Adjustable Ellipsoid Nanoparticles Assembled from Re-engineered Connectors of the Bacteriophage Phi29 DNA Packaging Motor

Feng Xiao,†, ‡ Ying Cai,†, ‡ Joseph Che-Yen Wang,§ Dominik Green,§ R. Holland Cheng,§ Borries Demeler,§ and Peixuan Guo†,*

†Department of Biomedical Engineering, College of Engineering and College of Medicine, University of Cincinnati, Cincinnati, Ohio 45221, ‡Department of Molecular & Cellular Biology, University of California, Davis, California 95616, and §Department of Biochemistry, The University of Texas Health Science Center at San Antonio, 7703 Floyd Curl Drive, San Antonio, Texas 78229. ‡Co-first-author. ¶ Current address: VGX Pharmaceuticals, Inc, 2700 Research Forest Drive, Suite 180, The Woodlands, Texas 77381.

Rational design of nanoparticles has become a prevalent trend requiring an in-depth comprehension of the chemical and physical characteristics of their building blocks. Taking advantage of the highly specific binding properties of biomolecules, the combination of biotechnology and nanotechnology has led to the development of hybrid nanobiomaterials. Biological materials, in the form of DNA, RNA, protein, and lipids, serve as models for the self-recognition and self-assembly of bionanoparticles.1–3 Peptides also play a unique role in nanostructure design, owing to their diversity, simplicity of synthesis, and ease of modifying them for a variety of functions. Understanding the self-assembly mechanism of these biomaterials enables us to design and engineer biomimetics on a nanoscale. Extensive investigations have been conducted and successful applications achieved with DNA,4 RNA,5–9 viral proteins,10–13 bacterial S-layer proteins,14–16 peptides, and peptidomimetics.17

Bacteriophage phi29 is one of several well-studied, small-tailed phages. It infects Bacillus subtilis. This double-stranded DNA virus utilizes a unique motor to package its 19.3 kb genome into a preformed procapsid.18 The connector protein is an essential component in the phi29 DNA-packaging motor. Its crystal structure, determined at high-resolution, is a 12-fold symmetric dodecamer that forms a truncated cone with a length of 7.5 nm, a diameter of 13.8 nm at the wide end and 6.8 nm at the narrow end, and contains a 3.6 nm central pore for DNA translocation.19,20 A crystal structure failed to reveal the N-terminal 13 amino acids because of their flexibility in structure. Inclusion of these 13 amino acid at the N-terminus of each wild-type gp10 subunit results in about 7.8 nm in diameter at the narrow end of the connector.12

Connector proteins of the bacteriophage phi29 can assemble into a rosette-like particle after interaction with pRNA.12,21,22 In this study, we report the construction of a distinct globular nanoparticle assembled from a re-engineered phi29 connector protein. The maximum diameter of the resultant nanoparticles is twice as large as the maximum diameter of the connector formed by native connector protein gp10, in addition to having a vastly different quaternary assembly. The structural transition between the connector and the nanoparticle can be easily controlled via the addition of N-terminal peptide extensions. Reversal of the 84-subunit ellipsoid nanoparticle to its dodecamer subunit was controlled by the cleavage of the extended N-terminal peptide with a protease. The 84 outward-oriented C-termini were conjugated with a streptavidin binding peptide which can be used for the incorporation of markers. This further extends the application of this nanoparticle to pathogen detection and disease diagnosis by signal enhancement.

**ABSTRACT** A 24 × 30 nm ellipsoid nanoparticle containing 84 subunits or 7 dodecamers of the re-engineered core protein of the bacteriophage phi29 DNA packaging motor was constructed. Homogeneous nanoparticles were obtained with simple one-step purification. Electron microscopy and analytical ultracentrifugation were employed to elucidate the structure, shape, size, and mechanism of assembly. The formation of this structure was mediated and stabilized by N-terminal peptide extensions. Reversal of the 84-subunit ellipsoid nanoparticle to its dodecamer subunit was controlled by the cleavage of the extended N-terminal peptide with a protease. The 84 outward-oriented C-termini were conjugated with a streptavidin binding peptide which can be used for the incorporation of markers. This further extends the application of this nanoparticle to pathogen detection and disease diagnosis by signal enhancement.

**KEYWORDS:** nanobiotechnology · bionanotechnology · viral DNA packaging · phi29 DNA packaging motor · protein nanoparticles · virus assembly · bacteriophage phi29 connector

*Address correspondence to guopn@ucmail.uc.edu.

Received for review February 24, 2009 and accepted July 07, 2009.

Published online July 17, 2009. 10.1021/nn900187k CCC: $40.75

© 2009 American Chemical Society

www.acsnano.org

VOL. 3 • NO. 8 • 2163–2170 • 2009

2163
or removal of the N-terminal peptide. This globular nanoparticle, with a surface exposed tag for functional conjugation with other molecules, has potential applications in nanotechnology.

RESULTS

Addition of the N-Terminal Peptide Produced a Novel Globular Nanoparticle. A 22-residue peptide was used to extend the N-terminus of gp10 protein of the phi29 connector, resulting in a modified N-Strep connector. This peptide included a Streptavidin binding peptide (WSHPQFEK, referred to as “Strep-II tag”)23,24 and a tobacco etch virus (TEV) protease cleavage site (ENLYFQG) (Figure 1). The Strep-tag, which bound the streptavidin protein, facilitated one-step Strep-Tactin (IBA GmbH, Germany) purification of protein,25 while the TEV protease cleavage-site enabled removal of the peptide when necessary. The purified N-Strep connector particles, first examined by transmission electron microscopy (TEM) (Figure 2A), revealed drastically different projections than typically isolated wild-type connectors, which form a dodecameric structure with a diameter of 7.8 nm at the narrow end12 and 13.8 nm at the wider end19,20 (Figure 1). However, the nanoparticles formed by the N-strep connector, which displayed an ellipsoid shape (see section on imaging the N-strep nanoparticles) were 30 nm (transverse, major axis) by 24 nm (lengthwise, minor axis), displaying a 5- or 10-fold rotational symmetry.

Switching between Connector and the Ellipsoid Nanoparticle by Cleaving the N-Terminal Peptide with Proteinase. The N-Strep connector protein preferentially assembles into ellipsoid nanoparticles within the cell during expression or within the crude lysate. These particles were purified to homogeneity since they were fairly stable under a range of salt concentrations, pH, and temperature. A TEV enzyme recognition site was inserted between the connector and the N-terminal extended peptide to allow the added peptide to be removed by protease treatment. Removal of the N-terminal peptide by TEV protease resulted in the dissociation of the ellipsoid particles into individual connectors, as revealed by TEM imaging. Gradient sedimentation with 15–35% glycerol further confirmed the similarity in the sedimentation rate of the TEV-processed particle and individual connectors (Figure 3). The N-Strep connector primarily centered at fraction 9 of 29 total fractions (Figure 3, dot line). However, removal of the N-terminus tag via TEV cleavage caused the peak to shift to fraction 25 where the wild-type dodecamer connectors can be found (Figure 3, solid line with open circle). These results suggest that adding the 22-residue peptide to the N-terminus of the connector contributes to the assembly of particles that are different from individual connectors in size, conformation, and mass, and subsequent removal of the N-terminus tag leads to dissociation of the nanoparticle.

Mass and Shape Analysis via Sedimentation Velocity by Analytical Ultracentrifugation Revealed Nanoparticles with Seven Connectors or 84 Copies of Monomeric gp10. Sedimentation velocity (SV) experiments can be used to estimate the molecular weight, partial concentration, and relative shape of multiple solutes in a mixture of macromolecules.26 A van Holde–Weischet analysis27 provides diffusion corrected sedimentation distributions and can be used to ascertain composition. Shape and molecular weight (MW) are derived from the sedimentation (s) and diffusion (D) coefficients fitted in the finite element solutions of the Lamm equation28 implemented in the 2-dimensional spectrum analysis (2DSA) or genetic algorithm (GA) analysis, and from the knowledge of the partial specific volume. Shape is parametrized with the frictional ratio, f/f0, a measurement of the globularity of the solute. An f/f0 value of 1.0 refers to a spherical particle, while values larger than 2.5 generally indicate a nonglobular, unfolded or extended, or chainlike molecule, such as DNA or fibrils. Values between 1.2–1.4 are typical for moderately globular proteins. The partial specific volume was estimated from the protein sequence to be 0.7265 cm³/g for both the N-Strep connector and the C-strep connector. The C-strep connector is a 22-residue peptide that includes a Streptavidin binding peptide (WSHPQFEK, referred to as “Strep-II tag”)23,24 and a tobacco etch virus (TEV) protease cleavage site (ENLYFQG) (Figure 1). However, the nanoparticles formed by the N-strep connector, which displayed an ellipsoid shape (see section on imaging the N-strep nanoparticles) were 30 nm (transverse, major axis) by 24 nm (lengthwise, minor axis), displaying a 5- or 10-fold rotational symmetry.
connector without N-terminus modification serves as a single connector control. A van Holde–Weischet sedimentation coefficient distribution indicated that N-Strep connectors are present as a mixture of three predominant species: 78.00 S (42%), 15.04 S (38%) and 3.72 S (20%), while N-terminal unmodified proteins, C-Strep connector, were homogeneous at 15.14 S (100%) (Figure 4). To further investigate the identity of the three peaks, a GA–MC (genetic algorithm–Monte Carlo) analysis was performed. This analysis can resolve a mixture of solutes according to size and shape, and provide partial concentrations for each species. Results are presented as plots of $s$ vs $f_0$. Partial concentration is measured in optical density units and is represented as a color gradient. The GA–MC analysis for C-Strep connector is shown in Figure 5, the same analysis for N-Strep connector is shown in Figure 6. Confidence intervals for sedimentation and diffusion coefficients, frictional ratios, partial concentration, and molecular weight of each species are summarized for the N-Strep connectors in Table 1 and the C-Strep connectors in Table 2. The MWs were consistent with monomeric (protein gp10), 12-meric (connector), and 84-meric (the ellipsoid nanoparticle) N-Strep connector. The $f_0$ values indicated an overall increase in globular shape from a monomeric unit to an 84-mer. An $f_0$ value of 1.1 (most globular) was determined for the 84-mer (seven dodecameric connectors), an $f_0$ value of 1.51 was determined for the dodecamer connector, and an $f_0$ value of 1.63 was determined for the monomeric protein gp10 subunit. The 12-mer dodecamer connector structure for the N-terminal unmodified nanoparticles, C-strep connector, was found to have an $f_0$ value of 1.38—similar to the frictional ratio obtained for the 12-mer of N-Strep connectors. The data agreed with the available information for the molecules and assemblies: protein gp10 is a linear molecule mainly composed of α-helix; the connector is a truncated cone shape structure with a 3.6 nm central channel, and the resulting nanoparticle has an ellipsoid shape, which shows a more globular conformation (as detailed in the next section).

Imaging the N-Strep Nanoparticles. It appears that most nanoparticles formed by N-strep connector within the data set exist as top-view only projections, since very few 2-fold symmetric side views could be detected (Figure 2B). Particles were found to preferentially lie along the major axis (transverse) as opposed to the minor axis (lengthwise). This finding, that the majority of particles displayed a size that corresponded to the major axis, supports the conclusion of the ellipsoid configuration (Figure 7). Each circular projection contains 10 arms of density that emanate radially outward from the particle center, suggesting the presence of a 10-fold symmetry axis. The length of each arm is approximately equivalent, all measuring between 7–9 nm in length. Additionally, nearly all of these projections contain a central ring of density from which the 10 radial arms extend. Within this central ring is a density-void pore measuring 3–4 nm in diameter, while the outermost diameter of the ring measures between 13–15 nm. The reprojection of the density of a modeled connector oriented perpendicular to its 12-fold axis measures ~8 nm along its edge, similar to that of the nanoparticle radial extension length. Also, the reprojection of a connector volume viewed down its 12-fold axis has a maximum diameter of 13.5 nm and a central pore measuring 3.5 nm in diameter. Both values match well with those determined for the nanoparticle central ring density. The projection image in Figure 7A has a measured length of 30 nm and a height of 24 nm. The length of this projection corresponds well to the maximum diameter of the nanoparticle projection (Figure 7B), suggesting that the view in Figure 7A corresponds to a projec-
The five equatorial connectors (Figure 7C). As seen from top, this nanoparticle configuration is expected to possess 5-fold rotational symmetry (although the view was not discernible from a 10-fold symmetry due to a low signal-to-noise ratio). Additionally, the side view projection of a nanoparticle (Figure 7A) reveals a horizontal mirror axis which suggests the presence of two polar connectors (one above and one below the equatorial connector plane) (Figure 7C). This yields a side dimension of 24 nm and an equatorial diameter of 30 nm (Figure 7A,C). Moreover, the vertical mirror axis suggests a view down the local 12-fold axis of an axial connector.

To further define the symmetry of the nanoparticles, rotational averaging was applied to each projection image. Application of n-fold rotational symmetry—n = 2, 5, and 10—to a single, well-resolved nanoparticle top-view resulted in the images depicted in Figure 7B. Although the imposition of rotational averaging generated images that closely resemble their parent projection, the image with imposed 5-fold rotational symmetry most accurately matches the structural interpretation of the nanoparticle described above. The 10-fold symmetry of individual nanoparticles seen within the data set most likely arises from the low signal-to-noise levels present in TEM imaging, thus causing the separation between individual equatorial connectors to be more difficult to resolve. The 2-fold rotational averaging mimics the 10-fold-imposed situation by simulating the superposition of a 36° rotation, inherently smearing the separation between neighboring connectors. The 2-fold symmetry was also applied to the side-view projection (Figure 7A). Enhancing the centralized equatorial connector hole and maintaining the same dimensions and relative pixel intensity measurements strongly support that a 2-fold symmetry (Figure 7A) is an accurate representation of the side-view (10-fold symmetry) projection depicted in Figure 7B.

### Table 1. Genetic Algorithm—Monte Carlo Results for Three Major Species in N-Strep Nanoparticles

<table>
<thead>
<tr>
<th>species</th>
<th>N-Strep, gp10 (monomer)</th>
<th>N-Strep connector (12-mer)</th>
<th>N-Strep ellipsoid particle (84-mer)</th>
</tr>
</thead>
<tbody>
<tr>
<td>sedimentation coefficient (×10^{-11} s)</td>
<td>2.46(2.45,2.47)</td>
<td>15.16(15.07,15.20)</td>
<td>78.11(77.95,78.31)</td>
</tr>
<tr>
<td>diffusion coefficient (×10^{-7} cm/sec^2)</td>
<td>6.01(5.94,6.06)</td>
<td>2.74(2.66,2.77)</td>
<td>1.93(1.73,2.07)</td>
</tr>
<tr>
<td>frictional ratio</td>
<td>1.63(1.62,1.64)</td>
<td>1.51(1.49,1.53)</td>
<td>1.10(1.04,1.18)</td>
</tr>
<tr>
<td>molecular weight (kilodalton)</td>
<td>36.3(36.0,36.8)</td>
<td>493.8(485.4,503.7)</td>
<td>3590 (3329, 3985)</td>
</tr>
<tr>
<td>molecular weight (theor.) (kilodalton)</td>
<td>38.6</td>
<td>463.3</td>
<td>3243</td>
</tr>
<tr>
<td>partial concentration (optical density, 230 nm)</td>
<td>0.294(0.293,0.295)</td>
<td>0.498(0.479,0.496)</td>
<td>0.586(0.572,0.603)</td>
</tr>
</tbody>
</table>

### Table 2. Genetic Algorithm—Monte Carlo Results for Single Component in C-Strep Nanoparticles

<table>
<thead>
<tr>
<th>species</th>
<th>N-terminal unmodified connector (12-mer)</th>
</tr>
</thead>
<tbody>
<tr>
<td>sedimentation coefficient (×10^{-11} s)</td>
<td>14.61(14.60,14.62)</td>
</tr>
<tr>
<td>diffusion coefficient (×10^{-7} cm/sec^2)</td>
<td>3.13(3.11,3.18)</td>
</tr>
<tr>
<td>frictional ratio</td>
<td>1.38(1.38,1.40)</td>
</tr>
<tr>
<td>molecular weight (kilodalton)</td>
<td>415.5(408.5,415.9)</td>
</tr>
<tr>
<td>molecular weight (theor.) (kilodalton)</td>
<td>447.5</td>
</tr>
</tbody>
</table>

**Figure 5. Genetic algorithm—Monte Carlo analysis:** Frictional ratio vs sedimentation coefficient for C-strep connector. This plot shows the globularity (as parametrized by the frictional ratio, see text) plotted against sedimentation coefficient. The relative concentration of each species is shown in a color gradient ranging from white (zero concentration) to black (1.43 optical density units) measured at 230 nm. Only one major species is apparent (the second and third unlabeled species are negligible contaminants).

**Figure 6. Genetic algorithm—Monte Carlo analysis:** Frictional ratio vs sedimentation coefficient for N-strep connector. This plot shows the globularity (as parametrized by the frictional ratio, see text) plotted against sedimentation coefficient. The relative concentration of each species is shown in a color gradient ranging from white (zero concentration) to black (0.7 optical density units) measured at 230 nm. Three major species are apparent (the fourth, unlabeled species is a negligible contaminant). Molecules become increasingly globular when oligomerizing from monomer to 84-mer via 12-mer.
Addition of a C-Terminal Tag Did Not Interfere with the Formation of the Ellipsoid Nanoparticle. The data described above suggests that the ellipsoid nanoparticle contains seven connectors that comprise a total of 84 gp10 molecules whose C-termini protrude from the nanoparticles. That is, the surface of the ellipsoid particle displayed 84 carboxyl groups. If these groups were used for conjugation, each particle would hold up to 84 markers, thus significantly enhancing the detection signals. However, it is important that the protein is able to refold to its original structure after modification, for example, the addition of a peptide to the C-terminus. To determine the possibility of C-terminal conjugation, a tag that binds streptavidin was fused to the C-terminus of each gp10 subunit (termed ‘C-Strep’). Our results, derived from TEM imaging and glycerol gradient sedimentation (Figure 2 and 3), demonstrated that the addition of the Strep-tag to the C-terminus of the connector proteins neither interfered with connector assembly nor hindered the formation of the ellipsoid nanoparticle, which suggests that adding biomarkers or fluorescent markers to the surface of the nanoparticle is feasible. However, without the N-terminal extension, the C-terminal extension alone did not lead to the formation of the ellipsoid particles.

**DISCUSSION**

In this study, we found that adding an N-terminal extended tag to the connector promotes the assembly of new ellipsoid nanoparticles that differ in size, shape, and geometry from the original cone-shaped connector. Glycerol gradient sedimentations, TEM, and analytical ultracentrifugation revealed that the new structure is composed of seven connectors, or 84 gp10 proteins. The connectors are arranged with the narrow N-terminus oriented toward the center of the particle and the wider C-terminus exposed at the surface. Removing the N-terminal extended tag reversibly converted the ellipsoid nanoparticle structures back into individual connectors. Although more studies are needed to determine the specific requirement of peptide sequence and length to facilitate nanoparticle formation, it is inferred that the col-
Electron Microscopy Imaging. Copper grids were coated with 400-mesh Formvar and carbon, and the glow was discharged prior to use. Purified protein samples were dialyzed and diluted, when necessary, before negative staining with 2% uranyl acetate. The samples were imaged with a Philips CM-100 TEM operating at 80 kV. CCD readout magnifications were either 39000× or 52000× for ellipsoid nanoparticles or connector particles, respectively. For the projection image analysis, images were collected on a JEOL-JEM2100 at 200 kV. Particles were embedded with 16% ammonium molybdate to prevent drying distortions. Micrographs were then screened to avoid image exhibiting astigmatism, drift, or charging in the analysis.

Analytical Ultracentrifugation (AUC) to Determine the Sedimentation Coefficient As Well As the Shapes and Sizes of the Particles. Purified proteins, C-strep connectors, and N-Strep connectors were studied by sedimentation velocity (SV) experiments performed in a Beckman XLA (Center for Analytical Ultracentrifugation of Macromolecular Assemblies, CAUMA, University of Texas Health Science at San Antonio, UTHSCSA) using absorbance optics measured at 230 nm (N-Strep connector: 0.70 OD, C-strep connector: 1.43 OD). The experiments were performed at 20 °C in 2-channel Epon centerpieces at 20k rpm in an AN60 TI rotor using a buffer containing 100 mM Tris, 500 mM NaCl, and 1 mM EDTA. All data were analyzed with the UltraScan software.26,27 Finite element simulations of the Lamm equation28 were performed according to methods described by Cao and Demeler.29 Hydrodynamic corrections were made on the basis of the known buffer composition, while partial specific volume was estimated on the basis of the peptide sequence,30 as implemented in UltraScan. All SV data were processed as follows: a preliminary van Holde–Weischet analysis was used to obtain a model-independent and diffusion-corrected range for the sedimentation coefficients of each sample;27 the value of the range was used to initialize a 2-dimensional spectrum analysis (2DSA).28 Simultaneously, all systematic noise was removed from this data.31 All subsequent analyses were performed on modified data sets that achieved the time-invariant noise corrected data set, which was achieved by performing a 2DSA Monte Carlo (MC) analysis with 50 iterations. This method attenuates stochastic noise contributions to the solution by enhancing the intrinsic signal through signal amplification.32 The resulting data were used to initialize the genetic algo-

**TABLE 3. Primer Sequence for the Construction of gp10 Vectors**

<table>
<thead>
<tr>
<th>location of extension</th>
<th>primer</th>
<th>primer sequence (S′−3′)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N-terminus</td>
<td>P1</td>
<td>CGTAACCTCGATATGTGGACCATGCAAGGTTTCAGTATGATACCTTCAGAC</td>
</tr>
<tr>
<td></td>
<td>P2</td>
<td>CTAGATCTCTGACGGTCTACATTGTGTTTACCCGTC</td>
</tr>
<tr>
<td>C-terminus</td>
<td>P3</td>
<td>CGCAAGCTGGAAATTCACGATGAC</td>
</tr>
<tr>
<td></td>
<td>P4</td>
<td>GGATGACCACACCCCTCCACCCGATCCTACGTTACAGG</td>
</tr>
<tr>
<td></td>
<td>P5</td>
<td>ATATGTCTGTGCTGGTCACCTTCCTTCGAGCAAGCCTCAACCTC</td>
</tr>
</tbody>
</table>
Acknowledgment. We thank Karen Rufus from Vanderbilt University for providing the TEV proteinase expressing strains, Debra Sherman from Purdue University for EM imaging, Anne Vanderheide from University of Cincinnati for valuable comments, and Virgil Schirf (UTHSCSA) for assistance with the analytical ultracentrifugation experiments at CAUMA. The research was mainly supported by NIH grants EB003730 (P.G.), GM59944 (P.G.), and by EY018230 “NIH Nanomedicine Development Center: Phi29 DNA packaging Motor for Nanomedicine (P.G.).” The development of the UltraScan software is supported by NIH Grant R022200 (B.D.). We thank the staff at TACC for support and acknowledge NSF Teragrid Allocation TG-MCB070038 (B.D.).

Note added after ASAP publication: The original Web version of this paper published on July 17, 2009 was corrected for two numerical values, a unit designation, and a reference citation, and the paper was reposted on the Web July 27, 2009.

REFERENCES AND NOTES


