Supplementary Data: Differentiating between gold nanorod samples

The supplementary information contains additional details on preparation of the gold nanorod samples, instruments used for analysis, the image acquisition process including thresholding, the raw data triage workflow, repeatability and reproducibility of the retained data, non-parametric comparisons of distributions for reproducibility, differentiating between the samples, and lab-mounted vs. user-mounted sample. A concordance table is included to clarify differences between descriptor definitions from common sources. The corresponding author for sample preparation is Dr. Xiaochun Wu (National Center for Nanoscience and Technology, Beijing, China; wuxc@nanoctr.cn). The corresponding author for the other sections is Prof. Eric A. Grulke (University of Kentucky, Lexington, KY: eric.grulke@uky.edu).

Table of contents

1. Sample preparation.

1.1. Case study background.

Gold nanorod test samples were prepared using a method similar to that used for the certified reference material, Au 720 (GSB 02-2994-2013), developed by the National Center for Nanoscience and Technology, Beijing, China.

The major objective of the study was to measure the particle size and shape distributions of discrete, rod-shaped nanoparticles using transmission electron microscopy. The methods and protocols developed should support the analyses of other two-dimensional nanoparticles, providing a complementary method for other methods of determining size and shape distribution, such as extinction spectrophotometric, small angle X-ray scattering, and dynamic light scattering.

1.2. Samples provided.

Two samples of gold nanoparticles coated with cetyltrimethylammonium bromide (CTAB) were received from Dr. Xiaochun Wu. These were provided as aqueous suspensions of gold nanoparticles and as 3-mm

transmission electron microscopy (TEM) grids that were pre-loaded with nanoparticles. Samples 1 and 2 were thought to have similar sizes but different size distributions.

1.3. Dispersion conditioning.
The gold nanorod aqueous dispersion test samples were coated with a layer of cetyl trimethylammonium bromide (CTAB, CAS Registry number 57-09-0). This surfactant is positively charged and stabilizes the nanorod aqueous dispersion. At the storage temperature of 4 °C, CTAB molecules partially crystallize and can separate from the solution. Prior to sample mounting on TEM grids, the samples were placed in a 30 °C water bath for 30 minutes to let CTAB crystals melt and redissolve in the solutions.

1.4. Mounting the sample on TEM grids.
Several types of TEM grids could be used, including carbon film grids. Five microliters of the undiluted gold nanorod suspension were deposited on the TEM with a micropipette. The TEM grid was then dried under a desk lamp.

1.5. Addendum to the mounting procedure.
Selecting mounting conditions for nanoparticle samples can be a separate study in itself. In this case, samples were mounted by the synthesis laboratory, which has extensive experience in generating monolayer dispersions of gold nanorods on TEM supports. There are a number of phenomena that need to be considered when mounting gold nanorods, including self-assembly, solvent evaporation, liquid crystalline behaviour of spherocylinder samples, and areal density of the nanorods on the support (Sharma *et al.*, 2009). Sharma et al. (Sharma *et al.*, 2009) provide a mini-review of gold nanorod capillary flow-driven self-assembly on TEM grids, evaporating a 1 μl sol droplet on carbon-coated TEM grids. Table SI-1 provides two-dimensional phase transformations of nanorods with different average aspect ratios, where W = width and L = length. Many variables can affect the self-assembly results, including number density in the dispersion, temperature, pH, drying rate, aspect ratio, polydispersity of shapes, and surfactants. In some cases, microphase separation regions can occur with different nanoparticle shapes. In general, areal densities of the nanorods of less than 20% on the support provide good numbers of non-touching particles. Lower areal densities usually result in lower numbers of touching particles, but increase the number of frames needed to count the desired number of particles.

Table SI-1. 2D phase transitions of nanorod self-assembly (Sharma *et al.*, 2009).

Variable	2D phase		
Aspect ratio	Low number density	Higher number density	Highest number density
W/L = 1/3	Isotropic	Smectic[a]-like	
W/L = 1/6	Isotropic	Nematic[b]-like	Smectic-like
W/L = 1/14	Isotropic	Nematic-like	
Evaporation rate	Slow	Fast	
	Nanorod agglomeration	Improved nanorod dispersion	

[a] Smectic: orientation order and layered structure
[b] Nematic: orientation order

2. Instruments and protocol factors.

Eric A. Grulke, University of Kentucky
Corresponding author: grulke@uky.edu.

Table SI-2 summarizes the instruments and their operating conditions for this ILC.

Table SI-2. Instruments used in this ILC.

Laboratory	Instrument	Model	Accelerating voltage
L1	TEM	EM922 (Zeiss)	200 kV
L2	TEM	Tecnai F30 (FEI)	300 kV
L3	TEM	JEM-1220 (JEOL)	80 kV
L4	TEM	JEM 2010F (JEOL)	130 kV
L5	TEM	Hitachi	300 kV
L6	STEM	Zeiss Supra 35 VP	30 kV

Some of the factors in table SI-3 might affect scale and shape distribution data. Three different calibration methods were used. The image scale varied from 0.15 to 0.6 nm per pixel. The number of data points per lab varied from 482 to 6406. Lab L6 automated the image acquisition process, taking 5 to 10 times the data of the other labs. The large difference in dataset size affected ANOVA analysis of the ensemble data, although the pair-wise ANOVA method provided good comparisons. The number of frames used to collect the data varied widely as well.

Table SI-3. Factors that might affect scale and shape distribution data.

Lab Code	Calibration method	Nm per pixel	Software	Sample 1		Sample 2	
				Particles	Images	Particles	Images
L1	Magical	0.150	ImageJ	840	25	753	20
L2	Cross grating, magical	0.204	ImageJ	643	13	643	13
L3	Grating	0.204	ImageJ	482	6	570	3
L4	Gold nanoparticles	0.6	ImageJ	600	2	512	2
L5	Gold nanoparticles	0.23	ImageJ	525	30	515	32
L6	144 nm 2D grating	0.225	ImageJ	6406	768	3489	657

3. Image acquisition, particle capture, and image thresholding.

The TEM calibration should be known and reported. Bright field images are preferred and each frame should contain a reference scale. All laboratories followed a robust protocol for capturing and thresholding gold nanoparticle images using ImageJ (NIOSH, 2012). This protocol was applied directly to the nanorod samples. In general, the automated settings for brightness/contrast and thresholding were near optimal for most labs. Figures SI-1-SI-4 show a progression of images developed during the processing of one image. Figure SI-1 is an image of the sample taken at 30 kV magnification. It includes the scale bar for setting the pixels per nm for the image. The background is generally uniform across the frame and there are some defects in the background near the lower left center of the image. These figures have about 20% coverage of nanoparticles, similar to areal density of images used by Sharma et al. (Sharma *et al.*, 2009)

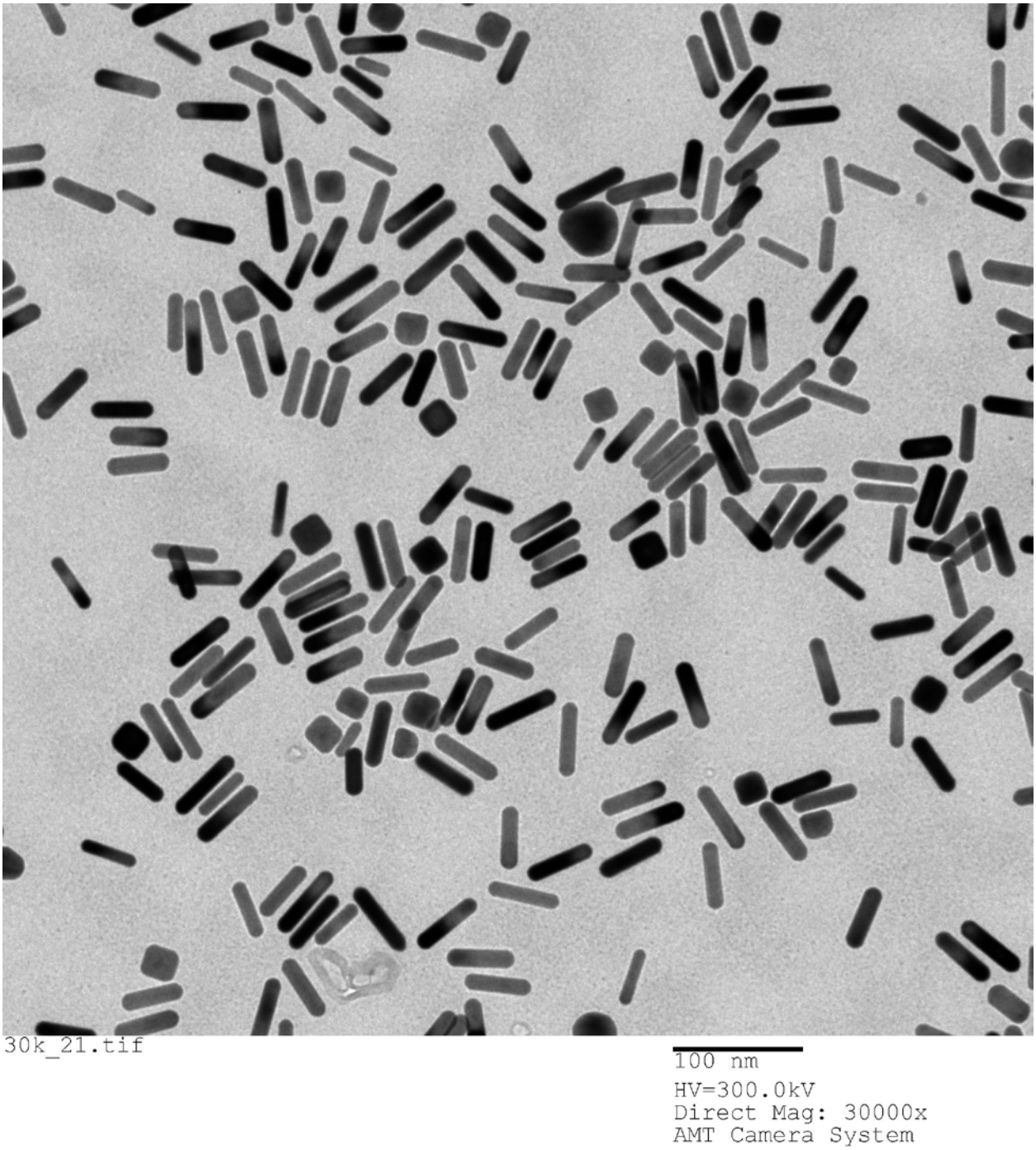

Figure SI-1. Example image of gold nanorod sample.

Figure SI-2 shows the cropped image. The scale bar was used to set the measurement scale. The rectangle feature of ImageJ has been used just inside the full frame, permitting almost all of the particles on the original frame to be included in analysis. Only the scale bar and metadata were removed. Had the scale bar been imprinted within the image frame, more of the image would have been cropped.

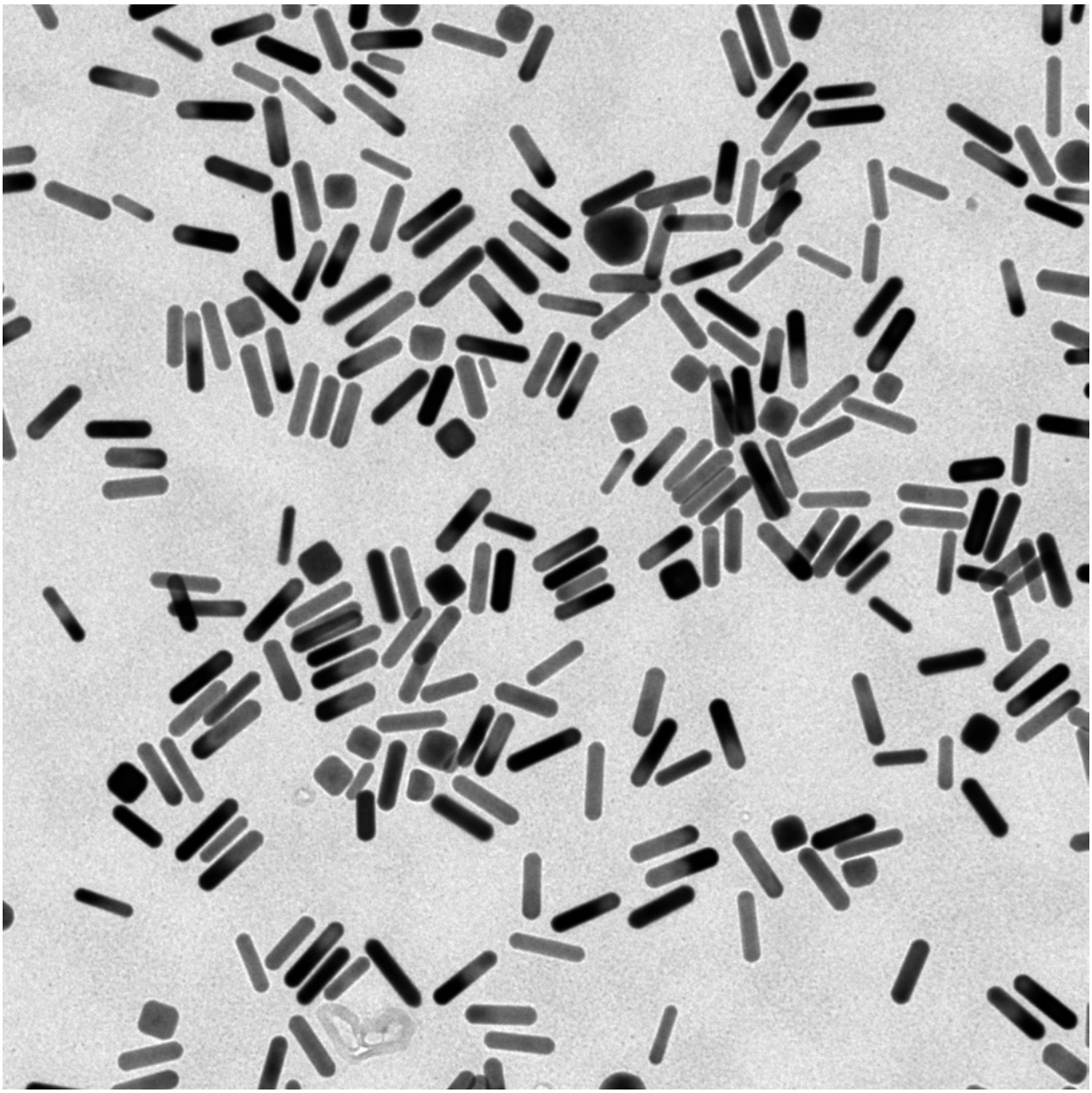

Figure SI-2. Example nanorod image after cropping.

Figure SI-3 shows the image after the brightness/contrast and the thresholding have been applied. ImageJ's automated settings can be close to optimal, particularly when image backgrounds are fairly uniform across the frame. The intensity of the particles in the thresholded image are more uniform than those in figure SI-1 and there is better contrast between the background and the particles. However, some of the particles that are discrete in the original image are 'touching' after this processing step. For example, compare the four aligned nanorods in the middle of figures SI-1 and SI-3: three of these nanorods are touching in figure SI-3 even though they have discrete edges in the original image.

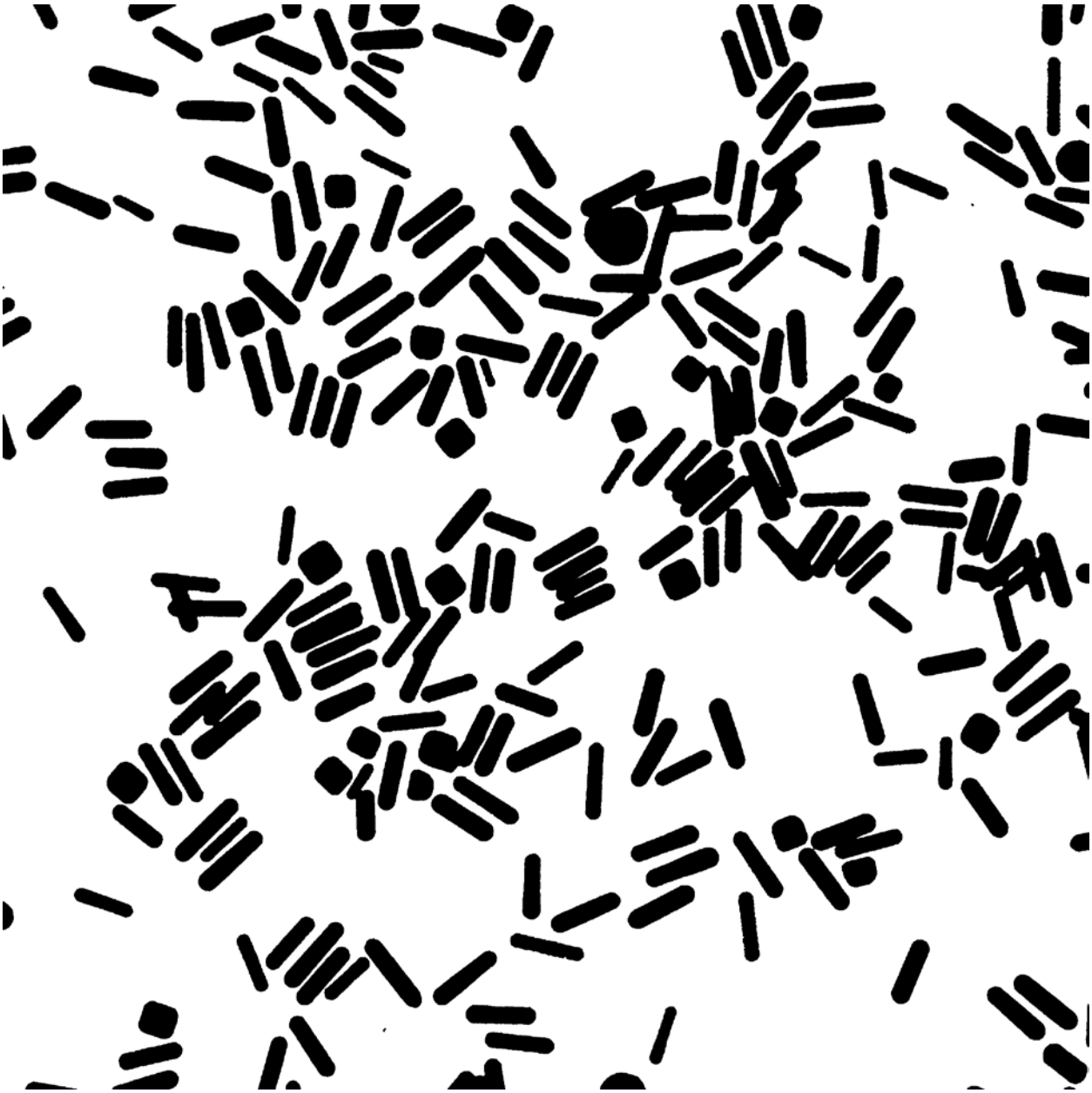

Figure SI-3. Example nanorod image after thresholding.

Figure SI-4 shows the particle outlines as reported by ImageJ in the Analyze Particles step. Particles whose areas cross the frame boundaries have not been outlined and were eliminated from analysis. Touching particles have unusually shaped outlines that are easily discriminated by using appropriate shape factor descriptors. As in the original frame, there are discrete nanorods, discrete nanocubes, and complex, touching particles. Figure SI-4 is identical to figure 4 of the article.

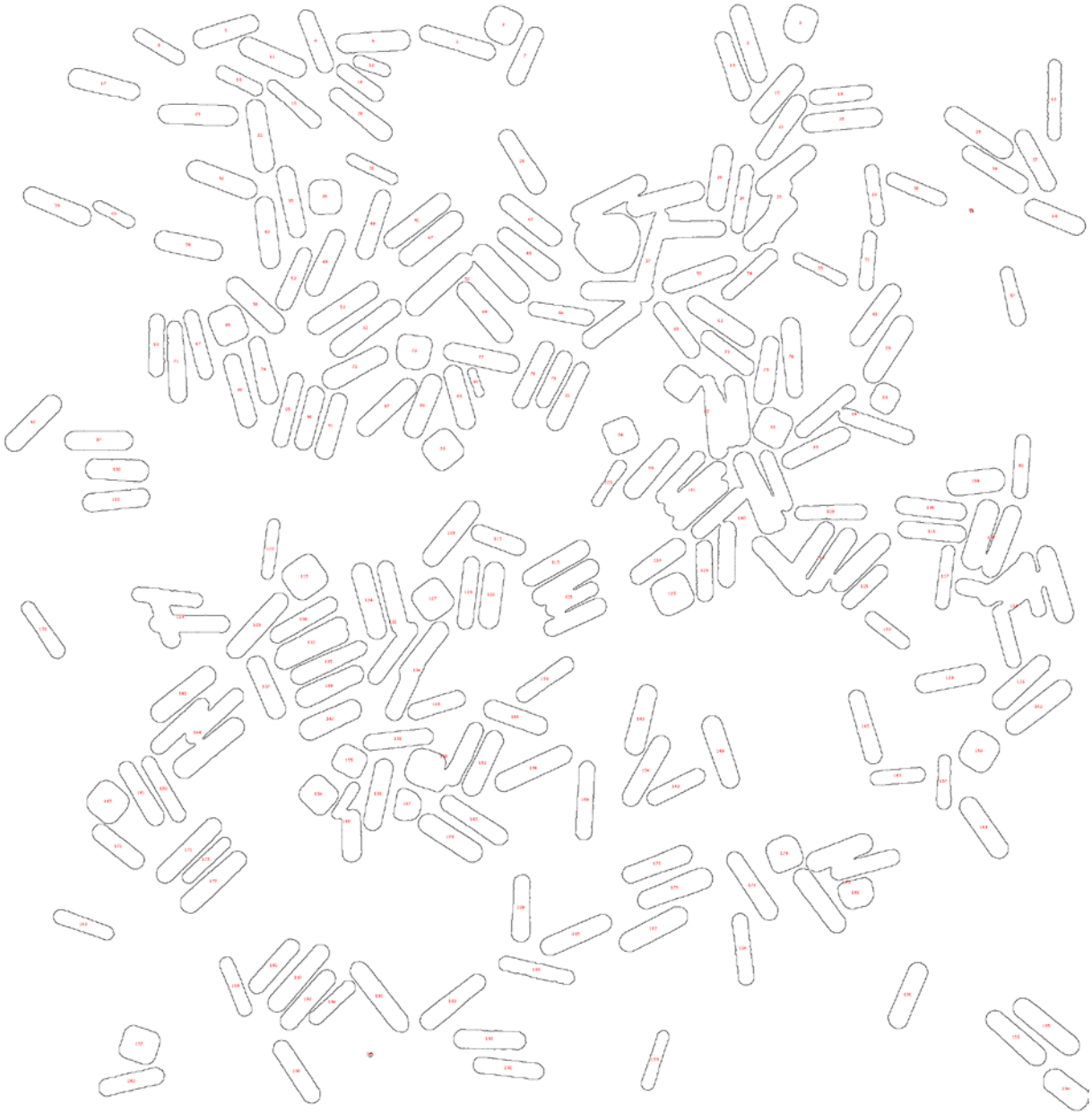

Figure SI-4. Example nanorod image after outlining by ImageJ.

4. Raw data triage.

The complex, touching particles and the nanocubes were removed from the analysis. The touching particles generally would have sizes (linear dimensions and areas) much larger than discrete nanorods, and skew these distributions if retained. The nanocubes have aspect ratios close to one and would greatly affect aspect ratio distributions, a key descriptor for shape. Figure SI-5 shows a generalized workflow for identifying both types and sorting the dataset to remove them. It was easiest to remove touching particles first, followed by the nanocubes. Although not discussed in detail in this article, it is important to remove particles that are below a specific size; as particle size becomes smaller, the uncertainty associated with length, perimeter, and area measurements increase. Typically, it is recommended that particles less than 200 pixels be ignored, as these can have as much as 5% error in measurements. Should this eliminate many of the particles in a frame, a higher magnification would be needed to image the smaller particles. This would require an additional calibration step for the new magnification.

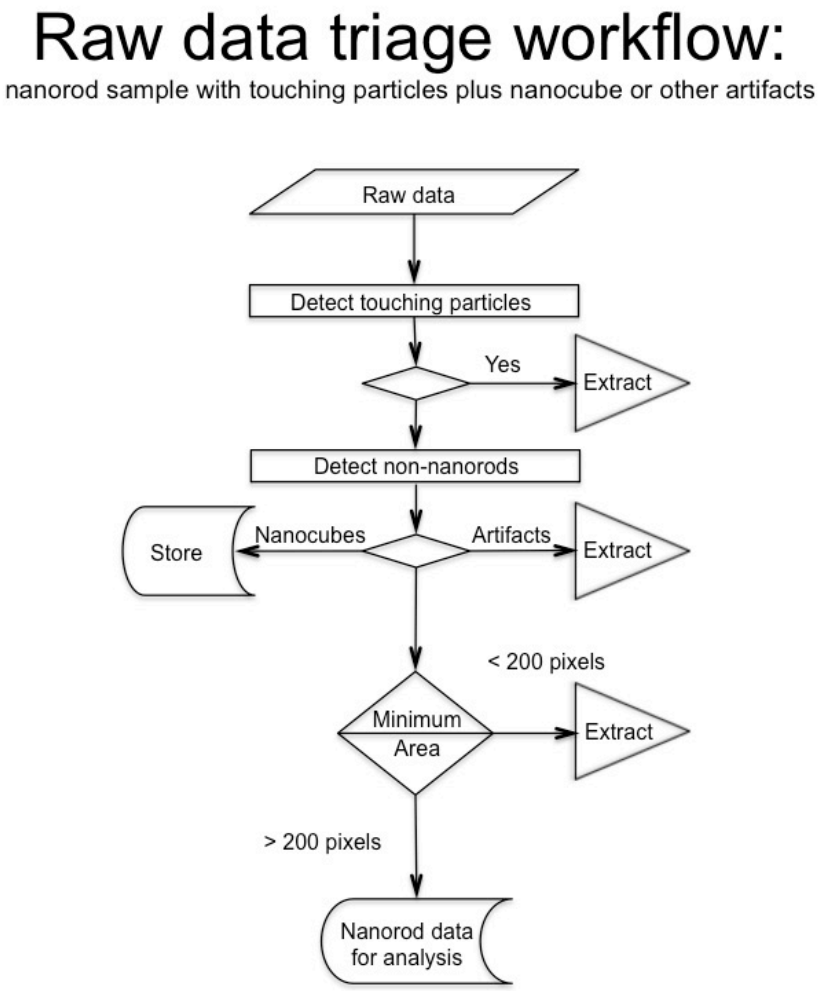

Figure SI-5. Raw data triage workflow: nanorod mixtures.

Figure SI-6 shows the ANOVA analysis of particle shape for the solidity descriptor. This shows that touching particles are within the range of solidities less than 0.9.

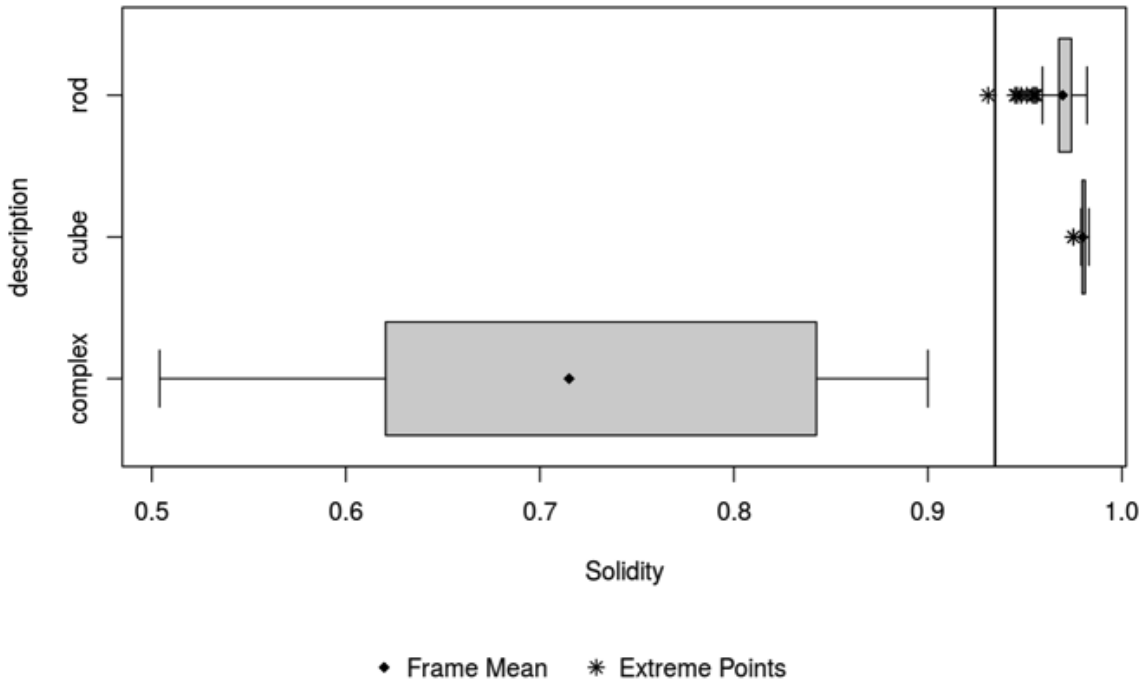

Figure SI-6. ANOVA of solidity vs. three shapes. Vertical black line = grand mean; complex = touching particles; cube = nanocubes; rod = nanorods; gray boxplots indicate +/- 1.5 standard deviations.

5. Repeatability and reproducibility of retained nanorod data.

Few of the datasets contained particles that might have been touching by the descriptor criteria developed in the study. However, most labs reported data that included some nanocubes. Table SI-4 shows the number of reported particles, nanocubes removed, other particles, and retained nanorod particles after raw data triage. The percentage of nanocubes removed ranged from 1% to 5%; this variation is likely due to differences in applying rejection criteria.

Table SI-4. Reported particles and retained nanorod particles. Sample 2.

Lab code	Reported particles	Nanocubes removed	Others	Retained nanorod particles
L1	753	31	2	720
L2	643	32	7	604
L3L	571	7	1	563
L3U	512	15	2	495
L4	513	9	1	503
L5	513	10	1	502

5.1 Repeatability.

Intralaboratory repeatability was assessed by using ANOVA to compare frame-wise descriptor means with the grand mean of the descriptor for the laboratory. Table SI-5 shows p-values for the repeatability assessments of Labs L2 (643 particles reported) and L6 (6159 particles reported). Values for four descriptors, solidity, aspect ratio, Feret diameter, and minimum Feret diameter, are shown. The p-value for the solidity descriptor of Lab L6 is 0.379, signifying that all frames have means similar to the grand mean. Clearly, touching particles were removed from this dataset. By contrast, Lab L2 had a p-value less than 0.001 for the solidity descriptor; this dataset had touching particles that were removed prior to further analysis. Evaluating intralaboratory repeatability by ANOVA of frame data provides one way to detect whether artefacts are present in the dataset.

It is also possible to identify descriptors with high repeatability. For example, the Feret diameter was the only descriptor for Lab L2 with p-values greater than 0.05 in a frame-wise ANOVA. Other descriptors had 15-23% of the frames with averages different from the grand mean. Generally, the Feret diameter descriptor had the highest intralaboratory repeatability in this ILC and would be a preferred choice to represent nanorod size.

Table SI-5. Intralaboratory repeatability, As-received data, L2 and L6, Sample 2.

Lab	L 2			L 6		
Descriptor	p-value	Frames with different average	% different means	p-value	Frames with different average	% different means
Solidity	<0.001	3	23%	0.379	51	6.6%
Aspect ratio	0.02	2	15%	<0.001	81	11%
Feret	0.998	0	0%	0.992	32	4.2%
minFeret	.002	3	23%	<0.001	64	8.3%

5.2.Reproducibility.

ANOVA would usually be a first choice for comparing interlaboratory reproducibility; it compares mean descriptor values for each dataset to the grand mean descriptor value for all datasets combined. However, Lab L6 has twice the number of points than the sum of labs L1-L5. Its descriptor values dominate grand averages in ANOVA or bivariate analyses, leading to low p-values by these tests.

6. Non-parametric comparison of distributions

6.1. Pair-wise reproducibility: ANOVA and bivariate analysis of aspect ratio datasets

ANOVA and bivariate analysis can be used to compare pairs of datasets. ANOVA results compare the aspect ratio means of the datasets. Bivariate analysis results compare aspect ratio datasets by using the empirical cumulative distributions. Table SI-6 shows p-values of these comparisons for aspect ratio descriptor datasets. Bivariate analysis showed that 9 pairs of datasets were similar for Sample 1 and 8 pairs of datasets were similar for Sample 2. ANOVA results showed that 4 pairs of datasets had similar means for Sample 1 and 3 pairs of datasets had similar means for Sample 2. Both methods have the advantage of comparing descriptor distributions with no reference to distribution models. However, they do not generate parameters that can be used to estimate measurement uncertainties.

Lab pair	Bivariate analysis of aspect ratio cumulative distributions, p-values		ANOVA of aspect ratio cumulative distributions, p-values	
	Sample 1	Sample 2	Sample 1	Sample 2
L1/L2	**0.805**	**0.214**	**0.636**	< 0.001
L1/L3L	**0.052**	**0.303**	< 0.001	< 0.001
L1/L3U	< 0.001	< 0.001	< 0.001	< 0.001
L1/L4	**0.206**	**0.203**	< 0.001	< 0.001
L1/L5	**1.000**	**0.923**	**1.00**	**0.657**
L2/L3L	0.016	0.020	< 0.001	< 0.001
L2/L3U	< 0.001	< 0.001	< 0.001	< 0.001
L2/L4	**0.675**	**1.000**	**0.257**	**1.00**
L2/L5	**0.941**	**0.180**	**0.60**	< 0.001
L3L/L3U	**0.406**	0.007	0.04	< 0.001
L3L/L4	0.003	0.022	< 0.001	< 0.001
L3L/L5	**0.067**	**0.488**	< 0.001	**0.062**
L3U/L4	< 0.001	< 0.001	< 0.001	< 0.001
L3U/L5	0.003	< 0.001	< 0.001	< 0.001
L4/L5	**0.281**	**0.198**	< 0.001	< 0.001

Table SI-6. p-values of ANOVA and bivariate pair-wise comparisons of aspect ratio. P-values greater than 0.05 are shown in bold font.

6.2. Differentiating between Samples 1 and 2: Kolmogorov-Smirnov two-sample test.

A key objective of the study was to determine whether Sample 1 and Sample 2 had different distributions. Since there was not uniform agreement between laboratory datasets, as shown in Section 6.1, it was decided to Sample 1 and Sample 2 data within each laboratory. ANOVA and bivariate analysis of Samples 1 and 2 for all laboratories demonstrated that neither the sample means nor the empirical cumulative distributions of the aspect ratio descriptor were similar.

The K-S test is designed to determine directly whether two cumulative distributions are dissimilar. It was applied to the aspect ratio datasets of each laboratory. The example shown here is for Laboratory 1, which reported 762 data points for the bad sample and 720 data points for the good sample. A number of commercial statistical software packages can provide analysis packages for the K-S test. A simplified version of the method was programmed in Excel. The cumulative distribution for each dataset was recomputed by dividing the data range (0.15 < aspect ratio < 0.50) into 140 equal steps. The binned (discrete open points) and empirical distributions (connected lines) are shown in figure SI-7. The binned data represent the empirical distributions well and were used directly to compute the difference between the samples' cumulative distributions at each value of the aspect ratio, generating the difference within the absolute value brackets on the right hand side of Eq. 1. For these two samples, the value of the K-S statistic for rejecting the null hypothesis is $D_{n,m} > 0.0707$. When the supremum difference between the two cumulative distributions is great than this value, they are shown to be different distributions. The supremum value is exceeded over the aspect ratio range, 0.335 < aspect ratio < 0.440. All labs reported aspect ratio distributions for Samples 1 and 2 that were found to be different by the Kolmogorov-Smirnov test.

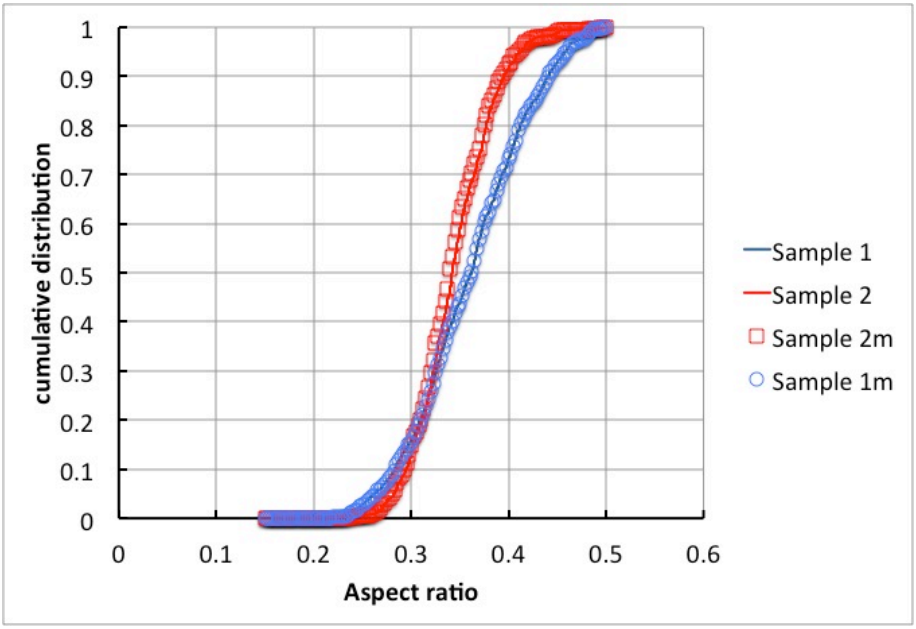

Figure SI-7 Comparison of empirical (solid lines) and binned (open form) cumulative distributions. Lab L1.

6.3 Lab-mounted vs. user-mounted samples.

Using the bivariate analysis tool, we can compare size-shape combinations statistically to identify descriptor pairs that yield distributions that meet the null hypothesis criterion. This approach was applied to compare the Lab L3 user-mounted sample data (L3U) to its data for lab-mounted samples (L3L). Feret diameter was used as the size descriptor. Aspect ratio and compactness were used as elongational descriptors, and form factor was used as the ruggedness shape descriptor. The bivariate p-values (Rizzo and Szekely, 2016) of various size-shape distributions are shown in Table SI-8.

Table SI-8 Bivariate analyses of size-shape distributions. P-value comparison of lab-mounted (L3L) and user-mounted (L3U) datasets

Sample	Size-shape descriptor distribution p-values		
	Feret/aspect ratio	**Feret/form factor**	**Feret/compactness**
Sample 1 (L3L/L3U)	0.054	0.053	0.053
Sample 2 (L3L/L3U)	0.218	0.184	0.221

The p-values are above 0.05 for all six cases, suggesting that the user-mounted and the lab-mounted distributions are similar. Bivariate analysis of size-shape distributions appears to be an alternative method for differentiating between nanorod sample distributions. Bivariate analysis does not reference a specific distribution and can be used as a tool to evaluate sample mounting procedures or other protocol modifications.

7. Descriptor concordance table

While users are encouraged to evaluate a wide variety of descriptors for particle size and shape distributions, there are differences in descriptor definitions used in the field. Table SI-8 shows the concordance of descriptor definitions for three references used in this work.

Table SI-8 Descriptor concordance table: Hentschel and Paige, ISO 9276-6 and ImageJ

Descriptor		Equation	Descriptor name per references		
Type	Category		Hentschel & Paige (Hentschel and Page, 2003)	ISO 9276-6 (ISO, 1998)	ImageJ (Ferreira and Rasband, 2012)
Elongation	1	$\frac{minFeret}{maxFeret}$	Aspect ratio	Aspect ratio[a,b]: its inverse is shape factor	Inverse of aspect ratio
	1	$\frac{minor}{major}$	-	Ellipse ratio[c,d]	Inverse of ellipse ratio
Ruggedness	1	$\frac{P_C}{Perim}$	-	Convexity[e,f]	-
	1	$\frac{\pi \cdot D_{avg}}{P}$	Convexity[g]	-	-
Elongation	2	$\frac{4 \cdot A}{\pi \cdot Feret^2}$	-	Roundness[h]	Roundness
	2	$\sqrt{\frac{4 \cdot A}{\pi \cdot Feret^2}}$	-	Compactness	-
Ruggedness	2	$\frac{4 \cdot \pi \cdot A}{Perim^2}$	-	Form factor	Circularity
	2	$\sqrt{\frac{4 \cdot \pi \cdot A}{Perim^2}}$	Form factor	Circularity	-
	2	$\frac{A}{A_C}$	-	Solidity[i]	Solidity
	2	$\frac{A}{maxFeret \cdot minFeret}$	-	Extent (bulkiness)	-

[a] minFeret = minimum Feret diameter
[b] Feret, maxFeret = Feret diameter
[c] minor = minor ellipse axis
[d] major = major ellipse axis
[e] Pc = perimeter of the convex hull envelope bounding the particle
[f] Perim = perimeter of the particle
[g] D_{avg} = mean of Feret diameters for all orientations
[h] A = area
[i] A_C = area of the convex hull envelope bounding the particle

List of references

Ferreira T and Rasband W 2012 ImageJ User Guide IJ 1.46m. ed NIH (Bethesday, MD: NIH)

Hentschel M L and Page N W 2003 Selection of descriptors for particle shape characterization *Part. Part. Syst. Charact.* **20** 25-38

ISO 1998 ISO 9276-1:1998 Representation of results of particle size analysis -- Part 1: graphical representation. In: *Geneva, Switzerland*: ISO) p 9

NIOSH 2012 NIOSH/DUNE Interlaboratory Study: Evaluation of a sample preparation technique for determination of TEM-based size distribution using NIST Reference Materials 8011, 8012, and 8013: gold nanoparticles. NIOSH/Dune Sciences) p 11

Rizzo M L and Szekely G J 2016 Energy distance *Wiley Interdisciplinary Reviews: Computational Statistics* **8** 27-38

Sharma V, Park K and Srinivasarao M 2009 Colloidal dispersion of gold nanorods: Historical background, optical properties, seed-mediated synthesis, shape separation and self-assembly *Mater. Sci. Eng., R* **R65** 1-38