Modular, Bioorthogonal Strategy for the Controlled Loading of Cargo into a Protein Nanocage

Lise Schoonen,†‡§ Selma Eising,§# Mark B. van Eldijk, † Jaleesa Bressleers, † Margo van der Pijl, † Roeland J. M. Nolte,† Kimberly M. Bonger,*‡§ and Jan C. M. van Hest,*‡†

†Department of Bio-Organic Chemistry and ‡Department of Biomolecular Chemistry, Institute for Molecules and Materials, Radboud University, Heyendaalseweg 135, 6525 AJ Nijmegen, The Netherlands
§Laboratory of Bio-Organic Chemistry, Eindhoven University of Technology, PO Box 513 (STO 3.31), 5600 MB Eindhoven, The Netherlands

Abstract: Virus capsids, i.e., viruses devoid of their genetic material, are suitable nanocarriers for biomedical applications such as drug delivery and diagnostic imaging. For this purpose, the reliable encapsulation of cargo in such a protein nanocage is crucial, which can be accomplished by the covalent attachment of the compounds of interest to the protein domains positioned at the interior of the cage. This approach is particularly valid for the capsid proteins of the cowpea chlorotic mottle virus (CCMV), which have their N-termini located at the inside of the capsid structure. Here, we examined several site-selective modification methods for covalent attachment and encapsulation of cargo at the N-terminus of the CCMV protein. Initially, we explored approaches to introduce an N-terminal azide functionality, which would allow the subsequent bioorthogonal modification with a strained alkyne to attach the desired cargo. As these methods showed compatibility issues with the CCMV capsid proteins, a strategy based on 2-pyridinecarboxaldehydes for site-specific N-terminal protein modification was employed. This method allowed the successful modification of the proteins, and was applied for the introduction of a bioorthogonal vinylboronic acid moiety. In a subsequent reaction, the proteins could be modified further with a fluorophore using the tetrazine ligation. The application of capsid assembly conditions on the functionalized proteins led to successful particle formation, showing the potential of this covalent encapsulation strategy.

Supporting Information

Introduction

From a nanotechnological perspective, viruses have been recognized as interesting carrier tools, due to their unique ability to protect their nucleic acid cargo and deliver it to cells in their infected host. In particular, virus-like particles (VLPs), viruses that are devoid of their viral nucleic acids, are used for nanocarrier purposes, as they are relatively harmless without their viral genetic information and have room in their interior for loading of cargo such as catalysts, drugs, or imaging agents.

A remarkable example of a VLP that has proven very useful is the cowpea chlorotic mottle virus (CCMV), which has its N-termini located at the inside of the capsid. In contrast to most other VLPs, CCMV capsids show reversible assembly and disassembly behavior, even in the absence of their genetic material. The capsid proteins exist as dimers in solution at a physiological pH and spontaneously form 28-nm-sized capsids with T = 3 symmetry out of 90 capsid protein dimers when the pH is lowered to pH 5.0.1,2 In order to make the capsids more stable at neutral pH for in vivo applications, we previously introduced an elastin-like polypeptide (ELP) block at the N-terminus of the CCMV capsid protein.3 ELPs are stimulus-responsive polypeptides, consisting of repeating Val-Pro-Gly-Xaa-Gly (VPGXG) pentapeptides containing any natural amino acid guest residue, Xaa (X), except proline.4,5 These peptides can reversibly switch from a water-soluble state to a collapsed, hydrophobic state upon a change in the environmental conditions. The introduction of this peptide at the N-terminus of the CCMV capsid protein resulted in a new ELP-induced assembly pathway that is triggered by increasing the temperature or salt concentration, yielding smaller T = 1 capsids with a diameter of 18 nm, composed of 30 capsid protein dimers.3 The endogenous assembly pathway leading to T = 3 capsids could still be triggered by lowering the pH to 5.0. The development of this ELP-CCMV variant resulted in a wider range of conditions at which the capsids remained stable, showing promise for in vivo applications.

Controlled encapsulation of cargo in the ELP-CCMV capsid interior is crucial for its successful application as a nanocarrier.
A classical way to encapsulate cargo in CCMV capsids is via statistical encapsulation, where cargo is added to the capsid proteins when they are in the dimer state, after which assembly conditions are applied to induce capsid formation. This, however, normally yields partial encapsulation, while most of the cargo material is lost. To increase the encapsulation efficiency, the cargo can be equipped with negatively charged moieties, mimicking the charge of the endogenous nucleic acid cargo and allowing interactions with the cationic N-terminal tails of the CCMV capsid proteins.11,12 The latter method, however, is not applicable for cargo encapsulation in ELP-CCMV capsids, as the ELP sequence was inserted at the N-terminus in such a way that it replaced the cationic RNA-binding motif. Alternative methods that have been used for cargo encapsulation in CCMV VLPs are (i) genetic cloning of a target protein onto the N-terminus of the capsid proteins;13,14 (ii) the genetic introduction of an N-terminal domain, which can form noncovalent interactions with a specific complementary domain;15,16 and (iii) enzymatic modification of the N-terminus using the enzyme Sortase A.13,14 The latter methodology is possibly the most optimal for versatile and reproducible cargo encapsulation, since it is modular, does not require extensive genetic engineering, and yields a robust covalent link between the cargo and the capsid proteins. Even though this technique is applicable to any cargo that can be equipped with the required Sortase recognition sequence, it lends itself best to the encapsulation of peptides and proteins, as these allow easy introduction of this recognition peptide. In order to further broaden the modularity and possibilities for cargo loading, we set out to develop a method for the selective modification of the ELP-CCMV capsid protein N-termini, which eliminates the need for incorporation of a peptide into the cargo and is therefore more easily applied to small molecules such as (organo)catalysts or imaging agents.

In order to achieve site-specific N-terminal modification, we aimed to introduce a reactive handle onto the N-terminus that can be used in a bioorthogonal reaction, a reaction that is inert to any natural occurring biological functionalities, such as the copper-catalyzed alkyne−azide cycloaddition (CuAAC), strain-promoted alkyne−azide cycloaddition (SPAAC), or the inverse electron-demand Diels−Alder (iEDDA) reaction with tetrazines.15−17 Fortunately, several methodologies are available to achieve the site-selective modification of the N-terminal amine over amines present in lysine side chains, ranging from pH-controlled reactions to reactions requiring side chain participation.18 Here, we describe our efforts to site-selectively modify the N-terminus of the CCMV capsid protein with a bioorthogonal reagent, and subsequently attach a model cargo in a modular fashion (Scheme 1).

RESULTS AND DISCUSSION
Introduction of an Azide Function at the Capsid Protein N-Terminus. We started exploring a route toward the selective N-terminal modification of the ELP-CCMV capsid proteins by introducing an azide functionality, which is widely used as a bioorthogonal handle. Azides have been introduced into proteins cotranslationally using a genetic engineering approach,19,20 or post-translationally via chemical or enzymatic modification.21,22 Initially, we attempted the post-translational chemical modification by applying a selective diazotransfer reaction to convert the α-amine at the N-terminus into an azide.23 In proteins, this primary amine often has the lowest pKᵢ and can therefore be modified selectively using the right conditions. The reaction of ELP-CCMV with imidazole-1-sulfonyl azide 1 was performed in a diethanolamine buffer of pH 8.5 (Figure 1A). Analysis by mass spectrometry of the N-terminal fragments, obtained by tryptic digestion, revealed a mass shift of 26 Da, corresponding to a successful diazotransfer reaction (Figure 1B). The other lysine-containing fragments were also detected and did not show a mass shift. Unfortunately, we observed partial precipitation of the CCMV protein during the reaction and a significant amount of residual starting material, even after further optimization of the reaction conditions.

The observed precipitation of CCMV in the diazotransfer reaction prompted us to explore other options for the introduction of an N-terminal azide. For this, we investigated the residue-specific biosynthetic incorporation of an unnatural amino acid by using azidohomoaanine (Aha) as a methionine surrogate.19 ELP-CCMV contains one additional methionine residue, Met137, positioned in the capsid protein part of the fusion protein. To prevent undesired incorporation of Aha at this position, we mutated Met137 into an alanine residue. In addition, we inserted an additional arginine residue following this position, we mutated Met137 into an alanine residue. In order to prevent undesired incorporation of Aha at this position, we mutated Met137 into an alanine residue. In addition, we inserted an additional arginine residue following this position.23,24 This new construct was first expressed in E. coli in the presence of methionine to confirm that the adjustments did not alter the properties of the capsid proteins and that the N-terminal methionine would indeed stay intact. Analysis of the affinity-purified Met-ELP-CCMV by SDS-PAGE and ESI-TOF mass spectrometry showed an acceptable pure...
protein sample (Figure 2A) with the expected molecular weight (Figure S1). Investigation of the endogenous assembly pathway (lowering the pH to 5 for $T=3$ particles) of this new protein by size exclusion chromatography coupled to multiangle laser light scattering (SEC-MALLS) showed that the VLPs had a molecular weight of $4.0 \pm 0.1$ MDa (Figure 2B), which is in good agreement with the $T=3$ particles containing 180 capsid proteins of 22 570.7 Da. Further analysis with transmission electron microscopy (TEM) showed monodisperse spherical particles with a diameter of 28 nm ($T=3$ particles), confirming that the introduced modifications did not affect the assembling properties of the ELP-CCMV (Figure 2C).

Next, we expressed the modified ELP-CCMV construct in the presence of Aha in a methionine auxotrophic E. coli strain. Following similar purification and analysis procedures as above we confirmed the efficient incorporation of N-terminal Aha, indicated by the 5 Da observed mass difference upon replacement of methionine with Aha (Figure S1). Unfortunately, the protein was expressed with a 10x lower yield and SDS-PAGE analysis revealed many impurities in the obtained Aha-ELP-CCMV (Figure 2A). TEM analysis after pH-induced assembly clearly demonstrated the presence of 28-nm-diameter particles for Aha-ELP-CCMV (Figure 2C), whereas SEC analysis showed that CCMV proteins were only partially assembled, with the majority of the proteins being present in their dimer form (Figure 2B).3 Unfortunately, attempts to improve the expression yield and the purity of the azide-modified capsid proteins were unsuccessful.

Despite the fact that both the diazotransfer reaction and the genetic modification of CCMV did not yield flawless formation of azide-functionalized ELP-CCMV, we subsequently attempted to react the modified proteins further in a cycloaddition reaction. As we observed partial aggregation of the protein upon the addition of copper, which is needed for the CuAAC reaction, we decided to use the copper-free SPAAC reaction for the subsequent modification step.30 To this end, both azide-modified proteins were reacted with a commercially available fluorescently labeled cyclooctyne (BCN-lissamine-rhodamine B), after which fluorescently imaged SDS-PAGE analysis showed successful modification of the N-terminus of ELP-CCMV (Figure S2). Unfortunately, we observed reoccurring precipitation of the protein during the SPAAC reaction, which led us to investigate yet another alternative approach for the site-selective modification of the capsid proteins.

Site-Specific N-Terminal Modification Using 2-Pyridinecarboxaldehydes. Recently, a bioconjugation method for specific N-terminal protein modification was described based on 2-pyridinecarboxaldehyde (2PCA) forming an N-terminal cyclic imidazolidinone condensation product with peptides and proteins (Figure 3A).31 This modification proceeds in aqueous environments under mild temperature and pH and as no specific amino acid residue at the N-terminus is required, it is generally applicable to many proteins. Furthermore, many new functionalities have been introduced using this method, such as affinity tags, MRI-contrasting chelators, targeting agents, and fluorophores. As this benign method might circumvent

![Figure 1](image1.png)  
**Figure 1.** (A) Schematic representation of N-terminal modification of ELP-CCMV with imidazole-1-sulfonyl azide 1 to obtain $N_2$-ELP-CCMV. (B) MALDI-TOF mass spectra of the N-terminal tryptic fragment before (top) and after (bottom) diazotransfer. A mass shift of +26 Da is observed, corresponding to the diazotransfer of the N-terminal amine.

![Figure 2](image2.png)  
**Figure 2.** Characterization of Met-ELP-CCMV and Aha-ELP-CCMV. (A) SDS-PAGE analysis of Met-ELP-CCMV and Aha-ELP-CCMV after expression and Ni$^{2+}$ affinity purification. Protein bands were visualized with Coomassie blue staining. (B) SEC-MALLS chromatograms of pH-induced assemblies of Met-ELP-CCMV (black) and Aha-ELP-CCMV (purple) measured at 215 nm. Dotted line (black) shows molecular mass data of the Met-ELP-CCMV particles. (C) Uranyl acetate-stained TEM micrographs of Met-ELP-CCMV and Aha-ELP-CCMV after pH-induced assembly. Scale bars correspond to 200 nm.
unfavorable precipitation, we aimed to evaluate this 2PCA-based modification strategy on our ELP-CCMV capsid proteins.

Initial experiments focused on the optimization of the reaction conditions for the modification of ELP-CCMV with 2PCA. Following the reaction conditions as used by Francis et al. resulted in precipitation of the ELP-CCMV protein, due to the high reaction temperatures (Figure S3). This corresponded well with previous observations in our lab regarding the increased instability of the capsid proteins at elevated temperatures of 30 °C or higher. Performing the modification at room temperature, however, with 10 equiv of commercially available 2PCA for 24 h in PBS buffer did result in the formation of the desired product 2PCA-ELP-CCMV, as shown by ESI-TOF analysis (Figure 3B). Furthermore, the modification did not result in visual precipitation of the protein, when the samples were centrifuged after the modification. SDS-PAGE analysis of the samples, of which supernatant was transferred to a clean tube before sample preparation, gave comparable intensities for both proteins, demonstrating the applicability of this modification strategy to CCMV capsid proteins (Figure 3D). To determine the highest achievable modification yield, a series of 0, 1, 5, 10, 50, and 100 equiv of 2PCA was added to the capsid proteins and the conversions were estimated using ESI-TOF analysis. Improved conversion was observed upon increasing addition of 2PCA reaching a plateau of approximately 65% (Figure 3E, Figures S4, S5). Our results compare well with the research conducted by Francis and co-workers on the 2PCA modification, where conversions of 43% to >95% were achieved at 37 °C depending on the type of protein.

After having established that the modification strategy using 2PCA was applicable to the ELP-CCMV capsid protein, we introduced a functional handle onto the N-terminus using this strategy. The inverse electron-demand Diels–Alder reaction of tetrazines with alkenes and alkynes is one of the most popular bioorthogonal reactions due to its selectivity and high reaction rate.32,33 We decided to use the reaction between vinylboronic acids (VBAs) and dipyridyl-tetrazines, as the VBA moiety is hydrophilic, readily accessible, and stable under physiological conditions.34 Additionally, it was shown previously that this bioorthogonal reaction was suitable for protein modification.

To this end, 2PCA-VBA was designed, containing a 2PCA-piperazine group31 coupled to the pinacol protected vinyl-
boronic acid via a short linker (Figure 4A, SI - experimental section). The protected boronic ester was used instead of the free boronic acid, since the ester is synthetically more accessible and hydrolyzes rapidly to the boronic acid in aqueous media within 15 min. The water solubility of 2PCA-VBA was found to be slightly lower than that of 2PCA, and thus the concentration of ELP-CCMV was lowered so that optimized conditions used for the coupling of 2PCA could be used (100 equiv of the small molecule). The modification of ELP-CCMV with 2PCA-VBA was analyzed by ESI-TOF and indicated successful formation of VBA-ELP-CCMV with a yield of approximately 92% (Figure S6). Next, we performed the two-step protein modification of ELP-CCMV with 2PCA-VBA, and subsequently with dipyridyl-s-tetrazine containing a Cy5 fluorophore as the model cargo (Figure 4A and B). The formation of Cy5-ELP-CCMV was analyzed by SDS-PAGE and showed a significant fluorescent signal for the two-step labeling using 2PCA-VBA and tetrazine, indicating that the coupled boronic acid is available for a subsequent reaction with a dipyridyl-s-tetrazine (Figure 4C). Control reactions in which either 2PCA or no 2PCA derivative (only DMSO) was used instead of 2PCA-VBA showed no or only very low fluorescent signal, eliminating aspecific reactions of tetrazine-Cy5 with the protein. The modification of both 2PCA-VBA and tetrazine caused a small mass shift of CCMV on SDS-PAGE-gel, which was too small for calculation of the modification yields. SDS-PAGE analysis of the reaction of ELP-CCMV with a concentration range of 2PCA-VBA and subsequent ligation with tetrazine showed that the highest achievable modification was achieved using ~50 equiv of 2PCA-VBA (Figures 4D).

Finally, we investigated the assembly behavior of the modified capsid proteins. To this end, ELP-CCMV was reacted with 2PCA-VBA and tetrazine-Cy5, after which the protein was washed against PBS buffer and subsequently transferred to pH 5.0 capsid buffer by centrifugal filtration, prompting pH-induced assembly of the capsid proteins. SEC analysis of the resulting solutions clearly showed a capsid peak around an elution volume of 11 mL, indicating the formation of T = 3 particles; no residual capsid protein dimers were observed (Figure 5). The capsid peak absorbed light of 646 nm, demonstrating coelution with the fluorescent Cy5 dye. A control, in which 2PCA-VBA had been left out of the initial modification reaction, only showed minor absorbance at 646 nm, which might result from statistical encapsulation of a small residual amount of tetrazine-Cy5.

■ CONCLUSIONS

N-terminal modification of the ELP-CCMV capsid proteins has proven to be challenging due to reoccurring precipitation in various reaction conditions. The diazotransfer reaction for selective modification of the N-terminal amine to an azide...
resulted in significant protein precipitation, and subsequent attempts to introduce an azide-containing unnatural amino acids were not successful. Fortunately, the modification method using 2-pyrindinecarboxaldehyde was found to be suitable for ELP-CCMV capsid proteins. No significant protein instability was observed during the reactions, while up to 92% of the proteins could be modified using this strategy. We applied this method to attach a bioorthogonal vinylboronic acid handle, which could be further modified with a diptyridyl-tetrazine moiety linked to a fluorescent dye as a model cargo. Proof-of-principle reactions showed that this modular two-step modification strategy was successful, demonstrating that this method is suitable for encapsulating cargo into ELP-CCMV nanocages.

**EXPERIMENTAL PROCEDURES**

**Synthesis of (E)-(4-(2-((6-formylpyridin-2-yl)-methyl)pyraperazin-1-yl)-2-oxoethoxy)styryl)boronic acid pinacol ester (3).** tert-Butyl 4-(((6-formylpyridin-2-yl)-methyl)piperazine-1-carboxylate 27 (25 mg, 82 µmol, 1.0 equiv) was dissolved in dry CH2Cl2 (1 mL) under N2 and 4 M HCl in dioxane (205 µL, 820 µmol, 10.0 equiv) was added. The mixture was stirred for 2 h, whereupon the volatiles were evaporated. The solid was dissolved in DMF and (E)-(4-(2-((2,5-dioxopyrrolidin-1-yl)oxy)-2-oxoethoxy)styryl)boronic acid pinacol ester 4 (39 mg, 98 µmol, 1.2 equiv) was added. Then, Et3N (34 µL, 250 µmol, 3.0 equiv) was added and the solution was stirred for 2 h. The volatiles were evaporated and the product was purified by column chromatography (0 to 5% MeOH in EtOAc) yielding 2PCA-VBA (4.4 mg, 21%) as a blue solid.1H NMR (500 MHz, CD3OD) δ 7.37 (d, J = 8.2 Hz, 1H), 7.30 (d, J = 8.3 Hz, 1H), 6.65 (t, J = 12.3 Hz, 1H), 6.39–6.35 (m, 1H), 6.29 (d, J = 13.7 Hz, 1H), 4.20–4.13 (m, 4H), 4.05 (s, 2H), 2.39–2.31 (m, 2H), 1.92–1.84 (m, 4H), 1.79–1.73 (m, 14H), 1.58–1.49 (m, 2H), 1.42–1.37 (m, 2H). 13C NMR (125 MHz, CD3OD) δ 175.1, 174.0, 173.6, 169.1, 163.3, 163.1, 154.9, 154.8, 150.1, 149.8, 144.1, 143.6, 143.1, 142.0, 141.9, 138.7, 138.2, 127.1, 126.8, 126.7, 126.2, 124.8, 124.2, 120.0, 119.9, 114.0, 114.0, 103.9, 103.5, 49.21, 49.20, 43.7, 42.7, 38.8, 34.8, 26.6, 26.4, 26.30, 26.29, 25.6, 24.8, 11.1. HRMS (ESI+) m/z calc. for C27H34BN3O5 [M + H]+ 492.26698, found: 492.26703.

**Expression of ELP-CCMV.** The pET-15b-G-H6-[V,L4G1-9]-CCMV(ΔN26) vector encoding for the hexahistidine-tagged ELP-CCMV protein was previously constructed as described by van Eldijk et al. The expression was performed according to a literature procedure. For a typical expression, LB medium (50 mL), containing ampicillin (100 mg/L) and chloramphenicol (50 mg/L), was inoculated with a single colony of E. coli BLR(DE3)pLysS containing the pET-15b vector encoding for the ELP-CCMV capsid protein, and was incubated overnight at 37 °C. This overnight culture was used to inoculate 2× TY medium (1 L), supplemented with ampicillin (100 mg/L). The culture was grown at 37 °C and protein expression was induced during logarithmic growth (OD600 = 0.4–0.6) by addition of IPTG (1 mM). After 6 h of expression at 30 °C, the cells were harvested by centrifugation (2700 g, 15 min, 4 °C) and the pellets were stored overnight at −20 °C.

After thawing, the cell pellet was resuspended in lysis buffer (50 mM NaH2PO4, 1.3 M NaCl, 10 mM imidazole, pH 8.0; 20 mL). The cells were lysed by ultrasonic disruption (3 times 30 s, 100% duty cycle, output control 3, Branson Sonifier 250, Marius Instruments). Then, the lysate was centrifuged (16 400 × g, 15 min, 4 °C) to remove the cellular debris. The supernatant was incubated with Ni-NTA agarose beads (3 mL) for 1 h at 4 °C. The suspension was loaded onto a column, the flow-through was collected and the beads were washed twice with wash buffer (50 mM NaH2PO4, 1.3 M NaCl, 20 mM imidazole, pH 8.0; 20 mL). Then, the protein of interest was eluted from the column with elution buffer (50 mM NaH2PO4, 1.3 M NaCl, 250 mM imidazole, pH 8.0; 1 time 0.5 mL, 7 times 1.5 mL). The purification was analyzed by SDS-PAGE. The fractions containing the desired protein were combined and dialyzed against pH 7.5 dimer buffer to obtain the capsid protein dimers. For storage, the proteins were assembled by dialysis against pH 5.0 capsid buffer. The pure protein was obtained with a yield of 100 mg/L of bacterial culture. The purity of the proteins was verified by SDS-PAGE. The assembly properties of the capsid proteins and the geometry of the resulting capsids were analyzed by SEC using a Superose 6 GL 10/300 column with pH 5.0 capsid buffer as the eluent and by TEM. ESI-TOF: calculated 22 253.4 Da, found 22 253.5 Da.

**Stability Studies of ELP-CCMV.** ELP-CCMV was dialyzed to PBS buffer by spin filtration (10 kDa MWCO, 3 × 10 min) and diluted to 50 and 10 µM. Next, the samples (25 µL) were incubated at 21, 25, 30, and 37 °C for 24 h (400 rpm), whereupon the samples were centrifuged (1 min, 13 000 rpm) and the supernatant was transferred to a clean Eppendorf tube. Loss of protein in the form of precipitation/aggregation was monitored by loading the soluble protein fraction onto an SDS-PAGE gel, the 50 µM samples were diluted 5× to be able to compare them to the 10 µM samples (Figure S3).
**Acknowledgments**

Mr. Floris van Dalen and Dr. Martijn Verdoes are kindly acknowledged for the synthesis and characterization of sulfo-Cyanine5 carboxylic acid. L. Schoonen, M. B. van Eldijk, R. J. M. Nolte, and J. C. M. van Hest thank the Ministry of Education, Culture and Science (NWO Gravitation program 024.001.035) for financial support. R. J. M. Nolte also thanks the European Research Council (ERC Advanced Grant, 740295 ENCOPOL). S. Eising and K. M. Bonger acknowledge funding from The Netherlands Research Institute for Chemical Biology (NRSB) and the Institute of Molecules and Materials (IMM) of the Radboud University in Nijmegen.

**References**


