Supplementary Material (ESI) for Chemical Communications

Glutathione-Bound Gold Nanoclusters for Selective-Binding and Detection of Glutathione S-Transferase-Fusion Proteins from Cell Lysates

Cheng-Tai Chen,[1] Wei-Jen Chen,[1] Chao-Zong Liu,[2,3] Ling-Ya Chang,[2,3] Yu-Chie Chen[1*]

[1]Department of Applied Chemistry, National Chiao Tung University, Hsinchu 300, Taiwan
[2]Department of Pharmacology and [3]Institute of Pharmacology & Toxicology, Tzu Chi University, Hualien 970, Taiwan

Supporting Information

Experimental

Reagents and Materials

L-Glutathione reduced, sodium phosphate dibasic heptahydrate, potassium phosphate monobasic anhydrous, Triton X-100, sinapinic acid, trypsin (proteomics grade), α-cyano-4-hydoxycinnamic acid, and ampicillin were purchased from Sigma (St. Louis, MO). Sodium chloride, potassium chloride, calcium chloride, ammonium hydrogen carbonate, glucose, and trifluoroacetic acid (TFA) were obtained from Riedel-deHaën (Seelze, Germany). Acetonitrile and thrombin were purchased from Merck (Darmstadt, Germany). Hydrogen tetrachloroaurate (Ⅲ) tetrahydrate was obtained from Showa (Tokyo, Japan), while isopropyl-beta-D-thiogalacto-pyranoside (IPTG) was obtained from Fluka (Seelze, Germany). Luria Bertani (LB) was obtained from BD (Franklin Lakes, NJ, USA). Millexn GS (pore size: 0.22 μm) and Amicon Ultra-4 (cut-off mass: 3000 Da) filters were purchased from Millipore Filter (Ireland).

NC-based isolation of GST and GST-tagged proteins from complex samples

Au@GSH NCs were generated by one-pot synthesis. The Au@GSH NCs used were generated by stirring GSH (2.5 mM, 1 mL) and $HAuCl_4$ (2.5 mM, 1 mL) at 32 ^oC for 36 h under room lighting, followed by centrifugation at 13000 rpm (rotor radius : 8.3 cm) (4 ^oC) for 30 min to remove large particles. The supernatant was then filtered by Amicon Ultra-4 filters under centrifugation at 1500 rpm for 1 h to remove non-reacted species. The nanoclusters remaining on the filter were then re-suspended in deionized water (1 mL). GST (originated from *Schistosoma japonicum*) and recombinant-GST fusion proteins (GST-$ClfA_{221-550}$), cell lysate containing GST-$ClfA_{221-550}$, and GST were used as the model samples for examination. *E. coli* competent cells (DH5α), was used as the control sample. The amounts of the total

protein in cell lysates were estimated using Bradford assay. Prior to starting the experiments, the concentrations of the total proteins were adjusted to 0.1 mg/mL by phosphate buffer solution (PBS, 10 mM, pH 7.3). The Au@GSH NCs (25 μL) were vortex-mixed with protein samples (0.5-mL) prepared in PBS buffer (10 mM, pH 7.3) for 15 min followed by standing in a refrigerator at 4 ^{o}C for 1 h. The samples were centrifuged at 13000 rpm (rotor radius: 8.3 cm) for 30 min. The sample vials were then examined under irradiation of a UV light (365 nm) with the naked eye. Fluorescence spectroscopy was employed for quantitative analysis. After removing the supernatant, the remaining conjugates were rinsed with PBS containing 1% Triton X-100 (100 μL × 3) and PBS (100 μL× 3) and then re-suspended in deionized water (10 μL). MALDI MS analysis was conducted to characterize the proteins in the conjugates. Sinapinic acid (15 mg/mL, 1 μL) prepared in acetonitrile/deionized water (2/1, v/v) containing 0.1% trifluoroacetic acid (TFA) was mixed with 1 μL of the suspension.

Characterization of the proteins trapped by the NCs

Proteomics strategies were employed for further characterization. Scheme S1 illustrates the steps of the analysis. In this strategy, on-NC-protein-tryptic digestion under microwave heating (power: 900 W, 40 s × 2), followed by MALDI-MS and MALDI MS/MS analysis, was conducted. The MS results, combined with protein database searching, were employed to characterize the protein identities. The suspension obtained above 2.5 μL was mixed with ammonium bicarbonate (50 mM, 2.5 μL) and trypsin (3 mg/mL, 2.0 μL) briefly before placing it in a domestic microwave oven to undergo microwave-assisted tryptic digestion. The resulting product (1 μL) was then mixed with CHCA (15 mg/mL, 1 μL) prepared in acetonitrile/deionized water (2/1, v/v) containing 0.1% trifluoroacetic acid (TFA) for MALDI MS and MS/MS analysis.

Database Searches

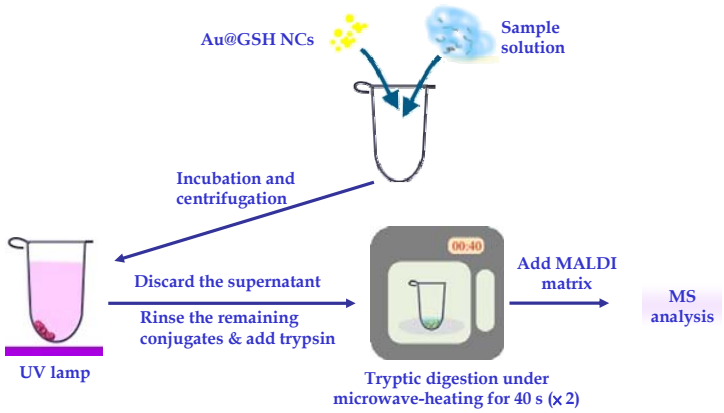
Fragment peaks (S/N > 3) resulting from precursor ions (S/N > 3) were submitted via Biotools (v. 3.0) to MASCOT (www.matrixscience.com) using the following search parameters: the database searched was NCBInr; taxonomy was set to all entries; the enzyme was trypsin; MS and MS/MS tolerances were set at ±0.3 and ±0.5 Da, respectively; the number of missed cleavages was set to 3. Methionine oxidation was set as variable modifications. MS/MS spectrum resulting from a

precursor ion was searched in the protein database search one at a time.

Instrumentation

Cell lysates were obtained by destroying cells using Misonix XL2000 ultrasonic homogenizer (Farmingdale, NY, USA). Absorption spectra were obtained from a Varian Cary50 UV-visible spectrophotometer (Melbourne, Australia), and fluorescence spectra were obtained from Horiba Jobin Yvon Fluomax-3 spectrofluorometer (Edison, NJ, USA). XPS data was obtained from an ESCA Phi 1600. TEM images were obtained from JEOL JEM-2100 high-resolution transmission electron microscope (Tokyo, Japan). A Bruker Daltonics Autoflex III (Bremen, Germany) TOF/TOF MS equipped with a 355-nm Nd:YAG laser was also employed for obtaining MS spectra. The following voltage parameters were employed: ion source 1, 19.06 kV; ion source 2, 16.61 kV; lens, 8.78 kV; reflector 1, 21.08 kV; reflector 2, 9.73 kV. The laser power was set to 50% with a frequency of 10 Hz, while each mass spectrum was collected from 500 laser shots. When MS/MS analyses were performed, the following voltage settings were used: ion source 1, 6.01 kV; ion source 2, 5.30 kV; lens, 3.01 kV; reflector 1, 27.10 kV; reflector 2, 11.74 kV; lift 1, 19.16 kV; lift 2, 4.42 kV. The laser power was set to 50% with a shot frequency of 10 Hz. Parent ions were obtained by collecting the mass spectra from 500 laser shots (50% power intensity) at a frequency of 20 Hz; daughter ions were obtained by collecting the mass spectra from 1000 laser shots (65% power intensity) at a frequency of 10 Hz.

Scheme S1. Cartoon representation of the procedures of using Au@GSH NCs as affinity probes to enrich their target species from protein samples. .



Figures S1a and b present the absorption spectrum and fluorescence spectrum of Au@GSH NCs, respectively. There is no plasmon absorption band in the visible wavelength region, but a broad band appears in the ultraviolet (UV) wavelength range (see Figure S1a). Furthermore, there is strong emission band appearing at $\lambda_{maximum}$~ 610 nm as shown in Figure S1b. Presumably, both the broad UV absorption and the strong emission result from S-Au charge transfer transitions of the Au@GSH NCs.[7] The formation of the Au(I)-ligand complexes also result in the appearance of liumescence of the as-prepared Au NCs

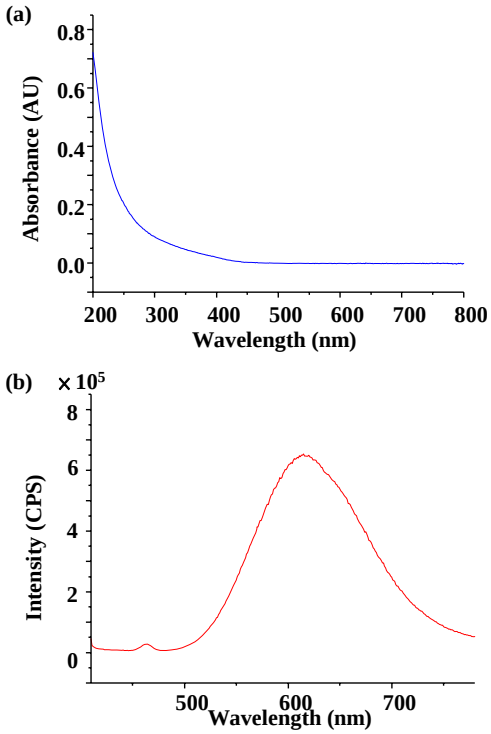


Figure S1. (a) Absorption spectrum and (b) fluorescence spectrum ($\lambda_{excitation}$= 400 nm) of Au@GSH NCs.

In general, fully reduced gold nanoclusters with particle sizes larger than 2 nm are difficult to be excited for fluorescence. However, it has been found that Au(I)-ligand complexes could give bright luminescence.[7] We employed X-ray photoelectron spectroscopy (XPS) to determine the oxidation state of the as-prepared Au@GSH NCs. The Au $4f_{7/2}$ binding energies for Au(0) and Au(I) are 84.0 eV (for ~2 nm fully reduced Au(0) nanoparticles) and 84.9~85.3 eV(AuS complexes),[7a] respectively. The Au $4f_{7/2}$ spectrum of the as-prepared Au@GSH NCs indicates the binding energy is >84.0 eV (see Figure S2), which suggests that both Au (0) and Au(I) exist in the NCs and therefore implies the possibility of the presence of Au-S

complexes by forming charge transfer bands. As a consequence, luminescent Au NCs are obtained.

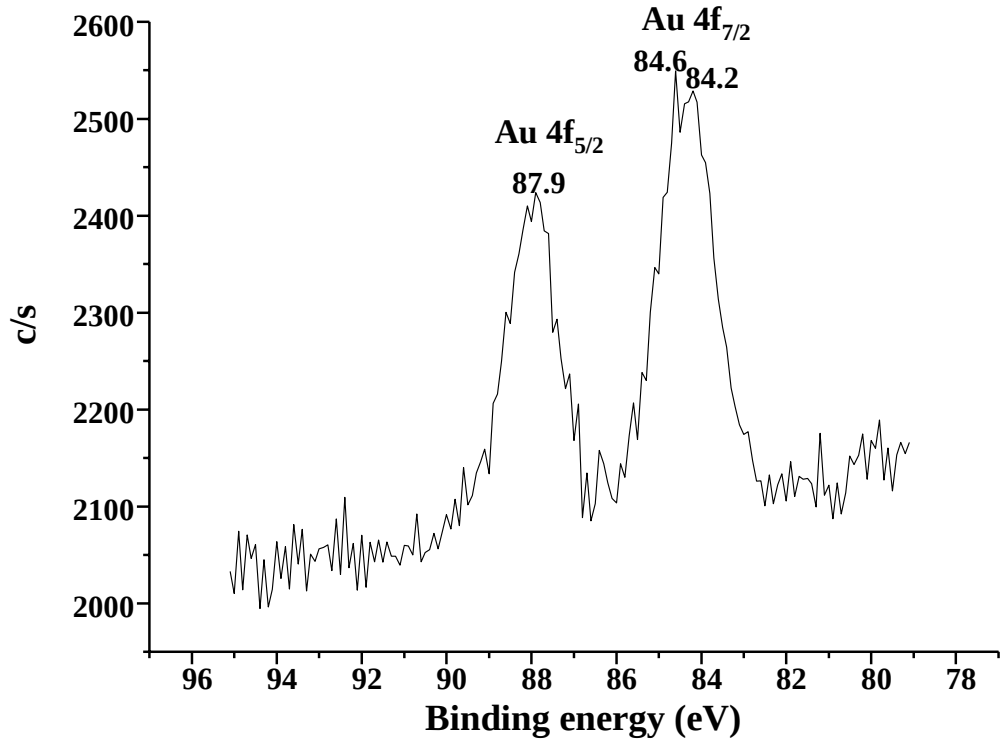


Figure S2. XPS spectrum of Au 4f for Au@GSH NCs (C 1s binding energy (285.3 eV) was used as an internal reference).

Figure S3a and b display the MALDI mass spectra obtained from the Au@GSH-target species conjugates from tubes a and b in Figure 4. The peaks at m/z~27 k appearing in panel a and 63 k appearing in panel b were presumably derived from GST and GST $ClfA_{221-550}$, respectively. To further confirm the identity of these proteins in Figure S3, we employed proteomics strategies to characterize the proteins trapped by the Au@GSH NCs in tubes a and b. Tryptic digestion and MALDI MS/MS analysis combined with protein database searches were employed for characterization of the proteins trapped by the NCs. The isolated Au@GSH-target species conjugates were first re-suspended in deionized water (10 μL). The suspension (2.5 μL) was mixed with ammonium bicarbonate (2.5 μL, 50 mM), and then trypsin was added for enzymatic digestion under microwave heating (power: 900 W) for 40 sec. After microwave heating, the sample was allowed to stand at room temperature for 5 min. The sample was then introduced into a microwave oven and heated for another 40 sec. The resultant products were analyzed by MALDI MS.

Figure S4a displays the MALDI mass spectrum of the tryptic digestion of the conjugates obtained from tube a. The peak profile was confirmed as GST with a score of 99 based on the peptide mapping fingerprinting (PMF) searching result. Nevertheless, to further identify each peak in Figure S4a, MALDI MS/MS was employed for characterization. These peaks were identified as the peptide residues from GST by combining these MS results with protein database searching. Table S1 lists the whole sequences of GST and GSTClfA$_{221\text{-}550}$, while Table S2 lists the corresponding amino acid sequence of each peak observed in Figure S4a. Figure S4b displays the MALDI mass spectrum of the tryptic digestion of the conjugate obtained from tube b. The peaks at m/z 1586.90 and 1638.97 appear in the mass spectrum. To further confirm the identity of these two peaks, these peaks were selected for MALDI MS/MS analysis. They were confirmed as the peptide residues from GST-ClfA$_{221\text{-}550}$ after protein database searching. The rest of the peaks were also characterized with MALDI MS/MS analysis combined with protein database searches. Table S3 lists the corresponding amino acid sequences of those peaks appearing in Figure S4b. These results indicate the Au@GSH NCs have the capability of selectively trapping proteins tagged with GST from complex samples such as cell lysates. This approach suggests an extremely simple and straightforward means to rapidly clarify the presence and the identity of GST fusion proteins.

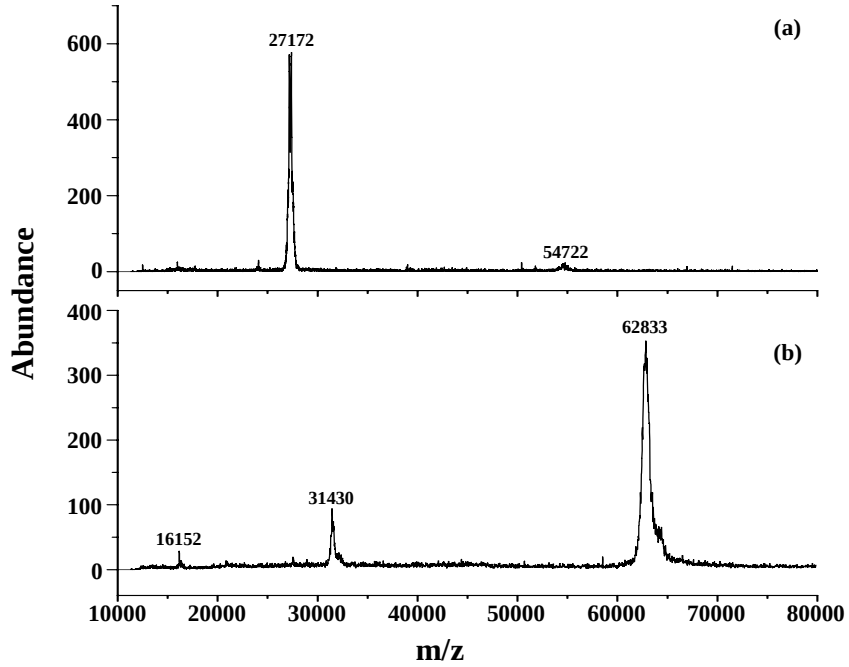


Figure S3. MALDI mass spectra of the samples obtained from the NC-target species isolated from the cell lysates containing (a) GST and (b) GST-ClfA$_{221-550}$.

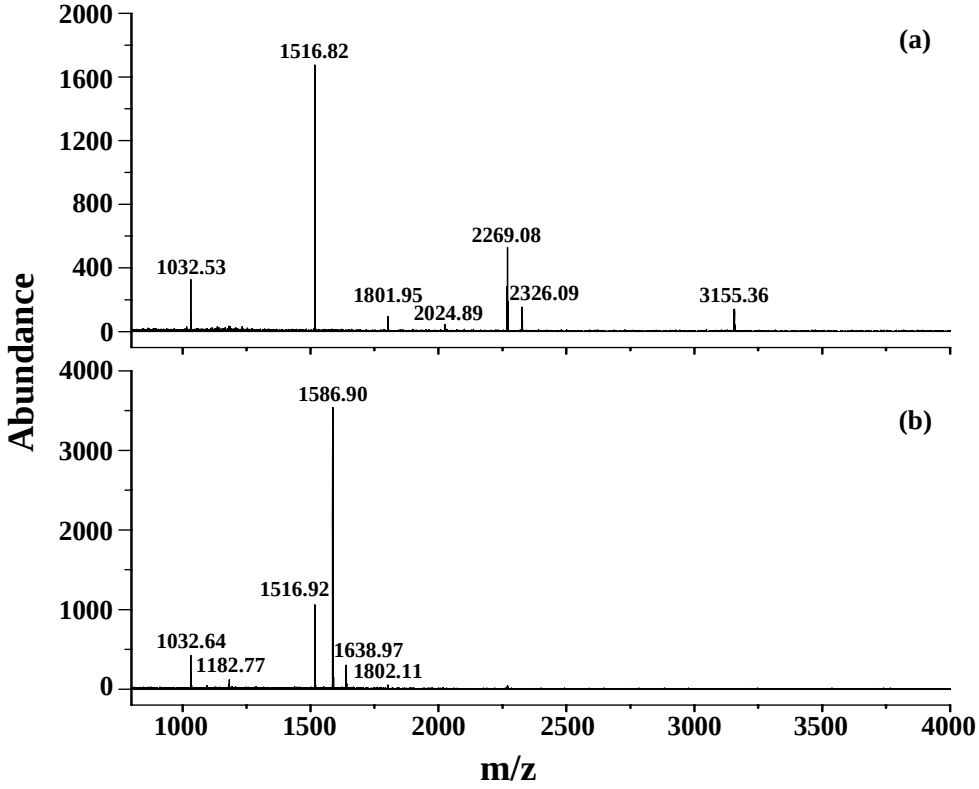


Figure S4. MALDI mass spectra of the samples obtained from the NC-target species isolated from the cell lysates containing (a) GST and (b) GST-ClfA$_{221-550}$ followed by tryptic digestions under microwave heating for 40 s (× 2).

Table S1. Complete sequences of GST and GST-ClfA$_{221-550}$

Protein	Sequence*
GST (*Schistosoma japonicum*)	MSPILGYWKI KGLVQPTRLL LEYPQEKYEE HLYERDEGDK WRNKKFELGL EFPNLPYYID GDVKLTQSMA IIRYIADKHN MLGGCPKERA EISMLEGAVL DIRYGVSRIA YSKDFETLKV DFLSKLPEML KMFEDRLCHK TYLNGDHVTH PDFMLYDALD VVLYMDPMCL DAFPKLVCFK KRIEAIPQID KYLKSSKYIA WPLQGWQATF GGGDHPPKSD LVPRGSPGIH RD
GST-ClfA$_{221-550}$	MSPILGYWKI KGLVQPTRLLL EYPQEKYEE HLYERDEGDK WRNKKFELGL EFPNLPYYID GDVKLTQSMA IIRYIADKHN MLGGCPKERA EISMLEGAVL DIRYGVSRIA YSKDFETLKV DFLSKLPEML KMFEDRLCHK TYLNGDHVTH PDFMLYDALD VVLYMDPMCL DAFPKLVCFK KRIEAIPQID KYLKSSKYIA WPLQGWQATF GGGDHPPKSD LVPRGS**VAAD APAAGTDITN QLTNVTVGID SGTTVYPHQA GYVKLNYGFS VPNSAVKGDT FKITVPKELN LNGVTSTAKV PPIMAGDQVL ANGVIDSDGN VIYTFTDYVN TKDDVKATLT MPAYINPENV KKTGNVTLAT GIGSTTANKT VLVDYEKYGK FYNLSIKGTI DQIDKTNNTY RQTIYVNPSG DNVIAPVLTG NLKPNTDSNA LIDQQNTSIK VYKVDNAADL SESYFVNPEN FEDVTNSVNI TFPNPNQYKV EFNTPDDQIT TPYIVVVNGH IDPNSKGDLA LRSTLYGYNS NIIWRSMSWD NEVAFNNGSG SGDGIDKPVV PEQPDE**

Table S2. **Corresponding sequences of the peptide ion peaks observed in Figure S4a obtained by combining MALDI MS/MS results with protein database searching.**

Observed $[M+H]^+$	Theoretical $[M+H]^+$	Sequences	Matching protein
1032.53	1032.58	LTQSMAIIR	Glutathione S-transferase (65-73)
1516.82	1516.80	AEISMLEGAVLDIR	Glutathione S-transferase (90-103)
1801.95	1801.94	ERAEISMLEGAVLDIR	Glutathione S-transferase (88-103)
2024.89	2024.90	YEEHLYERDEGDKWR	Glutathione S-transferase (28-42)
2269.08	2269.13	LLLEYLEEKYEEHLYER	Glutathione S-transferase (19-35)
2326.09	2326.13	YIAWPLQGWQATFGGGDHPPK	Glutathione S-transferase (198-218)
3155.36	3155.52	LLLEYLEEKYEEHLYERDEGDKWR	Glutathione S-transferase (19-42)

Table S3. **Corresponding sequences of the peptide ion peaks observed in Figure S4b obtained by combining MALDI MS/MS results with protein database searching.**

Observed $[M+H]^+$	Theoretical $[M+H]^+$	Sequences	Matching protein
1032.64	1032.58	LTQSMAIIR	Glutathione S-transferase (65-73)
1182.77	1182.68	RIEAIPQIDK	Glutathione S-transferase (182-191)
1516.92	1516.80	AEISMLEGAVLDIR	Glutathione S-transferase (90-103)
1586.90	1586.79	STLYGYNSNIIWR	Clumping factor A from S.A (306-318)
1638.97	1638.80	GTIDQIDKTNNTYR	Clumping factor A from S.A (181-194)
1802.11	1801.94	ERAEISMLEGAVLDIR	Glutathione S-transferase (88-103)