UC Berkeley
UC Berkeley Previously Published Works

Title

Evaluating Cryo-TEM Reconstruction Accuracy of Self-Assembled Polymer Nanostructures

Permalink

https://escholarship.org/uc/item/0kf312br

Authors

Luo, Xubo
Seidler, Morgan
Lee, Yen Jea
et al.

Publication Date

2024-09-12

DOI

10.1002/marc.202400589

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Evaluating Cryo-TEM Reconstruction Accuracy of Self-assembled Polymer Nanostructures

Xubo Luo[1,2], Morgan Seidler[1,3], Yen Jea Lee[1,2], Tianyi Yu [1,2], Ronald N. Zuckermann[2], Nitash P. Balsara[1,3], Brooks Abel[1,3], David Prendergast[2], and Xi Jiang[1*]

1. Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

2. Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

3. Department of Chemical and Biomolecular Engineering, University of California, Berkeley, CA 94720, USA

*Corresponding author: xijiang@lbl.gov;

Abstract:

Cryogenic transmission electron microscopy (cryo-TEM) combined with single particle analysis (SPA) is an emerging imaging approach for soft materials. However, the accuracy of SPA-reconstructed nanostructures, particularly those formed by synthetic polymers, remains uncertain due to potential packing heterogeneity of the nanostructures. In this study, we utilized the combination of molecular dynamics (MD) simulations and image simulations to validate the accuracy of cryo-TEM 3D reconstructions of self-assembled polypeptoid fibril nanostructures. Using CryoSPARC software, we performed image simulations, 2D classifications, *ab initio* reconstructions, and homogenous refinements. By comparing the results with atomic models, we were able to assess the recovery of molecular details, identify heterogeneous structures, and evaluate the influence of extraction location on the reconstructions. Our findings confirm the fidelity of single particle analysis in accurately resolving complex structural characteristics and heterogeneous structures, exhibiting its potential as a valuable tool for detailed structural analysis of synthetic polymers and soft materials.

Keywords: Cryogenic transmission electron microscopy; single particle analysis; peptoid; molecular dynamics simulation; image simulation.

Introduction:

Synthetic polymers composed of chemically distinct blocks can self-assemble into various amorphous or crystalline micelle-like nanostructures in solution, which exhibit an extensive range of structural variations from the molecular arrangement in the lattice at the atomic scale to the morphologies at the micrometer scale.[1-7] To design and engineer the self-assembled polymer nanostructures, revealing the structure details at the atomic level is of great importance.[8-10]

In addition to characterizing self-assembled polymer nanostructures in reciprocal space through widely-used X-ray and neutron scattering and diffraction methods, imaging in position space is of great importance for structural interpretation.[11-14] For example, atomic force microscopy (AFM) allows for direct observation of two-dimensional (2D) soft materials on substrates at the molecular level.[15-17] However, many self-assembled block copolymers form three-dimensional (3D) structures, whose internal structures cannot be observed by AFM. Cryogenic transmission electron microscopy (cryo-TEM) is a powerful tool for identifying structural characteristics of self-assembled block copolymers, conjugated polymers, and 2D polymers across different length scales using very low electron dose.[18-26] In this method, nanostructures in solution are rapidly frozen in liquid ethane to preserve their natural states as vitrified specimens.[27-31] Furthermore, at liquid nitrogen temperatures, damage to polymers by the electron beam is minimized, allowing for better preservation of atomic details in the micrographs. [32-38]

Despite its advantages, the 2D projections from 3D structures in cryo-TEM micrographs alone is not able to sufficiently recover all structural information. Therefore, 3D imaging techniques such as cryogenic electron tomography (cryo-ET),[39-43] helical reconstruction,[44-47] and

single particle analysis (SPA)[48-53] are employed to retrieve the 3D information from the nanostructures formed by proteins or synthetic polymers. Originally developed for structural biology, single particle analysis has been particularly effective in achieving atomic-scale 3D reconstructions of proteins by identifying subtle inhomogeneities in chain conformation through 2D and 3D classifications.[48] Recent advances in algorithm development have further enhanced its capabilities, enabling *ab initio* 3D reconstruction without prior structural knowledge.[54] This may especially beneficial for studying various polymer nanostructures like vesicles, helices, and nanofibers, which tend to have less ordered structures compared to proteins at atomic and nanoscales.

In a previous study, we synthesized the amphiphilic diblock copolypeptoid, poly(*N*-decylglycine)-block-poly(*N*-2-(2-(2-methoxyethoxy)ethoxy)ethylglycine) (Ndc_{10}-*b*-Nte_{10}), which is a sequence-defined bio-inspired polymer, as shown in Figure 1A.[55] It self-assembled into crystalline nanofibers comprising an ordered Ndc hydrophobic core and an amorphous Nte hydrophilic shell, in the presence of urea molecules, as illustrated in Figure 1B. Consequently, single particle analysis was employed to reconstruct the 3D structure of the nanofibers using the *ab initio* 3D reconstruction and eventually visualize the conformation of individual polymer chains within the nanofiber lattice in real space at a resolution of 3.6 Å.[55] The 3D reconstructions indicated that the backbones of the crystalline Ndc blocks adopt an extended, all-*cis* sigma strand conformation. Additionally, the reconstruction revealed the presence of two symmetrical crystalline domains with the arms of the V-shapes pointing toward each other within the nanofiber (as depicted in Figure 1C and Figure 2A).

Nevertheless, the accuracy of the SPA-reconstructed nanostructures formed by this polypeptoid remains uncertain due to potential reconstruction inaccuracy. These errors may arise

from mixing fiber structures with homogeneous crystalline domains in parallel V-shaped and antiparallel V-shaped chain configurations.[56, 57] Thus, testing the fidelity of the SPA reconstruction of these nanostructures is critical for the further application of the SPA approach to synthetic polymers. In this study, we first carried out molecular dynamics (MD) simulations to prepare three different atomic models of the fibril nanostructure comprising Ndc_{10}-*b*-Nte_{10}. The first model, based on previous experimental results, has peptoids arranged in two parallel V-shaped domains oriented in opposite directions. The second model exhibits a single parallel V-shaped domain, while the third model shows homogeneous antiparallel V-shaped chain configurations in the fibers. These models were then relaxed in water to reach equilibrium. Simulated cryo-TEM images containing different projection views at various angles of fibers in water were generated using CryoSPARC based on these models. The defocus values and spherical aberration used in the simulation were similar to those in the experimental images.[58] The simulated cryo-TEM images were sorted using 2D classification to remove those containing fibers with out-of-plane rotation. The orientation of the fibers in the remaining simulated images was limited to ranges close to the experimental conditions of frozen vitrified specimens. The *ab initio* reconstruction, performed using the selected simulation images, successfully recovered the molecular packing in the atomic model and revealed the presence of two opposing domains in the fiber. These results suggest that SPA *ab initio* reconstructions can be used to identify inhomogeneities and accurately recover the structural details of nanofibers formed by polypeptoids, with reasonable spatial resolution and sufficient fidelity.

Experimental Section:

Materials and cryo-TEM imaging:

The synthesis protocol and characteristics of the polypeptoid Ndc_{10}-*b*-Nte_{10} are elaborated in our previous work.[55] These polypeptoids were dissolved in tetrahydrofuran (THF) with 4M urea at 4 mg/mL, diluted to 2 mg/mL by adding equal volumes of Milli-Q water to form a 1:1 THF/water mix. The mixture was then allowed to evaporate slowly at 4˚C over 14 days in vials, forming self-assembled nanofibers. These nanofibers were vitrified and analyzed using the cryo-TEM, FEI Krios G2 equipped with a K3 camera and energy filter. The data was reconstructed using the softwre package CryoSPARC. Further details on the sample preparation, imaging, and SPA reconstructions are available in the reference 51.

MD Simulations:

The peptoid molecules being simulated were the Ndc_{10}-*b*-Nte_{10}, with a protonated N-terminus and trifluoroacetate counterions. The nanostructures were built with our best understanding of the cryo-TEM 3D reconstructions for the experimental nanocrystals.[55] The simulation box consisted of 6-stack nanofiber with the Ndc side chains in contact, which extended to infinite length with the periodic boundary condition. Each stack consisted of 12 peptoid molecules preassembled in *cis* conformation for the monolayer model (6 molecules for each layer in the bilayer model)[59]. In each stack, all peptoid molecules were placed in the same direction. To achieve a zero-charge system, each protonated N-terminus was paired with a trifluoroacetic acid. The chemical structures and initial peptoid structures are in Figure S1. In this work, the force field was identical to our prior works, [55, 60]which included a parameterized CGenFF-based force field for the peptoid backbone and standard CGenFF for the other chemicals. [61, 62]The MD simulations were conducted on GROMACS 2019.2.[63] The timestep was set to 2 fs. All bonds involving hydrogen atoms were fixed in length using the LINCS algorithm. The simulation started with a 10ns Langevin dynamics for peptoid and urea only. The

Ndc blocks were constrained in position to maintain the preassembled structure, and the amorphous Nte blocks were obtained at this stage. Subsequently, the simulation box was solvated with TIP3P water, and the equilibration run was performed for 90 ns in NpT ensemble at 300K and 1 atm. Afterward, several 50 ns runs were consecutively conducted using the same configuration until the averaged potential energy stopped decreasing. The relaxed peptoid nanocrystals are shown in Figure S2, including a bilayer model and two monolayer models with different stack-to-stack arrangements.

Image Simulation and Reconstruction:

Various software packages have been developed to simulate contrast transfer functions (CTF), electron dose, acceleration voltage, inelastic scattering, ice thickness, signal-to-noise ratio (SNR), and the detective quantum efficiency (DQE) of cameras for image simulations from atomic models. [64-67] In this study, image simulations and subsequent 2D classifications, *ab initio* reconstruction, and homogeneous refinement were performed using CryoSPARC V4.1.[58] A total of 15,000 images were generated from each atomic model, with a signal-to-noise ratio of 0.1 and a rotation randomness of 2 degrees. The defocus range was set from 1,000 to 10,000 Å at 300 KeV, with a spherical aberration of 2.7 mm, matching the FEI Krios G2 used in experimental data collection. High-order aberrations were not simulated. The default parameters in CryoSPARC were used for the 2D classifications of the simulated images. CTF phase flipping was applied based on the parameters used for image generation. The maximum resolution was set to 3 Å in the *ab initio* reconstruction and homogeneous refinement, with an initial low-pass filter set to 12 Å in the homogeneous refinement.

Results and discussion:

The chemical structures of Ndc_{10}-*b*-Nte_{10} and the urea molecule are shown in Figure 1A. The cryo-TEM micrograph in Figure 1B illustrates the morphology of self-assembled nanofibers embedded in vitreous ice. The fibers exhibit alternating bright and dark periodic regions along their long axis, with dark bands representing electron dense molecules. Figures 1C to 1E show orthogonal views of a SPA-reconstructed 3D reconstruction of this fibril structure. The details of cryo-TEM imaging and processing can be found in our previous work.[55] Figure 1C presents the *a*-*c* plane slice of the fiber, revealing four rows of molecular stacks in the reconstruction. The bright spots in the center of the rows represent the end views of the glycine backbones of polypeptoids (contrast is inverted compared to the image in Figure 1B). The V-shaped bright arms extending from the spots are n-decyl side chains. The molecules are arranged in rows along the *a* direction, and each row adopts a parallel V-shape relative to the adjacent rows along the c direction. In the center region of the top view, two segments of peptoid molecules in four rows are observable. It is noteworthy that the V-shaped arms in the two segments point toward the fibril's center along the *a* axis. The domain boundary is indicated by the black dashed line. Figures 1D and 1E show slices in the *b*-*c* and *a*-*b* planes, respectively, showing the side views of the backbone and side chains. It is evident that the arms of peptoid molecules in the same row are angled in opposite directions, indicating that the two segments have inverted symmetry along the *a* direction with the proposed geometry.

This discovery raises questions about the accuracy of SPA reconstruction for polymer nanostructures, which remains to be examined by more studies within the polymer community.[49, 68-70] Specifically, can the reconstruction process accurately identify inhomogeneity and reconstruct molecular structures without assuming symmetry, especially given its absence in these self-assembled polypeptoid nanostructures. Numerous high-resolution 3D structures of

proteins have been resolved using the SPA approach. It is known that structures of highly symmetrical proteins can be resolved at atomic resolution, up to 1.2 Å.[51] However, it is much more difficult for proteins with less symmetry, where the resolution is typically above 3 Å.[71, 72] In addition to low symmetry, which makes it challenging to average each molecule in the protein, the flexible chains in proteins leading to inhomogeneity also pose challenges for SPA reconstruction.[73] It is reasonable to forecast that these challenges are even more significant for the nanostructures formed by synthetic soft materials. This work represents the preliminary effort to establish the baseline for SPA reconstruction of synthetic polymers.

To answer this question, molecular dynamics simulations were conducted to build three atomic models of the nanofiber, as shown in Figure 2. The first model features peptoids arranged in two parallel V-shaped domains oriented in opposite directions. The second model displays a single parallel V-shaped domain, whereas the third model consists of homogeneous antiparallel V-shaped chain configurations within the fibers. Figure 2A displays the first model, which resembles the 3D reconstructed experimental results shown in Figure 1. This model features two segments with opposite orientations of parallel V-shaped packing, as indicated by two arrows. We refer to this model as the bilayer parallel model. The second and third models adopt a single domain configuration, each comprising either all parallel and all antiparallel V-shaped packing of rows, as depicted in Figures 2B and 2C, respectively. These are referred to as the single layer parallel and single layer antiparallel models. In all models, molecules adapt to form corrugated rows along the c-direction, which is particularly evident in the *b*-*c* plane. Figure 2D shows the *b*-*c* plane of the first model with two segments, where the overlap of side chains is observable. The single layer parallel and single layer antiparallel models exhibit the arrangement of rows is the *b*-

c plane as shown in Figures 2D and 2E. Detailed atomic structures are presented in Figure S2 in the supporting information.

Potential maps, which were generated based on the atomic models shown in Figure 2, are presented in Figure 3. In Figure 3A, a slice from the *a-c* plane shows the potential map of the bilayer parallel model depicted in Figure 2A. This map includes the water molecules surrounding the fiber structure, with the end view of glycine backbones in the row appearing brighter than the side chains. Two arrows point to two domains within the atomic model. Figure 3B presents *b-c* plane slices of the two domains, positioned as indicated by the arrows, clearly exhibiting the different orientations of molecules within the same row. Figures 3C and 3D illustrate the *a-c* and *b-c* planes of the potential map for the single layer parallel model shown in Figure 2B. This model features homogeneous V-shaped parallel packing of molecules in the *a-c* plane and a corrugated arrangement of molecules in rows in the *b-c* plane. The potential map for the single layer antiparallel model is displayed in Figures 3E and 3F, where it shows homogeneous V-shaped antiparallel packing of molecules in the *a-c* plane while maintaining the corrugated arrangement in the *b-c* plane.

The simulation strategy in this study is stated as follows: The first step involves generating simulated images from the relaxed atomic models in Figure 2. The second step tests the accuracy of the reconstructions in terms of recovering the atomic details of molecular arrangements using the *ab initio* reconstruction and homogeneous datasets that contain images generated from each model. The third step assesses whether heterogeneous structures can be identified and differentiated during the *ab initio* reconstruction without reference images. To facilitate this test, simulated images from the second and third models are mixed into one dataset to test the reconstruction's ability to identify heterogeneity induced by two homogeneous

nanostructures with different molecular packing. The final test examines whether the location from which images are extracted affects the accuracy in identifying heterogeneity and resolving correct molecular packing. This is achieved by simulating images from various locations on a V-shaped antiparallel model, as indicated by the yellow and blue dashed boxes in Figures 3E and 3F. The examples of simulated images, obtained after applying the appropriate experimental electron optical conditions and doses, are shown in Figures S3 to S5.

Figure 4 shows simulated images corresponding to the strategy for selecting images that mimic experimental conditions. The atomic model, depicted in Figure 4A, shows three cycles indicating the rotation directions used in the image simulation; water molecules are omitted for clarity. The image simulation involved assigning random rotation angles to the atomic model in CryoSPARC. Figure 4B shows the 3D angular distribution of the simulated images, which was exported from CryoSPARC and generated in Chimera.[74] Each cylinder represents a rotation angle, with the height varying proportionally to the relative number of images contain the structure projected at that angle. Colors indicate the image count at each angle: red signifies a higher number, as shown on the color bar.

The simulated images were processed in 2D classification to sort the orientations of the atomic model. In experimental conditions, the freedom of fibers in the thin vitreous ice is more restricted than in the simulations. Most fibers prefer the in-plane orientation with the rotation along the long axis, which is the *c* direction. Classes that contain projection views of the *a*-*c* plane, showing the featured spacings between two rows of molecules (24Å) as bright lines, are selected 2D classification, as shown in Figure 4D. Figure 4E displays the corresponding angular distribution of the selected classes, where projections of the *a*-*b* plane are excluded to better

replicate experimental conditions. Simulated images from three models, including surrounding water molecules, are used to test reconstructions of different models in subsequent steps.

Figure 5 presents the results of the 3D reconstruction of homogeneous structures after *ab initio* reconstruction and homogeneous refinement in CryoSPARC. The angular distribution and Fourier Shell Correlation (FSC) are detailed in Figure S6. The slice views of the bilayer model reconstruction are depicted in Figures 5A and 5B. In Figure 5A, the bright regions in the *a-c* plane represent the end views of glycine backbones and side chains arranged in V-shaped rows. Two arrows, labeled 1 and 2, point to the two layers, with corresponding *b-c* planes shown in Figure 5B. These slices clearly exhibit the distinct orientation of molecules in the two layers. The reconstructions of models with parallel and antiparallel V-shaped packing are shown in Figures 5C to 5F. The parallel V-shaped packing, where molecules appear as electron-dense bright regions, is clearly visualized in the *a-c* plane in Figure 5C, while the *b-c* plane shows the alignment of molecular rows, identical to the atomic model (Figures 2 and S2). Similarly, the rows of molecules arranged in antiparallel V-shaped packing are distinctly visible in the *a-c* plane in Figure 5E, with the same arrangement observable in the *b-c* plane in Figure 5F. The results from different homogeneous fiber structures suggest that single particle analysis can accurately reconstruct homogeneous nanostructures formed by polypeptoids using a limited number of projections from restricted projection angles. Compared to the angular distribution of simulated images, the angular distribution of projections used to reconstruct the final 3D map differs, (Figure S7). This indicates that the 3D reconstruction did not simply utilize the original Euler angles assigned to the simulated images but instead refined the angular distribution to achieve the final reconstruction and simulate experimental conditions. However, the 3D reconstruction is not able to completely recover the original orientations, leading to lower

resolution in the reconstructed structure, similar to scenarios encountered in experimental data processing which are usually less than ideal conditions.

Although the reconstruction of homogeneous structures demonstrates high fidelity and accuracy, the subsequent question concerns whether it can also identify the presence of heterogeneous fiber structures. To test this, simulated images generated from two single layer models, parallel and antiparallel V-shaped packings, were combined into a mixed dataset to simulate the prence of heterogeneity. The *ab initio* reconstruction of the structures with inhomogneity is displayed in Figure 6. Two initial structures were generated using CryoSPARC, with their angular distributions shown in Figure S8. As depicted in Figure 6A, the first initial reconstruction exhibits antiparallel packing in the a-c plane and opposing directions of neighboring rows in the *b-c* plane. The second initial reconstruction, clearly showing antiparallel packing in the *a-c* plane and opposing directions of neighboring rows in the *b-c* plane, is presented in Figure 6B. The *ab initio* reconstruction successfully identifies the heterogeneity and reconstructs the structures separately. This suggests that the experimental data primarily shows homogeneous fiber structures, as the *ab initio* reconstruction produced a single structure even though more initial structures were prompted at this stage.

The final test examines whether the location from which structures are extracted along fibers impacts the reconstruction results. Two sets of simulated images, generated from different locations on an antiparallel V-packing model, were mixed to form the test dataset (Figures S9 and S10). The same *ab initio* reconstruction process was applied to detect potential heterogeneity, with results shown in Figure 7. Figure 7A shows the antiparallel packing in the *a-c* plane, with the opposing directions of neighboring rows visualized in the *b-c* plane. The second initial structure is depicted in Figure 7B. No distinct structure emerged from the *ab initio*

reconstruction since the fiber structure is homogeneous and was correctly reconstructed in Figure 7A. The reconstruction result indicates that the location of fiber extraction does not significantly affect the ab initio 3D reconstruction.

Conclusions:

In this study, we successfully employed a combination of MD simulations and image simulations to validate the fidelity of cryo-TEM 3D reconstructions of self-assembled polypeptoid fibril nanostructures. This approach leverages advanced techniques to bridge the gap between theoretical models and experimental verification, demonstrating significant advancements in the field of cryo-TEM imaging of synthetic polymers. Three atomic models, designed based on prior cryo-TEM experimental results, featured different molecular packings and were constructed and relaxed to enhance the realism of image simulations. CryoSPARC, a well-established software package, was utilized for image simulation, 2D classifications, *ab initio* reconstruction, and homogeneous refinement, ensuring high accuracy and reliability in our simulation processes. We tested three key hypotheses. The first test is the recovery of molecular details. The *ab initio* reconstruction and homogeneous refinement successfully recovered complicated molecular details, including chain conformation and arrangement, from images generated at limited projection angles of fibril structures (Figure S11). This result proves the accuracy of the single particle analysis in capturing complex molecular geometries in self-assembled structures. The second test is the identification of heterogeneous structures. We examined the capability to detect heterogeneous fibril structures by reconstructing a mixture of images from two distinct molecular packings. The *ab initio* reconstruction effectively separated the mixed structures and accurately depicted two molecular packings with precise molecular details. The final test is the influence of extraction location. We explored whether the location of

structure extraction along the fibers impacts the reconstruction results. The findings indicate that the extraction location does not affect the ability to resolve the correct fibril structures.

These results highlight that the 3D reconstructions of fibril nanostructures using single particle analysis in CryoSPARC exhibit exceptional accuracy and fidelity in resolving structural details at the molecular level. This study validates our methodological approach and enhances the reliability of single particle analysis as a promising tool in the study of self-assembled crystalline fibers. Moreover, this work will shed light on the path for future applications of single particle 3D reconstruction across other polymer nanostructures and aid in the accurate and confident interpretation of complex molecular assemblies, where precise structural insights are required.

Acknowledgement:

Primary funding for this work was provided by the Soft Matter Electron Microscopy Program (KC11BN), supported by the Office of Science, Office of Basic Energy Science, US Department of Energy, under Contract DE-AC02-05CH11231. Work at the Molecular Foundry at Lawrence Berkeley National Laboratory was supported by user projects at these user facilities, supported by the Office of Science, Office of Basic Energy Sciences, of the US Department of Energy under Contract DE-AC02-05CH11231. Micrographs presented here were obtained at the Berkeley Bay Area Cryo-EM facility (BACEM) at the University of California, Berkeley.

Data Availability Statement:

Atomic models generated from MD simulations are available in the supporting information. For access to the cryo-TEM micrographs and metadata from the single-particle analysis, please contact the authors directly.

Conflict of interests:

The authors declare no competing financial interest.

Author contributions:

N.P.B and X.J designed the experiment. X.B.L and D.P carried the MD simulations. M.S and X.J carried the image simulations. Y.J. L, T.Y.Y, and R.N.Z. synthesized the polypeptoid. X.B.L and X.J wrote the manuscript. All authors contributed to the revision of manuscript.

Supporting information:

Figures S1 to S10

Atomic models 1, 2, and 3

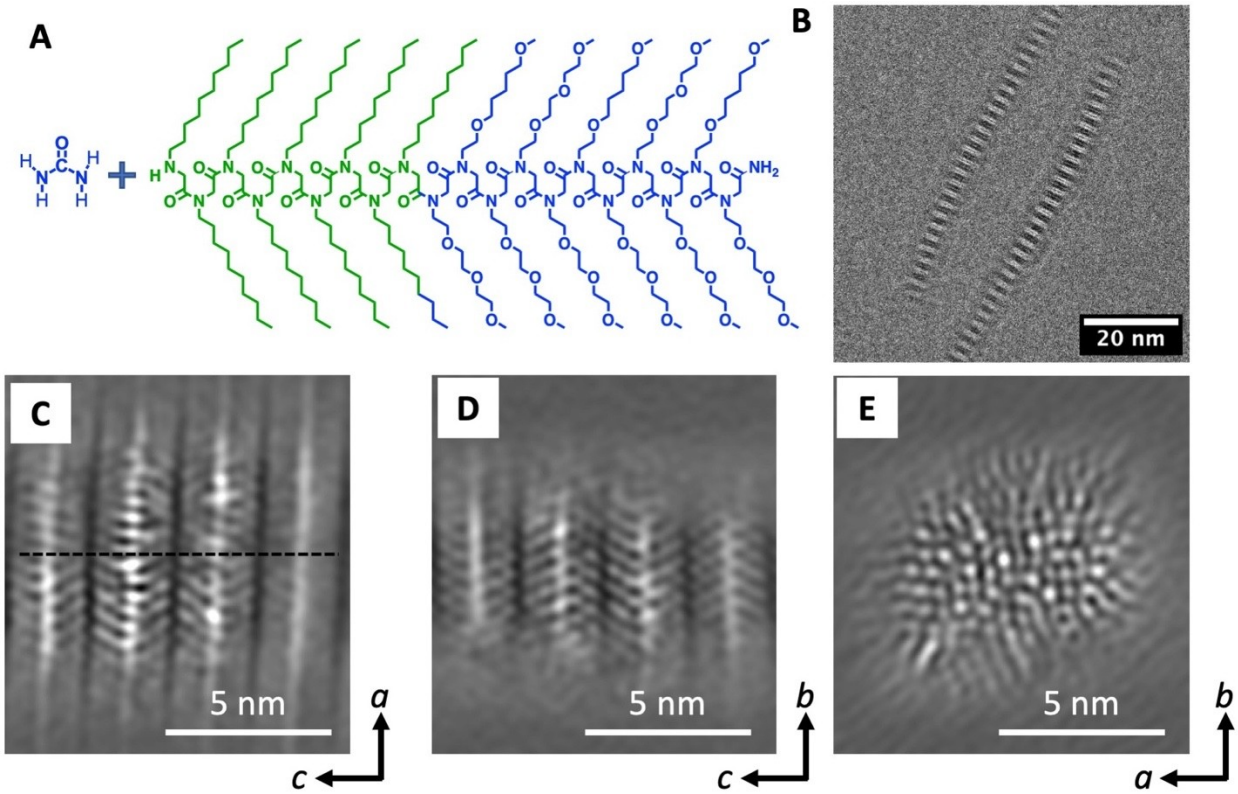


Figure 1. A. Chemical structure of peptoid Ndc_{10}-*b*-Nte_{10} with urea, **B**. Cryo-TEM image shows the morphology of nanofibers formed by Ndc_{10}-*b*-Nte_{10}. **C**. The slice view of the *a*-*c* plane in the 3D reconstruction of the Ndc_{10}-*b*-Nte_{10} nanofibers. **D**. The slice view of the *b*-*c* plane, and **E**. the slice view of the *b*-*a* plane. The bright region in slices represent the electron dense peptoid molecules (the crystallized Ndc block).

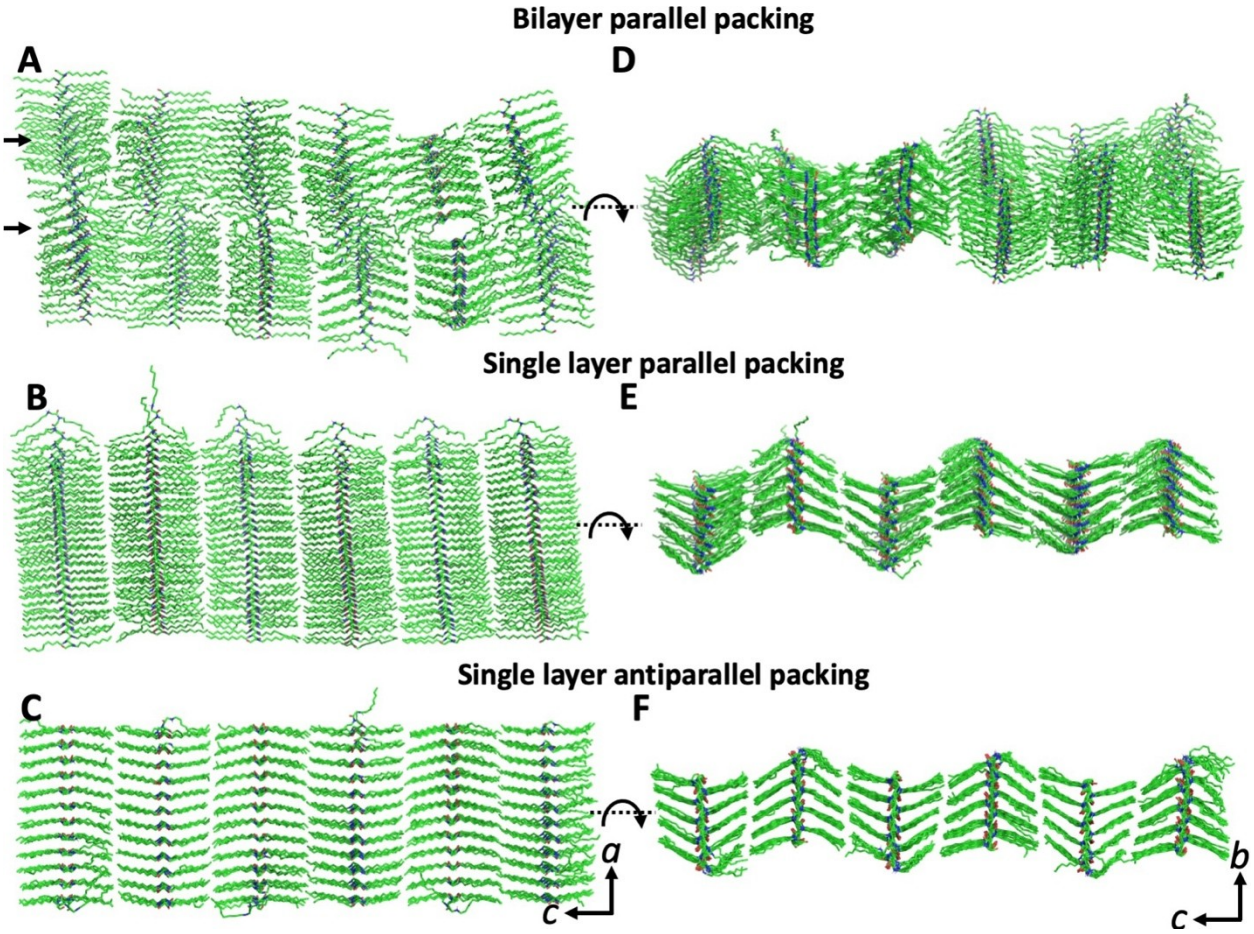


Figure 2. Three atomic models of the nanofiber structure obtained from MD simulations. **A**. The first configuration features a bilayer fibril structure composed of two segments, each formed by six rows of peptoid molecules arranged in parallel V-shaped packing with aliphatic side chains emanating from the *cis* glycine backbones. The two segments are oriented oppositely, with the V-shaped side chains at the end of one segment facing those at the end of the other, as indicated by the boxed region (close to experimental results). **B**. The second configuration is a single layer fibril structure comprising six rows of molecules arranged in parallel V-shaped packing. **C**. The third configuration is a single layer fibril structure with six rows of molecules arranged in an antiparallel V-shaped packing. Only the crystallized Ndc blocks are shown. The *b-c* planes of three models are displayed in the right columns as **D, E** and **F**, respectively.

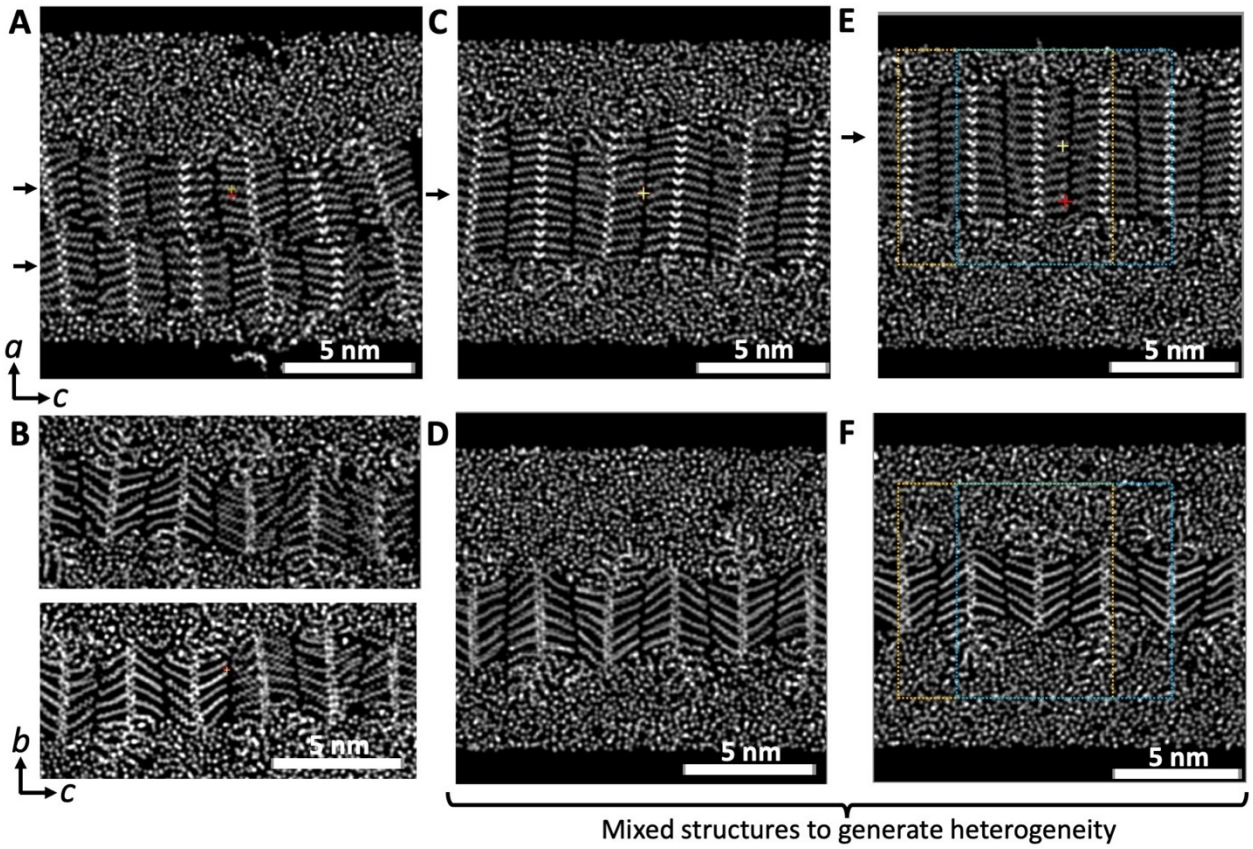


Figure 3. Slice views of potential maps obtained from different atomic models. **A**. The *a*-*c* and, **B.** *b*-*c* planes of the bilayer fiber structure, two slices of *b*-*c* planes corresponding to locations indicated by arrows, are shown in the bottom row. **C**. The *a*-*c* and, **D.** *b*-*c* planes of the single layer fiber structure with parallel V-shaped packing; a slice of *b*-*c* plane corresponding to the location indicated by the arrow is shown in the bottom row. **E**. The *a*-*c* and, **F.** *b*-*c* planes of the single layer fiber structure with antiparallel V-shaped packing; a slice of *b*-*c* plane corresponding to the location indicated by the arrow is shown in the bottom row. The black regions in the images represent vacuum areas, while the disordered bright spots around the ordered Ndc blocks in fiber structures are water molecules and Nte blocks, which are amorphous in water.

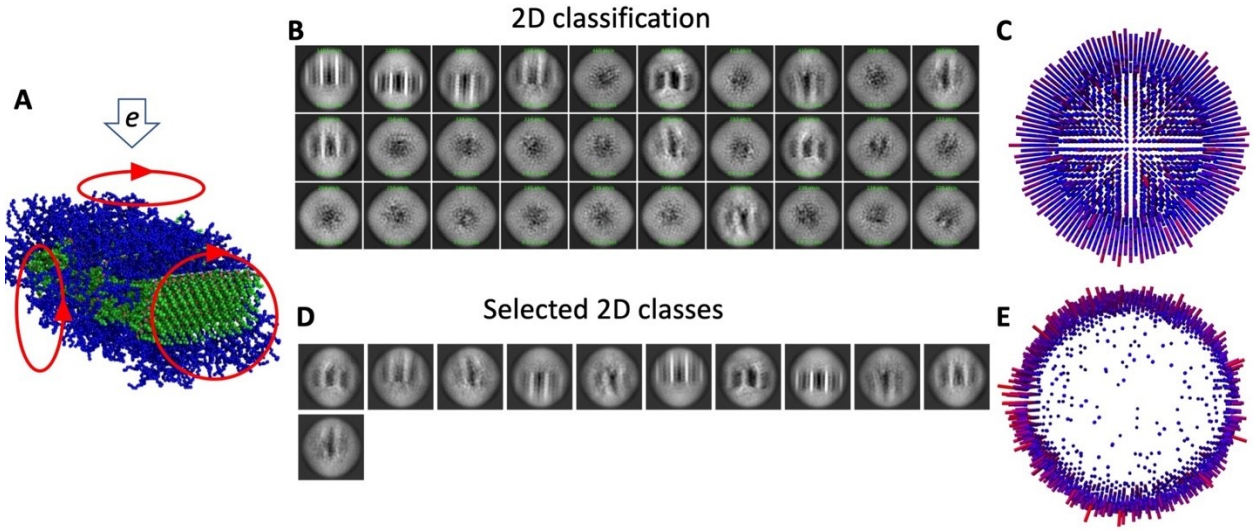


Figure 4. Strategy for image simulations. **A**. The schematic view illustrates the rotations of an atomic model during image simulation. The hydrophilic blocks (Nte) and hydrophobic blocks (Ndc) are depicted in blue and green, respectively. The arrow indicates the projection direction. **B**. The 2D classification result obtained after sorting the simulated images. **C**. Angular distribution of the simulated images, where the red column indicates a larger number of images and the blue column a smaller number of images at each projection angle. **D**. The selected classes from the 2D classification are shown in B. E. Angular distribution of the simulated images within the selected classes.

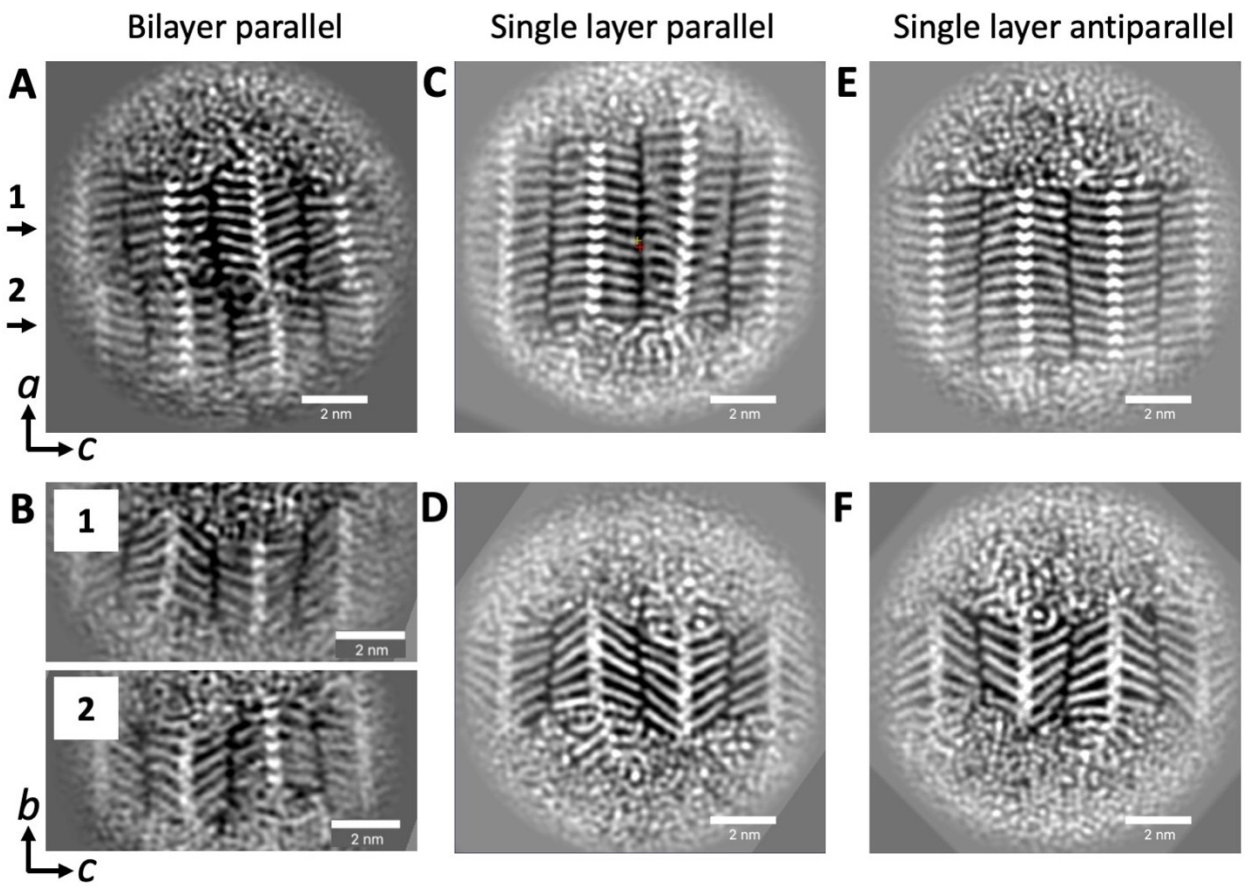


Figure 5. Slice views of reconstructed maps obtained from simulated images after homogenous refinement in Cryosparc. **A**. The *a-c* and, **B.** *b-c* planes of the bilayer fiber structure, two slices of *b-c* planes corresponding to locations indicated by arrows, are shown in the bottom row. **C**. The *a-c* and, **D.** *b-c* planes of the single layer fiber structure with parallel V-shaped packing; a slice of *b-c* plane corresponding to the location indicated by the arrow is shown in the bottom row. **E**. The *a-c* and, **F.** *b-c* planes of the single layer fiber structure with antiparallel V-shaped packing; a slice of *b-c* plane corresponding to the location indicated by the arrow is shown in the bottom row. The ordered Ndc blocks are brighter (electron dense) as compared to amorphous Nte blocks and surrounding water molecules. The gray corner regions in images are caused by the circular mask applied during reconstruction.

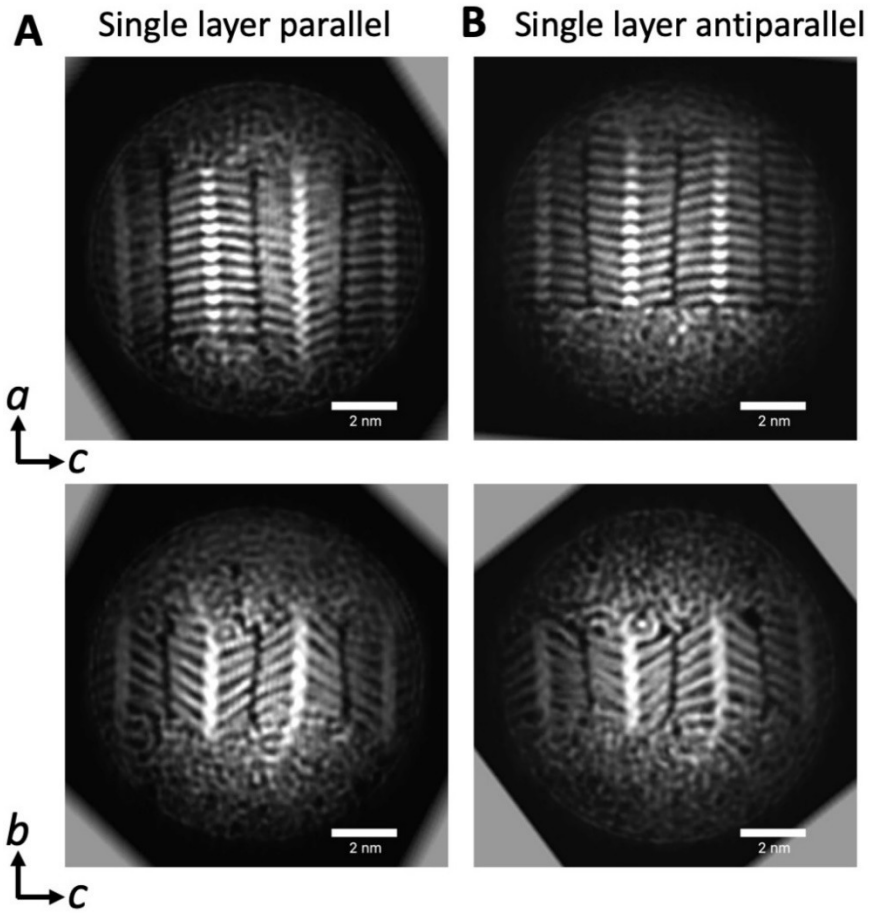


Figure 6. Slice views of reconstructed maps obtained from a mixture of simulated images obtained from single layer parallel and anti-parallel V-shaped fiber structures after the *ab initio* reconstruction. **A**. The *a-c* (top) and *b-c* (bottom) planes of the single layer fiber structure, two slices of *b-c* planes corresponding to locations indicated by arrows, are shown in the bottom row. **B**. The *a-c* (top) and *b-c* (bottom) planes of the single layer fiber structure with single layer antiparallel V-shaped packing; a slice of *b-c* plane corresponding to the location indicated by the arrow is shown in the bottom row. The ordered Ndc blocks are brighter (electron dense) as

compared to amorphous Nte blocks and surrounding water molecules. The gray corner regions in images are caused by rotating images to the same angle.

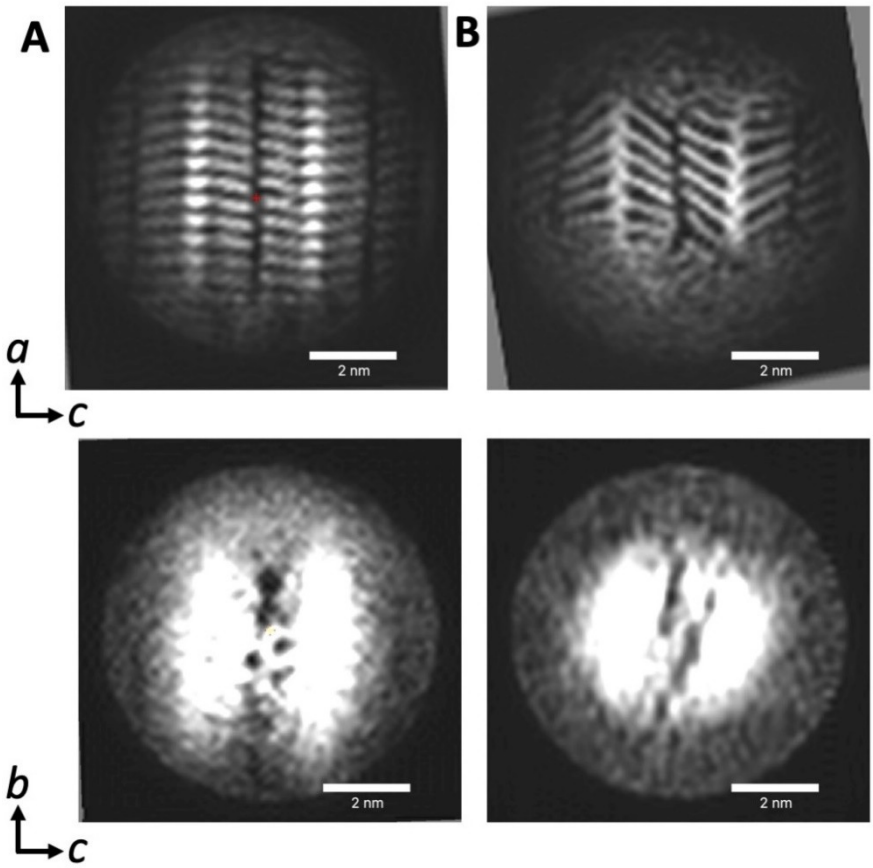


Figure 7. Slice views of reconstructed maps obtained from a mixture of simulated images generated at the different location on the fiber with single layer antiparallel V-shaped packing after the ab initio reconstruction. **A**. The *a*-*c* (top) and *b*-*c* (bottom) planes of the bilayer fiber structure, two slices of *b*-*c* planes corresponding to locations indicated by arrows, are shown in the bottom row. **B**. The *a*-*c* (top) and *b*-*c* (bottom) planes of the single layer fiber structure with parallel V-shaped packing; a slice of *b*-*c* plane corresponding to the location indicated by the arrow is shown in the bottom row. The ordered Ndc blocks are brighter (electron dense) than amorphous Nte blocks and surrounding water molecules. The gray corner regions in images are caused by rotating images to the same angle.

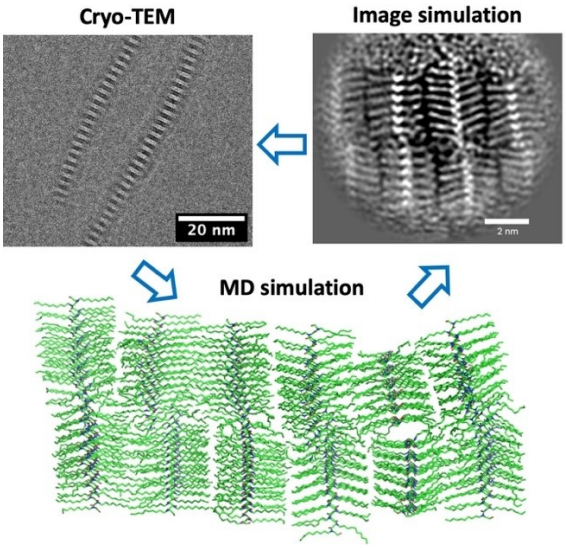
References:

[1] D. E. Discher, A. Eisenberg, *Science* **2002**, *297*, 967.
[2] Y. Y. Mai, A. Eisenberg, *Chemical Society Reviews* **2012**, *41*, 5969.
[3] M. Almgren, W. Brown, S. Hvidt, *Colloid Polym Sci* **1995**, *273*, 2.
[4] S. Jain, F. S. Bates, *Macromolecules* **2004**, *37*, 1511.
[5] K. R. Shull, K. I. Winey, E. L. Thomas, E. J. Kramer, *Macromolecules* **1991**, *24*, 2748.
[6] Y. Y. Won, H. T. Davis, F. S. Bates, *Science* **1999**, *283*, 960.
[7] J. Wu, E. M. Pearce, T. K. Kwei, A. A. Lefebvre, N. P. Balsara, *Macromolecules* **2002**, *35*, 1791.
[8] R. K. O'Reilly, M. J. Joralemon, K. L. Wooley, C. J. Hawker, *Chemistry of Materials* **2005**, *17*, 5976.
[9] R. C. Hayward, D. J. Pochan, *Macromolecules* **2010**, *43*, 3577.
[10] K. E. B. Doncom, L. D. Blackman, D. B. Wright, M. I. Gibson, R. K. O'Reilly, *Chemical Society Reviews* **2017**, *46*, 4119.
[11] C. V. Berney, R. E. Cohen, F. S. Bates, *Polymer* **1982**, *23*, 1222.
[12] B. Chu, B. S. Hsiao, *Chemical Reviews* **2001**, *101*, 1727.
[13] D. Uhrikova, N. Kucerka, A. Islamov, A. Kuklin, V. Gordeliy, P. Balgavy, *Biochimica Et Biophysica Acta-Biomembranes* **2003**, *1611*, 31.
[14] J. L. Putaux, E. Minatti, C. Lefebvre, R. Borsali, M. Schappacher, A. Deffieux, *Faraday Discussions* **2005**, *128*, 163.
[15] S. Scherb, A. Hinaut, R. Pawlak, J. G. Vilhena, Y. Liu, S. Freund, Z. Liu, X. L. Feng, K. Müllen, T. Glatzel, A. Narita, E. Meyer, *Commun Mater* **2020**, *1*.
[16] D. Ikeshima, K. Miyamoto, A. Yonezu, *Polymer* **2019**, *173*, 80.
[17] J. Kumaki, *Polymer Journal* **2016**, *48*, 3.
[18] Y. Zheng, Y. Y. Won, F. S. Bates, H. T. Davis, L. E. Scriven, Y. Talmon, *Journal of Physical Chemistry B* **1999**, *103*, 10331.
[19] Y. Y. Won, A. K. Brannan, H. T. Davis, F. S. Bates, *Journal of Physical Chemistry B* **2002**, *106*, 3354.
[20] D. Danino, A. Bernheim-Groswasser, Y. Talmon, *Colloids and Surfaces a-Physicochemical and Engineering Aspects* **2001**, *183*, 113.
[21] H. Cui, T. K. Hodgdon, E. W. Kaler, L. Abezgauz, D. Danino, M. Lubovsky, Y. Talmon, D. J. Pochan, *Soft Matter* **2007**, *3*, 945.
[22] S. Zhong, D. J. Pochan, *Polymer Reviews* **2010**, *50*, 287.
[23] C. J. Newcomb, T. J. Moyer, S. S. Lee, S. I. Stupp, *Current Opinion in Colloid & Interface Science* **2012**, *17*, 350.
[24] H. Weissman, B. Rybtchinski, *Current Opinion in Colloid & Interface Science* **2012**, *17*, 330.
[25] Y. Talmon, D. F. Evans, B. W. Ninham, *Science* **1983**, *221*, 1047.
[26] B. K. Liang, Y. Y. Zhang, C. Leist, Z. W. Ou, M. Polozij, Z. Y. Wang, D. Mücke, R. H. Dong, Z. K. Zheng, T. Heine, X. L. Feng, U. Kaiser, H. Y. Qi, *Nature Communications* **2022**, *13*.
[27] Y. Talmon, H. T. Davis, L. E. Scriven, E. L. Thomas, *Rev. Sci. Instrum.* **1979**, *50*, 698.

[28] J. R. Bellare, H. T. Davis, L. E. Scriven, Y. Talmon, *Journal of Electron Microscopy Technique* **1988**, *10*, 87.
[29] K. A. Taylor, R. M. Glaeser, *Journal of Ultrastructure Research* **1976**, *55*, 448.
[30] Y. Talmon, *Ultramicroscopy* **1985**, *17*, 167.
[31] J. Dubochet, M. Adrian, J. J. Chang, J. C. Homo, J. Lepault, A. W. Mcdowall, P. Schultz, *Quarterly Reviews of Biophysics* **1988**, *21*, 129.
[32] R. M. Glaeser, *Journal of Ultrastructure Research* **1971**, *36*, 466.
[33] D. T. Grubb, A. Keller, *Journal of Materials Science* **1972**, *7*, 822.
[34] D. T. Grubb, *Journal of Materials Science* **1974**, *9*, 1715.
[35] R. F. Egerton, P. Li, M. Malac, *Micron* **2004**, *35*, 399.
[36] R. F. Egerton, *Ultramicroscopy* **2013**, *127*, 100.
[37] R. F. Egerton, *Micron* **2019**, *119*, 72.
[38] B. Kuei, M. P. Aplan, J. H. Litofsky, E. D. Gomez, *Materials Science & Engineering R-Reports* **2020**, *139*.
[39] H. X. Sui, K. H. Downing, *Nature* **2006**, *442*, 475.
[40] A. L. Parry, P. H. H. Bomans, S. J. Holder, N. A. J. M. Sommerdijk, S. C. G. Biagini, *Angewandte Chemie-International Edition* **2008**, *47*, 8859.
[41] O. Le Bihan, P. Bonnafous, L. Marak, T. Bickel, S. Trepout, S. Mornet, F. De Haas, H. Talbot, J. C. Taveau, O. Lambert, *J Struct Biol* **2009**, *168*, 419.
[42] F. Nudelman, G. de With, N. A. J. M. Sommerdijk, *Soft Matter* **2011**, *7*, 17.
[43] P. Ercius, O. Alaidi, M. J. Rames, G. Ren, *Advanced Materials* **2015**, *27*, 5638.
[44] K. R. Pothula, D. Smyrnova, G. F. Schroder, *Ultramicroscopy* **2019**, *203*, 132.
[45] E. H. Egelman, *Methods in Enzymology, Vol 482: Cryo-Em, Part B: 3-D Reconstruction* **2010**, *482*, 167.
[46] E. H. Egelman, *Abstracts of Papers of the American Chemical Society* **2013**, *245*.
[47] F. B. Wang, O. Gnewou, A. Solemanifar, V. P. Conticello, E. H. Egelman, *Chemical Reviews* **2022**.
[48] J. Frank, "*Three-dimensional electron microscopy of macromolecular assemblies : visualization of biological molecules in their native state"*, 2nd edition, Oxford University Press, New York, 2006, p. xiv.
[49] H. D. Tagare, A. Kucukelbir, F. J. Sigworth, H. W. Wang, M. Rao, *Journal of Structural Biology* **2015**, *191*, 245.
[50] E. B. Pichkur, P. Kasatsky, T. N. Baymukhametov, Y. M. Chesnokov, M. Y. Presniakov, A. Miasnikov, A. L. Konevega, A. L. Vasiliev, M. V. Kovalchuk, *Acta Crystallographica a-Foundation and Advances* **2017**, *73*, C1297.
[51] T. Nakane, A. Kotecha, A. Sente, G. McMullan, S. Masiulis, P. M. G. E. Brown, I. T. Grigoras, L. Malinauskaite, T. Malinauskas, J. Miehling, T. Uchanski, L. B. Yu, D. Karia, E. V. Pechnikova, E. de Jong, J. Keizer, M. Bischoff, J. McCormack, P. Tiemeijer, S. W. Hardwick, D. Y. Chirgadze, G. Murshudov, A. R. Aricescu, S. H. W. Scheres, *Nature* **2020**, *587*, 152.
[52] M. Beckers, D. Mann, C. Sachse, *Prog Biophys Mol Bio* **2021**, *160*, 26.
[53] I. Lazic, M. Wirix, M. L. Leidl, F. de Haas, D. Mann, M. Beckers, E. V. Pechnikova, K. Müller-Caspary, R. Egoavil, E. G. T. Bosch, C. Sachse, *Nature Methods* **2022**, *19*, 1126.

[54] E. D. Zhong, A. Lerer, J. H. Davis, B. Berger, "CryoDRGN2: Ab initio neural reconstruction of 3D protein structures from real cryo-EM images", in *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, 20214046.
[55] T. Y. Yu, X. B. Luo, D. Prendergast, G. L. Butterfoss, B. Rad, N. P. Balsara, R. N. Zuckermann, X. Jiang, *Acs Nano* **2023**, *17*, 4958.
[56] X. Jiang, D. R. Greer, J. Kundu, C. Ophus, A. M. Minor, D. Prendergast, R. N. Zuckermann, N. P. Balsara, K. H. Downing, *Macromolecules* **2018**, *51*, 7794.
[57] X. Jiang, S. T. Xuan, J. Kundu, D. Prendergast, R. N. Zuckermann, N. P. Balsara, *Soft Matter* **2019**, *15*, 4723.
[58] A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, *Nature Methods* **2017**, *14*, 290.
[59] D. R. Greer, M. A. Stolberg, J. Kundu, R. K. Spencer, T. Pascal, D. Prendergast, N. P. Balsara, R. N. Zuckermann, *Journal of the American Chemical Society* **2018**, *140*, 827.
[60] X. B. Luo, T. Y. Yu, N. K. Li, R. N. Zuckermann, X. Jiang, N. P. Balsara, D. Prendergast, *Acs Nano* **2024**, *18*, 14917.
[61] K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, A. D. MacKerell, *Journal of Computational Chemistry* **2010**, *31*, 671.
[62] L. J. Weiser, E. E. Santiso, *Journal of Computational Chemistry* **2019**, *40*, 1946.
[63] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindahl, *SoftwareX* **2015**, *1-2*, 19.
[64] R. J. Hall, E. Nogales, R. M. Glaeser, *Journal of Structural Biology* **2011**, *174*, 468.
[65] Z. G. Shang, F. J. Sigworth, *Journal of Structural Biology* **2012**, *180*, 10.
[66] M. Vulovic, R. B. G. Ravelli, L. J. van Vliet, A. J. Koster, I. Lazic, U. Lücken, H. Rullgård, O. Öktem, B. Rieger, *Journal of Structural Biology* **2013**, *183*, 19.
[67] B. A. Himes, N. Grigorieff, *Iucrj* **2021**, *8*, 943.
[68] D. Messmer, C. Böttcher, H. Yu, A. Halperin, K. Binder, M. Kröger, A. D. Schlütert, *Acs Nano* **2019**, *13*, 3466.
[69] R. Hlushko, E. Pozharski, V. M. Prabhu, A. K. Andrianov, *Commun Mater* **2024**, *5*.
[70] X. Jiang, R. K. Spencer, J. Sun, C. Ophus, R. N. Zuckermann, K. H. Downing, N. P. Balsara, *Journal of Physical Chemistry B* **2019**, *123*, 1195.
[71] M. F. Liao, E. H. Cao, D. Julius, Y. F. Cheng, *Nature* **2013**, *504*, 107.
[72] G. Kang, A. T. Taguchi, J. Stubbe, C. L. Drennan, *Science* **2020**, *368*, 424.
[73] A. Punjani, D. J. Fleet, *Nature Methods* **2023**, *20*, 860.
[74] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *Journal of Computational Chemistry* **2004**, *25*, 1605.

Table of Contents



Molecular dynamics (MD) and image simulations were used to validate the accuracy of cryogenic transmission electron microscopy (cryo-TEM) 3D reconstructions of polypeptoid fibril nanostructures. We examined molecular detail recovery, structural heterogeneity, and the effect of extraction location. The findings demonstrate that single particle analysis effectively resolves complex geometries and heterogeneity in synthetic polymers.