JOURNAL OF APPLIED CRYSTALLOGRAPHY

Volume 56 (2023)

Supporting information for article:

Pair distribution function analysis of nano-object assemblies

Yugang Zhang and Oleg Gang

Supporting Information

S1. Debye equation of nano-assemblies

In the limit of kinematic scattering, for a sample of volume V with N scatterers, the scattering amplitude $\psi(\vec{q})$ from a set of elementary scatterers with electron density of $\rho_{\boldsymbol{n}}$ at position $\overrightarrow{r_n}$ is,

$$\psi(\vec{q}) = \sum_n^N \rho_{\boldsymbol{n}} \exp(i\overrightarrow{q} \cdot \overrightarrow{r_n}) \qquad (1)$$

The nano-assemblies can be considered to contain N_p particles at position r_p, and each particle contains N_a atoms at position r_a, and each atom contains N_e electrons at position r_e. Thus $r_n = r_p + r_a + r_e$,

$$\psi(\vec{q}) = \sum_p^{N_p} \exp\left(i\overrightarrow{q} \cdot \overrightarrow{r_p}\right) \sum_a^{N_a} \exp\left(i\overrightarrow{q} \cdot \overrightarrow{r_a}\right) \sum_e^{N_e} \rho_{\boldsymbol{e}} \exp\left(i\overrightarrow{q} \cdot \overrightarrow{r_e}\right) \quad (2)$$

The last term is known as atomic form factor ($f_a\left(\overrightarrow{q}\right)$ and the combined last two terms are known as particle form factor ($f(\vec{q})$.

$$f_a(\vec{q}) \equiv \sum_e^{N_e} \rho_{\boldsymbol{e}} \exp(i\overrightarrow{q} \cdot \overrightarrow{r_e}) \quad (3)$$

$$f(\vec{q}) \equiv \sum_a^{N_a} f_a\left(\overrightarrow{q}\right) \exp(i\overrightarrow{q} \cdot \overrightarrow{r_a}) \quad (4)$$

Considering scattering intensity $I(\vec{q}) = \psi^*(\vec{q})\psi(\vec{q})$, so

$$I(\vec{q}) = \left|\sum_p^{N_p} f(\vec{q}) \exp(i\overrightarrow{q} \cdot \overrightarrow{r_p})\right|^2 \quad (5)$$

Then,

$$I(\vec{q}) = \sum_j \sum_i f_j^*(\vec{q}) f_i(\vec{q}) \exp\left[i\overrightarrow{q} \cdot \left(\overrightarrow{r_i} - \overrightarrow{r_j}\right)\right]$$

$$= \sum_{ij} f_j^*(\vec{q}) f_i(\vec{q}) \exp(i\overrightarrow{q} \cdot \overrightarrow{r_{iJ}}) \qquad (6)$$

The scattering intensity can be separated into the self-scattering (i=j, namely, r_{ij}=0) and the distinct scattering (i≠j),

$$I(\vec{q}) = I_s(\vec{q}) + I_d(\vec{q})$$

$$= \sum_i f_i^*(\vec{q}) f_i(\vec{q}) + \sum_{i \neq j} f_j^*(\vec{q}) f_i(\vec{q}) \exp\left(i\overrightarrow{q} \cdot \overrightarrow{r_{iJ}}\right)$$

$$= N_p < f^2(\vec{q}) > + \sum_{i \neq j} f_j^*(\vec{q}) f_i(\vec{q}) \exp(i\overrightarrow{q} \cdot \overrightarrow{r_{iJ}}) \qquad (7)$$

The form factor intensity of the NPs is defined as:

$P(\vec{q}) = < f^2(\vec{q}) >$ (8)

The sample-averaged form factor is defined as:

$< f(\vec{q}) > = \frac{1}{N_p}\sum_i f_i(\vec{q})$ (9)

And the square of the averaged form factor becomes:

$< f(\vec{q}) >^2 = \frac{1}{{N_p}^2}\sum_{ij} f_j^*(\vec{q}) f_i(\vec{q})$ (10)

The beta factor is defined as,

$\beta(\vec{q}) = \frac{<f(\vec{q})>^2}{<f^2(\vec{q})>}$ (11)

The scattering intensity can be split into two parts, namely, $I_u(\vec{q}) = I(0 \le \vec{q} < \overrightarrow{q_m})$, the ultra-small angle scattering (USAXS) and the $I_s(\vec{q}) = I(\vec{q} > \overrightarrow{q_m})$, the small-angle scattering (SAXS). The USAXS part is usually not accessible by the conventional SAXS instruments, but it encodes the mesoscale structural information of nano-assemblies. Mathematically, $I_u(\vec{q})$ is a Fourier transformation of a shape factor, $\gamma_0(\vec{r})$, which is the characteristic function of the sample shape,

$I_u(\vec{q}) = \rho_0 N_p < f(\vec{q}) >^2 \int_0^\infty \gamma_0(\vec{r}) \exp(\overrightarrow{iq} \cdot \vec{r})\, d\vec{r}$ (12)

The shape factor is a self-convolution or autocorrelation of the shape function, $s(\vec{r})$, of the sample,

$\gamma_0(\vec{r}) = \frac{1}{V}\int s(\overrightarrow{r'}) s(\overrightarrow{r'} + \vec{r})\, d\overrightarrow{r'}$ (13)

So, by splitting the intensity into USAXS and SAXS, Debye equation for NP assemblies can be written as:

$$I(\vec{q}) = \begin{cases} N_p < f^2(\vec{q}) > + \sum_{i \ne j} f_j^*(\vec{q}) f_i(\vec{q}) \exp(\overrightarrow{iq} \cdot \overrightarrow{r_{ij}}) & \vec{q} > \overrightarrow{q_m} \\ \rho_0 N_p < f(\vec{q}) >^2 \int_0^\infty \gamma_0(\vec{r}) \exp(\overrightarrow{iq} \cdot \vec{r})\, d\vec{r} & 0 \le \vec{q} < \overrightarrow{q_m} \end{cases} \quad (14)$$

For powder diffraction, namely, assuming there are numerous identical domains covering all the possible orientations with equal probability, we can convert $\vec{q}$ to q by orientational average and $I(\vec{q})$ becomes:

$$I(q) = \begin{cases} N_p < f^2(\vec{q}) >_o + < \sum_{i \ne j} f_j^*(\vec{q}) f_i(\vec{q}) \exp(\overrightarrow{iq} \cdot \overrightarrow{r_{ij}}) >_o & q > q_m \\ \rho_0 N_p < f(\vec{q}) >^2_o < \int_0^\infty \gamma_0(\vec{r}) \exp(\overrightarrow{iq} \cdot \vec{r})\, d\vec{r} >_o & 0 \le q < q_m \end{cases} \quad (15)$$

The <>o corresponds to the orientational average.

For spherical nano-objects, $f(\vec{q}) = f(q)$. Also considering:

$$< \exp(\overrightarrow{iq} \cdot \vec{r}) >_o = \frac{\sin(qr)}{qr} \quad (16)$$

Now, we have the expression of I(q) for the assembly of spherical particles:

$$I(q) = \begin{cases} N_p < f^2(q) > + \sum_{i \neq j} f_j^*(q) f_i(q) \frac{\sin(qr_{ij})}{qr_{ij}} & \mathrm{q} > q_m \\ \rho_0 N_p < f(q) >^2 \int_0^\infty 4\pi \gamma_0(r) \frac{\sin(qr)}{qr} r^2 dr & 0 \leq q < q_m \end{cases} \quad (17)$$

Next, following the well-known equation defined in the Faber-Ziman formalism, the structure factor is written as:

$$S(q) = \frac{I(q)}{N_p < f(q) >^2} + [1 - \frac{< f(q)^2 >}{< f(q) >^2}]$$

$$= \frac{I(q)}{N_p < f(q) >^2} + \left[1 - \frac{1}{\beta(q)}\right] \quad (18)$$

Combining equation (17) and (18) and considering $\beta(q) = 1$ for $0 \leq q < q_m$, we get,

$$S(q) = \begin{cases} \sum_{i \neq j} \frac{f_j^*(q) f_i(q)}{N_p < f(q) >^2} \frac{\sin(qr_{ij})}{qr_{ij}} + 1 & \mathrm{q} > q_m \\ 4\pi\rho_0 \int_0^\infty r\gamma_0(r) \frac{\sin(qr)}{q} dr & 0 \leq q < q_m \end{cases} \quad (19)$$

In the particle assembly literature, the structure factor is defined as:

$$S'(q) = \frac{I(q)}{N_p < f(q)^2 >}$$

$$= \beta(q) S(q) + [1 - \beta(q)] \quad (20)$$

It seems like that there is a divergence in the definition of S(q) between atomic and nano-assembly literature. Actually, we will show later that equation (20) is the exact form of equation (18) by considering the polydispersity of particles.

S2. "Experimental" nano-pair distribution function

The reduced structure factor is defined as,

$$F(q) = q[S(q) - 1] \quad (21)$$

Substituting S(q) using equation (19), we get F(q) as,

$$F(q) = \begin{cases} \sum_{i \neq j} \frac{f_j^*(q) f_i(q)}{N_p < f^2(q) >} \frac{\sin(qr_{ij})}{r_{ij}} & \mathrm{q} > q_m \\ 4\pi\rho_0 \int_0^\infty r\gamma_0(r) \sin(qr)\, dr & 0 \leq q < q_m \end{cases} \quad (22)$$

For the case of q~0, the -q term is ignored and F(q~0) is in fact a FT of $4\pi\rho_0 r\gamma_0(r)$. By applying the iFT of F(q), we get the expression of the iFT-nPDF as,

$\mathcal{F}(r) = \frac{2}{\pi}\int_0^\infty F(q)\, sin(qr)\, dq$ (23)

From a practical viewpoint, considering the 'missing' USAXS scattering, it would be helpful to split the $\mathcal{F}(r)$ in two parts,

$\mathcal{F}(r) = \mathcal{N}(r) + G(r)$ (24)

The first term, $\mathcal{N}(r)$, is ascribed to the USAXS scattering,

$\mathcal{N}(r) = \frac{2}{\pi}\int_0^{q_m} F(0 \le q < q_m)\, sin(qr)\, dq$ (25)

The measured nPDF by conventional SAXS instruments is actually the second term, $G(r)$, which is well known as reduced pair distribution function in the atomic PDF literature and is written as,

$G(r) = \frac{2}{\pi}\int_{q_m}^\infty F(q > q_m)\, sin(qr)\, dq$ (26)

S3. "Computed" nano-pair distribution function

Combining equation (22-26), $\mathcal{F}(r)$ can be expressed as a function or r,

$$\mathcal{F}(r) = \frac{2}{\pi N_p}\sum_{i\neq j}\frac{1}{r_{ij}}\int_0^\infty \frac{f_j^*(q)f_i(q)}{<f^2(q)>}\sin(qr_{ij})\, sin(qr)\, dq$$

$$= \frac{2}{\pi N_p}\sum_{i\neq j}\frac{1}{r_{ij}}\,\phi\left[h_{ij}(q)\,\sin(qr_{ij})\right] \quad (27)$$

Where, ϕ denotes the iFT operation and $h_{ij}(q)$ is the normalized product of form factors of a pair of nano-objects, i and j,

$h_{ij}(q) \equiv \frac{f_j^*(q)f_i(q)}{<f^2(q)>}$ (28)

And define $H_{ij}(r)$ as the iFT of $h_{ij}(q)$,

$H_{ij}(r) \equiv \phi\left[h_{ij}(q)\right] = \int_0^\infty \frac{f_j^*(q)f_i(q)}{<f^2(q)>} sin(qr)\, dq$ (29)

The Fourier transformation of the $\sin(qr_{ij})$ part is:

$\phi\left[\sin(qr_{ij})\right] = \int_0^\infty \sin(qr_{ij})\, sin(qr)\, dq = \frac{\pi}{2}\delta(r - r_{ij})$ (30)

Applying the convolution theorem that the (inverse) Fourier transform of the product of two signals is a convolution of their (inverse) Fourier transforms, namely,

$\phi(f \cdot g) = \phi(f) \otimes \phi(g)$ (31)

Also considering, $f(x) \otimes \delta(x-a) = f(x-a)$ (32)

So, combing equations (27-32), we get $\mathcal{F}(r)$ as,

$$\mathcal{F}(r) = \frac{2}{\pi N_p}\sum_{i\neq j}\frac{1}{r_{ij}}\,\phi\left[h_{ij}(q)\right] \otimes \phi[\sin\left(qr_{ij}\right)]$$

$$= \frac{2}{\pi N_p}\sum_{i\neq j}\frac{1}{r_{ij}}\,H_{ij}(\mathrm{r}) \otimes \frac{\pi}{2}\delta\left(r-r_{ij}\right)$$

$$= \frac{1}{rN_p}\sum_{i\neq j} H_{ij}\left(r-r_{ij}\right) \qquad (33)$$

According to equation (24-26), the experimentally measured $G(r)$ the difference between $\mathcal{F}(r)$ and $\mathcal{N}(r)$. The equation (22) shows that $F(0 \le q < q_m)$ is a Fourier transformation of $4\pi\rho_0 r\gamma_0(r)$, and thus $\mathcal{N}(r)$, the inverse Fourier transformation iFT of $F(0 \le q < q_m)$ becomes,

$$\mathcal{N}(r) = 4\pi\rho_0 r\gamma_0(r) \quad (34)$$

Putting equations (33) and (34) together, we get the expression of the experimental accessible $G(r)$ as,

$$G(r) = \frac{1}{rN_p}\sum_{i\neq j} H_{ij}\left(r-r_{ij}\right) - 4\pi\rho_0 r\,\gamma_0(r) \qquad (35)$$

S4. nano-PDF for assembly of single-component monodisperse spheres

When the nano-assembly is built from single-component mono-sized, $f_i(q) = f_j(q)$ and $\beta(q) = 1$, and the equation for structure factor becomes the same as those for atomic systems and can be simplified as:

$$S_0(q) = \frac{1}{N_p}\sum_{i\neq j}\frac{\sin(qr_{ij})}{qr_{ij}} + 1 \qquad (36)$$

The reduced structure factor is:

$$F_0(q) = \frac{1}{N_p}\sum_{i\neq j}\frac{\sin(qr_{ij})}{r_{ij}} \qquad (37)$$

The inverse Fourier transformation of F(q), becomes:

$$\mathcal{F}_0(r) = \frac{1}{rN_p}\sum_{i\neq j}\delta\left(r-r_{ij}\right) \quad (38)$$

In atomic PDF literature, rF(r) is defined as the radial distribution function (RDF),

$$R(r) \equiv r\mathcal{F}(r) = \frac{1}{N_p}\sum_{i\neq j}\delta\left(r-r_{ij}\right) \quad (39)$$

RDF is related to the pair density, ρ(r), by:

$R(r) = 4\pi r^2 \rho(r)$ (40)

The ρ(r) is defined such that the number of total pairs equal to $\int 4\pi r^2 \rho(r) dr$. Accordingly, the equation (34) for G(r) becomes:

$G_0(r) = \frac{1}{rN_p} \sum_{i \neq j} \delta(r - r_{ij}) - 4\pi \rho_0 r \gamma_0(r)$ (41)

S5. nano-PDF for superlattices built from single-component monodisperse spheres with positional fluctuations

Here we consider a case where the single-component mono-sized spheres are arranged into a superlattice with positional fluctuations, namely, the particles displace randomly around the sites in the lattice. Such fluctuations are known to introduce a Debye-Waller (DW) factor, which attenuates the scattering intensity by a Gaussian-shape function (Feldman & Horton, 1963, Forster *et al.*, 2005), $D(q) = e^{-\sigma_d^2 L_a^2 q^2}$ (42)

Where σ_d, the DW factor, is the root mean square displacements of the particles divided by L_a, the nearest neighbor distance in the structure.

Due to the DW factor, the I(q) changes to:

$I_d(q) = N_p < f^2(q) > + \sum_{i \neq j} f_j^*(q) f_i(q) \frac{\sin(qr_{ij})}{qr_{ij}} D(q)$ (43)

Accordingly, the S(q) becomes,

$$S_d(q) = \frac{1}{\beta(q)} + \mathrm{D(q)} \frac{\sum_{i \neq j} f_j^*(q) f_i(q) \frac{\sin(qr_{ij})}{qr_{ij}}}{N_p < f(q) >^2} + \left[1 - \frac{1}{\beta(q)}\right]$$

$$= D(q) \frac{\sum_{i \neq j} f_j^*(q) f_i(q) \frac{\sin(qr_{ij})}{qr_{ij}}}{N_p < f(q) >^2} + 1 \quad (44)$$

Considering f_i(q) = f_j(q) due to monodisperse particles, we get,

$S_d(q) = 1 + D(q) \frac{1}{N_p} \sum_{i \neq j} \frac{\sin(qr_{ij})}{qr_{ij}}$

$= S_0(q) D(q) + [1 - D(q)]$ (45)

The reduced structure factor becomes,

$F_d(q) = q[S_d(q) - 1]$

$= F_0(q) D(q)$ (46)

D(q) plays a similar role as $h_{ij}(q)$ in the equation (27), and accordingly, the $H_{ij}(r)$ defined in the equation (28) changes to D(r):

$$D(r) \equiv \phi[D(q)] = \frac{1}{2\sqrt{\pi}\sigma_d L_a} e^{-\frac{x^2}{2(\sqrt{2}\sigma_d L_a)^2}} \quad (47)$$

D(r) is a Gaussian function with the standard deviation as $\sigma = \sqrt{2}\sigma_d L_a$. Combining equations (34) and (47), we get the expression of G(r) for the assembly of single-component monodisperse spheres with positional fluctuations as:

$$G_d(r) = \frac{1}{rN_p}\sum_{i\neq j} D(r - r_{ij}) - 4\pi\rho_0 r\,\gamma_0(r) \quad (48)$$

S6. nano-PDF for superlattice built from single-component polydisperse spheres with positional fluctuations

Practically, the synthetic particles are always polydisperse in the size. In this case, even for a single type of particle, the $f_i(q) \neq f_j(q)$, namely, $\beta(q) \neq 1$, due to size variations. Nevertheless, due to the assembly involving only one type of particle, the f(q) can be taken out of the sum of r_{ij} in the equation (18). We will demonstrate the validity of this operation by numerical simulations in section 3.5. Because of $\sum_{i\neq j} f_j^*(q) f_i(q) = \sum_j f_j^*(q) * \sum_{i\neq j} f_i(q) = <f(q)>^2$, the equation (44) changes to:

$$S_p(q) = D(q)\frac{<f(q)>^2\sum_{i\neq j}\frac{\sin(qr_{ij})}{qr_{ij}}}{N_p<f(q)>^2} + 1$$

$$= \frac{1}{N_p} D(q)\beta(q)\sum_{i\neq j}\frac{\sin(qr_{ij})}{qr_{ij}} + 1$$

$$= S_0(q)\beta'(q) + (1 - \beta'(q)) \quad (49)$$

where $\beta'(q) = D(q)\beta(q)$. The equation (49) reconciles the definitions of atomic structure factor by Faber-Ziman formalism (equation 17) and structural factor reported by nano-assembly literature (equation 19). Accordingly, the reduced structure factor is:

$$F_p(q) = q[S_p(q) - 1]$$

$$= F_0(q)\beta'(q) \quad (50)$$

So, in this case, $h_{ij}(q)$ and $H_{ij}(r)$ defined in the equations (27, 28) evolves into $\beta'(q)$ and B(r), respectively,

$$B(r) \equiv \phi[\beta'(q)] = \int_0^\infty \beta(q)D(q)\sin(qr)\,dq \quad (51)$$

Combining equations (34) and (51), we get the expression of G(r) for the single-component assemblies of polydisperse spheres with positional fluctuations as:

$$G_p(r) = \frac{1}{rN_p}\sum_{i\neq j} B\left(r - r_{ij}\right) - 4\pi\rho_0 r\,\gamma_0(r) \quad (52)$$

It has been demonstrated that the profile of β(q) can be estimated by a Gaussian function (Forster *et al.*, 2005, Yager *et al.*, 2014),

$$\beta(q) \sim e^{-\sigma_p^2 {R_p}^2 q^2} \quad (53)$$

where R_p is the mean radius of the particles. So, B(r) can be approximated by,

$$B(r) = \frac{1}{2\sqrt{\pi}\sigma_{dp}} e^{-\frac{x^2}{2(\sqrt{2}\sigma_{dp})^2}} \quad (54)$$

where $\sigma_{dp} = \sqrt{{\sigma_d}^2 {L_a}^2 + \sigma_p^2 {R_p}^2}$, and the B(r) is a Gaussian function with the standard deviation as $\sqrt{2}\sigma_{dp}$. Like the effects only caused by DW factors, the peak width of B(r) expands only a fraction of the pair distance.

S7. Estimation of q-low limit (q_m) in SAXS

Due to the X-ray beam size and divergency and other instrumental factors, there is always a minimum value of wave vector (q_m), below which SAXS patterns can't be assessed. The upper limit of the q_m can be estimated by considering the detector pixel size and scattering geometry. Assuming the following parameters in a SAXS experiment, the sample to detector as 5 meters and pixel size as 172 um (a typical Dectris 2D detector pixel size), and the smallest scattering angle to resolve the two adjacent pixels is 2θ = arctan(0.172/5000), thus the q_m = 4pi sin(θ)/ λ ~ 2pi*(0.172/5000) ~ 2*10-4 Å-1. The q_m would be larger by considering X-ray beam size and divergency, parasitic scattering, beam stop size, and X-ray polychromaticity. Typically, a modern SAXS beamline can give a q_m around 0.001.

S8. A numerical method to compute shape factor

Here, we demonstrate a four-step strategy for the numerical computation of the shape factor ($\gamma_0(r)$) of objects with arbitrary shape and size. Fig. 1S (a-c) illustrates this strategy for the calculation of $\gamma_0(r)$ for a sphere with a diameter of 1340 nm. First, we "decompose" the object of

interest into three-dimensional ordered arrays of nanospheres, namely, nanoparticle superlattices. Here, the interested sphere is "decomposed" into a superlattice comprised of 60 nm nanospheres. The superlattice itself is in a spherical shape with a diameter of 1340 nm. The nanospheres are arranged in a simple cubic (SC) lattice in the superlattice with a lattice constant of 83.8 nm. Using nanospheres of the same size and the same lattice parameters, we build another superlattice in a cube shape with a much larger size, e.g., the edge size is 10,000 nm. The small spherical and large cube-shaped superlattices are shown in the insets of Fig. 1S (a) and (b), respectively. Second, we calculate the $\mathcal{F}(r)$ of the two superlattices by applying the equation (38) in the main text. The results are plotted by black symbols in Fig. 1S (a) and (b), where $\mathcal{F}_S(r)$ and $\mathcal{F}_L(r)$ correspond to small and large superlattices, respectively. Next, the two $\mathcal{F}(r)$ are 'smeared' out by broadening the peak width to simulate the 'solid' object. Because our purpose is to calculate the shape factor of the 'solid' object, we 'smear' out all the structure details, such as lattice structure, lattice constant, etc., by broadening the peak width of $\mathcal{F}(r)$ (sort of increase particle displacements). The smeared the $F_S^s(r)$ and $F_L^s(r)$ are accordingly given as blue lines in Fig. 1S (a) and (b). Finally, the $\gamma_0(r)$ of the interested object is obtained by the division of the two smeared $\mathcal{F}(r)$, namely, $\gamma_0(r) = F_S^s(r)/F_L^s(r)$. We also calculate the γ_0(r) of the sphere using the well-known formula for the shape factor of spheres (Guinier, 1994). The resulted γ_0(r) by numerical simulations and by the analytical formula are shown as blue and red curves in Fig. 1S (c), respectively. The two curves agree well with each other especially for r larger than the lattice constant (83.8 nm). The slight difference in the small-r region is caused by the reason that there are no particle pairs with a distance smaller than the lattice constant in the "decomposition" simulation. The smallest particle distance in the simulation defines the resolution of the simulated γ_0(r). One can increase the resolution of γ_0(r) by using a smaller lattice constant if necessary. Furthermore, by this four-step strategy, we calculate the $\gamma_0(r)$ for different shapes, including a plate, rod, cube, and tube. All the shapes have a similar dimension as the sphere. The plate is a two-dimensional square with a lateral size of 1340 nm. The rod has a length and a diameter of 1340 nm. The cube has an edge length of 1340 nm. The tube has a length, an outer diameter, and an inner diameter of 1340, 1340, and 670 nm, respectively. The resulted γ_0(r) by numerical simulations for the four objects are shown as blue curves in Fig. 1S (d-g), respectively. Based on available analytic expressions for plate, rod, and cube (see literature for example (Usher *et al.*, 2018)), we also calculate γ_0(r) and plot the

results as red curves in Fig. 1S (d-f), which shows excellent agreements with the $\gamma_0(r)$ obtained by our numerical simulation methods.

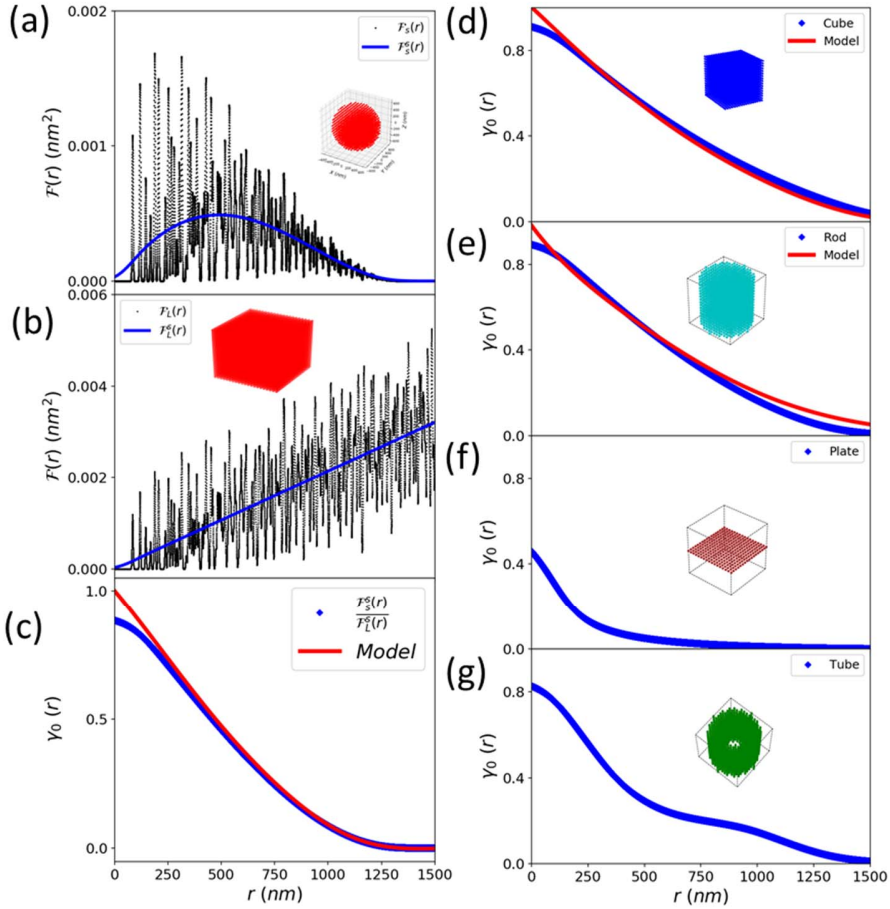

Fig 1S, Numerical computation of the shape factor, $\gamma_0(r)$, of arbitrary shapes. (a-c) illustrate the four-step procedures for the calculation of $\gamma_0(r)$ for a sphere. First, "decompose" the object of interest into three-dimensional ordered arrays of nanospheres, namely, nanoparticle superlattices. The inset in (a) shows the interested sphere is "decomposed" into a superlattice with a simple cubic (SC) lattice with a lattice constant of 83.8 nm. The superlattice, consisting of 60 nm nanospheres, itself is in a spherical shape with a diameter of 1340 nm. The inset in (b) shows another superlattice in a cube shape with an edge size of 10,000 nm. This large cube-shaped superlattice has nanospheres of the same size and the same lattice parameters as the small sphere-shaped superlattice. Second, $\mathcal{F}(r)$ of the two superlattices are calculated based on the equation (38) and shown as black symbols in (a) $\mathcal{F}_L(r)$ and $\mathcal{F}_s(r)$. Third, the two $\mathcal{F}(r)$ are 'smeared' out by broadening the peak width and plotted as the blue line in (a) $F_s^s(r)$ and $F_L^s(r)$. Finally, the shape factor is calculated by $\gamma_0(r) = F_s^s(r)/F_L^s(r)$ and shown by the blue line in (c). The shape factor of a sphere is also calculated by the analytical formula, $\gamma_0(r) = 1 - 1.5(r/D) + 0.5(r/D)^3$ (Guinier, 1994), where D is the diameter of the sphere. Using this method, shape factors of other shapes are calculated. The insets show nanoparticle superlattices "decomposed" from (d) a cube, (e) a rod, (f) a plate, and (g) a tube. The blue lines in (d-g) give the calculated $\gamma_0(r)$. The red lines in (d-e) shows the $\gamma_0(r)$ calculated by the analytical expressions (Usher *et al.*, 2018).

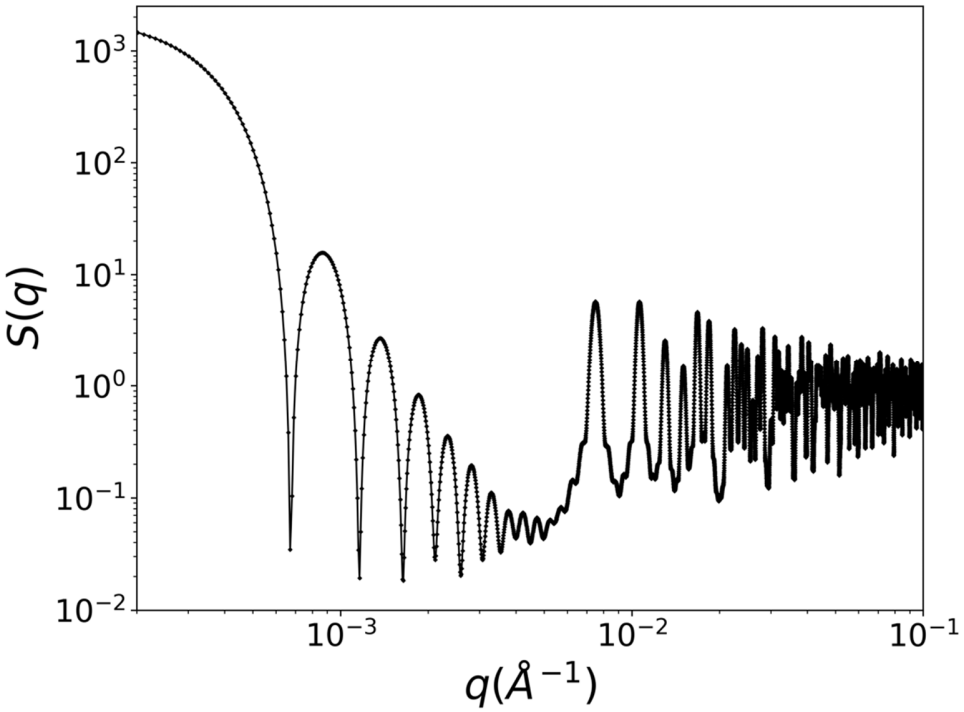

Fig 2S, Replot of Fig.1b, S(q) of the nanoparticle assembly (S_0), with axis in double logarithmic scale.

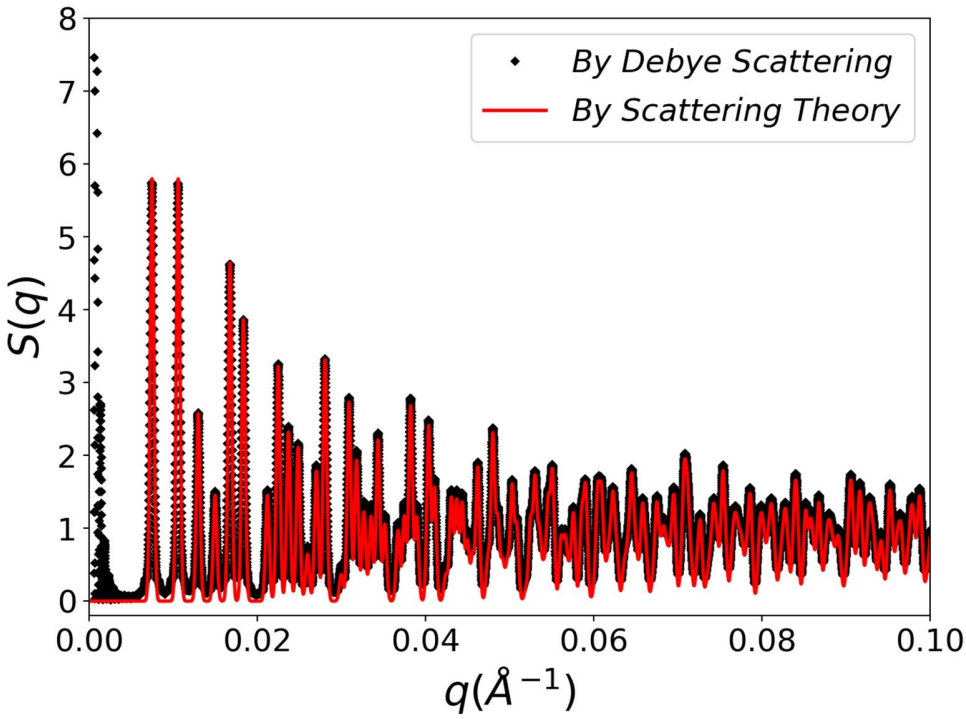

Fig 3S, A comparison of structure factors. Black symbols are calculated by Debye scattering method developed in this work (equation (18)). Redline is obtained by the crystallography scattering theory (Yager *et al.*, 2014). Both methods give similar scattering in the Bragg peak regions, while Debye scattering can capture the features in the small-q region, which encodes the size and shape information of the assembly.

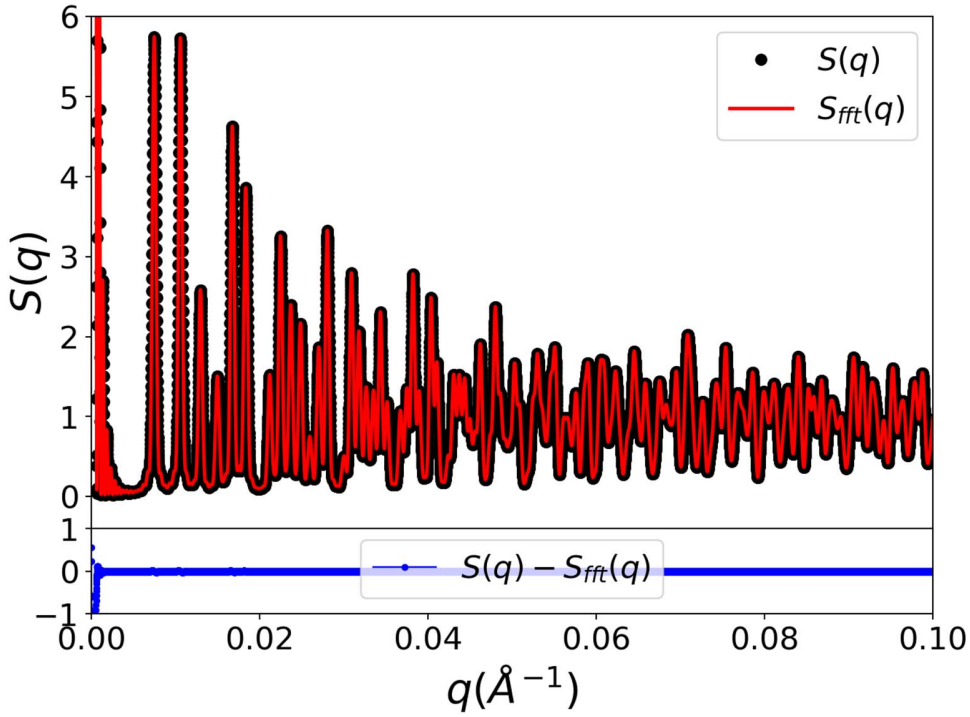

Fig 4S, Verification of (inverse) Fourier transformation algorithm used in this work. A comparison of the initial S(q) (black symbols) and the converted S(q) (red line) by the iFT-FT process excellently agrees with each other.

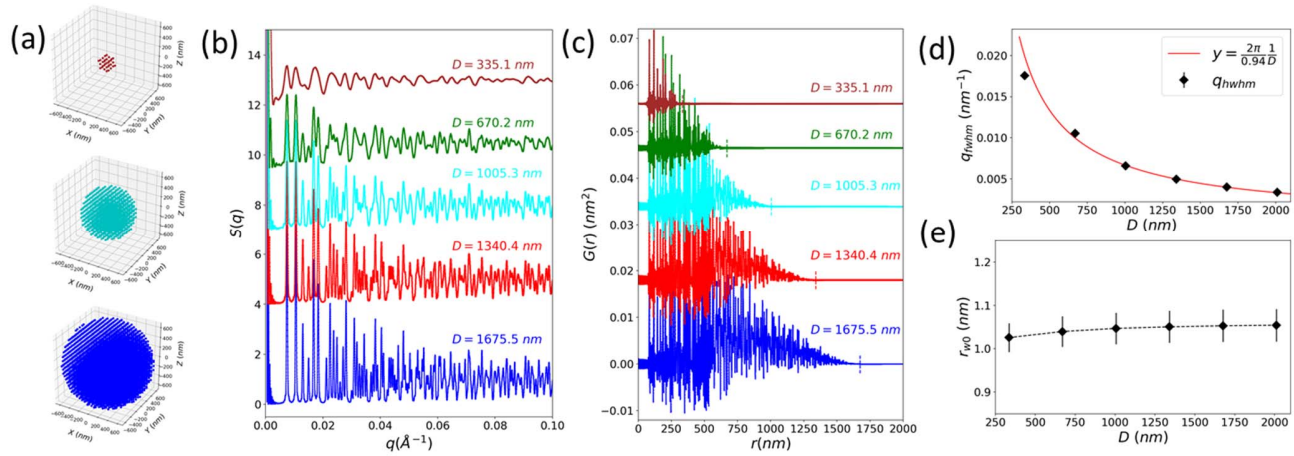

Fig 5S, S(q) and G(r) as a function of the size of nanoparticle assemblies. (a) (From top to bottom) The nanoparticle assembly with a diameter (D) from 335 to 1005 and 1675 nm. The calculated S(q) (b) and the corresponding G(r) (c) for assemblies with different D. The dash lines show that the intensity fluctuation of G(r) is on a background-level at R_M, which roughly equal to D. (d) (black symbols) the full width at half maximum of the first peak(q_{fwhm}) of S(q) as a function of D. (red line) the q_{fwhm} is fitted by $2\pi/(0.94D)$. (e) The reduced peak width (r_{w0}) of the first peak in G(r) as a function of D. The first peaks of S(q) and G(r) are both fitted by Gaussian functions. For S(q), $q_{fwhm}=2.35\sigma_q$, and for G(r), $r_{w0} = \sqrt{2}\sigma_{w0}$. σ_q and σ_{w0} are the standard

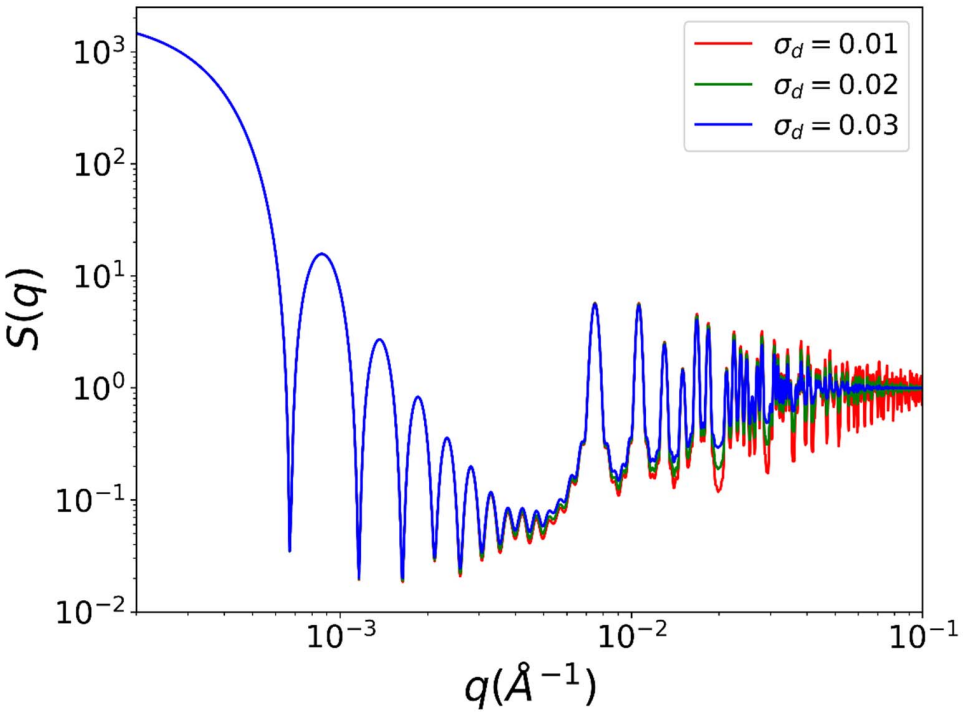

Fig 6S, Replot of Fig.3a, S(q) of the nanoparticle assembly as a function of positional fluctuation parameter, σ_d , with axis in double logarithmic scale.

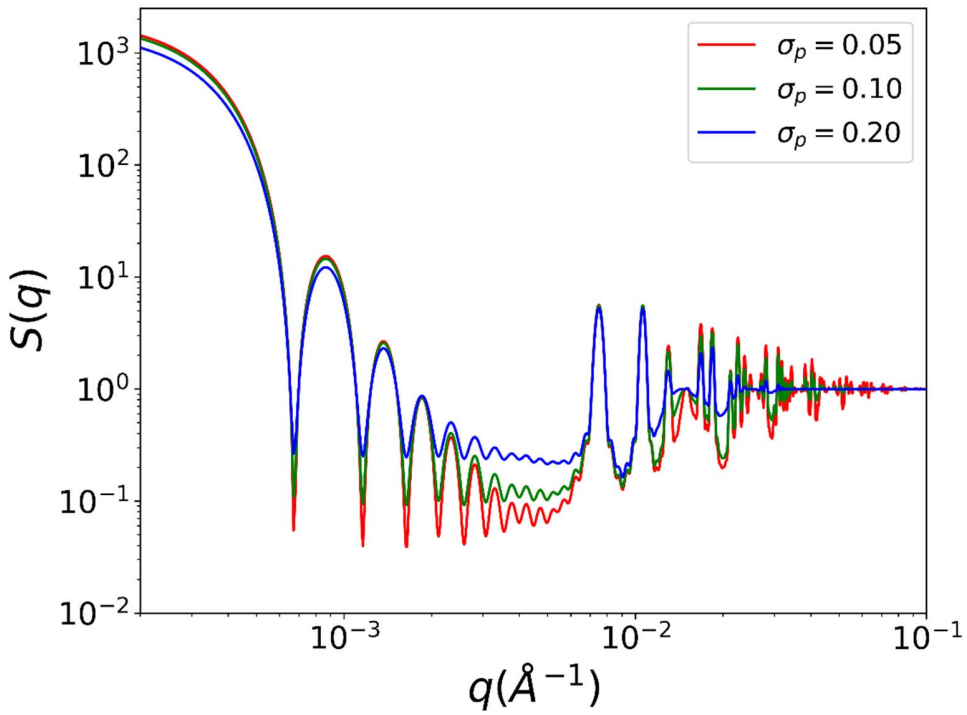

Fig 7S, Replot of Fig.4c, S(q) of the nanoparticle assembly as a function of particle polydispersity parameter, σ_p , with axis in double logarithmic scale.

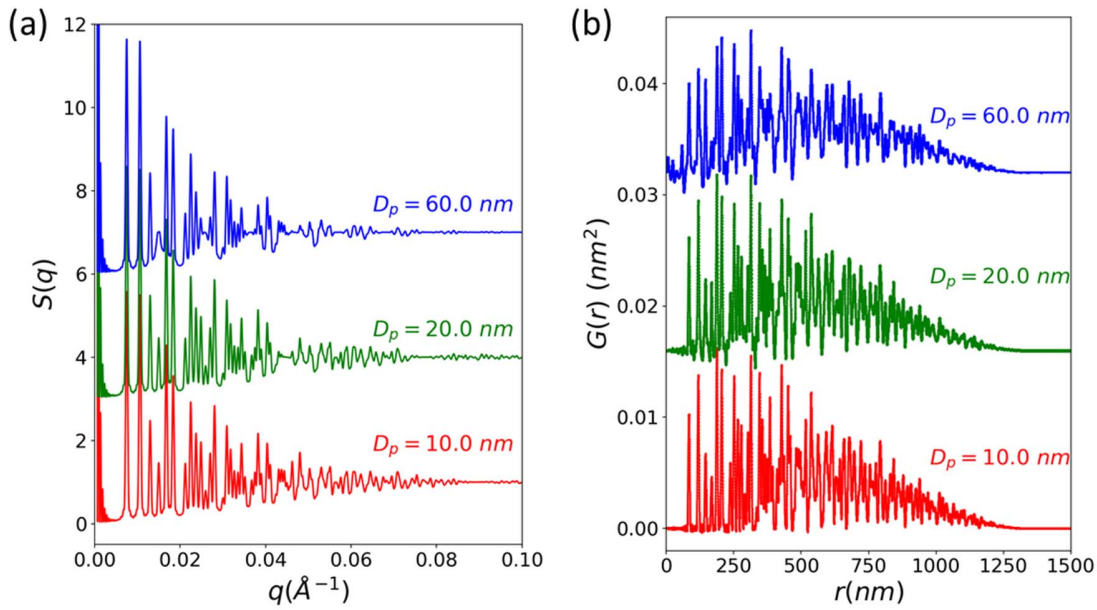

Fig 8S, (a) S(q) and (b) G(r) of nanoparticle assemblies comprised of polydisperse spheres and positional fluctuations. The size polydispersity satisfies a Gaussian distribution with mean diameter D_p and standard deviation $\sigma_p = 0.05$. The positional fluctuations give the DW factor $\sigma_d = 0.02$.

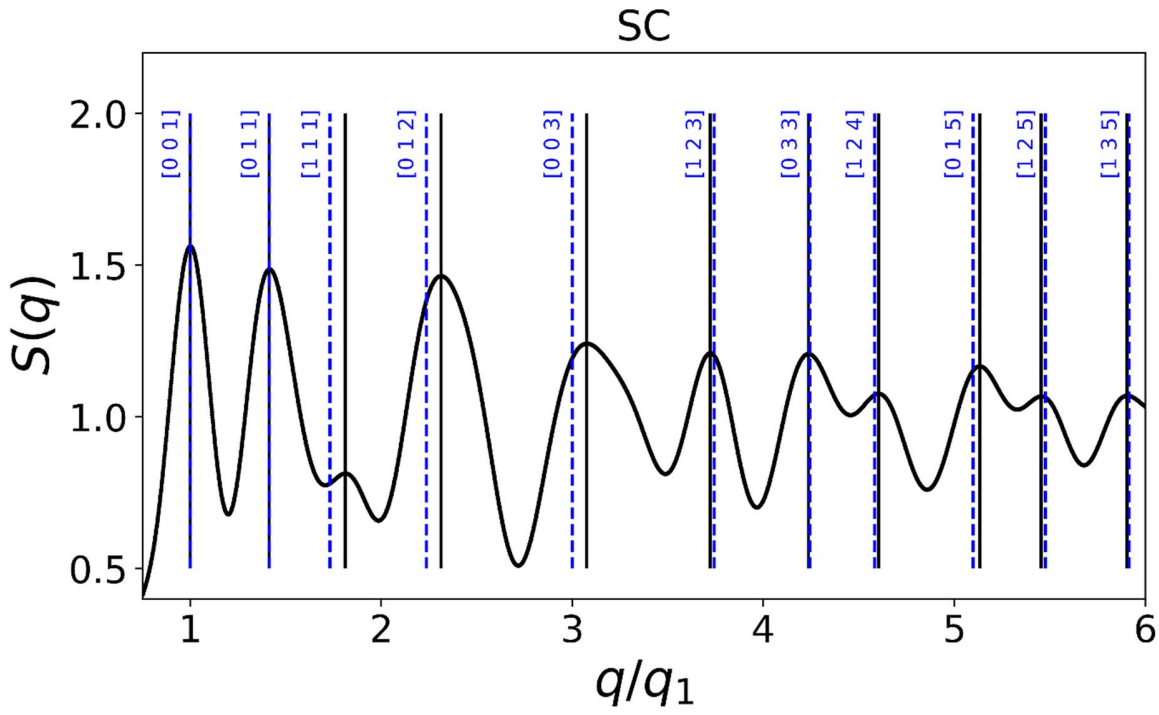

Fig 9S, S(q) (blue curve in Fig. 6b) of nano-assembly with a size of 330 nm in a simple cubic (SC) lattice. The x-axis is q normalized by the first q position, q_1. Vertical black solid lines represent the peak positions of this nano-assembly, while blue dash lines along with miller indexes indicate closed diffraction peaks in a perfect lattice, e.g, with infinite crystal size.

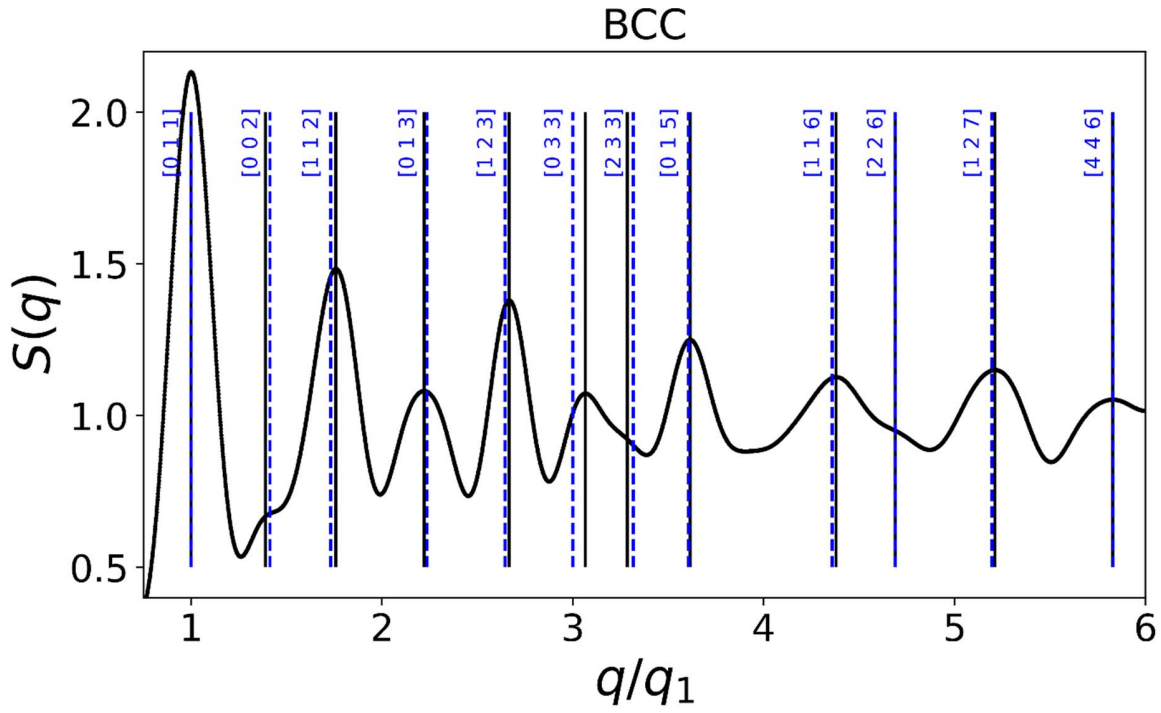

Fig 10S, S(q) (red curve in Fig. 6b) of nano-assembly with a size of 330 nm in a body centered cubic (BCC) lattice. The x-axis is q normalized by the first q position, q_1. Vertical black solid lines represent the peak positions of this nano-assembly, while blue dash lines along with miller indexes indicate closed diffraction peaks in a perfect lattice, e.g, with infinite crystal size.

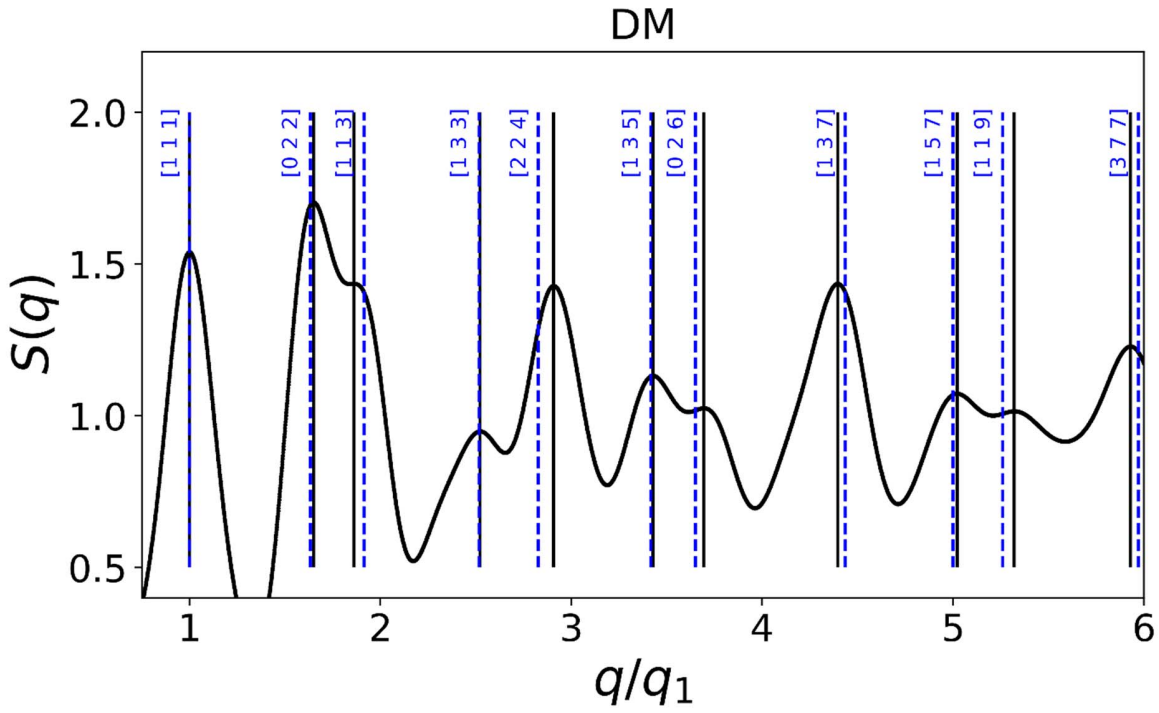

Fig 11S, S(q) (cyan curve in Fig. 6b) of nano-assembly with a size of 330 nm in a diamond (DM) lattice. The x-axis is q normalized by the first q position, q_1. Vertical black solid lines represent the peak positions of this nano-assembly, while blue dash lines along with miller indexes indicate closed diffraction peaks in a perfect lattice, e.g, with infinite crystal size.

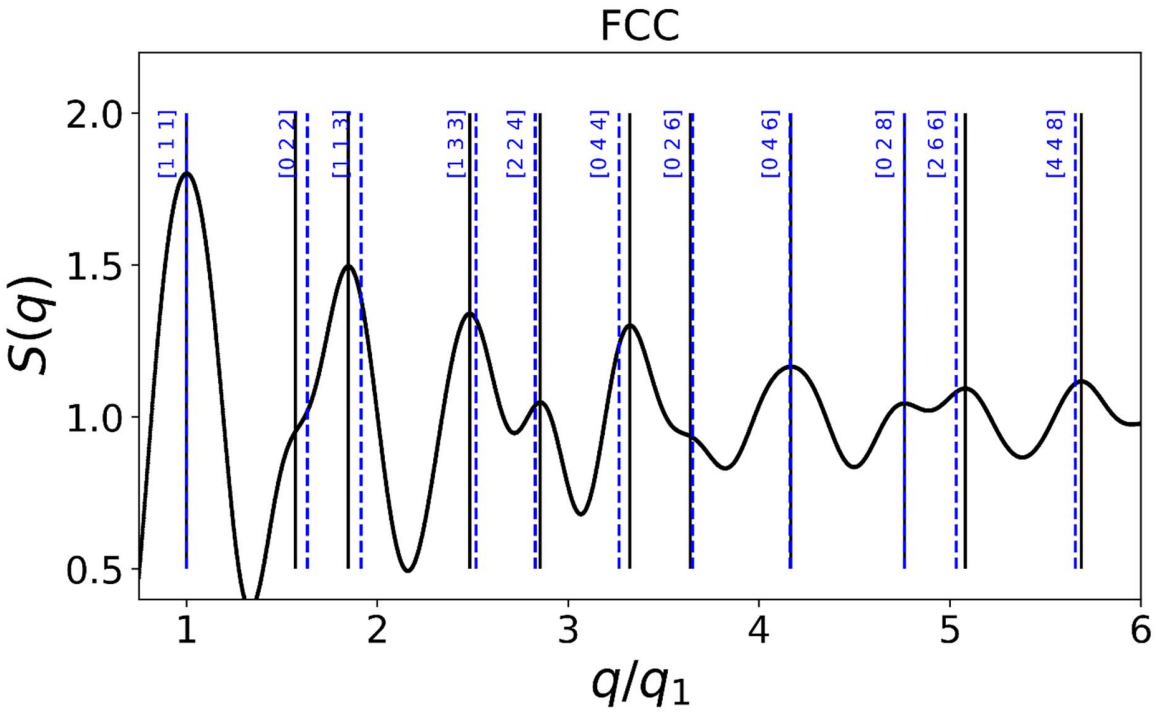

Fig 12S, S(q) (green curve in Fig. 6b) of nano-assembly with a size of 330 nm in a face centered cubic (FCC) lattice. The x-axis is q normalized by the first q position, q_1. Vertical black solid lines represent the peak positions of this nano-assembly, while blue dash lines along with miller indexes indicate closed diffraction peaks in a perfect lattice, e.g, with infinite crystal size.

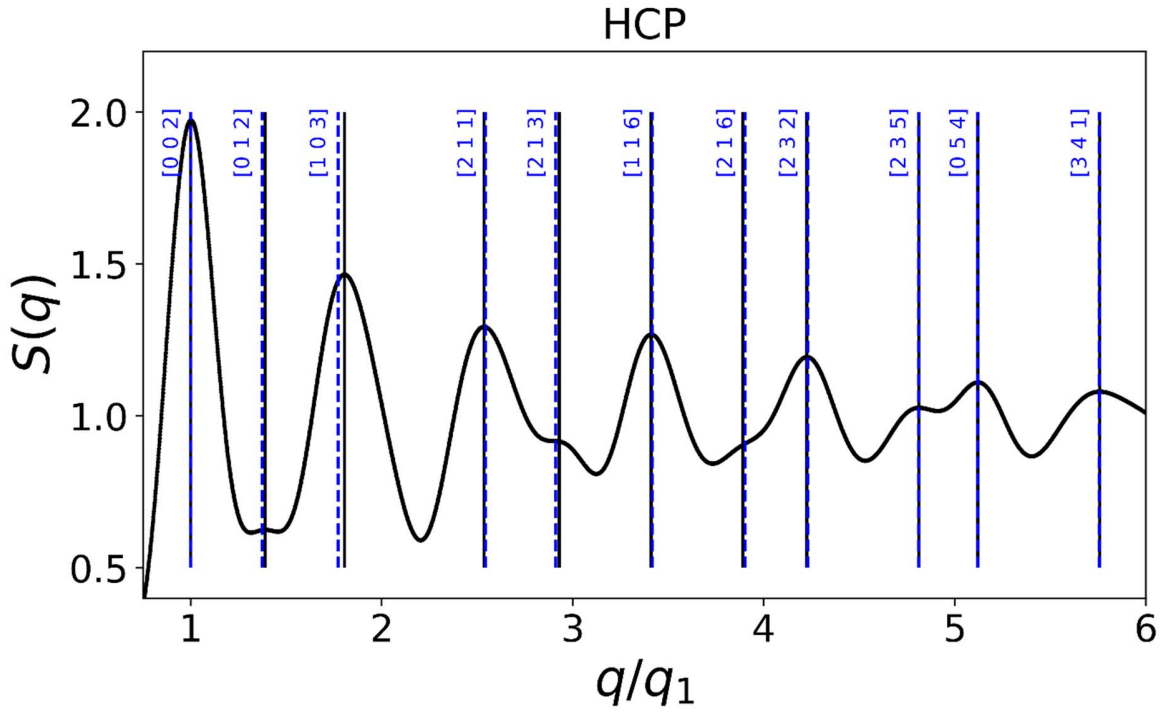

Fig 13S, S(q) (brown curve in Fig. 6b) of nano-assembly with a size of 330 nm in a hexagonal closed packed (HCP) lattice. The x-axis is q normalized by the first q position, q_1. Vertical black solid lines represent the peak positions of this nano-assembly, while blue dash lines along with miller indexes indicate closed diffraction peaks in a perfect lattice, e.g, with infinite crystal size.

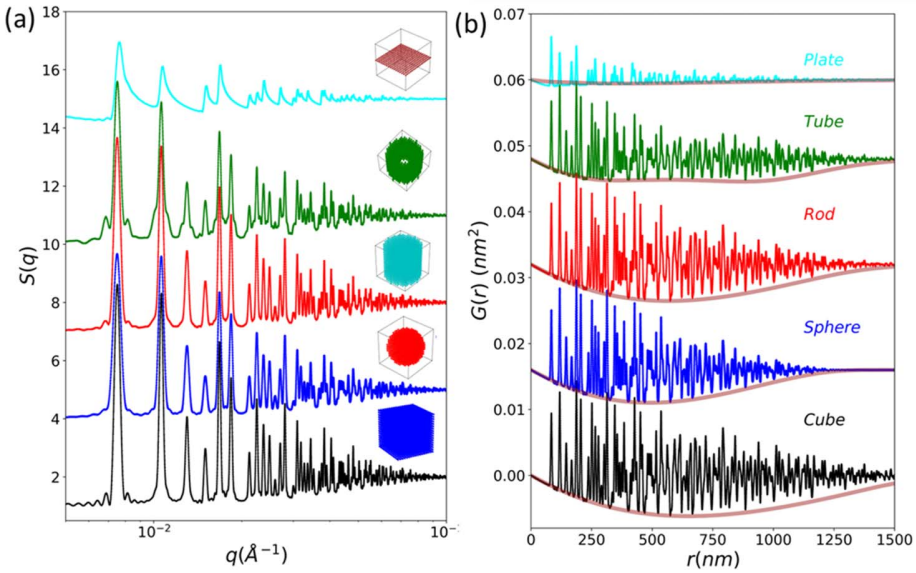

Fig 14S, (a) S(q) calculated by Debye equation (36, 45) with axis on double logarithmic scales and (b) the corresponding G(r) converted by the inverse Fourier transformation for nanoparticle assemblies with different shapes, including a plate, a tube, a rod, a sphere, and a cube (depicted in the insets of (a) from top to bottom). The envelops of each G(r) profile are given in brown

References

Billinge, S. J. L. (2019). *Philos T R Soc A* **377**.
Bodnarchuk, M. I., Shevchenko, E. V. & Talapin, D. V. (2011). *J Am Chem Soc* **133**, 20837-20849.
Boles, M. A., Engel, M. & Talapin, D. V. (2016). *Chem Rev* **116**, 11220-11289.
Dullens, R. P. A. & Petukhov, A. V. (2007). *Epl-Europhys Lett* **77**.
Egami, T. & Billinge, S. J. L. (2012). *Underneath the Bragg peaks : structural analysis of complex materials*, Second edition. ed. Amsterdam: Elsevier.
Faber, T. E. Z., J. M. (1965). *Philosophical Magazine* **11**, 153-173.
Farrow, C. L. & Billinge, S. J. L. (2009). *Acta Crystallogr A* **65**, 232-239.
Feldman, J. L. & Horton, G. K. (1963). *Phys Rev* **132**, 644-&.
Forster, S., Timmann, A., Konrad, M., Schellbach, C., Meyer, A., Funari, S. S., Mulvaney, P. & Knott, R. (2005). *J Phys Chem B* **109**, 1347-1360.
Gamez-Mendoza, L., Terban, M. W., Billinge, S. J. L. & Martinez-Inesta, M. (2017). *J Appl Crystallogr* **50**, 741-748.
Gilbert, B. (2008). *J Appl Crystallogr* **41**, 554-562.
Guinier, A. (1994). *X-ray diffraction in crystals, imperfect crystals, and amorphous bodies*. New York: Dover.
Guinier, A. & Fournet, G. r. (1955). *Small-angle scattering of X-rays*. New York,: Wiley.
Hudalla, G. A., Sun, T., Gasiorowski, J. Z., Han, H. F., Tian, Y. F., Chong, A. S. & Collier, J. H. (2014). *Nat Mater* **13**, 829-836.
Liu, C. H., Janke, E. M., Li, R. P., Juhas, P., Gang, O., Talapin, D. V. & Billinge, S. J. L. (2020). *J Appl Crystallogr* **53**, 699-709.
Lombardo, D., Kiselev, M. A., Magazu, S. & Calandra, P. (2015). *Adv Cond Matter Phys* **2015**.
Nie, Z. H., Petukhova, A. & Kumacheva, E. (2010). *Nat Nanotechnol* **5**, 15-25.
Nykypanchuk, D., Maye, M. M., van der Lelie, D. & Gang, O. (2008). *Nature* **451**, 549-552.
Olds, D., Wang, H. W. & Page, K. (2015). *J Appl Crystallogr* **48**, 1651-1659.
Senesi, A. J. & Lee, B. (2015). *J Appl Crystallogr* **48**, 1172-1182.
Usher, T. M., Olds, D., Liu, J. & Page, K. (2018). *Acta Crystallogr A* **74**, 322-331.
Warren, B. E. (1990). *X-ray Diffraction*. New York: Dover.
Yager, K. G., Zhang, Y. G., Lu, F. & Gang, O. (2014). *J Appl Crystallogr* **47**, 118-129.
Young, C. A. & Goodwin, A. L. (2011). *J Mater Chem* **21**, 6464-6476.
Zernike, F., Prins, J. A. (1927). *Z. Phys.* **41**, 184-194.