J|A|C|S
COMMUNICATIONS

Published on Web 03/22/2006

A Simple Tagging System for Protein Encapsulation

Florian P. Seebeck,[§] Kenneth J. Woycechowsky,[§] Wei Zhuang,[‡] Jürgen P. Rabe,[‡] and Donald Hilvert*,[§]

Laboratorium für Organische Chemie, ETH Zürich, Hönggerberg HCI F337, CH-8093, Zürich, Switzerland, and Department of Physics, Humboldt University Berlin, Newtonstrasse 15, D-12489 Berlin, Germany

Received December 9, 2005; E-mail: hilvert@org.chem.ethz.ch

The encapsulation of cargo molecules by proteinaceous containers represents a potentially powerful strategy for the manufacture of novel materials,[1−3] catalysts,[4] or drug delivery systems.[5] Indeed, many natural proteins self-assemble to form capsids that function as containers in a variety of biologically important processes. Examples include catalysis, iron storage by ferritin, the sequestration of non-native proteins by chaperones, and RNA/DNA transport by viral capsids.[6−8] Many protein capsids are constructed by the self-assembly of multiple identical subunits. These complexes recognize their guests through noncovalent interactions and exploit the cooperativity afforded by symmetry to gain affinity and specificity. Protein engineering has been used to achieve the encapsulation of novel guests, often by hijacking naturally evolved recognition elements to serve in a new context. In a typical manifestation, the encapsulation of non-natural guests by virus particles has been driven via the fusion of a capsid-targeting peptide (derived from the physiological guest) to small molecules or other proteins. By engineering both the capsid and the cargo, binding interactions can be convergently designed and thus alleviate the need to mimic natural systems.

Here we describe an encapsulation system (Figure 1, left) that utilizes a simple tagging scheme based on charge complementarity. The lumazine synthase from *Aquifex aeolicus* (AaLS) represents the container component, as it forms icosahedral capsids large enough to encapsulate guests such as nanoparticles[9] or other proteins.[7] To engineer the charge environment inside the capsid, four residues per monomer that project into the lumen—Arg[83], Thr[86], Thr[120], and Gln[123]—were mutated to glutamates (Figure 1, right). As AaLS capsids contain either 60 or 180 subunits, these mutations could produce up to 240 or 720 extra negative charges, respectively, spread over the interior surface. In addition, six histidines were appended to the C-terminus of the protein to facilitate purification. The resulting AaLS variant (AaLS-neg) was overproduced in *Escherichia coli* and purified by Ni^{2+}-affinity chromatography. Like hyperthermophilic wild-type AaLS,[10] AaLS-neg resists thermal unfolding up to 95 °C.

In principle, adding a short stretch of positively charged amino acids to a cargo protein should promote its specific encapsulation by AaLS-neg. We chose to implement this strategy by fusing a deca-arginine (R_{10}) tag to the C-terminus of the easily detectable green fluorescent protein (GFP). To test this design, we coproduced (1) GFP-R_{10} and AaLS-wt (wild-type AaLS with six histidines appended to the C-terminus), (2) GFP and AaLS-wt, (3) GFP-R_{10} and AaLS-neg, and (4) GFP and AaLS-neg. By measuring GFP-specific fluorescence before and after affinity purification for all four samples, we determined the fraction of GFP (R_{10}-tagged or untagged) encapsulated by AaLS-wt or AaLS-neg (Figure 2, left). When coproduced, a significant amount of GFP-R_{10} co-purified with

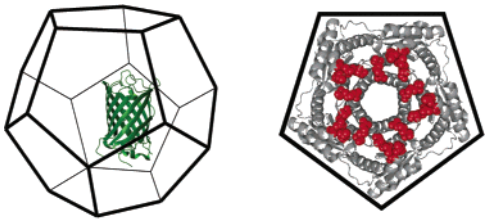

Figure 1. Left: The envisioned encapsulation of GFP inside a lumazine synthase capsid, based on the crystal structure of the $T = 1$ capsid formed by *A. aeolicus* lumazine synthase. Right: A pentamer of *A. aeolicus* lumazine synthase viewed from the capsid interior. The four residues per monomer that were mutated to glutamates are shown in red.

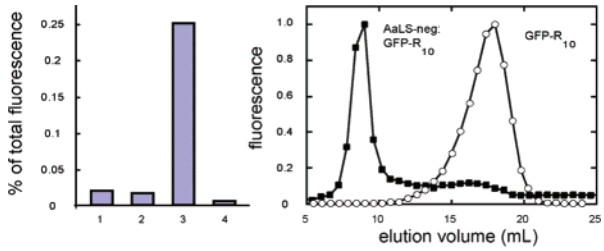

Figure 2. Left: The dependence of GFP:AaLS co-purification on electrostatic interactions. The GFP-specific fluorescence measured in protein samples purified by Ni^{2+}-affinity chromatography, expressed as a percentage of the total (GFP-specific) fluorescence in crude cell lysates, is shown for cells producing (1) GFP-R_{10} and AaLS-wt; (2) GFP and AaLS-wt; (3) GFP-R_{10} and AaLS-neg; (4) GFP and AaLS-neg. All values were measured in triplicate, and all errors are within 7% of the plotted values. Right: Analytical size-exclusion chromatography traces of GFP-R_{10} encapsulated by AaLS-neg (■) and free GFP-R_{10} (○). Fluorescence of collected fractions was measured at 500 nm after excitation at 450 nm.

AaLS-neg, consistent with a physical association between the tagged protein and the engineered capsid. Further, this association depends on both design features—negative charges in the container and positive charges on the guest. The electrostatic interaction between AaLS-neg and GFP-R_{10} does not involve nonspecific adsorption of guest molecules to charged residues on the capsid surface as coproduction and purification of this protein pair yields a much larger fluorescence signal compared to that of the other protein pairs.

We used size-exclusion chromatography to further examine the association of AaLS-neg and GFP-R_{10}. Following coproduction and affinity purification, AaLS-neg eluted from a Sephacryl-300 gel filtration column as a single peak that contained the bulk of the GFP-based fluorescence observed within the chromatogram (Figure 2, right). When produced in the absence of AaLS-neg, GFP-R_{10} elutes much later, and thus has a much lower apparent mass. Reinjection of the AaLS-neg:GFP-R_{10} complex onto the Sephacryl-300 column about 6 h after its initial elution gave a trace similar to the first injection of this complex, indicating slow dissociation of the guest from the container.

[§] ETH Zürich.
[‡] Humboldt University Berlin.

10.1021/ja058363s CCC: $33.50

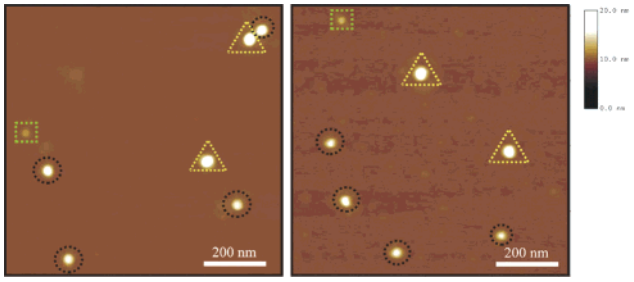

Figure 3. SFM tapping mode height images of AaLS-wt (left) and AaLS-neg:GFP-R_{10} (right). Particles in each sample were grouped into three size categories and are marked with squares (small particles), circles (medium particles) and triangles (large particles).

In addition to complex formation between the two components of the tagging system, the success of this design strategy also required that capsid assembly not be disrupted by the mutations to the AaLS scaffold. Although the X-ray structure of AaLS shows a 60-subunit assembly that forms a so-called $T = 1$ icosahedron with a mass of 1.1 MDa,[10] gel filtration chromatography suggested a mass of >2 MDa for AaLS-neg. Sedimentation equilibrium data for AaLS-neg:GFP-R_{10} fit well to a single species with a mass of 3.1 MDa, which is consistent with a $T = 3$ icosahedron built from 180 subunits. In comparison, a solution of AaLS-wt seems to include a mixture of the $T = 1$ and $T = 3$ states.[11] The larger average size of the engineered capsid might minimize charge density on the interior surface of the capsid, which would be favored because of Coulombic repulsion. A mutation-induced shift of the AaLS icosahedron toward the $T = 3$ state is not unprecedented.[12] More generally, switching between different icosahedral states in response to small genetic changes[13] or chemical stimuli[7,14] has been reported for a number viral or virus-like capsids.

To confirm that the AaLS-neg:GFP-R_{10} complexes do indeed form discrete structures that are similar to AaLS-wt, we visualized individual particles of both by scanning force microscopy (SFM) in tapping mode. Particles in three size categories (small, medium, and large) were seen for each of the samples, but with different distributions (Figure 3). For AaLS-wt, the average particle diameters[15] within each size category and their frequencies of occurrence are 10.7 ± 4.8 nm, 10% of the total population (small particles); 19.6 ± 9.0 nm, 67% (medium particles); and 36.1 ± 8.3 nm, 23% (large particles). For AaLS-neg:GFP-R_{10}, the average particle sizes within each size category and their frequencies of occurrence are 10.6 ± 5.2 nm, 16% of the total population (small particles); 19.2 ± 8.8 nm, 46% (medium particles); and 30.4 ± 8.8 nm, 38% (large particles). We attribute the medium and large size categories to the $T = 1$ and $T = 3$ state icosahedra, which have expected diameters of 15.4[10] and 29 nm,[16] respectively. The particles in the smallest size category are presumably capsids that partially decomposed during deposition on the SFM slides. The increase in the population of large particles seen for AaLS-neg:GFP-R_{10}, relative to AaLS-wt, is in qualitative agreement with the shift toward the $T = 3$ state seen in the sedimentation equilibrium experiments. For the engineered capsid, nearly half of the particles are medium-sized, which might reflect a shift of the equilibrium within the sample toward the $T = 1$ state during SFM sample preparation. AaLS-neg:GFP-R_{10} also exhibits more frequent deviations from a spherical shape, compared to AaLS-wt. These asymmetries could be caused by adsorption onto the sample slide or by the SFM tip tapping. For AaLS-neg:GFP-R_{10}, the higher general susceptibility to deformation, together with the higher percentage of particles populating the small size category (relative to AaLS-wt), suggests that the engineered capsids may be somewhat less stable than the wild-type capsids.

How many molecules of GFP-R_{10} does one AaLS-neg capsid accommodate? From a comparison of the total protein concentration in a purified sample of AaLS-neg:GFP-R_{10} with its GFP-R_{10} content, we estimate that one AaLS-neg capsid (in the $T = 3$ state) hosts 3.8 ± 0.8 molecules of GFP-R_{10}. This loading corresponds to a cargo concentration of about 22 mg/mL in the lumen of the capsid and suggests that considerably larger guests than GFP could be encapsulated by this host.[17]

In conclusion, we have described a simple design strategy to produce a proteinaceous host−guest complex in living cells. Since encapsulation by AaLS-neg depends on a C-terminal R_{10} sequence, this method should allow for the general encapsulation of diverse tagged cargo proteins. The simple design represents a promising starting point for further rounds of engineering or directed evolution to improve encapsulation efficiency. As such, this system may provide a valuable blue-print for future steps toward the design of multifunctional complexes of biomolecules.

Acknowledgment. We thank Prof. Dr. A. Bacher for the generous gift of the AaLS gene, and Joris Beld for assistance with the analytical ultracentrifugation studies. This work was supported by the ETH Zürich, the Schweizerischer Nationalfonds, and National Institutes of Health Fellowship 1 F32 GM071126-01A1 (to K.J.W.).

Supporting Information Available: Experimental details are provided for sample preparation and protein characterization. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Douglas, T.; Dickson, D. P. E.; Betteridge, S.; Charnock, J.; Garner, C. D.; Mann, S. *Science* **1995**, *269*, 54−57.
(2) Douglas, T.; Strable, E.; Willits, D.; Aitouchen, A.; Libera, M.; Young, M. *Adv. Mater.* **2002**, *14*, 415−418.
(3) Douglas, T.; Young, M. *Nature* **1998**, *393*, 152−155.
(4) Varpness, V.; Peters, J. W.; Young, M.; Douglas, T. *Nano Lett.* **2005**, *5*, 2306−2309.
(5) Abbing, A.; Blaschke, U. K.; Grein, S.; Kretschmar, M.; Stark, C. M. B.; Thies, M. J. W.; Walter, J.; Weigand, M.; Woith, D. C.; Hess, J.; Reiser, C. O. A. *J. Biol. Chem.* **2004**, *279*, 27410−27421.
(6) Kerfeld, C. A.; Sawaya, M. R.; Tanaka, S.; Nguyen, C. V.; Phillips, M.; Beeby, M.; Yeates, T. O. *Science* **2005**, *309*, 936−938.
(7) Bacher, A.; Ludwig, H. C.; Schnepple, H.; Ben-Shaul, Y. *J. Mol. Biol.* **1986**, *187*, 75−86.
(8) Goodsell, D. S.; Olson, A. J. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 105−153.
(9) Shenton, W.; Mann, S.; Colfen, H.; Bacher, A.; Fischer, M. *Angew. Chem., Int. Ed.* **2001**, *40*, 442−445.
(10) Zhang, X.; Meining, W.; Fischer, M.; Bacher, A.; Ladenstein, R. *J. Mol. Biol.* **2001**, *306*, 1099−1114.
(11) Among several different one- or two-component models tested, the sedimentation equilibrium data for AaLS-wt fit best to a monomer−trimer equilibrium, with the masses of the monomer and trimer corresponding to the $T = 1$ and $T = 3$ states, respectively. Still, the deviations between the data and the fitted curve suggest that it may not fully describe the complexity of this sample. AaLS-neg:GFP-R_{10} was also fit to a monomer−trimer equilibrium model, but the presence of capsids in the $T = 1$ state was not apparent. See also Supporting Information.
(12) Morgunova, E.; Meining, W.; Illarionov, B.; Haase, I.; Jin, G.; Bacher, A.; Cushman, M.; Fischer, M.; Ladenstein, R. *Biochemistry* **2005**, *44*, 2746−2758.
(13) Satheshkumar, P. S.; Lokesh, G. L.; Sangita, V.; Saravanan, V.; Vijay, C. S.; Murthy, M. R.; Savithri, H. S. *J. Mol. Biol.* **2004**, *342*, 1001−1014.
(14) Krol, M. A.; Olson, N. H.; Tate, J.; Johnson, J. E.; Baker, T. S.; Ahlquist, P. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 13650−13655.
(15) These values were corrected for broadening effects due to a finite SFM tip radius, assuming a spherical tip and a spherical particle (Butt, H.-J.; Guckenberger, R.; Rabe, J. P.; *Ultramicroscopy* **1992**, *46*, 375−393).
(16) The diameter of lumazine synthase from *Bacillus subtilis* in the $T = 3$ state has been previously measured by electron microscopy (ref 7).
(17) For comparison, we estimate that, theoretically, there is enough space to fit approximately 120 GFP molecules inside the $T = 3$ capsid, assuming a packing density for GFP of 0.7, a GFP volume of 41 nm^3, and a lumenal volume of 7200 nm^3 for the AaLS capsid in the $T = 3$ state. A loading capacity of 36 guest molecules per capsid, assuming one GFP-R_{10} binding site per AaLS-neg pentamer, may represent a more practical limit.

JA058363S