The ability to synthesize patterned arrays in a controllable fashion is of extensive interest for nanotechnology. Arrays can be engineered to serve as components in chips for a virtually inexhaustible list of applications ranging from disease diagnosis to ultra-high-density data storage. Phi29 motor dodecamer has been reported to form elegant multilayer tetragonal arrays. However, multilayer protein arrays are of limited use for nanotechnological applications which demand nanoreplica or coating technologies. The ability to produce a single layer array of biological structures with high replication fidelity represents a significant advance in the area of nanomimetics. In this paper, we report on the assembly of single layer sheets of reengineered phi29 motor dodecamer. A thin lipid monolayer was used to direct the assembly of massive sheets of single layer patterned arrays of the reengineered motor dodecamer. Uniform, clean and highly ordered arrays were constructed as shown by both transmission electron microscopy and atomic force microscopy imaging.

**KEYWORDS:** bacteriophage phi29 · portal vertex · dodecamer · single layer patterned arrays

© 2009 American Chemical Society

Received for review June 29, 2008 and accepted November 24, 2008.

Published online December 16, 2008.

10.1021/nn800409a CCC: $40.75

© 2009 American Chemical Society
logical molecules. Self-assembly of molecules on a surface can be a simple, versatile, and high-volume production approach for the construction of biological arrays. The ability to replicate biological shapes with nanoscale precision could have profound implications in tissue engineering, cell scaffolding, drug delivery, sensors, imaging, nanotechnology, and nanomedicine.

One particularly attractive candidate found in the viral DNA-packaging machinery, from which both protein and RNA bio-nanocomponents may be harvested, is the Bacillus subtilis bacteriophage phi29 DNA-packaging motor. This powerful motor comprises a portal vertex—a 12-subunit gp10 (dodecamer protein, also called as connector), pRNA, gp16, and ATP—that provides the chemical energy required for DNA packaging. These components can be combined in vitro to assemble one of the most powerful nanomachines constructed to date.

The class of dodecamer (connector) proteins which form varieties of portal vertex shares little sequence homology among different viruses, but the resulted portal vertex has considerable morphological similarity. Information from Cryo-EM and X-ray crystallographic studies revealed the portal vertex is a 12-fold symmetric dodecamer with a truncated cone shape about 7.5 nm long and with a diameter of 6.8 nm at the narrow end (N-terminus) and 13.8 nm at the wide end (C-terminus). The central channel has a diameter of 3.6 nm. The wide end of the dodecamer is embedded in the procapsid shell, while the narrow end of the dodecamer, protruding out of the procapsid, is the foothold for pRNA binding.

Previous work has shown that the wild-type dodecamer can be used to form arrays with a mixture of single, double, or multiple layered structures. Multilayer arrays could be easily produced from ordered aggregates of portal vertex. However, multilayer arrays are of limited use for nanotechnological applications such as replica, which demand uniform, single layer biomolecular arrays. It has been found that some multilayer crystals can be converted to single layer in solution over a period of a few weeks by gradually changing the salt concentration. However, such a step is time-consuming and only generates small sheets as a mixture with single, double, and multilayers. Due to the thinness and fragility, purification of the single layer from the mixture is almost impossible.

In this study, we demonstrate that we can easily produce huge two-dimensional single layer arrays of terminus-modified portal vertex using a lipid blend.
and reengineered dodecamer (Figure 1D inset) and imaged by TEM are shown in Figure 1C,D. As previously reported, the unmodified dodecamer generated multiple overlapping layers with tetragonal symmetry (Figure 1C). The different shades of gray represent the different layers formed and overlapping. The extent to which different layers overlap cannot be controlled, and thus it is difficult to reproduce the same multilayer structure. Interestingly, the reengineered dodecamer with the added N-terminal extension self-assembled into huge flat sheets that piled into three-dimensional stacks (Figure 1D). Such stacks are different from 3D crystals, of which the stacks are governed by specific interaction between different layers. However, in these stacks, the sheets arrange in random orientation (Figure 2), suggesting that each sheet formed or grew independently. It is understandable that, without a support, the fragile sheet of the thin layer could not stand alone (see next section for formation of single layer sheets on supporting lipid monolayer). From the EM images, it appeared that the size of the center channel was smaller (compare Figure 1C and D). This is possibly due to partial filling-up of the edge of the channel by the extended peptide at the N-terminal end.

The arrangement of the individual dodecamers was analyzed using statistical classification and averaging (Figure 1E inset). The truncated cone structure of a single dodecamer enables us to distinguish between the face-up and face-down orientations. The corresponding projection density map shows the alternating face-up and face-down arrangement of dodecamers (Figure 1E). The unit cell is a square with a dimension of ~20 nm. Due to the alternating orientations of the motor, the central face-up dodecamer displays a larger diameter (corresponding to the wider C-terminal end) than the surrounding face-down dodecamers located in each corner of the square. The alternative face-up and face-down data agree with the previous studies on EM imaging, X-ray crystallography, topological analysis, structure projection of the 2D crystals, 3D reconstruction, and computer modeling of the dodecamers. The individual 12 subunits can also be observed. Formation of single layer arrays was further confirmed by AFM imaging. Freshly cleaved muscovite mica was used as an alternative substrate. Figure 1F shows a typical monolayer as template. The method is simple and reproducible.

RESULTS AND DISCUSSION

Approaches to Redirect Formation of Multilayer Structures into Single Layer Patterned Arrays. The formation of multilayer arrays is driven by two distinct protein interaction mechanisms. First, horizontal side-by-side interactions between individual dodecamers allow for the extension and growth of a two-dimensional layer. Second, interactions between the narrow and the wide ends of dodecamer molecules promote the buildup of multiple layers vertically (Figure 1A). To facilitate the formation of a single layer and prevent the continuous growth of multiple layers, a short peptide sequence was introduced either into the gp10 N- or C-terminus, located at the narrow and wide end of the dodecamer, respectively (Figure 1B).

Strep-Tag Extension of the C- or N-Terminus Favors the Assembly of Dodecamer Sheets in Solution. The phi29 portal vertex is a truncated cone-shaped structure having the gp10 N- and C-terminus located at the narrow and wide end, respectively. Fusion of a simple 22 amino acid Strep-tag to the N-terminus of the portal vertex did not interfere with the assembly of the quartenary dodecamer structure. After expression in E. coli cells, the recombinant gp10 assembled into dodecamer particles with similar shape to the native portal vertex as shown by TEM (data not shown). Additionally, the Strep-tag extension facilitates purification of the dodecamer protein with high yield and homogeneity. It has been previously reported that two-dimensional dodecamer arrays could be grown in solution from concentrated native dodecamer (connector) solution following several weeks of incubation under a defined ionic strength gradient of the buffer. However, without this precise chemical treatment, the native portal vertex has the tendency to form patches of multiple layers. Figure 1A illustrates a multiple layer structure of native dodecamers. The individual dodecamers exhibit both lateral side-by-side and vertical head-to-tail interactions. We have used a reengineered motor dodecamer for the self-assembly of single layer dodecamer sheets (schematically shown in Figure 1B). Arrays constructed from both native (Figure 1C inset) and reengineered dodecamer (Figure 1D inset) and imaged by TEM are shown in Figure 1C,D. As previously reported, the unmodified dodecamer generated multiple overlapping layers with tetragonal symmetry (Figure 1C). The different shades of gray represent the different layers formed and overlapping. The extent to which different layers overlap cannot be controlled, and thus it is difficult to reproduce the same multilayer structure. Interestingly, the reengineered dodecamer with the added N-terminal extension self-assembled into huge flat sheets that piled into three-dimensional stacks (Figure 1D). Such stacks are different from 3D crystals, of which the stacks are governed by specific interaction between different layers. However, in these stacks, the sheets arrange in random orientation (Figure 2), suggesting that each sheet formed or grew independently. It is understandable that, without a support, the fragile sheet of the thin layer could not stand alone (see next section for formation of single layer sheets on supporting lipid monolayer). From the EM images, it appeared that the size of the center channel was smaller (compare Figure 1C and D). This is possibly due to partial filling-up of the edge of the channel by the extended peptide at the N-terminal end.

The arrangement of the individual dodecamers was analyzed using statistical classification and averaging (Figure 1E inset). The truncated cone structure of a single dodecamer enables us to distinguish between the face-up and face-down orientations. The corresponding projection density map shows the alternating face-up and face-down arrangement of dodecamers (Figure 1E). The unit cell is a square with a dimension of ~20 nm. Due to the alternating orientations of the motor, the central face-up dodecamer displays a larger diameter (corresponding to the wider C-terminal end) than the surrounding face-down dodecamers located in each corner of the square. The alternative face-up and face-down data agree with the previous studies on EM imaging, X-ray crystallography, topological analysis, structure projection of the 2D crystals, 3D reconstruction, and computer modeling of the dodecamers. The individual 12 subunits can also be observed. Formation of single layer arrays was further confirmed by AFM imaging. Freshly cleaved muscovite mica was used as an alternative substrate. Figure 1F shows a typical
AFM image of crystalline N-strep dodecamer single layers imaged in tapping mode in liquid. Patches of crystalline areas with submicrometer size can be observed. The large scan image shows a high surface coverage of the protein layer with only small imperfections. A line scan across the single layer sample (Figure 1F inset) indicates that the single layer thickness is \( \approx 7.5 \) nm, which is in excellent agreement with the height of the motor dodecamer determined from the three-dimensional crystal structures. Similar single layer arrays were also produced from dodecamers of gp10 with a peptide extension at the C-terminus which serves as a barrier for the vertical interactions (data not shown).

**Formation of Single Layer Dodecamer Arrays Directed by Lipid Monolayers.** Nanotechnological applications of arrays require the assembly of homogeneous broad and wide-ranging flat sheets. Due to the flexible and fragile nature of proteins, it would be desirable to employ a biological template to direct the assembly of single layer arrays. To explore this possibility, thin layers of biotinylated lipid mixtures were used to direct the assembly of N-strep dodecamers which carry a Strep-tag at the N-terminus of each gp10 subunit. The N-strep dodecamer was preincubated with streptavidin before the lipids were spread. Arrays were grown in situ on lipid monolayers at the air–water interface and then transferred on carbon-coated TEM grids. Saturated dipalmitoyl fatty acid chains (C_{16}) of biotinyl dAPPE and helper lipids egg PC were spread at the liquid–air interface to attract dodecamers to the surface of the liquid; dodecamers bound to the lipid via specific biotin–streptavidin interactions. (B) Negative-stain TEM image of a single layer array produced with the aid of a lipid monolayer. (C) Fourier transforms and (D) corresponding Fourier projection maps of lipid-directed N-strep dodecamer arrays.

![Figure 3](image.png)

**AFM image of crystalline N-strep dodecamer single layers imaged in tapping mode in liquid.** Patches of crystalline areas with submicrometer size can be observed. The large scan image shows a high surface coverage of the protein layer with only small imperfections. A line scan across the single layer sample (Figure 1F inset) indicates that the single layer thickness is \( \approx 7.5 \) nm, which is in excellent agreement with the height of the motor dodecamer determined from the three-dimensional crystal structures. Similar single layer arrays were also produced from dodecamers of gp10 with a peptide extension at the C-terminus which serves as a barrier for the vertical interactions (data not shown).

**Formation of Single Layer Dodecamer Arrays Directed by Lipid Monolayers.** Nanotechnological applications of arrays require the assembly of homogeneous broad and wide-ranging flat sheets. Due to the flexible and fragile nature of proteins, it would be desirable to employ a biological template to direct the assembly of single layer arrays. To explore this possibility, thin layers of biotinylated lipid mixtures were used to direct the assembly of N-strep dodecamers which carry a Strep-tag at the N-terminus of each gp10 subunit. The N-strep dodecamer was preincubated with streptavidin before the lipids were spread. Arrays were grown in situ on lipid monolayers at the air–water interface and then transferred on carbon-coated TEM grids. Saturated dipalmitoyl fatty acid chains (C_{16}) of biotinyl dAPPE and helper lipids egg PC were spread at the liquid–air interface to attract dodecamers to the surface of the liquid; dodecamers bound to the lipid via specific biotin–streptavidin interactions. (B) Negative-stain TEM image of a single layer array produced with the aid of a lipid monolayer. (C) Fourier transforms and (D) corresponding Fourier projection maps of lipid-directed N-strep dodecamer arrays.
respective. The angle between the x and y axis has been calculated to be \( \sim 71^\circ \). Even though the tetragonal arrangement previously observed was maintained, the slightly different and unequal unit cell dimensions suggest that the packaging unit of this type of crystal might be slightly different from that of the lipid-directed N-strep 2D crystal. The crystal lattice in Figure 4 is different from those in Figures 1 – 3 concerning the angle of the pattern. While asking whether the difference observed in the lattices was the consequence directly related to the mutation of the protein is very intriguing, still little is known. A rectangular lattice with unit cell dimensions of 18 x 18 nm\(^2\) has been observed for the C-strep mutant. The dodecamer orientation in the self-assembled layer on the mica surface was similar to that in the three-dimensional crystal and generated face-up and face-down arrangements. While occasionally the low force applied for imaging was sufficient to image what appeared to be the narrow ends of the dodecamer, most of the time we could only visualize the wide dodecamer domains due to the nature of sample interactions in AFM imaging.

Mica has more than 10 different phases concerning the surface lattice. The lattices of mica and dodecamer are of a different order of magnitude, in which mica is calculated about 6 Å compared to the protein with a unit cell in the regime of about 16–18 nm. In this AFM imaging, it is not clear whether the mica surface lattice played a role here in organizing the pattern of the proteins array, and whether the mica surface lattice and the protein crystal lattice are relevant or in a good match. However, previous studies have shown that the purified native dodecamer self-assembled into tetragonal arrays in solution without the mica support, guiding of these nanoparticles by the mica surface lattice to form the pattern in this report might not be necessary. Instead, the pattern of the lattice might have been dictated by the intrinsic property of the mutant dodecamer structure.

**CONCLUSIONS**

A short Strep-tag sequence modification of the N- or C-terminus of the phi29 portal vertex facilitates its purification with high yield and homogeneity. The modification did not interfere with the dodecamer assembly and function. The mutant protein exhibited favorable lateral interactions and led to the formation of large dodecamer sheets. In solution, the 2D dodecamer arrays interacted vertically to pile up into 3D stacks of protein sheets as revealed by TEM imaging. Large single layer sheets of highly ordered array have been constructed using a supporting lipid monolayer.

**METHODS**

**Reengineering of Phi29 Motor Dodecamer.** Two clones of portal vertex protein were engineered by attaching a Strep-tag II (WSH-PQFER) to either the N-terminus or the C-terminus of each gp10 subunit. Cloning methods of the N-strep and C-strep dodecamer were engineered by attaching a Strep-tag II (WSH-PQFER) to either the N-terminus or the C-terminus of each gp10 protein.

**Purification of the Reengineered Dodecamer.** A column packed with 1 mL of Strep-Tactin sepharose resin (IBA, St. Louis, MO) was equilibrated with 10 column bed volumes of buffer W (100 mM Tris-HCl, pH 8.0, 500 mM NaCl, 1 mM EDTA, 15% glycerol). After lysis of E. coli cells containing the reengineered gp10, the lysate was clarified and the supernatant was loaded onto the column, followed by washing with buffer W, the protein was eluted with buffer E (500 mM NaCl, 1 mM EDTA, 2.5 mM desthiobiotin, 100 mM Tris-HCl, pH 8.0, 15% glycerol).

**Assembly of Dodecamer Arrays.** Two approaches were used to construct dodecamer arrays: (1) self-assembly from concentrated solutions of purified native and N-strep or C-Strep motor dodecamer and (2) lipid-directed assembly of single layer dodecamer arrays. The schematic illustrations of multilayer arrays or single layer patterned sheets are shown in Figure 1A,B.

**Self-Assembled Dodecamer Arrays.** Concentrated solutions of purified, reengineered C- and N-strep dodecamer in buffer (100 mM Tris-HCl, 0.5 M NaCl, 1 mM EDTA, 0.02% sodium azide, 15% glycerol, pH 8.0) were stored at \(-20^\circ\)C. Protein solutions dialyzed and diluted if necessary to a stock of 1 mg/mL were kept at 4 °C for a few days and used for the construction of two-dimensional arrays. A 1:35 dilution of the protein stock solution in imaging buffer (10 mM Tris-HCl, pH 8.0, 500 mM KC1) was applied on freshly cleaved mica. The sample was placed in a humidified, closed Petri dish to avoid drying out. Following 2 h incubation at room temperature, the sample was rinsed with imaging buffer and kept at 4 °C overnight. The sample was allowed to reach room temperature prior to AFM imaging.

**Lipid-Directed Single Layer Dodecamer Arrays.** Two-dimensional dodecamer arrays were grown at the liquid–lipid interface as previously reported by Sun et al.\(^{71,72}\) The N-strep dodecamer at a concentration of 0.1 mg/mL was incubated in buffer (50 mM
Tris-HCl, 100 mM NaCl, 20 mM MgCl₂, pH 8) at room temperature with a 6-fold excess of streptavidin (Sigma). A volume of 15 μL of the N-strep dodecamer bound to streptavidin was placed into a custom-designed Teflon well of 4 mm in width and 1 mm in depth. The lipid mixture of 30 μg/mL, 1% uranyl acetate (UA) or 2% (w/v) ammonium molybdate. The two-dimensional arrays grown on the lipid matrix were transferred to a hydrophilic glow discharged carbon grids that were negatively stained with 1% uranyl acetate (UA) or 2% (w/v) ammonium molybdate. The V-shaped cantilevers were transferred into a JEOL-2100F TEM operated at 120 kV, and washed with distilled water, and negatively stained with 1% uranyl acetate (UA) or 2% (w/v) ammonium molybdate. The V-shaped cantilevers were transferred into a JEOL-2100F TEM operated at 120 kV, and washed with distilled water, and negatively stained with 1% uranyl acetate (UA) or 2% (w/v) ammonium molybdate. The V-shaped cantilevers were transferred into a JEOL-2100F TEM operated at 120 kV, and washed with distilled water, and negatively stained with 1% uranyl acetate (UA) or 2% (w/v) ammonium molybdate.

**Imaging of Single Layer Arrays by Transmission Electron Microscopy (TEM).** Samples were prepared by applying the protein stock solution on hydrophilic glow discharged carbon grids that were negatively stained with 1% uranyl acetate (UA) or 2% (w/v) ammonium molybdate. The V-shaped cantilevers were transferred into a JEOL-2100F TEM operated at 120 kV, and washed with distilled water, and negatively stained with 1% uranyl acetate (UA) or 2% (w/v) ammonium molybdate. The V-shaped cantilevers were transferred into a JEOL-2100F TEM operated at 120 kV, and washed with distilled water, and negatively stained with 1% uranyl acetate (UA) or 2% (w/v) ammonium molybdate.

**ACKNOWLEDGMENT.** We thank D. Green from University of California at Davis for his technical assistance and valuable comments, as well as the preparation of Figures 1D and 2. This work was supported by PN2 EY018230 from NIH Nanomedicine Development, as well as the preparation of Figures 1D and 2. This work was supported by PN2 EY018230 from NIH Nanomedicine Development.

**REFERENCES AND NOTES**


