PNAS
www.pnas.org

Supplementary Information for

Amyloid-like amelogenin nanoribbons template mineralization via a low energy interface of ion binding sites

S. Akkineni[1,2], C. Zhu[3], J. Chen[1,2], M. Song[1,2], S.E. Hoff[3], J.S. Bonde[4], J. Tao[2*], H. Heinz[3*], S. Habelitz[5*] and J.J. De Yoreo[1,2*]

*Corresponding Author Email: James.DeYoreo@pnnl.gov (JJDY), stefan.habelitz@ucsf.edu (SH), hendrik.heinz@Colorado.edu (HH) and Jinhui.Tao@pnnl.gov (JT)

This PDF file includes:

Legends for Movies S1 to S6

Figures S1 to S14

Methods and Notes 1 to 11

Tables S1 to S6

SI References (1-27)

Other supplementary materials for this manuscript include the following:

Movies S1 to S6

Movie S1: In situ AFM showing the nucleation and growth of amorphous calcium phosphate (ACP) on nanoribbons of 14P2 on HOPG at 1 frames/sec by continuous flow of supersaturated solution ($\sigma_{ACP} = 0.04$)

Movie S2: In situ AFM showing the nucleation and growth of amorphous calcium phosphate (ACP) on nanoribbons of 14P2Cterm on HOPG at 1 frames/sec by continuous flow of supersaturated solution ($\sigma_{ACP} = 0.04$)

Movie S3: In situ AFM showing the nucleation and growth of amorphous calcium phosphate (ACP) on nanoribbons of rH174 on HOPG at 1 frames/sec by continuous flow of supersaturated solution ($\sigma_{ACP} = 0.04$)

Movie S4: In situ AFM showing the nucleation and growth of amorphous calcium phosphate (ACP) on nanoribbons of p14P2 on HOPG at 1 frames/sec by continuous flow of supersaturated solution ($\sigma_{ACP} = 0.04$)

Movie S5: In situ AFM showing the nucleation and growth of amorphous calcium phosphate (ACP) on nanoribbons of p14P2Cterm on HOPG at 1 frames/sec by continuous flow of supersaturated solution ($\sigma_{ACP} = 0.04$)

Movie S6: In situ AFM showing the nucleation and growth of amorphous calcium phosphate (ACP) on nanoribbons of p14P2 on HOPG at 1 frames/sec. ACP particles are aligned in the direction of the ribbons. Supersaturation with respect to ACP (σ_{ACP}) is 0.138.

Supplementary Figures

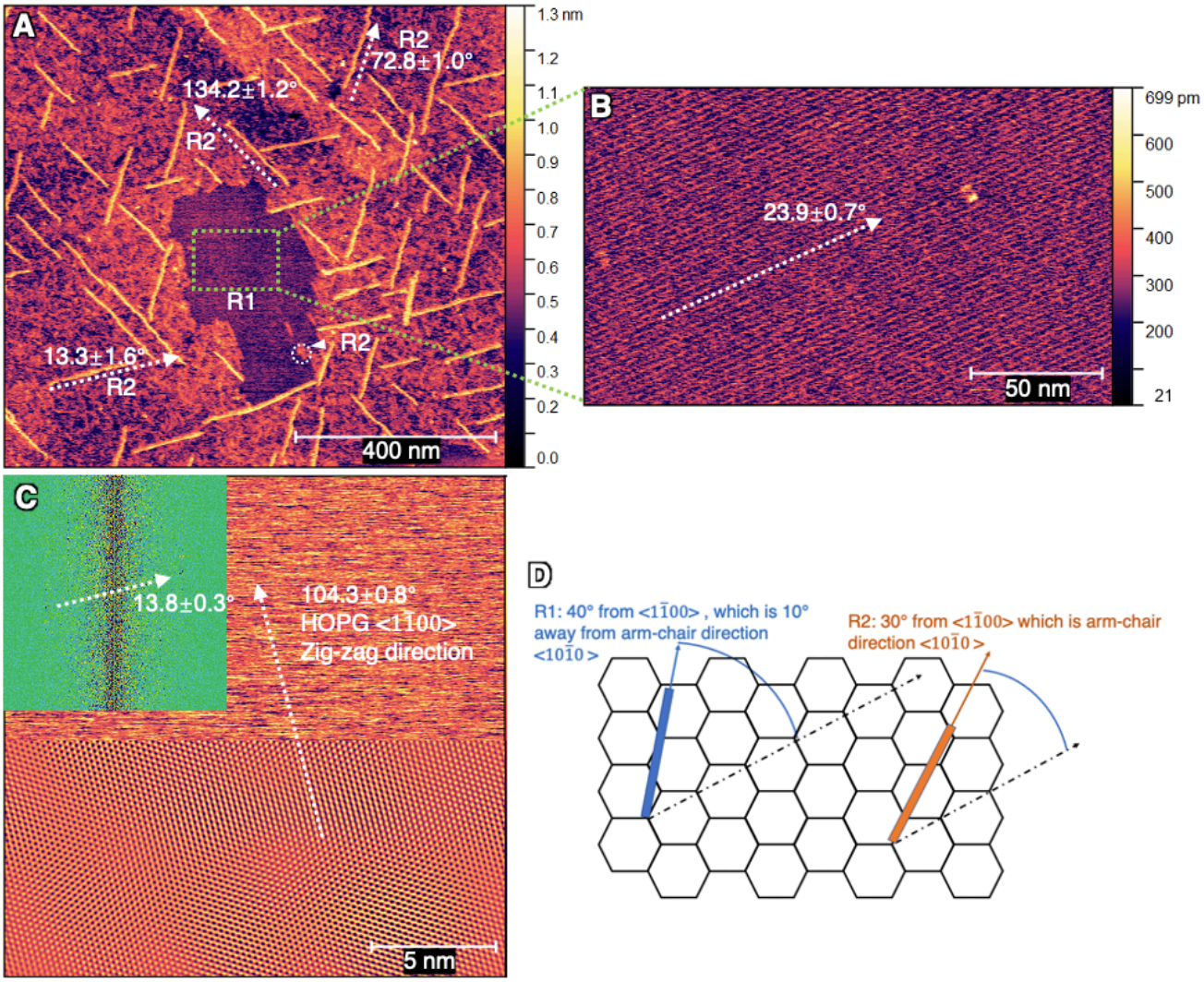


Fig. S1. Orientation of R1 and R2 NRs with respect to underlying HOPG (0001) plane from in situ AFM images of 14P2 using 0.01 mg/mL from freshly prepared 1 mg/mL stock solution (Fig. 1I). (**A**) Multiple layers are visible, white arrows indicate relative direction of R2 NRs which lie on the HOPG arm-chair direction, <10$\bar{1}$0> (~13.3˚), darkest area in the center consists of R1 NRs as seen in (B). (**B**) R1 NRs and direction relative to graphite lattice, (**C**) Top half is AFM image of underlying HOPG substrate collected after scratching the top layer with contact mode AFM and bottom half is the reconstructed image with noise removed by 2D-FFT filter using points shown in FFT inset indicating the HOPG zig zag direction <1$\bar{1}$00>. (**D**) R1 NRs lay 10˚ away from direction of R2 NRs (HOPG arm-chair direction <10$\bar{1}$0>).

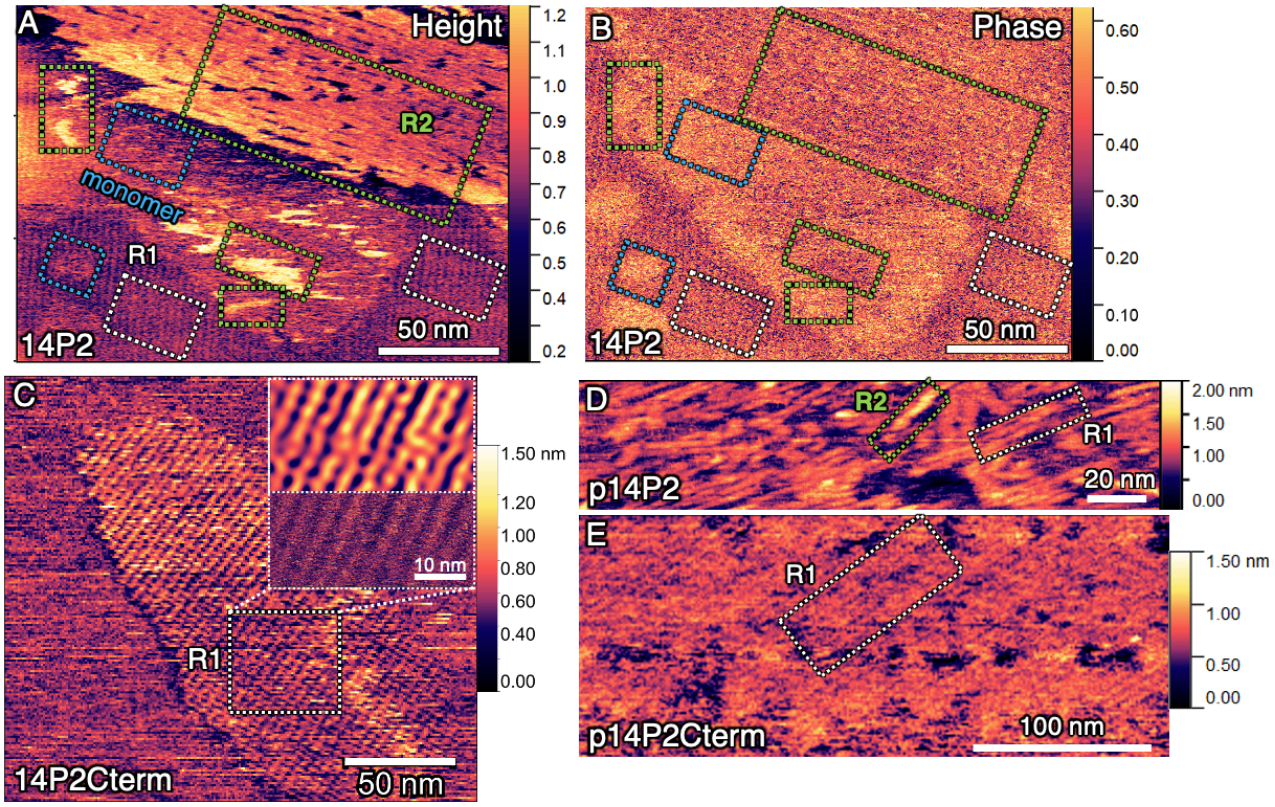


Fig. S2. In situ AFM images of R1 NRs on HOPG for 14P2, 14P2Cterm, p14P2 and p14P2Cterm observed with low peptide concentrations (≤0.01 mg/mL, fresh solutions). (**A** and **B**) Representative data of 14P2 showing the method for identifying NR type from AFM height and phase contrast data. Green boxes are R2 NRs which have a darker phase contrast and larger thickness (brighter features in height), white boxes are R1 NRs that also show a dark contrast but with roughly half the width as R2 NRs, and blue shows locations where amorphous structures or monomers (brightest features in phase images) bind to the surface. (**C**) 14P2Cterm R1 NRs with inset showing molecular resolution image of R1 NRs where bottom half is original image and top half with noise removed using 2D-FFT filter, **(D)** p14P2 R1 and R2 NRs and **(E)** p14P2Cterm R1 NRs.

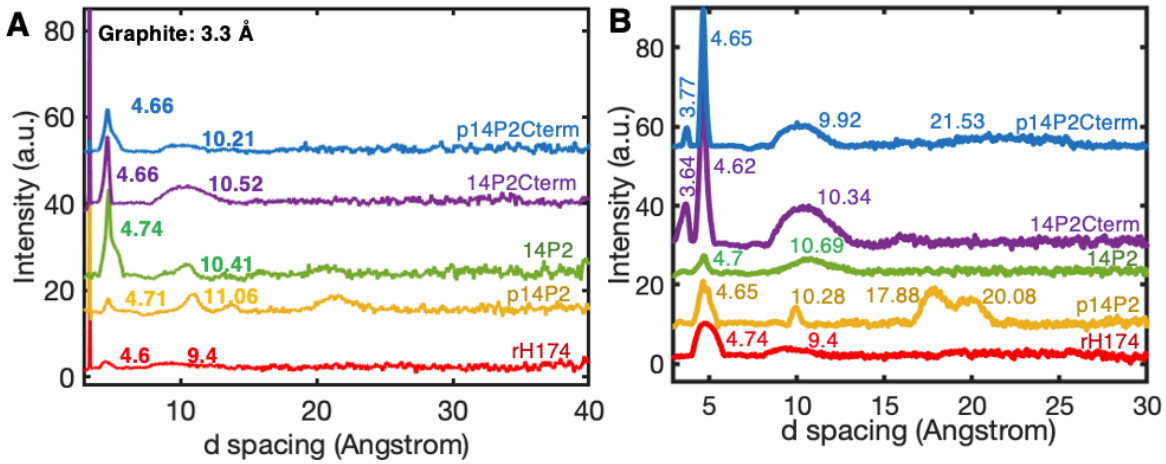
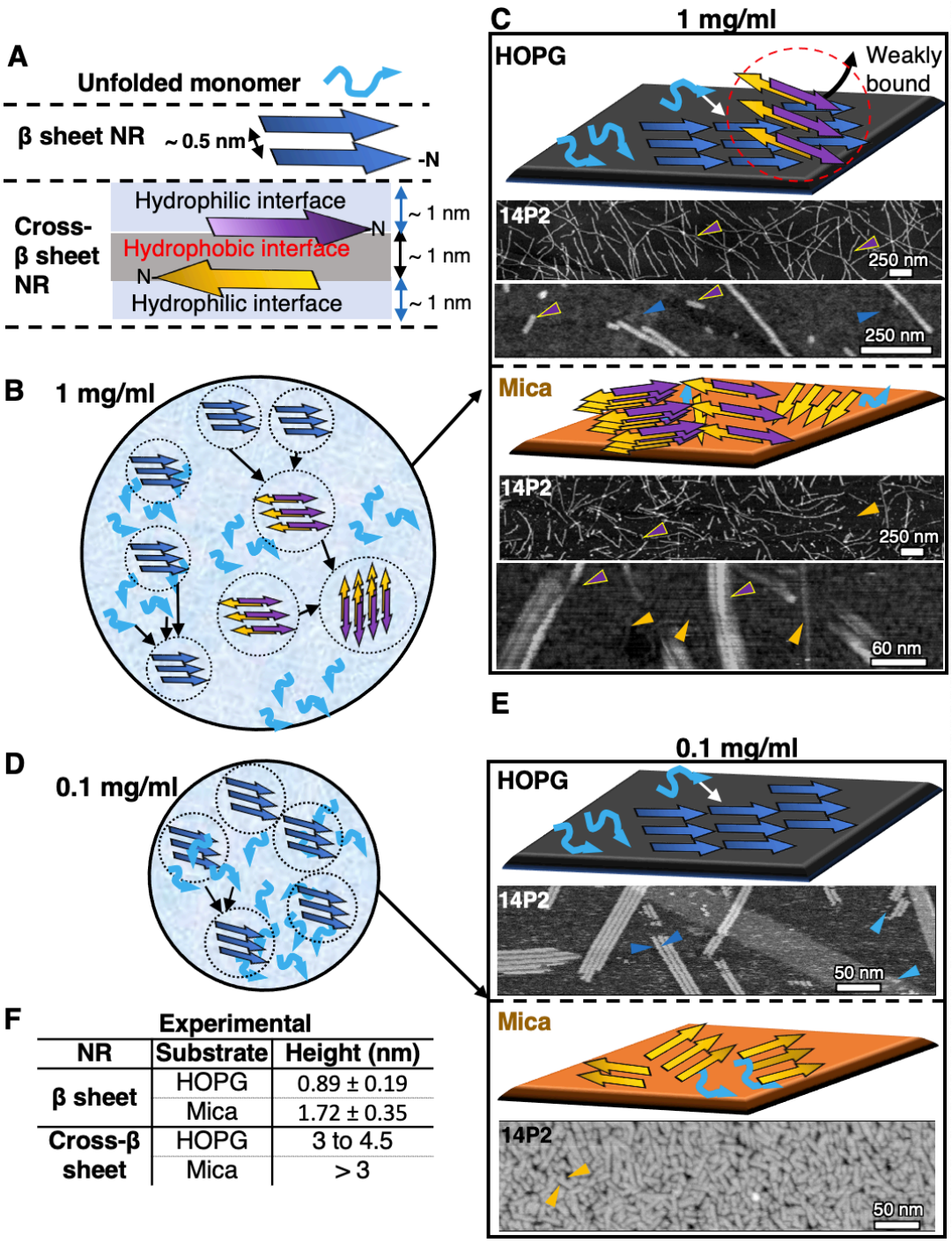


Fig. S3. In situ XRD patterns of all sequences at ≥1 mg/mL (pH 1.94) aged for 14 days (**A**) in presence of graphite prepared by using NRs as an adsorbate on HOPG and (**B**) without graphite. Both measurements show the characteristic ~4.7 Å d-spacing of β-sheet NRs, the broad 10 Å peak indicating presence of cross-β-sheet NRs, and absence of d-spacings from other conformations. Detailed methods and discussion are in *SI Appendix*, Method 3.1 and 3.2. The supramolecular structures here are further characterized in *SI Appendix*, Fig. S4.



F

Experimental

NR	Substrate	Height (nm)
β sheet	HOPG	0.89 ± 0.19
	Mica	1.72 ± 0.35
Cross-β sheet	HOPG	3 to 4.5
	Mica	> 3

Fig. S4. Schematic describing the composition of typical solutions used for XRD measurements and AFM-based mineralization experiments, and impact of surfaces on NR binding, orientation and self-assembly. (**A**) Arrow color legend for observed features and their expected configuration: Light blue - monomers; Dark blue - single β-sheet; Yellow: single β-sheet with hydrophilic interface facing mica; Yellow and Violet: Cross-β-sheet. (**B**) Cartoon of the constituents in high concentration solutions (≥1 mg/mL) aged for 14 days typically used for XRD, as seen in (**C**) for 14P2 on HOPG and muscovite mica showing cross-β-sheet NRs as bright features and single β-sheet NRs in the background. (**D**) Cartoon of the constituents in low concentration solutions (0.1 mg/ml diluted from 1 mg/mL stock aged for 48 – 336 hours), as seen in (**E**) for 14P2 on HOPG and muscovite mica showing only single-β-sheet NR. (**F**) Comparison of average height of NRs in C and E. See *SI Appendix*, Method 3.3 and 3.4 for detailed methods and discussion.

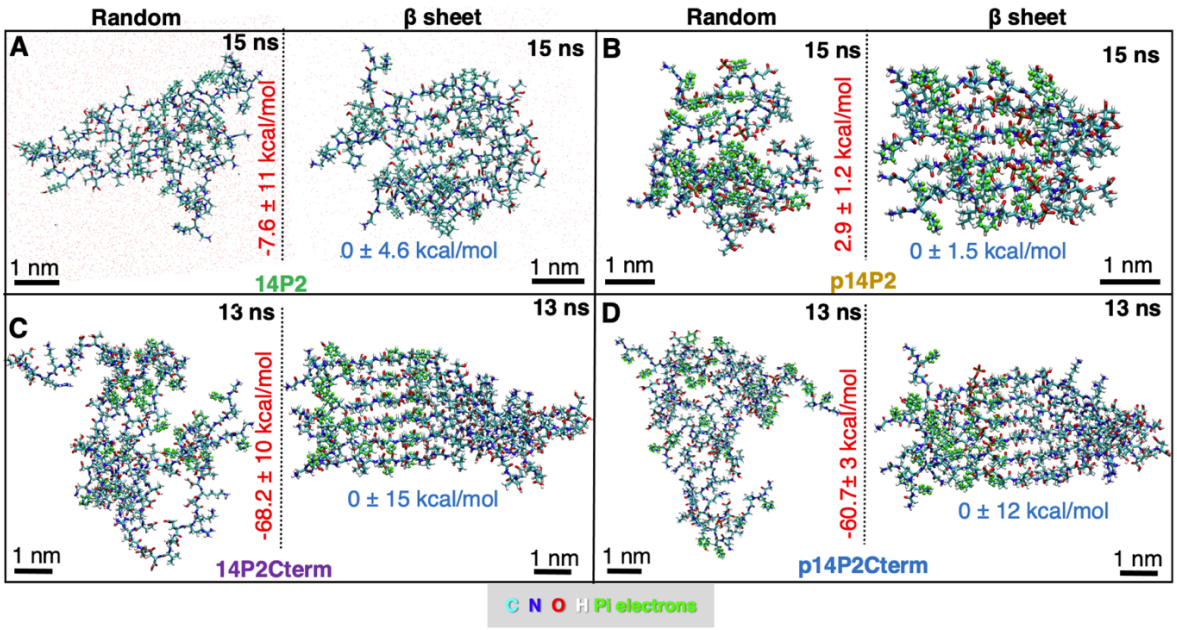


Fig. S5. Equilibrated structures and relative energies for 6-monomer peptides with random coil and β-sheet conformations in solution (10 mM HCl, pH 1.94) after 13-15 ns simulations. The average energy of the β-sheet structure was set to zero as a reference state including the uncertainty. (**A**) 14P2, (**B**) p14P2, (**C**) 14P2Cterm and (**D**) p14P2Cterm. Water molecules are hidden for clarity (except in A). See *SI Appendix*, Methods 4.1-4.3 for details.

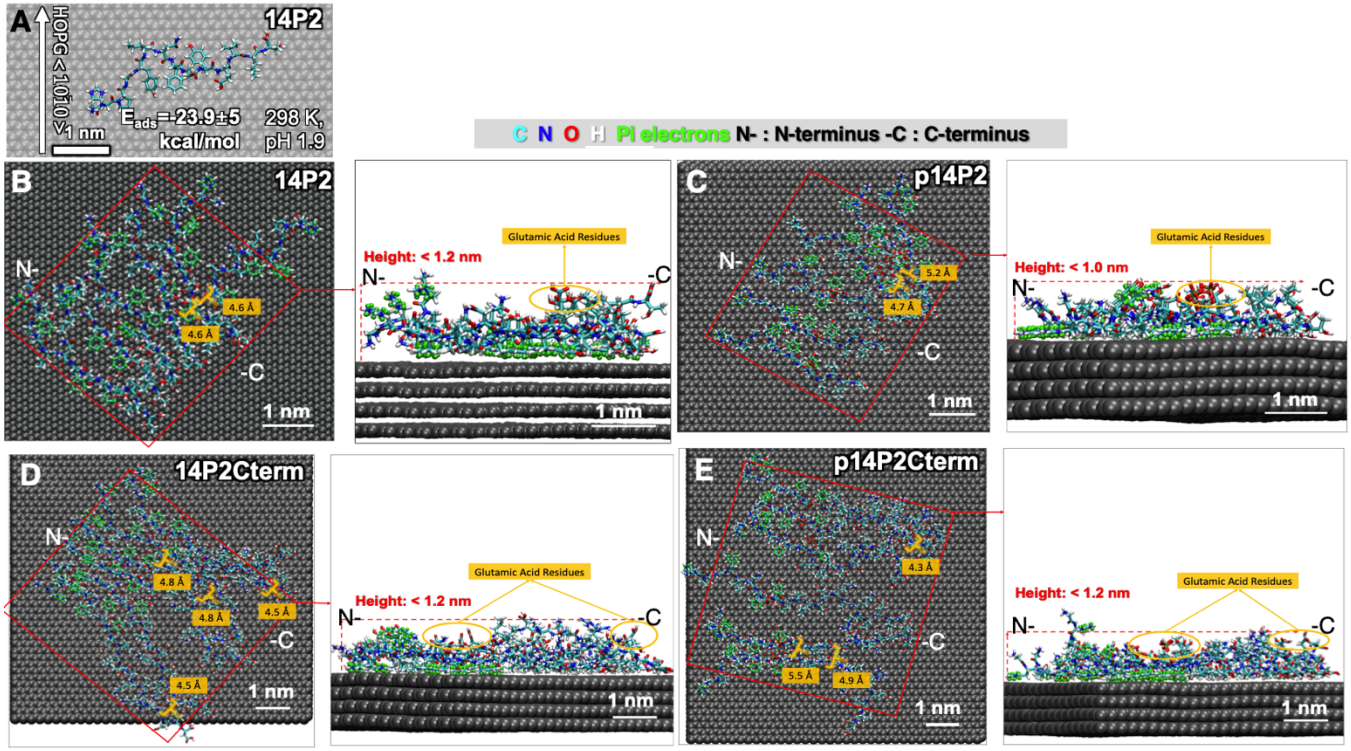


Fig. S6. Adsorption energy of monomer and simulation of β-sheet conformations of 6 peptide monomers on HOPG (water molecules are hidden for clarity). (**A**) Representative snapshot of a fully extended 14P2 (β strand) on HOPG at 8 ns (virtual π electrons are hidden). Snapshots of β-sheet conformation after 20 ns simulations in pH 1.94 on HOPG for (**B**) 14P2, (**C**) p14P2, (**D**) 14P2Cterm and (**E**) p14P2Cterm. Left: top view; right: side view. Distances between part of the backbone of the β-sheets and height of the peptides are marked in B-E. See *SI Appendix*, Methods 4.1, 4.2, 4.4 and 4.5 for details.

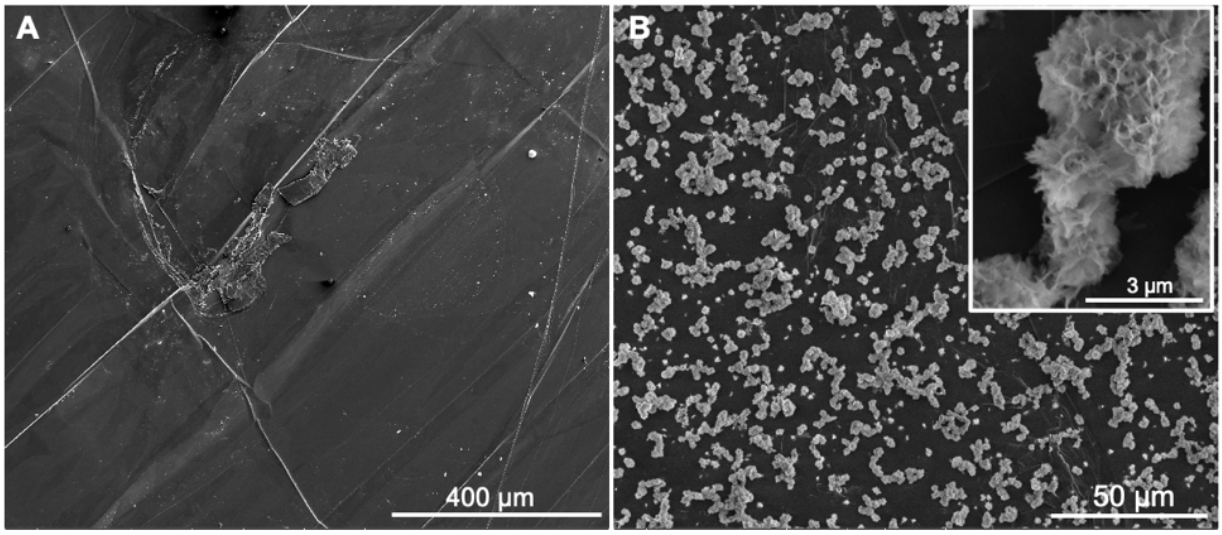


Fig. S7. SEM images of HOPG surfaces incubated in supersaturated calcium phosphate solutions σ_{ACP} = 0.221 for 1hr (**A**) without adsorbed peptide and (**B**) with 14P2 NRs. Control of bare HOPG shows no significant mineral, whereas 14P2 coated HOPG has plate-like mineral at 1 hr, morphologically identical to in situ AFM experiments (*SI Appendix*, Fig. S11 C). For detailed methods, refer to *SI Appendix*, Method 6.

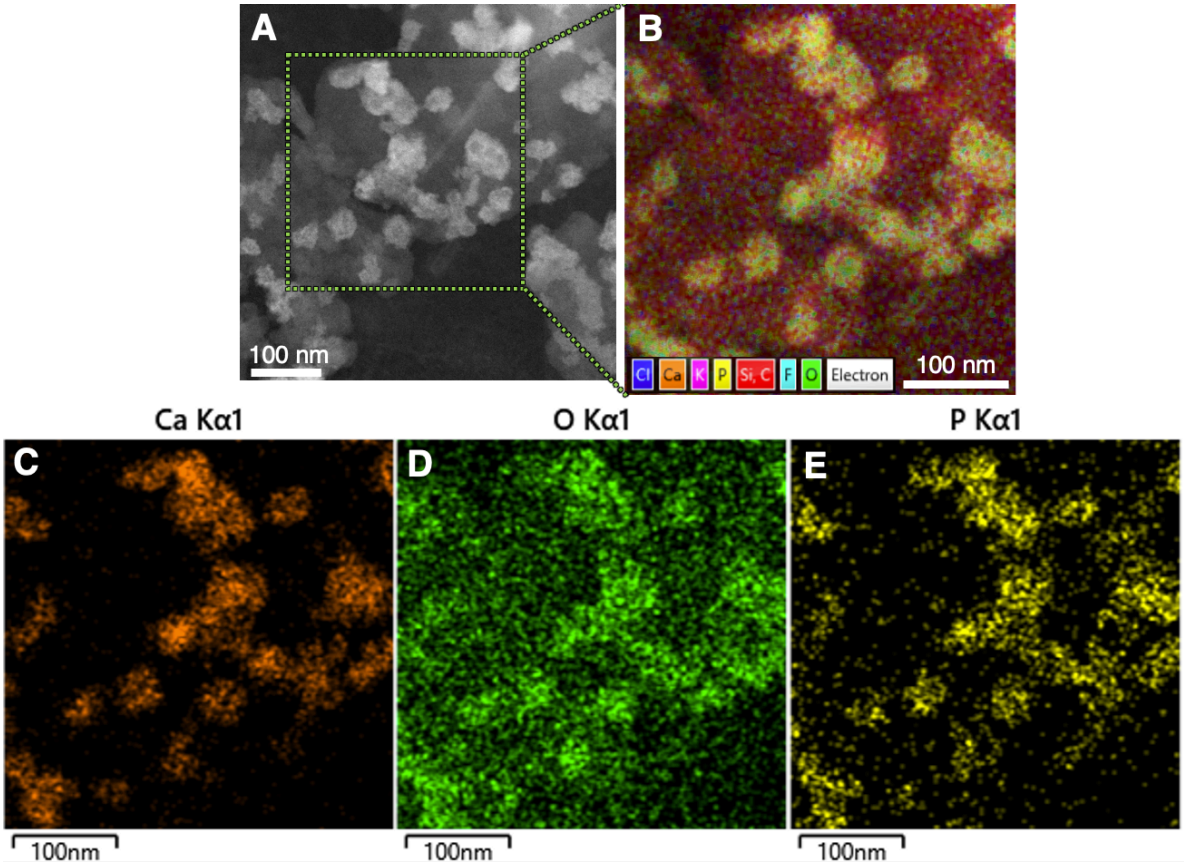


Fig. S8. Representative Scanning-TEM and Elemental Dispersive Spectroscopy (EDS) map of mineral particles on graphene grids functionalized with p14P2Cterm NRs then incubated in σ_{ACP} = 0.221 calcium phosphate solution at pH 7.4 for 20 min. (**A**) STEM image, (**B**) composite image of elements N, O, C, Si, Ca, K, P and electrons, (**C**) Map of Calcium, (**D**) Map of Oxygen, and (**E**) Map of Phosphate. The particles had high concentrations of Ca, P and O whereas the background had high C, O and N signal corresponding to the graphene and organic constituents.

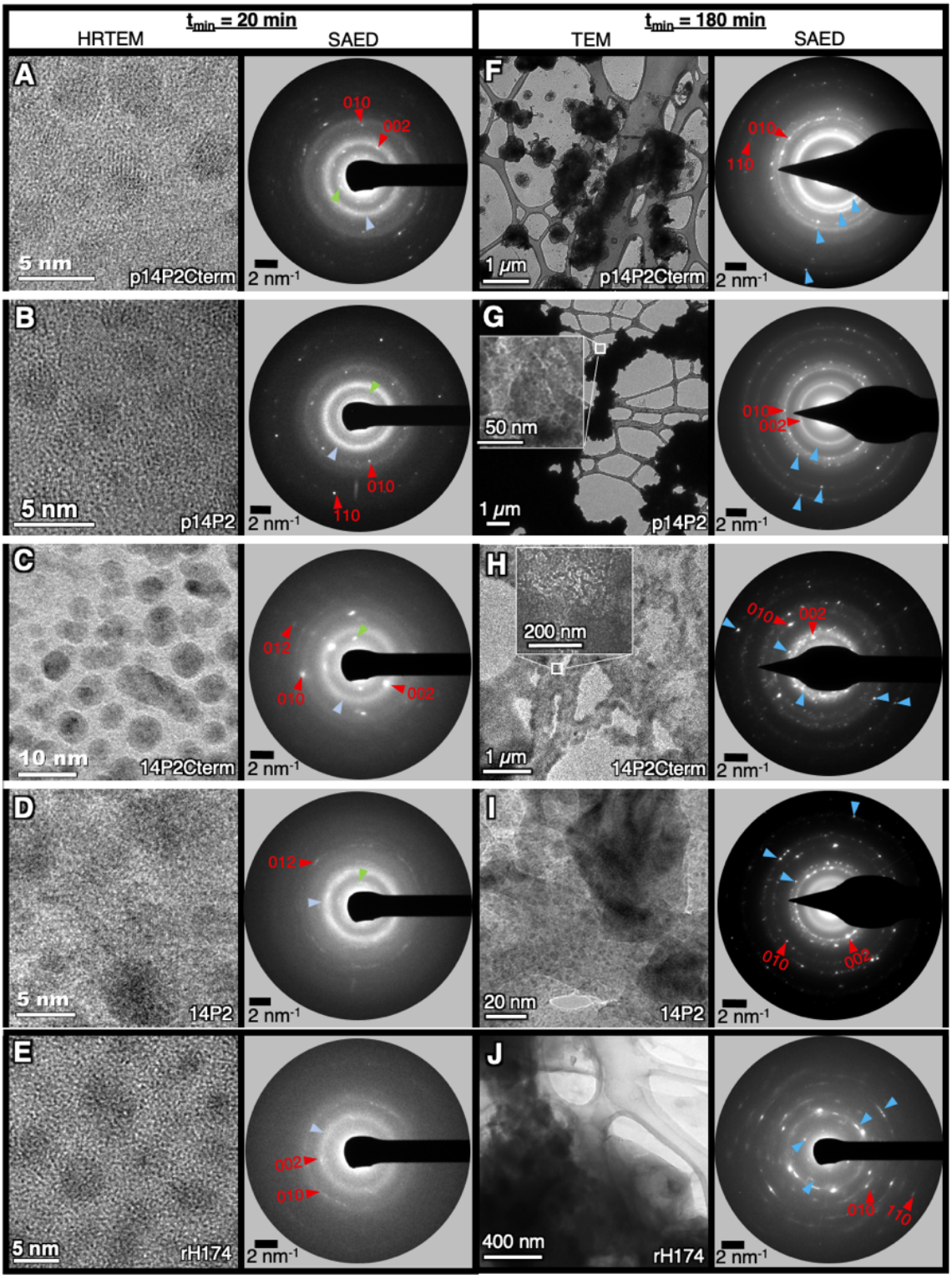


Fig. S9. TEM characterization of quenched graphene grids functionalized with NRs of (**A**) and (**F**) p14P2Cterm, (**B**) and (**G**) p14P2, (**C**) and (**H**) 14P2Cterm, (**D**) and (**I**) 14P2, and (**E**) and (**J**) rH174 after incubation in σ_{ACP} = 0.221 calcium phosphate solution for (**A**) to (**E**) 20 min and (**F**) to (**J**) 180 min. At 20 min, particles are amorphous, shown by SAED with diffraction spots from graphite (red), peptides (green) and diffuse ring of ACP (light blue) and further analyzed in Fig. S10. (F) to (J) shows higher crystallinity at 180 min and SAED confirms speckled ring diffraction with spots from AP or OCP (blue). For method and detailed discussion, refer to *SI Appendix,* Methods 7 and 8.

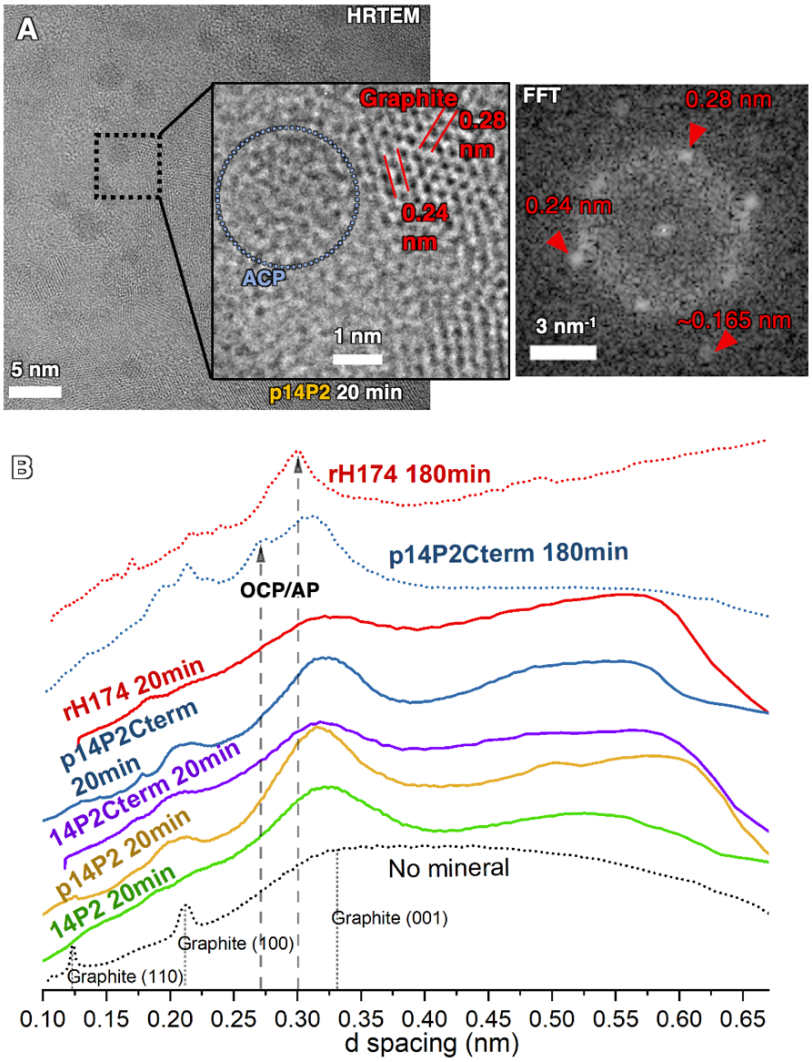


Fig. S10. Amorphous nature of particles mineralized for 20 min on 3-5 layer graphene substrates coated with NRs assessed through FFT of HRTEM and radially integrated SAED data. HRTEM for p14P2 in **(A)** shows several round particles (dark features). Further magnification of area in black dotted square shows the region marked as ACP (blue circle) lacks long range order and FFT of the image shows spacing correlated to graphene. 0.28 and 0.24 nm spacings match the C-C distance of graphene/graphite arm-chair <$10\bar{1}0$> and zig-zag <$1\bar{1}00$> directions (Cong et al 2013), respectively, while the weak ~0.165 nm matches the (004) plane of 3-5 layer graphene or graphite. **(B)** Radially integrated profiles of SAED on protein-coated grids after 20 min and 180 min mineralization (Fig. S9). At 20 min, all sequences show absence of the sharp crystalline peaks of OCP and AP at 0.26-0.33 nm d-spacing, (indicated by dashed vertical lines) previously reported using Cryo-TEM and Low Dose-SAED (Habraken et al. 2013), and also observed here at 180 min for rH174 and p14P2Cterm mineralization. Other sequences show similar diffraction patterns at 180 min. SAED from a 3-5 layer graphene grid coated with p14P2Cterm NRs but without mineral was used as control.

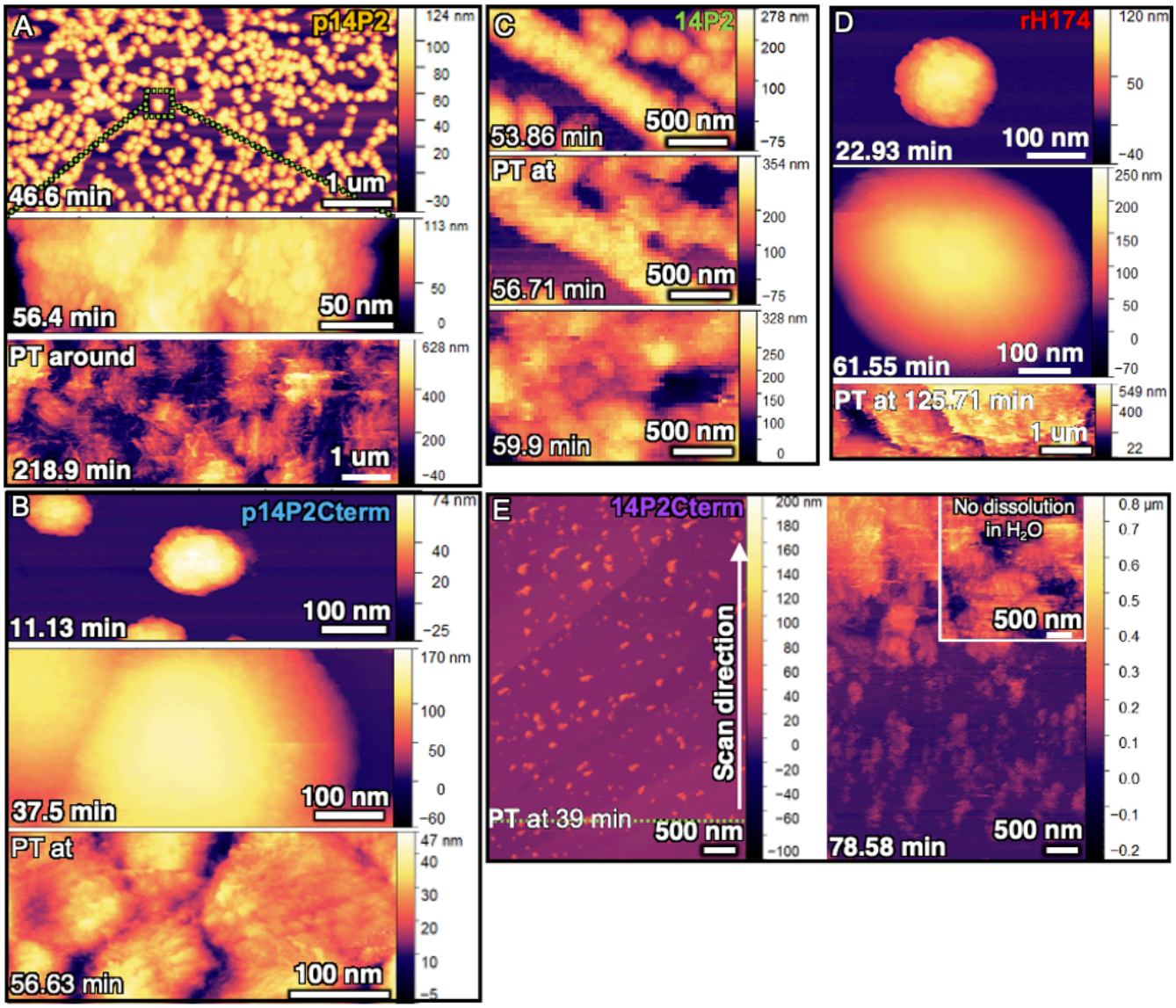


Fig. S11. ACP growth and phase transformation using height images from in situ AFM under constant composition (σ_{ACP} = 0.221), pH 7.4 and 25˚C on NRs of **(A)** p14P2, **(B)** p14P2Cterm, **(C)** 14P2, **(D)** rH174 and **(E)** 14P2Cterm**.** Surface of the particles are initially smooth but eventually becomes rougher, which indicates that the amorphous particles phase transformed into a crystalline phase with a nanoplatelet or fibrous morphology. Phase transformation was fastest for 14P2Cterm (~39 min) and slowest for p14P2 (>150 min), which correlates with the growth rates, implying that without phosphorylation, ions prefer the ACP-solution interface over NR-solution or NR-mineral interfaces. See *SI Appendix*, Method 9 for methods and additional details.

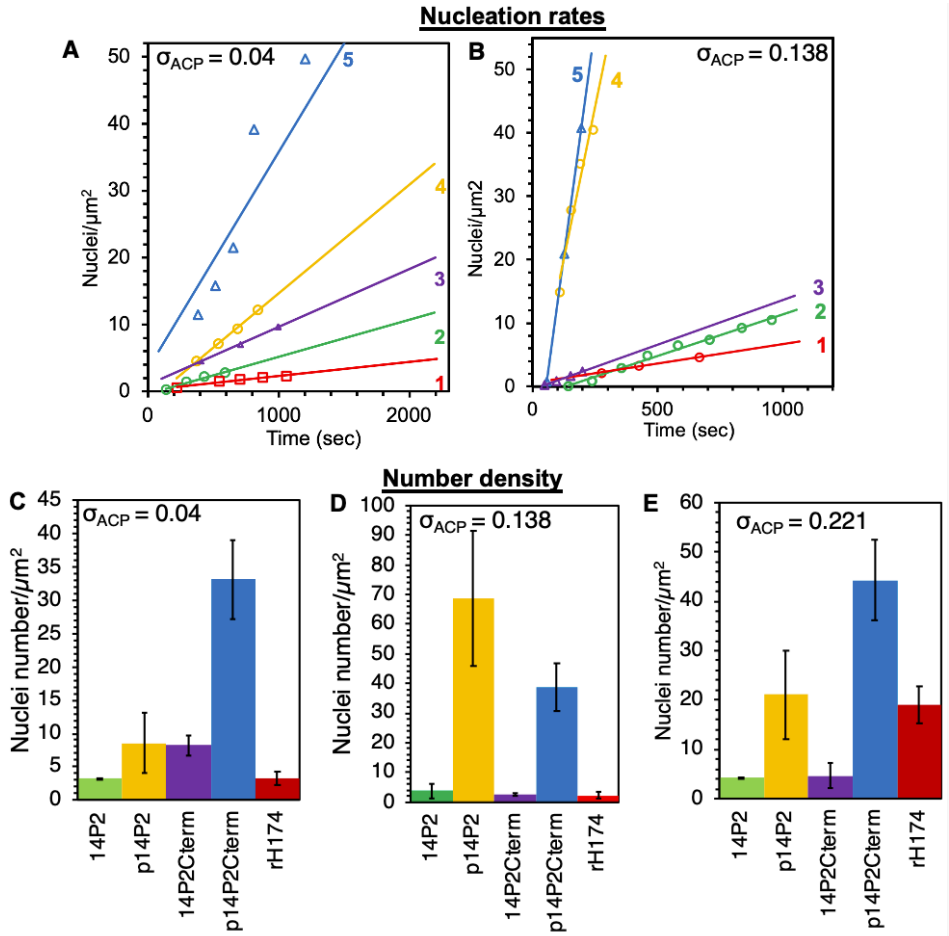


Fig. S12. Analysis of nucleation on various sequences at different supersaturations from in situ AFM data where (1) is rH174, (2) is 14P2, (3) is 14P2Cterm, (4) is p14P2, and (5) is p14P2Cterm. Representative nucleation rates (J_o) obtained from slopes of linear fits to number of nuclei per μm^2 over time at (**A**) $\sigma_{ACP} = 0.04$ and (**B**) $\sigma_{ACP} = 0.138$. Phosphorylated peptides have highest nuclei number density per μm^2 for various sequences at (**C**) $\sigma_{ACP} = 0.04$, at time = 15-18 minutes (**D**) $\sigma_{ACP} = 0.138$ at time = 5-6 minutes and (**E**) $\sigma_{ACP} = 0.221$ at time = 5-6 minutes. Error bar in C-E represents Std. dev, for sample size refer to Statistics section.

Supplementary Methods and Notes

Method 1: AFM image processing and analysis

1.1. Image processing: Raw height and phase images were processed using a polynomial flatten function and baselines corrected on Gwyddion SPM data analysis software. In addition, a conservative gaussian smoothing filter was applied to Fig 1 D to G and Movies S1 to S6 to remove pixelation from digital magnification. Noise reduction, reconstruction and deconvolution method for high-resolution images (Fig. 1 H, J and K and Fig. 2. B and F) are described in detail in *SI Appendix*, Method 2.

1.2. Height and width measurements: To obtain reasonable widths of NRs for all sequences that can be compared without using a deconvolution algorithm, the Full Width at Half Maximum (FWHM) height was averaged from images at similar magnifications and scan angle relative to ribbon long axis. Height measurements of soft matter from in situ AFM experiments are also subject to 0.5 – 1 nm error due to influences from the local environment (mobile proteins or monomers that skew the baseline), attractive and repulsive forces from the substrate or features, compression of the features, magnification, and AFM imaging setpoint and operation mode (1, 2). Therefore, the reported height is the apparent height measured using identical magnification, substrate and Tapping mode AFM

1.3. Nucleation rate and density analysis

Nucleation images were analyzed using Fiji/ImageJ (NIH). Strict classification rules were placed for definition of a nuclei, i.e., particles that appear in a frame and stay on to grow in the subsequent frames. However, in certain frames, nuclei disappeared or translated to another position after growth over the course of the experiment from the original locations. These nuclei were also counted and carefully tracked, frame by frame. The nuclei number densities in each frame were calculated by dividing number with frame sizes and plotted against time captured at end of the scan. The relationship of nuclei number density vs elapsed time was then linearly fit to get the slope, which is defined as nucleation rate for each supersaturation on a specific peptide or protein sequence (Fig. 3 F, *SI Appendix*, Fig. S12 and Table S4).

Method 2: High-resolution AFM image processing and NR structure analysis

Periodic features observed in the AFM images of 14P2 R1 and R2 were analyzed by radial integration of the noise filtered Fast Fourier transform (FFT) patterns and topography deconvolution (Fig. 2A and ***SI Appendix*, Fig. S13**).

Instrument generated noise was filtered out by masking the high intensity points in the raw FFT (*SI Appendix,* Fig. S13 B) and subtracting the background using 2D FFT-Filtering on Gwyddion. Subsequently, 2D FFT was performed on these 2D FFT-filtered images (*SI Appendix,* Fig. S13 C) to generate the noise-filtered FFT patterns (*SI Appendix,* Fig. S13 D). The noise filtered FFT patterns were then radially integrated using the Radial Profile plugin with spatial calibrations on Image J/Fiji (NIH) and converted to real-space (in nm) using $V(r) = 1/r_{INT}$, where r_{INT} is the reciprocal distance from the center (Fig. 2 A). The peak positions were identified by fitting using the Peak Analyzer tool on OriginPro software. Sharp peaks between 0.4-0.65 nm were clearly visible and identical to XRD pattern. The broad peak between 0.9-1.2 nm is not related to the cross-β-sheet structure since these are single β-sheets R2 NRs. For deconvolution of periodic features in the AFM topography, the high intensity spots in noise filtered FFTs were categorized into group 1 (~2 nm), group 2 (~1nm) and group 3 (~0.5 nm) (*SI Appendix,* Fig. S13 D). 2D FFT-filtering was performed again on the raw AFM data, but by masking combinations of the points marked in groups 1, 2 and 3 while subtracting the rest, until artifact-free features were generated (Fig. 2 B middle and bottom panel). Analysis showed group 3 (~0.5 nm) lies at the center of the NR and correlates with the distance between backbones of the monomers in β-sheet (Fig. 2 B middle panel). Deconvoluted features in group 2 (~1 nm) have an

unusual pattern and appear to coincide with the positions of N and C terminus in the overlaid model (Fig. 2 B bottom panel). Both periodicities were at different orientations and confirmed they are not duplicates. Group 1 (~2nm) features belong to the reduced width of the NR and may correlate to the ordered YINFSY domain.

A similar protocol was used for 14P2 R1 NRs. R1 NRs have intense peaks at periodicities of 0.78, 1.04, 1.24, 1.61, 1.76, 2.29 and 2.56 nm which are too large for a β-sheet structure and absent in the XRD measurements. Therefore, we infer that R1 may not form in solution and assembles only on graphite. The large distances (~0.7) suggest a structure with fully extended monomers held together by hydrogen bonds between the side chains (tyrosine) or between the side chain and the backbone. Such structures were previously reported to exist on atomically flat surfaces (3, 4).

For images without many high intensity FFT points, e.g. Fig. 2F, a radial mask fixed to the center and adjusted to include all ribbon features was used to filter out noise and streaks common in AFM images. This allowed an accurate reconstruction of the ribbon features.

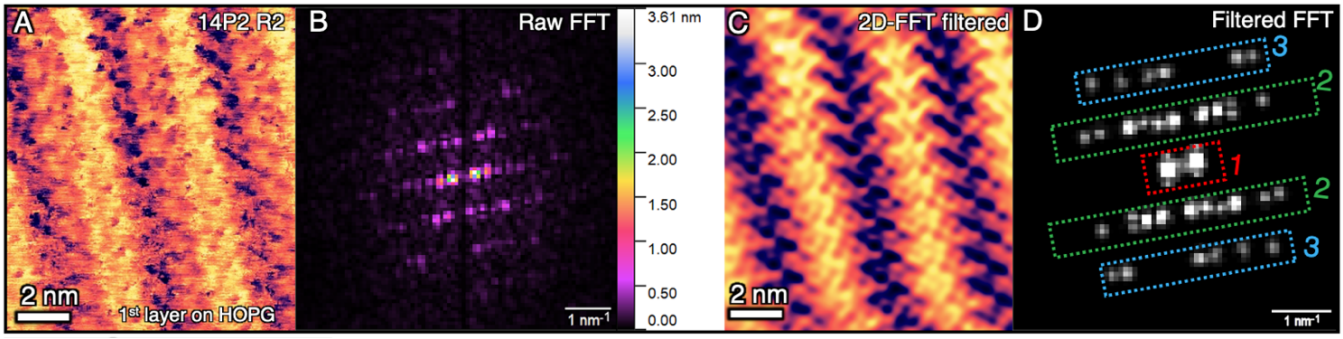


Fig. S13. FFT analysis of prominent periodic features in 14P2 R1 and R2 structures from AFM. Method for extracting periodic features from the AFM topography shown for (**A** -**E**) 14P2 R2 NRs in Fig. 2B. (**A**) High-resolution raw image of 14P2 R2 NR. (**B**) Raw FFT from un-filtered AFM image (A). (**C**) Image after 2D FFT-Filtering by masking and reconstructing high intensity points in (B). (**D**) FFT generated by 2D-FFT filtered images in C.

Method 3: Structure by in situ synchrotron XRD

3.1. Method:

Protein/peptide solutions were prepared according to the mineral-free self-assembly protocol (Materials and Methods) using higher concentrations to increase the XRD signal to noise ratio. Concentrations were 1 mg/ml (626.90 μM) for 14P2, 1.5 mg/ml (894.9 μM) for p14P2, 4 mg/ml (1.48 mM) for 14P2Cterm, 4.5 mg/ml (1.62 mM) for p14P2Cterm and 2 mg/ml (100.17 μM) for rH174. The stock solutions were aliquoted to prepare two samples, one with Amel NRs adsorbed on graphite and the other without graphite. To prepare suspensions containing graphite, 1-2 mm size freshly cleaved HOPG flakes were added to the freshly prepared stock solution aliquot then sonicated for an additional 30 min to disperse and exfoliate the HOPG flakes into smaller fragments for increasing the surface area and adsorption amount for Amel NRs. Both, with and without graphite, aliquots were then aged for 14 days and vortexed prior to use.

A 5 μl droplet of solution with or without graphite was drop cast onto the XRD mesh (M3-L18SP-10 MicroMesh or M7-L18SP-50 MicroGripper; MiTeGen, USA) and incubated for 40 minutes, then the excess solution was gently wicked away to leave a film covering the whole sample holder loop. Measurements were performed within 5 minutes of preparation to avoid dehydration and repeated at least three times. Powder diffraction measurements were performed with 0.11159-nm synchrotron radiation at Beamline 8.3.1 of the Advanced Light Source at Lawrence Berkley National Laboratory.

The 2D diffractograms were first converted into 1-D d-spacing scans by radial integration using FIT2D software. Due to low signal to noise ratio, background subtraction was performed with FullProf

software using two baselines, one from the sample due to water and second from a blank sample with comparable exposure time and detector distance to identify the NR d-spacings. These two controls identified the prominent peaks from the protein or peptide samples. Subsequently, the background signal was subtracted for each sample spectrum to get the refined spectrum.

3.2 XRD results and discussion

Previous analysis of the cross-β-sheet structure of amyloid ribbons by X-ray Diffraction (XRD) show that the dominant diffraction peaks lie at d-spacings of ~0.47 nm and ~1 nm (5–8). ~0.47 nm correlates to the distance between monomers arranged in parallel (N-N-N-…terminus) in the β-sheet backbone due to hydrogen bonding that resembles a tape or ribbon structure, while ~1 nm stems from the distance between two β-sheets in an antiparallel steric zipper conformation with N terminus end of one sheet facing C terminus end of the second and separated by a hydrophobic interface. A previous study also investigated the amyloid-like, cross-β-sheet of amelogenin and demonstrated that the 14P2 domain has a high propensity to form β-sheets in solution (9). Subsequent solid-state Nuclear Magnetic Resonance (ssNMR) and XRD studies performed on the LRAP segment of murine amelogenin (contains 14P2 except for F replaced with L), confirmed the presence of β-sheet structure within the 14P2 domain (6). The NRs of all sequences prepared using our protocol exhibit similar diffraction patterns in bulk solution (at concentrations ≥1 mg/mL and aged for 14 days) with and without HOPG, thus confirming the absence of other periodic conformations such as α-helix (***SI Appendix*, Fig. S3**). The structures in these solutions were further characterized using in situ AFM, detailed below (*SI Appendix*, Method 3.3-3.4). Muscovite mica, a charged hydrophilic surface, was also employed to compare the effect of surface chemistry on NR orientation.

3.3. Method to identify supramolecular structures in XRD samples

Using 14P2, solutions typically used for XRD (≥1 mg/mL, pH 1.94, aged for 14 days) and those used for mineralization experiments (0.1 mg/mL, pH 1.94, aged for 14 days then diluted to 0.1 mg/mL) were deposited on muscovite mica and HOPG to determine the composition and impact of surfaces on NR assembly, orientation, and dimensions (***SI Appendix*, Fig. S4**). Identical method was used to avoid sample preparation artifacts. 30 μL of the solution was drop cast onto the substrate (HOPG or muscovite mica) and left to incubate for 1 hour in a sealed chamber. After incubation on the substrate, the excess solution was carefully removed and exchanged with pure water then immediately characterized using in situ AFM.

3.4. Supramolecular structure characterization by AFM on mica and HOPG

1 mg/ml solutions of 14P2 aged for 14 days contained NRs with thickness 3-4.5 nm when deposited on muscovite mica or HOPG (*SI Appendix*, Fig. S4 B and C). The height and width had large variations due to aggregation of the multiple ribbons and non-specific binding, but the smallest observed width was ~6 nm and height was ~ 3 nm which match the dimensions of cross-β-sheet 14P2 NRs with XRD d-spacings of ~0.47 nm and ~1 nm, previously characterized from NRs prepared in physiologically relevant conditions (9). In addition to cross-β-sheet NRs, the solution also deposited R2 NRs with thickness ~0.65 nm on HOPG or ~1.7 nm on muscovite mica seen in the background as darker features.

In contrast, 0.1 mg/ml concentration solutions diluted from 1 mg/mL aged for 14 days mainly deposited R2 NRs with thickness ~0.65 nm on HOPG or ~1.7 nm on muscovite mica (*SI Appendix*, Fig. S4 D and E). Furthermore, these NRs assembled into highly aligned islands by binding on HOPG (0001) surface epitaxially, in the three symmetrical directions, similar to other β-sheet amyloids previously reported (10). Whereas mica had no direct evidence of assembly using our protocol. As mentioned in main text, R2 NRs are generally characterized as protofibrils with β-sheet structure that have a hydrophilic and hydrophobic interface. The variation in height on mica vs. HOPG for single β-sheet NR further confirms

this (*SI Appendix*, Fig. S4 F); binding of the hydrophilic side of the single β-sheet NR likely traps water between mica and NR, thus increasing the apparent height. Whereas binding of the hydrophobic side of single β-sheet NR excludes water to create a dry interface between HOPG and NR, and the π-π interactions with HOPG may further reduce the overall height (*SI Appendix*, Fig. S14 G and J). These observations of single β-sheet NR assembly on HOPG vs. mica show HOPG surface can act as a proxy for the buried hydrophobic interface of the cross-β-sheet dimer similar to previous studies on amyloids (11, 12). Cross-β-sheet NRs and bundles were also observed at 0.1 mg/mL in other data but they were relatively low in number and easily desorbed during buffer exchange (rinsing). The above findings were observed for other sequences as well, though a detailed investigation was not performed since the XRD results demonstrated there are NRs in 1 mg/mL solutions and single β-sheet NR assembled on HOPG with 0.1 mg/mL solutions (Fig. 1 D-H).

Method 4: Molecular dynamics (MD) simulations

4.1. Molecular models and force fields

Initial conformations of the peptides (14P2, p14P2, 14P2Cterm, and p14P2Cterm) assembled in solution and on HOPG in β-sheet structures for MD simulations were built using geometric parameters from synchrotron XRD (0.47 nm d-spacing, *SI Appendix,* Fig. S3), AFM morphology (Table 1, Fig 2A, B, F, *SI Appendix* Fig. S1 D) and previous studies (9). Virtual π electrons were introduced for graphite and C and N atoms in peptide Tyrosine, Histidine and Phenylalanine side chains (***SI Appendix,* Fig. S14 A** to **D**). These π electrons (normal to the plane of the ring) permit simulation of interactions between sp^2 orbitals of the aromatic rings in the peptide and the polycyclic aromatic rings on graphite (0001) surface. Considering the protonation state at pH = 1.94, partial charge was assigned to the amino acids according to IFF and CHARMM-36. In addition, the glycine, lysine, and arginine side groups were protonated, and H_2PO_4-Ser groups were half deprotonated into HPO_4^{2-}-Ser groups consistent with the approximate pKa value of 1.5-2.0. The entire system remained charge-neutral by compensating the positive charge of peptide with chloride anions, which was identical to experimental assembly environment. For HOPG surfaces, periodic cells with full 3D periodicity were constructed as needed using supercells from a 14.760×12.782×13.600 Å unit cell (4 layered graphite with 287 C atoms and 574 virtual π electrons). For surface-MD simulations on HOPG, peptide NRs were placed on graphite (0001) with an angle of 30° from $<1\bar{1}00>$ direction as indicated by the AFM results (*SI Appendix,* Fig. S1 D). TIP3P water molecules were used as the solvent in simulations. The graphite parameters from IFF accurately reproduce lattice parameters, surface energies, and hydration energies using standard water and biomolecular force field parameters, following the principle of thermodynamic consistency (13). The parameters for aromatic groups of the peptides were from IFF and the remaining parts were from CHARMM36 (14). Compatible IFF parameters were developed using the energy expressions of class I force fields CHARMM36 (14) and CVFF (15) for compatibility, which employ a 12-6 Lennard-Jones potential and Lorentz-Berthelot (or very similar) combination rules. Specifically, the force field parameters contain the harmonic bond stretching constants $\boldsymbol{K_r}$, equilibrium bond lengths $\boldsymbol{r_0}$, equilibrium nonbond diameters $\boldsymbol{r_m}$, and the nonbond potential well depth $\boldsymbol{\varepsilon}$. The bonded parameters $\boldsymbol{K_r}$ and $\boldsymbol{r_0}$ are needed for species with predominantly covalent bonds. The energy expressions used in this work are shown as ***SI Appendix,* Equation 1**:

$$E_{pot} = \sum_{ij\ bonds} K_{r,ij}(r_{ij} - r_{0,ij})^2 + \sum_{ijk\ angles} K_{\theta,ijk}(\theta_{ijk} - \theta_{0,ijk})^2 + \sum_{\substack{ijkl\ torsions \\ (in\ plane)}} K_{\phi,ijkl}[1 + cos(n\phi_{ijkl} - \phi_{0.ijkl})] + \sum_{\substack{ijkl\ impropers \\ (out\ of\ plane)}} K_{\chi,ijkl}[1 + cos(n \cdot \chi - \chi_{0.ijkl})] + \frac{1}{4\pi\varepsilon_0\varepsilon_r}\sum_{\substack{ij\ Coulomb \\ (1,2\ and\ 1,3\ excl)}} \frac{q_i q_j}{r_{ij}} + \sum_{\substack{ij\ VDW \\ (1,2\ \ and\ 1,3\ excl)}} \varepsilon_{0,ij}\left[\left(\frac{r_{m,ij}}{r_{ij}}\right)^{12} - 2\left(\frac{r_{m,ij}}{r_{ij}}\right)^{6}\right] \quad \textbf{(1)}$$

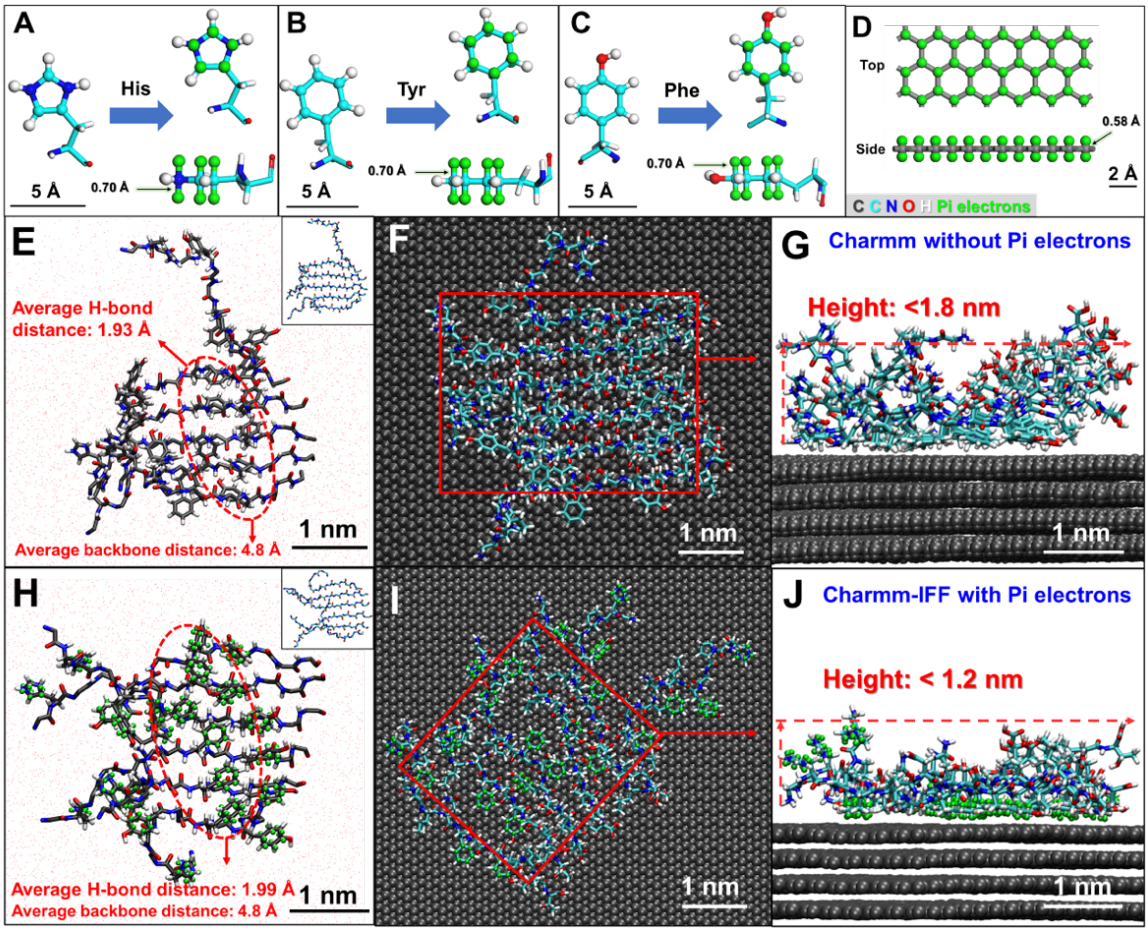


Fig. S14. Schematic and validation of the virtual π electrons introduced to the amino acids and graphite. (**A-C**) Features of the virtual π electrons on His, Tyr and Phe. (**D**) Top view and side view of a typical (0001) graphite layer with virtual π electrons on C atoms. (**E**) 6 monomer 14P2 β-sheet without π electrons at equilibrium state in solution; (**F**) Top and (**G**) side view of 14P2 without π electrons on HOPG. (**H**) 6 monomer 14P2 β-sheet with virtual π electrons on His, Tyr and Phe at equilibrium state in solution. (**I**) Top and (**J**) side view of the β-sheet including virtual π electrons on HOPG, showing height less than 1.2 nm, consistent with AFM data. Water molecules are hidden for clarity in F, G, I and J.

4.2. Settings for MD simulations

MD simulations were carried out using the Nanoscale Molecular Dynamics program (NAMD) (16). All atoms were allowed to move freely during the simulation. The isobaric-isothermal ensemble (NPT) was applied for the models with the HOPG surface, while the canonical ensemble (NVT) was employed for simulations in solutions. The pressure was set close to 101.3 kPa in the NPT ensemble, and the volume adjusted in the NVT ensemble to represent a pressure close to 101.3 kPa, to correlate with experimental conditions. The temperature was controlled at 298.15 K using the Langevin thermostat and a damping coefficient of 1 ps^{-1}. The average temperature remained within ±0.5 K of the target temperature with instantaneous fluctuations of ±10 K. A spherical cutoff of 12 Å was set for the summation of pair-wise Lennard-Jones interactions. The summation of the electrostatic interactions (Coulomb) was completed using the Particle Mesh Ewald (PME) method throughout the equilibration and production runs in high accuracy of 10^{-6}. All simulations were carried out for at least 8 ns with a timestep of 0.5 fs to ensure sufficient conformation sampling and energy evolution (17, 18). For all the energy calculations, simulations of each system were repeated three times to obtain average total energies, which were then corrected to the

exact target temperature of 298.15 K by utilizing the heat capacity of each system.

4.3. Peptide conformation in bulk solution

Pristine β-sheets (6 peptide monomers) of all peptide sequences (14P2, p14P2, 14P2Cterm and p14P2Cterm), with a backbone distance of ~0.47 nm, were simulated in water at pH 1.94 without HOPG (***SI Appendix,* Fig. S5**). For a comparative analyses of energy and stability over time for β-sheets, random coil structures of all sequences were also simulated at the same conditions. The box size was 80.000×60.000×50.000 $Å^3$ for 14P2 and p14P2, and 120.000×60.000×50.000 $Å^3$ for 14P2Cterm and p14P2Cterm. The simulations were carried out for 13 to 15 ns in the NVT ensemble to reach the equilibrium state. The obtained final configurations and total energies were then compared to understand the stability and preferences of the β-sheet structure with reference to the experimental results.

Results: During the simulations in solution without HOPG, all β-sheet conformations were found to remain relatively stable due to hydrogen bonding between the backbones (especially the ordered and closely packed central section, YINFSY) and π-π interactions between the aromatic rings (*SI Appendix,* Fig. S5 and S14). The position of aromatic rings between peptide strands was identical for conformations without and with π electrons due to similar backbone and H-bond distances (*SI Appendix,* Fig. S14 E and H). Moreover, the π-π interactions from the introduction of virtual π electrons resulted in a significant reduction in thickness of the assembled peptides on HOPG surface, clearly in better agreement with AFM data (*SI Appendix,* Fig. S14 G and J), thus validating the accuracy of IFF models and parameters. In addition, a small degree of twist was observed for all the β-sheets after simulations (*SI Appendix,* Fig. S5 A to D). The conformations of 14P2Cterm and p14P2Cterm show that their C terminus is relatively aperiodic and disordered compared to their YINFSY domain (*SI Appendix,* Fig. S5 C and D), which may account for the preference of the Cterm to adopt a random coil or flexible conformation in solution. In addition, the results also show that the Cterm of one monomer (DKTKREEVD) tends to pair with the Cterm of the adjacent monomer, likely due to electrostatic interactions among the charged side groups. Furthermore, modification of the single serine to phosphoserine had no distinct impact on the stability of the β-sheet structure (*SI Appendix,* Fig. S5 A to D).

For direct comparison of the energy, the β-sheets were assumed to be zero energy for reference within the uncertainties of 14P2: 0±4.6 kcal/mol, p14P2: 0±1.5 kcal/mol, 14P2Cterm: 0±15 kcal/mol and p14P2Cterm: 0±12 kcal/mol. Therefore, the energies of random coil conformations relative to the β-sheet structure, were 14P2: -7.6±11 kcal/mol, p14P2: 2.9±1.2 kcal/mol, 14P2Cterm: -68.2±10 kcal/mol and p14P2Cterm: -60.7±3 kcal/mol. The energy differences between random and β-sheet conformation for 14P2 and p14P2 are similar and within the error bar, which implies β-sheets can form in solution within the experimental conditions. Interestingly, 14P2Cterm and p14P2Cterm show propensity (energetically favorable) to adopt random conformation (*SI Appendix,* Fig. S9 C and D), however, this is likely related to the attraction of water molecules to the highly charged Cterm domain. Because molecular resolution AFM (Fig. 2F) and XRD (*SI Appendix,* Fig. S3) results have clearly demonstrated the formation and existence of peptide β-sheets in solution and on HOPG at concentrations from 6.26 μM to 1.6 mM. A previous study also suggested certain amyloid-like ribbon forming sequences can adopt a random-coil conformations in solution prior to nucleation and the existence of barrier associated with conformational switching to form β-sheets in solution. The energy change for protofibril (single β-sheet) formation with a similar structure (*Sup35*) suggested the energies are unlikely to be strongly negative since the amyloid-like fibrils are stabilized by protein concentration, formation of cross-β-sheets and hydrogen bonded backbones (8). But once a nucleus forms, high concentrations of protein drive the formation and contribute to a large barrier to dissolution of the protofibrils.

4.4 Adsorption energy of peptides on HOPG

For 14P2 and p14P2 the box size was 59.040×38.347×80.000 Å^3, whereas the box size for 14P2Cterm and p14P2Cterm was 88.560×63.912×80.000 Å^3. Two systems with the same atom number were established for each peptide sequence, namely the surface system and the away system. Typically, for the surface system, a fully extended β-strand was placed on the HOPG surface, while the peptide β-strand was floating in the solution and ~35 Å away from the surface in the away system. The detailed simulation settings were the same as described in Method 4.2. The E_{ads} was calculated using the following **Equation 2**:

$$E_{ads} = E_{away} - E_{sur} \quad (2)$$

where E_{ads} is the adsorption energy, E_{away} is the total energy of the away system, and E_{sur} represents the total energy of the surface system.

Results: The calculated E_{ads} are -23.9±5 kcal/mol for 14P2 (***SI Appendix,* Fig. S6 A**), -23.9±2 kcal/mol for p14P2, -26.9±8 kcal/mol for 14P2Cterm and -25.9±8 kcal/mol for p14P2Cterm, respectively. The negative E_{ads} implies peptide binding onto HOPG is thermodynamically favorable, but weak enough to allow for reversible binding of β-strands to the surface. This explains why peptides can self-assemble on the HOPG surface. The other contribution towards energy could be monomer-monomer binding energies within the β-sheet structure as measured previously by Dynamic Force Spectroscopy for 14P2 (9), but highly accurate structure and energy calculations by MD for a β-sheet on HOPG would need at least 100 monomers (3). The E_{ads} for β-sheets on HOPG was not carried out since the calculations of 100 monomers for our four peptide systems with 14 to 23 amino acids per peptide would require a monumental amount of computational resources. However, these calculations were performed previously for a 7 amino acid MoS_2 binding peptide (MoSBP1) that forms NRs on MoS_2 and HOPG. 14P2 and MoSBP1 (YSATFTY) have some sequence homology in the YINFSY domain. In MoSBP1 system, the E_{ads} for binding of one monomer on MoS_2 is -96 kcal/mol and -18 kcal/mol for ~100 monomer system. These energies corresponded to strong binding, consistent with in situ AFM experiments. Therefore, the results suggest similar trends for 14P2; E_{ads} for 14P2 β-sheets may be smaller in magnitude compared to monomers yet have sufficiently strong binding to HOPG.

4.5. Assembly and stability of peptide β-sheets on HOPG

14P2 was used as a model sequence to test the assembly and stability of the peptide β-sheet on HOPG. More than 50 starting conformations for 14P2 were simulated and ruled out due to their instability. This included dimers, parallel alignment (non-β-sheet), anti-parallel alignment (β-sheet and non-β-sheet), vertically stacked chains (with π-π stacking on aromatic rings), zigzag-folded backbones, and others. A 14P2 β-sheet consisting of 21 monomers was placed on HOPG (Fig. 2B) to represent the smallest NR length (~10 nm) observed on HOPG. Another system with a shorter β-sheet consisting of 6 monomers was also used for comparison to other sequences. The box size for 21-monomer system of 14P2 was 73.800×89.4768×50.000 Å^3. The box sizes for 6-monomer systems were 88.560×115.041×58.600 Å^3 for 14P2, 59.040×76.694×45.000 Å^3 for p14P2, and 88.560×102.250×45.000 Å^3 for 14P2Cterm and p14P2Cterm. The β-sheets were placed on the HOPG (0001) plane at 30° to <1$\bar{1}$00> direction. The simulation settings were described in Method 4.2.

Results: For the 21-monomer system of 14P2, the β-sheet conformation (Fig. 2E) remained stable with a d-spacing around 4.8 Å between the backbones after 13.5 ns simulation in NPT, indicating that long β-sheets on HOPG were indeed stable over time and consistent with the experimental results (Fig. 2B). As discussed in the main text, the simulated model coincides well with the AFM morphology. On the other hand, the β-sheet conformations consisting of 6 monomers, including 14P2, (***SI Appendix,* Fig. S6 B to E**) also remained reasonably stable during the 20 ns simulation in NPT with a d-spacing around 4.8 Å between

some sections of the backbone, while some monomers tend to separate and move away, thereby widening the gap to more than 1 nm, along with separation at the C and N terminus at the end of the simulation, especially for 14P2Cterm and p14P2Cterm whose long and charged Cterm domain tend to pair with the Cterm from an adjacent monomer in a single β-sheet and twist away from the 14P2 domain similar to solution simulations. The widening and twisting are likely caused by the short length of the ribbon as this effect was negligible for the 21-monomer system. We also observed 0.47-0.52 nm distances between the glutamic acid residues and periodic π-π interactions along the arm-chair direction <10$\bar{1}$0> which explains the 3-fold alignment on graphite surface. π-π stacking of the aromatic groups may also explain multiple layers of R2 NRs, wherein NR in the second layer lies in the gap region between two NRs in the bottom layer or directly on an NR while maintaining β-sheet structure. To conclude, only parallel (N to N terminus) aligned β-sheet structures are stable on HOPG and approaches a ribbon-like morphology, especially when the β-sheets increase in length.

Method 5: Vertical growth rate (*V*) analysis and calculation of supersaturation with respect to ACP

Equilibrium solubility values of ACP (K_{sp}) from most studies in literature and VISUAL MINTEQ calculations indicate that solutions used here are undersaturated with respect to all types of ACP. For such cases, ACP could nucleate, but not continue to grow, unless the previously reported solubilities do not reflect our solution conditions. Therefore, the relationship between growth velocity (V) and supersaturation in ***SI Appendix,* Equation 3** is used to extract the apparent solubilities, where ***IAP*** is ionic activity product for previously reported $(Ca_2(HPO_4)_3)^{2-}$ pre-nucleation cluster given by $(a(Ca^{2+})^2 \bullet a(HPO_4^{2-})^3)$, ω is volume of growth unit (ACP: 5 x 10^{-29} m^3), n= 5, a([x]) is ionic activity (in M), and β is the kinetic coefficient, was exploited to plot growth velocity vs $IAP^{1/5}$ (Fig. 3G).

$$V = \omega \beta (IAP^{1/5} - K_{sp}^{1/5}) \quad (3)$$

Ionic activity product of all three solutions was determined from free ion concentrations given by Visual MINTEQ for each solution conditions (Table 1 and ***SI Appendix,* Table S1**). Growth rates for each sequence and saturation were obtained from average slopes of linear fits to heights vs elapsed time for multiple particles (***SI Appendix,* Table S2)**. Then the relationship between growth rate vs $IAP^{1/5}$ was linearly fit with an error-weight and extrapolated to zero growth rate to determine IAP at which no growth occurs (V= 0, x-axis intercept), i.e., K_{sp}. From these measurements and *SI Appendix,* Equation 3, we found that solubility of ACP is 1.155 x 10^{-15} M^5 which lies between Visual MINTEQ (8.42 x 10^{-12} M^5) and our previous bulk solution measurements (8.03 x 10^{-17} M^5) (19). Using K_{sp} from growth rates extrapolated to zero or from previous measurements, supersaturations were calculated using ***SI Appendix,* Equation 4**. Thus, the supersaturations calculated using K_{sp} of ACP reported previously (19) gave σ_{ACP} = 0.57, 0.67 and 0.75. Whereas supersaturations calculated using growth rates are σ_{ACP} = 0.04; σ_{ACP} = 0.138; and σ_{ACP} = 0.221 (***SI Appendix,* Table S3**), which were primarily for discussions as they are specific to our AFM experiments.

$$\sigma_{ACP} = \frac{ln\, IAP - ln\, K_{sp}}{n} \quad (4)$$

Method 6: Mineralization control by Scanning Electron Microscopy

HOPG flakes coated with 14P2 NRs (identical to substrates used for AFM-based nucleation experiments) were placed in supersaturated calcium phosphate solution (σ_{ACP} = 0.221, identical to those used for in situ AFM experiments) for 1 hour. A control sample, a freshly cleaved HOPG flake, was also incubated in the σ_{ACP} = 0.221 supersaturated solution with a separate container for 1 hour. These solutions did not have any concomitant precipitation in the solution without the flakes for at least 3 hrs, measured by changes in pH and turbidity. The flakes were removed, dried and characterized with Scanning Electron

Microscopy (FEI Helios). The control sample had no mineral precipitates, whereas the flakes functionalized with 14P2 had spherulite-like mineral that resembled biomimetic apatite (***SI Appendix,* Fig. S7**).

Method 7: Sample preparation for Transmission Electron Microscopy (TEM) and Energy Dispersive Spectroscopy X-ray Spectroscopy (EDS)

To characterize the phase of the mineral nuclei formed on AFM substrates, sample preparation was kept identical to AFM substrates but using graphene grids. 5 µl of the 0.1 mg/mL protein or peptide solution was incubated on the grid (Pelco 3–5-layer graphene on lacey carbon 300 mesh copper grids, Ted Pella, USA) for 30 minutes. The grid was then rinsed to remove excess protein by immersion in 1 mM HCl followed by immersion in water. For mineralization, the protein or peptide functionalized graphene grids were incubated in a vial with freshly prepared supersaturated calcium phosphate solution (1 mM $CaCl_2$ and 9.5 mM KH_2PO_4 after mixing at pH 7.4 and 25 ºC). Each mineralization experiment was carried out for the required time in static condition (20 minutes or 180 minutes), then rapidly quenched in toluene. The quenched grid was then immediately dried in a vacuum desiccator prior to imaging. All grids were imaged with FEI Tecnai 12 TEM (Thermo Fisher Scientific, USA) or FEI Tecnai 20 TEM on the Titan-ES.

Method 8: High-resolution TEM (HRTEM) and Selected Area Electron Diffraction (SAED) of calcium phosphate particles at 20 min and 3hr

Particles mineralized on the NR-functionalized grids were confirmed as calcium phosphate through EDS (***SI Appendix,* Fig. S8**). To identify mineral phase, HRTEM and SAED were performed on the particles after 20 min mineralization (***SI Appendix,* Fig. S9 A** to **E**). FFT of HRTEM on several particle and all sequences revealed that the initial particles were amorphous from the absence of lattice fringes corresponding to crystalline calcium phosphate (***SI Appendix,* Fig. S10 A**). The analysis shows 0.28 and 0.24 nm spacings that match the C-C distance of graphene/graphite arm-chair $<10\bar{1}0>$ and zig-zag $<1\bar{1}00>$ directions (20) while ~0.17 nm only corresponds to the (004) plane of graphite or 3-5 layer graphene. SAED shows a broad diffraction band where d-spacing for crystalline calcium phosphate are expected (21). Absence of a sharp band or speckles in SAED suggests absence of nanocrystals smaller than 10 nm. The phase transformation of ACP particles was characterized with samples incubated for 180 min. TEM of these samples revealed large mineral films with fiber- or plate-shaped morphology (*SI Appendix,* Fig. S9 F to J), identical to in situ AFM observations in ***SI Appendix,* Fig. S11**, while SAED shows several d-spacings that corresponded to polycrystalline OCP and AP.

To further assess the crystallinity at both time points, radial integration of the SAED patterns in *SI Appendix,* Fig. S9 was performed for all sequences at 20 min and compared with those of rH174 and p14P2Cterm at 180 min, which are representative of crystalline calcium phosphate transformed from ACP (***SI Appendix,* Fig. S10 B**). Apart from the d-spacings of graphite, the analysis clearly shows the absence of any sharp peaks related to crystalline calcium phosphate phases at 20 min. Whereas 180 min samples show sharp peaks at a d spacing of 0.26-0.33 nm. Samples of other sequences show similar diffraction patterns at 180 min. These analyses using HRTEM and SAED, as well as AFM (*SI* Appendix, Fig. S11), are consistent with our previous work (29) in which our AFM-based nucleation study for calcium phosphate was developed and supported by Low Dose SAED in CryoEM and other spectroscopic and chemically analytical methods (FTIR, titration, zeta potential, and SAXS). The formation and lifetime of initially formed ACP before phase transformation was also monitored by in situ Raman in another study by us on ACP nucleation using Amel nanospheres (22). Moreover, the crystalline properties of particles at early and late stage monitored by in situ AFM (SI Appendix, Fig S11) have also been evaluated by ex situ TEM in our two previous studies. To summarize, the data from all these studies are consistent — the first formed particles show no evidence of crystallinity by any method, but they transform over time to one of the

crystalline phases (usually OCP) at which point they show crystallinity by all methods, regardless of whether the particle size is the same as that of the initially formed amorphous particles. In addition, a weak ~4.7 Å d-spacing was observed in SAEDs which matches the periodicity of the NR β-sheet structure.

Method 9: Phase transformation observed by in situ AFM

Mineralization experiments on NRs of all sequences using $\sigma_{ACP} = 0.221$ solutions were performed for longer durations to evaluate the changes in morphology and mineral phase of the ACP particles at different length and time scales. These AFM experiments were performed in two ways to ensure the results were free of instrument artifacts. First method involves fast, repeated scan of the surface (*SI Appendix,* Fig. S11 A to D). Second, uses a slow scan rate on a large area to reveal the transformation in a single frame (*SI Appendix,* Fig. S11 E).

For all sequences, the initial hemispherical particles grew in diameter and height. Eventually, the particles coalesce and become fiber- and plate-like in shape. Growth ceased when the imaging medium was switched to water (*SI Appendix,* Fig. S11 E inset) and absence of mineral dissolution after several minutes indicated that the particles were no longer amorphous but indeed crystalline. The results presented are consistent with previous reports on nucleation of calcium phosphate on collagen (19) and amelogenin nanospheres (30). The order of estimated ACP lifetimes observed on each sequence is 14P2Cterm: ~39 min < p14P2Cterm: 56.63min < 14P2: 56.71 min < rH174: ~125.71 min < p14P2: ~218.9 min. Interestingly, we observed that growth rates have nearly the same order except for rH174. This suggests a critical particle size must be reached before transformation could take place.

Method 10: Calculation of effective interfacial energies for nucleation of ACP

To analyze the nucleation rate data for interfacial energy in each case of peptides, we assume nucleation occurs through ion-by-ion mode at our supersaturations, i.e., classical nucleation theory given by ***SI Appendix,* Equation 5**, previously shown to effectively describe ACP nucleation kinetics (19). Briefly**,** the change in free energy required to create a hemispherical nuclei of radius $\boldsymbol{R}$ on a substrate is given by ***SI Appendix,* Equation 6**. Subscripts ***ML***, ***MS*** *and* ***LS*** refer to mineral-liquid, mineral-substrate and liquid-substrate interfaces, respectively. The critical radius $\boldsymbol{R_c}$ of nucleus is given by ***SI Appendix,* Equation 7** where $\boldsymbol{\alpha_{eff}}$ is given by ***SI Appendix,* Equation 8.** Substituting equation 7 in equation 6, then using equation 8, to simplify $\boldsymbol{\Delta G}$ in *SI Appendix,* Equation 6, gives ***SI Appendix,* Equation 9**. Substituting equation 9 back into equation 5, gives ***SI Appendix,* Equation 10**, and taking natural logarithm on both sides gives ***SI Appendix,* Equation 11.**

The volume of growth unit ($\boldsymbol{\omega}$**,** defined as the volume of complex/number of ions) of ACP ($Ca_2(HPO_4)_3)^{2-}$) used in this study is 5 x 10^{-29} m^3, which is an intermediate value between the volume of a Posner cluster with composition $Ca_9(PO_4)_6$ (2.993 x 10^{-29} m^3) and disk-like structure with volume of 9.5 x 10^{-29} m^3 that is the building block for OCP and AP (19). $\boldsymbol{\alpha_{ACP}}$, the effective interfacial energy of the nucleating phase, can be determined from plots of nucleation rates at different supersaturation using the classical nucleation theory (*SI Appendix,* Equation 11). The nucleation rates are obtained from nucleation events per μm^2 versus time (in seconds) then converted to $m^{-2}\,s^{-1}$ (***SI Appendix,* Table S4**) for plot of $\boldsymbol{\ln (J_o)}$ versus $\boldsymbol{\sigma_{ACP}}^{\boldsymbol{-2}}$ (Fig. 3 I). The values of α_{eff} obtained using $\boldsymbol{K_{sp}}$ from growth rates and previous report of $\boldsymbol{K_{sp}}$ from solution measurements are shown in ***SI Appendix,* Table 5** and ***SI Appendix,* Table 6**, respectively. The theoretical value of $\boldsymbol{\alpha_{ACP}}$ in bulk solution was reported as 150 mJ m^{-2}, and nominal value of 100 mJ m^{-2} for comparison of experimental data. The value of the equilibrium solubility values of ACP ($\boldsymbol{K_{sp}}$) affects the supersaturation, and thus the interfacial energies. However, the relative ratios of the energies for each sequence when compared to the other will remain unchanged for identical experimental conditions. Therefore, the ratios of the energies with respect to the basic 14P2 sequence without any functional group

modifications is used for comparison. The interfacial energy and kinetic pre-factor *A* ratios with respect to 14P2 in the analysis reflect $\alpha_{\text{ACP of Sequence}}/\alpha_{\text{ACP of 14P2}}$ and $A_{\text{Sequence}}/A_{14P2}$, respectively.

$$J_o = A.\, e^{\left(\frac{-\Delta G_c}{kT}\right)} \quad (5)$$

$$\Delta G = -\frac{2\pi R^3 kT\sigma}{3\omega} + 2\pi R^2 \alpha_{ML} + \pi R^2(\alpha_{MS} - \alpha_{LS}) \quad (6)$$

$$R_c = \frac{2\omega\alpha_{eff}}{kT\sigma} \quad (7)$$

$$\alpha_{eff} = \alpha_{ML} + \frac{1}{2}(\alpha_{MS} - \alpha_{LS}) \quad (8)$$

$$\Delta G_c = \frac{8\pi\omega^2\alpha_{eff}^3}{3(kT\sigma)^2} \quad (9)$$

$$J_o = Ae^{\left(\frac{-8\pi\omega^2\alpha_{eff}^3}{3(kT\sigma)^2}\right)} \quad (10)$$

$$\ln(J_o) = \ln(A) - \frac{B}{\sigma^2} \quad \text{where } B = \frac{8\pi\omega^2\alpha_{eff}^3}{3(kT)^3} \quad (11)$$

Method 11: Design of mineral-free self-assembly method

Prior *in vitro* studies on recombinant full-length amelogenin (rH174) and 14P2 peptide dissolved the lyophilized proteins/peptides in 1 mM HCl (pH 3.1), then Ca and PO_4 salts were added to assemble cross-β-sheet NRs at pH 4.5-5.5 (9). These initially added salts precipitate calcium phosphate in solution and the method produces a range of aggregated higher order ribbon structures. Both outcomes interfere with careful quantitative measurements of calcium phosphate nucleation rates on a single, identical interface. However, β-sheet interactions are intrinsic to the sequence and conformation compared to other interactions such as π-π bonding, van der Waal's forces or hydrophobicity that are involved in the formation of cross-β-sheet and aggregates. Therefore, single β-sheet protofibrils were the rational choice to obtain consistent interfaces. Hence, the protocol was redesigned to exclude Ca and PO_4 salts and isolate single β-sheets on a charge neutral surface that can coat 100% of our observed area under AFM (at least 20 x 20 μm scan area), which is also transferable to other sequences. This was performed by lowering pH from 3.1 to 1.94 and lowering concentrations from 1 mg/mL to 0.1 mg/mL for self-assembly on HOPG (8, 23–25). 14P2 was used as a model sequence for a detailed, high-throughput study due to commercial availability, prior data on structure of NRs, and self-assembly kinetics in the order of minutes relative to rH174 that takes several hours (9, 26). Most importantly, 14P2 provided a better map of the NR backbone structure since in situ AFM is unable to clearly resolve the packing of monomers in NR with long hydrophilic functional groups or flexible tail end domains (phosphoserine and Cterm domain). As observed for Cterm sequences compared to 14P2, the flexible tail likely makes the surface appear disordered. This same effect was observed in highly crystalline sheets of peptide-mimetic molecules possessing hydrophilic headgroups; despite the fact that both XRD and TEM demonstrate high crystallinity, the surface appears disordered (27).

Bulk σ_{AP}	Species	Reagent concentration (mM)	Free ion concentration (M)	Activity (M)	Exponents for ACP	Ionic Strength (M)
3.37	Ca^{2+}	1	4.94×10^{-4}	2.826×10^{-4}	2	0.0203
	HPO_4^{2-}	-	6.164×10^{-3}	3.526×10^{-3}	3	
	PO_4^{3-}	9.5	1.312×10^{-7}	3.735×10^{-8}	-	
3.28	Ca^{2+}	0.875	4.462×10^{-4}	2.618×10^{-4}	2	0.0182
	HPO_4^{2-}	-	5.499×10^{-3}	3.227×10^{-3}	3	
	PO_4^{3-}	8.475	1.134×10^{-7}	3.417×10^{-8}	-	
3.19	Ca^{2+}	0.75	3.963×10^{-4}	2.39×10^{-4}	2	0.016
	HPO_4^{2-}	-	4.8336×10^{-3}	2.916×10^{-3}	3	
	PO_4^{3-}	7.45	9.63×10^{-8}	3.089×10^{-8}	-	

Table S1. Species relevant to ACP and activities calculated by Visual MINTEQ for solutions supersaturated with respect to hydroxyapatite (σ_{AP}) at pH 7.4, with stoichiometry (log (activity of Ca^{2+}/activity of PO_4^{3-})) = 3.87. The phosphate reagent concentration (column 3) is the total concentration of phosphate species that includes HPO_4^{2-}, PO_4^{3-}, $H_2PO_4^-$ and all other phosphate containing species in solution.

		14P2Cterm	14P2	p14P2Cterm	p14P2	rH174
σ_{ACP}	$IAP^{1/5}$ (M)	V (nm/sec)	V (nm/sec)	V (nm/sec)	V (nm/sec)	V (nm/sec)
0.04	0.00107	1.89 ± 0.01	1.18 ± 0.04	1.18 ± 0.02	0.5 ± 0.02	1.05 ± 0.07
0.138	0.00118	5.24 ± 0.02	4.17 ± 0.08	3.86 ± 0.09	1.74 ± 0.05	4.3 ± 0.05
0.221	0.00128	10.16 ± 0.08	6.91 ± 1.96	6.93 ± 1.03	3.01 ± 0.06	6.78 ± 1.52
		14P2Cterm	14P2	p14P2Cterm	p14P2	rH174
Intercept ($K_{sp}^{1/5}$)		1.0202×10^{-3}	1.0265×10^{-3}	1.0264×10^{-3}	1.0284×10^{-3}	1.0284×10^{-3}
K_{sp} (M^5)		1.1052×10^{-15}	1.1395×10^{-15}	1.1392×10^{-15}	1.1501×10^{-15}	1.1503×10^{-15}
β		3.8×10^{-5}	2.72×10^{-5}	2.68×10^{-5}	1.18×10^{-5}	2.85×10^{-5}

Table S2. Vertical growth rates (*V*) of particles for various sequences at different supersaturation and intercepts and slopes obtained from error-weighted linear fits using *SI Appendix,* Equation 3. For detailed method of analysis refer to *SI Appendix,* Method 5. Error is std. dev., for sample size refer to Statistics (main text)

			K_{sp} from extrapolation of growth rates to zero			K_{sp} from Habraken et al		
Bulk σ_{AP}	Ln IAP	n	Ln K_{sp}	σ_{ACP}	$1/\sigma^2_{ACP}$	Ln K_{sp}	σ_{ACP}	$1/\sigma^2_{ACP}$
3.37	-33.285	5	-34.394	0.221	19.74	-37.06	0.75	1.75
3.28	-33.705	5	-34.394	0.138	50.15	-37.06	0.67	2.22
3.19	-34.19	5	-34.394	0.04	513.96	-37.06	0.57	3.03

Table S3. Supersaturations with respect to ACP calculated using equilibrium constant K_{sp} of $Ca_2(HPO_4)_3^{2-}$ from extrapolated growth rates in *SI Appendix,* Table S2 and from a previous study (19) at a different ionic strength (0.244 M). For detailed method of analysis refer to *SI Appendix,* Method 5.

		14P2	p14P2	14P2Cterm	p14P2Cterm	rH174
σ_{ACP}	σ_{ACP}^{-2}	Ln J_0	Ln J_0	Ln J_0	Ln J_0	Ln J_0
0.04	597.137	22.419 ± 0.021	23.408 ± 0.153	22.764 ± 0.164	24.165 ± 0.055	21.448 ± 0.011
0.138	52.502	23.075 ± 0.208	26.064 ± 0.162	23.089 ± 0.475	26.129 ± 0.323	22.447 ± 0.133
0.221	20.318	23.084 ± 0.313	25.362 ± 0.037	23.374 ± 0.179	26.021 ± 0.327	24.511 ± 0.180

Table S4. Nucleation rates (J_o, in nuclei $m^{-2}s^{-1}$) at different supersaturations used in calculations for interfacial energies of ACP nucleation on various sequences. Error is Std. dev., for sample size refer to Statistics (main text)

Sequence	Net charge per monomer at pH 7.4	Net charge relative to 14P2	Ln(*A*)	Ratios: $A_{sequence}$ / A_{14P2}	α_{eff} (mJ.m^{-2})	Ratios: $\alpha_{ACP\ of\ Sequence}/\alpha_{ACP\ of\ 14P2}$
14P2	-1.4	-	23.12 ± 0.01	-	1.58 ± 0.02	-
14P2Cterm	-2.4	1.71	23.27 ± 0.14	1.16 ± 0.18	1.41 ± 0.23	0.90 ± 0.14
rH174	-2.492	1.78	23.64 ± 1.04	1.69 ± 1.80	2.32 ± 0.73	1.47 ± 0.34
p14P2	-3.16	2.26	25.85 ± 0.45	15.28 ± 7.15	2.38 ± 0.26	1.51 ± 0.17
p14P2Cterm	-4.166	2.98	26.19 ± 0.12	21.62 ± 2.88	2.24 ± 0.07	1.42 ± 0.06

Table S5. Net charge, ln(*A*) and α_{eff} obtained from linear fits to nucleation rate data using Equation S11 and supersaturations calculated from extrapolated growth rates. Ratios of kinetic pre-factor *A* ($A_{Sequence}/A_{14P2}$), effective interfacial energy with respect to ACP α_{ACP} ($\alpha_{ACP\ of\ Sequence}/\alpha_{ACP\ of\ 14P2}$), and charge for each sequence with respect to 14P2 from analysis. For detailed method of analysis refer to *SI Appendix,* Method 10. Error is Std. error (range)

Sequence	Net charge per monomer at pH 7.4	Net charge relative to 14P2	Ln(*A*)	Ratios: $A_{sequence}$ / A_{14P2}	α_{eff} (mJ.m^{-2})	Ratios: $\alpha_{ACP\ of\ Sequence}/\alpha_{ACP\ of\ 14P2}$
14P2	-1.4	-	24.15 ± 0.49	-	12.23 ± 0.20	-
14P2Cterm	-2.4	1.71	28.97 ± 2.98	1.02 ± 0.43	11.57 ± 1.24	0.95 ± 0.13
rH174	-2.492	1.78	28.08 ± 2.00	51.19 ± 16.79	19.57 ± 0.83	1.60 ± 0.37
p14P2	-3.16	2.26	24.17 ± 0.13	123.86 ± 40.70	17.87 ± 0.06	1.46 ± 0.58
p14P2Cterm	-4.166	2.98	29.08 ± 1.65	138.49 ± 77.97	17.29 ± 0.69	1.41 ± 0.37

Table S6. Net charge, ln(*A*) and α_{eff} obtained from linear fits to nucleation rate data using Equation S11 and supersaturations calculated from a previous study (19). Ratios of kinetic pre-factor *A* ($A_{Sequence}/A_{14P2}$), effective interfacial energy with respect to ACP α_{ACP} ($\alpha_{ACP\ of\ Sequence}/\alpha_{ACP\ of\ 14P2}$), and charge for each sequence with respect to 14P2 from analysis. For detailed method of analysis refer to *SI Appendix,* Method 10. Error is Std. error (range)

References

1. A. Cerreta, *et al.*, FM-AFM constant height imaging and force curves: high resolution study of DNA-tip interactions. *Journal of molecular recognition : JMR* **25**, 486–493 (2012).
2. C.-Y. Lai, S. Santos, M. Chiesa, General interpretation and theory of apparent height in dynamic atomic force microscopy. *RSC Advances* **5**, 80069–80075 (2015).
3. J. Chen, *et al.*, Building two-dimensional materials one row at a time: Avoiding the nucleation barrier. *Science* **362**, 1135–1139 (2018).
4. L. Sun, *et al.*, Chiral Recognition of Self-Assembled Peptides on MoS2 via Lattice Matching. *Langmuir* **37**, 8696–8704 (2021).
5. S. A. Engelberth, *et al.*, Progression of Self-Assembly of Amelogenin Protein Supramolecular Structures in Simulated Enamel Fluid. *Biomacromolecules* **19**, 3917–3924 (2018).
6. C.-W. Ma, J. Zhang, X.-Q. Dong, J.-X. Lu, Amyloid structure of high-order assembly of Leucine-rich amelogenin revealed by solid-state NMR. *Journal of Structural Biology* **206**, 29–35 (2019).
7. B. Sanii, O. Martinez-Avila, C. Simpliciano, R. N. Zuckermann, S. Habelitz, Matching 4.7-Å XRD Spacing in Amelogenin Nanoribbons and Enamel Matrix. *Journal of Dental Research* **93**, 918–922 (2014).
8. R. Nelson, *et al.*, Structure of the cross-β spine of amyloid-like fibrils. *Nature* **435**, 773–778 (2005).
9. K. M. M. Carneiro, *et al.*, Amyloid-like ribbons of amelogenins in enamel mineralization. *Scientific Reports* **6**, 1–11 (2016).
10. N. Li, *et al.*, Graphite-Templated Amyloid Nanostructures Formed by a Potential Pentapeptide Inhibitor for Alzheimer's Disease: A Combined Study of Real-Time Atomic Force Microscopy and Molecular Dynamics Simulations. *Langmuir* **33**, 6647–6656 (2017).
11. T. Kowalewski, D. M. Holtzman, In situ atomic force microscopy study of Alzheimer's beta-amyloid peptide on different substrates: new insights into mechanism of beta-sheet formation. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 3688–93 (1999).
12. M. Arimon, *et al.*, Fine structure study of Aβ 1–42 fibrillogenesis with atomic force microscopy. *The FASEB Journal* **19**, 1344–1346 (2005).
13. C. Pramanik, J. R. Gissinger, S. Kumar, H. Heinz, Carbon Nanotube Dispersion in Solvents and Polymer Solutions: Mechanisms, Assembly, and Preferences. *ACS Nano* **11**, 12805–12816 (2017).
14. J. Huang, A. D. MacKerell, CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *Journal of computational chemistry* **34**, 2135–45 (2013).
15. P. Dauber-Osguthorpe, *et al.*, Structure and energetics of ligand binding to proteins: Escherichia coli dihydrofolate reductase-trimethoprim, a drug-receptor system. *Proteins* **4**, 31–47 (1988).
16. J. C. Phillips, *et al.*, Scalable molecular dynamics with NAMD. *Journal of computational chemistry* **26**, 1781–802 (2005).
17. J. M. Gottfried, E. K. Vestergaard, P. Bera, C. T. Campbell, Heat of adsorption of naphthalene on Pt(111) measured by adsorption calorimetry. *The journal of physical chemistry. B* **110**, 17539–45 (2006).
18. Hyeran Ihm, Henry M. Ajo, J. M. Gottfried, and P. Bera, C. T. Campbell*, Calorimetric Measurement of the Heat of Adsorption of Benzene on Pt(111)† (2004) https:/doi.org/10.1021/JP040159O.
19. W. J. E. M. Habraken, *et al.*, Ion-association complexes unite classical and non-classical theories for the biomimetic nucleation of calcium phosphate. *Nature Communications* **4**, 1507 (2013).
20. C. Cong, K. Li, X. X. Zhang, T. Yu, Visualization of arrangements of carbon atoms in graphene layers by Raman mapping and atomic-resolution TEM. *Scientific Reports* **3**, 1195 (2013).
21. J. Tao, H. Pan, Y. Zeng, R. Xu, R. Tang, Roles of amorphous calcium phosphate and biological additives in the assembly of hydroxyapatite nanoparticles. *Journal of Physical Chemistry B* **111**, 13410–13418 (2007).
22. J. Tao, *et al.*, Control of Calcium Phosphate Nucleation and Transformation through Interactions of Enamelin and Amelogenin Exhibits the "goldilocks Effect." *Crystal Growth and Design* **18**, 7391–7400 (2018).
23. Y. Wang, *et al.*, Formation of Amyloid Fibrils In Vitro from Partially Unfolded Intermediates of Human γC-Crystallin. *Investigative Ophthalmology & Visual Science* **51**, 672–678 (2010).
24. J. D. Harper, S. S. Wong, C. M. Lieber, P. T. Lansbury, Assembly of Aβ Amyloid Protofibrils: An in Vitro Model for a Possible Early Event in Alzheimer's Disease. *Biochemistry* **38**, 8972–8980 (1999).
25. L. C. Serpell, Alzheimer's amyloid fibrils: structure and assembly. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* **1502**, 16–30 (2000).
26. S. A. Engelberth, *et al.*, Progression of Self-Assembly of Amelogenin Protein Supramolecular Structures in Simulated Enamel Fluid. *Biomacromolecules* **19**, 3917–3924 (2018).
27. H. Jin, *et al.*, Highly stable and self-repairing membrane-mimetic 2D nanomaterials assembled from lipid-like peptoids. *Nature Communications* **7**, 12252 (2016).