Article

Computational Reverse-Engineering Analysis for Scattering Experiments of Assembled Binary Mixture of Nanoparticles

Christian M. Heil and Arthi Jayaraman*

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: In this paper, we describe a computational method for analyzing results from scattering experiments on dilute solutions of supraparticles, where each supraparticle is created by the assembly of nanoparticle mixtures. Taking scattering intensity profiles and nano-particle mixture composition and size distributions in each supraparticle as input, this computational approach called computational reverse engineering analysis for scattering experiments (CREASE) uses a genetic algorithm to output information about the structure of the assembled nanoparticles (e.g., real space pair correlation function, extent of nanoparticle mixing/segregation, sizes of domains) within a supraparticle. We validate this method by taking as input in silico scattering intensity profiles from coarse-grained molecular simulations of a binary mixture of nanoparticles, forming a close-packed structure and testing if our computational method can correctly reproduce the nanoparticle structure observed in those simulations. We test the strengths and limitations of our method using a variety of in silico scattering intensity profiles obtained from simulations of a spherical or a cubic supraparticle comprising binary nanoparticle mixtures with varying chemistries, with and without dispersity in sizes, that exhibit well-mixed to strongly segregated structures. The strengths of the presented method include its capability to analyze scattering intensity profiles even when the wavevector q range is limited, to handily provide all of the pairwise radial distribution functions, and to correctly determine the extent of segregation/mixing of the nanoparticles assembled in complex geometries.

KEYWORDS: *Small-angle scattering, structure, nanoparticle, assembly, computational analysis, segregation*
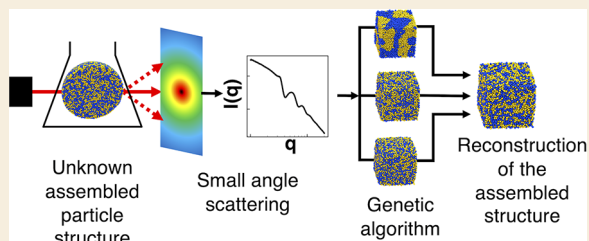
I. INTRODUCTION

Assembly of nanoparticles is relevant to engineering materials for a wide array of applications including optics, catalysis, biomedicine, and electronics.[1−5] Among the many different techniques[6−8] used for nanoparticle assembly, emulsion droplet and thin-film assembly are two common approaches that generate a micrometer-scale spherical or cuboidal "supraparticle" assembled from nanometer-scale particles or nanoparticles.[9−15] Supraparticles created using these ap-proaches have been used as optically active materials exhibiting colors that can be tuned with nanoparticle structure arrange-ment and size.[14,16−18] Photonic supraparticles, in particular, produce structural color due to the periodic arrangement of the nanoparticles,[19] which is more easily tunable and more resistant to degradation than typical absorption-based pig-ments. Structural characterization of assembled nanoparticles within these supraparticles is an important step during the design of materials with tailored optical or photonic properties.

Structural characterization of assembled nanoparticles has typically been performed using microscopy and scattering techniques, including small-angle X-ray and neutron scatter-ing.[6,20−27] Transmission electron microscopy (TEM) and scanning electron microscopy (SEM) are two common microscopy techniques for characterizing nanoparticle assem-blies.[6,21,22] TEM shoots an electron beam through a thin sample specimen (up to ∼200 nm thick), with the transmitted electrons used to create a two-dimensional representation of the three-dimensional sample.[6,21−23] TEM provides high-resolution images from the atomic to micron length scales.[21−23] SEM utilizes lower electron beam energies and scans the electron beam in raster fashion across the sample surface with the machine set to detect backscattered electrons.[6,22,23] SEM is typically used to probe the sample surface and is crucial for samples that are too thick for TEM imaging.[6] Due to the lower electron beam energy, SEM has a resolution on the nanometer to micron scale.[22] While microscopy is beneficial for obtaining images of nanoparticle assemblies, both TEM and SEM suffer from limited sample area and the tendency for sample bias of "nice looking" images as representative of the sample.[21,22,28] Furthermore, a microscopy image provides a single snapshot of the assembly
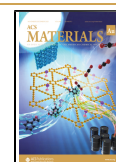
and not the ensemble average, whereas scattering techniques provide a method to characterize the entire material.[28] More advanced applications of SEM and TEM have investigated obtaining 3D data from the sample by combining a large set of correlated images into a 3D representation.[29−31] The focused ion beam (FIB) SEM system allows for the repeated collection of SEM images followed by a FIB milling process that progressively slices the sample to expose a new surface for imaging.[29−31] The 3D reconstruction using TEM requires obtaining many images at different tilt angles, which are then aligned to yield the 3D reconstruction with the sample thickness restriction imposed by TEM.[29] Both techniques require significant data acquisition times to obtain the many images as well as significant data processing and analysis to properly reassemble the images.[31] Additionally, the data processing and reconstruction step requires multicomponent mixtures or porous single-component systems to be sufficiently different visually to allow the reconstruction algorithm to correctly identify the different materials or pores: a nontrivial step that may require specific development of more advanced segmentation algorithms.[32]

Small-angle X-ray and neutron scattering (SAXS and SANS) are well-established techniques for probing assembled particle structures with a scope ranging from nanoparticle size (few nanometers) to the assembled domains scale (nanometers to microns), with the exact resolution depending on the instrument; for example, the commonly used NGB 30m SANS instrument at the National Institute of Standards and Technology is designed for distances from ∼1 to ∼700 nm.[33] Because SAXS intensity is based on the electron density difference between the nanoparticle and the solvent, SAXS is more sensitive to nanoparticles of heavier elements.[21,34] SANS intensity is based on the neutron scattering length density of a material that can vary based on element isotope, allowing for comparatively greater sensitivity to nanoparticles composed of lighter elements.[21,34] In scattering techniques, X-ray or neutron beams are used on samples, and the elastic scattering of those beams is measured.[6] The scattering intensity of the X-rays or neutrons (the output from SAXS and SANS experiments) depends on the nanoparticle form factor, which depends on the nanoparticle size and shape, and the structure factor, which captures the spatial correlations of the nanoparticles.[6] Contrast-matching of nanoparticle mixtures composed of multiple chemically distinct nanoparticles is used to reduce or eliminate the scattering contrast between one of the nanoparticle chemistries and the solvent leading to the scattering primarily from the non-contrast-matched nanoparticles, allowing for selective signal acquisition.[35,36] Contrast-matching the solvent to a nanoparticle chemistry in SAXS requires varying the electron density of the solvent by adding a contrast agent into the system.[37,38] Additionally, anomalous SAXS or ASAXS is an alternative technique to SAXS for complicated structures that requires the scattering intensity to be measured at least three times with X-rays of different energies.[39] By utilizing several different X-ray energies over multiple scattering experiments, ASAXS can be used to contrast-match as the scattering power of the components can vary depending on the X-ray energy used, possibly allowing isolation of individual components.[39] In SANS contrast-matching, the solvent or nanoparticle(s) is deuterated appropriately to remove the scattering contrast.[35,36,38] Regardless of whether one uses SAXS or SANS, the output from these techniques is a scattering intensity, $I(q)$, as a

function of scattering wavevector q, which is a function of the wavelength of radiation and scattering angle. The $I(q)$ versus q profile contains information about the particle−particle correlations and periodicity within the assembled structure between the different components comprising the sample.[36,38,40−42] Interpreting the $I(q)$ versus q typically requires fitting the scattering data with theoretical models that are applicable for nanoparticle systems such as the commonly used hard sphere[43,44] and sticky hard sphere[45,46] models. These models have been successfully applied to nanoparticle dispersions to reproduce experimental scattering results;[47,48] however, many of the theoretical models assume a fluid suspension, so those models gradually fail as the packing fraction exceeds 0.4.[49,50] Furthermore, the investigation of mixtures of multiple nanoparticle types (size, chemistry) and geometry (spherical, cubic) further complicates the scattering profile analysis. Thus, there has been growing interest in development of computational methods that side-step the choice of analytical model fits.

One computational method used to analyze nano- to micron-scale particle dispersions is reverse Monte Carlo (RMC) simulations. RMC simulations are a class of MC simulations in which randomly selected particle moves are accepted or rejected based on how much better the computed scattering profile of that post-move structure versus pre-move structure matches the target experimental scattering intensity profile.[51] The structures produced upon accepting MC particle moves evolve the system toward exhibiting a scattering intensity profile similar to the target experimental profile. RMC has been extensively applied to nanoparticle dispersions, forming aggregates as the method is applicable for low-density nanoparticle systems;[52−54] however, as RMC simulations rely on a good acceptance rate of the randomly selected particle moves, it is not as useful for high-density and close-packed nanoparticle assemblies where the acceptance rate for particle moves is low using the commonly implemented RMC algorithm.[55] Additionally, RMC simulations typically progressively iterate on one structure at a time, possibly converging toward a local optimum instead of the global optimum. Another method to analyze scattering from nanoparticle aggregates is fractal modeling using the Beaucage unified model, where the scattering profile is fit to provide information on the aggregate fractal dimension, aggregate radius of gyration, and aggregate subunit shape information.[56] Work comparing RMC and fractal modeling to determine the aggregate form factor of silica particles found both approaches to be similar in final results.[57] Oberdisse et al. state the fractal modeling approach was simpler to implement, but they noted the RMC approach was free from the fractal modeling approximations and provides a real space image of the aggregate.[57] We also note a recent study by Larsen et al. that presented a list of useful analytical structure factors to describe structural information about the aggregates.[58] Along the lines of RMC simulations, another method, MONSA, seeks to reconstruct the target structure by fitting experimental scattering curves. Reference 27 describes MONSA for a system of polymer-grafted nanoparticles. MONSA generates a close-packed structure of monodisperse beads and uses an MC-based approach to classify the beads as one of the components (in the specific case of that paper—nanoparticle, polymer graft, or solvent).[27] This method excelled for identifying different polymer-grafted nanoparticle morphologies;[27] however, its reliance on monodisperse beads prevents

application to nanoparticle assemblies with dispersity. For high-density and well-ordered crystalline nanoparticle assemblies, there are alternative computational methods for scattering analysis.[59,60] Previous work has developed a scattering formalism for superlattices comprising any arbitrarily shaped nanoparticles that can predict experimental superlattice structures.[59,60] That method requires fitting the lattice constant, lattice peak width, Debye–Waller factor (describes relative variation of particle positions from ideal lattice positions), a scaling factor, and the particle orientation matrix to the experimental scattering curve.[59,60]

There have been far fewer computational methods that can analyze structure of disordered nanoparticle assemblies at high density. One recent approach that characterizes nanoparticle assemblies varying from crystal-like to disordered structures is SASpdf.[61] This approach excelled at identifying the assembly structure, crystallite sizes of ordered assemblies, and the pair distribution function (or pdf). However, the authors state that this method "depends heavily on having good statistics (high signal-to-noise ratio) throughout the entire diffraction pattern $I_c(q)$ and the form factor $P(q)$." They illustrate this point by showing how noisy data can significantly impact the resulting pair distribution function.[61] Additionally, SASpdf requires the $I(q)$ to span a sufficiently large q range so that the Fourier transformation to obtain the pair distribution function is not too noisy. This large q range requirement may not be achieved if the scattering instrument resolution only spans a small q range corresponding to a few nanoparticle diameters; we discuss this shortcoming later in this paper by applying the SASpdf approach to one of the cases we examine to contrast SASpdf against our computational approach presented in this paper.

Our computational approach for analyzing small angle scattering profiles from assembled nanoparticles is inspired by a recently published computational method from our research laboratory, called CREASE (computational reverse engineering for scattering experiments), that demonstrated reliable reverse engineering of polymer micelle structure from experimental scattering data.[62] In that work, CREASE comprised two steps: the first step was a genetic algorithm step that provides macroscopic information about the micelle (e.g., core and corona sizes) for a given experimental scattering profile; the second step was a molecular reconstruction step that provided microscopic information (e.g., chain conformations, interfacial arrangement of monomers) by recreating the polymer chains within the micelle dimensions obtained from the first step.[62] For the given experimental scattering profile, the micelle core and corona diameters produced by CREASE were more in agreement with the cryo-TEM measurements than the analogous results from fitting the scattering profile with a commonly used polymer micelle analytical model.[62] This successful demonstration of CREASE motivates us to extend the method for analyzing scattering profiles obtained from dense, noncrystalline (amorphous), approaching close-packed arrangement of nanoparticle mixtures.

In this paper, we present a CREASE-inspired, genetic algorithm (GA) optimization-based approach that can link the $I(q)$ from a dilute solution of supraparticles to the structural arrangement of the assembled nanoparticles within a supraparticle. We first present the details and methodology for the case of a supraparticle that is made of a binary, chemically distinct, spherical nanoparticle mixture in either cubic or spherical confinement. We then validate our approach by taking as input the in silico experimental scattering profiles from prior simulations of supraparticles with binary mixtures of spherical nanoparticles with disperse nanoparticle sizes and varying extents of mixing and comparing the output of our GA approach to the known structure. This paper serves as a proof-of-concept for our approach, and we expect this approach to be easily extended to other nanoparticle systems with more complex nanoparticle shapes (core–shell, anisotropic), complex confinement (cylindrical), and gel or glassy systems.

II. APPROACH

We describe in this section the three main parts of our computational approach: the inputs, the steps within the genetic algorithm, and the outputs.

II.A. Required Inputs

The input information includes the experimental scattering intensity, $I_{target}(q)$ versus q, the nanoparticle size distribution(s), and the nanoparticle mixture composition.

For binary (A and B nanoparticles) mixtures explored in this work, the GA requires the $I_{target\,A}(q)$ and $I_{target\,B}(q)$ scattering profiles of the A-type nanoparticles and the B-type nanoparticles. In experiments, for solutions of binary nanoparticles, the $I_{target\,A}(q)$ can be obtained by contrast-matching nanoparticle B chemistries and the solvent, and vice versa for $I_{target\,B}(q)$.[24,35,36,38,63] For systems where the A and B nanoparticles have similar scattering densities or electron densities, we would use the total scattering intensity instead of the contrast-matched scattering profiles. We could also use $I_{target\,AB}(q)$ in place of $I_{target\,B}(q)$ as described in the Supporting Information. In all implementations described in this paper, we use the $I_{target\,A}(q)$ and $I_{target\,B}(q)$ scattering profiles of the A-type nanoparticles and the B-type nanoparticles.

The nanoparticle size distribution is input as the type of distribution (e.g., Gaussian, log-normal) and the parameters of that distribution (e.g., average diameter, standard deviation). In experiments, this information would be available from the characterization of the nanoparticle sizes after their synthesis. The nanoparticle mixture composition is input as volume fraction of each nanoparticle chemistry.

In this work, we mainly focus on equal volume fraction (i.e., symmetric 1:1 A/B binary mixture) of type A and type B nanoparticles with results for one asymmetric 1:3 A/B composition shown in the Supporting Information, each with a log-normal distribution of nanoparticle diameters with average nanoparticle diameter of 220 nm and standard deviation of 9, 15, or 30%. The log-normal distribution of nanoparticle diameters is discretized by utilizing 11 distinct groups of nanoparticles, with the groups' diameters drawn to match the log-normal distribution (Supporting Information Figure S1). For the average nanoparticle (220 nm diameter), the Guinier regime is for q less than 0.00117 Å^{-1}, and the close contact correlation peak occurs at $q = 0.001428$ Å^{-1}. Even though we focus on densely packed symmetric binary mixtures of solid, spherical nanoparticles, our method can be extended to incorporate other input nanoparticle shapes or types (e.g., cylinders, hollow spheres), mixture compositions (e.g., ternary mixture), and density regimes (i.e., densities below close-packed).

II.B. Genetic Algorithm

A GA is an adaptive heuristic search method utilizing population genetics (natural selection and genetic combinations/mutations) to identify optimal solutions.[64] We use a GA

to optimize the assembled nanoparticle structure whose computed scattering intensity, $I_{comp}(q)$, most closely matches $I_{target}(q)$ for both contrast-matched scattering profiles. Figure 1
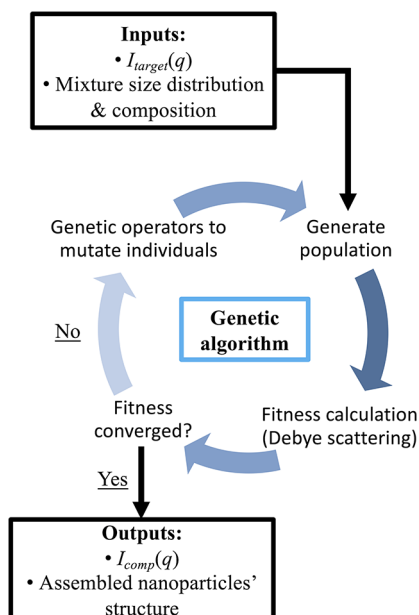

Figure 1. GA method workflow for reconstructing assembled nanoparticles' structure from scattering intensity profiles of nanoparticle mixtures.

provides an overview of the GA workflow. In short, the GA algorithm begins with a population of N different possible assembled nanoparticles' structures (termed as "individuals" in the rest of this paper). We evaluate the "fitness" of every individual by calculating how closely its $I_{comp}(q)$ matches the $I_{target}(q)$. In each generation, we identify the best individual defined as the assembled nanoparticle structure with high fitness; the probability of being selected for the next generation is proportional to the individual's fitness. Each selected individual undergoes a *mutation* (i.e., change in structure) to generate a new individual, and these new individuals then form the new generation that proceeds to the next iteration of the algorithm (evaluation, selection, and mutation). The GA continues this process for multiple generations until the fitness values of the top 10% of individuals over all generations reach a plateau value. The output of the GA is the structure corresponding to the individual with the highest fitness (i.e., the structure with an $I_{comp}(q)$ that best matches $I_{target}(q)$). Due to the expected variations between GA runs, we conduct three independent GA runs for each $I_{target}(q)$ and average the best structures from these three GA runs and report the standard deviation between the three runs as the error.

We describe each of the above steps in more detail below.

Step 1: Generating Individuals (Initial Generation) and Calculating $I_{comp}(q)$ for Each Individual. We begin each GA run by generating a population of 36 random individuals where each individual is an assembled structure of 20,000 nanoparticles matching the nanoparticle mixture composition and nanoparticle size distribution input information. When deciding how many individuals we should have in each population, we arrive at a number somewhere between having a small, computationally efficient population size but likely lacking in diversity and/or requiring more generations to

find the "best" structure and having a large, diverse population that requires more computational resources evaluating all individuals in that population. For our system, we found that one individual per CPU core in a 36-CPU core machine provided a sufficiently large population with balanced computational efficiency and diversity. The choice of 20,000 nanoparticles per individual is arbitrary, but it is sufficient to ensure suitable resolution for both $I_{comp}(q)$ and the pairwise radial distribution functions while balancing the computational effort required for each step in the GA.

Each individual's structure is generated by randomly placing the 20,000 nanoparticles with the input size distribution, in a simulation box at an initial occupied volume fraction (η) of 0.3. These nanoparticles are brought to a close-packed structure with $\eta \approx 0.55-0.6$ (depending on nanoparticle dispersity) using a conjugate gradient energy minimization technique,[65] implemented in the LAMMPS software package. To bring the nanoparticles to the close-packed state near the center of the box, we apply an additional energy potential to all nanoparticles based on the nanoparticle's x, y, and z position such that the nanoparticles further from center of the box (marked as origin) have an energy higher than that of the nanoparticles closer to the origin. During this step, the nanoparticles interact via the colloidal Lennard-Jones (cLJ) potential[66] with the Hamaker constant for all pairwise interactions (H_{A-A}, H_{A-B}, H_{B-B}) initially set to a weak 0.25 k_BT to prevent nanoparticle overlap. In simpler terms, the close-packed structures are generated by biasing the nanoparticles to move toward the origin with the conjugate gradient method used to minimize that energy gradient (creating the close-packed nanoparticle structure). We then maintain that energy gradient while performing a short molecular dynamics (MD) simulation in the NVT ensemble at $T^* = 1.0$ for ~10,000 time steps with a time step size of 0.005τ. This short MD simulation allows the nanoparticles to better pack especially at high nanoparticle diameter dispersity. While this conjugate gradient approach more quickly generates close-packed structures than performing molecular dynamics simulations in a shrinking confinement[9,14,15] or at high density,[67] this structure generation step is a rate-limiting step in the GA. Reducing the number of nanoparticles per simulation, reducing nanoparticle size dispersity, or starting at higher initial η are all ways to speed up this step.

After generating the close-packed structures for each individual, we perform multiple nanoparticle swaps to generate a population of diverse initial structures ranging from randomly mixed nanoparticles to strongly segregated nanoparticles. Having a diverse initial population is critical for the GA to prevent premature convergence to a "trapped" solution that is not the optimal structure. We use two different methods to swap nanoparticle locations: (a) swap with fixed nanoparticle distance (FD) and (b) swap with variable nanoparticle distance (VD). We perform GA runs with each nanoparticle swap method separately to confirm that the two methods independently produce similar "best" structures from the GA run. For both methods, we randomly generate H_{A-A}, H_{A-B}, and H_{B-B} Hamaker constants to use when selecting nanoparticles to swap by performing nanoparticle swaps (these Hamaker constants are different from the 0.25 k_BT used when creating the close-packed structure described earlier). These Hamaker constants are initially randomly selected in the interval $(0,1)$ as the GA method does not require a priori knowledge about the underlying particle interactions. The nanoparticles within each

individual structure undergo a randomly chosen 5000 to 5,000,000 number of nanoparticle swap steps. In each swap step, one randomly selected A nanoparticle is swapped in position with one randomly selected B nanoparticle. As done in Monte Carlo simulations, we use the Metropolis acceptance criterion[68] to accept or reject the nanoparticle swap. The pre-swap and post-swap energies needed for the Metropolis acceptance criterion are calculated based on the nanoparticles and their neighbors interacting via the cLJ potential with the Hamaker constants that were randomly assigned to all the nanoparticles in that individual. We note that while the cLJ potential is used in our implementation, other interaction potentials can be substituted as long as the interaction potential allows for differentiation of the neighborhood of the nanoparticles to be swapped.

The two separate swap methods—FD method and VD method—differ only in their calculation of the nanoparticle separation, visually depicted in Supporting Information Figure S2. The FD method sets the nanoparticle separation between adjacent neighbors as the sum of the two nanoparticles' radii plus 0.5 nm; the 0.5 nm separation distance is the average separation distance between close-packed nanoparticles for the 9% dispersity systems. The VD method maintains the separation distance between the original nanoparticle and its adjacent neighbors for both the original nanoparticle and the nanoparticle attempting to swap with the original nanoparticle. For example, if the separation between the original nanoparticle and two of its adjacent neighbors is 0.4 and 0.6 nm, then the separation between the nanoparticle attempting to swap and those two neighbors is 0.4 and 0.6 nm when calculating the energy for the Metropolis acceptance criteria. Both methods seek to consider the nanoparticle neighborhoods when attempting the swap. If the nanoparticle swap move is accepted, the nanoparticle packing fraction is decreased to $\eta = 0.3$ to remove nanoparticle overlap before being compressed back to close-packed using the same conjugate gradient method described above.

After generating the first generation of 36 individuals (i.e., 36 different structures), we calculate the computed scattering intensity, $I_{comp}(q)$, for each individual. We calculate the scattering for A nanoparticles, $I_{comp\ A}(q)$, and scattering of B nanoparticles, $I_{comp\ B}(q)$, using the Debye scattering equation.[69,70]

$$I_{comp\ X}(q) = \frac{1}{V_{sample}} \left[\sum_{i=1}^{N_x} f_i(q)^2 + \sum_{i=1}^{N_x} \sum_{j \neq i}^{N_x} f_i(q) \right. $$
$$\left. f_j(q) \frac{\sin(qr_{ij})}{qr_{ij}} \right] \tag{1}$$

We define $I_{comp\ X}(q)$ as the scattering intensity for nanoparticle X (either nanoparticle A or B) over all N_X nanoparticles in the system. The term r_{ij} is the distance between each pair of nanoparticles, and V_{sample} is defined so $I_{comp\ X}(q)$ is 1.0 at the lowest q considered to allow easy comparison between $I_{target}(q)$ and $I_{comp}(q)$. The spherical form factor amplitude f_i for each nanoparticle is defined as[71,72]

$$f_i(q) = \Delta \rho_i V_i \frac{3(\sin(qR_i) - qR_i \cos(qR_i))}{(qR_i)^3} \tag{2}$$

where V_i is the volume of nanoparticle i, R_i is the radius of nanoparticle i, and $\Delta \rho_i$ is the scattering length density

difference between the nanoparticle and solvent. For our case, we assume $\Delta \rho_i$ is the same for both nanoparticle types A and B (and set to 1.0 as V_{sample} already scales the scattering profile); however, they are easily adjusted for the system of interest. For example, $\Delta \rho_i$ can be set between 0.0 and 1.0 to correspond to scattering experiments with varying degrees of contrast-matching from perfectly contrast-matched to no contrast-match. Additionally, we do not incorporate instrument smearing into our $I_{comp}(q)$ or $I_{target}(q)$ calculations because our comparisons are against simulated scattering profiles, but for experimental applications, $I_{comp}(q)$ should be adjusted to account for smearing that is typically found in experimental scattering experiments.[73] We calculate $I_{comp}(q)$ for q values between 0.001 and 0.0035 Å^{-1} that correspond to distances between ~630 and ~180 nm (i.e., distances three times the average nanoparticle diameter and slightly below the average nanoparticle diameter, respectively). These values are chosen to ensure fitting was not performed at distances much smaller than the nanoparticle diameter where the nanoparticle form factor will dominate the scattering profile or at q values not easily achievable on experimental scattering instruments (for example, the commonly used NGB 30m SANS instrument at the National Institute of Standards and Technology).[33] While we choose this q range when fitting, we are not limited computationally, and we plot the $I_{target}(q)$ and $I_{comp}(q)$ over an extended q range from 0.0002 to 0.02 Å^{-1} to demonstrate that the method is not limited to a specific q range and can provide insight beyond experimentally measured q values. We found that, despite fitting over the limited q range, the GA is able to reproduce correct structures for the cases considered, unlike the alternative nanoparticle assembly characterization method, SASpdf,[61] described in the Introduction, that fails when considering a limited q range (Supporting Information Figure S3). While the $I_{comp}(q)$ calculation is not the rate-limiting step in the GA, it does require significant computational time, as the calculation of eq 1 involves going through all nanoparticles and their pairwise distances with all other nanoparticles. As the calculation time increases with the number of q values considered, the selection of the q range for fitting benefits from focusing on the regime most impacted by the assembled nanoparticles' structure. Addition of machine learning algorithms (e.g., neural networks) should allow us to compute the scattering intensity more rapidly and lead to significant reduction in the GA generation time.[74,75]

Step 2: Computing Individual Fitness and Selecting Individuals for Next Generation. For each individual in every generation, we calculate the individual's fitness. We define high fitness as a close similarity between the individual's $I_{comp}(q)$ and $I_{target}(q)$ and low fitness as a poor similarity. To determine similarity, we define a sum of log difference squared error (SSE) between each contrast-matched $I_{comp}(q)$ and $I_{target}(q)$.[62]

$$SSE = \sum \ln(\Delta q) \left(\ln \left(\frac{I_{target\ A}(q)}{I_{comp\ A}(q)} \right) \right)^2$$
$$+ \sum \ln(\Delta q) \left(\ln \left(\frac{I_{target\ B}(q)}{I_{comp\ B}(q)} \right) \right)^2 \tag{3}$$

We choose this form because of the multiple orders of magnitude of $I(q)$ values to ensure more equal weighting to all $I(q)$ values. The $\ln(\Delta q)$ term attempts to adjust for unequal q

spacing between evaluations of $I_{target}(q)$ typically found in experimental systems. The SSE decreases as the similarity between $I_{comp}(q)$ and $I_{target}(q)$ increases. For cases where $I_{target}(q)$ comes from experimental scattering with associated errors, one could also calculate the SSE with a commonly used χ^2 goodness of fit.[76] We note that the exact form of SSE can be varied as long as the error function ensures appropriate error weighting across the q range to drive the algorithm to low error structures. We then define the fitness function based on the SSE.[62]

$$\text{fitness} = X(\text{SSE}_{max} - \text{SSE}) + Y \qquad (4)$$

We define SSE_{max} as the maximum value of SSE for any individual in the current generation. The definitions for X and Y are

$$X = (c_s - 1)\frac{\max(\text{SSE}_{max} - \text{SSE})}{\max(\text{SSE}_{max} - \text{SSE}) - \text{average}(\text{SSE}_{max} - \text{SSE})} \qquad (5)$$

$$Y = (1 - X)\text{average}(\text{SSE}_{max} - \text{SSE}) \qquad (6)$$

where c_s is a scaling constant chosen to be 10. Our definition of fitness ensures that a high fitness corresponds to a $I_{comp}(q)$ closely matching $I_{target}(q)$. To ensure that low fitness solutions are not immediately discarded, causing the GA to converge prematurely, the X and Y variables are used to scale the fitness.

The probability that an individual is selected, or preserved, for the next generation is proportional to the individual's fitness. Akin to natural selection, our GA selection enables the most fit individuals to continue while removing the least fit individuals; though, we scale the fitness to ensure the least fit individuals are more likely to be eliminated, not guaranteed to be eliminated. Our GA has a population of 36 individuals, where 27 are selected as described above based on their fitness. The remaining nine individuals are composed of three copies of the top three individuals with the highest fitness from any generation up to that point. This approach seeks to enrich each generation with high fitness individuals while removing low fitness individuals. While the addition of multiple copies of the top three individuals with the highest fitness in any generation may appear to bias or cause early convergence, this approach is taken because multiple generations are required to optimize the individuals to produce the highest overall fitness. Thus, by supplementing the GA with the best individuals up to that point, we provide multiple opportunities for the future generation's individuals to improve during subsequent iterations of the GA.

After selecting the next population of individuals, those individuals must undergo a process to increase diversity and ensure that no two individuals' structures are the exact same in either the current or preceding generations. We accomplish this by forcing each individual to undergo a mutation in its structure.

Step 3: Mutating Individuals. In our GA, we utilize mutations to create a new population of individuals from the previously selected individuals with the goal of generating new structures that may have an improved fitness. Before the structure can be mutated by swapping A and B nanoparticles, the two swap methods (VD and FD methods) described earlier must be updated. As previously mentioned, these nanoparticle swap methods use a unique set of Hamaker constants set by the GA to determine energetics when attempting to swap nanoparticles, as detailed in Step 1. To allow the Hamaker

constants to possibly vary between generations, we define a step where the Hamaker constants themselves can undergo a mutation or combination event. The mutation step has a probability of occurrence, PM, whereby one of the individual's Hamaker constants is randomly changed within the range (0,1). The combination step has a probability PC, and the combination step mixes all three Hamaker constants from two individuals to create a new sequence of Hamaker constants. The mutation step allows for the emergence of novel Hamaker constants, and the combination step creates the opportunity to generate individuals with a mixture of desirable Hamaker constants. We allow PM and PC to vary throughout the GA run to promote continual improvements in individuals' fitness where PC begins the GA run at 0.5 and PM at 0.01. We maintain those values until the 20th generation; after the 20th generation, if the preceding two generations have had an improvement in the top three individuals from any generation then PM decreases by 20% and PC increases by 25%. If the preceding two generations both failed to have an improvement in the top three individuals from any generation, then PM increases by 25% and PC decreases by 20%. These adjustments seek to balance the introduction of new Hamaker constants with the refinement of Hamaker constants to converge to a final optimal structure.

We mutate individual structures by randomly selecting N_{swap} A-type nanoparticles and N_{swap} B-type nanoparticles and swapping the nanoparticles utilizing the Metropolis acceptance criterion.[68] If the selected nanoparticles do not pass the Metropolis acceptance criterion,[68] the selection process is repeated by randomly selecting a different group of N_{swap} A-type nanoparticles and N_{swap} B-type nanoparticles. To calculate the energies used in the Metropolis acceptance criteria, the GA run utilizes either the fixed nanoparticle distance swap or the variable nanoparticle distance swap. N_{swap} is initially set to 200, but every 10 generations, we halve that value to allow for higher-resolution refinement as the GA proceeds. If the nanoparticle swap move is accepted, to accomplish the nanoparticle swap without a potential overlap, we first expand the simulation box and nanoparticles to an $\eta = 0.3$, where any nanoparticle overlap after swapping is minimal or nonexistent. We then compress the simulation box to achieve a close-packed final structure with $\eta \approx 0.55-0.6$ using the conjugate gradient protocol explained previously in Step 1. After obtaining the close-packed structure of nanoparticles for each individual, the GA checks if the program should terminate (see Step 4), and if not, the GA returns these new individuals to Step 2 for fitness calculations.

Step 4: Terminating the GA. The GA step is terminated when the fitness values of the top 10% of individuals from any generation plateaus for at least 10 generations. For the work described here, we find this to occur before the 120th generation. The $I_{comp}(q)$ of the A and B nanoparticles and pairwise radial distribution functions (A−A, A−B, and B−B) are reported as the average of the individuals with the highest fitness from three separate GA runs, using each nanoparticle swap method, with the reported error being the standard deviation.

II.C. Structures for GA Validation

To prove that the GA can produce the correct structure for a given scattering profile, we use the GA algorithm to analyze a range of scattering profiles from in silico experiments (i.e., molecular simulations). The procedure for generating in silico

experimental structures is based on our previously developed model and simulation methods for assembly of nanoparticle mixtures in spherical confinement to form a large assembled supraparticle.[14,15] The details are available in those papers, and a shortened procedure is provided in the Supporting Information accompanying this paper.

We show that the GA can correctly produce the A and B nanoparticle mixture structure for input scattering profiles obtained from *simulated* structures when the A and B nanoparticle mixture is randomly mixed, weakly segregated, and strongly segregated, as shown in Supporting Information Figure S4. We note that this method aims to determine the extent of nanoparticle segregation rather than the exact locations of individual nanoparticles. As the input scattering profiles are from *simulated* structures instead of real experiments, we expect the chosen simulation box size (i.e., supraparticle diameter) to limit how low in q value one can calculate a reliable $I_{target}(q)$; this is a well-known issue of simulation box's finite-size effects.[77] As a result, the lowest q value we use in the GA is also chosen based on this limitation.

The GA is applied to analyze $I_{target}(q)$ from simulations of polydisperse nanoparticle size mixtures with log-normal distributions with 9, 15, or 30% size dispersity (Supporting Information Figure S1), which are realistic in experiments. We note that the GA method as described is not suitable for monodisperse systems forming crystals for which other methods described in the Introduction may be more suitable. If one wanted to apply the GA method for such ordered crystals, modifications to the close-packed structure forming process (Step 1) would be required to form a crystalline structure instead of amorphous structure, but the remaining GA components would remain applicable. Unless otherwise mentioned, results are shown for the 9% dispersity. Finally, we apply the GA method to analyze $I_{target}(q)$ from simulations of assembled nanoparticles in spherical and cubic confinement (i.e., spherical supraparticle or cubic supraparticle) to exemplify the GA's applicability to complex geometries. For all cases, we confirm the structure returned by the GA matches the "target" structure corresponding to the $I_{target}(q)$ by ensuring the local structure matches using the radial distribution function (RDF) from simulations.

III. RESULTS AND DISCUSSION

We evaluate our GA method's performance on scattering profiles obtained from simulated experiments because in these cases we know a priori the true nanoparticle mixture structure (i.e., the "target" structure) corresponding to that $I_{target}(q)$. By comparing the output structure from our GA method against the target structure, we can demonstrate how well the GA method is able to reproduce the target (known) local packing and domain (mixed or segregated nanoparticles) structure.

III.A. Evaluation of GA Method on Spherical Supraparticle with Varying Extent of Nanoparticle Segregation and Nanoparticle Size Dispersity within the Supraparticle

We focus on assembled nanoparticles exhibiting a range of A and B nanoparticle mixing from random mixing to weak particle segregation to strong particle segregation (see snapshots in Supporting Information Figure S4) to serve as three different tests of the GA method's capabilities. Additionally, these assembled nanoparticles (or supraparticle) are in a spherical geometry, similar to supraballs formed from an emulsion assembly process.[14,15]

The random nanoparticle mixing spherical supraparticle (Figure 2) is obtained when A and B nanoparticles have the
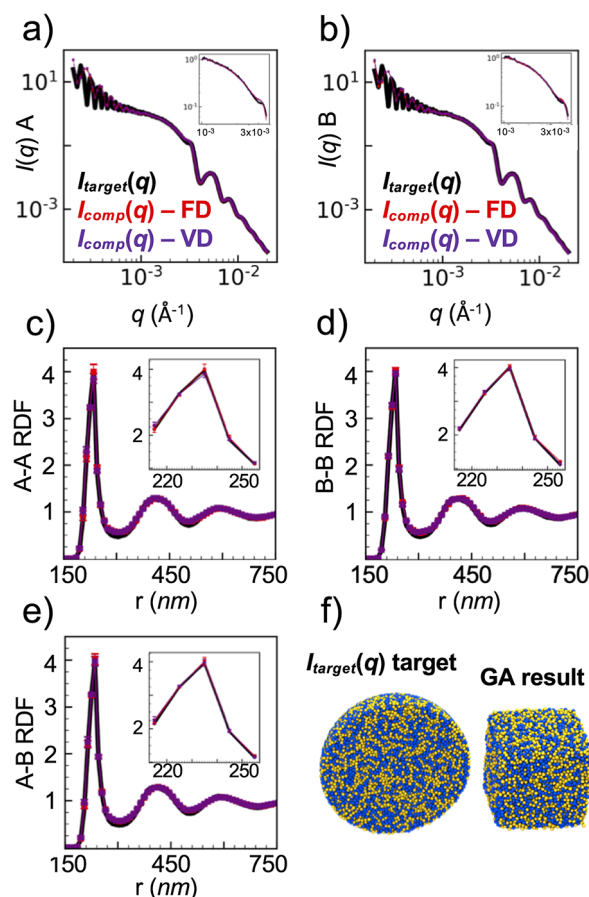

Figure 2. Evaluation of GA method for scattering intensity profiles from a simulated spherical supraparticle with random nanoparticle mixing and symmetric composition. (a,b) Scattering intensity from the simulated structure, $I_{target}(q)$, in black lines and from the GA method $I_{comp}(q)$ for both nanoparticle swap methods, VD and FD, in red and purple lines. The inset images in (a,b) provide the scattering profiles over the q range considered for the GA method. (c–e) A–A, B–B, and A–B radial distribution functions for the target structure in black, the GA method result from the FD method in red, and the VD method in purple. The inset images in (c–e) provide a zoomed-in view of the primary RDF peak. (f) Renderings of the target structure (left) and a representative GA result obtained from the FD method (right). For the GA method results, for each swap method, we present the average and standard deviation from three independent GA runs.

same like and unlike pairwise nanoparticle attractions. The $I_{target\,A}(q)$ and $I_{target\,B}(q)$ calculated from the target structure are nearly identical, with both experiencing a gradual decline at low to medium wavevector q followed by a slight oscillation at high q corresponding to the nanoparticle form factor (Figure 2a,b). As stated before, we observe large oscillations in $I_{target\,A}(q)$ and $I_{target\,B}(q)$ at the lowest q, which we attribute to the finite size of the target (simulated) structure.[77] The entire $I_{target}(q)$ shown in the inset images of Figure 2a,b is used as an input for the GA method. In the main plots in Figure 2a,b, we show the averaged computed scattering profile, $I_{comp}(q)$, from the "best" predicted structures from three separate GA runs each using the FD and VD nanoparticle swap methods. Overall, the scattering profiles from the GA method match closely to the target scattering profile for a large range of q

values even for q outside of the range used for fitting the GA method. The local nanoparticle packing is quantified using the pairwise radial distribution function, and the comparison between the target structure and the GA results is shown in Figure 2c–e. The inset in Figure 2c–e zooms in on the initial RDF peak to illustrate the match more clearly between the GA method results and the target structure. For visual reference, the best structure returned by the GA with the FD swap method is shown in Figure 2f.

In Supporting Information Figure S5, we present the $I_{comp}(q)$ and SSE during the GA run. Supporting Information Figure S5a,b shows $I_{comp}(q)$ from the best individual in the first, 25th, and 50th generation as well as the best overall individual for both the FD and VD swap methods (for clarity, only a single GA run is shown). For both FD and VD, there is a significant improvement in the scattering profile fit from the first to the 25th generation with only minor improvements afterward. Supporting Information Figure S5c depicts the SSE from the best, worst, and average individual for both FD and VD swap methods for each generation (again only a single GA run is shown for clarity). Supporting Information Figure S5c shows that the best individual SSE decreases during the GA run by about 2 orders of magnitude; the average and worst individual SSE also decrease during the GA run, but they are more constant for the first 25–50 generations. This difference is because, during the GA run, lower SSE individual structures are selected for future generations, whereas structures with larger SSE are rejected, improving both the average and worst individuals over time. Supporting Information Figure S5d shows the SSE from the best individual for both the FD and VD swap methods, and each line is the best individual for a separate GA run. Overall, both FD and VD swap methods plateau to one of the two SSE values after which the GA is not able to improve the best individual structure further; however, the minor difference in SSE of the best structure from the three separate GA runs does not significantly alter the final structure from the three GA runs as their RDFs are similar (seen as a small standard deviation in Figure 2c–e). As seen in Figure 2, the local structure (RDF) and the global structure $I_{comp}(q)$ from the GA match well with the target RDF and $I_{target}(q)$ for random nanoparticle mixing in spherical confinement. We quantify the difference between the target RDF and the RDF from the GA using both swap methods in Supporting Information Figure S5e–g. The value of these differences between target and GA output RDF is small (all values lie well within the range of −0.15 to +0.15 for RDF contact values ~4); this small difference reinforces that the GA method output structure closely matches the target structure.

Next, we test the GA for a spherical supraparticle that exhibits slight demixing of A and B nanoparticles forming small, segregated domains (Figure 3). The $I_{target}(q)$ from this slightly demixed target structure and $I_{comp}(q)$ output from the GA method are shown in Figure 3a,b. The $I_{target}(q)$ is characterized by similar features as the random nanoparticle mixing with a steeper low to medium q decline. Like Figure 2, the $I_{target}(q)$ at the lowest q in Figure 3a,b shows oscillations related to the finite system size used in the simulated experiment. As can be seen in Figure 3a,b, the GA method output possesses a global structure closely matching $I_{target}(q)$; the best structure returned by the GA method with the FD nanoparticle swap method is shown in Figure 3f, and the RDF comparison is displayed in Figure 3c–e, with the inset providing a clearer illustration of the primary peak values.
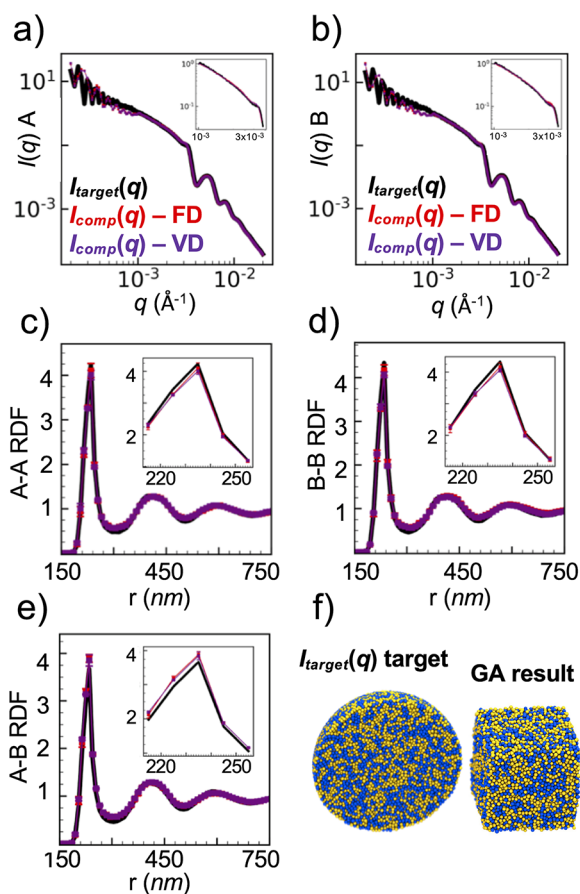

Figure 3. Evaluation of GA method for scattering intensity profiles from a simulated spherical supraparticle with weak nanoparticle segregation and symmetric composition. (a,b) Scattering intensity from the simulated structure, $I_{target}(q)$, in black lines and from the GA method, $I_{comp}(q)$, for both nanoparticle swap methods, VD and FD, in red and purple lines. The inset images in (a,b) provide the scattering profiles over the q range considered for the GA method. (c–e) A–A, B–B, and A–B radial distribution functions for the target structure in black, the GA method result from the FD method in red, and the VD method in purple. The inset images in (c–e) provide a zoomed-in view of the primary RDF peak. (f) Renderings of the target structure (left) and a representative GA result obtained from the FD method (right). For the GA method results, for each swap method, we present the average and standard deviation from three independent GA runs.

The primary peak in the A–B RDF obtained from the GA structure is slightly overestimated as compared to the target structure, whereas the A–A and B–B RDFs exhibit relatively good agreement. The difference in RDFs between the target and the GA structure in Supporting Information Figure S6e–g shows that differences between the target structure's RDF and GA structure's RDF are slightly larger than the random nanoparticle mixing especially for the primary contact peak; nonetheless, these RDF differences are small and confirm our assessment that the target structure is reproduced by the GA method.

By examining Supporting Information Figure S6, we can also understand how the GA progresses in this case. Supporting Information Figure S6a,b is similar to Supporting Information Figure S5a,b, with the computed scattering profiles drastically improving their match to experimental scattering profile from the first to 25th generation; however, Supporting Information Figure S6a,b shows a more pronounced improvement from the

25th to 50th generation, suggesting the GA run requires a few more generations to achieve a close scattering match with $I_{target}(q)$. Additionally, Supporting Information Figure S6c shows that the worst and average individual structures improve with each generation, but they are more dynamic between generations with rapid changes occurring through the GA run, unlike Supporting Information Figure S5c. Finally, Supporting Information Figure S6d is similar to Supporting Information Figure S5d as both figures show that we can achieve slightly different plateau SSE values for the best structure from three different GA runs on the same $I_{target}(q)$ and yet achieve similar final structures (Figure 3).

Next, the GA method is applied to a system with strong nanoparticle segregation (Figure 4). This system is defined by a rapid decline in $I_{target}(q)$ at low q, followed by subsequent increase at q corresponding to the nanoparticle diameter, and finally a decrease at higher q, as seen in Figure 4a,b. This
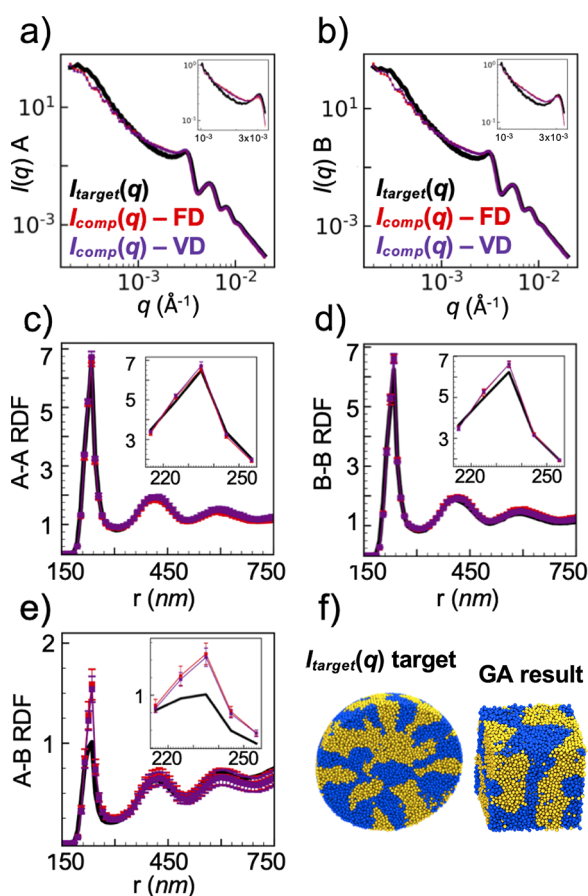

Figure 4. Evaluation of GA method for scattering intensity profiles from a simulated spherical supraparticle with strong nanoparticle segregation and symmetric composition. (a,b) Scattering intensity from the simulated structure, $I_{target}(q)$, in black lines and from the GA method, $I_{comp}(q)$, for both nanoparticle swap methods, VD and FD, in red and purple lines. The inset images in (a,b) provide the scattering profiles over the q range considered for the GA method. (c−e) A−A, B−B, and A−B radial distribution functions for the target structure in black, the GA method result from the FD method in red, and the VD method in purple. The inset images in (c−e) provide a zoomed-in view of the primary RDF peak. (f) Renderings of the target structure (left) and a representative GA result obtained from the FD method (right). For the GA method results, for each swap method, we present the average and standard deviation from three independent GA runs.

scattering profile is representative of the formation of large nanoparticle domains as seen in the image of the target structure (Figure 4f). After applying the GA method to the $I_{target}(q)$, the resulting $I_{comp}(q)$ profiles are shown in Figure 4a,b. The $I_{comp}(q)$ is not able to fully converge to $I_{target}(q)$ for either A or B scattering profile as both $I_{comp}(q)$ profiles form a shallower valley at mid-q. The corresponding B−B RDF from the best structure from GA method also has a small mismatch at the peak position with a slight overprediction of the primary peak; the A−B RDF significantly overpredicts the primary peak (Figure 4c−e). The corresponding differences between the target RDFs and the GA method's RDFs presented in Supporting Information Figure S7e−g show larger difference values as compared to that seen for random mixing and weak segregation cases presented in Figures S5e−g and S6e−g. The significant overprediction of the A−B RDF primary peak in the GA method's best structure as compared to the target structure for this strong segregation case is because the GA method determines a phase-separated structure without clean (defect-free) domains with just one nanoparticle type (Figure 4f).

The generation-to-generation changes during the GA runs (Supporting Information Figure S7a,b) show the largest improvement in SSE occurring from the first to the 25th generation. However, Supporting Information Figure S7c,d shows that the best individuals at the end of the GA run have SSE that have barely improved from the first generation, further confirming the GA method's inability to create phase-separated structures correctly and to generate clean (defect-free) segregated nanoparticle domains. The GA runs on randomly mixed and weakly segregated structures (Supporting Information Figures S5 and S6), with the GA output structures quantitatively similar to the target structure achieved SSEs on the order 10^{-2}, whereas the strongly segregated system (Supporting Information Figure S7) GA runs' SSEs are only on the order 10^{-1}. Although Supporting Information Figure S7d shows the VD swap method to have the lowest SSE for the runs, the output structure is similar to the output from the FD swap method as the difference in SSE is small.

While the results in Figure 4 and Supporting Information Figure S7 illustrate the GA method's shortcoming in matching strongly demixed systems. Practically, in experiments, such clean, defect-free, strongly demixed, close-packed systems are not always observed, and experiments may require a biasing field to achieve such defect-free domains.[78] Additionally, while we have not fine-tuned the GA method for the strongly demixed case, if one wishes, an alternative (computationally intensive) approach using neighbor lists to specifically swap nanoparticles in the wrong nanoparticle domain may result in defect-free domains. All nanoparticles are ranked based on the number of neighbors of each type (A or B), and A-type nanoparticles that are highest ranked as being surrounded by B-type nanoparticles are swapped with B-type nanoparticles that are highest ranked as being surrounded by A-type nanoparticles. The primary drawback would be the need to have a priori knowledge of the system that our current implementation of the GA method does not require.

We compare the performance of the GA on both the random nanoparticle mixing system and the strong nanoparticle segregating system to explore the difference in the GA method's output best structure match to the target structure. Figure 5 shows the sum of log squared errors from the best overall individual as the GA proceeds, with each run of the GA shown as a separate line. For the better performing random
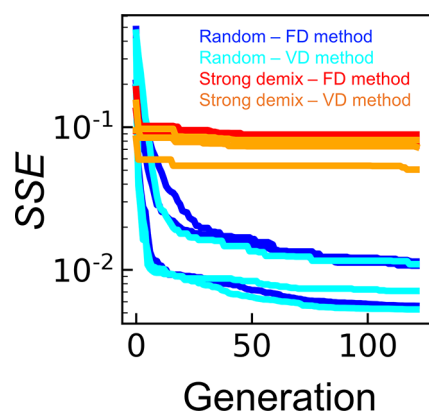
Figure 5. Comparison between sum of log squared error for the best individual per GA run for a simulated spherical supraparticle with random nanoparticle mixing and strong nanoparticle segregation. For the random nanoparticle systems, blue lines are the SSE for the best individual using the FD nanoparticle swap method (shown for each of the three GA runs), and cyan lines are the SSE for the best individual using the VD nanoparticle swap method (shown for each of the three GA runs). For the strong nanoparticle segregating systems, red lines are the SSE for the best individual using the FD nanoparticle swap method (shown for each of the three GA runs), and orange lines are the SSE for the best individual using the VD nanoparticle swap method (shown for each of the three GA runs). For all cases, each GA run is shown as a separate line, and all systems are symmetric in composition.

nanoparticle mixing system, the GA initially starts at a higher SSE (poorer $I_{comp}(q)$ and $I_{target}(q)$ match); however, over the course of the run, the GA can significantly reduce the SSE allowing for a close final match between the $I_{comp}(q)$ and $I_{target}(q)$ and ultimately between the target structure and GA method's best structure. In the case of the strong nanoparticle segregating system, the GA method begins at a lower initial SSE, but the GA method does not further reduce the SSE comparatively.

Overall, as expected, the results in Figures 2−5 show that the GA method does find the assembled nanoparticles' structure whose $I_{comp}(q)$ matches up with the input $I_{target}(q)$, and the better the match between the $I_{comp}(q)$ and $I_{target}(q)$, the closer is the RDF match between the target structure [that gave the $I_{target}(q)$] and GA method's determined structure.

The results in Figures 2−5 were for A and B nanoparticle size dispersity of 9%; next, we increase the A and B nanoparticle dispersity to 15% to examine the performance of the GA method as nanoparticle size dispersity increases. As before, we examine the same three nanoparticle mixing types (random mixing, weak segregation, and strong segregation). The random mixing with 15% nanoparticle dispersity possesses a similar $I_{target}(q)$ as the random mixing structure with 9% dispersity with dampened oscillations at high q due to the increased dispersity (Figure 6a,b). Figure 6a,b presents the $I_{comp}(q)$ obtained from the GA method, revealing the excellent match to the $I_{target}(q)$. The local packing in the best structures from the GA method likewise attain high similarity with the local packing in the target structure (Figure 6c−e). The RDF difference between target and GA output structure shown in Supporting Information Figure S8e−g is only slightly larger for 15% dispersity than the corresponding values for 9% dispersity (Supporting Information Figure S5e−g). Overall, Figure 6 illustrates that increasing the nanoparticle dispersity does not
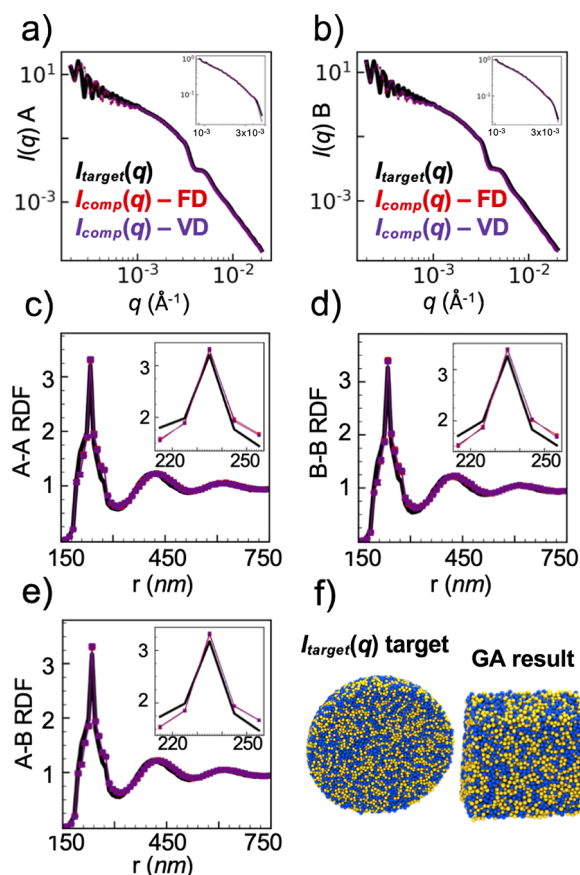

Figure 6. Evaluation of GA method for scattering intensity profiles from a simulated spherical supraparticle with random nanoparticle mixing, 15% nanoparticle dispersity, and symmetric composition. (a,b) Scattering intensity from the simulated structure, $I_{target}(q)$, in black lines and from the GA method, $I_{comp}(q)$, for both nanoparticle swap methods, VD and FD, in red and purple lines. The inset images in (a,b) provide the scattering profiles over the q range considered for the GA method. (c−e) A−A, B−B, and A−B radial distribution functions for the target structure in black, the GA method result from the FD method in red, and the VD method in purple. The inset images in (c−e) provide a zoomed-in view of the primary RDF peak. (f) Renderings of the target structure (left) and a representative GA result obtained from the FD method (right). For the GA method results, for each swap method, we present the average and standard deviation from three independent GA runs.

adversely impact the ability of the GA method to analyze $I_{target}(q)$ for structures with random nanoparticle mixing.

Interestingly, Supporting Information Figure S8c depicts that the GA SSE per generation for the average and worst individual for 15% dispersity follows a drastically different trend than the GA runs for the analogous system at 9% dispersity (Supporting Information Figure S5), where the SSE for both trends down toward smaller error. Instead, Supporting Information Figure S8c shows that the average and worst individuals appear to maintain a higher SSE than the GA runs begin with. This oddity is related to the increased dispersity having a larger impact on the structure formation part of the GA causing an increased SSE as individuals are not as closely packed with the larger dispersity. Despite that issue, the best individuals have low SSE, as can be seen in Supporting Information Figure S8c,d, and Figure S8d shows that all GA runs approach a single similar SSE value. Thus, while the increased nanoparticle dispersity impacts the average and worst

individuals, the GA is still able to improve the best structure to closely match the target structure.

The weakly segregating nanoparticle system with 15% nanoparticle size dispersity (Figure 7) shows that increasing
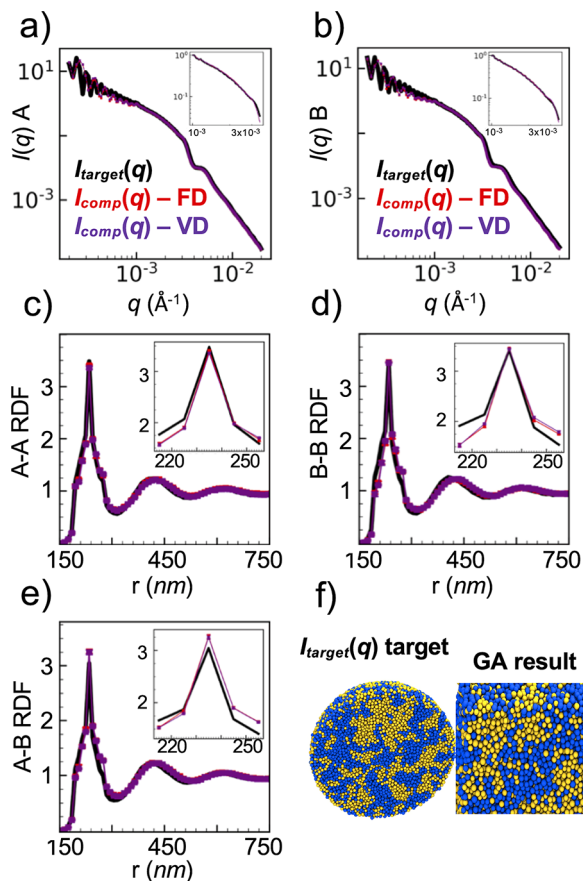

Figure 7. Evaluation of GA method for scattering intensity profiles from a simulated spherical supraparticle with weak nanoparticle segregation, 15% nanoparticle dispersity, and symmetric composition. (a,b) Scattering intensity from the simulated structure, $I_{target}(q)$, in black lines and from the GA method, $I_{comp}(q)$, for both nanoparticle swap methods, VD and FD, in red and purple lines. The inset images in (a,b) provide the scattering profiles over the q range considered for the GA method. (c−e) A−A, B−B, and A−B radial distribution functions for the target structure in black, the GA method result from the FD method in red, and the VD method in purple. The inset images in (c−e) provide a zoomed-in view of the primary RDF peak. (f) Renderings of the target structure (left) and a representative GA result obtained from the FD method (right). For the GA method results, for each swap method, we present the average and standard deviation from three independent GA runs.

the dispersity almost entirely removes the oscillation at high q compared to the analogous system with 9% dispersity. After running the GA method, the resulting structures generate $I_{comp}(q)$ that obtains an agreeable match with the $I_{target}(q)$, as seen in Figure 7a,b. The GA method determines that the best structures have nanoparticles slightly less segregated than the target structure has (Figure 7c−e). This is seen with the primary peak in both the A−A and B−B RDFs being nearly identical to the target structure and the A−B RDF primary peak being slightly larger for the GA method's best structure compared to the target structure. The RDF differences between target and GA output structure shown in Supporting Information Figure S9e−g are only slightly larger for this

weakly segregating, 15% dispersity system than the corresponding values for 9% dispersity (Supporting Information Figure S6e−g) and are similar to those seen for 15% dispersity random mixing system (Supporting Information Figure S8e−g). Overall, the GA method RDFs show good agreement with the target RDF for 15% dispersity and weakly segregating nanoparticles system, and the visual reference also confirms this idea (Figure 7f).

Supporting Information Figure S9 provides insight into the inner workings of the GA run. Unlike in Supporting Information Figure S8c, Figure S9c shows that the worst and average individuals per generation achieve lower SSE during the GA run. Taken together, Supporting Information Figure S9 illustrates that the GA method ability to provide consistently correct structures does not change with increasing nanoparticle dispersity to 15%.

The final spherical supraparticle system considered is one with strong nanoparticle segregation with 15% nanoparticle dispersity (Figure 8). Based on the previous results from the strong nanoparticle segregating system at lower dispersity, we expect the GA method to not be able to determine a structure with defect-free segregated nanoparticle domains required to match the target structure. After applying the GA method to the $I_{target}(q)$, the resulting $I_{comp}(q)$ is shown for comparison against the $I_{target}(q)$ in Figure 8a,b. The $I_{comp}(q)$ and $I_{target}(q)$ differ largely because the $I_{comp}(q)$ does not exhibit the dip in $I(q)$ at mid-q values that is seen in $I_{target}(q)$. However, the local structure shows a reasonable match between the target structure and the GA method's structures. The A−A and B−B RDFs from the GA method's best structure and the target structure achieve good agreement even with the primary peak height, as depicted in Figure 8c,d. From Figure 8e, we see that the A−B RDF primary peak height is higher for the GA method's best structure versus the target structure, similar to the results we see with strong nanoparticle segregation system at 9% dispersity with both mismatches caused by nonclean nanoparticle domains (Figure 4e). As also seen for randomly mixed and weakly segregating systems, the quantitative differences between the GA method's structure and target structure are slightly larger at 15% dispersity than at 9% dispersity for this strong nanoparticle segregation, as well.

Unlike the previous systems, Supporting Information Figure S10a,b shows that the scattering profile from the best individual has noticeable changes from the first to the 25th to the 50th generation, whereas all previous systems studied have little to no changes after the 25th generation. Additionally, this 15% dispersity, strongly segregating nanoparticle system suffers from the best individual's SSE remaining relatively constant throughout the GA run, as can be seen in Supporting Information Figure S10c,d.

We have also considered using as input $I_{target\ A}(q)$ and $I_{target\ AB}(q)$, where $I_{target\ AB}(q)$ is the A−B cross-term scattering profile. We detail the modifications to incorporate the $I_{target\ AB}(q)$ instead of the $I_{target\ B}(q)$ in the Supporting Information, with Figure S11 demonstrating that this approach does not improve our results for the case of strong nanoparticle segregation.

For these strongly segregating systems, we quantify the impact of nanoparticle dispersity specifically on the performance of the GA method by comparing the SSE of the best structures from the GA method applied to systems with 9 and 15% nanoparticle dispersity. Overall, the 9% nanoparticle dispersity system achieves a lower SSE than the 15% dispersity
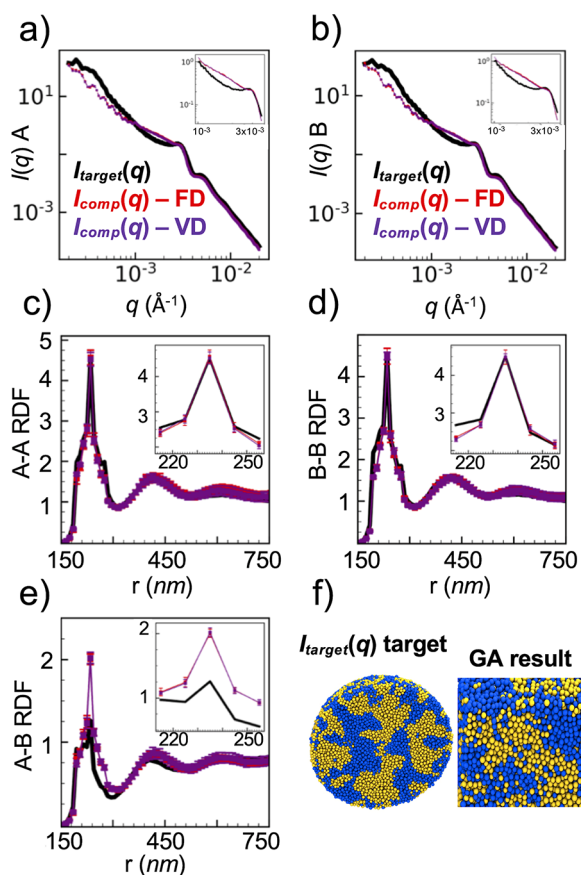
Figure 8. Evaluation of GA method for scattering intensity profiles from a simulated spherical supraparticle with strong nanoparticle segregation, 15% nanoparticle dispersity, and symmetric composition. (a,b) Scattering intensity from the simulated structure, $I_{target}(q)$, in black lines and from the GA method, $I_{comp}(q)$, for both nanoparticle swap methods, VD and FD, in red and purple lines. The inset images in (a,b) provide the scattering profiles over the q range considered for the GA method. (c–e) A–A, B–B, and A–B radial distribution functions for the target structure in black, the GA method result from the FD method in red, and the VD method in purple. The inset images in (c–e) provide a zoomed-in view of the primary RDF peak. (f) Renderings of the target structure (left) and a representative GA result obtained from the FD method (right). For the GA method results, for each swap method, we present the average and standard deviation from three independent GA runs.
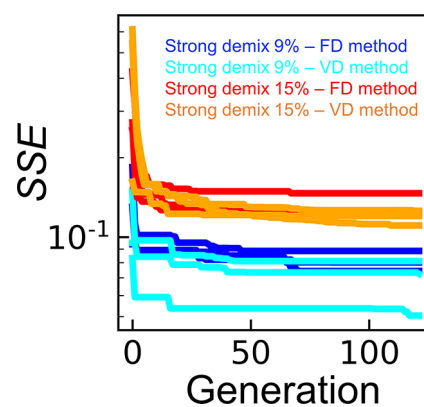

Figure 9. Comparison between sum of log squared error for the best individual per GA run for a simulated spherical supraparticle with strong nanoparticle segregation and nanoparticle dispersity of 9 and 15%. For the 9% nanoparticle dispersity systems, blue lines are the SSE for the best individual using the FD nanoparticle swap method (shown for each of the three GA runs), and cyan lines are the SSE for the best individual using the VD nanoparticle swap method (shown for each of the three GA runs). For the 15% nanoparticle dispersity systems, red lines are the SSE for the best individual using the FD nanoparticle swap method (shown for each of the three GA runs), and orange lines are the SSE for the best individual using the VD nanoparticle swap method (shown for each of the three GA runs). For all cases, each GA run is shown as a separate line, and all systems are symmetric in composition.

system (Figure 9); however, the magnitude of the SSE difference between the 9 and 15% dispersity systems is minor, especially compared to the difference in performance between the strong nanoparticle segregating system and the random mixing system (Figure 5). Overall, the change in nanoparticle dispersity has a minor impact on the ability of the GA for strong nanoparticle segregating systems, with the major issue remaining its inability to create clean nanoparticle domains. We resolve this issue at the end of the next section.

We have also considered increasing the nanoparticle size dispersity to 30% for random nanoparticle mixing (Supporting Information Figures S12 and S13), weak nanoparticle segregation (Supporting Information Figures S14 and S15), and strong nanoparticle segregation (Supporting Information Figures S16 and S17). We note that the GA method for the random and weakly segregating cases determines a structure with a close match between $I_{comp}(q)$ and $I_{target}(q)$. Similar to the cases of 9 and 15% dispersity, the strongly segregating

system with 30% dispersity overpredicts the A–B RDF due to the "clean" nanoparticle domains in the target structure. We note that the error bars for both GA output RDFs are larger for 30% dispersity than seen for previous cases of lower dispersity, indicating that as dispersity increases, multiple runs are required to average to the target structure. Furthermore, the differences between the GA method's output RDF and the corresponding target RDF are also larger for 30% dispersity, as compared to the corresponding 9 and 15% dispersity, confirming that the GA method's ability to determine the target structure gets worse with increasing particle dispersity.

While all the results so far were for a symmetric nanoparticle composition, we also explore one asymmetric nanoparticle composition of 1:3 A/B and 9% dispersity in Supporting Information Figures S18–S23 and present a brief discussion with those figures.

III.B. Evaluation of GA Method to Analyze Structure of Assembled Nanoparticles within a Film (or a Cubic Supraparticle)

After applying the GA method to spherical supraparticles, we also seek to illustrate the method's applicability to analyzing nanoparticles' structure within a cubic geometry, which would correspond to a thin-film experiment. As done in the previous section, the input $I_{target}(q)$ for all results in this section are from *simulated* experiments. We choose to focus our attention only on the weakly and strongly nanoparticle segregating cases as the GA method produced the lesser agreement with target structures in those cases as compared to the randomly mixed case in the spherical geometry.

Figure 10 describes the results for a weakly segregating cubic supraparticle with 9% nanoparticle dispersity. Figure 10a,b displays the scattering profile comparisons demonstrating the close match between the $I_{target}(q)$ and the $I_{comp}(q)$. The three RDFs shown in Figure 10c–e indicate that the structure
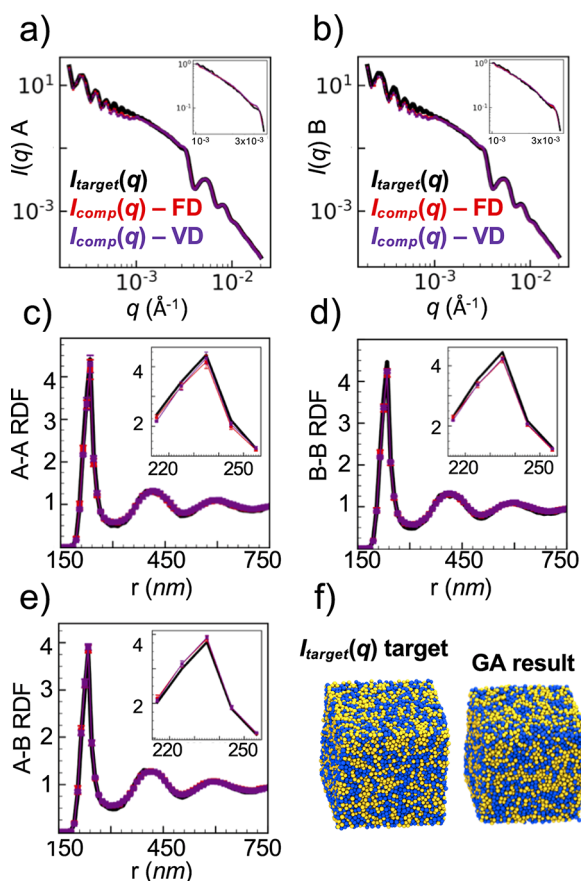
Figure 10. Evaluation of GA method for scattering intensity profiles from a simulated cubic supraparticle with weak nanoparticle segregation, 9% nanoparticle dispersity, and symmetric composition. (a,b) Scattering intensity from the simulated structure, $I_{target}(q)$, in black lines and from the GA method, $I_{comp}(q)$, for both nanoparticle swap methods, VD and FD, in red and purple lines. The inset images in (a,b) provide the scattering profiles over the q range considered for the GA method. (c−e) A−A, B−B, and A−B radial distribution functions for the target structure in black, the GA method result from the FD method in red, and the VD method in purple. The inset images in (c−e) provide a zoomed-in view of the primary RDF peak. (f) Renderings of the target structure (left) and a representative GA result obtained from the FD method (right). For the GA method results, for each swap method, we present the average and standard deviation from three independent GA runs.

determined by the GA method agrees with the target. This is also confirmed visually in Figure 10f.

Interestingly, the GA method when applied to $I_{target}(q)$ from a cubic supraparticle (Supporting Information Figure S24) takes more generations to converge to the optimal structure as compared to the analogous spherical supraparticle (Supporting Information Figures S6 and S9). This can be seen by comparing the generation where the SSE plateaus in Supporting Information Figure S24d with that in Figures S6d and S9d; the thin film (cubic supraparticle) case requires nearly double the number of generations as the spherical supraparticle to converge. The extra generations are required because the simulated thin-film target structure possesses similar number of nanoparticles as are in each GA run unlike the simulated spherical supraparticle, which has roughly 3× the number of nanoparticles. While more generations are required before the SSE value plateaus, Supporting Information Figure S24a,b illustrates that the scattering profile does not vary much

after the 25th generation, suggesting subsequent generations have only minor impact on the GA method's best structure. Ultimately, despite taking more generations to converge, the GA methods perform as well in analyzing the $I_{target}(q)$ from a cubic target supraparticle as it does for a spherical supraparticle; this is based on the similar value of differences between the target RDF and the GA method's RDF for both cubic (Supporting Information Figure S24e−g) and spherical (Supporting Information Figure S6e−g) supraparticles with weakly segregating nanoparticles.

Figure 11 shows the GA method results for the strong nanoparticle segregating cubic supraparticle with 9% nanoparticle size dispersity. The scattering profile for this case exhibits a rapid decline in $I_{target}(q)$ at low q, followed by the increase at q near the nanoparticle diameter, and a further decline at the highest q (Figure 11a,b). The major difference between this input $I_{target}(q)$ and the analogous $I_{target}(q)$ for
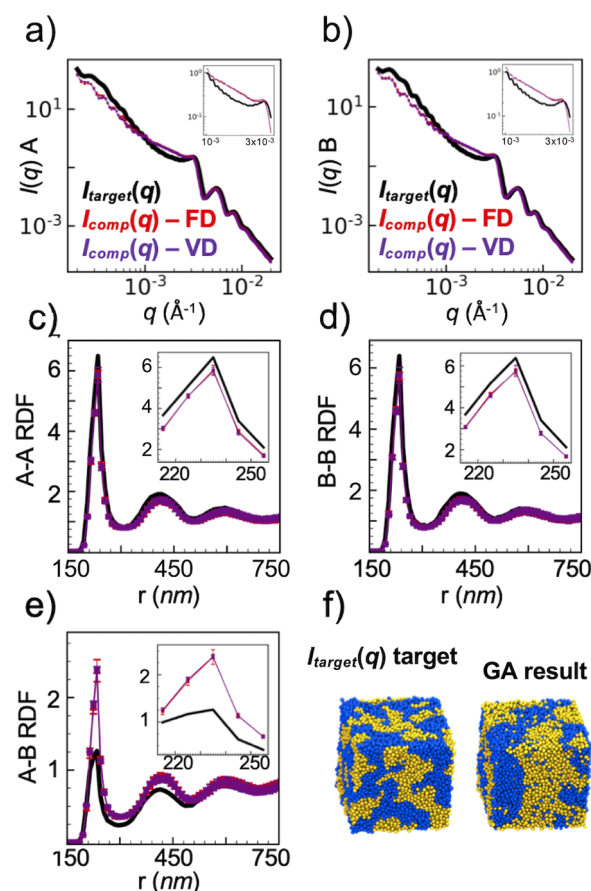

Figure 11. Evaluation of GA method for scattering intensity profiles from a simulated cubic supraparticle with strong nanoparticle segregation, 9% nanoparticle dispersity, and symmetric composition. (a,b) Scattering intensity from the simulated structure, $I_{target}(q)$, in black lines and from the GA method, $I_{comp}(q)$, for both nanoparticle swap methods, VD and FD, in red and purple lines. The inset images in (a,b) provide the scattering profiles over the q range considered for the GA method. (c−e) A−A, B−B, and A−B radial distribution functions for the target structure in black, the GA method result from the FD method in red, and the VD method in purple. The inset images in (c−e) provide a closer examination of the primary RDF peak. (f) Renderings of the target structure (left) and a representative GA result obtained from the FD method (right). For the GA method results, for each swap method, we present the average and standard deviation from three independent GA runs.

strongly segregating spherical supraparticle is the increased fluctuation for most of the low- and middle-q regions due to the smaller system size in the simulated experiment (i.e., few number of nanoparticles in the simulation of the nanoparticles in cubic geometry). After the GA method was applied to the $I_{\text{target}}(q)$, the resulting $I_{\text{comp}}(q)$ are displayed in Figure 11a,b. Like the analogous spherical supraparticle results, the GA method outputs a structure with a computed scattering profile that deviates slightly from $I_{\text{target}}(q)$. The GA method's RDFs have significant deviations for all pairs A−A, B−B, and A−B (Figure 11c−e) from that of the target structure; as seen in the analogous spherical case, there is a large overprediction by the GA method for the A−B RDF. The differences between the GA method's RDF and the target RDF are quantified in Supporting Information Figure S25e−g. As compared to the 9% dispersity spherical geometry, these differences are larger for the cubic supraparticle for all three RDFs, likely a result of the smaller number of nanoparticles in the cubic geometry compared to the spherical supraparticle. This inability of the GA method to reproduce the target structure arises from the difficulty to create defect-free nanoparticle domains without the presence of the other nanoparticle type. Figure 11f shows the defect-free target structure and the GA method's structure with defects (e.g., yellow particles in blue particle domains). Supporting Information Figure S25 also provides the progression of the GA method with generations. The thin-film (cubic) geometry of the simulated structure that provided input $I_{\text{target}}(q)$ does not improve the GA method's ability to generate structures with neat/pure nanoparticle domains. As seen in analogous spherical systems with strong nanoparticle segregation (Supporting Information Figures S7 and S10), Figure S25c,d shows that the best individuals only marginally reduce SSE over generations. However, the SSE for these GA runs are an order of magnitude above the 10^{-2} value that gives an output structure with a close match to the target structure (based on all systems considered in this work). This shortcoming is not a function of the supraparticle geometry but instead the strong nanoparticle segregation in the target structure.

We test if the GA method can perform better if the target (simulated experiment) structure exhibits strong segregation with defects (e.g., structures with some nanoparticles of B type in the A-rich domains, as seen in Figure 12f). This strongly segregated system with defects is likely more practical than a clean, defect-free segregated close-packed nanoparticle structure, the latter requiring biasing fields to achieve.[78] Figure 12a,b features the scattering profiles and strikingly reveals a significantly improved match between $I_{\text{target}}(q)$ and the GA method's generated best $I_{\text{comp}}(q)$. The local packing in the target structure and the GA determined best structure match very well (Figure 12c−e), especially considering the relatively poorer performance of the GA method with strong nano-particle segregation without defects (see Figures 4 and 8). The A−A, A−B, and B−B RDFs from the target structure and the GA produced structure are nearly identical, with the inset image further elucidating the primary peak position agreement (Figure 12c−e). This is confirmed quantitatively in Supporting Information Figure S26e−g with differences between target RDF and GA method's RDF being significantly smaller than that in Supporting Information Figure S25e−g (analogous system with clean, defect-free target structure). The close match between the target and the GA best result for this target structure with strong segregation and defects is illustrated over
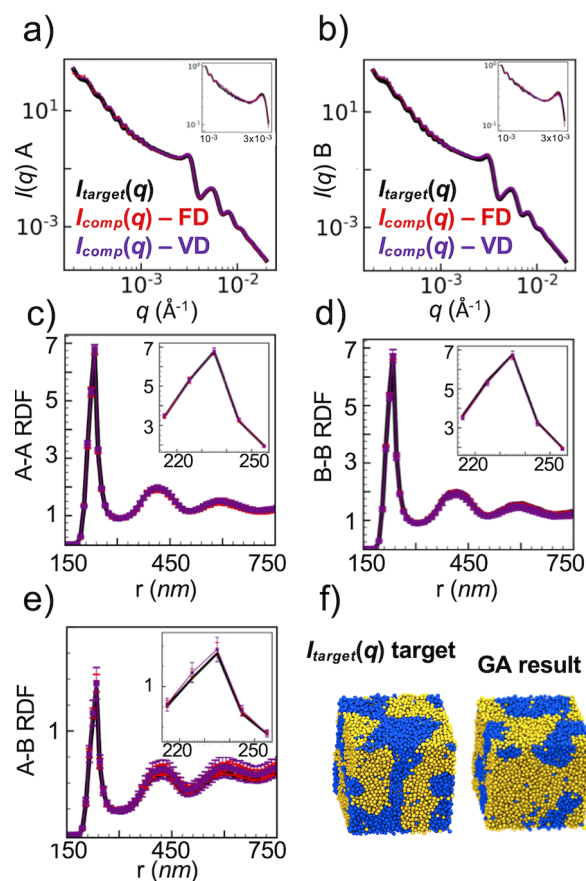

Figure 12. Evaluation of GA method for scattering intensity profiles from a simulated cubic supraparticle with 9% nanoparticle dispersity, strong nanoparticle segregation, imperfect nanoparticle domains, and symmetric composition. (a,b) Scattering intensity from the simulated structure, $I_{\text{target}}(q)$, in black lines and from the GA method, $I_{\text{comp}}(q)$, for both nanoparticle swap methods, VD and FD, in red and purple lines. The inset images in (a,b) provide the scattering profiles over the q range considered for the GA method. (c−e) A−A, B−B, and A−B radial distribution functions for the target structure in black, the GA method result from the FD method in red, and the VD method in purple. The inset images in (c−e) provide a closer examination of the primary RDF peak. (f) Renderings of the target structure (left) and a representative GA result obtained from the FD method (right). For the GA method results, for each swap method, we present the average and standard deviation from three independent GA runs.

the course of the GA run in Supporting Information Figure S26. For this case, Supporting Information Figure S26a,b shows that the scattering profile from the first generation nearly matches with the final, best computed scattering profile. This is because one of the many initial individuals provides a scattering profile that is a relatively close match to $I_{\text{target}}(q)$, which is confirmed by Supporting Information Figure S26d, achieving a very low SSE within the first few generations and maintaining that value throughout the GA run. Thus, the GA method rapidly converges to a structure that is similar to the target structure if the target does not have defect-free nanoparticle domains.

IV. CONCLUSION

In this paper, we present a new computational method for analyzing experimental scattering profiles from dilute solutions of supraparticles where each supraparticle comprises binary mixtures of nanoparticles to determine the extent of nano-

particle segregation and the local packing of the nanoparticles within each supraparticle. This genetic-algorithm-based computational analysis method overcomes the limitations of typical analysis techniques for scattering experiments that either erroneously extend approximate models to close-packed nanoparticles or require high-resolution scattering profiles over a much larger q range. Additionally, the GA method as presented allows for easy calculation of the cross-term RDF (A−B RDF) despite only using contrast-matched scattering profiles.

We evaluate this GA method by analyzing the scattering profiles from simulated experiments of a single supraparticle with varying degrees of nanoparticle mixing and size dispersity for both spherical supraparticle geometry and cubic supraparticle geometry. For all of the cases considered here, we demonstrate comparable structures between the simulated target structure, whose scattering profile is used as input to the GA, and the output structure determined by the GA method. The only target structure that exhibits weak agreement with the GA method's output structure is the cubic/spherical supraparticle with strongly segregated pure domains of each nanoparticle type. The GA method shows remarkable improvement in determining a structure close to the target structure if the target structure has strongly segregated domains with a few defects (i.e., presence of the nanoparticle of other type).

The demonstration of successful development and implementation of this GA-based computational method for scattering analysis of assembled mixtures of spherical nanoparticles in cubic/spherical geometries promises future avenues include extending this GA method to analyze anisotropic nanoparticles mixtures, blends of polymers in confinement, and structure of spherical micelles packing. For example, by maintaining a similar GA framework as our previous work which determined the form factor of complex micelle structures,[62,79,80] we foresee the ability to combine this paper's approach to our earlier papers' approaches and determine both the form factor, $P(q)$, of a complex micelle as well as the structure/correlation of multiple micelles, $S(q)$.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsmaterialsau.1c00015.

> Simulation model and method to generate the target supraparticle structures; adjustments to GA method to incorporate the A−B cross-term scattering profile; nanoparticle diameter distribution used; visual schematics of particle swapping methods; comparison of GA method to SASpdf method; GA method progression across the multiple generations illustrating the $I_{comp}(q)$ and SSE changes; GA method applied for systems with 30% nanoparticle size dispersity; GA method applied for systems with 1:3 A/B nanoparticle composition (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Arthi Jayaraman − *Department of Chemical and Biomolecular Engineering, University of Delaware, Newark, Delaware 19716, United States; Department of Materials Science and Engineering, University of Delaware, Newark, Delaware 19716, United States;* ⓞ orcid.org/0000-0002-5295-4581; Email: arthij@udel.edu

Author

Christian M. Heil − *Department of Chemical and Biomolecular Engineering, University of Delaware, Newark, Delaware 19716, United States*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acsmaterialsau.1c00015

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Grzelczak, M.; Vermant, J.; Furst, E. M.; Liz-Marzán, L. M. Directed self-assembly of nanoparticles. *ACS Nano* **2010**, *4* (7), 3591−3605.

(2) McGorty, R.; Fung, J.; Kaz, D.; Manoharan, V. N. Colloidal self-assembly at an interface. *Mater. Today* **2010**, *13* (6), 34−42.

(3) Böker, A.; He, J.; Emrick, T.; Russell, T. P. Self-assembly of nanoparticles at interfaces. *Soft Matter* **2007**, *3* (10), 1231−1248.

(4) Liu, J.; Wickramaratne, N. P.; Qiao, S. Z.; Jaroniec, M. Molecular-based design and emerging applications of nanoporous carbon spheres. *Nat. Mater.* **2015**, *14* (8), 763−774.

(5) Lemaster, J. E.; Wang, Z.; Hariri, A.; Chen, F.; Hu, Z.; Huang, Y.; Barback, C. V.; Cochran, R.; Gianneschi, N. C.; Jokerst, J. V. Gadolinium doping enhances the photoacoustic signal of synthetic melanin nanoparticles: a dual modality contrast agent for stem cell imaging. *Chem. Mater.* **2019**, *31* (1), 251−259.

(6) Boles, M. A.; Engel, M.; Talapin, D. V. Self-assembly of colloidal nanocrystals: from intricate structures to functional materials. *Chem. Rev.* **2016**, *116* (18), 11220−11289.

(7) Xu, Z.; Wang, L.; Fang, F.; Fu, Y.; Yin, Z. A review on colloidal self-assembly and their applications. *Curr. Nanosci.* **2016**, *12* (6), 725−746.

(8) Kretzschmar, I.; Song, J. H. K. Surface-anisotropic spherical colloids in geometric and field confinement. *Curr. Opin. Colloid Interface Sci.* **2011**, *16* (2), 84−95.

(9) Liu, W.; Midya, J.; Kappl, M.; Butt, H.-J.; Nikoubashman, A. Segregation in Drying Binary Colloidal Droplets. *ACS Nano* **2019**, *13* (5), 4972−4979.

(10) Cho, Y.-S.; Yi, G.-R.; Lim, J.-M.; Kim, S.-H.; Manoharan, V. N.; Pine, D. J.; Yang, S.-M. Self-organization of bidisperse colloids in water droplets. *J. Am. Chem. Soc.* **2005**, *127* (45), 15968−15975.

(11) Schulz, M.; Keddie, J. A critical and quantitative review of the stratification of particles during the drying of colloidal films. *Soft Matter* **2018**, *14* (30), 6181−6197.

(12) Tang, Y.; Grest, G. S.; Cheng, S. Stratification in drying films containing bidisperse mixtures of nanoparticles. *Langmuir* **2018**, *34* (24), 7161−7170.

(13) Bai, F.; Wang, D.; Huo, Z.; Chen, W.; Liu, L.; Liang, X.; Chen, C.; Wang, X.; Peng, Q.; Li, Y. A versatile bottom-up assembly approach to colloidal spheres from nanocrystals. *Angew. Chem.* **2007**, *119* (35), 6770−6773.

(14) Xiao, M.; Hu, Z.; Gartner, T. E., 3rd; Yang, X.; Li, W.; Jayaraman, A.; Gianneschi, N. C.; Shawkey, M. D.; Dhinojwala, A. Experimental and theoretical evidence for molecular forces driving surface segregation in photonic colloidal assemblies. *Sci. Adv.* **2019**, *5* (9), eaax1254.

(15) Gartner, T. E.; Heil, C. M.; Jayaraman, A. Surface composition and ordering of binary nanoparticle mixtures in spherical confinement. *Molecular Systems Design & Engineering* **2020**, 5 (4), 864−875.

(16) Kawamura, A.; Kohri, M.; Yoshioka, S.; Taniguchi, T.; Kishikawa, K. Structural color tuning: Mixing melanin-like particles with different diameters to create neutral colors. *Langmuir* **2017**, 33 (15), 3824−3830.

(17) Dutta, H. S.; Goyal, A. K.; Srivastava, V.; Pal, S. Coupling light in photonic crystal waveguides: A review. *Photonics and Nanostructures-Fundamentals and Applications* **2016**, 20, 41−58.

(18) Echeverri, M.; Patil, A.; Hu, Z.; Shawkey, M. D.; Gianneschi, N. C.; Dhinojwala, A. Printing a Wide Gamut of Saturated Structural Colors Using Binary Mixtures, With Applications in Anticounterfeiting. *ACS Appl. Mater. Interfaces* **2020**, 12 (17), 19882−19889.

(19) Zhao, Y.; Xie, Z.; Gu, H.; Zhu, C.; Gu, Z. Bio-inspired variable structural color materials. *Chem. Soc. Rev.* **2012**, 41 (8), 3297−3317.

(20) Shipway, A. N.; Katz, E.; Willner, I. Nanoparticle arrays on surfaces for electronic, optical, and sensor applications. *ChemPhysChem* **2000**, 1 (1), 18−52.

(21) Mourdikoudis, S.; Pallares, R. M.; Thanh, N. T. Characterization techniques for nanoparticles: comparison and complementarity upon studying nanoparticle properties. *Nanoscale* **2018**, 10 (27), 12871−12934.

(22) Modena, M. M.; Rühle, B.; Burg, T. P.; Wuttke, S. Nanoparticle characterization: What to measure? *Adv. Mater.* **2019**, 31 (32), 1901556.

(23) Berbel Manaia, E.; Paiva Abucafy, M.; Chiari-Andreo, B. G.; Lallo Silva, B.; Oshiro-Junior, J. A.; Chiavacci, L. Physicochemical characterization of drug nanocarriers. *Int. J. Nanomed.* **2017**, 12, 4991.

(24) Melnichenko, Y. B.; Wignall, G. D. Small-angle neutron scattering in materials science: Recent practical applications. *J. Appl. Phys.* **2007**, 102 (2), 021101.

(25) Chu, B.; Liu, T. Characterization of nanoparticles by scattering techniques. *J. Nanopart. Res.* **2000**, 2 (1), 29−41.

(26) Hollamby, M. J. Practical applications of small-angle neutron scattering. *Phys. Chem. Chem. Phys.* **2013**, 15 (26), 10566−10579.

(27) Luo, Z.; Marson, D.; Ong, Q. K.; Loiudice, A.; Kohlbrecher, J.; Radulescu, A.; Krause-Heuer, A.; Darwish, T.; Balog, S.; Buonsanti, R.; Svergun, D. I.; Posocco, P.; Stellacci, F. Quantitative 3D determination of self-assembled structures on nanoparticles using small angle neutron scattering. *Nat. Commun.* **2018**, 9 (1), 1343.

(28) Li, T.; Senesi, A. J.; Lee, B. Small angle X-ray scattering for nanoparticle research. *Chem. Rev.* **2016**, 116 (18), 11128−11180.

(29) Schryvers, D.; Cao, S.; Tirry, W.; Idrissi, H.; Van Aert, S. Advanced three-dimensional electron microscopy techniques in the quest for better structural and functional materials. *Sci. Technol. Adv. Mater.* **2013**, 14, 014206.

(30) Nan, N.; Wang, J. FIB-SEM three-dimensional tomography for characterization of carbon-based materials. *Adv. Mater. Sci. Eng.* **2019**, 2019, 8680715.

(31) Holzer, L.; Cantoni, M.Review of FIB-tomography. *Nanofabrication using focused ion and electron beams: Principles and applications*; Oxford University Press, 2012, pp 410−435.

(32) Čalkovský, M.; Müller, E.; Meffert, M.; Firman, N.; Mayer, F.; Wegener, M.; Gerthsen, D. Comparison of segmentation algorithms for FIB-SEM tomography of porous polymers: Importance of image contrast for machine learning segmentation. *Mater. Charact.* **2021**, 171, 110806.

(33) Glinka, C.; Barker, J.; Hammouda, B.; Krueger, S.; Moyer, J.; Orts, W. The 30 m small-angle neutron scattering instruments at the National Institute of Standards and Technology. *J. Appl. Crystallogr.* **1998**, 31 (3), 430−445.

(34) Gosecka, M.; Gosecki, M. Characterization methods of polymer core−shell particles. *Colloid Polym. Sci.* **2015**, 293 (10), 2719−2740.

(35) Blin, J.-L.; Impéror-Clerc, M. Mechanism of self-assembly in the synthesis of silica mesoporous materials: in situ studies by X-ray and neutron scattering. *Chem. Soc. Rev.* **2013**, 42 (9), 4071−4082.

(36) Ballauff, M. Analysis of Polymer Colloids by Small-Angle X-Ray and Neutron Scattering: Contrast Variation. *Adv. Eng. Mater.* **2011**, 13 (8), 793−802.

(37) Garcia-Diez, R.; Gollwitzer, C.; Krumrey, M. Nanoparticle characterization by continuous contrast variation in small-angle X-ray scattering with a solvent density gradient. *J. Appl. Crystallogr.* **2015**, 48 (1), 20−28.

(38) Lombardo, D.; Calandra, P.; Kiselev, M. A. Structural Characterization of Biomaterials by Means of Small Angle X-rays and Neutron Scattering (SAXS and SANS), and Light Scattering Experiments. *Molecules* **2020**, 25 (23), 5624.

(39) Sakurai, S. Recent developments in polymer applications of synchrotron small-angle X-ray scattering. *Polym. Int.* **2017**, 66 (2), 237−249.

(40) Sen, D.; Khan, A.; Bahadur, J.; Mazumder, S.; Sapra, B. Use of small-angle neutron scattering to investigate modifications of internal structure in self-assembled grains of nanoparticles synthesized by spray drying. *J. Colloid Interface Sci.* **2010**, 347 (1), 25−30.

(41) Harada, T.; Matsuoka, H.; Yamamoto, T.; Yamaoka, H.; Lin, J.; Agamalian, M. M.; Wignall, G. D. The structure of colloidal alloy crystals revealed by ultra-small-angle neutron scattering. *Colloids Surf., A* **2001**, 190 (1−2), 17−24.

(42) Carr, A. J.; Liu, W.; Yager, K. G.; Routh, A. F.; Bhatia, S. R. Evidence of stratification in binary colloidal films from microbeam X-ray scattering: Toward optimizing the evaporative assembly processes for coatings. *ACS Applied Nano Materials* **2018**, 1 (8), 4211−4217.

(43) Blum, L.; Stell, G. Polydisperse systems. I. Scattering function for polydisperse fluids of hard or permeable spheres. *J. Chem. Phys.* **1979**, 71 (1), 42−46.

(44) Salacuse, J.; Stell, G. Polydisperse systems: Statistical thermodynamics, with applications to several models including hard and permeable spheres. *J. Chem. Phys.* **1982**, 77 (7), 3714−3725.

(45) Menon, S.; Manohar, C.; Rao, K. S. A new interpretation of the sticky hard sphere model. *J. Chem. Phys.* **1991**, 95 (12), 9186−9190.

(46) Baxter, R. Percus−Yevick equation for hard spheres with surface adhesion. *J. Chem. Phys.* **1968**, 49 (6), 2770−2774.

(47) Ottewill, R.; Rennie, A.; Johnson, G. Small-angle neutron scattering studies on binary mixtures of polystyrene and perfluorinated particles. *Adv. Colloid Interface Sci.* **2003**, 100, 585−611.

(48) De Kruif, C.; Briels, W.; May, R.; Vrij, A. Hard-sphere colloidal silica dispersions. The structure factor determined with SANS. *Langmuir* **1988**, 4 (3), 668−676.

(49) Schmidt, P. Modern Aspects of Small-Angle Scattering. *NATO Science Series C (Springer, New York, 1994)* **1995**, 1−55.

(50) Zemb, T.; Lindner, P. *Neutrons, X-rays and light: scattering methods applied to soft condensed matter*; North-Holland, 2002.

(51) McGreevy, R.; Pusztai, L. Reverse Monte Carlo simulation: a new technique for the determination of disordered structures. *Mol. Simul.* **1988**, 1 (6), 359−367.

(52) Gereben, O.; Pusztai, L.; McGreevy, R. RMCSANS—modelling the inter-particle term of small angle scattering data via the reverse Monte Carlo method. *J. Phys.: Condens. Matter* **2010**, 22 (40), 404216.

(53) Musino, D.; Genix, A.-C.; Chauveau, E.; Bizien, T.; Oberdisse, J. Structural identification of percolation of nanoparticles. *Nanoscale* **2020**, 12 (6), 3907−3915.

(54) Hagita, K.; Tominaga, T.; Sone, T. Large-scale reverse Monte Carlo analysis for the morphologies of silica nanoparticles in end-modified rubbers based on ultra-small-angle X-ray scattering data. *Polymer* **2018**, 135, 219−229.

(55) Tóth, G. Simultaneous Monte Carlo determination of particle size distribution and pair-correlation function of spherical colloids from a diffraction experiment. *Langmuir* **1999**, 15 (20), 6718−6723.

(56) Beaucage, G. Approximations leading to a unified exponential/power-law approach to small-angle scattering. *J. Appl. Crystallogr.* **1995**, 28 (6), 717−728.

(57) Oberdisse, J.; Hine, P.; Pyckhout-Hintzen, W. Structure of interacting aggregates of silica nanoparticles in a polymer matrix:

small-angle scattering and reverse Monte Carlo simulations. *Soft Matter* **2007**, *3* (4), 476−485.

(58) Larsen, A. H.; Pedersen, J. S.; Arleth, L. Assessment of structure factors for analysis of small-angle scattering data from desired or undesired aggregates. *J. Appl. Crystallogr.* **2020**, *53* (4), 991.

(59) Lu, F.; Vo, T.; Zhang, Y.; Frenkel, A.; Yager, K. G.; Kumar, S.; Gang, O. Unusual packing of soft-shelled nanocubes. *Science advances* **2019**, *5* (5), eaaw2399.

(60) Yager, K. G.; Zhang, Y.; Lu, F.; Gang, O. Periodic lattices of arbitrary nano-objects: modeling and applications for self-assembled systems. *J. Appl. Crystallogr.* **2014**, *47* (1), 118−129.

(61) Liu, C.-H.; Janke, E. M.; Li, R.; Juhás, P.; Gang, O.; Talapin, D. V.; Billinge, S. J. sasPDF: pair distribution function analysis of nanoparticle assemblies from small-angle scattering data. *J. Appl. Crystallogr.* **2020**, *53* (3), 699.

(62) Beltran-Villegas, D. J.; Wessels, M. G.; Lee, J. Y.; Song, Y.; Wooley, K. L.; Pochan, D. J.; Jayaraman, A. Computational Reverse-Engineering Analysis for Scattering Experiments on Amphiphilic Block Polymer Solutions. *J. Am. Chem. Soc.* **2019**, *141* (37), 14916−14930.

(63) Borsali, R.; Pecora, R. *Soft-matter characterization*; Springer Science & Business Media, 2008.

(64) Kumar, M.; Husain, M.; Upreti, N.; Gupta, D. Genetic algorithm: Review and application. *SSRN J.* **2010**, *2* (2), 451−454.

(65) Polak, E.; Ribiere, G. Note sur la convergence de méthodes de directions conjuguées. *ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique* **1969**, *3* (R1), 35−43.

(66) Everaers, R.; Ejtehadi, M. Interaction potentials for soft and hard ellipsoids. *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.* **2003**, *67* (4), 041710.

(67) Farr, R. S.; Groot, R. D. Close packing density of polydisperse hard spheres. *J. Chem. Phys.* **2009**, *131* (24), 244104.

(68) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. Equation of state calculations by fast computing machines. *J. Chem. Phys.* **1953**, *21* (6), 1087−1092.

(69) Debye, P. Zerstreuung von röntgenstrahlen. *Ann. Phys.* **1915**, *351* (6), 809−823.

(70) Scardi, P.; Billinge, S. J.; Neder, R.; Cervellino, A. Celebrating 100 years of the Debye scattering equation. *Acta Crystallogr., Sect. A: Found. Adv.* **2016**, *72* (6), 589.

(71) Pedersen, J. S. Analysis of small-angle scattering data from colloids and polymer solutions: modeling and least-squares fitting. *Adv. Colloid Interface Sci.* **1997**, *70* (0), 171−210.

(72) Rayleigh, L. Form factor of a homogenous sphere. *Proc. R. Soc. London, Ser. A* **1911**, *84*, 25−38.

(73) Barker, J.; Pedersen, J. Instrumental smearing effects in radially symmetric small-angle neutron scattering by numerical and analytical methods. *J. Appl. Crystallogr.* **1995**, *28* (2), 105−114.

(74) Patra, T. K.; Meenakshisundaram, V.; Hung, J.-H.; Simmons, D. S. Neural-network-biased genetic algorithms for materials design: Evolutionary algorithms that learn. *ACS Comb. Sci.* **2017**, *19* (2), 96−107.

(75) Wessels, M. G.; Jayaraman, A. Machine Learning Enhanced Computational Reverse Engineering Analysis for Scattering Experiments (CREASE) to Determine Structures in Amphiphilic Polymer Solutions. *ACS Polym. Au* **2021**, DOI: 10.1021/acspolymer-sau.1c00015.

(76) Doucet, M.; Cho, J. H.; Alina, G.; Bakker, J.; Bouwman, W.; Butler, P.; Campbell, K.; Gonzales, M.; Heenan, R.; Jackson, A.; Juhas, P. *SasView*, version 5.0.4, 2021.

(77) Olds, D. P.; Duxbury, P. M. Efficient algorithms for calculating small-angle scattering from large model structures. *J. Appl. Crystallogr.* **2014**, *47* (3), 1077−1086.

(78) Furst, E. M. Directed self-assembly. *Soft Matter* **2013**, *9* (38), 9039−9045.

(79) Wessels, M. G.; Jayaraman, A. Computational Reverse-Engineering Analysis of Scattering Experiments (CREASE) on Amphiphilic Block Polymer Solutions: Cylindrical and Fibrillar Assembly. *Macromolecules* **2021**, *54* (2), 783−796.

(80) Lee, J. Y.; Song, Y.; Wessels, M. G.; Jayaraman, A.; Wooley, K. L.; Pochan, D. J. Hierarchical Self-Assembly of Poly (d-glucose carbonate) Amphiphilic Block Copolymers in Mixed Solvents. *Macromolecules* **2020**, *53* (19), 8581−8591.