Electronic Supplementary Material (ESI) for RSC Advances.

Supporting Information

Insight into N-terminal localization and dynamics of engineered virus-like particles

Daan F. M. Vervoort, [+][a] Chiara Pretto [+][a] and Jan C. M. van Hest* [a]

[a] Eindhoven University of Technology, Institute for Complex Molecular Systems PO Box 513 (STO 3.41), 5600 MB Eindhoven, The Netherlands. Email: j.c.m.v.hest@tue.nl

[+] These authors contributed equally to this work.

Contents

1 Materials and methods

1.1 Materials

Ampicillin, Kanamycin and chloramphenicol were obtained from Sigma-Aldrich. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was purchased from UBPBio. Ni-NTA agarose beads were obtained from Qiagen. Restriction enzymes and T4 DNA ligase were obtained from New England Biolabs. Buffers

Table S1 – Composition of buffers

Name	Composition
pH 5.0 capsid buffer	50 mM NaOAc, 500 mM NaCl, 10 mM $MgCl_2$, 1 mM EDTA, pH 5.0
pH 7.5 100 mM NaCl buffer	50 mM Tris·HCl, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM EDTA, pH 7.5
pH 7.5 2000 mM NaCl buffer	50 mM Tris·HCl, 2500 mM NaCl, 10 mM $MgCl_2$, 1 mM EDTA, pH 7.5
PBS	137 mM NaCl, 2.7 mM KCl and 10 mM phosphate buffer solution
PBS 500 mM NaCl	500 mM NaCl, 2.7 mM KCl and 10 mM phosphate buffer solution
PBS 2000 mM NaCl	2000 mM NaCl, 2.7 mM KCl and 10 mM phosphate buffer solution
Sortase buffer	50 mM Tris·HCl, 500 mM NaCl, 10 mM $MgCl_2$, 5 mM $CaCl_2$, pH 7.5
Native Page loading buffer (2x)	62.5 mM Tris, 40% glycerol, 0.01% bromophenol blue
Native PAGE running buffer	25 mM Tris, 192 mM Glycine
Blotting buffer	25 mM Tris, 192 mM Glycine, 20% methanol
PBST	PBS, 0.1% tween
Blocking buffer	PBS, 5% milk, 0.1% tween

All buffers were filtered over a 0.2-micron filter prior to use.

1.2 Mass spectrometry

Protein mass characterization was performed using a High-Resolution LC-MS system consisting of a Waters ACQUITY UPLC I-Class system coupled to a Xevo G2 Quadrupole Time of Flight (Q-TOF). The system consisted of a Binary Solvent Manager and a Sample Manager with Fixed-Loop (SM-FL). Proteins were separated (0.3 mL/min) on the column (Polaris C18A reverse phase column, 2.0 x 100 mm, Agilent) using an acetonitrile gradient in water (15% to 75%, v/v) supplemented with formic acid (0.1%, v/v) before analysis in positive mode in the mass spectrometer. Deconvoluted mass spectra were obtained using the MaxENT1 algorithm in the Masslynx v4.1 (SCN862) software. Isotopically averaged molecular weights were calculated using the 'Protein Calculator v3.4' at http://protcalc.sourceforge.net. Protein samples were desalted by spin-filtration (Amicon® Ultra 0.5 mL) in Milli-Q prior to measurement.

1.3 General protocol for the expression of His_6-ELP-CCMV variants

The pET-15b-H_6-ELP-CCMV(ΔN26) and pET-15b-H_6-VY1-VY8-ELP-CCMV(ΔN26) vectors encoding for the two ELP-CCMV variants used in this paper were previously constructed.[1,2] ELP-CCMV capsid protein variants were expressed according to a standard expression procedure. Briefly, *E.coli* BLR(DE3)pLysS glycerol stocks containing the pET-15b-H_6-ELP-CCMV(ΔN26) and pET-15b-H_6-VY1-VY8-ELP-CCMV(ΔN26) vectors were used for an overnight culture at 37 °C in lysogeny broth medium (LB , 50 mL), containing ampicillin (100 mg/L) and chloramphenicol (50 mg/L). 2xTY medium (1 L), containing ampicillin (100 mg/L) was inoculated with the overnight culture and grown at 37 °C till an optimal optical density was reached (OD_{600} between 0.4 and 0.6). Protein expression was subsequently induced by addition of IPTG (1 mM) and the culture was incubated at 30°C for 6 hours. Cells were harvested by centrifugation (2700 g, 15 min, and 4 °C) and pellets were stored overnight at -20 °C. The cell pellet was resuspended in lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, pH 8.0; 25 mL). Cell lysis was performed by ultrasonic disruption (7-10 times 30 seconds, 70% amplitude, Branson Sonifier 150). Cell lysate was then centrifuged (16.400 g, 15 min, 4 °C) to remove bacterial debris. The supernatant was incubated with Ni-NTA

agarose beads (3 mL) for 1 hour at 4 °C, followed by column loading. The flow-through was collected and the column was washed twice with wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole, pH 8.0; 20 mL). The proteins of interest were eluted from the column with elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole, pH 8.0; 1 time 0.5 mL, 7 times 1.5 mL) and fractions containing histidine-tagged ELP-CCMV were combined and dialyzed against pH 7.5 dimer buffer (50 mM Trizma base, 500 mM NaCl, 10 mM $MgCl_2$, 1 mM EDTA, pH 7.5; 3 times 30 minutes using a 12-14 kDa tubing). Proteins were dialyzed against pH 5.0 capsid buffer (50 mM Trizma base, 500 mM NaCl, 10 mM $MgCl_2$, 1 mM EDTA, pH 5.0; 2 times 30 minutes, followed by overnight dialysis using a 12-14 kDa tubing) for stable storage at 4°C. Proteins were characterized by SDS-PAGE, Q-TOF, DLS and SEC (Fig. S1 and Fig. S2). Protein expression yields as well as Q-TOF results are listed in Table S2.

1.4 General protocol for the expression of Sortase A

E.coli BL21 glycerol stocks containing the pQE30 vector encoding for the Sortase A gene[3] were used for an overnight culture at 37°C in LB medium (50 mL), containing ampicillin (100 mg/L). The overnight culture was used to inoculate 2xTY medium (1 L) containing ampicillin (100 mg/L) and grown at 37 °C till an optimal optical density was reached (OD_{600} between 0.4 and 0.6). Protein expression was then induced by addition of IPTG (1 mM) and the culture was incubated overnight at 18°C. Cells were harvested by centrifugation (2700 g, 15 min, and 4 °C). The cell pellet was resuspended in lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, pH 8.0; 25 mL). Cell lysis was performed by ultrasonic disruption (7-10 times 30 seconds, 70% amplitude, Branson Sonifier 150). Cell lysate was then centrifuged (16,400 g, 15 min, 4 °C) to remove bacterial debris. The supernatant was incubated with Ni-NTA agarose beads (3 mL) for 1 hour at 4 °C, followed by column loading. The flow-through was collected and the column was washed twice with wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole, pH 8.0; 20 mL). The proteins of interest were eluted from the column with elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole, pH 8.0; 1 time 0.5 mL, 7 times 1.5 mL). Elution fractions containing the protein of interest were combined and dialyzed towards Sortase A storage buffer at 4° (50 mM HEPES, 150 mM NaCl, 5 mM $CaCl_2$ for 2x 45 mins, 1x overnight). Glycerol was added to get a 12% glycerol solution in Sortase A storage buffer and stored at -20°C. The protein was characterized by SDS-PAGE and Q-TOF (Fig. S3). Protein expression yields as well as Q-TOF results are listed in Table S2.

1.5 General protocol for the expression of (super folder) GFP

Gblocks containing the (superfolder) GFP sequence were ordered from Integrated DNA Technologies. Gblocks and the pET28a+ vector were treated with XhoI and NcoI restriction enzymes (1 hours, 37°C). Ligated constructs, using T4 DNA ligase (1 hour, 21°C), were transformed into XL1-blue competent cells and used for vector replication. Correct incorporation was confirmed by sanger sequencing (BaseClear). The pET28a+ vector encoding for the (superfolder) GFP was transformed into BL21(DE3) competent cells (Novagen) for protein expression. An overnight culture at 37°C (LB, 50 mL, 100 mg/L kanamycin) was used to inoculate 2xTY medium (1 L, 100 mg/L kanamycin) and grown (37 °C) till an optimal optical density was reached (OD_{600} between 0.4 and 0.6). Protein expression was then induced by addition of IPTG (1 mM) and the culture was incubated overnight at 18°C. Cells were harvested by centrifugation (2700 g, 15 min, and 4 °C) and pellets were resuspended in lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, pH 8.0) and lysed using an EmulsiFlexC3 High-Pressure homogenizer (Avestin) at 15000 psi three times. Cell lysate was then centrifuged (16,400 g, 15 min, 4 °C) to remove bacterial debris. The supernatant was incubated with Ni-NTA agarose beads (3 mL) for 1 hour at 4 °C, followed by column loading. The flow-through was collected and the column was washed twice with wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole, pH 8.0; 20 mL). The protein of interest was eluted from the column with elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole, pH 8.0; 1 time 0.5 mL, 7 times 1.5 mL). Elution fractions containing the protein of interest were combined and dialyzed towards PBS. For storage, protein was flash frozen

in liquid N_2 and stored at -80°C. The protein was characterized by SDS-PAGE and Q-TOF (Fig. S4). Protein expression yields as well as Q-TOF results are listed in Table S2.

Table S2 – Expression yields and Q-TOF results of the ELP-CCMV variants, Sortase A and GFP

Name	Yield (mg/L)	Q-TOF results	
		Calculated MW (Da)	Observed MW (Da)
Standard ELP-CCMV	47	22253.4	22253
VY1-VY8 ELP-CCMV	49	22381.5	22381
Sortase A	28	21947.5	21947
(super folder) GFP	81	29716.3*	29696

* mass calculated with Protein Calculator v3.4 of GFP results in a mass of 29716.3, however due to chromophore maturation the mass of the expressed GFP is 20 Da lower (29696.3).

1.6 Protein sequences

Table S3 – Amino Acid sequences of the ELP-CCMV variants

Name	Sequence
Standard ELP-CCMV	GHHHHHHVPGVGVPGLGVPGVGVPGLGVPGVGVPGLGVPGGGVPGVGVPGLGLEVVQPVIVEPIASGQGKAIKAWTGYSVS KWTASCAAAEAKVTSAITISLPNELSSERNKQLKVGRVLLWLGLLPSVSGTVKSCVTETQTTAAASFQVALAVADNSKDVVAAMY PEAFKGITLEQLTADLTIYLYSSAALTEGDVIVHLEVEHVRPTFDDSFTPVY
VY1-VY8 ELP-CCMV	GHHHHHHVPGYGVPGLGVPGVGVPGLGVPGVGVPGLGVPGGGVPGYGVPGLGLEVVQPVIVEPIASGQGKAIKAWTGYSVSK WTASCAAAEAKVTSAITISLPNELSSERNKQLKVGRVLLWLGLLPSVSGTVKSCVTETQTTAAASFQVALAVADNSKDVVAAMYP EAFKGITLEQLTADLTIYLYSSAALTEGDVIVHLEVEHVRPTFDDSFTPVY
Sortase A	TGSHHHHHHGSKPHIDNYLHDKDKDEKIEQYDKNVKEQASKDKKQQAKPQIPKDKSKVAGYIEIPDADIKEPVYPGPATPEQLNR GVSFAEENESLDDQNISIAGHTFIDRPNYQFTNLKAAKKGSMVYFKVGNETRKYKMTSIRDVKPTDVGVLDEQKGKDKQLTLITC DDYNEKTGVWEKRKIFVATEVK
(super folder) GFP	MVKMGASKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMK QHDFFKSAMPEGYVQERTISFKDDGTYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNFNSHNVYITADKQKNGIKANF KIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYKTLPETGENLYF QSGGSHHHHHH

Table S4 – Degree of labeling of either Cy5-labeled or Alexa532-labeled VY1-VY8 ELP-CCMV

Name	Degree of labelling per CP	Average absorbance (Au)*			Dye properties	
		280 nm	530 nm	646 nm	Correction factor	Extinction coefficient
Alexa532-labeled capsids	0.6	0.189	0.548	ND.	0.04	81000
Cy5-labeled capsids	1.3	0.184	ND.	0.184	0.09	271000

* Absorbance values of labeled VY1-VY8 ELP-CCMV were measured in triplicate at indicated wavelengths using a ND-1000 spectrophotometer.

2 Supplemental Figures

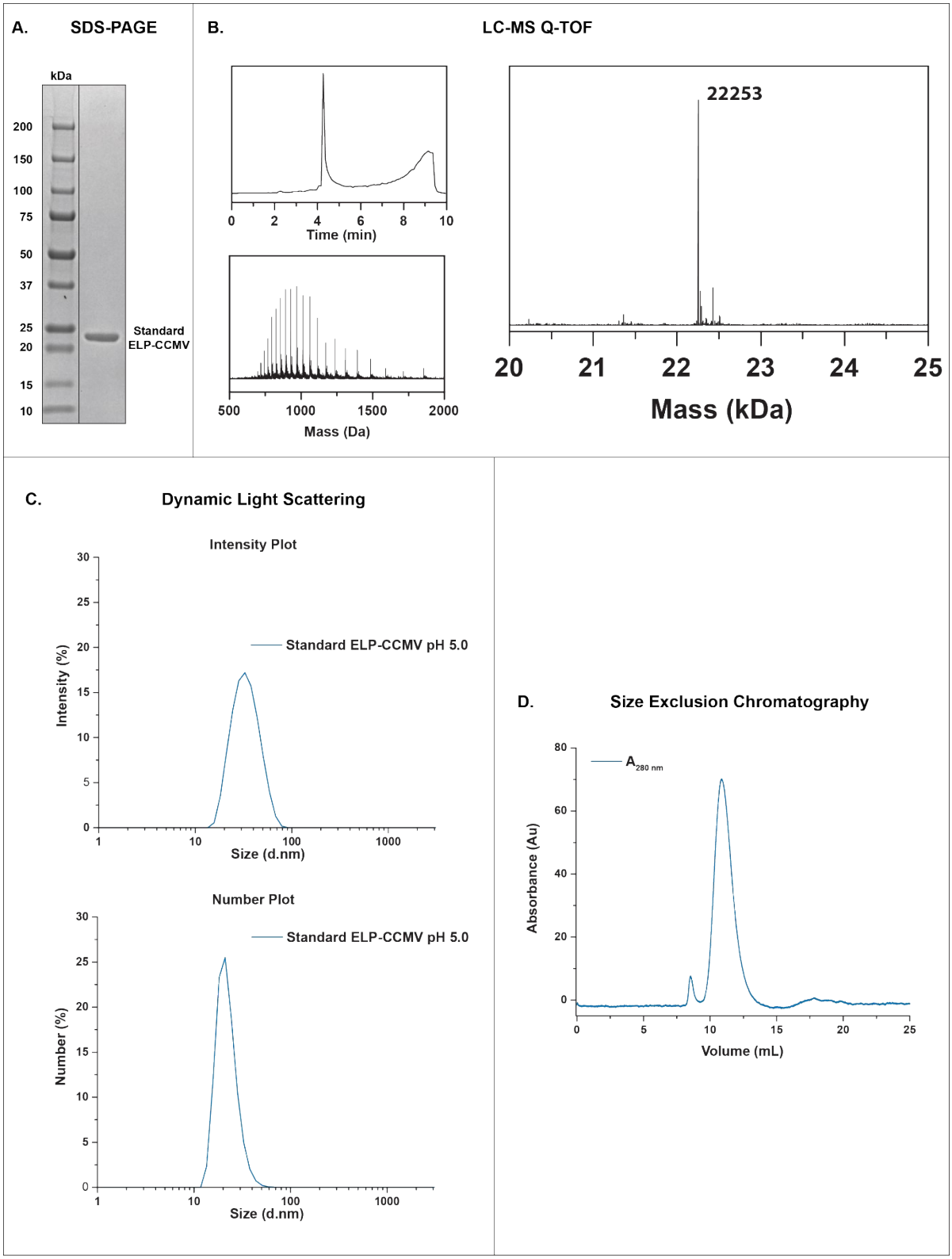

Fig. S1: Characterization of purified standard ELP-CCMV (calculated molecular weight is 22253.4 Da): A. SDS-PAGE; B. LC-MS Q-TOF, total ion count chromatogram (top left), m/z spectrum (bottom left) and mass spectrum (right); C. Dynamic Light Scattering, Intensity plot (top), Number plot (bottom); D. Size Exclusion Chromatography (pH-induced VLPs, elution volume ~11 mL).

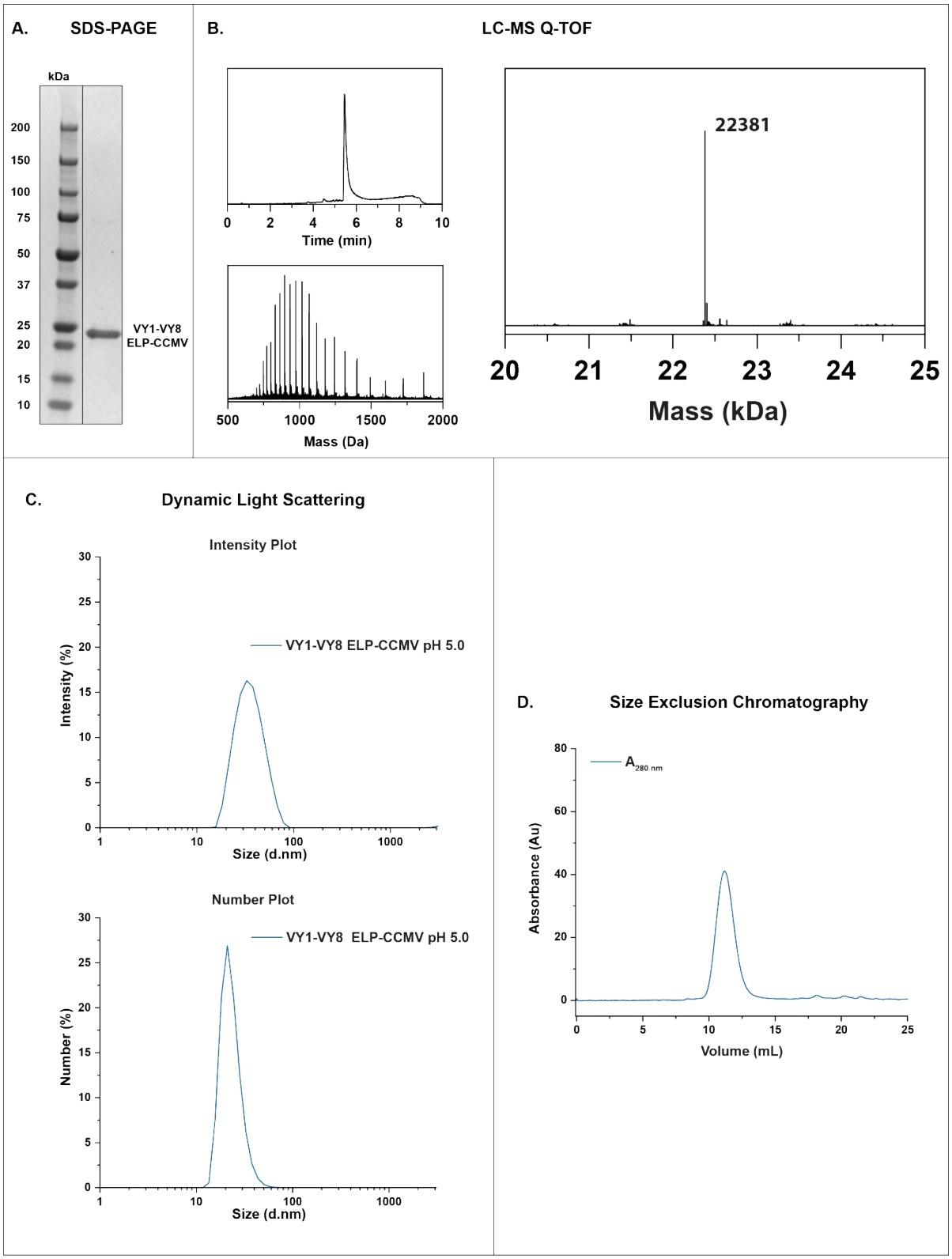

Fig. S2: Characterization of purified VY1-VY8 ELP-CCMV (calculated molecular weight is 22381.5 Da): A. SDS-PAGE; B. LC-MS Q-TOF, total ion count chromatogram (top left), m/z spectrum (bottom left) and mass spectrum (right); C. Dynamic Light Scattering, Intensity plot (top), Number plot (bottom); D. Size Exclusion Chromatography (pH-induced VLPs, elution volume ~11 mL).

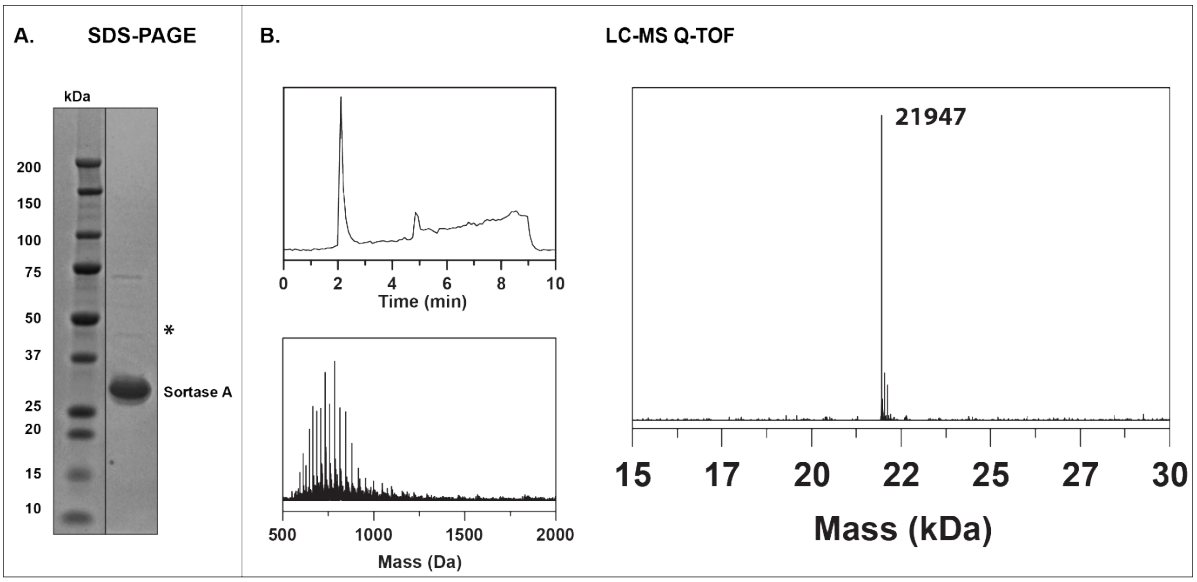

Fig. S3: Characterization of purified Sortase A (calculated molecular weight is 21947.5 Da): A. SDS-PAGE, *Impurity; B. LC-MS Q-TOF, total ion count chromatogram (top left), m/z spectrum (bottom left) and mass spectrum (right).

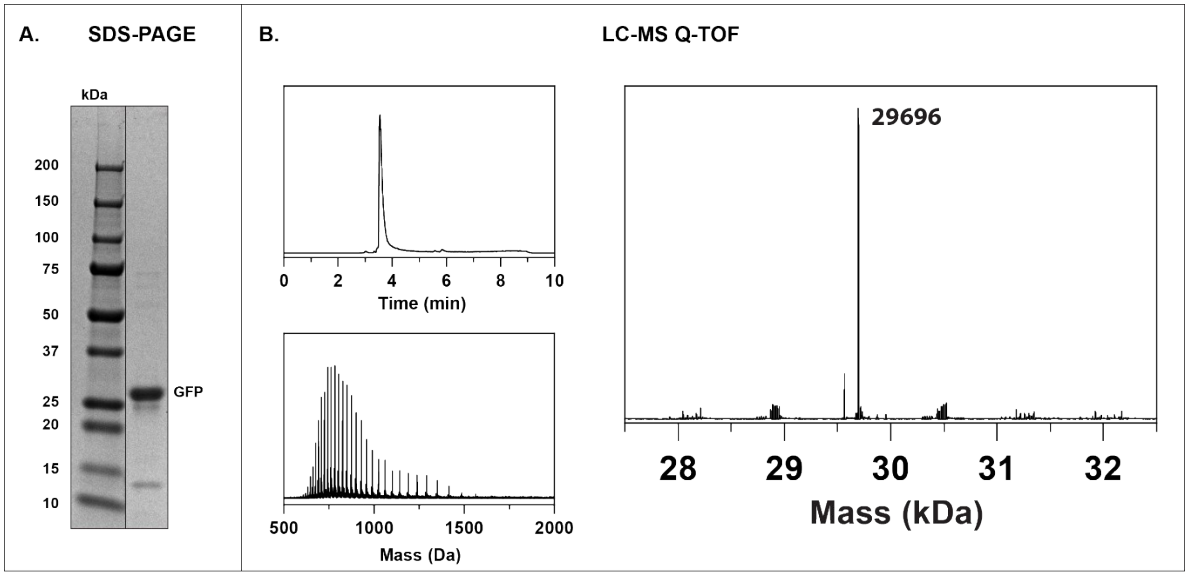

Fig. S4: Characterization of purified GFP (calculated molecular weight corrected for chromophore maturation is 29696.3 Da): A. SDS-PAGE; B. LC-MS Q-TOF, total ion count chromatogram (top left), m/z spectrum (bottom left) and mass spectrum (right). The experimental molecular weights are 20 Da lower due to chromophore maturation.

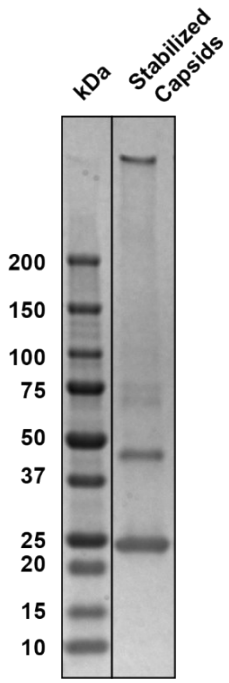

Fig. S5: SDS-PAGE analysis of stabilized VY1-VY8 ELP-CCMV capsids: CP bands appear at 22kDa, dimers and trimers at 44 and 66 kDa respectively, larger structure and crosslinked capsids on top of the gel.

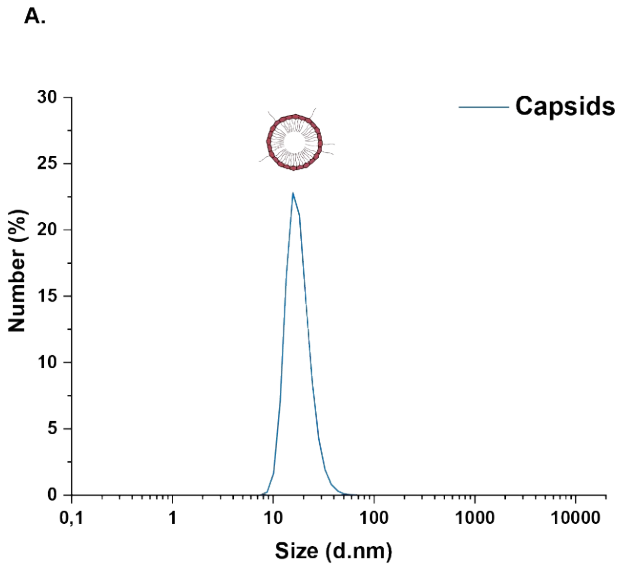

Fig. S6: DLS analysis of VY1-VY8 ELP-CCMV capsids in PBS (2 M NaCl). The Number distribution plot displays capsids of ~18 nm.

3 References

1 M. B. Van Eldijk, J. C. Y. Wang, I. J. Minten, C. Li, A. Zlotnick, R. J. M. Nolte, J. J. L. M. Cornelissen and J. C. M. Van Hest, *J. Am. Chem. Soc.*, 2012, **134**, 18506–18509.

2 L. Schoonen, R. J. M. Maas, R. J. M. Nolte and J. C. M. van Hest, *Tetrahedron*, 2017, **73**, 4968–4971.

3 L. Schoonen, J. Pille, A. Borrmann, R. J. M. Nolte and J. C. M. Van Hest, *Bioconjug. Chem.*, 2015, **26**, 2429–2434.