J|A|C|S COMMUNICATIONS

Published on Web 11/19/2009

Controlled Encapsulation of Multiple Proteins in Virus Capsids

Inge J. Minten, Linda J. A. Hendriks, Roeland J. M. Nolte, and Jeroen J. L. M. Cornelissen*,†

Institute for Molecules and Materials, Radboud University Nijmegen, Heyendaalseweg 135, 6525 AJ Nijmegen, The Netherlands

Received September 15, 2009; E-mail: j.j.l.m.cornelissen@tnw.utwente.nl

Viruses are masters of encapsulating nucleic acids; they are typically composed of a shell of highly organized protein molecules, which surround densely packed DNA or RNA chains. Recently, these protein shells (also called capsids) have also been used to package other guest compounds.[1−4] A virus that is particularly suitable for this purpose is the Cowpea Chlorotic Mottle Virus (CCMV), an RNA plant virus 28 nm in size,[5] which can be disassembled and reassembled reversibly after removal of the RNA by adjusting the pH. At pH 7.5 it is dissociated into 90 separate dimers, while at pH 5.0 the capsid is formed. This makes it possible to encapsulate compounds by mixing them with CCMV dimers at pH 7.5 and subsequently lowering the pH to entrap the material inside the capsid.[1,6]

Since the CCMV capsid interior is positively charged, encapsulation works especially well with negatively charged guest molecules such as inorganic salts and negatively charged polymers.[1,7] However, potentially useful cargos such as enzymes are usually not negatively charged at acidic pH, resulting in a low encapsulation efficiency of these enzymes. Low encapsulation efficiency is not always a disadvantage, as has been shown in our study of a virus-based enzyme nanoreactor[6] that has provided new insight in the behavior of single-enzyme molecules.

As a follow-up to this work, it is of great interest to study the behavior of multiple enzymes in the confined space of a virus nanoreactor. For example, cascade reactions catalyzed by a series of enzymes might take place more efficiently. In addition, such systems may provide new insights into the way living cells work, since viral nanoreactors could serve as a model system for simple cell organelles.

In order to achieve this goal, we need to improve and control the encapsulation efficiency of enzymes in the CCMV capsid relative to the statistical encapsulation strategy previously reported by us.[6] In this paper, we describe a novel encapsulation method that makes use of a noncovalent anchoring moiety to attach the target protein to the capsid protein prior to assembly of the latter. In this way, the guest encapsulation efficiency is significantly improved, while to a large extent, control over the amount of encased proteins is obtained.

Noncovalent anchoring was chosen over covalent anchoring methods because it eliminates the need for chemical reactions (which are often not bio-orthogonal) to bind the two proteins together. As a noncovalent anchor, we used a heterodimeric coiled-coil protein.[8] Coiled-coil motifs are oligomerization domains that are frequently found in nature[9,10] and are also used in the laboratory, in, for example, affinity chromatography,[11] biosensors,[12] hydrogels,[13] and the study of liposome fusion.[14] Coiled-coil motifs consist of a seven-residue repeat, denoted as $(abcdefg)_n$, in which residues a and d are usually hydrophobic amino acids that align when the peptide sequence folds into a helical conformation. A hydrophobic strip is subsequently formed at one side of the coiled coil, which is the primary driving force for the self-assembly of two such species.[15] While nature commonly uses homodimeric coiled coils, we chose a heterodimeric coiled coil to avoid the problem of self-dimerization of either the capsid protein or the guest protein. To minimize the risk that the introduction of a coiled coil would alter the structure and assembly properties of the capsid protein, a short coiled coil that nonetheless still has a small dissociation constant ($K_d = 7 \times 10^{-8}$ M) was chosen.[8]

Heterodimerization can be achieved by the placement of charged residues (glutamic acid or lysine) at the e and g positions the coiled-coil motif. The resulting coiled coil with glutamic acid at these positions is called the E-coil, and the one with lysine at these positions is called the K-coil. In this particular case, we denote the seven-residue repeat as $(gabcdef)_n$ in order to stick with a and d as the hydrophobic amino acids; the complete amino acid sequences of these peptides are $(EIAALEK)_3$ for the E-coil and $(KIAALKE)_3$ for the K-coil.

Since the capsid interior is positively charged, we chose to attach the positively charged K-coil to the capsid protein, ensuring that the coiled coil does not stick to the capsid protein surface, which presumably would make it less available for binding with the E-coil. In order to attach the K-coil to the capsid protein, we cloned the gene encoding for the coiled-coil sequence in an expression plasmid containing the capsid protein gene, introducing the K-coil at the N-terminus of this protein. The N-terminus was chosen for modification since the C-terminal part of the capsid protein is involved in the dimerization and assembly of the capsid proteins. Since the E-coil and K-coil bind in a parallel fashion, the E-coil was introduced at the C-terminus of the enhanced green fluorescent protein (EGFP) by introducing it in an expression plasmid containing the gene sequence for EGFP and an N-terminal His-tag sequence. The modified proteins were brought to expression in *Escherichia coli* cells. In order to assemble and purify the EGFP−capsid protein complex, we used the N-terminal His-tag sequence of the EGFP to immobilize it on a Ni-NTA column. The E-coil is attached to the C-terminus of the EGFP and is therefore available for binding to the modified capsid protein, which was added in excess to the column containing the bound EGFP (Figure 1).

The modified capsid protein was allowed to bind for at least 1 h before all of the nonbound protein was washed away and the complex was eluted from the column with an excess of imidazole. To remove the imidazole, the complex was immediately dialyzed to a pH 7.5 buffer. In principle, only the complete capsid−EGFP complex should elute from the column, but small impurities and any EGFP protein that did not bind to a capsid protein could have coeluted from the column as well. The eluate was therefore further

† Present address: Laboratory for Biomolecular Nanotechnology, MESA+ Institute for Nanotechnology, University of Twente, P.O. Box 217, NL-7500 AE Enschede, The Netherlands.

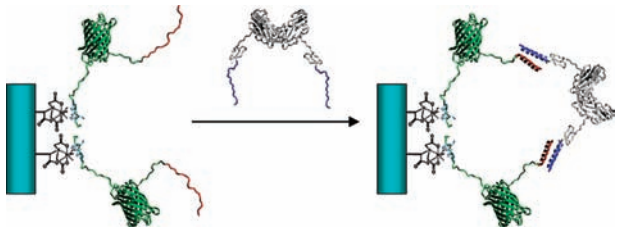

Figure 1. Purification of the EGFP−capsid protein complex. Bacterial lysate containing the EGFP is added to a Ni-NTA column. Only EGFP binds to the Ni-NTA with its N-terminal His-tag. A wash step removes all other proteins that lack the His tag. The lysate containing the capsid protein is then added. The capsid protein binds to the C-terminal coiled coil (red) of the EGFP with its N-terminal coiled coil (blue). After another wash step, the entire complex is eluted from the column using an excess of imidazole.

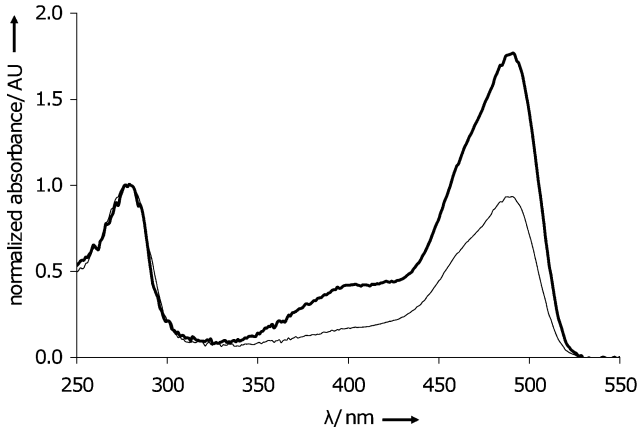

Figure 2. UV−vis spectra of EGFP (thick line) and EGFP complexed with capsid protein (thin line) at pH 7.5. Both absorption spectra have been normalized to an absorbance of 1.0 at $\lambda = 280$ nm.

purified on a fast-performance liquid chromatography (FPLC) system equipped with a Superdex 200 column.

The ratio with which EGFP was bound to the capsid protein could be determined by comparing the UV−vis spectra of pure EGFP with the spectra of the capsid−EGFP complex (Figure 2). Besides the normal protein absorption at $\lambda = 280$ nm, EGFP also has a specific absorption at $\lambda = 490$ nm at pH 7.5 that is approximately twice as intense as the $\lambda = 280$ nm absorption. However, when EGFP was bound to the capsid protein, which does not absorb at 490 nm, the absorption intensity at 490 nm was about equal to that for the 280 nm absorption. This suggests that the EGFP and capsid proteins are approximately bound in a 1:1 ratio (for a detailed calculation of this ratio, see the Supporting Information). Because the capsid protein always exists as a dimer, it is probable that a complex of a capsid protein dimer bound to two EGFP proteins was formed. This was further confirmed by FPLC analysis of the complex. At low concentrations, the EGFP−capsid protein complex eluted from the FPLC column at $V = 1.4$ mL, which is the expected elution volume for a complex of this molecular weight (see the Supporting Information).

In order to encapsulate EGFP, the EGFP−capsid protein complex was mixed with wild-type capsid protein isolated from the CCMV virus grown in plants (for procedures, see the Supporting Information). This was done in various ratios to allow control over the amount of encapsulated EGFP and prevent overcrowding of the capsid. After a short incubation period, this mixture was dialyzed to pH 5.0 (Figure 3) and analyzed by FPLC (Figure 4).

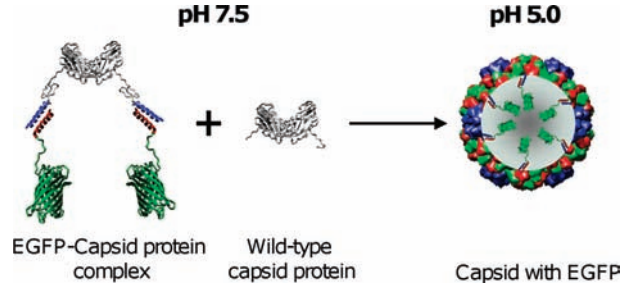

Figure 3. Schematic representation of EGFP encapsulation. The EGFP−capsid protein complex is mixed with wild-type capsid at pH 7.5 and subsequently dialyzed to pH 5.0 to induce capsid formation.

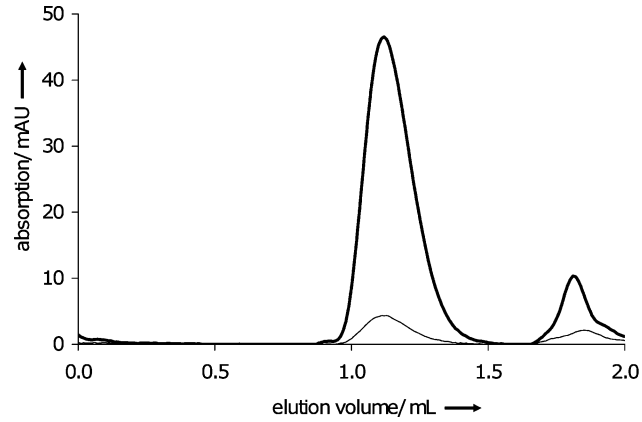

Figure 4. FPLC trace of a mixture of EGFP−capsid protein complex with wild-type capsid at pH 5.0. The FPLC system was equipped with a Superose 6 column. The thick line represents the protein absorption at $\lambda = 280$ nm and the thin line the EGFP-specific absorption at $\lambda = 395$ nm. The capsid with encapsulated EGFP eluted at $V = 1.1$ mL, and the unassembled capsid dimers and EGFP−capsid protein complex eluted at ~1.8 mL.

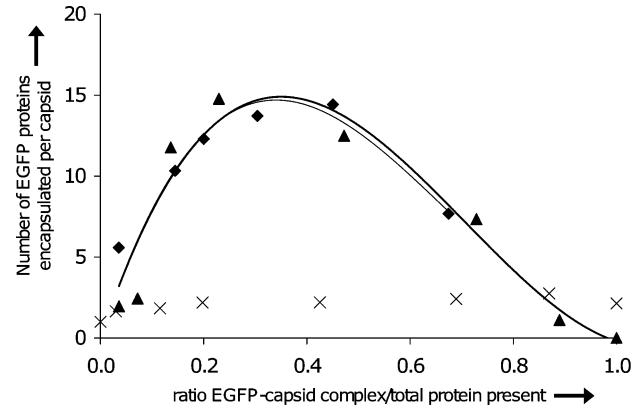

Figure 5. Number of encapsulated EGFP proteins per capsid as a function of the EGFP−capsid complex/total protein ratio. Diamonds and triangles represent data points of duplicate experiments. Crosses represent negative control experiments in which unbound EGFP with the E-coil was used instead of the EGFP−capsid protein complex.The thick line represents the polynomial trend line through the data points depicted with a triangular shape; the thin line represents the polynomial trend line through the diamond-shaped data points.

The amount of EGFP encapsulated per capsid can be calculated using the 280 nm/395 nm absorbance ratio of the capsid peak at $V = 1.1$ mL. Since the EGFP protein absorbs at both 280 and 395 nm at pH 5.0 while the capsid protein absorbs only at 280 nm, the

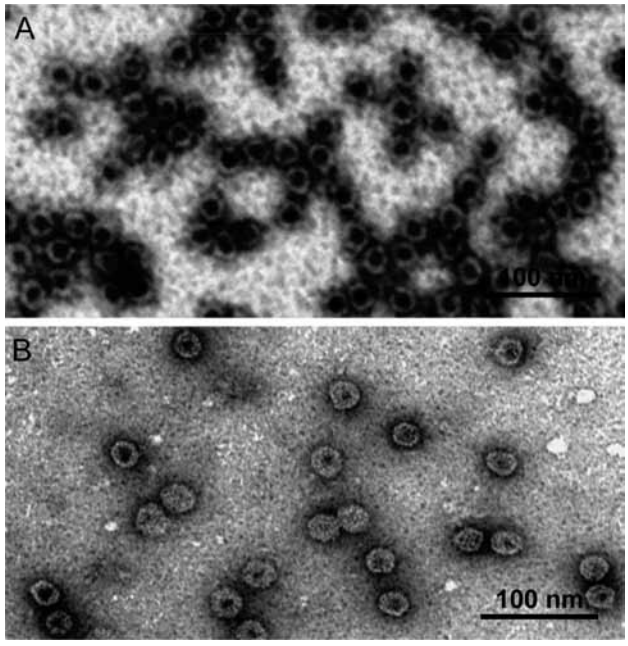

Figure 6. TEM micrographs of uranyl acetate-stained capsids: (A) empty wild-type capsids; (B) capsids filled with ~15 EGFP proteins per capsid.

capsid protein/EGFP ratio was determined from these absorption ratios using the respective extinction coefficients (for determination of the extinction coefficients and calculation of the number of encapsulated EGFP molecules, see the Supporting Information).

The data shows that it is possible to encapsulate up to 15 EGFP proteins per capsid (Figure 5) and that the amount of encapsulated EGFP to a certain extent can be controlled by varying the ratio between the wild-type capsid protein and the EGFP−capsid protein complex.

The experiments were highly reproducible. Transmission electron microscopy (TEM) analysis of the EGFP-filled capsids supported the notion that the capsids are largely filled with EGFP. Wild-type empty capsids showed up on TEM images as white spots with black interiors. The black interior was caused by accumulation of the staining agent (uranyl acetate) in the empty interior of the capsids (Figure 6a). When the capsids were filled with EGFP, however, the dark interiors largely disappeared, indicating that their cavities were no longer empty (Figure 6b). Not all capsids seemed to contain the same amount of EGFP, as indicated by the difference in dark appearance of the capsid interiors. These TEM images also provide evidence that the EGFP molecules are located on the inside of the capsids. This finding was further supported by dynamic light scattering data, which revealed that the EGFP-filled capsids are the same size as the empty capsids, and by the inability of anti-EGFP antibodies to bind to the outside of the EGFP-loaded capsids (see the Supporting Information).

From the data above, we may conclude that it is possible to bind proteins inside the CCMV capsid in an efficient and controlled fashion, opening the way to the encapsulation of different types of enzymes in one capsid. Our study furthermore shows that viral capsid assembly is a powerful tool for the controlled formation of bionanostructures with diverse functions.

Acknowledgment. This work was supported by the Chemical Council of The Netherlands Organization for Scientific Research (NWO-CW) (a TOP Grant to R.J.M.N. and a Vidi Grant to J.J.L.M.C.), the Royal Netherlands Academy for Arts and Sciences (KNAW) (grants to R.J.M.N. and J.J.L.M.C.), and the European Science Foundation (ESF) (EURYI Grant to J.J.L.M.C). A plasmid-encoding EGFP was a generous gift of R. Teeuwen.

Supporting Information Available: Full experimental details, extensive discussion of the location of EGFP in the capsids, and details of the calculations performed. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Douglas, T.; Young, M. *Nature* **1998**, *393*, 152.
(2) Loo, L.; Guenther, R. H.; Lommel, S. A.; Franzen, S. *Chem. Commun.* **2008**, 88.
(3) de la Escosura, A.; Verwegen, M.; Sikkema, F. D.; Comellas-Aragones, M.; Kirilyuk, A.; Rasing, T.; Nolte, R. J. M.; Cornelissen, J. J. L. M. *Chem. Commun.* **2008**, 1542.
(4) For an overview, see: Young, M.; Willits, D.; Uchida, M.; Douglas, T. *Annu. Rev. Phytopathol.* **2008**, *46*, 361.
(5) Speir, J. A.; Munshi, S.; Wang, G. J.; Baker, T. S.; Johnson, J. E. *Structure* **1995**, *3*, 63.
(6) Comellas-Aragones, M.; Engelkamp, H.; Claessen, V. I.; Sommerdijk, N.; Rowan, A. E.; Christianen, P. C. M.; Maan, J. C.; Verduin, B. J. M.; Cornelissen, J. J. L. M.; Nolte, R. J. M. *Nat. Nanotechnol.* **2007**, *2*, 635.
(7) Sikkema, F. D.; Comellas-Aragones, M.; Fokkink, R. G.; Verduin, B. J. M.; Cornelissen, J. J. L. M.; Nolte, R. J. M. *Org. Biomol. Chem.* **2007**, *5*, 54.
(8) Litowski, J. R.; Hodges, R. S. *J. Biol. Chem.* **2002**, *277*, 37272.
(9) Burkhard, P.; Stetefeld, J.; Strelkov, S. V. *Trends Cell Biol.* **2001**, *11*, 82.
(10) Siegert, R.; Leroux, M. R.; Scheufler, C.; Hartl, F. U.; Moarefi, I. *Cell* **2000**, *103*, 621.
(11) Tripet, B.; Yu, L.; Bautista, D. L.; Wong, W. Y.; Irvin, R. T.; Hodges, R. S. *Protein Eng.* **1996**, *9*, 1029.
(12) Contarino, M. R.; Sergi, M.; Harrington, A. E.; Lazareck, A.; Xu, J.; Chaiken, I. *J. Mol. Recognit.* **2006**, *19*, 363.
(13) Wang, C.; Stewart, R. J.; Kopeček, J. *Nature* **1999**, *397*, 417.
(14) Robson Marsden, H.; Elberss, N. A.; Bomans, P. H. H.; Sommerdijk, N. A. J. M.; Kros, A. *Angew. Chem., Int. Ed.* **2009**, *48*, 2330.
(15) Dong, H.; Hartgerink, J. D. *Biomacromolecules* **2006**, *7*, 691.

JA907843S