Supporting Information (SI)

Direct observation of kinetic traps associated with structural transformations leading to multiple pathways during S-layer assembly

Seong-Ho Shin*[1,3,5], Sungwook Chung*[1,2], Babak Sanii[1,3], Luis R. Comolli[4], Carolyn R. Bertozzi[§3,5] and James J. De Yoreo[§1,3]

[1]*Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, CA 94720*

[2]*Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720*

[3]*Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720*

[4]*Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720*

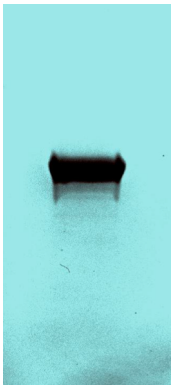
[5]*Department of Chemistry, University of California, Berkeley, Berkeley, CA 94720*

* These authors contributed equally.

§ To whom correspondence should be addressed. E-mail: jjdeyoreo@lbl.gov, crbertozzi@lbl.gov

Contents

Figure S1. Native polyacrylamide gel electrophoretic (PAGE) analysis of the purified monomeric S-layer protein.



The purified monomeric SbpA, ~6 µg was mixed with Native gel sample buffer (Bio-Rad Laboratories, 161-0738) and then the sample was run in 7 % Tris-Acetate gel (Bio-Rad Laboratories, 345-0135) with Tris-Glycine buffer (Bio-Rad Laboratories, 161-0734). The gel was stained in Imperial Protein Stain solution (Thermo Scientific, 24615) and destained.

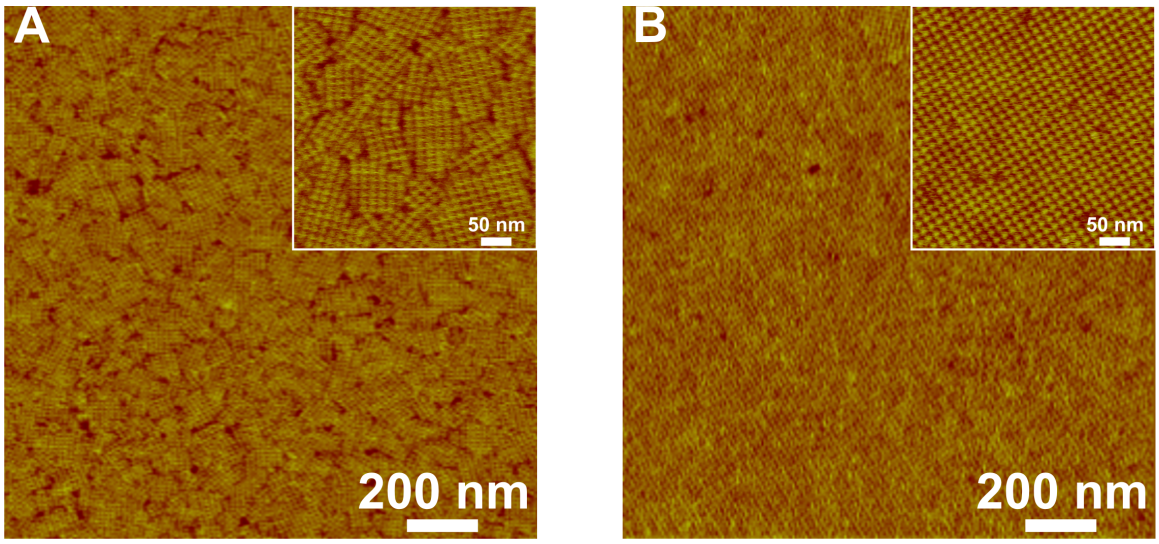


Figure S2. AFM height images of S-layer assembly on different surfaces. (*A*) AFM height image of crystalline S-layers grown on supported lipid bilayer (SLB) showing only single phase of randomly oriented crystalline domains. SLB of 1-palmitoyl-2-oleoyl-sn-glycero-3 phosphocholine (POPC) lipid were prepared on mica surfaces. Wild type (WT) SbpA (C_P = 140 µg/mL) was incubated for 3 hr at 4 °C. (*B*) AFM height image of crystalline S-layer grown on poly-Lysine (Poly-L) polymer coated mica surfaces showing a single crystalline domain with a few point defects over the length scale of ~ 1 µm. Freshly peeled mica was treated with 20 µL, aqueous poly-L solution (0.1% w/v, Ted Pella, Inc) for 30 seconds and rinsed with double distilled water. WT SbpA (C_P = 41 µg/mL) was incubated for 4.5 hr at 25 °C. In both experiments, the same buffer (10 mM Tris pH 7.2, 50 mM $CaCl_2$, 100 mM NaCl) was used for the growth.

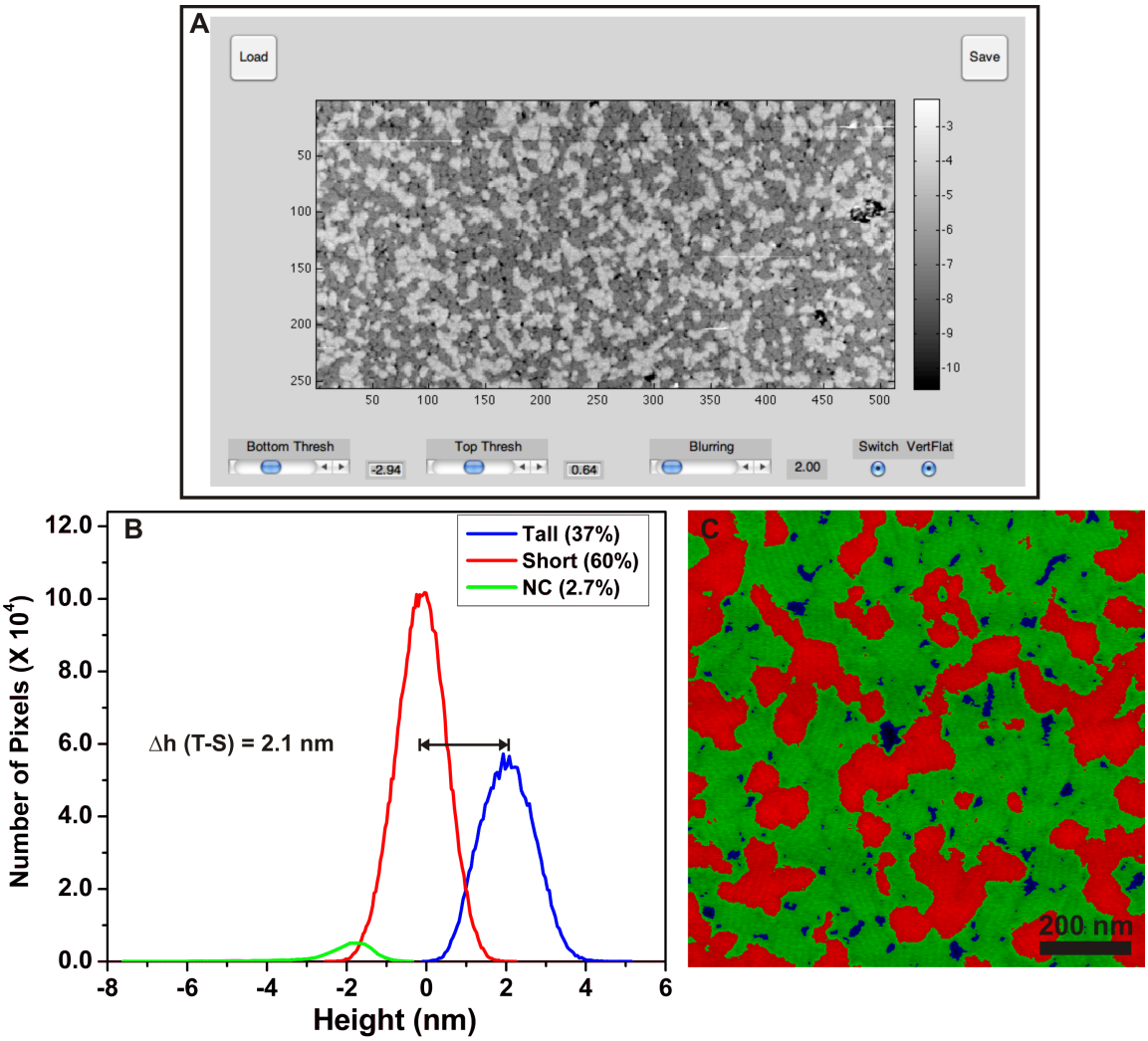


Figure S3. AFM image analysis for determining the height distribution of crystalline domains. Figs. S3*A*, S3*B*, and S3*C*. Each AFM image was analyzed in a custom Matlab® (Mathworks, Inc.) interface environment where human input guides the analysis (Fig. S3*A*). The human inputs consist of determining whether to flatten the images to account for AFM tip-changes during scanning, the degree of blurring to introduce into the image prior to thresholding, and the height thresholds for the three layers (NC, Short, Tall). All human inputs are recorded with the analysis output for reproducibility and verification. Raw AFM images are imported and flattened if needed to account for AFM tip-changes during scanning. Automatic thresholding is not sufficient as protein structures have internal variations in heights that are sometimes apparent in AFM images, but the images vary in resolution from tip-to-tip and scan-to-scan. Thus human input is

required to blur the image appropriately so that high and low crystal regions can be properly identified (see Fig. S3*C*). After thresholding for each region, the original (un-blurred) image was masked for each type of region (NC/Short/Tall) and a histogram for each developed (see Fig. S3*B*). The height difference between the low and high regions was determined from the separation in the Gaussian fits to the histograms, and the relative areas of each type of region were determined from the areas of the masks.

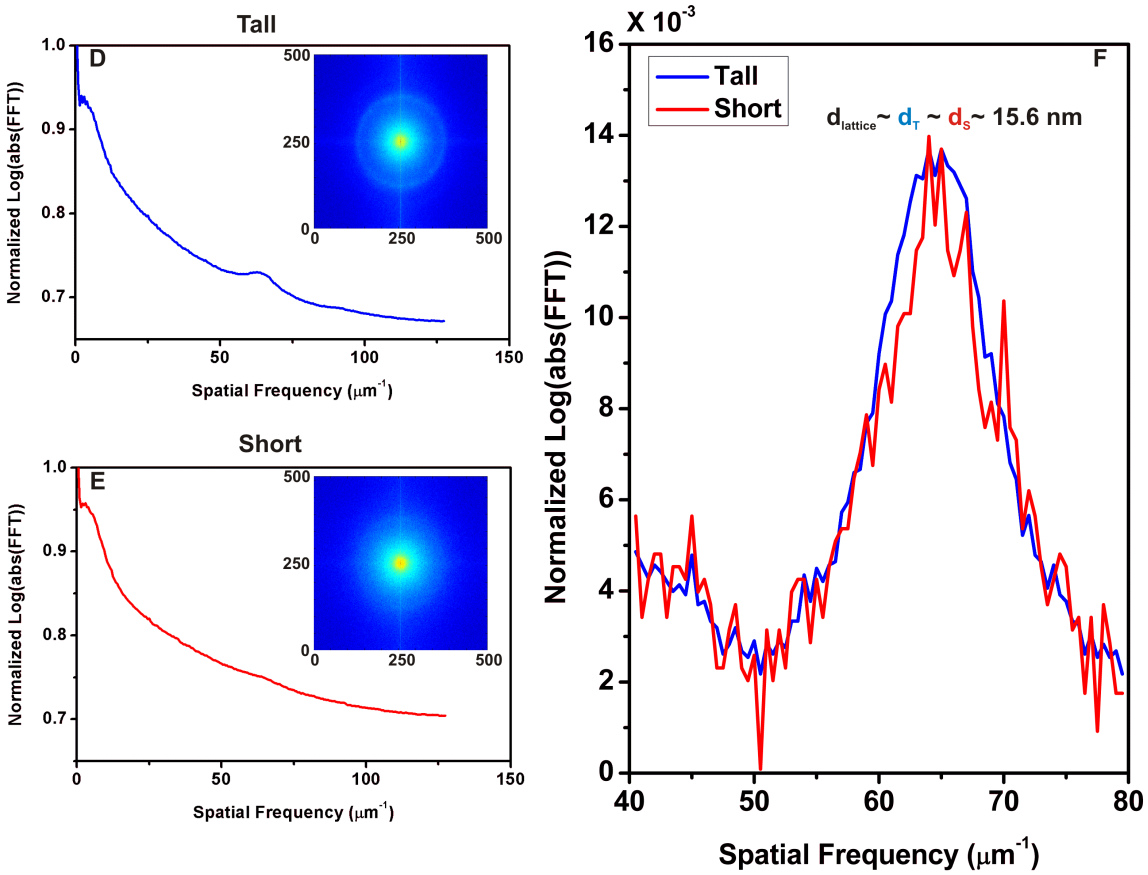


Figure S3 (cont.) AFM image analysis for determining the height distribution of crystalline domains (continued). Figs. S3*D*, S3*E*, and S3*F*. Fourier spatial analysis of domains of crystalline T and S phases. To determine the spatial spectrum of the each region, the region masks were used to separate a set of identical-scale images into NC/Short/Tall images. The images from each region were aggregated into a large image that was padded with sufficient zeros as to form a square matrix, and passed through a fast 2D fourier transform (2D FFT) using custom Matlab® scripts. The resulting spectrum was radially averaged, to produce the broad traces seen in the inset image of

Figs. S3*D* and S3*E*. To isolate the peak of interest, a second order polynomial that excluded the region of interest was subtracted from both high and low traces (features were between k = 55 – 70 μm^{-1}, fits were from 40 – 55 μm^{-1} and 70 – 80 μm^{-1}). The resulting peaks were normalized to their maximum.

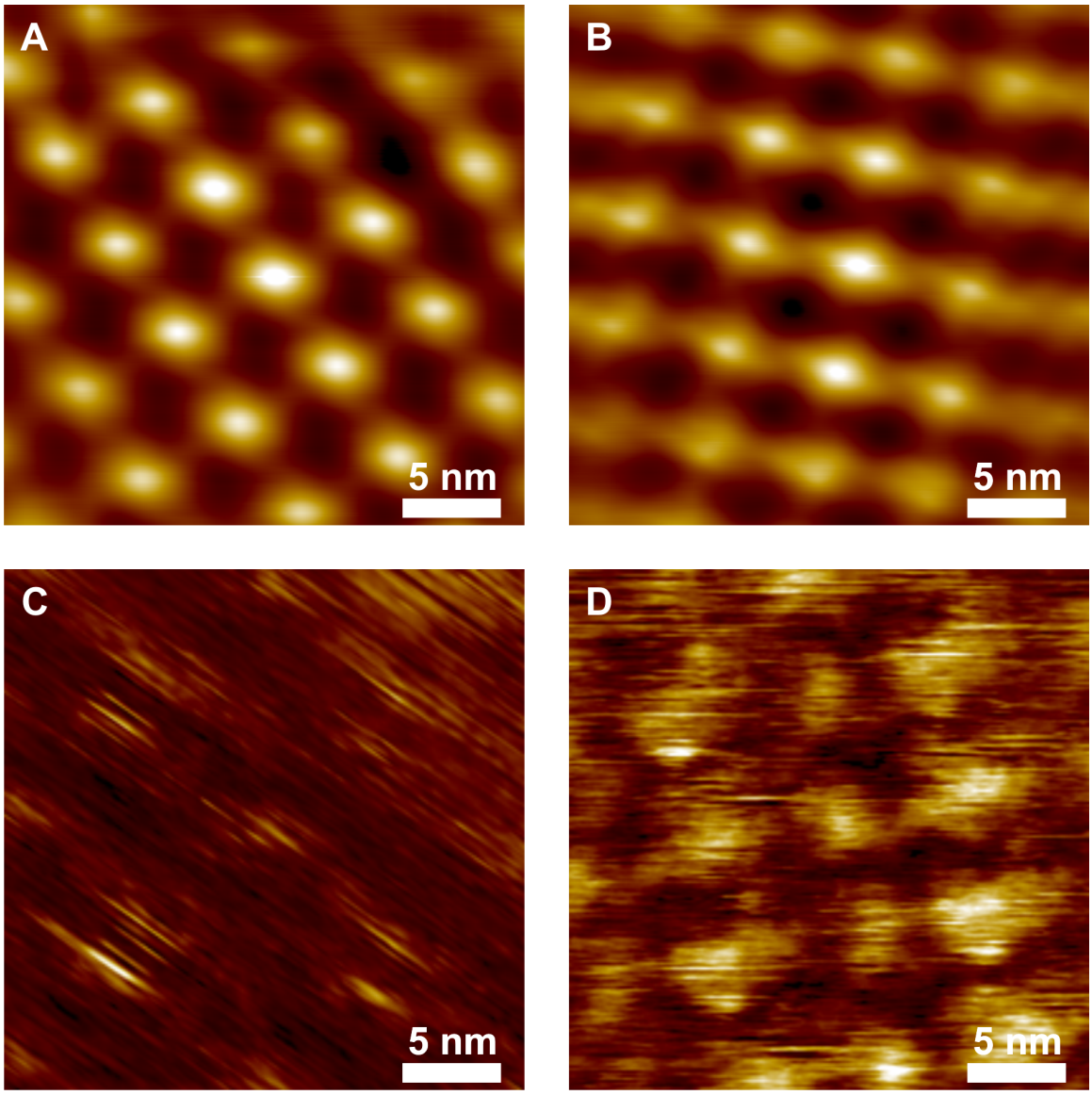


Figure S4. Cross correlation (CC) and standard deviation (SD) images calculated from correlation averaging (CA) of Figure 4C and 4D AFM images. CA imaging process creates cross correlation (CC) images (*A* and *B*) displaying CC peaks associated with the area of 4 tetramers of each T and S phase (Figs. 4*C* and 4*D*) over the entire raw

AFM image areas for averaging. From CC peaks and the raw AFM images (Figs. 4*A* and 4*B*), CA images (Figs. 4*C* and 4*D*) and SD images (*C* and *D*) were calculated. CC images (*A* and *B*) show that S phase tetramers have a more spatially extended configuration due to higher peak-to-peak overlapping correlation than that of the T phase. SD images (*C* and *D*) imply that S phase tetramers have less uniform structure than T phase tetramers.

Energy barriers to formation of T and S phases of S-layer crystals:

The relative magnitudes of the energy barriers to formation of T and S phases of S-layer crystals can be determined directly from the relative numbers of clusters that transform into the S (ϕ_S) and T (ϕ_T) phases. The number of crystalline domains of T phase (n_T) and of S phase (n_S) were counted (Figs. 1*B–D*) before the transformation process from S to T phase. The difference between the energy barriers to formation of T and S phases can be expressed as

$$\frac{\phi_S}{\phi_T} = \frac{\frac{n_S}{n_S + n_T}}{\frac{n_T}{n_S + n_T}} = \exp[-\frac{E_S - E_T}{kT}] \qquad \text{(Eq. S1)}$$

Temperature (K)	Ave. ϕ_S	Ave. ϕ_T	ϕ_S/ϕ_T	SD
298	0.343	0.657	0.522	±0.018

Table S1 Values of ϕ_S, ϕ_T, and ϕ_S/ϕ_T with standard deviation.

From Eq. S1, we get the difference between the two energy barriers ($\Delta = E_S - E_T$) of 1.6 ± 0.1 (kJ/mol) at T = 298 K.

Energy barrier to transformation from the S to T phase of S-layer crystals:

If the process of transformation from the S to T phase is the result of thermally-driven fluctuations that occur through independent events for each domain as indicated by Fig. 3, then the transformation from the S to T phase should depend exponentially on time according to:

$$N_S(t) = N_S(0) \times \exp\left(-\frac{t}{\tau}\right) \qquad \text{(Eq. S2)}$$

where the rate constant ($1/\tau$) follows an Arrhenius dependence: $1/\tau = \nu \exp(-E_B/kT)$.
Here, ν is the characteristic attempt frequency associated with conformational fluctuations and E_B is the energy barrier to transformation. Based on the condition that the sum of $N_S(t)$ and $N_T(t)$ is constant, the relative number ratio of S phase (f_S) was calculated from the data set of Figure 3a.

$$f_S(t) = \frac{N_S(t)}{N_S(t) + N_T(t)} = \frac{N_S(0)}{N_S(t) + N_T(t)} \times \exp\left(-\frac{t}{\tau}\right) \qquad \text{(Eq. S3)}$$

Time (min)	f_S	SD
271	0.515	±0.05
411	0.421	±0.03
439	0.409	±0.03
529	0.341	±0.03
601	0.323	±0.04

Table S2 Values of f_S with standard deviation at different times from Figure 3a.

Fig. 2*B* shows the time dependence of log (f_S). Based on the best linear fit (Figure 2*C*), we get $1/\tau = 0.0015 \pm 0.0001$ (min^{-1}) ($R^2 = 0.99$). From the Arrhenius dependence of $1/\tau$ and the characteristic attempt frequency of $\sim 1.0 \times 10^6$ (Hz), the magnitude of the energy barrier (E_B) of the transformation process is 61 ± 0.3 (kJ/mol).

Movie S1: Phase transformation of a single crystal domain on mica.

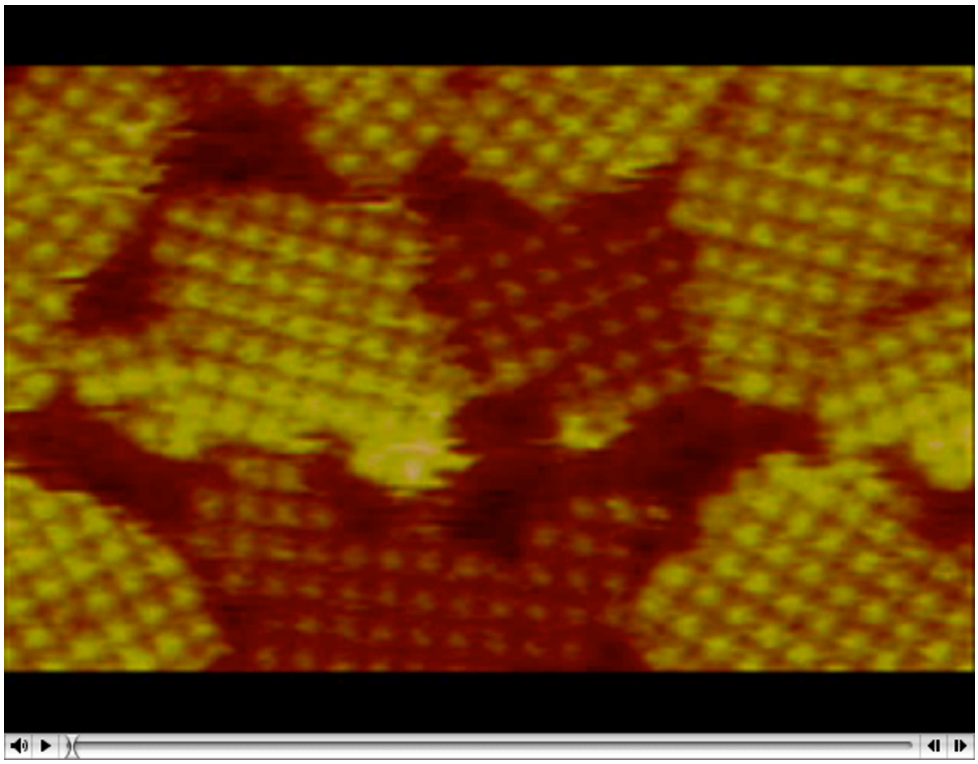


Figure S5. Still image from Movie S1 showing the transformation of a single S-layer domain on mica surface observed in solution *in situ* by AFM. The growth of 2D S-layer crystals on the mica surface was initiated by injecting the monomeric S-layer protein dissolved in the growth buffer into the fluid cell of the AFM. Movie S1 captures the transformation of a single S domain whose boundary consists of three T domains and one S domain. Once initiated at a single tetramer, the transformation occurs tetramer-by-tetramer and propagates from one edge to the other (The first frame was captured at ~ 180 min elapsed since the initiation of the experiment and time between frames is about 4.6 min). Movie S1 file can be accessed at:

https://webspace.lbl.gov/xythoswfs/webui/_xy-11142104_1-t_o1pXM0aU

To view this movie, please download a QuickTime viewer.

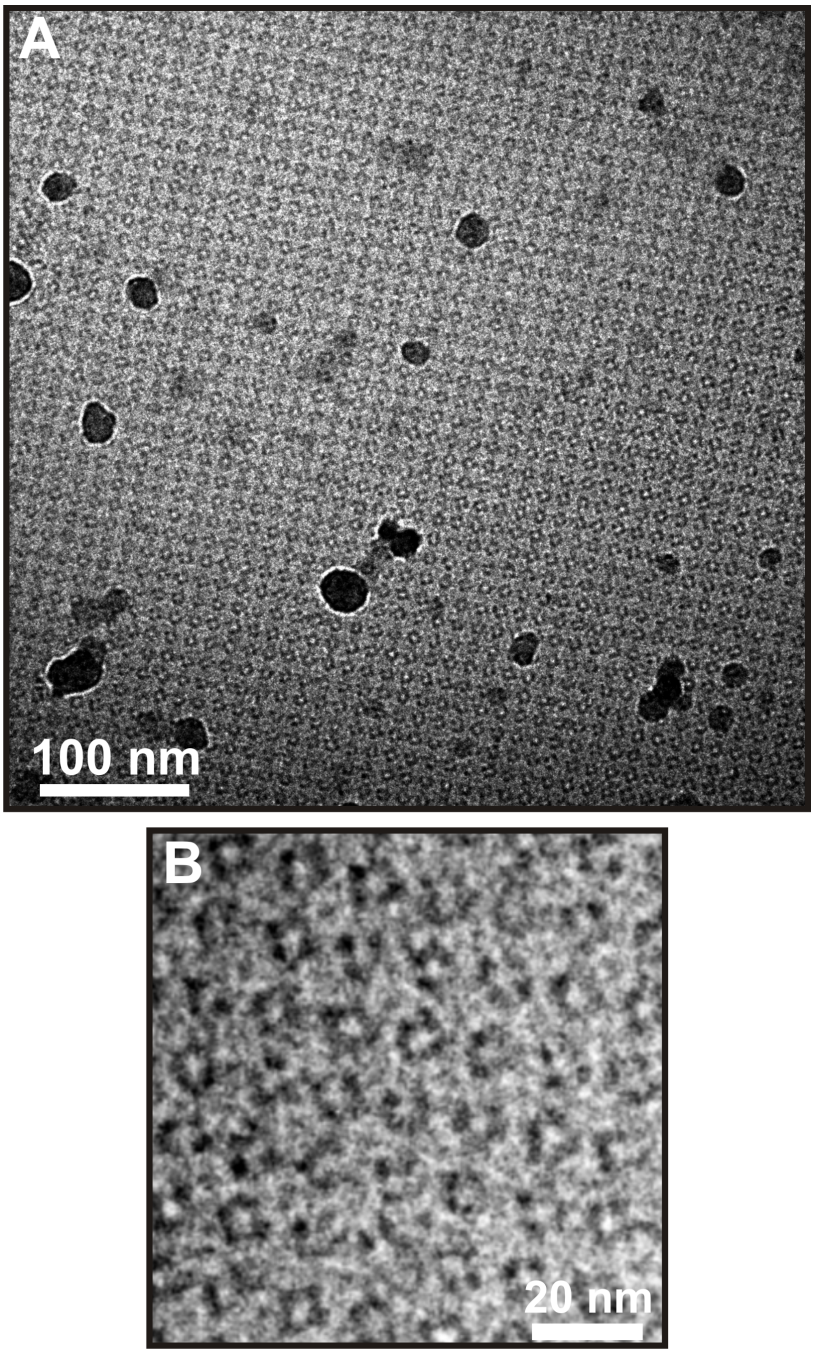


Figure S6. Cryo-TEM images of 2D S-layers from the mica-free-system. (*A*) low resolution cryo-TEM of 2D crystals of S-layers from the mica free system. (*B*) high resolution cryo-TEM of square lattice of S-layers. Each lattice unit is a tetramer, which in projection is doughnut shaped.

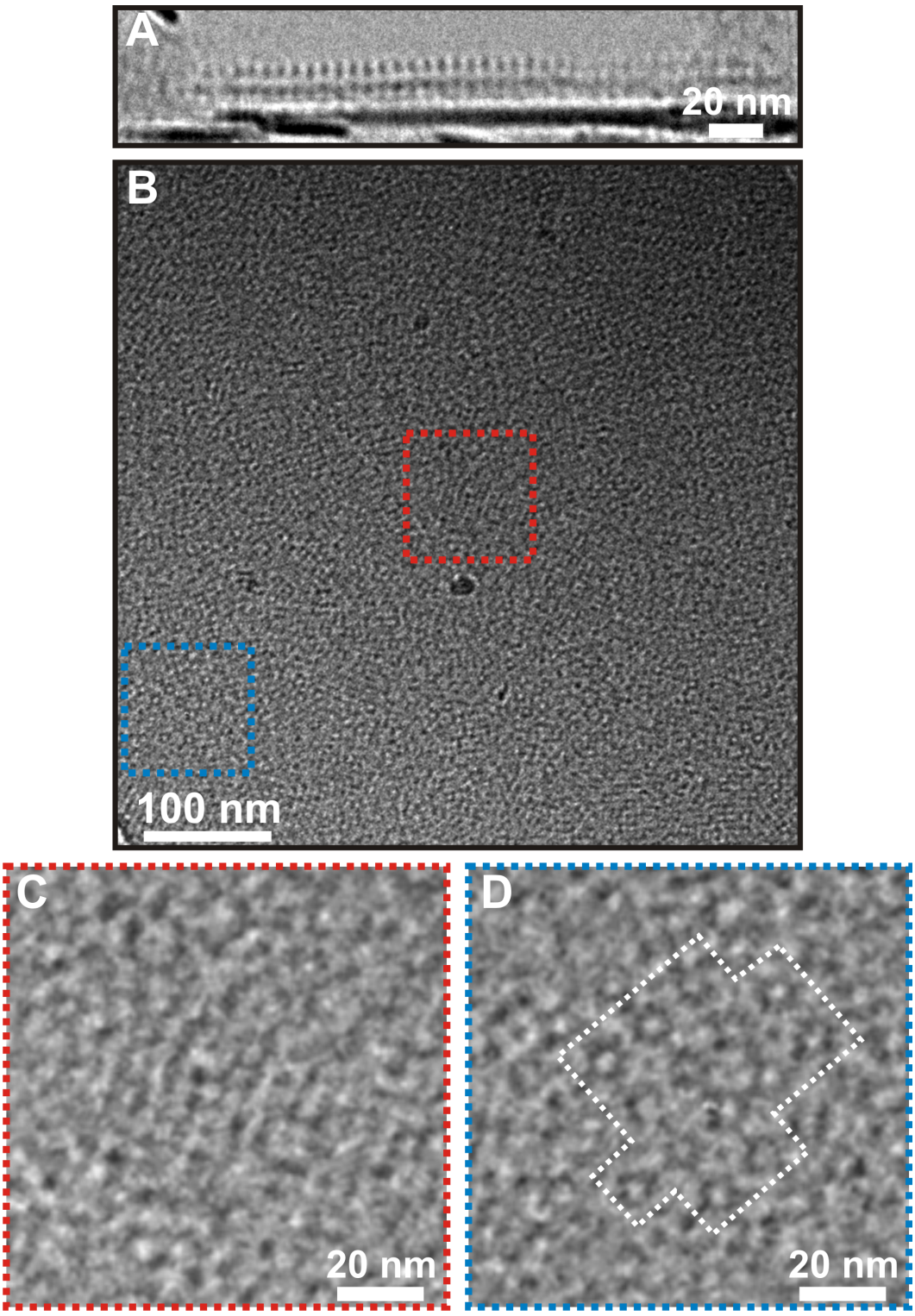


Figure S7. Cryo-TEM images of 2D S-layers formed on mica flakes. (*A*) cross-sectional side view showing a single layer of S-layer proteins. (*B*) low resolution cryo-TEM image

of 2D crystalline S-layers on mica. There are two distinct crystalline domains, which are highlighted by the dotted red and blue boxes. High resolution cryo-TEM image of domains in (*C*) the red dotted box and (*D*) the blue dotted box. Cryo-TEM image in (*D*) shows a crystalline domain (perimeter highlighted by white dotted line) composed of ~14 tetramers that are very similar to the ones in the mica-free S-layer of Fig. S6*B*.