Supporting Information: Data-Driven Exploration of Silver Nanoplate Formation in Multidimensional Chemical Design Spaces

Huat Thart Chiang, Kiran Vaddi, and Lilo Pozzo*

Department of Chemical Engineering, University of Washington, Seattle, WA 98105.

E-mail: dpozzo@uw.edu

NOTE: Equation or Figure references without a preceding "S" letter refer to the main manuscript

Fast Spectroscopic Exploration

Histograms of the concentrations of the reagents used to synthesize the silver nanoparticles are shown. These histograms support the data obtained from the 5d plots in Figure 3.

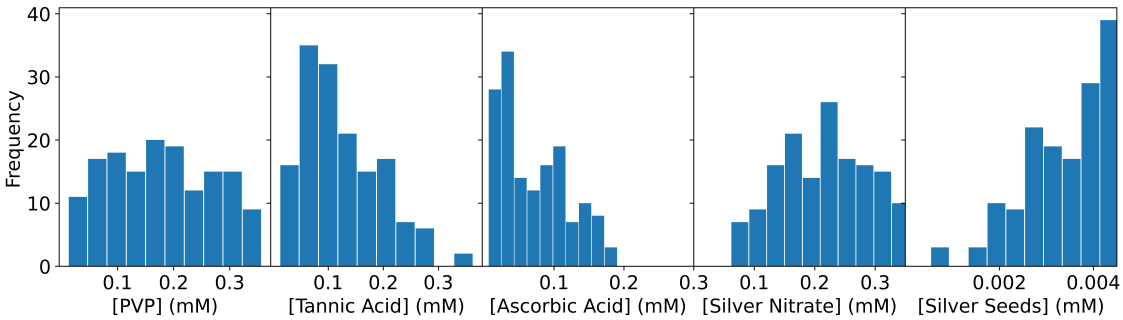

Figure S1: A histogram of the concentrations of the reagents that were used to synthesize samples that had a distance below the chosen threshold distance. These samples should have the characteristics of small, monodisperse, colloidally stable, plate-like nanoparticles.

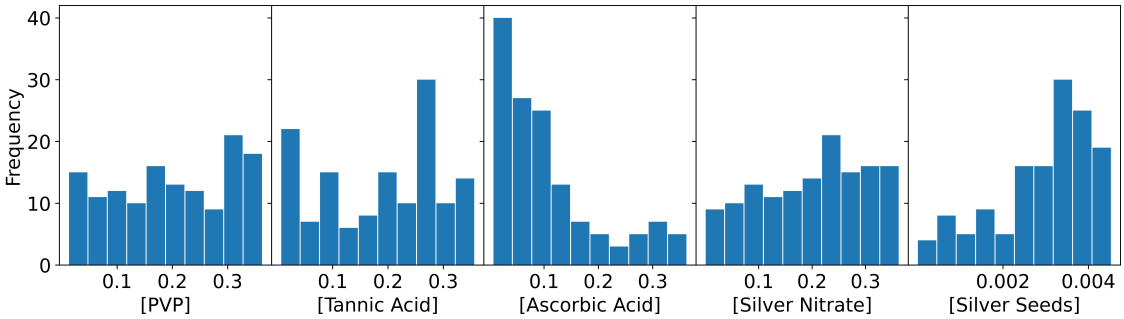

Figure S2: A histogram of the concentrations of the reagents that were used to synthesize the samples that had distances above the chosen threshold distance.

To test the reproducibility of the experiment, it was repeated with the same distance metric with the same distance threshold. The only difference was the randomly sampled coordinates from the first iteration. In Figure S3, the results are similar to that of Figure 3 in the main manuscript.

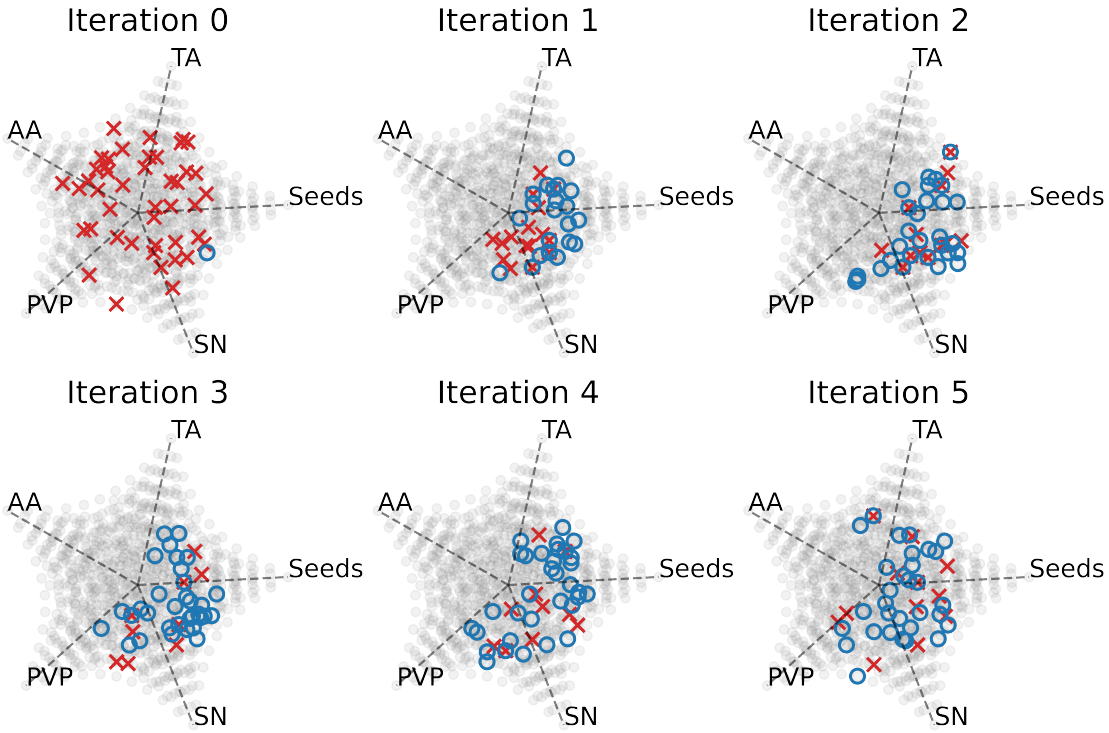

Figure S3: A repeated experiment showing the 5-dimensional plots for the volume fractions of each of the reagents in each iteration in the Fast Spectroscopic Exploration. This repetition of the experiment used a newly selected set of randomly sampled compositions for Iteration 0. The experiment shows good agreement with the one in Figure 3 in the main text. In iteration 0 there are 1/48 samples labeled "Below Threshold", 30/48 in iteration 1, 34/48 in iteration 2, 38/48 in iteration 3, 35/48 in iteration 4, and 35/48 in iteration 5.

The weight factors that were chosen for each of the terms of the distance metric (i.e., amplitude-phase, peak position, area under the curve, and peak intensity) affect the classification. These weights are adjustable and are a mechanism to tune the distance metric to bias the selection towards a certain particle structure. In this work, the weights were chosen to synthesize small, colloidally stable, monodisperse, plate-like particles, but they can be modified to include more of a sample of a specific characteristic (i.e., increasing the weight of the area under the curve term would increase the presence of monodisperse particles).To investigate the effect of each term on the classification, Figure S4 was created. This figure shows how the classification of the spectra would change if the weight factor of each term was set to zero. The spectra in gray are the original 137 spectra that had distances below the selected threshold distance using the weight factors described in the main manuscript. The spectra that had a distance above the threshold were not plotted because there are too many curves and plotting all of them in the same color would oversaturate the figure. The spectra in other colors are the spectra that would have a distance below the threshold if the weight factor of a term was set to zero. This means that the colored spectra were assigned high distance scores by the corresponding term of the distance metric. Figure S4 (A) shows that the amplitude-phase metric assigned high distance scores to spectra that had peaks and high intensities at low wavelengths. (B) shows that the peak position term filtered out spectra that had peaks at low wavelengths. (C) shows that the area under the curve term filtered out spectra that had broad peaks. (D) shows that the intensity below 450 nm term filtered out spectra that had high intensities in this wavelength range. For example, in Figure S4 C, the area under the curve term is set to zero, so we expect the colored curves to have large areas under the curve, which is exactly what is observed. In Figure S4 D, the intensity below 450 nm is set to zero, so we expect the colored curves to have a high intensity below 450 nm, which is what is observed. Therefore, Figure S4 shows that the weight of each term can be adjusted accordingly to exclude or include the colored spectra from the classification.

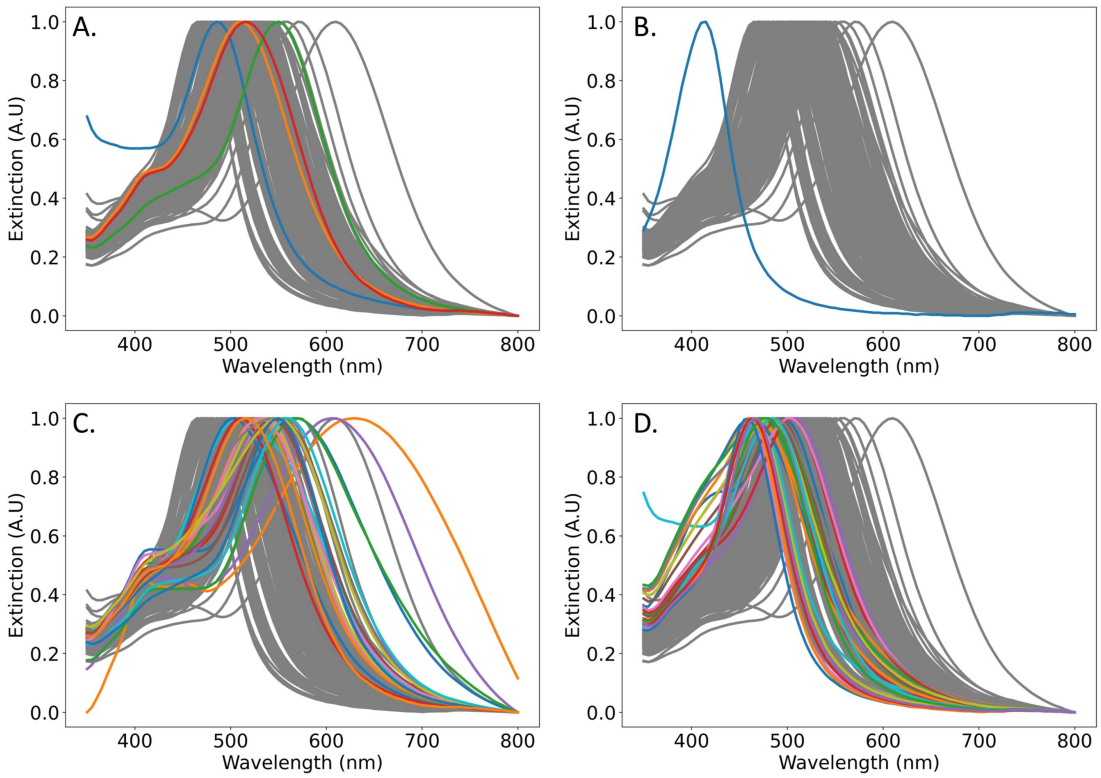

Figure S4: This figure shows the effect of each distance term on the classification of the spectra. The spectra in gray are the spectra that had distances below the selected threshold using the distance metric with the original weight factors. The spectra in other colors are the spectra that would have a distance below the selected threshold if the weight factor of the (A) amplitude-phase, (B) peak position, (C) the area under the curve, and (D) intensity below 450 nm, were set to zero.

Gaussian Process Classifier

A Gaussian process (GP) classifier was used to explore the nanoparticle morphologies in the experimental design space. A Gaussian process classifier is a supervised, non-parametric, probabilistic model that can be trained from data and perform classifications. It works by fitting linear combinations of kernels to data points. Kernels, a hyperparameter of a GP, are functions (e.g., gaussian, exponential, quadratic) used to measure the covariance or the similarity between two data points. The kernels are placed in arbitrary points in the parameter space, and the weighted sum of these kernels makes up an estimation function which is used to make predictions of the outcomes of the input parameter set. The weight coefficients are sampled from a Gaussian distribution, which results in a distribution of different estimator functions called Gaussian Process Priors. A loss function, composed of the sum of a similarity term and a regularization term, is minimized to obtain the best estimation function. The similarity term is composed of the difference between the actual training data points and the estimation function's predictions of the points the regularization term accounts for the smoothness of the function which is important to prevent overfitting. The loss function is minimized by changing the weight coefficients of the kernels and a parameter that affects the smoothness of the fit, obtaining an estimation function that best represents the data.[1]

Amplitude Phase Distance

The amplitude-phase distance measures the shape-similarity between two one-dimensional functions that map a domain x to a scalar value $f(x)$. A key component of this distance measurement is the process of *aligning* any query function to a target function against which we want to measure the similarity. The alignment can be performed using a variety of optimization processes presented in[2] and involves finding a domain warping function that results in the least **L**2 distance between the query and target function. Figure S5 depicts the alignment process with the target function selected as a 1D curve with a single peak (dotted black line). A set of query functions (grey color) are then aligned to the target function using their respective domain warping function (optimized individually). We can then compute an **L**2 distance between aligned query functions to measure the similarity on the y-scale of a 1D function–which we refer to as the amplitude distance. The alignment

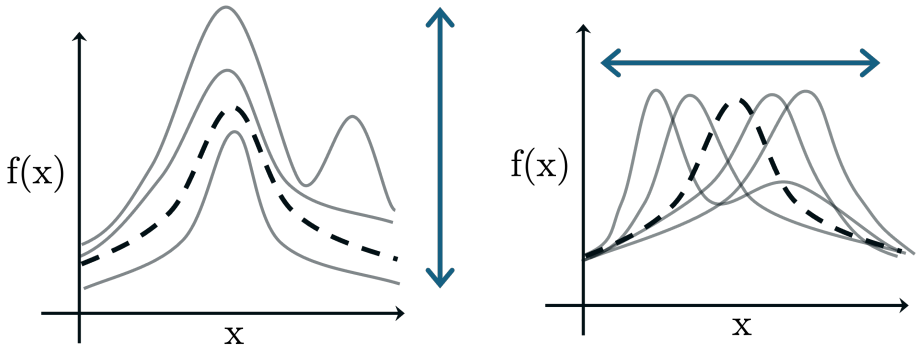

Figure S5: Amplitude-Phase distance measures shape-similarity using two components : (left) similarity in amplitude when all the functions have been 'aligned' such that the only variation between them is on the intensity scale; (right) phase component measures the 'distortion' added into any shape to align it to a reference (shown in dotted black line). Grey-colored solid lines are representative functions to depict the kind of variations each distance component would capture.

process itself has removed some of the dis-similarities between the query and target function that need to be accounted for to provide an accurate distance measure. Figure S6 depicts distortion imparted on the function shown on the left resulting in a new function (middle) and the comparison between the no-warping state (solid-blue) of the function with its warped

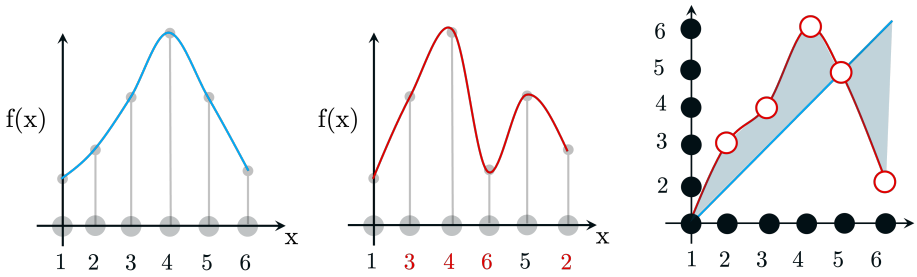

Figure S6: Computing distance between domain warping functions : (left) target curve; (middle) query curve; (right) domain warping functions to align the query and target functions and the distortion measured using the area between the no-warping (solid-blue) and optimized warping of query function (solid-red). The shaded region depicts the distortion added to the query function to align it to the target.

state (solid-red) in the right-most panel. To explain the warping conceptually, we identify each function with five discrete locations shown as grey circles on the x-axis and distort the left curve in Figure S6 by swapping intensity values with those identified in red font on the curve in the middle panel of Figure S6. In Figure S6, the left curve was distorted by the following swaps of the intensity values $(2,3),(3,4),(4,6),(6,2)$ where each tuple (a,b) represents the swapping of intensity value at a with that of b. This swapping (or warping when done continuously) can be represented as a map on a 2D plot as shown on the right-most panel of Figure S6 where each white circle represents a tuple identifying the swapping of intensities at those locations. If we do not swap any points, that would correspond to no-warping that can be represented as a line going through the origin at a 45° angle (solid-blue line in the right-most panel of Figure S6). The amount of distortion (i.e. the phase distance) added by the warping roughly corresponds to the area between these maps on the 2D plane shown on the right-most panel of Figure S6. Mathematically, the phase distance is computed using a Riemannian geometric method since the domain warping functions are optimized to be a class of functions that preserve the norm of original functions giving rise to a spherical geometry i.e. a curved manifold. We defer the readers interested in the detailed mathematical derivations of amplitude-phase distance to[2] and.[3]

SAXS Structural Exploration

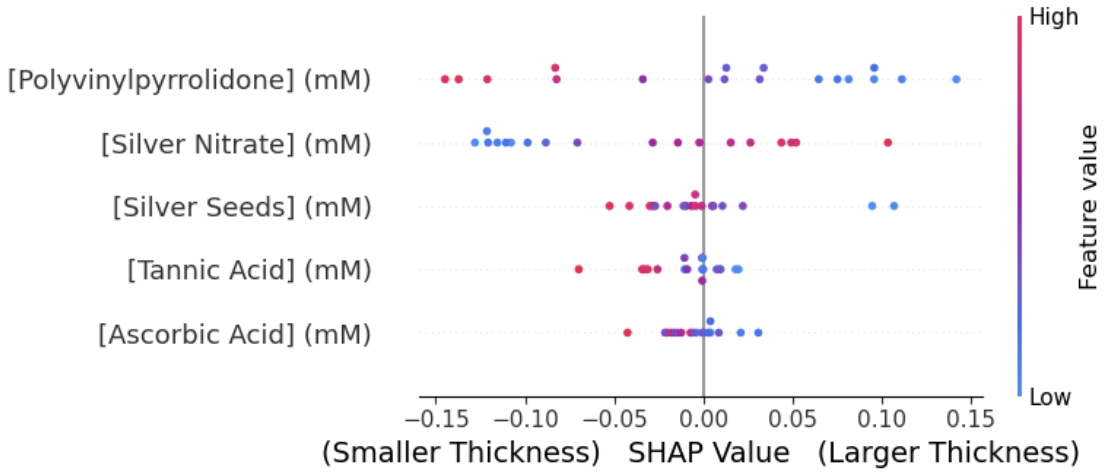

Figure S7: The SHAP summary plot of the effect of the reagents on plate thickness. The order of the reagents, from top to bottom, indicates the importance of them on the plate thickness.

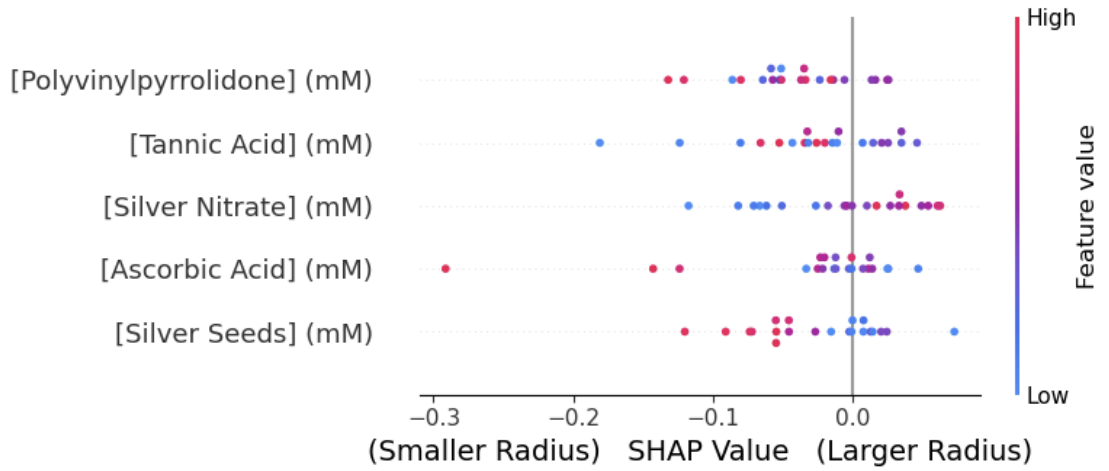

Figure S8: The SHAP summary plot of the effect of the reagents on plate radius. The order of the reagents, from top to bottom, indicates the importance of them on the plate radius.

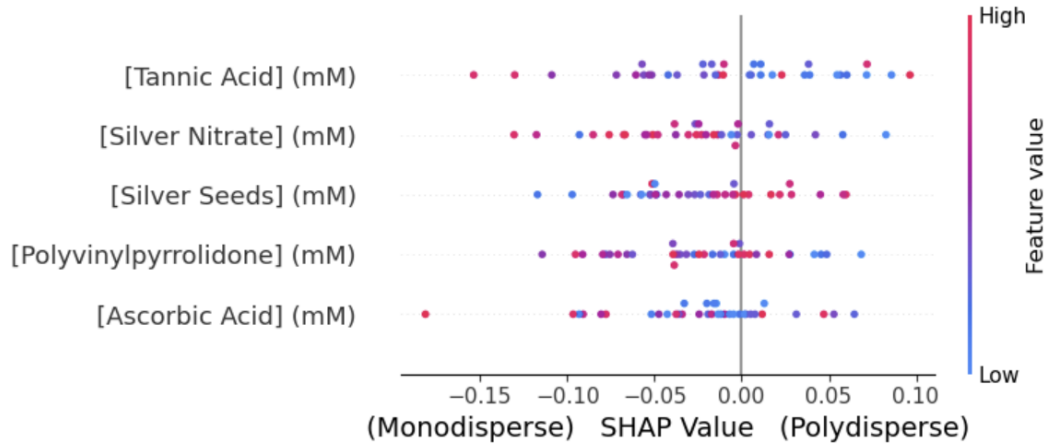

Figure S9: The SHAP summary plot of the effect of the reagents on polydispersity, which was determined by the model used to fit the SAXS data. The order of the reagents, from top to bottom, indicates the importance of them on the polydispersity.

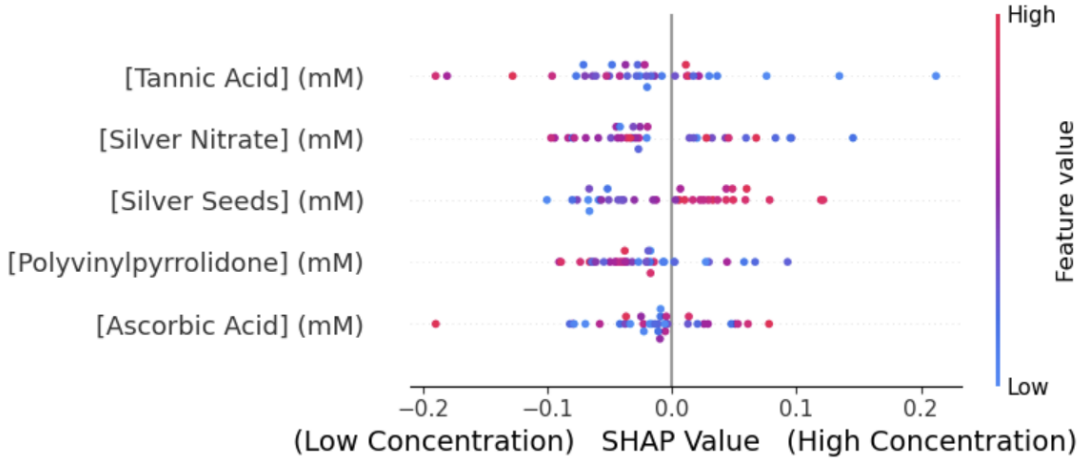

Figure S10: The SHAP summary plot of the effect of the reagents on the concentration of the particles, which was determined by the scale parameter of the model used to fit the SAXS data. The order of the reagents, from top to bottom, indicates the importance of them on the concentration.

TEM

Additional TEM images of the sample that was classified as "below threshold" are shown.

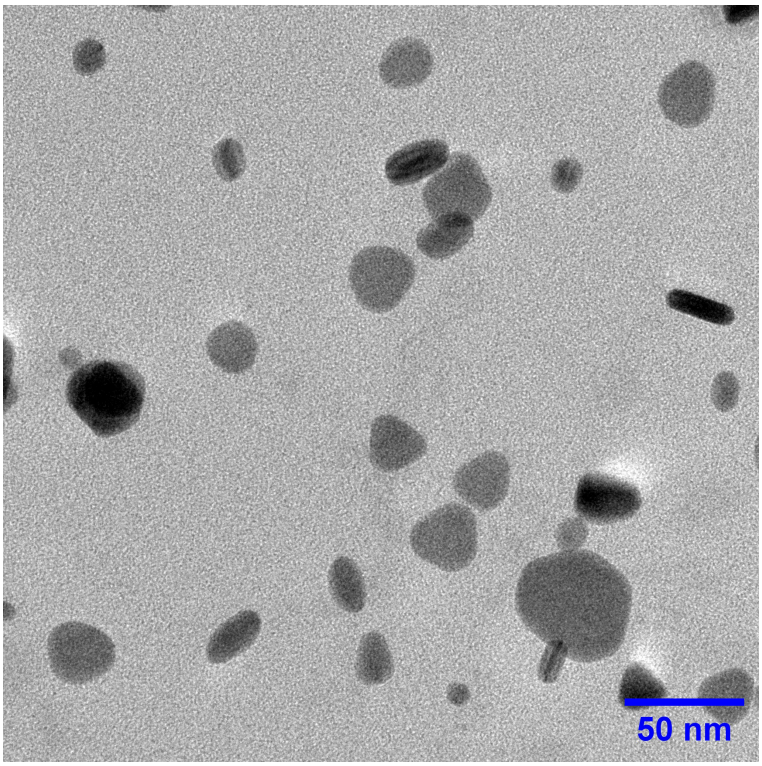

Figure S11: Additional image of the sample that was classified as "below threshold".

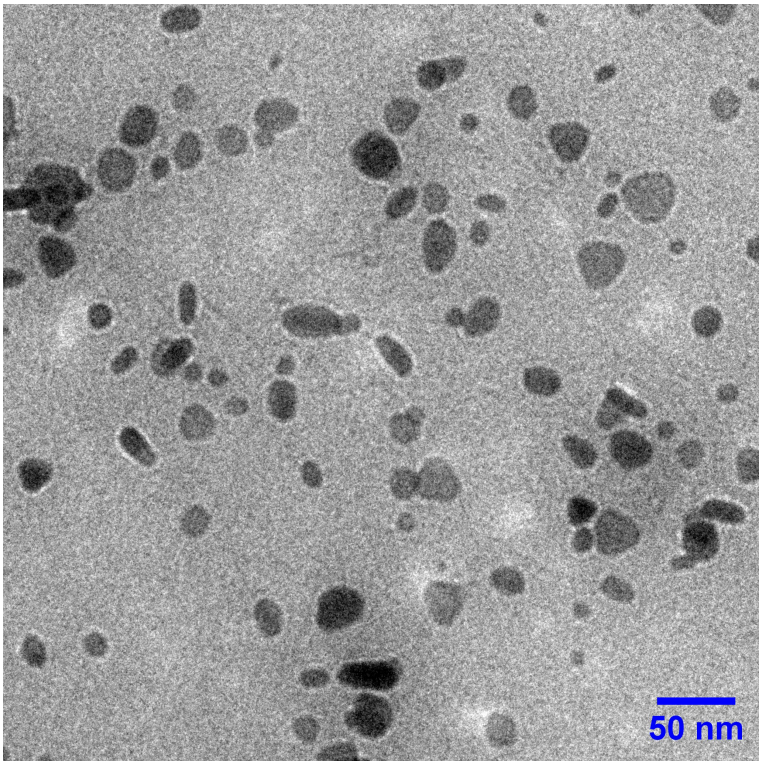

Figure S12: Additional image of the sample that was classified as "below threshold".

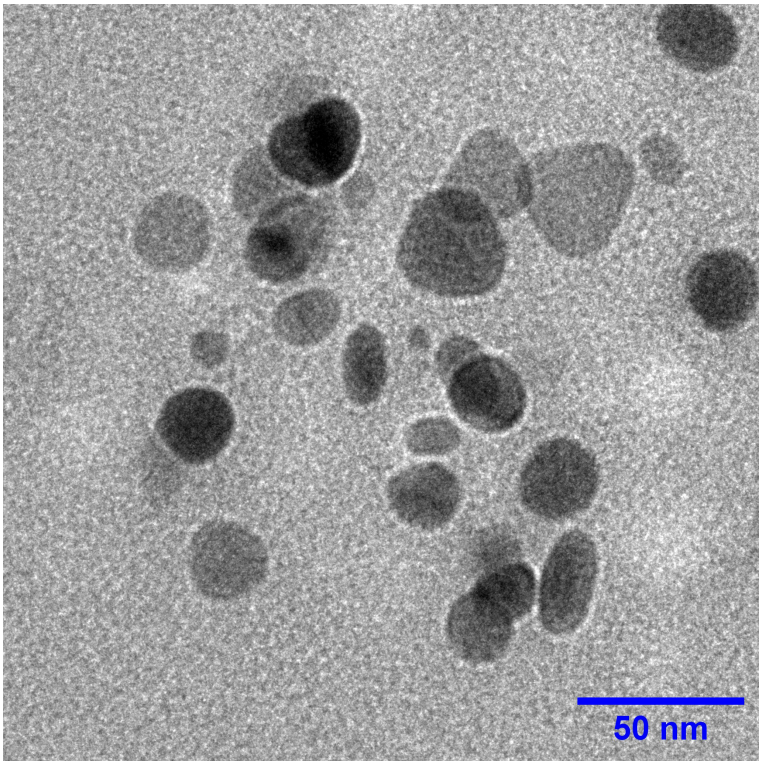

Figure S13: Additional image of the sample that was classified as "below threshold".

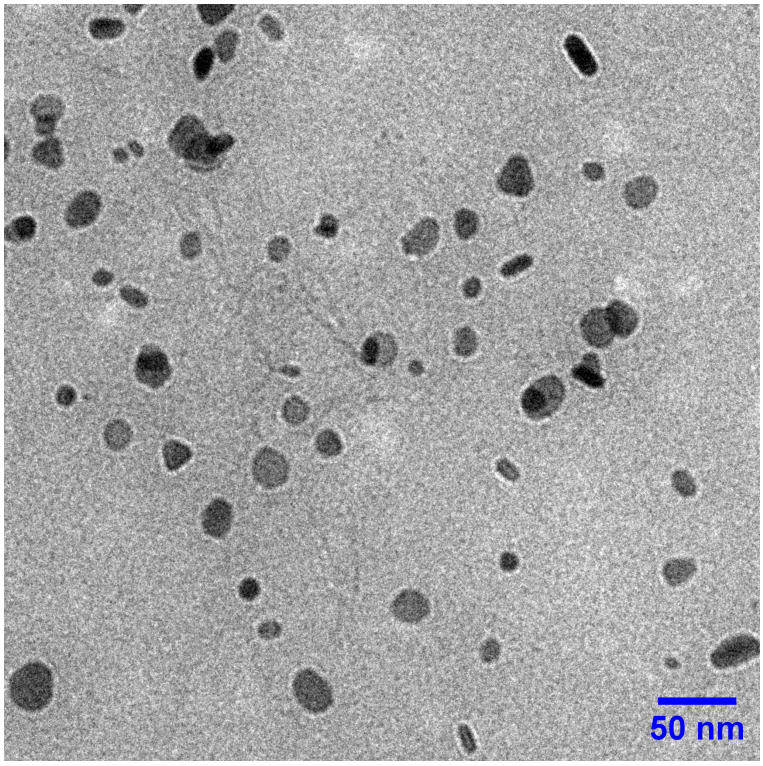

Figure S14: Additional image of the sample that was classified as "below threshold".

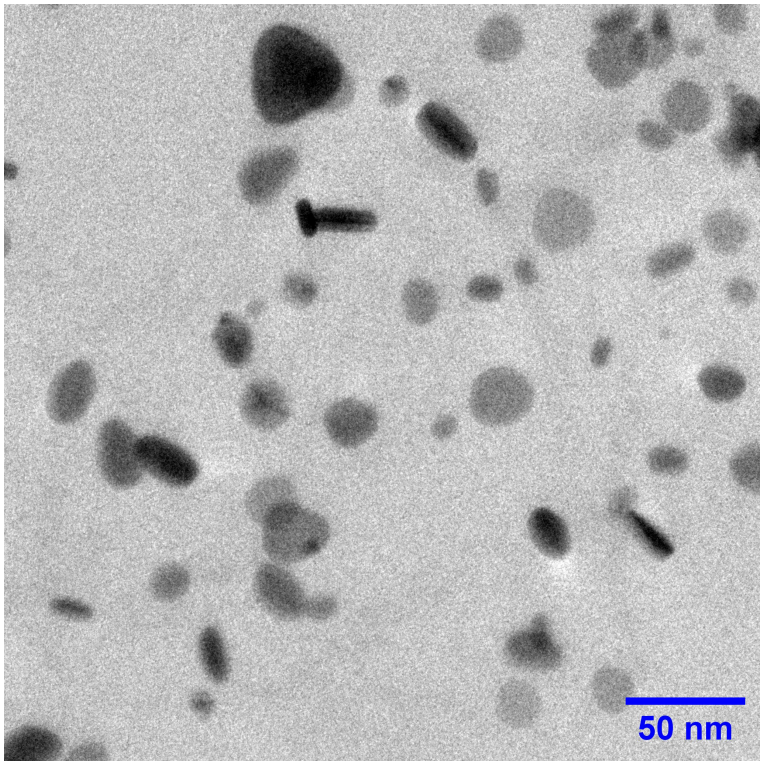

Figure S15: Additional image of the sample that was classified as "below threshold".

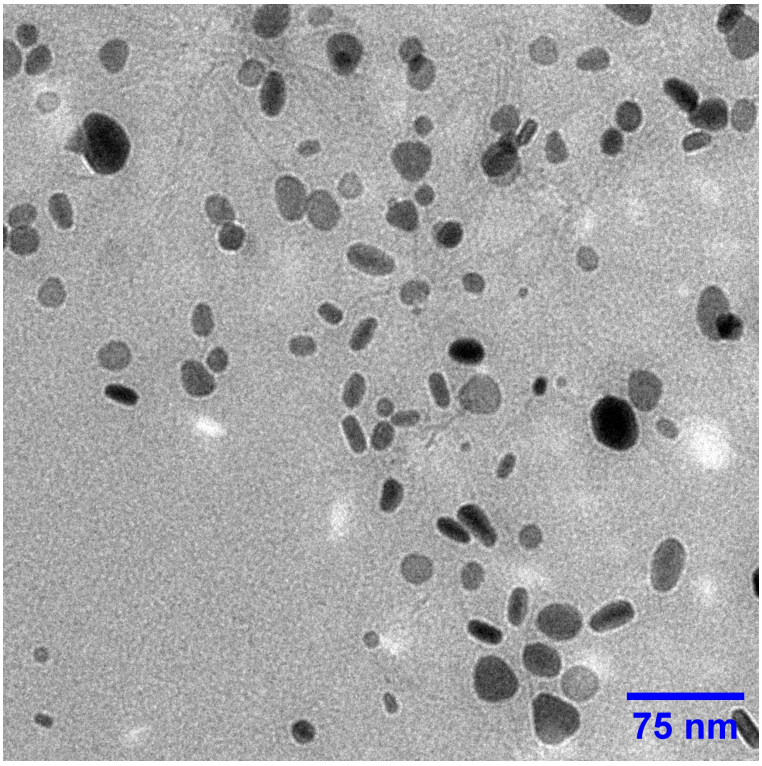

Figure S16: Additional image of the sample that was classified as "below threshold".

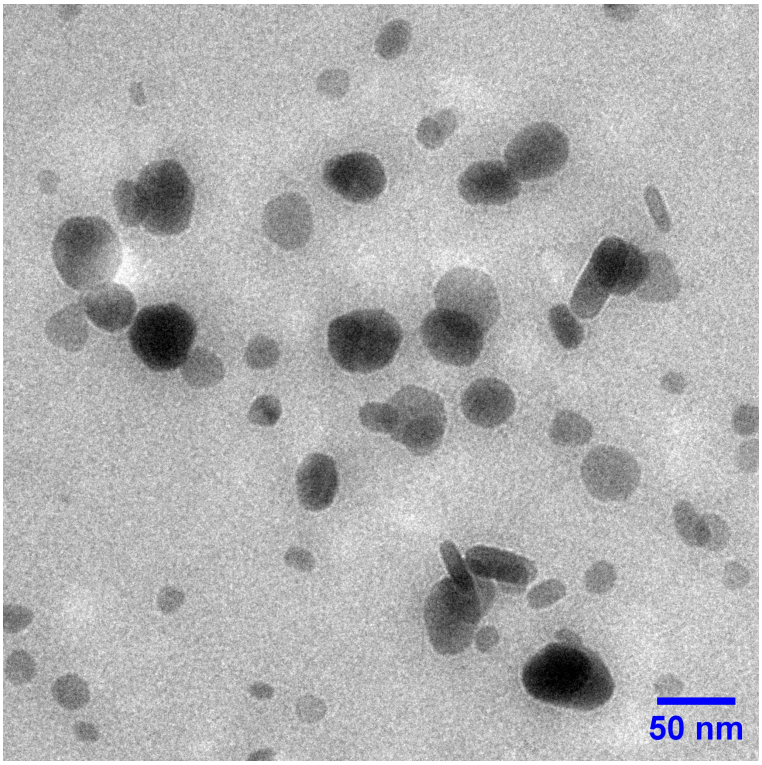

Figure S17: Additional image of the sample that was classified as "below threshold".

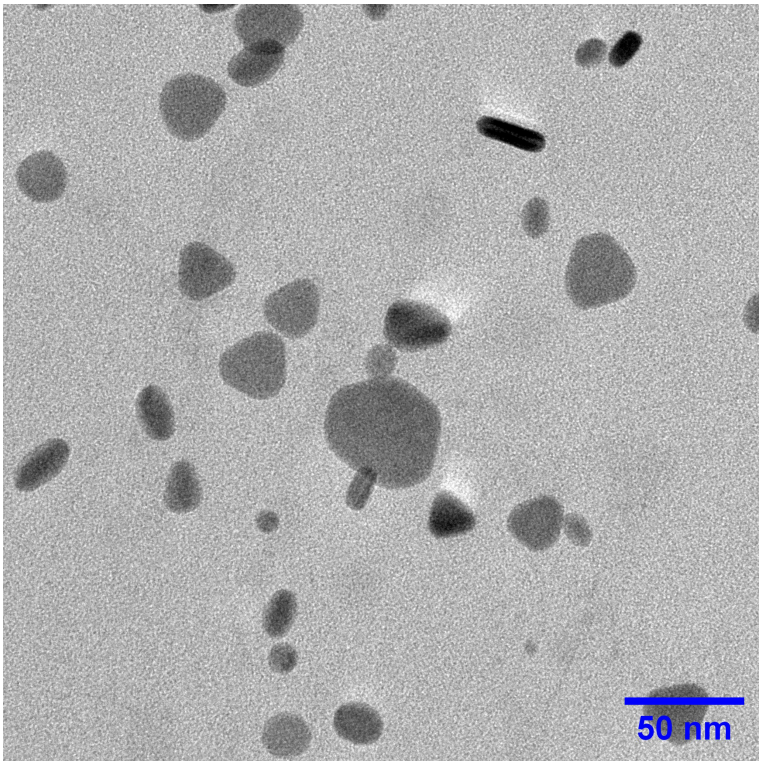

Figure S18: Additional image of the sample that was classified as "below threshold".

Effect of Organic Compounds on the Refractive Index

It was hypothesized that the presence of organic compounds such as PVP could affect the refractive index of the solution, which then affects the observed UV-Vis spectroscopy curve. To investigate this, nanoDDSCAT simulations of a silver nanoplate of 7 nm in thickness and 20 nm in diameter were performed in solution refractive indexes of 1.33 (water) and 1.55 (PVP).[4] We expect the refractive index of the solution surrounding the particle to be between 1.33 and 1.55, since it was hypothesized that PVP gets adsorbed onto the surface of the nanoparticles.

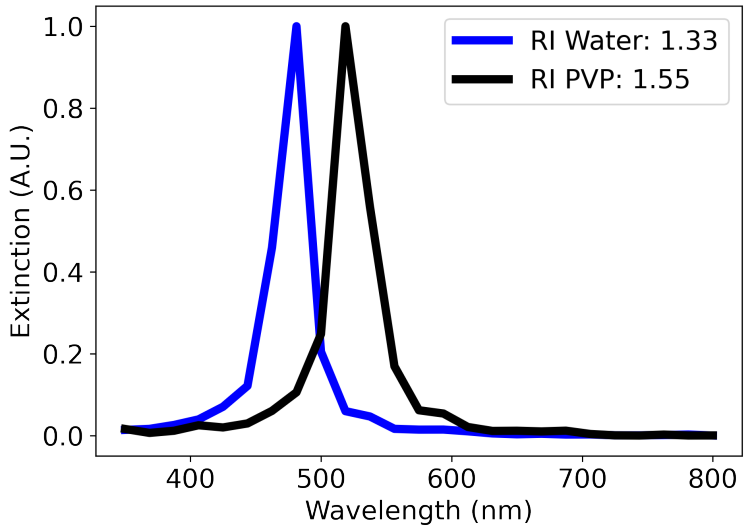

Figure S19: Effect of the solution's refractive index on the nanoDDSCAT simulation of a silver nanoplate of 7 nm in thickness and 20 nm in diameter.

Effect of Reagent Order and Time of Addition

An experiment was performed to determine the effect of different orders of addition and times of addition of the reagents on the synthesis of silver nanoparticles. Using code from (https://github.com/pozzo-research-group/otto), the volumes, order, and times of addition of each reagent could be varied. The compositions of five randomly selected samples that were classified as "Below Threshold" were used for the experiment. Four different orders of adding the reagents were tested, and four different delay times of adding a reagent were tested.

Table 1: Concentrations (mM) of the five randomly selected "Below Threshold" samples that were tested for different orders of addition and times of addition.

Sample	PVP	Tannic Acid	Ascorbic Acid	Silver Nitrate	Silver Seeds
A	0.26	0.36	0.024	0.34	0.0043
B	0.13	0.35	0.012	0.11	0.0043
C	0.14	0.36	0.012	0.35	0.0042
D	0.11	0.12	0.049	0.23	0.0037
E	0.18	0.28	0.049	0.15	0.0037

The following orders were tested. Reagents in **bold** indicate the reagent where a time delay of 1, 2, 8, and 15 minutes was performed before adding the reagent.

- Order 1: PVP, Water, Tannic Acid, Ascorbic Acid, Silver Nitrate, **Seeds**
- Order 2: Water, Silver Nitrate, Seeds, PVP, **Ascorbic Acid**, Tannic Acid
- Order 3: Water, Silver Nitrate, Seeds, Ascorbic Acid, **Tannic Acid**, PVP
- Order 4: Water, Silver Nitrate, Seeds, Tannic Acid, **PVP**, Ascorbic Acid

Order 1 was the same order used for all the experiments in the main manuscript. A delay time was added before Silver Seeds to test the hypothesis that adding reducing agents before silver nitrate could lead to a large number of nuclei and the possibility of clumping of nanoclusters. This mechanism results in polydisperse particles. Orders 2, 3, and 4 were chosen to determine the effect of adding a reducing agent immediately after Silver Seeds.

The results from the experiment, in fig. S20, show that the order in which the reagents are added affects the final structure. Order 1, which is the first column, is the original order used for all the experiments in the main manuscript. The spectra from the samples made using this order have a single peak, which is expected of samples with plate-like characteristics. In addition, the time delay of silver seeds does not affect the structure of the nanoparticle. This refutes the hypothesis that adding reducing agents before silver nitrate could lead to a large number of nuclei clusters which result in polydispersity. A more likely explanation is that a large number of nuclei are formed, but do not start assembling into clusters until silver seeds are added.

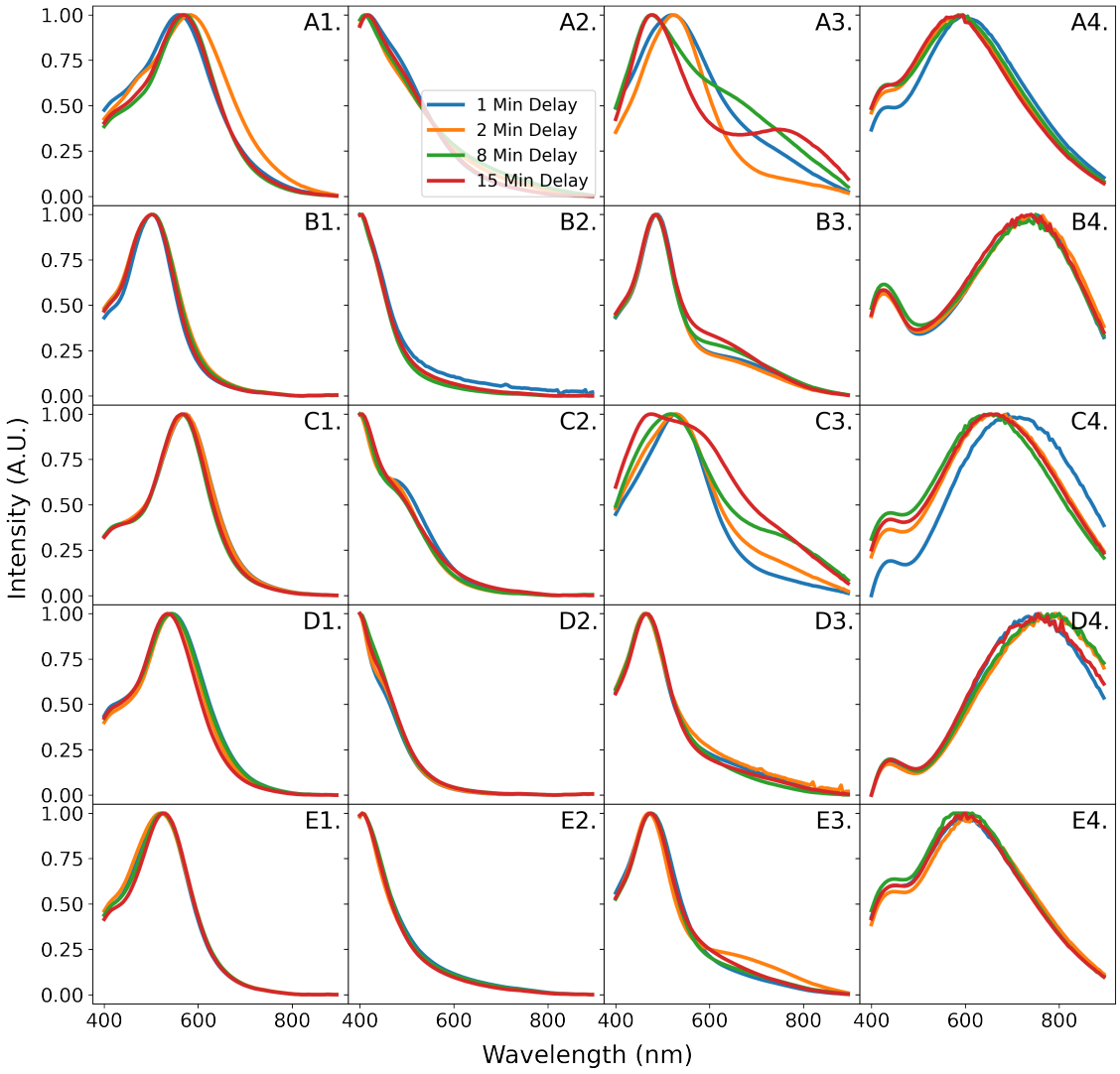

Figure S20: Effect of the order and time of addition of the reagents on the UV-Vis spectroscopy curves of the silver nanoparticles. Spectroscopy curves are normalized from 0 to 1. Five samples of different concentrations (A-E) were created with different orders (1-4) and delay times (blue, orange, green, red).

Order 2, the second column of fig. S20, tests the effect of adding PVP immediately after silver nitrate and silver seeds. The spectra of all the samples have peaks at low wavelengths which suggests the presence of small particles. An explanation could be that PVP stabilizes the small particles, preventing them from growing into larger nanoparticles. The effect of the time delay of ascorbic acid does not seem to affect the structure since the particles are already stabilized.

Order 3, the third column of fig. S20, tests the effect of adding ascorbic acid after silver nitrate and silver seeds. PVP is then added with a variable time delay. This order is interesting because the time delay seems to affect the nanoparticle's structure seen in A3, B3, and C3 in fig. S20. The addition of tannic acid at different times results in high absorbance at high wavelengths, which is a sign of large particles or aggregation. Since ascorbic acid is the strongest reducing agent, a high concentration of silver atoms is expected to be produced. The addition of tannic acid at different delay times acts as a reaction quencher, which results in different nanoparticle structures.

Finally, Order 4, the fourth column of fig. S20, tests the effect of adding tannic acid after silver nitrate and silver seeds. PVP is added with a variable time delay. The time delay does not seem to affect the structure of the nanoparticles. However, most of the spectra have peaks at high wavelengths which indicates the presence of large nanoparticles. This could be attributed to the addition of tannic acid immediately after silver seeds, which acts as a reducing agent as well as a stabilizer.

In summary, this experiment showed that the order in which the reagents are added, as well as the time of addition of the reagents, affect the final nanoparticle structure. Depending on the desired nanoparticle shape and size, a specific order can be used to achieve it. For example, if large nanoparticles are desired, Order 4 can be used, which adds tannic acid after the silver seeds. If small nanoparticles are desired, Order 2 can be used which adds PVP after the silver seeds. There are many orders and times of addition that can be used to synthesize nanoparticles. In our manuscript, it was decided that the order and time of addition would remain constant since the design space was already very large. Other experiments could use our high-throughput setup to explore different orders and times of addition. We would like to emphasize how experiments on the order or time of reagent addition are something that can be easily tested and explored using robotic systems like the one we used but are challenging to explore manually due to the manual labor involved. This means that the data-driven workflow that we used can be applied to other systems where time affects the

final material structure.

Effect of Mixing on the Synthesis

To test if the mixing in our protocol was adequate, a few samples were resynthesized manually using a magnetic stir bar at 900 rpm to mix the sample. The UV-Vis spectra was compared to the samples that were synthesized using the OT2. The results show good agreement between the samples using the two methods.

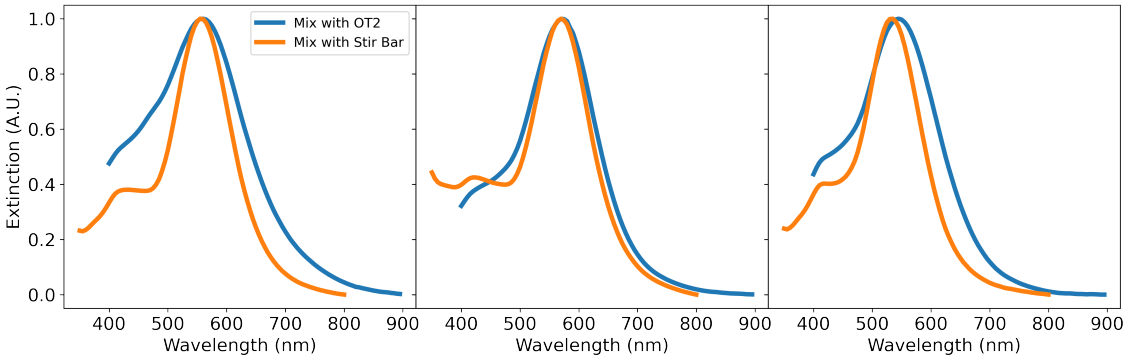

Figure S21: The UV-Vis spectra of three samples that were synthesized using the OT2 with a volume of 325 μL and synthesized manually with a volume of 10 mL. The sample synthesized using the OT2 was mixed by aspirating and dispensing 100 μL of solution 3 times. The sample synthesized manually was mixed using a magnetic stir bar at 900 rpm for 4 hours. The UV-Vis Spectra of all samples were taken 24 hours after synthesis.

UV-Vis Spectroscopy as a Proxy for Particle Size

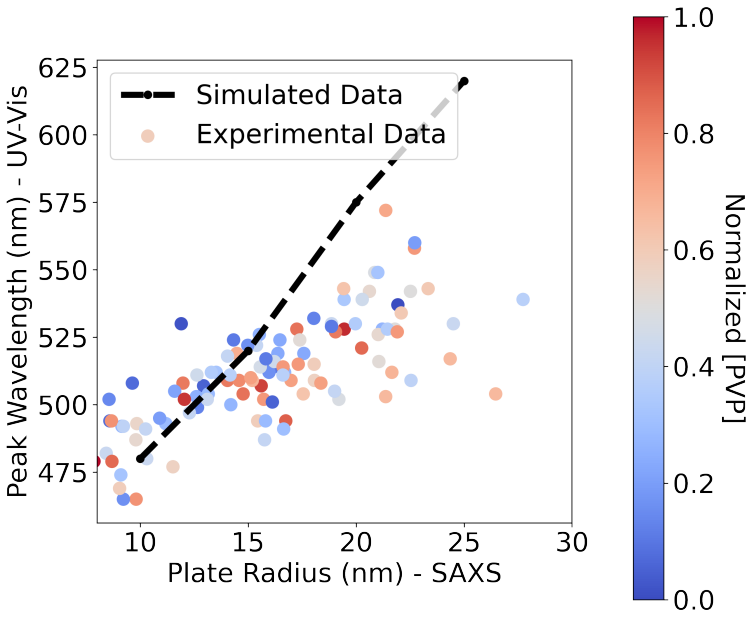

Figure S22: The effect of PVP on the size measured by UV-Vis spectroscopy and by SAXS.

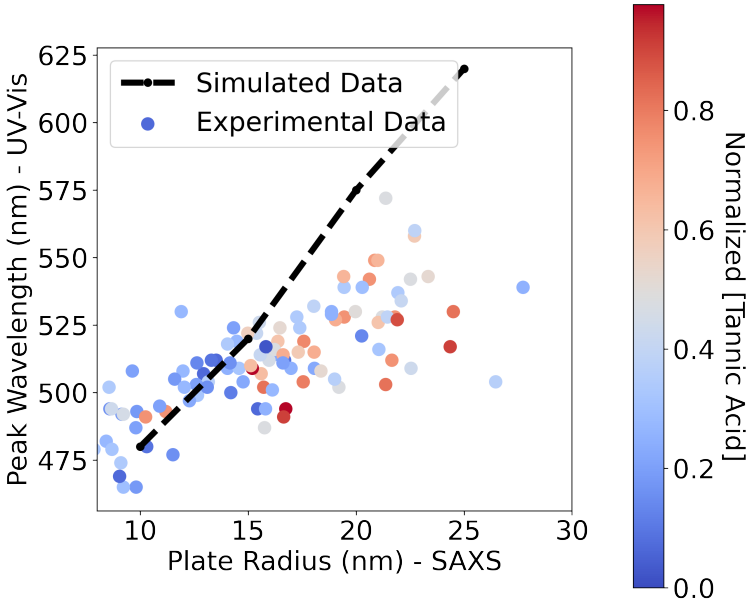

Figure S23: The effect of tannic acid on the size measured by UV-Vis spectroscopy and by SAXS.

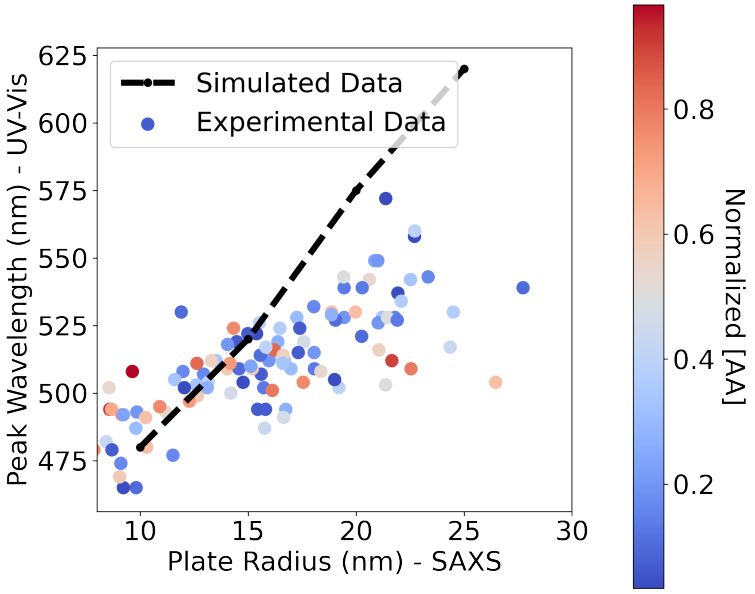

Figure S24: The effect of ascorbic acid on the size measured by UV-Vis spectroscopy and by SAXS.

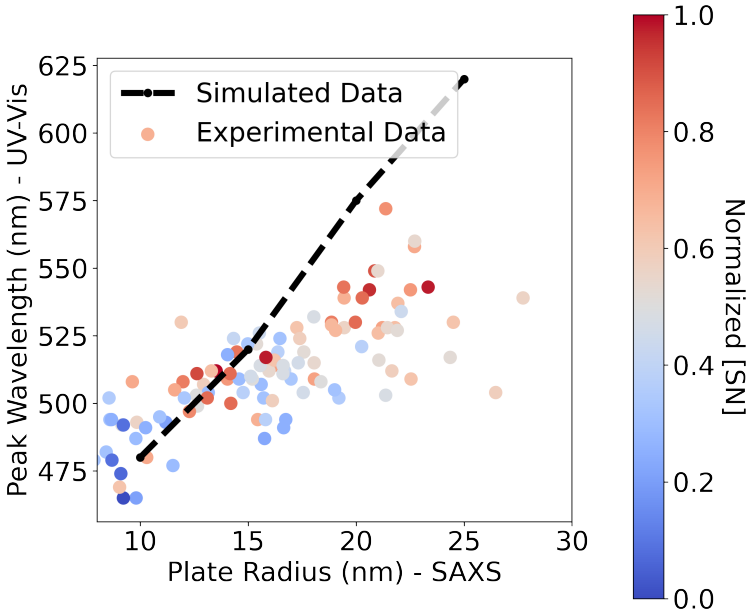

Figure S25: The effect of silver nitrate on the size measured by UV-Vis spectroscopy and by SAXS.

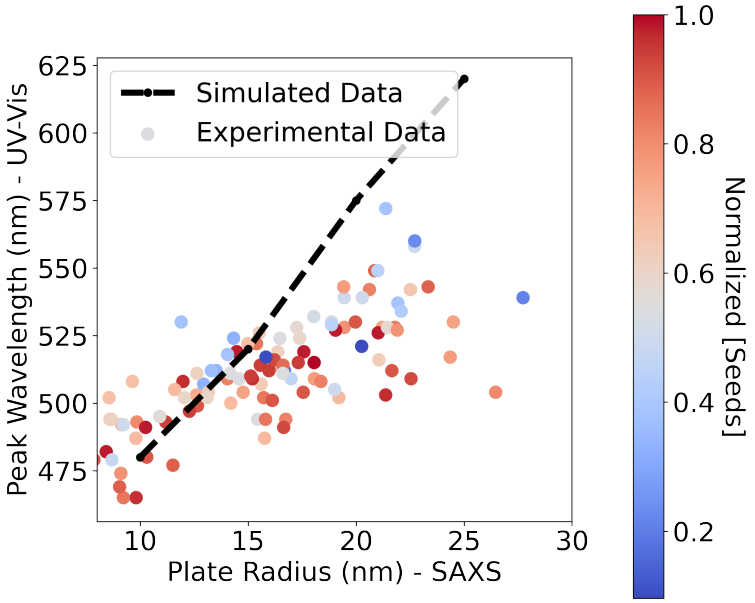

Figure S26: The effect of silver seeds on the size measured by UV-Vis spectroscopy and by SAXS.

SEM

Scanning electron microscopy (SEM) was used to determine differences in the samples classified as "above threshold" and "below threshold". Three samples of each classification were synthesized and SEM images were obtained for each sample. Two images for each sample are shown here. The full set of images can be found in the online repository. The sizes of the particles in the images shown below were measured by ImageJ and the mean and standard deviation of the measurements are shown in the caption of the images. The first three figures (fig. S27, fig. S28, fig. S29) are from the samples labeled "below threshold" and the figures (fig. S30, fig. S31, fig. S32) are from the samples labeled "above threshold".

From the SEM images, the particles from the samples labeled "below threshold" resemble plate-like particles, are small (25-30 nm), and are monodisperse, which was the objective of the Gaussian process classifier in the Fast Spectroscopic Exploration section of this work. On the other hand, the particles from the samples labeled "above threshold" are large ($\geq$ 85 nm) and polydisperse.

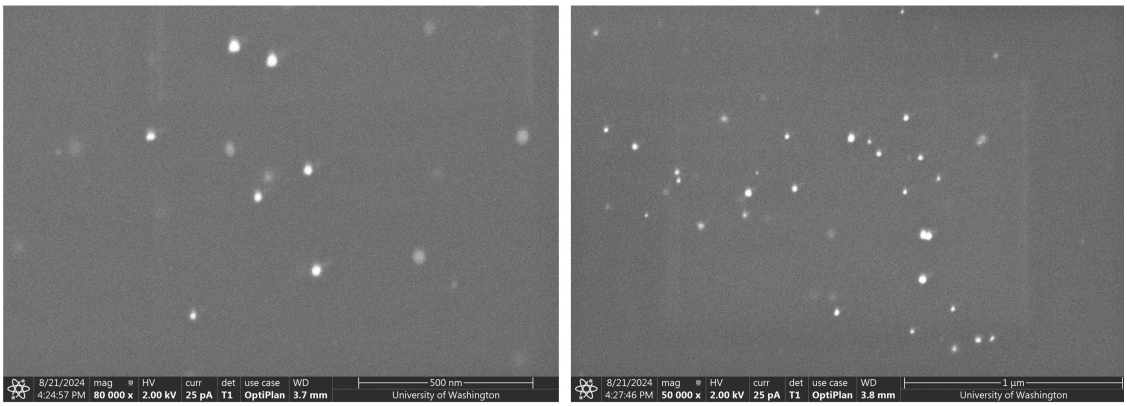

Figure S27: Images of the first sample classified as "below threshold". The mean size of all the particles in these images is 30.25 nm with a standard deviation of 6.20 nm.

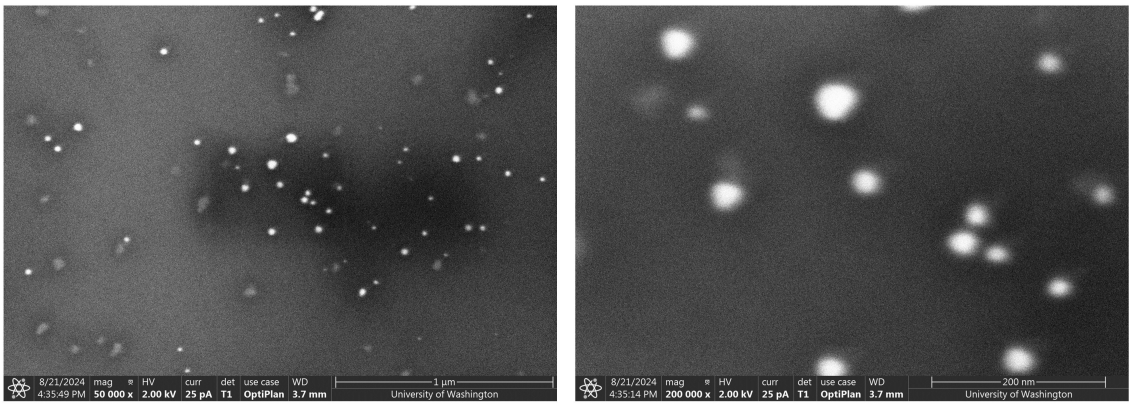

Figure S28: Images of the second sample classified as "below threshold". The mean size of all the particles in these images is 26.46 nm with a standard deviation of 4.75 nm.

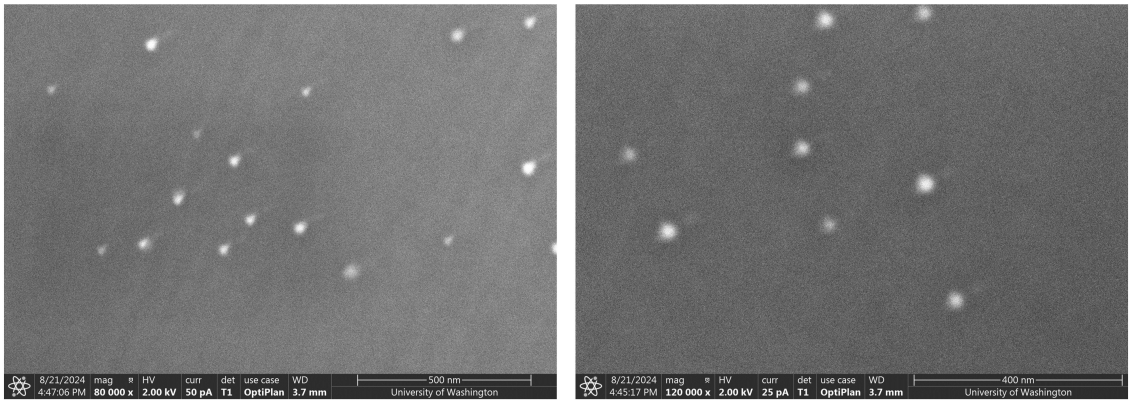

Figure S29: Images of the third sample classified as "below threshold". The mean size of all the particles in these images is 25.14 nm with a standard deviation of 3.14 nm.

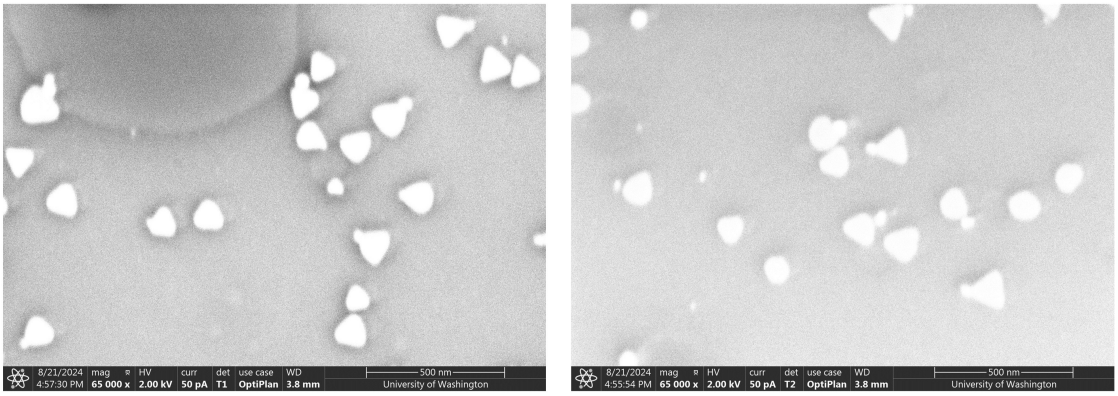

Figure S30: Images of the first sample classified as "above threshold". The mean size of all the particles in these images is 112.19 nm with a standard deviation of 13.46 nm.

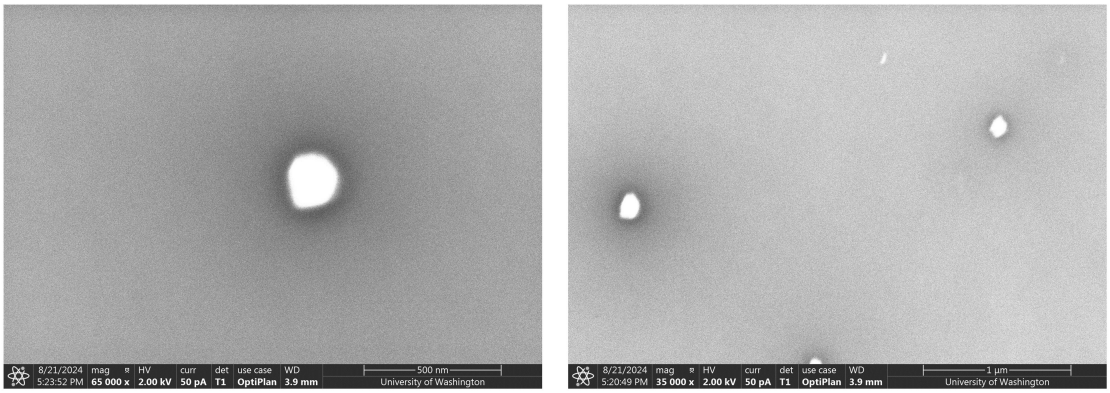

Figure S31: Images of the second sample classified as "above threshold". The mean size of all the particles in all the images taken of this sample is 173.64 nm with a standard deviation of 27.60 nm. This sample has a low concentration of particles in the images. The mean size was calculated using all the particles taken from all the images of this sample, which can be found in the online repository.

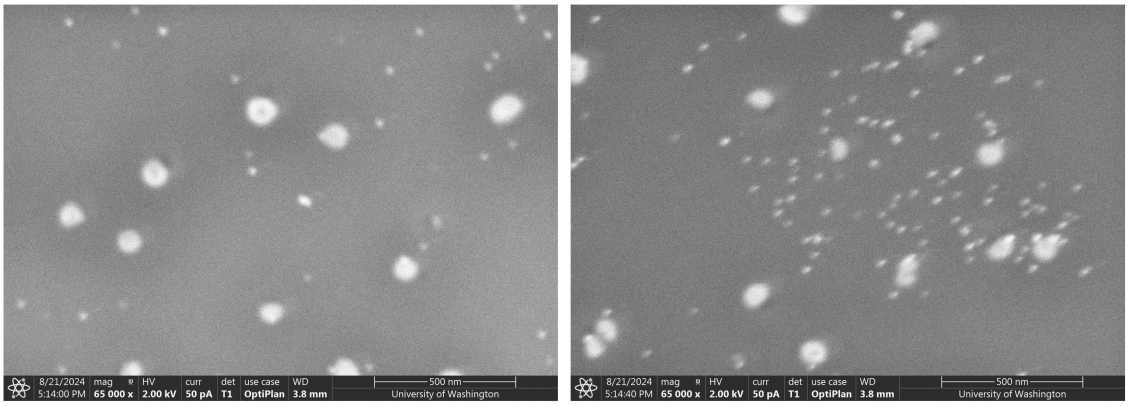

Figure S32: Images of the third sample classified as "above threshold". It was observed that this sample has a bimodal distribution of particles. The mean size of all the large particles in these images is 89.03 nm with a standard deviation of 10.38 nm. The mean size of all the small particles is 27.96 nm with a standard deviation of 3.03 nm.

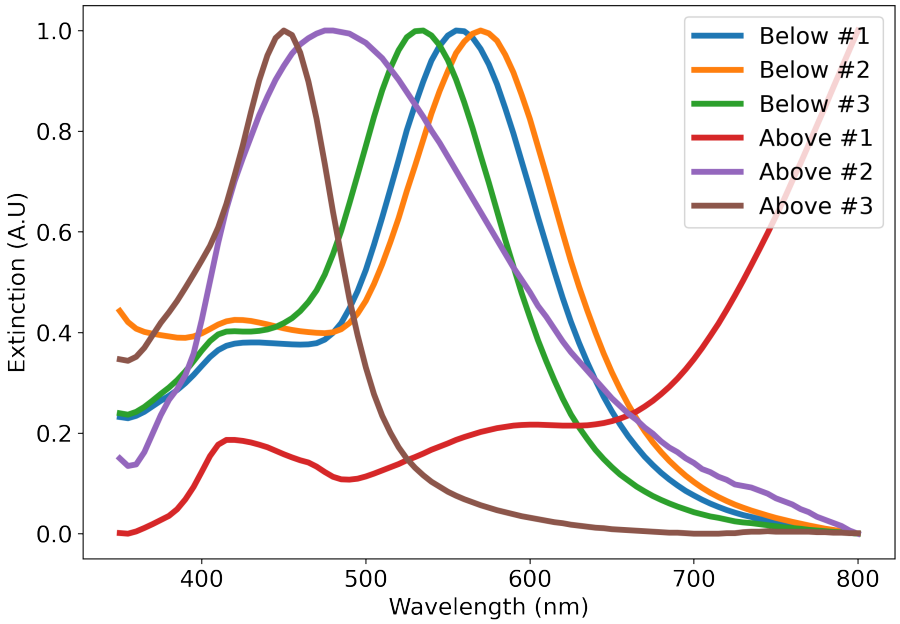

Figure S33: The UV-Vis Spectra of the samples that were analyzed with SEM. The samples labeled "Below" refer to the ones classified as "Below Threshold" and the ones labeled "Above" refer to the ones classified as "Above Threshold".

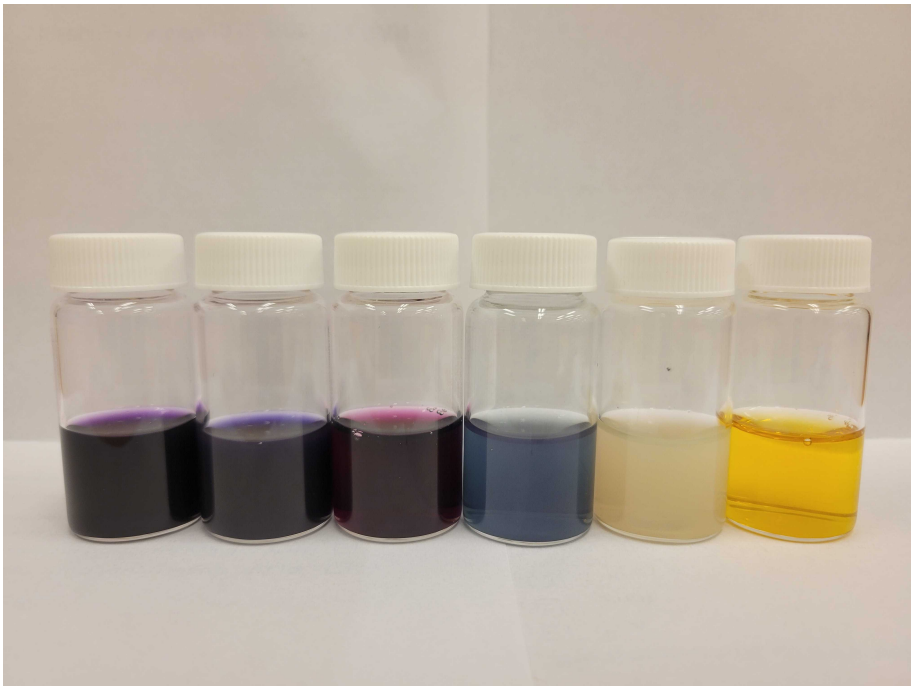

Figure S34: An image of the six samples that were analyzed using SEM. From left: "Below Threshold 1, 2, 3, "Above Threshold" 1, 2, 3.

References

(1) Deringer, V. L.; Bartók, A. P.; Bernstein, N.; Wilkins, D. M.; Ceriotti, M.; Csányi, G. Gaussian Process Regression for Materials and Molecules. *Chemical Reviews* **2021**, *121*, 10073–10141.

(2) Kiran, V.; Huat Thart, C.; Lilo D., P. Autonomous retrosynthesis of gold nanoparticles via spectral shape matching. *Digital Discovery* **2022**, *1*, 502–510.

(3) Srivastava, A.; Klassen, E. P. *Functional and shape data analysis*; Springer, 2016; Vol. 1.

(4) König, T. A. F.; Ledin, P. A.; Kerszulis, J.; Mahmoud, M. A.; El-Sayed, M. A.; Reynolds, J. R.; Tsukruk, V. V. Electrically Tunable Plasmonic Behavior of Nanocube–Polymer Nanomaterials Induced by a Redox-Active Electrochromic Polymer. *ACS Nano* **2014**, *8*, 6182–6192, PMID: 24870253.