Synthesis of SWCNT-based nanosensors

Single-stranded DNA with single-walled carbon nanotube (ssDNA-SWCNT) suspensions were prepared with 1 mg of mixed-chirality SWCNTs (small diameter HiPco™ SWCNTs, NanoIntegris) and 1 mg of ssDNA (custom ssDNA oligos with standard desalting, Integrated DNA Technologies, Inc.) in 1 mL of 0.1 M phosphate-buffered saline (PBS). Solutions were bath sonicated for 10 min (Branson Ultrasonic 1800) and probe-tip sonicated for 10 min in an ice bath (3 mm probe tip at 50% amplitude, 5-6 W, Cole-Parmer Ultrasonic Processor). Samples were incubated at room temperature for 30 min then centrifuged to pellet insoluble bundles and contaminants (16.1 krcf, 30 min). Supernatant containing the product was collected. ssDNA-SWCNTs were spin-filtered to remove free ssDNA (Amicon Ultra-0.5 mL centrifugal filters with 100 kDa MWCO, Millipore Sigma) by washing with Milli-Q water two times (8 krcf, 5 min) then reversing the spin filter and centrifuging to recover sample (1 krcf, 5 min). ssDNA-SWCNT concentration was determined via sample absorbance at 632 nm (NanoVue Plus, GE Healthcare Life Sciences) and the extinction coefficient ε_{632nm}=0.036 L mg^{-1} cm^{-1}.[1] ssDNA-SWCNTs were stored at 4°C until use and then diluted to a working concentration of 100 mg L^{-1} in 0.1 M PBS.

Nanoparticle characterization

100 nm polystyrene nanoparticles were purchased from Polysciences, Inc. (PNPs; Fluoresbrite® yellow-green fluorophore-labeled). Size was confirmed with DLS to be in accordance with manufacturer specifications (**Figure S2b**) and measured zeta potential is -59.7 ± -3.24 mV **(Figure S3**; Zetasizer Nano, Malvern Panalytical; 1.67 g L^{-1} in 0.1 M PBS, 700 µL volume).

ssDNA-SWCNTs were synthesized as described above. Mixed-chirality HiPco™ SWCNTs are reported by the manufacturer (NanoIntegris) to have diameters 0.8-1.2 nm (average 1 nm; measured by Unidym from TEM) and lengths 100-1,000 nm (measured by Unidym from AFM). Upon suspension with ssDNA, previous AFM work informs a diameter of ~1 nm and length distribution centered around ~500 nm,[2] yet AFM sample deposition is known to influence such measurements in a DNA sequence-dependent manner.[3] Our previous work depicts ssDNA-SWCNT morphology by TEM.[4] Measured zeta potential of $(GT)_{15}$-SWCNTs is -19.4 ± 0.945 mV (**Figure S3**; Zetasizer Nano, Malvern Panalytical; 28.67 mg L^{-1} in 0.1 M PBS, 700 µL volume). Absorbance and fluorescence spectra of $(GT)_{15}$-SWCNTs are presented in **Figure S1**, confirming formation of a stable SWCNT dispersion[5] and in agreement with previous literature.[6] Absorbance of 30 mg L^{-1} $(GT)_{15}$-SWCNTs in 0.1 M PBS was measured in a 700 µL volume, black-sided quartz cuvettes (Thorlabs, Inc.) with a UV-VIS-nIR spectrophotometer (Shimadzu UV-3600 Plus). Fluorescence was obtained with an inverted Zeiss microscope (Axio Observer.D1, 10x objective) coupled to a Princeton Instruments spectrometer (SCT 320) and liquid nitrogen cooled Princeton Instruments InGaAs detector (PyLoN-IR). Fluorescence was measured in a glass-bottom 384 well-plate format (30 µL volume sample, 10 mg L^{-1} concentration in 0.1 M PBS), with a 721 nm laser (OptoEngine LLC) excitation light source and 800 – 1400 nm emission wavelength range.

We have previously determined approximately 140 $(GT)_{15}$ molecules or 364 $(GT)_6$ molecules adsorbed per SWCNT.[7] Using $(GT)_{15}$ and $(GT)_6$ contact areas from MD simulations,[6] this translates to ssDNA surface coverages of 2.1% and 6.5%, respectively. Previous work reports ~20-25% surface coverage of ssDNA on SWCNTs in the saturation regime (i.e. when further ssDNA adsorption is sterically unfavorable).[8–10] To capture the differing experimental and modeling conditions, we report the full range of ~1-25% initial ssDNA surface coverage on the SWCNT.

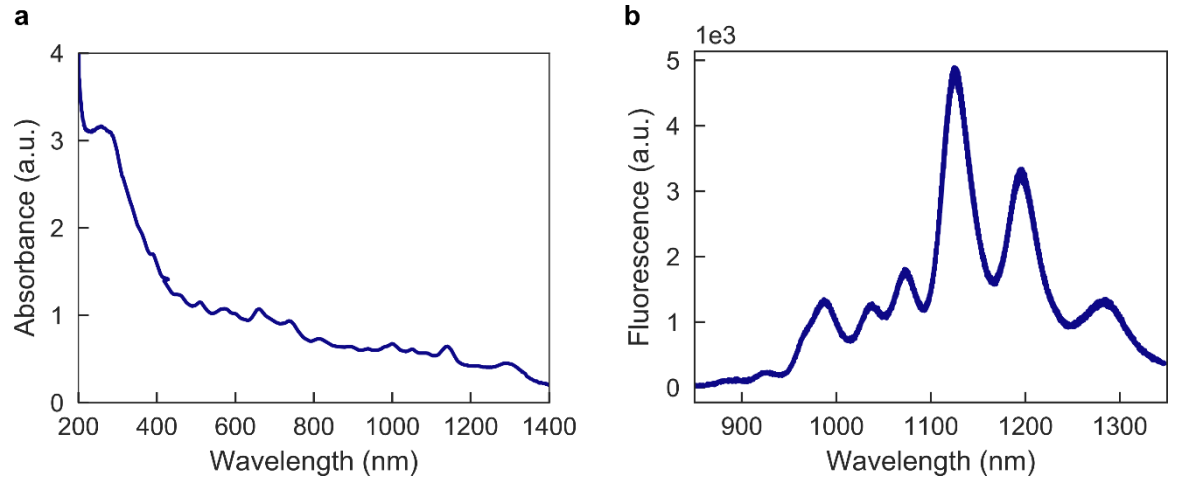


Figure S1. Optical characterization of $(GT)_{15}$-SWCNTs. (a) Absorbance spectrum of 30 mg L^{-1} $(GT)_{15}$-SWCNTs in 1X PBS. **(b)** Fluorescence spectrum of 10 mg L^{-1} $(GT)_{15}$-SWCNTs in 0.1 M PBS. Stable SWCNT suspension in aqueous medium is confirmed by absorbance peaks across the visible and near-infrared range and fluorescence emission that would otherwise be quenched in a SWCNT-aggregated state.

Isolation and characterization of protein-nanoparticle complexes

Protein corona composition was studied on PNPs, $(GT)_{15}$-SWCNTs, and $(GT)_6$-SWCNTs. PNPs were vortexed prior to use (1 min in 5 s pulses). Biofluids studied were human blood plasma and human CSF (**Table S6**), obtained with informed consent from all donors and in acordance with the relevant guidelines and regulations. CSF was concentrated 10X prior to incubation to match protein to nanoparticle ratios under volume constraints (14 krcf, 30 min; Amicon Ultra-0.5 mL centrifugal filters with 3 kDa MWCO, Millipore Sigma). The ratio of protein concentration to nanoparticle surface area was maintained constant for each respective nanoparticle in different biofluids, with 26 g L^{-1} protein per m^2 nanoparticle surface area for PNPs (from previous literature[11]) and 200 g L^{-1} protein per m^2 nanoparticle surface area for $(GT)_6$- and $(GT)_{15}$-SWCNTs. Based on experimental optimization, an 8-fold higher ssDNA-SWCNT surface area relative to PNP was selected to collect enough protein material from the SWCNT corona for downstream characterization, due to significantly lower protein adsorption on SWCNTs compared to PNPs. These incubation ratios translate to 1.67 g L^{-1} PNPs with 2.67% (v/v) plasma; 0.4 g L^{-1} PNPs with 8.67% (v/v) 10X CSF; 28.67 mg L^{-1} $(GT)_{15}$-SWCNTs with 2.67% (v/v) plasma; and 12.67 mg L^{-1} $(GT)_{15}$-SWCNTs with 16% (v/v) 10X CSF. Biofluid percentages are nominal, and were adjusted on a mass basis to match the target protein per surface area ratios. Nanoparticles were incubated with biofluids in 0.1 M PBS, 750 µL total volume, for 1 h at ambient temperature (**Figure S2a**). Protein-nanoparticle complexes were pelleted by centrifugation (16.1 krcf, 20 min). Supernatant containing unbound proteins was removed, the pellet resuspended in 0.1 M PBS, and the pellet broken up by pipetting. Washing was repeated three times to ensure removal of unbound proteins.

Each step was validated for polystyrene nanoparticles (PNPs) exposed to blood plasma proteins as follows: (i) incubation of proteins with nanoparticles induced an increase in nanoparticle hydrodynamic radius as determined by dynamic light scattering (DLS), where the number distribution shifted to a larger peak center and broadened out due to nonuniform aggregate formation as protein to nanoparticle loading was increased (**Figure S2b**); (ii) proteins initiated nanoparticle aggregation, as shown by solution absorbance before and after initial pelleting (**Figure S2c**), thus facilitating nanoparticle recovery for analysis; (iii) three washing steps were sufficient to remove unbound proteins by quantifying proteins remaining in the supernatant (**Figure S2d**; also valid for all nanoparticle/biofluid combinations); and (iv) proteins were fully eluted from nanoparticles by boiling in solutions of sodium dodecyl sulfate/β-mercaptoethanol (SDS/βME, for 2D PAGE analysis; **Figure S2e**) and urea/dithiothreitol (urea/DTT, for LC-MS/MS analysis). The equivalent verification was performed with $(GT)_{15}$-SWCNTs, yet the high aspect ratio of SWCNTs precluded accurate DLS measurement. Zeta potentials of the

nanoparticle/plasma mixtures were determined as a proxy of the nanoparticle-protein complex surface charge, although this measurement captures a convolution of any free proteins, free nanoparticles, and nanoparticle-protein complexes (**Figure S3**). Zeta potential measurements of plasma proteins alone and nanoparticles alone reveal that the separate entities were initially negatively charged, whereby mixing results in a broadened zeta distribution of lower average magnitude than the nanoparticles alone. The measured reduction in effective surface charge implies some degree of protein adsorption to the nanoparticles and lowering of electrostatic repulsion, contributing to the experimentally observed colloidal instability upon combining nanoparticles with plasma, in agreement with previous literature.[12]

Towards (i), the incubation solution was characterized by dynamic light scattering and zeta potential measurements in folded capillary zeta cell disposable cuvettes (Zetasizer Nano, Malvern Panalytical; 700 µL volume). PNPs are negatively charged as a result of initiator fragments from the polymerization process, yet these PNPs are conventionally considered to be a model plain nanoparticle due to no explicit functionalization.[11] $(GT)_{15}$-SWCNTs are slightly negatively charged due to the presence of the ssDNA on the surface, with the phosphate backbone extending into solution. Towards (ii), absorbance spectra were measured in a 700 µL volume, black-sided quartz cuvettes (Thorlabs, Inc.) with a UV-VIS-nIR spectrophotometer (Shimadzu UV-3600 Plus). For (iii), free protein remaining in the supernatant after centrifugation was quantified during subsequent wash steps using the Qubit Protein Assay (Thermo Fisher Scientific). Note that PNPs contribute minimally to the Qubit signal (~2%), therefore the protein mass calculated for wash 0 is slightly inflated. For (iv), eluted protein from the nanoparticle was quantified using the Pierce 660nm Assay (with Ionic Detergent Compatibility Reagent; Thermo Fisher Scientific). Elution buffer was modified from SDS/βME for 2D PAGE to urea/DTT for LC-MS/MS analysis due to SDS interference with trypsin digestion, reverse-phase HPLC, and electrospray ionization efficiency.[13] The profile of eluted proteins was confirmed to be invariable to the elution system by 2D PAGE and S-trap (Protifi) LC-MS/MS analysis, although total eluted protein amount decreased.

Nanoparticle mass loss during pelleting and washing was estimated by measuring solution absorbance of each collected supernatant after centrifugation. This measured mass loss serves as a maximum estimate due to scattering of solubilized proteins and any remaining protein-nanoparticle aggregates that increase the absorbance baseline and impede fully accurate quantification of the nanoparticles alone. After each centrifugation step as shown in **Figure S2a** (four total), the supernatant was removed and absorbance was measured (NanoDrop™ One/OneC Microvolume UV-Vis Spectrophotometer). For PNPs, absorbance was measured at the excitation maximum of the fluorophore (441 nm) and a standard curve over the relevant absorbance range (linear fit, R^2 = 0.9986) was used to convert this to concentration using Beer-Lambert's Law. For $(GT)_{15}$-SWCNTs, absorbance was measured at 632 nm and the known extinction coefficient was applied similarly. Results are presented in **Table S1**, with standard deviations of technical triplicate measurements and "0" denoting absorbance reading at the noise level of the instrument (e.g. absorbance ≤ 0.0133, read for buffer). The mass loss percentage is calculated as the ratio of this measured total mass removed to the calculated initial mass added to solution. We conclude that the maximum mass loss estimates of ~12% for PNPs and ~32% for $(GT)_{15}$-SWCNTs in each biofluid are not a significant portion of the population.

As a control, in the absence of nanoparticles in the incubation step, no measurable protein was present after pelleting and denaturation. This result implies that we are measuring selective protein adsorption to nanoparticles, not merely to the container, nor simply seeing the high background of proteins in biofluids. This latter point is further confirmed by the result that protein corona abundance does not scale as a function of native abundance on ssDNA-SWCNTs (**Figure S11**). Contamination of the isolated protein corona with bio-nanoparticles, such as extracellular vesicles and lipoproteins,[14] was inferred by the aforementioned control (no "protein corona" measurable in the absence of nanoparticles) and the absence of large peaks in the plasma-alone DLS (**Figure S2b**).

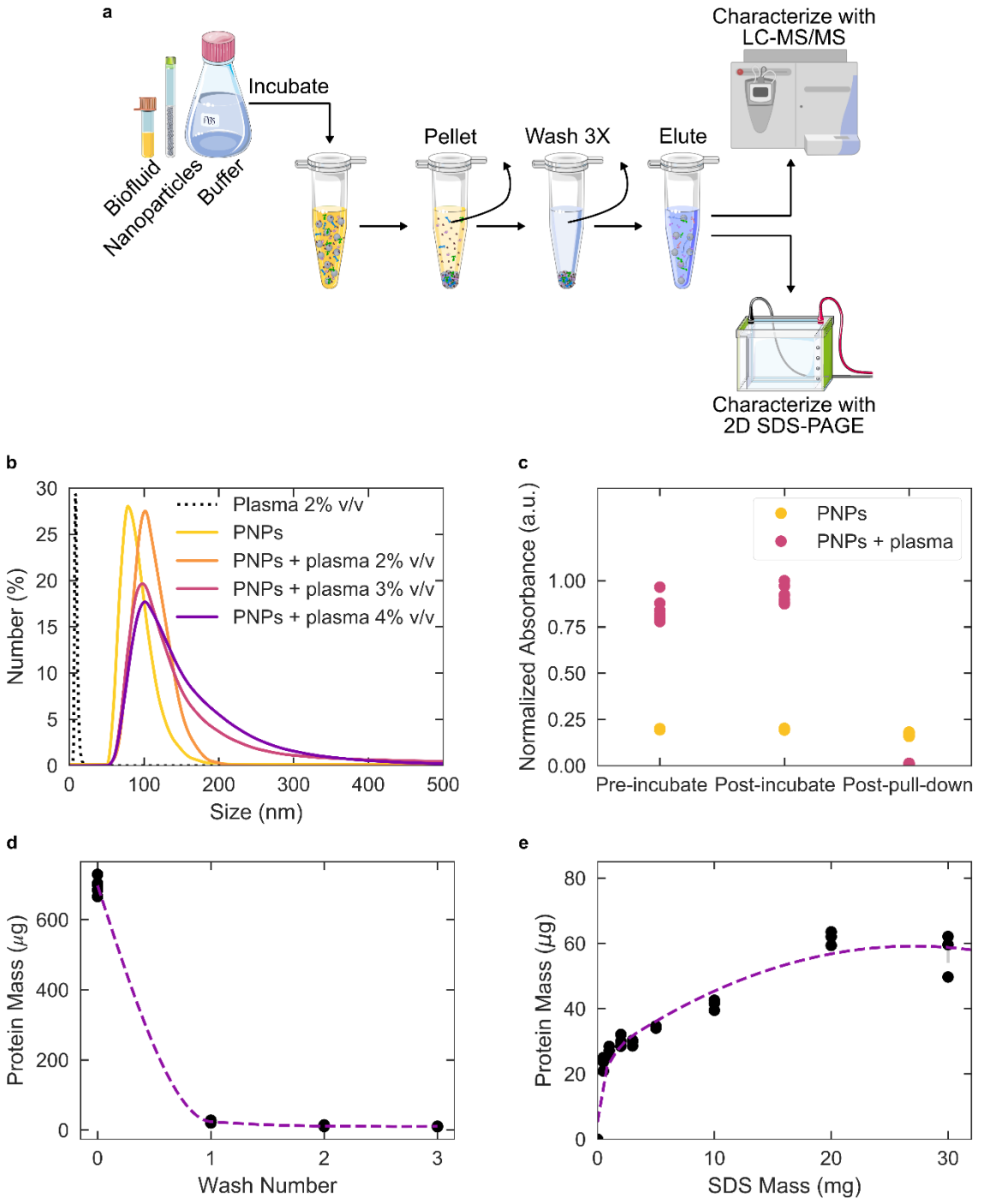


Figure S2. Isolation and characterization of protein-nanoparticle complexes to determine protein corona composition on nanoparticles. (a) Schematic detailing experimental procedure: nanoparticles are incubated with the desired biofluid in buffered solution, nanoparticle-protein complexes are pelleted by centrifugation and washed three times to remove non-selectively pelleted proteins, and corona proteins are eluted and characterized by two-dimensional polyacrylamide gel electrophoresis (2D PAGE) or liquid chromatography-tandem mass spectrometry (LC-MS/MS). **(b)** Dynamic light scattering (DLS) reveals that plasma protein corona formation induces an increase in the hydrodynamic radius of the PNPs (1.67 g L^{-1} in 0.1 M PBS) via peak shifting and broadening. **(c)** Absorbance at PNP excitation max (441 nm) immediately after adding plasma to incubation solution, incubating for 1 hour, and after the first pelleting step demonstrates the presence of proteins facilitates isolation of nanoparticles from solution in the initial pelleting step. **(d)** Quantification of free protein in solution via Qubit Protein Assay for varying wash number shows nearly complete depletion of free protein by three washes. **(e)** Quantification of eluted protein from nanoparticles via Pierce 660 nm Protein Assay with increasing SDS reducing buffer confirms complete elution of bound proteins from

nanoparticle surface prior to characterization. Error bars on (b)-(d) are ± standard error for experimental replicates of N = 6, 6, and 3, respectively.

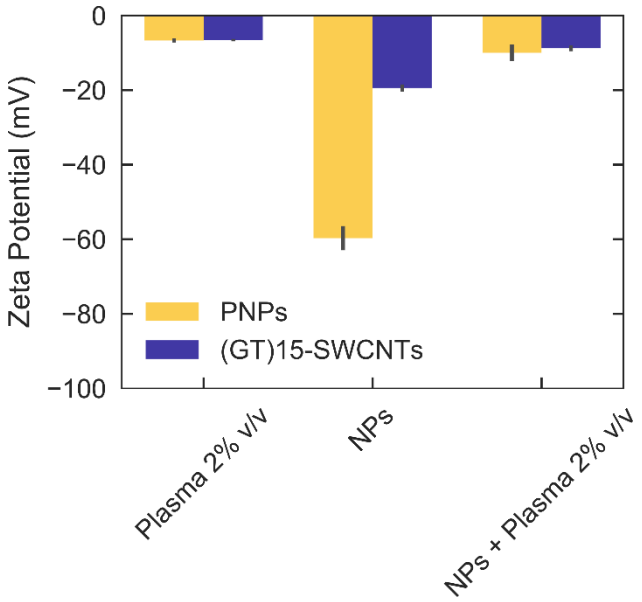


Figure S3. Surface charge changes induced by plasma protein corona formation. Zeta potential of native plasma, nanoparticles alone (PNPs yellow, $(GT)_{15}$-SWCNTs purple), and plasma protein-nanoparticle complexes. Lower magnitude zeta potential of protein-nanoparticle complexes indicates reduction in colloidal stability in the presence of surface-adsorbed proteins, as expected by visible aggregates formed. PNPs are 1.67 g L^{-1} and $(GT)_{15}$-SWCNTs are 28.67 mg L^{-1}, in 0.1 M PBS, 700 µL volume. Error bars are ± standard deviation for technical replicates (N = 3).

Table S1. Nanoparticle mass loss during corona isolation.

Wash	Mass [µg] PNPs (Plasma)	Mass [µg] PNPs (CSF)	Mass [µg] $(GT)_{15}$-SWCNTs (Plasma)	Mass [µg] $(GT)_{15}$-SWCNTs (CSF)
0	103.45 ± 29.40	37.51 ± 0.00	6.25 ± 0.21	3.33 ± 0.21
1	19.33 ± 1.97	0.00	0.00	0.00
2	11.37 ± 1.97	0.00	0.00	0.00
3	10.23 ± 0.00	0.00	0.00	0.00
Total Mass Removed	144.37	37.51	6.25	3.33
Initial Calculated Mass	1250	300	215	9.5
Estimated Mass Loss %	*11.55%*	*12.50%*	*29.07%*	*35.09%*

Composition studies by two-dimensional polyacrylamide gel electrophoretic separation (2D PAGE)

2D PAGE was performed to identify proteins via separation by isoelectric point in the first dimension and molecular weight in the second dimension. For analysis by 2D PAGE, bound proteins were eluted from nanoparticles by heating at 95°C for 10 min in SDS/BME reducing buffer (2% SDS, 5% β-mercaptoethanol, 0.066 M Tris-HCl). 1D separation was run according to the O'Farrell protocol[15] (adapted for Bio-Rad Mini-PROTEAN Tube Cell). Briefly, 1D sample buffer (8 M urea, 2% Triton X-100, 5% β-mercaptoethanol, 2% total carrier ampholytes - 1.6% Bio-Lyte 5/7, 0.4% Bio-Lyte 3/10) was added to samples in a 1:1 or 0.07:1 volume ratio (relative to initial plasma and CSF volumes, respectively) and incubated for 10 min. 1D separation was carried out in capillary tube PAGE with gel composition of 4% acrylamide (total monomer), 8 M urea, 2% Triton X-100, 2% total carrier ampholytes, 0.02% ammonium persulfate (APS), and 0.15% Tetramethylethylenediamine (TEMED). 25 µL sample and 25 µL 1D sample overlay buffer (4 M urea, 1% total carrier ampholytes) was loaded per capillary tube gel. Upper and lower chamber buffers were 100 mM sodium hydroxide and 10 mM phosphoric acid, respectively. 1D separation was run at 500 V for 10 min, 750 V for 3.5 h. Nanoparticles were filtered from the eluted

proteins by the gel itself. Capillary gels were extruded and loaded onto 2D gels. 2D separation was run according to the Laemmli protocol[16] (adapted for Bio-Rad Mini-PROTEAN Tetra Cell). Briefly, SDS/BME reducing buffer was added to the 2D well to cover the capillary gel and incubated for 10 min. 2D separation was carried out in 1 mm vertical mini gel format with a discontinuous buffer system under denaturing conditions. Gel composition was 12% acrylamide (total monomer), 0.375 M Tris-HCl, 0.1% SDS, 0.05% APS, 0.05% TEMED for the resolving gel and 12% acrylamide (total monomer), 0.125 M Tris-HCl, 0.1% SDS, 0.05% APS, 0.1% TEMED for the stacking gel. Electrode buffer was 25 mM Tris, 192 mM glycine, and 3.5 mM SDS (pH 8.3). 2D separation was run at 200 V for 1 h. Gels were extracted and silver stained according to Bio-Rad's Silver Stain Plus protocol and identified with ExPASy's SWISS-2DPAGE database (**Figure S4**).[17]

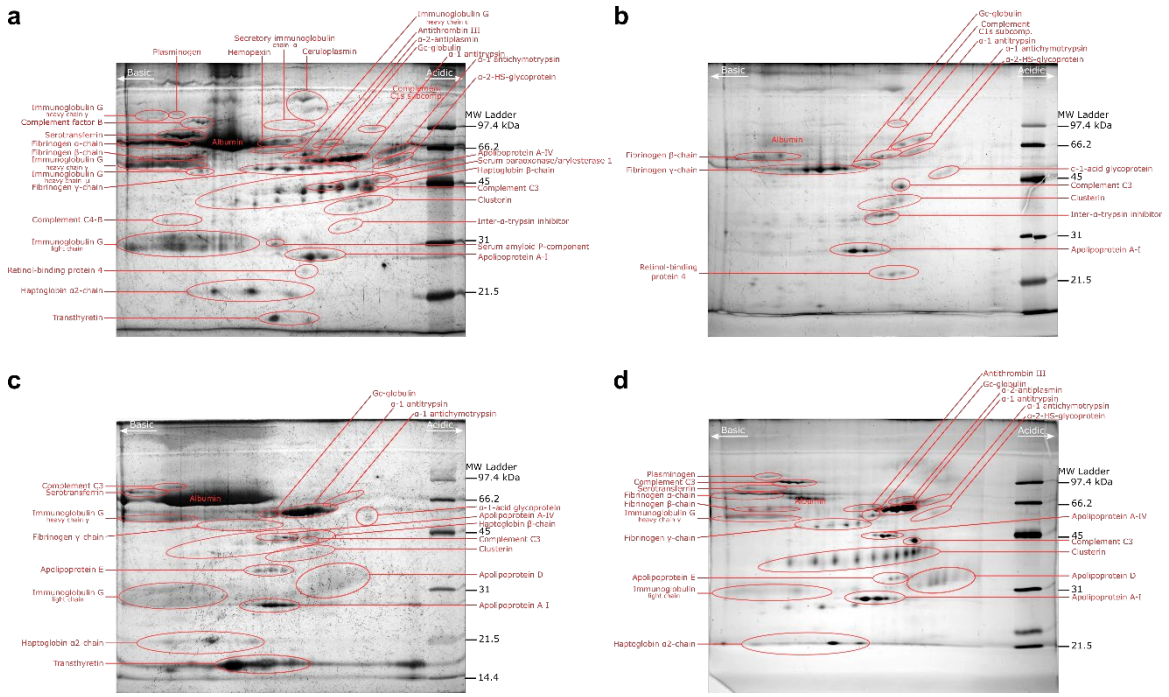


Figure S4. Representative 2D PAGE gels. (a) Plasma alone, **(b)** Plasma protein corona composition formed on $(GT)_{15}$-SWCNTs, **(c)** CSF alone, and **(d)** CSF protein corona composition formed on PNPs.

Composition studies by liquid chromatogrpahy-tandem mass spectrometry (LC-MS/MS)

Incubation amounts were scaled up by factors of 2, 4, 3, and 8 for PNPs/plasma, $(GT)_x$-SWCNTs/plasma, PNPs/CSF, and $(GT)_{15}$-SWCNTs/CSF, respectively, to ensure enough protein mass for analysis while retaining the original protein to nanoparticle ratios. Bound proteins were eluted from nanoparticles by heating at 37°C for 60 min in urea/DTT reducing buffer (8 M urea, 5 mM DTT, 50 mM Tris-HCl, pH 8). Eluted protein concentration was determined with the EZQ Protein Quantitation Kit (Thermo Fisher Scientific). Protein solution was centrifuged to pellet the majority of nanoparticles (16 krcf, 20 min) and this supernatant was spin-filtered to concentrate and remove impurities (14 krcf, 30 min; Amicon Ultra-0.5 mL centrifugal filters with 3 kDa MWCO, Millipore Sigma; pre-rinsed). Proteins were alkylated with 15 mM iodoacetamide for 30 min in the dark. 500 mM DTT was added to quench excess iodoacetamide in a volume ratio of 3:1 and incubated for 20 min. The reaction was diluted 1:1 with 50 mM Tris-HCl pH 8 to allow enzymatic protein digestion. In-solution protein digestion was done with a ratio of 1:25 w/w Trypsin/Lys-C (Mass Spectrometry Grade, Promega) to protein, overnight at 37°C. Any remaining nanoparticles were removed by spin filtering (14 krcf, 30 min; Amicon Ultra-0.5 mL centrifugal filters with 30 kDa MWCO, Millipore Sigma; pre-rinsed). Nanoparticle removal was done after protein digestion into peptides due to the otherwise very similar sizes of nanoparticles and proteins. Peptide concentration was determined with the Pierce Peptide Quantitation Kit (Thermo Fisher Scientific) and samples were

normalized to 0.1 g L^{-1} in 100 µL total volume. Peptide solutions were spiked with 50 fmol of E. coli housekeeping peptide (Hi3 Ecoli Standard, Waters) per 5 µL sample volume to allow for protein quantification. Digestion was terminated by freezing samples to -20°C. Note that biofluid-alone samples underwent these same processing steps, from denaturation to tryspin digestion. The preceding isolation steps of pelleting and washing were only necessary for nanoparticle-protein complexes, and were accordingly omitted for biofluids alone. An alternative mass spectrometry preparation technique was pursued, using S-traps (Protifi), confirming our results were not biased by the sample preparation protocol.

Proteolytically digested proteins were analyzed using a Synapt G2-S*i* mass spectrometer equipped with a nanoelectrospray ionization source and connected directly in line with an Acquity M-class ultra-performance liquid chromatography system (UPLC; Waters, Milford, MA). This instrumentation is in the California Institute for Quantitative Biosciences (QB3)/College of Chemistry Mass Spectrometry Facility at UC Berkeley. Data-independent, ion mobility-enabled mass spectra and tandem mass spectra[18–20] were acquired in the positive ion mode. Data acquisition was controlled with MassLynx software (version 4.1) and tryptic peptide identification and quantification using a label-free approach[21–23] were performed with Progenesis QI for Proteomics software (version 4.0, Waters).

Proteomic mass spectrometry data interpretation

Prior to LC-MS/MS analysis, all samples were normalized on a total protein mass basis (where normalizing on a total molar basis is experimentally not feasible due to the complexity of biofluid samples). Consequently, the reported abundance of each protein species *i,* b_i, is the ratio of mole number of protein *i,* n_i, to the total protein mass:

$$b_i = \frac{n_i}{\sum_j n_j MW_j}$$

where MW_j is the molecular weight of each protein species *j.* LC-MS/MS data is then expressed as the fold change ε_i between the abundance of protein species *i* in the corona on the nanoparticle surface (phase *s*) to that in the bulk biofluid (phase *f*):

$$\varepsilon_i = \frac{b_i^s}{b_i^f} = \left(\frac{n_i^s}{n_i^f}\right)\left(\frac{\sum_j n_j^f MW_j}{\sum_j n_j^s MW_j}\right)$$

Here, the second term in parentheses is equal to 1 because all samples have the same total protein mass. Therefore, the reported fold change is the molar abundance ratio of a particular protein in the corona phase to that in the bulk biofluid phase.

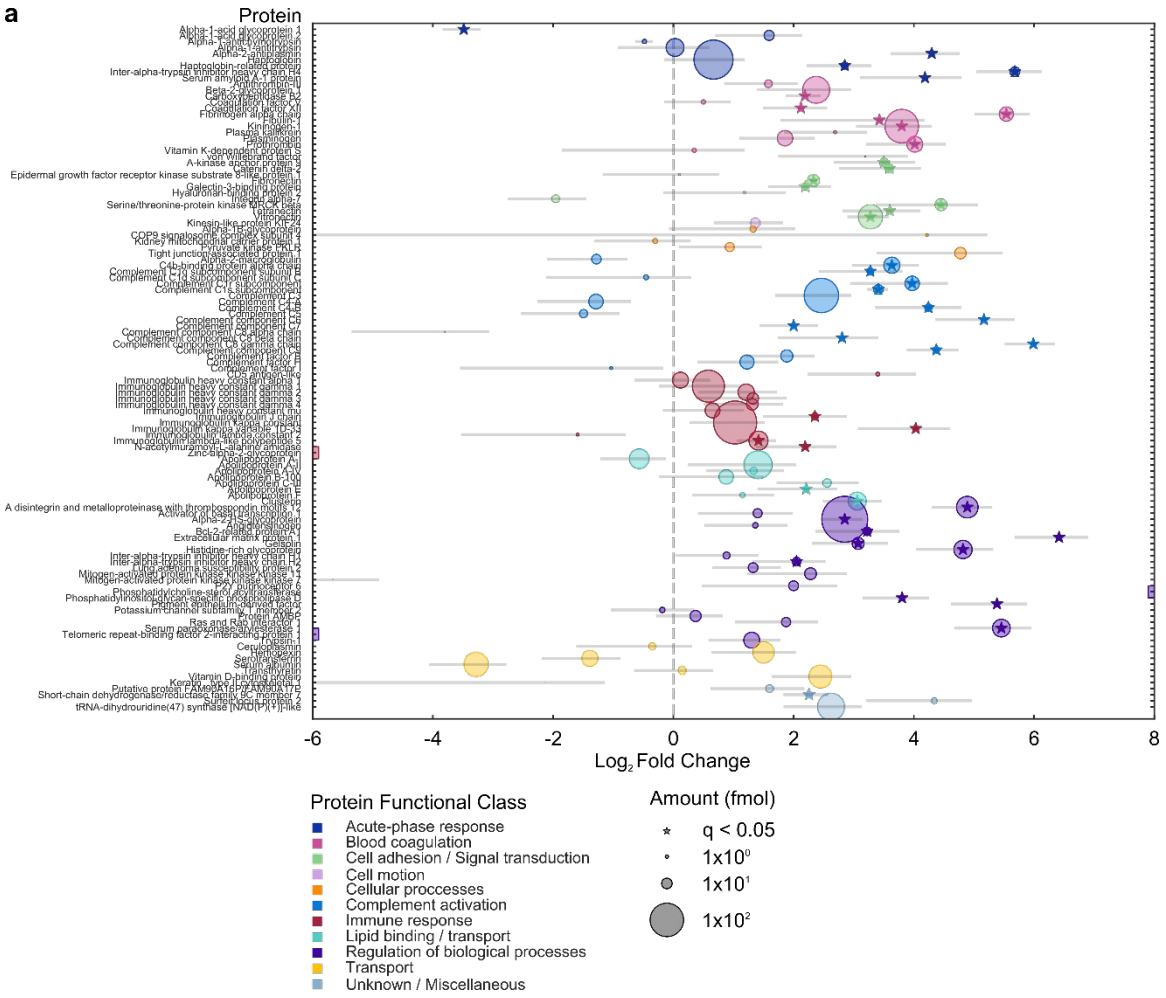
Extended discussion on protein corona constituents identified by proteomic mass spectrometry

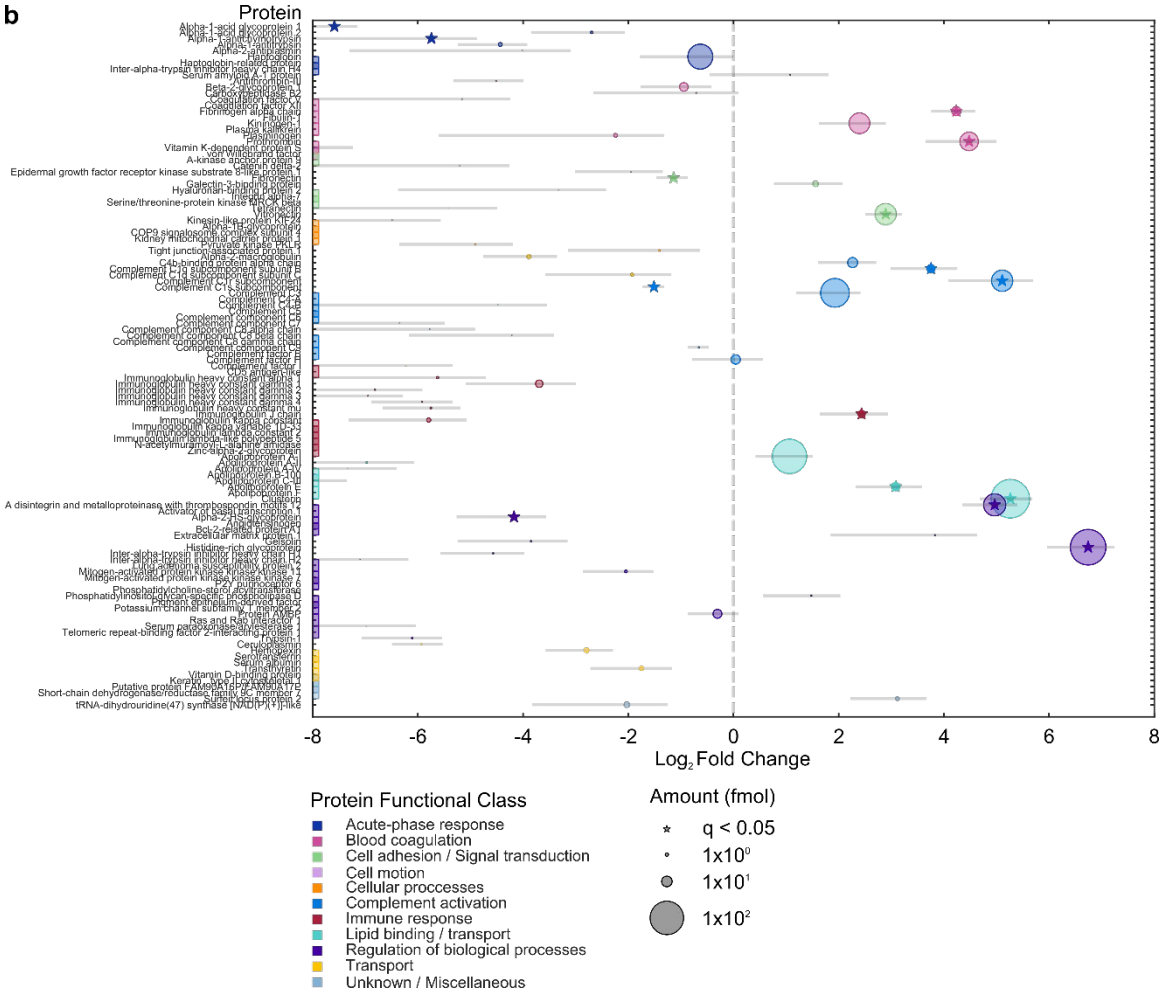
Albumin composes 55% (w/v) of proteins in blood plasma, corresponding to 35-50 g L^{-1},[24] and is often assumed to comprise a representative constituent in nanoparticle protein coronas. In consequence, many nanotechnologies are tested for functionality in serum instead of plasma.[1,47,48] Although albumin alone has been known to disperse SWCNTs in aqueous solution under sonication conditions,[28,29] here we note that albumin is likely unable to outcompete higher affinity proteins when in the presence of a complex biofluid. Accordingly, we hypothesize that albumin plays a minimal role in the full plasma protein corona and subsequent *in vivo* trafficking and fate.

The most highly enriched plasma protein, histidine-rich glycoprotein (107-fold enrichment), has been found as a representative plasma corona protein in prior studies on silica nanoparticles and is hypothesized to interact with other plasma proteins to enter the corona.[12] Another greatly enriched plasma protein, unreported in prior carbon nanoparticle corona literature, is ‘a disintegrin and metalloproteinase with thrombospondin motifs 12’ (ADAMTS12), which appeared in high abundance on both $(GT)_{15}$-SWCNTs (6^{th} highest abundance) and $(GT)_6$-SWNCTs (1^{st} highest abundance). ADAMTS12 is a metalloprotease with a zinc cofactor thought to possess anti-tumorigenic properties and to play a role in cell adhesion, pointing to potential applications in protein-SWCNT construct design.

Plasma proteins displaying high abundance and enrichment on $(GT)_{15}$-SWCNTs have functional roles in (i) lipid binding/transport (150% fold change over the average of all protein classes) and (ii) complement activation (140% fold change) (**Figure 1** and **Figure S7**). A key example of a lipid binding/transport corona protein is clusterin (a.k.a. apolipoprotein J) as the most abundant plasma protein in the $(GT)_{15}$-SWCNT corona (38-fold enrichment), adsorption of which has shown promise in reducing non-specific cellular uptake of other types of nanocarriers.[30] Apolipoproteins broadly act as dysopsonins that promote prolonged circulation in the blood.[12] Hydrophobic interactions are posited to drive apolipoprotein adsorption in mimicry of native functions, such as apolipoprotein A-I (3rd most abundant) that binds and transports hydrophobic fats through aqueous environments.[31,32] We expect apolipoprotein binding, including clusterin and apolipoprotein A-I, to have a considerable impact on intracellular trafficking and fate of nanoparticles *in vivo*.[11,33] Regarding group (ii), complement C3 is the 4th most abundant plasma protein on $(GT)_{15}$-SWCNTs, with a 14-fold enrichment. This result is in agreement with previous literature that SWCNTs activate the complement system.[32,34,35] Binding of complement proteins leads to nanoparticle opsonization if the bound proteins are intact, which could be useful in developing targeted therapies, yet detrimental if longer bloodstream circulation time is desired.[12,36] The ssDNA-SWCNT surface may present an array of adsorbed corona proteins that the complement system deems as “foreign”, thus activating complement systems and leading the coated SWCNT to act as an adjuvant that increases immune response. Complement proteins may bind either directly to the SWCNT surface (as was found for complement component 1q, or C1q, on double-walled carbon nanotubes[32]), or interact with other plasma proteins adsorbed on the SWCNT (where, for example, C1q binds to immunoglobulins, IgG and IgM, or fibronectin). Unfortunately, even a low degree of C1q binding can activate the complement system due to the amplification steps involved in the pathway. Yet, if the SWCNT serves to either locally sequester proteins that initiate complement activation (such as complement C3, 4th most abundant) or the corona contains down-regulators in their native state (such as complement factor H, 12th most abundant), this could in turn bypass recognition and complement activation. In contrast to the high representation of complement proteins in the corona that serve a role in the innate immune response, it is interesting to note the low corona representation of immunoglobulins (81% lower fold change than average, proteins involved in the adaptive immune response.

In addition to groups (i) and (ii), blood coagulation proteins are enhanced 76% on $(GT)_{15}$-SWNCTs in plasma, but the wide distribution of regression coefficients for these proteins precludes making statistically significant conclusions about this class (**Figure S8**). A notable enriched blood coagulation protein is fibrinogen, with 19-fold enrichment in the $(GT)_{15}$-SWCNT corona relative to concentrations in plasma. Fibrinogen’s presence in the corona is unfavorable, as fibrinogen is responsible for eliciting inflammatory responses and nanoparticle aggregation.[12,37,38] Fibrinogen is a large rod-like multimeric protein with alpha, beta, and gamma subunits. Although identified on 2D PAGE, fibrinogen beta and gamma chains were absent from not only the nanoparticle-biofluid LC-MS/MS results, but also the native plasma samples. Additionally, there was no improvement using elution with 5% SDS, 50mM TEAB and purification with the S-trap mini column (Protifi). However, the fibrinogen alpha chain was present and enriched from plasma. Based on the reproducible involvement of fibrinogen in the corona from 2D PAGE results and representation of the alpha chain in LC-MS/MS results, fibrinogen was concluded to bind to $(GT)_{15}$-SWCNTs.


a
Protein
-6
-4
-2
0
2
4
6
8
Log2 Fold Change
Protein Functional Class
Acute-phase response
Blood coagulation
Cell adhesion / Signal transduction
Cell motion
Cellular proccesses
Complement activation
Immune response
Lipid binding / transport
Regulation of biological processes
Transport
Unknown / Miscellaneous
Amount (fmol)
q < 0.05
1x10^0
1x10^1
1x10^2


b
Protein
-8
-6
-4
-2
0
2
4
6
8
Log2 Fold Change
Protein Functional Class
Acute-phase response
Blood coagulation
Cell adhesion / Signal transduction
Cell motion
Cellular proccesses
Complement activation
Immune response
Lipid binding / transport
Regulation of biological processes
Transport
Unknown / Miscellaneous
Amount (fmol)
q < 0.05
1x10^0
1x10^1
1x10^2

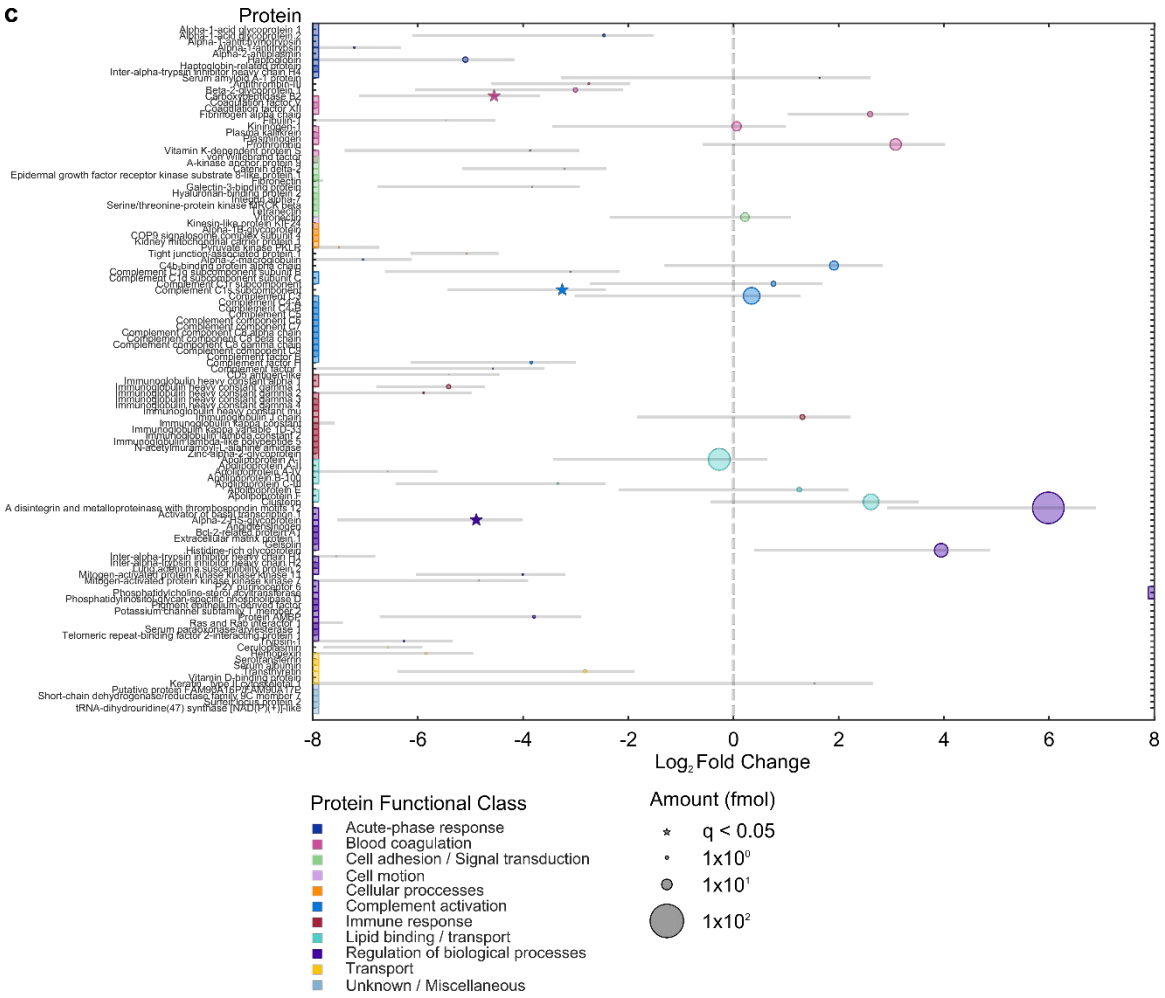
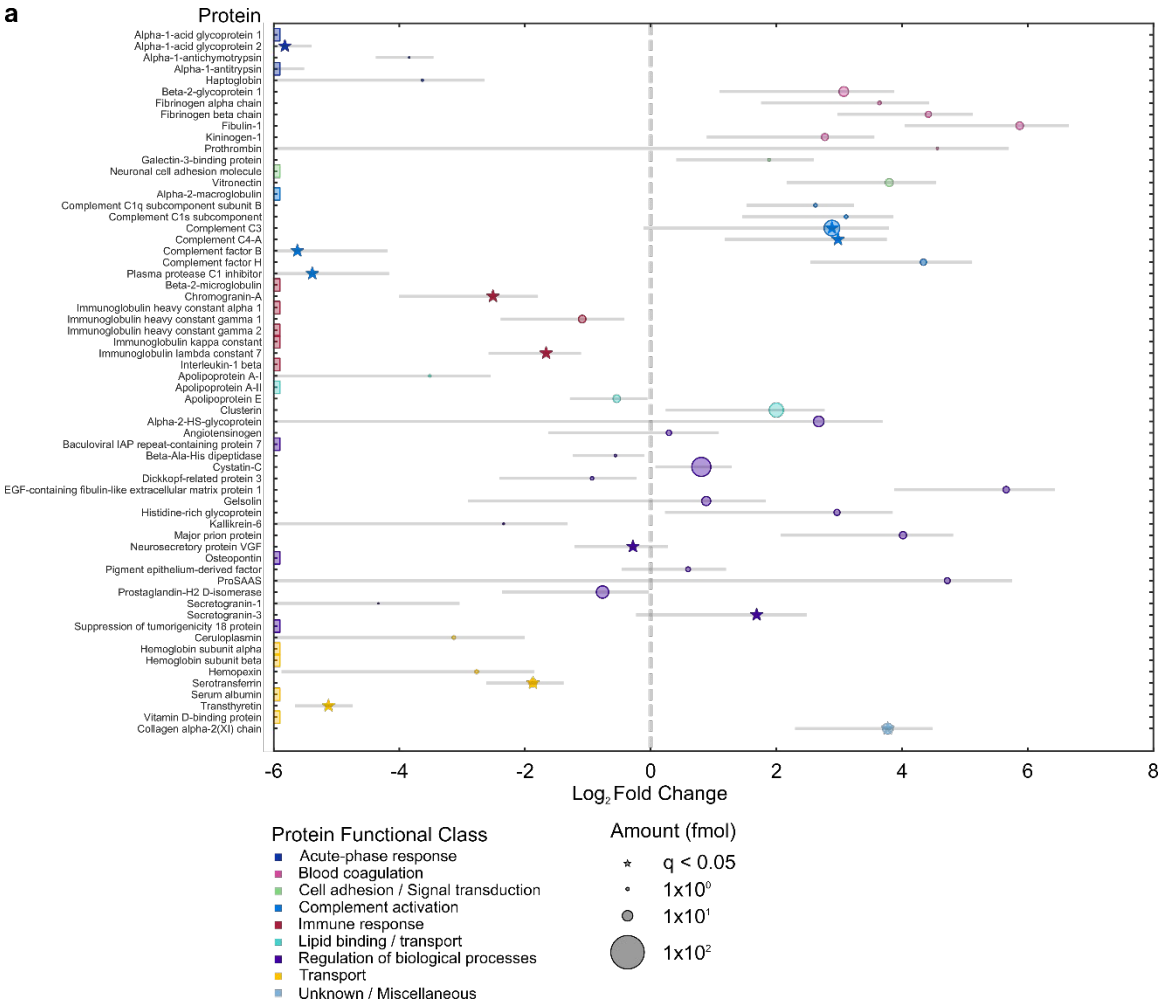


Figure S5. Blood plasma protein corona compositional map determined by proteomic mass spectrometry, full results. Protein corona formed from blood plasma on **(a)** PNPs, **(b)** $(GT)_{15}$-SWCNTs, and **(c)** $(GT)_6$-SWCNTs. Circle size corresponds to protein abundance (femtomolar). Proteins are grouped by functional class according to color (PANTHER).[39] Log_2 fold change is in comparison to the biofluid alone, e.g. log_2 fold change of zero indicates the same relative amount of protein exists in the corona as in bulk solution of the native biofluid, while < 0 is depletion and > 0 is enrichment. Names are included for proteins of interest or used for subsequent experiments. Colored boxes at x-axis limits indicate no protein detected in either corona ($x < 2^{-6}$ or 2^{-8}) or biofluid ($x > 2^8$). Error bars indicate standard error of fold change between experimental replicates (N = 3).

Table S2. Top 20 most abundant proteins identified by proteomic mass spectrometry in plasma $(GT)_{15}$-SWCNT and $(GT)_6$-SWCNT coronas.

Plasma	$(GT)_{15}$-SWCNTs in plasma	$(GT)_6$-SWCNTs in plasma
Serum albumin	Clusterin	A disintegrin and metalloproteinase with thrombospondin motifs 12
Haptoglobin	Histidine-rich glycoprotein	Apolipoprotein A-I
Ig kappa constant	Apolipoprotein A-I	Complement C3
Ig heavy constant gamma	Complement C3	Clusterin
Serotransferrin	Haptoglobin	Histidine-rich glycoprotein
Apolipoprotein A-I	A disintegrin and metalloproteinase with thrombospondin motifs 12	Prothrombin
Complement C4	Complement C1r subcomponent	Kininogen-1
Telomeric repeat-binding factor 2-interacting protein	Vitronectin	C4b-binding protein alpha chain
Alpha-1-antitrypsin	Kininogen-1	Vitronectin
Alpha-2-HS-glycoprotein	Prothrombin	Haptoglobin
Apolipoprotein A-II	C4b-binding protein alpha chain	Fibrinogen alpha chain
Ig heavy constant alpha 1	Complement factor H	Ig J chain
Integrin alpha-7	Fibrinogen alpha chain	Complement C1r subcomponent
Alpha-2-macroglobulin	Protein AMBP	Apolipoprotein E
Complement C3	Beta-2-glycoprotein 1	Beta-2-glycoprotein 1
Complement C5	Apolipoprotein E	Ig heavy constant gamma 1
Hemopexin	Complement C1q subcomponent subunit B	Alpha-2-HS-glycoprotein
Alpha-1-acid glycoprotein 1	Ig heavy constant gamma 1	Transthyretin
Ig heavy constant mu	Ig J chain	Protein AMBP
Beta-2-glycoprotein 1	Galectin-3-binding protein	Alpha-1-acid glycoprotein 2


a
Protein
Alpha-1-acid glycoprotein 1
Alpha-1-acid glycoprotein 2
Alpha-1-antichymotrypsin
Alpha-1-antitrypsin
Haptoglobin
Beta-2-glycoprotein 1
Fibrinogen alpha chain
Fibrinogen beta chain
Fibulin-1
Kininogen-1
Prothrombin
Galectin-3-binding protein
Neuronal cell adhesion molecule
Vitronectin
Alpha-2-macroglobulin
Complement C1q subcomponent subunit B
Complement C1s subcomponent
Complement C3
Complement C4-A
Complement factor B
Complement factor H
Plasma protease C1 inhibitor
Beta-2-microglobulin
Chromogranin-A
Immunoglobulin heavy constant alpha 1
Immunoglobulin heavy constant gamma 1
Immunoglobulin heavy constant gamma 2
Immunoglobulin kappa constant
Immunoglobulin lambda constant 7
Interleukin-1 beta
Apolipoprotein A-I
Apolipoprotein A-II
Apolipoprotein E
Clusterin
Alpha-2-HS-glycoprotein
Angiotensinogen
Baculoviral IAP repeat-containing protein 7
Beta-Ala-His dipeptidase
Cystatin-C
Dickkopf-related protein 3
EGF-containing fibulin-like extracellular matrix protein 1
Gelsolin
Histidine-rich glycoprotein
Kallikrein-6
Major prion protein
Neurosecretory protein VGF
Osteopontin
Pigment epithelium-derived factor
ProSAAS
Prostaglandin-H2 D-isomerase
Secretogranin-1
Secretogranin-3
Suppression of tumorigenicity 18 protein
Ceruloplasmin
Hemoglobin subunit alpha
Hemoglobin subunit beta
Hemopexin
Serotransferrin
Serum albumin
Transthyretin
Vitamin D-binding protein
Collagen alpha-2(XI) chain
-6
-4
-2
0
2
4
6
8
Log2 Fold Change
Protein Functional Class
Acute-phase response
Blood coagulation
Cell adhesion / Signal transduction
Complement activation
Immune response
Lipid binding / transport
Regulation of biological processes
Transport
Unknown / Miscellaneous
Amount (fmol)
q < 0.05
1x10^0
1x10^1
1x10^2

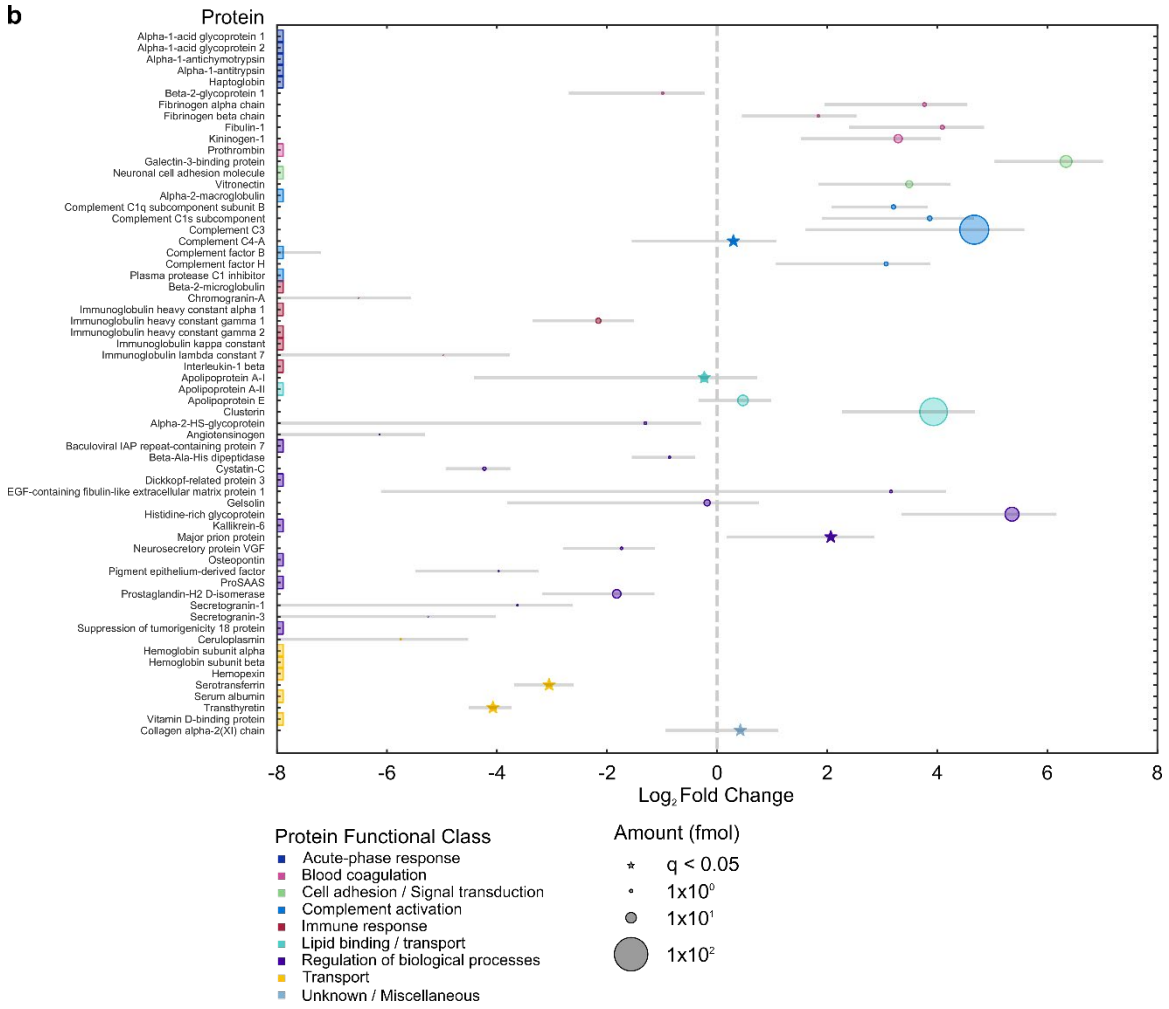


Figure S6. Cerebrospinal fluid (CSF) protein corona compositional map determined by proteomic mass spectrometry, full results. Protein corona formed from CSF on **(a)** PNPs and **(b)** $(GT)_{15}$-SWCNTs. Circle size corresponds to protein abundance (femtomolar). Proteins are grouped by functional class according to color (PANTHER).[39] Log_2 fold change is in comparison to the biofluid alone, e.g. fold change of zero indicates the same relative amount of protein exists in the corona as in bulk solution of the native biofluid, while < 0 is depletion and > 0 is enrichment. Names are included for proteins of interest or used for subsequent experiments. Colored boxes at x-axis limits indicate no protein detected in either corona ($x < 2^{-6}$ or 2^{-8}) or biofluid ($x > 2^{8}$). Error bars indicate standard deviation of fold change between technical replicates (N = 3).

Linear regression models for corona composition

We linearly regressed[40] the natural log of the fold change of proteins for each nanoparticle-biofluid pairing using two sets of protein descriptors. The first set of descriptors are categorical variables denoting what class a protein is in (*i.e.* 1 for a protein in a given class and 0 otherwise), namely, involved in acute-phase response, blood coagulation, cell adhesion/signal transduction, complement activation, immune response, lipid binding/transport, regulation of biological processes, transport, or miscellaneous/unknown (**Figure S7** and **Figure S8**; grouped according to PANTHER[39]). The variables were sum-effect coded such that the coefficients quantify how a protein class deviates from the grand mean of all protein classes and the intercept of the regression is the grand mean. Because each protein is grouped into one and only one class, the categorical variables are not linearly independent and one class is excluded from the regression;[40] we chose the miscellaneous class.

The second set of descriptors are molecular and biophysical properties of the proteins: protein mass, fraction of amino acids that are non-aromatic hydrophobic (sum of alanine, valine, isoleucine, leucine, and methionine content), hydrophilic (sum of serine, threonine, asparagine, glutamine content), arginine (R), histidine (H), lysine (K), acidic (sum of aspartic acid and glutamic acid content), phenylalanine (F), tyrosine (Y), tryptophan (W), number of glycosylated sites, number of ligand binding sites, number of metal binding sites, and number of disulfide binds **Figure 3**, **Figure S9**, and **Figure S10**). Each of these descriptors is a continuous variable. The regression coefficients quantify the fractional difference in the fold change for a unit increase in the independent variable.[40] Protein-specific information was acquired from UNIPROT.[41] Note that these particular descriptors were chosen after primary analyses that eliminated highly co-dependent descriptors. An example was choosing to include percentage of acidic/basic amino acids rather than protein isoelectric point (from ExPASy Compute pI/MW), where the isoelectric point was deemed less exact because it relies on a theoretical calculation, omits protein fragments, and necessitates an average value for multicomponent proteins. Other examples were including number of disulfide bonds as an estimate of protein stability rather than protein instability index and segmenting to percentage of hydrophobic/aromatic amino acids rather than grand average hydropathy (GRAVY) score, in both cases due to the involvement of arbitrarily set scales (from ExPASy ProtParam).

For each regression, we included the measured protein fold changes for each replicate of a nanoparticle-biofluid system and controlled for sample-to-sample variability by including a categorical variable for the specific replicate. Protein abundances that fell below the lower limit of detection in the samples from the protein corona were set to $1x10^{-5}$ fmol, corresponding to the lowest detected protein abundance of all systems. Left-censoring the data in this way provides a conservative estimate of the regression coefficients by underestimating of the magnitude and significance.[40] Calculated variance inflation factors for all variables in each independent regression was <4, indicating negligible multicollinearity between the independent variables. To avoid overestimating the statistical significance of independent variables, p-values were adjusted using the Benjamini-Hochberg false discovery rate procedure.[42] All statistical analysis was implemented in Python using the StatsModels V0.10.1 package [43](0.27-0.39). **Table S3** and **Table S4** provide coefficients, standard errors, false discovery rate corrected (FDRC) p-values, and R-squared values for each regression. The median R-squared of the first and second regression models for the nanoparticle-biofluid systems are 0.29 and 0.34, respectively, indicating the statistical models are descriptive rather than predictive. Nonlinear or decision tree algorithms provide more precise prediction of corona composition,[44] however, these approaches were not considered because they are not readily interpretable, which is a principle goal of our analysis.

Protein properties that were controlled for but that did not show a statistically significant effect on fold change for any nanoparticle in any biofluid include: the number of disulfide bonds (used as a proxy for protein stability), number of biomolecular binding sites, number of metal binding sites, and percentage of histidine or tryptophan. The lack of dependence on disulfide bond content and also instability index is surprising in the context of previous corona literature, which suggests that less structurally stable proteins are more surface active.[38]

Connecting linear regression model to thermodynamics

The ideal solution chemical equilibrium constant K_P of a protein in bulk solution P adsorbing onto a nanoparticle into the protein corona P^* is equal to the ratio of the surface concentration of the protein on the nanoparticle Γ_P to the concentration of the protein in solution c_P: [45]

$$K_P = \frac{\Gamma_P}{c_P}$$

The logarithm of the chemical equilibrium constant is related to ΔG_P^o, the change in standard state Gibbs free energy for a protein adsorbing from solution to the nanoparticle surface,[45]

$$\ln K_{\mathrm{P}} = -\frac{\Delta G_P^o}{RT}$$

where R is the ideal gas constant and T is temperature. The molar abundances of P measured by LC-MS/MS from the native biofluid, n_P^f, and eluted from the nanoparticle surface, n_P^s, are related to c_P and Γ_P according to:

$$c_P = \frac{n_P^f f_P^f}{V}$$

$$\Gamma_P = \frac{n_P^s f_P^s}{S}$$

where V is the volume and S is the total surface area of the nanoparticle in the nanoparticle-biofluid solution. f_P^f and f_P^s are the fraction of moles of P that enter the LC-MS/MS relative to amount in the fluid and on the nanoparticle surface, respectively. Dilution and steps in isolating the protein-nanoparticle complexes cause f_P^f and f_P^s to vary from unity. This analysis neglects changes to the solution protein concentration due to corona formation. Rearrangement puts the LC-MS/MS measured log-fold change, $\ln(n_P^s/n_P^f)$, in terms of dilution factors and free energy changes:

$$\ln\left(\frac{n_P^s}{n_P^f}\right) = -\ln\left(\frac{V f_P^s}{S f_P^f}\right) - \frac{\Delta G_P^o}{RT}$$

In comparison, linear regression of the log-fold change gives:

$$\ln\left(\frac{n_P^s}{n_P^f}\right) = \beta_0 + \sum_{i=1} \beta_i X_{i,P} + \epsilon_P$$

where β_i is a regression coefficient corresponding to $X_{i,P}$, the ith independent variable of protein P, and ϵ_P is the disturbance term that accounts for any factors other than $X_{i,P}$ controlling the fold changes.[40] Because LC-MS/MS sample preparation should not impact proteins differently, f_P^s and f_P^f are the same for all proteins for a set nanoparticle-biofluid system. Consequently, we can relate the chemical and statistical parameters:

$$\beta_0 = -\ln\left(\frac{V f_P^s}{S f_P^f}\right)$$

$$\sum_{i=1} \beta_i X_{i,P} + \epsilon_P = -\frac{\Delta G_P^o}{RT}$$

The regression coefficients β_i, therefore, relate the protein properties X_i to the Gibbs free energy change of proteins binding to the nanoparticle ΔG_P^o.

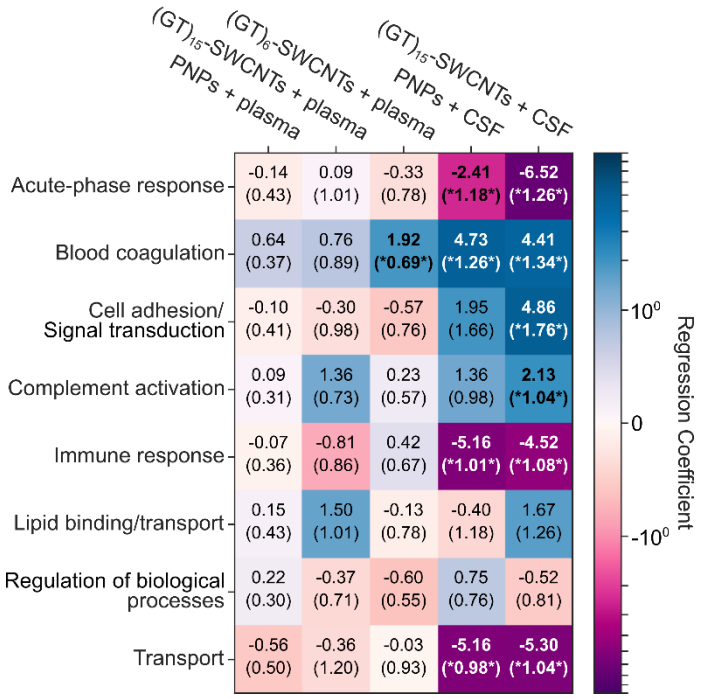


Figure S7. Role of protein functional class in protein corona formation for each nanoparticle-biofluid pairing. Ln-fold change, effect-coded regression coefficients of protein classes (rows) for each nanoparticle-biofluid pairing (columns). Cells are colored from dark purple (lower than the average fold change) to white (average fold change) to dark blue (higher than average fold change). Standard errors of the coefficients are given in parentheses. Results that have false-discovery-rate-corrected p-values of below 0.1 are bolded and noted with asterisks.

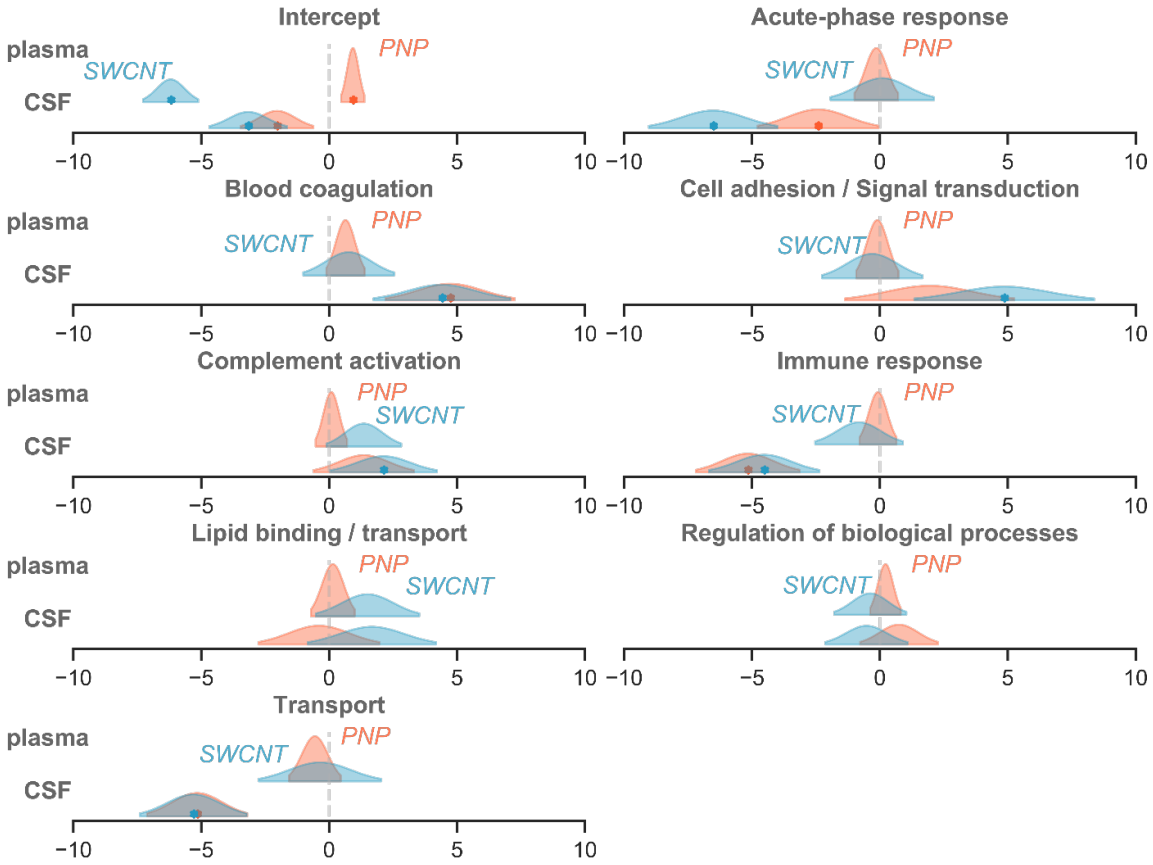


Figure S8. Distribution for protein class mean regression coefficients in each nanoparticle-biofluid pairing. Stars indicate false-discovery-rate adjusted p-values<0.1.

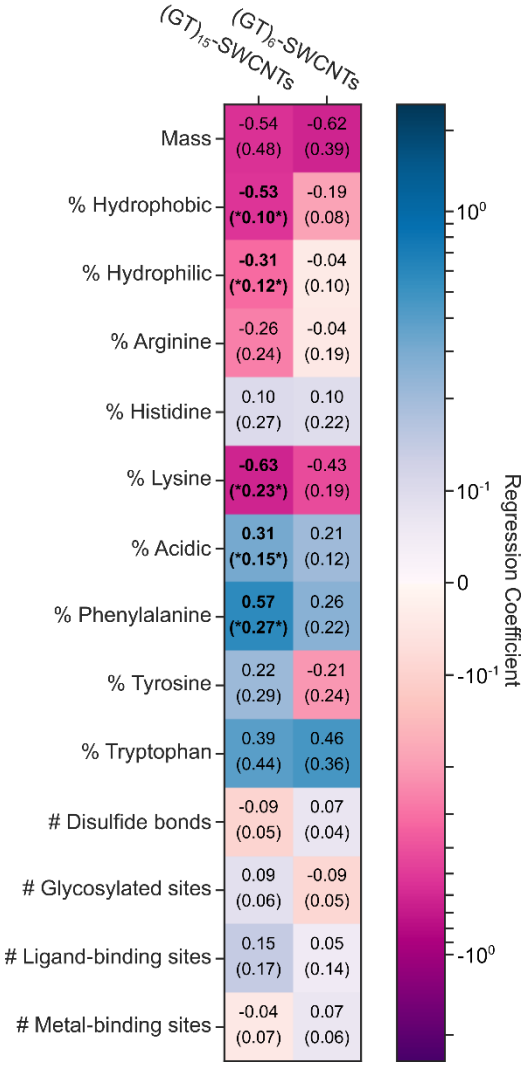


Figure S9. Molecular attributes of proteins that govern protein corona formation for $(GT)_x$-SWCNTs in plasma. Ln-fold change regression coefficients for molecular attributes of proteins (rows) for each nanoparticle-biofluid pairing (columns). Cells are colored from dark purple (negative effect on fold change) to white (no effect) to dark blue (positive effect). Standard errors of the coefficients are given in parenthesis. Results that have false-discovery-rate-corrected p-values below 0.1 are bolded and noted with asterisks. Amino acid groupings include: non-aromatic hydrophobic (sum of alanine, valine, isoleucine, leucine, and methionine content), hydrophilic (sum of serine, threonine, asparagine, glutamine content), and acidic (sum of aspartic acid and glutamic acid content).

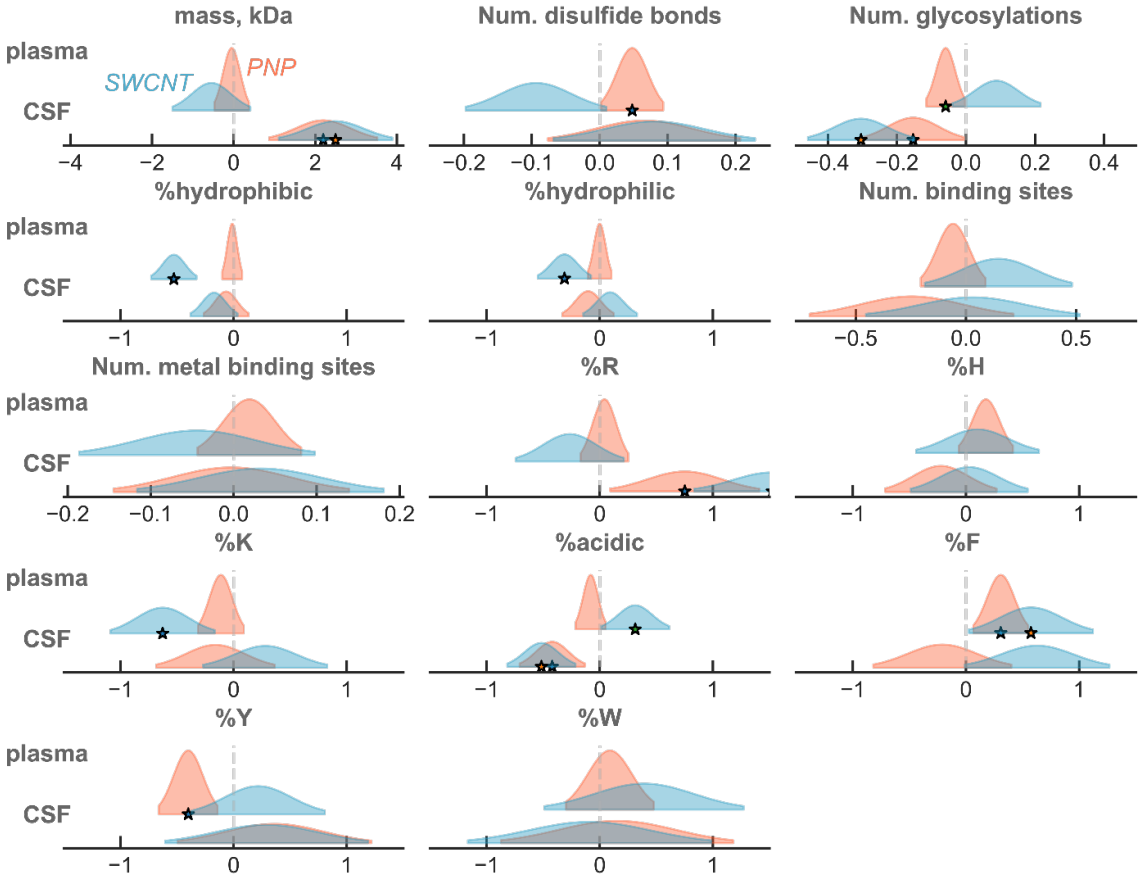


Figure S10. Distribution for microscale mean regression coefficients in each nanoparticle-biofluid pairing. Stars indicate false-discovery-rate adjusted p-values<0.1.

Table S3. Protein class regression results for each nanoparticle-biofluid pairing.

	PNPs in plasma			$(GT)_{15}$-SWCNTs in plasma		
	R-squared	**Adjusted R-squared**		**R-squared**	**Adjusted R-squared**	
	0.26	0.24		0.1	0.13	
	Parameter	**Standard Error**	**FDRC p-values**	**Parameter**	**Standard Error**	**FDRC p-values**
Intercept	0.9277	0.2270	0.0003	-6.1942	0.5380	0.0000
Sample 1	2.8743	0.3253	0.0000	4.0556	0.7711	0.0000
Sample 2	-0.0358	0.3146	0.9095	0.3037	0.7457	0.8412
Acute-phase response	-0.1393	0.4271	0.9095	0.0916	1.0124	0.9279
Blood coagulation	0.6394	0.3749	0.3267	0.7639	0.8887	0.7163
Cell adhesion / Signal transduction	-0.0981	0.4133	0.9095	-0.3003	0.9797	0.8412
Complement activation	0.0854	0.3086	0.9095	1.3564	0.7316	0.2372
Immune response	-0.0727	0.3620	0.9095	-0.8149	0.8580	0.7163
Lipid binding / transport	0.1467	0.4271	0.9095	1.5017	1.0124	0.3823
Regulation of biological processes	0.2203	0.2985	0.9095	-0.3749	0.7075	0.8412
Transport	-0.5563	0.5047	0.7460	-0.3584	1.1965	0.8412

	PNPs in CSF			$(GT)_{15}$-SWCNTs in CSF		
	R-squared	**Adjusted R-squared**		**R-squared**	**Adjusted R-squared**	
	0.32	0.28		0.35	0.31	
	Parameter	**Standard Error**	**FDRC p-values**	**Parameter**	**Standard Error**	**FDRC p-values**
Intercept	-2.0347	0.7108	0.0131	-3.1619	0.7551	0.0001
Sample 1	0.9624	0.8879	0.3850	-0.0962	0.9431	0.9939
Sample 2	0.8264	0.8915	0.3909	-0.0073	0.9470	0.9939
Acute-phase response	-2.4143	1.1830	0.0943	-6.5197	1.2566	0.0000
Blood coagulation	4.7294	1.2588	0.0009	4.4095	1.3371	0.0026
Cell adhesion / Signal transduction	1.9484	1.6552	0.3785	4.8610	1.7583	0.0117
Complement activation	1.3603	0.9806	0.3067	2.1252	1.0417	0.0675
Immune response	-5.1599	1.0141	0.0000	-4.5191	1.0772	0.0001
Lipid binding / transport	-0.3967	1.1830	0.7378	1.6716	1.2566	0.2548
Regulation of biological processes	0.7459	0.7590	0.3909	-0.5229	0.8062	0.6325
Transport	-5.1597	0.9806	0.0000	-5.3045	1.0417	0.0000

Table S4. Microscale regression results for each nanoparticle-biofluid pairing.

	PNPs in plasma			$(GT)_{15}$-SWCNTs in plasma		
	R-squared	**Adjusted R-squared**		**R-squared**	**Adjusted R-squared**	
	0.32	0.29		0.27	0.23	
	Parameter	**Standard Error**	**FDRC p-values**	**Parameter**	**Standard Error**	**FDRC p-values**
Intercept	2.7690	3.6300	0.7585	19.6116	8.2339	0.0607
Sample 1	2.8759	0.3147	0.0000	4.0662	0.7137	0.0000
Sample 2	-0.0427	0.3036	0.9436	0.3186	0.6886	0.6842
Mass	-0.0475	0.2109	0.9314	-0.5446	0.4785	0.4153
% hydrophobic residues (nonaromatic)	-0.0126	0.0442	0.9314	-0.5286	0.1003	0.0000
% hydrophilic residues	-0.0008	0.0520	0.9877	-0.3122	0.1179	0.0363
% arginine	0.0409	0.1051	0.9121	-0.2643	0.2385	0.4153
% histidine	0.1762	0.1197	0.4030	0.1028	0.2716	0.7053
% lysine	-0.1128	0.1015	0.5678	-0.6276	0.2302	0.0363
% acidic residues	-0.0806	0.0662	0.5447	0.3131	0.1502	0.0921
% phenylalanine	0.3077	0.1209	0.0648	0.5747	0.2743	0.0921
% tyrosine	-0.4013	0.1297	0.0184	0.2165	0.2942	0.5615
% tryptophan	0.0893	0.1940	0.9121	0.3912	0.4400	0.4942
Number of disulfide bonds	0.0477	0.0229	0.1288	-0.0938	0.0519	0.1527
Number of glycosylated sites	-0.0588	0.0280	0.1288	0.0885	0.0636	0.3124
Number of ligand binding sites	-0.0584	0.0736	0.7585	0.1473	0.1668	0.4942
Number of metal binding sites	0.0190	0.0313	0.8425	-0.0442	0.0710	0.6052

	PNPs in CSF			$(GT)_{15}$-SWCNTs in CSF		
	R-squared	**Adjusted R-squared**		**R-squared**	**Adjusted R-squared**	
	0.35	0.29		0.4	0.34	
	Parameter	**Standard Error**	**FDRC p-values**	**Parameter**	**Standard Error**	**FDRC p-values**
Intercept	-18.8198	9.7672	0.1897	-32.4292	10.2373	0.0063
Sample 1	0.7932	0.8845	0.5921	-0.4235	0.9271	0.8692
Sample 2	0.7215	0.8883	0.5921	-0.2622	0.9311	0.9097
Mass	2.1956	0.6640	0.0199	2.4973	0.6960	0.0025
% hydrophobic residues (nonaromatic)	-0.0670	0.0987	0.6048	-0.1726	0.1034	0.2361
% hydrophilic residues	-0.1034	0.1133	0.5921	0.0913	0.1187	0.7529
% arginine	0.7511	0.3295	0.1358	1.5253	0.3453	0.0003
% histidine	-0.2217	0.2464	0.5921	0.0293	0.2583	0.9097
% lysine	-0.1607	0.2616	0.6118	0.2798	0.2742	0.5836
% acidic residues	-0.4209	0.1441	0.0341	-0.5162	0.1511	0.0034
% phenylalanine	-0.2088	0.3052	0.6048	0.6294	0.3199	0.1441
% tyrosine	0.3624	0.4282	0.5921	0.2910	0.4488	0.8001
% tryptophan	0.1523	0.5128	0.8149	-0.0914	0.5375	0.9097
Number of disulfide bonds	0.0653	0.0708	0.5921	0.0801	0.0742	0.5836

Number of glycosylated sites	-0.1529	0.0733	0.1644	-0.3039	0.0769	0.0010
Number of ligand binding sites	-0.2462	0.2313	0.5921	0.0309	0.2424	0.9097
Number of metal binding sites	-0.0028	0.0709	0.9685	0.0323	0.0744	0.8692

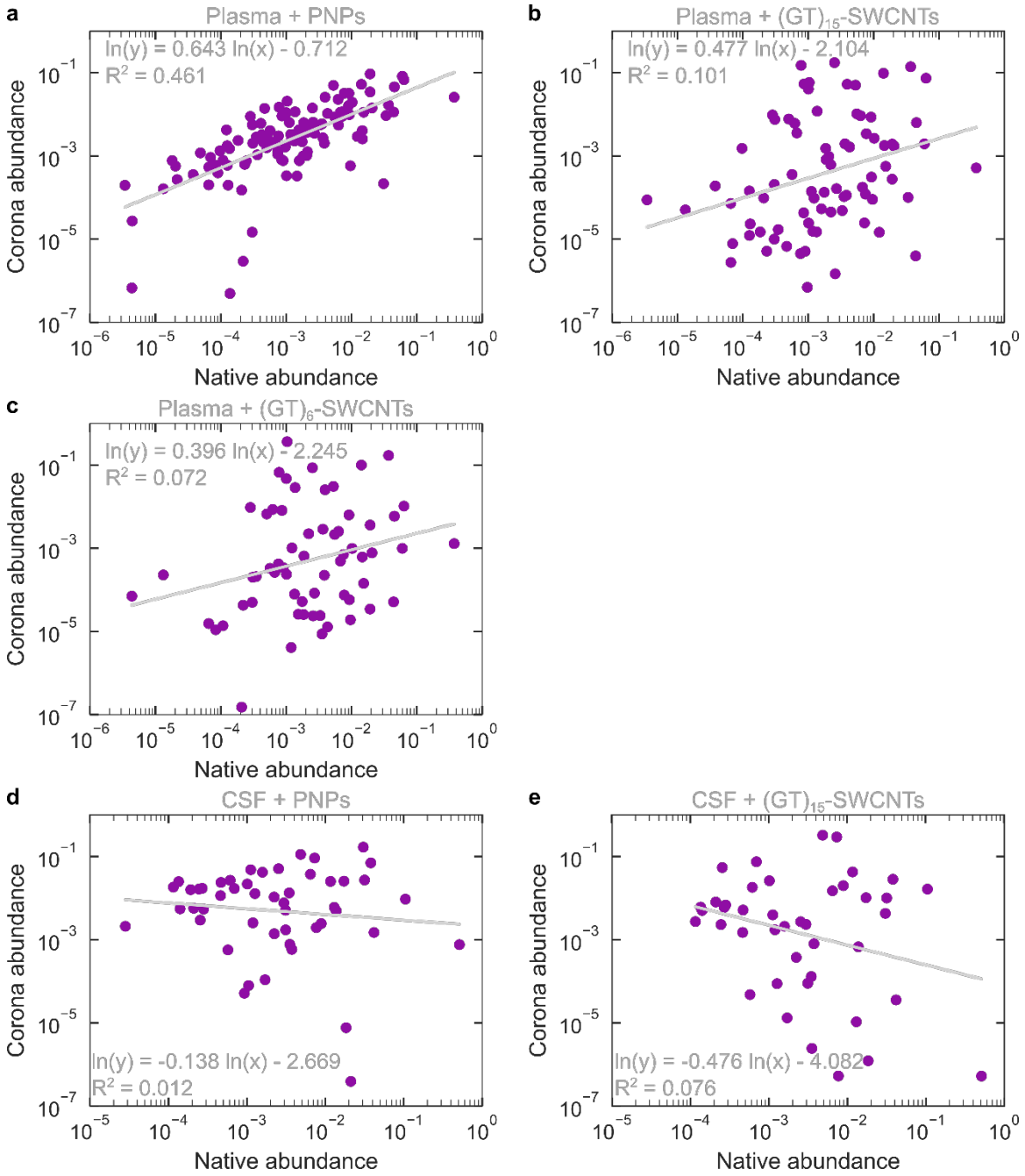


Figure S11. Scaling of protein abundance in corona vs. in native biofluid. Protein mole fraction of plasma proteins in corona of **(a)** PNPs, **(b)** $(GT)_{15}$-SWCNTs, and **(c)** $(GT)_6$-SWCNTs, vs. protein mole fraction of plasma proteins in native biofluid. Corona abundance scaling is approximately linear for plasma proteins on PNPs (R^2 = 0.461) vs. highly scattered for $(GT)_{15}$-SWCNTs (R^2 = 0.101) and $(GT)_6$-SWCNTs (R^2 = 0.072). Protein mole fraction of CSF proteins in corona of **(d)** PNPs and **(e)** $(GT)_{15}$-SWCNTs vs. protein mole fraction of CSF proteins in native biofluid. Corona abundance displays a weak negative correlation with native abundance for CSF proteins on both PNPs (R^2 = 0.012) and $(GT)_{15}$-SWCNTs (R^2 = 0.076). All mole fractions are on a solvent-free basis. Note that proteins with zero corona abundance are excluded from the analysis for clarity, but the same conclusions hold when included.

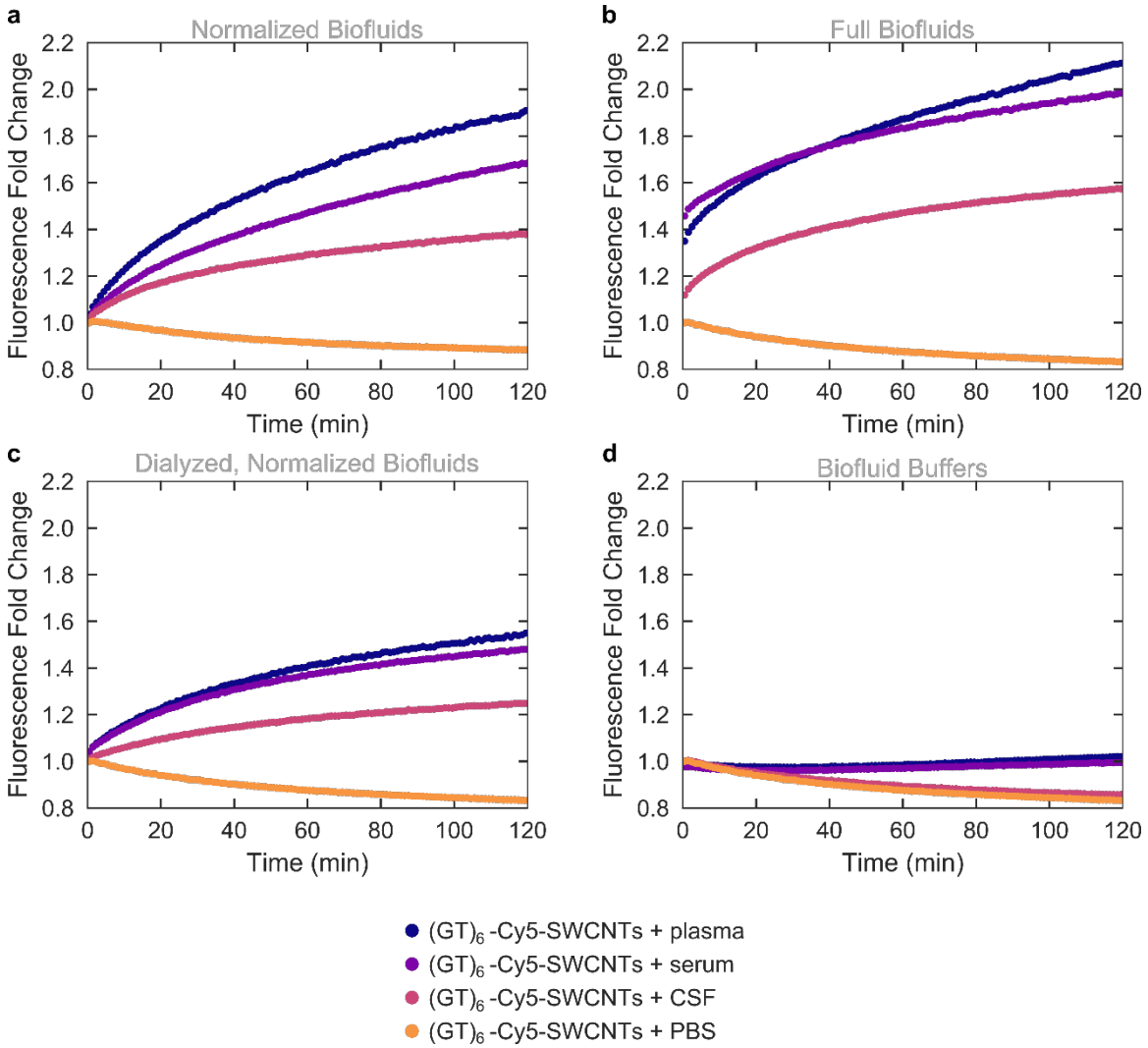


Figure S12. Protein corona dynamics assessed for binding of proteins in varying biofluids to $(GT)_6$-SWCNTs. A corona exchange assay is employed to track binding of proteins from the biofluids: blood plasma, blood serum, and cerebrospinal fluid (CSF) to $(GT)_6$-Cy5-SWCNTs (5 mg L^{-1} final concentration). ssDNA desorption from SWCNTs is observed as increasing fluorescence of Cy5-labeled $(GT)_6$, and taken as a proxy for protein binding.[7] **(a)** Biofluids normalized to 40 mg L^{-1} final protein concentration. **(b)** Biofluids injected as full, as-received solutions. **(c)** Biofluids dialyzed, with > 3 kDa portion re-suspended in 1X PBS and normalized to 40 mg L^{-1} final protein concentration. **(d)** Biofluids dialyzed, with < 3 kDa portion considered the biofluid buffer. Data points are the average of experimental replicates (N = 3).

Varying incubation parameters

Protein corona composition was studied under varied incubation conditions in comparison to the reference state incubation (static, 0.1 M phosphate-buffered saline, and 25°C incubation). Varying incubation conditions include dynamic (on orbital shaker at medium speed; to test corona stability), no salt (water; to test ionic effects), and high temperature (37°C; to test entropic contributions). **Figure 4** summarizes data from 2D PAGE gels, with experimental replicates of 6 for plasma alone, 10 for reference, 4 for dynamic, 6 for no salt, and 3 for high temperature. Note that the numerous replicates for the reference state were due to the frequent repetition of this condition in comparison to the varied conditions.

Isothermal titration calorimetry (ITC) methods

ITC measurements were performed with a NanoITC (TA Instruments). Prior to each experiment, samples and buffer were degassed for 10 min and the reference cell was filled with fresh Milli-Q water.

Equilibration time was set to 1 h before the experiment start and the initial and final baselines were collected for 300 s. For each experiment, 1.2 g L^{-1} protein in 0.1 M PBS was titrated from the syringe (250 µL total volume) into 0.1 g L^{-1} $(GT)_{15}$-SWCNTs in 0.1 M PBS in the cell (1 mL total volume) under constant stirring (250 rpm) at 25 °C. 10 µL of protein titrant was injected into the nanoparticle solution in the cell every 7 min, with a total of 24 injections. By standard practice, every run was initiated with a 5 µL injection to ensure no artifacts due to bubbles and was removed from analysis. All protein-nanoparticle binding experiments were accompanied by three heat-of-dilution control experiments: (1) protein injected into buffer, (2) buffer injected into nanoparticles, and (3) buffer injected into buffer (where buffer is 0.1 M PBS). Heat of binding of protein to nanoparticles was then calculated as: (heat from titration of protein into nanoparticles) – (1) – (2) + (3). Data processing was completed with NanoAnalyze software (TA Instruments). Baseline correction was done using the auto-fit routine. An independent binding model was applied to fit the fibrinogen data set, suitable to model weak nonspecific interactions such as those present in the system under study,[46] and a blank (constant) model was applied to fit the albumin data set.

Protein and nanoparticle concentrations and ITC setup parameters were varied in attempt of obtaining binding curves for both proteins to $(GT)_{15}$-SWCNTs. However, for albumin this was not possible within the ITC instrument's operational range, therefore albumin was concluded to not bind to $(GT)_{15}$-SWCNTs.

Extended discussion on ITC

ITC was employed to extract relative binding parameters of protein-nanoparticle association. ITC was performed at constant pressure such that the heat absorbed or released is equivalent to the change in enthalpy (ΔH^o) upon binding. The binding curve can be fit to determine the equilibrium dissociation constant (K_d) and molar binding stoichiometry (n). This enables subsequent calculation of changes in standard state Gibbs free energy (ΔG^o) and entropy (ΔS^o) as follows:

$$\Delta G^o = RT\ln K_d = \Delta H^o - T\Delta S^o$$

where R is the ideal gas constant and T is temperature. The optimized run parameters to measure heats of binding for this system require relatively high protein and nanoparticle concentrations: for each run, 10 µL of 1.2 g L^{-1} protein was added for each of 24 injections from the syringe into 1 mL of 0.1 g L^{-1} $(GT)_{15}$-SWCNTs in the cell. At these concentrations, addition of fibrinogen causes visible sample aggregation, presumably due to polymer bridging interactions of proteins adsorbed on one nanoparticle interacting with another nanoparticle. One of the key assumptions of ITC is that the system is equilibrated during each titration step. Yet, aggregation is a kinetically controlled, non-equilibrium process. As the key assumption is not held, these binding values are actually the convolution of protein binding to individual SWCNTs, fibrinogen binding to aggregated SWCNTs, and SWCNTs aggregating. We can compensate for this limitation in data processing by applying the Lumry-Eyring model,[47] in which an equilibrium reaction is coupled to a self-association reaction (i.e. aggregation), and the heats measured are separated out accordingly. This encompasses subtracting out baseline aggregation heats and arriving at an apparent binding heat. Therefore, the thermodynamic parameters are reported with consideration of these higher order processes taking place simultaneously. A further note is that baseline drift/shift were observed during these ITC experiments involving $(GT)_{15}$-SWCNTs. These changes in baseline often indicate slow non-equilibrium processes in action, further confirming the presence of aggregation. In conclusion, ITC is not a suitable methodology to study nanoparticle-protein corona formation for all systems, and these limitations must be considered during experimental design and reporting of results.

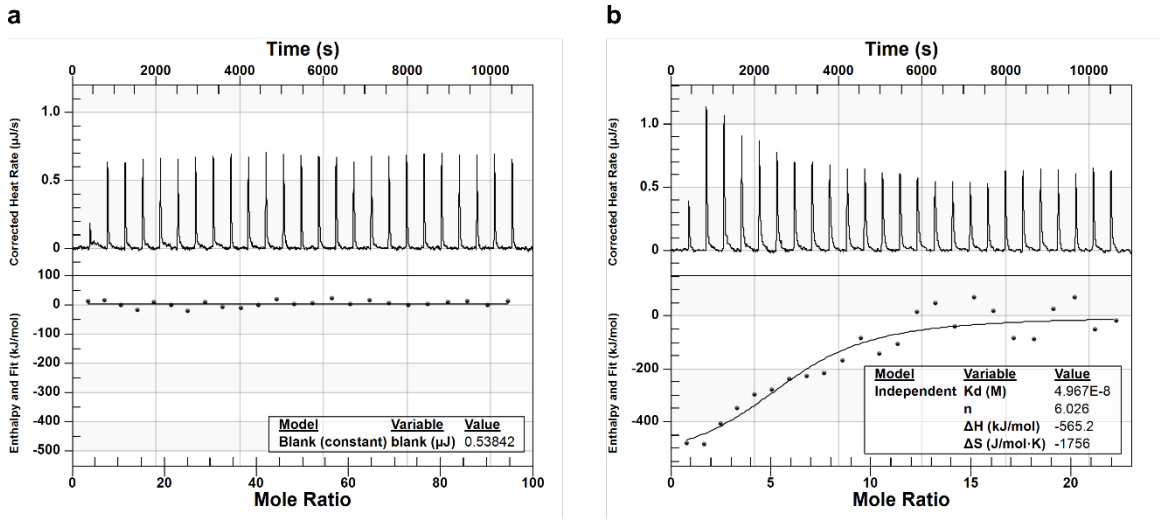


Figure S13. Protein corona thermodynamics assessed with ITC for binding of key proteins to $(GT)_{15}$-SWCNTs. Isothermal titration calorimetry (ITC) is employed to determine binding thermodynamics of **(a)** albumin and **(b)** fibrinogen to $(GT)_{15}$-SWCNTs. Albumin does not bind to $(GT)_{15}$-SWCNTs within experimentally accessible limits of this instrument, whereas fibrinogen does, in agreement with the corona compositional analyses from proteomic mass spectrometry and gel electrophoresis.

Corona exchange assay

Corona dynamic studies were completed as described previously.[7] Briefly, the same suspension protocol was employed for preparation of fluorophore-labeled ssDNA-SWCNT complexes, using ssDNA-Cy5 (3' Cy5-labeled custom ssDNA oligos with HPLC purification, Integrated DNA Technologies, Inc.) in place of unlabeled ssDNA. Lyophilized proteins were purchased (see details in **Table S6**) and reconstituted by adding 5 mg to 1 mL of 0.1 M PBS, tilting to dissolve for 15 min, filtering with 0.45 µm syringe filter (cellulose acetate membrane, VWR International), and quantifying with the Qubit Protein Assay (Thermo Fisher Scientific). Because of variation in amine-labeling of proteins, fluorescently labeled ssDNA was solely tracked, and the displacement of ssDNA from the SWCNT surface was taken as a proxy for protein adsorption. Equal volumes of 10 mg L^{-1} $(GT)_{15}$- or $(GT)_6$-Cy5-SWCNTs and 160 mg L^{-1} protein were added to a 96-well PCR plate (Bio-Rad) to a total volume of 50 µL. The plate was sealed with an optically transparent adhesive seal (Bio-Rad) and spun down on a benchtop centrifuge. Fluorescence time series readings were taken in a Bio-Rad CFX96 Real Time qPCR System, scanning the Cy5 channel every 2 min at 22.5℃. Fluorescence time series were analyzed without default background correction. Fluorescence values were converted to mass concentration using linear standard curves for ssDNA-Cy5. Note that in the case of the control, ssDNA adsorption to the SWCNT is observed, in line with previous studies.[7]

Structure studies by small-angle x-ray scattering (SAXS)

SAXS data was collected at SIBYLS beamline (bl12.3.1) at the Advanced Light Source of Lawrence Berkeley National Laboratory, Berkeley, California.[48] X-ray wavelength was set at λ = 0.1127 nm and the sample-to-detector distance was 2.1 m, resulting in scattering vector (q) ranging from 0.1–4 nm^{-1}. The scattering vector is defined as $q = 4\pi\sin\theta/\lambda$, with scattering angle 2θ. Data was collected using a Dectris PILATUS3X 2M detector at 20°C and processed as described previously.[49]

Immediately prior to data collection, 30 µL of each sample was added to 96-well plates kept at 10°C and transferred to the sampling position via a Tecan Evo liquid handling robot with modified pipetting needles acting as sample cells as described previously.[50] Samples were exposed to X-ray synchrotron radiation for 30 s at a 0.5 s frame rate for a total of 60 images. Each collected image was circularly integrated and normalized for beam intensity to generate a one-dimensional scattering profile by beamline-specific software. Buffer subtraction was performed for the one-dimensional scattering profile of each sample

using each of two 0.1 M PBS buffer wells to ensure the subtraction process was not subject to instrument variations. Scattering profiles over the 30 s exposure were sequentially averaged together to eliminate any potential radiation damage effects. Averaging was performed with web-based software FrameSlice (sibyls.als.lbl.gov/ran).

Extended experimental and modeling details and discussion on SAXS

Experimental SAXS profiles were collected for 0.5 g L^{-1} $(GT)_{15}$-SWCNTs with and without albumin or fibrinogen, each at 0.5 g L^{-1} final concentrations (**Figure 5c**). The linear combination of $(GT)_{15}$-SWCNTs and albumin standard curves produced a SAXS profile identical to the mixed sample of $(GT)_{15}$-SWCNTs with albumin, suggesting no interaction between the species. Dissimilarly, no calculated linear combination of the $(GT)_{15}$-SWCNTs and fibrinogen standard curves could be produced to fit the SAXS profiles of the mixed sample, indicating formation of unique form factors and thus complexation. Additionally, a clear concentration dependence is observed with an increase in the ratio of fibrinogen to $(GT)_{15}$-SWCNTs by two-fold, while albumin shows no additional binding at elevated concentrations (**Figure S14a-b**). Control SAXS profiles of albumin, fibrinogen, and $(GT)_{15}$-SWCNTs alone were collected at identical concentrations to those of the mixing experiments (**Figure S14c**). Data was collected at elevated concentrations (0.5 g L^{-1} both protein and $(GT)_{15}$-SWCNTs) to enhance SAXS signal, however, a concentration series was also performed for $(GT)_{15}$-SWCNTs to ensure that the scattering profiles do not deviate under more relevant nanoparticle conditions down to 0.01 g L^{-1} (**Figure S14d**).[51]

All $(GT)_{15}$-SWCNT samples with and without proteins were determined to be intrinsically disordered and experimental SAXS profiles were accordingly fit using mass fractal geometries. These fits were complemented by calculating power-law dependencies from the Porod region, and were both calculated using the SasView software package (www.sasview.org). Scattering intensity as a function of scattering vector $I(q)$ calculations for the mass fractal modeling (**Figure 5**) was done as follows:[52]

$$I(q) = \text{scale} * P(q)S(q) + \text{background}$$

$$P(q) = F(qR)^2$$

$$F(\text{x}) = \frac{3[\sin(x) - x\cos(x)]}{x^3}$$

$$S(q) = \frac{\Gamma(Dm-1)\zeta^{Dm-1}}{[1+(q\zeta)2]^{\frac{Dm-1}{2}}} \frac{\sin[(Dm-1)\tan^{-1}(q\zeta)]}{q}$$

$$\text{scale} = \text{scale factor} * N\left(\frac{4}{3}\pi R^3\right)^2 (\rho_{\text{particle}} - \rho_{\text{solvent}})^2$$

where R is the radius of the building block, Dm is the mass fractal dimension, ζ is the cut-off length, N is number of scatters, $\rho_{solvent}$ is the scattering length density of the solvent, and $\rho_{particle}$ is the scattering length density of particles. Dm relates the mass (m) to the radius as $m \sim R^{Dm}$ and is analogous to $I(q) \sim q^{-p}$ from the power-law calculations (with power-law exponent p), where $Dm = p$ when $q\zeta >> 1$.

The power-law dependencies were determined by fitting the experimental SAXS profiles (**Figure S14e**), where $0.3 \leq q \leq 1$ nm^{-1} with the following:[53]

$$I(q) = \text{scale} * q^{-p} + \text{background}$$

These power-law dependencies (fits listed in **Figure S14e**) recapitulate the calculated Dm values from the mass fractal model fits.

Three main values are derived from these mass fractal and power-law calculations: (i) radius R (nm), (ii) fractal dimension Dm, and (iii) cutoff length ζ (nm) (**Table S5**).[53–55] The radius R in the mass fractal

analysis is traditionally defined as the radius of the uniform sphere used to cover the fractal. The fractal dimension *Dm* and analogous power-law exponent *p* estimate the overall bulk geometries of the mass fractals, where the integer values of these variables represent the three dimensions in Euclidean space. Thus, *Dm* or *p* = 1, 2, or 3 represent rod, disk, or sphere geometries, respectively. The cutoff length ζ defines the maximum distance between any two points of the mass fractal.

As another control, carboxylic acid functionalized SWCNTs (COOH-SWCNTs) were also examined via power-law scattering obtaining $p \sim 3.3$ (**Figure S14e**). This fit suggests that without ssDNA functionalization, COOH-SWCNTs form roughly spherical aggregates better modeled as a uniform density as opposed to a polymeric mass fractal. Thus, it may be inferred that ssDNA provides some semblance of order to the fine molecular structure of the system and should be the subject of further investigation.

The effect of aggregation on the scattering vector at very small angles ($q < 1$ nm^{-1}) precluded the use of the Guinier approximation and subsequent calculated metrics such as the radius of gyration (*Rg*), and the scattering intensity at q=0, $I(0)$, which is proportional to the molecular weight.[56] Additionally, while mathematically possible to calculate a pair-distribution function, *P(r)*, from the indirect Fourier transformation, the level of aggregation leads to non-zero values for *r*=*Dmax*.[57] Accordingly, we fit the whole SAXS profile to a specific mass fractal model, providing an estimate for the average cutoff length ζ, superseding the need to calculate the analogous *Dmax* value which we determined to be less accurate.

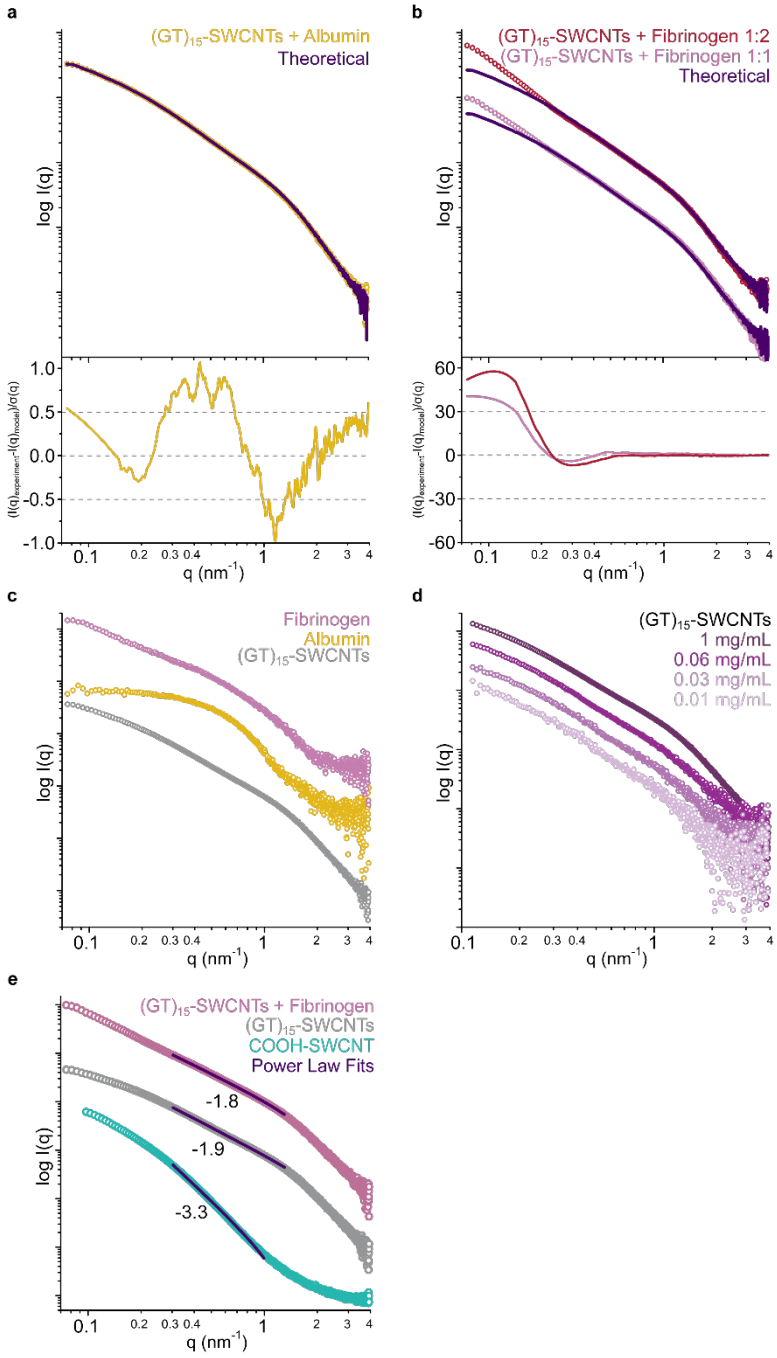


Figure S14. Protein corona structure assessed with SAXS for binding of key proteins to $(GT)_{15}$-SWCNTs. The linear combination of respective standard curves from panel c in purple and fit-residuals below, fit against the curves produced by the potential complexes of $(GT)_{15}$-SWCNTs with **(a)** albumin or **(b)** fibrinogen, at two different ratios of $(GT)_{15}$-SWCNTs to fibrinogen (1:1 is 0.5 g L^{-1} final concentrations of $(GT)_{15}$-SWCNTs and fibrinogen; 1:2 is 0.25 g L^{-1} $(GT)_{15}$-SWCNTs and 0.5 g L^{-1} fibrinogen). **(c)** Experimental SAXS profiles for standards of albumin, fibrinogen, and $(GT)_{15}$-SWCNTs alone, at identical concentrations to the mixing experiments (all 0.5 g L^{-1}). **(d)** SAXS profiles for concentration series of $(GT)_{15}$-SWCNTs alone, 0.01 – 1 g L^{-1}. **(e)** SAXS profiles fit to show power law dependencies in the Porod regions, including the COOH-SWCNT control without surface-adsorbed ssDNA.

Table S5. SAXS mass fractal modeling parameters.

Sample	Radius (nm)	Fractal Dimension (*Dm*)	Cutoff Length (nm)
$(GT)_{15}$-SWCNTs + Fibrinogen	1.05 ± 0.003	1.77	103.34 ± 9.70
$(GT)_{15}$-SWCNTs + Albumin	1.05 ± 0.003	1.90	10.60 ± 0.05
$(GT)_{15}$-SWCNTs	1.01 ± 0.002	1.89	10.91 ± 0.04

Transmission electron microscopy (TEM) methods

Holey carbon-coated grids (EMS Electron Microscopy Science) were surface-treated by glow discharge to make the support hydrophilic. Samples of $(GT)_{15}$-SWCNTs with fibrinogen or plasma were negatively stained with 1% uranyl acetate solution. For the $(GT)_{15}$-SWCNTs alone sample, no negative staining was done. 5 µl of 10 mg L^{-1} solution was drop-cast onto the grid. FEI ThemIS 60-300 STEM/TEM (National Center of Electron Microscopy, Molecular Foundry) with acceleration voltage of 60kV was used to acquire TEM images by video recording. A low acceleration voltage was chosen to minimize sample damage and increase sample contrast.

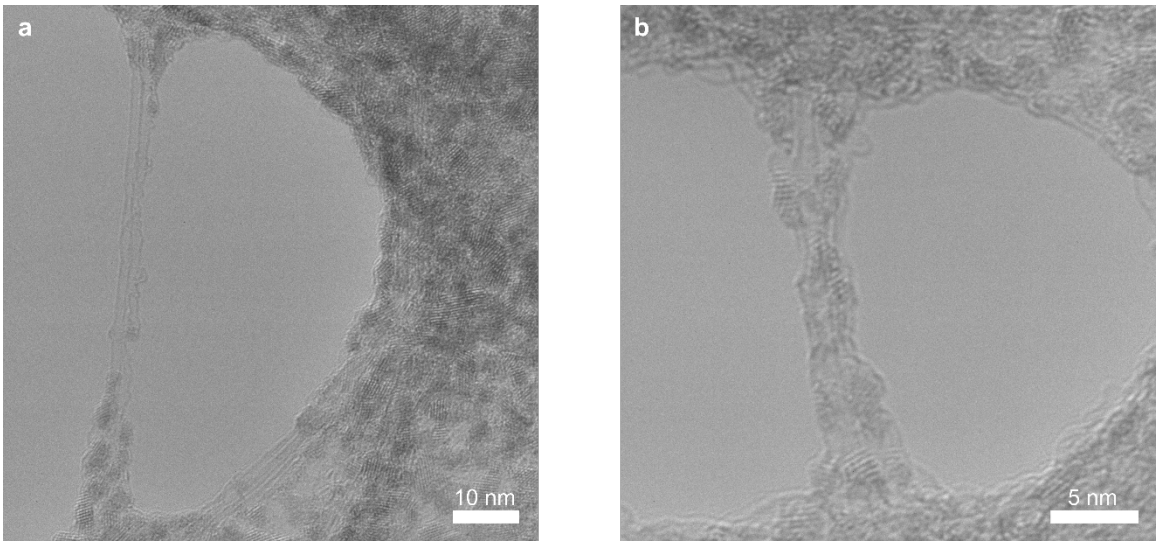


Figure S15. Protein corona morphology visualized by TEM for adsorption of plasma proteins to $(GT)_{15}$-SWCNTs. Transmission electron microscopy (TEM) of **(a)** plasma protein corona and **(b)** fibrinogen corona on $(GT)_{15}$-SWCNTs.

Table S6. Purchased biofluid and protein specifications.

Protein	Manufacturer	Lot #	Source	Form
Blood plasma	Innovative Research Inc.	#23791	Pooled normal human plasma	Biofluid
Cerebrospinal fluid	Lee Biosolutions	#07C5126	Pooled normal human CSF, from remnant lumbar puncture	Biofluid
Albumin	Sigma-Aldrich	#SLBZ2785	Human plasma	Lyophilized
Alpha-2-HS glycoprotein	Biovision Inc.	#4C08L75480	Human plasma	Lyophilized
Apolipoprotein A-I	Alfa Aesar	#927J17A	Human plasma	1 g L^{-1} in 10mM ammonium bicarbonate buffer, pH 7.4
Clusterin	R&D Systems	NEV1519031	Mouse myeloma cell line, NS0-derived human; Asp23-Arg227 (beta) & Ser228-Glu449 (alpha) with a C-terminal 6-His tag	Lyophilized
Complement C3	Mybiosource Inc.	#N30/20170	Human plasma	5 g L^{-1}
Fibrinogen	Millipore Sigma	#3169957	Human plasma	Lyophilized
Immunoglobulin G	Lee Biosolutions	#06B2334	Human plasma	Lyophilized

References

[1] J. Zhang, M. P. Landry, P. W. Barone, J.-H. Kim, S. Lin, Z. W. Ulissi, D. Lin, B. Mu, A. A. Boghossian, A. J. Hilmer, A. Rwei, A. C. Hinckley, S. Kruss, M. A. Shandell, N. Nair, S. Blake, F. Şen, S. Şen, R. G. Croy, D. Li, K. Yum, J.-H. Ahn, H. Jin, D. A. Heller, J. M. Essigmann, D. Blankschtein, M. S. Strano, *Nature Nanotechnology* **2013**, *8*, 959–968.

[2] R. Nißler, F. A. Mann, P. Chaturvedi, J. Horlebein, D. Meyer, L. Vuković, S. Kruss, *J. Phys. Chem. C* **2019**, *123*, 4837–4847.

[3] M. M. Safaee, M. Gravely, C. Rocchio, M. Simmeth, D. Roxbury, *ACS Appl. Mater. Interfaces* **2019**, *11*, 2225–2233.

[4] A. A. Alizadehmojarad, X. Zhou, A. G. Beyene, K. E. Chacon, Y. Sung, R. L. Pinals, M. P. Landry, L. Vuković, *Advanced Materials Interfaces* **n.d.**, *n/a*, 2000353.

[5] M. J. O'Connell, S. M. Bachilo, C. B. Huffman, V. C. Moore, M. S. Strano, E. H. Haroz, K. L. Rialon, P. J. Boul, W. H. Noon, C. Kittrell, J. Ma, R. H. Hauge, R. B. Weisman, R. E. Smalley, *Science* **2002**, *297*, 593–596.

[6] A. G. Beyene, A. A. Alizadehmojarad, G. Dorlhiac, N. Goh, A. M. Streets, P. Král, L. Vuković, M. P. Landry, *Nano Lett.* **2018**, *18*, 6995–7003.

[7] R. L. Pinals, D. Yang, A. Lui, W. Cao, M. P. Landry, *J. Am. Chem. Soc.* **2019**, DOI 10.1021/jacs.9b09617.

[8] E. S. Jeng, A. E. Moll, A. C. Roy, J. B. Gastala, M. S. Strano, *Nano Lett.* **2006**, *6*, 371–375.

[9] F. Schöppler, C. Mann, T. C. Hain, F. M. Neubauer, G. Privitera, F. Bonaccorso, D. Chu, A. C. Ferrari, T. Hertel, *J. Phys. Chem. C* **2011**, *115*, 14682–14686.

[10] F. K. Brunecker, F. Schöppler, T. Hertel, *J. Phys. Chem. C* **2016**, *120*, 10094–10103.

[11] M. Lundqvist, J. Stigler, G. Elia, I. Lynch, T. Cedervall, K. A. Dawson, *Proceedings of the National Academy of Sciences* **2008**, *105*, 14265–14270.

[12] M. P. Monopoli, D. Walczyk, A. Campbell, G. Elia, I. Lynch, F. Baldelli Bombelli, K. A. Dawson, *J. Am. Chem. Soc.* **2011**, *133*, 2525–2534.

[13] J. R. Wiśniewski, A. Zougman, N. Nagaraj, M. Mann, *Nat Methods* **2009**, *6*, 359–362.

[14] J. B. Simonsen, R. Münter, *Angewandte Chemie International Edition* **2020**, *59*, 12584–12588.

[15] P. H. O'Farrell, *Journal of biological chemistry* **1975**, *250*, 4007–4021.

[16] U. K. Laemmli, *Nature* **1970**, *227*, 680–685.

[17] "World-2DPAGE Constellation: SWISS-2DPAGE," can be found under https://world-2dpage.expasy.org/swiss-2dpage/, **n.d.**

[18] P. V. Shliaha, N. J. Bond, L. Gatto, K. S. Lilley, *J. Proteome Res.* **2013**, *12*, 2323–2339.

[19] U. Distler, J. Kuharev, P. Navarro, Y. Levin, H. Schild, S. Tenzer, *Nature Methods* **2014**, *11*, 167–170.

[20] R. S. Plumb, K. A. Johnson, P. Rainville, B. W. Smith, I. D. Wilson, J. M. Castro-Perez, J. K. Nicholson, *Rapid Communications in Mass Spectrometry* **2006**, *20*, 1989–1994.

[21] K. A. Neilson, N. A. Ali, S. Muralidharan, M. Mirzaei, M. Mariani, G. Assadourian, A. Lee, S. C. van Sluyter, P. A. Haynes, *PROTEOMICS* **2011**, *11*, 535–553.

[22] S. Nahnsen, C. Bielow, K. Reinert, O. Kohlbacher, *Mol Cell Proteomics* **2012**, mcp.R112.025163.

[23] J. C. Silva, M. V. Gorenstein, G.-Z. Li, J. P. C. Vissers, S. J. Geromanos, *Molecular & Cellular Proteomics* **2006**, *5*, 144–156.

[24] N. L. Anderson, N. G. Anderson, *Molecular & Cellular Proteomics* **2002**, *1*, 845–867.

[25] Z. Liu, S. Tabakman, K. Welsher, H. Dai, *Nano Res.* **2009**, *2*, 85–120.

[26] G. Bisker, N. A. Bakh, M. A. Lee, J. Ahn, M. Park, E. B. O'Connell, N. M. Iverson, M. S. Strano, *ACS Sens.* **2018**, *3*, 367–377.

[27] J. D. Harvey, H. A. Baker, M. V. Ortiz, A. Kentsis, D. A. Heller, *ACS Sens.* **2019**, *4*, 1236–1244.

[28] D. Dutta, S. K. Sundaram, J. G. Teeguarden, B. J. Riley, L. S. Fifield, J. M. Jacobs, S. R. Addleman, G. A. Kaysen, B. M. Moudgil, T. J. Weber, *Toxicol Sci* **2007**, *100*, 303–315.

[29] B. D. Holt, K. N. Dahl, M. F. Islam, *Small* **2011**, *7*, 2348–2355.

[30] S. Schöttler, G. Becker, S. Winzen, T. Steinbach, K. Mohr, K. Landfester, V. Mailänder, F. R. Wurm, *Nature Nanotechnology* **2016**, *11*, 372–377.

[31] S. Tenzer, D. Docter, J. Kuharev, A. Musyanovych, V. Fetz, R. Hecht, F. Schlenk, D. Fischer, K. Kiouptsi, C. Reinhardt, K. Landfester, H. Schild, M. Maskos, S. K. Knauer, R. H. Stauber, *Nat Nanotechnol* **2013**, *8*, 772–781.

[32] C. Salvador-Morales, E. Flahaut, E. Sim, J. Sloan, M. L. H. Green, R. B. Sim, *Molecular Immunology* **2006**, *43*, 193–201.

[33] L. Digiacomo, F. Cardarelli, D. Pozzi, S. Palchetti, M. A. Digman, E. Gratton, A. L. Capriotti, M. Mahmoudi, G. Caracciolo, *Nanoscale* **2017**, *9*, 17254–17262.

[34] K. M. Pondman, M. Sobik, A. Nayak, A. G. Tsolaki, A. Jäkel, E. Flahaut, S. Hampel, B. Ten Haken, R. B. Sim, U. Kishore, *Nanomedicine* **2014**, *10*, 1287–1299.

[35] M. J. Rybak-Smith, R. B. Sim, *Advanced Drug Delivery Reviews* **2011**, *63*, 1031–1041.

[36] A. E. Nel, L. Mädler, D. Velegol, T. Xia, E. M. V. Hoek, P. Somasundaran, F. Klaessig, V. Castranova, M. Thompson, *Nature Materials* **2009**, *8*, 543–557.

[37] Z. J. Deng, M. Liang, M. Monteiro, I. Toth, R. F. Minchin, *Nature Nanotechnology* **2011**, *6*, 39–44.

[38] T. A. Horbett, *Journal of Biomedical Materials Research Part A* **2018**, *0*, DOI 10.1002/jbm.a.36460.

[39] P. D. Thomas, M. J. Campbell, A. Kejariwal, H. Mi, B. Karlak, R. Daverman, K. Diemer, A. Muruganujan, A. Narechania, *Genome Res.* **2003**, *13*, 2129–2141.

[40] W. H. Greene, *Econometric Analysis*, Pearson Education India, **2003**.

[41] *Nucleic Acids Res* **2019**, *47*, D506–D515.

[42] Y. Benjamini, Y. Hochberg, *Journal of the Royal Statistical Society: Series B (Methodological)* **1995**, *57*, 289–300.

[43] S. Seabold, J. Perktold, **2010**, 6.

[44] M. R. Findlay, D. N. Freitas, M. Mobed-Miremadi, K. E. Wheeler, *Environmental Science: Nano* **2018**, *5*, 64–71.

[45] E. A. Guggenheim, *Thermodynamics - An Advanced Treatment for Chemists and Physicists (7th Edition)*, **1985**.

[46] N. Gal, M. Schroffenegger, E. Reimhult, *J. Phys. Chem. B* **2018**, *122*, 5820–5834.

[47] A. Schön, B. R. Clarkson, M. Jaime, E. Freire, *Proteins* **2017**, *85*, 2009–2016.

[48] S. Classen, G. L. Hura, J. M. Holton, R. P. Rambo, I. Rodic, P. J. Mcguire, K. Dyer, M. Hammel, G. Meigs, K. A. Frankel, J. A. Tainer, *Journal of Applied Crystallography* **2013**, *46*, 1–13.

[49] K. N. Dyer, M. Hammel, R. P. Rambo, S. E. Tsutakawa, I. Rodic, S. Classen, J. A. Tainer, G. L. Hura, *High-Throughput SAXS for the Characterization of Biomolecules in Solution : A Practical Approach*, Springer International Publishing, **2014**.

[50] G. L. Hura, A. L. Menon, M. Hammel, R. P. Rambo, F. L. Poole II, S. E. Tsutakawa, F. E. Jenney Jr, S. Classen, K. A. Frankel, R. C. Hopkins, S. Yang, J. W. Scott, B. D. Dillard, M. W. W. Adams, J. A. Tainer, *Nature Methods* **2009**, *6*, 606–612.

[51] A. G. Beyene, K. Delevich, J. T. D. Bonis-O'Donnell, D. J. Piekarski, W. C. Lin, A. W. Thomas, S. J. Yang, P. Kosillo, D. Yang, G. S. Prounis, L. Wilbrecht, M. P. Landry, *Science Advances* **2019**, *5*, eaaw3108.

[52] D. F. R. Mildner, P. L. Hall, *J. Phys. D: Appl. Phys.* **1986**, *19*, 1535–1545.

[53] D. W. Schaefer, *Science* **1989**, *243*, 1023–1027.

[54] D. F. R. Mildner, P. L. Hall, *J. Phys. D: Appl. Phys.* **1986**, *19*, 1535–1545.

[55] P. W. Schmidt, *J. Appl. Cryst.* **1991**, *24*, 414–435.

[56] C. D. Putnam, M. Hammel, G. L. Hura, J. A. Tainer, *Quarterly Reviews of Biophysics* **2007**, *40*, 191–285.

[57] A. G. Kikhney, D. I. Svergun, *FEBS Letters* **2015**, *589*, 2570–2577.