Machine Learning Predicts the Functional Composition of the Protein Corona and the Cellular Recognition of Nanoparticles

Zhan Ban[#], Peng Yuan[#], Fubo Yu, Ting Peng, Qixing Zhou, Xiangang Hu*

Key Laboratory of Pollution Processes and Environmental Criteria (Ministry of Education)/Tianjin Key Laboratory of Environmental Remediation and Pollution Control, College of Environmental Science and Engineering, Nankai University, Tianjin 300350, China

# Zhan Ban and Peng Yuan contributed equally to this paper.

* Corresponding author: Xiangang Hu, huxiangang@nankai.edu.cn

Fax: 0086-022-23507800

Tel.: 0086-022-23507800

METHODS AND EXPERIMENTS

Data Extraction

The literature references were obtained from the PubMed, Scopus and ISI Web of Knowledge databases on November 26, 2017, with the following search criteria: (nanoparticle[Title/Abstract] OR nanomaterial[Title/Abstract] OR "quantum dot"[Title/Abstract] OR nanosphere[Title/Abstract] OR fullerene[Title/Abstract] OR nanodiamond[Title/Abstract] OR nanowire[Title/Abstract] OR nanotube[Title/Abstract] OR MWCNT[Title/Abstract] OR nanofiber[Title/Abstract] OR graphene[Title/Abstract] OR nanocrystal[Title/Abstract] OR nanopowder[Title/Abstract] OR nanosheet[Title/Abstract] OR "metal organic framework"[Title/Abstract] OR MOF[Title/Abstract] OR nanocapsule[Title/Abstract] OR dendrimer[Title/Abstract] OR nanocomposite[Title/Abstract]) AND (protein[Title/Abstract] AND corona[Title/Abstract]). There were no search limitations on the publication date or language. With this extensive search, 2634 publications were obtained, including 1065 duplications with the same title, authors and publication date. A total of 719 studies were removed because of a lack of relevance to the topic of the work. Finally, 56 papers[1-56] satisfied the following selection criteria and were selected for further analyses: (i) the full text was available online; (ii) the topic was the protein corona of nanoparticles (NPs); (iii) the protein corona was formed in blood, serum, plasma or cell culture medium; (iv) the paper

provided information about the physicochemical properties of the NPs (e.g., the core/shell, shape, and surface ligands or other modifications); (v) the paper provided information about protein corona formation in healthy plasma, serum, blood or cell culture medium (e.g., exposure time, plasma concentration and species of the plasma or blood); (vi) the paper provided information about the isolation of the protein corona from NPs by centrifugal capture (e.g., centrifugation time and relative centrifugation force); (vii) the paper estimated the relative protein abundance (RPA) with normalized spectral counts (NSCs) of proteins in the protein corona by the following equation:

$$MWNSC_K = \frac{(NSC/MW)_K}{\sum_{i=1}^{n}(NSC/MW)_i} \times 100\%$$

where $MWNSC_k$ is the percentage molecular weight normalized by the NSC for protein *k,* and the molecular weight (MW) is in kDa; (viii) the paper listed the protein names and the RPA (or spectral count or NSC) of each protein with an RPA greater than 1%. This work extracted 21 important factors related to NP properties (NP type, NP core, surface modification, modification type, size measured by transmission electron microscopy (TEM) ($size_{TEM}$), size measured by dynamic light scattering (DLS) ($size_{DLS}$), zeta potential, polydispersity index (PDI, NP shape, dispersion medium, dispersion medium pH), corona formation (protein source, incubation culture, incubation plasma concentration, incubation NP concentration, incubation time, incubation temperature), and corona isolation (centrifugation speed,

centrifugation time, centrifugation temperature and centrifugation repetitions). The DLS technique is well known to struggle with providing meaningful data on aggregated particles, and $size_{TEM}$ instead of $size_{DLS}$ was used in corona composition prediction models. To control the heterogeneity and reveal the adsorption patterns of numerous proteins on various NPs, the following treatment methods were applied to the dataset.

Special Data Treatment

The present work calculated the missed RPA values of proteins in the literature publications using equation 1. The information (e.g., protein description, protein entry, length, MW, keywords and sequence analysis) on individual proteins (with RPA greater than 1%) was extracted from UniProt.org. The theoretical isoelectric point (pI) of proteins was calculated on web.expasy.org, and the grand average of hydropathicity (GRAVY) score, aliphatic index and cysteine content of proteins were calculated using the ProtParam tool at https://web.expasy.org/protparam/. To compare the protein adsorption of various NPs from different studies, the analysis transformed the proteins from different organisms (e.g., rat, mouse and cattle) to humans. Finally, 593 individual proteins and 652 data pieces were collected for individual prediction models. To further evaluate the attachment mechanism of proteins on NPs, this work classified the overall proteins in terms of their protein properties (e.g., MW, pI, length and GRAVY). According to the molecular properties and biological function of

proteins identified from the UniProt database, the functional composition was classified as immunity proteins, apolipoproteins, complement proteins, coagulation proteins, clusterin and other proteins. Because some studies did not list the RPA values of all proteins, the analysis selected the data on proteins with an overall RPA value of more than 70% and normalized the RPA values of each class according to the overall RPA value of each data point. Finally, 567 data points and 73 protein classes were collected for composition prediction models.

For the selected factors affecting the components of the protein corona, the present work converted special data points into common ones that could be made available as training input for the machine learning model (random forest, RF). For the NP modifications, the analysis combined some modifications as listed in Table S1 because the RF model could not handle categorical factors with more than 50 categories. The missing values for NP shape were treated as spheres. For other NP shapes (e.g., tube and sheet), the widths and thicknesses were recorded. The "overnight" incubation time and the indoor temperature were set to 12 h and 25℃, respectively. Regarding other factors with a small number of missing values, the present work replaced the missing value with the most frequent value (e.g., phosphate-bu□ered saline (PBS) for solution and water for other solution cultures) for categorical factors and with the median value (e.g., 20 min for centrifugation time and 1 h for incubation time) for numeric factors.

Tenfold Cross-validation Analysis

Given the high heterogeneity of complex protein coronas with numerous factors and the limited datasets for the analysis, tenfold cross-validation was applied to further avoid overfitting. The original dataset was randomly partitioned into ten folds. Nine folds were utilized to train the model as the training set, and the remaining one-fold model was evaluated as the test set. The average R^2 and RMSE were applied to measure model performance.

Factor Selection

To select the key factors determining the various proteins on NPs, the intersections of key factors selected by variable importance were regarded as key factors determining the NP-protein interaction patterns. According to variable importance from models containing overall data pieces and factors, ten significant and independent factors (i.e., NP without modifications, modification, $size_{TEM}$, zeta potential, incubating protein source, incubating plasma concentration, incubating NP concentration, centrifugation speed, centrifugation time and centrifugation temperature) were identified for further analyses.

Protein identification by Mass Spectrometry

The protein corona was analyzed using LC-MS/MS (Orbitrap Fusion and Easy-nLC 1000 system, Thermo Fisher Scientific, USA). Prior to analysis, the protein corona was diluted in solution A (0.1% formic acid in water), and the peptide fragments were

loaded on a reversed-phase column (precolumn: packing material size, 3 μm; aperture, 120 Å; 2 cm×100 μm ID; analytical column: packing material size, 1.9 μm; aperture, 120 Å; 15 cm×150 μm ID) for preparation with a 600 nL/min flow rate and 60 min elution. The Orbitrap full scan was conducted in data-dependent acquisition mode with the following parameters: scanned area, 300-1400; scanning resolution, 120000; automatic gain control (AGC) targets, 5×10^5 ions; max injection time, 100 ms. Coupled MS with high-energy collision dissociation was employed to break parent ions. Then, the broken ions were detected in an ion trap with the following parameters: first mass, 120; resolution rate, 30000; AGC targets, 5×10^3 ions; max injection time, 35 ms; and exclusion duration, 18 s.

The raw data from LC-MS/MS were searched in the UniProt database (2019-02) using Proteome Discoverer 2.1 (Thermo Scientific). To obtain quantitative and qualitative information on proteins and peptides, the search parameters were set as follows: mass deviation of parent ions, 20 ppm; mass deviation of fragment ions in coupled MS, 0.5 Da; the number of restriction enzyme cutting sites of trypsin, 2; variable modifications, protein acetyl (protein N-term), oxidation, and carbamidomethyl; and false discovery rate of peptides, 1% by the decoy database. Over 1000 proteins were detected in each sample. The normalized spectral counts of each protein were utilized to measure the RPA values of the identified proteins.

Cytokine Analysis

The murine macrophage cell line RAW264.7, was obtained from Shanghai Cell Bank of the Type Culture Collection of China. Cells were grown in Dulbecco's modified Eagle medium (DMEM, high glucose, Ding Guo, China) supplemented with 10% fetal bovine serum (FBS, AusGeneX, Australia) and a final concentration of 100 units /mL penicillin-streptomycin in a humidified incubator with 5% CO_2 at 37 °C. To analyze cytokine secretion by macrophages (RAW264.7), dendritic cell line (DC2.4) and human leukemic cell line (THP-1), cells were seeded in 96-well plates (1.5×10^4 cells/well) for 12 h. NP dispersions (Fe_3O_4, Fe_3O_4-CIT, Ag, TiO_2, Au-NH_2 and Au-COOH NPs) at 25 mg/L with or without protein coronas were cultured with cells in serum-free medium. After 6 h of exposure, the supernatant was collected, and the cytokines (TNF-α and IL-6) were measured using ELISA kits (Dakewe, Shenzhen, China). To analyze the immune effect of the corona, the cells were cultured with NPs and seeded on 96-well plates (1.5×10^4 cells/well) for 12 h. After 6 h of exposure, the supernatants were removed, and the cells were washed with PBS. Then, cells were treated with LPS (20 ng/mL) for another 6 h. Finally, the supernatants were collected, and the cytokines were detected with ELISA kits.

To measure the cellular recognition effects of the protein corona on cellular uptake, IL-6 secretion and TNF-α secretion, the following formula was applied to calculate the recognition index of the protein corona; the physiological level of iron as the control was detected. The concentrations of Fe_3O_4 NP uptake were much higher

(approximately 2-6-fold higher) than the physiological level of iron, and the physiological level of iron was subtracted during the calculation ($n \geq 3$):

$$recognition\ index = \frac{R_{PC}\text{-}R_{NO}}{R_{NO}} \times 100\%$$

where the recognition index represents normalized changes in cellular uptake or cytokine release due to the presence of the corona; R_{PC} and R_{NO} represent the measured cellular uptake or cytokine release with and without the protein corona, respectively.

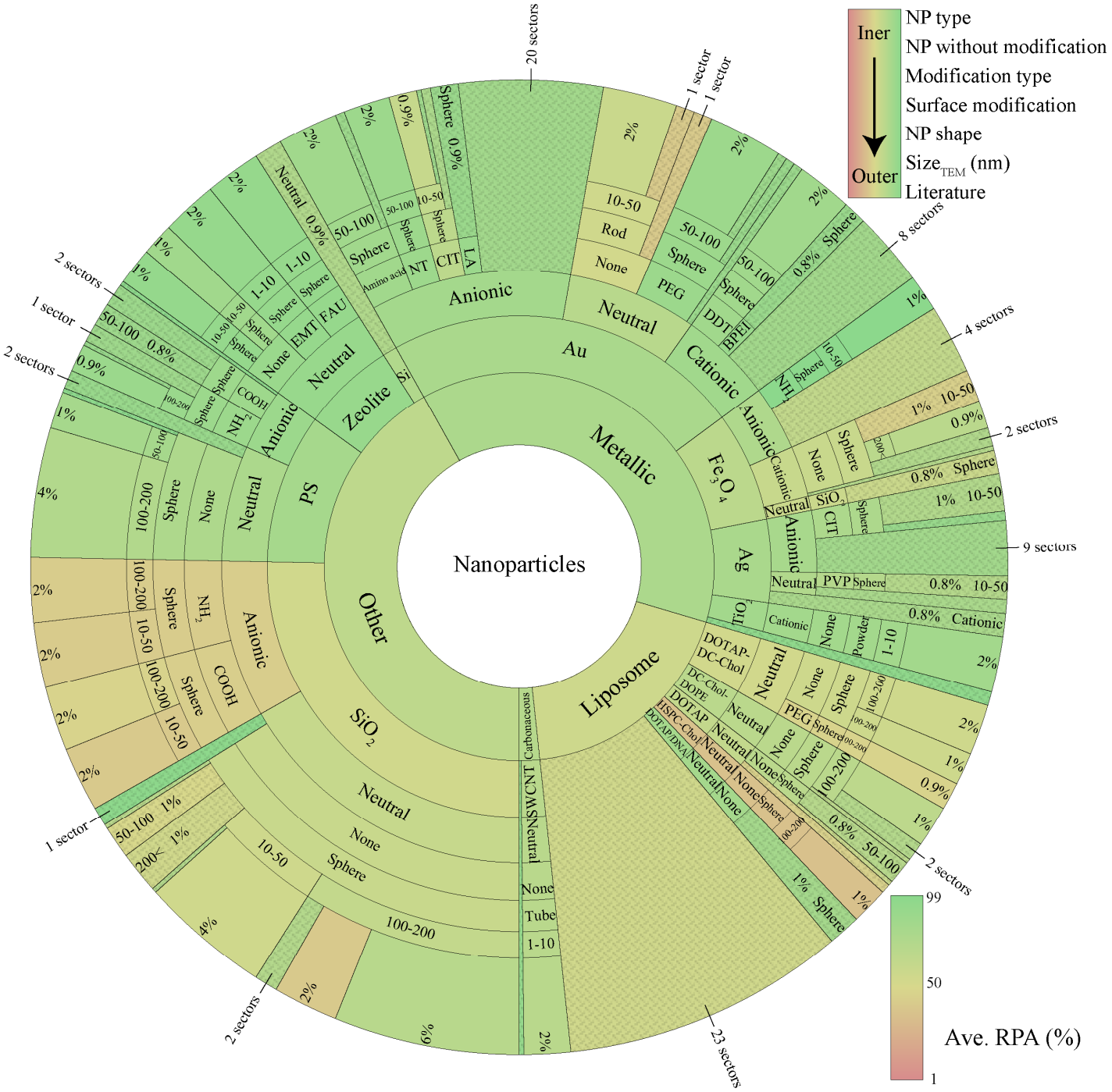


Figure S1. Distributions of 652 mined data items. The Krona chart subdivides each sector into smaller sectors according to the data components. The values of the outermost sectors are the proportions of NPs with a given characteristic among the overall 652 data items. The subgroups with small angles are presented by collapsed

sectors. RPA, relative protein abundance; PS, polystyrene; SWCNT, single-walled carbon nanotube; DOTAP, 1,2-dioleoyl-3-trimethylammonium-propane; HSPC, hydrogenated soy phosphatidylcholine; Chol, cholesterol; DC, 3β-[N-(N′,N′-dimethylaminoethane)-carbamoyl]; DOPE, dioleoylphosphatidylethanolamine. The abbreviations of the modifications are listed in Table S1.

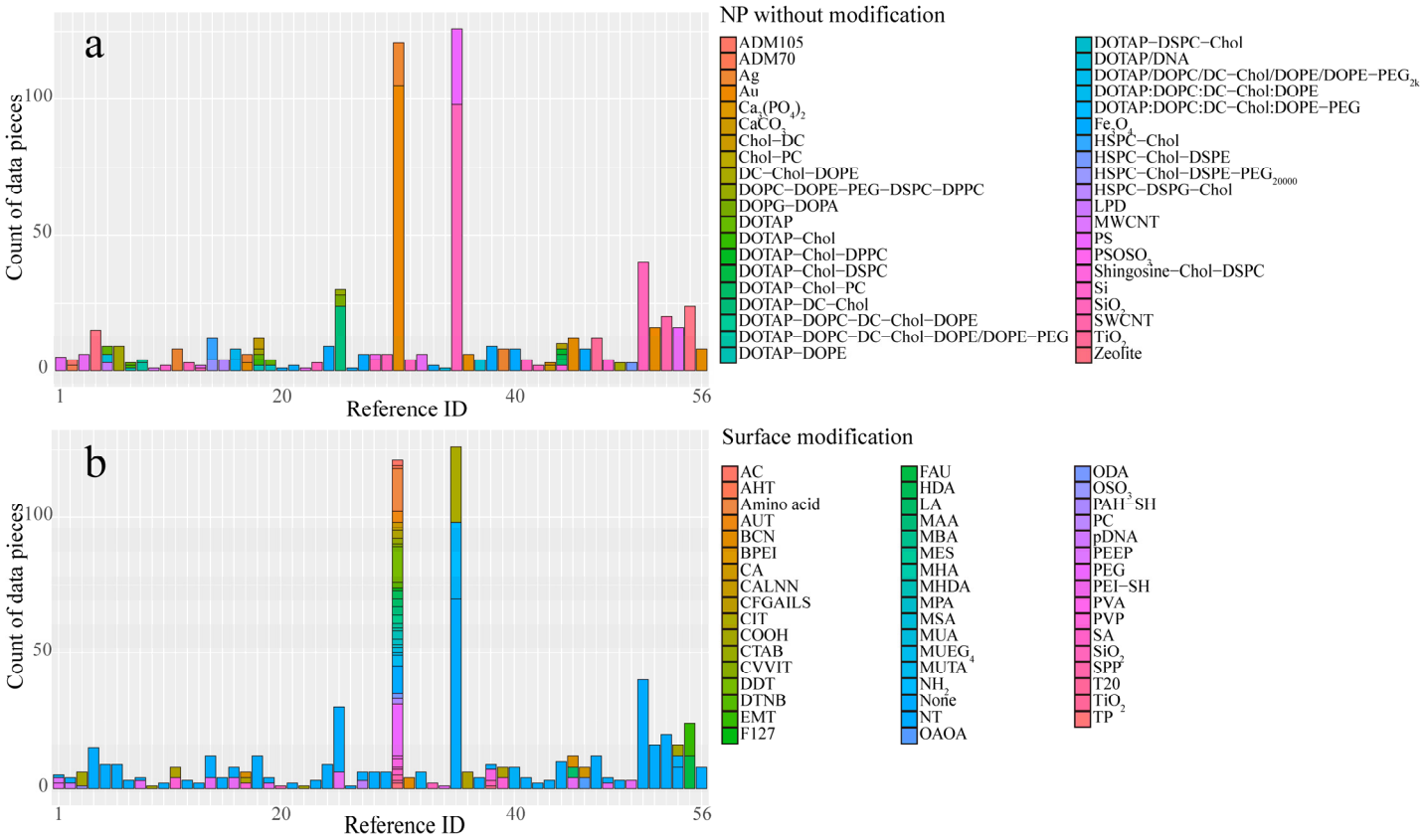


Figure S2. Distribution of data points on NP cores and surface modifications extracted from the literature. The abbreviations of the modifications are listed in Table S1. The references are listed as # 1-56.

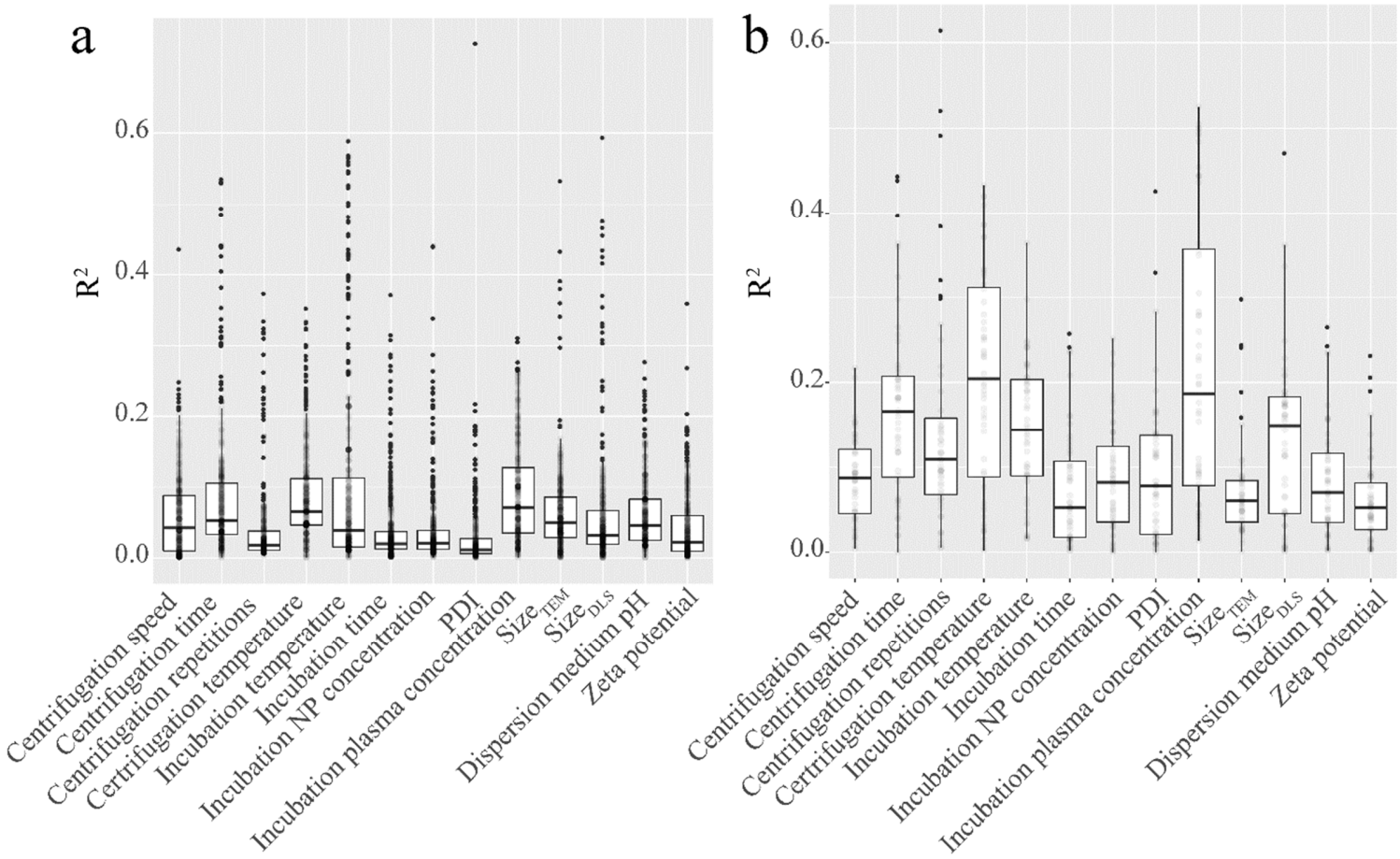


Figure S3. Performance of the linear regression model to reveal the relationships between quantitative factors and protein corona composition. a, Linear regression model based on 593 individuals; b, linear regression model based on 73 protein classes. The analysis estimated the model performance from the correlation coefficient (R^2) between the observations and the predictions. PDI, polydispersity index.

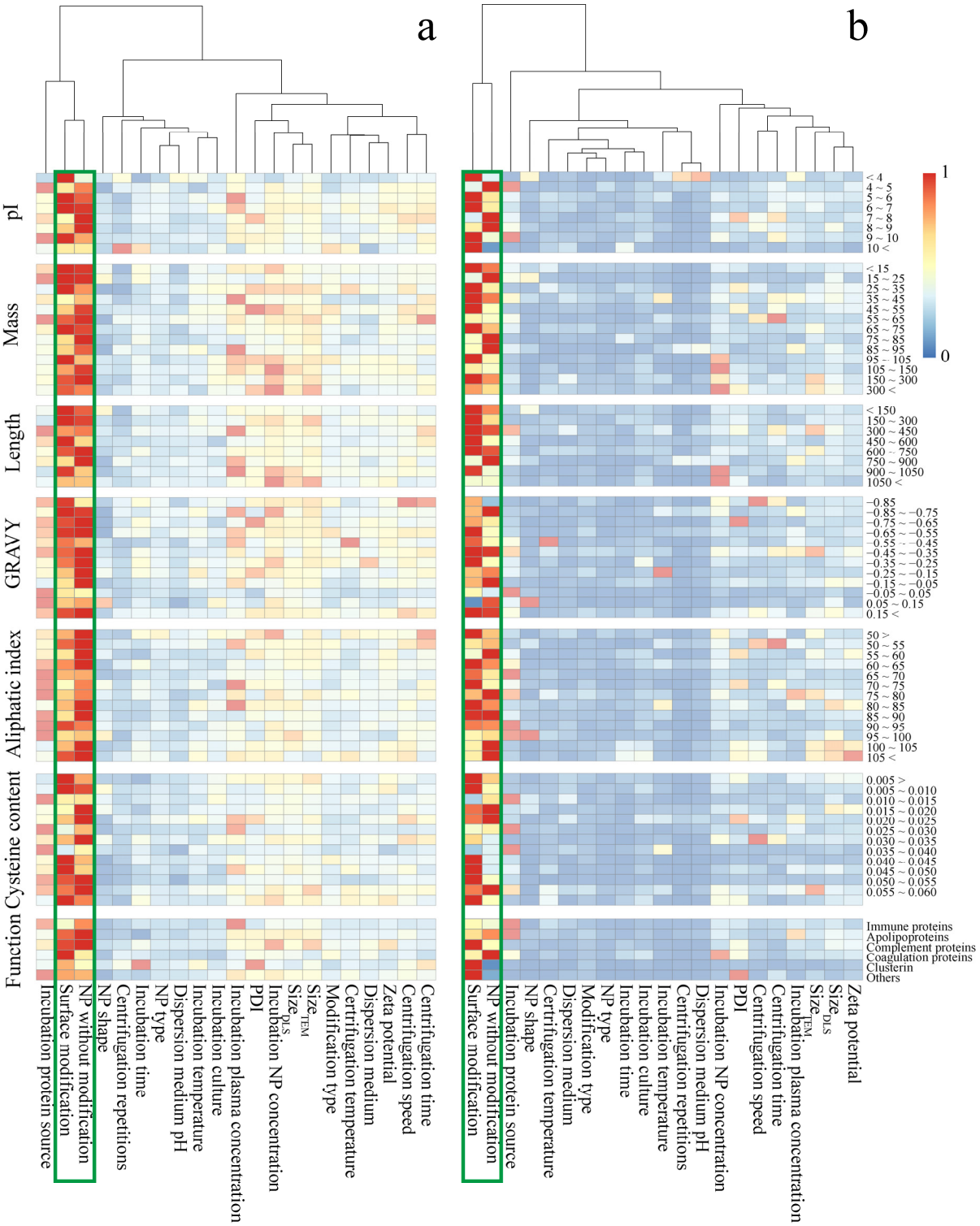


Figure S4. Variable importance of the physicochemical and functional composition prediction

models. The normalized percentages of the increase in mean square error (a) and the normalized increase in node purity (b) were employed to measure the variable importance.

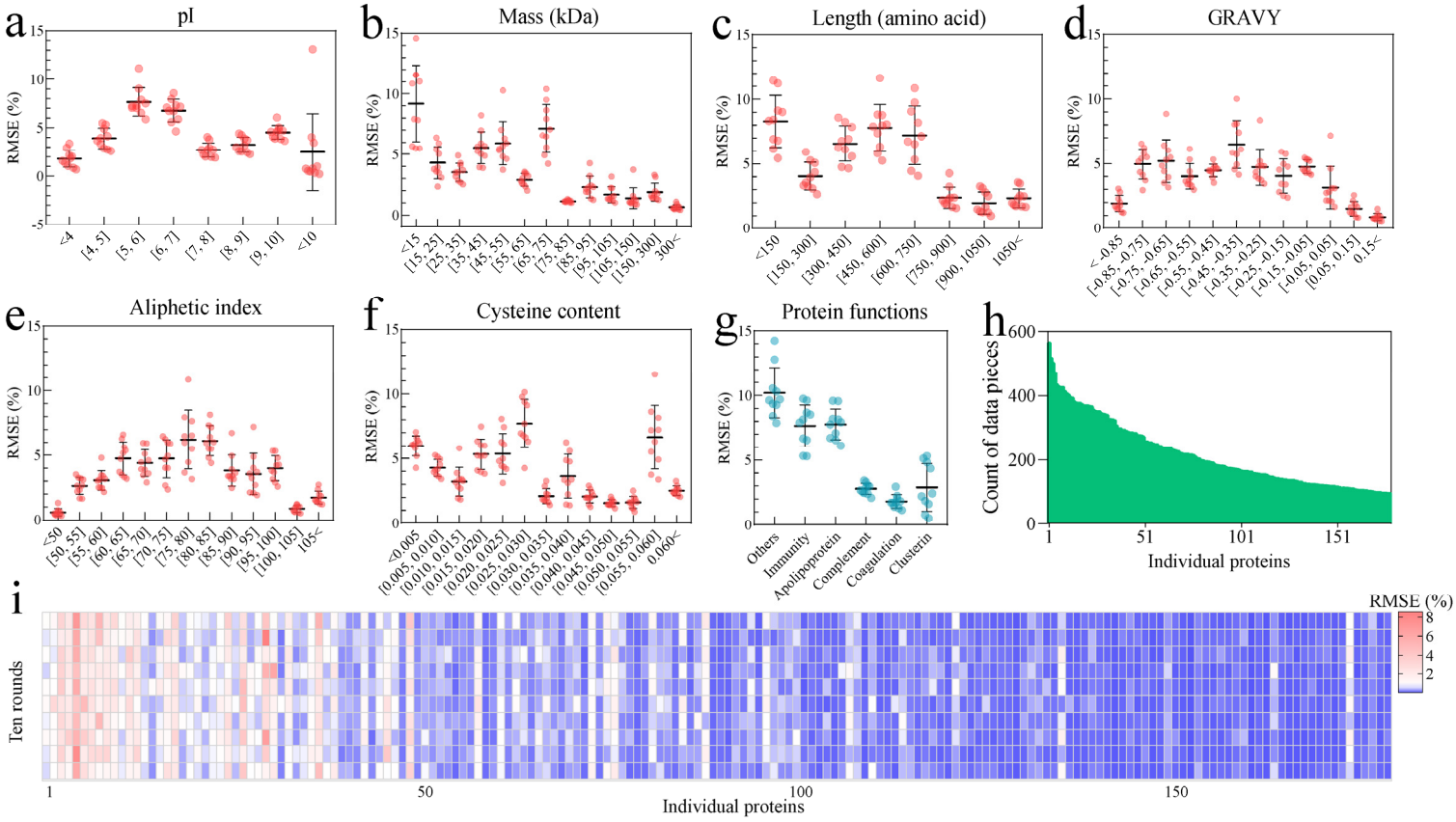


Figure S5. Model performance on functional, physicochemical and individual compositions of protein corona predictions. A tenfold cross-validation was used to evaluate the prediction accuracy without overfitting. More information about individual proteins is shown in the supporting dataset.

Figure S6. Variable importance of the physicochemical and functional composition prediction models. The normalized percentage of the increase in mean square error (a) and the normalized increase in node purity (b) were employed to measure the variable importance.

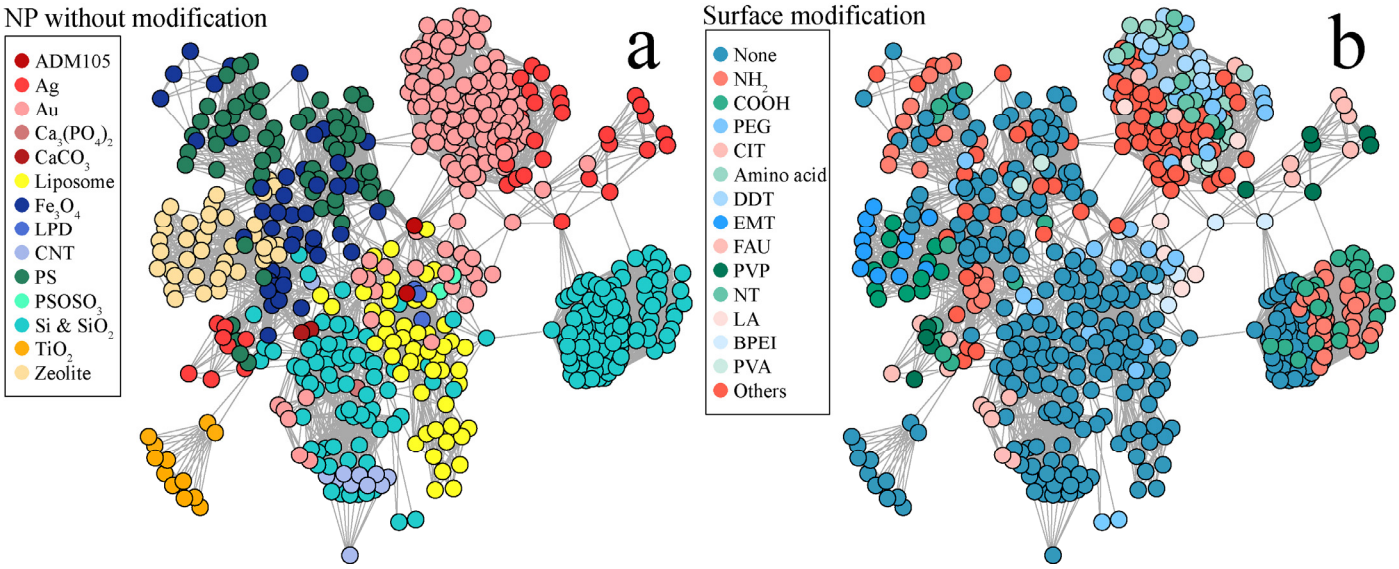


Figure S7. Heterogeneity distribution of priority factors from clusterin composition prediction models visualized by a similarity network. Each node represents a data piece in the functional composition models. The nodes are colored according to the priority factors, NP without modification (a) and surface modification (b). In the well-performed RF models, the connected nodes are more than four times the average in each proximity matrix. Tight connections in each cluster indicate the high homogeneity of nodes for the factor-response dependence learned by the RF models. The sparse connections represent the heterogeneity of nodes in terms of NP properties and experimental conditions in the cluster. The full forms of the abbreviations are listed in Table S1.

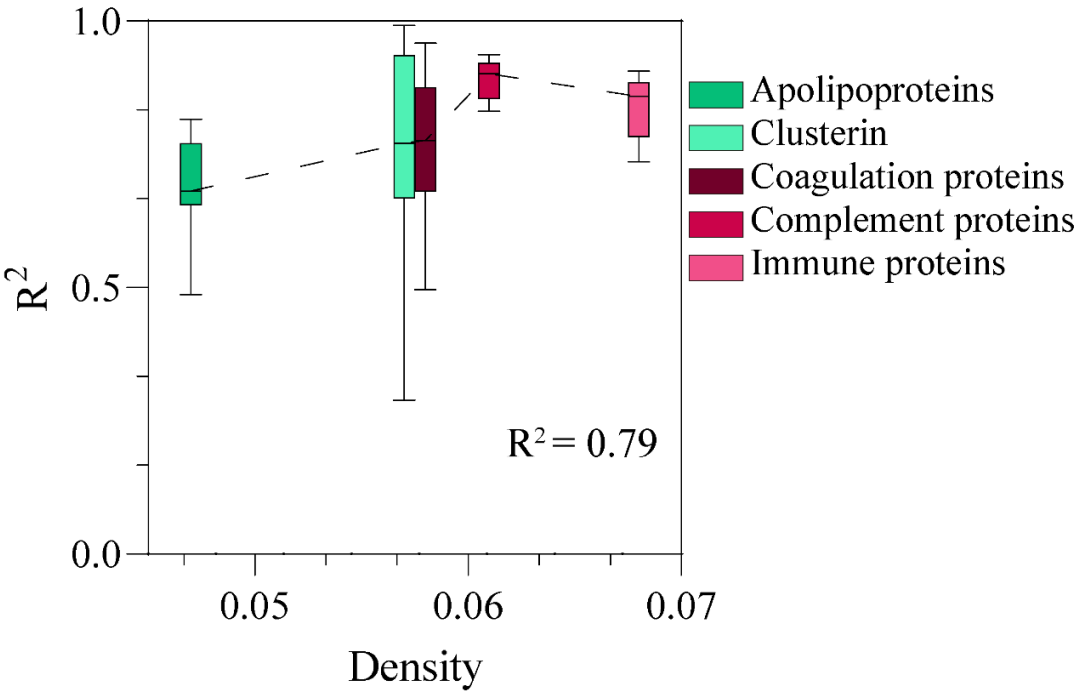
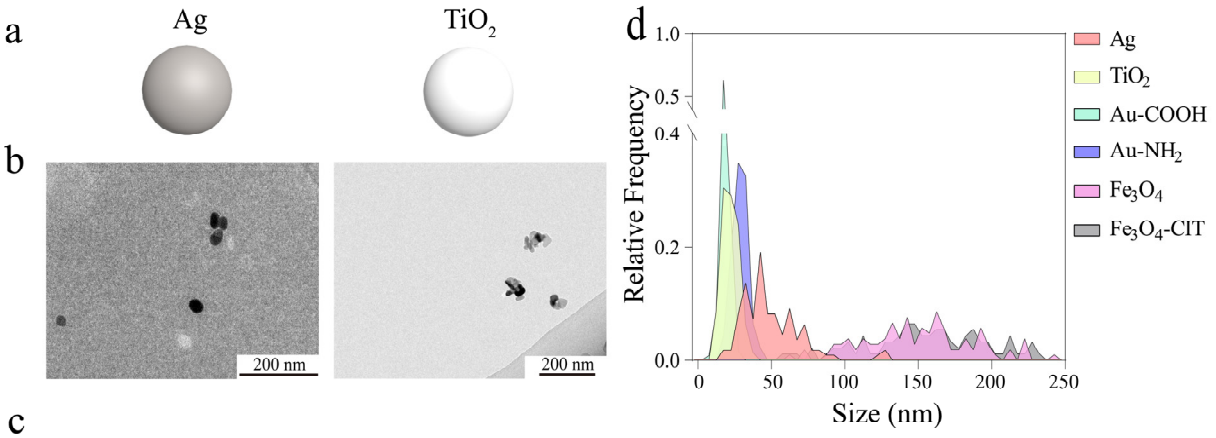


Figure S8. Relationship between the network density and model performance. A tenfold cross-validation was utilized to measure the model performance. The correlation coefficient (R^2) was applied to assess the relationship between the cluster density of the network and the model performance (average R^2).



c

	Ag	TiO_2	Au-COOH	$Au-NH_2$	Fe_3O_4	Fe_3O_4-CIT
ζ_{water}[mV]	-27.1	-19.2	-43.9	-38.3	-5.2	-20.3
PDI_{water}	0.37	0.58	0.23	0.51	0.55	0.6
ζ_{PBS} [mV]	-	-21.5	-17.6	-19.9	-27.4	-23.3
PDI_{PBS}	-	0.45	0.44	0.38	0.61	0.62

Figure S9. Physicochemical characterization of the tested NPs. a, Ag and TiO_2 NPs with no surface modifications; b, TEM micrograph showing the morphologies and sizes of Ag and TiO_2 NPs; c, physicochemical characterization of NPs dispersed in water and phosphate-buffered saline (PBS); d, size distribution measured by TEM. $\zeta_{water/PBS}$, zeta potential of NP dispersion in water or PBS; PDI, polydispersity index.

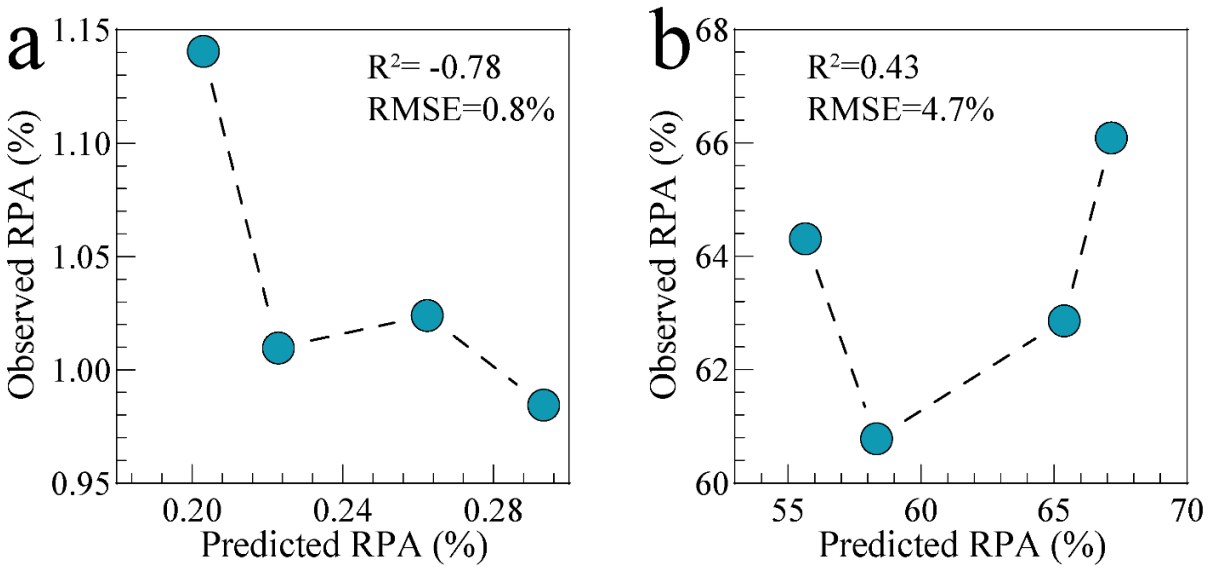


Figure S10. Experiments evaluating the predictions of functional and physicochemical compositions of protein coronas. The observed functional components of the protein coronas on four NPs (Fe_3O_4, Fe_3O_4-CIT, Au-COOH, and Au-NH_2) were measured by LC-MS/MS. The model prediction accuracy was measured by the correlation coefficient (R^2) and root-mean-square error (RMSE) between observations and predictions for clusterin (a) and other protein (b) compositions of selected NPs.

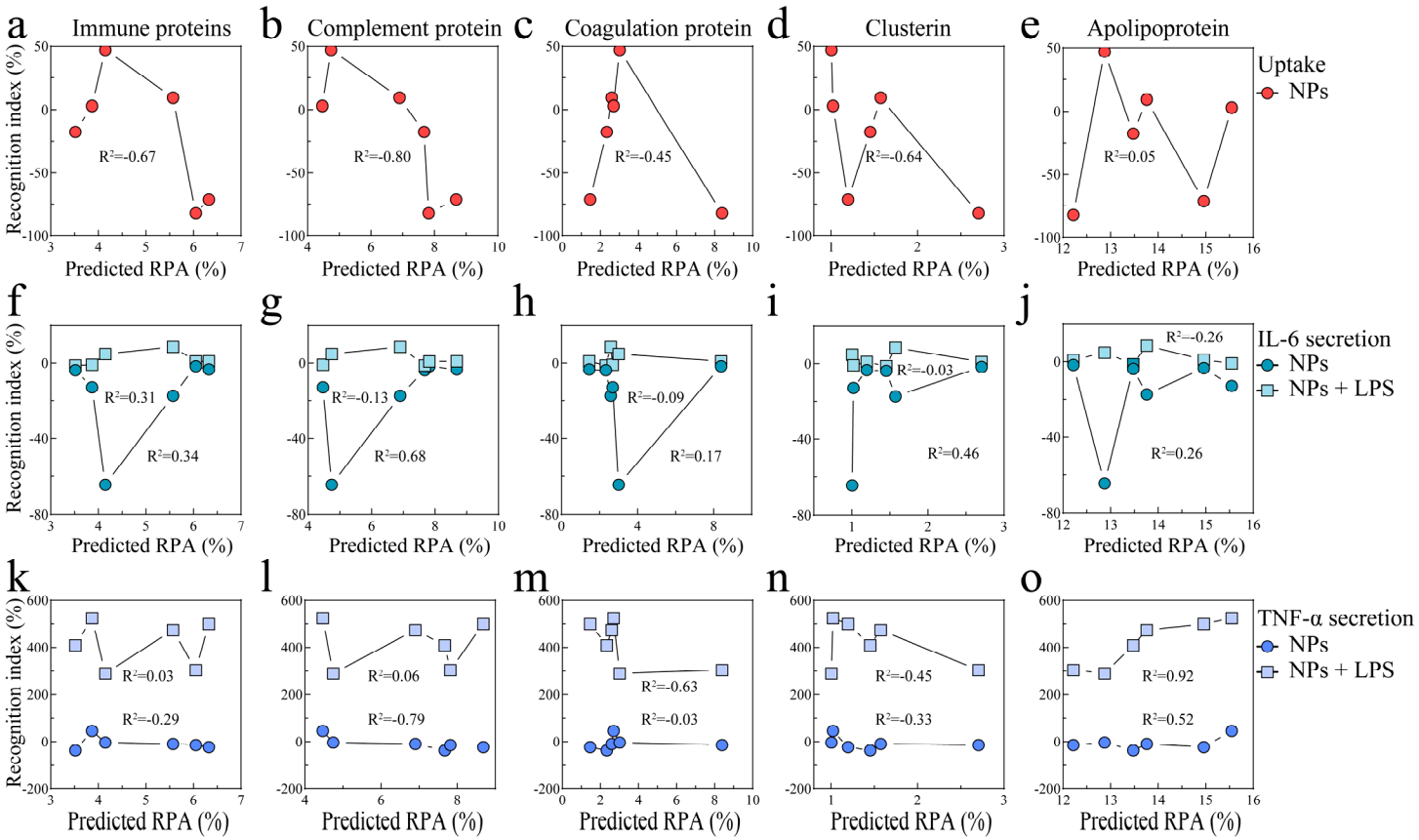
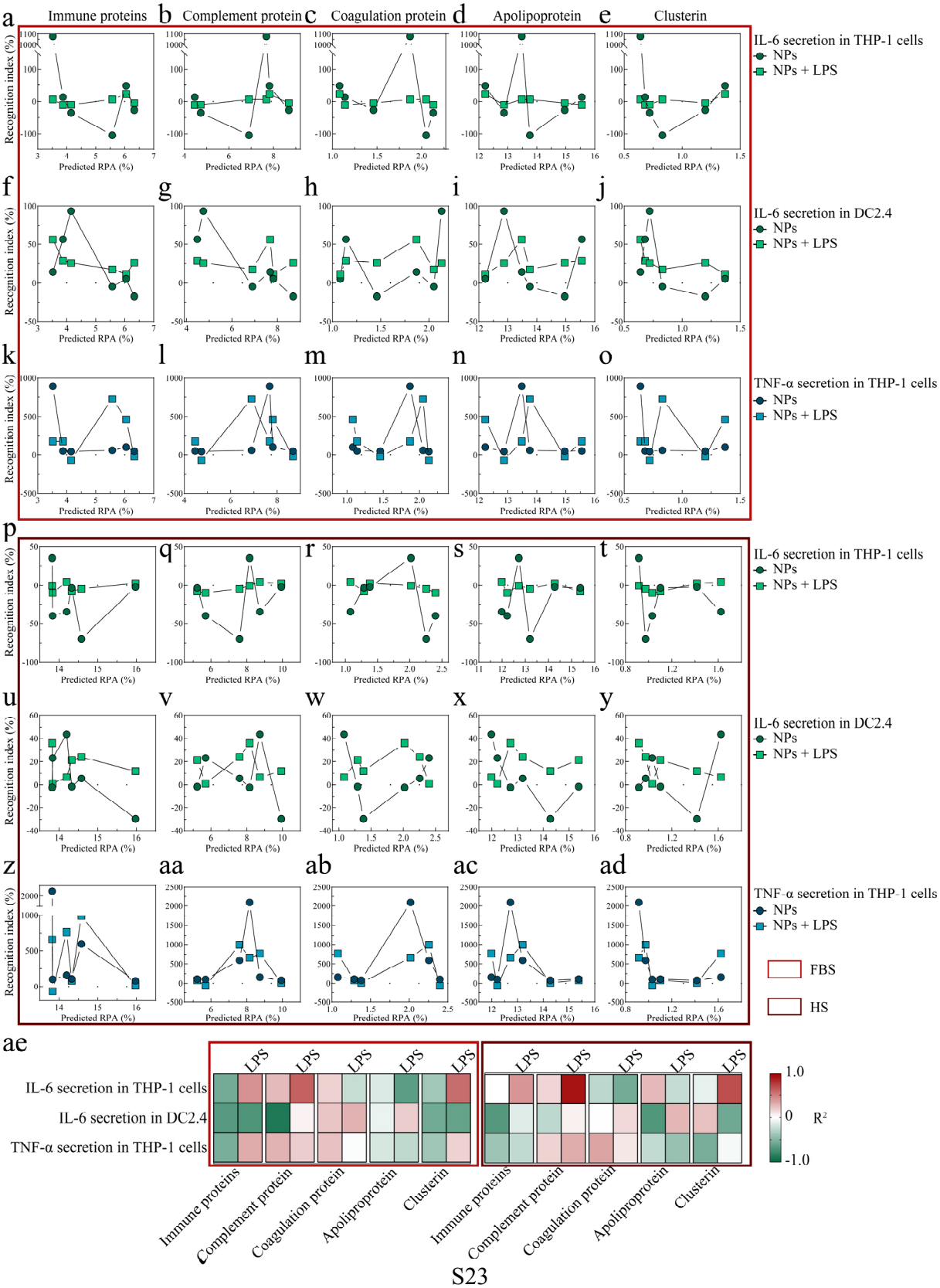


Figure S11. Uptake of NPs and immune responses of macrophages mediated by the protein corona. a-e, Uptake of NPs in macrophages (RAW264.7). Macrophages were treated with NPs (Fe_3O_4, Fe_3O_4-CIT, Au-COOH, Au-NH_2, TiO_2, and Ag NPs) with or without a protein corona at a concentration of 50 mg/L in serum-free medium for 4 h. f-j or k-o, IL-6 or TNF-α generated from macrophages treated with 25 mg/L NPs (Fe_3O_4, Fe_3O_4-CIT, Au-COOH, Au-NH_2, TiO_2, and Ag NPs) with or without a protein corona in serum-free medium for 6 h. Macrophages were treated with LPS for 6 h after NP exposure as a positive control. Relationships in the recognition index (defined in Methods) due to the presence of a corona, with the five functional components (apolipoprotein, complement protein, coagulation protein, immune protein and clusterin) of the

protein corona predicted by functional prediction models, as measured by the correlation coefficient (R^2).


a
Immune proteins
b
Complement protein
c
Coagulation protein
d
Apolipoprotein
e
Clusterin
Recognition index (%)
Predicted RPA (%)
IL-6 secretion in THP-1 cells
NPs
NPs + LPS
f
g
h
i
j
IL-6 secretion in DC2.4
k
l
m
n
o
TNF-α secretion in THP-1 cells
p
q
r
s
t
u
v
w
x
y
z
aa
ab
ac
ad
FBS
HS
ae
LPS
IL-6 secretion in THP-1 cells
IL-6 secretion in DC2.4
TNF-α secretion in THP-1 cells
Immune proteins
Complement protein
Coagulation protein
Apoliproprotein
Clusterin
1.0
0
-1.0
R²

S23

Figure S12. Immune responses of THP-1 and DC2.4 cells mediated by the protein corona. Protein coronas were formed in fetal bovine serum (FBS in panels a-o) or human serum (HS in panels p-ad). The expression of IL-6 (a-j and p-y) or TNF-α (k-o and z-ad) was detected in THP-1 (a-e, k-t and z-ad) or DC2.4 (f-j and u-y) cells treated with 25 mg/L NPs (Fe_3O_4, Fe_3O_4-CIT, Au-COOH, Au-NH_2, TiO_2, and Ag NPs) for 6 h. Cells were treated with LPS for 6 h after NPs exposure as a positive control. Relationships between the recognition index (defined in the method section) and the functional components (apolipoprotein, complement protein, coagulation protein, immune protein and clusterin) of the protein corona predicted by functional prediction models, were measured by the correlation coefficient (R^2), as shown in panel ae.

Table S1. Classes of NP modifications.

Original modification	Transformed modification	Full name	Class
AC	AC	N-Acetyl-L-cysteine	Anionic
AHT	AHT	6-Amino-1-hexanethiol	Cationic
Ala-SH	Amino acid	Thiolated L-alanine	Anionic
Asn-SH	Amino acid	Thiolated L-asparagine	Anionic
AUT	AUT	11-Amino-1-undecanethiol	Cationic
BCN	BCN	Bicyclononyne	Neutral
BPEI	BPEI	Branched polyethyleneimine	Neutral
CA	CA	Citric acid	Anionic
CALNN	CALNN	Peptide sequence 'CALNN'	Anionic
CFGAILS	CFGAILS	Peptide sequence 'CFGAILS'	Anionic
COOH	COOH	Carboxyl	Anionic
$cPEG_{5K}$-SH	PEG	Carboxymethyl-poly(ethylene glycol)-thiol(5kDa)	Neutral
CTAB	CTAB	Hexadecyltrimethylammonium bromide	Cationic
CVVIT	CVVIT	Peptide sequence 'CVVIT'	Anionic
DDT@BDHDA	DDT	1-Dodecanethiol @ benzyldimethylhexadecylammonium bromide	Cationic
DDT@CTAB	DDT	1-Dodecanethiol @ hexadecyltrimethylammonium bromide	Cationic
DDT@DOTAP	DDT	1-Dodecanethiol @ 1,2-dioleoyl-3-trimethylammonium-propane	Cationic
DDT@HDA	DDT	1-Dodecanethiol @ hexadecylamine	Cationic
DDT@ODA	DDT	1-Dodecanethiol @ octadecylamine	Cationic
DDT@SA	DDT	1-Dodecanethiol @ stearic acid	Anionic
DDT@SDS	DDT	1-Dodecanethiol @ sodium dodecyl sulfate	Anionic
DTNB	DTNB	5,5'-Dithiobis (2-nitrobenzoic acid)	Anionic
EMT	EMT	$Na_{88}(AlO_2)_{88}(SiO_2)_{104}$	Neutral
F127	F127	Pluronic F-127	Anionic
FAU	FAU	$Na_{73}(AlO_2)_{73}(SiO_2)_{119}$	Neutral
Gly-SH	Amino acid	Thiolated L-glycine	Anionic
Gold	None	Gold	Neutral
CIT	CIT	Citrate	Anionic
HDA	HDA	Hexadecylamine	Cationic
LA	LA	α-Lipoic acid	Anionic
MAA	MAA	Mercaptoacetic acid	Anionic
MBA	MBA	4-Mercaptobenzoic acid	Anionic
MES	MES	2-Mercaptoethanesulfonate	Anionic
Met-SH	Amino acid	Thiolated L-methionine	Anionic
MHA	MHA	6-Mercaptohexanoic acid	Anionic
MHDA	MHDA	16-Mercaptohexadecanoic acid	Anionic
MPA	MPA	3-Mercaptopropionic acid	Anionic
$mPEG_{1K}$-SH	PEG	Methoxy-poly(ethylene glycol)-thiol (1kDa)	Neutral

$mPEG_{20K}$-SH	PEG	Methoxy-poly(ethylene glycol)-thiol (20kDa)	Neutral
$mPEG_{2K}$-SH	PEG	Methoxy-poly(ethylene glycol)-thiol (20kDa) (low density)	Neutral
$mPEG_{5K}$	PEG	Methoxy-poly(ethylene glycol)-thiol (2kDa)	Neutral
$mPEG_{5K}(NH_2)$-SH	PEG	Methoxy-poly(ethylene glycol)-thiol (5kDa)	Neutral
$mPEG_{5K}$-SH	PEG	Thiolated amino-poly(ethylene glycol) (methoxy-terminated) (5kDa)	Neutral
MSA	MSA	Mercaptosuccinic acid	Anionic
MUA	MUA	11-Mercaptoundecanoic acid	Anionic
MUEG4	MUEG4	(11-Mercaptoundecyl)tetra(ethylene glycol)	Neutral
MUTA	MUTA	(11-Mercaptoundecyl)-N,N,N-trimethylammonium	Cationic
NH_2	NH_2	Amino	Cationic
None	None	None	Neutral
$nPEG_{5K}$-SH	PEG	Amino-poly(ethylene glycol)-thiol (5kDa)	Neutral
NT@DCA	NT	2-Napthalenethiol @ deoxycholic acid	Neutral
NT@F127	NT	2-Napthalenethiol @ Pluronic F-127	Anionic
NT@PSMA-AAP	NT	2-Napthalenethiol @ (4'-aminoacetophenone)-modified poly(styrene-co-maleic anhydride)	Neutral
NT@PSMA-AP	NT	2-Napthalenethiol @ aminopropanol-modified poly(styrene-co-maleic anhydride)	Anionic
NT@PSMA-EA	NT	2-Napthalenethiol @ ethanolamine-modified poly(styrene-co-maleic anhydride)	Anionic
NT@PSMA-EDA	NT	2-Napthalenethiol @ ethylenediamine-modified poly(styrene-co-maleic anhydride)	Anionic
NT@PSMA-Urea	NT	2-Napthalenethiol @ urea-modified poly(styrene-co-maleic anhydride)	Anionic
NT@PVA	NT	2-Napthalenethiol @ poly(vinyl alcohol)	Anionic
OAOA	OAOA	Oleic acid double layer	Anionic
ODA	ODA	Octadecylamine	Neutral
OSO_3	OSO_3	Sulpho	Anionic
P80	None	Polysorbate 80	Neutral
PAH-SH	PAH-SH	Thiolated poly(allylamine)	Cationic
PC	PC	Phosphatidylcholine	Neutral
pDNA	pDNA	Plasmid deoxyribonucleic acid	Anionic
$PEEP_{49}$	PEEP	Poly(ethyl ethylene phosphate)49	Neutral
$PEEP_{92}$	PEEP	Poly(ethyl ethylene phosphate)92	Neutral
PEG	PEG	Polyethylene glycol	Neutral
$PEG_{3K}(NH_2)$-SH	PEG	Thiolated amino-poly(ethylene glycol) (3kDa)	Neutral
PEI-SH	PEI-SH	Thiolated poly(ethyleneimine)	Cationic
Phe	Amino acid	L-Phenylalanine	Anionic
Phe-SH	Amino acid	Thiolated L-phenylalanine	Anionic
PLL-SH	Amino acid	Thiolated poly(L-lysine)	Cationic
PVA	PVA	Poly(vinyl alcohol)	Anionic
PVA (COOH)	PVA	Poly(vinyl alcohol)	Anionic

PVA (NH_2)	PVA	Poly(vinyl alcohol)	Anionic
PVA (NH_2/OH)	PVA	Poly(vinyl alcohol)	Anionic
PVA (OH)	PVA	Poly(vinyl alcohol)	Anionic
PVP	PVP	Poly(vinylpyrrolidone)	Anionic
SA	SA	Stearic acid	Anionic
Ser-SH	Amino acid	Thiolated L-serine	Anionic
SiO_2	SiO_2	Silicon dioxide	Neutral
SiO_2 (APTES)	SiO_2	Silicon dioxide	Neutral
SiO_2 (TEOS)	SiO_2	Silicon dioxide	Neutral
SPP	SPP	Bis(p-sulfonatophenyl)phenylphosphine	Anionic
T20	T20	TWEEN20	Anionic
Thr-SH	Amino acid	Thiolated L-threonine	Anionic
TiO_2	TiO_2	Titanium dioxide	Neutral
TP	TP	N-(2-Mercaptopropionyl)glycine	Anionic
Trp-SH	Amino acid	Thiolated L-tryptophan	Anionic

References

1. Schöttler, S.; Becker, G.; Winzen, S.; Steinbach, T.; Mohr, K.; Landfester, K.; Mailänder, V.; Wurm, F. R., Protein adsorption is required for stealth effect of poly(ethylene glycol)- and poly(phosphoester)-coated nanocarriers. *Nat. Nanotechnol.* **2016,** *11* (4), 372-377.

2. Martíneznegro, M.; Caracciolo, G.; Palchetti, S.; Pozzi, D.; Capriotti, A. L.; Cavaliere, C.; Laganà, A.; Ortiz, M. C.; Benito, J. M.; García Fernández, J. M., Biophysics and protein corona analysis of Janus cyclodextrin-DNA nanocomplexes. Efficient cellular transfection on cancer cells. *Biochim. Biophys. Acta Gen. Subj.* **2017,** *1861* (7), 1737-1749.

3. 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), 15268-15276.

4. Laurent, S.; Ng, E. P.; Thirifays, C.; Lakiss, L.; Goupil, G. M.; Mintova, S.; Burtea, C.; Oveisi, E.; Hébert, C.; De Vries, M.; Motazacker, M. M.; Rezaee, F.; Mahmoudi, M., Corona

protein composition and cytotoxicity evaluation of ultra-small zeolites synthesized from template free precursor suspensions. *Toxicol. Res.* **2013,** *2* (4), 270-279.

5. Capriotti, A. L.; Caracciolo, G.; Caruso, G.; Foglia, P.; Pozzi, D.; Samperi, R.; Lagan, A., Differential analysis of "protein corona" profile adsorbed onto different nonviral gene delivery systems. *Anal. Biochem.* **2011,** *419* (2), 180-189.

6. Capriotti, A. L.; Caracciolo, G.; Cavaliere, C.; Foglia, P.; Pozzi, D.; Samperi, R.; Laganà, A., Do plasma proteins distinguish between liposomes of varying charge density? *J. Proteomics* **2012,** *75* (6), 1924-1932.

7. Caracciolo, G.; Pozzi, D.; Capriotti, A. L.; Cavaliere, C.; Laganà, A., Effect of DOPE and cholesterol on the protein adsorption onto lipid nanoparticles. *J. Nanopart. Res.* **2013,** *15* (3), 1498.

8. Pozzi, D.; Colapicchioni, V.; Caracciolo, G.; Piovesana, S.; Capriotti, A. L.; Palchetti, S.; De Grossi, S.; Riccioli, A.; Amenitsch, H.; Laganà, A., Effect of polyethyleneglycol (PEG) chain length on the bio-nano- interactions between PEGylated lipid nanoparticles and biological fluids: From nanostructure to uptake in cancer cells. *Nanoscale* **2014,** *6* (5), 2782-2792.

9. De Paoli, S. H.; Diduch, L. L.; Tegegn, T. Z.; Orecna, M.; Strader, M. B.; Karnaukhova, E.; Bonevich, J. E.; Holada, K.; Simak, J., The effect of protein corona composition on the interaction of carbon nanotubes with human blood platelets. *Biomaterials* **2014,** *35* (24), 6182-6194.

10. 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.

11. Huang, H.; Lai, W.; Cui, M.; Liang, L.; Lin, Y.; Fang, Q.; Liu, Y.; Xie, L., An evaluation of blood compatibility of silver nanoparticles. *Sci. Rep.* **2016,** *6*, 25518.

12. 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.

13. Lo Giudice, M. C.; Herda, L. M.; Polo, E.; Dawson, K. A., In situ characterization of nanoparticle biomolecular interactions in complex biological media by flow cytometry. *Nat. Commun.* **2016,** *7*, 13475.

14. Hadjidemetriou, M.; Al-Ahmady, Z.; Mazza, M.; Collins, R. F.; Dawson, K.; Kostarelos, K., *In vivo* biomolecule corona around blood-circulating, clinically used and antibody-targeted lipid bilayer nanoscale vesicles. *ACS Nano* **2015,** *9* (8), 8142-8156.

15. Amici, A.; Caracciolo, G.; Digiacomo, L.; Gambini, V.; Marchini, C.; Tilio, M.; Capriotti, A. L.; Colapicchioni, V.; Matassa, R.; Familiari, G.; Palchetti, S.; Pozzi, D.; Mahmoudi, M.; Laganà, A., In vivo protein corona patterns of lipid nanoparticles. *RSC Adv.* **2017,** *7* (2), 1137-1145.

16. Palchetti, S.; Pozzi, D.; Capriotti, A. L.; Barbera, G. L.; Chiozzi, R. Z.; Digiacomo, L.; Peruzzi, G.; Caracciolo, G.; Laganà, A., Influence of dynamic flow environment on nanoparticle-protein corona: From protein patterns to uptake in cancer cells. *Colloids Surf. B Biointerfaces* **2017,** *153*, 263-271.

17. Lai, W.; Wang, Q.; Li, L.; Hu, Z.; Chen, J.; Fang, Q., Interaction of gold and silver nanoparticles with human plasma: Analysis of protein corona reveals specific binding patterns. *Colloids Surf. B Biointerfaces* **2017,** *152*, 317-325.

18. Caracciolo, G.; Pozzi, D.; Capriotti, A. L.; Cavaliere, C.; Piovesana, S.; Amenitsch, H.; Lagana, A., Lipid composition: a "key factor" for the rational manipulation of the liposome-protein corona by liposome design. *RSC Adv.* **2015,** *5* (8), 5967-5975.

19. Caracciolo, G.; Pozzi, D.; Capriotti, A. L.; Cavaliere, C.; Piovesana, S.; La Barbera, G.; Amici, A.; Lagana, A., The liposome-protein corona in mice and humans and its implications for in vivo delivery. *J. Mater. Chem. B* **2014,** *2* (42), 7419-7428.

20. Bertoli, F.; Davies, G. L.; Monopoli, M. P.; Moloney, M.; Gun'Ko, Y. K.; Salvati, A.; Dawson, K. A., Magnetic nanoparticles to recover cellular organelles and study the time resolved nanoparticle-cell interactome throughout uptake. *Small* **2014,** *10* (16), 3307-3315.

21. Liu, Z.; Zhan, X.; Yang, M.; Yang, Q.; Xu, X.; Lan, F.; Wu, Y.; Gu, Z., A magnetic-dependent protein corona of tailor-made superparamagnetic iron oxides alters their biological behaviors. *Nanoscale* **2016,** *8* (14), 7544-7555.

22. Kelly, P. M.; Åberg, C.; Polo, E.; O'Connell, A.; Cookman, J.; Fallon, J.; Krpetić, Ž.; Dawson, K. A., Mapping protein binding sites on the biomolecular corona of nanoparticles. *Nat. Nanotechnol.* **2015,** *10* (5), 472-479.

23. Tenzer, S.; Docter, D.; Rosfa, S.; Wlodarski, A.; Kuharev, J.; Rekik, A.; Knauer, S. K.; Bantz, C.; Nawroth, T.; Bier, C.; Sirirattanapan, J.; Mann, W.; Treuel, L.; Zellner, R.; Maskos, M.; Schild, H.; Stauber, R. H., Nanoparticle size is a critical physicochemical determinant of the human blood plasma corona: A comprehensive quantitative proteomic analysis. *ACS Nano* **2011,** *5* (9), 7155-7167.

24. Hu, Z.; Zhang, H.; Zhang, Y.; Wu, R.; Zou, H., Nanoparticle size matters in the formation of plasma protein coronas on Fe_3O_4 nanoparticles. *Colloids Surf. B Biointerfaces* **2014,** *121*, 354-361.

25. Palchetti, S.; Digiacomo, L.; Pozzi, D.; Peruzzi, G.; Micarelli, E.; Mahmoudi, M.; Caracciolo, G., Nanoparticles-cell association predicted by protein corona fingerprints. *Nanoscale* **2016,** *8* (25), 12755-12763.

26. Hu, Z.; Zhao, L.; Zhang, H.; Zhang, Y.; Wu, R.; Zou, H., The on-bead digestion of protein corona on nanoparticles by trypsin immobilized on the magnetic nanoparticle. *J. Chromatogr. A* **2014,** *1334*, 55-63.

27. Groult, H.; Ruiz-Cabello, J.; Lechuga-Vieco, A. V.; Mateo, J.; Benito, M.; Bilbao, I.; Martínez-Alcázar, M. P.; Lopez, J. A.; Vázquez, J.; Herranz, F. F., Phosphatidylcholine-coated iron oxide nanomicelles for *in vivo* prolonged circulation time with an antibiofouling protein corona. *Chem. - Eur. J.* **2014,** *20* (50), 16662-16671.

28. Monopoli, M. P.; Walczyk, D.; Campbell, A.; Elia, G.; Lynch, I.; Baldelli Bombelli, F.; Dawson, K. A., Physical-Chemical aspects of protein corona: Relevance to *in vitro* and *in vivo* biological impacts of nanoparticles. *J. Am. Chem. Soc.* **2011,** *133* (8), 2525-2534.

29. Clemments, A. M.; Botella, P.; Landry, C. C., Protein adsorption from biofluids on silica nanoparticles: Corona analysis as a function of particle diameter and porosity. *ACS Appl. Mater. Interfaces* **2015,** *7* (39), 21682-21689.

30. Walkey, C. D.; Olsen, J. B.; Song, F.; Liu, R.; Guo, H.; Olsen, D. W. H.; Cohen, Y.; Emili, A.; Chan, W. C. W., Protein corona fingerprinting predicts the cellular interaction of gold and silver nanoparticles. *ACS Nano* **2014,** *8* (3), 2439-2455.

31. Mirshafiee, V.; Mahmoudi, M.; Lou, K.; Cheng, J.; Kraft, M. L., Protein corona significantly reduces active targeting yield. *Chem. Commun.* **2013,** *49* (25), 2557-2559.

32. Schöttler, S.; Klein, K.; Landfester, K.; Mailänder, V., Protein source and choice of anticoagulant decisively affect nanoparticle protein corona and cellular uptake. *Nanoscale* **2016,** *8* (10), 5526-5536.

33. Vogt, C.; Pernemalm, M.; Kohonen, P.; Laurent, S.; Hultenby, K.; Vahter, M.; Lehtiö, J.; Toprak, M. S.; Fadeel, B., Proteomics analysis reveals distinct corona composition on magnetic nanoparticles with different surface coatings: Implications for interactions with primary human macrophages. *PLoS One* **2015,** *10* (10), e0129008.

34. Pozzi, D.; Caracciolo, G.; Capriotti, A. L.; Cavaliere, C.; Piovesana, S.; Colapicchioni, V.; Palchetti, S.; Riccioli, A.; Laganà, A., A proteomics-based methodology to investigate the protein corona effect for targeted drug delivery. *Mol. Biosyst.* **2014,** *10* (11), 2815-2819.

35. Tenzer, S.; Docter, D.; Kuharev, J.; Musyanovych, A.; Fetz, V.; Hecht, R.; Schlenk, F.; Fischer, D.; Kiouptsi, K.; Reinhardt, C.; Landfester, K.; Schild, H.; Maskos, M.; Knauer, S. K.; Stauber, R. H., Rapid formation of plasma protein corona critically affects nanoparticle pathophysiology. *Nat. Nanotechnol.* **2013,** *8* (10), 772-781.

36. Albanese, A.; Walkey, C. D.; Olsen, J. B.; Guo, H.; Emili, A.; Chan, W. C. W., Secreted biomolecules alter the biological identity and cellular interactions of nanoparticles. *ACS Nano* **2014,** *8* (6), 5515-5526.

37. Caracciolo, G.; Cardarelli, F.; Pozzi, D.; Salomone, F.; Maccari, G.; Bardi, G.; Capriotti, A. L.; Cavaliere, C.; Papi, M.; Laganà, A., Selective targeting capability acquired with a protein corona adsorbed on the surface of 1,2-dioleoyl-3-trimethylammonium propane/dna nanoparticles. *ACS Appl. Mater. Interfaces* **2013,** *5* (24), 13171-13179.

38. Sakulkhu, U.; Mahmoudi, M.; Maurizi, L.; Coullerez, G.; Hofmann-Amtenbrink, M.; Vries, M.; Motazacker, M.; Rezaee, F.; Hofmann, H., Significance of surface charge and shell material

of superparamagnetic iron oxide nanoparticle (SPION) based core/shell nanoparticles on the composition of the protein corona. *Biomater. Sci.* **2015,** *3* (2), 265-278.

39. Shannahan, J. H.; Lai, X.; Ke, P. C.; Podila, R.; Brown, J. M.; Witzmann, F. A., Silver nanoparticle protein corona composition in cell culture media. *PLoS One* **2013,** *8* (9), e74001.

40. Mahmoudi, M.; Shokrgozar, M. A.; Behzadi, S., Slight temperature changes affect protein affinity and cellular uptake/toxicity of nanoparticles. *Nanoscale* **2013,** *5* (8), 3240-3244.

41. Pisani, C.; Rascol, E.; Dorandeu, C.; Gaillard, J. C.; Charnay, C.; Guari, Y.; Chopineau, J.; Armengaud, J.; Devoisselle, J. M.; Prat, O., The species origin of the serum in the culture medium influences the in vitro toxicity of silica nanoparticles to HepG2 cells. *PLoS One* **2017,** *12* (8), e0182906.

42. Caracciolo, G.; Palchetti, S.; Colapicchioni, V.; Digiacomo, L.; Pozzi, D.; Capriotti, A. L.; La Barbera, G.; Laganà, A., Stealth effect of biomolecular corona on nanoparticle uptake by immune cells. *Langmuir* **2015,** *31* (39), 10764-10773.

43. Wu, C.-Y.; Young, D.; Martel, J.; Young, J. D., A story told by a single nanoparticle in the body fluid: demonstration of dissolution-reprecipitation of nanocrystals in a biological system. *Nanomedicine* **2015,** *10* (17), 2659-2676.

44. Pozzi, D.; Caracciolo, G.; Capriotti, A. L.; Cavaliere, C.; La Barbera, G.; Anchordoquy, T. J.; Laganà, A., Surface chemistry and serum type both determine the nanoparticle-protein corona. *J. Proteomics* **2015,** *119*, 209-217.

45. Chandran, P.; Riviere, J. E.; Monteiro-Riviere, N. A., Surface chemistry of gold nanoparticles determines the biocorona composition impacting cellular uptake, toxicity and gene expression profiles in human endothelial cells. *Nanotoxicology* **2017,** *11* (4), 507-519.

46. Jedlovszky-Hajdú, A.; Bombelli, F. B.; Monopoli, M. P.; Tombácz, E.; Dawson, K. A., Surface coatings shape the protein corona of SPIONs with relevance to their application in vivo. *Langmuir* **2012,** *28* (42), 14983-14991.

47. Kuruvilla, J.; Farinha, A. P.; Bayat, N.; Cristobal, S., Surface proteomics on nanoparticles: a step to simplify the rapid prototyping of nanoparticles. *Nanoscale Horiz.* **2017,** *2* (1), 55-64.

48. Nissinen, T.; Näkki, S.; Laakso, H.; Kučiauskas, D.; Kaupinis, A.; Kettunen, M. I.; Liimatainen, T.; Hyvönen, M.; Valius, M.; Gröhn, O.; Lehto, V. P., Tailored dual pegylation of inorganic porous nanocarriers for extremely long blood circulation *in vivo*. *ACS Appl. Mater. Interfaces* **2016,** *8* (48), 32723-32731.

49. Barrán-Berdón, A. L.; Pozzi, D.; Caracciolo, G.; Capriotti, A. L.; Caruso, G.; Cavaliere, C.; Riccioli, A.; Palchetti, S.; Laganaì, A., Time evolution of nanoparticle-protein corona in human plasma: Relevance for targeted drug delivery. *Langmuir* **2013,** *29* (21), 6485-6494.

50. Hadjidemetriou, M.; Al-Ahmady, Z.; Kostarelos, K., Time-evolution of in vivo protein corona onto blood-circulating PEGylated liposomal doxorubicin (DOXIL) nanoparticles. *Nanoscale* **2016,** *8* (13), 6948-6957.

51. Pisani, C.; Gaillard, J. C.; Odorico, M.; Nyalosaso, J. L.; Charnay, C.; Guari, Y.; Chopineau, J.; Devoisselle, J. M.; Armengaud, J.; Prat, O., The timeline of corona formation around silica nanocarriers highlights the role of the protein interactome. *Nanoscale* **2017,** *9* (5), 1840-1851.

52. Mahmoudi, M.; Lohse, S. E.; Murphy, C. J.; Fathizadeh, A.; Montazeri, A.; Suslick, K. S., Variation of protein corona composition of gold nanoparticles following plasmonic heating. *Nano Lett.* **2014,** *14* (1), 6-12.

53. Raghavendra, A. J.; Fritz, K.; Fu, S.; Brown, J. M.; Podila, R.; Shannahan, J. H., Variations in biocorona formation related to defects in the structure of single walled carbon nanotubes and the hyperlipidemic disease state. *Sci. Rep.* **2017,** *7* (1), 8382.

54. Kokkinopoulou, M.; Simon, J.; Landfester, K.; Mailänder, V.; Lieberwirth, I., Visualization of the protein corona: Towards a biomolecular understanding of nanoparticle-cell-interactions. *Nanoscale* **2017,** *9* (25), 8858-8870.

55. Rahimi, M.; Ng, E. P.; Bakhtiari, K.; Vinciguerra, M.; Ahmad, H. A.; Awala, H.; Mintova, S.; Daghighi, M.; Bakhshandeh Rostami, F.; De Vries, M.; Motazacker, M. M.; Peppelenbosch, M. P.; Mahmoudi, M.; Rezaee, F., Zeolite nanoparticles for selective sorption of plasma proteins. *Sci. Rep.* **2015,** *5*, 17259.

56. Garcíaalvarez, R.; Hadjidemetriou, M.; Sánchezíglesias, A.; Lizmarzan, L. M.; Kostarelos, K., *In vivo* formation of protein corona on gold nanoparticles: The effect of size and shape. *Nanoscale* **2017**, *10* (3), 1256-1264.