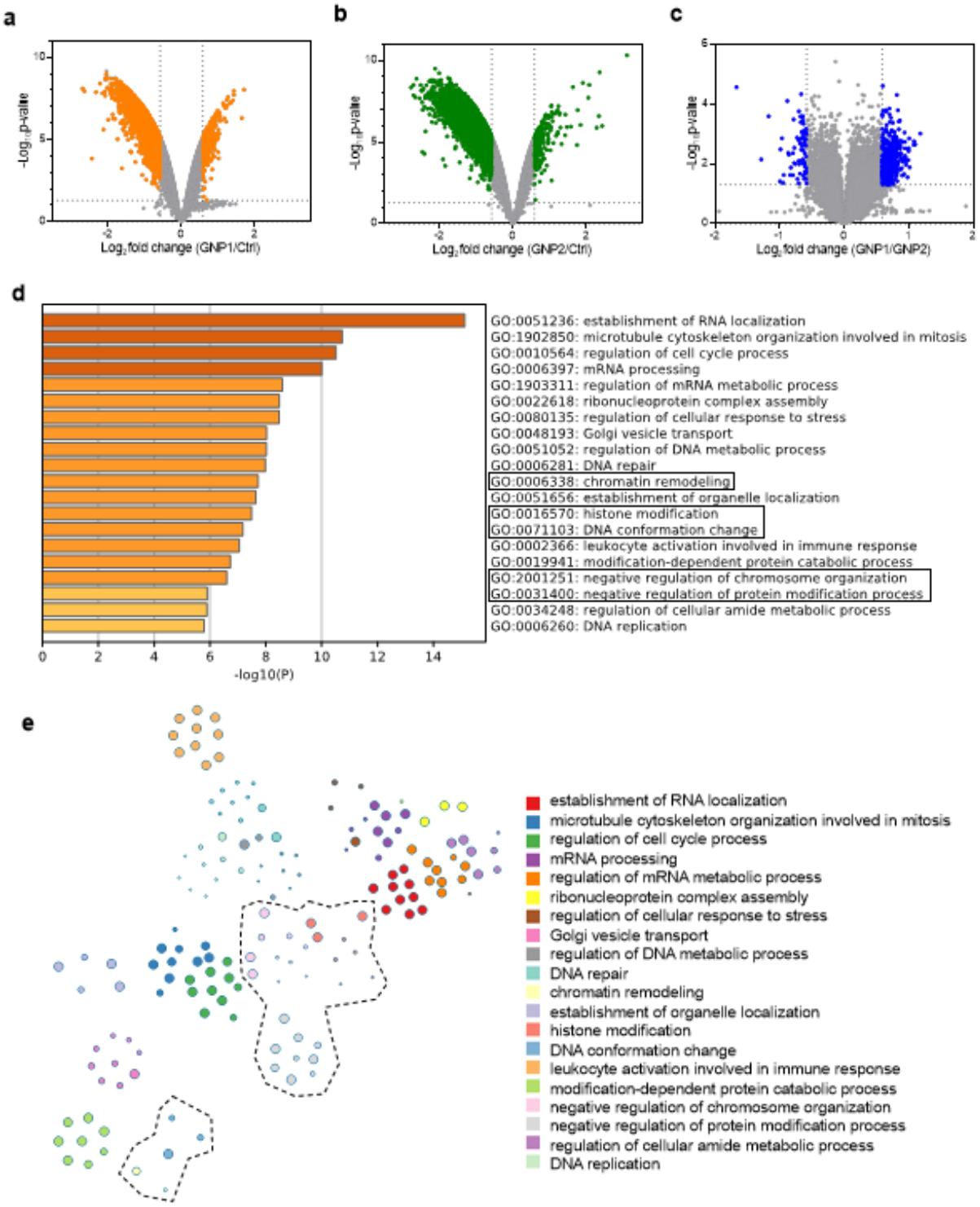


Figure 3. 2 Transcriptome analysis of dTHP-1 treated with GNP1 or GNP2 for 24 h.

Volcano plots indicating differentially expressed genes between untreated and GNP1 treated cells (a), untreated and GNP2 treated cells (b), and GNP1 and GNP2 treated cells (c). Each coloured dot indicates a differentially expressed gene which fulfils the threshold: $\log_2|FC|>\log2(1.5)$ with P-value<0.05. (d) GO enrichment analysis (biological processes) of DEGs between GNP1 and GNP2. GO terms were plotted against $-\log_{10}$ false discovery rate

(FDR) value. (e) Cluster analysis (Metascape) of the significantly enriched biological processes.

Differential Effects of GNP1 and GNP2 on Histone Methylation and Acetylation

Methylation and acetylation are the two most studied histone modifications that are intimately connected with chromosome organization[24]. Hence, we employed a bottom-up mass-spectrometry (MS) based strategy to obtain a comprehensive profile of methylation and acetylation motifs on histones (Figure 3.3A). Briefly, after treatment with GNP1 or GNP2 at concentration of 1e10 GNPs/mL, histones (H2A, H2B, H3, and H4) were extracted and purified as described previously[25]. The purified histones were firstly derivatised with deuterated acetic anhydride (D3AA), where unmodified and mono-methylated lysine were decorated with deuterated acetyl groups, then trypsinized, and finally injected to liquid chromatography (LC)-MS device. The abundances of the peptides were compared to control group and expressed in heatmap (Figure 3.3B). Firstly, we found that all the detected significant changes in methylation and acetylation upon the treatment with the GNPs were harboured in H3. Secondly, GNP1 and GNP2 both showed rather broad effects on H3 PTMs, contributing to many motifs. Notably, overall GNP1 exhibited fewer changes and/or to a less extent in comparison to GNP2. GNP1 exhibited 7 motifs that significantly differed from control group, whereas GNP2 showed 13 motifs (Table 3.2 for p-value). Furthermore, the treatment of GNP2 illustrated a consistent demethylation trend on K27, as 5 out of 13 motifs were associated with demethylation of K27 that were also absent in GNP1 treatment. As an example, the MS2 spectrum of H3K27me3 and the extracted ion signal of different treatment groups are shown in Figure 3.3C. Next, we employed H3K27me3 ELISA to validate the decrease of H3K27me3 induced by GNP2. As a quality control, a known methyl-transferase inhibitor, GSK, was used. Consistently, around 30% reduction of H3K27me3 was illustrated with the treatment of GNP2, whereas no significant change in the GNP1 group (Figure 3.3D).

H3K27me3 deposition is mainly controlled by the lysine methyltransferase of the Polycomb Repressive Complex 2 (PRC2)[26-30]. Enhancer of Zeste 2 (EZH2) is the enzymatic member of PRC2, which catalyses the trimethylation of H3K27[31]. To investigate the role of PRC2 in the decreased H3K27me3, we treated the cells with the GNPs at 1×10^{10} NPs/mL for 16 and 24 h, respectively, and then examined the abundance of EZH2 and H3K27me3 by western blot. In the cells treated with GNP2, a decrease in EZH2 and H3K27me3 was observed for both time intervals (Figure 3.3E). The mild decrease of H3K27me3 was confirmed by a number of biological repeats (Figure S3.8). Taken together, the down regulation of H3K27me3 is a result of GNP2 treatment which is not apparent in the GNP1 treated cells. This modification likely involves a change in the activity of PRC2.

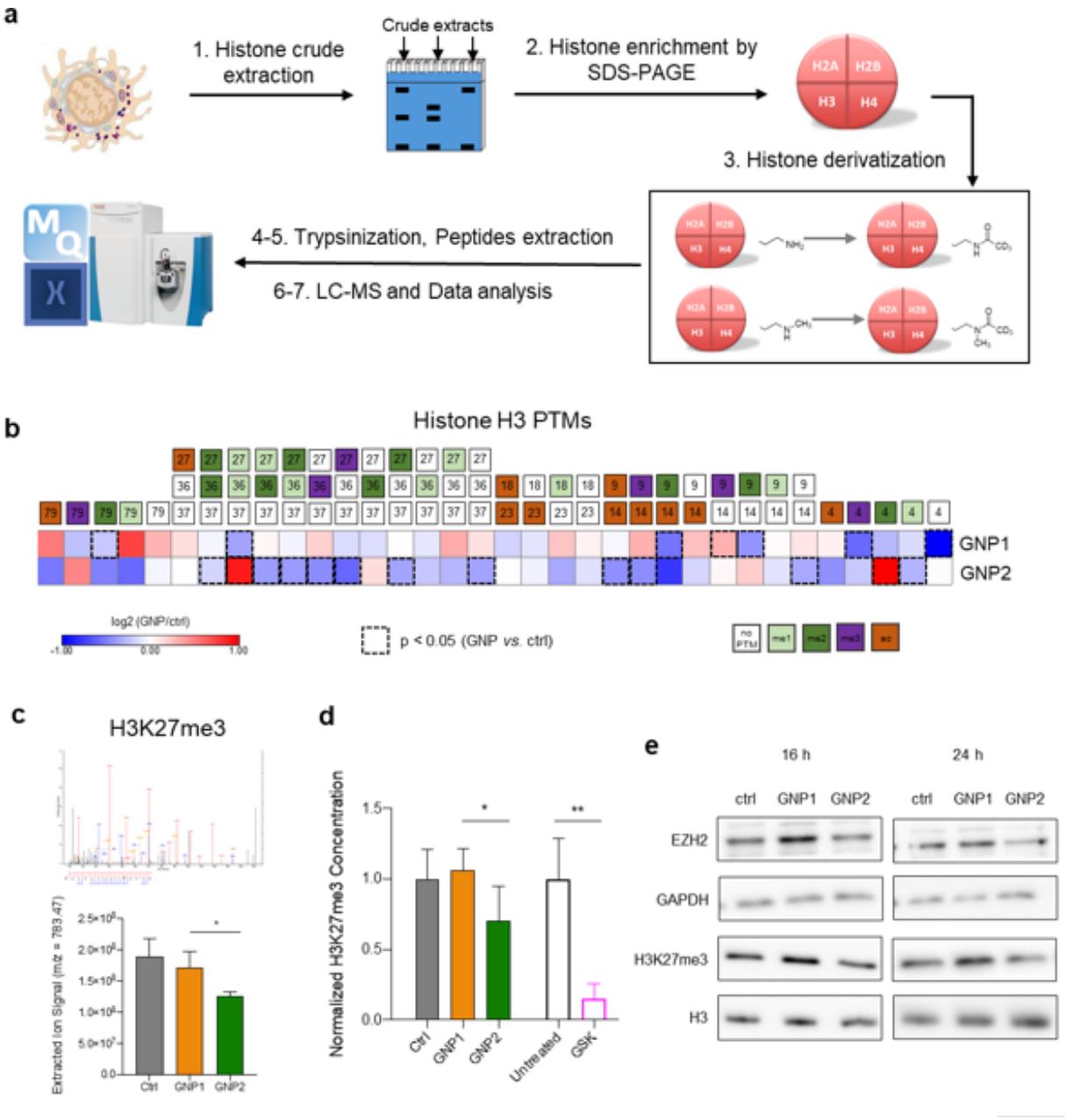


Figure 3. 3 Global Changes of histone methylation and acetylation after treatment with GNP1 or GNP2.

(a) Scheme showing proteomic-based histone modification analysis on identification and quantification of modified histone H3 peptides. (b) Heatmap of proteomics showing hierarchical clustering analysis of GNP1 and GNP2-induced the post-translational modifications on histone H3. (c) MS/MS spectrum of -27KSAPATGGVKKPHR40- peptide with tri-methylation on Lys27 (H3K27me3) and the extracted Ion signals of H3K27me3 peptides. (d) Concentration of H3K27me3 quantified by ELISA. Data of untreated or ctrl was set as 1. GSK treated cell lysate was used as a positive control. (e) Western blot of EZH2 and H3K27me3 after dTHP-1 cells were treated with GNP1 or GNP2 for 16 h or 24 h.

Distinct subsets of genes marked with H3K27me3 between GNP1 and GNP2

Considering that differentiation of THP-1 cells from monocytes to macrophages is controlled by reprogramming of histone methylation (*e.g.*, H3K27me3)[32], the dTHP-1 cells could harbor highly heterogeneous H3K27me3 mark prior to the nanoparticle treatment, which would give rise to significant challenges for precise mapping locations of H3K27me3 modification specifically induced by nanoscale shape. To address this, we firstly used western blot to confirm that a drastic decrease of H3K27me3 modification in A549 cells was induced by GNP2 treatment, while a much subtler decrease was observed with GNP1 treatment (Figure 3.4A). We then used A549 cells to profile H3K27me3 modifications that are differentially changed between untreated, GNP1- or GNP2-treated by chromatin immunoprecipitation sequencing (ChIP-seq).

We optimized the crosslink and sonication steps to generate 200-500 bp small DNA fragments (SI9a). Subsequently, we employed ChIP-qPCR with known H3K27me3 negative and positive primers to determine the optimal antibody concentration to improve signal to background ratio (SI9b). We chose 50 million reads for each sample as H3K27me3 involves large enrichment domains and employed MACS[33] for peak calling. The correlation heatmap showed that the samples (two independent replicates for each treatment) were clustered before and after ChIP indicating a successful ChIP (Figure S3.10).

To get robust binding sites of H3K27me3, we identified consensus peaks (peaks present at least three times in the four biological replicates) for each treatment group. 25141, 16254 and 15025 consensus peaks were found in untreated (control), GNP1-, or GNP2- treated cells, respectively (Figure 3.4B). The fewest number of peaks in the GNP2-treated cells is consistent with the substantial decrease of H3K27me3 in the western blot. We then determined the overall distribution pattern of H3K27me3 peaks in genomic regions. H3K27me3 peaks were significantly enriched in intergenic regions (nearly 40%) (Figure S3.11), which is consistent with the strong bias of H3K27me3 towards intergenic regions observed in many studies[34-36]. The treatment of GNP1 or GNP2 did not greatly alter the overall genomic distribution of H3K27me3, however, a notable increase in promotor region was observed for GNP1 and GNP2 treatment groups. After this we did the differential binding cites analysis with a threshold fold change > 2 and p-value < 0.05 (Figure 3.4C and Figure S3.12A). We investigated this further by annotating these differential binding sites to their closet gene, 1311, 1260, 1094 genes we identified for GNP1 vs. GNP2 group, GNP1 vs. control group, GNP2 vs. control group respectively (Figure 3.4D). Subsequently, we performed functional enrichment analysis to identify predominant biological themes based on GO ontology. The top 10 enriched pathways were visualized using the bar plot (Figure 3.4E and Figure S3.12B). The genome

track of representative genes which were annotated to the top 10 GO terms was shown in Figure 3.4F and Figure S3.12C. All the above suggests that H3K27me3 mark was specifically changed in the genes associated with these pathways. More importantly, GNP1 and GNP2 show different regulation of neuron and synapse related signaling pathways.

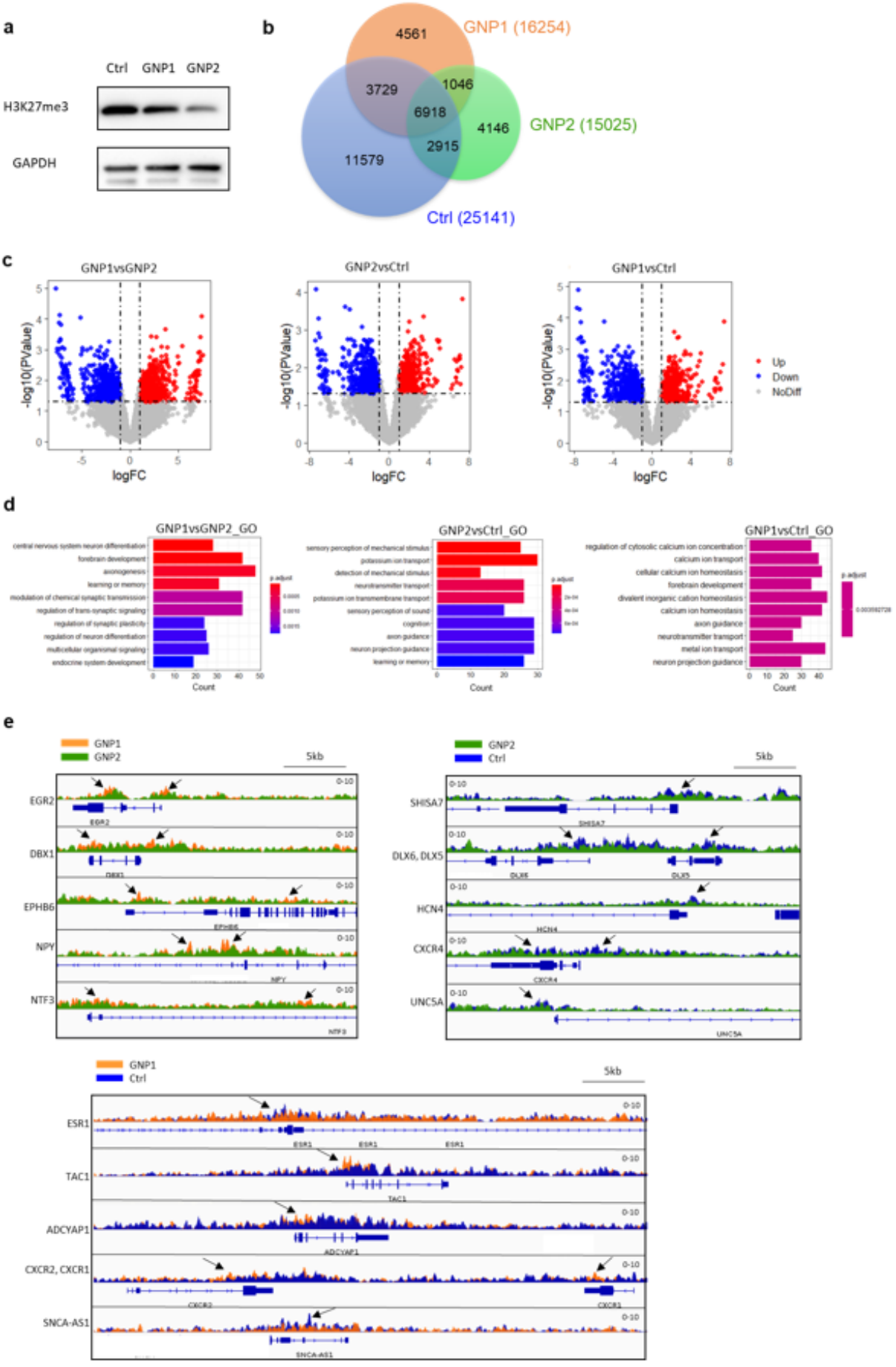

a
Ctrl GNP1 GNP2
H3K27me3
GAPDH
b
GNP1 (16254)
GNP2 (15025)
Ctrl (25141)
4561
3729
1046
6918
4146
2915
11579
c
GNP1vsGNP2
GNP2vsCtrl
GNP1vsCtrl
-log10(PValue)
logFC
Up
Down
NoDiff
d
GNP1vsGNP2_GO
GNP2vsCtrl_GO
GNP1vsCtrl_GO
Count
e
GNP1
GNP2
Ctrl
5kb
EGR2
DBX1
EPHB6
NPY
NTF3
SHISA7
DLX6, DLX5
HCN4
CXCR4
UNC5A
ESR1
TAC1
ADCYAP1
CXCR2, CXCR1
SNCA-AS1

Figure 3. 4 Comparison of H3K27me3 locations and the functions of associated genes with H3K27me3 mark after the treatment of GNP1 or GNP2.

(a) Western blot showing the expression of H3K27me3 and (b) Venn diagram showing the number of peaks marked by H3K27me3 in different treatment groups by H3K27me3 ChIP-seq with two biological replicates. (c) Volcano plot showing differential binding sites of H3K27me3 ChIP-seq. (d) GO pathway enrichment analysis of the annotated differential expression genes in H3K27me3 ChIP-seq. (e) Representative genes IGV tracks showing the differences between different treatment groups in H3K27me3 ChIP-seq.

Conclusion

In this paper we used two distinctive shapes, spherical (GNP1) and star-like (GNP2) gold nanoparticles, illustrated the shape effects on epigenome. To avoid any potential contamination which can lead to the biological process changes of cells, all of the nanoparticles used in biological investigation were strictly synthesized in a clean room with high standard operation protocol and the LPS level of all the nanoparticles was detected before exposing to cells, only nanoparticles with LPS level lower than 0.5 EU/ml were used for cellular exposure. Then various characterization methods were used to identify the physiochemical properties of the nanoparticles and also ensure the reproducibility of nanoparticles used in biological investigations. Afterwards, the two different nano-shapes with similar size were used to explore their impacts on cellular profile. Transcriptome study on GNP1 and GNP2 treated macrophage cells revealed that GNP1 and GNP2 showed a different regulation of genes related with histone modifications biological process. Proteomics analysis confirmed the finding that the nano-shape plays an important role in the histone modifications, which GNP2 showed a global decrease in the methylation and acetylation of histone H3. Western blot and ELISA results also validated this discovery. Further ChIP-seq study revealed that the H3K27me3 mark is closely correlated with neuron and synapse related signalling pathways, which GNP1 and GNP2 showed a significant difference in the regulation of genes associated with these biological pathways.

When it comes to the nano-shape effects in the biological process, it is an intriguing field with many issues unsolved. For instance, it becomes to be widely accepted that the biomolecular corona gives a new identity of nanoparticles when they interact with biological environment[2, 6], which monitors the downstream biological outcomes induced by nanoparticles. However, how the cells recognise the nano-shape and give a new identity of the shape which plays an important role in the following biological process are still unclear. Questions like these are

subtle and need time to figure out. We believe decoupling the surface and shape effects of nanoparticles is the next step for the investigation of nanoscale shape biology.

Detailed and elaborated exploration needed to fully understand how the shape regulate the epigenetic modifications and further link these modifications to specific phenotype or biological process. Still, the shape-dependent histone modifications discovery in this paper is potentially highly significant, revealing that some of the previously reported shape induced biological outcomes could be due to the regulation of these epigenetic markers which was never investigated, and this discovery greatly stimulated the clarification of the mechanisms behind the nanoscale shape caused biological outcomes. The results reported here also raise questions regarding the human health and environmental risk assessment of our almost unlimited ability to produce huge varieties of the nanoscale shapes. To establish a well-defined risk assessment of the safety of nanoparticles especially the giant universe of nano-shapes, it would be necessary to do a carefully and fully investigation of the shape effects on epigenome. The central purpose of this paper is to trigger the investigation of nanoscale shape effects on the epigenome, which didn't receive adequate due attention in the nanoscale shape biology study field during the past two decades. We believe that the work presented in the paper will set a model for the exploration of the shape-dependent epigenetic effects from the nanoparticles' synthesis to characterization and finally the biological impacts investigation, enable further correlation of the shape-dependent epigenetic modifications to specific biological outcomes, validation of the biosafety of nanoscale shapes and potential nanomedicine application. All of the above will finally frame a sound foundation for future research in the field.

Methods

Chemicals

Two shapes of gold nanoparticles were prepared following the methods described in literature[37]. After synthesis, ligand exchange reaction was processed on GNP1 and GNP2 with Bis(*p*-sulfonatophenyl)phenylphosphine dehydrate dipotassium salt (BSPP). The following chemicals: Hydrogen tetrachoroaurate trihydrate ($HAuCl_4 \cdot 3H_2O$), Trisodium citrate dehydrate ($C_6H_9Na_3O_9$), Tannic acid, ($C_{76}H_{52}O_{46}$), Potassium Carbonate (K_2CO_3), Glycerol ($C_3H_8O_3$), Silver nitrate ($AgNO_3$), phorbol 12-myristate 13-acetate (PMA), clean water (CHROMASOLV Plus, for HPLC) and (LC−MS Ultra CHROMASOLV, tested for UHPLC−MS), Acetic anhydride-d6 [$(CD_3CO)_2O$], Hydrochloric acid (HCl), Ammonium bicarbonate (NH_4HCO_3), Acetonitrile (CH_3CN), Trypsin (TPCK-treated), Tri-fluoroacetic acid (TFA), Formic Acid (FA), Trizma base ($NH_2C(CH_2OH)_3$)，DL-dithiothreitol (DTT), Sodium dodecyl sulphate (SDS), Sodium

phosphate dibasic dehydrate ($Na_2HPO_4 \cdot 2H_2O$) were purchased from Sigma-Aldrich. Rabbit anti-histone H3 antibody (ab1791), Rabbit anti-histone H3 (tri-methyl K4) antibody (ab8580), were purchased from Abcam. Rabbit anti-GAPDH antibody (2118), Rabbit anti-Tri-methyl-Histone H3 (Lys27) antibody (9733S), Rabbit anti-Tri-methyl-Histone H3 (Lys9) antibody (13969S), Normal Rabbit IgG antibody (7074S) were purchased from Cell Signalling Technology.

NPs Characterization

Characterisation of the gold nanoparticles were performed by Transmission electron microscopy (TEM), Nanoparticle Tracking Analysis (NTA), Ultraviolet-visible spectroscopy (UV-Vis), Dynamic Light Scattering (DLS) and Differential Centrifugal sedimentation (DCS). The shape and size of the nanoparticles were confirmed by TEM images. Computational identification of shapes were performed following a published procedure[21]. Surface area was calculated based on the extracted contour during the computational shape analysis. The concentration of gold nanoparticles was defined as number of particles in certain volume (particles/ml), which was determined by NTA. Distribution and stability of gold nanoparticles in biological medium was checked by DLS and DCS analysis.

Endotoxin contamination control

Endotoxin level in all materials involved in cell culture was examined by Pierce LAL Chromogenic Endotoxin Quantitation Kit (Thermo Fisher). 250 µl of cell culture were centrifuged at 300 rcf and the supernatant were collected for testing. Gold nanoparticles were diluted to 1×10^{11} particles/ml with endotoxin-free water prior to testing. The testing was performed according to manufacturer's instruction.

Cell Culture

THP-1 cell line (DSMZ) were maintained at 1×10^5 to 8×10^5 cells/ml in RPMI 1640 medium (Gibco) supplied with 10% heated-inactivated (56 ℃, 30 min) Fetal Bovine Serum (FBS, Gibco), 100 units/ml of penicillin and 100µg/ml of streptomycin (Gibco) in 5% CO_2 at 37 ℃. THP-1 cells (4×10^5 cells/ml) were differentiated using 100 nM PMA for 48 h. The PMA-treated cells were washed with fresh growth medium and incubated in growth medium for 24 h prior to nanoparticle treatment. Gold nanoparticles were pre-coated by incubating with 100% heated-inactivated human serum (Sigma-Aldrich) for 1 hour at 37 ℃. Pre-coated gold nanoparticles were re-dispersed in macrophage serum-free medium (Gibco) to working concentration before applied to the cells. A549 cell line (ATCC CCL-185) were cultured in MEM (Life Technologies) supplied with 10% FBS, 50 units/ml of penicillin and 50µg/ml of streptomycin (referred to hereafter as cMEM) in 5% CO_2 at 37 ℃.

Cell viability of RAW 264.7 and dTHP-1 was tested by CellTiter 96® AQueous One Solution Cell Proliferation Assay (MTS assay, Promega) and CytoTox 96® Non-Radioactive Cytotoxicity Assay (LDH assay, Promega). Cell culture medium was removed after 24 hour incubation with different concentrations of gold nanoparticles. The cells were incubated with CellTiter 96® AQueous One reagent contained medium for 30 min in 5% CO_2 at 37 ℃. The absorbance at 492 nm was recorded by 96-well plate reader VarioSkan (Thermo, USA). The medium removed from the plate was tested with LDH assay according to the manufacture's instruction. The absorbance at 492 nm was recorded by 96-well plate reader. Cell viability of A549 was tested by tetramethylrhodamine, methyl ester (TMRM) assay and Live/Dead cell count assay. A549 cells were treated with GNP1 or GNP2 for 4 h pulse 24 h chase, with different concentrations. Negative control is untreated A549, positive control is A549 treated with toxic PS-NH_2 with a concentration of 10mg/ml. TMRM intensities were normalized by negative control, which was set as 100%. Hoechst was used for staining the nuclei, TOPRO-3 was used for staining dead cells in the 24h live/dead cell count. Experiment was carried out by high content, all conditions performed in triplicate.

Cellular Uptake of NPs

The internalised and free gold content in cell culture was quantified by ICP-MS. 2.7×10^5 cells/ml of differentiated THP-1 macrophages were exposed to medium containing 1×10^{11} particles/ml gold nanoparticles for 2 h, 4 h, 16 h and 24 h, respectively. The cells were washed with PBS for twice. Both medium and washes were collected and centrifuged at 20817 rcf for 40 min to recover all free gold nanoparticles in cell culture. The nanoparticle pellet was re-suspended in 100 µl phosphate buffered saline (PBS, Gibco). The cells were detached with Accutase and collected in 100 µl PBS. The samples were digested with aqua regia and diluted for measurements.

TEM

Cells were differentiated in 24-well plate with glass coverslip at the bottom. Nanoparticle-treated cells were washed with PBS and fixed with 2.5% glutaraldehyde in 0.1 M Sorensen's phosphate buffer at 4 ℃ and preliminarily stained with 0.1% OsO_4 for 1 hour. Dehydrating of cells were processed by exposing to 30%, 50%, 70%, 90% and 100% (twice) ethanol for 10 min each. The glass coverslip holding cells were dip to propylene oxide and covered with TAAB Epon resin 812, which was prepared by mixing 12 g Agar 100 resin, 4.75 g Dodecenyl Succinic Anhydride (DDSA), 8.25 g Methyl Nadic Anhydride (MNA) and 0.5 g 2,4,6-tri (dimethylaminoethyl) phenol (DMP-30) and incubated at 37 ℃ for 2 hour to allow Epon infiltration. The Epon was polymerised by baking at 60 ℃ overnight. 80 nm sections were produced by Ultramicrotome (Leica EM UC6) and diamond knife (DiATOME). The sections

were post-stained with 2% uranyl acetate for 20 min and lead citrate for 2 min before viewing under FEI Tecnai 120.

Transcriptome Microarray

THP-1 momocytes were seeded at 2.5 × 10^5 cells/ml in 6-well plate and differentiated into THP-1 macrophages as described above, before nanoparticles exposure. Heated-inactivated human serum coated gold nanoparticles were dispersed in medium at concentration of 1 × 10^{11} particles/ml before applying to cells. After 4-hour and 24-hour treatments, the cells were washed with ice-cold PBS for twice. RNA of each well was extracted by Invitrap® Spin cell RNA mini kit following manufacturer's instruction. The concentration and purity of RNA was examined by Nanodrop. RNA integrity numver (RIN), which indicates the quality of RNA was determined by Bioanalyser (Aligent 2100) on Aligent RNA 6000 Nano kit. 20 µl of RNA was transferred into the center of each RNAstable tube (Merck), which is designed for long-term preservation of RNA samples at room temperature. The samples were then dried using a vaccum concentrator without heat for around 30 min and stored in a heat-sealed pouch with a desiccant packet provided with the tubes. Transcriptome microarray was performed by Inge Nelissen (VITO NV, Belgium).

Proteomics

Histone post-translational modification analysis was carried out with a quantitative proteomic strategy described by Feller, et.al. in 2015[38]. Nanoparticle treated cells were collected by cell scrapper after washing with ice-cold PBS. Crude extraction of histones with other acid-soluble proteins were processed by suspending cell pellets in 0.2 M sulphuric acid overnight. Histone extracts were precipitated with 26% TCA and re-suspended in neutralised buffer (pH 8.0) after 3 times washes with ice-cold acetone. Further histone purification was carried out by separating crude extracts on 15% SDS-PAGE gel. The gels were stained with Coomassie blue and bands between 22 kDa and 11 kDa were isolated from the gels. Gel pieces were washed twice with HPLC water and twice with 100 mM Ammonium bicarbonate buffer (ABC buffer). Destaining of the gel pieces were processed by washing three times with 50 mM ABC/50% ACN buffer for 10 min with shaking at 37 ℃. The decoloured gels were incubated with 100 mM and 20 mM ABC for 15 min respectively and incubated three times with 100% ACN for 10 min to be dehydrated. 10 µl of D3AA were added to dehydrated gels followed by another 20 µl of 100 mM ABC buffer. Another 70 µl of 100 mM ABC were added to cover the gel pieces and incubated for 45 min at 37 ℃ for chemical acetylation. The pH of the reaction was adjusted to pH 7.0-8.0 with 1 M ABC buffer at 5 min after the start of incubation. After acetylation, the gel pieces were washed and incubated with 100 Mm ABC buffer and ACN for dehydration. Digestion was carried out by incubating the dehydrated gel pieces with Trypsin

overnight at 37 ℃. Peptides were recovered by incubating two times with 70% ACN/0.25% TFA for 10 min at 37 ℃ followed by two times with 100% ACN. The harvested peptides were concentrated and harvested I 0.1% TFA. The purification and desalting of the peptides were carried out with 100 µL bed Pierce™ C18 tips (Thermo Fisher Scientific). The samples were injected into the Dionex Ultimate 3000 UPLC (Thermo Fisher scientific) connected to a Thermo Scientific Q-Exactive mass spectrometer. LC-MS data was analysed by Maxquant software (version 1.5.3.3) and Xcalibur software (Thermo Fisher Scientific).

Western blot

The nanoparticle treated cells were lysed in Laemmli buffer. Cell lysates were incubated at 97 ℃ for 2 min followed by 30 seconds sonication and another 5 min incubation at 97 ℃. The lysates were then centrifuged at 20,817 g for 10 min to remove nanoparticles and insolubilities. 10 µg lysate was separated on 15% SDS-PAGE gel and transferred to PVDF membrane with pore size of 0.2 µm. The membranes were then blocked with 5% skim milk in TBS-T buffer (150 mM NaCl, 10 mM Trizma base, 0.1% Tween-20, pH 7.5) for 1 hour. The membranes were further incubated with primary antibody diluted in blocking buffer at 4 ℃ overnight. After 4 times washes with TBS-T buffer for 10 min, the membranes were incubated with HRP-labelled secondary antibody for 1 hour and soaked four times in TBS-T buffer for 10 min. The membranes were developed with PierceTM ECL Western blotting substrate mix solution and visualised in Syngene G:BOX imaging system.

ELISA

The nanoparticle treated cells were washed with pre-cooled PBS and collected with cell scraper. The histones were extracted by re-suspended the cell pellets in 0.4 M HCl and incubated on ice with occasionally inverting for 30 min. The supernatant fraction containing acid soluble proteins were harvested by centrifuging at 11,000 g for 10 min and neutralised by 2/5 total volume of 1 M sodium phosphate, dibasic (pH 12.5) containing 2.5 mM DTT and protease inhibitor. The ELISA assay was processed following manufacturer's protocol.

ChIP and ChIP-seq

ChIP experiment was done following the protocol described in literatures[39-41]. Briefly, a total of 10^8 A549 cells were 1% formaldehyde-fixed, lysed and sonicated for 16 cycles (30 s sonication 55 s rest) in a Bioruptor sonicator (Diagenode, Denville, NJ, USA). Supernatants were precleared with pierce™ protein A/G magnetic beads (Thermo Scientific). Chromatin fragments were immune precipitated by using rabbit H3K27me3 antibody or normal rabbit IgG antibody (Cell Signalling) and purified using QIAquick PCR purification columns (Qiagen, Hilden, Germany). DNA samples were shipped to BGI on dry ice for ChIP-seq libraty

preparation and sequencing. Samples' quality was checked by bioanalyzer 2100 before sequencing, sequencing was performed using DNBseq system (BGI) with each library sequenced in 50 base pairs single read run with 50 million reads for each sample. Reads were aligned to the Homo sapiens genome (hg 38) using Bowtie 2[42] with default parameters. Peak calling was performed by using MACS[33] with command for broad peaks. ChIP-seq was performed in two biological replicates and Diffbind[43-44] was used to find the consensus peaks and the differential binding sites for each treatment. The differential binding sites annotated genes were used for further network and ontogeny analysis by ChIPseeker[45] and clusterProfiler[46] in R package.

Author Information

Y.Y. and K.A.D. conceived and developed the overall concept of the paper. Y.Y. and W.Z. wrote the paper. Y.Y. led the project overall and interpreted and supervised the experiments. W.Z. was central to the nanoparticles' synthesis, characterization, A549 cell line related experiments including cytotoxicity, western blot and ChIP-seq, and performed the data analysis. J.L. carried out the THP-1 cell lines related experiment, including ICPMS, transcriptome, proteome, western blot and ELISA, with V.A.G.'s assistant. Q.C. and C.P.S supplied some of the GNPs used in the THP-1 cell line.

§. These authors contributed equally: Wei Zhang and Jingji Li.

Supporting Figures

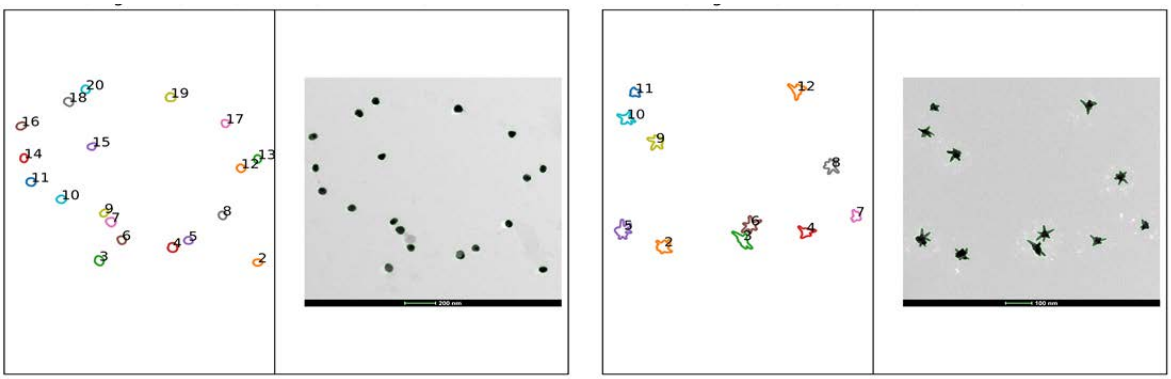


Figure S3. 1 Examples of contour extraction from TEM images.

Each nanoparticle is labelled with a number and recorded as a file.dat which contains the coordinates of the contour coding the shape.

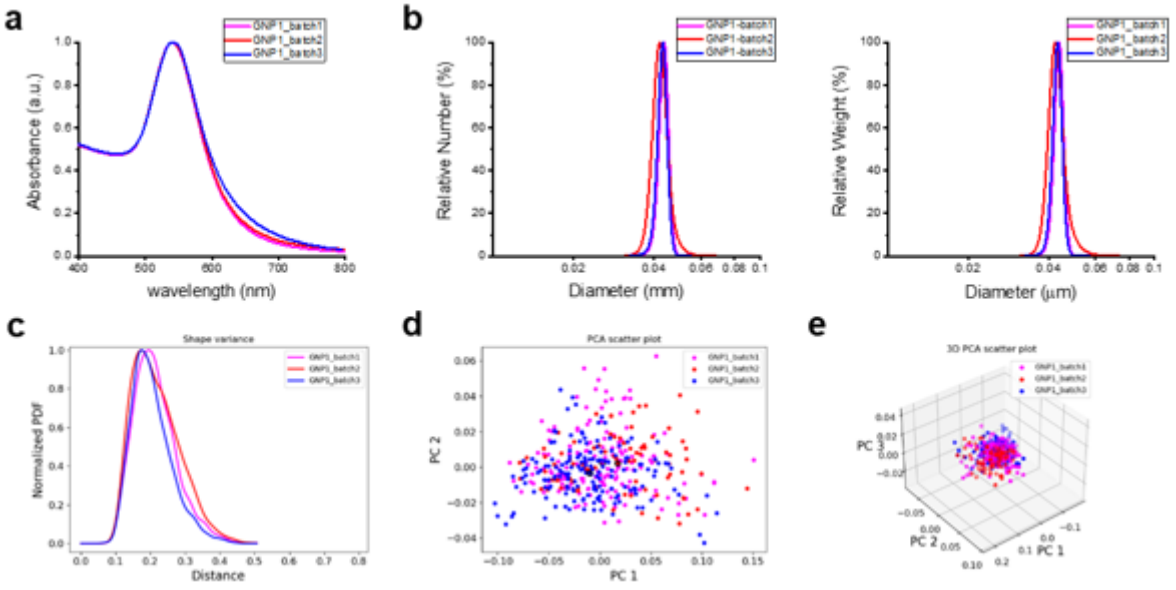


Figure S3. 2 Reproducibility between independent batches of GNP1.

(a) UV-vis absorption spectra. (b) DCS analysis by relative weight and number. (c) Intra-batch shape variance, (d) 2D scatter plot and (e) 3D scatter plot of the principal components from the contour's analysis.

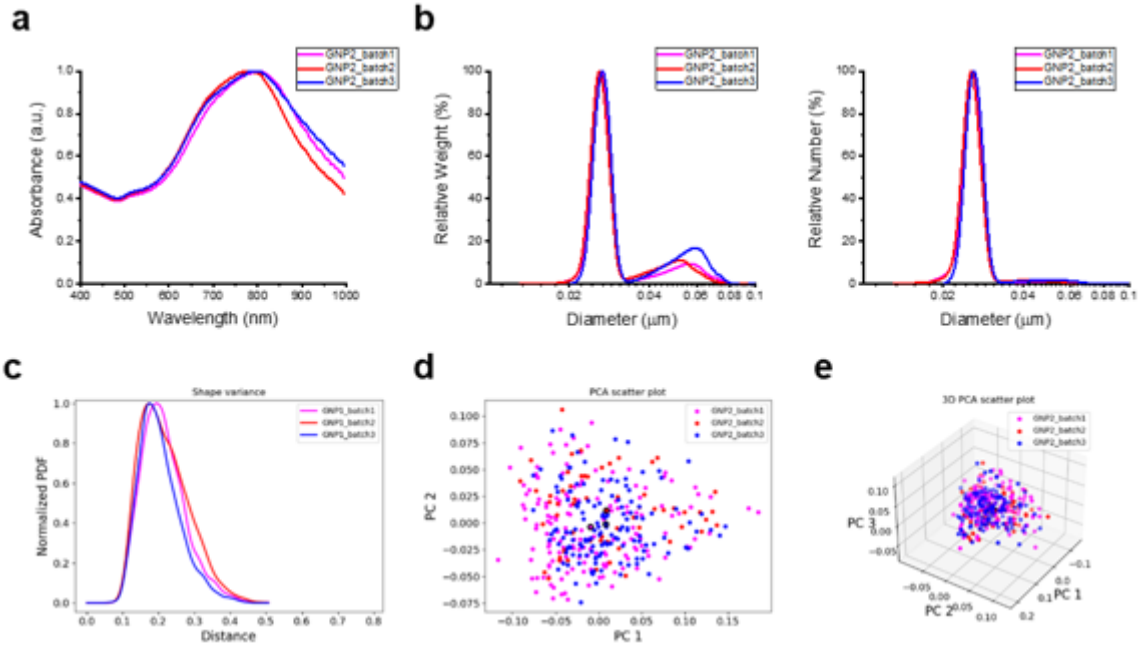


Figure S3. 3 Reproducibility between independent batches of GNP2.

(a) UV-vis absorption spectra. (b) DCS analysis by relative weight and number. (c) Intra-batch shape variance, (d) 2D scatter plot and (e) 3D scatter plot of the principal components from the contour's analysis.

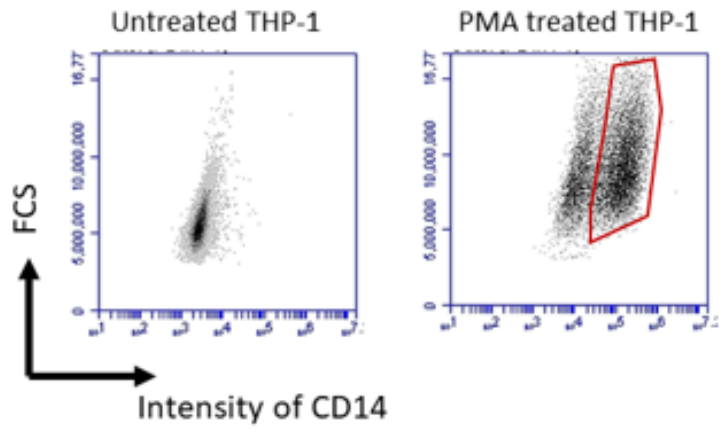


Figure S3. 4 The macrophage-like differentiation identified by flow cytometry for the surface marker CD14.

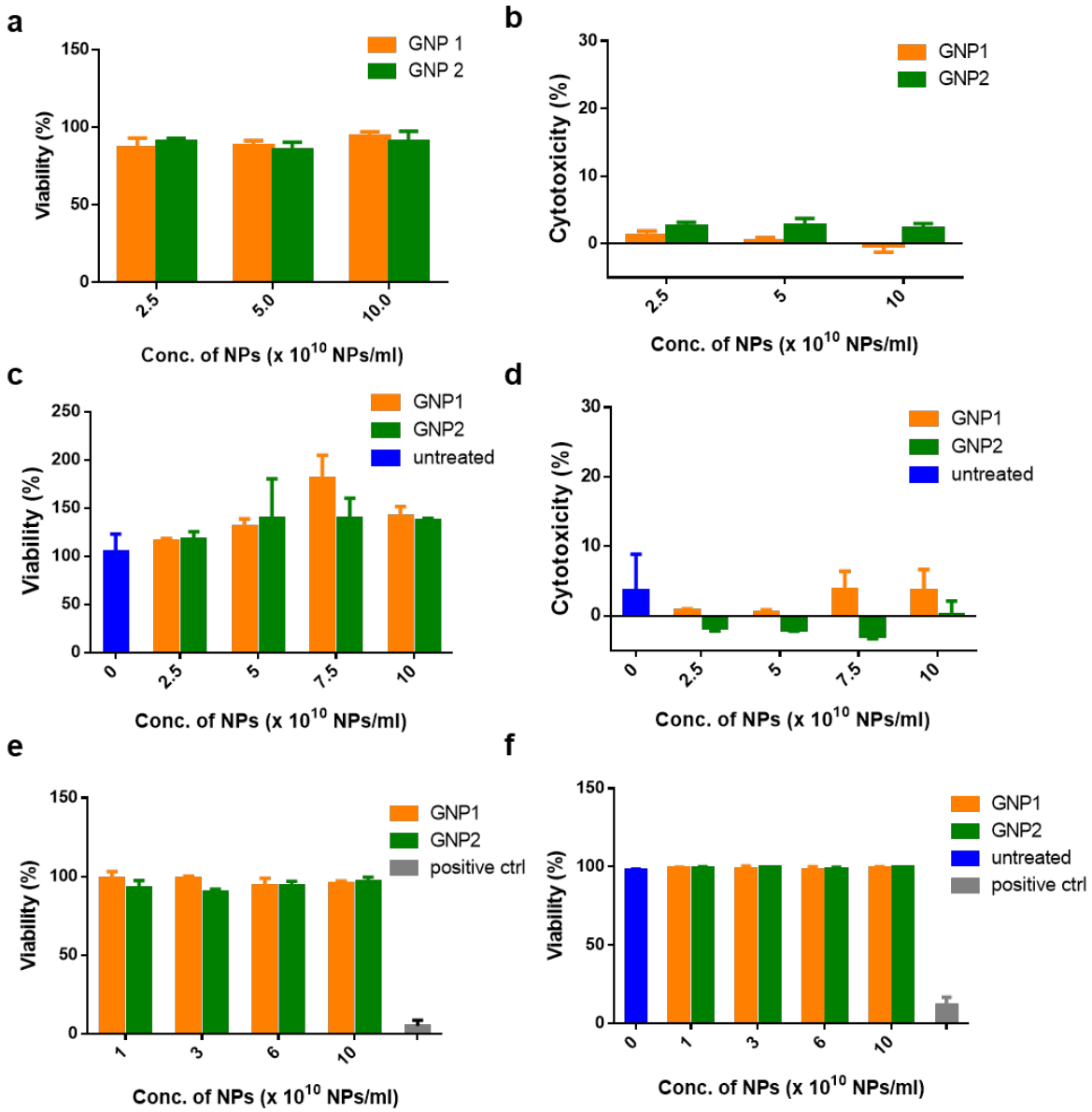


Figure S3. 5 cytotoxicity of GNP1 and GNP2 in different cell lines.

(a) MTS cell assay and (b) LDH-release assay for GNP1 and GNP2 cytotoxicity in dTHP-1 cell line. (c) MTS cell assay and (d) LDH-release assay in RAW 264.7 cell line. (e) TMRM assay and (f) Live/dead cell count assay for GNP1 and GNP2 cytotoxicity in A549 cell line.

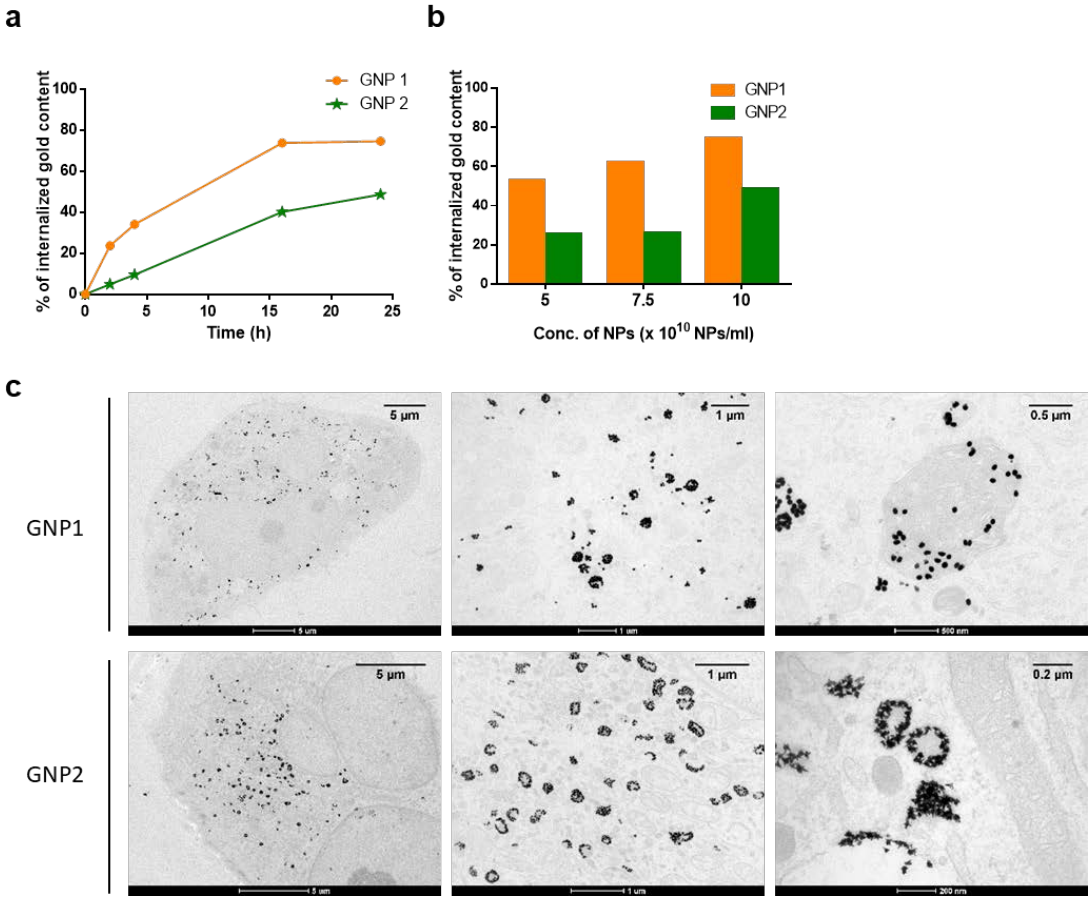


Figure S3. 6 Cellular uptake of GNP1 and GNP2 in dTHP-1 cell line.

(a) Kinetics of GNPs internalization. (b) Cellular uptake amount of GNPs with different concentrations. (c) Representative TEM images of the internalized GNPs in dTHP-1.

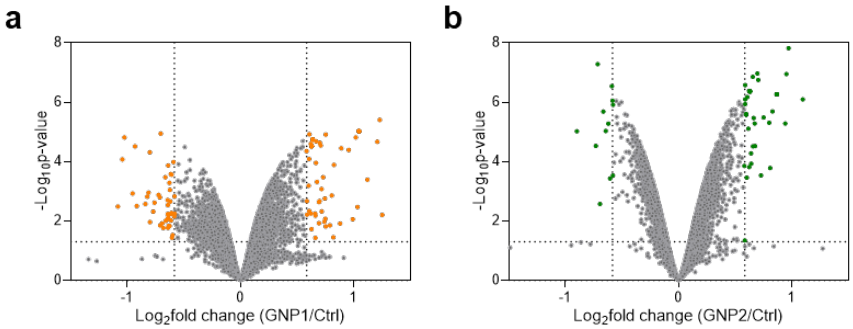


Figure S3. 7 Volcano plot of the DEGs from transcriptome after 4h treatment of GNPs in dTHP-1 cell line.

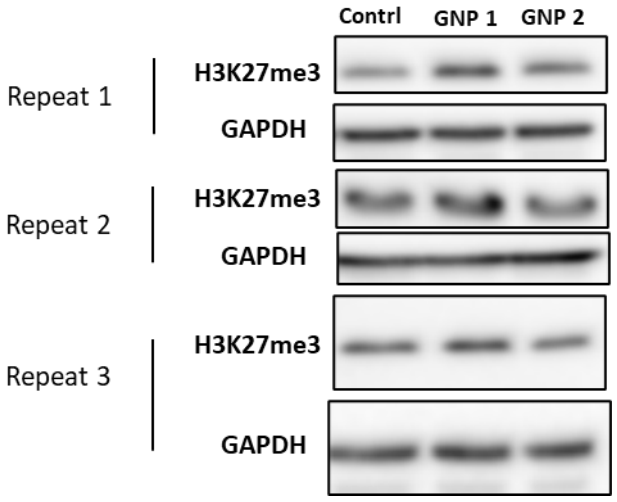


Figure S3. 8 Biological replicates for the WB of H3K27me3 in dTHP-1.

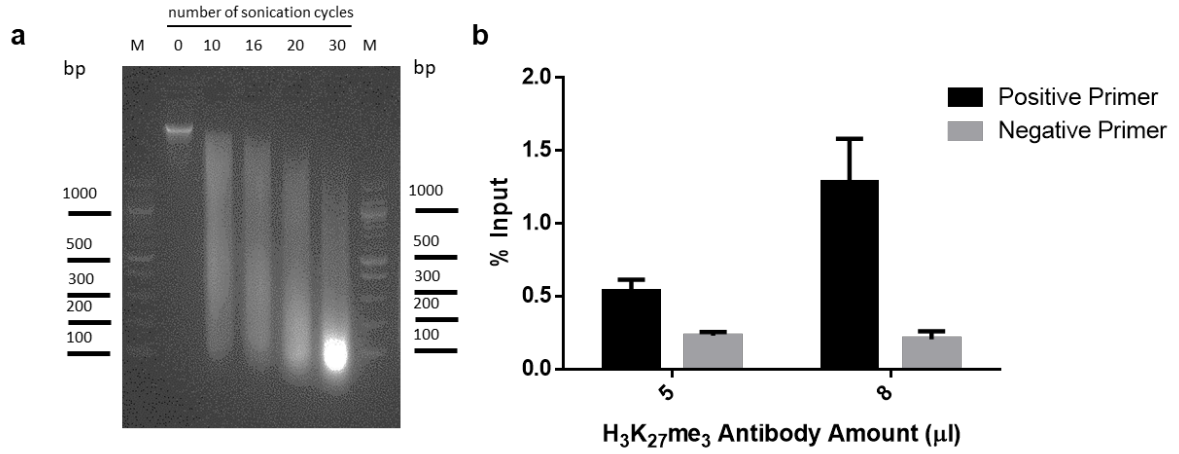


Figure S3. 9 Optimization of ChIP experiment.

(a) 2% agarose gel result for the fragment size of different sonication cycles showing 16 cycles of sonication provide a good degree of sonication. (b) ChIP-qPCR result of the antibody amount influence on ChIP yield, 8 μl antibody per 5 μg chromatin gives a better signal to background ratio.

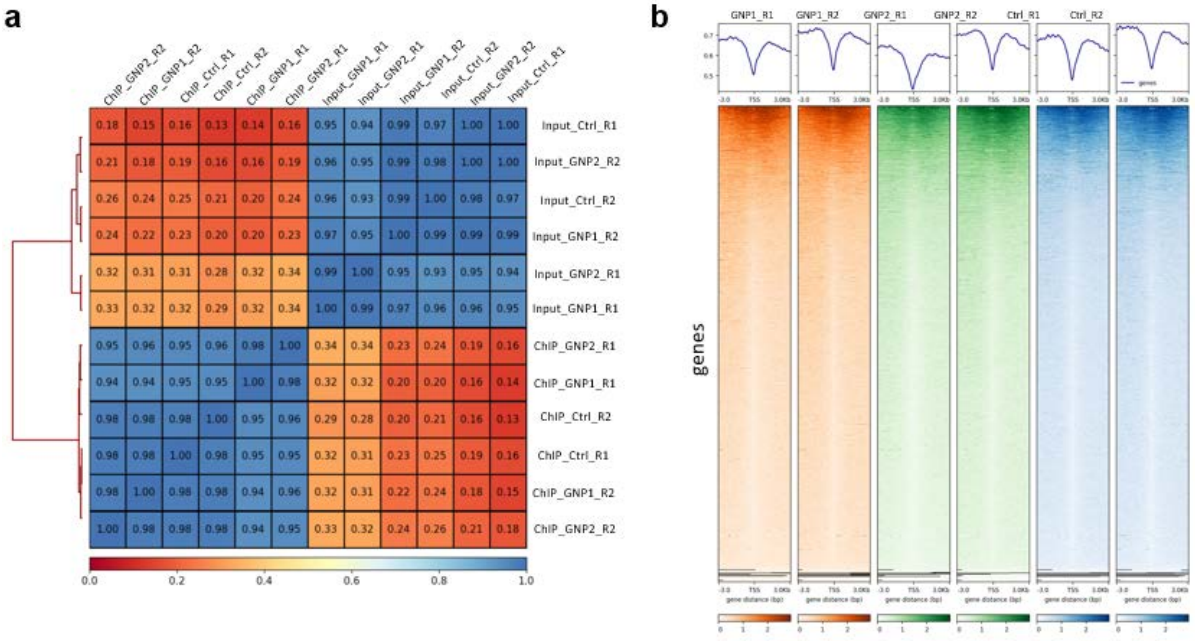


Figure S3. 10 Correlation heatmap plot (a) and profile heatmap plot (b) of ChIP-seq data with two biological replicates.

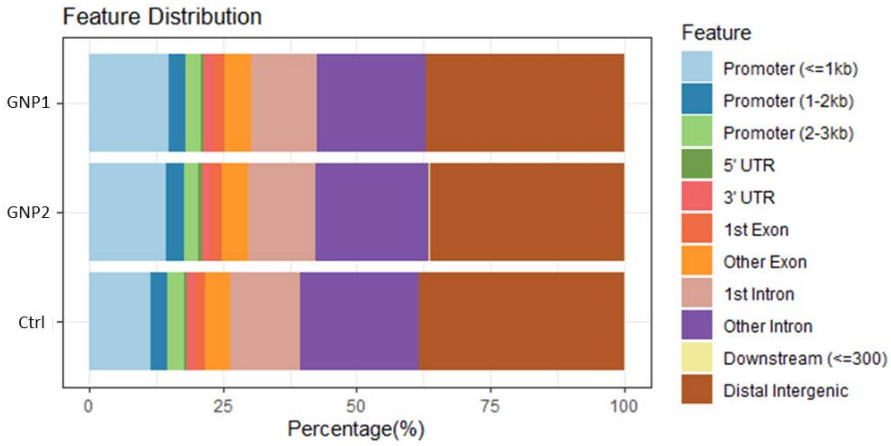


Figure S3. 11 Feature distribution of annotated ChIP-seq consensus peaks in genomic regions for different treatment groups.

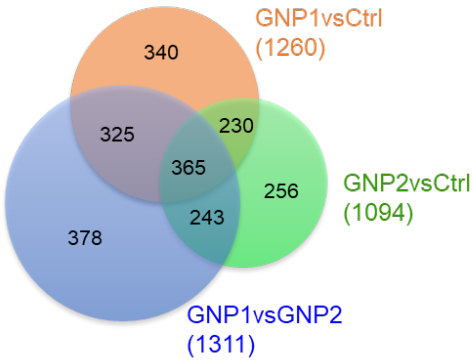


Figure S3. 12 Veen diagram for the ChIP-seq differential binding sites annotated genes.

	Hydrodynamic diameter (nm)	Polydispersity index (PDI)	Zeta potential in PBS (pH = 7.4) (mV)	Estimated surface area (nm^2)
GNP1	49 ± 9	0.26 ± 0.142	-38 ± 1	9402.5 ± 1250.5
GNP2	59 ± 2	0.30 ± 0.102	-34 ± 2	7628.7 ± 218.6

Table 3. 1 Hydrodynamic diameter, zeta potential and estimated surface area of GNP1 and GNP2.

modification	Ctrl_1	Ctrl_2	Ctrl_3	GNP1_1	GNP1_2	GNP1_3	GNP2_1	GNP2_2	GNP2_3	GNP1vs ctrl		GNP2 vs ctrl	
K4unmod	1620000000	1640000000	1640000000	573000000	806000000	874000000	1680000000	1610000000	1570000000	0.00064363	****	0.70542958	
K4me1	367000000	351000000	332000000	324000000	324000000	302000000	298000000	294000000	263000000	0.05593088		0.01228625	*
K4me2	23400000	34600000	51300000	11400000	43200000	35500000	90100000	84700000	82200000	0.63687366		0.00429373	***
K4me3	119000000	105000000	101000000	78500000	71600000	66300000	103000000	96100000	80400000	0.00509317	**	0.15370861	
K4Ac	1290000	1130000	718000	1650000	880000	733000	647000	554000	1120000	0.90603734		0.32760783	
K9unmod	2310000000	2010000000	1940000000	2570000000	2030000000	2110000000	1700000000	1710000000	1550000000	0.50087104		0.02548148	*
K9me1	898000000	863000000	847000000	951000000	819000000	863000000	997000000	739000000	672000000	0.85108696		0.54227873	
K9me2	8140000	7540000	7230000	5390000	5240000	6480000	15800000	5240000	5170000	0.01504021	*	0.77163358	
K9me3	609000000	561000000	517000000	699000000	687000000	702000000	740000000	547000000	435000000	0.00771 61	**	0.90617327	
K9unmod/K14Ac	483000000	979000000	822000000	119000000	532000000	986000000	897000000	404000000	739000000	0.59259353		0.71346988	
K9me2/K14Ac	112000000	101000000	960000000	718000000	711000000	688000000	172000000	830000000	793000000	0.00252503	**	0.11953039	
K9me3/K14Ac	31800000	29500000	28700000	43100000	33300000	34800000	23800000	21400000	20900000	0.09081594		0.00349323	***
K9Ac/K14Ac	107000000	97800000	100000000	111000000	98200000	88800000	77800000	74700000	72600000	0.76250179		0.00109328	***
K18/23unmod	5820000000	5600000000	5970000000	6350000000	5890000000	5780000000	5830000000	5150000000	5060000000	0.36354271		0.16564473	
K18me1/K23unmod	109000000	109000000	86100000	131000000	118000000	107000000	91300000	77900000	70800000	0.16878294		0.09273194	
K18unmod/K23Ac	2170000000	1960000000	1870000000	2240000000	1740000000	1590000000	2430000000	1790000000	1560000000	0.54265422		0.80293329	
K18/23Ac	189000000	158000000	149000000	190000000	135000000	115000000	245000000	138000000	110000000	0.50451842		0.98251223	
K27/K36/K37unmod	568000000	625000000	521000000	586000000	667000000	612000000	450000000	494000000	386000000	0.26008099		0.04211 8	*
K27me1	503000000	481000000	434000000	689000000	588000000	500000000	436000000	379000000	349000000	0.10924 18		0.06041156	
K36me1	86300000	64500000	78200000	107000000	48500000	58300000	64300000	73800000	58000000	0.80465986		0.23473487	
K27me2	1280000000	1190000000	1120000000	1450000000	1000000000	1160000000	987000000	924000000	890000000	0.96420106		0.00839598	**
K36me2	87500000	93400000	73800000	105000000	84800000	42200000	108000000	83200000	88900000	0.71636633		0.42245078	
K27me3	201000000	209000000	158000000	155000000	160000000	200000000	119000000	134000000	123000000	0.45343271		0.01771281	*
K36me3	127000000	122000000	122000000	181000000	114000000	126000000	106000000	80200000	85100000	0.46573669		0.01472307	*
K27me2/K36me1	849000000	842000000	689000000	1020000000	730000000	560000000	672000000	612000000	524000000	0.87919804		0.04784511	*
K27me1/K36me2	684000000	692000000	656000000	884000000	575000000	624000000	607000000	494000000	456000000	0.86872512		0.02741 15	*
K27me1/K36me1	170000000	208000000	206000000	166000000	152000000	137000000	414000000	366000000	309000000	0.04490608	*	0.00680146	**
K27me2/K36me2	109000000	101000000	98800000	103000000	105000000	92300000	90400000	92000000	92300000	0.60225657		0.02269598	*

K27Ac/K36unmod	7130000	7020000	7260000	7340000	7990000	8160000	7490000	7710000	6220000	0.05559111		0.99467541	
K79unmod	379000000	368000000	291000000	496000000	391000000	384000000	429000000	336000000	245000000	0.16366474		0.88372218	
K79me1	282000000	230000000	11400000	354000000	269000000	252000000	181000000	151000000	16500000	0.25692969		0.58064924	
K79me2	82600000	90400000	79700000	71000000	70700000	76100000	42200000	68700000	68600000	0.03313458	*	0.05992801	
K79me3	22300000	19000000	25700000	20500000	18500000	17500000	23700000	30200000	39100000	0.17501365		0.14940992	
K79Ac	41300000	30700000	14300000	63900000	14900000	45100000	18100000	11300000	30300000	0.4845884		0.40892696	

Table 3. 2 Significantly different motifs in untreated, GNP1 and GNP2 treated cells proteomics analysis. $p<0.05$, *; $p<0.01$, **; $p<0.005$, ***; $p<0.001$, ****.

References

1. Dawson, K. A.; Yan, Y., Current understanding of biological identity at the nanoscale and future prospects. *Nat Nanotechnol* **2021,** 16 (3), 229-242.
2. Walczyk, D.; Bombelli, F. B.; Monopoli, M. P.; Lynch, I.; Dawson, K. A., What the Cell "Sees" in Bionanoscience. *Journal of the American Chemical Society* **2010,** 132 (16), 5761-5768.
3. Murugan, K.; Choonara, Y. E.; Kumar, P.; Bijukumar, D.; du Toit, L. C.; Pillay, V., Parameters and characteristics governing cellular internalization and trans-barrier trafficking of nanostructures. *International journal of nanomedicine* **2015,** 10, 2191.
4. Duan, X.; Li, Y., Physicochemical characteristics of nanoparticles affect circulation, biodistribution, cellular internalization, and trafficking. *Small* **2013,** 9 (9-10), 1521-1532.
5. Mitchell, M. J.; Billingsley, M. M.; Haley, R. M.; Wechsler, M. E.; Peppas, N. A.; Langer, R., Engineering precision nanoparticles for drug delivery. *Nature Reviews Drug Discovery* **2021,** 20 (2), 101-124.
6. Monopoli, M. P.; Aberg, C.; Salvati, A.; Dawson, K. A., Biomolecular coronas provide the biological identity of nanosized materials. *Nat Nanotechnol* **2012,** 7 (12), 779-86.
7. Lynch, I.; Dawson, K. A., Protein-nanoparticle interactions. *Nano Today* **2008,** 3 (1-2), 40-47.
8. Waddington, C. H., *The epigenotype.* **1942.** (1464-3685 (Electronic)).
9. Bernstein, B. E.; Meissner, A.; Lander, E. S., The mammalian epigenome. *Cell* **2007,** 128 (4), 669-681.
10. Murr, R., Interplay between different epigenetic modifications and mechanisms. *Advances in genetics* **2010,** 70, 101-141.
11. Portela, A.; Esteller, M., Epigenetic modifications and human disease. *Nature biotechnology* **2010,** 28 (10), 1057-1068.
12. Handy, D. E.; Castro, R.; Loscalzo, J., Epigenetic modifications: basic mechanisms and role in cardiovascular disease. *Circulation* **2011,** 123 (19), 2145-2156.
13. Villeneuve, L. M.; Reddy, M. A.; Natarajan, R., Epigenetics: deciphering its role in diabetes and its chronic complications. *Clinical and Experimental Pharmacology and Physiology* **2011,** 38 (7), 451-459.
14. Jennifer, M.; Maciej, W., Nanoparticle Technology as a Double-Edged Sword: Cytotoxic, Genotoxic and Epigenetic Effects on Living Cells. *Journal of Biomaterials and Nanobiotechnology* **2013,** 04 (01), 53-63.
15. Pogribna, M.; Hammons, G., Epigenetic Effects of Nanomaterials and Nanoparticles. *J Nanobiotechnology* **2021,** 19 (1), 2.
16. Sierra, M. I.; Valdes, A.; Fernandez, A. F.; Torrecillas, R.; Fraga, M. F., The effect of exposure to nanoparticles and nanomaterials on the mammalian epigenome. *Int J Nanomedicine* **2016,** 11, 6297-6306.
17. Deng, R.; Shen, N.; Yang, Y.; Yu, H.; Xu, S.; Yang, Y. W.; Liu, S.; Meguellati, K.; Yan, F., Targeting epigenetic pathway with gold nanoparticles for acute myeloid leukemia therapy. *Biomaterials* **2018,** 167, 80-90.
18. Surapaneni, S. K.; Bashir, S.; Tikoo, K., Gold nanoparticles-induced cytotoxicity in triple negative breast cancer involves different epigenetic alterations depending upon the surface charge. *Sci Rep* **2018,** 8 (1), 12295.
19. Ng, C. T.; Dheen, S. T.; Yip, W. C.; Ong, C. N.; Bay, B. H.; Lanry Yung, L. Y., The induction of epigenetic regulation of PROS1 gene in lung fibroblasts by gold nanoparticles and implications for potential lung injury. *Biomaterials* **2011,** 32 (30), 7609-15.
20. Brzoska, K.; Gradzka, I.; Kruszewski, M., Silver, Gold, and Iron Oxide Nanoparticles Alter miRNA Expression but Do Not Affect DNA Methylation in HepG2 Cells. *Materials (Basel)* **2019,** 12 (7).

21. Boselli, L.; Lopez, H.; Zhang, W.; Cai, Q.; Giannone, V. A.; Li, J.; Moura, A.; de Araujo, J. M.; Cookman, J.; Castagnola, V.; Yan, Y.; Dawson, K. A., Classification and biological identity of complex nano shapes. *Communications Materials* **2020,** 1 (1).
22. Nomura, F.; Akashi, S.; Sakao, Y.; Sato, S.; Kawai, T.; Matsumoto, M.; Nakanishi, K.; Kimoto, M.; Miyake, K.; Takeda, K., Cutting edge: endotoxin tolerance in mouse peritoneal macrophages correlates with down-regulation of surface toll-like receptor 4 expression. *The Journal of Immunology* **2000,** 164 (7), 3476-3479.
23. Zhou, Y.; Zhou, B.; Pache, L.; Chang, M.; Khodabakhshi, A. H.; Tanaseichuk, O.; Benner, C.; Chanda, S. K., Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nature communications* **2019,** 10 (1), 1523-1523.
24. Allfrey, V. G.; Faulkner, R.; Mirsky, A., Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proceedings of the National Academy of Sciences of the United States of America* **1964,** 51 (5), 786.
25. Shechter, D.; Dormann, H. L.; Allis, C. D.; Hake, S. B., Extraction, purification and analysis of histones. *Nature protocols* **2007,** 2 (6), 1445-1457.
26. Schwartz, Y. B.; Pirrotta, V., A new world of Polycombs: unexpected partnerships and emerging functions. *Nature Reviews Genetics* **2013,** 14 (12), 853-864.
27. Di Croce, L.; Helin, K., Transcriptional regulation by Polycomb group proteins. *Nature structural & molecular biology* **2013,** 20 (10), 1147-1155.
28. Delest, A.; Sexton, T.; Cavalli, G., Polycomb: a paradigm for genome organization from one to three dimensions. *Current opinion in cell biology* **2012,** 24 (3), 405-414.
29. Margueron, R.; Reinberg, D., The Polycomb complex PRC2 and its mark in life. *Nature* **2011,** 469 (7330), 343-349.
30. Simon, J. A.; Kingston, R. E., Mechanisms of polycomb gene silencing: knowns and unknowns. *Nature reviews Molecular cell biology* **2009,** 10 (10), 697-708.
31. Cao, R.; Zhang, Y., The functions of E (Z)/EZH2-mediated methylation of lysine 27 in histone H3. *Current opinion in genetics & development* **2004,** 14 (2), 155-164.
32. Zheng, Q. F.; Wang, H. M.; Wang, Z. F.; Liu, J. Y.; Zhang, Q.; Zhang, L.; Lu, Y. H.; You, H.; Jin, G. H., Reprogramming of histone methylation controls the differentiation of monocytes into macrophages. *FEBS J* **2017,** 284 (9), 1309-1323.
33. Zhang, Y.; Liu, T.; Meyer, C. A.; Eeckhoute, J.; Johnson, D. S.; Bernstein, B. E.; Nusbaum, C.; Myers, R. M.; Brown, M.; Li, W.; Liu, X. S., Model-based Analysis of ChIP-Seq (MACS). *Genome Biology* **2008,** 9 (9), R137.
34. Luo, J.; Mitra, A.; Tian, F.; Chang, S.; Zhang, H.; Cui, K.; Yu, Y.; Zhao, K.; Song, J., Histone Methylation Analysis and Pathway Predictions in Chickens after MDV Infection. *PLOS ONE* **2012,** 7 (7), e41849.
35. Wei, G.; Wei, L.; Zhu, J.; Zang, C.; Hu-Li, J.; Yao, Z.; Cui, K.; Kanno, Y.; Roh, T.-Y.; Watford, W. T., Global mapping of H3K4me3 and H3K27me3 reveals specificity and plasticity in lineage fate determination of differentiating CD4+ T cells. *Immunity* **2009,** 30 (1), 155-167.
36. Cui, K.; Zang, C.; Roh, T.-Y.; Schones, D. E.; Childs, R. W.; Peng, W.; Zhao, K., Chromatin signatures in multipotent human hematopoietic stem cells indicate the fate of bivalent genes during differentiation. *Cell stem cell* **2009,** 4 (1), 80-93.
37. Talamini, L.; Violatto, M. B.; Cai, Q.; Monopoli, M. P.; Kantner, K.; Krpetić, Z. e.; Perez-Potti, A.; Cookman, J.; Garry, D.; Silveira, C. P., Influence of size and shape on the anatomical distribution of endotoxin-free gold nanoparticles. *ACS nano* **2017,** 11 (6), 5519-5529.
38. Feller, C.; Forné, I.; Imhof, A.; Becker, P. B., Global and specific responses of the histone acetylome to systematic perturbation. *Molecular cell* **2015,** 57 (3), 559-571.
39. Blecher-Gonen, R.; Barnett-Itzhaki, Z.; Jaitin, D.; Amann-Zalcenstein, D.; Lara-Astiaso, D.; Amit, I., High-throughput chromatin immunoprecipitation for genome-wide mapping of in vivo protein-DNA interactions and epigenomic states. *Nat Protoc* **2013,** 8 (3), 539-54.

40. Mukhopadhyay, A.; Deplancke, B.; Walhout, A. J.; Tissenbaum, H. A., Chromatin immunoprecipitation (ChIP) coupled to detection by quantitative real-time PCR to study transcription factor binding to DNA in Caenorhabditis elegans. *Nat Protoc* **2008,** 3 (4), 698-709.
41. Lee, T. I.; Johnstone, S. E.; Young, R. A., Chromatin immunoprecipitation and microarray-based analysis of protein location. *Nature Protocols* **2006,** 1 (2), 729-748.
42. Langmead, B.; Salzberg, S. L., Fast gapped-read alignment with Bowtie 2. *Nature Methods* **2012,** 9 (4), 357-359.
43. Ross-Innes, C. S.; Stark, R.; Teschendorff, A. E.; Holmes, K. A.; Ali, H. R.; Dunning, M. J.; Brown, G. D.; Gojis, O.; Ellis, I. O.; Green, A. R.; Ali, S.; Chin, S.-F.; Palmieri, C.; Caldas, C.; Carroll, J. S., Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* **2012,** 481 (7381), 389-393.
44. Stark R, B. G., DiffBind: differential binding analysis of ChIP-Seq peak data. **2011.**
45. Yu, G.; Wang, L.-G.; He, Q.-Y., ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **2015,** 31 (14), 2382-2383.
46. Yu, G.; Wang, L.-G.; Han, Y.; He, Q.-Y., clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **2012,** 16 (5), 284-287.

Chapter IV

Horizontal Cellular Transfer of Biomolecular Condensates

Laurent Adumeau [a,§], Mura McCafferty [a,§], Silvia Vercellino [a,b,§], Xiaoliang Yang [a,§], Wei Zhang [a,§], Cara Gaffney [a,b], Ying Ling Dee [a,b], Linlin Song [a], Yuchen Lin [a], Koen Evers [a], Lorenzo Cursi [a], Vanya Petseva [a], Zengchun Xie [a], Aisling Fleming [a], Emily Sheridan [a], Ingrid Morera [a], Yan Yan [a,b,], & Kenneth A. Dawson [a,*]*

a. Centre for BioNano Interactions, School of Chemistry, University College Dublin, Belfield, Dublin 4, Ireland

b. School of Biomolecular and Biomedical Science, UCD Conway Insti ute of Biomolecular and Biomedical Research, University College Dublin, Belfield, Dublin 4, Ireland

§ These authors contributed equally and are ranked alphabetically. They will be putting their name first on the citation in their CVs.

*Email: Kenneth.a.dawson@cbni.ucd.ie; yan.yan@cbni.ucd.ie

Abstract

Functional intracellular contact domains constitute a highly dynamic network of membrane-bounded and membrane-free organelles that mediate biomolecular transfer and communication[1-9]. Biological recognition on the nanoscale is governed by analogous domains of coordinated molecular interactions at apposed surfaces[10-13]. Our working hypothesis is that such domain interactions gate protective mechanisms and grant privileged intracellular and transcellular access, otherwise elusive, to certain nanostructures and their cargoes[14-17]. However, while the unravelling of these processes is key to therapeutic advances, including RNA medicines, the mechanisms involved are poorly understood. Here we show how nanostructures undergo cellular processing by a hierarchy of gated cellular sorting mechanisms that selectively block some components, while allowing highly efficient horizontal transfer of protein-RNA condensate complexes from one cell to another. Thus, synthetic nanostructure scaffolds undergoing contact interactions, primarily with endoplasmic reticulum and mitochondria, are re-exported as hybrid complexes containing protein-RNA condensates that are distinct from established concepts of “exosomes”. When extracellular, these particle-condensate complexes are solid-like and resist degradation. While the condensate is dominated by ribosomal and mitochondrial RNP granules, other components play significant roles in a remarkably coordinated program involving condensate melting, escape from the endo-lysosomal pathway, and intracellular redistribution. The programmed transfer of protein and RNA between cells is strikingly efficient, enabling the delivery of foreign RNA into recipient cells. In effect nanoscale granule condensates mediate highly efficient gated intercellular transfer of RNA assemblies with sub-organelle resolution, providing a mechanism, and potentially control over, intercellular communication and coordination of cell population status.

Introduction

We have long suspected that there exist highly protected nanostructure-mediated intercellular communication pathways allowing cytosolic access[12,15,17,18]. Still, the idea that secreted RNA nanostructures (for example 'exosomes') enable cell-to cell communication, even regulating tissue biology at distant sites (*e.g.*, in tumor metastases[19-21]), while evolving into a compelling narrative[22,23], is not fully settled. Challenges to observe, isolate and maintain (in functional form) sub-populations of extracellular nanoscale species often obscured the mechanisms and lead to divergent views on function, ranging from key regulators to debris released from the cell[24]. For instance, careful mechanistic studies have questioned the whole concept of exosome mediated intercellular transfer of RNA, suggesting instead that mRNAs pass along membrane nanotubes[25]. Indeed, recent measured reviews suggest that firm conclusions on the role of extracellular vesicles in RNA transfer remain elusive[26,27]. In the context of the present article, unorthodox ideas of "stable extracellular ribosomes" carrying RNA, including in the bloodstream are particularly intriguing[28-30]. In summary suppositions that endogenous nanostructure populations could selectively access nanoscale pathways, efficiently and safely transfer information against an overwhelming background ranging from cellular debris, to dust, to pathogens, whilst maintaining cellular messaging fidelity had limited mechanistic evidence.

Here we show that, in fact, endogenous cellular processing can resolve all of these challenges using stable extracellular condensates that are processed via a remarkable coordinated hierarchy of gated cellular sorting mechanisms. The evolution of these ideas has occurred over time, during which evidence has progressively grown of the existence of a highly selective but efficient endogenous nanoscale access pathway to the cytosol[15,17,18,31,32]. Thus, whilst most synthetic nanostructures traffic to lysosomes and are degraded there, we have long known that a very small proportion of heterogeneously coated ("biomolecular corona") particles are processed differently, leading them to traverse unconventional non-endo-lysosomal pathways, and subsequently be actively re-exported[17]. Now, further investigation of these rare (at most several percent) intracellular particle sub-populations has shown that re-export is preceded by sporadic clusters of collisions with intracellular targets including mitochondria (Figure 4.1A, blue, Mitotracker), endoplasmic reticulum (Figure 4.1B, blue, EGFP-Calnexin), stress granules (Figure 4.1C, blue, EGFP-G3BP1), and P bodies (Figure 4.1D, blue, EGFP-DDX6). Both the composition and architecture of the cell derived coating on exported particle complexes derived during this biogenesis process could therefore reflect those key intracellular contact interactions.

While in principle the release of these particle complexes into the extracellular milieu enables their capture, in practice (as with exosomes) typical methods of isolation lead to multi-component mixtures, in this case containing unprocessed particles, extracellular vesicles, cell-derived debris and particle coacervates with all of these components. We therefore designed fluorescently labelled silica coated particles with magnetic cores (physicochemical characterization see Figure S4.1-4.5) and used appropriately chosen particle pulse-chase regimes involving exposure (τ1), wash (τ2), and chase-collection (τ3) periods. Since individual events in uptake, intracellular processing and re-export are asynchronous, a compromise between duration of exposure pulses, washes and collection times was required to optimize the yield and limit the sub-population mixing (Figure 4.1E). Then, using focused magnetic field extraction we were able to isolate re-exported particle complexes with limited background contamination[33] (Figure S4.6). These complexes were found to be structurally stable allowing undegraded RNA to be extracted, even many hours later.

Results and Discussions

Structure and Composition of Extracellular Granule Complexes

Standard methodologies showed the particle complexes to have a solid-like cell-derived molecular coating overlaid on the remnant of the original core particle surface corona (Figure 4.1F-H, Figure S4.7-4.9). Treatment in various physiologically relevant conditions, including those typical of the late endocytic environment and other conditions, failed to detach the coating (Figure S4.10). Metabolic isotope labelling of the producer cells followed by mass spectrometry analysis (Figure S4.11) allowed the cell derived layers to be reliably distinguished from the residue of the original corona[34] (Figure 4.2A-E, Figure S4.12,4.13). Then, assigning these using the Cell Atlas[35] we determined that, while the identified proteins traffic between different locations, as a group they are most strongly correlated to endoplasmic reticulum and mitochondria. This is consistent with the coating having been derived from the contacts observed during biogenesis and suggests their role as key hubs in horizontal transfer (Figure 4.2A). Notably most of the proteins (around 70%) had previously been associated with mesoscopic intracellular RNP granules36 such as P and Q bodies (Figure 4.2B). Examples of well-known granule proteins (including LSM14A, SYNCRIP, FUS, several heterogeneous nuclear ribonucleoproteins) were found to be abundant (Figure 4.2B). Particle complexes produced by cells stably expressing EGFP fusion proteins (for example FUS and several heteronuclear proteins) carried the labelled proteins, and when these were studied with large scale particle-by-particle imaging it was observed that fusion proteins were incorporated into nearly all of the exported particle complexes, rather than only sub- populations. mRNA coding

the fusion protein was also detected. Analysis identified key protein complexes involved in control of RNA transport (*e.g.*, Emerin complex 52) and metabolism (*e.g.*, ribosomes and spliceosomes, Figure 4.2D).

We then extended the study to encompass different cell types, culture conditions, core particle surface modifications, including oriented grafted layers (Figure S4.14-4.16), and identified a conserved group of condensate-linked proteins (present across all particle complexes) that were both abundant and significantly enriched against the producer cell background (Figure 4.2E, Figure S4.17). Still, each complex was found to harbor different proteins, for instance, variation of the core particle surface presentation did lead to some differentiated encoding of the exported complexes. Around 73% of the 117 common proteins (indicated in the frame on heatmap) were identified in the RNA granule database. Notably, using the preparative procedures described here, commonly cited exosome markers were found to be absent, and cited exosome 'exclusion' biomarkers[37], were unambiguously present (Figure 4.2E). We hypothesized a link between some of these common proteins and overarching host-triggered biological functions reported later. For instance, the representation of chaperones and co-chaperones (6%) using manually curated human chaperone families[38] is significant because these have been previously implicated in the reversibility of granule condensation-melting, and maintenance of protein structure during functional insertion into target organelles[39].

RNA and DNA were extracted from particle complexes (Figure S4.18,4.19), and RNA was sequenced and analyzed using several different approaches. We identified a range of RNA types (including mRNA, miRNA, lncRNA, snRNA, rRNA, and tRNA) and significant coverage of the producer cell transcriptome (Figure S4.20), and mitochondrial RNA is significantly enriched (Figure 4.2F, G). The presence of cytoplasmic rRNA 28s, 5.8s, 5s and 18s, and mitochondrial rRNA 12s, 16s (mt-RNR1, mt-RNR2, Figure 4.2F, Figure S4.18) is consistent with our identification of both types of ribosomal protein complexes in the condensate (Figure 4.2D). Furthermore, small RNA sequencing identified the snRNAs U1(U11), U2(U12) and U4-U7 and that, combined with the identification of protein spliceosome complex (Figure 4.2D,G), suggested condensates also carry (major and minor) spliceosome complexes required for normal and alternate splicing. We also noted that, while in situ RNase footprinting led longer RNAs to be fully degraded, there remained RNA fragment lengths (around 20 nt) consistent with protected fragments familiar in ribosome profiling[40] (Figure S4.21). Indeed, various interaction network analyses involving key groups of proteins and RNA (Figure 4.2H, Figure S4.22) illustrates known and predicted interactions between the top proteins and most abundant mRNAs (Figure 4.2D, F). Our data therefore point towards the presence of an

extensive overlapping network of protein-protein, protein-RNA and RNA-RNA interactions, involving both rRNP and snRNPs and other complexes.

Taken together our data suggested that nanoscale biogenesis leads to particle-cell-derived biomolecular complexes possessing many properties of RNA granule condensates with some embedded functionalities in a stable extracellular form.

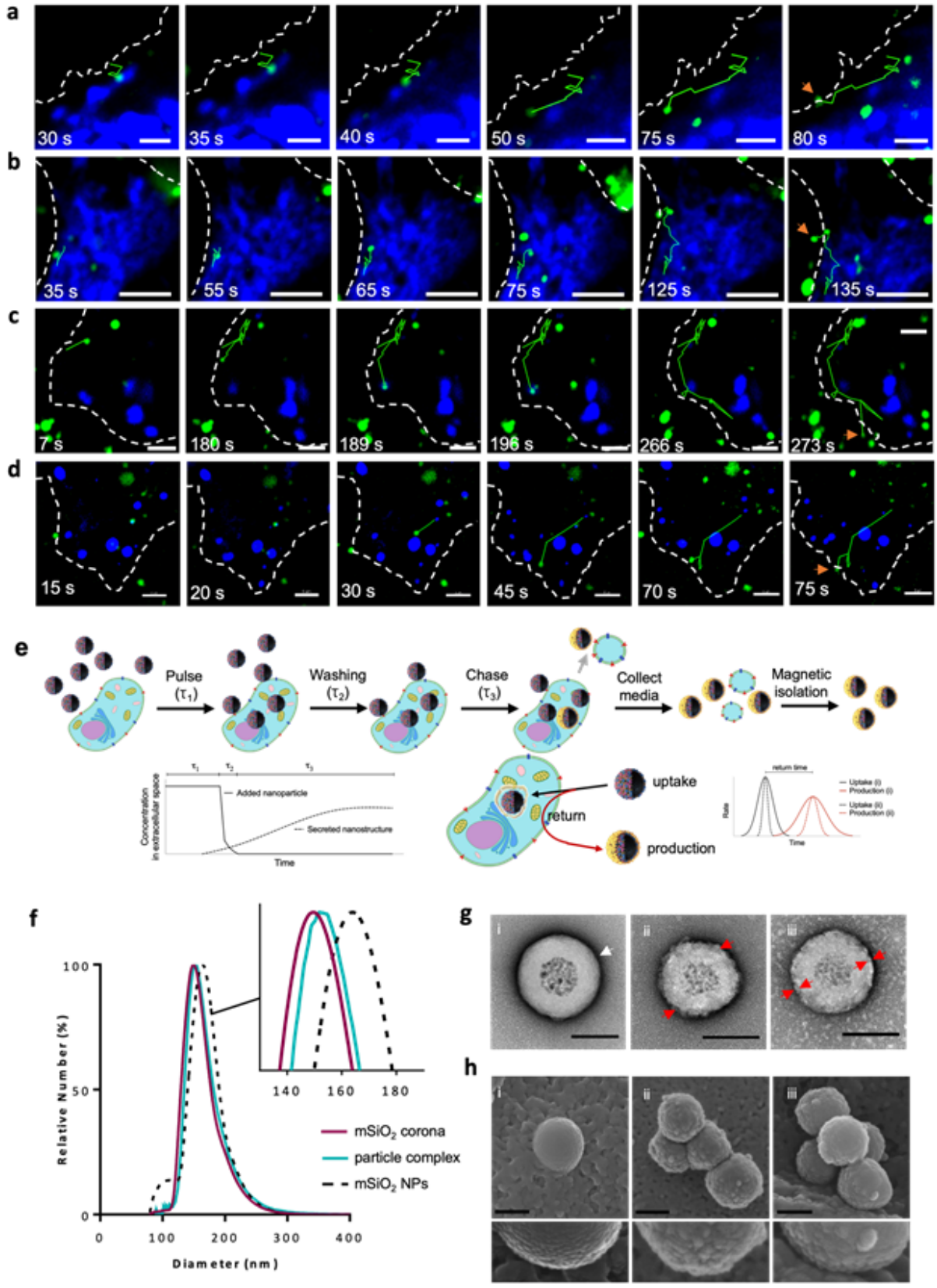


Figure 4. 1 Biogenesis, isolation and characterization of extracellular particle complexes.

(a-d) Trajectories of particles that are re-exported as particle complexes from HEK, captured by live cell confocal microscopy. These particles are 'directed' towards export but have prolonged pauses at contacts with several cellular structures (particularly mitochondria, ER and granules) where they may 'pick up' biomolecular assemblies. White dotted line: plasma membrane; Green: $mSiO_2$ NPs; Blue: Mitotracker (a), EGFP-Calnexin (b), EGFP-G3BP1 (c), and EGFP- DDX6 (d, colocalized with EGFP-DDX6 at 15s, 20s). The arrows identify the point at which the complex is released from the cell. Scale bar 5 µm. (e) Schematic showing production and isolation of particle complexes (not in scale). Contemporaneous particle uptake and release of complexes is limited when cells are exposed to NPs (*e.g.*, t1=10 min), rapidly washed (*e.g.*, t2= 3x1-5min), and complexes collected over limited times (*e.g.*, t3= 60-240min) using magnetic isolation. (f) DCS shows addition of a cell-derived biomolecular layer in complex compared to $mSiO_2$ and serum corona coated ($mSiO_2$ corona) particles. (g) Representative TEM micrographs of negatively stained bare $mSiO_2$ (i) $mSiO_2$ corona (ii), and complex (iii). The white arrow indicates the accumulation of uranyl acetate stain around the bare $mSiO_2$ particle; red arrows indicate ii) the formation of biomolecular corona on $mSiO_2$, and iii) coat on complex. Scale bar, 100nm. (h) Representative SEM images of $mSiO_2$ (i), $mSiO_2$ corona (ii), and complex (iii) deposited on cell surfaces. The smooth surface of $mSiO_2$ is distinguishable, but structural differences between $mSiO_2$ corona and complex are more subtle.

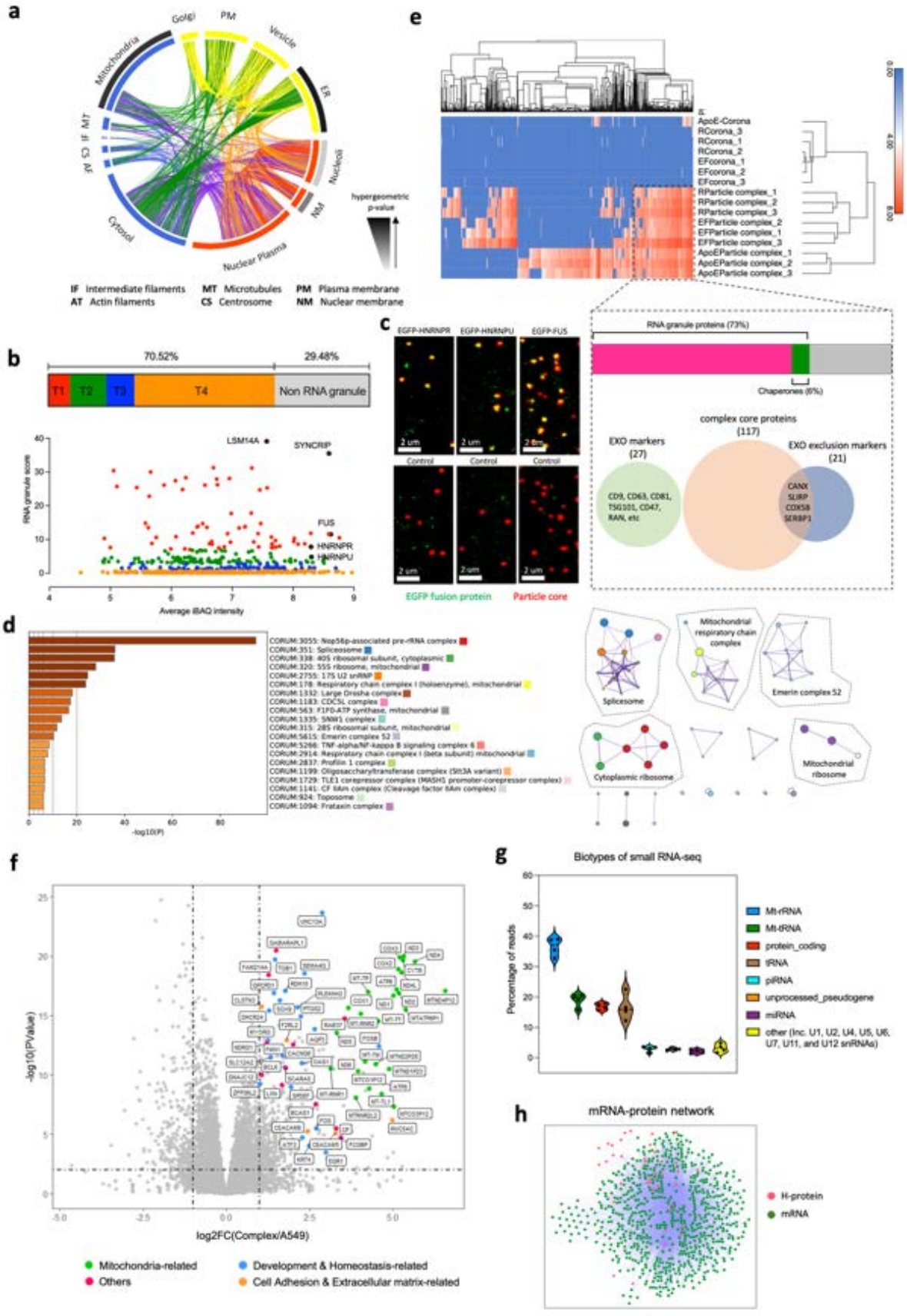


Figure 4. 2 Proteome and transcriptome of cell-derived component of particle complex.

Complexes have fingerprints linked to mitochondria and ER. Their overall composition suggests biomolecular granule condensates supporting ribosome, spliceosome and other complexes, that could become functional after intracellular melting and sorting. (a) Stable isotope labeling by amino acids in cell culture (SILAC) proteomics predicted subcellular locations from Human Protein Atlas. Proportion represented by inner color-coded segments; colored arcs identify multi-location proteins. Subcellular enrichment correlations (outer greyscale) segmented curves. (b) Proteomics RNA granule protein abundance vs. granule curation score. (c) Fluorescent confocal imaging of complexes generated from cells expressing EGFP-fusions with proteins that are abundant in complex's proteome illustrates population-wide presence of granular proteins. (d) CORUM enrichment analysis for the proteomics data identifies cytoplasmic and mitochondrial ribosomes, spliceosomes, and other mitochondrial and nuclear related complexes. (e) Proteomics heatmap of abundant proteins for three classes of complexes' proteome (produced from $mSiO_2$ corona, ApoE modified corona-NP, and exosome-free sera corona) compared to producing cells. Of proteins common to all complexes, 73% are in the RNA Granule Database and 6% in manually curated human chaperone families. Inset; some common 'exosome' markers are absent and known exosome exclusion markers detected. (f) Enriched mRNA in complexes compared to A549 cell transcriptome; dotted lines P-value = 0.01, $\log_2$FC(Complex/A549) = ±1. Of 64 abundant mRNAs (in color), many are mitochondria-associated. (g) Small RNA biotype and bioanalyser (Figure S4.18) are consistent with presence of ribosomes and spliceosomes. (h) Predicted mRNA-protein interaction networks between the top 50 H-proteins and abundant mRNAs (count >10,000). Blue edges, known interactions; green nodes, mRNAs; red nodes, H-proteins.

Horizontal Cellular Transfer of Biomolecular Complexes

While in situ such complexes would be generated and exchanged between cells, in the present study intermediate isolation and subsequent re-introduction into cell cultures allowed studies of various combinations of producer and recipient cell types, engineered cells and different conditions, and various chemical modifications. This revealed common highly elaborated functionalities of efficient energy-dependent cellular uptake and processing (Figure S4.23). Most striking, beginning at tens of minutes after recipient cell uptake, cell-derived condensate portions of the complex (fluorescently labelled) began to detach from the core particle over a brief interval of several tens of seconds (Figure 4.3A) and thereafter gain direct access to the cytosol. Certainly, in its extracellular (solid) form, separation of the residual serum corona-condensate interface would require disjoining forces far in excess of cellular interaction energies, so we studied the collective discontinuous nature of this intracellular detachment

process in some detail. While the final outcome for every particle complex was near complete detachment of particle core and condensate (green, particle core, red, condensate Figure 4.3B), often this occurred via multiple steps that could be categorized into recognizable stages.

Firstly, within the endocytic environment condensate melting was found to precede condensate detachment events, and time-resolved imaging showed that prior to, and at that point, the condensate behaved as a liquid (Figure 4.3C t=0-90s, 4.3D t=0-55s). This melting effect could not be recreated outside the cell (Figure S4.10). That suggested more elaborated endocytic prompts are required to trigger melting, possibly involving ATP-dependent chaperones and co-chaperones that have previously been associated with modulation of condensate melting[41].

The second stage (sorting) progressed via numerous repetitive 'contact' interactions involving fluidized condensate collisions with intracellular targets, notably endoplasmic reticulum, mitochondria and intracellular condensates. Some interactions were found to lead directly to complete transfer (Figure 4.3A, C) via the event sequence, polarization (Figure 4.3C frames t=0- 30s), rapid collective flow of the entire condensate across a narrow-elongated tunnel in the endocytic environment (frames t=90-195s), and deposition onto the target sites (Figure 4.3C, t=208s). The freed condensates remained associated with distinctive Rab5 peri-condensate organizations.

However, most interactions, while leading to multiple transient condensate polarizations within the original endocytic region, failed to immediately produce such complete condensate detachment. Instead, multiple melt polarization events eventually lead to a series of partial sorting events (Figure 4.3D, E). Intermediate events of partial detachment and transfer of condensate occurred, leaving the particle with a residual condensate coating that advanced to further steps and only finally (after numerous rearrangements) this culminated in complete condensate removal. Interesting condensate exchange events between formerly distinct complexes containing endocytic regions appeared to end in more effective detachment for one of the partners (Figure 4.3E). The requirement for multiple polarized fluid condensate collisions and condensate rearrangements before escape suggested a highly gated rate limiting cytosolic access process, possibly based on the assembly of appropriately mediated contact recognition domains between the endocytic fluid condensate and prospective target. Notably, the primary targets of successful detachment and deposition were observed to be endoplasmic reticulum (Figure 4.3F, blue EGFP-calnexin), mitochondria (Figure 4.3G, blue mitotracker), and RNA granules (Figure 4.3H, blue EGFP-DDX6), suggesting that for this type of horizontal transfer, endoplasmic reticulum and mitochondria, besides acting as hubs for ('pick up') biogenesis are also primary sites of ('drop off') deposition. The tendency for

detached (in distinction to particle associated) condensates to vanish during various types of cells permeabilization suggested a distinct intracellular environment for escaped condensates (Figure S4.24, 4.25).

Interactions with endogenous DDX6 positive intracellular condensates ('P bodies') were distinctive, typically transient, reversible and often short lived. Frequently events involved multiple partners in which the detached condensates transitioned between the particle, P body and target site before finally transferring irreversibly to the target. For instance, Figure 4.3h (blue EGFP-DDX6; grey mitotracker) illustrates how condensates were shuttled between particle, P body and mitochondria before final deposition on mitochondria. P body interactions with condensates had 'dwell times' of many tens of milliseconds and typically these lead to some small component being deposited on the P body, and a modified (now reduced) condensate fragment departing to another target for final deposition (Figure 4.3I, Images of Trajectory C, white arrow). The mesoscopic P bodies themselves were also observed to evolve significantly during these interactions (Figure 4.3J, other trajectories Figure S4.26), and some small particle derived condensate fragments became integrated into the core of P bodies, while other small (EGFP-DDX6 labelled) structures budded from the highly mobile liquid P body surface (Figure 4.3I, Trajectory C). Together such observations could suggest that P bodies act as a local storage and sorting reservoir in which components of the detached condensates are augmented or removed before the final deposition and dispersion of the particle condensate at a suitable target. While the underlying drivers of these P body interactions are not yet understood, previous studies of P body interactions and exchanges with microscopic granules raise the possibility that condensate rearrangement is linked to RNA status[42].

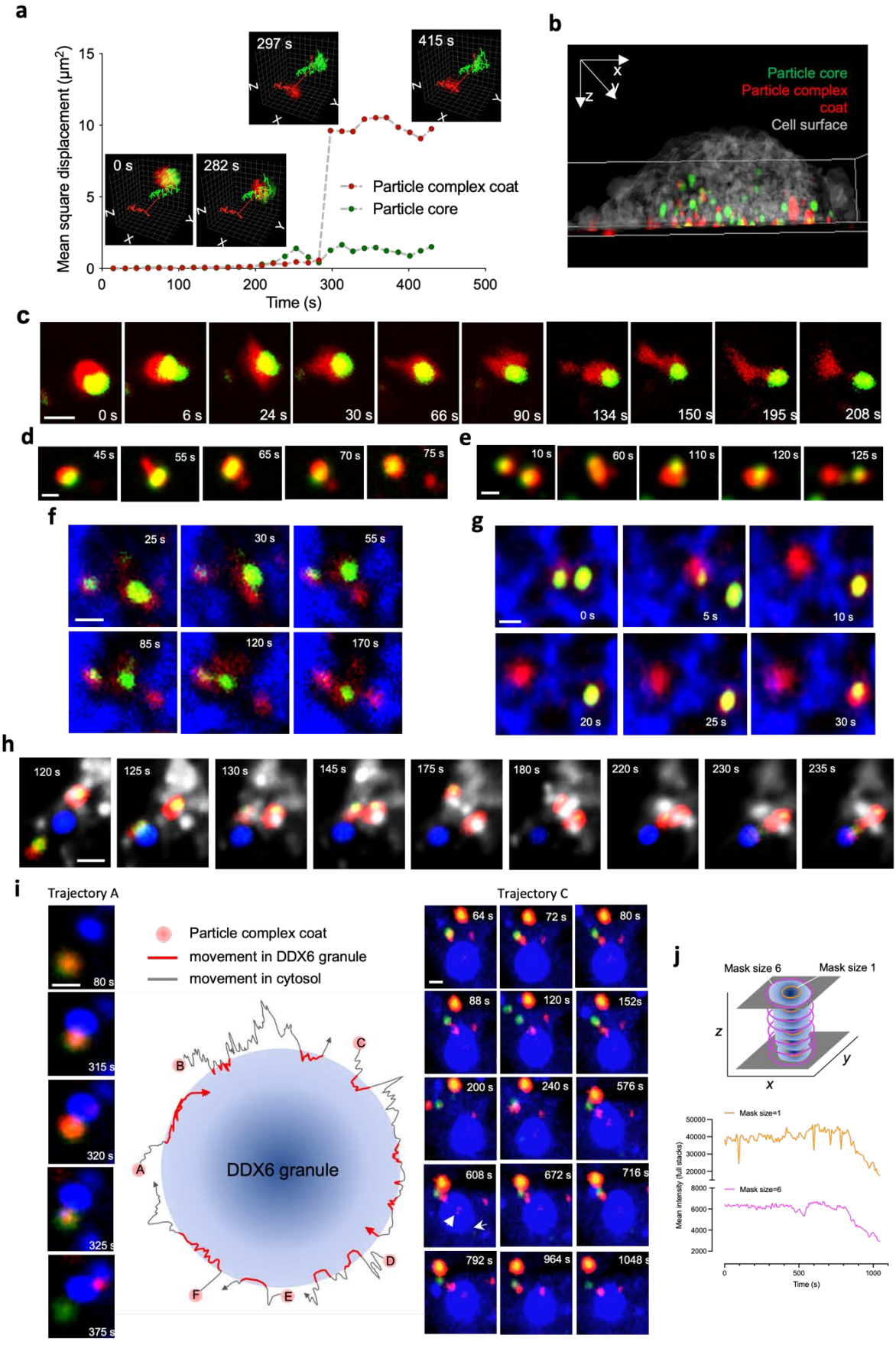

a
Mean square displacement (μm²)
Time (s)
297 s
415 s
0 s
282 s
Particle complex coat
Particle core
b
Particle core
Particle complex coat
Cell surface
c
0 s
6 s
24 s
30 s
66 s
90 s
134 s
150 s
195 s
208 s
d
45 s
55 s
65 s
70 s
75 s
e
10 s
60 s
110 s
120 s
125 s
f
25 s
30 s
55 s
85 s
120 s
170 s
g
0 s
5 s
10 s
20 s
25 s
30 s
h
120 s
125 s
130 s
145 s
175 s
180 s
220 s
230 s
235 s
i
Trajectory A
Particle complex coat
movement in DDX6 granule
movement in cytosol
DDX6 granule
Trajectory C
j
Mask size 6
Mask size 1
Mask size=1
Mask size=6
Mean intensity (full stacks)
Time (s)

Figure 4. 3 Detachment and sorting of particle complex coat.

After recipient cell uptake, coats melt, undergo fluid-like polarization, contact induced selective detachment from particle core and thereby escape endo-lysosomal pathway. Red: coat; green: core $mSiO_2$. Scale bar, 2 μm; transfers A549-A549 (a, b) and A549-HEK (c-j). (a) Discontinuous detachments typically conclude within several 100 seconds after initiation. (b) After 14 hours chase, 3D projection of remaining detached mesoscopic condensate coats, others having already dispersed. (c) opening of endocytic escape gateway, condensate flow and deposition on target organelle. (d) Sometimes only part of the coat detaches. (e) Coats from two complexes shared and sorted. Condensate deposition on ER (f, Blue: EGFP-calnexin) and mitochondria (g, Blue: mitotracker). (h) Deposition often involves multi-centered contacts, *e.g.*, Complex, P body, mitochondria (Blue, EGFP-DDX6; Grey, mitotracker). (i) Dynamic associations between P bodies and coat have multiple modes linked to the overall sorting and condensate distribution program. Trajectory is color coded reflecting condensate's approximate location (cytosol- grey, P body-red). Trajectories: A detachment onto the DDX6 granule; B, C, D repeated transient contacts; A, E, F prolonged contacts where P body itself rearranges. Trajectory C partial coating detachment, some reaching core of the granule (filled arrow 608s) and subsequent blebbing of to create new granules (unfilled arrow, 608s). (j) (Trajectory C) intensity of EGFP-DDX6 shows a dramatic decrease after 800s, masks of granule (mask size 6) and granule core (mask size 1) suggest molecular rearrangements reach the granule core.

Condensate Distribution in Recipient Cell

Condensate granule redistribution processes were found to be generic across different combinations of cell and particle complexes[43]. Post endocytic escape and deposition, detached particle condensates were observed to adopt a stable independent existence during which their movements could be imaged (Figure 4.4A, B) prior to dispersion. They were observed to traffic along different elements of the cytoskeleton (intermediate filaments, microtubules and microfilaments, Figure S4.27-4.29) or move along with tubules of endoplasmic reticulum (Figure S4.30, 4.31). Some condensates appear to undergo fusion or fission, others appear to move along the mitochondria and to be associated with the structural rearrangements (Figure S4.32). Overall morphological changes in the mitochondria caused by exposure to condensate complex were quantifiable by high content imaging (Figure S4.33). Some condensates associated with the perinuclear endoplasmic reticulum, at first appeared to pause (Figure 4.4B, t=100s) near the endoplasmic reticulum- nuclear membrane boundary and then precipitously dispersed within the nucleus (Figure 4.4B, t=170-295s). Quantitative analysis from high content imaging suggested that while the overall distribution between

mitochondria and nucleus reached stationarity relatively quickly (Figure 4.4C), some condensates could remain up to a single cell cycle (Figure 4.3B, 4.4D), during which local rearrangements continued. The distribution of condensates appeared to be more heterogeneous within the nucleus, than mitochondria. We observed that dispersion events appeared to be initiated at specific times and places, after which condensate granules vanished very rapidly possibly suggesting they may undergo a site-specific-triggered rate limiting processes.

While these concepts of actively gated redistribution are consistent overall with our study of condensate fate, some detailed biomolecular (protein and RNA) outcomes were found to be more specific to cell and complex type. Firstly, after uptake of labelled condensate complexes in recipient cells, distinct bands in western blots against the fluorescent dye label (AF405) were detected in the cytosol and nuclei up to one cell cycle, whereas (as expected) dye labelled corona proteins had been fully degraded44 (Figure 4.4E). We therefore studied the fate of isotope- labelled condensate protein complexes (derived from isotope labelled producer cells) in recipient cells. Mass-spectrometry analysis detected isotopically labelled proteins in recipient cells even up to one cell cycle, and Metascape analysis identified spliceosomes and DNA repair linked complexes in the nucleus and cytosolic ribosomes in the cytosol with high confidence, and abundance (Figure 4.4F, G, more analysis in Figure S4.34, S4.35). While conventional checks cannot fully exclude cross contamination or (at later stages, isotope reprocessing) there was evidence for some complexes (*e.g.*, cytosolic ribosomes) to be present in both cytosol and nucleus. This could arise from some of the transferred proteins being integrated into the known processing, assembly and maturation cycles operating between the cytosol and nucleus[45,46]. While these biodistribution studies will need to be extended in future, it is worth remarking that a component of fluorescently labelled condensate observed in the nucleus was punctuated and partially colocalized with Ki-67 protein (Figure S4.36, typical of nucleolar and Cajal body condensate structures) and that could be consistent with integration into Sm-like snRNPs.

The fate of (unlabeled and unperturbed) condensate RNAs passing between cells of the same (*e.g.*, A549 cell-derived particle complexes to A549 cells, Figure S4.37, S4.38) and different cell types (*e.g.*, A549 cell-derived particle complexes to HEK cells, Figure 4.4H, I) was also explored using time-resolved RNA sequencing and PCR in recipient cells. In general, we found that condensate mRNAs derived from a given cell type were quite rapidly eliminated (within the first two hours) in recipient cells of a different type (Figure 4.4H, I). While we had indeed expected that rapidly initiated degradative mechanisms would provide a (possibly RNA binding protein mediated) barrier to transfer between cell phenotypes, in fact transfer lead to some

more persistent changes in the transcriptome even after six hours (Figure 4.4H). Detailed time-resolved studies clarified those more nuanced outcomes (Figure 4.4I). For instance, of four genes (CECAM6, OAS1, MUC5AC, and KRT4) that are abundant in A549 cell-derived condensates but of low abundance in HEK cells, time-resolved RT-qPCR found that for two genes (MUC5AC, KRT4) the mRNA returned to their resting levels in HEK cells over several hours, while the other two (CECAM6, OAS1) at first decayed rapidly, but then remained somewhat elevated even after chases of twelve hours (Figure 4.4I).

Significantly, while mRNA derived from different cells was eliminated, we found that some fusions with endogenous mRNAs already detected in condensate complexes could be transferred between cells. As an example, Figures 4.4J, K illustrate how (FUS-EGFP mRNA containing) particle condensate complexes derived from stably expressing FUS-EFGP mRNA HEK cell lines allowed the transfer to naïve HEK recipient cells, where it remained undegraded over an entire cell cycle (Figure 4.4K). We note carefully that the details of these observations of RNA transfer and degradation kinetics, while they could be made highly reproducible, required careful preparative control, a matter that is likely relevant for future studies linked to RNA therapeutics (Figure S4.39).

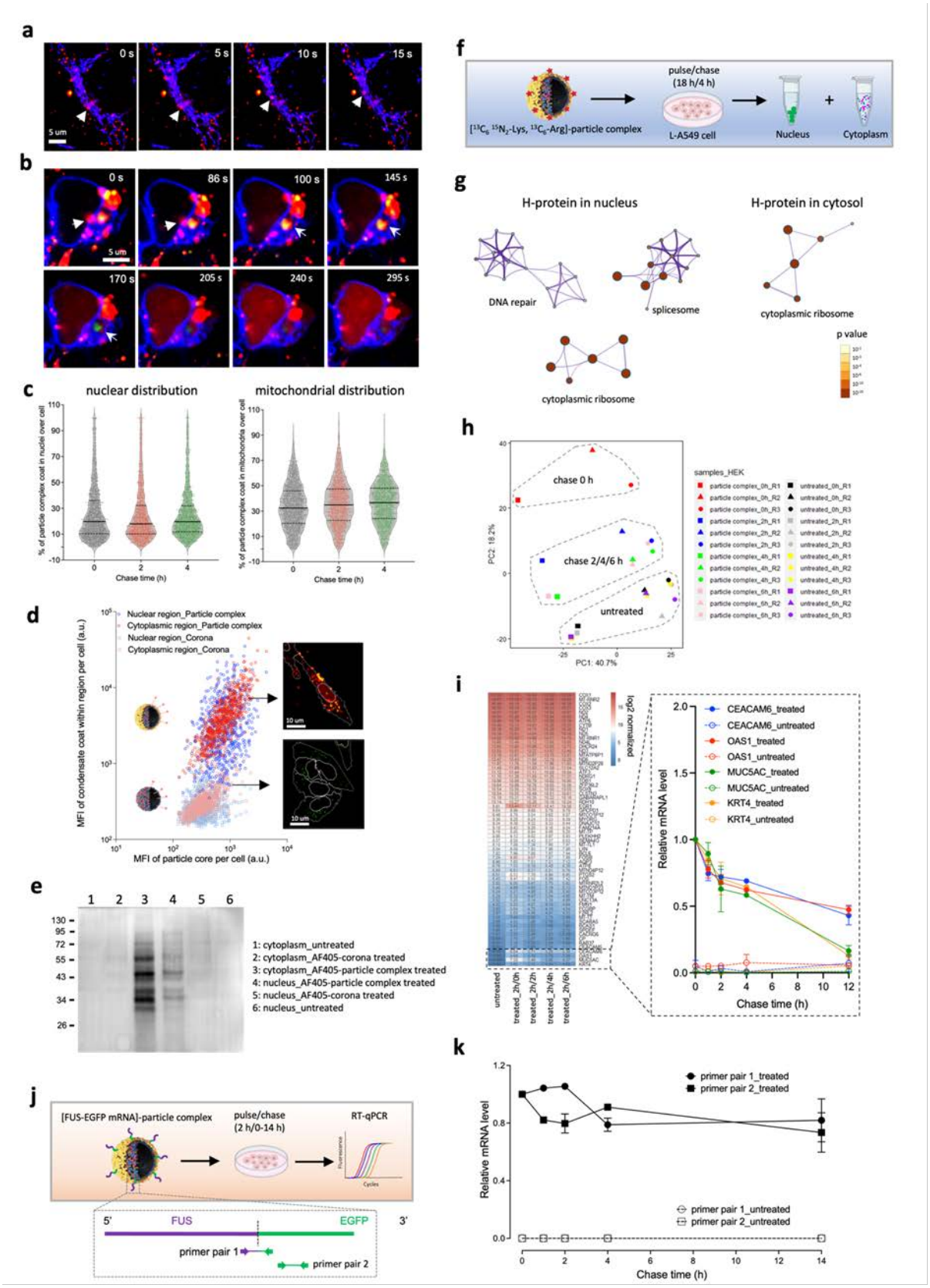


Figure 4. 4 Condensate coat deposition and subsequent biodistribution.

Confocal imaging shows after several (ca. four) hours of transient exchanges coats accumulate near mitochondria, ER and endogenous granules. Some suddenly disperse, others remain as mesoscopic punctuates that traffic further. Red, coat; Green, mSiO2. (a) condensates traffic along mitochondria (Blue, mitotracker). (b) detach at the ER reach nuclear envelope (closed arrows 0-86s, open arrow 100-170s), and suddenly disperse into the nucleus (Blue, EGFP-calnexin). (c) Single cell analysis of confocal images shows the heterogenous inter-organellar distribution (nuclei, 20%; mitochondria, 35%), becoming stationary after several hours. (d) Intracellular condensate vs. particle fluorescence is proportional for cytosol, but independent for nucleus after 18h exposure and 4h chase. That and a wider cell-to-cell spread suggest nuclear gating of condensate is independently regulated from uptake (insets: red, coat, green, particle core, white lines, masks). (e) Western blots (@AF405) from cells exposed to AF405-labelled complex for 18h pulse and 4h chased show long-lived intact coat proteins. (f, g) SILAC proteomics used to identify those coat proteins. Cytosolic ribosomes, spliceosome, and DNA repair complexes were enriched by CORUM. (h) PCA of transcriptome of HEK cells treated with A549 cell-derived particle complexes. Most mRNAs reached stationary after 2h chase, with only few mRNAs persist many hours. (i) Kinetics of enriched and abundant mRNA in the complexes confirmed by RT-qPCR. Groups of mRNA decay as distinct clustered classes. (j) Scheme for complexes carrying fusion mRNA time resolved fate via RT-qPCR. (k) RT-qPCR validation of mRNA coding EGFP-FUS protein in HEK cell-produced complexes transferred to normal HEK cells and retained.

Conclusion

We here report a paradigm in which stable solid like extracellular condensates melt on entering the target cell and then undergo a remarkable coordinated program of contact driven spatiotemporal gating and sorting mechanisms. This leads to highly regulated horizontal transfer between the cytosol and nucleus of one cell to another. Further study of the gated recognition-based access mechanisms in the present model system will now open the pathway to determine and catalogue (at molecular resolution) gateway domain interactions, ultimately enabling nanoscale architectural control of highly protected cellular processes[47,48]. Notably though, we also show that sophisticated gated sorting and other protective mechanisms involved in horizontal transfer can still be selectively subverted, and engineered messages transferred between cells.

Nevertheless, we believe the scope revealed by the current results is just a starting point. We note carefully that the present extracellular condensates transfer a rich biomolecular tapestry, including proteins, different types of nucleic acids. Networks of proteins, regulatory and coding

RNAs and other components found in virial infection and cancer metastases strategies remind us of the importance of system level cooperative forms of biological regulation that lead to enduring remodeling of the biological landscape. That is more difficult to achieve only via the transfer of several biomolecules. While the practical significance of these results for development of protein and RNA therapeutics is immediately apparent, we anticipate the present observations will catalyze and enable more far-reaching studies.

Methods

Synthesis of magnetic multicore fluorescent silica shell nanoparticles

The magnetic multicore encapsulated in the fluorescent silica shell were synthesized following a previously described procedure[33]. Briefly, magnetite nanoparticles were synthesized by coprecipitation and then coated with oleic acid to form stable dispersion in cyclohexane. Then, the magnetite nanoparticles were assembled to form multicores by evaporating an emulsion of the organic ferrofluid in an aqueous solution of sodium dodecyl sulfate prepared by sonication. The prepared multicores were consolidated by the formation of a silica shell following a sol-gel process in a basic hydro-alcoholic media using Tetraethyl orthosilicate as precursor of the silica. The fluorescent silica shell was then grown in similar conditions, by adding a conjugate of FITC with an aminosilane to incorporate the dye within the silica matrix. In order to obtain endotoxin free nanoparticles, the dispersion of particles was first sterilized by adding ethanol to reach ratio of 70% ethanol, 30% water (v/v), and the particles were then washed twice with LPS free water. Finally, under aseptic working conditions, using pyrogen free consumable and glassware, a dense layer of silica was grown at the surface of the particles in an aqueous dispersion in order to reduce the leakage of the dye over time in biological media due to silica dissolution.

The endotoxin level of all finished batches was measured with the Pierce LAL Chromogenic Endotoxin Quantitation Kit (cat. no. 88282), following manufacturer's instructions. Briefly, the dispersion of particles at 0.8 g/L, was incubated at 37 °C with the limulus amebocyte lysate (LAL), followed by incubation with the chromogenic substrate before reading the absorbance at 405 nm. Standards (between 0.1 and 1 EU/mL) as well as blanks containing the same nanoparticle concentration (to account for any interference), were measured at the same time.

It was initially confirmed that the presence of the nanoparticles in the test was not preventing the detection of the endotoxin by adding a known concentration of LPS with the nanoparticles.

Particle complex production

To collect the exported particle complexes, 6 x 10^6 cells were seeded into a T175 flask and incubated for 48 hours at 37˚C in a humidified atmosphere of 5% CO_2. Nanoparticles were incubated in complete media for 1 hour at 37˚C prior to exposure to cells at 800 ug/ml of $mSiO_2$. Cells were incubated with 10 ml of the nanoparticle suspension for 10 mins; referred to as the pulse time. The suspension was removed, and cells were washed twice with complete media and twice with PBS to remove non-internalized particles, before 20 ml of complete media containing 10% (v/v) FBS that was previously centrifuged at 20,000 g for 30 mins was added to the cells and incubated for 4 hours; this is referred to as the cha e period. After the 4- hour chase, the cell supernatant was collected, and cell debris was removed by centrifugation at 700 x g for 5 minutes. The exported particle complexes were concentrated from the cell supernatant by centrifugation at 4°C at 12,000 x g, 40 minutes. The pellets of particle complexes were resuspended in PBS.

Magnetic purification

For the particle complexes produced with $mSiO_2$ NPs, they were isolated from other cell products (such as multivesicular bodies, apoptotic bodies, exosomes *etc.*) by magnetic extraction using our in-house designed magnetic rack (Supplementary Figure S4.4). Particle complexes were quantified by measurement of the core particle fluorescent signal against a standard curve of particles.

Live cell imaging of particle complex biogenesis

Cell were treated with 800 ug/ml of $mSiO_2$ corona-NPs for 10 minutes then washed twice with complete media and twice with PBS. The cells were either then stained with CellMask™ Orange (5 ug/ml) for 20 minutes, or with MitoTracker™ Red CMXRos (500 nM) for 30 minutes, or with CellTrace™ far red (2 µM) for 20 minutes. The dye was removed, and the cells then washed twice in PBS. The cells were incubated in fresh phenol-free MEM and imaged with an Opera Phenix™ HCS in live cell mode. The imaging chamber was pre warmed at 37° C, the CO2 concentration was 5 %, and the acquisition frequency was 0.33 frames/second.

Fluorescent dye labelling proteins in particle complexes

Particle complexes and corona-NP controls were incubated with Alexa Fluor™ 405 NHS Ester (Succinimidyl Ester) or Alexa Fluor™ 568 NHS Ester (Succinimidyl Ester) (Thermo Fisher Scientific) at a w/w ratio of 1:4 dye:corona-NP or particle complex, for 1 h at room temperature with gentle shaking. To quench any remaining unreacted dye, glycine was added at a w/w ratio of 1:2 glycine:dye and incubated for 30 mins at room temperature with gentle shaking. The stained mSiO2 corona and particle complexes were collected from the dye solution by

centrifugation and resuspended in PBS. They were then washed 3 times in PBS by centrifugation to remove unbound dye.

Live cell imaging of detachment of particle complex coat from nanoparticle core

Particle complexes were stained using an Alexa Fluor™ NHS Es er (Succinimidyl Ester) dye. Cells were incubated with 40 ug/ml of stained Particle complexes for 2 hours, then washed twice in PBS and twice in fresh media to remove non-internalized particles. The cells were incubated in fresh phenol-free media and after 1 hour of chase time, timelapses were acquired with a Nikon Eclipse TI spinning disc confocal microscope equipped with a 100x oil immersion objective. The chamber was pre warmed at 37°C and the CO2 concentration was set at 5 %. The frequency of acquisition was 0.1 frame/s. Time-lapses were analyzed with Imaris 1.7.6 imaging software (Bitplane), using to the spots tracking tool.

High content image analysis (HCA)

HCA experiments were performed using an Opera Phenix® High Content Screening System (PerkinElmner), equipped with 4 lasers (405, 488, 561, 647 nm) and several objectives, including a water immersion 40 x and 63 x that were used in this work. The image sets were analyzed with the Harmony® high-content analysis software provided by the same company. The analysis pipelines were set according to the experiment goal, with each pipeline composed of several building blocks to define the cell compartments of interest and markers used for the staining.

Stable isotope labelling by amino acids

A549 cells were isotopically labelled using an established stable isotope labelling by amino acids in cell culture (SILAC) protocol[49]. Briefly, cells were incubated with DMEM for SILAC with 10% (v/v) dialyzed FBS, 0.8 mM L-Arg 13C6, and 0.4 mM L-Lys 13C6 15N2). The cells were passaged every 2-3 days and after several passages, the incorporation of the amino acids was checked by LC-MS. Isotope labelled Particle complexes (heavy labelled-particle complexes) were produced from labelled A549s using the previously described protocol.

Sample preparation of particle complexes for mass spectrometry

Particle complexes were washed once in PBS by centrifugation before the pellet was resuspended in 8 M urea (in 50 mM NH_4HCO_3, 1 mM EDTA plus protease inhibitors) with 2 cycles of 30s sonication in an ultrasonic bath then ProteaseMAX™ Surfactant was added to a final concentration of 0.05%. Dithiothreitol (DTT) was added to a final concentration of 10 mM and samples were incubated for 1 hour at room temperature. Iodoacetamide (IAA) was added to a final concentration of 40 mM and samples were incubated for 1 hour at room

temperature, protected from light. DTT was added to a final concentration of 40 mM to consume excess IAA and samples were incubated for 30 mins at room temperature. Lysyl Endopeptidase (Lys-C) was added at a w/w enzyme to substrate ratio of 1:100 and samples were incubated for 4 hours at 37°C with mild shaking. Samples were diluted with 3 x volumes of 50 mM NH_4HCO_3 to reduce the urea concentration to < 2M. Trypsin was added at a w/w enzyme to substrate ratio of 1:100 and samples w re incubated for 18 hours at 37°C with mild shaking. Trifluoroacetic acid (TFA) was added to a final concentration of 0.1% (v/v) and the samples were centrifuged at 16,000 g for 10 mins to remove any particles or precipitates. Peptide suspensions were then concentrated and desalted using Pierce™ C18 Tips following manufacturer's instructions and then resuspended in 2% acetonitrile + 0.15% TFA in HPLC grade water. Peptide concentration was quantified via nanodrop to normalize samples for mass spectrometry analysis.

Label-free MS method for particle complexes

Digested peptide samples (Regular Particle complex, Exo-Free Particle complex and ApoE Partic complex, and corresponding corona) were analyzed on a Dionex™ Ultimate 4000 HPLC system (Thermo Fisher Scientific) coupled to an Orbitrap Fusion™ Tribrid™ Mass Spectrometer (Thermo Fisher Scientific). The column employed was a nanoViper Acclaim PepMap 100 column (Thermo Fisher Scientific; id 75 µm x 50 cm, 3 µm particle size and 100 Å pore size). Peptides were eluted over 60 minutes using a linear A/B solvent gradient (solvent A = 100% water, 0.1% TFA, solvent B = 20% water, 80% ACN, 0.08% TFA) and each run was performed with 0% to 50% solvent B used over the first 45 mins followed by 50% - 90% solvent B for 5 min. Peptides were then analyzed in data-dependent mode. For MS, the resolution was 120k, MS scan was 380-1500 m/z, AGC (Automatic Gain Control) target was 4 × 105 and the injection time was 50 ms. Higher-energy collision dissociation (HCD) was employed using 28% collision energy, and precursor ions were accumulated using an isolation width of 1.6 m/z. Charge states from 2 to 7 were included for MS/MS. Dynamic exclusion was set to 60 s. For MS/MS, resolution was 30k, AGC (Automatic Gain Control) target was 5 × 104 and an injection time was 300 ms.

MS method for heavy labelled particle complexes

Digested peptide samples were analyzed on a Dionex™ Ultimate 4000 HPLC syste (Thermo Fisher Scientific) coupled to an Orbitrap Fusion™ Tribrid™ Mass Spec rometer (Thermo Fisher Scientific). The column employed was a nanoViper Acclaim PepMap 100 column (Thermo Fisher Scientific; id 75 µm x 50 cm, 3 µm particle size and 100 pore size). Peptides of 'Heavying labelling Particle Complex' and corresponding 'Corona' were eluted over 60

minutes using a linear A/B solvent gradient where Solvent A=100% water, 0.1% TFA, Solvent B=20% water, 80% ACN, 0.08% TFA; each run was performed with 0% to 50% solvent B used over the first 45 mins followed by 50% - 90% B for 5 min. Peptides of 'Heavy labelled Particle Complex Treated Cytoplasm' and 'Heavy labelled Particle Complex Treated Nuclear' were then analyzed in data-dependent mode. Peptides were eluted over 120 minutes using a linear A/B solvent gradient (solve A = 100% water, 0.1% TFA and Solvent B = 100% ACN, 0.08% TFA) and each run was performed with 2% to 27.5% solvent B used over the first 100 mins followed by 27.5% - 55% solvent B for 10 min and by 55% - 90% for 5 min. For MS, the resolution was 120k, MS scan w s 380-1500 m/z, AGC target was 4 × 105 and an injection time was 50 ms. HCD was employed using 28% collision energy, and precursor ions were accumulated using a isolation width of 1.6 m/z. Charge states from 2 to 7 were included for MS/MS. Dynamic exclusion was set to 60 s. The MS/MS was conducted with a rapid 'Ion Trap Scan Rate', an AGC target of 2 × 104 ions with an injection time of 35 ms.

Nuclear and cytosolic fractionation

The nuclei of cells were extracted for following the protocol of Hengstschläger *et al*[50]. The untreated, corona and particle complex treated cells were washed once with PBS, collected by trypsinization and centrifuged at 1,500 RPM for 3 mins to pellet, then washed once with PBS. The cell pellet was then resuspended in 5 pellet volumes (250 µl) of cytoplasmic extraction buffer (10 mM HEPES, pH 7.4, 2 mM MgCl2, 120 mM sucrose and cOmplete™ Protease Inhibitor) and incubated at room temperature for 2 minutes, then on ice for 10 minutes to induce hypotonic swelling. To break the cell membrane, 25 µl of 10% (v/v) Nonident P-40 was added to a final concentration of 1% (v/v) and homogenized by repeat pipetting. Nuclei were separated from the cytoplasmic extract by centrifugation at 500 g for 3 mins at 4°C. The supernatant (cytoplasmic extract) was collected and stored at -80°C for further analysis. The pellet (containing nuclei) was washed by resuspension in 10x pellet volume of nuclei wash buffer (10 mM HEPES, pH 7.4, 5 mM MgCl2, 320 mM Sucrose, 1 % (v/v) Nonident P-40 and cOmplete™ Protease Inhibitor). The washed nuclei were collected by centrifugation at 500 g for 3 mins at 4°C, and then soluble proteins were extracted from isolated nuclei by resuspension in 5x pellet volume of freshly prepared nuclear extraction buffer (20 mM HEPES, pH 7.9, 1.5 mM MgCl2, 0.42 M NaCl, 25% (v/v) Glycerol, 1 mM DTT fresh) and cOmplete™ Protease Inhibitor. Nuclei were vortexed then incubated at 4°C for 30 mi while shaking. Soluble nuclear proteins were separated from the non-soluble fraction by centrifugation at 20,000 g for 20 mins at 4°C. The supernatant was collected and stored at 80°C for further analysis.

Preparation of nuclear and cytosolic proteins for mass spectrometry analysis

For both cytosolic and nuclear extracts, 100 µg of protein was prepared for mass spectrometry analysis using a filter-aided sample preparation (FASP) protocol by Coleman *et al*[51]. Peptide concentration was quantified via nanodrop to normalize samples for mass spectrometry analysis.

MS method for heavy labelled particle complex treated cells

Digested peptide samples were analyzed on a Dionex™ Ultimate 3000 HPLC system (Thermo Fisher Scientific) coupled to a quadrupole Orbitrap mass spectrometer (QExactive, Thermo Scientific). The column employed was a C18 home-made column (Thermo Fisher Scientific; id 75 µm x 15 cm, 3 µm particle size and 100 Å pore size). Peptides of 'Heavy labelled Particle Complex' were eluted over 60 minutes using a linear A/B solvent gradient where Solvent A=2.5% ACN, 0.5% Acetic Acid, Solvent B=97.5% CAN, 0.5% Acetic Acid; each run was performed with 2% to 25% solvent B used over 58 min. Peptides of 'Heavy labelled Particle Complex Treated Cytoplasm' and 'Heavy labelled Particle Complex Treated Nuclear' were eluted over 120 minutes using a linear A/B solvent gradient where Solvent A=2.5% ACN, 0.5% Acetic Acid, Solvent B=97.5% CAN, 0.5% Acetic Acid; each run was performed with 2% to 28% solvent B used over 120 min. Peptides were then analyzed in data-dependent mode using a top 12 method. For MS, resolution was 70k, MS scan was 350-1600 m/z, AGC target was 3 × 106 and an injection time was 60 ms. HCD was employed using 27% collision energy, and precursor ions were accumulated using an isolation width of 2 m/z. Charge states from 2 to 7 were included for MS/MS. Dynamic exclusion was set to 30 s. For MS/MS, resolution was 17.5k, AGC (Automatic Gain Control) target was 5 × 104 and an injection time was 250 ms.

LC-MS data analysis

The data was analyzed using MaxQuant software (version 1.6.5.3) via built-in Andromeda search engine. For the datasets from Label-free samples, the initial maximum allowed mass deviation was set to 10 ppm for precursor ions and 0.02 Da for MS/MS peaks. For the datasets from SILAC samples which performed by DCU facility, the initial maximum allowed mass deviation was set to 10 ppm for precursor ions and 0.6 Da for MS/MS peaks. For the datasets from SILAC samples which performed by UCD facility, the initial maximum allowed mass deviation was set to 10 ppm for precursor ions and 0.5 Da for MS MS peaks. Enzyme specificity was set to Trypsin/P, defined as C-terminal to arginine and lysine excluding proline, and a maximum of two missed cleavages were allowed. Carbamidomethyl cysteine was set as a fixed modification, N-terminal acetylation and methionine oxidation as variable modifications.

The corona data spectra were searched by the Andromeda search engine against the Bovine Uniprot sequence database (2015.05 release combined with common contaminants and concatenated with the reversed versions of all sequences. The particle complex data spectra were searched by the Andromeda search engine against the Human Uniprot sequence database (2015.06 release) and the Bovine Uniprot sequence database (2015.05 release). Protein identification required at least one unique or razor peptide per protein group. For the label-free datasets, 'Match True between Run' was chosen with 1 min in 'Alignment time window' for building a reference within datasets. Quantification in MaxQuant was performed using iBAQ algorithm co idering facility difference and time gaps between samples. The required false positive rate was set to 1% at the peptide and 1% at the protein level, and the minimum required peptide length was set to 6 amino acids.

The proteins which are identified in at least two biological repeats will be taken into further analysis. Venn plot[52] (http://www.interactivenn.net/) was used to show the identified proteins, Perseus 1.6.14.0 was used to perform volcano plot and heatmap, GO terms analysis was performed by Metascape[53] (https://metascape.org/gp/index.html#/main/step1) and PANTHER54 (http://pantherdb.org/), Circlize package[55] in R was used to plot circular chart of multilocated proteins referring to HUMAN PROTEIN ALTAS database[35] (https://www.proteinatlas.org/humanproteome/subcellular).

RNA extraction from particle complexes

All RNA extraction were performed in a Biosafety laboratory Level 3. The air entering the room is filter through HEPA filters in order to eliminate particles and reduce the risk of contamination. Magnetically separated particle complexes (extraction time: 90 mins) were resuspended in 100 ul RNase, DNase-free PBS. RNA was then extracted from complexes using one of the following kits: InviTrap Kit (INVITEK; cat. no. 1060100300), RNeasy Micro Kit (Qiagen; cat. no. 74034), miRNeasy Plasma Kit (Qiagen; cat. no. 217204), miRNeasy Cell Kit (Qiagen Cat. NO. 217684). Extracted RNA was quantified and RNA integrity was determined using the 2100 Agilent Bioanalyzer instrument using the RNA 6000 Nano and mall RNA assays. Samples with RNA integrity number (RIN) > 7 were used for downstream experiments.

RNA sequencing and data analysis

RNA sequencing was performed at BGI Genomics, China. mRNA-seq with paired end 150 bp sequence and small RNA-seq with single end 50 sequence were performed on the MGISEQ-2000 platform. Clean sequence data quality w s evaluated via FastQC prior to downstream bioinformatic analysis. Clean mRNA-seq reads were aligned to human reference genome GRCH38 via HISAT256 and then assigned mapped reads to genomic features by

featurecounts[57]. Differential gene expression analysis, volcano plot and heat map assimilation were performed by EdgeR[58,59], ggplot2[60] and pheatmap. Venn diagrams were produced using InteractiVenn[61]. Clean small RNA-seq data biotype analysis was performed using exceRpt[62] and then plotted via graphPad Prism 9. The biological replicates exhibit excellent reproducibility between different operators (Figure S4.40).

Author Information

Corresponding Authors

*Email: Kenneth.a.dawson@cbni.ucd.ie; yan.yan@cbni.ucd.ie

Author contributions

L.A, M.M., S.V., X.Y., and W.Z. contributed equally and are listed alphabetically. They will be putting their name first on the citation in their CVs. S.L., X.Y., and Y.D. conducted confocal imaging experiments with critical help from M.M. C.C., M.M., and L.S. prepared the proteomic samples, and X.Y. analyzed the original mass spectrometry data. M.M. and W.Z. performed RNA extraction. W.Z. and X.Y. analyzed RNA-seq data. W.Z. performed RT-qPCR experiments and analyzed the data. L.A. established critical reagents and characterization methodology. L.C. and Z.X. performed magnetic nanoparticle synthesis. Y.L. and L.S. established plasmids and stable cell lines. K.E. and A.F. performed fluorescent protein mapping experiment. K.E. and V.P. performed SEM and TEM experiments, respectively. E.S. and I.M. supported the workflow. Y.Y. and K.A.D. conceived and planned this project, co-wrote the manuscript with input from all co-authors. The authors declare no competing interests.

Supporting Information

Supporting methods

Materials

Fluorescein 5-isothiocyanate (≥90%); tetraethyl orthosilicate (reagent grade, 98%), (3 Aminopropyl)-trimethoxysilane (97%), ferric chloride hexahydrate (≥99%), ferrous chloride tetrahydrate (≥99.0%), oleic acid (technical grade, 90%) and cyclohexane (99.5%) were purchased from Sigma-Aldrich. Ethanol (99.8%, HPLC grade), acetone (Reagent grade, 99%), hydrochloric acid 37% w/w (ACS reagent, Honeywell Fluka™) and ammonia (35 % w/w) were purchased from Fisher scientific. Formaldehyde (HT501128), urea (U5378), 40% acrylamide/bis-acrylamide solution (A7802), trizma base (T1503), ammonium persulfate

(APS, A3678), N,N,N,N-tetramethylethylenediamine (TEMED, T9281), Triton X-100 (T9284), Tween 20 (P1379), skim milk powder (70166), glycine (G8898), Sucrose (50389), sodium dodecyl sulfate (SDS, L3771), ethylenediaminetetraacetic acid (EDTA, E4884), and DL-Dithiothreitol (DTT, D5545) were all purchased from Merck-Sigma Aldrich. Ethanol (32294) and methanol (34860) were purchased from Honeywell.

Antibodies

Rabbit anti-Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) antibody (14C10), mouse anti-Lamin A/C antibody (4C11), Rabbit anti-NUP98 antibody (C39A3) were purchased by Cell Signalling technologies.

Rabbit anti-Ki67 antigen (Ki67) antibody (ab92742), Rabbit anti-Lysosomal associated membrane protein 1 (LAMP1) antibody (ab24170), Mouse anti- Vimentin antibody (ab20346), Rabbit anti-Rab11 antibody (ab3612), Rabbit anti-beta COP antibody (ab289), Rabbit anti-REEP5 antibody (ab167405), Mouse anti- TOMM20 antibody (ab56783), HRP-functionalised Goat anti-Mouse IgG (ab97023) and anti-Rabbit IgG (ab6721) were purchased from Abcam.

Anti-alpha Tubulin antibody (236-10501), AlexaFluor® 546-functionalised Goat anti-Rabbit IgG (A11035), AlexaFluor® 546-functionalised Goat anti-Mouse IgG (A11003), Rabbit Anti–Alexa Fluor 405/Cascade Blue Dye antibody (A5760), anti-Calnexin antibody (GT1563) were purchased from Thermofisher.

Nanoparticles colloidal characterization

Size distribution of the nanoparticles was determined by DLS measurement after dispersion in water, differential centrifugal sedimentation (DCS) in a sucrose gradient (8-24% w/w in water) considering a colloidal SiO2 density of 2 g/cm3 (in the case of the composite magnetic particles the same density was considered meaning that the measured size doesn't correspond exactly to the real size), and transmission electron microscopy (TEM). Measurement of the zeta potential of the nanoparticles dispersed in HEPES 5 mM pH 7.4 was performed on a Malvern Zetasizer ZS series. In the case of the particles coated with biomolecules (Corona, Particle Complex), the DLS and the DSC were performed in PBS.

Negative staining protocol for transmission electron microscopy

After magnetic pull-down, the samples were fixed with a mixture of 2.5% Glutaraldehyde + 2% Paraformaldehyde (in 0.1 M phosphate buffer, pH 6.4) for 90 min at 4 °C in the dark. formvar/carbon copper grids (200 mesh) were hydrophilized (PELCO easiGlow™, negative charge, 1 min, 1 mA) and were immediately incubated with the fixed sample (0.4 µg) for 15

mins. The grids were washed six times with Milli-Q water (to remove any salt residue) prior to negative staining with 1% uranyl acetate + 1% (or 0.1% when indicated) trehalose for 2 mins. The last step was performed on ice and protected from light. The grids were left to air dry and convectional TEM images were obtained on a JEOL 2100 LaB TEM or FEI Tecnai G2 20 Twin both operating at 200 kV.

Cell culture

A549 non-small lung carcinoma (ATCC® CCL-185™), and HEK-293T (ATCC® CRL-11268™) cell lines were purchased from ATCC. A549s were cultured in MEM supplemented with 10% (v/v) FBS, 50 units/mL penicillin and 50 µg/mL streptomycin; referred to hereafter as cMEM. HEK-293T cells were cultured in DMEM supplemented with 10% (v/v) FBS, 50 units/mL penicillin and 50 µg/mL streptomycin; referred to hereafter as cMEM Cells were grown at 37˚C in a humidified atmosphere of 5% CO2 and sub-cultured at 70-80% confluence using trypsin (0.05% in EDTA). Cells were screened monthly for mycoplasma contaminations using the MycoAlert™ Mycoplasma Detection Kit (Lonza), and all cultures were free of contamination for the duration of experiments reported.

Lipid staining

BNS and corona-NP controls (2 µg/ml) were incubated with CellMask™ Orange Plasma membrane Stain (Thermo Fisher Scientific; 5 µg/ml working concentration) for 30 minutes at 37°C. The BNS and corona-NPs were collected by centrifugation and washed by resuspension in PBS and centrifugation. The corona-NP and BNS pellets were finally resuspended in 0.22 µm filtered PBS, for analysis by flow cytometry and microscopy.

Total protein staining

BNS and corona-NP controls were incubated with Alexa Fluor™ 405 NHS Ester (Succinimidyl Ester) or Alexa Fluor™ 568 NHS Ester (Succinimidyl Ester) (Thermo Fisher Scientific) at a w/w ratio of 1:4 dye:corona-NP or particle complex, for 1 h at room temperature with gentle shaking. To quench any remaining unreacted dye, glycine was added at a w/w ratio of 1:2 glycine:dye and incubated for 30 mins at room temperature with gentle shaking. The stained mSiO2 corona and particle complexes were collected from the dye solution by centrifugation and resuspended in PBS. They were then washed 3 times in PBS by centrifugation to remove unbound dye.

Particle complex imaging

Nanoparticle and particle complex confocal images were acquired with a Nikon Eclipse TI spinning disc confocal microscope equipped with a 100x objective. The particle complexes

were stained with the CellMask™ orange or Alexa Fluor™ NHS Ester (Succinimidyl Ester) dyes as previously described, and after the last wash step, the particles were resuspended in mounting media then 10 µl of this suspension was mounted on a microscopy slide for image analysis.

Flow cytometry

Nanoparticles and Particle Complexes were analysed using a Beckman Coulter CytoFLEX LX flow cytometer equipped with 6 lasers (UV-Violet-Blue-Yellow-Green-Red). Data was analyzed using the CytExpert 2.3.0.84 and FlowJo™ X 10.0.7r2 software. For nanoscale detection, the filters were configured so that the Violet SSC (VSSC) 405/10 nm channel served as the trigger channel to discriminate the noise, and the FITC channel was used to discriminate the core particle population.

Immunocytochemistry

For immunocytochemical localization of lamina A/C and Ki67, the cells were firstly fixed with a 4 % v/v paraformaldehyde (PFA) solution for 15 minutes at room temperature. They were then permeabilized with a 0.1% v/v Triton-X100 solution for 10 minutes at room temperature, followed by blocking with a 1 % w/v bovine serum albumin (BSA) solution for 1 hour at room temperature. The samples were then incubated for 1 hour with the primary antibody (2 to 5 ug/ml, in a 1 % w/v BSA solution), washed twice and incubated for 1 hour at RT with the secondary antibody (2 ug/ml, in a 1 % w/v BSA solution). All the steps were carried out keeping the samples protected from light.

Immunocytochemistry with optimized permeabilization

To preserve the Particle complexcargo during immunocytochemical localization experiments, a mild permeabilization protocol was developed to limit the detergent-induced dissolution that was observed using the standard (previously described) Triton-X100 permeabilization method. The cells were washed in PBS and fixed with 4 % v/v PFA for 15 minutes at room temperature, followed by two washes with fresh PBS. The permeabilization was performed with a 0.1 % v/v saponin solution in a 1 % w/v BSA solution for 10 mins. The samples were then incubated for 1 hour with the primary antibody (2 to 5 ug/ml, in 1 % w/v BSA), washed twice in BSA and incubated for 1 hour at room temperature with the secondary antibody (2 ug/ml, in 1 % w/v BSA). All the steps were carried out keeping the samples protected from light.

Fixed cell imaging

Fixed cell imaging experiments were carried out in glass bottom 96-well plates, seeded 24 hours prior with 5,000 cells per well. The cells were stained with DAPI (2 ug/ml) for 15 minutes at RT, or with AlexaFluor® 594-conjugated wheat germ agglutinin (5 ug/ml) for 20 minutes at RT, followed by immunostaining as previously described.

Particle complex cargo detachment live cell imaging

Particle complexes were stained using an Alexa Fluor™ NHS Ester (Succinimidyl Ester) dye as previously described. Cells were incubated with 40 ug/ml of stained Particle Complexes for 2 hours, then washed twice in PBS and twice in fresh media to remove non-internalized particles. The cells were incubated in fresh phenol-free media and after 1 hour of chase time, timelapses were acquired with a Nikon Eclipse TI spinning disc confocal microscope equipped with a 100x oil immersion objective. The chamber was pre warmed at 37°C and the CO_2 concentration was set at 5 %. The frequency of acquisition was 0.1 frame/s. Time-lapses were analyzed with Imaris 1.7.6 imaging software (Bitplane), using to the spots tracking tool.

Super-resolution radial fluctuations (SRRF) imaging

SRRF imaging was performed using a Nikon Eclipse TI spinning disc confocal microscope equipped with a 100x objective. The images were acquired and reconstructed with the Fusion software (Andor) according to the manufacturer instruction. The samples were prepared for SRRF imaging as outlined in the “fixed cell imaging” section.

Quantitative reverse transcription PCR (RT-qPCR)

The following genes were validated by qPCR: CEACAM6, OAS1, MUC5AC, KRT4 and EGFP-FUS; GAPDH was used as the housekeeping gene. After RNA extraction, cDNA was obtained by reverse transcription using the high-capacity cDNA reverse transcription kit (ThermoFisher; cat. no. 4368813). qPCR was performed using PowerUp™ SYBR™ Green Master Mix (ThermoFisher cat. no. A25742) together with specific gene primers (Table 4.1) and analysed on a QuantStudio™ Real-Time PCR System. The comparative ΔCt ($\Delta\Delta$Ct) method was used to calculate the fold change of target mRNA expressions.

Supporting figures

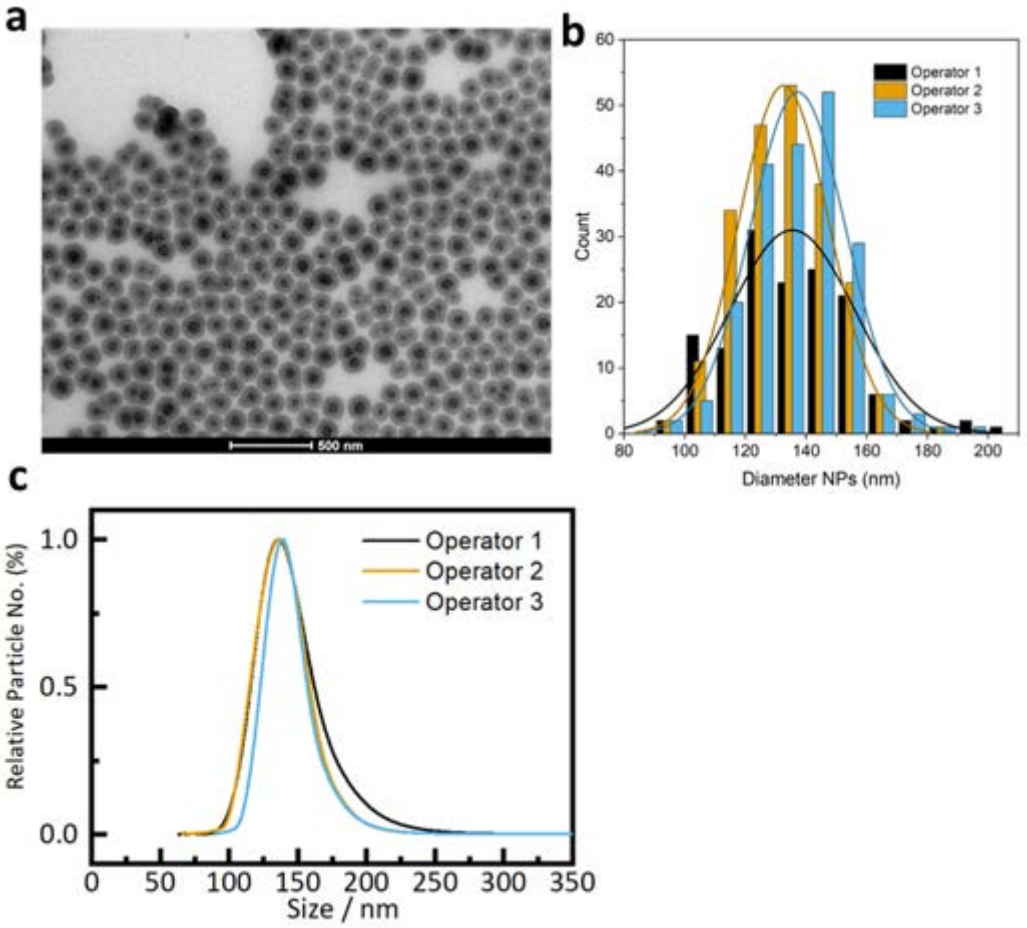
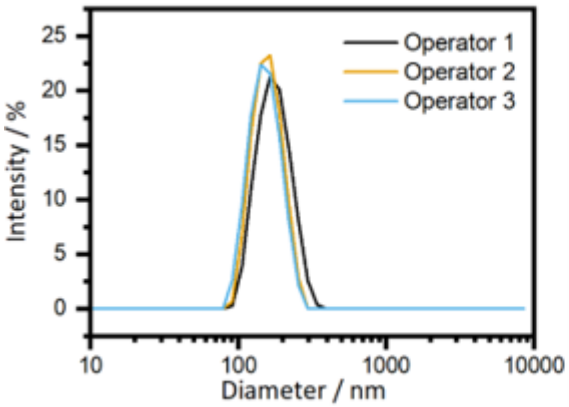


Figure S4. 1 Size characterization and batch-to-batch reproducibility of $mSiO_2$ NPs.

(a) Representative TEM images of the $mSiO_2$ particles illustrating the magnetic multicore encapsulated in a silica shell. (b) Size distribution of 3 independent batches, based on manual counts of particles in ImageJ software from ($n \geqslant 140$ total), and corresponding Gaussian distribution fits. The mean and standard deviation for the batches from Operator 1, 2 and 3 are 136 ± 21, 133 ± 15 and 137 ± 16, respectively. (c) The average particle size for the same 3 independent batches is ca. 140 nm by DCS, in good agreement with TEM measurements and demonstrating good reproducibility between separate batches by independent operators.



Sample	Z-Ave (nm)	PDI	Zeta Potential (mV)
Operator 1	168 ± 4	0.04 ± 0.02	-38 ± 2
Operator 2	153 ± 1	0.03 ± 0.02	-41 ± 1
Operator 3	148 ± 1	0.02 ± 0.02	-38 ± 1

Figure S4. 2 Colloidal dispersion of $mSiO_2$ NPs Colloidal stability for the particles was assessed via DLS measurements.

DLS measurements for the same independent batches show that the particles are all of a similar hydrodynamic size around 150-170 nm; the hydrodynamic size is larger than the true diameter determined from TEM, as expected due to the solvation of the particles, but also due to the stronger contribution of the larger particles in the measured scattered intensity. Zeta potential measurements of the particles in 5 mM HEPES pH 7.4 also indicate that the particles have a large negative zeta potential value attributed to the surface silanolate groups.

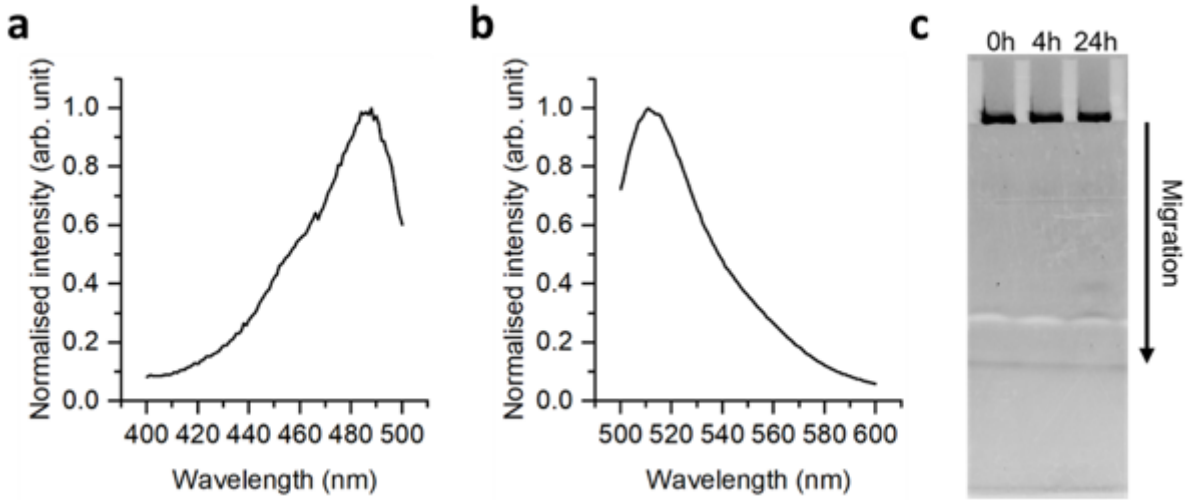


Figure S4. 3 Fluorescence characterization of $mSiO_2$ NPs.

Typical excitation (a) and emission (b) spectra of mSiO2 NPs at a concentration of 0.5 g/L in PBS. The excitation peak maximum of 488 nm when measuring emission intensity at 515 nm and emission peak maximum at ca. 511 nm when measuring excitation at 488 nm are typical of FITC dye and indicate that the dye was successfully incorporated into the silica shell. (c) Dye leaking test: to both ensure dye stability and verify that incorporated dye does not leach from the particles under experimental conditions, $mSiO_2$ particles were immersed in cMEM at

a concentration of 0.1 g/L and incubated at 37°C for different times, to simulate incubation conditions for the particles in cell exposure experiments. Subsequent dilution of exposed particles in SDS loading buffer and PAGE at 130 V for 45 min in 10% acrylamide gel was used to test for dye leakage, as the particles are too large to enter the gel matrix. No FITC dye leakage from the particles could be observed and all detected fluorescence (in black in the figure) was co-localized with particles in the wells.

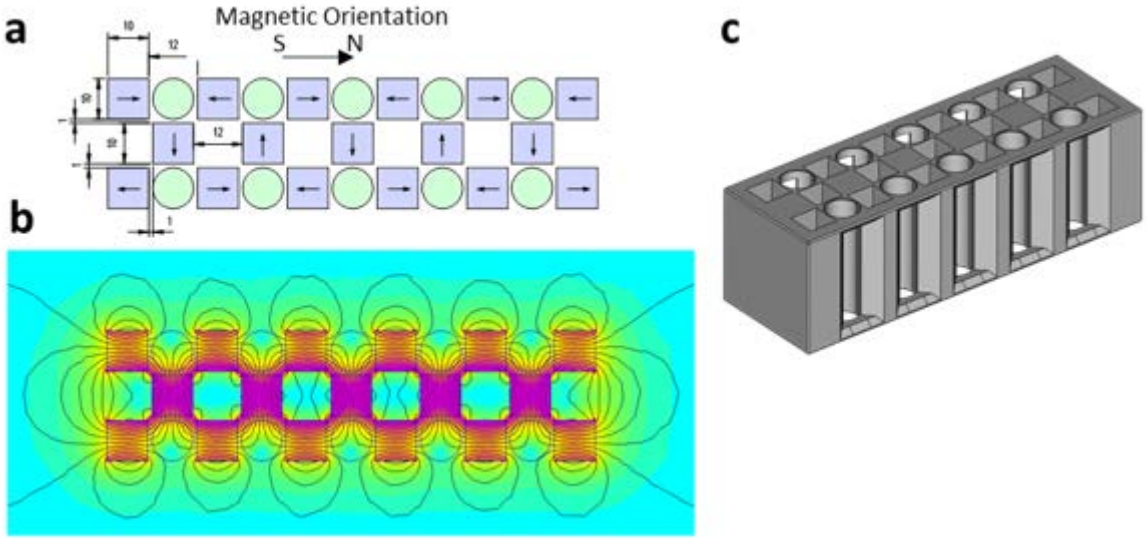


Figure S4. 4 Development of the in-house magnetic rack used for $mSiO_2$ NP extraction.

(a) Schematic of the organization of the magnet (blue) (Speromagnet, Germany), Material NdFeB, Magnetization N42, 40 x 10 x 10 mm), and position of the tubes (green), dimensions are in millimeter. (b) Simulation of the magnetic flux lines and density, realized with FEMM 4.2. This magnet configuration allows the formation of high magnetic gradients that will attract the particles on one side of the tube. (c) Three-Dimensional representation of the designed magnetic rack. Gaps of the 3D structure have been included to facilitate the visibility of the samples in the tube during extraction.

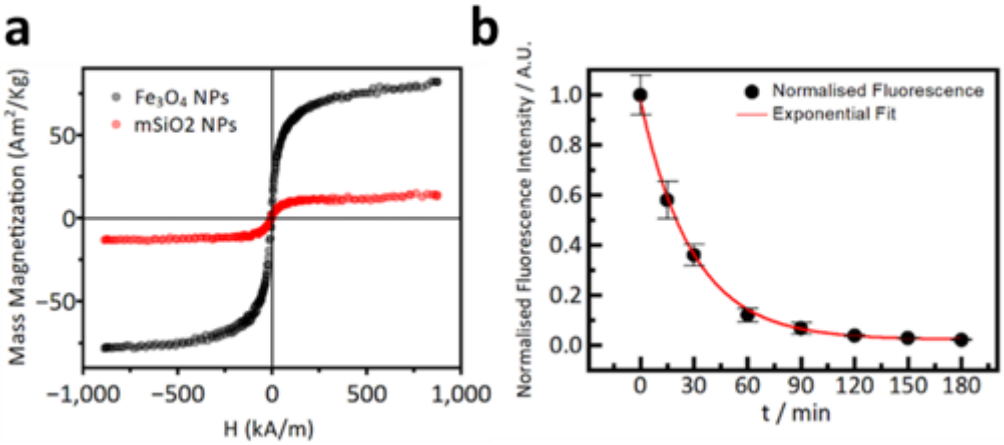


Figure S4. 5 Magnetic Properties of $mSiO_2$ NPs.

(a) Magnetization curve: the sample presents no coercivity and no remanent magnetization at room temperature indicating that the particles are superparamagnetic. The mass magnetic saturation is lower for $mSiO_2$ compared to the magnetite NPs forming the core because of the presence of the silica shell. Reproduced from Ref. 33 (Cursi *et al.* 2021) with permission from the Royal Society of Chemistry. (b) Magnetic pull-down kinetics of particles in the magnetic rack. The particles were diluted to a concentration of 0.5 mg/mL in a 1 mL Eppendorf tube. The progression of the pull down was monitored by measuring the fluorescence of the supernatant exciting at 488 nm and measuring at 515 nm at a Horiba Fluorolog. Fluorescence intensities are reported normalized to the value at t = 0 i.e., before the start of the pull-down experiment. Values are average of 3 independent measurements ± standard error. As it has already been demonstrated (Figure S4.3) that FITC dye does not leak from the silica shell of the MCSP@$fSiO_2$, a decrease in the supernatant fluorescence can be used to assess magnetic precipitation of particles from solution. As shown in the figure, the normalized fluorescence intensity decrease is well-described by an exponential decay curve with the pull down being more than 50% within the first 30 minutes and essentially 100% pull down after t = 120 minutes.

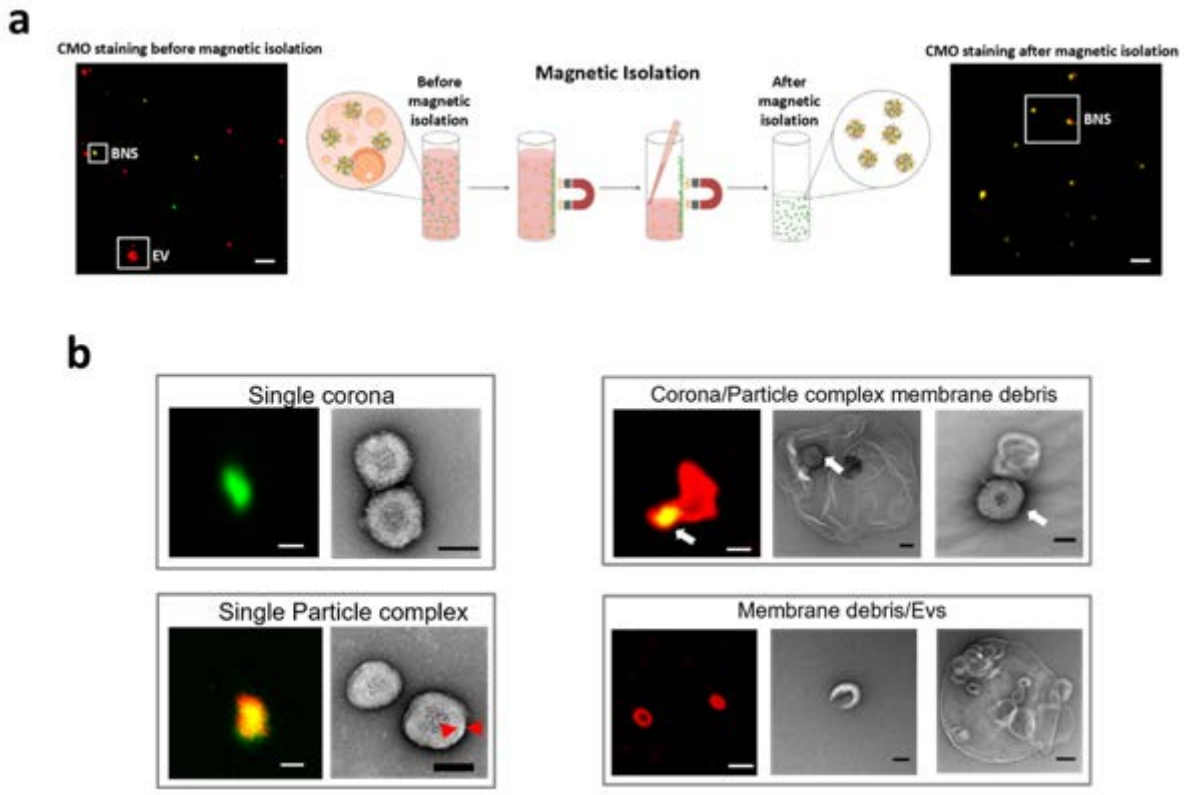
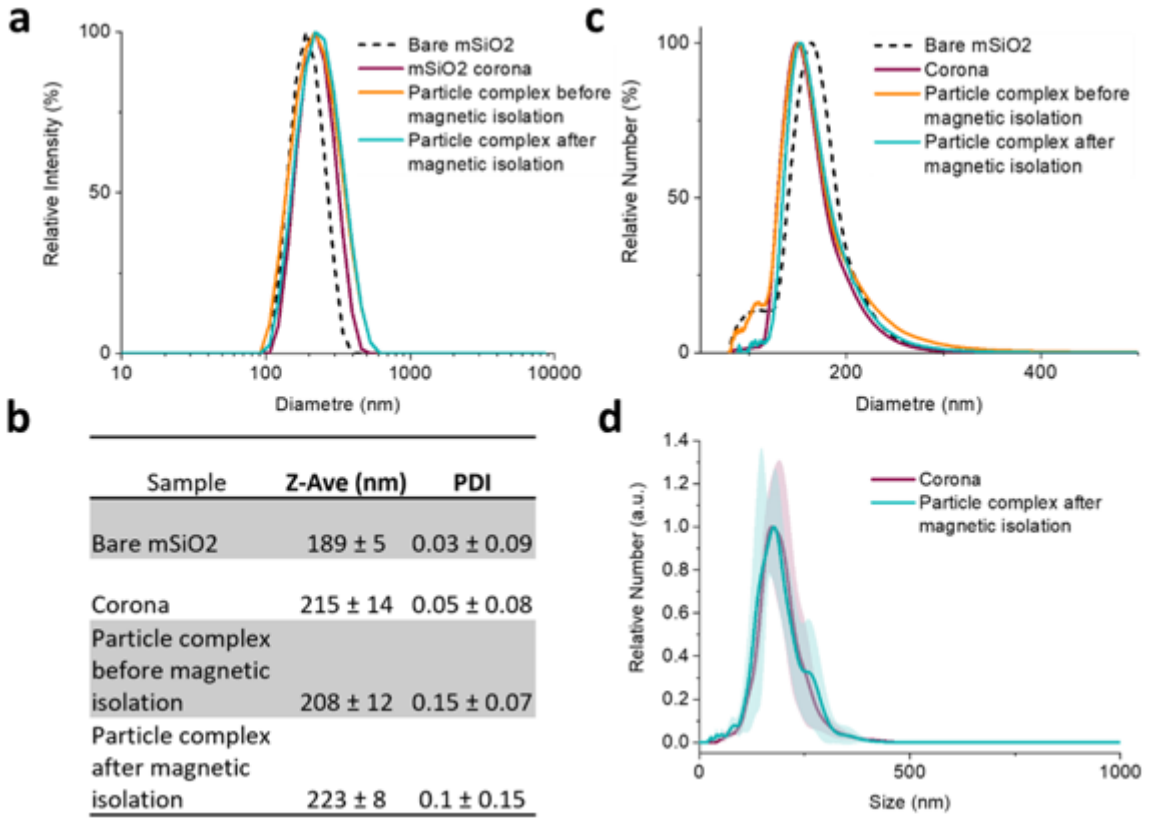


Figure S4. 6 Quantification of aggregation and extracellular vesicles before and after magnetic isolation.

(a) Confocal image analysis of particle complexes produced using $mSiO_2$ nanoparticles (green) and stained with CMO (red) reveal that there are membrane debris/extracellular vesicles (EV) products also collected with the particle complexes during the centrifugation step; these contaminations are stained red, while the particle complexes appear yellow. (b) Confocal and TEM analysis (staining no.1) of the different populations of particles, including particles aggregated with membrane debris, and extracellular vesicles.



Sample	Z-Ave (nm)	PDI
Bare mSiO2	189 ± 5	0.03 ± 0.09
Corona	215 ± 14	0.05 ± 0.08
Particle complex before magnetic isolation	208 ± 12	0.15 ± 0.07
Particle complex after magnetic isolation	223 ± 8	0.1 ± 0.15

Figure S4. 7 Characterization of the colloidal dispersion of the Particle complex.

(a) and (b) DLS distribution and Cumulant results, respectively, after dispersion of the samples in PBS. The sample “Corona” corresponds to particles incubated one hour in cMEM at 37 °C before being diluted by 10 with PBS for the measurement. Average of 3 measurement ± 4.3*SD. (c) Size distribution obtained by DCS. (d) Size distribution obtained by NTA. Measurements were realized at 1 µg/mL in PBS Representation of the average of three independent measurements and the standard deviation. The shift of the dominant population toward the smaller sizes in DCS is an indication of the adsorption of biomolecules on the surface of the nanoparticles reducing the overall density of the particle leading to a slower sedimentation and a smaller apparent diameter.

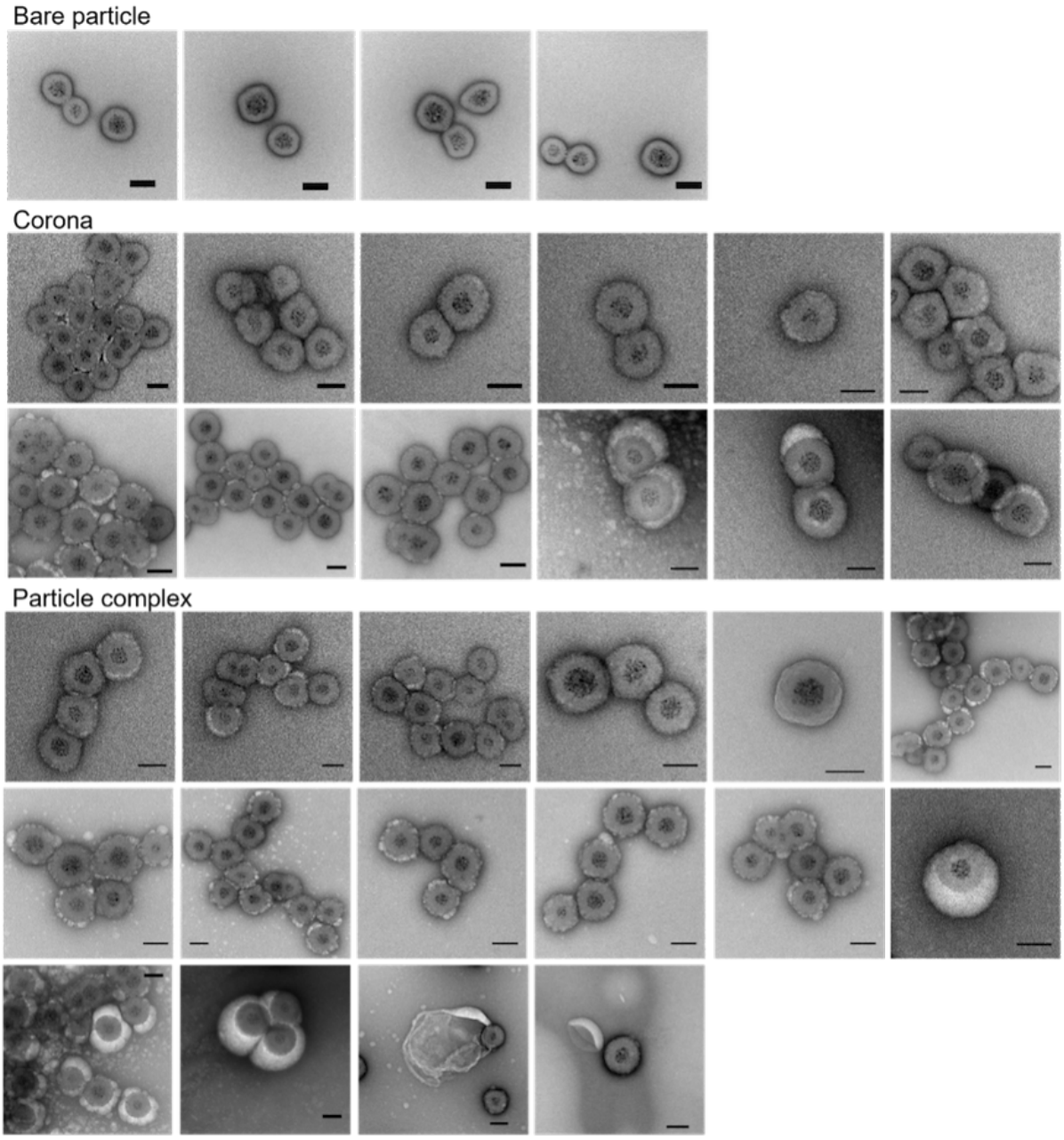


Figure S4. 8 TEM analysis of negatively stained (staining no.2 with 1% Trehalose) $mSiO_2$ NPs, corona and particle complex.

For the bare particles, it is possible to observe the electron dense magnetic core and the silica shell, but no evident structure is observed at the surface of the nanoparticles. For the nanoparticles with a corona and for the particle complexes, globular structures can be observed at the surface of the nanoparticles. These structures may correspond to the biomolecular corona; however, they could be attributed in certain cases (in particular in the case of think staining) to beam damage. It must be noticed the occasional presence of particle complexes aggregated with membranous objects (2 last images). Scale bar: 100 nm.

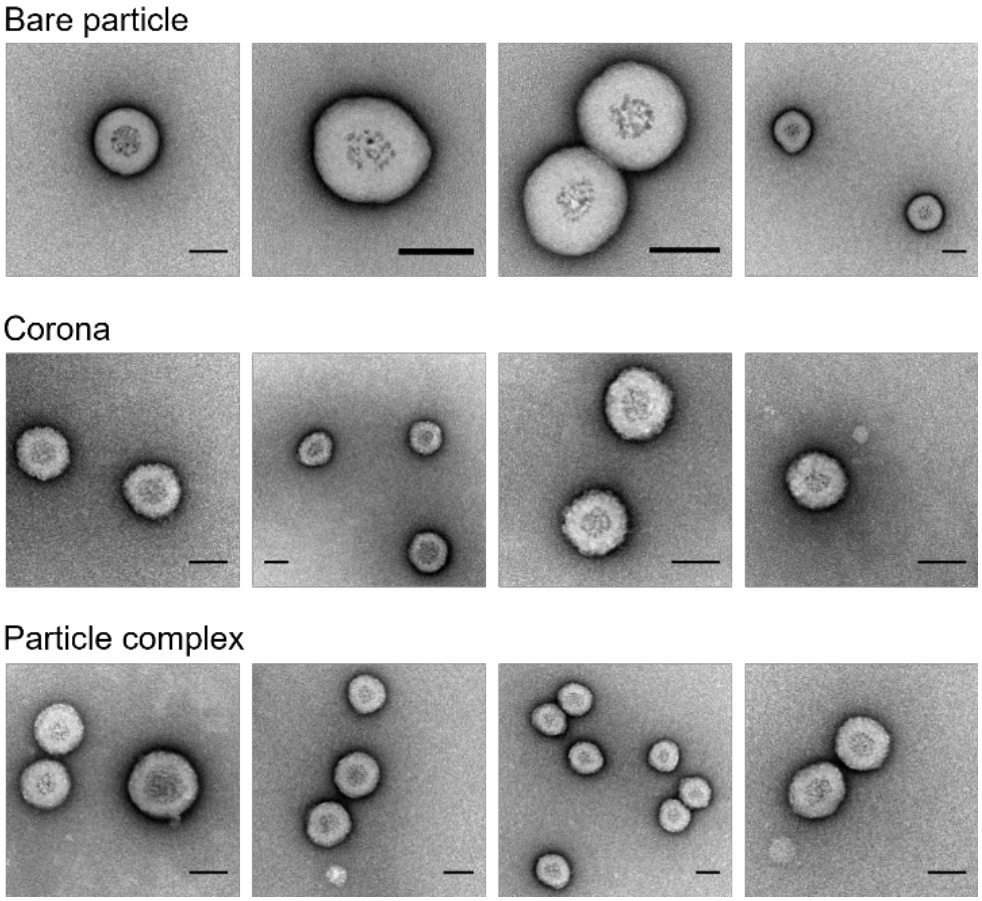


Figure S4. 9 TEM of negatively stained (staining No.1) bare $mSiO_2$, corona and particle complex.

Trehalose was not used as an additive to the staining agent to prevent artefacts from thick staining and “bubbling”. No structure can be observed at the surface of the bare $mSiO_2$. Globular structures can be observed at the surface of the corona and protein complexes. Scale bar: 100 nm.

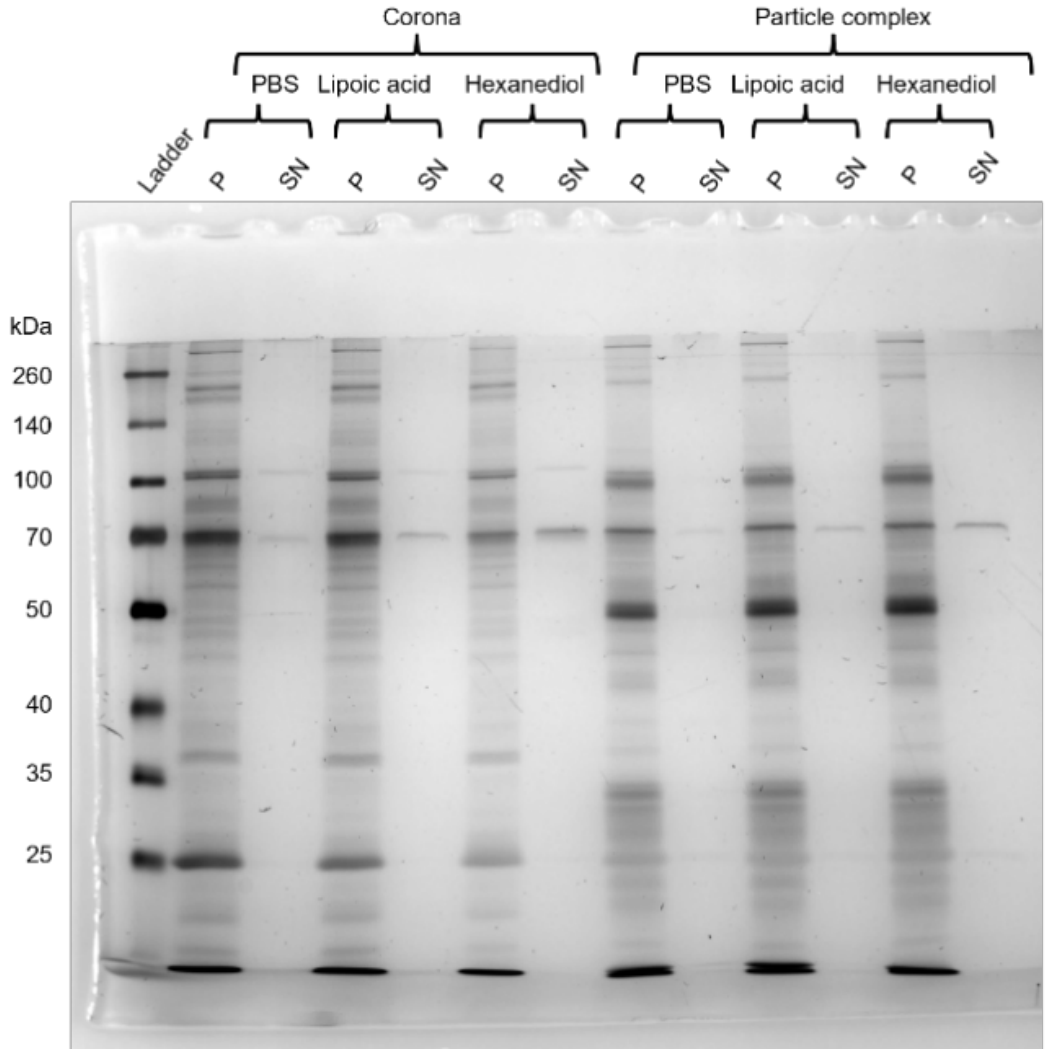


Figure S4. 10 Test of detachment of the biomolecular layer.

The biomolecular layer was exposed to Lipoic acid or 1,6-hexanediol, known for dissolving liquid–liquid phase separated biomolecules condensates. Silver staining of SDS-PAGE of 5 µg of particles after 12 h incubation in 1mM Lipoic acid (in PBS), or 5% w/v of 1,6-hexanediol in PBS. The potential proteins released in solution were separated from the nanoparticles by centrifugation and the pellet (P) and supernatant (SN) were loaded separately on the gel after being heated at 95°C for 10 min in presence of loading buffer and DTT. No protein was detected in the supernatant, suggesting that the Particle complexbiomolecular layer didn't detach for the different conditions tested. The visible band in all the supernatants is attributed to a residual contamination from FBS (BSA).

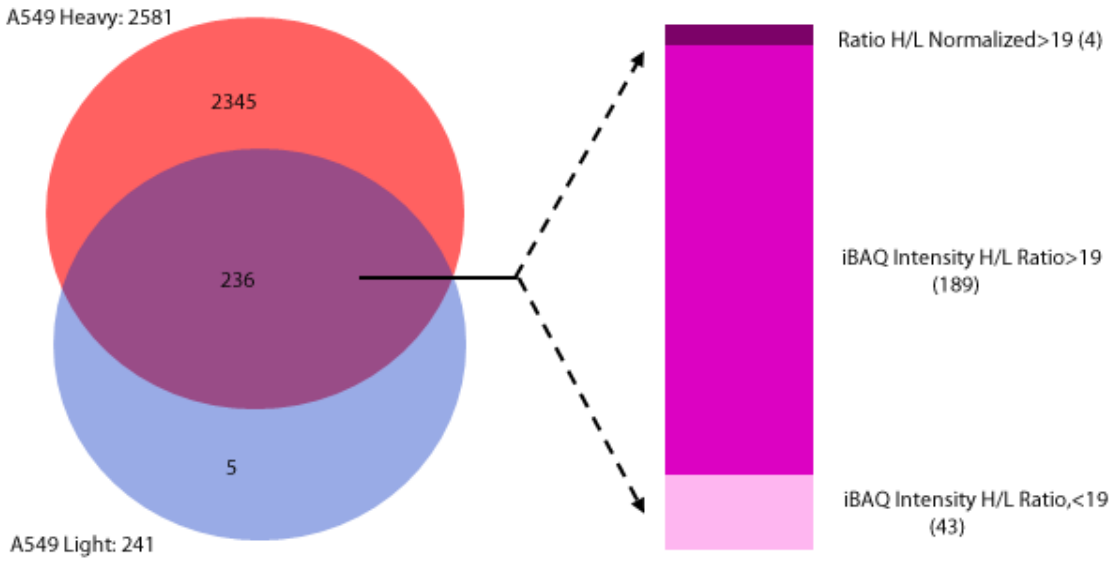


Figure S4. 11 A549 cell labelling efficiency by SILAC proteomics.

Blue circle is the light proteins, red circle is the heavy proteins, the purple circles are common proteins. In which, the 4 represents the proteins have normalized H/L ratio over 19 (95% fully labelled for a protein), and 189 represents the proteins have the iBAQ intensity ratios over 19, so in total there are 98.14% fully labelled proteins.

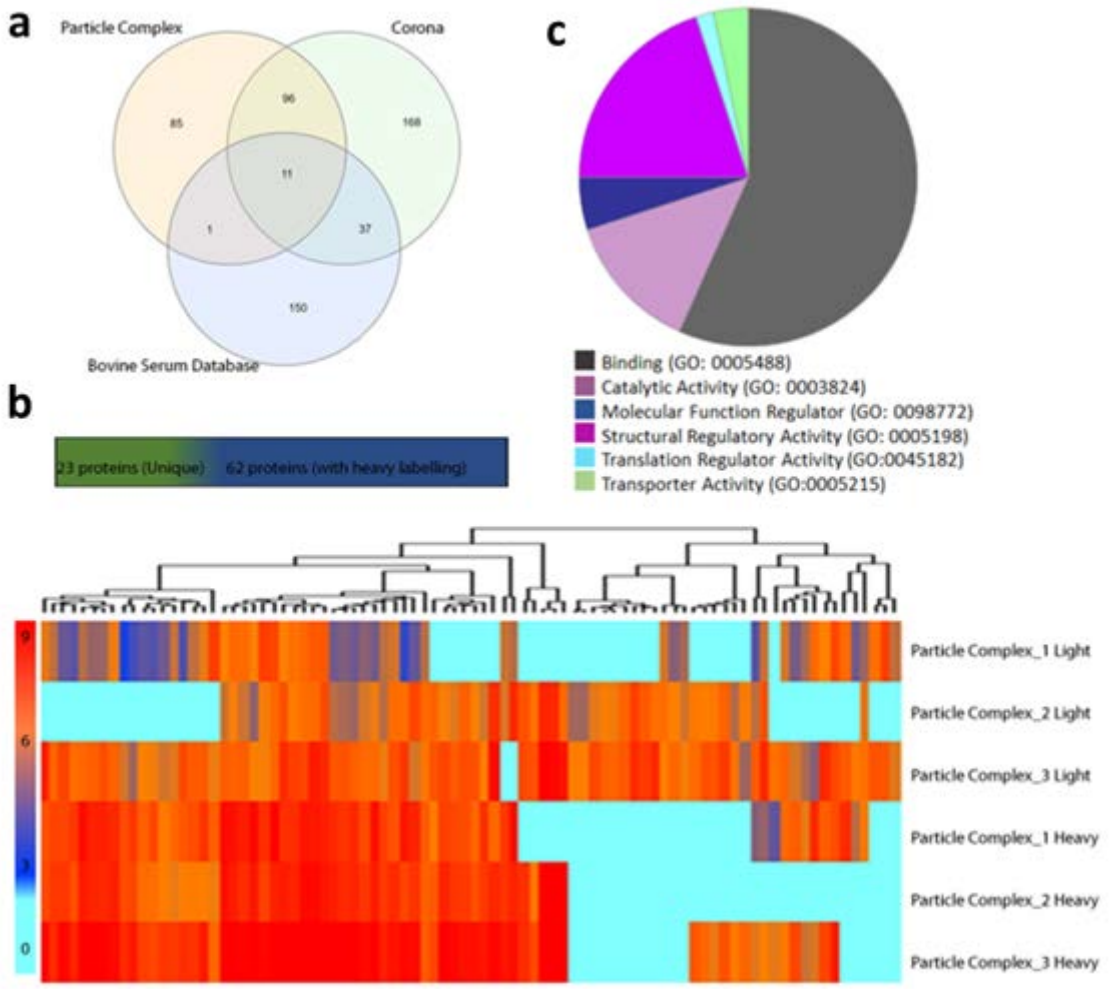


Figure S4. 12 Light proteins in particle complex and corona from SILAC proteomics.

(a) Venn plot of light proteins between particle complex and corona. The light proteins are obtained by firstly running against Bovine database and then selection of 2 out 3 biological repeats. 'Bov Serum Database' is a database containing bovine serum proteins. Venn plot was performed by Interactivenn (http://www.interactivenn.net/) by Uniprot ID. There are 85 light proteins which are only identified in CP. (b) Heatmap of light proteins and corresponding heavy proteins of these 85 proteins. Colours from blue to orange to red represents the protein iBAQ intensity from low to high, and the cyan is unidentified. 62 proteins which both have heavy and light labelling probably are the common proteins in bovine and human, however they aren't identified in corona samples, meaning they are probably human proteins. (c) GO term analysis of 62 proteins. The GO term analysis was performed by PANTHER (http://pantherdb.org/), and the shown terms are terms of Molecular Function.

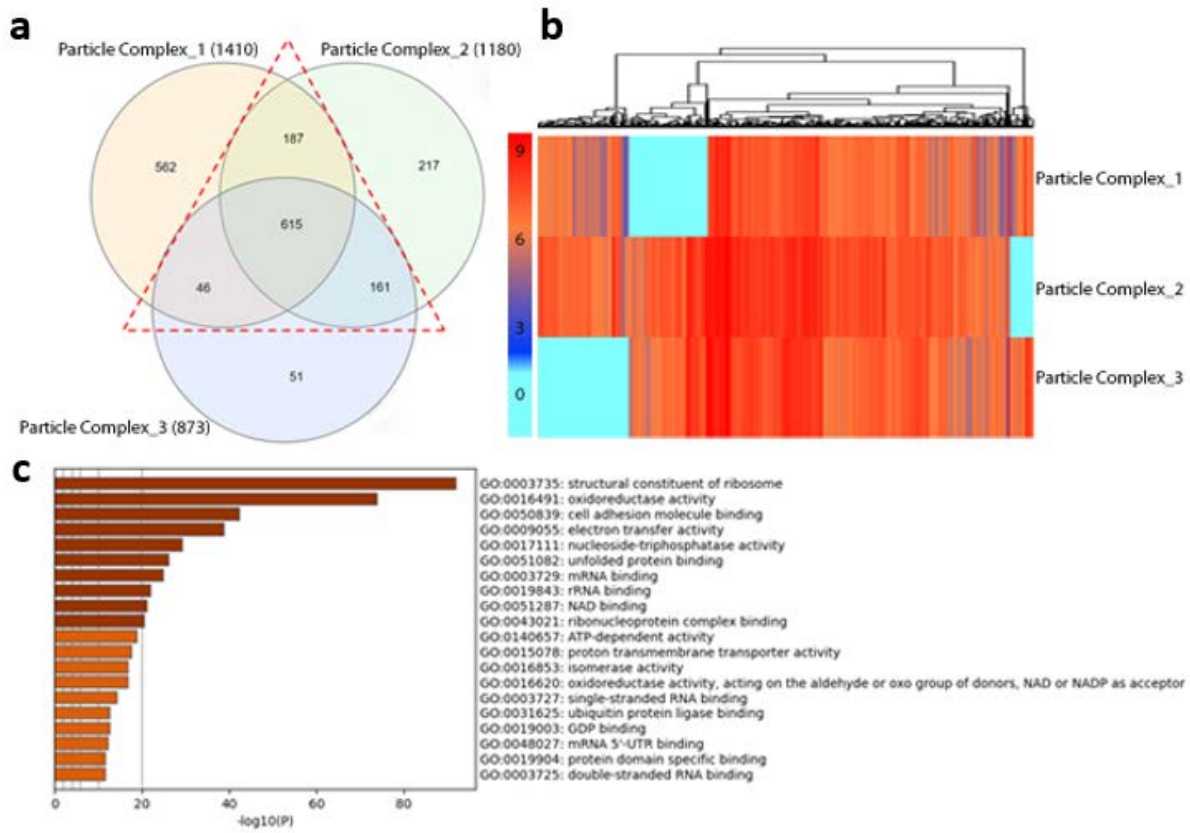


Figure S4. 13 H-proteins in particle complex from SILAC proteomics.

(a) Venn plot of heavy proteins in particle complex. Heavy proteins are obtained by running against human database and performed by Interactivenn (http://www.interactivenn.net/) by Entry name. (b) Heatmap of heavy proteins. Heavy proteins are obtained by selection of 2 out 3 biological repeats (in triangle in a). Colors from blue to orange to red represents the protein iBAQ intensity from low to high, and the cyan is unidentified. (c) GO term analysis of heavy proteins. The GO term analysis was performed by Metascape (https://metascape.org/), and the shown terms are terms of Molecular Function.

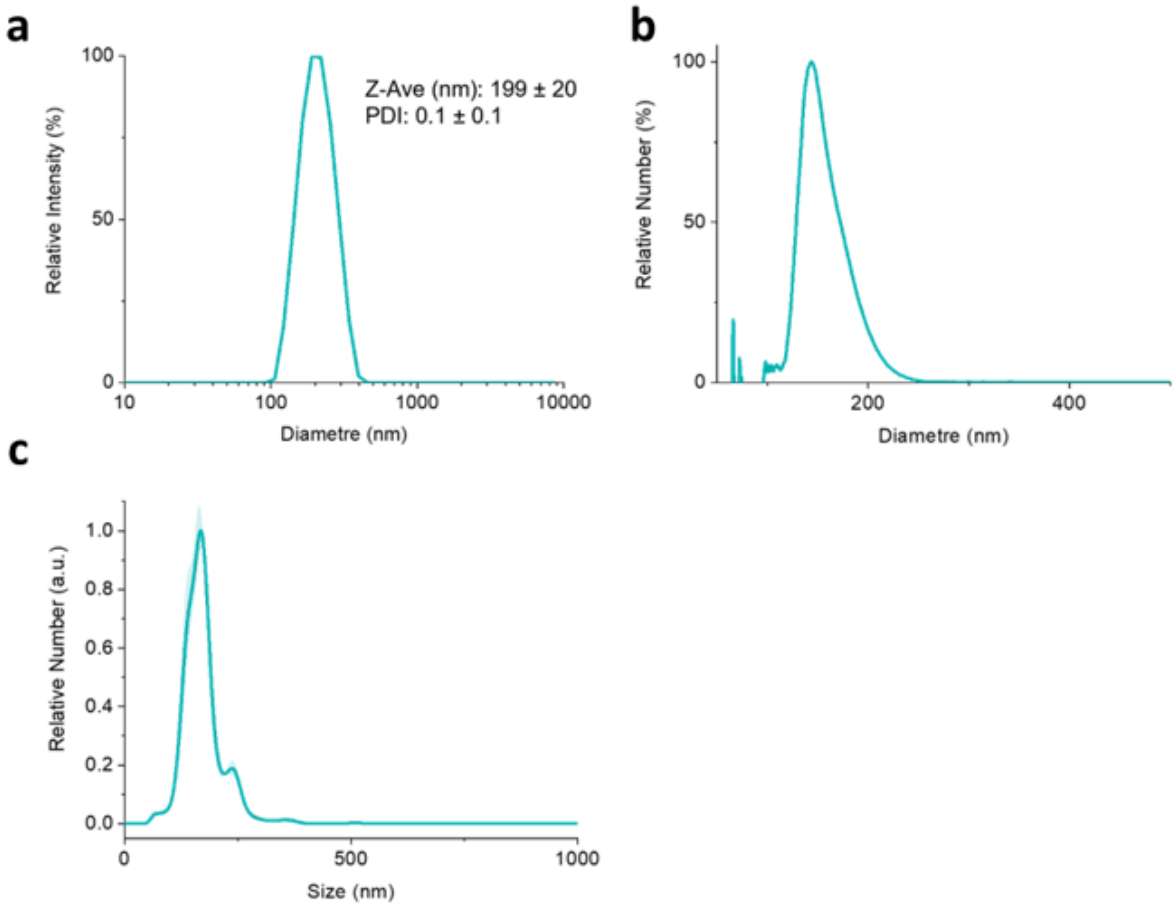
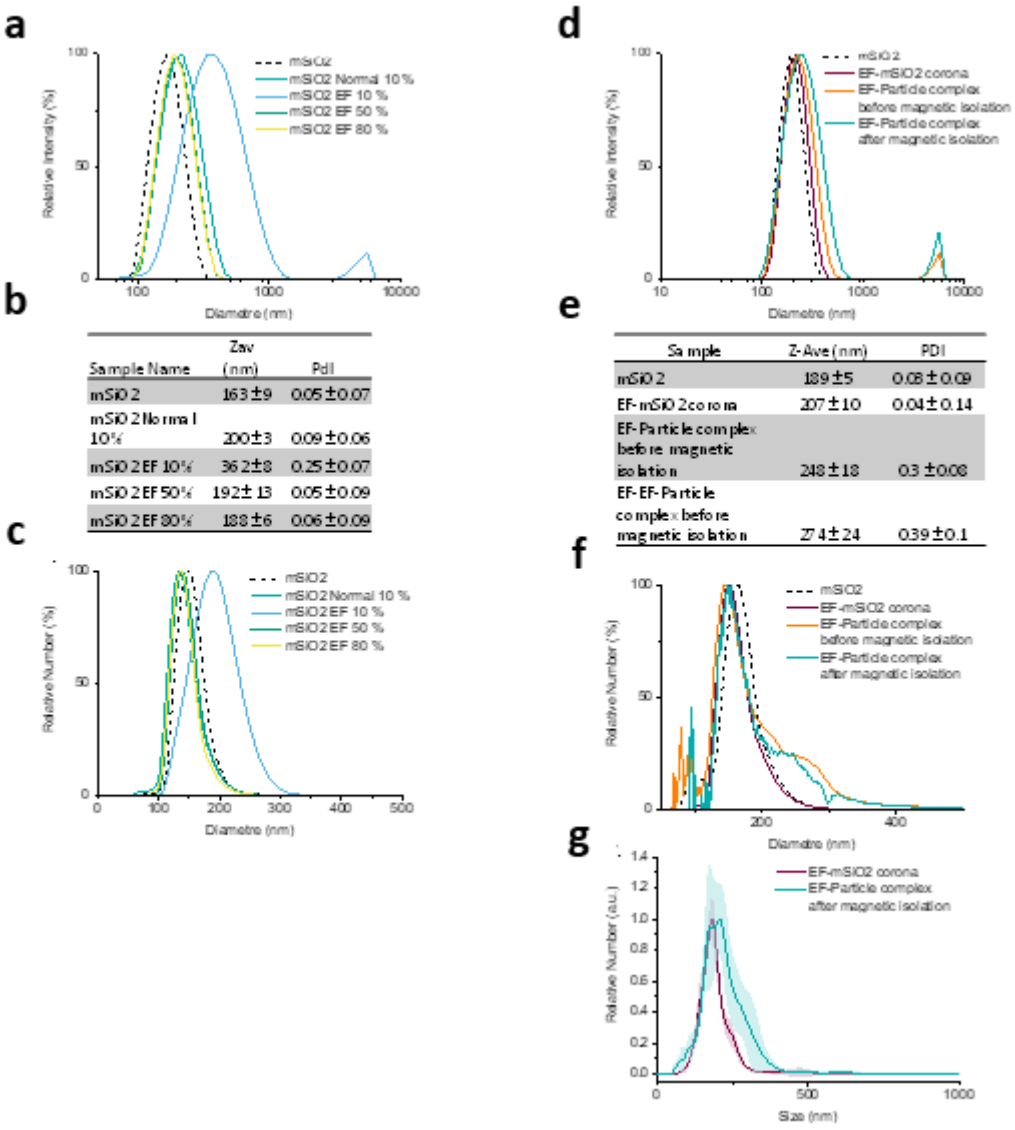


Figure S4. 14 Characterization of particle complex prepared from HEK-T cells.

(a) DLS distribution and Cumulant results after dispersion of the sample in PBS. Average of 3 measurement ± 4.3*SD. (b) DCS distribution, (c) Size distribution obtained by NTA. Measurements were realised at 1 µg/mL in PBS Representation of the average of three independent measurements and the standard deviation. The particle complex collected from HEK-T cells are colloidally stable.



Sample Name	Zav (nm)	PdI
mSiO2	163±9	0.05±0.07
mSiO2 Normal 10%	200±3	0.09±0.06
mSiO2 EF 10%	362±8	0.25±0.07
mSiO2 EF 50%	192±13	0.05±0.09
mSiO2 EF 80%	188±6	0.06±0.09

Sample	Z-Ave (nm)	PDI
mSiO2	189±5	0.08±0.09
EF-mSiO2 corona	207±10	0.04±0.14
EF-Particle complex before magnetic isolation	248±18	0.3±0.08
EF-EF-Particle complex before magnetic isolation	274±24	0.39±0.1

Figure S4. 15 Colloidal stability of exosome-free vs FBS mSiO2 corona (a-c) and Characterization of EF-Particle complexprepared from A549 cells (d-f).

Colloidal stability of exosome-free vs FBS $mSiO_2$ corona (a) DLS distribution and (b) Cumulant results. After their incubation for one hour in the media at 37 °C at 0.8 g/L, nanoparticles were centrifuged 5 min at 8000 g and resuspended in PBS at 80 mg/L for the measurement. Average of 3 measurement ± 4.3*SD. (c) size distribution obtained by DCS. Particles in the media were injected without dilution. The results indicate that the concentration of Exosome Free (EF) serum had to be increased to 50% to reach the same colloidal stability as with 10% of standard FBS. This observation agrees with the fact that the total protein concentration of the EF serum was 5 times lower than the standard FBS (as measured by BCA). Characterization of EF-particle complex prepared from A549 cells. (d) DLS distribution and (e) Cumulant results, respectively, after dispersion of the samples in PBS. Average of 3 measurement ± 4.3*SD. (f) Size distribution obtained by DCS. (g) Size distribution obtained

by NTA. Measurements were realized at 1 µg/mL in PBS representation of the average of three independent measurements and the standard deviation. The different results indicate a significant colloidal aggregation, not observed for the regular particle complex.

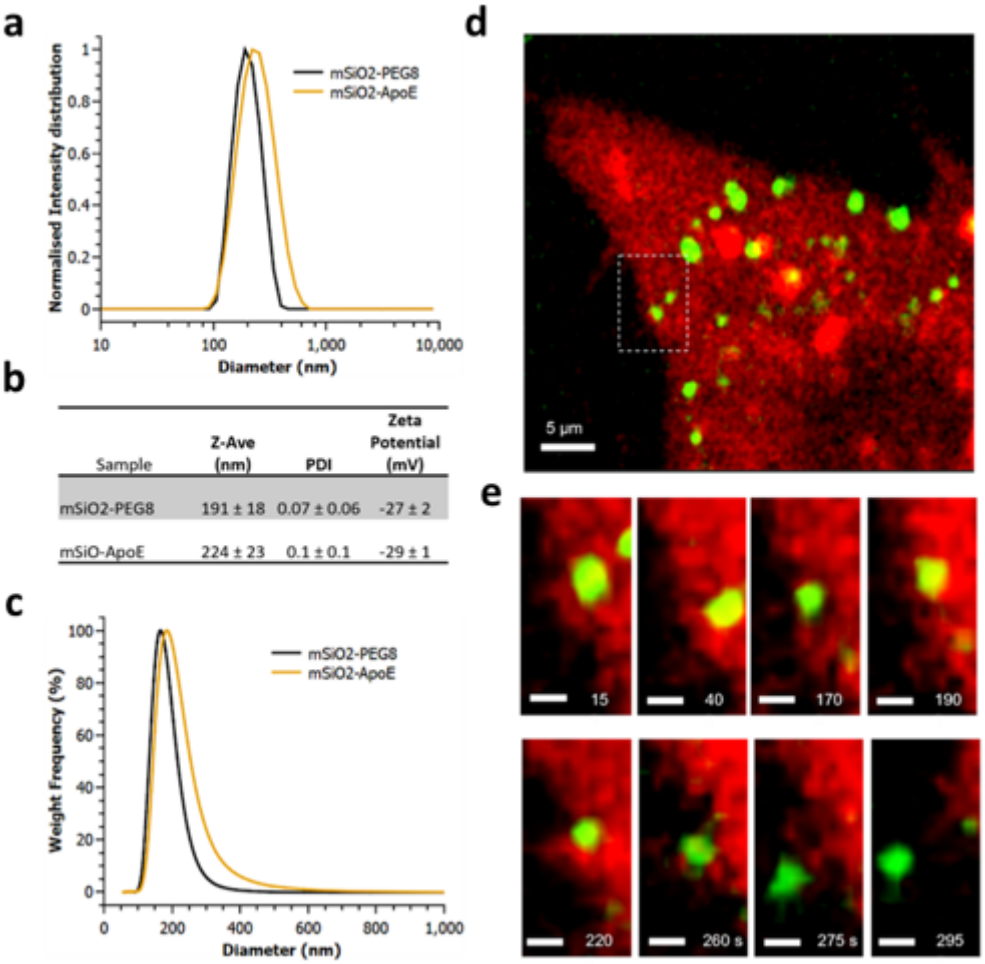


Figure S4. 16 Characterization of ApoE-mSiO$_2$.

(a) DLS size distribution (in PBS), (b) DLS cumulant results (in PBS) and Zeta potential (in HEPES buffer 5 mM, pH 7.4). Average of 3 measurement ± 4.3*SD. (c) Size distribution obtained by DCS of the particles dispersed in PBS. The amount of ApoE grafted on the nanoparticles was quantified by Bicinchoninic Acid (BCA) protein assay (Micro BCA Protein Assay Kit, Thermo scientific), It was found that these is 6 mg of protein grafted per g of particle. (d) Live cell imaging frames showing the ApoE-particle complex biogenesis in U87 MG cells after pulse with 800 ug/ml mSiO$_2$-ApoE for 10 minutes. (e) time frames showing the release of ApoE-particle complexes. Red: cell membrane; green: PS-COOH NPs. Inset scale bar: 2 µm.

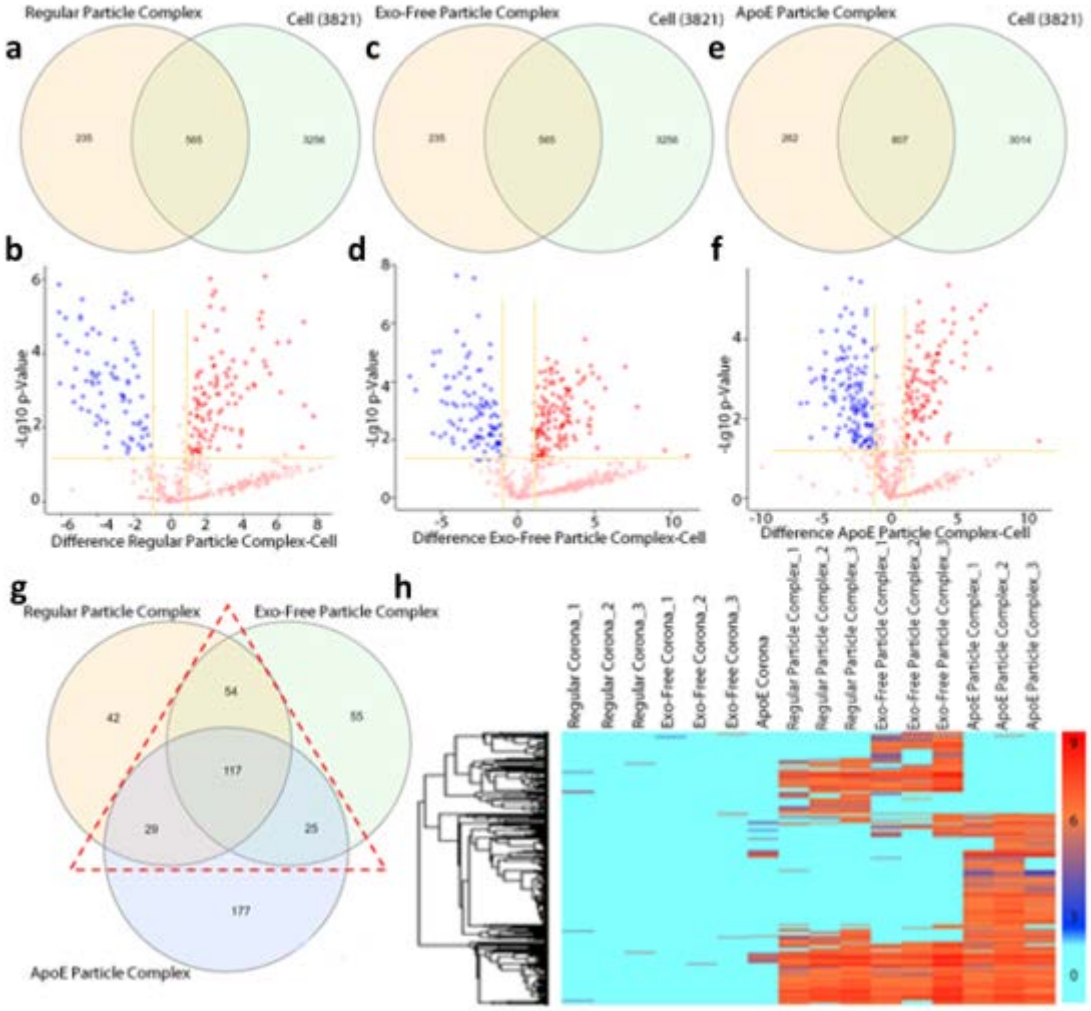


Figure S4. 17 Proteomic analysis of commonly enriched proteins in particle complex.

(a), (c) and (e) Venn plot of label-free MS analysis. Regular particle complex is the particle complex produced from the cells treated with corona made by complete MEM media, Exo-Free particle complex is the particle complex produced by cells treated with corona made by exosome free media, and the ApoE-particle complex is the particle complex produced by the cells treated with ApoE grafted corona. And cell is the proteomics of corona free A549 cells. (b), (d) and (f) Volcano plot between particle complexes and cells. Red are the significant enriched proteins in cells, and blue are the significant decreased proteins in the particle complex. P-value of 0.05 and 2-fold change were used for thresholds. (g) Venn plot of selected proteins. There are two selection rules: 1. particle complex proteins were identified at least two times in one of particle complex types; 2. there are less than 2 (<2) corresponding corona proteins identified. (h) Heatmap of total proteins shown in the triangle in g. The log10(iBAQ intensity) was used to correlate with colors. Values are increased from light blue to red. The cyan means no protein was identified.

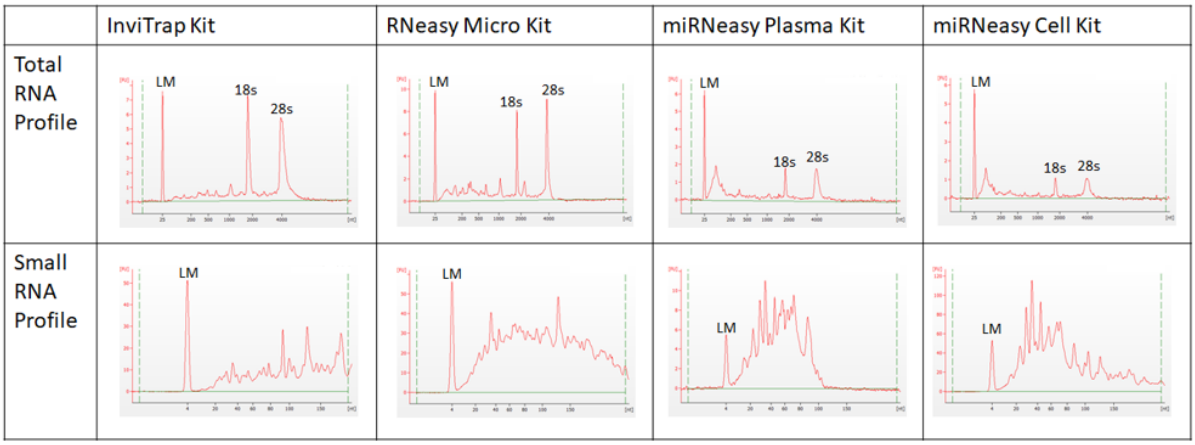


Figure S4. 18 Small and total RNA profiles for the particle complex, extracted by different kits: InviTrap Kit (INVITEK, Cat. NO. 1060100300), RNeasy Micro Kit (Qiagen Cat. NO. 74034), miRNeasy Plasma Kit (Qiagen Cat. NO. 217204), miRNeasy Cell Kit (Qiagen Cat. NO. 217684), LM: ladder marker. Analysis was performed using the Agilent 2100 Bioanalyzer with the small and RNA 6000 Nano assays.

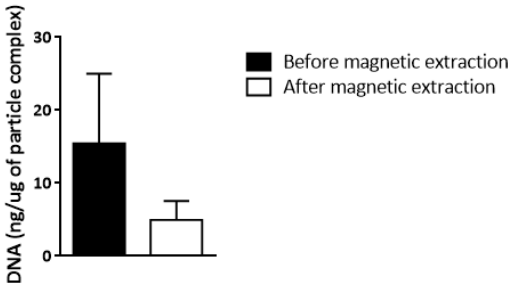


Figure S4. 19 The particle complex contains dsDNA that is retained after magnetic extraction, measured by the Quant-iT™ dsDNA Assay Kit (high sensitivity). n = 2 biological replicates ± SD.

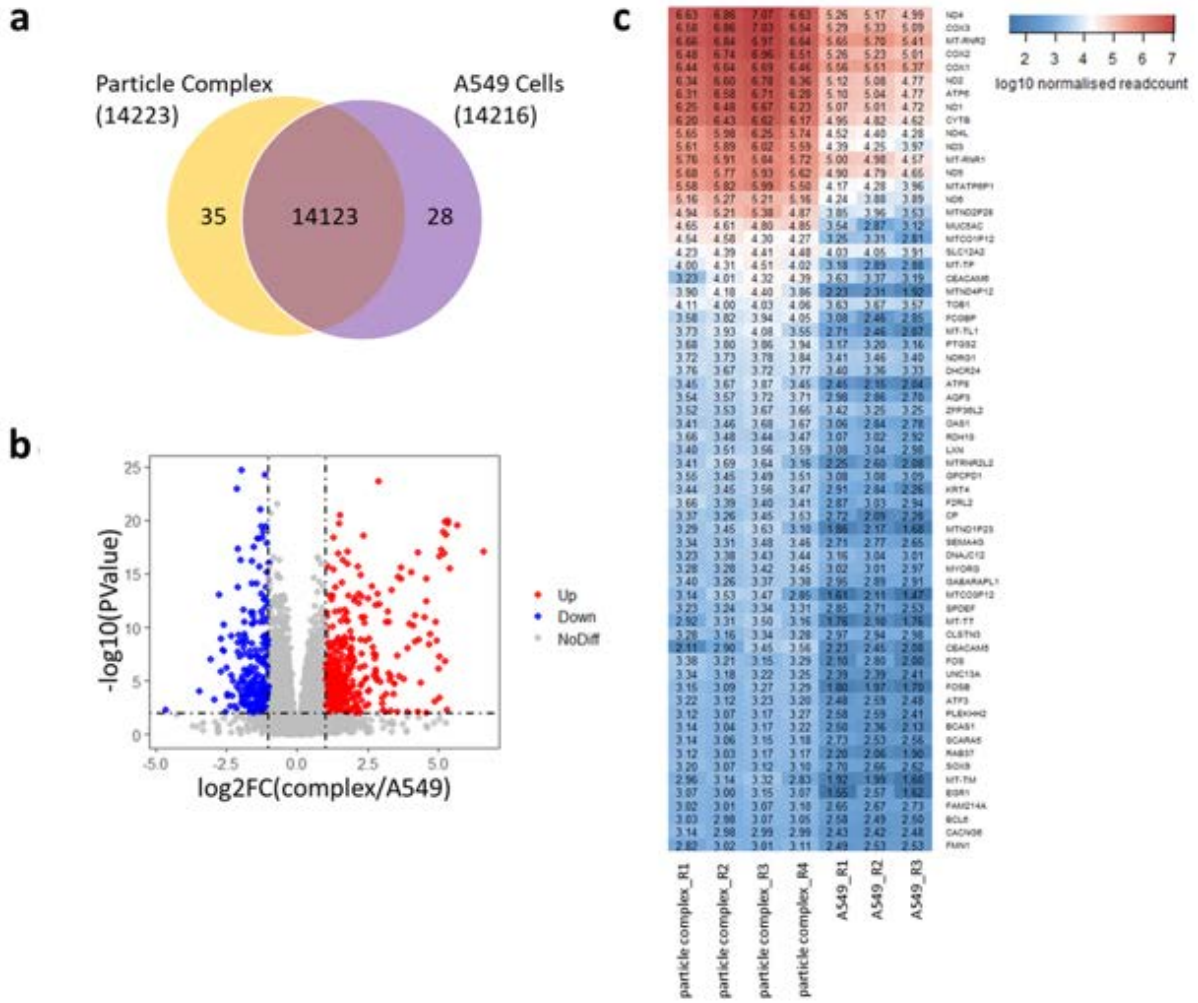


Figure S4. 20 Analysis of particle complexes and A549 cells mRNA-seq annotated genes.

(a) Venn diagram of particle complex and A549 cells mRNA-seq annotated genes. (b) Volcano plot of particle complex vs. A549, dashes lines are $log_2FC = \pm 1$, P-value = 0.01. (c) Heatmap of the top 64 abundant and enriched genes in particle complex.

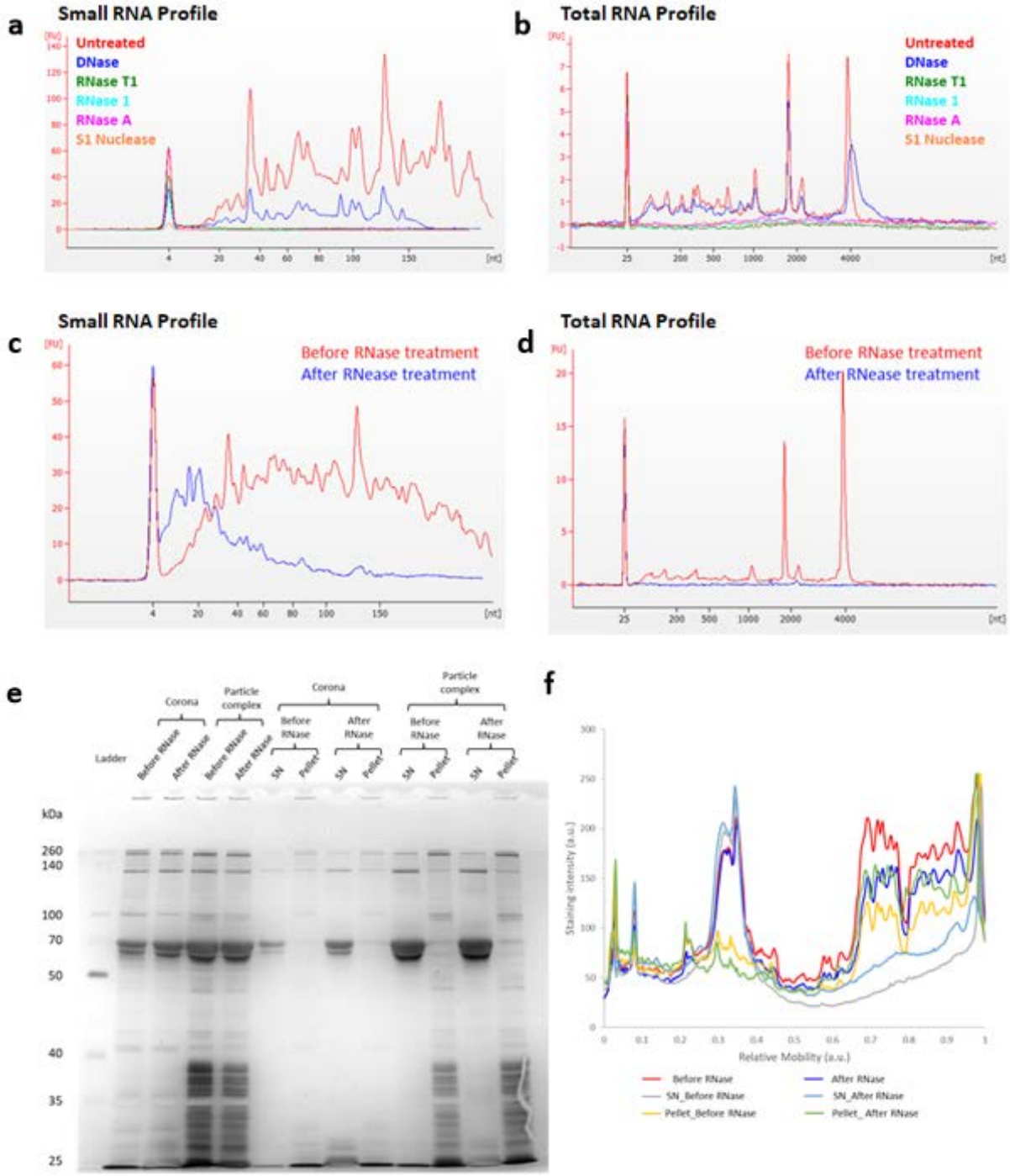


Figure S4. 21 RNA footprinting of particle complex.

(a) Small and (b) total RNA profiles of extracted particle complex RNA treated with various nucleases to confirm purity of RNA, analyzed using the Agilent 2100 Bioanalyzer. (c) Small and (d) total RNA profiles are completely degraded in particle complexes treated in situ with RNase, suggesting the RNA is accessible on the surface of the biomolecular layer. (e) Protein profile of particle complexes treated in situ with RNase with associated densitometry analysis (f) of the silver stained SDS-PAGE gel (10%, 20 µg of particles per condition), confirming there

is no disruption to the protein composition of the biomolecular layer with degradation of the RNA. The intense band visible in the supernatants (SN) is attributed to free BSA coming from a residual contamination by the FBS.

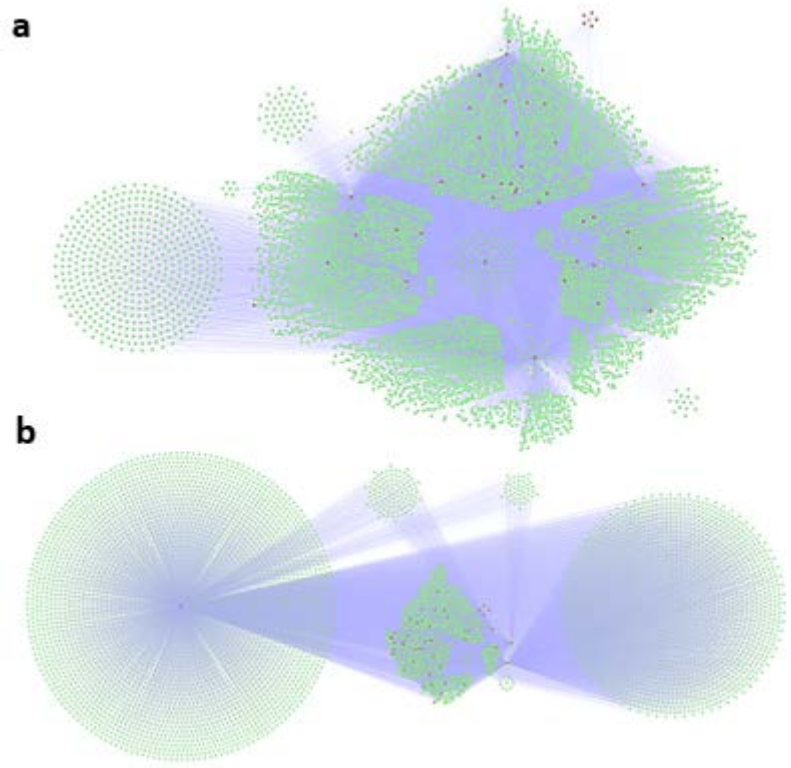


Figure S4. 22 Protein-RNA network.

(a) Proteins come from Heavy labelled particle complex. (b) Proteins come from the label-free particle complex. The network was referred to the RNAInter database. Proteins with high abundance and also recorded in the Granule database will be selected for network plot. Red is protein and Green is RNA (Count > 1000), and the blue line is the link between protein and RNA.

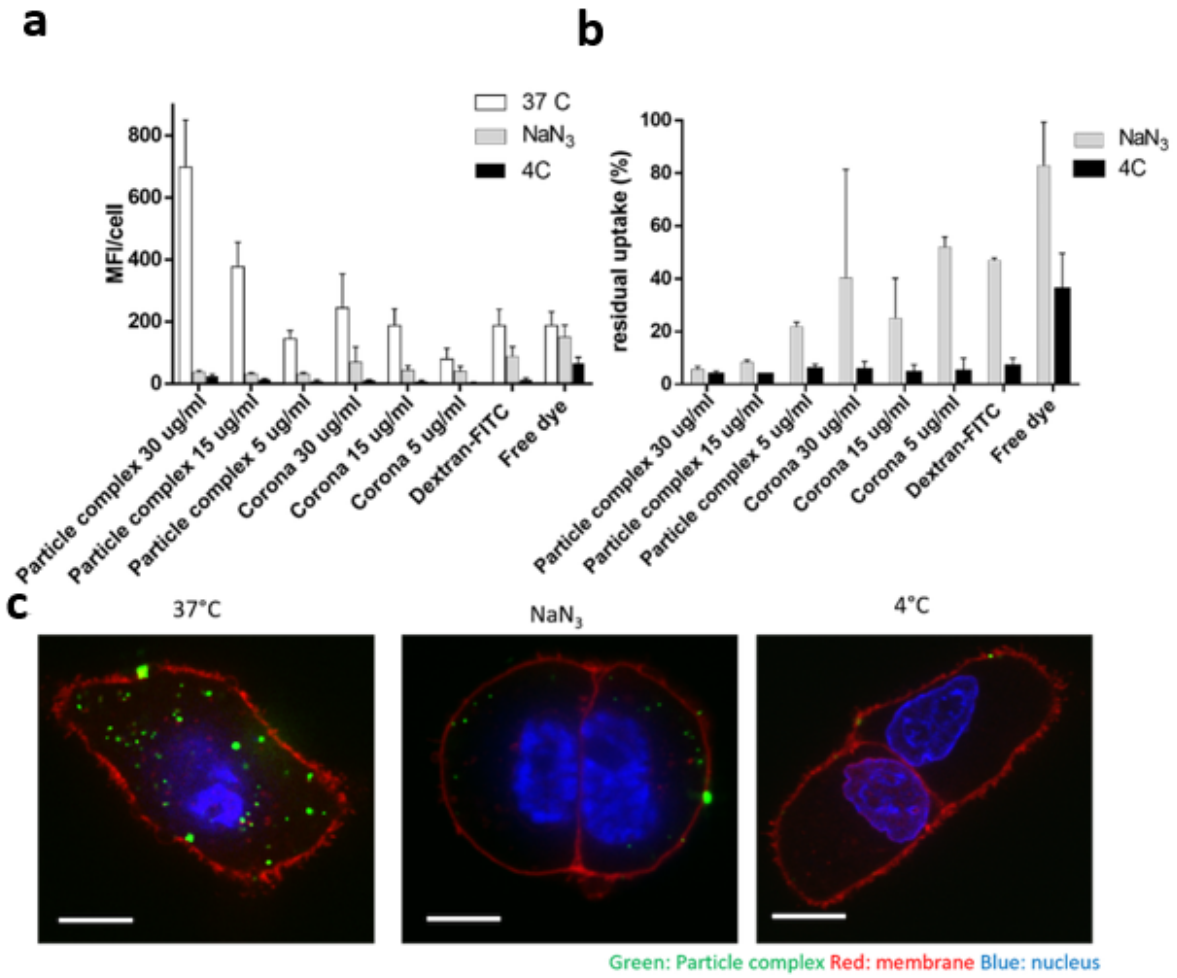


Figure S4. 23 The particle complex uptake is receptor mediated and energy dependent.

The high content analysis (HCA) experiment was conducted on cells depleted in ATP by treatment with sodium azide, and the uptake of particle complex or corona-NPs was compared to the one of cell cultured in regular cMEM (no depletion). The HCA analysis revealed that the uptake is greatly inhibited by the sodium azide treatment (a, b); the results were validated by confocal imaging, showing that only few particles were internalized by the NaN_3 treated cells (c, central image). As a positive control for the inhibition, a set of cells were incubated at 4°C during the particle exposure, and neither the particle complex nor the corona-NPs were internalized, but some particles are stuck on the cell membrane. (n=3, P-value < 0.005).

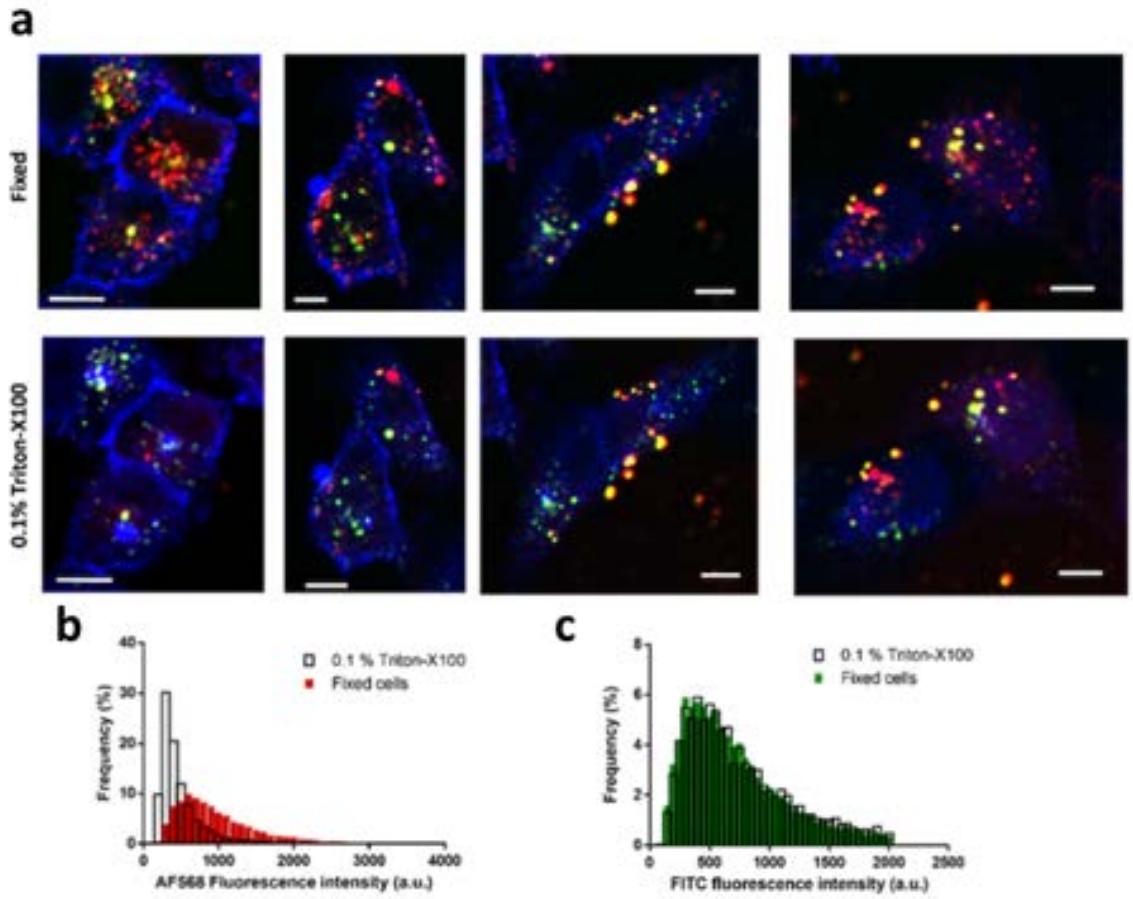


Figure S4. 24 The particle complex biomolecular cargo is dissolved by detergents.

HCA images (a) and analysis (b, c) showing cells treated with 100 µg/ml AF568_particle complex after fixation and the comparison with the same cells after subsequent permeabilization. The standard permeabilization (0.1% Triton-X100, 10 mins) dissolved the vast majority of particle complex cargoes after detachment (red, fluorescence intensity quantification in b), and also partially affected the particle complex (appearing yellow). The fluorescence signal of the particle complex core (green, fluorescence intensity quantification in c) is not affected by the treatment. Blue: cell membrane, green: particle complex core, red: particle complex cargo. Scale bar: 10 µm.

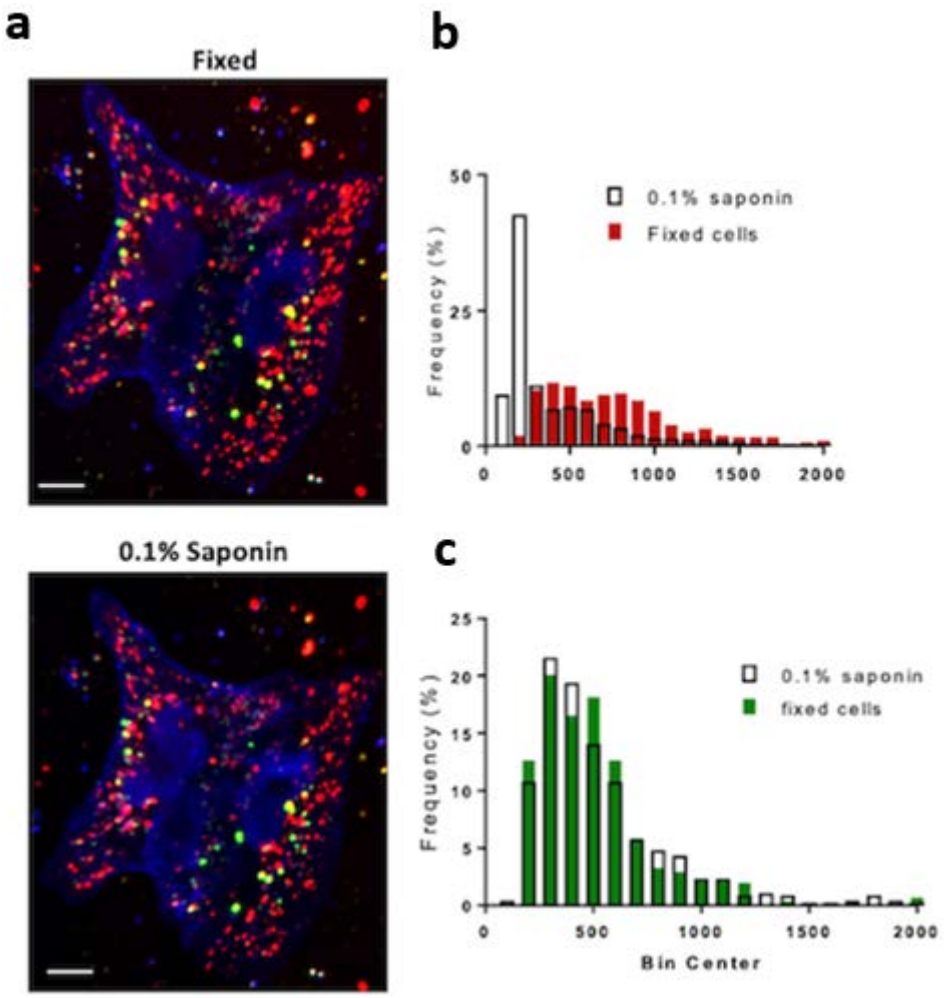


Figure S4. 25 The particle complex biomolecular layer is preserved after permeabilization with saponin.

HCA images (a) and single cell analysis (b) of cells treated with 100 µg/ml AF- particle complex after fixation (red bars) or after saponin permeabilization (white bars). The mild permeabilization had a lighter impact on the particle complex cargoes loss. Blue: cell membrane, green: particle complex core, red: particle complex cargo. Scale bar: 10 µm.

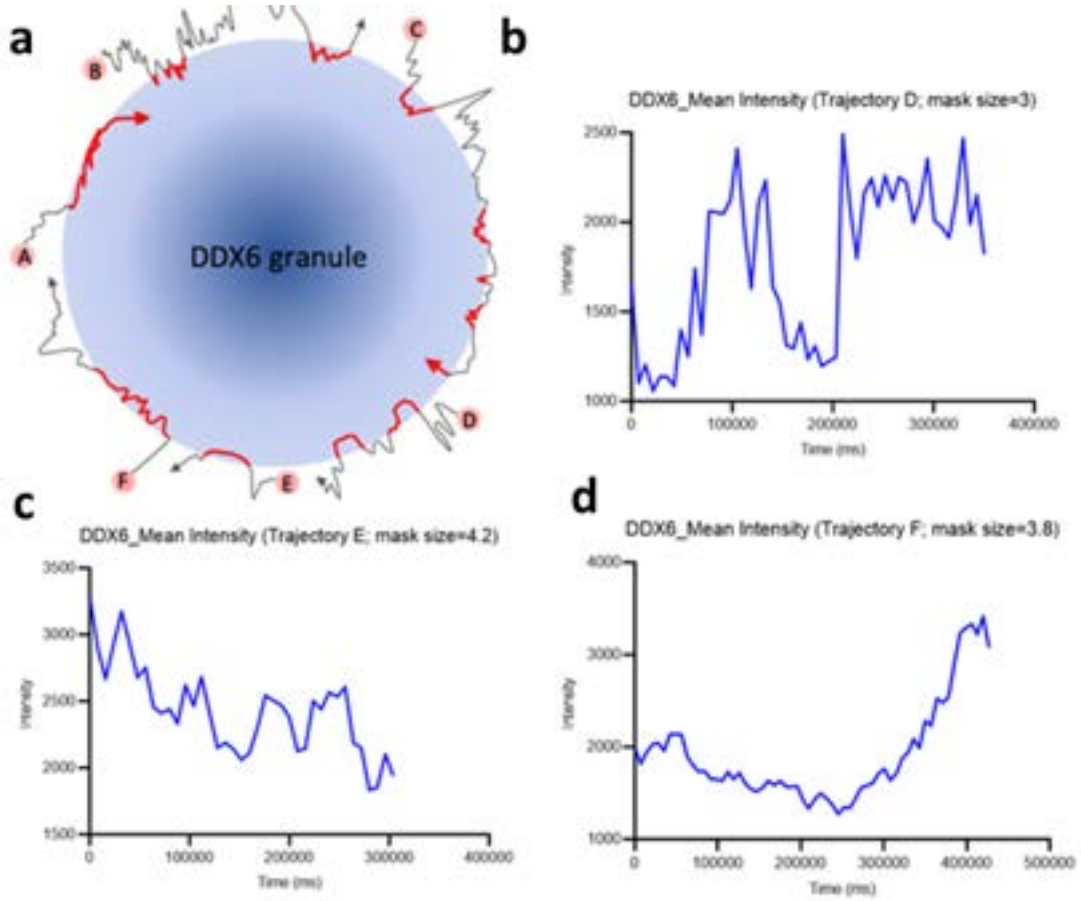


Figure S4. 26 3D mean fluorescence intensity of EGFP-DDX6 over the time course.

The masks were including the entire granules. The mean fluorescence intensity of EGFP-DDX6 was shown to increase or decrease after multiple contacts with particle complex coat, suggesting that molecular exchange between the coat and DDX6 RNA granule.

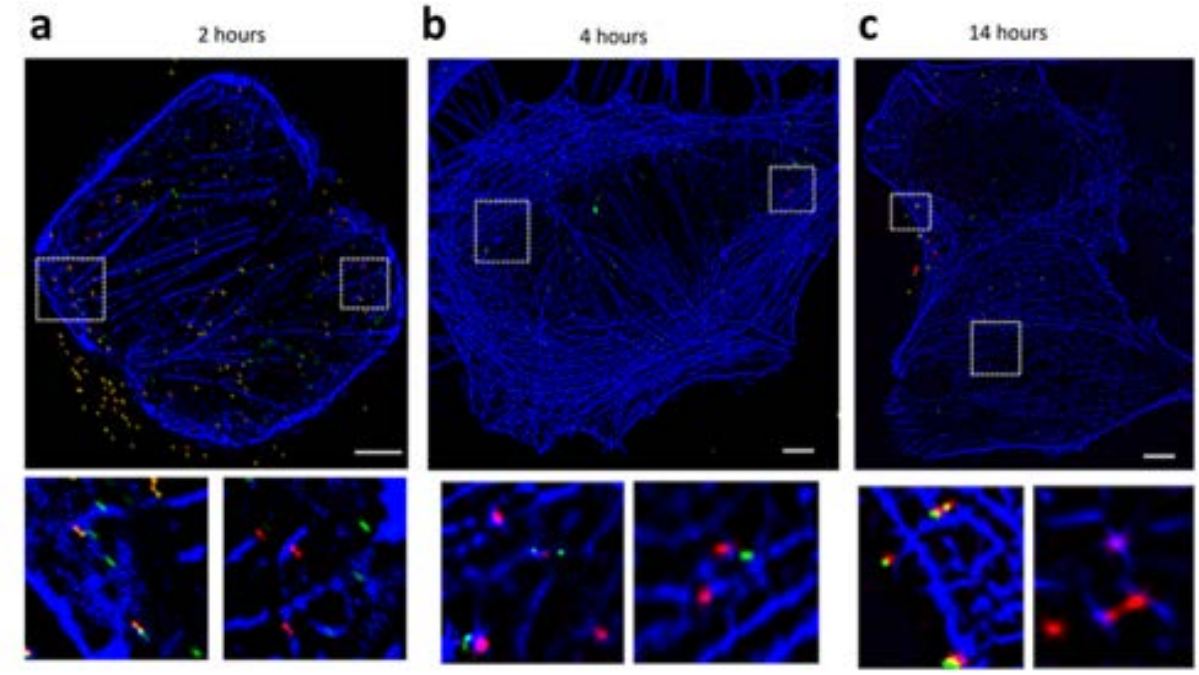


Figure S4. 27 Co-localization analysis of particle complex and actin.

Co-localization analysis of particle complex and actin. A549 were treated with 40 μg/ml of AF-particle complex for 2 hours and then chased for different times. The cells were thein stain with Phalloidin-Texas Red and imaged in super resolution mode (SRRF). The co-localization degree of the particle complexes and the cargo alone is significant within the first 4 chase hours. At late timepoint, 14 hours chase, a limited number of cargoes were observed while interacting with the actin. Blue: actin, green: particle complex core, red: particle complex cargo. Scale bar: 10 μm.

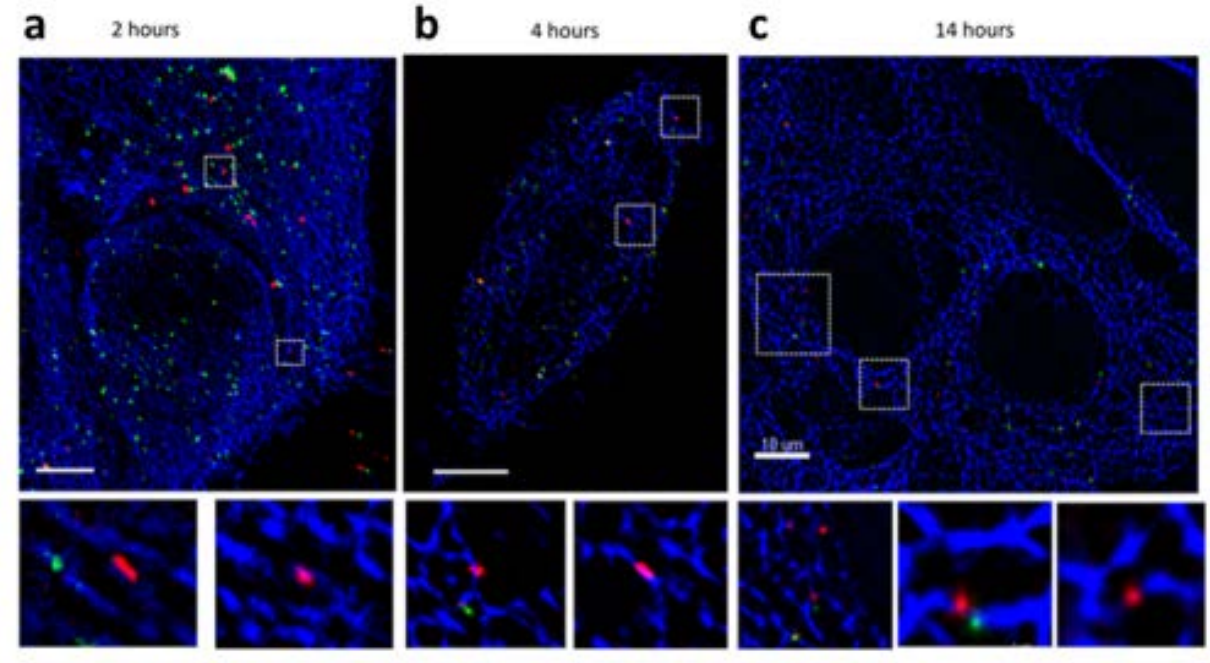


Figure S4. 28 Co-localization analysis of particle complex and tubulin.

A549 were treated with 40 ug/ml of AF- particle complex for 2 hours and then chased for different times (2, 4, 14 hours). The cells were subsequently stained with an anti-tubulin antibody and imaged in super resolution (SRRF). The co-localization degree of the cargo was constant for the first timepoints (2 and 4), as for the core particles. At late timepoint the co-localization degree increased. Blue: tubulin, green: particle complex core, red: particle complex cargo. Scale bar: 10 μm.

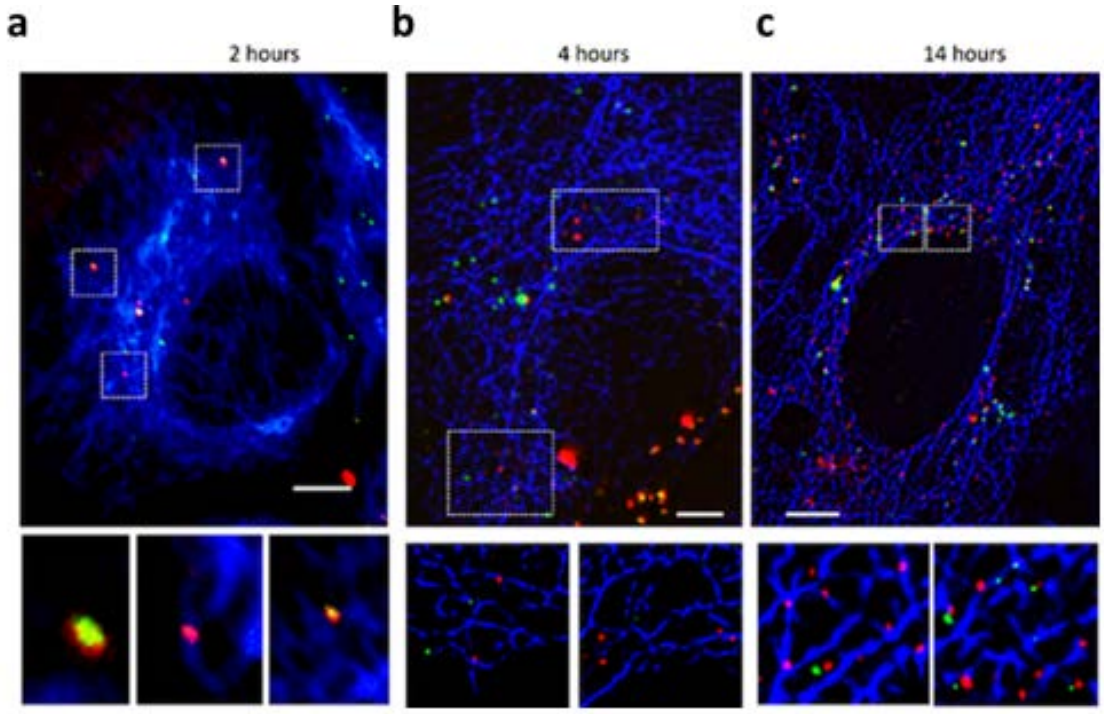


Figure S4. 29 Co-localization analysis of particle complex and vimentin.

A549 were treated with 40 μg/ml of AF-particle complex for 2 hours and then chased for different times (2, 4, 14 hours). (a) after 2 hours chase most of the particle complex complexes did not appear co-localized with the intermediate filaments (IF) if not sporadically; (b) after 4 hours, most of the complexes are split into cargoes and cores, both partially co-localized. (c) After 14 hours chase the localization of the particle complex cargoes appeared to be increased if compared to the previous timepoints. Blue: vimentin, green: particle complex core, red: particle complex cargo. Scale bar: 10 μm. The cells were imaged in SRRF mode.

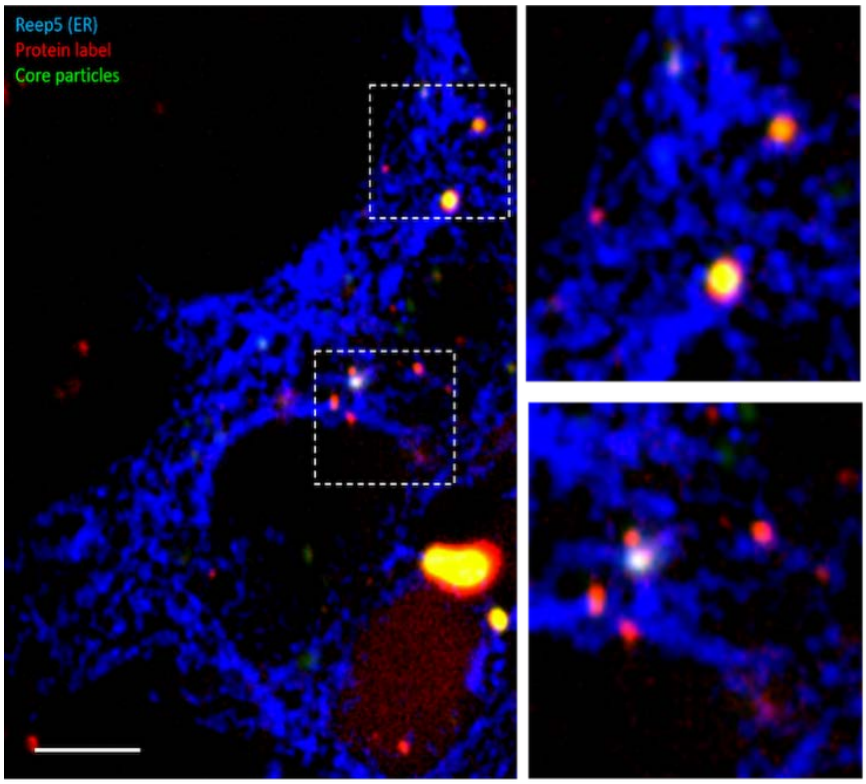


Figure S4. 30 Particle complex co-localization analysis with REEP5, ER marker.

A549 cells were treated with 40 μg/ml for 2 hours and placed in fresh media for further 2 hours. The samples were then imaged with a fully automated microscope, after staining with an anti-Reep5 antibody to visualize the ER structure. Many particle complex cargoes were co-localized with the filaments, while the particle complexes were mainly distributed nearby the structure. Blue: ER, green: particle complex core, red: particle complex cargo. Scale bar: 10 μm. The cells were imaged in SRRF mode.

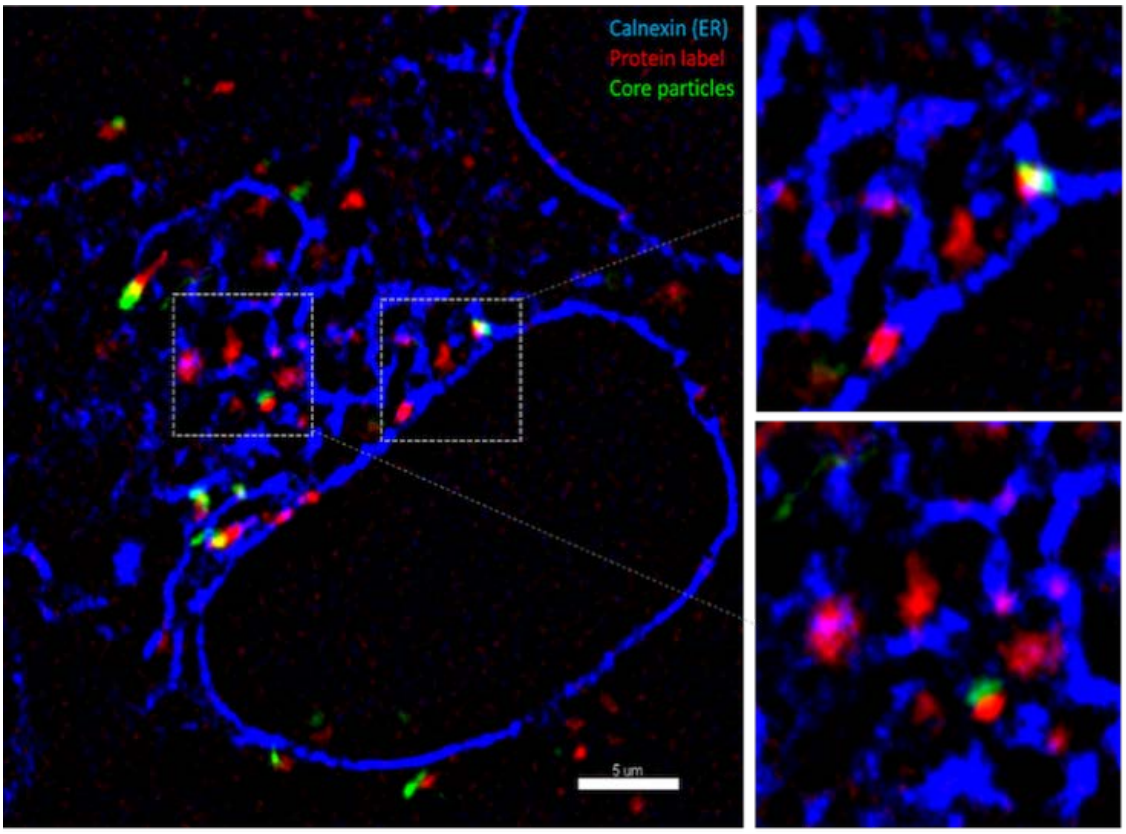


Figure S4. 31 Particle complex co-localization with Calnexin.

EGFP-calnexin HEK293 were exposed to 40 μg/ml AF-particle complex for 2 hours, chased for 2 hours and immediately imaged in live cell imaging, SRRF mode. The super resolution frames showed that the particle complexes were mainly located on the side of the ER structures, probably enclosed into vesicles interacting with the filaments (top inset). The particle complex cargoes were often visualized as enclosed in the ER structure instead (bottom inset). Blue: Calnexin, green: Particle complex core, red: Particle complex cargo. Scale bar: 5 μm. The cells were imaged in SRRF mode.

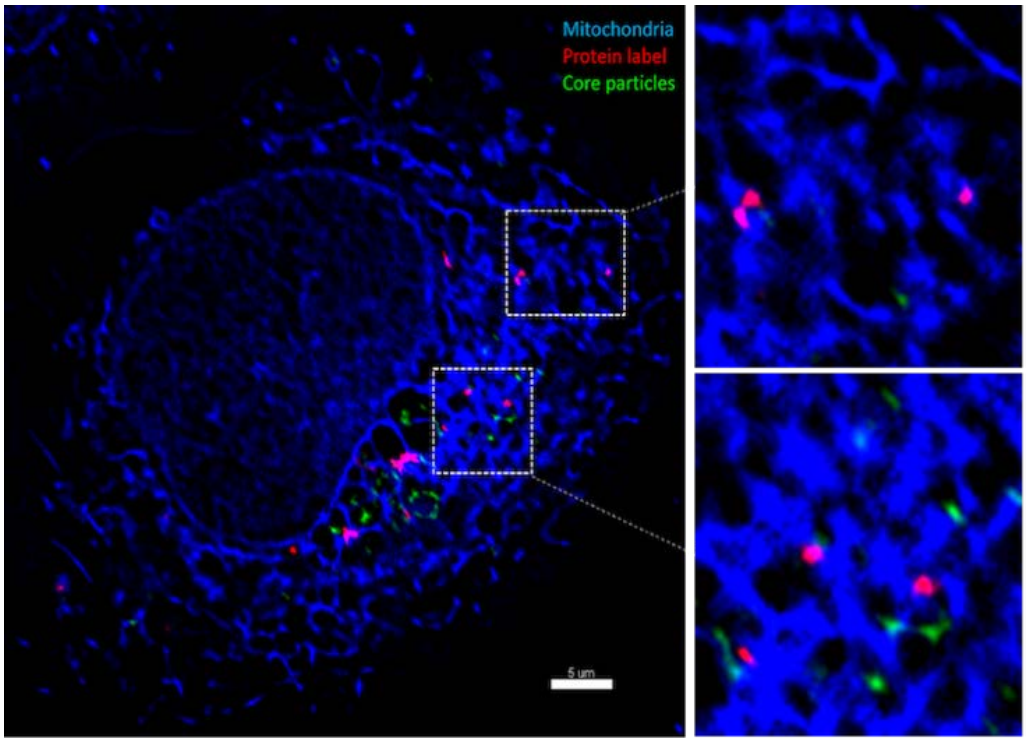


Figure S4. 32 Particle complex co-localization analysis with mitochondrial outer membrane.

A549 cells were treated with AF-particle complex for 2 hours and chased for 14, to allow the particle complexes to split and to be fully trafficked in the intracellular environment. The mitochondria were visualized with anti-TOMM20 antibody (mitochondrial outer membrane) to avoid harsh permeabilization. The particle complex cargoes resulted to be interacting with the structure, while the particle complex cores were found nearby, but not co-localized. Blue: mitochondria, green: Particle complex core, red: Particle complex cargo. Scale bar: 5 μm. The cells were imaged in SRRF mode.

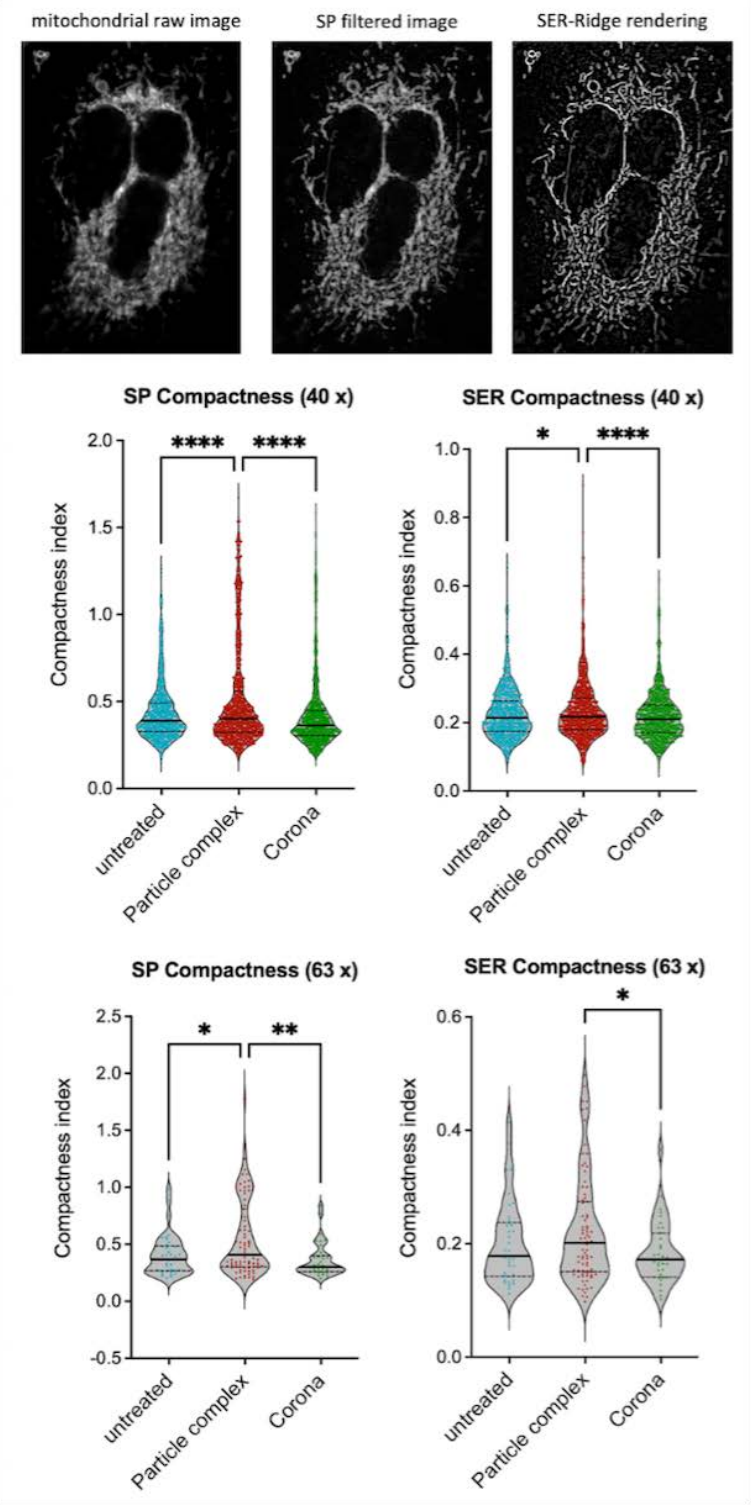


Figure S4. 33 Particle complex induced mitochondrial morphology change HCA analysis. A549 cells were treated with 200 μg/ml of particle complexes or corona-NPs for 6 hours. The cells were fixed and stained with an anti-TOMM20 antibody to visualize the mitochondria and imaged with an Opera Phenix HCA. The morphological analysis was showed significant variation in the mitochondria compactness index for the cells treated with the particle complexes when compared to untreated cells or corona-NPs treated cells.

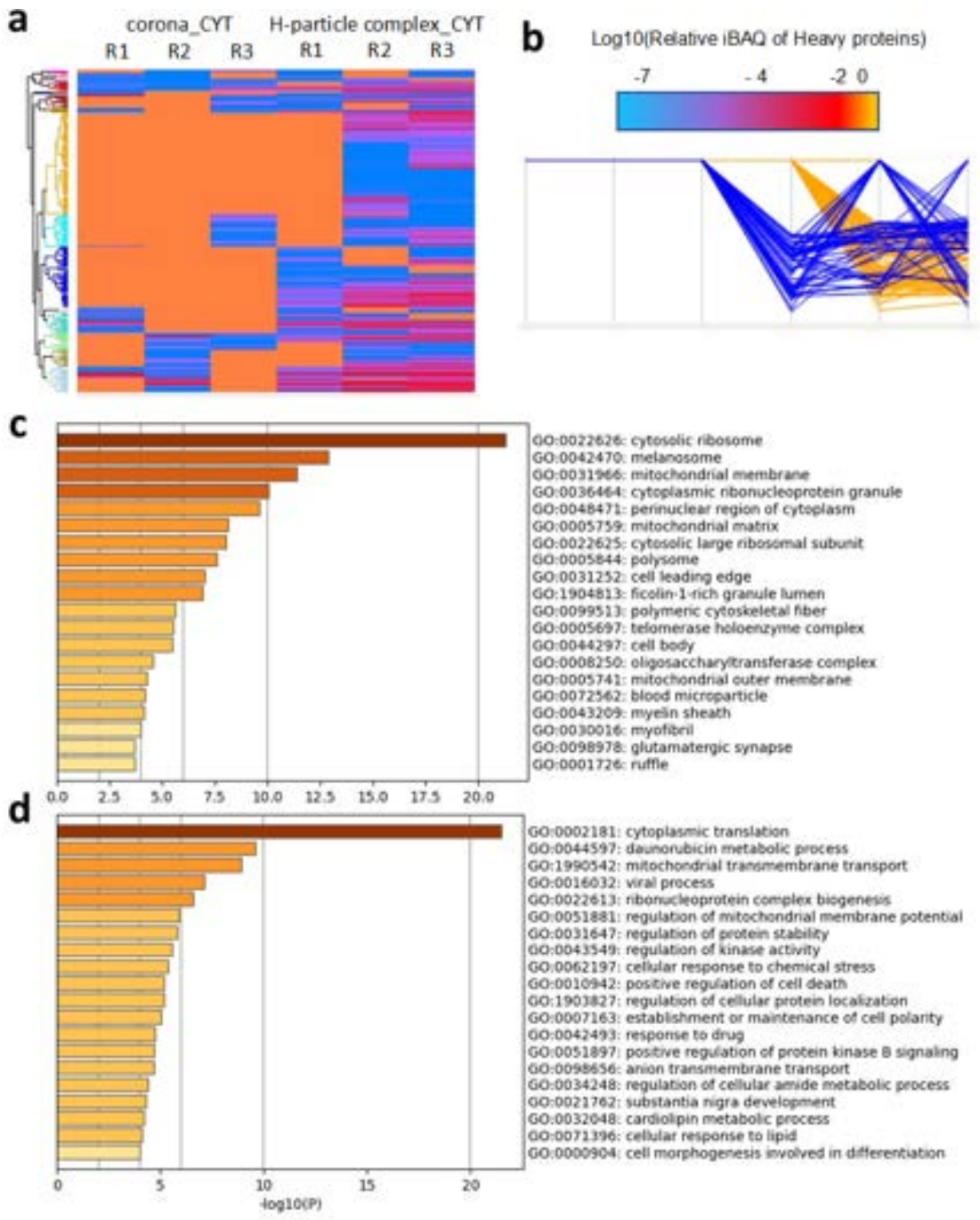


Figure S4. 34 Proteomics analysis of corona and particle complex_CYT.

(a) Heatmap of heavy proteins in Heavy labeled particle complex_CYT and corona_CYT. The color gradient is digitalized against the relative iBAQ Intensity of heavy proteins. The Orange represents proteins were not identified correspondingly. (b) Heavy protein clusters only identified in H-particle complex_CYT. Each color represents one cluster which is consistent with heatmap. Cellular Component (c) and Biological Process (d) of unique heavy proteins in H-particle complex_CYT by Metascape.

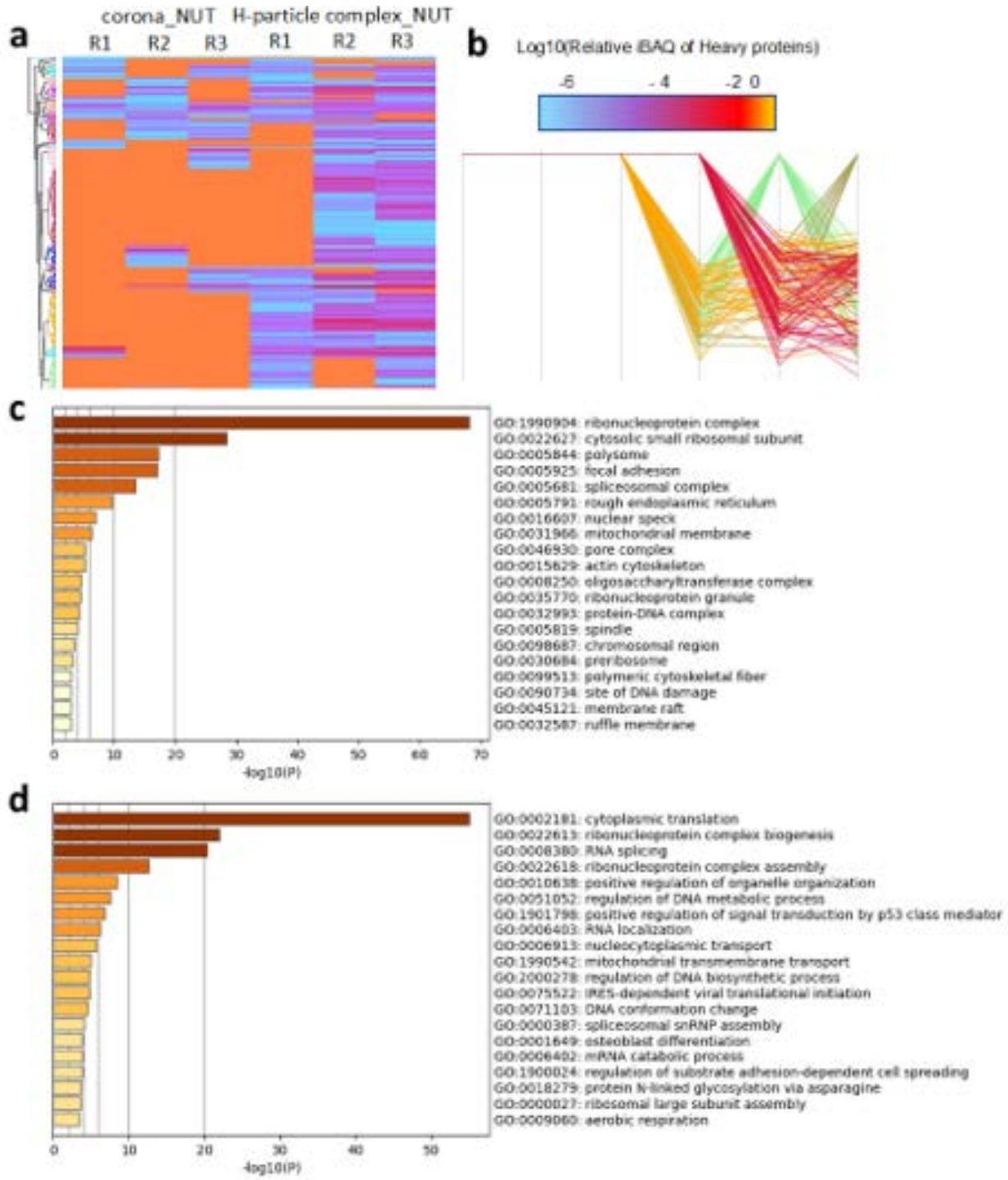


Figure S4. 35 Proteomics analysis of corona and particle complex_NUT.

(a) Heatmap of heavy proteins in Heavy labelled particle complex_NUT and corona_NUT. The color gradient is digitalized against the relative iBAQ Intensity of heavy proteins. The Orange represents proteins were not identified correspondingly. (b) Heavy protein clusters only identified in Heavy labelled Particle Complex _NUT. Each color represents one cluster which is consistent with heatmap. Cellular Component (c) and Biological Process (d) of unique heavy proteins in Heavy labelled Particle Complex _NUT by Metascape.

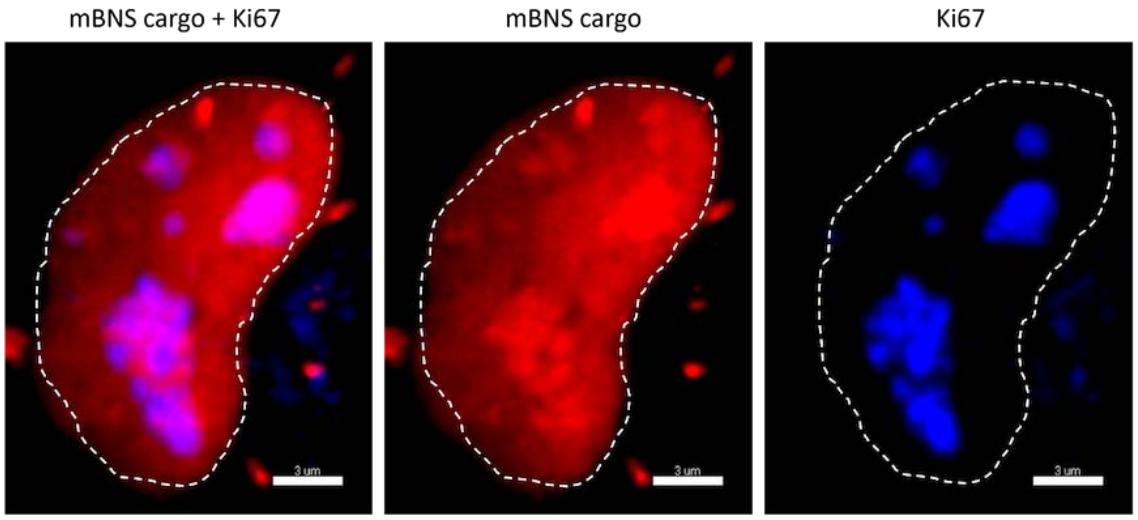


Figure S4. 36 The nuclear delivery pattern is denser in the nucleolar area.

In part of the cells displaying the particle complex nuclear delivery patten (red), some denser areas corresponding to the cell nucleolus (blue) can be identified. The distribution of the two signals appeared specular within the nucleolus. Scale bar: 3 µm.

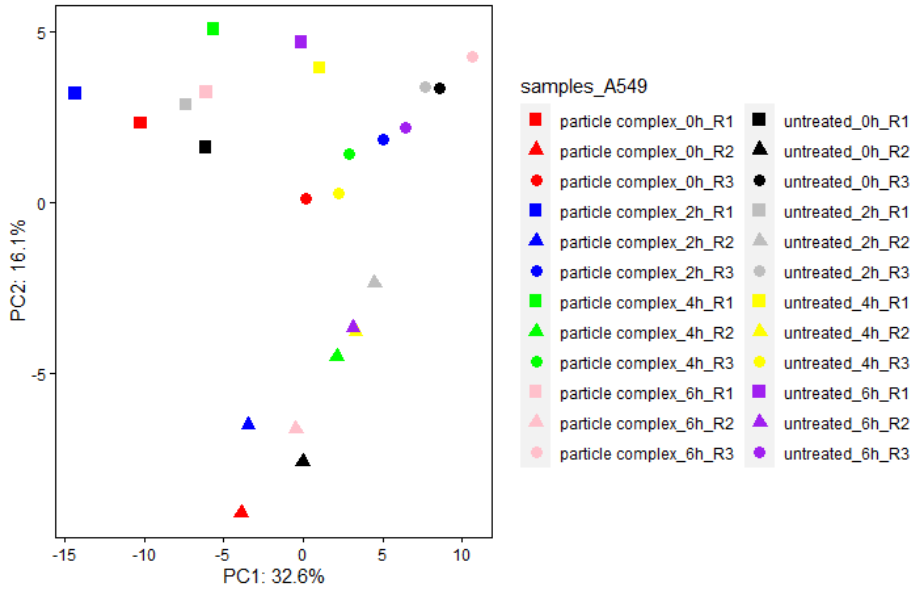


Figure S4. 37 PCA analysis for the transcriptome of particle complex treated and untreated A549 cells.

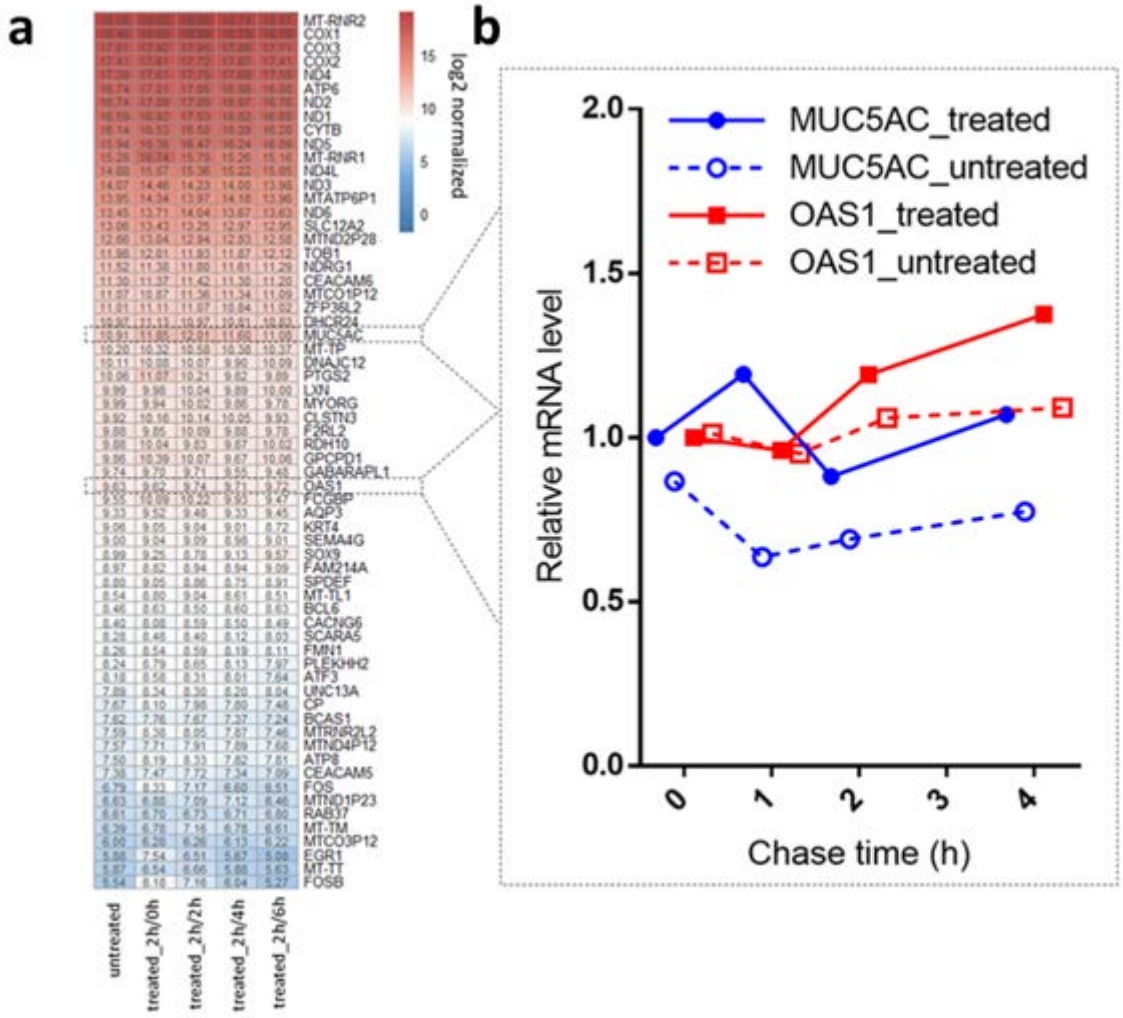


Figure S4. 38 Relative mRNA expression level of particle complex treated and untreated A549 cells.

(a) Heatmap of the 64 significantly enriched and abundant mRNAs (colored in Figure 4.2F) in particle complex treated and untreated A549 cells, data from mRNA-seq. (b) The abundance of MUC5AC and OAS1 mRNAs were validated by RT-qPCR.

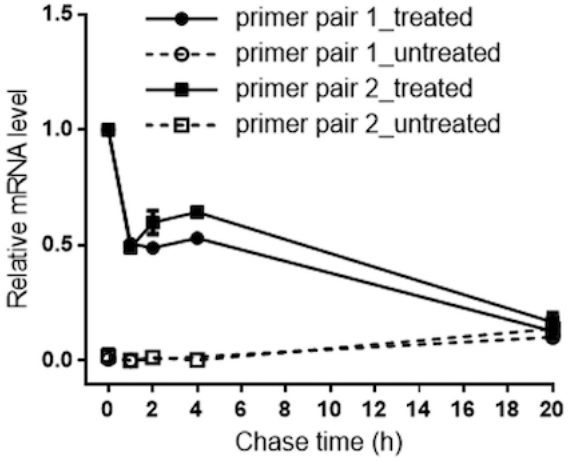


Figure S4. 39 RT-qPCR results for the overnight pull-down particle complex treated HEK cells.

Particle complex was produced from EGFP-FUS transfected HEK cells with overnight pull down at 4 °C by adding RNase inhibitor to prevent the degradation of RNA. RNase inhibitor was removed by spinning down at 12,000 rcf for 15 mins at 4 °C before adding the particle complex to the recipient HEK cells.

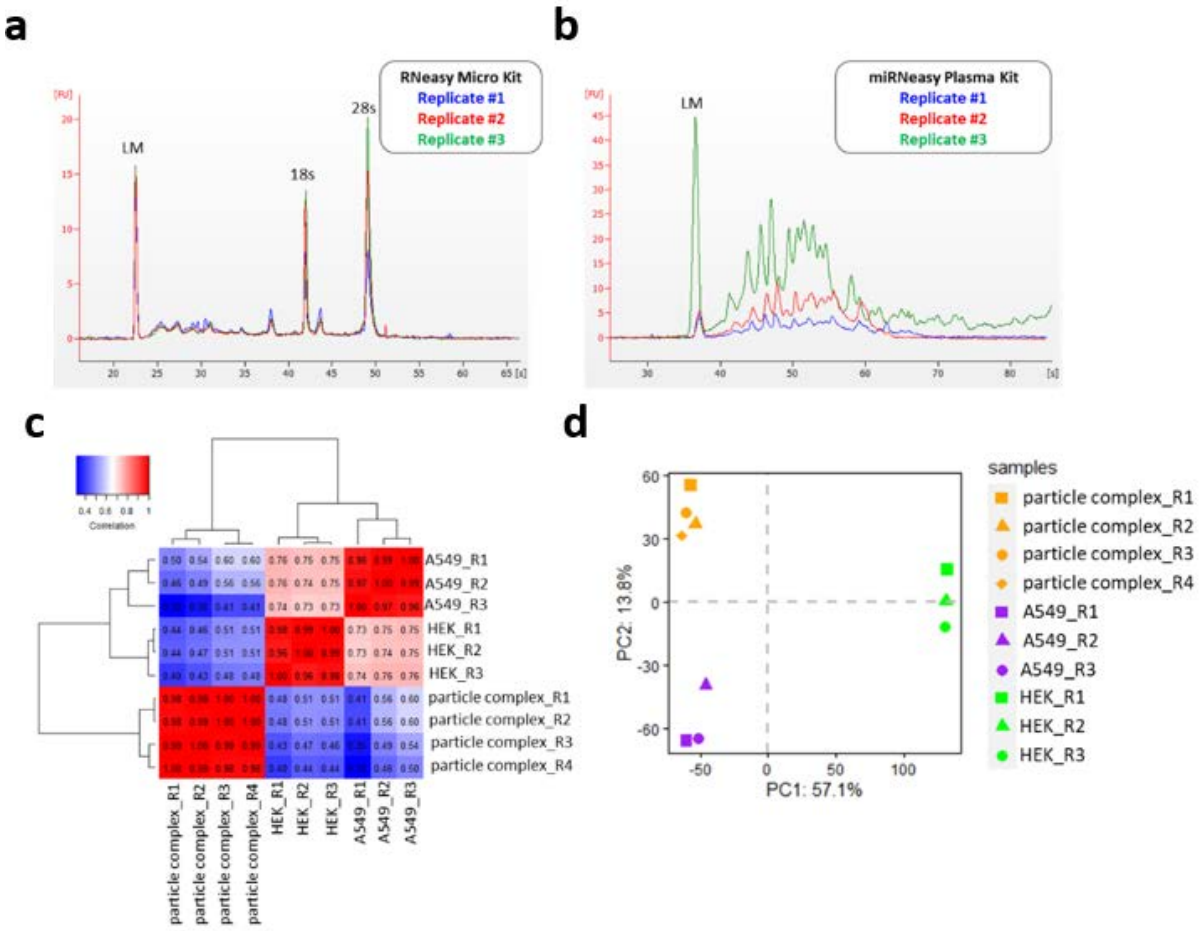


Figure S4. 40 Reproducibility of RNA extraction and RNA-seq.

(a-b) Bioanalyzer results of three biological replicates of the RNA extracted from particle complexes by RNeasy Micro Kit (a) and miRNeasy Plasma Kit (b) operated performed by two independent operators. (b) Correlation and (c) PCA plot of the mRNA-seq result for A549 cells, HEK cells, and particle complexes.

Gene	Forward Primer	Reverse Primer
CEACAM6	ACGATGATCACAGTCTCTGGAAG	AGCACTCCAATCGTGATGCC
KRT4	TCCTTCATCGACAAGGTGCAG	GGGCTCAAGGTTTTTGCTGG
MUC5AC	CCTTCGACGGACAGAGCTAC	TCTCGGTGACAACACGAAAG
OAS1	AGGAAAGGTGCTTCCGAGGTAG	GGACTGAGGAAGACAACCAGGT
EGFP FUS primer pair 1	CTATGATCGAGGCGGCTACC	CGGTGGCTTATCGTCGTCAT
EGFP FUS primer pair 2	CGACGTAAACGGCCACAAG	GACTTGAAGAAGTCGTGCTGC
GAPDH	TGCACCACCAACTGCTTAGC	GGCATGGACTGTGGTCATGAG

Table 4. 1 RT-qPCR primer sequences.

References

1. Vance, J. E. Phospholipid synthesis in a membrane fraction associated with mitochondria. *J Biol Chem* **1990,** 265, 7248-7256.
2. Prinz, W. A., Toulmay, A. & Balla, T. The functional universe of membrane contact sites. *Nat Rev Mol Cell Biol* **2020,** 21, 7-24, doi:10.1038/s41580-019-0180-9.
3. Phillips, M. J. & Voeltz, G. K. Structure and function of ER membrane contact sites with other organelles. *Nat Rev Mol Cell Biol* **2016,** 17, 69-82, doi:10.1038/nrm.2015.8.
4. Scorrano, L. *et al.* Coming together to define membrane contact sites. *Nat Commun* **2019,** 10, 1287, doi:10.1038/s41467-019-09253-3.
5. Shai, N. *et al.* Systematic mapping of contact sites reveals tethers and a function for the peroxisome-mitochondria contact. *Nat Commun* **2018,** 9, 1761, doi:10.1038/s41467-018-03957-8.
6. Valm, A. M. *et al.* Applying systems-level spectral imaging and analysis to reveal the organelle interactome. *Nature* **2017,** 546, 162-167, doi:10.1038/nature22369.
7. Lyon, A. S., Peeples, W. B. & Rosen, M. K. A framework for understanding the functions of biomolecular condensates across scales. *Nat Rev Mol Cell Biol* **2021,** 22, 215- 235.
8. Lee, J. E., Cathey, P. I., Wu, H., Parker, R. & Voeltz, G. K. Endoplasmic reticulum contact sites regulate the dynamics of m neless organelles. *Science* **2020,** 367.
9. Cohen, S., Valm, A. M. & Lippincott-Schwartz, J. Interacting organelles. *Curr Opin Cell Biol* **2018,** 53, 84-91.
10. Bowie, A. G. & Unterholzner, L. Viral eva ion and subversion of pattern-recognition receptor signalling. *Nat Rev Immunol* **2008,** 8 911-922.
11. Kouhestani, D. *et al.* Variant signaling topology at the cancer cell–T-cell interface induced by a two-compone T-cell engager. *Cellular & Molecular Immunology* **2020,** 18, 1568-1570, doi:10.1038/s41423-020-0507-7.
12. Dawson, K. A. & Yan, Y Current understanding of biological identity at the nanoscale and future prospects. *Nat Nanotechnol* **2021,** 16, 229-242, doi:10.1038/s41565-021-00860-0.
13. Monopoli, M. P., Aberg, C., Salvati, A. & Dawson, K. A. Biomolecular coronas provide the biological identity of nanosized materials. *Nat Nanotechnol* **2012,** 7, 779-786.
14. Para W. J., Nel, A. E. & Weiss, P. S. Grand Challenges for Nanoscience and Nanote hnology. *ACS Nano* **2015,** 9, 6637-6640.
15. Sahay, G. *et al.* Efficiency of siRNA delivery by lipid nanoparticles is limited by endocytic recycling. *Nat Biotechnol* **2013,** 31, 653-U119.
16. Balint, S. *et al.* Supramolecular attack particles are autonomous killing entities released from cytotoxic T cells. *Science* **2020,** 368, 897-901.
17. Sandin, P., Fitzpatrick, L. W., Simpson, J. C. & Dawson, K. A. High-speed imaging of Rab family small GTPases reveals rare events in nanoparticle trafficking in living cells. *ACS Nano* **2012,** 6, 1513-1521.
18. Gilleron, J. *et al.* Image-based analysis of lipid nanoparticle-mediated siRNA delivery, intracellular trafficking and endosomal escape. *Nat Biotechnol* **2013,** 31, 638-U102.
19. Hoshino, A. *et al.* Tumour exosome integrins determine organotropic metastasis. *Nature* **2015,** 527, 329-+.
20. Luga, V. *et al.* Exosomes Mediate Stromal Mobilization of Autocrine Wnt-PCP Signaling in Breast Cancer Cell Migration. *Cell* **2012,** 151, 1542-1556.
21. García-Silva, S. *et al.* Melanoma-derived small extracellular vesicles induce lymphangiogenesis and metastasis through an NGFR-dependent mechanism. *Nature Cancer* **2021.**
22. Moller, A. & Lobb, R. J. The evolving translational potential of small extracellular vesicles in cancer. *Nat Rev Cancer* **2020,** 20, 697-709.

23. Nieuwland, R., Falcón-Pérez, J. M., Théry, C. & Witwer, K. W. Rigor and standardization of extracellular vesicle research: Paving the road towards robustness. *Journal of Extracellular Vesicles* **2020,** 10.
24. Takahashi, A. *et al.* Exosomes maintain cellular homeostasis by excreting harmful DNA from cells. *Nature Communications* **2017,** 8.
25. Haimovich, G. *et al.* Intercellular mRNA trafficking via membrane nanotube-like extensions in mammalian cells. *P Natl Acad Sci USA* **2017,** 114, E9873-E9882.
26. Gruner, H. N. & McManus, M. T. Examining the evidence for extracellular RNA function in mammals. *Nat Rev Genet* **2021,** 22, 448-458.
27. Raposo, G. & Stahl, P. D. Extracellular vesicles: a new communication paradigm? *Nat Rev Mol Cell Bio* **2019,** 20, 509-510.
28. Tosar, J. P. *et al.* Assessment of small RNA sorting into different extracellular fractions revealed by high-throughput sequencing of breast cell lines. *Nucleic Acids Res* **2015,** 43, 5601-5616.
29. Tosar, J. P. *et al.* Dimerization confers increased stability to nucleases in 5 ' halves from glycine and glutamic acid RNAs. *Nucleic Acids Res* **2018,** 46, 9081-9093.
30. Tosar, J. P. *et al.* Fragmentation of extracellular ribosomes and tRNAs shapes the extracellular RNAome. *Nucleic Acids Res* **2020,** 48, 12874-12888.
31. Hinde, E. *et al.* Pair correlation microscopy reveals the role of nanoparticle shape in intracellular transport and site of drug release. *Nat Nanotechnol* **2017,** 12, 81-89.
32. Maugeri, M. *et al.* Linkage between endosomal escape of LNP-mRNA and loading into EVs for transport to other cells. *Nature Communications* **2019,** 10.
33. Cursi, L. et al. Multifunctional superparamagnetic nanoparticles with a fluorescent silica shell for the in vitro study of bio-nano interactions at the subcellular scale. *Nanoscale* **2021,** 13, 16324-16338.
34. Monopoli, M. P. *et al.* Physical-Chemical Aspects of Protein Corona: Relevance to in Vitro and in Vivo Biological Impacts of Nanoparticles. *J Am Chem Soc* **2011,** 133, 2525- 2534.
35. Thul, P. J. *et al.* A subcellular map of the human proteome. *Science* **2017,** 356.
36. Youn, J.-Y. *et al.* Properties of Stress Granule and P-Body Proteomes. *Mol Cell* **2019,** 76, 286-294.
37. Kugeratski, F. G. *et al.* Quantitative proteomics identifies the core proteome of exosomes with syntenin-1 as the highest abundant protein and a putative universal biomarker. *Nat Cell Biol* **2021,** 23, 631-641.
38. Shemesh, N. *et al.* The landscape of molecular chaperones across human tissues reveals a layered architecture of core and variable chaperones. *Nature Communications* **2021,** 12.
39. Yu, H. Y. *et al.* HSP70 chaperones RNA-free TDP-43 into anisotropic intranuclear liquid spherical shells. *Science* **2021,** 371, 585-+.
40. Lareau, L. F., Hite, D. H., Hogan, G. J. & Brown, P. O. Distinct stages of the translation elongation cycle revealed by sequencing ribosome-protected mRNA fragments. *Elife* **2014,** 3, e01257.
41. Alberti, S. & Hyman, A. A. Biomolecular condensates at the nexus of cellula stress, protein aggregation disease and ageing. *Nat Rev Mol Cell Bio* **2021,** 22, 196-213.
42. Pitchiaya, S. *et al.* Dynamic Recruitment of Single RNAs o Processing Bodies Depends on RNA Functionality. *Mol Cell* **2019,** 74, 521-+.
43. Shin, Y. & Brangwynne, C. P. Liquid phase condensation in cell physiology and disease. *Science* **2017,** 357.
44. Bertoli, F., Garry, D., Monopoli, M. P., Salvati, A. & Dawson, K. A. The Intracellular Destiny of the Protein Corona: A St on its Cellular Internalization and Evolution. *ACS Nano* **2016,** 10, 10471-10479.
45. Klinge, S. & Woolford, J. L. Ribosom assembly coming into focus. *Nat Rev Mol Cell Bio* **2018,** 20, 116-131.

46. Matera, A. G. & Wang, Z. A day i the life of the spliceosome. *Nat Rev Mol Cell Biol* **2014,** 15, 108-121.
47. Hou, X. C., Zaks, T., Langer, R. Dong, Y. Z. Lipid nanoparticles for mRNA delivery. *Nat Rev Mater* **2021**.
48. Poon, W., Kingston, B. R., Ouyang, B., Ngo, W. & Chan, W. C. W. A framework for designing de very systems. *Nat Nanotechnol* **2020,** 15, 819-829.
49. Ong, S.-E. & Mann, M. A practical recipe for stable isotope labeling by amino acids in cell ulture (SILAC). *Nature Protocols* **2007,** 1, 2650-2660.
50. Rosner, M., Schipany, K. & Hengstschläger, M. Merging high-quality biochemical fractionation with a refined flow cytometry approach to monitor nucleocytoplasmic protein expression throughout the unperturbed mammalian cell cycle. *Nature Protocols* **2013,** 8, 602-626.
51. Coleman, O., Henry, M., Clynes, M. & Meleady, P. in Heterologous Protein Production in CHO Cells. *Methods in Molecular Biology Ch. Chapter* **2017,** 12, 187-194.
52. Heberle, H., Meirelles, G. V., da Silva, F. R., Telles, G. P. & Minghim, R. InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. *Bmc Bioinformatics* **2015,** 16, 169.
53. Zhou, Y. *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nature Communications* **2019,** 10.
54. Thomas, P. D. *et al.* PANTHER: a library of protein families and subfamilies indexed by function. *Genome Res* **2003,** 13, 2129-2141.
55. Gu, Z., Gu, L., Eils, R., Schlesner, M. & Brors, B. circlize implements and enhances circular visualization in R. *Bioinformatics* **2014,** 30, 2811-2812.
56. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* **2019,** 37, 907-+.
57. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **2014,** 30, 923-930.
58. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **2009,** 26, 139- 140.
59. McCarthy, D. J., Chen, Y. & Smyth, G. K. Differential expressi analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res* **2012,** 40, 4288-4297.
60. Wickham, H. ggplot2: Elegant Graphics for Data Analys . *Use R* **2009**, 1-212.
61. Heberle, H., Meirelles, G. V., da Silva, F. R., Telles, G. P. & Minghim, R. InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinformatics* **2015,** 16.
62. Rozowsky, J. *et al.* exceRpt: A Comprehensive Analytic Platform for Extracellular RNA Profiling. *Cell Systems* **2019,** 8.

Chapter V

Conclusion and Perspectives

Conclusion

With the development of nanotechnologies, we are now confronted by a vast, almost "unlimited" universe of nanoscale shapes ensembles, among which some of the nanostructures cannot be simply identified by only few parameters. Gaining more depth of knowledge regarding the nanoscale shapes regulated biological pathways becomes imperative not only for safety purposes, but also for the design of better nanocarriers for medical treatments. To ensure complex nanostructures especially different shape ensembles can be meaningfully reproduced and communicated, an imaging-based computational method to capture, analyse, store and process objectively structures in digital form for the categorization of nanoscale shapes was proposed.

Based on the computational shape analysis method, a framework for systematic study of the biological effects of nanoscale shape ensembles was established, which was described in chapter II. The first part of the framework is a novel microfluidic reactor-based synthesis platform, which can achieve high reproducibility, narrow mono-dispersity, well-defined tunability, various coating ability synthesis of GNP library. Secondly, combining the computational shape analysis method with traditional physiochemical characterization methods, GNPs were well identified and classified. Then selected GNPs with high quality and distinguish shape populations were used for biological studies both *in vitro* and *in vivo* to link the shape trajectory to their biological impacts. Finally, the relationship between the shape trajectory and biological impacts can further guide the synthesis of novel nano-shapes to achieve specific therapeutic purpose.

While working on the establishment of the above framework described in chapter II, a study of nano-shape effects on the epigenome was performed at the same time. In chapter III, the benchtop synthesis method was used to synthesis two distinctive nano-shapes for the study. For the synthesis part, although the benchtop synthesis method is not as stable and reproducible as the microfluidic method, it was the chosen method for this study (microfluidic method was not ready during the epigenome study), several batches were synthesized at the same time and the best quality one was picked up and also loads of efforts were made to ensure all of the batches used in the biological experiments had good reproducibility to avoid batch bias on the biological effects.

It has been highlighted in the thesis that the critical effect of any potential contamination (*e.g.*, LPS) on the biological outcomes and this could make the results difficult to be trusted. Therefore, an endotoxin-free synthesis platform was established during this period and applied to all the GNPs used in the biological studies. Moreover, the biological stability, surface

area, surface charge and so on of GNPs was thoroughly checked prior to the biological experiment. With a high standard control of the synthesis and characterization of GNPs, we found that there is a shape-dependent histone modifications for GNPs treated cells (THP-1 and A549). Although GNPs induced histone modifications have attracted researcher's attention in recent years, our finding regarding the shape-dependent histone modifications of GNPs is the first reported shape effects in epigenetic regulation, which gives an indication of the mechanism how shape regulates biological pathways.

For the third part of my PhD work, during the bio-nanoscale recognition and interaction process, while typically nanoparticles traffic along the endo-lysosomal pathway and accumulate in lysosomes; there is evidence that other rarer subpopulations of nanoparticles may adopt alternative pathways, within which their surface is modified. Investigation of the properties of these escaped nanoparticles would be of great benefit for the artificial design of nanostructures to achieve specific targeting in nano-therapy. Based on the above consideration, I moved to the re-engineering of nanoparticles to investigate their intracellular fate and gene delivery efficiency and is presented in chapter IV.

Considering the difficulty to track the intracellular trafficking of GNPs and separate GNPs from extracellular vesicles, we used magnetic silica NPs for the study, which can be easily labelled with different dyes for imaging, flow cytometry detection and can be easily purified by magnet pull down. In chapter IV, we reported on the isolation and purification of a specific, well-defined particle complex which is re-engineered by the cellular endogenous process to form a new biomolecular surface, and subsequently secreted from the cell into the extracellular milieu. After isolation, the cell-engineered biomolecular layer is found to be composed of proteins, membrane lipids and, notably, different types of nucleic acids with the nanoparticle making up the core. Then, using flow cytometry, proteomic and RNA sequencing (RNA-seq) techniques, we determined that most of the biomolecules on the particle complex are producer cell derived. By modifying the genome of producer cells, we successfully loaded additional genes into the particle complex. Further studies found that these particle complexes evaded the lysosome destination and were able to deliver loaded genes into recipient cells. By controlling the producer cells and process, diverse possible modifications of these constructs can lead to striking functionalities quite distinct from the original nanoparticles and make them become promising RNA delivery machines.

Perspectives

For the future studies, based on the conclusions of the above three chapters, there are three main directions needing further exploration:

I. How the cell recognises different nano shapes and long-term shape effects on the cellular and molecular level.

In chapter II and chapter III, nanoscale shape-dependent biological effects in immunomodulation and epigenome were discussed. However, how the cells recognise different shapes including the molecules involved in the recognition process, the information transferred during the recognition procedure and the related downstream cellular processing are still unclear. A thorough understanding of these questions is the next step for the nanoscale shape effects study.

Meanwhile, a systematic investigation of the long-term shape effects on the cellular and molecular level is also an important direction to make up for the shortcomings of the traditional *in vitro* studies which usually focused on the short-term effects. Relying on the 3D cell culture platform established in our group, the long-term shape effects in cellular and molecular level investigation is now ongoing and current results are described in appendix II.

II. *In vivo* effects of these tuned nanoscale shapes

In chapter II preliminary *in vivo* exploration was done for three nanoscale shapes and we found that there is a shape-dependent expression in IgG, IgM and anti-c1q, which indicates that shapes can influence the immune system. A deeper and thorough *in vivo* study will be needed to fully understand how the shapes influence the immunomodulation. For example, the biodistribution differences between these shapes and also high-throughput sequencing of B-cell and T-cell immunoglobulin repertoires to gain insights into the immune responses' differences induced by different shapes.

III. Artificial designed nanostructures for potential high efficiency RNA/DNA/protein delivery

To achieve this goal, we need to figure out how the RNA (DNA) is transferred to different types of recipient cells and whether they can keep "alive" once they are internalized, which means the kinetics of this delivered RNA in recipient cells needs to be studied. Then validation as to whether the delivered RNA can be translated or not, using techniques like RNA labelling, confocal imaging, molecular cloning *etc.* If current particle complexes cannot achieve a successful translation, then the next step is to investigate how to do an advanced artificial design to achieve the translation. Once this is solved, it will be a primary contribution to the nanomedicine and nano therapy fields.

In summary, in the thesis we introduced a framework for a systematic study of the nano-shape biological effects, validated the existence of shape-dependent histone modifications, and demonstrated a cellular endogenous process for re-engineering nanoparticles which can

escape the lysosome destination and achieve efficient RNA delivery. The implementation of the above-mentioned perspectives will lead to a methodical understanding of the mechanisms of nano-shapes induced biological responses and open up a new window for the artificial design of nanostructures to achieve highly efficient RNA delivery or target nano therapy.

Appendix

Appendix I: biological stability, cytotoxicity and intracellular locations of GNPs

1.1 Stability of GNPs in biological milieu

The stability of gold NPs in biological milieu was investigated *in situ* by using Differential Centrifugal Sedimentation (DCS) (Figure A1.1). Briefly, aliquots of 100 μl for each diluted sample ($1x10^{12}$ NPs/ml) were incubated with 900 μl of human serum or complete MEM (cMEM) at 37 °C, 300rpm for 1, 4, 24 and 48 hours by using a shaking incubator. The samples were finally gentle mixed by pipettes before injecting in the DCS containing an 8-24% sucrose gradient in PBS. From the DCS result, GNP1 and GNP2 are stable in biological milieu after 24 hours.

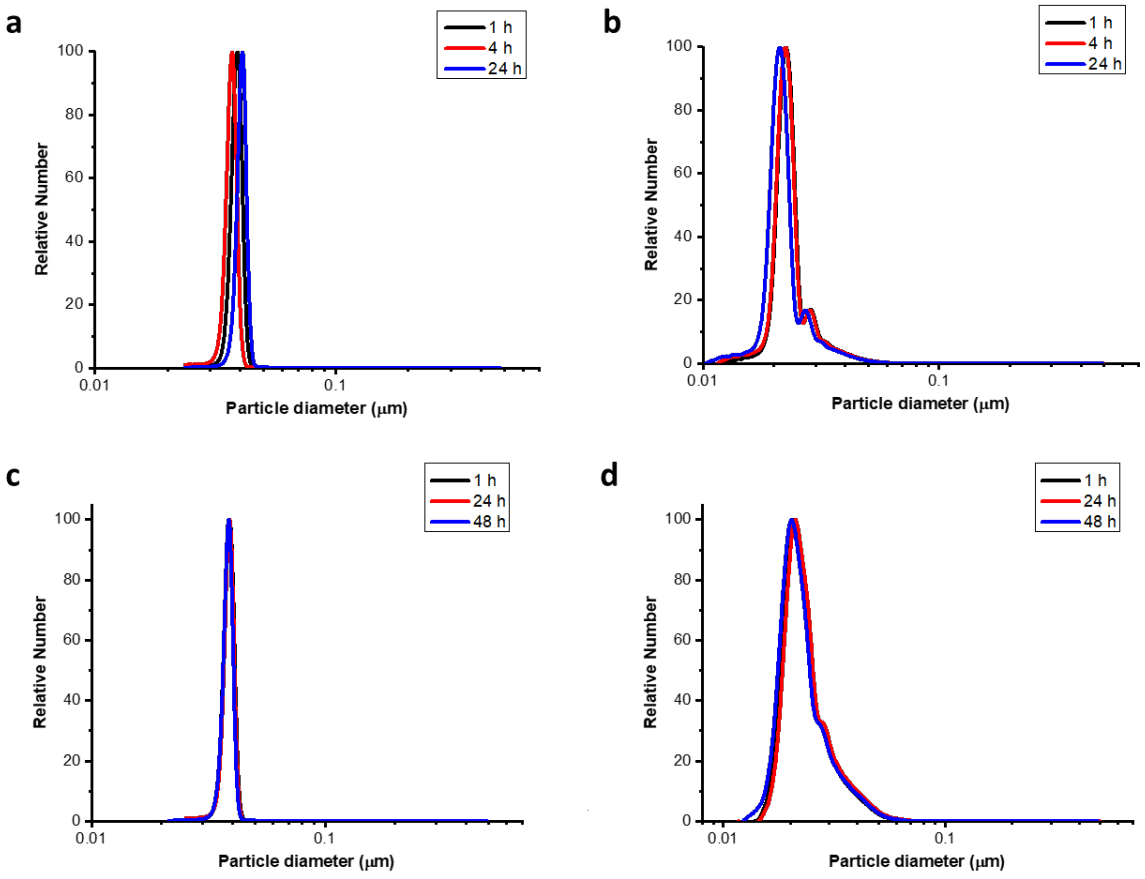


Figure A 1. 1 Stability test performed by DCS analysis.

(a) GNP1, (b) GNP2 were suspended in human serum at 37 °C, 300rpm for 1, 4 and 24 h.
(c) GNP1, (d) GNP2 were suspended in cMEM medium at 37 °C, 300rpm for 1, 24 and 48 h.

1.2 Cytotoxicity of GNPs

The cytotoxic of two shapes of GNPs in A549 was determined by examining the cells with CellTiter 96® AQ_{ueous} One Solution Cell Proliferation Assay (MTS assay), Tetramethylrhodamine methyl ester (TMRM) assay and live/dead cell count assay separately.

A549 was treated with GNP and GNS with concentrations ranges from 1×10^{10} NPs/ml to 1×10^{11} NPs/ml. After 4 h pulse, 1, 6, 16 and 24 hours chase there was no significant effects on cell viability for both shapes of good NPs even at the maximum concentration (1×10^{11} NPs/ml). Therefore, it was confirmed that neither of the shape could cause severe cytotoxicity to A549, which is suitable for down-stream studies on gold nanoparticle internalization.

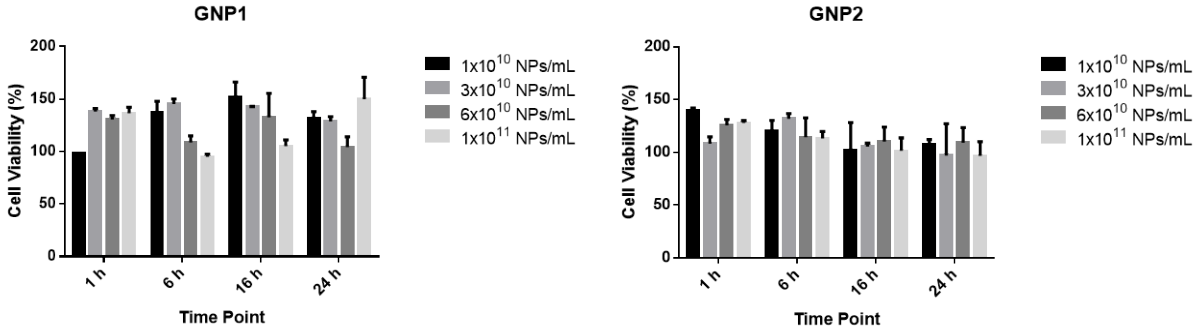


Figure A 1. 2 Relative cell viability of A549 cells after 1, 6, 16, 24 h exposure to a series of concentrations of GNP1 and GNP2 respectively.

Tetramethylrhodamine methyl ester (TMRM) assay and live/dead cell count assay were performed on high content, show GNPs are nontoxic after 4h pulse 24h chase, even at the concentration of 1×10^{11} NPs/ml. Negative control is untreated cell, positive control is cells treated with toxic PS-NH_2 with a concentration of 10mg/ml. TMRM intensities were normalized by negative control, which was set as 100%. Hoechst was used for staining the nuclei, TO-PRO-3 was used for staining dead cells in the 24h live/dead cell count. Experiment was carried out by High Content Assay, all conditions performed in triplicate.

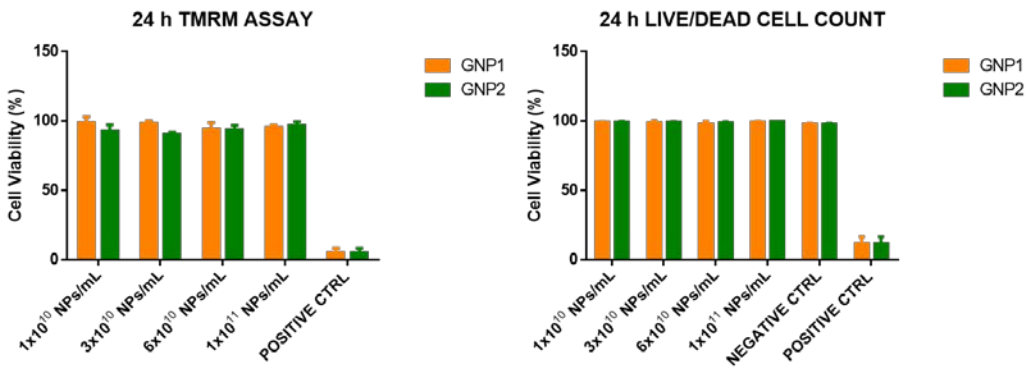


Figure A 1. 3 Relative cell viability of A549s after 4 h pulse 24 h chase exposure to a series of concentrations of GNP1 and GNP2.

1.3 Intracellular uptake and locations of GNPs

GNPs uptake analysis was performed by high content analysis (HCA), the intensity is based on the brightfield images from the cells treated for 24h with GNP 1 or GNP2, the concentration 0 is for untreated cells. Average intensity result shows GNP1 has higher uptake than GNP2 in higher treated concentration. This is consistent with TEM results. (Figure A 1.4 and A 1.5)

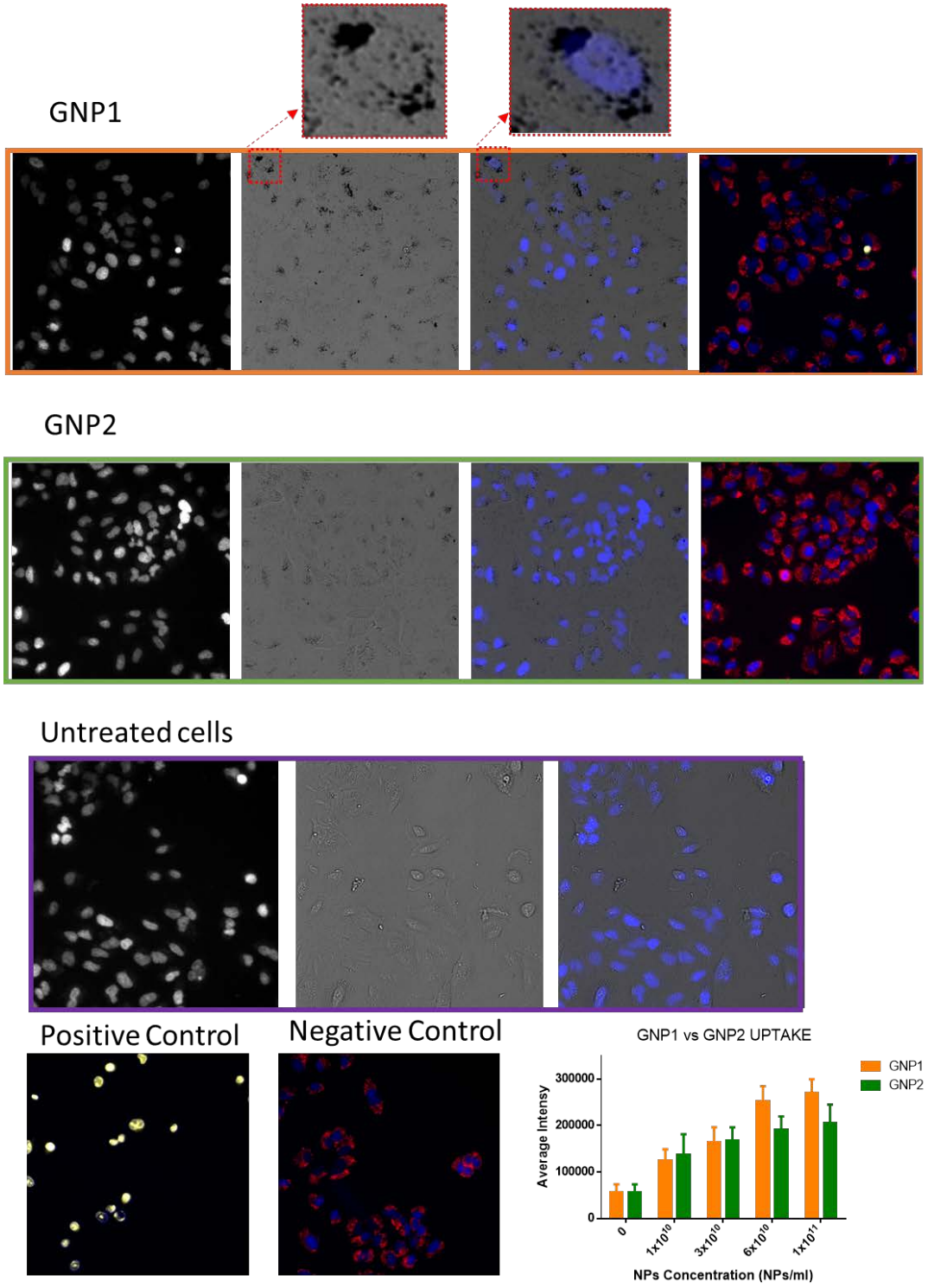


Figure A 1. 4 High content assay for GNPs uptake.

Blue is Hoechst for nuclear staining, red is TO-PRO-3 for cell membrane staining, GNPs are clustered as black granules.

In order to visualize and localize the intracellular trafficking of GNP1 and GNP2, A549 were incubated with 1×10^{11} NPs/ml of GNP1 and GNP2 for 4 h pulse, 1, 6,16,24 h chase before

fixation and further prepared for TEM imaging. At 1 h, most of the GNPs were attached on the membrane, some GNPs were found in the early endosomes; at 6-16 h, GNPs began to traffic to various vesicles; after 24 h, the majority of GNP1 and GNP2 were clustered to the lysosomes, while GNP2 were also found in multilamellar bodies.

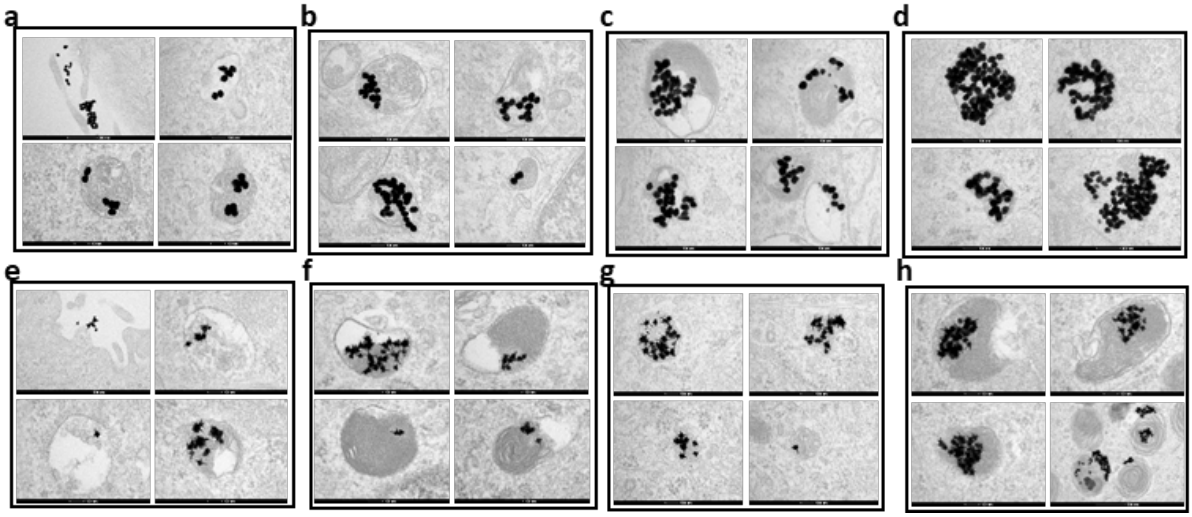


Figure A 1. 5 Time resolved intracellular locations of GNPs.

(a-d) 1, 6, 16, 24 h intracellular locations of GNP1, (e-h) 1, 6, 16, 24 h intracellular locations of GNP2.

Appendix II: Investigation of the long-term intracellular shape effects based on 3D cell culture platform

2.1 Introduction

When nanoscale materials are exposed to cells and *in vivo*, they are (guided by biomolecular coronas derived from the biological environment) incorporated and ultimately processed within endogenous intra-, trans- and cellular trafficking and processing pathways. There they may persist for extended periods of weeks, months (or longer) potentially either leading to persistent signalling dysregulation or other untoward biological processes. While cell-level studies extending over many hours have shed much light on the role, and ultimately removal of the biomolecular corona, the consequences of nanoparticle accumulation over weeks and months have not been hitherto been accessible to cellular mechanistic analyses. Such limited information that we have is largely observational and being derived from highly skilled and specialized *in vivo* studies, is limited in scope. To allow the long-term observations of the bio-nanoscale interactions, our group proposed a three-dimensional (3D) cell culture platform that, while remaining uniformly accessible to nanoparticle exposure, possess some of the basic

phenotypical features as those *in vivo*.[1] They re-form expected inter-cellular junctions and other organizations, and restore various functions known to predominate in the original tissues. Cells within them are (reversibly) quiescent, while remaining metabolically active. This allows the study of NPs intracellular effects in a long-term and avoids confounding effects (due to cell division) derived from the splitting of accumulated intracellular particle loads between daughter cells, while new organelles undergoing biogenesis are exposed to fresh particles. Within these assemblies most cells are accessible from the cluster surface leading them all to be viable (rather than partly hypoxic) over many months, while remaining available for NPs exposure and accumulation in a relatively uniform manner.

2.2 Formation of 3D clusters

Based on our group's previously reported work[1], the A549 3D cell cluster was formed by using a commercially available suspension media which can slow the cell motion due to its low viscosity while allowing for diffusion of nutrients and NPs. (Figure A 2.1).

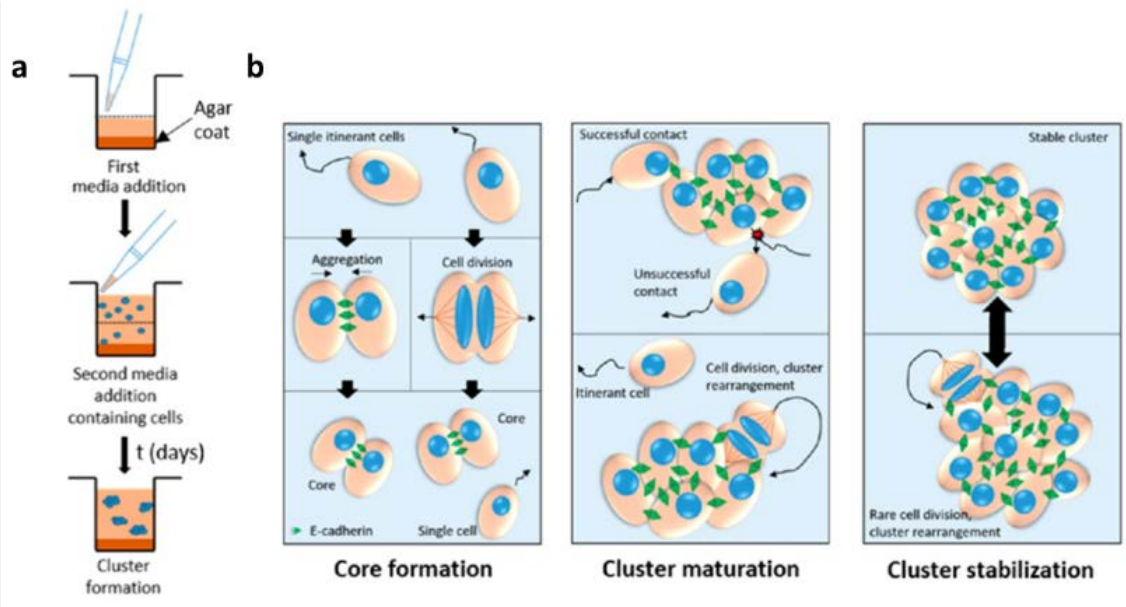


Figure A 2. 1 Schematic illustration of the key steps for cluster formation and evolution.

(a) An agar coating is first deposited to avoid cell attachment and then a layer of suspension media is added; cells are seeded on top of this layer to avoid excessive precipitation. (b) When cells are suspended within the suspension media, several stages of cluster growth occur. In the first hours, most cells are singlets and can move and divide more often. Some cells eventually come in contact and aggregate and so in the second stage some clusters begin to form and become less mobile. In the third stage, most clusters are fully formed, and further growth is limited.[1]

The 3D clusters were characterized by high content microscopy and the size distribution confirms A549 clusters' formation. (Figure A 2.2)

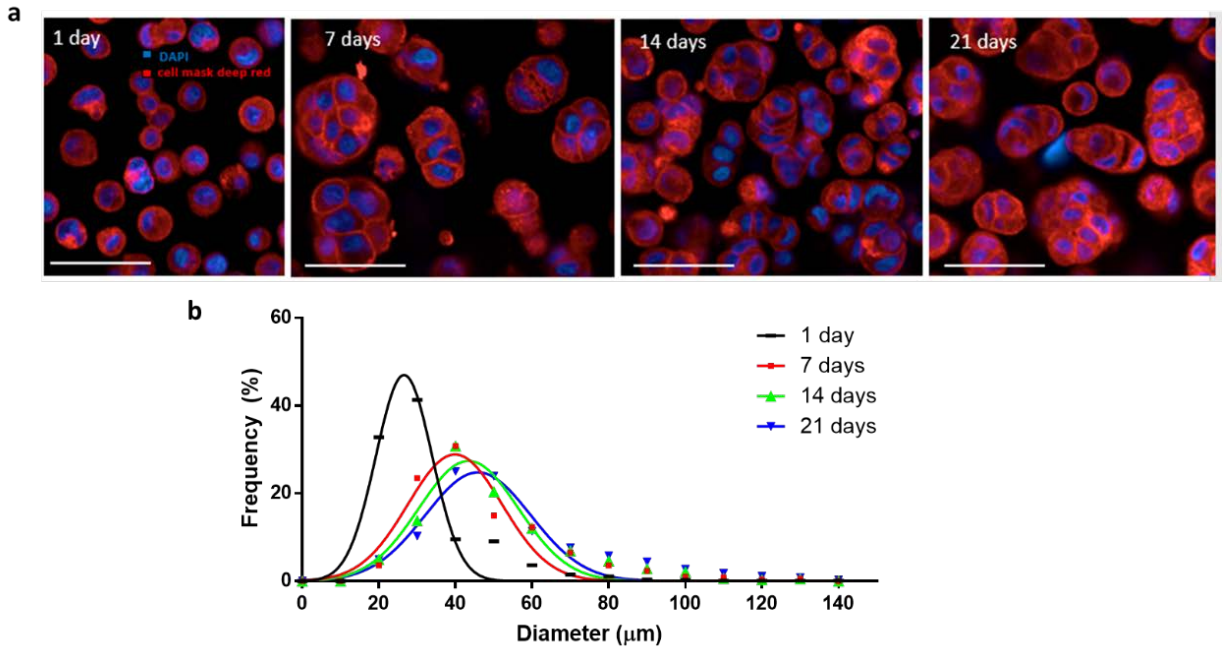


Figure A 2. 2 Formation of 3D clusters.

(a) Representative images of the cell clusters at different time points showing the initial growth and then a more stationary phase. Images were determined by high content microscopy analysis. Cells were harvested and stained with cell mask deep red for cell membrane, followed by fixing in paraformaldehyde and stained with DAPI (blue) for nuclei, then seeded in a 96 well optical plate and imaged with a 40X water objective. All scale bars are 50 μm. (b) Gaussian fitting of the size distribution of cell cluster determined by the Harmony software, more than 1000 clusters were counted.

2.3 Cytotoxicity of GNPs for 3D clusters

After the formation of the 3D clusters, the study started from the basic cell viability of GNPs treated A549 cell line. For the GNPs treatment, A549 cells were pre-seeded in monolayer culture flasks for 24 h and then cells were treated with GNPs with the concentration of 1×10^{11} NPs/ml (GNPs were pre-incubated with cMEM for 1 h to get corona coating on the surface) for 4 h pulse and then cells were washed with 2x cMEM and 2x PBS and seeded on the 3D culture platform for time course cell viability analysis. Cytotoxicity of GNPs were detected by TMRM assay, TMRM intensity for each time point was normalized by negative control (untreated A549 cells), which was set as 100%. More than 10,000 events were counted for each condition, experiment was carried out by flowcytometry, all conditions performed in

triplicate. GNPs are nontoxic at the concentration of 1×10^{11} NPs/ml for 3D clusters up to 3 weeks.

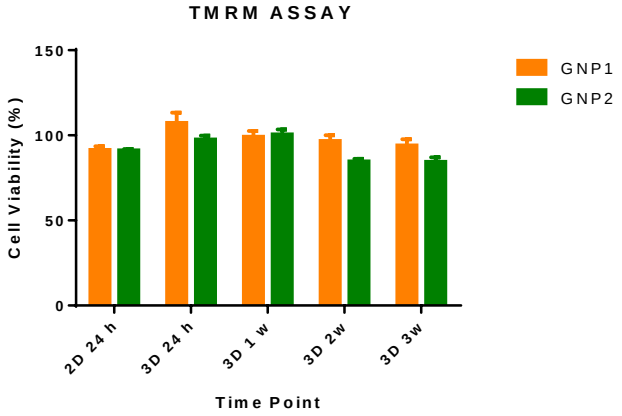


Figure A 2. 3 Cell viability assay for A549 3D cell clusters treated with GNP1 and GNP2 respectively.

2.4 Pilot study of long-term shape effects

For the long-term shape effects investigation, we first explored the influence of GNPs on lysosome functions. As mentioned above for the GNPs treatment, A549s were pre-seeded in monolayer culture flasks for 24 h and then cells were treated with GNPs with the concentration of 1×10^{11} NPs/ml (GNPs were pre-incubated with cMEM for 1 h to get corona coating on the surface) for 4 h pulse and then cells were washed with 2x cMEM and 2x PBS and seeded on the 3D culture platform for long term lysosome function study. For this exploration we used DQ RED-BSA (ThermoFisher, D12051) for the detection of the proteolytic activity and Magic Red Cathepsin B Assay (ImmunoChemistry TECHNOLOGIES, 937) for the monitoring of intracellular cathepsin-B activity. No obvious influence of proteolytic and cathepsin-B activity function observed for GNPs treatment.

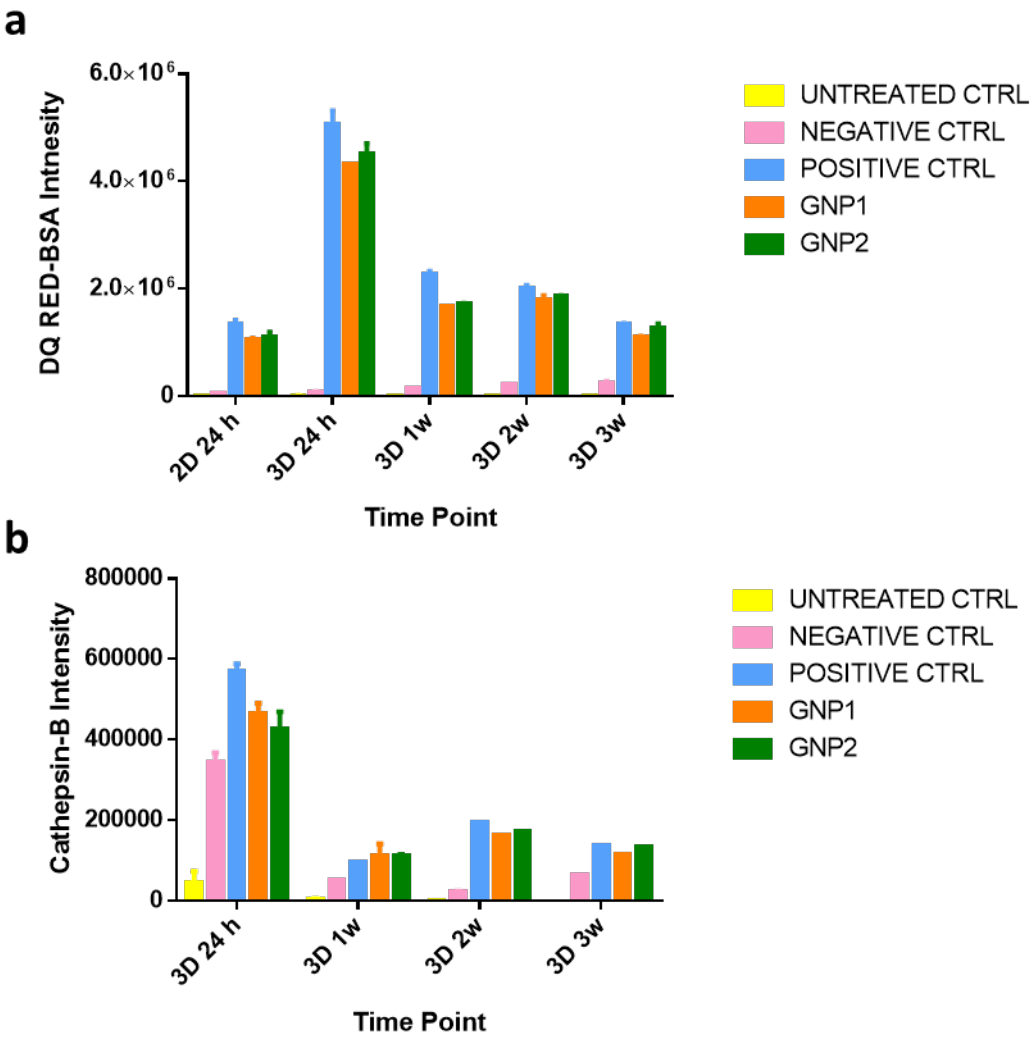


Figure A 2. 4 Lysosome function assay for A549 3D cell clusters treated with GNP1 and GNP2 respectively. UNTREATED CTRL: untreated A549s. NEGATIVE CTRL: A549s treated with 200 nm Bafilomycin A1 and then treated with 10ug/ml DQ RED-BSA for 8h. POSITIVE CTRL: A549s treated with 1×10^{11} NPs/ml GNP1/GNP2 for 4h pulse, 24h, 1w, 2w, 3w chase and then treated with 10ng/ml DQ RED-BSA for 8h. Experiment was carried out by CytoFLEX, all conditions performed in triplicate.

After this, to further investigate the shape-dependent histone modifications, we used $H_3K_{27}me_3$ antibody western blot to check its intensity in the long-term cell culture. This part of work was done in cooperation with Dr. Jingji Li. From the western blot results we observed after 3 weeks treatment GNP2 treated A549s showed significant higher expression on H3K27me3 while for the 2D culture and 3D short term culture (24 h to 2 weeks) GNP1 treated A549s have higher intensity for H3K27me3. Questions like what's the reason for this phenomenon, whether it is universal to other histone modification markers and so on are still unclear. The long-term investigation is still ongoing to fully understand these questions.

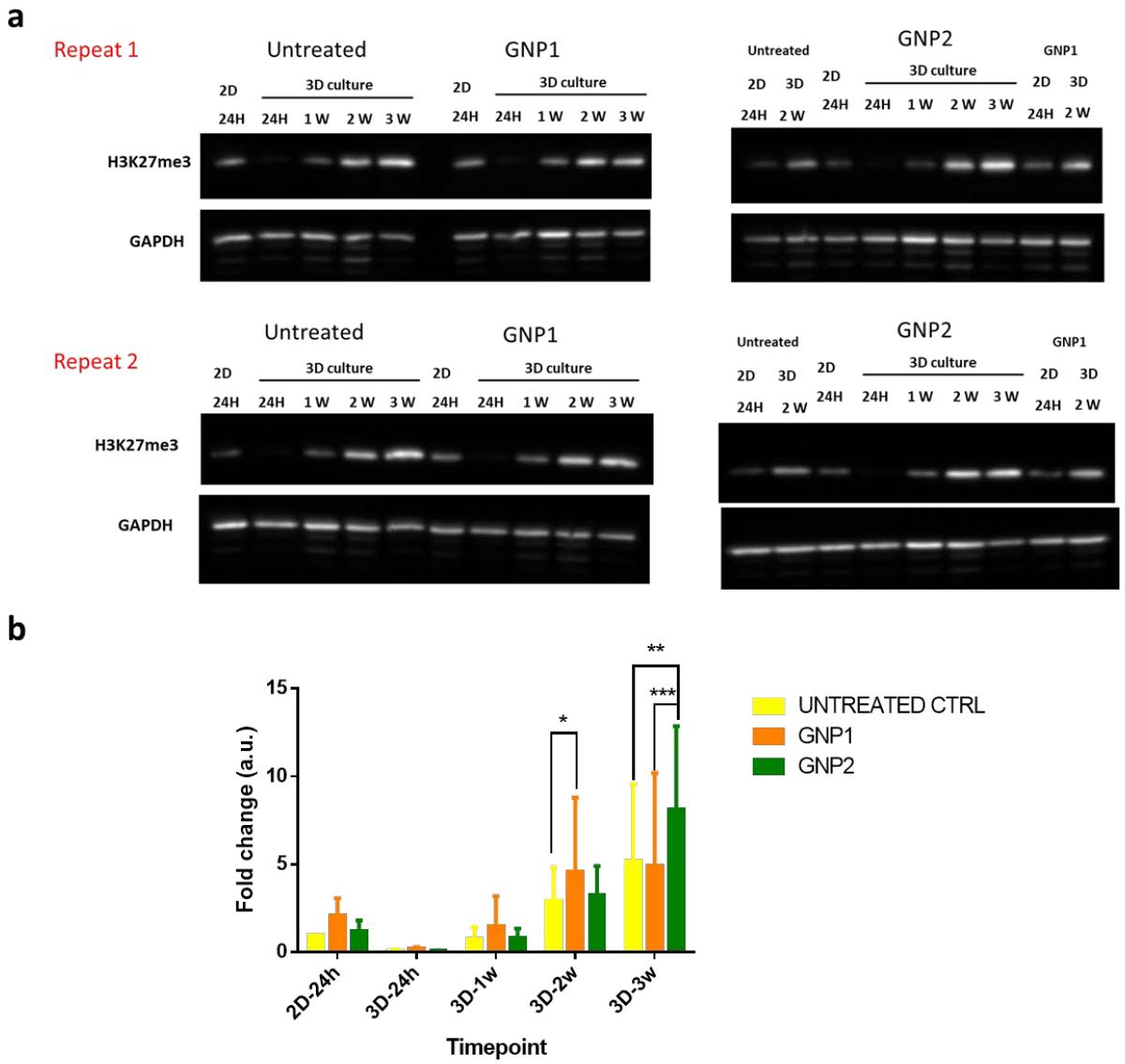


Figure A 2. 5 H3K27me3 expression measured by western blot.

(a) Western blot results for untreated, GNP1 and GNP2 treated A549s' $H_3K_{27}me_3$ expression level up to 3 weeks with two biological replicates. (b) Quantification of the western blot intensity was done by imageJ.

References

1. Muraca, F.; Alahmari, A.; Giannone, V. A.; Adumeau, L.; Yan, Y.; McCafferty, M. M.; Dawson, K. A., A Three-Dimensional Cell Culture Platform for Long Time-Scale Observations of Bio-Nano Interactions. ACS Nano 2019, 13 (11), 13524-13536.

Appendix III: Paper as published

COMMUNICATIONS MATERIALS

ARTICLE

https://doi.org/10.1038/s43246-020-0033-2 OPEN

Classification and biological identity of complex nano shapes

Luca Boselli[1,6], Hender Lopez[1,2,6], Wei Zhang[1,3], Qi Cai[1], Valeria A. Giannone[1], Jingji Li[1], Alirio Moura[4], João M. de Araujo[1,4], Jennifer Cookman[1], Valentina Castagnola[1✉], Yan Yan[1,5✉] & Kenneth A. Dawson[1,3✉]

Everywhere in our surroundings we increasingly come in contact with nanostructures that have distinctive complex shape features on a scale comparable to the particle itself. Such shape ensembles can be made by modern nano-synthetic methods and many industrial processes. With the ever growing universe of nanoscale shapes, names such as "nano-flowers" and "nanostars" no longer precisely describe or characterise the distinct nature of the particles. Here we capture and digitise particle shape information on the relevant size scale and create a condensed representation in which the essential shape features can be captured, recognized and correlated. We find the natural emergence of intrinsic shape groups as well-defined ensemble distributions and show how these may be analyzed and interpreted to reveal novel aspects of our nanoscale shape environment. We show how these ideas may be applied to the interaction between the nanoscale-shape and the living universe and provide a conceptual framework for the study of nanoscale shape biological recognition and identity.

[1] Centre for BioNano Interactions, School of Chemistry, University College Dublin, Belfield, Dublin 4, Ireland. [2] School of Physics and Optometric & Clinical Sciences, Technological University Dublin, City Campus, Kevin Street, Dublin 8, Ireland. [3] Guangdong Provincial Education Department Key Laboratory of Nano-Immunoregulation Tumor Microenvironment The Second Affiliated Hospital, Guangzhou Medical University, Guangzhou, 510260 Guangdong, P.R. China. [4] Departamento de Física Teórica e Experimental, Universidade Federal do Rio Grande do Norte, 59078-970 Natal, RN, Brazil. [5] School of Biomolecular and Biomedical Science, University College Dublin, Belfield, Dublin 4, Ireland. [6] These authors contributed equally: Luca Boselli, Hender Lopez. ✉email: cast.valentina@gmail.com; yan.yan@cbni.ucd.ie; kenneth.a.dawson@cbni.ucd.ie

It is now clear that man-made nanoscale objects are processed by the whole range of endogenous biological machinery, in part via their in situ biomolecular corona surface[1–6]. The idea that "biology sees nanoscale shape" and that certain shape features could also specifically regulate (or even target) given biological pathways is an intriguing but still emerging concept[7–13]. Various aspects of shape at different length scales have always been implicitly involved in nanoscience[14–20]; nanostructures are never smooth and spherical (on an atomic or few-atom scale) by virtue of their synthetic or other origin, and also some highly non-spherical (for example rod-like) objects are quite well studied[21–28]. However, everywhere in our surroundings we increasingly make (purposefully or by accident, as in 3D printing) nanoscale shape ensembles, that possess complex features on length-scales comparable to the particle itself[29–35]. Also, advancing synthetic methods now potentially offer "unlimited" freedom to controllably make a new universe of such complex geometrical ensembles that cannot be described, recognized or characterized with only a few length parameters[36–42]. We currently find ourselves unable to even name and share information, let alone carry out many systematic investigations of them. These general limitations become very significant barriers to understanding the link between nanoscale shape ensembles and biology, and progress there now requires new ways of thinking, new concepts, as well as practical innovations.

As a basic step we require a quantitative statistical definition of nanoscale shape that can ensure complex shape ensembles are meaningfully reproduced and their properties communicated. Such a quantitative framework with the capacity to describe the relevant aspects of shape ensembles might also link nanoparticle shape statistics to biological function in new and hitherto undiscovered ways. Thereby, arenas extending from the science of immunological adjuvants to the biological and clinical impacts of environmental dusts (currently phenomenological in approach, lacking any substantive quantitative basis for development) may be advanced.

While many disciplines will value a quantitative framework for nanoscale shape, here we discuss the two interwoven arenas of nanoscale shape statistics, nanoparticle synthetic control (framed by quantitative measures of shape), and the resulting linkage of those well-controlled shape ensembles to cellular read-outs. We will use that discussion to raise the question of "biological shape identity".

We first introduce the concept of statistically defined nanoscale geometries that go beyond the current practice of borrowing "family" names from common objects ("nanoflowers"' "nanostars", "urchin-like", etc.)[43–46]. It should be stressed that the ideas we describe do not replace conventional experimental nanoparticle fingerprints such as differential centrifugal sedimentation, and spectroscopic methods (i.e. single-particle extinction and scattering), all of which still give valuable basic information[47–51]. However, by using digitized transmission electron microscopy-derived images of shapes, combined with concepts of computational geometrical analysis, we are now able to capture particle shape information that leads to a quantitative definition of shape ensembles. Based on this approach we then define simplified parameters that generalise familiar concepts of particle averages, measures of heterogeneity, and correlation for nanoparticle shape ensembles.

Those developments, when implemented, impose significant discipline on particle preparation, even for well-established particle shape syntheses (such as the gold particle system used here), and lead to optimization of shape homogeneity and reproducibility.

We illustrate the overall approach by analysing whole transcriptome readouts of cells that have been exposed to sufficiently well-defined and homogeneous shape ensembles. We find that distinct shapes (nominally composed of the same materials, prepared under the same conditions) can elicit distinct and non-overlapping biological responses, raising the possibility of shape-targeted regulation of biological pathways. We finally discuss the potential to isolate canonical nanoscale (structural) biological control parameters.

Results

Nanoparticle population shape identity and shape groups. In contrast to molecules, nanoparticle (NP) shapes do not have a single well-defined and unique geometry. For fixed average reaction conditions they are formed by largely irreversible nucleation and growth processes, with particle-to-particle shape fluctuations derived from local growth kinetics around a population of initial seeds[52–57]. Progressive methodological improvement of the synthetic method eliminates many factors (such as mesoscopic mixing heterogeneities) ultimately leading to relatively reproducible populations in which each particle, while being slightly different, possesses an apparently recognizable "typical" shape (see an illustrative selection of NPs in Fig. 1a–c).

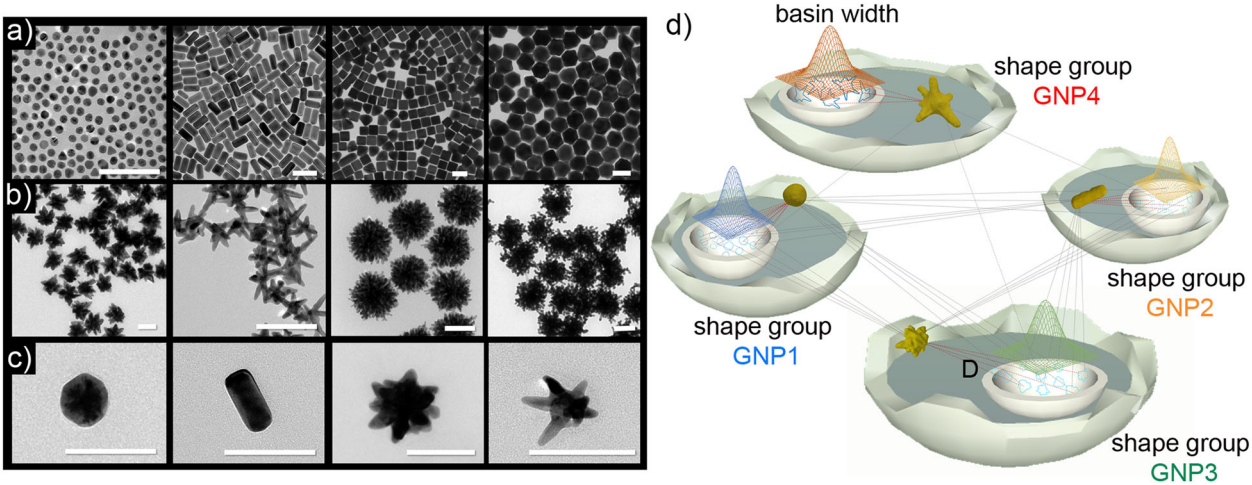


Fig. 1 Nanoparticle shape groups. TEM micrographs illustrating a library of different gold NPs with **a** regular geometrical shapes and **b** branched shapes. **c** Four representative NPs shapes used in this work. From the left: GNP1 ("nanospheres"), GNP2 ("nanorods"), GNP3 ("nanourchins"), and GNP4 ("nanostars"). Scale bars are 100 nm. **d** Schematic representation of the hyperspace geometrical identity and definition of shape groups (basins of similarity).

Transmission electron microscopy (TEM) allows the capture of NP shapes that are appropriately (and homogenously) presented on grids, followed by image digitization in three dimensions (tomography) or two (contour) dimensions (see Supplementary Figs. 1–4)[58–60]. Various mathematical representations of the volume, surface or the contour of the NPs may then be used, but given the typical length scales and organization of synthetic surface features, we can efficiently capture information in a Discrete Fourier Transform (DFT, see more details in "Methods" and Supplementary Table 1)[61–66]. There is nothing unique about such a representation, and in the future others will be found useful. Still, it is noteworthy that (for the shape families studied here, and likely many others) this representation allows contour descriptions of shape using relatively few (two or three) parameters in a low dimensional setting, instead of the many hundreds of points defining the original contour.

We first introduce the concept of shape probability distributions (along axes in the generalized hyperspace of shape parameters) and discuss the idea of population shape distinctiveness (identity) and measures of shape difference (see Fig. 1d for a schematic illustration of the concept).

In this picture, "similar" nanoparticles within a given ensemble (possibly derived from identical syntheses, or other common origins) are manifested as distinct probability maxima (or "basins" in an equivalent entropic-free-energy description). If sufficiently homogeneous, and well separated from others, such ensembles may be considered as distinct pure "shape groups". In some cases, these probability maxima certainly map onto known terms (e.g., the so-called "nanoflowers", "nanostars" or "urchin-like" particles)[44,45,67] but there is no necessity that they do so, and any ensemble can be statistically characterized leading to nanoscale shape-substance definitions. Evidently, the specifics of the geometrical description and presentation somewhat affects the details of this picture, and different applications (from biology to plasmonics) may require different conventions about what constitutes a separate shape group. For instance, within the Fourier representation, it is convenient to normalise with the first (zeroth-order) coefficient, which (in the biological examples discussed later) means that particles with similar features such as "tip sizes", inter-tip separations (tip density), are considered to be from the same shape group, even if the particles are of different overall average size. That appears to be an appropriate choice for some of the biological readouts explored below.

We should also differentiate between complete 3D information that must be captured to fully describe the shapes involved (see Supplementary Figs. 3, 4 and Supplementary Methods), and much more practicable reduced representations that may be sufficient to classify and frame "shape identity", and (possibly) the biological identity of shape. Here we will show that to statistically distinguish some shape groups, it is sufficient to capture, digitise and analyse only the 2D projected contours from particle populations. This approach essentially samples smaller segments of correlated surface curvature via their statistically presented contours. In general then, the aim is to create minimal (low-dimensional) shape descriptors by progressive simplifications in both data collection and analysis, providing the outcome still parameterises those differences (and similarities) between shape ensembles that are relevant to the purpose.

Nevertheless, one has to be attentive to the choices made. For some classes of structure specific over-simplifications of geometrical representation may subordinate (obscure) too much biologically relevant information, and the underlying biological control parameters may not be captured. Within that (under-determined) representation, materials could then be (wrongly) identified as similar while eliciting quite distinct biological readouts. In our current studies, this is not relevant, but in future one should bear in mind the general concepts of shape definition presented here, rather than only the specific choices made.

Distance in shape space and identification of groups. To meaningfully attribute concepts of "shape ensemble difference" one must choose appropriate definitions of "distance" between particles in shape ensembles using the geometrical coefficients that were used to capture and describe them. In the present case, a generalized (complex coordinate) shape descriptor is defined using tens or hundreds of Fourier coefficients obtained by a DFT (see Supplementary Fig. 5, "Methods" and Supplementary Methods for details) of the contours. The Fourier representation is of course equivalent to the full spatial (discrete) contour itself, but for many classes of shape a given representation may require only a few coefficients (instead of hundreds of real-space points) to sufficiently capture the key information. We, therefore, now define a distance (so-called $L_n$norm, see "Methods" and Supplementary Fig. 6) based on differences between these Fourier coefficients. Using this concept of "distance" between particle contours (for example for the material ensembles of Fig. 1) "spatial proximity" becomes a proxy for geometrical similarity, and particles visually identified as members of the same group indeed cluster with each other. Those recognizably different populate other separate clusters. Also, the same particle syntheses (when sufficiently optimized) are reproducibly assigned to the same location in shape space, and "spatial" proximity (in this space) maps well onto shared particle group identity, as well as measures of ensemble quality and reproducibility. This all now imposes some conceptual order on nanoscale shape ensembles.

More striking is the a priori identification of shape groups (Fig. 2a) starting only with a large contour database of different shapes derived from numerous different syntheses but randomized so that our visually based prejudices are eliminated. To illustrate, we then apply an efficiently connected cluster algorithm (see Supplementary Methods) in which this concept of distance can be applied computationally to progressively identify all near neighbors, neighbors of neighbors, and so forth, until the process terminates, and there are no further unidentified neighbors. We can then extract well-defined "connected clusters" (of identity), each element of which is within some chosen threshold distance from one or more other elements of the same cluster. One can then check that the overall outcome is not highly sensitive to the threshold (see "Methods" and Supplementary Figs. 7 and 8). Several variants of the approach are discussed in Supplementary Fig. 9, but the basic outcome is that when we apply the method (Fig. 2a) to particle contour databases (from which all knowledge of particle synthetic origin has been eliminated) shape families (highlighted there with separate colors) readily re-emerge in a low dimension space of principal components composed of the Fourier coefficients.

Interestingly (by visual inspection) we see that for a priori clustering, rare aberrant particles (arising from anomalous synthetic events in a single batch) now appear to be more correctly assigned to shape groups than by reliance only on their synthetic origin. We can thereby fingerprint (to a level appropriate to the given application) the geometrical variability of the ensemble, batch-to-batch reproducibility, and detect (within the limits of the contour sampling itself) "shape impurities", making contact with all of the familiar ideas of particle "characteristics and quality" commonly applied for spherical particle ensembles.

Besides the overview provided by scatter plots along principal component axes (Fig. 2a) we can also define simpler analogs to familiar experimental "probability distribution indices" (Fig. 2b). There the "PDI" widths no longer represent the dispersion in

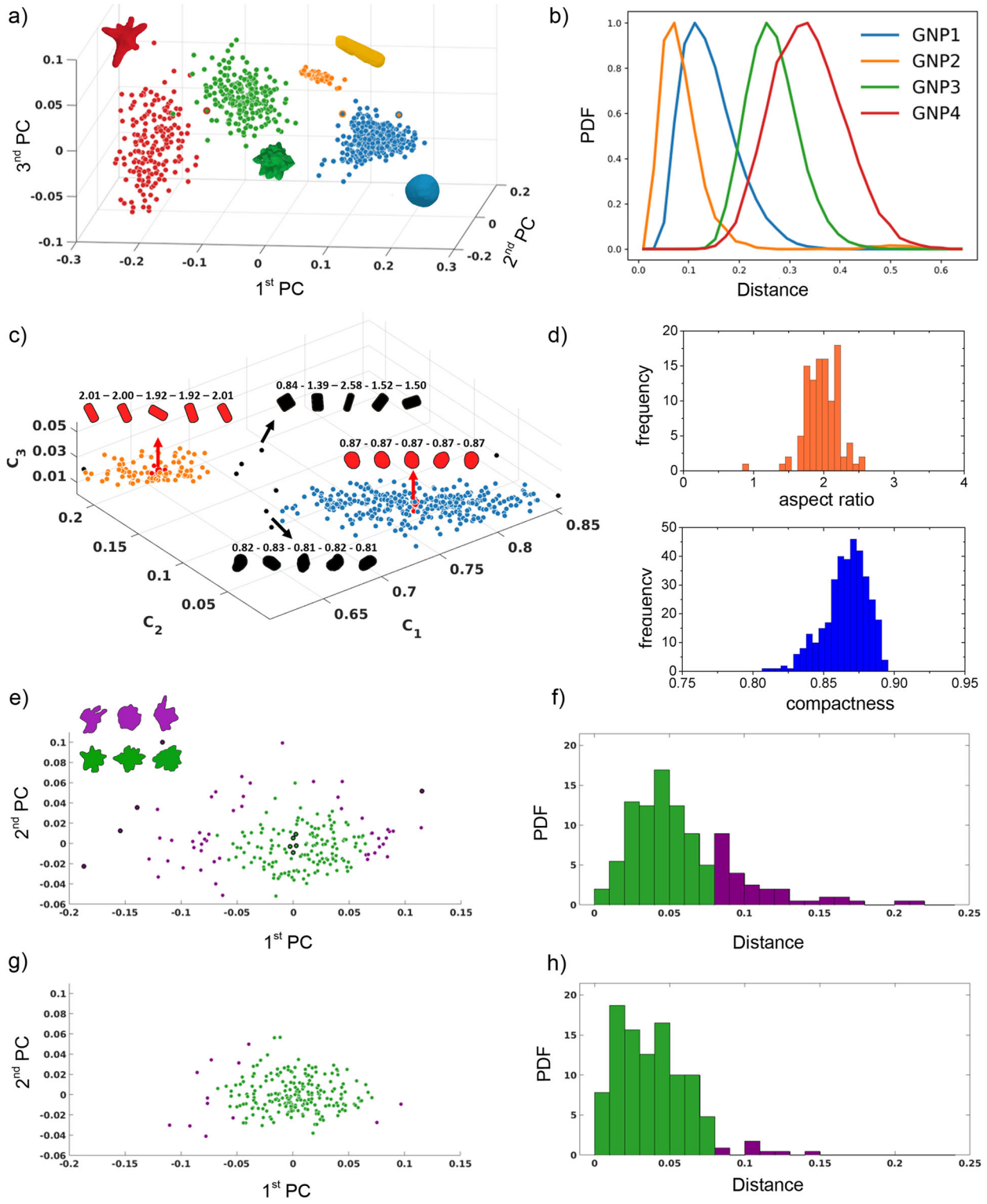


Fig. 2 Shape group assignment. a Clustering using center of gravity method (CGCM—see "Methods") represented as scatter plots for the first three principal components (PCs). Each dot represents one individual NP. Shape group colors are GNP1 (blue), GNP2 (orange), GNP3 (green), GNP4 (red). The first three PCs with magnitude and phase are represented in the expanded space (variance = 88%); **b** normalized probability distribution function (PDF) of shape variability for each shape group. **c** Graphical representation of the contours of 5 particles from the center (red) and 5 particles from the periphery (black) of the cluster of GNP1 and GNP2. On top of each GNP1 particle, the relative values of compactness (1 for a perfect sphere) and on top of each GNP2 particle the relative values of aspect ratio. **d** Distribution of aspect ratio for GNP2 and compactness for GNP1. **e** Clustering and **f** normalized PDF histogram for a "poorly controlled batch" of GNP3. **g** Clustering and **h** normalized PDF histogram for a "good synthetic batch" of GNP3. Green bars/dots define the particles belonging to GNP3 shape group and purple bars/dots define the particles that are at the edge of the basin of similarity (the cut-off distance is calculated as the half height of the distribution). The inset in **e** illustrates three representative contours from the center of the cluster (green) and three representative contours outside the cluster (purple), related to the dots highlighted with a black border in the scatter plot.

radius (as for spheres) but the ensemble spread of important shape related properties, ranked within the principal component analysis, or important in the specific application. Since that "width" (Fig. 2b) of such shape distributions is not a familiar concept we use a specific example to illustrate the physical meaning of geometrical fluctuations (or dispersity) along a geometrical principal component axis. Thus, we use connected cluster analysis to identify the vicinities of the clusters ("basins") of spheres (GNP1) and rods (GNP2). In Fig. 2c we show representative particles from the center (red) and outliers at the periphery (black) of each cluster (GNP1-blue and GNP2-orange). As expected, the actual structures derived from the center of each cluster are (respectively) nearly ideal rods and spheres, while the fluctuations from these shapes are localized outside the two main clusters. Moreover, these fluctuations represent the tails of the distributions of aspect ratio and sphericity (compactness), respectively for GNP2 and GNP1 (Fig. 2d). Further detailed insights into the nature of sub-ensemble shape fluctuations could be obtained by calculating an average shape of these objects (see "Methods" and Supplementary Fig. 10). This could be very useful for small groups of outliers having some specific biological action (absent in the larger ensemble) as one searches for the conceptual (shape) equivalent of "impurity" classifications.

We now also apply these ideas (Fig. 2e–h) to more disordered particle ensembles that could, for instance, model environmentally derived or process derived nanoscale dusts. As an example, we show a poorly controlled synthesis of the same type (GNP3) as used throughout the paper, in which many of the resulting particles have no discernible "group name" such as those commonly used in particle synthesis. Applying the same methods used in Fig. 2a, with appropriate principal component analysis (PCA, see Fig. 2e, g) and probability functions (PDF—Fig. 2f, h), we detect a well-defined sub-population cluster sharing similar properties to GNP3 admixed with many separated more isolated and apparently random structures. Interestingly, for comparison (Fig. 2f, h), we can restore the synthetic conditions closer to those for "good" batches of GNP3, and again see emerging the strong representation of those particle types. These results are quite typical and illustrate how well-defined sub-populations often occur against a significant background of (un-noticed) "broad-band shape noise". One should be aware that many syntheses are carried out without the benefit of these analyses and, from our typical experience, the results can be quite heterogenous, when analyzed by the measures described here.

Other characteristic phenomena can be observed in synthesis of nanoparticle shapes that extend beyond population dispersity or "impurity", and touch on the question of how many shapes groups may arise from a given class of material syntheses. For instance, efforts to tune some reaction conditions between well-defined shape groups appears to form a type of macroscopic distinct shape group coexistence (often regarded as undesirable), rather than a smooth evolution to new shape groups. These effects have not been systematically investigated, but it is possible that they arise from constraints on specific seed faceting that, along with fixed reaction parameters and concentrations, lead to a "reaction-induced shape phase-separation", and thereby essentially "shape transitions" between some shape groups. As systematic shape studies become more quantitative, we may expect such phenomena to be classified in reaction mixtures, and described as mixtures derived from "pure shape" basins (see Fig. 1a), mirroring aspects of traditional macroscopic phase behavior.

Biological functional responses derived from distinct basins. We stress one aim in this work is to establish a conceptual and practicable framework to enable shape identity to play a role in different arenas, including biology. The biological results we describe below are intended to illustrate the idea that ensemble geometry can be meaningfully fixed and causally linked to biological outcomes. We are therefore using high-dimensional cellular (transcriptomic) read-outs as measures of "material identity" rather than performing an exhaustive biological study that claims downstream consequences. There is a great deal of detailed exhaustive biological, surface, and materials science to be done before such claims, carefully made, would be appropriate. However, we hope the present work will frame such studies, and help ensure that they become a durable contribution to the science of shape in biology.

To link biological outcomes to nanoscale geometry there are several key issues that should be addressed. Firstly, particle shape heterogeneity is potentially of such variety and complexity that (using methods of the type described here) we should capture and report on the reproducibility of shape samples including "average shape", shape dispersity and potentially levels of "shape impurity". Obviously, these new issues do not lessen the importance of conventional requirements such as maintaining chemically and immunologically clean conditions (Supplementary Fig. 11), quality and stability of protein (or other) particle dispersion method (Supplementary Figs. 12 and 13), nor the numerous other physiochemical methods in use. However, alone, those are insufficient to prescribe the material properties, in the biological context.

We present data (Figs. 3, 4) for three characteristic shape groups (GNP1, 3, 4) that illustrate their structural properties, basic global biological read-outs, and the link between them. The (occasional) use of legacy "family" names may temporarily assist the reader to make connections to previous literature, but in reality their structural and biological identity lies in the geometrical parameterization, presented in the PC scatter plots, Fig. 3b, c. Nanospheres (GNP1) are included as a conventional reference point, as well as two types of shape ensemble with similar architectures (many tips protruding from a central core but with different overall sizes; "large nano-urchins"—GNP3a and "small nano-urchins"—GNP3b), and another branched shape family ("nanostar"—GNP4) with far fewer and much more developed tips, though of comparable overall size to GNP3b. The principal component scatter plots (Fig. 3b, c) for the different optimized particle ensembles reproducibly lead to probability distributions that are well separated. Recall that (since Fourier coefficients were normalized by the zeroth order-roughly speaking a mean radius) GNP3a and GNP3b, ensembles in this representation overlap, implying they have nearly identical architectures, with different overall (average) sizes. Conventional characterization of these shape ensembles were performed using differential centrifugation measurements (Fig. 3e) and UV-Vis-NIR absorption spectroscopy, showing the typical localized plasmon resonance peaks (Fig. 3f) that "fingerprints" the size and some aspects of architecture of the nanomaterials.

It is worth stressing that the discipline and constraints imposed by these quantitative probability distribution analyses bring shape syntheses to a new level of reproducibility that is suitable for biological study. This occurs in part because the syntheses, even when already known in general outline, are now optimized with an approach (more reminiscent of conventional synthetic organic chemistry) in which we seek to reproduce the shape statistics, rather than conventional macroscopic physiochemical measurement. This guides the nature of the optimizations, but also presents a much stronger reproducibility constraint than merely overlapping of macroscopic measurements. Thus, (Fig. 3h, i) principal component shape scatter plots of three independent preparations of GNP3b-type particles have no discernible

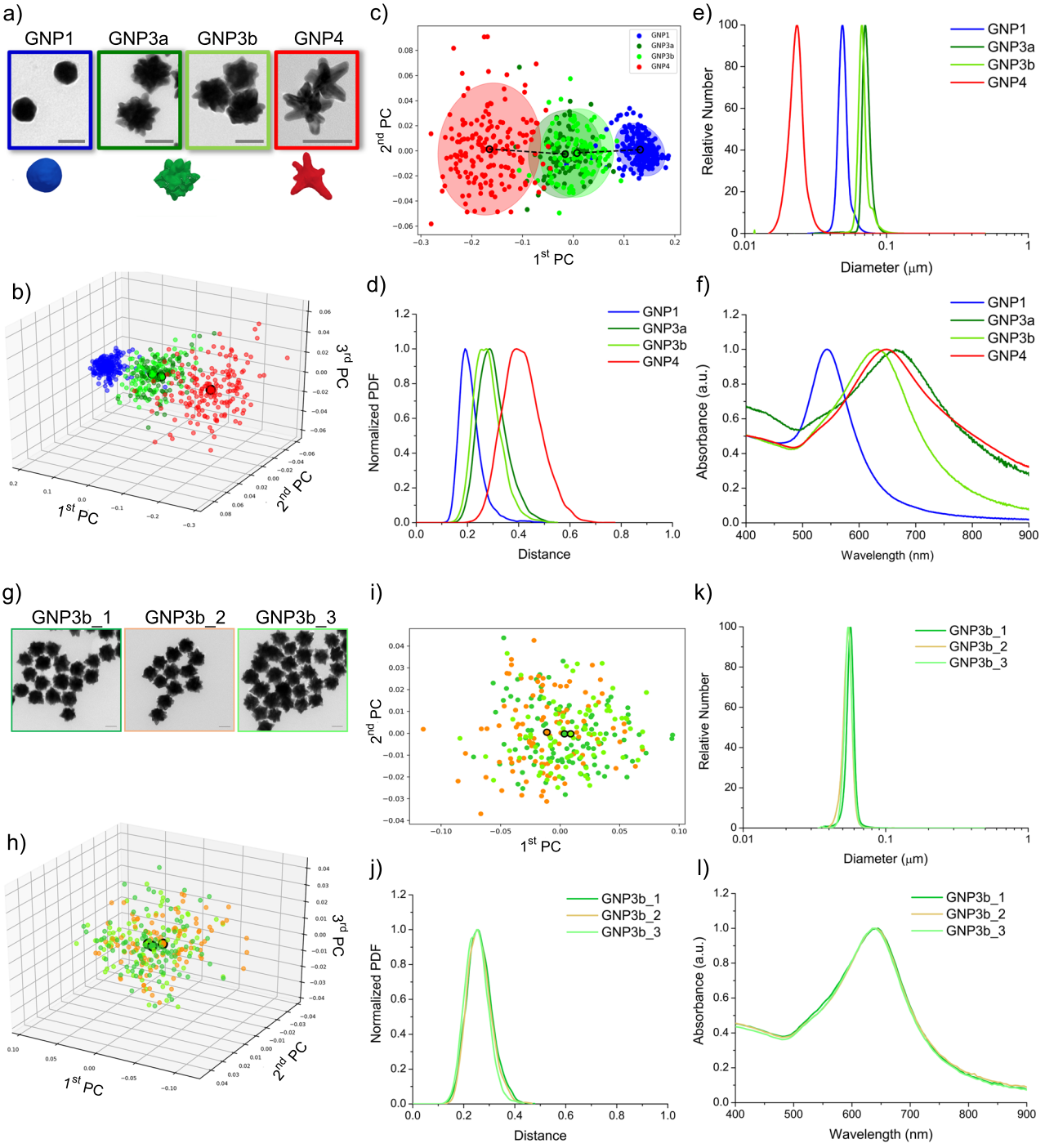


Fig. 3 Full characterization for the GNPs used in the biological study. a Representative TEM images and cartoon for GNP1, 3a, 3b, and 4 (scale bars are 50 nm). **b** 3D and **c** 2D scatter plot of the principal components, **d** intra-batch probability density distribution, **e** differential centrifugal sedimentation (DCS) analysis expressed as relative number, **f** normalized UV–Vis–IR absorption spectra. Qualitatively, GNP3a and GNP3b have 15–20 tips (GNP3a tip length = 15–25 nm, GNP3b tip length = 10–15 nm) while GNP4 have 3–5 tips (tip length = 15–25 nm). **g** Representative TEM images for three independent batches of GNP3b (scale bars are 50 nm), **h** 3D and **i** 2D scatter plot of the principal components, **j** intra-batch probability density distribution, **k** DCS analysis expressed as relative number, and **l** normalized UV–Vis–IR absorption spectra.

structural differences, nor any measurable difference in conventional physiochemical characterization. Other results (Supplementary Figs. 12 and 13) show a high level of time-dependent stability in different protein-stabilized dispersions used in biological studies. Such considerations allow us to design studies with the expectation that biological experiments should attain their intrinsic level of reproducibility. Then the reproducible biological read-outs can legitimately be causally connected to (the equally reproducible) nanoscale shape ensembles.

Shape ensembles differentiate biological responses at cell level. For illustration, we screen high-dimensional biological read-outs using the transcriptome of a murine dendritic cell line (JAWS II)

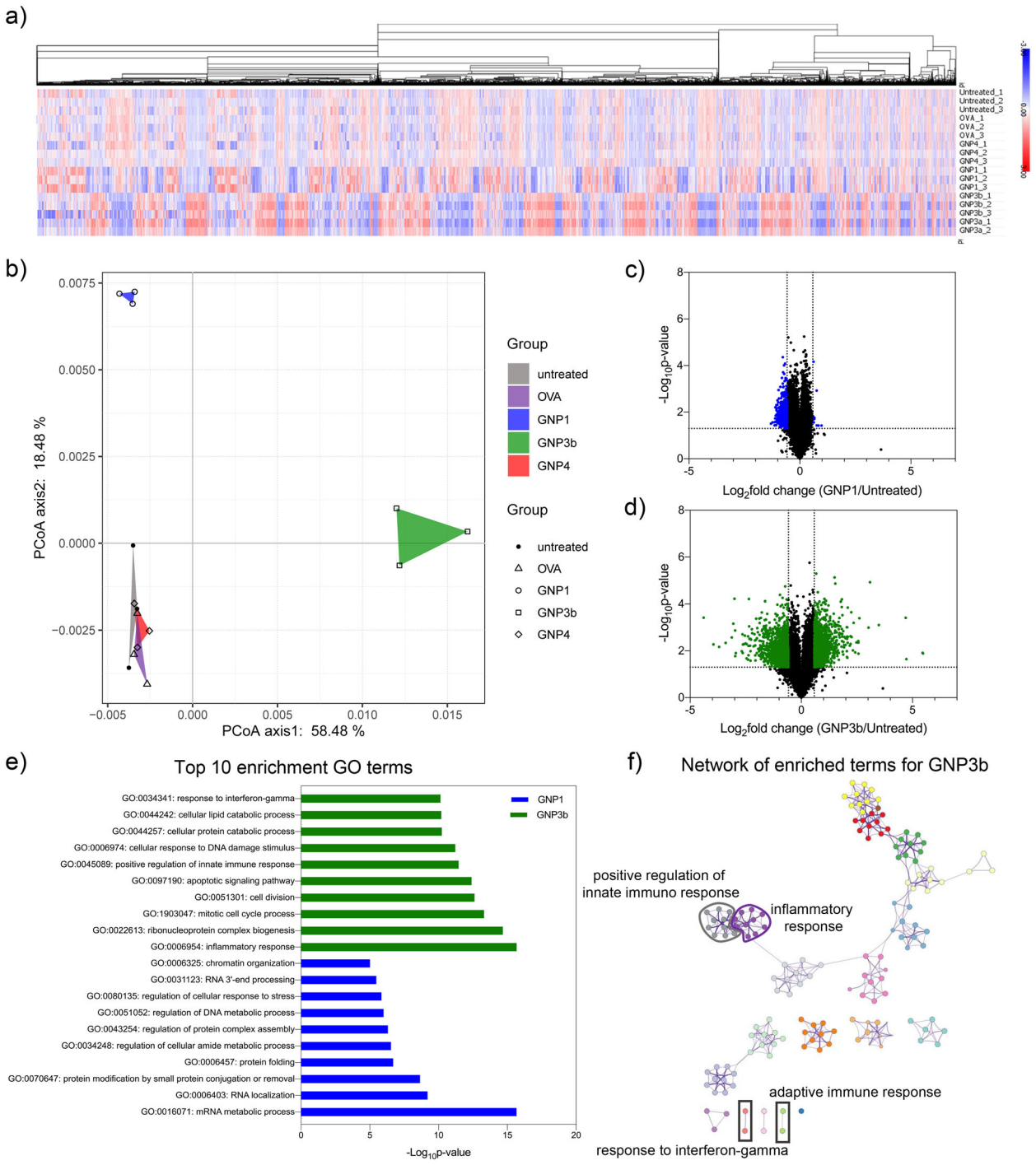


Fig. 4 Transcriptome analysis for shape-dependent biological responses in murine dendritic cell line JAWS II. The cells were treated with GNPs for 24 h. **a** Heatmap of genes identified in the microarray. Z-score hierarchical clustering based on Euclidean distance and average linkage. **b** Principal coordinate analysis (PCoA) illustrating the difference in transcriptome between treatment groups. The PCoA was conducted using Bray-Curtis dissimilarity matrices. **c**, **d** Volcano plot showing differentially expressed genes in comparison to untreated cells after incubation with GNP1 (**c**, in blue) and GNP3b (**d**, in green). **e** Gene ontology (GO) enrichment analysis for differentially expressed genes by Metascape. Bar graph showing top ten GO terms (biological process aspect) for GNP1 and GNP3b. **f** Network enrichment analysis by Metascape. Each term represents a node and colored by its cluster. Terms with similarity score >0.3 are linked by an edge, and the thickness of the edge represents the similarity score. The immunological response-related clusters impacted by GNP3b are highlighted.

exposed to quantified shape distributions. We use highly reproducible shape ensembles (Fig. 3) of fixed particle number, possessing narrow shape distribution, stably dispersed in ovalbumin, prior to presentation in the relevant mouse dendritic cell medium (see Supplementary Fig. 13). The whole cell transcriptomes for the three shape families (Fig. 4) have a high level of biological reproducibility in entirely independent experiments (Fig. 4a) and reveal considerable differentiation in the transcriptome between the particle shapes and the relevant controls. From the heatmap the particle ensembles exhibit three distinct profiles (Fig. 4a). The

shape groups GNP1, GNP3a, and GNP3b result in significant overall changes in the cellular transcriptomes, but even with the more pronounced "tip" features, GNP4 has relatively few detectable impacts. Strikingly, despite their difference in overall size, the particle ensembles of GNP3a and GNP3b elicited similar cellular responses, suggesting that they may be considered (as they were defined structurally) to have similar biological identity. Principal coordinate analysis (PCoA) graphically illustrates the relationships between the transcriptomic profiles induced by the particle ensembles (Fig. 4b). Again, we see the high level of biological reproducibility (reflected in the repeat triangles) and separation (in the major principal component sub-space) in their biological profiling. Significantly, the near-orthogonality of regulation induced by groups GNP1 and GNP3 points towards a qualitatively different biological outcome. Volcano plots further highlight the differentially expressed genes (DEGs) in comparison to the untreated cells (Fig. 4c and d). With the thresholds of p-value <0.05 and fold-change of 1.5, the DEGs (colored dots) for GNP1 and GNP3b are 561 and 2725, respectively. Furthermore, when Metascape is used to infer enriched biological pathways and networks of the DEGs (Fig. 4e and f) the top ten gene ontology (biological process) enrichment terms show no overlap between the two shape groups, and these results together suggest quite distinct (orthogonal) biological shape-dependent outcomes.

We note carefully the stipulation that the purpose of such studies is to establish causal relationships between shape ensembles and cellular read-outs, and not to draw general conclusions for the biology. However, it is noted that four of the unique gene ontology (GO) terms in the dendritic cells response to GNP3b treatment are strongly associated with immunological processes. In Fig. 4f the intra- and inter-cluster interactions are further visualized by network enrichment analysis where GO terms are represented as nodes (colors based on their clusters) and their interactions represented by the strength of the links between the nodes. The results suggest that GNP3 ensembles stimulated three distinct immunological processes. Thus, while the two clusters of "inflammatory response" and "positive regulation of innate immune response" form a compact local large cluster, the other two enrichments of "response to interferon-gamma" and "adaptive immune response" are not linked, suggesting separate regulated processes of GNP3.

Whatever their broader biological implications, such results clearly suggest the potential for differentiated causal shape-dependent impact on sub-groups of biological pathways. It might be argued that, since the particles are nominally of the same material (gold), with same initial stabilizing protein (ovalbumin) and exchangeable cell culture medium dispersant, the results shed light also on the mechanism. However, to ensure clarity, and cast light on future investigations, we suggest such interpretations are premature, and the results do not even establish the level at which the causal linkage occurs. While one can speculate that aspects of the shape directly induce biological differentiation (via receptor clustering or local cytoskeleton stressing) it is also feasible that the shape is causally linked to other (for instance material surface) or surface corona organization aspects. Indeed, the fact that group GNP3 has a high density of features (and gaps between them) of the same size scale as the proteins adsorbed to them, may lead to an in situ corona not previously envisaged, and not presently understood. These are therefore complex questions that require significant efforts to unravel them, and in-depth in vivo investigation to understand the broader significance. Hopefully, the investment in such studies will be justified by the well-defined conceptual framework, and highly reproducible causal link between shape and biology.

Discussion

Since the work here represents quite new and possibly unfamiliar concepts in the quantitative linkage of nanoscale shape to biological (and potentially other) outcomes it is also worth considering thoughtfully and critically the issues that will frame the future. We see the absence of any such framework as currently limiting the fields capacity to create (meaningful) "reproducible" scientific research, to protect intellectual property or advance towards regulation as medical products.

However, it is fair to ask what level of detail we need to "define" material shapes (say for biological applications) beyond existing experimental approaches such as plasmonics, differential centrifugation times, and various forms of scattering. Evidently, such methods, currently used to minimally fingerprint materials in an effort to make them reproducible, are not useful proxies for population shape differences (especially subtle ones). They instead pick up generic particle differences (mostly quite unrelated to shape) and while it is reassuring if material batches have overlapping spectra, they cannot describe shape, capture its biologically relevant characteristics nor track evolution of shape distributions that correlate to biological outcomes. We, therefore, propose that shape features on this scale, if they are control factors in biological function, must now be captured by structural information that goes beyond such experimental measurements, or the "visual impressions" of electron microscope images.

Succinctly put, currently published results that cite "names of shapes" may (even unknowingly to the authors) refer in detail to complex and irreproducible mixtures, or label as "identical" quite different shape distributions that significantly and differentially affect the biology. That could confound the future development of the field.

That situation is very evident for crude heterogeneities. They, and their biological impacts, can be identified, eliminated, and those facts communicated within the framework presented here. For instance, a simple mixtures of quite distinct structures (an extreme example would be GNP1, GNP3b) can be quantified by (multiple peaks or anomalous line shapes) in even the simplest representation of the present method and limits of detection are set by (controllable) image capture statistics rather than resolution or representation limitations. Heterogenous populations composed of distinct sub-populations, each with different functional biological outcomes (for example GNP1,3b mixtures) when correlated with macroscopic averages of the biological endpoints (such as whole transcriptomes, proteomes, etc.) will be confounded by appearing as multiple additive pathway activation, or worse, interpreted in bioinformatics as non-existent connected pathways and should be eliminated. That much is clear.

There are, however, more subtle questions on which the broader future of the biological study of shape will depend. Thus, given the type of information captured here, to what level can quasi-homogenous (single shape group) but well spread shape distributions (parameterized by some median or average and principal component widths) be structurally distinguished from a "nearby" distribution with comparable widths, and how relevant will that resolution be in the exploration of shape-dependent biology. The answer to the first question is already relatively clear within the current paper, and populations that look "similar" can be definitively (and reproducibly) made, captured and recorded as distinct by virtue of the methods outlined here. Indeed, that is the power of high dimensional (e.g., Fourier) representations condensed (e.g., via principal components) to reduced representational dimensions for they have the capacity to capture, detect and quantify complex population differences that are otherwise hard to describe, or even observe.

The second question is whether that capacity to recognise multiple subtle, and (qualitatively) difficult-to-define changes will

materially affect the destiny of nanoscale-shape-dependent biology. The answer to that question resides partly in understanding whether nanoscale shape-dependent biology will itself allow quite new and independent biological interventions. For example, if overall changes in shape are found to differentially and specifically target distinct biological outcomes (pathways, systems) then rather narrow and well-controlled (shape-surface) populations will amount to active agents ("drugs"), in a background of rather similar but less interesting nanostructures of the same chemical identity. We believe that is indeed the paradigm to expect, and that in the absence of methods such as those presented here we face the impossible task of searching for, identifying, and reporting on an "active agent" in a shape landscape by qualitative visual inspection. Evidently, the causal assignment linking those structural changes to biological (e.g., immunological or other) outcomes would also be challenging, and it would be impossible in a synthesis to recover identical shape populations. More fundamental (and with more profound long-term significance) the lack of a low dimensional quantitative structural "order parameter" that can be correlated with the biological outcome currently inhibits the development of "structure-function" hypotheses in which specific (low dimensional) aspects of the structure can be (meaningfully) associated with the biology, and act as the basis for further improvement on design.

Besides these specific interests, we may also credibly hypothesize a significant (if still poorly understood) role for particle shape in other important arenas. This includes environmental pollution, and the many associated (and growing) mysteries of human health (including immune regulation) involved there. Durable advances throughout will require the same quantitative approach to environmental dusts in terms of "substance definition" to help frame the strategy of investigations and reliably report the outcomes.

We believe, in time, these points will be established with growing clarity, as future research develops along more quantitative lines.

Methods

NPs contour extraction. All the TEM images were analyzed using the Open Source Computer Vision Library (https://opencv.org/). In particular, for the contour extraction the function cv2.findContours was employed. This function is an implementation of the algorithm for contour extraction proposed by Suzuki et al.[68] Before extracting the contours, the original TEM grayscale images were transformed to binary and filtered (using again OpenCV). A threshold of 160 was used for the binary transformation and to smooth the image contours a blurring low-pass filter kernel was applied using a blur size of 16. Care was taken to manually confirm that all the contours extracted corresponded to NPs. Objects that were not NPs (e.g., the scalebars, see Supplementary Fig. 1) were manually removed from the database. The contours extracted, considering the resolution of the TEM images contain on average between 500 and 800 points. An example of an image and the resulting contour extraction is shown in Supplementary Figs. 1 and 2.

Contour normalization and resampling. The contours obtained are closed curves in 2D with coordinates ordered counter-clockwise. The coordinates are integer numbers (corresponding to the position of the pixels that make the contour in the image) connected horizontally, vertically or diagonally. Each shape was centered so that its center of mass lies on the origin, and the contours were transformed so that each pair of consecutive points were one-unit distance apart. This required adding a point between two diagonally connected points to form an L-shaped connection of three equally spaced points.

Discrete Fourier Transform (DFT) calculations. Each contour is made of N points. The n^{jh} point of the curve has coordinates x_n, y_n which are used to define the complex coordinate as $z_n = x_n + iy_n$ (where i is imaginary number). We then use the complex coordinates to perform a Discrete Fourier Transform (DFT) of a contour defined as:

$$\begin{aligned} \hat{z} &= F\{z\} \\ \hat{z}_k &= \textstyle\sum_{j=1}^{N} z_j e^{\frac{-2\pi i(j-1)(k-1)}{N}} \end{aligned} \tag{1}$$

where $\hat{z}_k$ is k^{th} Fourier coefficient and $1<k<N$.

Average shape calculation. For each contour, the N Fourier coefficients calculated by using Eq. 1 were normalized for unitary L_1norm. Also, a global phase was applied to all coefficients to make the first non-zero coefficient a pure real number. The number of coefficients thus obtained is the same as the number of points on the original contours. Since this number was not constant, the number of coefficients was decreased or increased by eliminating or adding zero-valued coefficients on the high frequency part of the Fourier transform, around the position $N/2$ and renormalizing to unity. If not stated otherwise, 1024 Fourier coefficients were used to describe the shapes.

For a set of N shapes belonging to a single group, a representation of the average shape of this group was extracted by finding a Fourier representation for each one that minimizes the sum over the whole set of the L_1norm distance to the average (see Eq. 3). There are three operations that do not alter the shape itself but affect the phases of the Fourier coefficients: rotation, starting point, and mirroring. Any rotation in real space is equivalent to multiplying all the coefficients in Fourier space by a single imaginary number of unit norm, which affects the phase of each coefficient equally. This global phase was fixed by requiring, on the normalization procedure, that z_2 is real.

The second operation, changing the starting point, results in a cyclical permutation of the ordered set $\{z_n\}$ which affects the phases of $\{\hat{z}_n\}$ in the following manner.

Let T_m be the operator that changes the starting point from z_1 to z_m then the new set of Fourier coefficient is given by:

$$\begin{aligned} T_m\{z\} &= \{z_m, z_{m+1}, \ldots ..z_N, z_1, \ldots ..z_{m-1}\} \\ F\{T_m\{z\}\}_k &= e^{\frac{-2\pi i}{N}(k-1)(m-1)}\hat{z}_k \end{aligned} \tag{2}$$

Which only differ from the original set by their phase.

Finally, mirroring is equivalent to taking the complex conjugate of the coefficients. For these last two operations (mirroring and starting point), we iteratively tried every combination to obtain the desired coefficients, by minimizing the sum of the distances between each particle in the known (or training) set to the average shape of its group.

Average shape classification method (ASCM). After calculating the average shape for each group, we compare the distance from the unknown shape's Fourier coefficients to each average shape using the complex-valued coordinates of the Fourier space. The same restriction of real-valued $\hat{z}_N(z_2)$ was applied for rotation, and the distance was calculated for every combination of mirroring plus starting point permutation of the unknown particle. The particles were classified as belonging to the group with the minimum distance to its average shape.

In order to classify an unknown shape, we compare its L_1norm distance to the center of each basin and assign it to the closest group. The L_1norm on the complex coefficients is defined as usual. Let $\{\hat{a}_n\}$ and $\left\{\hat{b}_n\right\}$ be the N Fourier coefficients representing two different shapes, A and B. The L_1norm distance between them is given by:

$$\|A - B_1\| = \left(\sum_{n=1}^{N}\left|\hat{a}_n - \hat{b}_n\right|\right) = \sum_{n=1}^{N}\sqrt{\Re\left(\hat{a}_n - \hat{b}_n\right)^2 + \Im\left(\hat{a}_n - \hat{b}_n\right)^2} \tag{3}$$

where $\Re(z)$ and $\Im(z)$ are the real and imaginary parts of the complex number z.

Nearest neighbor classification method (NNCM). The Fourier coefficients obtained by using Eq. 1 are not directly useful for classification as they are scale, rotational and starting point dependent as pointed out before. One way to avoid this limitation is by the alignment explained above. Another option is to perform a normalization as the one reported by Sokic *et al.*[64] This normalization first starts by calculating the normalization factor γ,

$$\gamma = \sum_{i=2}^{N}|\hat{z}_i| \tag{4}$$

Then the normalized Fourier coefficients are:

$$\hat{c}_i = \frac{|\hat{z}_i|}{\gamma}, 2<i<N \tag{5}$$

Notice that the normalized coefficients $\hat{c}_i$ are now real numbers and that the element $n = 1$ is not used. This shape signature was chosen as it is invariant under translation, under rotation and it does not depend on which point of the contour is selected as the initial point. In addition, the coefficients are also scale invariant.

Similar to the ASCM, to compare two contours using the normalized coefficients defined in Eq. 5, the L_1norm distance between their set of $\hat{c}_i$ was calculated. Considering two different shapes labeled as 1 and 2 and described by their set of normalized Fourier coefficients c_i^1 and c_i^2 where $i = 2....N$, the L_{1norm} between them is calculated as follow:

$$D = \sum_{i=2}^{N}\left|c_i^a - c_i^b\right| \tag{6}$$

From the experimental procedure and the image analysis developed, we obtained a number of contours for each type of NP studied (4 types in total). With

the obtained contours a database for each NP type is generated by randomly selecting K contours of each type. Then an unknown NP is classified based on the following procedure: (1) The distances D (Eq. 6) between the unknown NP and each of the elements of the database are calculated. This will give $4K$ distances. (2) The unknown NP is then classified as the type for which the D is minimum.

Center of gravity classification method (CGCM). We will call type A a database of any type of GNPs. This database is composed of M contours. We calculate the average normalized Fourier coefficient $\hat{c}_i$ as:

$$\langle \hat{c}_i \rangle_A = \frac{1}{M} \sum_{j=1}^{M} \hat{c}_i^j \tag{7}$$

The set of $\langle \hat{c}_i \rangle (i = 2 \ldots .N)$ correspond to the Fourier coefficients' Center of Gravity (CG) for the NP type A. Then the distance of any contour, say contour j, to the CG of the NPs of type A is given by:

$$D_A = \sum_{i=2}^{N} \left| \hat{c}_i^j - \langle \hat{c}_i \rangle_A \right| \tag{8}$$

Similar to the NNCM, a database of K contours for each NP type is generated. Then for each type of NP, the average normalized Fourier coefficient (Eq. 7) are calculated. Then an unknown NP is assigned to a shape group as follow: (1) The distances D_A (Eq. 8) between the unknown NP and the average normalized Fourier coefficient for the 4 types of NPs are calculated. This gives 4 distances. (2) The unknown NP is then classified as the shape group for which the D_A is minimum.

Principal component analysis (PCA). In order to evaluate the particle classification methods, we performed a principal component analysis (PCA) on the space of the Fourier coefficients. In this way we can measure which coefficients contained more of the information necessary for the classification of the shapes. With the particles represented by complete Fourier coefficients in the way done for the calculation of the average shapes, in a 1024-dimensional space with complex coordinates, PCA is able to find a suitable base set of axes in this multidimensional space ranked in order of projections containing most of the statistical variance between samples, which in other words means the most information.

For example, one particular coefficient that has the same value for all particles obviously is of no help when trying to distinguish between them, and can safely be ignored, whereas a coefficient that has a wide range of values over all the particles may have a particular range of values associated with each particle group, and be instrumental in separating them. We are also interested in measuring how much information is lost by using only the magnitude of the Fourier coefficients (see above for details) as real-valued coordinates for the particles.

By interpreting each particle as a sample, and each Fourier coefficient as an independent variable, first, we subtract the center of mass of the whole set of particles in order to make them distributed around the origin with zero mean, then we calculate the covariance matrix of the entire population and find its eigenvectors which will be the principal components, in decreasing order of eigenvalue. Principal components are always orthogonal and normalized. In order to facilitate the comparisons and plotting between the different geometrical spaces, the full complex Fourier coefficients versus real-valued magnitudes, we treated each part of the complex number for the full coefficients as different variables $\left\{\hat{\zeta}_n\right\}$ by separating each Fourier coefficient into its real and imaginary parts, which doubles the number of coordinates but now both spaces have a similar range of values, and the same scale, so their variances are directly comparable for further analysis.

$$\hat{\zeta}_n = \Re(\hat{z}_n), \ 1 \le n \le 1024 \tag{9}$$

$$\hat{\zeta}_n = \Im(\hat{z}_{n-1024}), \ 1025 \le n \le 20148 \tag{10}$$

where $\Re(z)$ and $\Im(z)$ are the real and imaginary parts of the complex number z.

It is remarkable that a small number of principal components are able to accurately separate the particles into groups, even if in the analysis we considered the whole set of particles as "database" for classification. Because the phase of the Fourier coefficients of the particles have been aligned explicitly to calculate the representation with the minimal spread around the center of gravity for each group, the number of misclassified particles is very small, and composed of some malformed particles that are not very good representatives of their respective groups. It is also noticeable (see Supplementary Table 2) that the principal components are almost entirely aligned with just two Fourier coefficients, suggesting that these two coefficients would be enough to classify an unknown particle in a computationally cheap way that involves very little calculation.

The first two principal components (Fig. 2a) for the $\hat{\zeta}$ add up to 95% of their respective magnitudes:

$$PC_1\left(\hat{\zeta}\right) = 0.97\vec{\zeta}_{1024} - 0.13\vec{\zeta}_4 + \vec{\epsilon} \tag{11}$$

$$PC_2\left(\hat{\zeta}\right) = 0.93\vec{\zeta}_2 - 0.28\vec{\zeta}_{1022} - 0.13\vec{\zeta}_4 + \vec{\epsilon} \tag{12}$$

where $\vec{\zeta}$ is to be interpreted as a unitary vector in the direction of this coefficient, and ϵ is a small contribution from other coefficients. Likewise, for the $|\hat{z}|$ space (real normalized coefficients), the first 2 principal components are:

$$PC_1(|\hat{z}|) = 0.97\vec{z}_{1024} - 0.11\vec{z}_2 + \vec{\epsilon} \tag{13}$$

$$PC_2(|\hat{z}|) = 0.93\vec{z}_2 - 0.28\vec{z}_{1022} - 0.13\vec{z}_4 + \vec{\epsilon} \tag{14}$$

As we can see, they are practically the same, and furthermore they are practically aligned with the coefficients $|\hat{z}_2|$ and $|\hat{z}_{1024}|$ that we labeled 1st PC and 2nd PC in Fig. 2a, which shows that we can achieve a good separation in clusters without losing much information by using only these two descriptors for each particle.

Transcriptome analysis. A mouse dendritic cell line, JAWS II, was purchased from ATCC (ATCC® CRL-11904™). The cells were maintained in complete culture medium consisting of Alpha MEM with ribonucleotides, deoxyribonucleosides (Gibco) with 10% heated-inactivated fetal bovine serum (FBS) (Gibco), 4 mM l-glutamine (Gibco), 10 U mL^{-1} penicillin and 100 µg mL^{-1} streptomycin (Gibco), 1 mM sodium pyruvate (Gibco), and 5 ng mL^{-1} murine GM-CSF (pre-Protech).

For treatment with GNPs, 1.5×10^5 cells were seeded in 12-well plates in the complete culture media for 48 h prior to the treatment. 3×10^{10} OVA-coated GNPs were dispersed in 1 mL of serum free mouse dendritic cell culture media (Celprogen), added to each well, and incubated with the cells for 24 h. The GNP treatment was performed in replicates and repeated independently 2 or 3 times. Subsequently, the cells were washed with 500 µL of PBS once, and total RNA was extracted according to the manufactory instructions of InviTrap® Spin Universal RNA Mini Kit. The concentration and purity of extracted RNA were confirmed by Bioanalyzer. The samples were then transferred to RNA stable tubes (Sigma Aldrich), dried out in a vacuum, and ready for Whole-Mouse Genome One-Color Microarray (Agilent) analysis.

After correcting background and filtering low intensity signals, 16418 probes were identified. Data are expressed in the mean of replicates in $\log_2$ scale. p-value was calculated by one sample T-Test. Differentially Expressed Genes (DEGs) in comparison to untreated were determined by fold-change >= 1.5 and p-value < 0.05. The heatmap was generated by Morpheus (https://software.broadinstitute.org/morpheus). The Gene ontology (GO) enrichment analysis was performed by Metascape (http://metascape.org/gp/index.html#/main/step1).

Data availability

The datasets generated and analyzed during the current study are available in the Figshare repository, https://doi.org/10.6084/m9.figshare.11948886. The deposit is entitled: Supplementary dataset for "Classification and Biological Identity of Complex Nano Shapes"—a submission to *Communications Materials*.

Code availability

The code used to generate the data produced in this study is available in the Figshare repository, https://doi.org/10.6084/m9.figshare.11948886. The deposit is entitled: Supplementary dataset for "Classification and Biological Identity of Complex Nano Shapes"—a submission to *Communications Materials*.

Received: 11 February 2020; Accepted: 12 March 2020;
Published online: 12 June 2020

References

1. Salvati, A. et al. Transferrin-functionalized nanoparticles lose their targeting capabilities when a biomolecule corona adsorbs on the surface. *Nat. Nanotechnol.* **8**, 137–143 (2013).
2. Lundqvist, M. et al. Nanoparticle size and surface properties determine the protein corona with possible implications for biological impacts. *Proc. Natl Acad. Sci. USA* **105**, 14265–14270 (2008).
3. Monopoli, M. P., Aberg, C., Salvati, A. & Dawson, K. A. Biomolecular coronas provide the biological identity of nanosized materials. *Nat. Nanotechnol* **7**, 779–786 (2012).
4. Monopoli, M. P. et al. Physical−chemical aspects of protein corona: relevance to in vitro and in vivo biological impacts of nanoparticles. *J. Am. Chem. Soc.* **133**, 2525–2534 (2011).
5. Lara, S. et al. Identification of receptor binding to the biomolecular corona of nanoparticles. *ACS Nano* **11**, 1884-1893 (2017).
6. Alnasser, F. et al. Graphene nanoflakes uptake mediated by scavenger receptors. *Nano Lett.* **19**, 1260-1268 (2019).
7. Talamini, L. et al. Influence of size and shape on the anatomical distribution of endotoxin-free gold nanoparticles. *ACS Nano*, **11**, 5519-5529 (2017).
8. Castagnola, V. et al. Towards a classification strategy for complex nanostructures. *Nanoscale Horiz.* **2**,187–198 (2017).

9. Xu, M. et al. How entanglement of different physicochemical properties complicates the prediction of in vitro and in vivo interactions of gold nanoparticles. *ACS Nano* **12**, 10104–10113 (2018).
10. Chithrani, B. D., Ghazani, A. A. & Chan, W. C. Determining the size and shape dependence of gold nanoparticle uptake into mammalian cells. *Nano Lett.* **6**, 662–668 (2006).
11. Wang, J. et al. Physical activation of innate immunity by spiky particles. *Nature Nanotechnol.* **13**, 1078-1086 (2018).
12. Decuzzi, P. et al. Size and shape effects in the biodistribution of intravascularly injected particles. *J. Control. Release* **141**, 320–327 (2010).
13. Rivera Gil, P., Oberdörster, G. N., Elder, A., Puntes, V. C. & Parak, W. J. Correlating physico-chemical with toxicological properties of nanoparticles: the present and the future. *ACS Nano* **4**, 5527–5531 (2010).
14. Tang, M. L., Liu, N., Dionne, J. A. & Alivisatos, A. P. Observations of shape-dependent hydrogen uptake trajectories from single nanocrystals. *J. Am. Chem. Soc.* **133**, 13220–13223 (2011).
15. Glotzer, S. C. Nanotechnology: shape matters. *Nature* **481**, 450 (2012).
16. Libbrecht, K. G. The physics of snow crystals. *Rep. Prog. Phys.* **68**, 855 (2005).
17. Kelly, K. L., Coronado, E., Zhao, L. L. & Schatz, G. C. The optical properties of metal nanoparticles: the influence of size, shape, and dielectric environment. *The J. Phys. Chem. B* **107**, 668–677 (2003).
18. Zhao, Y. et al. Shape-controlled deterministic assembly of nanowires. *Nano Lett.* **16**, 2644–2650 (2016).
19. Nakayama, K., Tanabe, K. & Atwater, H. A. Plasmonic nanoparticle enhanced light absorption in GaAs solar cells. *Appl. Phys. Lett.* **93**, 121904 (2008).
20. Horton, M. R., Manley, S., Arevalo, S. R., Lobkovsky, A. E. & Gast, A. P. Crystalline protein domains and lipid bilayer vesicle shape transformations. *J. Phys. Chem. B* **111**, 880–885 (2007).
21. Yang, P. & Lieber, C. M. Nanorod-superconductor composites: a pathway to materials with high critical current densities. *Science* **273**, 1836–1840 (1996).
22. Osberg, K. D. et al. Dispersible gold nanorod dimers with sub-5 nm gaps as local amplifiers for surface-enhanced Raman scattering. *Nano Lett.* **12**, 3828–3832 (2012).
23. Horsch, M. A., Zhang, Z. & Glotzer, S. C. Self-assembly of polymer-tethered nanorods. *Phys. Rev. Lett.* **95**, 056105 (2005).
24. Nikoobakht, B. & El-Sayed, M. A. Preparation and growth mechanism of gold nanorods (NRs) using seed-mediated growth method. *Chem. Mater.* **15**, 1957–1962 (2003).
25. Murphy, C. J. et al. Gold nanoparticles in biology: beyond toxicity to cellular imaging. *Acc. Chem. Res.* **41**, 1721–1730 (2008).
26. Durr, N. J. et al. Two-photon luminescence imaging of cancer cells using molecularly targeted gold nanorods. *Nano Lett.* **7**, 941–945 (2007).
27. Grabinski, C. et al. Effect of gold nanorod surface chemistry on cellular response. *ACS Nano* **5**, 2870–2879 (2011).
28. Ye, X. et al. Improved size-tunable synthesis of monodisperse gold nanorods through the use of aromatic additives. *ACS Nano* **6**, 2804–2817 (2012).
29. Lin, H. et al. Clathrate colloidal crystals. *Science* **355**, 931–935 (2017).
30. Grzelczak, M., Pérez-Juste, J., Mulvaney, P. & Liz-Marzán, L. M. Shape control in gold nanoparticle synthesis. *Chem. Soc. Rev.* **37**, 1783–1791 (2008).
31. Cozzoli, P. D. et al. Colloidal synthesis and characterization of tetrapod-shaped magnetic nanocrystals. *Nano Lett.* **6**, 1966–1972 (2006).
32. Ahmadi, T. S., Wang, Z. L., Green, T. C., Henglein, A. & El-Sayed, M. A. Shape-controlled synthesis of colloidal platinum nanoparticles. *Science* **272**, 1924 (1996).
33. Sun, Y. & Xia, Y. Shape-controlled synthesis of gold and silver nanoparticles. *Science* **298**, 2176–2179 (2002).
34. Xia, Y. & Halas, N. J. Shape-controlled synthesis and surface plasmonic properties of metallic nanostructures. *MRS Bull.* **30**, 338–348 (2005).
35. Li, N., Zhao, P. & Astruc, D. Anisotropic gold nanoparticles: synthesis, properties, applications, and toxicity. *Angew. Chem. Int. Ed.* **53**, 1756–1789 (2014).
36. Lee, H.-E. et al. Amino-acid-and peptide-directed synthesis of chiral plasmonic gold nanoparticles. *Nature* **556**, 360 (2018).
37. Han, D. et al. DNA origami with complex curvatures in three-dimensional space. *Science* **332**, 342–346 (2011).
38. Gentleman, D. J. & Chan, W. C. A systematic nomenclature for codifying engineered nanostructures. *Small* **5**, 426–431 (2009).
39. Bigdeli, A., Hormozi-Nezhad, M. R., Jalali-Heravi, M., Abedini, M. R. & Sharif-Bakhtiar, F. Towards defining new nano-descriptors: extracting morphological features from transmission electron microscopy images. *RSC Adv.* **4**, 60135–60143 (2014).
40. Rice, K. P., Saunders, A. E. & Stoykovich, M. P. Classifying the shape of colloidal nanocrystals by complex Fourier descriptor analysis. *Cryst. Growth. Des.* **12**, 825–831 (2012).
41. Miernicki, M., Hofmann, T., Eisenberger, I., von der Kammer, F. & Praetorius, A. Legal and practical challenges in classifying nanomaterials according to regulatory definitions. *Nature Nanotechnol.* **14**, 208 (2019).
42. Ling, H. & Jacobs, D. W. Shape classification using the inner-distance. *IEEE T Pattern Anal. Mach. Intell.* **29**, 286–299 (2007).
43. Liang, H. et al. Asymmetric silver "nanocarrot" structures: Solution synthesis and their asymmetric plasmonic resonances. *J. Am. Chem. Soc.* **135**, 9616–9619 (2013).
44. Yuan, H. et al. Gold nanostars: surfactant-free synthesis, 3D modelling, and two-photon photoluminescence imaging. *Nanotechnology* **23**, 075102 (2012).
45. Zhao, L. et al. Formation and stability of gold nanoflowers by the seeding approach: the effect of intraparticle ripening. *J. Phys. Chem. C* **113**, 16645–16651 (2009).
46. Bakr, O. M., Wunsch, B. H. & Stellacci, F. High-yield synthesis of multi-branched urchin-like gold nanoparticles. *Chem. Mater.* **18**, 3297–3301 (2006).
47. Wang, R. et al. Experimental determination and analysis of gold nanorod settlement by differential centrifugal sedimentation. *RSC Adv.* **6**, 43496–43500 (2016).
48. Carney, R. P. et al. Determination of nanoparticle size distribution together with density or molecular weight by 2D analytical ultracentrifugation. *Nat. Commun.* **2**, 335 (2011).
49. Sharma, V., Park, K. & Srinivasarao, M. Shape separation of gold nanorods using centrifugation. *Proc. Natl Acad. Sci. USA* **106**, 4981–4985 (2009).
50. Brown, P. H. & Schuck, P. Macromolecular size-and-shape distributions by sedimentation velocity analytical ultracentrifugation. *Biophys. J.* **90**, 4651–4661 (2006).
51. Potenza, M. et al. Detecting the shape of anisotropic gold nanoparticles in dispersion with single particle extinction and scattering. *Nanoscale* **9**, 2778–2784 (2017).
52. Finney, E. E. & Finke, R. G. Nanocluster nucleation and growth kinetic and mechanistic studies: a review emphasizing transition-metal nanoclusters. *J. Colloid Interf. Sci.* **317**, 351–374 (2008).
53. Frens, G. Controlled nucleation for the regulation of the particle size in monodisperse gold suspensions. *Nature* **241**, 20–22 (1973).
54. Thanh, N. T., Maclean, N. & Mahiddine, S. Mechanisms of nucleation and growth of nanoparticles in solution. *Chem. Rev.* **114**, 7610–7630 (2014).
55. Turkevich, J., Stevenson, P. C. & Hillier, J. A study of the nucleation and growth processes in the synthesis of colloidal gold. *Discuss. Faraday Soc.* **11**, 55–75 (1951).
56. Gast, A. P. & Monovoukas, Y. A new growth instability in colloidal crystallization. *Nature* **351**, 553 (1991).
57. Jana, N. R., Gearheart, L. & Murphy, C. J. Seed-mediated growth approach for shape-controlled synthesis of spheroidal and rod-like gold nanoparticles using a surfactant template. *Adv. Mater.* **13**, 1389 (2001).
58. Bals, S., Goris, B., Liz-Marzán, L. M. & Van Tendeloo, G. Three-dimensional characterization of noble-metal nanoparticles and their assemblies by electron tomography. *Angew. Chem. Int. Ed.* **53**, 10600–10610 (2014).
59. Levin, B. D. et al. Nanomaterial datasets to advance tomography in scanning transmission electron microscopy. *Sci. Data* **3**, 160041 (2016).
60. Midgley, P. A., Ward, E. P., Hungría, A. B. & Thomas, J. M. Nanotomography in the chemical, biological and materials sciences. *Che.Soc. Rev.* **36**, 1477–1494 (2007).
61. Bowman, E. T., Soga, K. & Drummond, W. Particle shape characterisation using Fourier descriptor analysis. *Geotechnique* **51**, 545–554 (2001).
62. Sokic, E. & Konjicija, S. in *Content-Based Multimedia Indexing (CBMI), 2014 12th International Workshop on Content-Based Multimedia Indexing (CBMI)*. 1–4 (IEEE).
63. Vranic, D. & Saupe, D. 3D Shape Descriptor Based n 3D Fourier Transform in *Proceedings of the EURASIP Conference on Digital Signal Processing for Multimedia Communications and Services, Budapest, Hungary*, pp. 271–274, (2001).
64. Zhang, D. & Lu, G. A comparative study of curvature scale space and Fourier descriptors for shape-based image retrieval. *J. Visual Commun. Image R.* **14**, 39–57 (2003).
65. Zhang, D. & Lu, G. Shape-based image retrieval using generic Fourier descriptor. *Signal Process. Image Commun.* **17**, 825–848 (2002).
66. Zhang, H. & Fiume, E. in *Shape Modeling International, 2002. Proceedings.* 261-268 (IEEE).
67. Li, J. et al. Controllable synthesis of stable urchin-like gold nanoparticles using hydroquinone to tune the reactivity of gold chloride. *J. Phys. Chem. C* **115**, 3630–3637 (2011).
68. Suzuki, S. Topological structural analysis of digitized binary images by border following. *Comput. Vis. Graph. Image Process.* **30**, 32–46 (1985).

Acknowledgements

This work was supported by Science Foundation Ireland under SIRG (Funding Agency Ref. No: 15/SIRG/3423) (Y.Y., V.A.G.) and SFI-NSFC Partnership (Funding Agency Ref. No: 17/NSFC/4898) (K.A.D.). K.A.D. acknowledges the funding of Guangdong Provincial Education Department Key Laboratory of Nano-Immunoregulation Tumor

Microenvironment (2019KSYS008). H. L., V.C., and Q.C. acknowledge financial support from the Irish Research Council (H.L., EPSPD/2015/5; V.C., GOIPD/2016/128; Q.C., GOIPG/2014/874). Q.C. and W.Z and J.L. acknowledge Chinese Scholarship Council (agreement no. 201408300003 and 201706220062 and 201508300009). J.M.A. acknowledges CNPq for support through Bolsa de Produtividade em Pesquisa No. 313431/2018-3. The work described in this article, prior to reaching this level, evolved over some years with contributions from Authors and other previous researchers at the Center for Bio-Nano Interactions (CBNI), University College Dublin. They were helpful in bringing the work to the present stage of development. Inge Nelissen performed the gene array measurements. Marco Monopoli and Zeljka Krpetic were involved in early phases when syntheses and dispersions of shape particles were quite irreproducible, and helped parameterize the nature of the problems involved. Marco Monopoli played a significant role in the early thinking on dispersion stability. Camila P. Silveira also played a significant role in collaboration with Q.C. in unravelling various aspects of the irreproducibility. More recently, James Behan played a significant role, working with L.B. and V.C., in supporting the transition to fully reproducible syntheses. Ban Zhan (visiting student to CBNI) collaborated with Y.Y. on running the PCoA analysis. Zengchun Xie and Eoin O'Neill both extracted and analyzed contours.

Author contributions

Y.Y. and K.A.D. conceived and developed the overall concept of shape specification and its relationship to biology, developed the methods and approaches, lead the overall project. Y.Y. lead the project overall, including biological development, established working methods for the materials in biology, interpreted and supervised the experiments. Both wrote the paper. L.B. and V.C. helped develop the ideas behind the work and played a leading role in that. V.C. was central to design of overall approach to synthesis, including implementing numerous technical developments that lead to reproducibility, supervised the synthesis experiments for part of the projects evolution, and wrote the paper. V.C. and L.B. supported and carried out some of the synthetic and characterization experiments, and Q.C. persisted with many of the key experiments during the transition to more reproducible phase of synthesis, during which the problems were more fully clarified. L.B. and W.Z. worked together to achieve highly reproducible syntheses, and W.Z. finally achieved almost complete reproducibility for the shapes reported. All three extracted contours of nanoparticles. H.L. was central to the design and writing the codes, and perform the computational experiments, and contributed to the paper. H.L, A.M. and J.M.A. helped design some of the computational experiments and interpretations. H.L, A.M., and J.M.A. tested the reproducibility of the computational analysis. All of the above contributed to the writing of the paper. V.A.G. and J.L. independently extracted the RNA, and worked with Y.Y. to reproduce the biological data. J.C. performed the 3D tomography.

Competing interests

The authors declare no competing interests.

Additional information

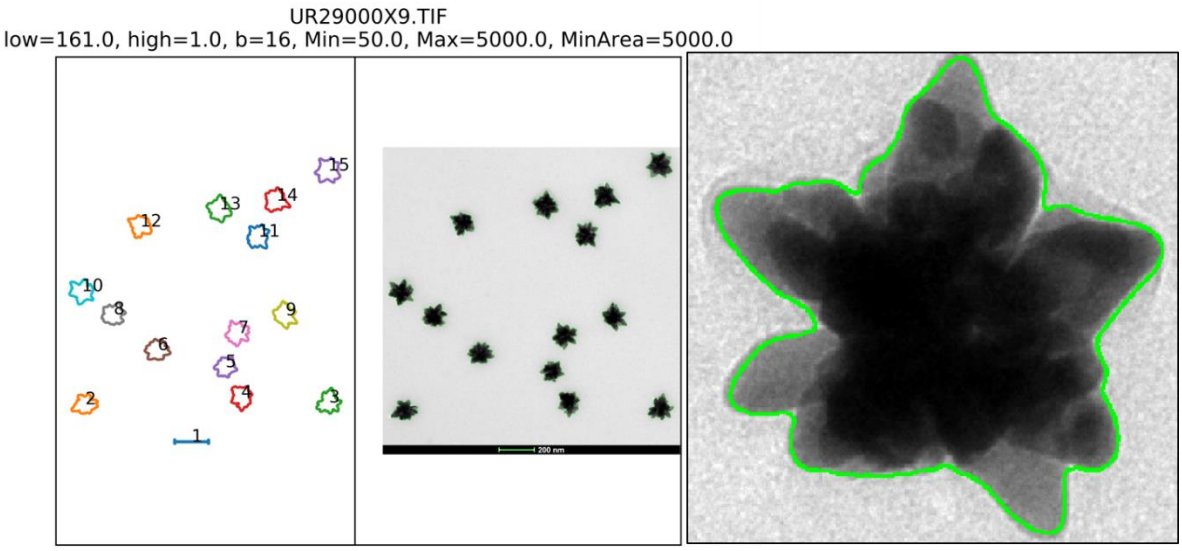
Supplementary information is available for this paper at https://doi.org/10.1038/s43246-020-0033-2.

Correspondence and requests for materials should be addressed to V.C., Y.Y. or K.A.D.

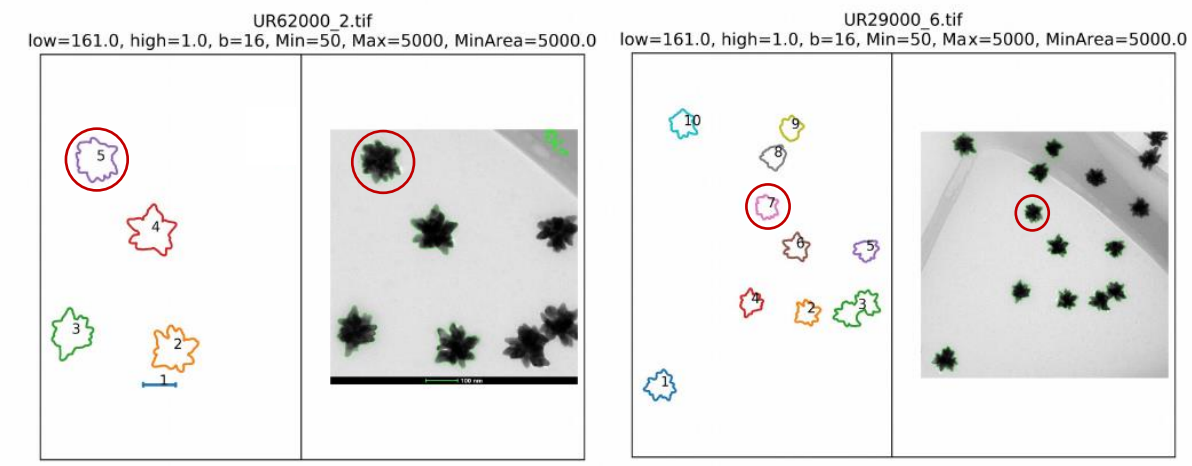
Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

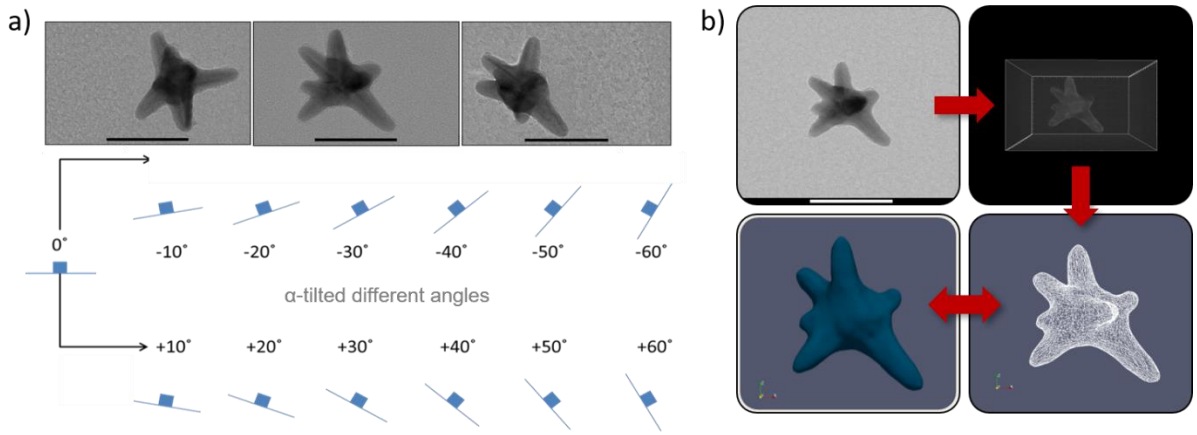
Supplementary figures



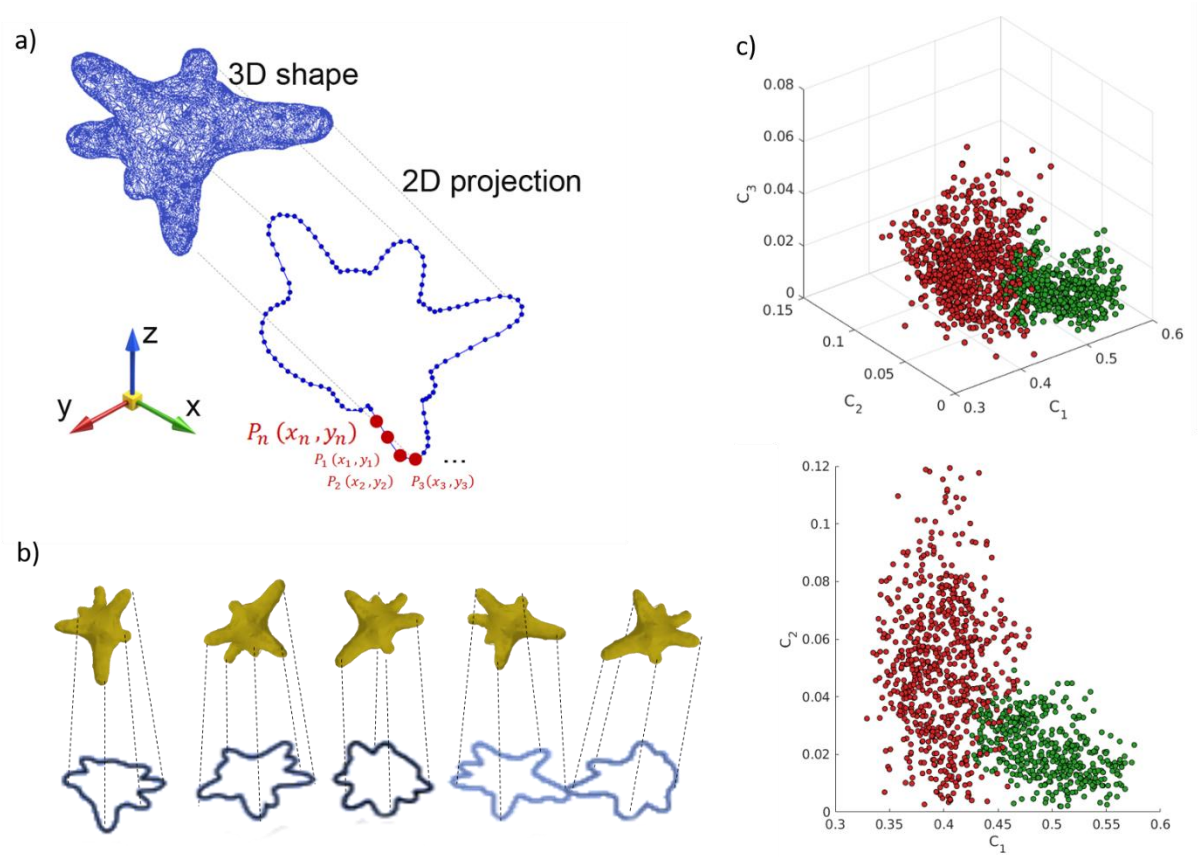
Supplementary Figure 1. Contour extraction resulting from the in-house developed Python script (see Methods). On the right magnification of a single particle. The image generated shows a series of numbered NPs contours adjacent to the original TEM image. The numbers correspond to a series of files containing the xy coordinates of each NP contour. If a contour generated does not correspond to a NP, as in the case of contour 1 (scale bar), the corresponding xy file can be simply identified and removed from the database.



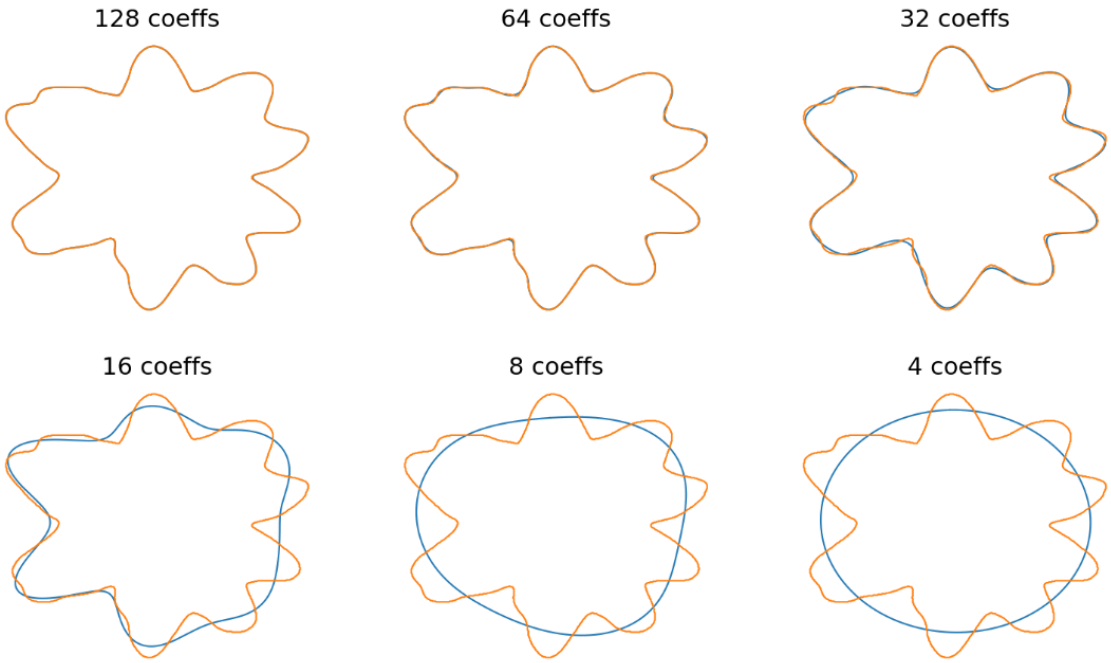
Supplementary Figure 2. Contours extracted for the same GNP (belonging to GNP3 group) from different TEM magnifications: 62000x (left) and 29000x (right). The same particle contour label as 5 on the left has 2415 points while the contour label as 7 on the right has 1068. The ratio of the number of points (2.3) corresponds to the ratio of the magnification (2.1). Different image format (.tiff, .jpeg and .png) did not influence the number of points obtained for each contour.



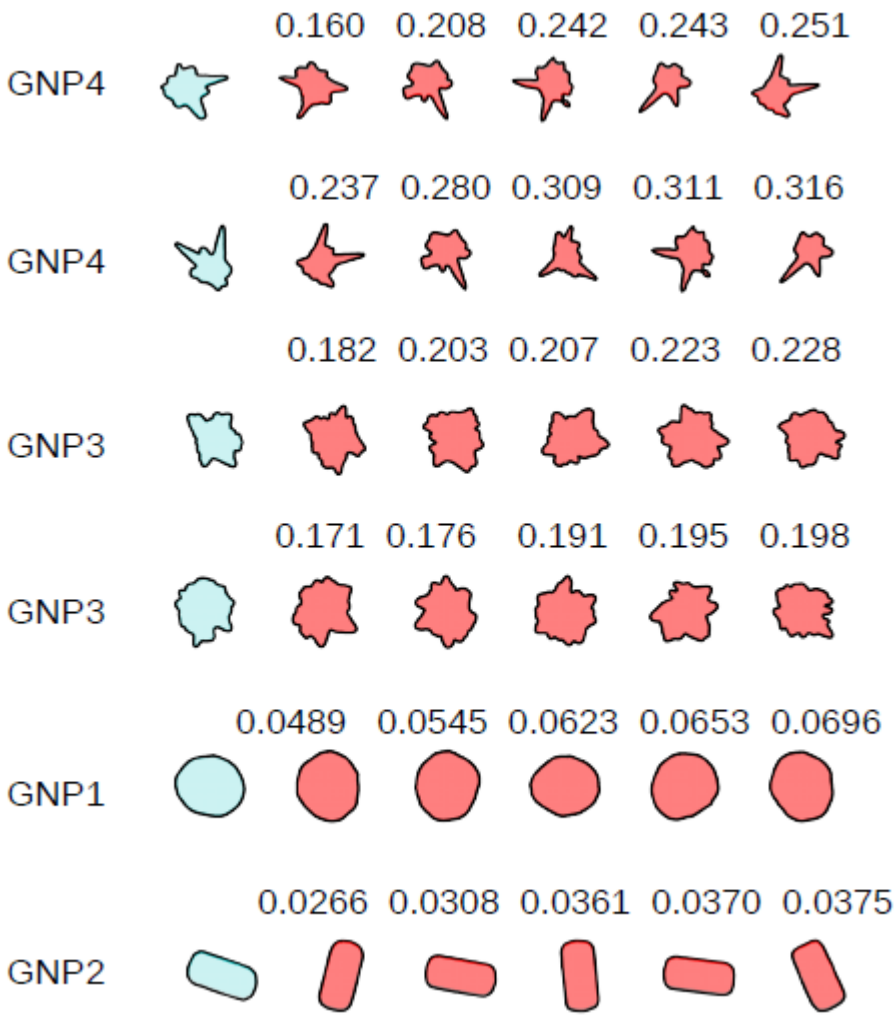
Supplementary Figure 3. Tomography protocol. a) TEM micrograph of the very same NP at a set of α-tilted different angles and scheme of the tilting angles used for the TEM tomography process to reconstruct the 3D model. b) Example of workflow of the TEM Tomography protocol starting with 2D TEM micrograph of the NP at 0˚ and visualisation of the NP in a 3D volume with Imaris (before being decimated with EM3D prior to segmentation), the resultant 3D model and wireframe model after segmentation with TurtleSeg which are interchangeable depending on function (see Supplementary Methods for more details).



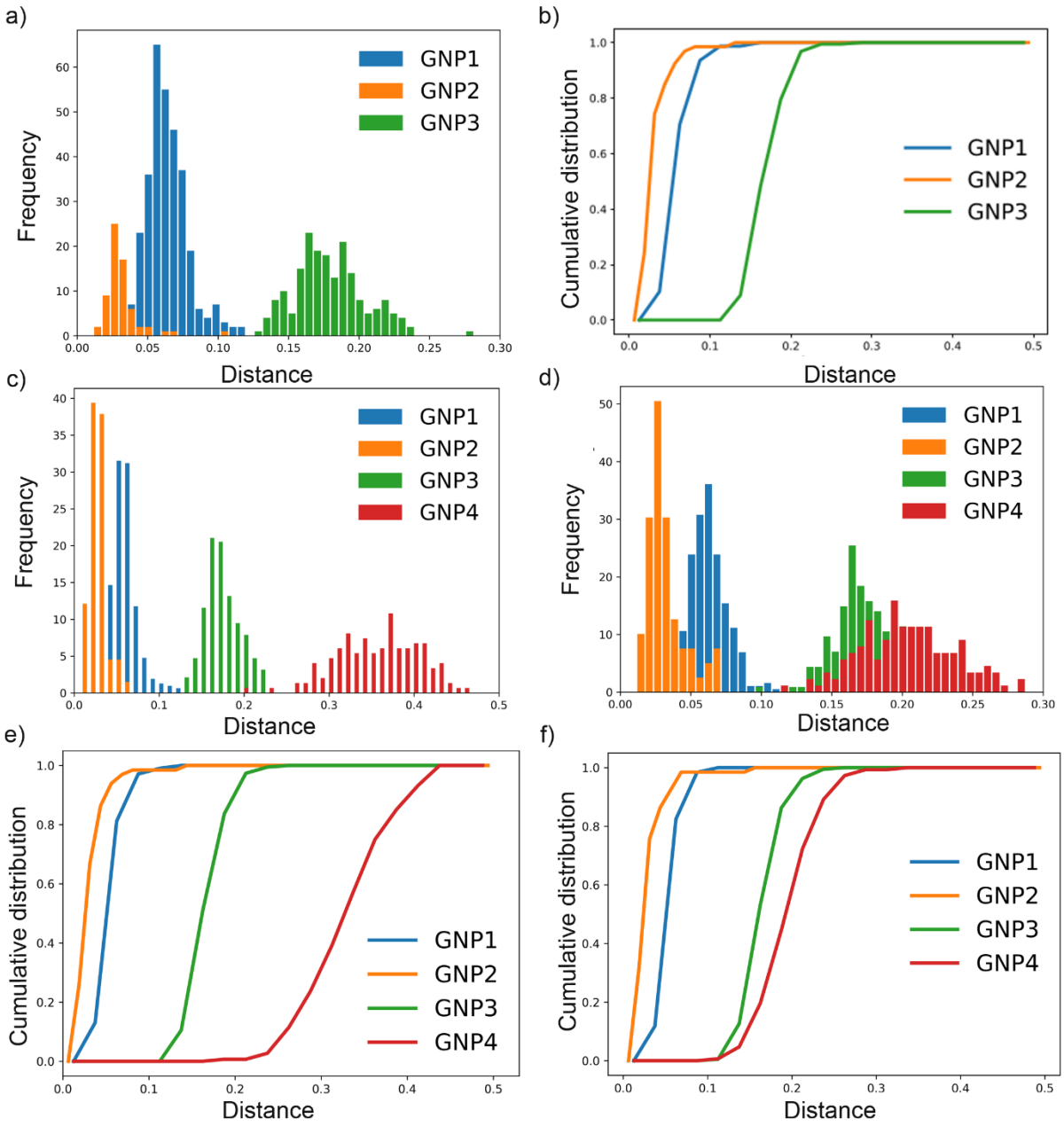
Supplementary Figure 4. 2D projections of 3D tomography reconstruction. a) Schematic of a 2D contour generated from a 3D object (GNP4 is taken as example). b) Examples of five 2D projections generated from 3D reconstruction of GNP4 as captured by TEM tomography. The 2D projections used are generated from four 3D reconstructions of GNP3 and six 3D reconstructions of GNP4. 120 projections for each 3D reconstruction were used for a total of 1200 projections. c) Scatter plot presenting the cluster analysis using the first 3 (top) or 2 (bottom) Fourier coefficients. Each dot represents one contour. The cluster plots show that two different populations were generated, and two shape groups can be distinguished based on 2D projections.



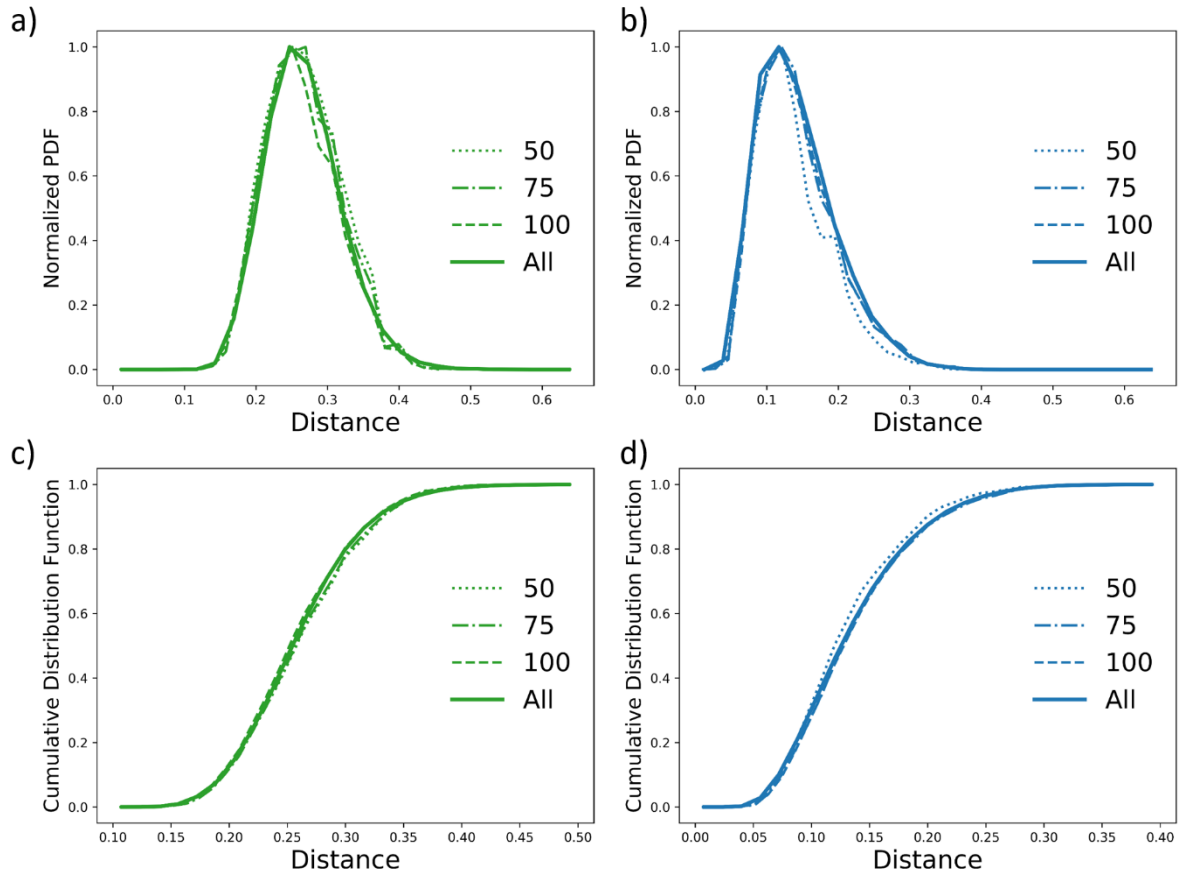
Supplementary Figure 5. Contour reconstruction from DFT coefficients. The reconstructions were performed after setting some coefficients to zero (higher indices) and keeping only a certain number of non-zero coefficients (the first "frequencies"). About 128 coefficients are needed to obtain an accurate reconstruction. Using more than 128 coefficients does not noticeably change the distance between contours.



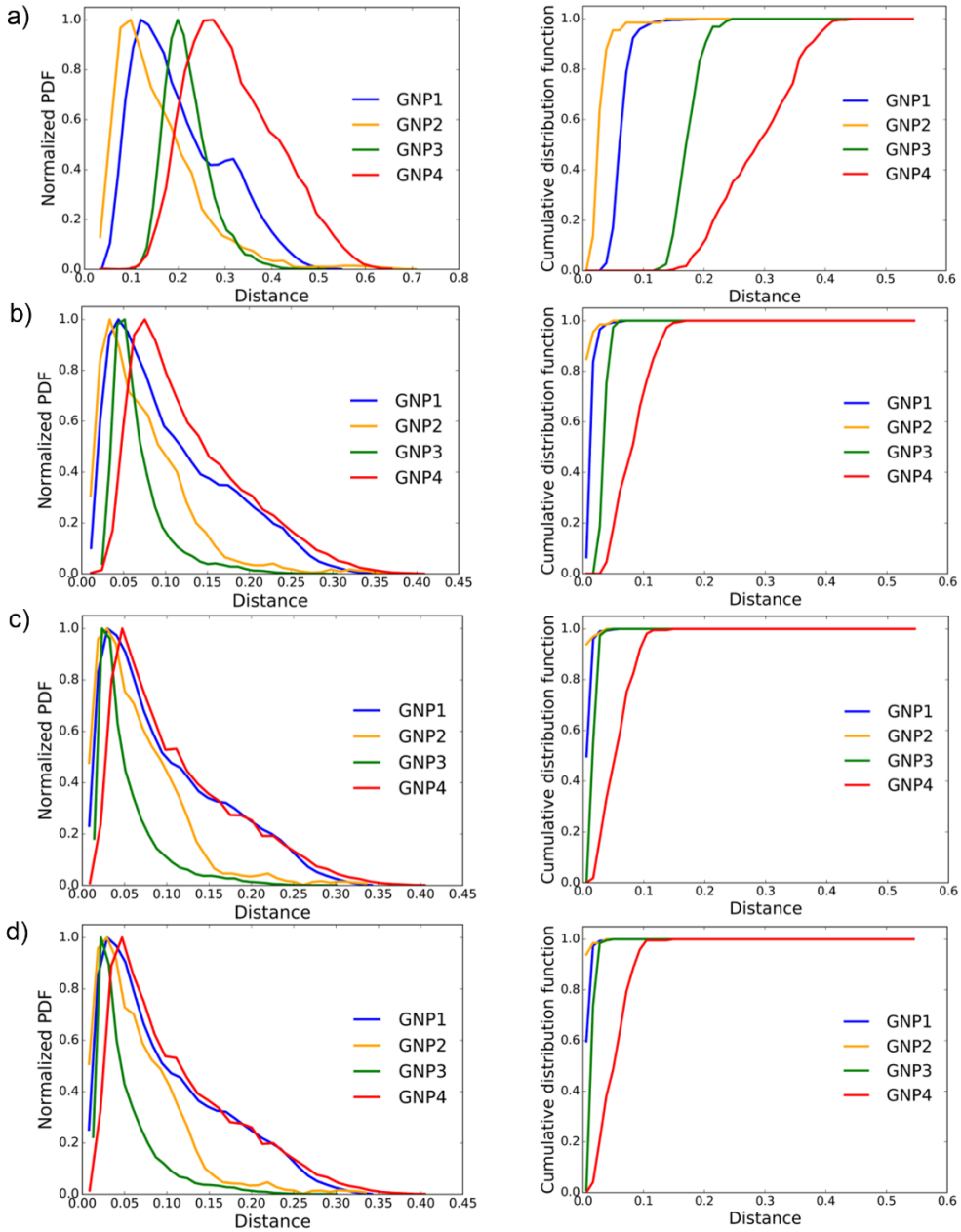
Supplementary Figure 6. Examples of the NPs contours (in red) related to the shortest distances calculated between a given NPs (contour in light blue) used as test belonging to the shape groups GNP1-4. Distance values are displayed above the NPs contours.



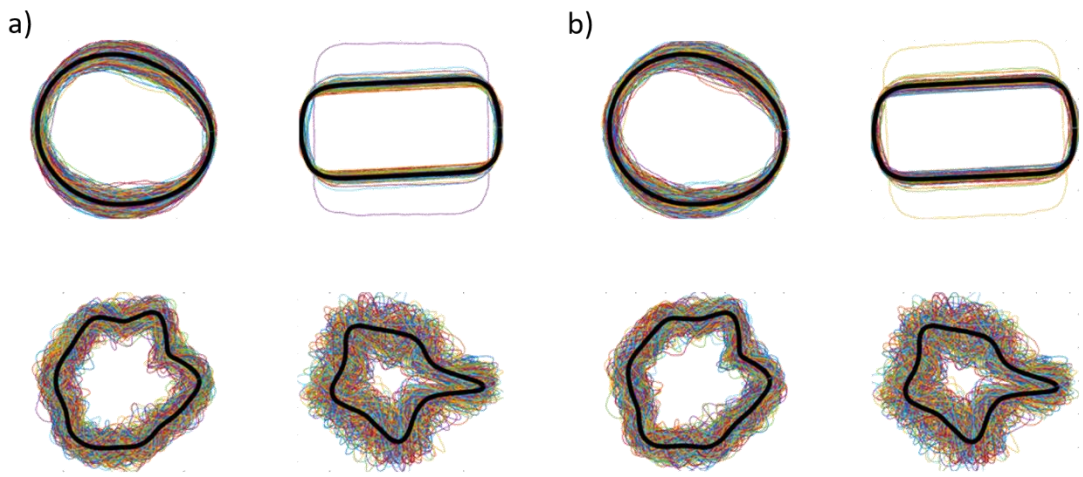
Supplementary Figure 7. a) Frequency distribution and b) cumulative probability distribution of the minimum distances for specific batches of NPs (belonging to different shape groups) presented to the database. Specifically, the number of NPs in the test batches were: GNP1 = 314, GNP2 = 66 and GNP3 = 190. c) Frequency distribution and e) cumulative probability distribution of the minimum distances when GNP4 is not present in the database. d) Frequency distribution and f) cumulative probability distribution of the minimum distances when GNP4 is included in the database.



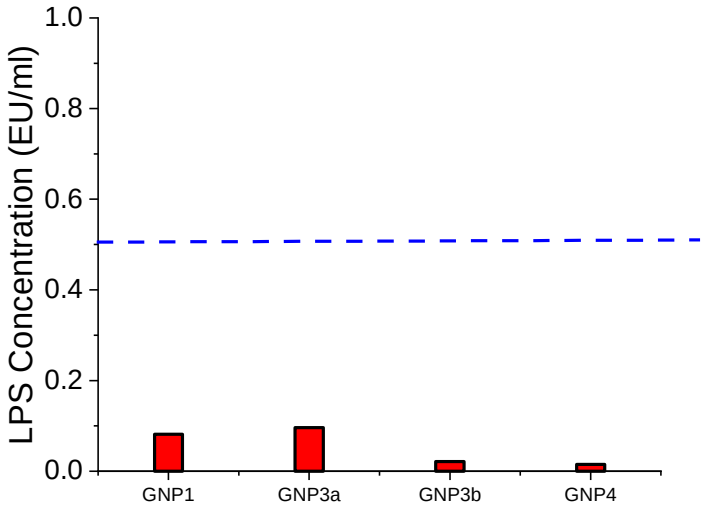
Supplementary Figure 8. Convergence plots for two representative shape groups. Probability distribution of shape variability for increasing number of a) GNP1 and b) GNP3 in a batch. In both cases the distributions converge when more than 100 NPs are used. However increasing number of NPs (above 100) reduces the noise. Cumulative distribution for c) GNP1 and d) GNP3 for increasing number of NPs in the database.



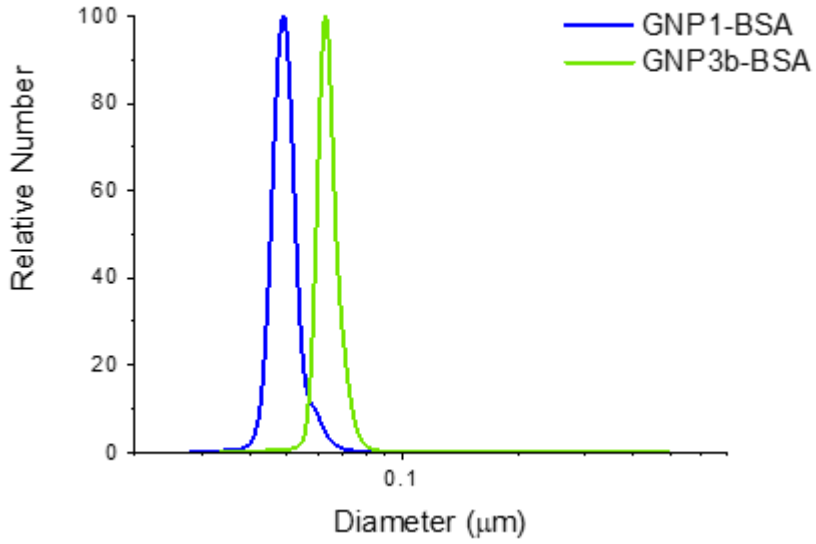
Supplementary Figure 9. Normalized probability density function (PDF - left) and cumulative probability distribution (right) of the minimum distance for 40 test NPs (belonging to different shape groups) presented to a database of 120 objects and calculated based on a) $L_1\,norm$, b) $L_2 norm$, c) $L_5 norm$ and d) $L_{10} norm$ metrics.



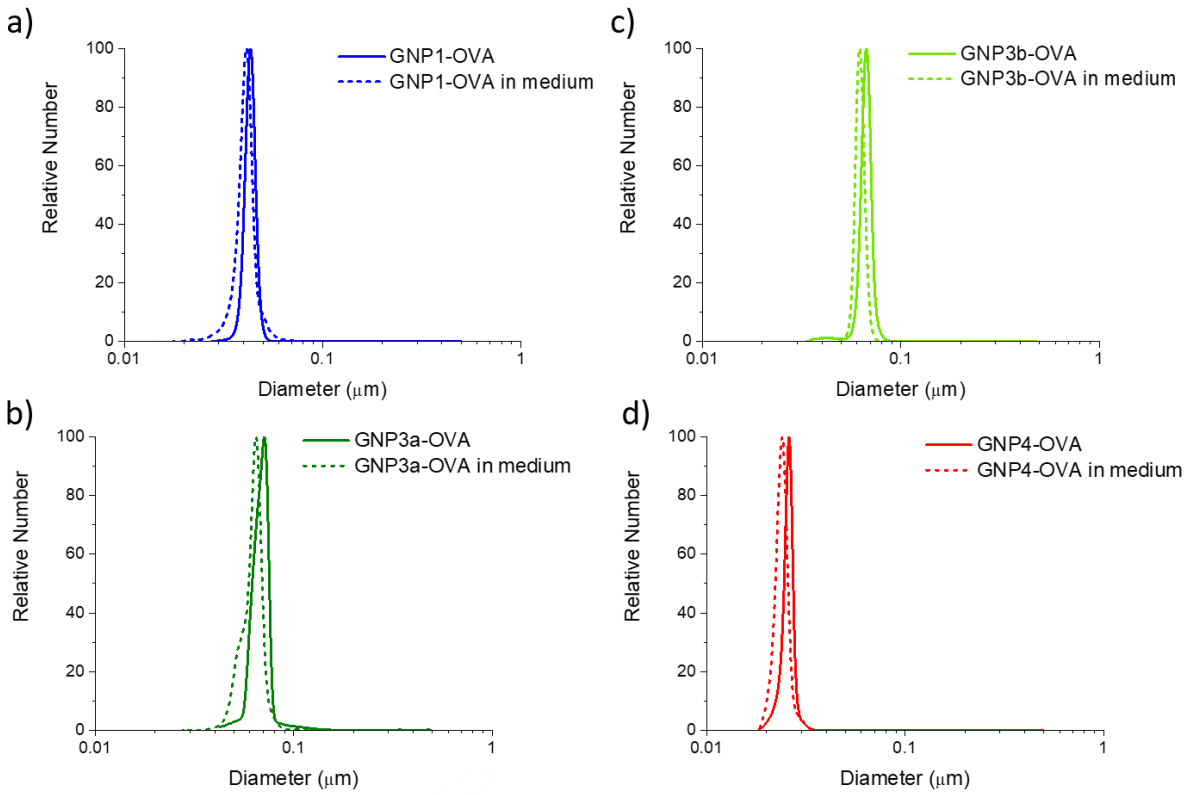
Supplementary Figure 10. Average shape reconstruction for the 4 GNPs using resampling of the original contours with a) 1024 and b) 512 contour points. The resampling is obtained by placing the desired number of points equally spaced over the original curve. With these resampled curves the averaging procedure is performed, using all particles, and aligning them on Fourier space by phase-shifting (see Methods).



Supplementary Figure 11. Quantification of LPS in the GNP dispersions by Limulus Amebocyte Lysate (LAL) assay. The concentration of GNPs in the LAL assay was at 1 x 10^{11} NPs/mL, which is 3.3 times higher than the concentration used for transcriptome experiment.



Supplementary Figure 12. Differential centrifugal sedimentation (DCS) analysis of GNP1 and GNP3b coated bovine serum albumin (BSA).



Supplementary Figure 13. DCS analysis of GNP1, GNP3a, GNP3b, and GNP4 coated ovalbumin (OVA), as well as the OVA-coated GNP1-4 dispersed in dendritic cell culture medium.

Supplementary Tables

Supplementary Table 1. Accuracy of assignment for objects belonging to GNP1-4 shape groups using both real and imaginary part of the Fourier coefficients (full DFT), the magnitude of the Fourier coefficients after resampling (Norm DFT – resampling) or without resampling (Norm DFT).

	Full DFT	Norm DFT - resampling	Norm DFT
Groups	Accuracy %	Accuracy %	Accuracy %
GNP1	100.0	99.1	99.0
GNP2	99.1	100.0	100.0
GNP3	99.6	97.4	98.9
GNP4	94.5	83.7	99.3

Supplementary Table 2. PCA explained variances for each of the first 5 PCs. In parentheses are the relative percentages to the total variance.

	PC_1	PC_2	PC_3	PC_4	PC_5
$\hat{\zeta}$	2.00 (70.04%)	0.38 (13.38%)	0.11 (3.80%)	0.06 (2.00%)	0.04 (1.43%)
$\lvert\hat{z}\rvert$	2.02 (70.73%)	0.29 (10.06%)	0.04 (1.42%)	0.03 (0.91%)	0.02 (0.60%)
C_0	1.90 (66.52%)				
C_1	0.40 (14.42%)				
	Total variance = 2.85				

Supplementary Methods

NPs synthesis

The 4 NPs shape groups used in this work were synthetized according to previously reported protocols.[1,2]

For biological investigation, all NPs used were synthesized in endotoxin-free condition under a B2 laminar fumehood. All the reagents were only opened under laminar fumehood and all the glassware cleaned with aqua regia and washed multiple times with Ultrapure water (Optima, Fisher chemical) prior each synthesis. All the plastic consumables used were non pyrogenic.

After synthesis, the NPs were suspended in a solution of 1 mg/mL of endotoxin-free OVALBUMIN (Endograde®), then washed by centrifugation (550 rpm for 20 min) to remove the excess of protein. The endotoxin level of the nanoparticle dispersion was confirmed by LAL assay as previously described.[1]

Sample preparation for TEM imaging and tomography

To obtain reliable contours of the NPs on the TEM grid, the samples were diluted to the concentration of 1-3x10^{11} NPs/mL and incubated at R.T. in 0.2% methylcellulose. Ultrathin carbon on holey 400 mesh Cu grids (Ted Pella Inc.) were used. The grids were previously treated with a glow discharger and 0.5-1 µL of NPs solution was drop casted on the grid. TEM analyses were performed on a FEI Tecnai G2 20 Twin operating at 120 kV. Images were taken between 29000x and 62000x magnification.

TEM Tomography protocol

3D reconstruction from TEM tomography was performed for the four types of GNP.[3] Briefly, the tomography and reconstruction process involved acquiring a number of micrographs of the same object and subsequently reconstructing the projections to create a 3D volume encapsulating the object. The input for the tomography process was a set of α-tilted 2D images obtained from TEM with the output as a 3D volume. Amongst the various software available, eTomo, a graphical user interface included in the IMOD Tomography Package,[4,5] was employed.

Single-axis geometry was completed in the tilt series, containing 121 micrographs obtained from the angles -60° to +60°. There are three main steps within the eTomo for the calculations of the 3D volume: image alignment, reconstruction and visualisation. The image alignment was necessary due to the specimen shift from the instrumental limitations. For the reconstruction, the weighted back-projection (WBP) algorithm was employed in this work. To segment an object inside a 3D volume TurtleSeg software[6,7] was employed. The segmentation process software can be done in a semi-automatic fashion.

References

1 Talamini, L. *et al.* Influence of size and shape on the anatomical distribution of endotoxin-free gold nanoparticles. *ACS nano*, 5519-5529 (2017).

2 Potenza, M. *et al.* Detecting the shape of anisotropic gold nanoparticles in dispersion with single particle extinction and scattering. *Nanoscale* **9**, 2778-2784 (2017).

3 Bals, S., Goris, B., Liz-Marzán, L. M. & Van Tendeloo, G. Three-Dimensional Characterization of Noble-Metal Nanoparticles and their Assemblies by Electron Tomography. *Angewandte Chemie International Edition* **53**, 10600-10610 (2014).

4 Kremer, J. R., Mastronarde, D. N. & McIntosh, J. R. Computer visualization of three-dimensional image data using IMOD. *Journal of structural biology* **116**, 71-76 (1996).

5 Mastronarde, D. Tomographic reconstruction with the IMOD software package. *Microscopy and Microanalysis* **12**, 178-179 (2006).

6 Hamarneh, G., Yang, J., McIntosh, C. & Langille, M. in *Medical imaging.* 1597-1603 (International Society for Optics and Photonics).

7 Memiş, A., Albayrak, S. & Bilgili, F. in *2018 Medical Technologies National Congress (TIPTEKNO).* 1-4 (IEEE).