Title

Effect of processing and end groups on the crystal structure of polypeptoids studied by cryogenic electron microscopy at atomic length scales

Journal

Soft Matter, 15(23)

ISSN

1744-683X

Authors

Jiang, Xi
Xuan, Sunting
Kundu, Joyjit
et al.

Publication Date

2019-06-12

DOI

10.1039/c9sm00633h

Peer reviewed

Effect of Processing and End Groups on the Crystal Structure of Polypeptoids Studied by Cryogenic Electron Microscopy at Atomic Length Scales

Xi Jiang[1&], Sunting Xuan[2&], Joyjit Kundu[2#], David Prendergast[2], Ronald N. Zuckermann[2], Nitash P. Balsara[1,3]**

[1] Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, United States.

[2] Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, United States.

[3] College of Chemistry, University of California, Berkeley, CA 94720, United States.

& Authors contributed equally to this work.

Present address: Department of Chemistry, Duke University, Durham, North Carolina 27708 United States

* *Corresponding authors: rnzuckermann@lbl.gov; nbalsara@berkeley.edu.*

Submission to Soft Matter

ABSTRACT

Cryogenic electron microscopy at atomic length scales was used to study the structure of self-assembled crystalline nanosheets obtained from a series of polypeptoids with the same chain architecture but with different end groups. While long-range order is enhanced by slowing down the self-assembly process, the dominant crystalline motif was found to be a sensitive function of both processing details and end group chemistry. In some cases, adjacent rows of polypeptoid molecules adopt anti-parallel V-shaped side chain conformations. In other cases, adjacent rows of polypeptoid molecules adopt parallel V-shaped side chain conformations. Interestingly, the unit cell is rectangular in both cases with dimensions a = 4.5 Å and c = 50 Å. In all cases, long-range order, quantified by the average number of concatenated unit cells of the same type, is more prevalent along the a direction.

INTRODUCTION

Polypeptoids are bio-inspired polymers comprising repeating N-substituted glycine monomer units. They are similar to polypeptides except for the fact that the side chain is appended to the nitrogen atom rather than the α-carbon[1]. Polypeptoids have demonstrated the ability to form various nanostructures, such as ribbons,[2] helices,[3, 4] tubes,[5, 6] and two-dimensional (2D) crystalline nanosheets.[2, 7-13] Polypeptoid crystalline nanosheets serve as functional protein mimetics. They have been effectively functionalized for the recognition of specific protein targets[9, 12], for templated mineral growth,[14] and

for controlling catalytic activity.[15] Both the internal structure and surface properties of the nanosheets can be designed at the atomic level as they can be synthesized by automated solid-phase synthesis.[16, 17]

Nanosheets can be generated by the self-assembly of polypeptoids in dilute solutions[5, 18] or by compressing polypeptoid molecules at either the air-water or the oil-water interface.[8, 9] Ultra-low polydispersity of molecular weight and the absence of chirality make polypeptoids an ideal platform for probing molecular self-assembly principles and crystallization.[1, 19] Molecular self-assembly depends on several factors such as the hydrophilicity of side chains, end groups, and processing details. The purpose of this paper is to shed light on this dependence.

The unit cell of a variety of crystalline polypeptoids is orthorhombic cell[5, 12, 18, 20-23]. Following the nomenclature in ref.22, the *a* dimension represents the inter-chain spacing (4.5 Å for all crystalline polypeptoids), the *b* dimension corresponds to the chain length in an extended conformation, and the *c* dimension is related to the width of the side chains. Recent studies based on X-ray scattering[5, 22, 23], MD simulation[24], NMR[25], and high resolution electron microscopy imaging[18], indicate that the chains adopt an all-*cis* conformation in the crystalline state.

Transmission electron microscopy (TEM) is a powerful tool for studying 2D crystalline structures comprising soft materials, such as polymer single crystals and proteins membranes.[26-38] Like proteins, polypeptoids are unstable to electron beam exposure.[18] Minimizing the damage caused by the

electron beam is essential when atomic level information of polypeptoids is desired. This is enabled by low-dose cryogenic electron microscopy (cryo-EM).[39-41] The effect of radiation damage is minimized when the specimen is cooled to cryogenic temperatures leading to the preservation of high spatial frequency information.[42, 43] Vitrification results in the preservation of the natural state of the specimen. In a previous study, we determined the structure of polypeptoid crystalline nanosheets at around 2 Å resolution by using cryo-EM. Low-dose images were sorted and averaged using methods developed for structural biology.[18] The analysis indicated the presence of numerous distinct crystalline motifs. Other studies on semi-crystalline polymers such as poly(paraxylylene)[27, 44] and syndiotactic polystyrene,[45-47] also suggested the presence of structural heterogeneity.

In this work, we study three different polypeptoids with the same monomer sequence but with different end groups. Nanosheets were obtained by dissolving polypeptoids in a mixture of solvents and then removing the more volatile solvent by evaporation. We determine the structure of the vitrified, hydrated nanosheets using cryo-EM at atomic length scales, as reported in our previous study.[18] The nanosheets obtained by slow evaporation exhibit less structural heterogeneity. Changes in end group chemistry and processing result in surprising changes of the dominant crystal motif. Correlations between neighboring unit cells were determined from the distribution maps obtained from cryo-EM image analysis.

EXPERIMENTAL SECTION

Polypeptoid synthesis

The 2-(2-(2-methoxyethoxy)ethoxy)ethylamine was purchased from Peptide Solutions, Inc. (98% purity). Decylamine and 4-methylpiperidine were purchased from TCI (>98% purity). All submonomers were used without further purification. *N*,*N*'-diisopropylcarbodiimide was purchased from Chem-Impex International, Inc. (≥99% purity). Bromoacetic acid was purchased from ACROS Organics (99% purity). Trifluoroacetic acid was purchased from Sigma Aldrich (99% purity). Acetic anhydride (98% purity), chloroacetic anhydride (95% purity), and iodoacetic anhydride (95% purity) were purchased from Sigma-Aldrich. All other solvents (HPLC grade) used in this study were purchased from Millipore. Automated solid-phase submonomer synthesis of the diblock copolypeptoids was performed on a Symphony X peptide synthesizer at a scale of 200 mg Rink amide resin (0.64 mmol/g, Novabiochem, San Diego, CA) following published procedures.[22, 23] Bromoacetylation reactions were performed on the resin with bromoacetic acid (0.8 M) and *N*,*N*'-diisopropylcarbodiimide (DIC, 0.8 M) in DMF for 20 min at room temperature. Displacement of the bromide with the submonomers were performed at amine concentrations of 1 M in DMF for 30 min at room temperature. The crude peptoids were cleaved from resin by the addition of an acid cleavage cocktail of dichloromethane/trifluoroacetic acid/water (50/45/5, v/v) for 10 min at room temperature. The resin was filtered and

washed with DCM. The filtrate was evaporated to obtain the crude peptoid, followed by lyophilization from acetonitrile/water (1:1, v/v).

N-terminal acetylation was performed on the crude peptoid (~200 mg). The polypeptoids were dissolved in 2 mL THF followed by the addition of the anhydride (20 eq.) and pyridine (40 eq.). The mixture was stirred at room temperature for 20 minutes followed by evaporation of the solvents. The obtained crude product was lyophilized from acetonitrile/water (1:1, v/v).

The crude acetylated polypeptoids were purified on a reverse phase HPLC Waters prep system using a XSelect HSS cyano column (5 µm, 18 x 150 mm) with 50 – 95% B gradient over 20 minutes at a flow rate of 20 mL/min (where solvent A was 10% isopropyl alcohol (IPA) in water with 0.1% TFA, and solvent B was 10% IPA in acetonitrile with 0.1% TFA). The pure fractions were collected, evaporated, and lyophilized from ACN/water (1:1, v/v) to afford a fluffy white powder. The purified polypeptoid was then analyzed by reverse phase HPLC equipped with an analytical cyano column and MicroTOF electrospray mass spectrometer. The results of LC-MS analysis are shown in Figure S1-S3 in SI. The molar mass of the monodispersed polypeptoids are given in Table 1.

Self-assembly

To form the nanosheets, the polypeptoids were dissolved in tetrahydrofuran (THF) /water (1:1, v/v) at a concentration of 2 mg/mL. The solutions were made in glass vials. The caps of the vials were left loose to

enable evaporation of THF. We use three protocols to evaporate THF from the solutions as shown in Table 1. In the first protocol, THF was removed by fast evaporation under a vacuum of 500 Torr at room temperature, over a period of 24 hours. We refer to this as the fast evaporation process. In the second protocol, THF was removed by placing the vial in a fume hood at room temperature, over a period of 72 hours. We refer to this as the slow evaporation process. In the third protocol, THF was removed by placing the vial in a refrigerator at 4 ^{o}C, over a period of 72 hours. We refer to this as the very slow evaporation process. All solutions turned turbid after THF removal, indicating the formation of nanosheets as described in our previous study.[18] All solutions were equilibrated at room temperature for extended period of time ranging from weeks to months prior to EM imaging. The extent to which molecular self-assembly is affected by processing details such as evaporative cooling, sample environment (vacuum oven at room temperature versus refrigerator at 4 ^{o}C) is outside the scope of present study.

Electron microscopy data collection

The atomic scale low-dose cryo-EM results reported in this paper are based on vitrified, hydrated nanosheets. These specimens were prepared on C-Flat grids (Protochips Inc.) which were covered by a thin, continuous carbon film (3 nm). A 3 µl droplet of the desired nanosheet-containing aqueous solution was transferred to the grid and blotted using two filter

papers. The grid was then plunged into liquid ethane to obtain vitrified hydrated specimen using Vitrobot (FEI Inc.). The specimens were imaged with a Titan Krios (FEI Inc.) operated at 300 kV with a K2 Summit direct detection camera and post-column energy filter (slit width at 25 eV) (Gatan Inc.) using defocus values from -1.0 to -0.2 µm. Dose-fractionation movies comprised 20 frames. The accumulated dose for each movie was about 20 *e*/Å^2. Movie frames were aligned and summed by MotionCorr2 to eliminate beam-induced motion.[48] Contrast transfer function (CTF) estimations were carried out using gCTF.[49] CTF correction was applied to all micrographs analyzed in this study.

Low magnification images of nanosheets were obtained from dry specimens. These specimens were prepared by depositing a 3 µl droplet of the desired nanosheet-containing aqueous solution on a continuous carbon film that was supported on a copper grid. The droplet was blotted from the edge of the grid using a filter paper. The grid was transferred to a TEM cryo holder (914 Gatan Inc.) and micrographs were collected on a Philips CM200 at 200 kV using a Gatan US1000 CCD camera at liquid nitrogen temperature in low-dose mode to minimize radiation damage. We expect the nanosheets to be completely dry due to exposure to vacuum, first in the airlock and then in the column of the TEM.

Image processing

Images of 2D crystals are generally not perfect due to dislocations, distortions from stress and image distortion within the microscope, which cause high resolution diffraction spots to be smeared out. In order to recover the high spatial frequency signal, a crystal unbending process was conducted on all micrographs. The motion-corrected and summed low-dose micrographs were imported into 2dx, an image processing package for 2D electron crystallography.[50] Details of the principle of unbent processing can be found elsewhere.[32] Briefly, the position of each unit cell in the image is found by cross correlation with a small reference area, and a smooth function is defined for displacements from the ideal lattice. This function is then used to re-interpolate the image onto the regular lattice. The defocus values, astigmatism, and specimen tilt geometry were determined using the gCTF program and corrected after unbending.

Overlapping small square boxes (90 Å long on each side), were extracted from the micrographs. The centers of the boxes coincide with the centers of the locations of unit cells as described in the previous work.[18] The extracted boxes were sorted into image classes using the Relion software package.[51] Relion uses the maximum-likelihood approach for sorting the small boxes extracted from cryo-EM micrographs.[51, 52] This approach has been proven to be particularly useful in the classification of structurally heterogeneous data from biomolecules. The intensity in the boxes is first normalized. Reference-free class averages are obtained in a completely unsupervised manner by starting multiple references from average images

of random sets of the normalized images in the extracted boxes. All images are compared to all references in all possible orientations and probability weights are calculated for each possibility instead of assigning images to one particular class or orientation. Class averages are then calculated as weighted averages over all possible assignments. The number of classes requested is set by the user: fewer boxes participate in each class with increasing number of classes. Averaging over a small number of boxes leads to noisy averages, which result in suboptimal alignment and classification. The number of classes requested can be set within a wide range, and we found that 12 classes gave the most informative results. To display the structural changes more clearly, in this study, the averaged images in 12 classes comprising similar or nearly identical motifs were separated into 3 groups.

The electron dense areas are dark in the bright field cryo-EM micrographs recorded by camera. After electron crystallography and classification processing in 2DX and Relion software packages, the contrast of electron dense region is inverted in order to produce the electron density maps (bright areas represent electron dense regions)

RESULTS AND DISCUSSION

The polypeptoids diblock copolymers used in this study are listed in Figure 1A. They comprise a hydrophobic poly-N-decylglycine (pNdc) block with decyl side chains and a hydrophilic poly-N-2-(2-(2-methoxyethoxy)

ethylglycine) (pNte) block with ethyleneoxy side chains. The number of repeating units in each block is 9. Three different end groups were used in this study: $COCH_3$, $COCH_2Cl$ and $COCH_2I$ and we refer to the polypeptoids as $pNdc_9$-$pNte_9$, Cl-$pNdc_9$-$pNte_9$ and I-$pNdc_9$-$pNte_9$. These polypeptoid diblock copolymers form nanosheets in aqueous solutions, similar to those reported in our previous study.[18] The pNdc blocks crystallize and form the core of the nanosheets. The pNte blocks dissolve in water and stabilize the nanosheet. In the crystalline state, decyl side chains are tilted relative to the backbone due to angle between the N-C and C-C bonds (see Figure 1A). Figure 1B shows a schematic of the structure of a $pNdc_9$-$pNte_9$ crystalline nanosheet containing interdigitated chains. The hydrophilic $pNte_9$ chains associated with adjacent crystallied $pNte_9$ chains emanate from opposite side of the crystal and they are not shown in this schematic for clarity. Only the crystallized pNdc blocks are shown with V-shaped motifs along in-plane (*b*-*c* plane) and through-plane (*a*-*c* plane) directions. The nanosheets in cryo-EM micrographs were obtained from the projection view of nanosheets through plane (or *a*-*c* plane). A typical chain conformation of $pNdc_9$ in the crystalline nanosheet obtained by MD simulations is shown in Figure 1C.[18] This conformation shows the V-shaped motif. The crystal structure of $pNdc_9$-$pNte_9$ nanosheets contains a glide plane, and it belongs to the plane group $P12_1$ with unit cell of a=4.7 Å, b=50 Å.[18]

The $pNdc_9$-$pNte_9$ molecules are first dissolved in a THF/water mixture and the THF is evaporated at different rates as described in the Experimental

Section. Since the molecules self-assemble in a form of nanosheets during THF evaporation, the kinetics of self-assembly are controlled by evaporation rate. We use three different evaporation processes labeled fast, slow and very slow as shown in Table 1. Low magnification images of the $pNdc_9$-$pNte_9$ nanosheets obtained by the three processes are shown in Figure 2.

Figure 2A shows the morphology of the $pNdc_9$-$pNte_9$ nanosheets obtained by fast evaporation. These nanosheets have irregular shapes and are relatively small (1 to 2 μm). A magnified view of one of the nanosheets and its Fourier transform (FFT) are shown in Figure 2B. The FFT shows a reflection corresponding to 25 Å which is the spacing between adjacent rows of pNdc backbones. This gives the orientation along the *c* direction of the rectangular unit cell.[18, 22] The reflection is in the form of an arc indicating that the crystal is imperfect. Figure 2C shows the morphology of the $pNdc_9$-$pNte_9$ nanosheets obtained by slow evaporation. These nanosheets are much larger than those obtained by fast evaporation. The nanosheet shown in Figure 2C is 2.5 μm wide and 6 μm long. A magnified view of the corner of this sheet and its Fourier transform (FFT) are shown in Figure 2D. The FFT obtained from the nanosheet prepared by slow evaporation shows a strong reflection corresponding to the spacing between adjacent rows of pNdc backbones. This reflection, which gives the orientation along the *c* direction of the rectangular unit cell,[14] is much stronger than that obtained from the nanosheet prepared by fast evaporation (compare Figures 2D and 2B). This suggests that better long-range order is obtained by slow evaporation.

However, the order obtained in the nanosheet prepared by slow evaporation is not perfect. The Bragg reflection contains a radial streak indicating the presence of stacking faults in the *c* direction. Figure 2E shows the morphology of the $pNdc_9$-$pNte_9$ nanosheets obtained by very slow evaporation. These nanosheets are similar to those obtained by slow evaporation. In this case, the Bragg reflections seen in the FFT appear as spots (Figure 2F) indicating excellent long-range order in the *c* direction. Since we know that the unit cell is rectangular, we assume that the *a* direction of the unit cell is in the plane of the specimen and perpendicular to the *c* direction in all samples.

Micrographs obtained at low resolution (Figure 2) reveal the morphological changes of nanosheets but not structural changes at atomic length scales. A section of a typical high resolution low-dose cryo-EM micrograph obtained from the vitrified hydrated $pNdc_9$-$pNte_9$ nanosheet obtained by fast evaporation is shown in Figure 3A. The dark vertical lines represent the electron dense backbones of the pNdc blocks, but the signal-to-noise ratio is so low that no other structural details are evident. The FFT of the micrograph in Figure 3A that shows the presence of strong reflections along the 4.7 Å rows (*a* spacing) is shown in Figure 3B. The FFT was filtered by applying a mask with 1 pixel diameter at each reflection and the FFT filtered image was generated by inverting the masked FFT. The Fourier filtered image, shown in the Figure 3C, contains structural information (phase and amplitude) averaged over the whole micrograph. The contrast of

electron dense regions is inverted in the FFT filtered image as compared to that in the cryo-EM micrographs. The bright areas represent electron dense regions in the nanosheet. Although the filtered image (Figure 3C) contains many features that were not evident in the micrograph (Figure 3A), interpretation of these features in terms of atomic scale structure is not possible. The presence of highly symmetrical features in Figure 3C is attributed to artifacts produced by averaging different crystalline motifs in the nanosheet.[18] The reason for these artifacts are discussed below.

The analysis described above was repeated for $pNdc_9$-$pNte_9$ nanosheets obtained by slow evaporation, and the results are shown in Figures 3D through F. The *a* spacing in these nanosheets is 4.5 Å, slightly smaller than that obtained by fast evaporation. The filtered image also contains highly symmetrical features that cannot be interpreted in terms of atomic scale structure. Figures 3G through I show the results of repeating this analysis for $pNdc_9$-$pNte_9$ nanosheets obtained by very slow evaporation. The *a* spacing obtained in this case is 4.5 Å. Interestingly, the filtered image (Figure 3I) in this case shows V-shaped crystalline motifs that are distinctly different from the other filtered images in Figure 3 (Figures 3C and 3F). Even in the case of very slow evaporation, the filtered image contains artifacts produced by averaging different crystalline motifs in the nanosheet.

In order to resolve the atomic scale structure in the nanosheets, we use the approach demonstrated in our previous work.[18] The parameters used to obtain the low-dose cryo-EM micrographs are shown in Table S1. The

locations of all unit cells in the micrograph were determined using the unbending step as described in the Experiment Section. We then extracted 90 x 90 Å boxes from the micrographs containing a pre-specified number of unit cells. These boxes were sorted into 12 classes and averaged using the Relion software package. The set of averaged images in each class is based on 10,000 – 100,000 nearly identical boxes.

The 12 classes obtained from the nanosheets prepared by fast evaporation are shown in Figure 4A. We divide the 12 classes into two groups that we label group 1 and group 3. (Group 2 will be defined shortly.) In group 1, we see vertically aligned bright spots running along the apex of a V-shaped structure with arms extending symmetrically out on both sides of the spots. Adjacent V-shaped structures are anti-parallel. The spacing between the rows of bright spots is consistently 25 Å and the distance between adjacent bright spots is 4.7 Å. In group 3, we see complex geometries. It seems unlikely they represent single crystals.

The 12 classes obtained from the nanosheets prepared by slow evaporation are shown in Figure 4B. In this case, we also see that the 12 classes can be divided into two groups. As was the case in the nanosheets prepared by fast evaporation, group 1 contains anti-parallel V-shaped motifs. The spacing between the rows of bright spots is 25 Å and the distance between adjacent bright spots is 4.5 Å. Group 3 motifs appear highly disordered.

The 12 classes obtained from the nanosheets prepared by very slow evaporation are shown in Figure 4C. We divide the classes into 3 groups. Group 1 contains anti-parallel V-shaped motifs. In addition to this, we observe new motifs with vertically aligned bright spots running along the apex of a V-shaped structure. However, adjacent V-shaped structures are parallel. Group 2 comprise these motifs. The spacing between the rows of bright spots is 25 Å and the distance between adjacent bright spots is 4.5 Å. Group 3 motifs have poorly defined features. The FFTs of motifs in $pNdc_9$-$pNte_9$ nanosheets obtained by difference processes are shown in Figure S5.

If all of the motifs in a given sample are averaged the resulting images would have high symmetry artifacts similar to those seen in Figures 3C, 3F and 3I.

A minimum of 70,000 boxes were obtained from each sample; the exact numbers of boxes analyzed for each sample are given in Table 1. We focus on groups 1 and 2 which show clearly defined motifs. The averaged images corresponding to anti-parallel V-shaped (group 1) and parallel V-shaped (group 2) motifs are shown in Figures 5A and B. These particular averaged images in these figures were extracted from the nanosheets prepared by very slow evaporation. As seen in Figure 5A, the pNdc chains are arranged in rows along the *a* direction. A schematic of the crystalline nanosheet is shown in Figure 5B. Within each row, we see bright spots that represent the end view of the glycine backbones in the *cis* conformation as illustrated in Figure 1A and 1C. The V-shaped arms emanating from these

spots in the averaged image (Figure 5A), correspond to the locations of the side chains. The side chains in adjacent rows are anti-parallel. A thin vertical dark band between adjacent Vs represents low electron density gap between adjacent chains. As expected, there is also a thin dark band between adjacent pNdc side chains. The side chains from one row appear to point towards the gap in the adjacent row. This is shown schematically in Figure 5B. The doted box in Figure 5A shows the unit cell of the anti-parallel V-shaped motif. The FFT of the averaged image in Figure 5A, shown in Figure 5C, contains the expected reflections at 4.5 Å along the meridians (*a* direction) and at 25 Å along the equator (*c* direction) as indicated by a dashed circle. Several additional reflections are evident up to 2.1 Å along the equator. The highest resolution of Thon rings is 1.9 Å indicating that the phase information in the averaged image is reliable after CTF correction (see the Table S1 in SI).

The averaged image of the parallel V-shaped motif is shown in Figure 5D. Figure 5D is similar to Figure 5A except for the fact that adjacent rows of pNdc chains are in the parallel arrangement. A schematic of the parallel V-shaped crystalline nanosheet is shown in Figure 5E. Note that adjacent rows of pNdc chains are displaced relative to each other along the *a* direction. The unit cell of the parallel V-shaped motif thus contains two chains along the *c* direction. The unit cell size of both parallel and anti-parallel V-shaped motifs is identical; see doted boxes in Figures 5A and 5D. The FFT of the averaged image in Figure 5D, shown in Figure 5F, is similar to Figure 5C. Subtle

differences in the intensity and spatial frequency of the reflections within the rows at 4.5 Å along equator in Figures 5C and F give rise to the parallel and anti-parallel V-shaped motifs.

Figure 6A shows a distribution map of the different motifs in a $pNdc_9$-$pNte_9$ nanosheet obtained by fast evaporation. Only a small portion of nanosheet is shown in Figure 6A for purpose of clarity. The distribution map for the entire nanosheet is shown in Figure S9 in SI. The rectangles in distribution maps represent positions of unit cells. The local motif in each box is indicated by the color of the box: the parallel V-shaped motifs are colored green, the anti-parallel V-shaped motifs are colored blue, the disordered motifs are colored yellow, and boxes that couldn't be analyzed are left blank. Contiguous collections of boxes of the same color can be viewed as coherently ordered grains of a given motif. Figure 6A shows that fast evaporation results in relatively small grains with anti-parallel V-shaped motifs that coexist with disordered grains. The distribution map of the different motifs in a $pNdc_9$-$pNte_9$ nanosheet obtained by slow evaporation is shown in Figure 6B. This figure is similar to Figure 6A except for the fact that the grains with anti-parallel V-shaped motifs are larger. For reasons that not clear, the number of boxes could not be analyzed was higher in the slow evaporation case. The distribution map of the different motifs in a $pNdc_9$-$pNte_9$ nanosheet obtained by very slow evaporation is shown in Figure 6C. Here the map is dominated by very large grains with the parallel V-shaped motif.

The coherent order in the nanosheets described in the preceding paragraph was quantified by calculating the distribution of concatenated unit cells of given type along the specified direction. In Figure 7A, we show the distribution of motifs obtained along the *a* direction in a $pNdc_9$-$pNte_9$ nanosheet obtained by fast evaporation in the form of histogram. This histogram applies to concatenated unit cells with the anti-parallel V-shaped motif. The x-axis in Figure 7A indicates the number of concatenated unit cells along the *a* direction (N_a) while the y-axis indicates the number of times a particular number of concatenated unit cells (P_a) was observed in the nanosheet. The data in Figure 7A can be considered as an ordered pair ($P_{a,j}$, $N_{a,j}$) where j goes from 1 to 44 ($N_{a,max}$ is 44). Figure 7B, where we show the distribution of motifs obtained along the *c* direction in a $pNdc_9$-$pNte_9$ nanosheet obtained by fast evaporation, is analogous to Figure 7A. One can define a mean coherence length as

$$\acute{N}_i = \frac{\sum_{j=1}^{N_{i,max}} P_{i,j} N_{i,j}}{\sum_{j=1}^{N_{i,max}} P_{i,j}} \qquad (i=a \text{ or } c) \qquad \text{(eq 1)}$$

where $N_{i,max}$ is the largest value of N_i where P_i is non zero.

The values of $\bar{N}_a$ and $\bar{N}_c$ are 4.3 and 2.1, respectively in Figures 7A and 7B. Note that $\bar{N}_i$ represents coherence length in units of unit cell dimensions along *a* and *c* directions. It is evident that the coherence length in these units along the *a* direction is larger than that along *c* direction. Note however the unit cell is a 4.5 x 50 Å rectangle, implying that the mean physical

coherence lengths in the *a* and *c* directions are 19 and 105 Å, respectively. In other words, the physical coherence length along the *c* direction is larger than that along *a* direction. In this paper, we mainly discuss $\bar{N}_i$ in units of unit cell dimensions.

In Figure 7A, we see that P_a is a monotonically decreasing function of N_a in the range $1 \leq N_a \leq 30$. In the range $5 \leq N_a \leq 30$, P_a is, to a good approximation, an exponentially decaying function. More generally, we write,

$$P_i = P_{i0}\exp(-N_i/L_i) \quad (i=a \text{ or } c) \quad \text{(eq 2)}$$

where L_i is another measure of the coherence length along the *a* or *c* directions in units of the unit cell dimensions along the particular direction. The straight line in Figure 7A is the fit of eq 2 through the data in the range of $5 \leq N_a \leq 30$. This fit gives $L_a = 5.8$. In the range $1 \leq N_a \leq 5$, the decrease P_a and N_a is more rapid than that predicted by eq 2 with $L_a = 5.8$. In the range $30 \leq N_a \leq 50$, we have poor statistics due to the finite size of the nanosheet. The parameter L_a does not reflect data outside the range $5 \leq N_a \leq 30$. It should be clear from this discussion that long-range order between unit cells is governed by a complex correlation function. For simplicity, we characterize this function by two parameters, $\bar{N}_a$ and L_a. For the data in Figure 7A, $\bar{N}_a$ and L_a are similar, 4.3 and 5.8, respectively. From analysis of the data in Figure 7B we conclude that $\bar{N}_c$ and L_c are 2.1 and 1.5. The values of $\bar{N}_i$ and L_i are similar in all cases, and they are presented in Tables 2 and 3.

In Figures 7C and D, we show the distributions of motifs obtained along the *a* and *c* directions in a pNdc$_9$-pNte$_9$ nanosheet obtained by slow evaporation. Slowing down evaporation rate increases both $\bar{N}_a$ and $\bar{N}_c$ (see Table 2). Not surprisingly, slowing down evaporation rate also increases both L_a and L_c (see Table 3). In Figures 7E and F, we show the distributions of motifs obtained along the *a* and *c* directions in a pNdc$_9$-pNte$_9$ nanosheet obtained by very slow evaporation. In this case, L_a and L_c for the anti-parallel V-shaped motifs are very small. This is because that the nanosheet is dominated by parallel V-shaped motifs. The distribution of this motif along the *a* direction is shown in Figure 7G. The distribution function is extremely broad; we have substantial number of rows with 400 concatenated unit cells along the *a* direction ($L_{a,max}$). Due to this, we have divided the data into two sets in Figure 7G. The distribution of the parallel V-shaped motif along the *c* direction is shown in Figure 7H. We are unable to estimate L_a and L_c because the data do not follow eq (2). It is evident that very slow evaporation results in highly coherent crystal structure. The values of $\bar{N}_a$ and $\bar{N}_c$ are large, 54.1 and 19.2, respectively, as given in Table 2.

The analysis described above was repeated for the nanosheets obtained from Cl-pNdc$_9$-pNte$_9$ and I-pNdc$_9$-pNte$_9$. To facilitate comparisons with the discussion above, these nanosheets were prepared by very slow evaporation. The low-resolution images of nanosheets obtained from these systems, shown in Figures 8A and B, are similar to the nanosheets obtained

from pNdc$_9$-pNte$_9$ prepared by very slow evaporation. The boxes from these systems belong to the same 3 groups identified in Figure 3. The results (cryo-EM micrographs, sorting and averaging) are given in Figures S6 through S8 in SI. Figure 9 shows distribution maps of nanosheets of Cl-pNdc$_9$-pNte$_9$ and I-pNdc$_9$-pNte$_9$ obtained by very slow evaporation. Both distribution maps are dominated by parallel V-shaped motif. A few anti-parallel V-shaped motifs are seen in the Cl-pNdc$_9$-pNte$_9$ sample (Figure 9A) but not in the I-pNdc$_9$-pNte$_9$ sample (Figure 9B). In Figure 10, we show the distributions of motifs (P_i versus N_i) obtained along the *a* and *c* directions in Cl-pNdc$_9$-pNte$_9$ and I-pNdc$_9$-pNte$_9$ nanosheets. The correlation functions of the anti-parallel V-shaped motif in Cl-pNdc$_9$-pNte$_9$ are characterized by L_a =6.0 and L_c =1.1 (Table 3). Correlations of the parallel V-shaped motifs along the *a* direction shown in Figures 10C and 10E are broad, similar to Figure 7G. Correlations of the parallel V-shaped motifs along the *c* direction are shown in Figures 10D and 10F. For the Cl-pNdc$_9$-pNte$_9$ nanosheet, we obtained L_c=3.8 for this motif. The correlation function for I-pNdc$_9$-pNte$_9$ nanosheet is complex and does not decay over accessible length scales.

The coherence lengths obtained from the analysis described above are summarized in Tables 2 and 3. It is evident that the nanosheets prepared by slow evaporation exhibit more ordered arrangement as compared to the nanosheets prepared by fast evaporation in terms of the appearance of longer coherence lengths along *a* and *c* direction. The presence of the dominant parallel V-shaped motifs in the pNdc$_9$-pNte$_9$ nanosheet prepared by

very slow evaporation suggests that processing condition has a significant effect on the crystal motifs. In addition, Na bar is 54.1, which is about 4 times larger than that of the anti-parallel V-shaped motifs in the $pNdc_9$-$pNte_9$ nanosheet prepared by slow evaporation (see Table 2). The Cl-$pNdc_9$-$pNte_9$ nanosheet shows the coexistence of anti-parallel V-shaped and parallel V-shaped motifs. The Ni bar values for the parallel V-shaped motifs (13.9 along *a* direction and 3.5 along *c* direction) are obviously larger than those for the anti-parallel V-shaped motifs (5.9 along *a* direction and 1.5 along *c* direction). Interestingly, the I-$pNdc_9$-$pNte_9$ nanosheet show homogeneity with the absence of anti-parallel motifs.

We conclude this section by discussing the composition of the nanosheets obtained from the 5 samples listed in Table 1. The total number of boxes examined for the samples are also given in Table 1. We define f to be the fraction of boxes belonging to a particular group (f_a, anti-parallel, f_p, parallel and f_d, disordered) in the sample of interest. Figure 11A shows the effect of the evaporation rate on the crystal structure of $pNdc_9$-$pNte_9$ nanosheets. In nanosheets prepared by fast evaporation, $f_a = 0.47$, $f_p = 0$ and $f_d = 0.53$. In nanosheets prepared by slow evaporation, f_a increases to 0.65 while f_d decreases to 0.35 (f_p is still zero). It is evident that slow processing increases the fraction of ordered boxes in the nanosheets. In the nanosheets prepared by very slow evaporation, $f_p = 0.78$ while both f_a and f_d are less than 0.2. Figure 11B shows the effect of the end groups on the crystal structure of $pNdc_9$-$pNte_9$ nanosheets prepared by very slow evaporation. Ordered motifs

dominate all samples; f_d is less than 0.2 in all cases. $pNdc_9$-$pNte_9$ and I-$pNdc_9$-$pNte_9$ nanosheets have f_p values greater than 0.75 while Cl-$pNdc_9$-$pNte_9$ nanosheets exhibit both parallel and anti-parallel motifs (f_p = 0.55 and f_a = 0.38). Figure 11 shows that subtle changes in processing conditions and end group chemistry can have a large effect on crystal structure. Under very slow evaporation conditions, parallel motifs are dominant irrespective of end group chemistry (Figure 11B). This suggests that the equilibrium morphology of all three polypeptoids is the parallel motif. However, subtle changes in processing and end group chemistry lead to the formation of significant fractions of anti-parallel motifs. This suggests the free energies of formation of parallel and anti-parallel motifs are similar. In ref. 18, we attempted to quantify the energy difference between parallel and anti-parallel motifs using MD simulations. We found that this difference was within the error of the simulations. Further work is needed to determine the underpinnings of the observed crystalline motifs in polypeptoids.

CONCLUSIONS

In conclusion, we have studied the effect of the evaporation rate and end group chemistry on self-assembly of $pNdc_9$-$pNte_9$ nanosheets. Crystal structure within the nanosheets was determined using cryo-EM at atomic length scales. By slowing down the evaporation rate, we were able to obtain nanosheets with excellent long-range order. Two dominant motifs were

identified: the anti-parallel V-shaped motif wherein adjacent rows of $pNdc_9$ polypeptoid molecules adopt anti-parallel side chain conformations, and the parallel V-shaped motif wherein adjacent rows of $pNdc_9$ polypeptoid molecules adopt parallel side chain conformations. In both cases, the unit cells are rectangular with a = 4.5 Å and c = 50 Å. Well-ordered motifs were maintained when the acetyl end groups were replaced by chemically-reactive haloacetyl groups, indicating the robustness of the polypeptoid lattice structure and opening up the possibility of functionalizing the nanosheet surface without disruption of the crystalline hydrophobic core. Correlations between neighboring unit cells, determined from distribution maps, differ substantially in the a and c directions. Long-range order, measured in terms of the number of identical unit cells, along the a direction was more pronounced than that along the c direction. As many as 400 contiguous unit cells with the same motif were observed in nanosheets obtained by very slow evaporation.

SUPPLEMENTARY MATERIALS: LC-MS traces of solutions, TEM micrographs, parameters of analyzed cryo-EM micrographs, crystal Motifs, FFTs of motifs and distribution maps of motifs in the nanosheets.

CONFLICTS OF INTEREST

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

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 was supported by the Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. Micrographs presented here were obtained at the Donner Cryo-EM facility in Lawrence Berkeley National Laboratory and the Berkeley Bay Area Cryo-EM facility in University of California Berkeley.

REFERENCES

1. J. Sun and R. N. Zuckermann, *ACS nano*, 2013, **7**, 4715-4732.
2. J. A. Crapster, I. A. Guzei and H. E. Blackwell, *Angew Chem Int Edit*, 2013, **52**, 5079-5084.
3. J. R. Stringer, J. A. Crapster, I. A. Guzei and H. E. Blackwell, *J Am Chem Soc*, 2011, **133**, 15559-15567.
4. C. W. Wu, T. J. Sanborn, K. Huang, R. N. Zuckermann and A. E. Barron, *J Am Chem Soc*, 2001, **123**, 6778-6784.
5. H. B. Jin, Y. H. Ding, M. M. Wang, Y. Song, Z. H. Liao, C. J. Newcomb, X. P. Wu, X. Q. Tang, Z. Li, Y. H. Lin, F. Yan, T. Y. Jian, P. Mu and C. L. Chen, *Nat Commun*, 2018, **9**.
6. J. Sun, X. Jiang, R. Lund, K. H. Downing, N. P. Balsara and R. N. Zuckermann, *P Natl Acad Sci USA*, 2016, **113**, 3954-3959.
7. H. B. Jin, F. Jiao, M. D. Daily, Y. L. Chen, F. Yan, Y. H. Ding, X. Zhang, E. J. Robertson, M. D. Baer and C. L. Chen, *Nat Commun*, 2016, **7**.
8. E. J. Robertson, G. K. Olivier, M. L. Qian, C. Proulx, R. N. Zuckermann and G. L. Richmond, *P Natl Acad Sci USA*, 2015, **112**, E338-E338.
9. G. K. Olivier, A. Cho, B. Sanii, M. D. Connolly, H. Tran and R. N. Zuckermann, *ACS nano*, 2013, **7**, 9276-9286.
10. B. Sanii, R. Kudirka, A. Cho, N. Venkateswaran, G. K. Olivier, A. M. Olson, H. Tran, R. M. Harada, L. Tan and R. N. Zuckermann, *Biophys J*, 2012, **102**, 269a-269a.
11. R. Kudirka, H. Tran, B. Sanii, K. T. Nam, P. H. Choi, N. Venkateswaran, R. Chen, S. Whitelam and R. N. Zuckermann, *Biopolymers*, 2011, **96**, 586-595.
12. K. T. Nam, S. A. Shelby, P. H. Choi, A. B. Marciel, R. Chen, L. Tan, T. K. Chu, R. A. Mesch, B. C. Lee, M. D. Connolly, C. Kisielowski and R. N. Zuckermann, *Nat Mater*, 2010, **9**, 454-460.
13. X. Ma, S. Zhang, F. Jiao, C. J. Newcomb, Y. L. Zhang, A. Prakash, Z. H. Liao, M. D. Baer, C. J. Mundy, J. Pfaendtner, A. Noy, C. L. Chen and J. J. De Yoreo, *Nat Mater*, 2017, **16**, 767-+.
14. J. M. V. Jun, M. V. P. Altoe, S. Aloni and R. N. Zuckermann, *Chem Commun*, 2015, **51**, 10218-10221.
15. N. A. Merrill, F. Yan, H. B. Jin, P. Mu, C. L. Chen and M. R. Knecht, *Nanoscale*, 2018, **10**, 12445-12452.
16. R. N. Zuckermann, J. M. Kerr, S. B. H. Kent and W. H. Moos, *J Am Chem Soc*, 1992, **114**, 10646-10647.
17. C. Proulx, S. Yoo, M. D. Connolly and R. N. Zuckermann, *J Org Chem*, 2015, **80**, 10490-10497.
18. X. Jiang, D. R. Greer, J. Kundu, C. Ophus, A. M. Minor, D. Prendergast, R. N. Zuckermann, N. P. Balsara and K. H. Downing, *Macromolecules*, 2018, **51**, 7794-7799.
19. E. J. Robertson, A. Battigelli, C. Proulx, R. V. Mannige, T. K. Haxton, L. S. Yun, S. Whitelam and R. N. Zuckermann, *Accounts Chem Res*, 2016, **49**, 379-389.

20. J. Sun, A. A. Teran, X. X. Liao, N. P. Balsara and R. N. Zuckermann, *J Am Chem Soc*, 2014, **136**, 2070-2077.
21. H. Jin, F. Jiao, M. D. Daily, Y. Chen, F. Yan, Y. H. Ding, X. Zhang, E. J. Robertson, M. D. Baer and C. L. Chen, *Nat Commun*, 2016, **7**, 12252.
22. D. R. Greer, M. A. Stolberg, J. Kundu, R. K. Spencer, T. Pascal, D. Prendergast, N. P. Balsara and R. N. Zuckermann, *J Am Chem Soc*, 2018, **140**, 827-833.
23. D. R. Greer, M. A. Stolberg, S. T. Xuan, X. Jiang, N. P. Balsara and R. N. Zuckermann, *Macromolecules*, 2018, **51**, 9519-9525.
24. J. R. Edison, R. K. Spencer, G. L. Butterfoss, B. C. Hudson, A. I. Hochbaum, A. K. Paravastu, R. N. Zuckermann and S. Whitelam, *P Natl Acad Sci USA*, 2018, **115**, 5647-5651.
25. B. C. Hudson, A. Battigelli, M. D. Connolly, J. Edison, R. K. Spencer, S. Whitelam, R. N. Zuckermann and A. K. Paravastu, *J Phys Chem Lett*, 2018, **9**, 2574-2578.
26. R. M. Glaeser, G. McMullan, A. R. Faruqi and R. Henderson, *Ultramicroscopy*, 2011, **111**, 90-100.
27. P. Pradere and E. L. Thomas, *Ultramicroscopy*, 1990, **32**, 149-168.
28. D. L. Dorset and F. Zemlin, *Ultramicroscopy*, 1990, **33**, 227-236.
29. D. L. Dorset, *P Natl Acad Sci USA*, 1990, **87**, 8541-8544.
30. A. Uemura, M. Tsuji, A. Kawaguchi and K. I. Katayama, *J Mater Sci*, 1988, **23**, 1506-1509.
31. J. F. Revol and R. S. Manley, *J Mater Sci Lett*, 1986, **5**, 249-251.
32. R. Henderson, J. M. Baldwin, K. H. Downing, J. Lepault and F. Zemlin, *Ultramicroscopy*, 1986, **19**, 147-178.
33. B. Moss, D. L. Dorset, J. C. Wittmann and B. Lotz, *J Polym Sci Pol Phys*, 1984, **22**, 1919-1929.
34. S. Giorgio and R. Kern, *J Polym Sci Pol Phys*, 1984, **22**, 1931-1951.
35. M. Tsuji, S. Isoda, M. Ohara, A. Kawaguchi and K. Katayama, *Polymer*, 1982, **23**, 1568-1574.
36. R. Henderson and P. N. Unwin, *Nature*, 1975, **257**, 28-32.
37. P. H. Till, *J Polym Sci*, 1957, **24**, 301-306.
38. K. H. Storks, *J Am Chem Soc*, 1938, **60**, 1753-1761.
39. J. Frank, *Three-dimensional electron microscopy of macromolecular assemblies : visualization of biological molecules in their native state*, Oxford University Press, New York, 2nd edn., 2006.
40. R. M. Glaeser, *Electron crystallography of biological macromolecules*, Oxford University Press, Oxford ; New York, 2007.
41. J. Frank, *Single-particle cryo-electron microscopy : the path toward atomic resolution : selected papers of Joachim Frank with commentaries*, World Scientific, New Jersey, 2018.
42. R. M. Glaeser, *Annu Rev Phys Chem*, 1985, **36**, 243-275.
43. E. Nogales and W. Chiu, *Journal of structural biology*, 1999, **128**, 1-2.
44. P. Pradere and E. L. Thomas, *Philos Mag A*, 1989, **60**, 177-203.
45. N. Hamada, M. Tosaka, M. Tsuji, S. Kohjiya and K. Katayama, *Macromolecules*, 1997, **30**, 6888-6892.

46. M. Tosaka, N. Hamada, M. Tsuji and S. Kohjiya, *Macromolecules*, 1997, **30**, 6592-6596.
47. M. Tosaka, N. Hamada, M. Tsuji, S. Kohjiya, T. Ogawa, S. Isoda and T. Kobayashi, *Macromolecules*, 1997, **30**, 4132-4136.
48. S. Q. Zheng, E. Palovcak, J. P. Armache, K. A. Verba, Y. F. Cheng and D. A. Agard, *Nat Methods*, 2017, **14**, 331-332.
49. K. Zhang, *Journal of structural biology*, 2016, **193**, 1-12.
50. B. Gipson, X. Zeng, Z. Y. Zhang and H. Stahlberg, *Journal of structural biology*, 2007, **157**, 64-72.
51. S. H. Scheres, *Journal of structural biology*, 2012, **180**, 519-530.
52. S. H. Scheres, *J Mol Biol*, 2012, **415**, 406-418.

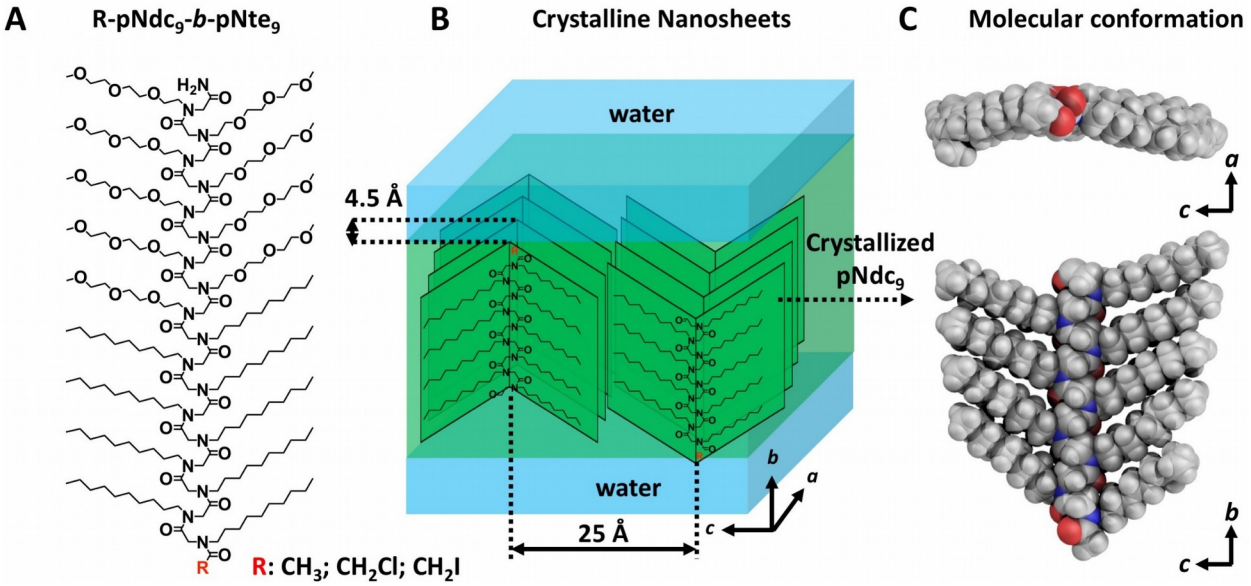

Figure 1. A. Chemical structures of $pNdc_9$-$pNte_9$ polypeptoid block copolymers with different end groups R used in this study. **B**. A schematic of

the $pNdc_9$-$pNte_9$ crystalline nanosheet. Side chains are shown in V-shaped conformation in both *a*-*c* plane and *b*-*c* plane. **C.** Atomic model of $pNdc_9$-$pNte_9$ obtained from MD simulations; only one crystallized $pNdc_9$ block is shown. Side chains adopt the V-shaped conformation in both *a*-*c* plane and *b*-*c* plane after 300 ns relaxation, similar to the V-shaped conformations shown in the schematic. Nitrogen atoms are blue, oxygen atoms are red and sidechains are grey.

Table 1. Characteristics of synthesized crystalline polpypeptoids, self-assembly conditions, and size of data sets.

Polypeptoids	Molar mass [1]	Self-assembly process	Number of boxes [5]
$pNdc_9$-$pNte_9$	3663.9	Fast [2]	166117
		Slow [3]	70902
		Very slow [4]	262826
Cl-$pNdc_9$-$pNte_9$	3698.7	Very slow [4]	346113
I-$pNdc_9$-$pNte_9$	3789.6	Vert slow [4]	56860

[1] Determined by LC-MS, as shown in Figure S1-S3; [2] Fast evaporation of THF at room temperature with 500 torr vacuum; [3] Slow evaporation of THF at room temperature; [4] Very slow evaporation of THF at 4 °C. [5] Number of boxes extracted from the micrographs for analysis.

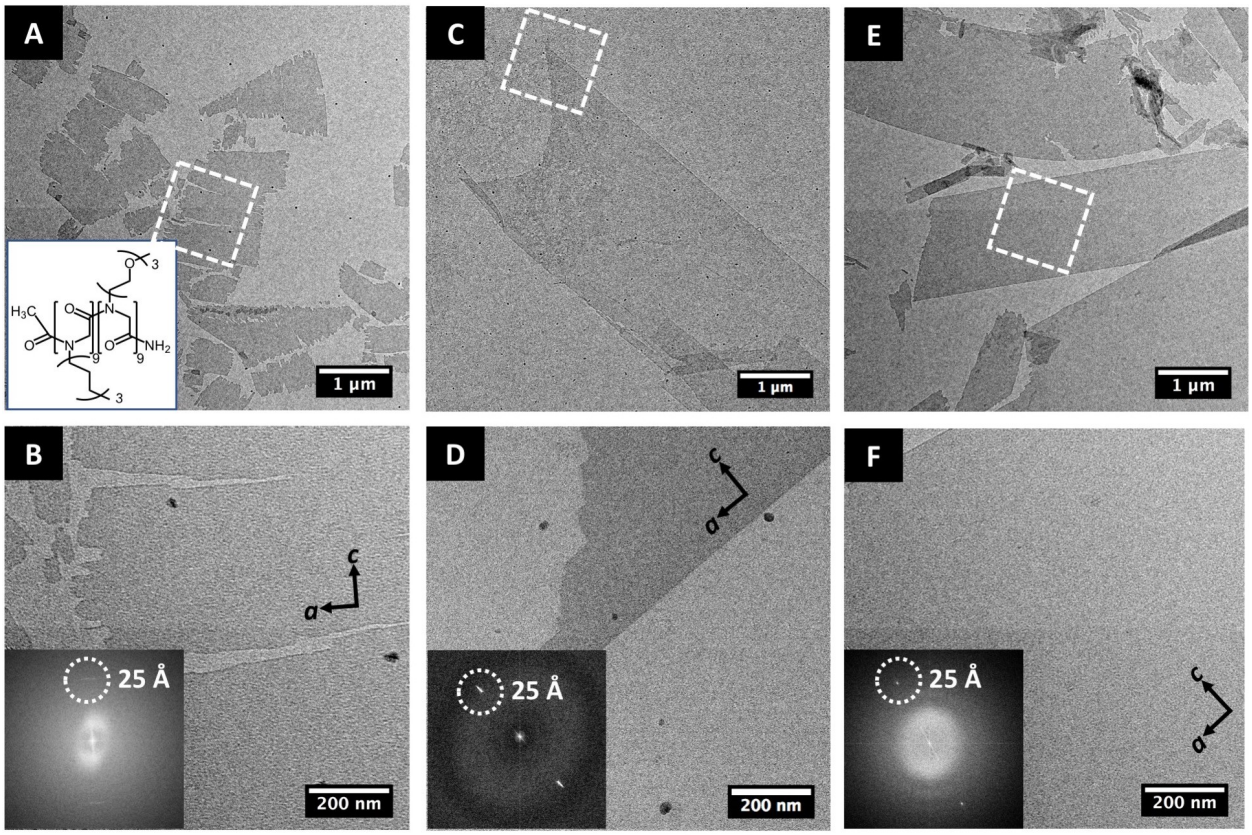

Figure 2. Morphology of the self-assembled $pNdc_9$–$pNte_9$ nanohseets prepared by different processes. **A.** TEM micrograph of $pNdc_9$–$pNte_9$ nanosheets obtained by fast evaporation. Inset shows the chemical structure of the polypepotid. **B.** Magnified micrograph of one of the sheets identified as the dashed box in A. Inset shows the corresponding Fourier transform (FFT). **C.** TEM micrograph of $pNdc_9$–$pNte_9$ nanosheets obtained by slow evaporation. **D.** Magnified micrograph of one of the sheets identified as the dashed box in C. Inset shows the corresponding FFT. **E.** TEM micrograph of $pNdc_9$–$pNte_9$ nanosheets obtained by very slow evaporation. **F.** Magnified micrograph of one of the sheets identified as the dashed box in E. Inset shows the corresponding FFT. The FFTs show reflections corresponding to

25Å which is the spacing between adjacent rows of pNdc backbones. This gives the orientation of the *c* axis in each nanosheet. The *a* axis is orthogonal to the *c* axis. The diffuse rings in the FFTs are due to Thon rings.

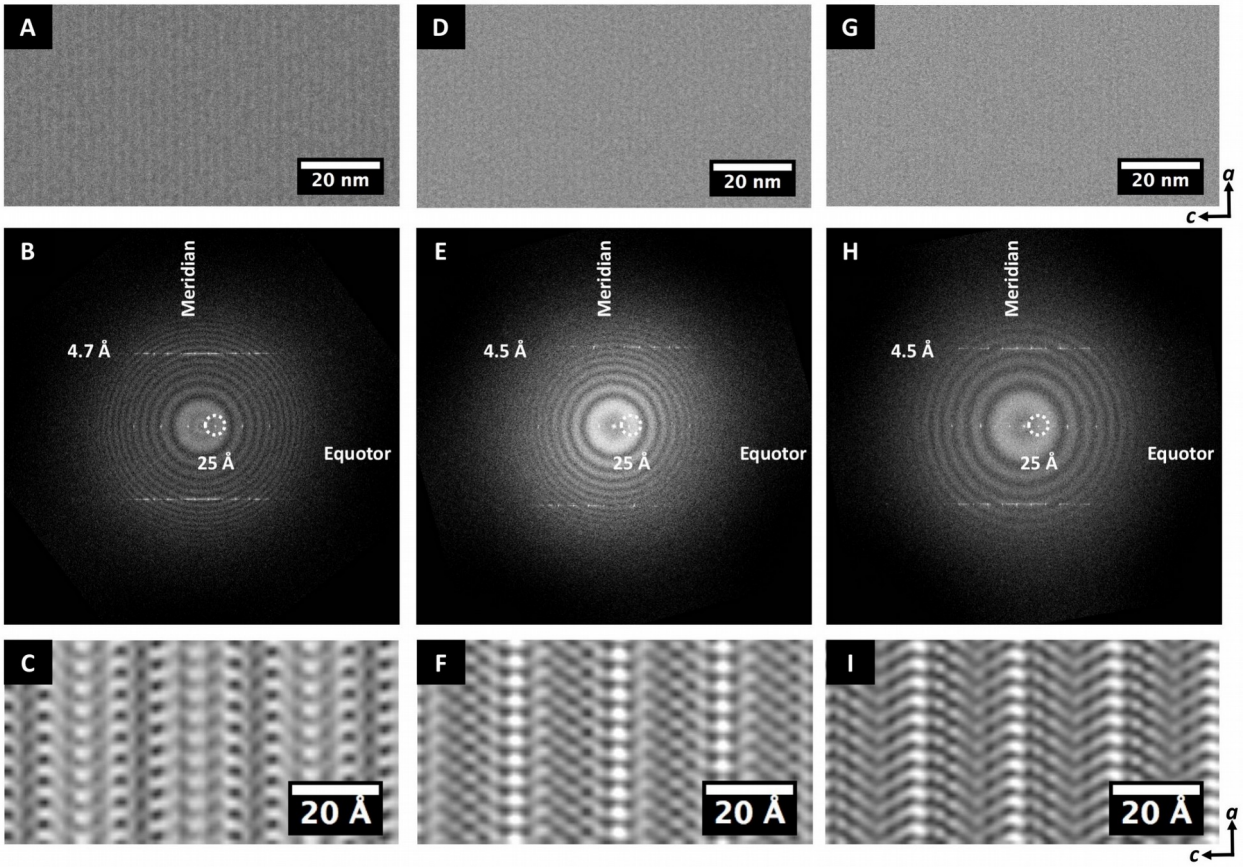

Figure 3. Structure of the self-assembled $pNdc_9$–$pNte_9$ nanosheets prepared by different processes. A. A section of a low-dose cryo-EM micrograph of a vitrified hydrated $pNdc_9$–$pNte_9$ nanosheet obtained by fast evaporation. **B.** The FFT of the low-dose cryo-EM micrograph in A showing reflections and Thon rings due to the thin carbon support. The dashed circle

indicates the reflection corresponding to 25 Å which is the spacing between adjacent rows of pNdc backbones (*c* direction). The reflections at 4.7 Å indicate the spacing between two adjacent pNdc backbones (*a* direction). **C.** A section of a Fourier filtered image of the FFT in B. Reflections in B are filtered by applying a mask with 1 pixel diameter at each reflection. **D** through **F**: Analogous to Figures A through C for a vitrified hydrated $pNdc_9$–$pNte_9$ nanosheet obtained by slow evaporation. The only difference is the reflections at 4.5 Å which indicate the spacing between two adjacent pNdc backbones (*a* direction). **G** through **I**: Analogous to Figures D through F for a vitrified hydrated $pNdc_9$–$pNte_9$ nanosheet obtained by very slow evaporation. Bright areas in C, F and J represent the electron dense regions. The high symmetry images in C, F, and J are artifacts due to averaging over distinct motifs.

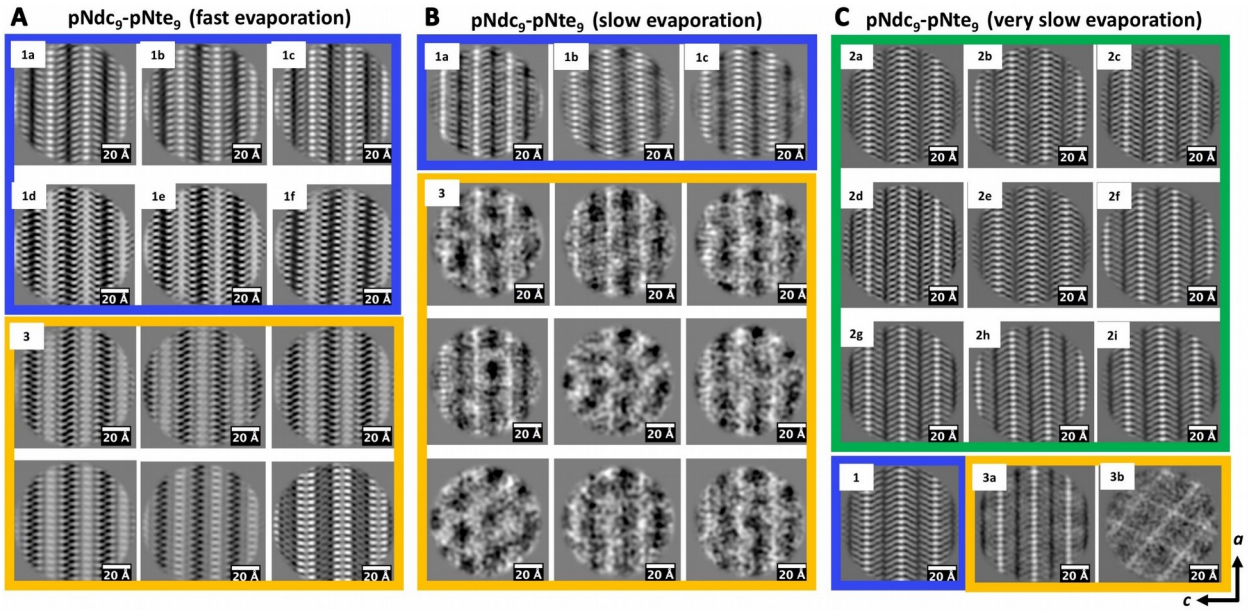

Figure 4. Motifs in the $pNdc_9$-$pNte_9$ nanosheets prepared by different processes. Boxes were sorted and averaged assuming the presence of 12 classes. These classes fall naturally into two distinct groups that contain ordered motifs: anti-parallel V-shaped motifs (group 1 identified by a blue outline), and parallel V-shaped motifs (group 2 identified by a green outline). Classes that were inconsistent with the molecule shown in Figure 1C are collectively referred to as disordered (group 3 identified by a yellow outline). Motifs of $pNdc_9$-$pNte_9$ nanosheets obtained by **A**: fast evaporation, **B**: slow evaporation, **C**: very slow evaporation. The *a* and *c* directions are specified for ordered motifs . Bright areas in all images represent the electron dense regions.

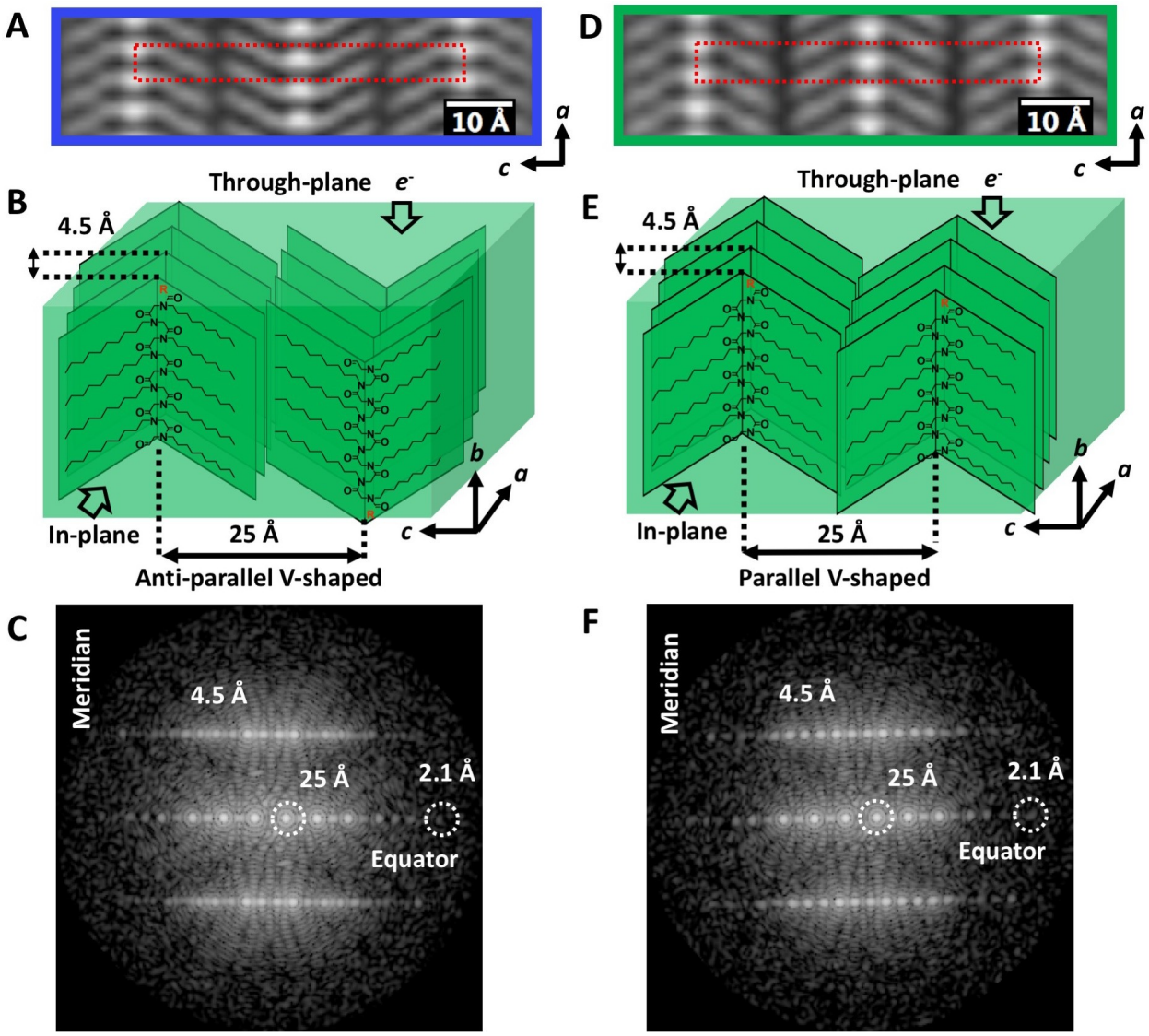

Figure 5. Atomic-scale images and models of V-shaped motifs. A. A section of an averaged image obtained from pNdc$_9$-pNte$_9$ nanosheets (very slow evaporation) showing the anti-parallel V-shaped motif. **B.** Schematic of the arrangement of anti-parallel pNdc blocks. Note the V-shaped arrangement of adjacent molecules within both *a*-*c* and *b*-*c* planes in the crystallized nanosheet. **C.** FFT of the averaged image in A with detectable reflections up to 2.1 Å along equator (*c* direction). **D.** A section of an averaged image obtained from pNdc$_9$-pNte$_9$ nanosheets (very slow

evaporation) showing the parallel V-shaped motif. **E.** Schematic of the arrangement of parallel pNdc blocks. **F.** FFT of the averaged image in A with detectable reflections up to 2.1 Å along equator (*c* direction). The red dashed boxes in A and D indicate identical rectangular unit cells with a = 4.5 Å and c = 50 Å. In schematics B and E, R indicates the positions of end groups. Arrows with electrons indicate the projection direction in the electron microscope (*a*-*c* plane). Bright areas in A and D represent the electron dense regions.

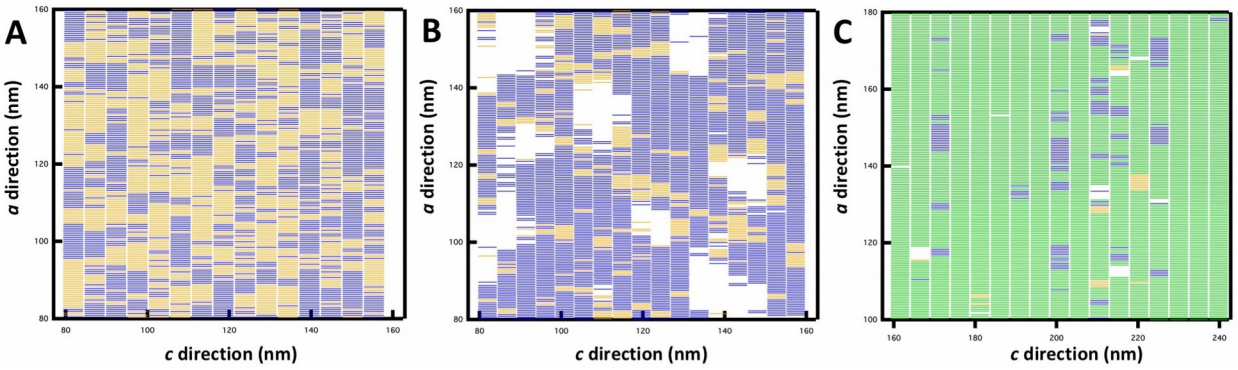

Figure 6. Distribution maps of motifs in the $pNdc_9$-$pNte_9$ nanosheets prepared by different processes. Each individual rectangle in the maps represents a unit cell. Anti-parallel V-shaped motifs are shown as blue boxes. Parallel V-shaped motifs are shown as green boxes. Disordered motif are shown as yellow boxes. White areas within the maps represent the regions that could not be analyzed, i.e., they could not be placed in one of the 12 classes shown in Figure 4. **A**. Nanosheet obtained by fast evaporation. **B**. Nanosheet obtained by slow evaporation. **C**. Nanosheet obtained by very slow evaporation.

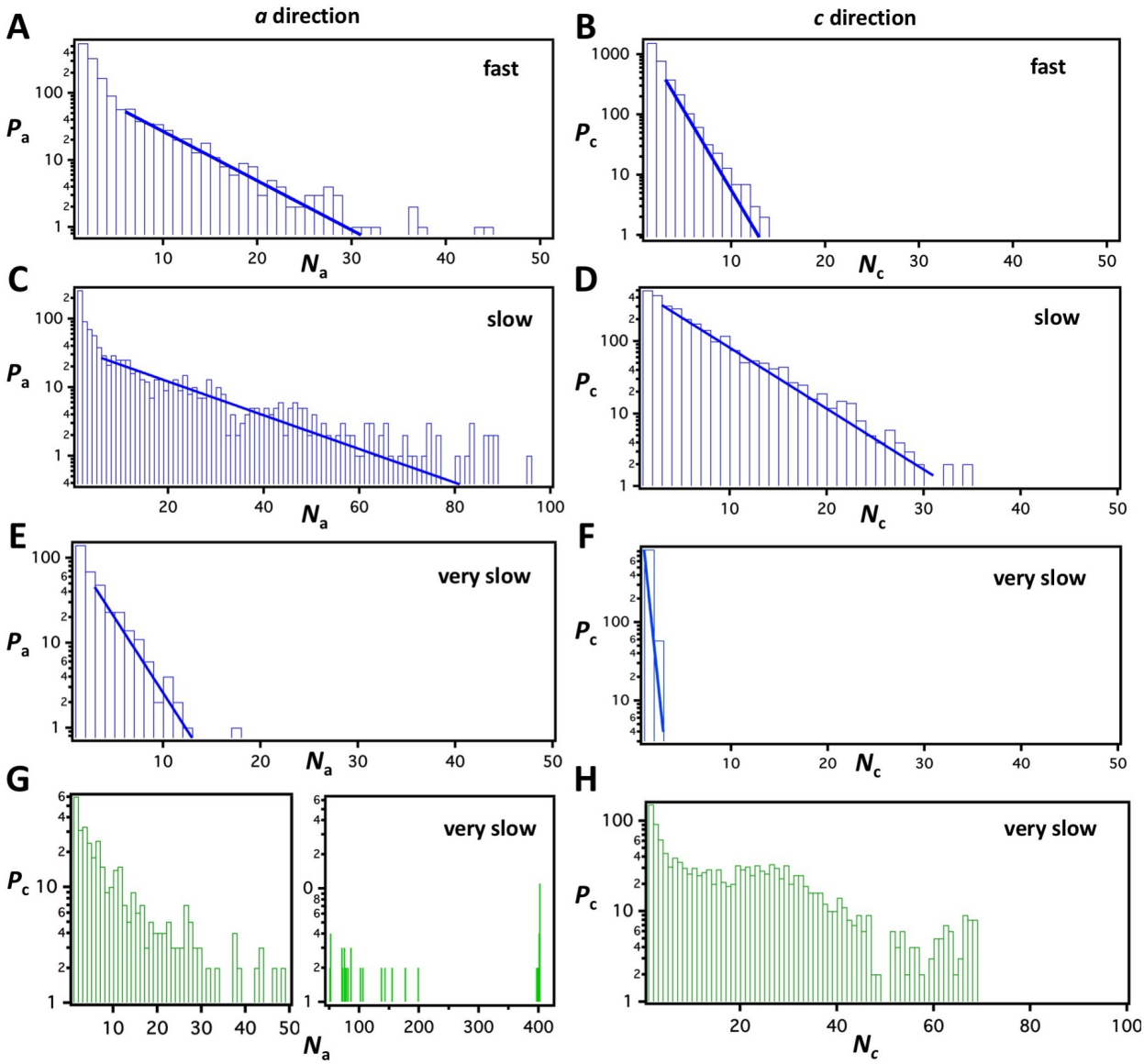

Figure 7. Quantitative analysis of the distributions of motifs in the pNdc$_9$–pNte$_9$ nanohseets.

P_a and P_c represent the population of concatenated boxes with the same motif along *a* and *c* directions. N_a and N_c represents the coherence length of concatenations along *a* and *c* directions. When applicable, eq 1 was used to fit the data and the lines in the figures represent these fits. Blue bars represent correlations between adjacent anti-parallel V-shaped motifs. Green

bars represent correlations between adjacent parallel V-shaped motifs. **A.** Histogram showing the distribution of anti-parallel V-shaped motifs along *a* direction in the nanosheet prepared by fast evaporation. For example, in the nanosheet analyzed in A, there were 28 rectangular regions with 10 concatenated boxes along the *a* direction with the anti-parallel V-shaped motif (P_a = 28 at N_a = 10). **B.** Histogram showing the distribution of the anti-parallel V-shaped motifs along *c* direction in the nanosheet prepared by fast evaporation. **C through H.** Histograms showing the distributions of the anti-parallel and parallel V-shaped motifs along *a* and *c* directions. **C, D:** anti-parallel V-shaped motifs obtained during slow evaporation. **E, F:** anti-parallel V-shaped motifs obtained during very slow evaporation. **G, H:** parallel V-shaped motifs obtained during very slow evaporation.

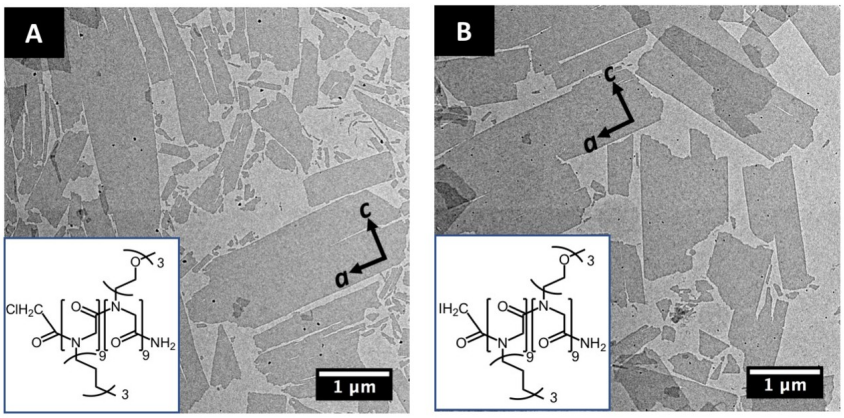

Figure 8. Morphology of the self-assembled Cl-pNdc$_9$–pNte$_9$ and I-pNdc$_9$–pNte$_9$ nanohseets obtained by very slow evaporation. **A**. Low-dose TEM micrograph of Cl-pNdc$_9$–pNte$_9$ nanosheets. **B.** Low-dose TEM micrograph of I-pNdc$_9$–pNte$_9$ nanosheets. Insets show the chemical structures of polypeptoids; end groups at the left end are different. The *a* and *c* directions for particular nanosheets are shown.

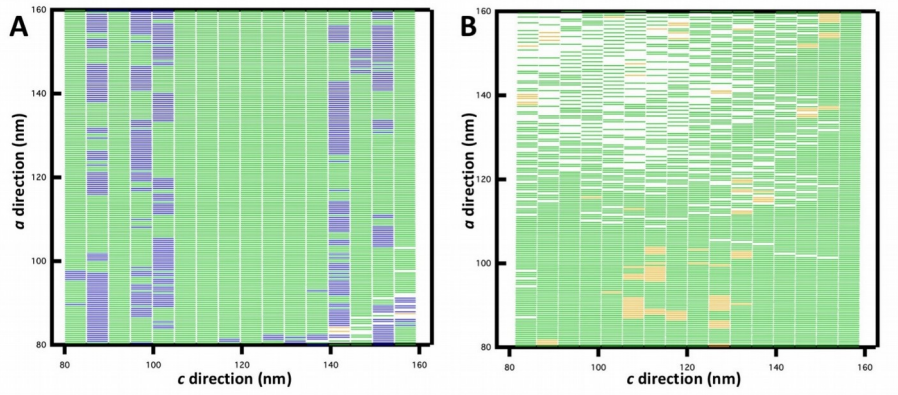

Figure 9. Distribution maps of motifs in the Cl-pNdc$_9$-pNte$_9$ and I-pNdc$_9$-pNte$_9$ nanosheets prepared by very slow evaporation. Each individual rectangle in the maps represents a unit cell. Anti-parallel V-shaped motifs are shown as blue boxes. Parallel V-shaped motifs are shown as green boxes. Disordered motifs are shown as yellow boxes. White areas within map represent the regions that were not analyzed. **A**: Cl-pNdc$_9$-pNte$_9$ nanosheet, **B**: pNdc$_9$-pNte$_9$ nanosheet.

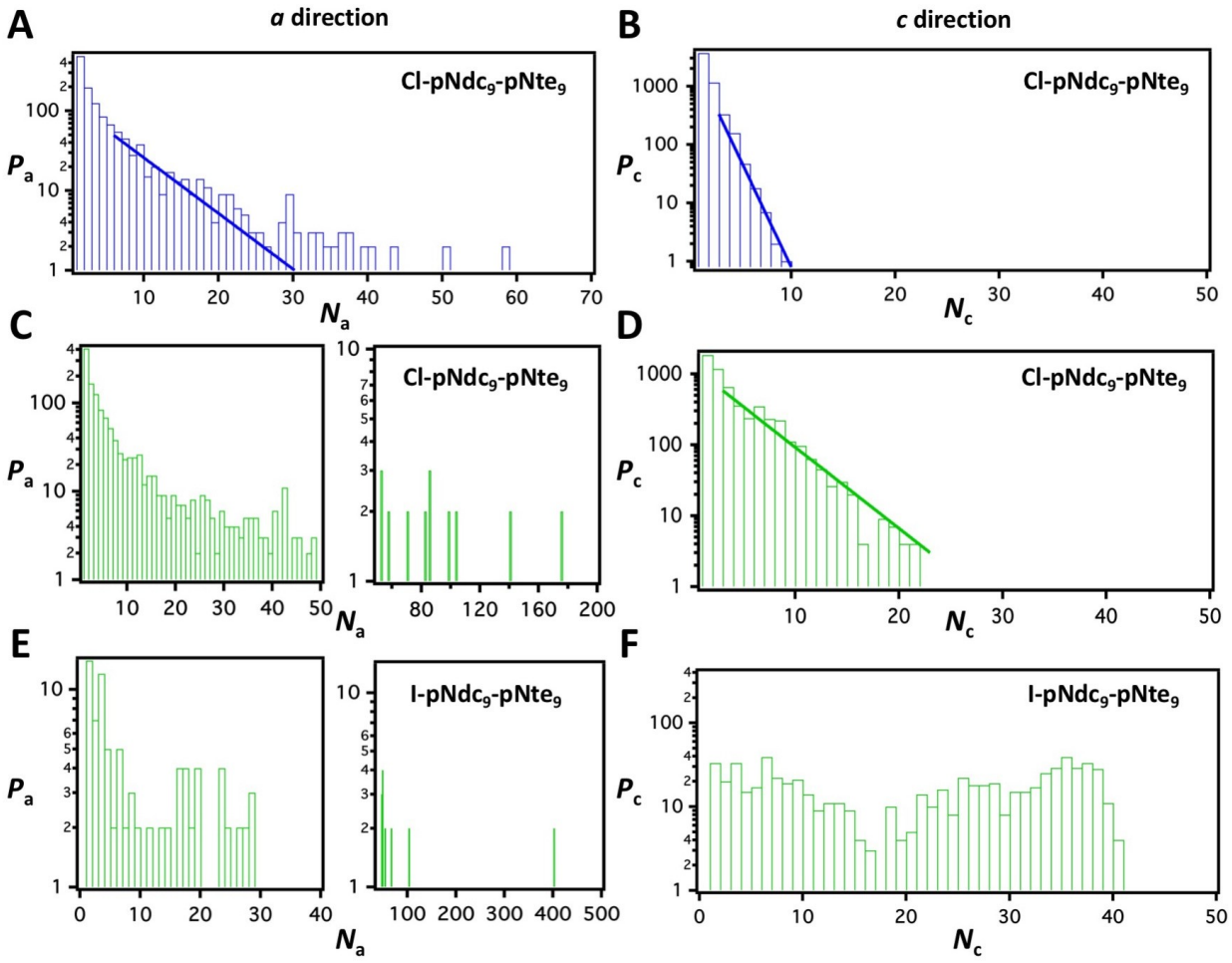

Figure 10. Quantitative analysis of the distributions of motifs in Cl-pNdc$_9$–pNte$_9$ and I-pNdc$_9$–pNte$_9$ nanohseets prepared by very slow evaporation. P_a and P_c represent the population of concatenated boxes with the same motif along *a* and *c* directions. N_a and N_c represent the coherence length of concatenations along *a* and *c* directions. When applicable, eq 1 was used to fit the data and the lines in the figures represent these fits. Blue bars represent correlations between adjacent anti-parallel V-shaped motifs. Green bars represent correlations between adjacent parallel V-shaped motifs. **A.** Histogram showing the distribution of the anti-parallel V-shaped motifs along *a* direction in the Cl-pNdc$_9$–pNte$_9$ nanosheet. **B.** Histogram showing the

distribution of the anti-parallel V-shaped motifs along *c* direction in the Cl-pNdc$_9$–pNte$_9$ nanosheet. **C through F.** Histograms showing the distributions of the parallel V-shaped motifs along *a* and *c* directions. **C, D:** Cl-pNdc$_9$–pNte$_9$. **E, F:** I-pNdc$_9$–pNte$_9$.

Table 2. Mean coherence length in units of unit cells along *a* and *c* directions in the nanosheets based on eq 1.

Mean length $\bar{N}_i$	pNdc$_9$-pNte$_9$ (fast)	pNdc$_9$-pNte$_9$ (slow)	pNdc$_9$-pNte$_9$ (very slow)	Cl-pNdc$_9$-pNte$_9$ (very slow)	I-pNdc$_9$-pNte$_9$ (very slow)
$\bar{N}_a$ (anti-parallel)	4.3	14.4	2.7	5.9	NO
$\bar{N}_c$ (anti-parallel)	2.1	5.7	1.0	1.5	NO
$\bar{N}_a$ (parallel)	NO	NO	54.1	13.9	78.7
$\bar{N}_c$ (parallel)	NO	NO	19.2	3.5	20.8

NO: not observed

Table 3. Coherence length in units of unit cells along *a* and *c* directions in the nanosheets based on an exponential distribution function (eq 2).

	$pNdc_9$-$pNte_9$ (fast)	$pNdc_9$-$pNte_9$ (slow)	$pNdc_9$-$pNte_9$ (very slow)	Cl-$pNdc_9$-$pNte_9$ (very slow)	I-$pNdc_9$-$pNte_9$ (very slow)
L_a (anti-parallel)	5.8	17.8	2.4	6.0	NO
L_c (anti-parallel)	1.5	5.2	0.4	1.1	NO
$L_{a,max}$ (anti-parallel)	45	95	18	58	NO
$L_{c,max}$ (anti-parallel)	14	34	2	9	NO
L_a (parallel)	NO	NO	NEBD	NEBD	NEBD
L_c (parallel)	NO	NO	NEBD	3.8	NEBD
$L_{a,max}$ (parallel)	NO	NO	400	176	400
$L_{c,max}$ (parallel)	NO	NO	69	22	41

L_a represents the average coherence length along the *a* direction; L_c represents the average coherence length along the *c* direction; $L_{a,max}$ represents the maximum coherence length along the *a* direction; $L_{c,max}$ represents the maximum coherence length along the *c* direction. NO: not observed. NEBD: non-exponential broad distribution.

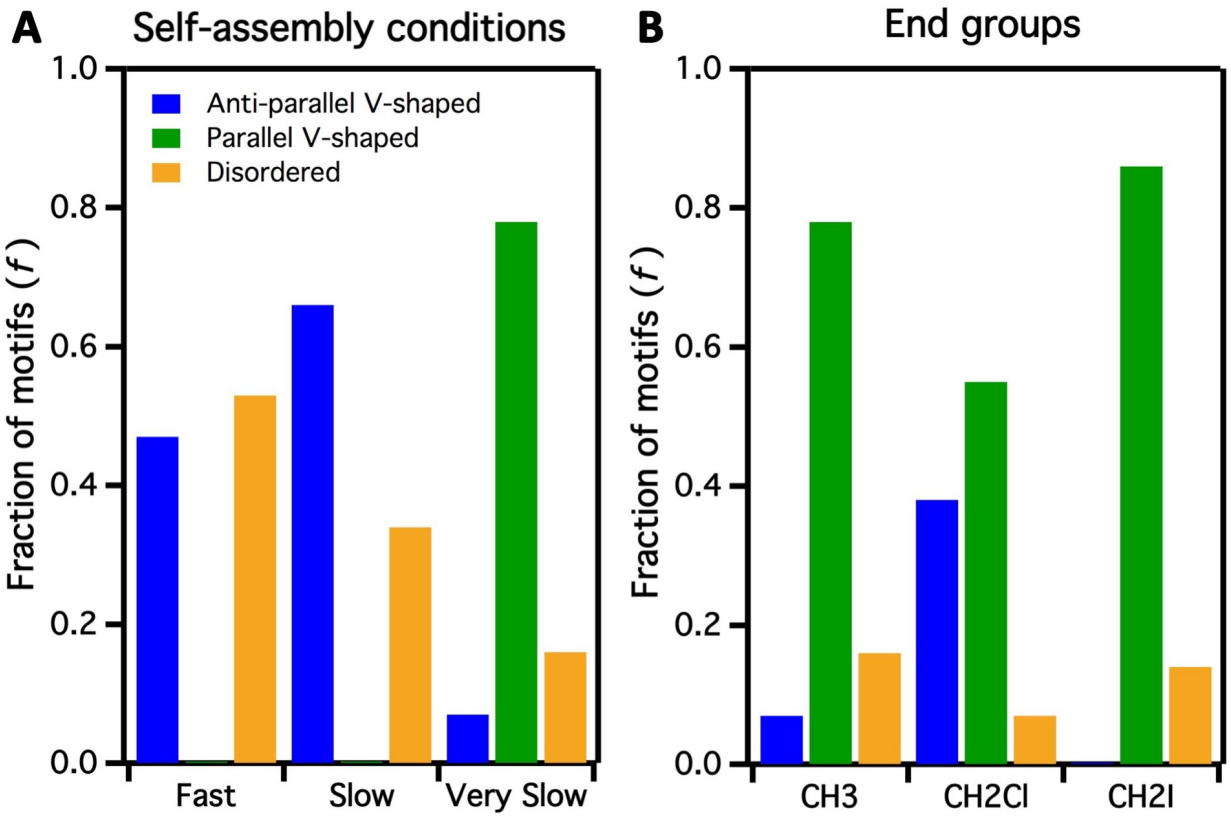

Figure 11. Effect of evaporation rate and end group chemistry on the proportion of crystal motifs. Fraction of motifs, *f*, are given as a function of **A**: evaporation rate for pNdc$_9$-pNte$_9$, **B**: end group chemistry under very slow evaporation conditions. Blue, green and yellow bars represent anti-parallel, parallel, and disordered motifs in both panels.