Electronic Supporting Information for:

Machine Learning-Assisted Surface-Enhanced Raman Spectroscopy for the Rapid Determination of the Glutathione Redox Ratio

Wilson A. Garuba,[a] *Brian A. Barth,*[b] *Adam E. Imel,*[c] *and Bhavya Sharma*[a*]

[a]Dept. of Chemistry, University of Tennessee, Knoxville, TN 37996, USA

[b]Dept. of Chemical and Biomolecular Engineering, University of Tennessee, Knoxville, TN 37996 USA

[c]University of Tennessee – Oak Ridge Innovation Institute, Knoxville, TN 37996 USA

*Corresponding author: bhavya.sharma@utk.edu

Table of Contents

I. Nanoparticle Characterization

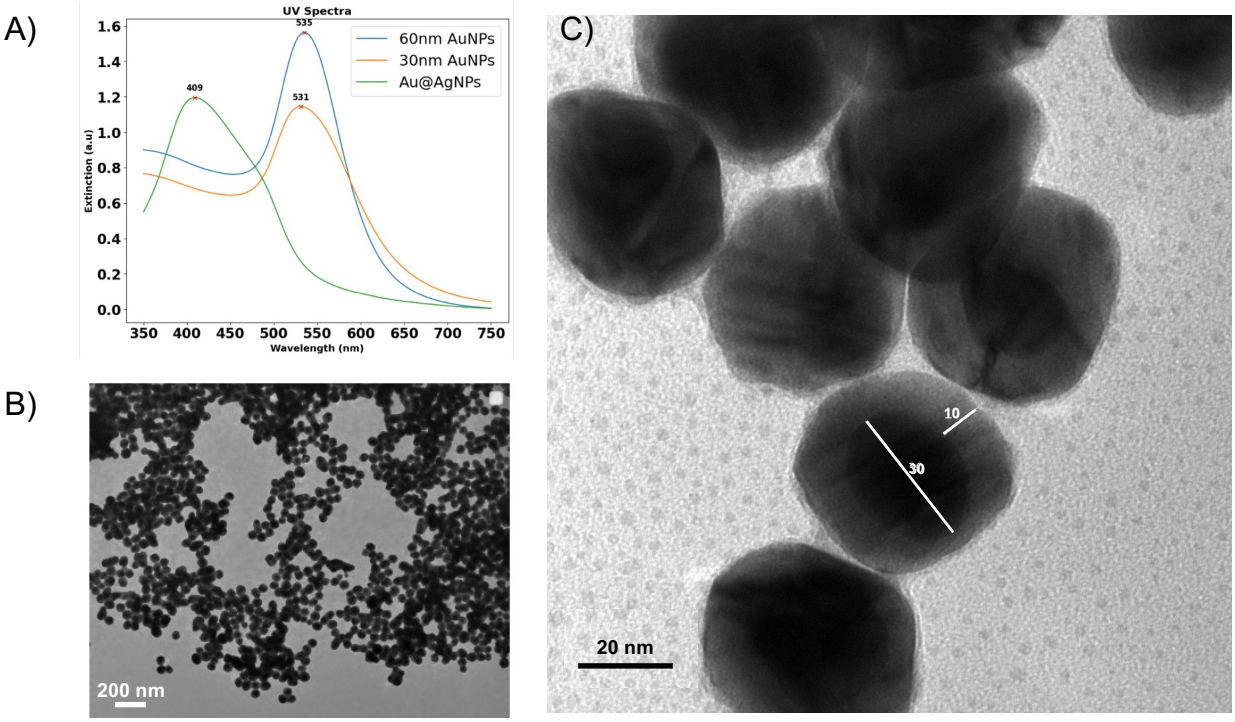

Figure ESI1. Characterization of the SERS substrate. (A) UV-vis extinction (LSPR) spectra of 60 nm (blue), 30 nm (orange) AuNPs, and the Au@AgNPs (green). (B) TEM image of Au@AgNPs. (C) Higher magnification TEM image where the silver shell (10 nm) is seen surrounding the gold core nanoparticle (diameter = 30 nm).

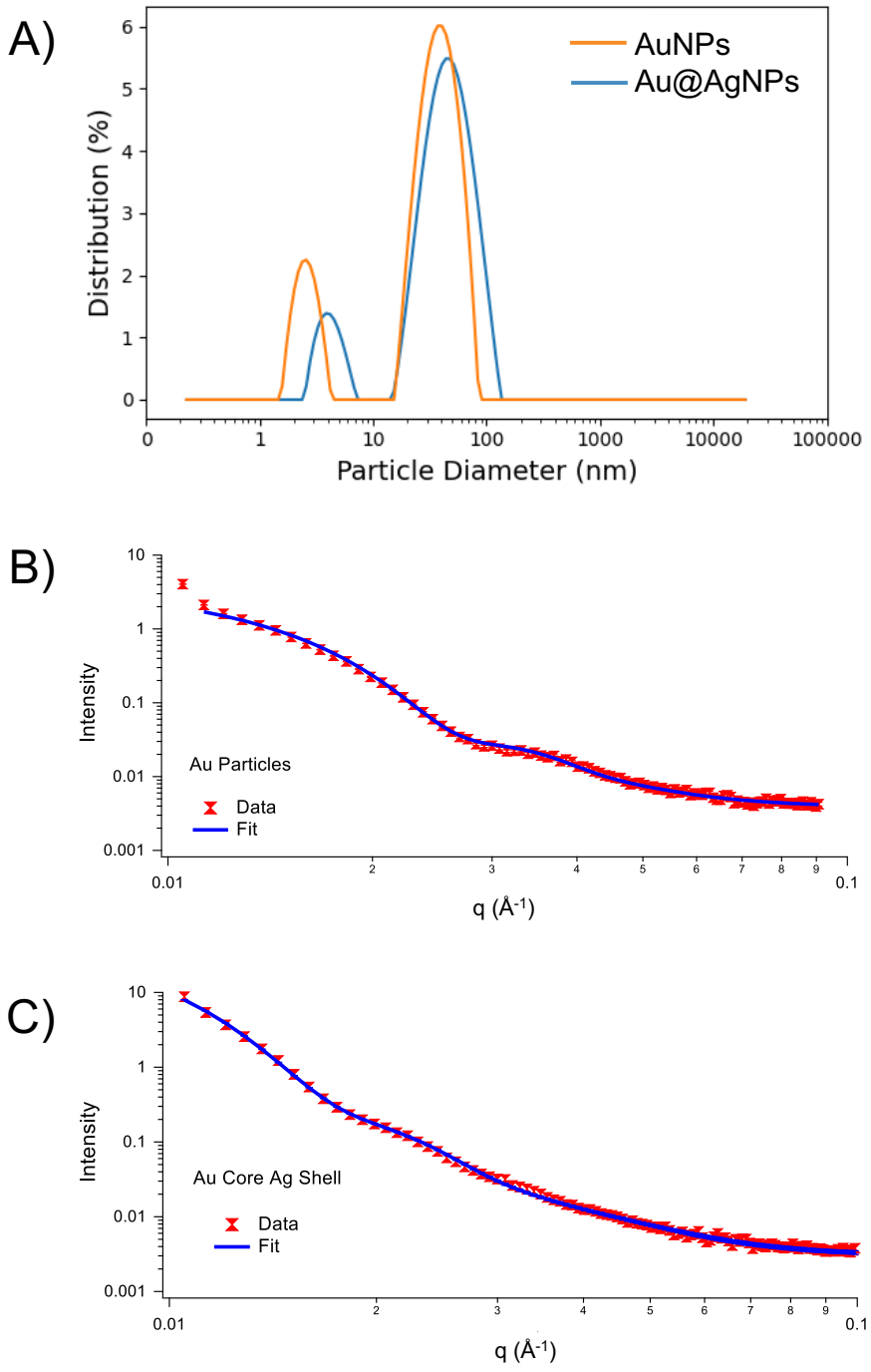

Figure ESI2. Characterization of Au and Au@Ag nanoparticles using dynamic light scattering (DLS), zeta potential measurements, and small-angle X-ray scattering (SAXS). (A) DLS-derived size distributions for Au nanoparticles (orange) and Au@Ag nanoparticles (blue) and, highlighting differences in hydrodynamic diameter and size distribution. SAXS intensity profile (red crosses) and corresponding fit (blue line) for (B) Au nanoparticles and (C) Au-core/Ag-shell nanoparticles, demonstrating scattering behavior.

II. Dimensionality Reduction

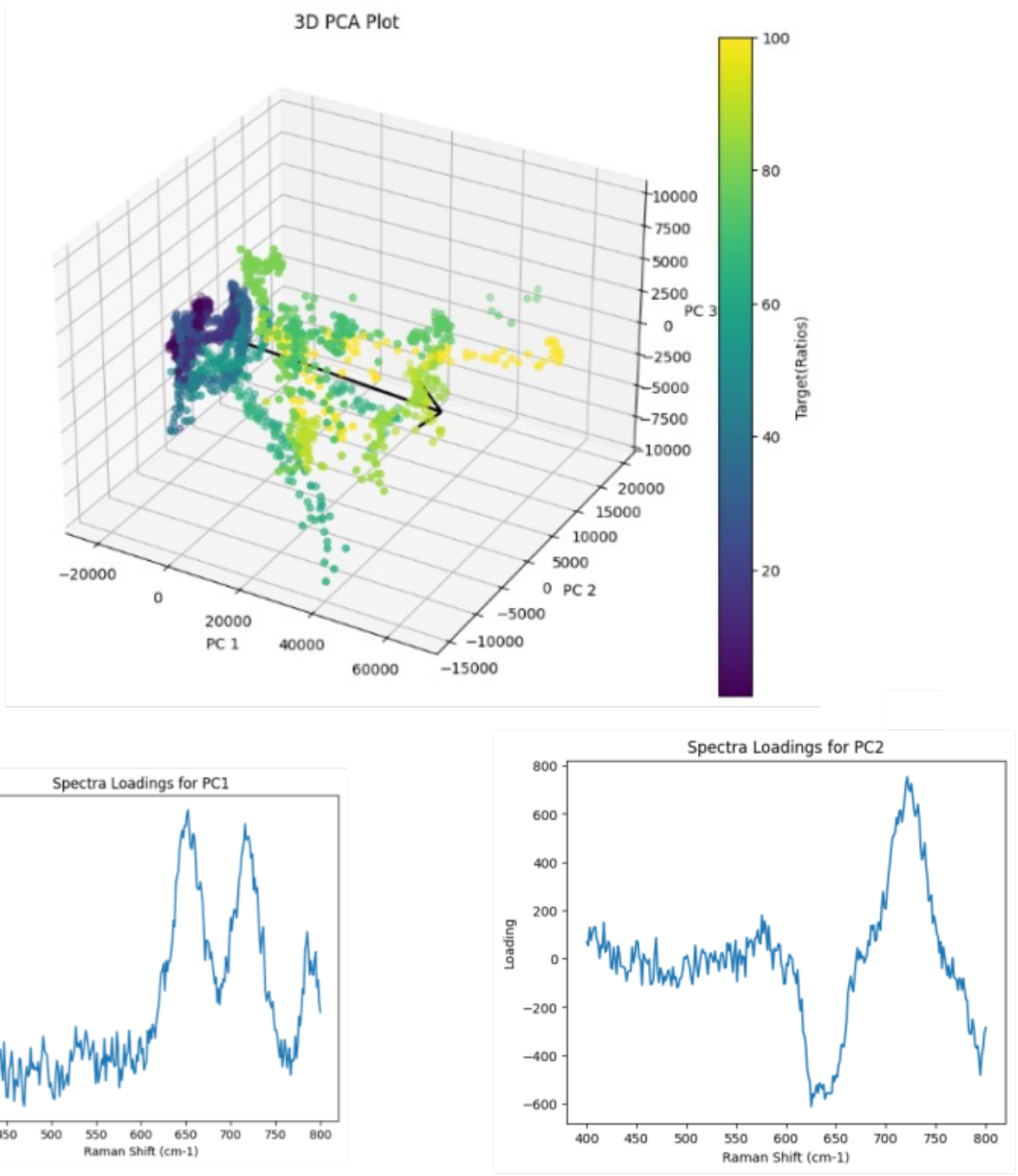

Figure ESI3. Principal Component Analysis (PCA) visualization plot for the entire dataset. The black arrow indicates the direction of increasing data values along the first principal component (PC1), which exhibits a slightly non-linear relationship. The loadings plot for PC1 confirms that the observed increase is primarily attributed to the C-S stretching vibrations in GSH.

III. Normalization Comparison: SNV vs Max-Min

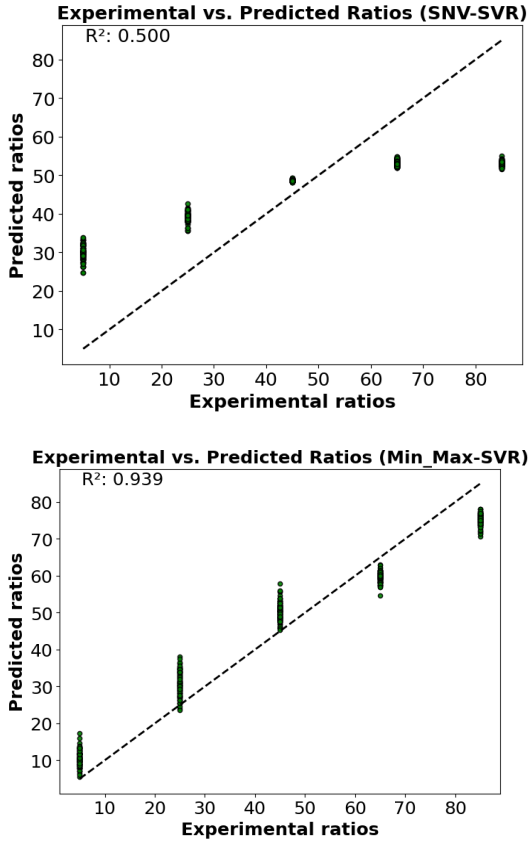

Figure ESI4. Comparison of experimental and predicted ratios for the standard normal variate (SNV) and max-min normalization methods combined with the support vector regression (SVR) model. (A) SNV scales the spectra with a mean of 0 and a standard deviation of 1. When combined with the SVR model, there is a low R^2 value (0.500) and high RMSE value (20.005) indicating poor model performance. (B) Max-min normalization scales the data to between 0 and 1. The plot shows the correlation between experimental and predicted values, with a high R^2 value (0.939) and low RMSE (6.970) indicating robust model performance.