Supporting Information: Resolving non-equilibrium shape variations amongst millions of gold nanoparticles

Zhou Shen,[†,‡,¶] Paul Lourdu Xavier,[‡,§,‖] Richard Bean,[⊥] Johan Bielecki,[⊥] Martin Bergemann,[⊥] Benedikt J Daurer,[#,@] Tomas Ekeberg,[△] Armando D. Estillore,[‖] Hans Fangohr,[⊥] Klaus Giewekemeyer,[⊥] Mikhail Karnevskiy,[⊥] Richard A. Kirian,[∇] Henry Kirkwood,[⊥] Yoonhee Kim,[⊥] Jayanath C. P. Koliyadu,[⊥] Holger Lange,[††,‡‡,¶¶] Romain Letrun,[⊥] Jannik Lübke,[††,‖,§§] Abhishek Mall,[‡,¶] Thomas Michelat,[⊥] Andrew J. Morgan,[‖‖] Nils Roth,[‖,§§] Amit K Samanta,[††,‖] Tokushi Sato,[⊥] Marcin Sikorski,[⊥] Florian Schulz,[¶¶] Patrik Vagovic,[‖,⊥] Tamme Wollweber,[‡,¶,††] Lena Worbs,[‖,§§] Filipe Maia,[△,⊥⊥] Daniel Alfred Horke,[††,‖,##] Jochen Küpper,[††,‖,§§,@@] Adrian P. Mancuso,[⊥,△△,@] Henry Chapman,[††,‖,§§] Kartik Ayyer,[‡,¶,††] and N. Duane Loh[*,#,†]

†*Department of Physics, National University of Singapore, Singapore 117551, Singapore*
‡*Max Planck Institute for the Structure and Dynamics of Matter, 22761 Hamburg, Germany*
¶*Center for Free-Electron Laser Science, 22761 Hamburg, Germany*
§*The Hamburg Center for Ultrafast Imaging, Universität Hamburg, 22761 Hamburg, Germany*
‖*Center for Free-Electron Laser Science, Deutsches Elektronen-Synchrotron DESY, 22607 Hamburg, Germany*
⊥*European XFEL, 22869 Schenefeld, Germany*
#*Center for BioImaging Sciences, National University of Singapore, Singapore 117557, Singapore*
@*Diamond Light Source, Harwell Campus, Didcot, OX11 0DE, UK*
△*Department of Cell and Molecular Biology, Uppsala University, 75124 Uppsala, Sweden*
∇*Department of Physics, Arizona State University, Tempe, Arizona 85287, USA*
††*The Hamburg Center for Ultrafast Imaging, Universität Hamburg, 22761 Hamburg, Germany*
‡‡*Institute of Physics and Astronomy, Universität Potsdam, Karl-Liebknecht-Str. 24, 14476 Potsdam, Germany*
¶¶*Institute of Physical Chemistry, Universität Hamburg, 20146 Hamburg, Germany*
§§*Department of Physics, Universität Hamburg, 22761 Hamburg, Germany*
‖‖*University of Melbourne, Physics, Melbourne, VIC 3010, Australia*
⊥⊥*NERSC, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*
##*Radboud University Institute for Molecules and Materials, 6525 AJ Nijmegen, The Netherlands*
@@*Department of Chemistry, Universität Hamburg, 20146 Hamburg, Germany*
△△*Department of Chemistry and Physics, La Trobe Institute for Molecular Science, La Trobe University, Melbourne, VIC 3086, Australia*

E-mail: duanenloh@nus.edu.sg

1 Match 2D EMC clusters with ideal 2D intensity slices

In the fig. S1, we manually match several typical 2D EMC clusters with 2D intensity slices of an ideal octahedra from different orientations. The shape of this octahedra is given by the average model we reconstructed.

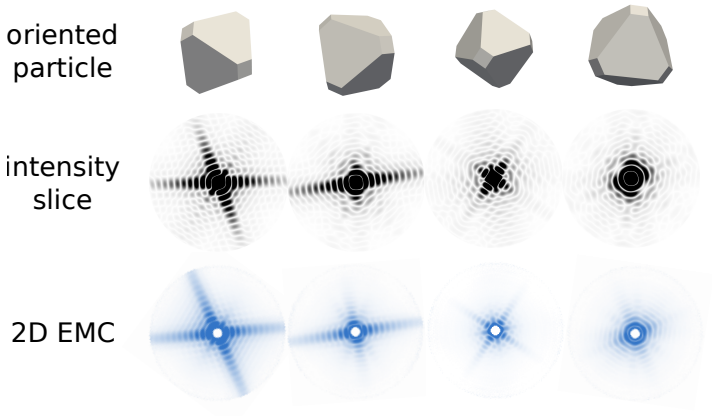

Figure S1: Here, we select four typical 2D EMC clusters (blue). For each one, a 2D intensity (black) slice is manually matched, and the corresponding oriented particle is showed in grey.

2 Finite volume Fourier Transformation

Comparing to biomolecules, nanoparticles are also a common kind of sample in XFEL SPI experiment but has a much simpler structure. Usually, we could assume its density is uniform and has a polyhedron shape. In this section, we propose a numerical Fourier transformation scheme for any uniform polyhedron S. This method could avoid the voxelization of S. As the difficulty of the Fourier transformation of S , $\iiint_S \exp(\mathrm{i}\mathbf{k}\cdot\mathbf{x})\ \mathrm{d}\mathbf{x}$ comes from the complexity of the shape of S, we convert this one big integral into several smaller integrals over tetrahedrons. The surface (boundary) of S, ∂S, can be triangulated into a list of n triangular faces

$$\left\{[v_1^{(i)}, v_2^{(i)}, v_3^{(i)}] \,\middle|\, i = 1, 2, \ldots, n\right\},$$

where $v^{(i)}$s are the vertices of triangular i, $[v_1^{(i)}, v_2^{(i)}, v_3^{(i)}]$. Suppose S is convex and the origin v_O is inside S. Then $S = \sum_i [v_1^{(i)}, v_2^{(i)}, v_3^{(i)}, v_O]$ where $[v_1^{(i)}, v_2^{(i)}, v_3^{(i)}, v_O]$ is the tetrahedron with base $[v_1^{(i)}, v_2^{(i)}, v_3^{(i)}]$ and apex v_O. Immediately, we get

$$\iiint_S \exp(\mathrm{i}\mathbf{k}\cdot\mathbf{x})\ \mathrm{d}\mathbf{x} = \iiint_{\sum_i [v_1^{(i)}, v_2^{(i)}, v_3^{(i)}, v_O]} \exp(\mathrm{i}\mathbf{k}\cdot\mathbf{x})\ \mathrm{d}\mathbf{x}. \tag{1}$$

Equation (1) retains its validity even when these two assumptions are relaxed. A straightforward argument in support of this is that both sides of eq. (1) exhibit continuity across all vertices, and the integral remains independent of the choice of origin. The critical factor here is ensuring the correct orientation of a surface, which is determined by the sequential selection of three vertices, $\{v_1^{(i)}, v_2^{(i)}, v_3^{(i)}\}$, on the surface following the right-hand rule. It is imperative that the orientation, represented by the "thumb" direction, points outward from the S. A more rigorous mathematical statement pertaining to this concept, "orientability", can be found in most algebraic topology

textbooks.

The region covered by $[v_1^{(i)}, v_2^{(i)}, v_3^{(i)}, v_O]$ and $[v_x, v_y, v_z, v_O]$ can be converted into each other by a linear transform, A,

$$A^{-\mathrm{T}}[\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_O] = \begin{bmatrix} 1 & & & 0 \\ & 1 & & 0 \\ & & 1 & 0 \end{bmatrix} = [\mathbf{v}_x, \mathbf{v}_y, \mathbf{v}_z, \mathbf{v}_O] \tag{2}$$

where

$$A = \begin{bmatrix} \mathbf{v}_1^{\mathrm{T}} \\ \mathbf{v}_2^{\mathrm{T}} \\ \mathbf{v}_3^{\mathrm{T}} \end{bmatrix}, \tag{3}$$

v_x, v_y, and v_z are three unit points of x, y, and z-axis, and the $\mathbf{v}$ emphasizes that it is a column vector of vertex v. The integrals over $[v_1^{(i)}, v_2^{(i)}, v_3^{(i)}, v_O]$ in eq. (1) can be converted to integrals over the same trirectangular triangular pyramid $[v_x, v_y, v_z, v_O]$.

$$\begin{aligned} \iiint_S \exp(\mathrm{i}\mathbf{k}\cdot\mathbf{x})\,\mathrm{d}\mathbf{x} &= \sum_i \iiint_{[v_1^{(i)}, v_2^{(i)}, v_3^{(i)}, v_O]} \exp(\mathrm{i}\mathbf{k}\cdot\mathbf{x})\,\mathrm{d}\mathbf{x} \\ &= \sum_i \det A^{(i)} \iiint_{[v_x, v_y, v_z, v_O]} \exp(\mathrm{i}A^{(i)}\mathbf{k}\cdot\mathbf{x})\,\mathrm{d}\mathbf{x} \\ &= \sum_i \det A^{(i)} \cdot F^\star(A^{(i)}\mathbf{k}) \end{aligned}$$

where $F^\star$ is the Fourier transformation of $[v_x, v_y, v_z, v_O]$. The closed-form expression for $F^\star$ can be found by

$$\begin{aligned} F^\star(\mathbf{k}) &= \int_0^1 \int_0^{1-x} \int_0^{1-x-y} \exp\left[\mathrm{i}(k_x x + k_y y + k_z z)\right]\,\mathrm{d}z\,\mathrm{d}y\,\mathrm{d}x \\ &= \frac{\mathrm{i}\sum_{xyz}\left[\exp(\mathrm{i}k_x) - 1\right] k_y k_z (k_z - k_y)}{\prod_{xyz} k_x \prod_{xyz}(k_x - k_y)}, \end{aligned} \tag{4}$$

where $\sum_{xyz}$ and $\prod_{xyz}$ are the index-rolling summation and production, for instance,

$$\sum_{xyz} k_x = k_x + k_y + k_z,\ \prod_{xyz}(k_y - k_z) = (k_y - k_z)(k_z - k_x)(k_x - k_y).$$

In the numerical calculation of eq. (4), zero denominator in eq. (4) causes zero division error. Those cases are supposed to be handled separately by calculating the limits of eq. (4). It should be pointed out that as the right-triangular pyramid has $\mathrm{C_{3v}}$ symmetry, all permutations $(\sigma(x)\ \sigma(y)\ \sigma(z))$, like $(y\ x\ z)$, to the index list $(x\ y\ z)$ in $F^\star(k_x, k_y, k_z)$ give the same values, which also could be verified directly from eq. (4). Therefore, all zero-division cases are divided into six classes, and the condition label for each class is a representative of its index-permutated class.

1. $k_x = k_y = k_z = 0$

$$F^\star(0, 0, 0) = \frac{1}{6};$$

2. $k_x = k_y = k_z \neq 0$

$$F^\star(k_x, k_x, k_x) = \frac{-2\mathrm{i} + \exp(\mathrm{i}k_x)(2\mathrm{i} + 2k_x - \mathrm{i}k_x^2)}{2k_x^3};$$

3. $k_x \neq 0,\ k_y = k_z = 0$

$$F^\star(k_x, 0, 0) = \frac{1}{k_x^2} + \frac{\mathrm{i}\left[-2 + 2\exp(\mathrm{i}k_x) + k_x^2\right]}{2k_x^3};$$

4. $k_z = k_x \neq 0,\ k_y = 0$

$$F^\star(k_x, 0, k_x) = -\frac{-2\mathrm{i} + k_x + \exp(\mathrm{i}k_x)(2\mathrm{i} + k_x)}{k_x^3};$$

5. $k_x = 0,\ k_y \neq 0,\ k_z \neq 0,\ k_y \neq k_z$

$$F^\star(0, k_y, k_z) = \frac{\mathrm{i}\left[\exp(\mathrm{i}k_y) - 1\right]}{k_y^2(k_y - k_z)} + \frac{\mathrm{i}\left[\exp(\mathrm{i}k_z) - 1\right]}{k_z^2(k_z - k_y)} - \frac{1}{k_y k_z};$$

6. $k_x = k_y \neq 0,\ k_z \neq 0$

$$\begin{aligned} F^\star(k_x, k_x, k_z) = \frac{1}{k_y^2(k_y - k_z)^2 k_z}\Big\{ & \mathrm{i}\left[\exp(\mathrm{i}k_y) - 1\right]k_z^2 + \\ & \mathrm{i}k_y^2\left[\exp(\mathrm{i}k_z) - 1 + \mathrm{i}\exp(\mathrm{i}k_y)k_z\right] + \\ & k_y k_z\left[2\mathrm{i} + \exp(\mathrm{i}k_y)(-2\mathrm{i} + k_z)\right]\Big\}. \end{aligned}$$

3 Characterization of Synthesis Results: UV-VIS and TEM Analysis of Nanostructures

UV-VIS and TEM analyses were performed to characterize the synthesis results of the samples. UV-VIS spectroscopy (fig. S2(a)) was utilized to assess the optical properties, while transmission electron microscopy (TEM) images (fig. S2(b,c)) provided detailed morphological information. Specifically, TEM imaging of the smaples depicted predominantly octahedral shapes ($> 95\%$), alongside minor by-products such as triangular prisms, tetrahedra, decahedra, and dodecahedra. These observations offer insights into the morphology and crystallinity of the synthesized samples. It is important to mention that these shapes, rather than the expected octahedral forms, were initially observed from diffraction patterns. However, they were subsequently eliminated during the dataset purification process using 2D-EMC(see section "*In silico* filtration with 2D EMC" of the paper).

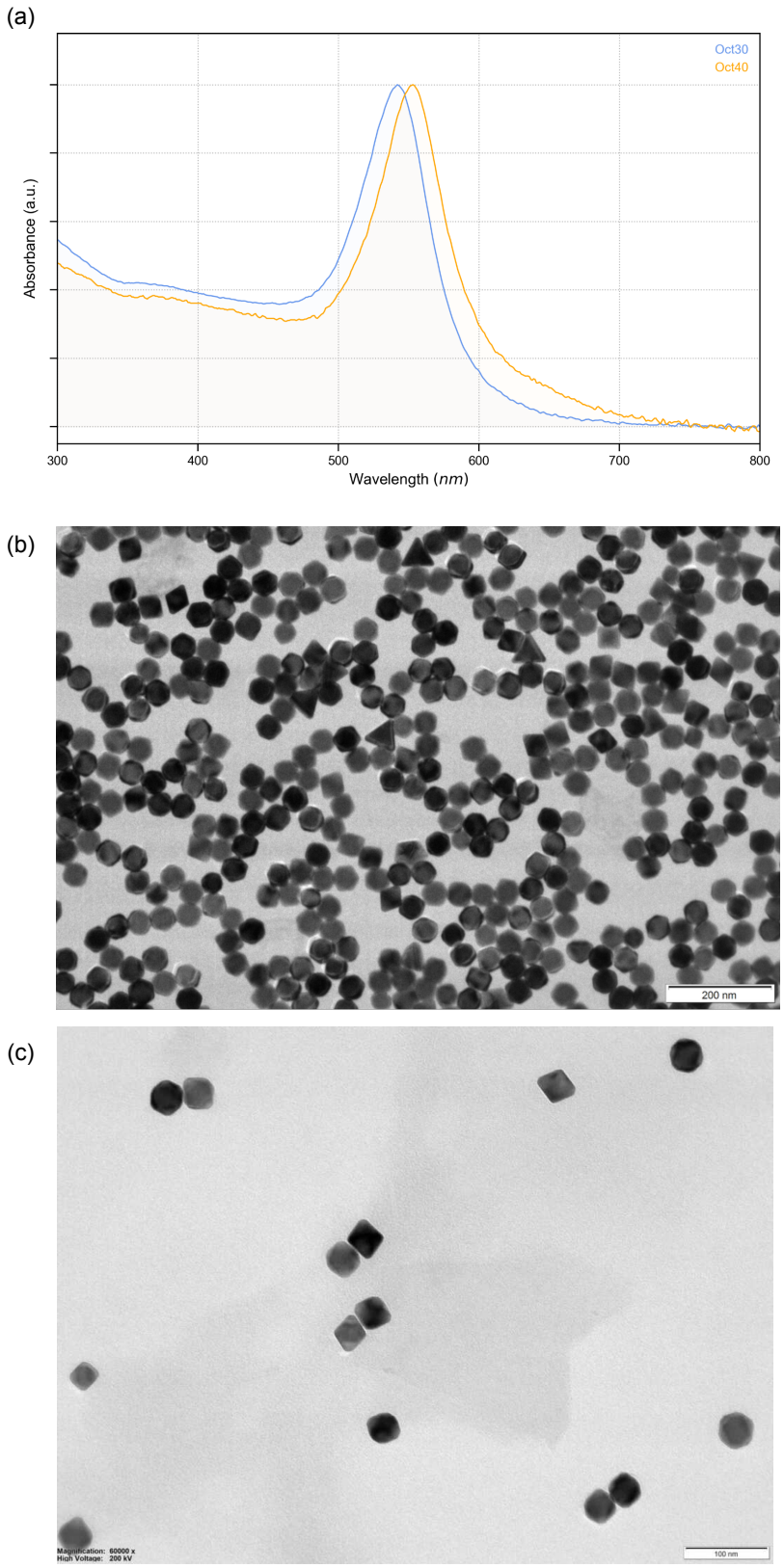

Figure S2: (a) UV-VIS spectroscopy results. (b,c) Transmission electron microscopy images revealing the nanostructure morphology of the samples: oct40 in panel (b), oct30 in panel (c).