JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Communication

pubs.acs.org/JACS

In Vivo Encapsulation of Nucleic Acids Using an Engineered Nonviral Protein Capsid

Seth Lilavivat, Debosmita Sardar, Subrata Jana, Geoffrey C. Thomas, and Kenneth J. Woycechowsky*

Department of Chemistry, University of Utah, 315 South 1400 East, Salt Lake City, Utah 84112, United States

Supporting Information

ABSTRACT: In Nature, protein capsids function as molecular containers for a wide variety of molecular cargoes. Such containers have great potential for applications in nanotechnology, which often require encapsulation of non-native guest molecules. Charge complementarity represents a potentially powerful strategy for engineering novel encapsulation systems. In an effort to explore the generality of this approach, we engineered a nonviral, 60-subunit capsid, lumazine synthase from *Aquifex aeolicus* (AaLS), to act as a container for nucleic acid. Four mutations were introduced per subunit to increase the positive charge at the inner surface of the capsid. Characterization of the mutant (AaLS-pos) revealed that the positive charges lead to the uptake of cellular RNA during production and assembly of the capsid *in vivo*. Surprisingly, AaLS-pos capsids were found to be enriched with RNA molecules approximately 200–350 bases in length, suggesting that this simple charge complementarity approach to RNA encapsulation leads to both high affinity and a degree of selectivity. The ability to control loading of RNA by tuning the charge at the inner surface of a protein capsid could illuminate aspects of genome recognition by viruses and pave the way for the development of improved RNA delivery systems.

Many proteins self-assemble to form icosahedral capsids, in which a nearly spherical protein shell encloses a hollow interior. In Nature, such architectures are commonly used as molecular containers for the encapsulation of guest molecules such as inorganic compounds, other proteins, or nucleic acids. Consequently, protein capsids have recently attracted much attention from nanotechnologists as promising scaffolds for the development of applications including inorganic synthesis, nanoscale reaction chambers, bioimaging, and drug delivery.[1,2] Tailoring protein capsids to encapsulate desired cargo molecules is fundamental to exploiting these scaffolds and remains a major challenge.[3]

Viruses possess perhaps the most well-known class of protein capsids. In simple RNA viruses, protein capsids act as both storage containers and delivery vehicles for the nucleic acid molecule(s) that constitute the viral genome. RNA viruses seem to exploit charge complementarity as a common chemical strategy to build host–guest complexes between the protein capsid and viral genomic RNA.[4] For example, both structural and thermodynamic studies identify Coulombic interactions between positively charged arginines and the negatively charged phosphate backbone of RNA as playing a crucial role in RNA encapsulation by the bacteriophage MS2 capsid.[5]

Indeed, many RNA virus capsids possess an arginine-rich motif (ARM) at their interiors.[6,7] Various biophysical, computational, and mutagenesis studies provide support for the hypothesis that ARMs play a crucial role in RNA encapsulation by many virus capsids.[8−15] This notion is consistent with the observation that "supercharging" the surface of proteins, by installing a high density of positively charged residues, can confer affinity for nucleic acids *in vitro*.[16] These observations suggest that it should be possible to convert a nonviral protein capsid into a container for nucleic acids by generating a supercharged, arginine-rich environment at the inner surface.

The capsid formed by *Aquifex aeolicus* lumazine synthase (AaLS) provides an appealing starting point for engineering a nonviral nucleic acid encapsulation system. This protein self-assembles into hollow 60-subunit dodecahedral capsids, which can be heterologously overproduced in *Escherichia coli*.[17] Examination of the high-resolution structure of wild-type AaLS reveals that the interior surface possesses a net negative charge at pH 7, as the number of acidic residues (D90, D119, E122, and E126) is larger than the number of basic residues (R86, R127, and K131). Previously, the number of negative charges at the interior surface of AaLS was increased to generate a protein encapsulation system.[18] By substituting glutamates for three neutral residues and one basic residue on the capsid interior, a variant (AaLS-neg) was engineered that encapsulated guest proteins bearing a positively charged deca-arginine tag upon co-production in *E. coli*. The directed evolution of AaLS-neg led to a new capsid variant (AaLS-13) that more efficiently encapsulated positively charged guest proteins, presumably as a consequence of mutations that further increased the density of negatively charged residues on the interior.[19] Indeed, up to 100 positively supercharged green fluorescent protein molecules can be loaded into a single AaLS-13 capsid, demonstrating the power of charge complementarity as a strategy for biomolecular encapsulation.[20] In this study, we invert this charge complementarity approach to target RNA guests.

To generate affinity for negatively charged nucleic acid guests, the interior of the AaLS capsid was modified to increase positive charge (Figure 1). Four amino acid substitutions were introduced per monomer: T86R, D90N, T120R, and E122R. Assuming a 60-subunit assembly, these substitutions would

Received: March 21, 2012
Published: July 24, 2012

   dx.doi.org/10.1021/ja302743g | *J. Am. Chem. Soc.* 2012, 134, 13152–13155

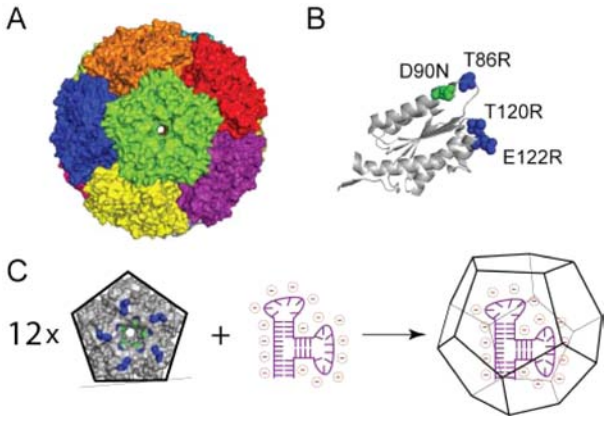

Figure 1. Design of AaLS-pos. (A) AaLS forms a 60 subunit, 1 MDa, dodecahedral capsid. The crystal structure is displayed as a surface model and the otherwise identical pentameric building blocks are colored differently (PDB accession 1HQK). (B) Ribbon diagram of a monomer extracted from the AaLS capsid with the amino acids that were mutated to create AaLS-pos highlighted. (C) Scheme for encapsulation of nucleic acids by AaLS-pos. A pentameric building block of AaLS is shown as a gray surface with mutated residues colored as in panel B. Cellular RNA is shown as a purple cartoon. Fully formed capsid is represented as a wire frame of a dodecahedron.

yield a net charge of up to +240 inside the capsid, compared to a net charge of −60 in the wild-type capsid. In addition to these inner surface mutations, six histidines were appended to the C-terminus of this variant (AaLS-pos), as well as the wild-type protein (AaLS-wt), to facilitate affinity purification.

The ability of AaLS-pos to associate with cellular nucleic acid was tested by producing the protein in *E. coli* BL-21(DE3) cells. Following Ni^{2+}-NTA affinity and gel filtration chromatography purification, the UV spectrum of AaLS-pos displayed a dramatically different shape than that of AaLS-wt and AaLS-neg (Figure 2). The most striking feature of the AaLS-pos UV spectrum is a pronounced peak at 260 nm which suggests the presence of nucleic acids in the protein sample. Indeed, AaLS-pos exhibits an A_{260}/A_{280} ratio of 1.3 which corresponds to a proportion of nucleic acid of approximately 7.5%, whereas pure protein has an A_{260}/A_{280} ratio of approximately 0.6.[21,22] In contrast, AaLS-wt and AaLS-neg have A_{260}/A_{280} ratios of 0.8 and 0.7, respectively, indicating that nucleic acid associated with AaLS-pos is a consequence of the engineered positive charges.

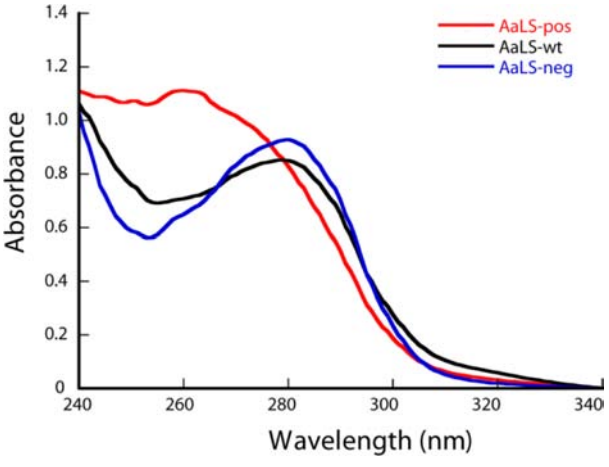

Figure 2. UV spectra of AaLS-wt (black), AaLS-pos (red), and AaLS-neg (blue) following purification by Ni^{2+}-NTA affinity chromatography.

The gel filtration chromatogram of AaLS-pos closely resembles that of AaLS-wt, suggesting that capsid assembly was unaffected by the mutations. The capsid structure of AaLS-pos was further investigated by transmission electron microscopy (TEM). The images reveal that AaLS-pos assembles as spherical particles that are uniform in size and distribution, and have nearly the same outer diameter, 16.7 ± 1.7 nm, as the AaLS-wt capsids, 16.8 ± 2.1 nm (Figure 3A,B).

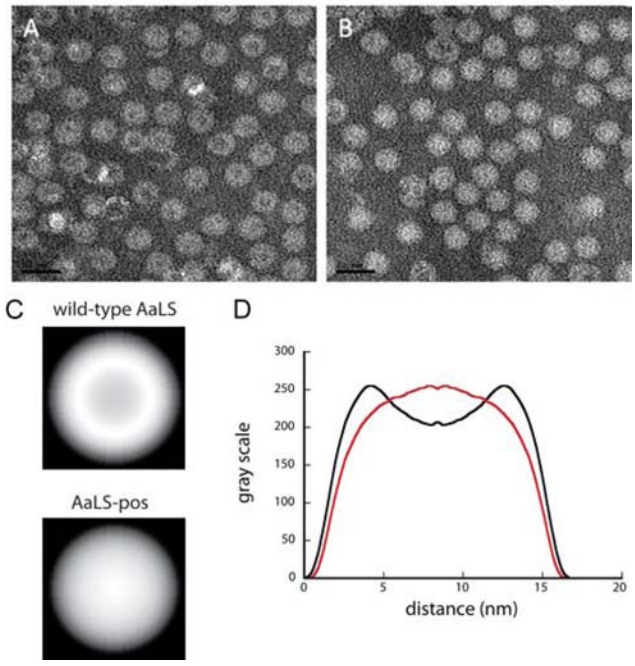

Figure 3. TEM analysis of capsids. (A and B) Micrographs of AaLS-wt capsids (A) and AaLS-pos capsids (B) stained with phosphotungstic acid. Scale bars are 20 nm. (C) Aligned and circularly averaged images of 230 AaLS-wt capsids (top image) and 171 AaLS-pos capsids (lower image). (D) Gray scale vs capsid diameter. The gray scale of the average images in panel C was measured and plotted against distance in nm. Curves for AaLS-wt and AaLS-pos are shown in black and red, respectively. A higher gray scale value represents a lighter color, whereas a lower gray scale value represents a darker color.

This analysis is consistent with the size distribution indicated by gel filtration chromatography (Figure S1). Therefore, it is likely that AaLS-pos forms a 60-subunit dodecahedron much like AaLS-wt. The assembly of AaLS-pos provides an interesting contrast to AaLS variants with negatively charged interiors (AaLS-neg and AaLS-13), which show large expansions in the capsid structure.[18,19] Indeed, the unperturbed assembly of AaLS-pos may stem from efficient charge saturation provided by the RNA guests. For some virus capsids, proper assembly has been shown to depend on the presence of negatively charged guest molecules to overcome Coulombic repulsion between the capsid protein subunits.[23−25]

The TEM experiments also reveal that the nucleic acid associated with AaLS-pos is localized inside the capsid. In these negatively stained images, dark signal in the capsid interior indicates a hollow space in which the stain has accumulated, whereas those particles with light centers are presumably already filled with guest molecules that prevent penetration of the stain. The majority of AaLS-wt particles have dark interiors,

dx.doi.org/10.1021/ja302743g | J. Am. Chem. Soc. 2012, 134, 13152−13155

while the centers of AaLS-pos assemblies appear to be relatively light (Figure 3A,B). To verify this difference, hundreds of capsid images were aligned and circularly averaged. The signal intensity was measured as a function of capsid diameter, confirming that the interiors of AaLS-wt capsids are significantly darker than those of AaLS-pos (Figure 3C,D). Thus, the AaLS-wt capsids are empty, whereas the AaLS-pos capsids are filled. On the basis of this difference, we conclude that the nucleic acid in the AaLS-pos sample has been encapsulated.

Upon agarose gel electrophoresis, AaLS-wt and AaLS-pos both ran as single, compact bands, confirming their similar capsid assemblies. While both samples were stained by Coomassie blue dye, AaLS-pos could also be visualized by SYBR green II staining (Figure 4A,B, lane 2 vs 3), further

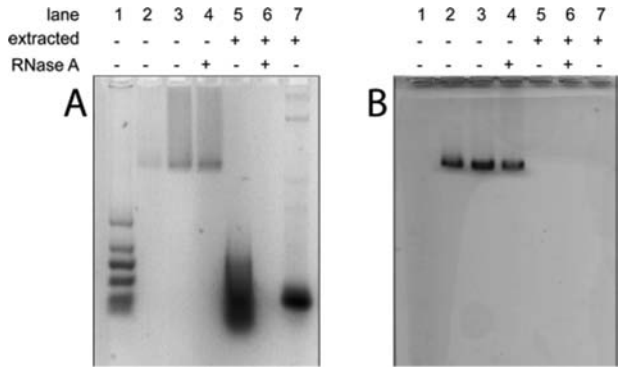

Figure 4. Agarose gel electrophoresis of encapsulated and unencapsulated RNAs. (A) Agarose gel stained with SYBR green II. Lane 1 contains RNA standards (1000, 500, 300, 150, 80, and 50 bases in length). AaLS-wt is in in lane 2. AaLS-pos samples are in lanes 3−6. RNA extracted from whole *E. coli* cells are in lane 7. Lane numbers and sample treatments are indicated above the gel. An "extracted" sample was subjected to an organic extraction that separates nucleic acid from protein, and the aqueous layer (containing nucleic acid) was loaded onto the gel. "RNase A" refers to incubation of the sample with RNase A prior to loading onto the gel. (B) The same gel shown in panel A was stained in Coomassie blue to reveal the mobility of the AaLS-wt and AaLS-pos proteins.

confirming its association with nucleic acid. The nucleic acid in the AaLS-pos sample was resistant to nuclease degradation (Figure 4, lane 4, and Figure S2). An additional nuclease protection assay was performed in which the complex was incubated with nuclease and then repurified by gel filtration chromatography. Following this treatment, the nucleic acid content of the capsid did not change appreciably, as assessed by A_{260}/A_{280} (Table S1). We also attempted to disassociate the nucleic acid from the capsid by increasing the ionic strength of the buffer.[26] Following dialysis and gel filtration chromatography in 2 M NaCl, AaLS-pos retained an elevated A_{260}/A_{280} ratio (Table S1). While the high salt concentration should destabilize the protein−nucleic acid interactions in the complex, AaLS-pos prevents nucleic acid from diffusing away under these conditions because the RNA is physically trapped within the capsid walls. The ability of this complex to withstand treatment with either nuclease or high salt concentration further supports encapsulation of nucleic acid by AaLS-pos.

The associated nucleic acid was isolated from purified AaLS-pos capsids by organic extraction. The extraction method segregates RNA into the aqueous phase, while protein and DNA partition into the organic phase and interphase, respectively. A substantial amount of material was isolated in the aqueous phase and was found to have an A_{260}/A_{280} ratio in the range of 1.8−2.0, characteristic of pure nucleic acid[21,22] (Figure S3). Approximately 40 μg of nucleic acid was obtained from 1 mg of protein. Analysis by agarose gel electrophoresis and SYBR green II staining indicates that the isolated nucleic acid undergoes a dramatic gel shift relative to the unextracted AaLS-pos sample, giving a smear of much higher mobility. Interestingly, this smear also has a much higher intensity than the band of the intact AaLS-pos sample (Figure 4A, lane 3 vs 5), suggesting that the nucleic acid associated with intact capsids is relatively inaccessible to the dye or binds the dye weakly due to conformational constraints. The extracted nucleic acid is sensitive to degradation by ribonuclease A (RNase A) (Figure 4A, lane 5 vs 6), but not DNase I (Figure S4). Thus, the isolated nucleic acid is RNA rather than DNA.

The simplest model for nucleic acid encapsulation by AaLS-pos assumes that RNA molecules are localized inside the capsid via purely electrostatic interactions. Since all RNA molecules have the same negative charge density, AaLS-pos should show little-to-no guest specificity apart from a size limit; all RNA molecules small enough to fit inside the capsid should be encapsulated with equal efficiencies. Consistent with this model, a comparison of RNA extracted from purified AaLS-pos capsids with RNA extracted from *E. coli* whole-cell lysates by agarose gel electrophoresis (Figure 4A, lane 5 vs 7) reveals that the high molecular weight RNAs from whole cells are absent in the sample extracted from AaLS-pos and that the RNAs extracted from AaLS-pos range from approximately 30 to 500 bases in length.

To test this model further, we performed urea-PAGE analysis of the nucleic acid isolated from AaLS-pos and this sample shows a long smear between 50 and 400 bases (Figure S5), similar to that seen in the agarose gel (Figure 4, lane 5). Within this smear, however, the majority of the signal occurs at roughly 200−350 bases. In contrast, many distinct bands are resolved within the whole cell lysate sample, and most of the signal occurs either below 100 bases or above 500 bases. The AaLS-pos sample contains very little signal above 400 bases, supporting the notion that a size limit exists for guest encapsulation. We estimate that the theoretical size limit for RNA guests of AaLS-pos is around 900 bases, using a spherical capsid interior with a diameter of 9 nm, a packing density of 0.7, and the reported volumes of the four nucleotides.[27] Interestingly, the whole-cell lysate sample displays relatively little signal between 200 and 350 bases. The different pattern of signal displayed between the samples reflects different relative abundances, suggesting that simple charge-complementarity is not sufficient to explain guest binding by AaLS-pos. Perhaps this length of RNA provides an optimum in which the cargo is sufficiently long to saturate the interior charges on the capsid by itself but not long enough to require extensive folding of the RNA in order to fit inside the capsid. Additional factors might also contribute to the observed specificity pattern, including intrinsic preferences of the capsid interior for binding certain RNA structures or sequences, competition with cellular factors for RNA binding, and the timing and subcellular location of RNA production relative to capsid assembly.

To investigate the importance of the protein capsid architecture for binding nucleic acid, we engineered a variant of lumazine synthase from *Saccharomyces cerevisiae* (ScLS-pos) that contained the mutations T96R, E100N, T130R, and E132R (ScLS numbering), which are homologous to those

made in AaLS-pos (Figure S6A). ScLS shares homology to AaLS but assembles as a homopentamer rather than as a capsid.[28] Following the standard protein production in *E. coli*, ScLS-pos also co-purified with nucleic acid (Table S1 and Figure S6B), but eluted from the gel filtration column much earlier than pentameric, wild-type ScLS. TEM imaging of ScLS-pos shows heterogeneous, amorphous particles (Figure S7) rather than the homogeneous, well-ordered capsids seen with AaLS-pos. Thus, the capsid structure of AaLS-pos is not essential for binding nucleic acid. Interestingly, ScLS-pos does not shield its bound RNA against degradation by RNase A (Table S1 and Figure S6B). This observation suggests that the capsid structure of AaLS-pos plays an important role in providing nuclease resistance (see Supporting Information for a more detailed analysis of AaLS-pos vs ScLS-pos).

It is important to understand the supramolecular chemistry of nucleic acid encapsulation both to gain insight into how viruses work[4,7] and to enable the development of novel molecular encapsulation systems.[29−31] We have reported here the design and characterization of an engineered protein capsid that can act as a container for cellular nucleic acids. Host−guest affinity in this system is based on charge-complementarity, a general strategy that potentially suits a broad range of applications. Guest specificity is broad but significant, and depends on RNA size as well as other, more subtle, factors. The elucidation of these specificity factors will require further study and might illuminate important aspects of discrimination by viral capsids between genomic and host-cell RNA. Further, the ability to control guest selection by protein capsids should spur the development of next-generation delivery vehicles for RNA therapeutics.

■ ASSOCIATED CONTENT

Supporting Information

Complete experimental procedures as well as tables and figures showing RNA−protein association and protein assembly. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

kwoycech@chem.utah.edu

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank M. Standing of the BYU microscopy lab for assistance collecting the EM images and D.M. Belnap for help with EM data analysis. Funding for this work was provided by the University of Utah.

■ REFERENCES

(1) Douglas, T.; Young, M. *Science* **2006**, *312*, 873−875.
(2) Flenniken, M. L.; Uchida, M.; Liepold, L. O.; Kang, S.; Young, M. J.; Douglas, T. *Curr. Top. Microbiol. Immunol.* **2009**, *327*, 71−93.
(3) Papapostolou, D.; Howorka, S. *Mol. BioSyst.* **2009**, *5*, 723−732.
(4) Rossmann, M. G.; Johnson, J. E. *Annu. Rev. Biochem.* **1989**, *58*, 533−573.
(5) Peabody, D. S. *EMBO J.* **1993**, *12*, 595−600.
(6) Lazinski, D.; Grzadzieska, E.; Das, A. *Cell* **1989**, *59*, 207−218.
(7) Rao, A. L. N. *Annu. Rev. Phytopathol.* **2006**, *44*, 61−87.
(8) Nassal, M. *J. Virol.* **1992**, *66*, 4107−4116.
(9) Tang, L.; Johnson, K. N.; Ball, L. A.; Lin, T.; Yeager, M.; Johnson, J. E. *Nat. Struct. Biol.* **2001**, *8*, 77−83.
(10) Krishna, N. K.; Marshall, D.; Schneemann, A. *Virology* **2003**, *305*, 10−24.
(11) Zhang, D.; Konecny, R.; Baker, N. A.; McCammon, J. A. *Biopolymers* **2004**, *75*, 325−337.
(12) Annamalai, P.; Apte, S.; Wilkens, S.; Rao, A. L. N. *J. Virol.* **2005**, *79*, 3277−3288.
(13) Pogam, S. L.; Chua, P. K.; Newman, M.; Shih, C. *J. Virol.* **2005**, *79*, 1871−1887.
(14) Devkota, B.; Petrov, A.; Lemieux, S.; Burak Boz, M.; Tang, L.; Schneemann, A.; Johnson, J. E.; Harvey, S. C. *Biopolymers* **2009**, *91*, 530−538.
(15) Venter, P. A.; Marshall, D.; Schneemann, A. *J. Virol.* **2009**, *83*, 2872−2882.
(16) Lawrence, M. S.; Phillips, K. J.; Liu, D. R. *J. Am. Chem. Soc.* **2007**, *129*, 10110−10112.
(17) Zhang, X.; Meining, W.; Fischer, M.; Bacher, A.; Ladenstein, R. *J. Mol. Biol.* **2001**, *306*, 1099−1114.
(18) Seebeck, F. P.; Woycechowsky, K. J.; Zhuang, W.; Rabe, J. P.; Hilvert, D. *J. Am. Chem. Soc.* **2006**, *128*, 4516−4517.
(19) Wörsdörfer, B.; Woycechowsky, K. J.; Hilvert, D. *Science* **2011**, *331*, 589−592.
(20) Wörsdörfer, B.; Pianowski, Z.; Hilvert, D. *J. Am. Chem. Soc.* **2012**, *134*, 909−911.
(21) Nelson, D. L. , Cox, M. M. *Lehninger Principles of Biochemistry*; 5th ed.; W.H. Freeman and Company: New York, 2008.
(22) Glasel, J. A. *BioTechniques* **1995**, *18*, 62−63.
(23) Sikkema, F. D.; Comellas-Aragonès, M.; Fokkink, R. G.; Verduin, B. J. M.; Cornellisen, J. J. L. M.; Nolte, R. J. M. *Org. Biomol. Chem.* **2007**, *5*, 54−57.
(24) Minten, I. J.; Ma, Y.; Hempenius, M. A.; Vancso, G. J.; Nolte, R. J. M.; Cornellisen, J. J. L. M. *Org. Biomol. Chem.* **2009**, *7*, 4685−4688.
(25) Kwak, M.; Minten, I. J.; Anaya, D.-M.; Musser, A. J.; Brasch, M.; Nolte, R. J. M.; Müllen, K.; Cornellisen, J. J. L. M.; Herrmann, A. *J. Am. Chem. Soc.* **2010**, *132*, 7834−7835.
(26) Minten, I. J.; Wilke, K. D. M.; Hendriks, L. J.; van Hest, J. C. M.; Nolte, R. J.; Cornelissen, J. J. *Small* **2011**, *7*, 911−919.
(27) Voss, N. R.; Gerstein, M. *J. Mol. Biol.* **2005**, *346*, 477−492.
(28) Meining, W.; Mörtl, S.; Fischer, M.; Cushman, M.; Bacher, A.; Ladenstein, R. *J. Mol. Biol.* **2000**, *299*, 181−197.
(29) Fiedler, J. D.; Brown, S. D.; Lau, J. L.; Finn, M. G. *Angew. Chem., Int. Ed.* **2010**, *49*, 9648−9651.
(30) Lau, J. L.; Baksh, M. M.; Fiedler, J. D.; Brown, S. D.; Kussrow, A.; Bornhop, D. J.; Ordoukhanian, P.; Finn, M. G. *ACS Nano* **2011**, *5*, 7722−7729.
(31) Lu, X.; Thompson, J. R.; Perry, K. L. *J. Gen. Virol.* **2012**, *93*, 1120−1126.