Supporting Information

Self-Assembly of a Fluorescent Virus-Like Particle for Imaging in Tissues with High Autofluorescence

Ikeda Trashi,[a] Mateusz Z. Durbacz,[b] Orikeda Trashi,[a] Yalini H. Wijesundara,[a] Ryanne N. Ehrman,[a] Alyssa C. Chiev,[a] Cary B. Darwin,[a] Fabian C. Herbert,[a] Jashkaran Gadhvi,[c] Nicole J. De Nisco,[c] Steven O. Nielsen,[a] and Jeremiah J. Gassensmith*[a,d]

[a] Department of Chemistry and Biochemistry, [d] Department of Bioengineering, [c] Department of Biological Science, The University of Texas at Dallas, Richardson, Texas 75080, United States

[b] Department of Molecular Biology and Hamon Center for Regenerative Science and Medicine, University of Texas Southwestern Medical Center, Dallas, Texas

Methods and Materials

Dynamic Light scattering (DLS)

Measurements were conducted on a Malvern Zetasizer Nano ZS at 25℃. All samples were measured after purification in a 1.5 ml cuvette with a 633 nm laser and 175 scattering angle, with a medium refractive index of 1.51.

UV-VIS

UV-VIS characterization was conducted using a Shimadzu UV-2401 PC in 1.5 mL quartz cuvettes.

Transmission electron microscopy (TEM)

All TEM samples were performed using a JEOL JEM-1400 plus microscope. First, 5 µL of sample was incubated on a 300 mesh formvar-coated copper grid. Then it was stained with 2% uranyl acetate solution as a negative stain for 60s. To remove the excess solution a filter paper was used and then was left to dry in air before imaging. All images were taken with an accelerating voltage of 120 kV.

Gel Electrophoresis

18% non-reducing SDS-page was run using a BIORAD kit, and samples were prepared using gel loading dye that contains β-mercaptoethanol and boiled 5 min before loading. The protein ladder used was Bio-Rad precision plus protein standards. The gel was run at 180 V for 75 min with Tris-Glysine pH 8.5. 1% agarose gel was run under 100 V for 30 min with 1× TBE buffer. Fluorescent agarose gel was imaged using Biomolecular Imager-GE Typhoon FLA 9000 for Cy5 fluorescence. Cells were stained with Coomassie brilliant blue and destained after before imaging.

Fluorescence spectra

Samples were run in a Horiba IHR320 Fluorometer with 1.5 ml quartz cuvettes.

BALB/c mice (n=3) were treated endotracheally with s@Qβ, GFP@Qβ, and saline samples. After 4 h, mice were sacrificed, and lungs were extracted, washed, fixed in 4% paraformaldehyde (PFA) for 48 h, and then they were washed 3× with phosphate-buffered saline (PBS). Then, the lungs were put into a cassette and washed for 12 h using a LEICA ASP3000S tissue processor. Parafilm embedding was performed using a heated paraffin-embedded station form HistoCore Aracdia H. After embedding, the parafilm was left to solidify, and once cooled, it was cut thinly with a manual microtome (LEICA RM2235).

Protein expression of s@Qβ

A double plasmid transformation was performed using E. *coli* BL21 (DE3) competent cells of pJF15-CP-smURFP and pBAD-HO-1 plasmids. Single colonies were picked and grown in LB media overnight under shaking conditions at 37℃. Then, an overnight starter culture was used to inoculate larger LB cultures, supplementing it with 50 µg/mL streptomycin and ampicillin. When the bacteria reach OD_{450} ~1.0, protein expression was initiated by inducing with 1 mM IPTG and 1% w/v L-arabinose. Expression happened overnight at 37 ℃ under shaking conditions. The next day, the cell cultures were pelleted down at 10000 ×g for 30

min in a Sorvall Lynx 4000 centrifuge. The resulting bright blue pellet was resuspended in 100 ml of 100 mM potassium phosphate (KP) buffer (pH = 7). Cells are lysed using a Qsonica sonicator at 80% amplitude for 4 min (intervals of 30 sec off and 30 sec on). The cell debris were pelleted down at 10000 ×g for 20 min, and the supernatant was supplied with 0.256 g/mL ammonium sulfate and stirred overnight at 4 ℃ to precipitate the VLP. The blue pellet was resuspended in 20 mL KP buffer. An organic extraction with 1:1 chloroform/butanol was performed and centrifuged at 10000 ×g for 10 min. After collecting the aqueous phase, the VLP was purified in a 10–40% sucrose gradient for 16 h at 61852 ×g in a Sorvall wx+ ultra-centrifuge. Finally, a light blue band, which is visualized by illuminating the tube from underneath with a light, was collected and pelleted at 400000 ×g for 2 h.

Protein expression of GFP@Qβ

The same purification method as described above was used using plasmids gifted from the Finn Lab at the Georgia Institute of Technology. The bacteria were grown in LB media supplemented with 50 µg/mL kanamycin at 37 ℃ and induced using 1 mM IPTG.

Protein purification of smURFP.

Plasmid pBad-smURFPRBS-HO-1(Addgene) was transformed using E. *coli* BL21 (DE3) competent cells and plated in LB agar plates supplemented with 50 µg/mL ampicillin. Single colonies were selected and grown overnight as a starter culture in LB media with 50 µg/mL ampicillin. The next day, the starter culture was used to inoculate larger cultures that were grown until bacteria reached an OD_{450} ~1.0 overnight at 37℃ and later induced with 1% w/v L-arabinose. Bacteria cells were pelleted down at 10000 ×g for 40 min in Sorvall Lynx 4000 centrifuge. The bacteria pellet was resuspended in PBS pH 8, then sonicated using Qsonica sonicator at 80% amplitude for 4 min (intervals of 30 sec off and 30 sec on). The bacteria cells were pelleted down to remove the cell debris at 10000 ×g for 40 min. The supernatant was purified using a NGC Quest 10 FPLC equipped with a 5 mL Bio-scale Profinity IMAC cartage. The samples were loaded, washed with 10 mM imidazole, and eluted with 200 mM imidazole in 1× PBS pH 8. The protein was dialyzed for 3 days against Milli-Q water to clean it further. Then the protein was lyophilized using Labconco Freezone 2.5 Lyophilizer.

Cell Viability.

A549 carcinoma lung cell line was grown in RPMI supplemented with 20% FBS and 1% L-GLU. The cells were seeded at a concentration of 2×10^{-4} cells/mL in a 4-well MatTek culture slide (1mL/well) and allowed to adhere overnight. All incubations took place in a 37 °C CO_2 incubator. The following day the cells were treated with the s@Qβ and GFP@Qβ at the same fluorescence intensity for 4 h before being washed 3× with serum-free media and 3 with PBS. The cells were stained with DAPI for nucleus staining and imaged by confocal microscopy.

Cellular uptake

The cells were seeded at a concentration 3×10^5 cells/mL in a 6-well plate(2000µL/well) and allowed to adhere overnight in a 37 °C, 5% CO_2 incubator. The next day cells were treated with similar fluorophore intensity s@Qβ and GFP@Qβ. Cells were incubated for 4h, washed, fixed and then imaged with CLSM to determine the distribution in cells.

Flow cytometry.

To quantify the cellular uptake of all our VLPs, flow cytometry was performed on A549 cells after 4 h incubation with 4 mg/mL s@Qβ and GFP@Qβ. Cells were washed 2× with PBS and stained for 20 minutes with 1:500 zombie UV (Biolegend) to differentiate the live and dead cells. Cells were washed with 2× with FACS buffer (PBS 10% FBS and 2 mM EDTA). Cy5 excitation and emission filter sets were used for s@Qβ and FITC for GFP@Qβ. Dead cells were excluded, and Qβ positive cells were determined using media alone as the baseline.

Estimating the number of smURFP proteins that can fit inside the Qβ capsid.

Goal: To estimate the maximum number of smURFP proteins that can fit inside the Qβ capsid.

Method: We will only consider the RNA attached to smURFP (*e.g.,* assume the protein volume is negligible compared to the RNA volume), and we will estimate how much RNA can fit inside the Qβ capsid by analyzing the void volume of the Qβ VLP. This will give us an upper bound on the number of smURFP.

Simple example to illustrate the Monte Carlo method of determining volume:

Let us assume we can draw uniformly random numbers between 0 and 1, and call these u. We can then draw uniformly random

numbers between $-a$ and a, call these x, by $x = \left(u - \frac{1}{2}\right) * 2a$. Now, let us draw three such numbers and call them x,y,z. Regarding the

triplet (x,y,z) as a point in three-dimensional Cartesian space, we have selected a random point inside the cube centered at the origin with side length $2a$. This cube has volume $V_{cube} = (2a)^3 = 8a^3$. Next, we check if this point is inside the sphere of radius a by asking if $\sqrt{x^2 + y^2 + z^2} < a$. Denote the total number of triplets by N, and denote the number of triplets inside the sphere by S. Then, we can estimate the volume of the sphere simply by $V_{sphere} = (S/N)\ V_{cube}$, that is, the sphere volume is the cube volume times the ratio of accepted points (inside the sphere) to total points. Some numbers to illustrate how well this works are shown:

Use $a = 10$, so that $V_{\text{cube}} = 8000$			
The true sphere volume is then $V_{\text{true}} = \frac{4}{3}\pi a^3 = 4188.79$			
N	S	estimated sphere volume V_{estimate}	error $= \lvert V_{\text{estimate}} - V_{\text{true}} \rvert$
$100 = 10^2$	45	3600.00	588.79
$1000 = 10^3$	549	4392.00	203.21
$10000 = 10^4$	5227	4181.60	7.19
10^5	52440	4195.20	6.41
10^6	524151	4193.20	4.42
10^7	5232501	4186.00	2.79
10^8	52357546	4188.60	0.19
10^9	523593234	4188.75	0.04

Void volume estimate for the Qβ capsid and the Qβ virion:
Similar to the simple example, we determine the void volume of the Qβ capsid (coat protein only) and the Qβ virion (coat protein and genomic RNA). The only difference is that we use the atoms of these objects to determine the void space—specifically, if our random triplet is within $r_{threshold}$ of *any* atom, this point is *not* in the void space since it overlaps with a protein and/or rna atom. The value of $r_{threshold}$ is somewhat arbitrary, but certainly, it should be in the range 2.0 Å to 3.0 Å. We used 11 different values to give some idea of what difference it makes.
The ratio of the void volumes tells us how much void space there is in the viron compared to the (empty) capsid. One minus this ratio is the fraction of the capsid volume occupied by the genomic RNA. Since there are 4127 nucleotides in the genomic RNA, dividing 4127 by the fraction occupied gives us the total number of nucleotides that can fit inside Qβ. Then, we divide this number by 1218 which is the number of nucleotides per smurfp protein (between the T7 promotor and the terminator sequence). This is how we estimate the maximum number of smURFP proteins that can fit inside Qβ. From the last column of the table, we predict the the maximum number to be in the range 10-17, depending on the threshold value.

$r_{\text{threshold}}$ (Å)	Qβ capsid void volume (Å^3)	Qβ viron void volume (Å^3)	ratio	# smurfp
2.0	7204088	5790300	0.803	17.3
2.1	7099696	5596077	0.788	16.0
2.2	6999195	5406670	0.772	14.9
2.3	6905057	5231550	0.758	14.0
2.4	6823565	5067559	0.743	13.2
2.5	6747963	4907113	0.727	12.4
2.6	6676400	4763013	0.713	11.8
2.7	6610059	4628900	0.700	11.3
2.8	6552401	4495583	0.686	10.8
2.9	6491764	4377279	0.674	10.4
3.0	6440548	4259740	0.661	10.0

To understand why the Qβ capsid void volume decreases with increasing $r_{threshold}$, essentially, the protein “rim” gets fatter.

Primers used for amplification and Sanger sequencing.

smuRFP Forward	5'– GCGCCGGCCGGTATGAAAACTTCTGAACAACGTG – 3'
smuRFP Reverse	5'– GCGCCTCGAGTTAGAATTCGGACATAGCCTTGA – 3'
HO-1 Forward	5'– AAGGAGATATGGATCCATGAGTGTCAACTTAGCTTCC– 3'
HO-1 Reverse	5'– AGCTGGAGACCGTTTAAACCTAGCCTTCGGAGGTGG– 3'
HO-1 sequencing Forward	5'– ATGCCATAGCATTTTTATCC – 3'
HO-1 sequencing Reverse	5'– GATTTAATCTGTATCAGG – 3'
pJF51-CP- sequencing Forward	5'–GGAATTGTGAGCGGATAACA– 3'
pJF51-CP- sequencing Reverse	5'–GCTAGTTATTGCTCAGCGG– 3'

Table S1. Shown in the table are primers used for Sanger sequencing to confirm ligations and amplifications for PCR reactions.

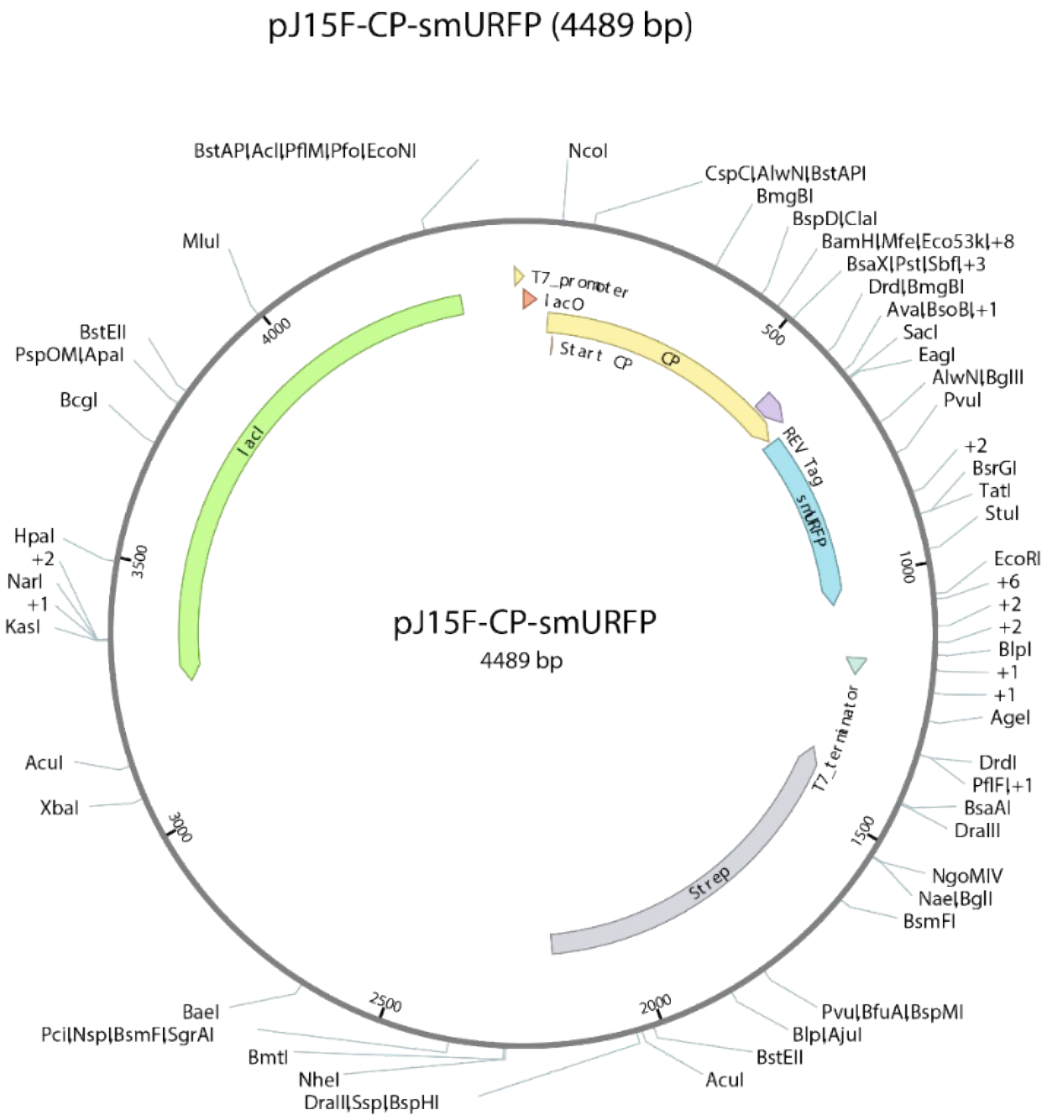


Fig S1. The plasmid map for the pJF15-CP-sMURFP was generated from Benchling, showing the restriction sites, genes of interest, and antibiotic resistance.

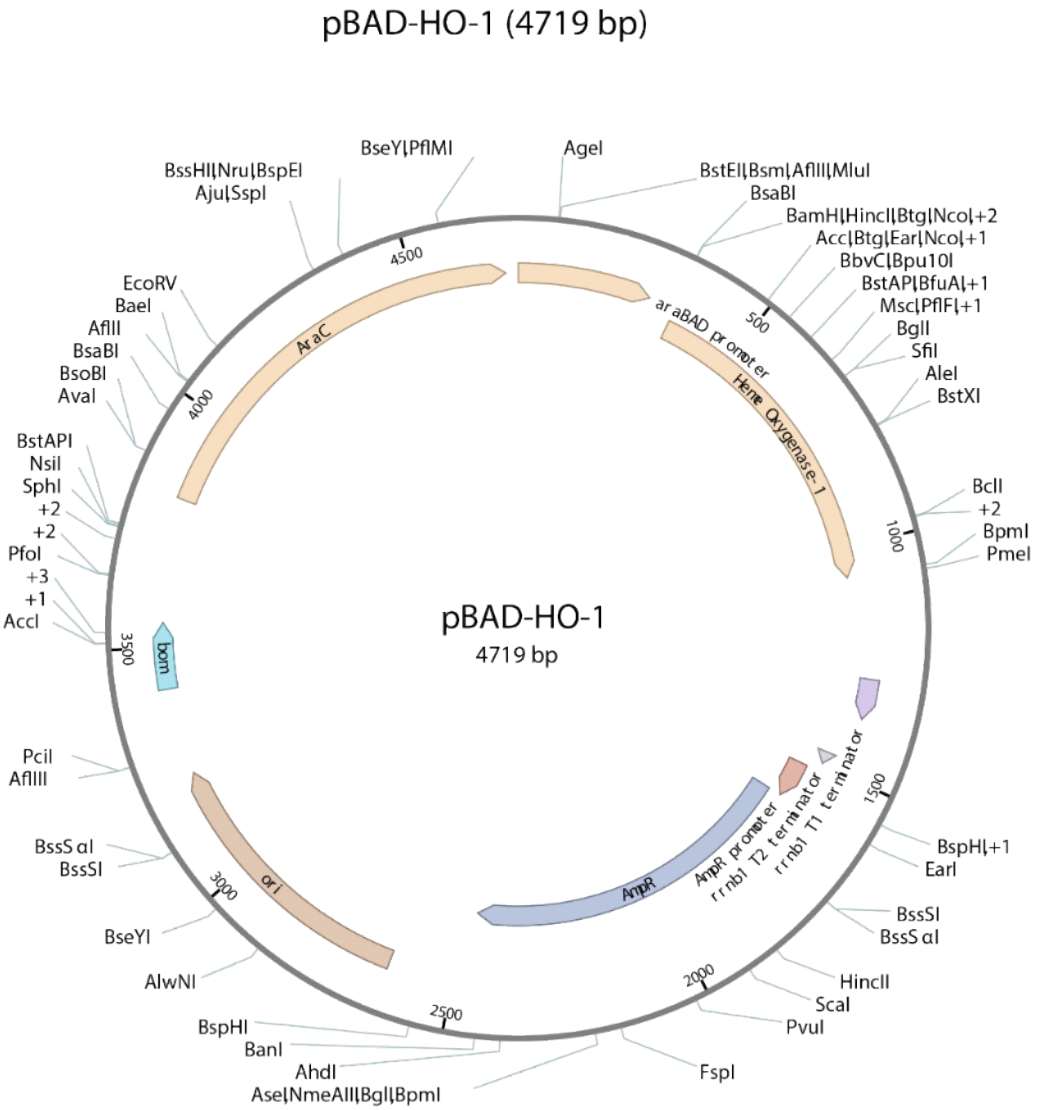


Fig S2. Plasmid map for the pBAD-HO-1 used for chromophore Bv expression. The plasmid map was generated from Benchling, showing the restriction sites, genes of interest, and antibiotic resistance.

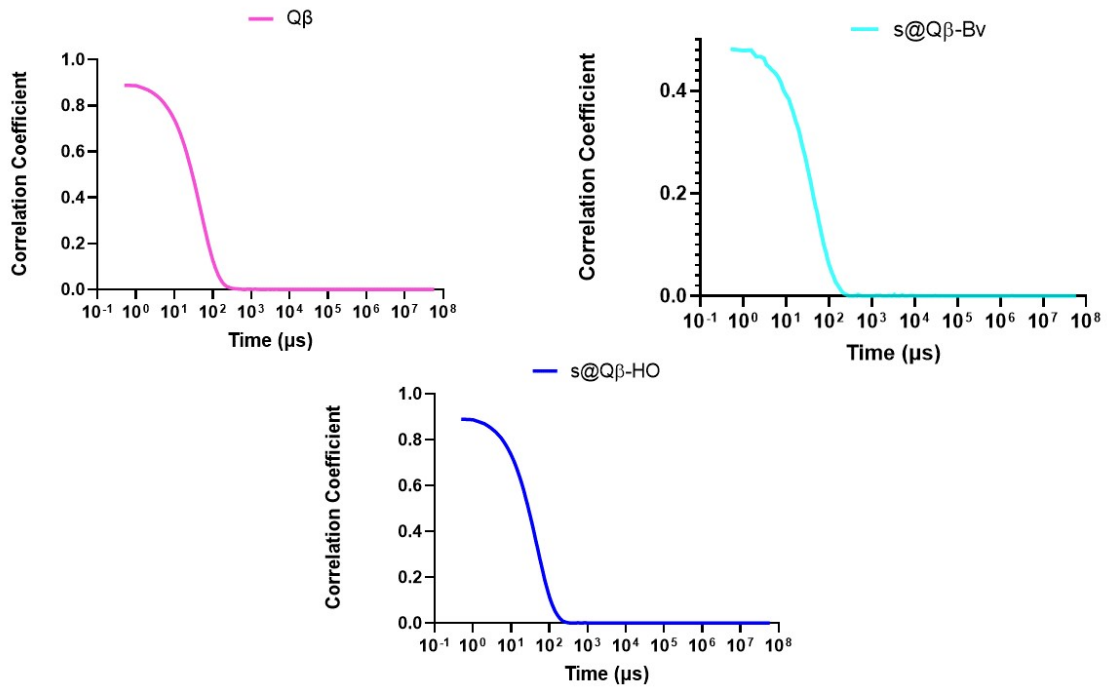
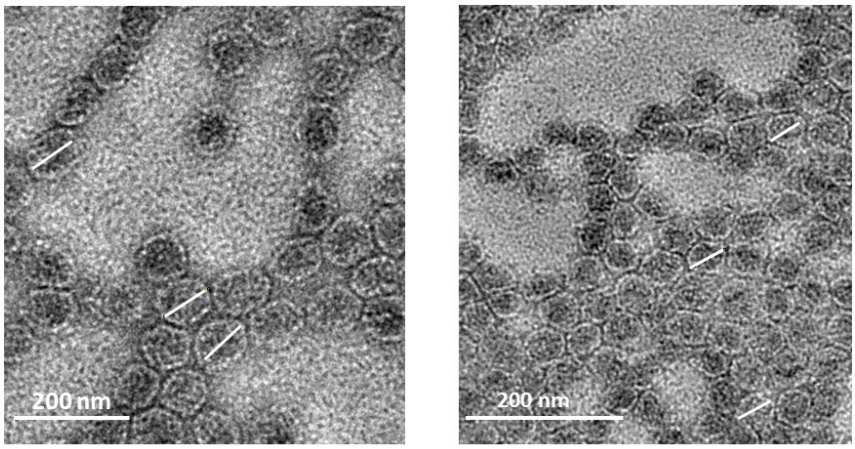


Fig S3. Autocorrelation graph after size measurement using DLS indicating good quality of data.



Qβ s@Qβ

Qβ size (nm)	s@Qβ size (nm)
34.444	33.128
31.378	32.303
28.687	27.581

Fig S4. TEM images used to determine size for Qβ and s@Qβ using Image J.

Gel densitometry.

Gel densitometry was performed using 18% SDS-Page and plotting a standard curve of bovine serum albumin (BSA) with concentrations ranging from **1** (0.1563 mg/mL), **2**(0.07812 mg/mL, **3** (0.03906 mg/mL), **4** (0.01953 mg/mL)**, 5** (0.009765 mg/mL)

to **6** (0.004883 mg/mL). Then s@Qβ-HO and s@Qβ-Bv sample with an unknown concentration were loaded, and the concentration of smURFP and Qβ were calculated.

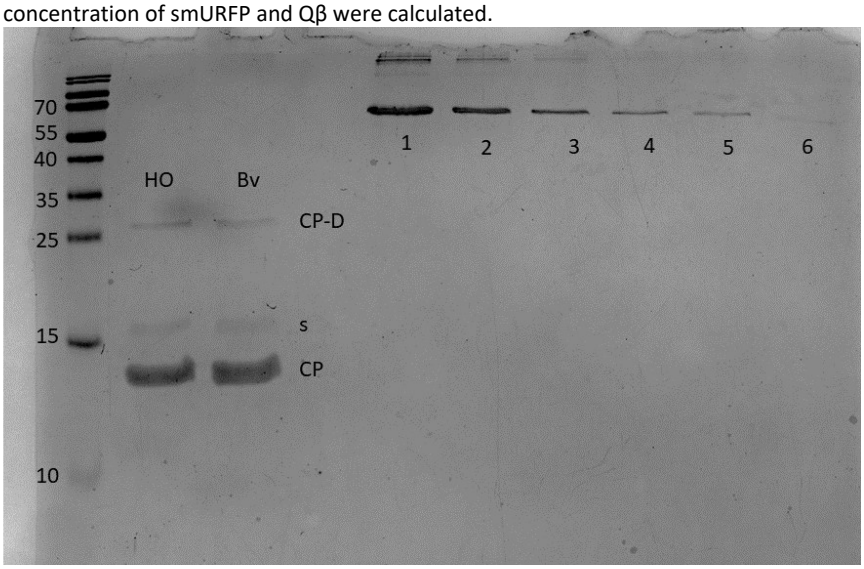


Fig S5. 18% SDS-Page showing bands of **CP** monomer, dimer (**CP-D**), and smURFP (**s**) for s@Qβ-HO and s@Qβ-Bv shown as HO and BV in gel respectively. Serial dilutions of BSA protein are loaded starting from **1** (0.1563 mg/mL) down to **6** (0.004883 mg/mL).

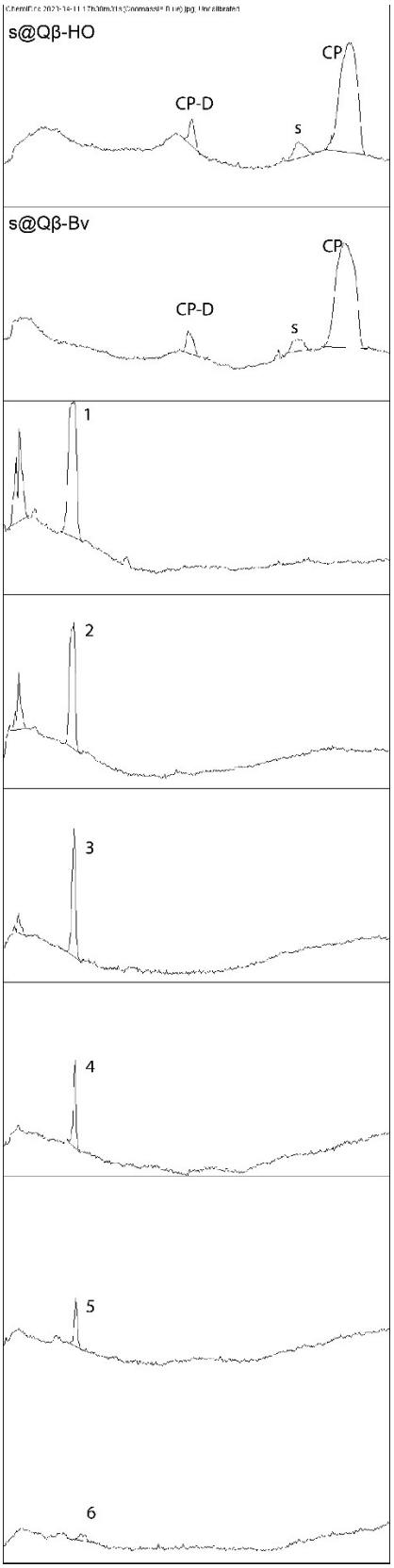


Fig S6. Intensity plots for the BSA standards and s@Qβ bands generated from ImageJ.[1]

To calculate the amount of smURFP's packed inside the VLP fluorescence intensity was used. First, a standard curve from Qβ was prpared to determine Qβ concentration in the VLP and then a standard curve based on fluorescence intensity was performed to determine smURFP intensity. To calculate the amount of smURFP's packed inside the VLP fluorescence intensity was used. Knowing the molecular weight of Qβ and Smurfp their molar ratio was found. For s@Qβ -HO, Qβ concentration was 2.5 mg/mL and smURFP concentration was 0.1484 mg/mL. For s@ Qβ -Bv Qb concentration was found out to be 5.3 mg/mL and smURFP concentration was 0.249 mg/mL. From this calculations resulted in approximately ~10 smURFP's encapsulated in s@Qβ-HO and ~8 smURFP's in s@Qβ-Bv

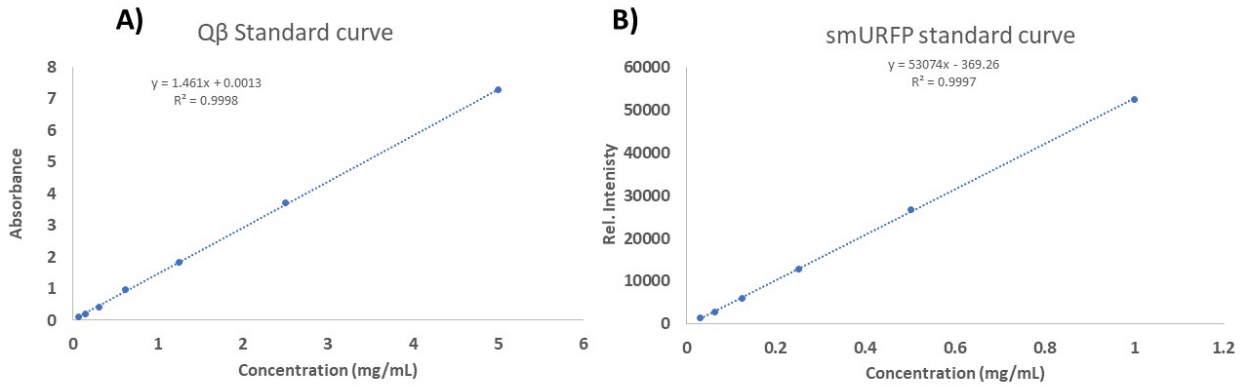


Fig S7. A) Standard curve of Qβ VLP and B) smURFP used to determine concentration.

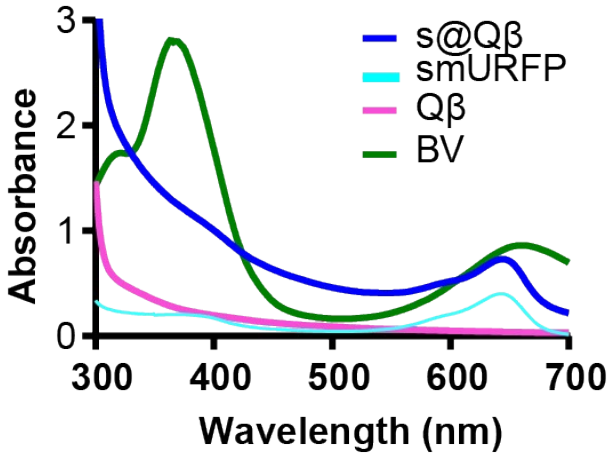


Fig S8. UV-Vis spectra of s@Qβ, smURFP, Qβ, and Biliverdin hydrochloride.

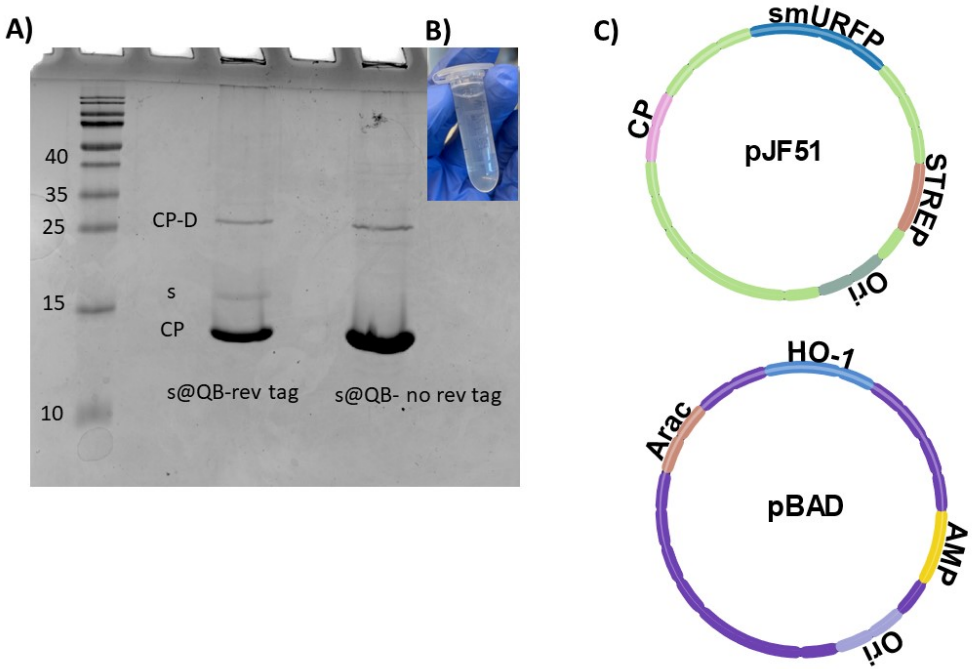
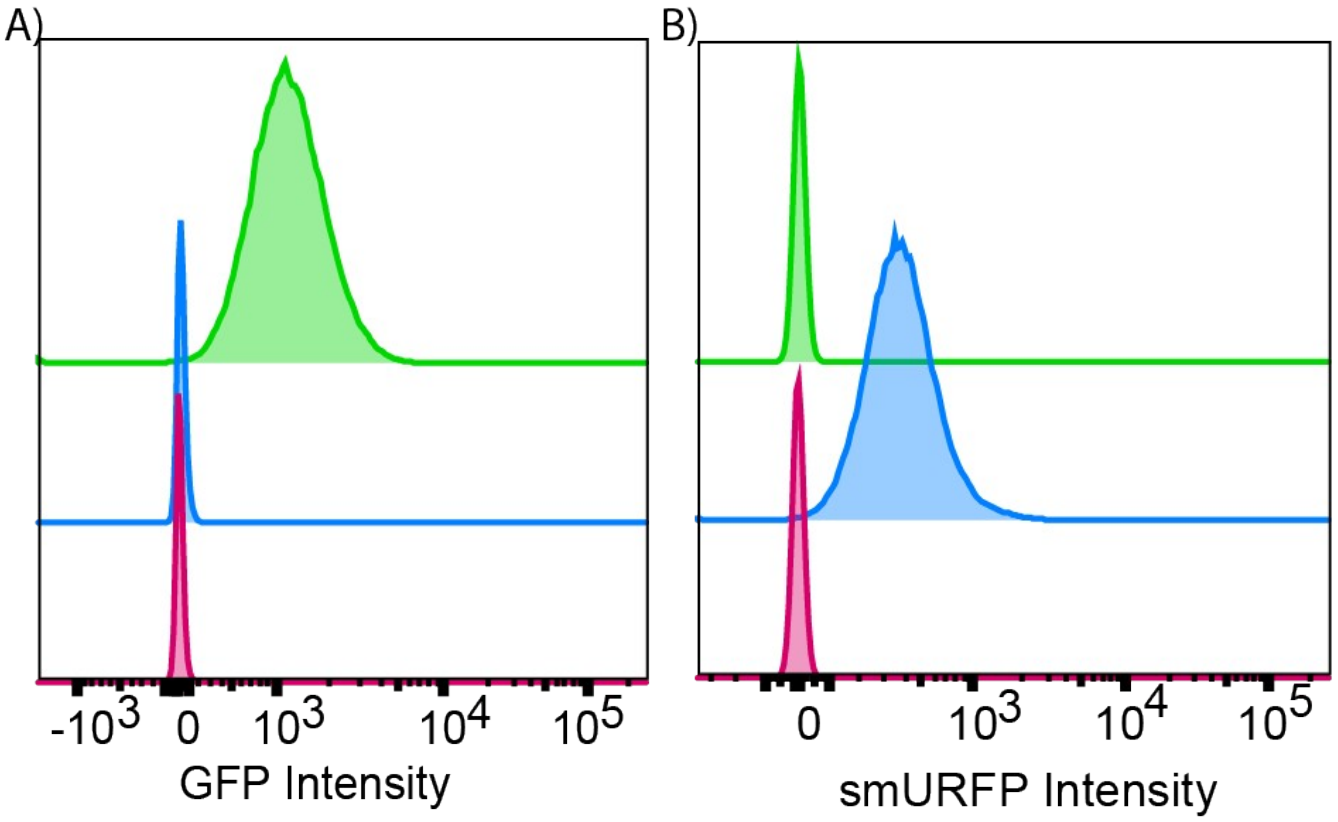


Fig S9. A) SDS page of s@QB with Rev tag and s@QB after removing the REV tag. B) s@QB no REV tag solution. C) Plasmid map after removing the REV tag.



	Sample Name	Subset Name	Count
	GFP@Qβ_E3_E03_015.fcs	LIVE	81956
	s@Qβ_B2_B02_007.fcs	LIVE	61383
	Media_A1_A01_001.fcs	LIVE	26821

	Sample Name	Subset Name	Count
	GFP@Qβ_E3_E03_015.fcs	LIVE	81956
	s@Qβ_B2_B02_007.fcs	LIVE	61383
	Media_A1_A01_001.fcs	LIVE	26821

Fig S10. Representative histogram of the uptake in A549 cells for GFP@Qβ and s@Qβ in either the A) green channel or B) Cy5 channel.

1. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nature methods. 2012;9(7):671-5.