Supporting Information

Array-based Sensing of Metastatic Cells and Tissues Using Nanoparticle-Fluorescent Protein Conjugates

Subinoy Rana,[†,§] Arvind K. Singla,[‡,§] Avinash Bajaj,[†,#] S. Gokhan Elci,[†] Oscar R. Miranda,[†] Rubul Mout,[†] Bo Yan,[†] Frank R. Jirik,[‡] and Vincent M. Rotello[†,*]

[†]Department of Chemistry, University of Massachusetts, Amherst, Massachusetts, USA.
[‡]Department of Biochemistry and Molecular Biology, The McCaig Institute for Bone and Joint Health, University of Calgary, Calgary, Alberta, Canada.
[#]Regional Centre for Biotechnology, Gurgaon, Haryana, India.

[§]These authors contributed equally.

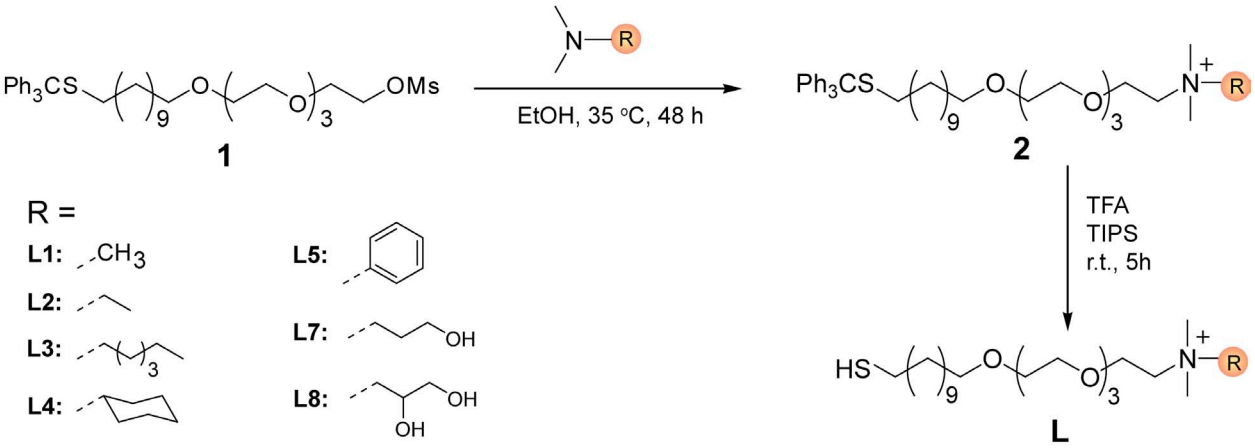
* Correspondence should be addressed to V.R. (rotello@chem.umass.edu)

Table of contents

1. Nanoparticle synthesis

(a) Ligand synthesis

Scheme 1:



General procedure: The starting compound for functionalization with various headgroups was the compound **1**, synthesized according to reported procedure.[1] Compound **2** bearing quaternary ammonium headgroups were synthesized by the reaction of compound **1** with corresponding substituted N,N-dimethylamines for 48 hours at 35 °C. The trityl protected thiol compound **2** was dissolved in dry dichloromethane (DCM) and an excess of trifluoroacetic acid (TFA, 20 equivalents) was added. The color of the solution turned into yellow immediately. Subsequently, triisopropylsilane (TIPS, 1.2 equivalents) was added to the reaction mixture. The reaction mixture was stirred for 5 h under Ar-atmosphere at room temperature. After evaporating the solvent, the pale yellow residue was further dried in high vacuum and characterized by NMR.

Compound **L1**: ^{1}H NMR (400MHz, $CDCl_3$, TMS): δ 3.95 (br, 2H, -CH_2N-), 3.70-3.58 (m, 14H, -CH_2O- + -OC$\underline{H_2}$-(CH_2N)-), 3.49 (t, 2H, -CH_2O-), 3.25 (s, 9H, -N(CH_3)$_3$-), 2.90 (s, 3H, -$CH_3SO^-_3$-), 2.52 (q, 2H, -CH_2S-), 1.64-1.51 (m, 4H, (SCH_2)C$\underline{H_2}$ + -C$\underline{H_2}$(CH_2O)-) , 1.36-1.22 (m, 15H, -SH + -CH_2-).

Compound **L2**: ^{1}H NMR (400MHz, $CDCl_3$, TMS): δ 3.95 (br, 2H, -CH_2N-), 3.70-3.60 (m, 14H, -CH_2O- + -OC$\underline{H_2}$-(CH_2N)-), 3.57 (q, 2H, -N(CH_3)$_2$-C$\underline{H_2}$-CH_3), 3.42 (t, 2H, -CH_2-C$\underline{H_2}$O-), 3.20

(s, 6H, -N(C<u>H</u>$_3$)$_2$-), 2.52 (q, 2H, -CH_2S-), 1.68-1.25 (m, 22H, -(SCH_2)C<u>H</u>$_2$- + -N(CH_3)$_2$-C<u>H</u>$_2$-CH_3 + S<u>H</u>).

Compound **L3**: ^{1}H NMR (400MHz, $CDCl_3$, TMS): δ 3.95 (br, 2H, -CH_2N-), 3.68-3.56 (m, 14H, -CH_2O- + -OC<u>H</u>$_2$-(CH_2N)-), 3.46 (t, 2H, -CH_2O-), 3.40-3.33 (m, 2H,-NCH_2-), 3.19 (s, 6H, -(CH_3)$_2$N-), 2.87 (s, 3H, -$CH_3SO^-_3$-), 2.52 (q, 2H, -CH_2S-), 1.76-1.53 (m, 6H, -(NCH_2)C<u>H</u>$_2$-) + (SCH_2)C<u>H</u>$_2$ + -C<u>H</u>$_2$(CH_2O)-) , 1.41-1.22 (m, 21H, -SH + -(NCH_2CH_2-)C<u>H</u>$_2$-) + -CH_2-), 0.89 (t, 3H, - CH_3-).

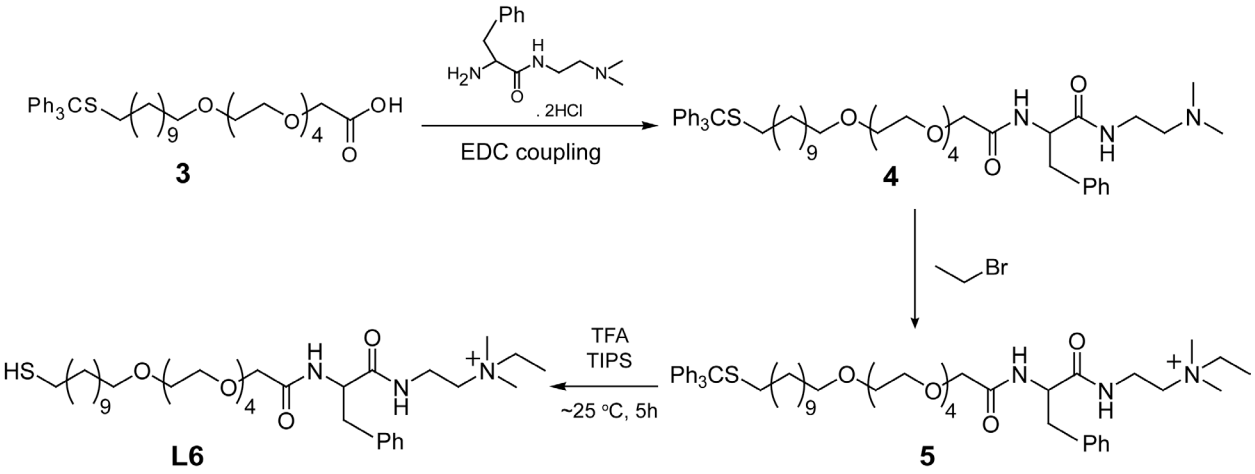
Compound **L4**: ^{1}H NMR (400MHz, $CDCl_3$, TMS): δ3.95 (br, 2H, -CH_2N-), 3.81-3.72 (m, 1H, H_{Cyclo}), 3.69-3.53 (m, 14H, -CH_2O- + -OC<u>H</u>$_2$-(CH_2N)-), 3.49 (t, 2H, -CH_2O-), 3.11 (s, 6H, -(CH_3)$_2$N-), 2.91 (s, 3H, -$CH_3SO^-_3$-), 2.52 (q, 2H, -CH_2S-), 2.23 (d, 2H, H_{Cyclo}), 1.99 (d, 2H, H_{Cyclo}), 1.78-1.52 (m, 4H, -(SCH_2)C<u>H</u>$_2$ + -C<u>H</u>$_2$(CH_2O)-), 1.51-1.12 (m, 21H, SH + -CH2- + H_{Cyclo}).

Compound **L5**: ^{1}H NMR (400MHz, $CDCl_3$, TMS): δ 7.82 (d, 2H, H_{Ar}), 7.66-7.51 (m, 3H, H_{Ar}), 4.24 (br, 2H, -CH_2N-), 3.78 (s, 6H, -(CH_3)$_2$N-), 3.68-3.52 (m, 14H, -CH_2O- + -OC<u>H</u>$_2$-(CH_2N)-), 3.47-3.36 (m, 2H, -CH_2O-), 2.87 (s, 3H, -$CH_3SO^-_3$-), 2.52 (q, 2H, -CH_2S-), 1.70-1.46 (m, 4H, -(SCH_2)C<u>H</u>$_2$ + -C<u>H</u>$_2$(CH_2O)-), 1.42-1.1.16 (m, 15H, -SH + -CH_2-).

Compound **L7**: ^{1}H NMR (400MHz, $CDCl_3$, TMS): δ 3.94 (br, 2H, -CH_2N-), 3.75-3.52 (m, 16H, -CH_2O- + -OC<u>H</u>$_2$-(CH_2N)- + -C<u>H</u>$_2$-OH), 3.48 (t, 2H, -CH_2O-), 3.39-3.31 (m, 2H,-NCH_2-), 3.25 (s, 6H, -(CH_3)$_2$N-), 3.2 (br, 1H, -OH), 2.89 (s, 3H, -$CH_3SO^-_3$-), 2.52 (q, 2H, -CH_2S-), 2.35-2.26 (m, 2H, -(NCH_2)C<u>H</u>$_2$-), 1.70-1.52 (m, 4H, + (SCH_2)C<u>H</u>$_2$ + -C<u>H</u>$_2$(CH_2O)-), 1.36-1.21 (m, 15H, -SH + -CH_2-).

Compound **L8**: ^{1}H NMR (400MHz, $CDCl_3$, TMS): δ 4.78 (br, 1H, -CHO<u>H</u>(CH2OH)-), 4.59 (br, 1H, -CH2O<u>H</u>-), 4.50-4.45 (m, 1H, -C<u>H</u>OH(CH2OH)-), 4.43 (d and br, 2H,-NCH_2-), 3.95 (d and br, 2H, -CH_2N-), 3.86-3.76 (d and br, 2H, -C<u>H</u>$_2$-OH), 3.75-3.55 (m, 14H, -CH_2O- + -OC<u>H</u>$_2$-(CH_2N)-), 3.48 (t, 2H, -CH_2O-), 3.34 (s, 6H, -(CH_3)$_2$N-), 2.99 (s, 3H, -$CH_3SO^-_3$-), 2.52 (q, 2H, -CH_2S-), 1.71-1.51 (m, 4H, + (SCH_2)C<u>H</u>$_2$ + -C<u>H</u>$_2$(CH_2O)-), 1.42-1.21 (m, 15H, -SH + -CH_2-).

Scheme 2:

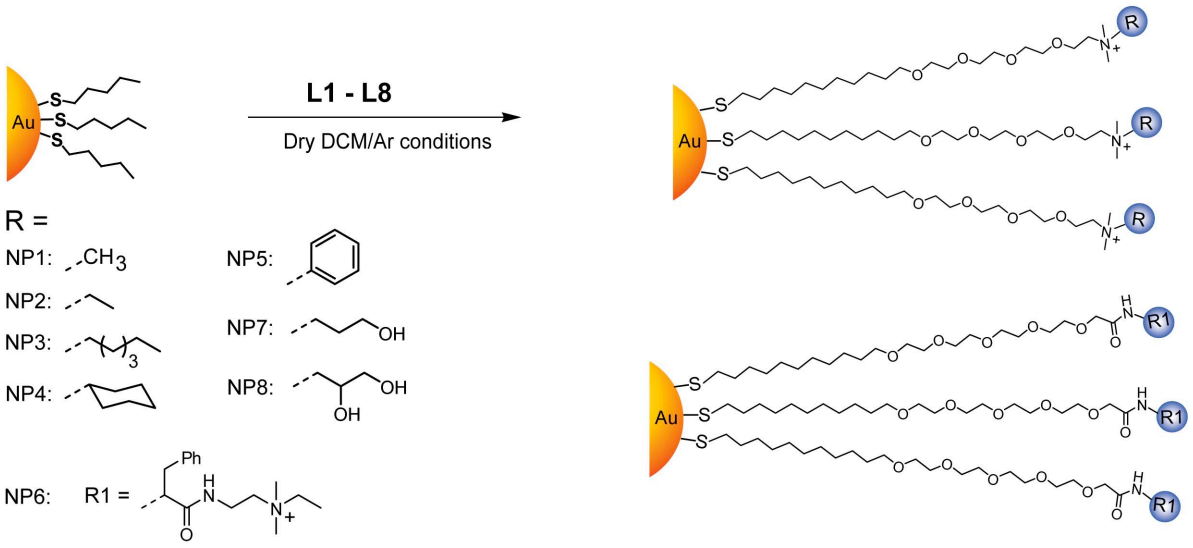


Procedure: Compound **4** bearing L-Phe group was synthesized through the reaction of compound **3** with corresponding 2-amino-N-(2-(dimethylamino)ethyl)-3-phenylpropanamide using EDC [1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide] as the coupling agent. The reaction was done in a mixture of dichloromethane (DCM) and N, N-dimethylformamide. After 2 days, the solution was poured into ethyl acetate (EtOAc) and water. The organic layer was washed with brine, and dried with Na_2SO_4. The crude product was purified on silica column using EtOAc/MeOH (90:10) and EtOAc/MeOH/NH_4OH (90:10:1) as gradient eluent. Compound **5** was obtained through nucleophilic substitution of compound **4** with bromoethane. Next, the trityl protection of the compound **5** was removed following the above mentioned procedure.

Compound **L6**: ^{1}H NMR (400MHz, $CDCl_3$, TMS): δ 8.47 (s, 1H, -N<u>H</u>(CH)-), 7.65 (s, 1H, -N<u>H</u>(CH_2)-), 7.29 (d of d, 1H, H_{Ar}), 7.23 (d of d, 1H, H_{Ar}), 4.62-4.59 (m, 1H, -NH(C<u>H</u>)-), 4.11-3.94 (m, 2H, -NH(C<u>H</u>$_2$)-), 3.76-3.51 (m, 20H, -CH_2O- + -C<u>H</u>$_2$(NCH_3)$_2$- + - (NCH_3)$_2$C<u>H</u>$_2$), 3.47 (t, 2H, -CH_2O-), 3.23 (d of d, 2H, -CH-C<u>H</u>$_2$-Ar), 3.10 (s, 6H, -(CH_3)$_2$N-), 2.9 (s, 2H, -OCH_2C=O-), 2.52 (q, 2H, -CH_2S-), 1.60-1.52 (m, 4H, -(SCH_2)C<u>H</u>$_2$ + -C<u>H</u>$_2$(CH_2O)-), 1.41-1.15 (m, 18H, -CH3 + -SH + -CH_2-).

(b) Nanoparticle synthesis

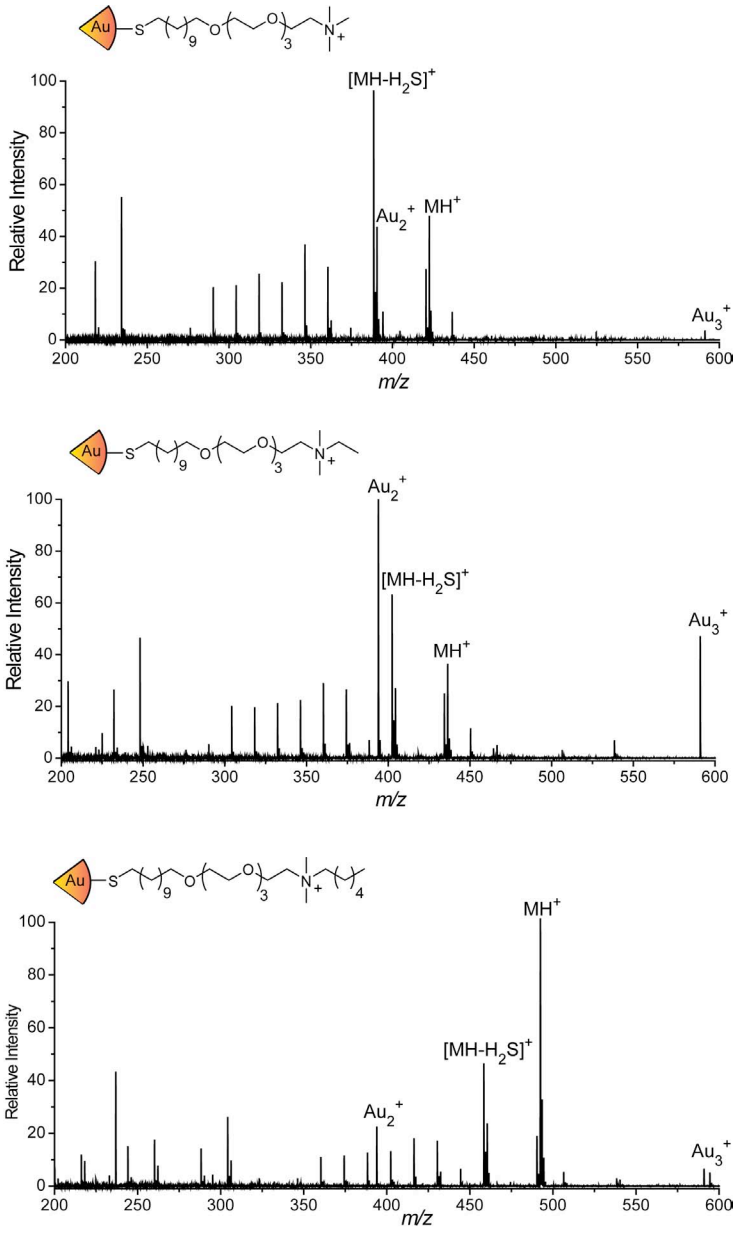
Scheme 3:

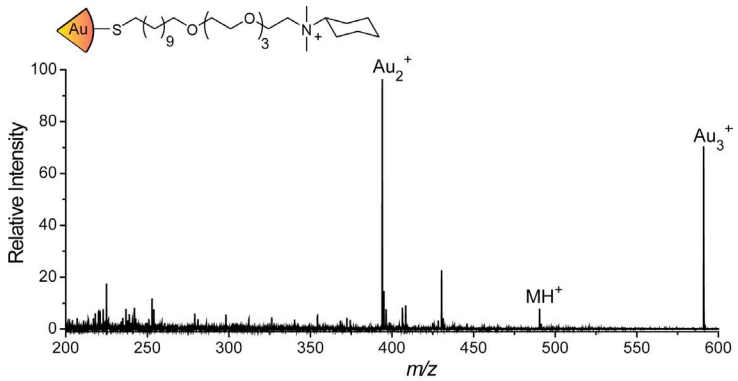


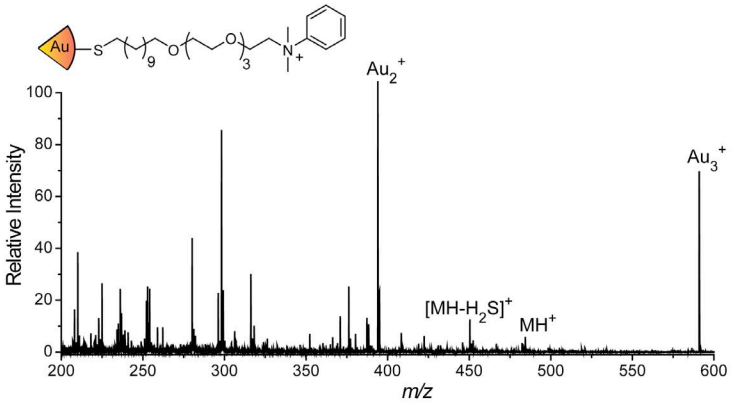
General procedure: 1-Pentanethiol coated gold nanoparticles (*d* = ~2 nm) were prepared according to the previously reported protocol.[2] Place-exchange reaction[3] of compound **Ln** (**n** = 1 – 8) dissolved in DCM with pentanethiol-coated gold nanoparticles (d = ~2 nm) was carried out for 3 days at environmental temperature. Then, DCM was evaporated under reduced pressure. The residue was dissolved in a small amount of distilled water and dialyzed (membrane MWCO = 1,000) to remove excess ligands, acetic acid and the other salts present with the nanoparticles. After dialysis, the particles were lyophilized and subsequently redispersed in deionized water (18 MΩ-cm). ^{1}H NMR spectra in D_2O showed substantial broadening of the proton signals and no free ligands were observed. Ligands on the nanoparticles were further characterized by mass spectrometry as described below.

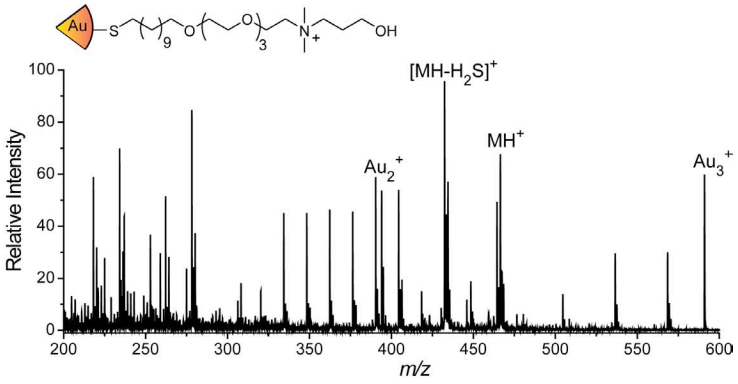
Laser Desorption/Ionization Mass Spectrometry (LDI-MS): The LDI-MS characterization of these NPs was performed following the reported method[4] on a Bruker Autoflex III mass spectrometer (Bruker Daltonics, Bremen, Germany) (Autoflex III). Operating conditions were as follows: ion source 1 = 19.00 kV, ion source 2 = 16.60 kV, lens voltage = 8.44 kV, reflector voltage = 20.00 kV, reflector voltage 2 = 9.69 kV, pulsed ion extraction time = 10 ns, suppression = 100 Da, and positive mode. The LDI mass spectra showed the fragmentation patterns of NP1 – 5, 7, and 8 that were consistent with previously published results. The

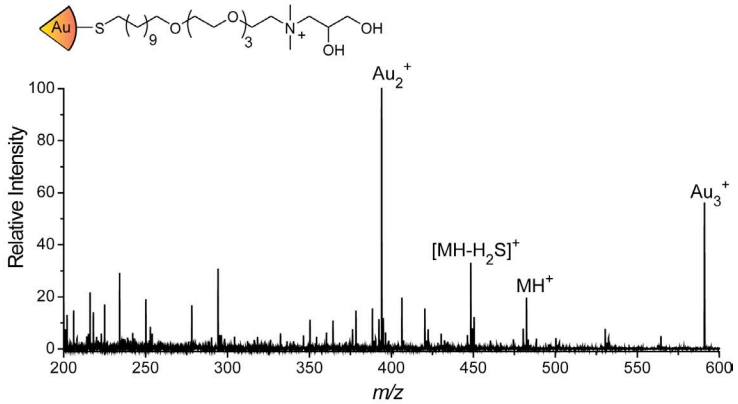
molecular ions (MH^+), major fragments ($[MH\text{-}H_2S]^+$), and gold cluster ions (Au_2^+ and Au_3^+) were observed. The LDI-MS spectra for the NPs are presented below:



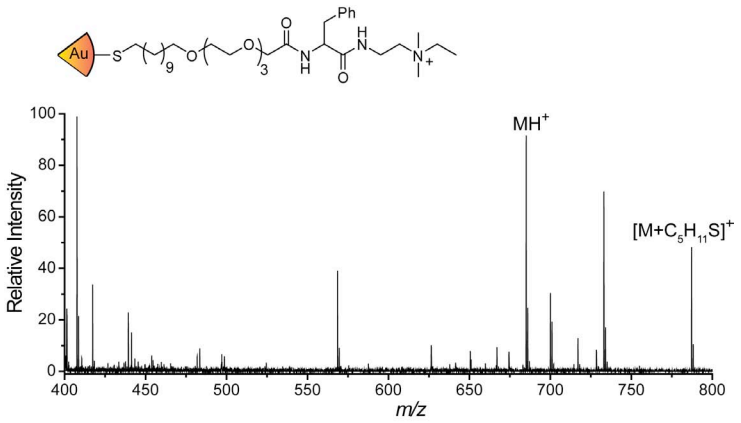

Au
S
9
O
O
3
N+
Au2+
Au3+
MH+
Relative Intensity
m/z


Au
S
9
O
O
3
N+
Au2+
Au3+
[MH-H2S]+
MH+
Relative Intensity
m/z


Au
S
9
O
O
3
N+
OH
[MH-H2S]+
MH+
Au2+
Au3+
Relative Intensity
m/z


Au
S
9
O
O
3
N+
OH
OH
Au2+
Au3+
[MH-H2S]+
MH+
Relative Intensity
m/z

The surface ligand on NP6 has a larger molecular weight and is easier to get fragmented under LDI-MS conditions. Thus, a softer ionization method, such as matrix assisted LDI-MS (MALDI-MS) has been performed. A saturated α-CHCA stock solution was prepared in 70% acetonitrile and 30% H_2O. An equal volume of NP6 solution was added to the matrix stock solution. 2.5 µL of this mixture was applied to the sample carrier, and then the MALDI-MS analysis was performed. The molecular ion (MH^+, *m/z* = 685) has been readily observed in the mass spectrum. The MALDI-MS spectrum is as follows:



2) Expression of GFP and spectral characterization.

Green fluorescence protein (GFP) was expressed according to the known procedure.[5] Briefly, starter cultures from a glycerol stock of GFP [enhanced GFP (EGFP) was cloned into the pET21d vector (Novagen) where 6×His tag is located at C-terminus] in BL21(DE3) was grown overnight in 50 ml of 2×YT culture media (16 g tryptone, 10 g yeast extract, 5 g NaCl in 1 L water) with 50 µL of 100 mg/mL ampicillin at a final concentration of 100 µg/mL. The cultures were shaken overnight at 250 rpm at 37 ^oC. On the following day, 5 ml of the starter cultures was added to a Fernbach flask containing 1 L of 2×YT and 1 ml 1000x ampicillin; and flask was shaken until the OD_{600} reaches ~0.6. The culture was then induced by adding Isopropyl β-D-1-thiogalactopyranoside (IPTG, 1 mM final concentration) and shaken at 28 ^oC. After three hours, the cells were harvested by centrifugation (5000 rpm for 15 minutes at 4 ^oC). The pellet was then

resuspended in lysis buffer (2 mM Immidizole, 50 mM NaH_2PO_4, 300 mM NaCl). The cells were lysed using a microfluidizer and once lysed, the solution was spun down at 15000 rpm for 45 minutes at 4 ^oC. The supernatant was further purified using HisPur Cobalt columns from Pierce (cat. Number 89969) followed by dialysis in 5 mM sodium phosphate buffer (pH 7.4).

The absorbance and emission spectra (at $\lambda_{exc.}$= 475 nm) of EGFP is shown below:

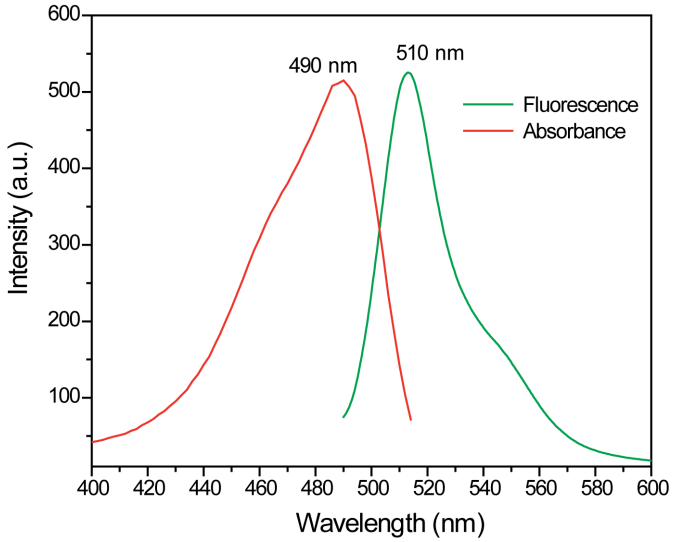


Figure S1. Absorbance and fluorescence spectra of GFP in 5 mM sodium phosphate buffer at pH 7.40.

3. Fluorescence titrations

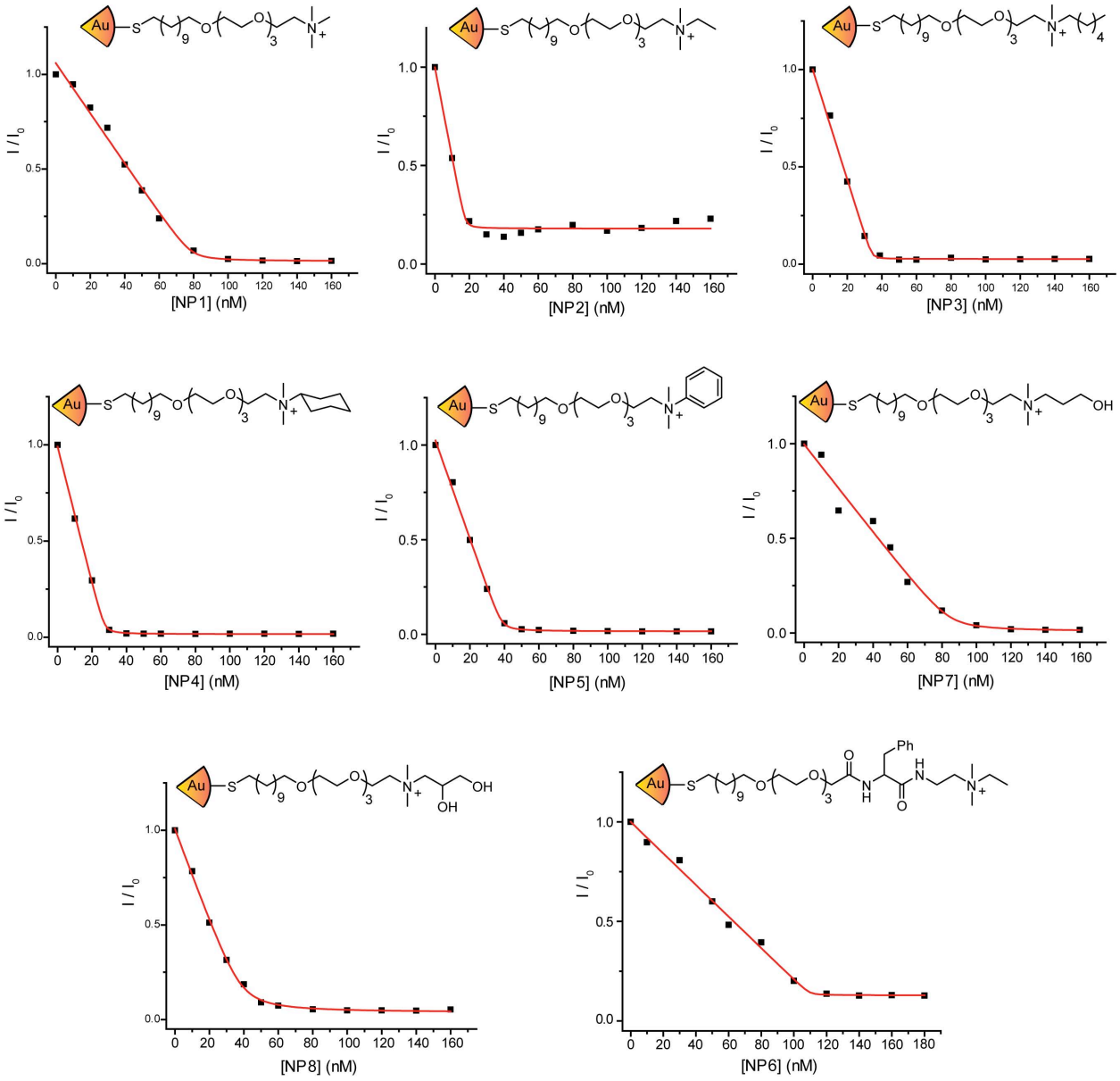


Figure S2. Fluorescence titration curves for the complexation of GFP with 8 different cationic gold nanoparticles (NP1 – NP8). The changes of fluorescence intensity at 510 nm were measured following the addition of cationic nanoparticles (0-180 nM) with an excitation wavelength of 475 nm. The red solid lines represent the best curve fitting using the model of single set of identical binding sites.

Table S1. Binding constant (K_a), and binding stoichiometry (n, GFP/NP) between GFP and cationic nanoparticles (NP1 – NP8) as determined from the fluorescence titrations.

Nanoparticle	K_a (10^9 M^{-1})	n
NP1	3.41	1.3
NP2	4.03	3.6
NP3	6.8	4.6
NP4	5.99	1.9
NP5	3.77	2.1
NP6	18.6	1.6
NP7	1.53	1.1
NP8	0.37	2.5

4. Histopathology of the metastases

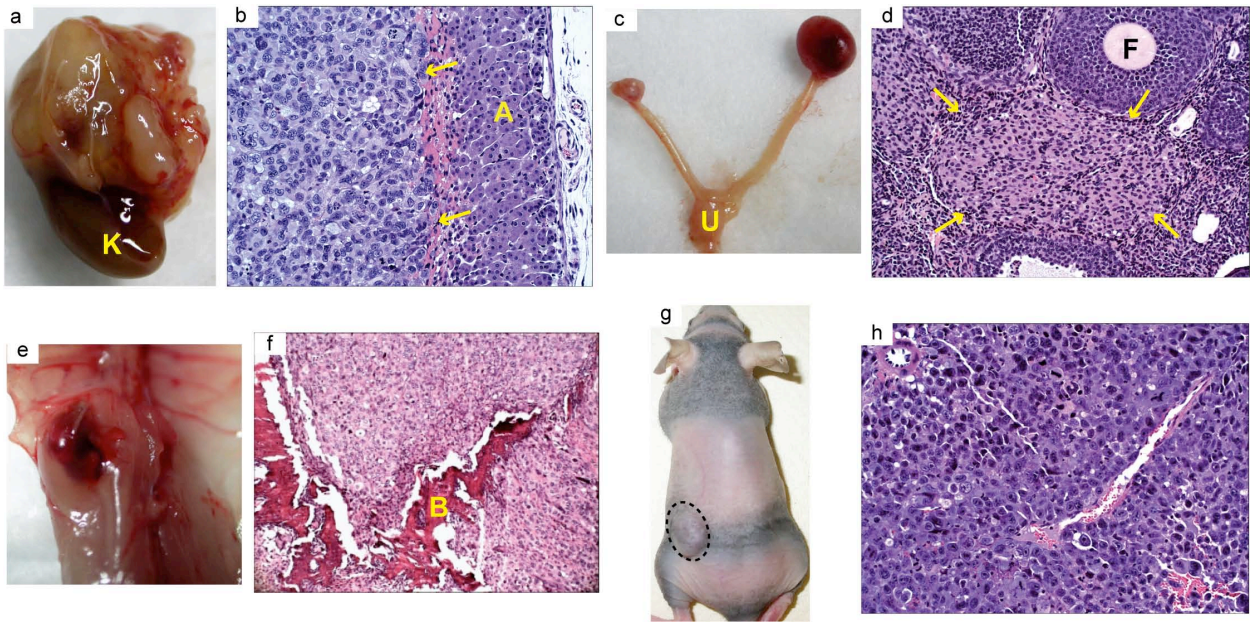


Figure S3. Establishing an *in vivo* metastasis model. Gross anatomy and histopathology of adrenal (a, b), ovarian (c, d), and vertebral (e, f) metastases (indicated by arrows), as well as tumor resulting from the subcutaneous injection of H1299 cells (g, h) respectively. K = kidney, A = adrenal tissue, U = uterus, F = ovarian follicle, B = vertebral bone.

5. Fluorescence response data

Table S2. Fluorescence response (I / I_0) data set for the parental and organ-specific metastatic subline cell lysates using NP-GFP complexes (corresponding to Fig. 3a in the main text).

Subline cells	Fluorescence response							
	NP1	NP2	NP3	NP4	NP5	NP6	NP7	NP8
Adrenal	6.130	1.084	0.990	3.097	4.736	1.721	1.686	1.521
Adrenal	6.760	1.083	0.979	3.107	4.702	1.768	1.568	1.579
Adrenal	5.680	1.085	0.967	3.353	4.347	1.744	1.643	1.556
Adrenal	6.699	1.125	0.964	3.115	4.328	1.748	1.569	1.479
Adrenal	6.909	1.106	0.986	3.174	4.546	1.865	1.593	1.457
Adrenal	6.331	1.100	0.982	3.045	4.335	1.813	1.484	1.513
Bone	6.439	1.173	1.027	3.783	4.713	1.951	1.746	1.660
Bone	5.769	1.169	1.019	3.529	4.724	1.924	1.661	1.734
Bone	6.452	1.175	1.003	3.155	4.536	1.877	1.803	1.585
Bone	6.625	1.163	1.009	3.292	4.607	1.984	1.735	1.613
Bone	5.948	1.157	1.018	3.425	4.578	1.910	1.667	1.614
Bone	6.502	1.156	0.992	3.817	4.609	2.000	1.707	1.695
Ovary	5.117	1.127	1.024	3.034	3.919	1.778	1.587	1.600
Ovary	5.335	1.137	1.002	2.990	3.993	1.776	1.640	1.483
Ovary	5.053	1.149	1.029	2.954	4.042	1.780	1.654	1.577
Ovary	5.332	1.126	1.002	3.050	3.880	1.892	1.594	1.549
Ovary	4.890	1.122	1.040	2.857	4.080	1.751	1.602	1.453
Ovary	4.695	1.125	1.017	2.890	3.944	1.814	1.535	1.563
Parental	5.590	1.126	0.996	3.622	4.479	1.992	1.774	1.631
Parental	5.618	1.122	0.998	3.943	4.246	1.973	1.725	1.575
Parental	5.884	1.173	1.018	3.513	4.264	2.073	1.729	1.678
Parental	5.550	1.166	0.993	3.967	4.498	2.150	1.629	1.636

Parental	5.606	1.139	1.005	3.708	4.277	2.021	1.644	1.663
Parental	4.841	1.139	1.024	3.982	4.136	1.968	1.702	1.657

Table S3. Fluorescence response (I / I_0) data set for the tumor tissue lysates using NP-GFP complexes (corresponding to Fig. 4a).

Tumor tissues	**Fluorescence responses**							
	NP1	NP2	NP3	NP4	NP5	NP6	NP7	NP8
Bone	5.535	1.010	0.912	6.665	5.362	2.592	1.846	1.885
Bone	5.976	1.024	0.897	7.009	5.633	2.731	1.794	1.992
Bone	5.731	0.991	0.886	7.064	5.463	2.655	1.786	2.000
Bone	5.535	0.992	0.886	6.982	4.945	2.600	1.747	1.894
Bone	5.346	0.986	0.910	6.138	5.015	2.609	1.802	1.875
Bone	5.435	1.024	0.899	6.260	5.061	2.733	1.799	1.934
Subcutaneous	6.723	1.010	0.933	7.524	5.661	2.724	2.265	1.833
Subcutaneous	7.060	1.066	0.926	7.417	5.501	2.795	2.121	1.939
Subcutaneous	6.819	1.052	0.928	7.387	5.063	2.802	2.276	1.844
Subcutaneous	6.645	1.091	0.935	7.047	5.046	2.767	2.094	1.778
Subcutaneous	6.470	1.072	0.945	6.919	5.110	2.776	2.119	1.809
Subcutaneous	6.894	1.074	0.938	6.970	5.107	2.922	2.055	1.921
Adrenal	6.554	1.035	0.921	7.675	5.166	2.885	2.032	1.821
Adrenal	6.509	1.019	0.916	7.779	5.107	2.773	2.062	1.872
Adrenal	6.200	1.013	0.911	7.387	4.872	2.908	1.947	1.831
Adrenal	6.119	1.015	0.905	7.023	4.623	2.917	1.891	1.859
Adrenal	6.379	1.058	0.920	7.881	4.809	2.829	1.950	1.875
Adrenal	6.440	1.019	0.948	7.377	5.337	2.903	1.883	1.888
Ovary	7.922	1.045	0.987	9.303	6.377	3.397	2.109	1.712
Ovary	7.998	1.051	0.974	8.205	5.828	3.238	1.961	1.705
Ovary	8.259	1.033	1.020	8.322	5.390	3.080	1.991	1.744

Ovary	8.062	1.027	1.009	8.378	5.853	3.085	1.919	1.722
Ovary	6.986	1.051	0.971	8.998	5.421	3.159	1.905	1.749
Ovary	8.351	1.035	0.994	9.560	5.942	3.026	1.940	1.787

Table S4. Fluorescence response (I / I_0) data set for the normal tissue lysates using NP-GFP complexes (corresponding to Fig. 4a).

Normal tissues	**Fluorescence responses**							
	NP1	NP2	NP3	NP4	NP5	NP6	NP7	NP8
Adrenal	3.937	1.255	0.983	3.35	5.426	2.096	1.666	1.383
Adrenal	3.760	1.106	0.973	3.120	4.712	2.115	1.694	1.361
Adrenal	4.039	1.007	0.969	3.052	6.016	2.027	1.677	1.388
Adrenal	4.076	1.102	0.957	2.947	4.920	2.146	1.775	1.365
Adrenal	3.982	1.156	0.964	3.183	5.871	2.120	1.805	1.388
Adrenal	4.226	1.061	0.998	3.179	5.797	2.020	1.790	1.402
Lung	4.367	1.030	0.989	3.209	6.372	2.871	1.871	1.282
Lung	5.360	1.081	0.980	3.382	6.074	2.831	1.881	1.304
Lung	5.220	1.130	0.975	3.243	5.823	2.794	1.869	1.286
Lung	5.526	1.060	1.007	3.256	6.355	3.012	1.909	1.252
Lung	4.764	1.099	0.992	3.694	6.600	2.902	1.843	1.358
Lung	4.455	1.042	0.994	3.395	5.811	2.877	1.881	1.369
Ovary	5.723	0.996	1.002	4.913	7.254	3.156	2.409	1.476
Ovary	5.442	1.072	0.996	5.868	7.145	3.238	2.214	1.478
Ovary	6.153	1.037	1.011	5.144	6.854	3.196	2.169	1.538
Ovary	5.880	1.161	0.982	5.059	6.975	3.080	2.190	1.585
Ovary	5.848	1.108	0.988	5.626	6.649	3.145	2.102	1.528
Ovary	6.032	1.025	1.006	5.195	6.691	3.268	2.014	1.506
Skin	4.427	1.077	0.981	3.984	6.964	2.894	2.113	1.466
Skin	4.370	1.080	1.039	3.944	7.139	2.989	2.054	1.420

Skin	4.510	1.053	1.073	3.964	6.424	2.841	1.910	1.443
Skin	4.535	1.062	1.034	3.761	6.113	2.850	2.051	1.470
Skin	4.572	1.047	0.993	3.908	6.024	3.122	1.940	1.662
Skin	4.339	1.077	0.976	3.774	6.572	2.800	1.922	1.614

6. Description of linear discriminant analysis

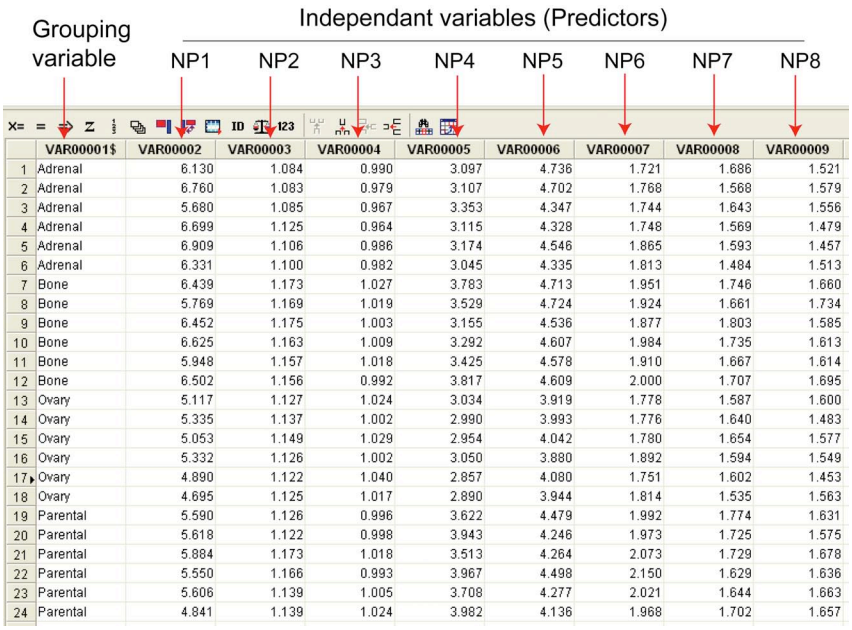
Linear discriminant analysis (LDA), a supervised multivariate method, is used to separate classes of objects or assign new objects to appropriate classes. The analysis generates a new space given by the canonical discriminant **factors** (also called canonical functions or discriminant axes), which describes best similarities and differences between groups under consideration. In the discriminant score plot each response pattern generated can be reduced to a single score and plotted in the new discriminant space. The discriminants are linear combinations of the measured (independent) variables such as the sensor responses. Discriminant functions (also called canonical functions for multiple group case) are calculated with the objective of maximizing the distance between classes relative to the variation within classes. The first discriminant is the linear combination of the variables that best discriminates among the groups; the second discriminant is orthogonal to the first and is the next best combination of variables, and so on. First discriminant function:

$$L_{DF} = c + a_1x_1 + a_2x_2 + \ldots + a_nx_n$$

where, a_1 through a_n are discriminant coefficients, x_1 through x_n are discriminating variables, and c is a constant. The coefficients for each variable in each discriminant function can be interpreted as: the larger the standardized coefficient, the greater is the contribution of the respective variable to the discrimination between groups. Computationally, a canonical correlation analysis determines the successive functions and canonical *roots* (the term root refers to the eigenvalues that are associated with the respective canonical function). The maximum number of functions will be equal to the number of groups minus one, or the number of variables in the analysis, whichever is smaller.

The different statistical terms are described using the parameters obtained for data in Table 1.

Below is the data input section in the software, SYSTAT:



	VAR00001$	VAR00002	VAR00003	VAR00004	VAR00005	VAR00006	VAR00007	VAR00008	VAR00009
1	Adrenal	6.130	1.084	0.990	3.097	4.736	1.721	1.686	1.521
2	Adrenal	6.760	1.083	0.979	3.107	4.702	1.768	1.568	1.579
3	Adrenal	5.680	1.085	0.967	3.353	4.347	1.744	1.643	1.556
4	Adrenal	6.699	1.125	0.964	3.115	4.328	1.748	1.569	1.479
5	Adrenal	6.909	1.106	0.986	3.174	4.546	1.865	1.593	1.457
6	Adrenal	6.331	1.100	0.982	3.045	4.335	1.813	1.484	1.513
7	Bone	6.439	1.173	1.027	3.783	4.713	1.951	1.746	1.660
8	Bone	5.769	1.169	1.019	3.529	4.724	1.924	1.661	1.734
9	Bone	6.452	1.175	1.003	3.155	4.536	1.877	1.803	1.585
10	Bone	6.625	1.163	1.009	3.292	4.607	1.984	1.735	1.613
11	Bone	5.948	1.157	1.018	3.425	4.578	1.910	1.667	1.614
12	Bone	6.502	1.156	0.992	3.817	4.609	2.000	1.707	1.695
13	Ovary	5.117	1.127	1.024	3.034	3.919	1.778	1.587	1.600
14	Ovary	5.335	1.137	1.002	2.990	3.993	1.776	1.640	1.483
15	Ovary	5.053	1.149	1.029	2.954	4.042	1.780	1.654	1.577
16	Ovary	5.332	1.126	1.002	3.050	3.880	1.892	1.594	1.549
17	Ovary	4.890	1.122	1.040	2.857	4.080	1.751	1.602	1.453
18	Ovary	4.695	1.125	1.017	2.890	3.944	1.814	1.535	1.563
19	Parental	5.590	1.126	0.996	3.622	4.479	1.992	1.774	1.631
20	Parental	5.618	1.122	0.998	3.943	4.246	1.973	1.725	1.575
21	Parental	5.884	1.173	1.018	3.513	4.264	2.073	1.729	1.678
22	Parental	5.550	1.166	0.993	3.967	4.498	2.150	1.629	1.636
23	Parental	5.606	1.139	1.005	3.708	4.277	2.021	1.644	1.663
24	Parental	4.841	1.139	1.024	3.982	4.136	1.968	1.702	1.657

The important statistical outputs are as follows:

Group frequencies

Adrenal	Bone	Ovary	Parental
6	6	6	6

Group means

	Adrenal	Bone	Ovary	Parental
VAR00002	6.418	6.289	5.070	5.515
VAR00003	1.097	1.166	1.131	1.144
VAR00004	0.978	1.011	1.019	1.006
VAR00005	3.148	3.500	2.962	3.789
VAR00006	4.499	4.628	3.976	4.317
VAR00007	1.776	1.941	1.798	2.030
VAR00008	1.591	1.720	1.602	1.701
VAR00009	1.518	1.650	1.537	1.640

Classification functions

	Adrenal	Bone	Ovary	Parental
CONSTANT	-8333.996	-9310.458	-8721.498	-9249.783

VAR00002	37.136	28.559	19.594	17.413
VAR00003	4523.237	4823.019	4610.813	4650.215
VAR00004	6967.687	7212.380	7245.597	7268.904
VAR00005	218.819	228.673	205.273	227.559
VAR00006	53.271	49.848	14.405	25.213
VAR00007	705.366	806.622	820.732	916.694
VAR00008	1094.119	1206.462	1171.652	1255.724
VAR00009	479.127	515.166	472.138	490.321

Jackknifed classification matrix

	Adrenal	Bone	Ovary	Parental	%correct
Adrenal	6	0	0	0	100
Bone	0	6	0	0	100
Ovary	0	0	6	0	100
Parental	0	0	0	6	100
Total	6	6	6	6	100

The Jackknifed classification shows no overlap between the groups at all, and all the cases are classified with 100% classification accuracy.

Eigenvalues

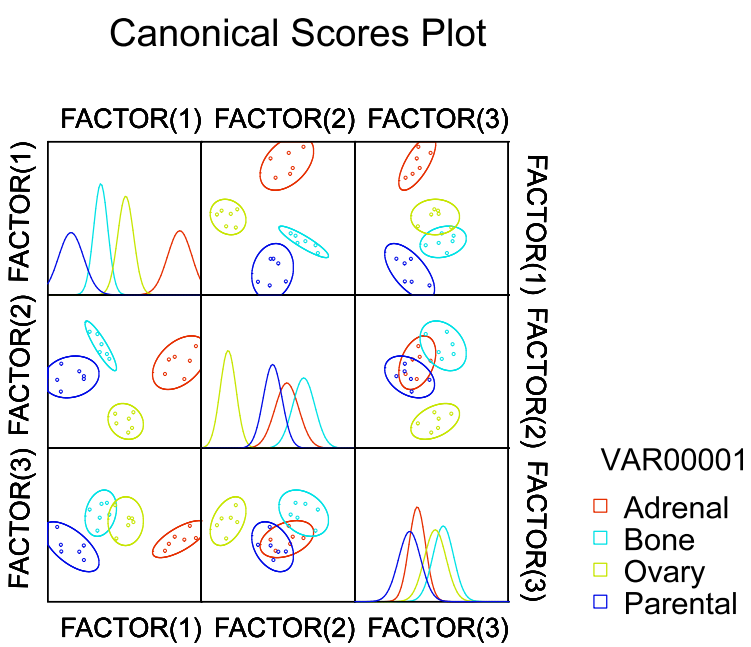
18.250	9.035	2.077

The Eigenvalues or roots are noteworthy. The first eigenvalue (18.250) is very large relative to the second, indicating that the first canonical function (or the Factor (1) in the LDA figures) captures most of the difference among the groups. Also, it indicates that three discriminants are derived for classification of the groups.

Cumulative proportion of total dispersion

0.622	0.929	1.000

The contribution from for each factor towards the total classification is computed from the above result, where the first term corresponds to Factor (1), second term corresponds to Factor (1)+Factor (2), and the third is the combination of all three. It is clear that the factors have contribution of 62.2%, 30.7%, and 7.1% for Factor (1), Factor (2), and Factor (3) respectively.



The scores corresponding to each factor are then plotted as scatter plot with a certain confidence ellipse.

Jackknifed classification:

The Jackknifed classification (leave-one-out cross-validation) is used to test the predictability of the sensor array by leaving one observation out of the set at the time, and uses the rest of the data as a training set to generate the linear discriminant (LD) function. The LD function is then used to place the excluded observation (data point) within the correct cluster. This is performed for each observation, and the overall ability to classify the observations describes the quality and predictability of the array.

The results from the Jackknifed analysis are listed below (Table S5 – S8). The classification matrices indicate the differential binding of the NPs with the lysates and the contributions of the NPs in discrimination. It should be noted that the minimum number of sensor element varies from 3 (metastatic cells) to 6 (metastatic tissues) depending on the analyte contents. Increasing the number of sensors, however, increases the robustness of the sensor system. It should be noted that based on the percent accuracy of identification by a particular set, the next optimal set with

higher classification accuracy is determined. In order to have a general set of sensors, we reported the result from the Jackknifed analysis with a common set of sensor elements (NP1 to NP8).

7. Jackknifed analyses of the fluorescence responses

Table S5. Percent accuracy of classification of the *metastatic cells* using Jackknifed analysis (leave-one-out validation) using a NP or combinations.

NP1	NP2	NP3	NP4	NP5	NP6	NP7	NP8	Classification Accuracy, %
■								50
	■							75
		■						63
			■					63
				■				63
					■			54
						■		42
							■	46
■	■							71
	■	■						67
		■	■					83
			■	■				79
				■	■			83
					■	■		75
						■	■	67
■	■	■	■					88
	■	■	■	■				88
		■	■	■	■			88
			■	■	■	■		83
				■	■	■	■	83
					■	■	■	79
■	■	■			■			100
	■	■			■			83
		■		■	■			96
					■	■		83
					■		■	79
■	■	■	■		■			96
■	■	■	■	■	■			100
■	■	■		■	■			100

								100
								100
								100

Table S6. Percent accuracy of classification of the *metastatic tissues* using Jackknifed analysis (leave-one-out validation) using a NP or combinations.

NP1	NP2	NP3	NP4	NP5	NP6	NP7	NP8	Classification Accuracy, %
								83
								54
								79
								54
								29
								71
								58
								54
								83
								88
								92
								83
								83
								88
								88
								83
								88
								92
								96
								96
								92
								88
								96
								88
								92
								96
								96
								92
								96
								92
								88

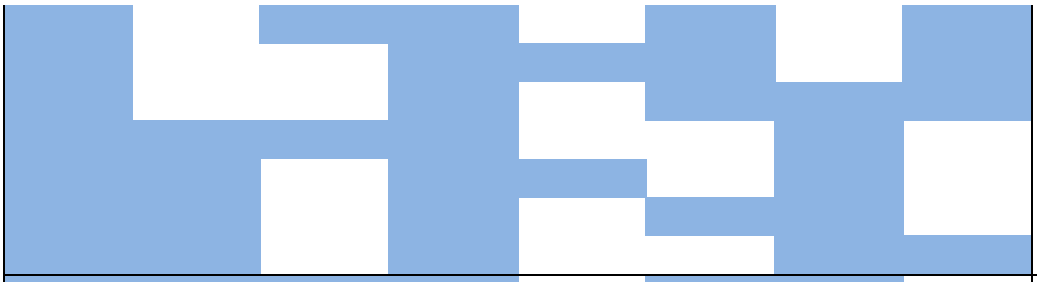
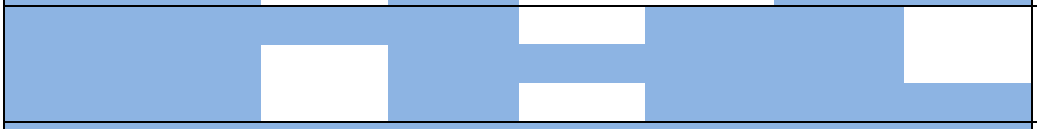
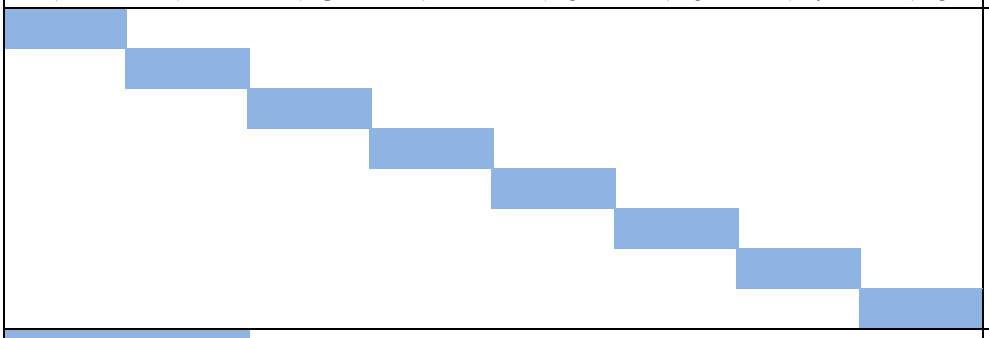
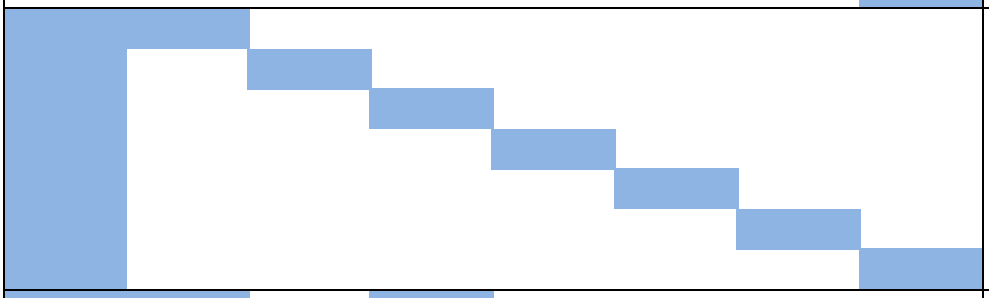
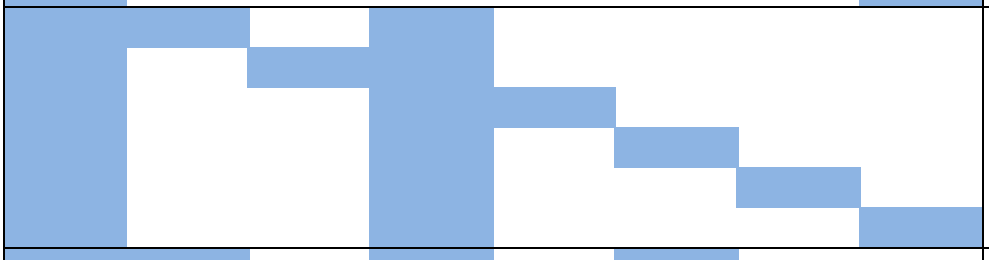
	92 83 92 92 92 96 96
	100 96 100
	100

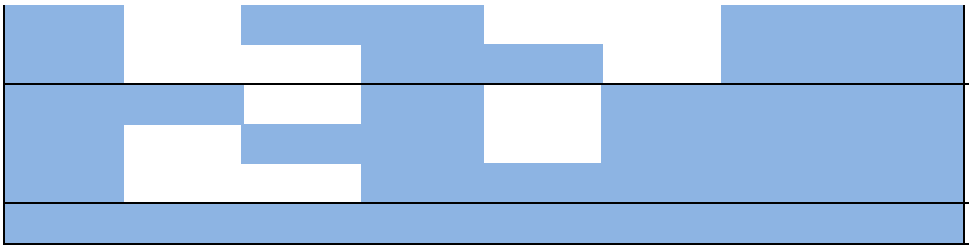
Table S7. Percent accuracy of classification of the *normal tissues* using Jackknifed analysis (leave-one-out validation) using a NP or combinations.

NP1	NP2	NP3	NP4	NP5	NP6	NP7	NP8	Classification Accuracy, %
								79 21 46 79 42 71 79 50
								75 67 92 79 88 92 88
								88 92 88 96 96 92
								92 88

								92
								100
								100
								100

Table S8. Percent accuracy of classification of the *normal* and *metastatic tissues* using Jackknifed analysis (leave-one-out validation) using a NP or combinations.

NP1	NP2	NP3	NP4	NP5	NP6	NP7	NP8	Classification Accuracy, %
								71
								29
								50
								67
								29
								42
								46
								52
								73
								77
								90
								77
								88
								85
								88
								88
								90
								90
								94
								96
								94
								96
								94
								94
								98
								98
								96
								96
								94
								98
								98



	98
	98
	100
	100
	98
	100

8. Identification of unknown samples

Table S9. Fluorescence response (I / I_0) data set for identification of unknown tumor tissue lysates selected from the training set (Figure 4) using NP-GFP complexes.

#	Fluorescence responses								Identified as	Verification
	NP1	NP2	NP3	NP4	NP5	NP6	NP7	NP8		
1	8.030	1.066	1.029	7.860	5.846	2.523	2.153	1.322	Ovary	Ovary
2	7.601	1.054	1.031	8.536	6.127	2.296	2.570	1.283	Ovary	Ovary
3	6.501	1.068	0.940	7.779	4.800	2.965	1.863	1.780	Adrenal	Adrenal
4	5.901	1.045	0.997	4.431	5.714	2.293	2.863	1.592	Subcutaneous	Subcutaneous
5	5.379	1.056	0.936	5.161	5.426	2.326	3.124	1.538	Subcutaneous	Subcutaneous
6	5.431	1.047	0.931	6.143	4.849	2.055	1.766	1.706	Bone	Bone
7	5.498	1.077	0.953	7.444	5.842	2.665	2.890	1.699	Subcutaneous	Subcutaneous
8	6.322	1.097	0.962	7.221	4.982	2.657	1.921	1.697	Adrenal	Adrenal
9	6.582	1.266	0.929	7.182	5.023	2.443	1.906	1.780	Adrenal	Subcutaneous
10	5.143	1.028	0.900	6.091	4.313	2.072	1.570	1.709	Bone	Bone
11	4.836	1.112	0.928	5.708	5.167	2.583	1.627	1.863	Bone	Bone
12	5.632	0.975	0.915	6.145	4.124	2.274	1.474	1.597	Bone	Bone
13	7.208	1.071	0.950	7.694	6.630	2.195	2.285	1.330	Ovary	Ovary
14	8.357	1.084	0.986	7.848	5.727	2.251	2.085	1.360	Ovary	Ovary
15	5.557	1.124	0.928	5.611	4.731	2.097	2.689	1.836	Subcutaneous	Subcutaneous

16	7.543	1.108	0.952	5.065	5.400	2.219	2.709	1.661	Subcutaneous	Subcutaneous
17	6.697	1.104	0.967	5.485	4.538	2.095	2.880	1.746	Subcutaneous	Subcutaneous
18	6.101	1.109	0.946	7.626	5.004	2.772	1.858	1.750	Adrenal	Adrenal
19	6.594	1.061	0.923	7.732	4.786	2.708	1.964	1.758	Adrenal	Adrenal
20	5.414	1.087	0.958	4.276	4.397	2.962	2.563	1.507	Subcutaneous	Subcutaneous
21	6.724	1.104	0.948	3.801	4.321	2.369	2.643	1.583	Subcutaneous	Subcutaneous
22	5.217	1.021	0.916	6.003	3.995	2.165	1.508	1.602	Bone	Bone
23	5.369	1.155	0.968	5.138	4.413	1.657	1.496	1.758	Bone	None
24	8.144	1.182	0.960	7.565	6.091	2.824	2.738	1.358	Ovary	Ovary
25	5.647	1.094	0.982	7.913	5.202	2.112	2.195	1.464	Adrenal	Adrenal
26	7.153	1.086	0.951	7.885	5.313	2.158	2.121	1.712	Adrenal	Adrenal
27	8.034	1.117	1.002	7.180	5.297	2.210	2.040	1.520	Ovary	Ovary
28	6.192	1.086	1.003	7.235	5.126	2.657	1.970	1.630	Adrenal	Adrenal
29	6.648	1.084	0.968	7.769	4.932	2.354	1.833	1.719	Adrenal	Adrenal
30	5.875	1.174	0.940	7.628	5.003	2.286	1.991	1.717	Adrenal	Adrenal
31	5.312	1.057	0.938	5.687	4.426	1.886	1.532	1.749	Bone	Bone
32	5.415	1.023	0.969	6.571	4.525	1.715	1.431	1.783	Bone	Bone

9. Three-dimensional LDA plot for normal and tumor tissues

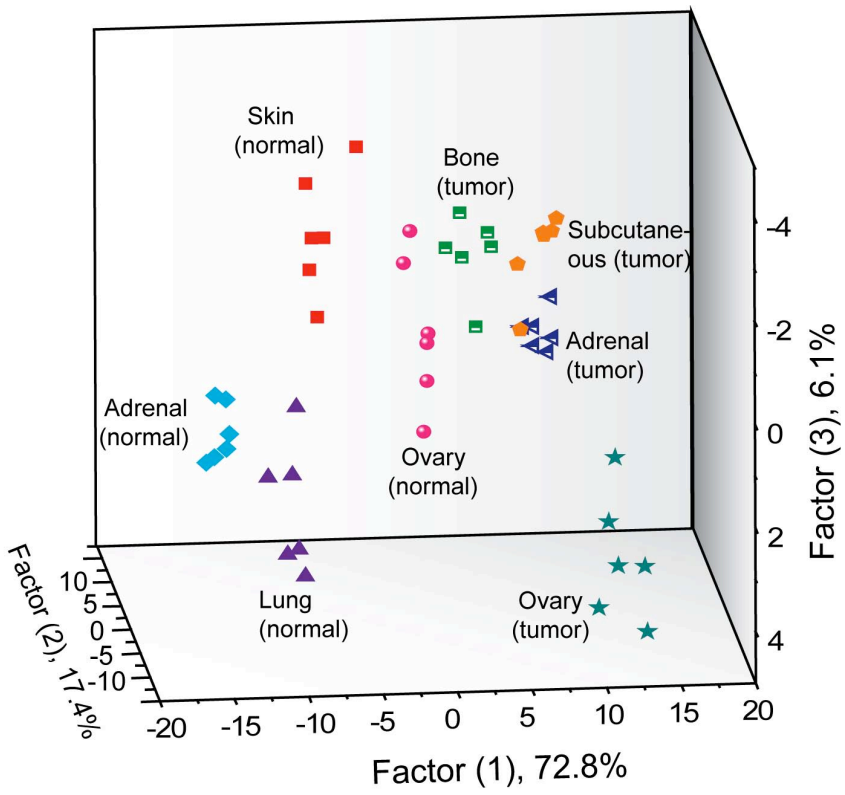


Figure S4. A 3-Dimensional LDA plot of the combined fluorescence response patterns of tumor and healthy tissues including the three most significant variances amongst the seven, with 95% confidence ellipses.

References

1. Miranda, O. R.; Chen, H.; You, C.; Mortenson, D. E.; Yang, X.; Bunz, U. H. F.; Rotello, V. M. Enzyme-Amplified Array Sensing of Proteins in Solution and in Biofluids. *J. Am. Chem. Soc.* **2010**, *132*, 5285-5289.

2. Brust, M.; Walker, M.; Bethell, D.; Schiffrin, D. J.; Whyman, R. Synthesis of Thiol-Derivatised Gold Nanoparticles in a Two-Phase Liquid–Liquid System. *J.Chem. Soc. – Chem. Comm.* **1994**, 801.

3. Hostetler, M. J.; Templeton, A. C.; Murray, R. W. Dynamics of Place-Exchange Reactions on Monolayer-Protected Gold Cluster Molecules. *Langmuir* **1999**, *15*, 3782-3789.

4. Yan, B.; Zhu, Z. J.; Miranda, O. R.; Chompoosor, A.; Rotello, V. M.; Vachet, R. W. Laser Desorption/Ionization Mass Spectrometry Analysis of Monolayer-Protected Gold Nanoparticles. *Anal. Bioanal. Chem.* **2010**, *396*, 1025-1035.

5. De, M.; Rana, S.; Rotello, V. M. Nickel-Ion-Mediated Control of the Stoichiometry of His-Tagged Protein/Nanoparticle Interactions. *Macromol. Biosci.* **2009**, *9*, 174-178.