Cloning gpB and gpC for the *In Vitro* Assembly of Bacteriophage Lambda Procapsids for use as Viral Nanoparticles

Bethany Arabe, Madelyn Stueve, Tobias Wakefield-Carl
Under the mentorship of Dr. Kristopher Koudelka

Department of Biology, Point Loma Nazarene University, San Diego, CA

Abstract

Bacteriophage lambda procapsids have shown promising uses as viral nanoparticles, as its structure allows for the possibility of drug encapsulation. The icosahedral shape of the procapsid consists of several individual proteins: the major capsid protein - gpE, head decoration protein - gpD, scaffolding protein - gpNu3, portal protein - gpB, and protease - gpC. These viral capsids function to protect and transport the genome of viral DNA to a host cell, and have shown potential in use for targeted drug delivery. Production, identification, and purification of these proteins were done to isolate the proteins for *in vitro* assembly through techniques that also use glutathione-S-transferase (GST) tag affinity chromatography. Different ratios between the individual proteins were combined for assembly and then the combinations were evaluated by size exclusion chromatography (SEC) to determine the uniformity of the assembly and the success of procapsid formation. In previous research, the *in vitro* assembly of the procapsids was attempted using only gpE, gpD, and gpNu3, with success in aggregation but not uniformity. The purpose of this project is to determine the effect of the presence of gpB and gpC on assembly, by cloning them into expression plasmids so the proteins have a glutathione-S-transferase (GST) affinity tag. After purification and incorporation of gpB and gpC, uniformity of procapsid formation increases, which shows potential for improvement in the *in vitro* assembly of procapsids.

Introduction

The development and study of nanoparticles and nanotechnology has opened the door for novel ways of treating various diseases. Instead of flooding the entire body with drugs to treat a disease, nanoparticles can be utilized to target specific areas of the body. Currently, the conventional drugs that are widely used are hypothesized to have a low bioavailability and solubility, which drastically reduces its retention in biological systems.[1] Due to the versatility of nanoparticles, much research is being conducted to discover ways in which nanoparticles can be used in the field of medicine. Nanoparticle drug delivery systems are even more valuable because they are capable of delivering drugs to areas of the body that other delivery systems cannot.[2]

One particular area of nanoparticles that has shown potential for use as a drug delivery system is viral nanoparticles (VNPs). VNPs can come from a wide variety of sources, including both human and plant viruses as well as

bacteriophages. Manufactured VNPs, examples of which are shown in **Figure 1,** are biocompatible and biodegradable, and can be used to house cancer therapies, imaging particles, immunotherapies, and more.[3] With the ability to modify the nanoparticles to target specific sites, researchers are able to direct drugs or other treatments to affected areas with a much higher precision and lower exposure than conventional drugs.

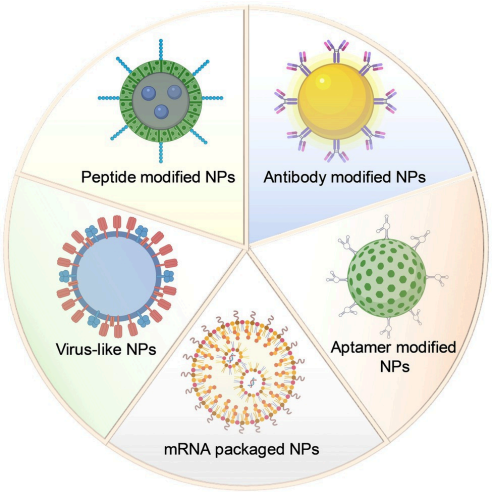


Figure 1: Examples of viral nanoparticles and the different ways to manufacture their structures for use as drug delivery systems. Much of the versatility of these VNPs lies in the antibody modifications.
Image from Frontiers in Microbiology 2021:
https://www.frontiersin.org/journals/microbiology/articles/10.3389/fmicb.2021.767104/full

An example of VNPs are bacteriophages. Bacteriophages have been recognized as one of the most abundant biological agents and all strains are made up of nucleic acid encased in a shell of phage-encoded proteins that protect the genetic material during its transport and delivery to the next host.[4] One specific strain of bacteriophages that has been extensively studied is bacteriophage lambda. Bacteriophage lambda infects the bacteria *E. coli*, which has made it extremely useful as a model system for the study and manipulation of DNA and gene expression. Because of its known ability to hold and deliver material to *E. coli* bacterial cells, this bacteriophage has the potential to use such capabilities for encapusulating and transporting drugs in mammalian systems.

Bacteriophage lambda, as shown in **Figure 2,** possesses a viral icosahedral head (known as the procapsid), a sheath-like body and also a tail complete with tail fibers that allow for the targeted infection of host cells.

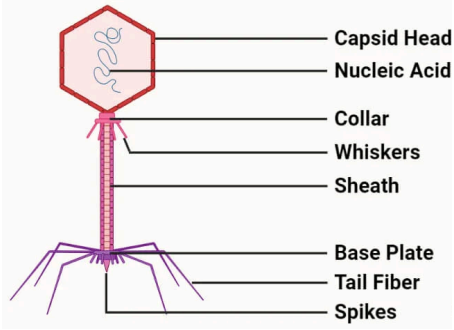


Figure 2: General structure of the bacteriophage lambda, including the procapsid head, where nucleic acid is contained, the main body (sheath), as well as the tail end with the infection mechanisms.
Image from Microbe Notes:
https://microbenotes.com/bacteriophage

The procapsid is where phage lambda's genetic information is stored. This procapsid shell is made up of five known proteins. Gene product (gp) B, C, D, E, and Nu3. gpE is the major capsid protein, and around 400 copies are combined with the stabilization protein gpD, present in equal ratios to gpE, to make up the shell of the procapsid.[5] The *in vivo* assembly of the procapsid shell is accompanied by ~70-200 copies of the scaffolding protein gpNu3, which serves as a chaperone for the formation of the procapsid that is later degraded.[6] The degradation of gpNu3 is performed by gpC, which is a viral protease. The protease gpC is

required for the breakdown of the scaffolding post assembly, and is also responsible for the proteolysis of the portal complex, which is essential to the later steps in particle assembly.[7] The last major protein that makes up the procapsid is gpB, which forms the portal complex and functions as a nucleator for the assembly[8]. Bacteriophage lambda procapsid assembly mechanism is shown below in **Figure 3.**

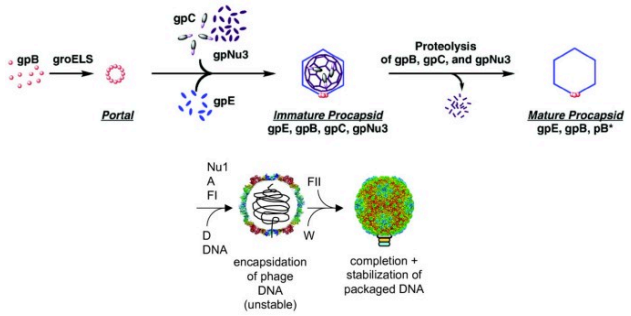


Figure 3: Overview of the assembly of the mature procapsid head. Assembly begins with the formation of the portal through gpB. Then, gpC, gpE, and gpNu3 are added to form an immature procapsid. The proteolysis of the proteins shown form a mature procapsid. Not pictured is gpD, the stabilizing protein for gpE.
Image from BMC Microbiology:
https://bmcmicrobiol.biomedcentral.com/articles/10.1186/1471-2180-11-213

The focus of this project is to take these five necessary protein components and attempt to control the formation of the procapsid *in vitro*.

The ability to control formation of the procapsid head would allow for the encapsulation of drugs or other particles that need to be delivered to specific sites within the body. Previous research has shown that the proteins that the procapsid is composed of will naturally congregate into larger structures. However, previous attempts at *in vitro* assembly of the bacteriophage lambda procapsid did not yield uniform icosahedral shells, as displayed in **Figure 4**. It is important to note that in those trials, gpB and gpC had not yet been cloned into an expression vector and thus were not used in the assembly process for these Cryo-EM images. Thus, the main goal of this research project is to clone both gpB and gpC into expression vectors with a GST-tag that will allow for growth and purification of the proteins that can later be incorporated into the *in vitro* assembly of bacteriophage lambda procapsids.

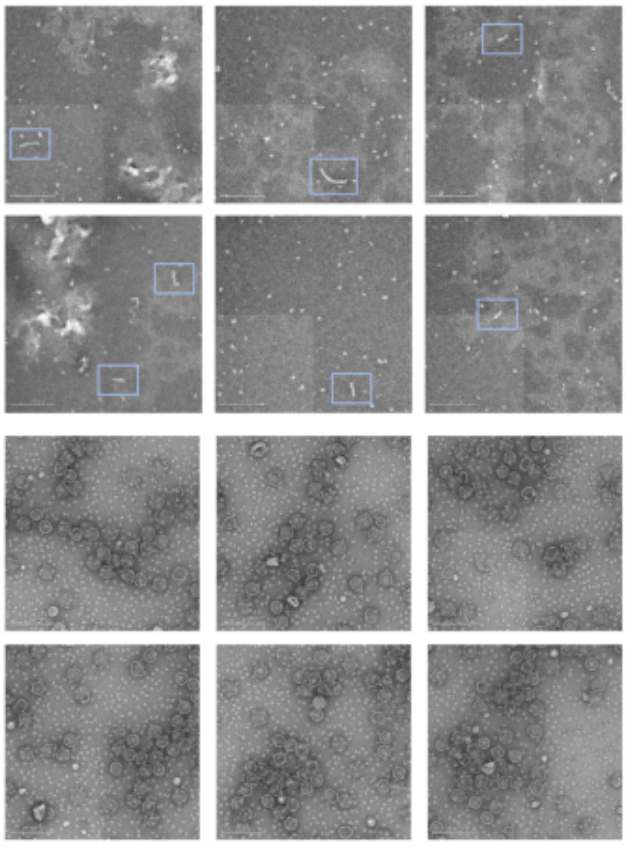


Figure 4: Cryo-EM images of attempted *in vitro* assembly of procapsids compared to control procapsids. In the top series of pictures, procapsids assembled with a ratio of 1:1:1 ratio of gpD, gpE, and gpNu3 did not form uniform icosahedrons. Inside the boxes, long tubule structures can be seen, indicating that the proteins are congregating to form a larger molecular structure, but are not uniformly shaped. The bottom series of images displays the lambda control. The icosahedral shapes represent fully formed and intact procapsids generated through *in vivo* assembly. These procapsid samples were made in Summer 2022 *in vivo* in *E. coli* bacteria cells and

imaged through Cryo-EM. These cells produce the procapsids fully intact, and these procapsids function as a control for the research.

Methods

gpB and gpC Protein Preparation

Bacteriophage lambda capsid proteins gpE, gpD, gpNu3 had previously been cloned and purified. This first part of this research focused on the cloning and purification of bacteriophage lambda capsid proteins gpB and gpC with GST tags.

Cloning, Restriction Digest and Ligation

To prepare the expression vectors for transformation, they first needed to be digested and ligated with the proper inserts, the sequences for gpB and gpC. The vector was combined in a polymerase chain reaction (PCR) tube with rCutSmart Buffer (10X), restriction enzymes SmaI and XhoI, and nuclease free water. This solution was incubated for one hour at 37℃ in a thermocycler. To deactivate the restriction enzymes, the solution was incubated for one hour at 80℃. The digested vector was transferred from the PCR tube to a microcentrifuge tube and combined with 3M sodium acetate and 100% ethanol, then placed in a -80℃ freezer for 15 minutes. This was then spun at 1,4000 rpm for 15 minutes and the supernatant was carefully removed. The tube was allowed to air dry for ten minutes and then the pellet was resuspended in 30µL of nuclease free water. After nano-dropping the solution to determine the concentration of the vector, a 10X dephosphorylation buffer was added, along with 1µL of diluted CIAP. To deactivate the phosphorylase, 3µL of 0.001M EDTA was added. The final solution was incubated at 65℃ for 15 minutes.

Ligation of the vector and insert was conducted at a ratio of 0.015pmol of Vector to 0.045pmol of Insert. The vector and insert DNA were combined in a PCR tube along with 2µL of T4 ligase buffer and nuclease free water. After this, 1µL of T4 ligase was added to the solution and mixed by pipetting up and down. After incubating the tube at room temperature for one hour, the vectors were ready for transformation.

Transformation

Following the preparation and ligation of the expression vectors containing the GST tag sequence, the resulting plasmids were ready for transformation into *E. coli* cells. For each expression plasmid, one vial of OneShot cells was thawed on ice and 5-10ng of the DNA (plasmid) was added to the vial (1-5 µL). This solution was incubated on ice for 30 minutes. Post incubation, the cells were heat shocked for exactly 30 seconds in a 42℃ water bath and immediately placed on ice. Next, 250 µL of warmed SOC medium was added, which is a Super Optimal broth with Catabolite repression medium used for maximum transformation efficiency. The vials of cells and media were secured to a microcentrifuge rack and placed in a shaking incubator for one hour at 225 rpm and 37℃. After the incubation period, two different volumes from the respective vials were plated onto LB agar plates (containing ampicillin). Around 20µL of cells were added to one plate, while the remaining 230µL were added to another plate. The transformants that were selected through this process were then

available for further cultivation and analysis to determine if the transformation was successful.

PCR Check

In order to clone both gpB and gpC (both with GST-tags) into new expression vectors, a PCR was performed to create the new genes and insert them into *E. coli* plasmids. The gene sequence of Bacteriophage Lambda was analyzed to create forward and reverse primers, allowing for the insertion of the gene of interest. After both primers were obtained, they were combined in a PCR tube along with 4x Reliance One-Step Multiplex Supermix, the RNA template (plasmid), as well as nuclease free water. PCR tubes containing gpB and gpC tests as well as corresponding controls were run in a thermocycler following the protocol below in **Figure 5**.

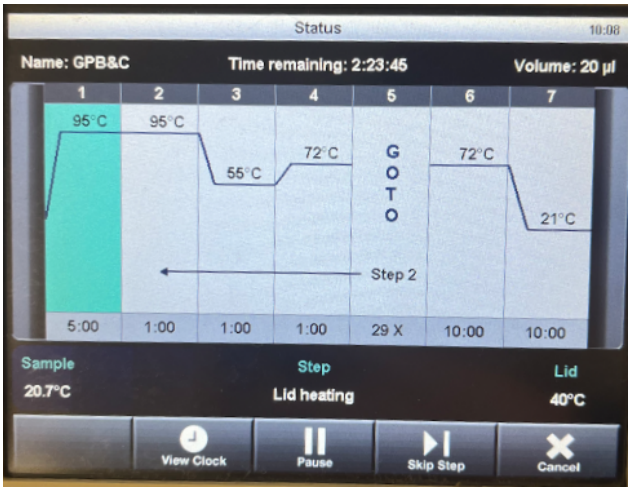


Figure 5: Thermocycler protocol for the cloning of gpB and gpC. The first step is an activation and denaturation at 95℃. Next is an amplification sequence starting with 95℃ for one minute, 55℃ for one minute, and 72℃ for one minute. The amplification sequence was repeated 29 times.

Once the PCR was complete, the samples were then ready to be analyzed by gel electrophoresis to confirm the presence of cloned gpB and gpC.

Gel Visualization

Throughout this experiment, both agarose gels and SDS-Page gels were used to visualize proteins contained within a sample. For the agarose gels, 1X TAE Buffer was combined with agarose powder to make a suitably sized gel, and once the solution was cooled, ethidium bromide was added. The agarose solution was then poured into a gel box and a comb was placed to make the wells. Once the gel was fully set, the ladder, as well as the PCR product along with the proper controls, were loaded into wells. The gels were run at 100 V until the bands approached the end of the gel (~1 hour). The finished gel was removed from the gel box and placed in a UV box to visualize all of the bands. For SDS-Page, samples were prepared by combining lysate samples with Bolt LDS sample buffer and Milli Q water. The ladder, samples, and controls were loaded into the SDS gel (bis-tris-plus), and the gel box chambers were filled with a 1X running buffer along with a small amount of Bolt antioxidant. The gel was run at 200V until the bands approached the bottom of the gel. The gel was removed from its casing and visualized according to the SimplyBlue staining protocol.

Plasmid Sequencing

To analyze the contents of plasmids generated in this experiment, overnight solutions of the samples were prepared by placing an *E. coli* colony into a conical vial containing LB Broth and left shaking in an incubator overnight to promote production of the sample plasmid. The next day, aqueous samples containing a plasmid

were first mini-prepped to purify the sample using the QIAprep Spin Miniprep Kit, and sent to Eurofins Laboratories for final sequencing. The protocol contained in the Miniprep Kit was performed as instructed with all spins being run at 13,000 rpm. Following the successful miniprep of the samples, each one designated to be sent out for sequencing was nano dropped to determine the concentration of the purified sample. After obtaining the concentrations of the samples, calculations were utilized to bring the samples to the proper concentrations required by Eurofins for the sequencing. Each sample needed to have a final volume of 15 µL at a plasmid concentration of 30 ng/µL. Once the samples were fully prepared, they were sent to Eurofins Laboratories. Upon completion of sequencing, Eurofins returned the data and it was analyzed using the SnapGene Viewer, allowing for the visualization of the plasmid components.

Production of gpB and gpC proteins with GST Tags

Overnight cultures were prepared by adding 20 mL of autoclaved lysogeny broth (LB), 20 µL of sterile-filtered ampicillin and respective samples of transfected *E. coli* into conical vials and placed into a shaking incubator at 37°C at 200 rpm, overnight. Next, to prepare for induction 2 flasks were prepared with 500 mL of 2X YT and supplemented with 7.75 mL 1M sodium phosphate buffer, 2.5 mL 1M glucose and 0.5 mL ampicillin. Before adding the respective overnight cultures to their flasks, 5 mL was removed as a blank from each beaker. 10 mL of each overnight culture was added to a flask and optimal density (OD) readings were recorded as a measure of bacterial growth using a spectrophotometer. The culture was induced with sterile-filtered, 1M IPTG once the OD reached 0.6-0.7. The flasks were placed in a shaking incubator for 2 hours at 37°C. The two samples were portioned into conicals for centrifugation and spun at 5500 rpm for 15 minutes in the SLA-3000 centrifuge at 4°C. The supernatant was decanted, and pellets stored at 4°C overnight.

The gpB and gpC pellets were then resuspended in 5mL procapsid buffer and 2 µL DNase and transferred to 15mL conicals to be sonicated for 10 seconds with 10 seconds rest in between for a total of 10 repetitions. 55 µL of EDTA and PIC were added to both conicals containing either gpB-GST or gpC-GST proteins. The solutions were poured into lipped Nalgene tubes, balanced and placed into a SA-600 rotor to be centrifuged at 7300 rpm for 24 minutes at 4°C. The supernatant was decanted and sterile filtered using a 0.2 µm filter and stored in the fridge at 4°C.

GST Tags Cleaved from gpB-GST and gpC-GST Proteins

Proteins gpB and gpC are tagged with GST which will be utilized in a column cleavage reaction to isolate proteins from the GST tags and to further purify our desired proteins from the sample. This purification process utilizes the affinity between the GST tag and the GST column. To start the column was equilibrated with 2-4 resin bed volumes of 1x HRV 3C reaction buffer, and the column was placed in the centrifuge and spun at 3200 rpm for 2 minutes to remove the buffer. The buffer was discarded and the column was placed in the centrifuge tube and the steps to equilibrate the column were repeated 3 times. After the washes, the proteins with GST tags were added with 1x HRV 3C reaction buffer to make 2-3 resin bead

volumes and loaded on the column and was incubated for 30-120 minutes at 4ºC with end-over-end mixing using a rocking platform to give the GST tags adequate time to bind to the column. After binding, the bottom plug was removed and was centrifuged at 3200 rpm for 2 minutes. The flow-through was collected and stored for later analysis to evaluate the binding capacity of the column. The resin beds were washed with 2 resin bed volumes of buffer and centrifuged twice.

At this point in the protocol, the GST tagged proteins bound to the column should be the only thing remaining. The next step is to perform a cleavage reaction to isolate the protein of interest from the GST tag. The cleavage was accomplished by creating a resin slurry in which 2 resin bed volumes of 1x reaction buffer were loaded onto the column. The HRV 3C protease was then added and the column was capped to be incubated overnight at 4ºC with end over end mixing. After incubation the column was centrifuged at 3200 rpm for 3 minutes to collect the desired proteins gpB and gpC without GST tags.

Protein/Capsid Purification

Gel purification was used to extract any desired protein from the solution. To do so, an agarose gel was prepared (see gel visualization for the protocol), but with larger wells made by taping together multiple teeth of the comb. In the larger wells, samples of the proteins (gpB, gpC, etc.) were loaded in the gel and it was run at 100V until the bands neared the end (~1 hour). After the gel was complete, it was placed under a UV light and the band corresponding to the protein was excised, using a scalpel and placed in a microcentrifuge tube. Following this, the excised gel band was dissolved according to the QIAquick gel extraction kit protocol. All spins in the protocol were run at 13,000 rpm. The final product resulted in a purified sample of the desired protein.

To assemble procapsids, the purified protein samples were combined at a series of ratios to test for the uniformity of congregation. After each combination was incubated and mixed for the allotted time, the final solutions were run over a column using fast protein liquid chromatography (FPLC). Utilizing a size exclusion column, the mixed samples were run through the column in buffer solution, and the results were visualized through the chromatography system.

Results

Cloning of gpB and gpC

In order to determine the effect of the presence of gpB and gpC on the assembly and formation of the procapsid, these individual protein plasmids were cloned to have a GST affinity tag for purification after production. The process of ligation and transformation were done, followed by a PCR check with the inserts in the plasmids. The inserts were not cut back out of the plasmid for the PCR. The PCR check gel electrophoresis was done to determine if the cloning produced the plasmids of the correct size, indicating the inserts were successfully taken up by the plasmid. With the known size of gpB being 1602 bp and gpC being 1320 bp, bands at approximately 1600 bp and 1300 bp show promising results to be the desired proteins (**Figure 6a and 6b**).

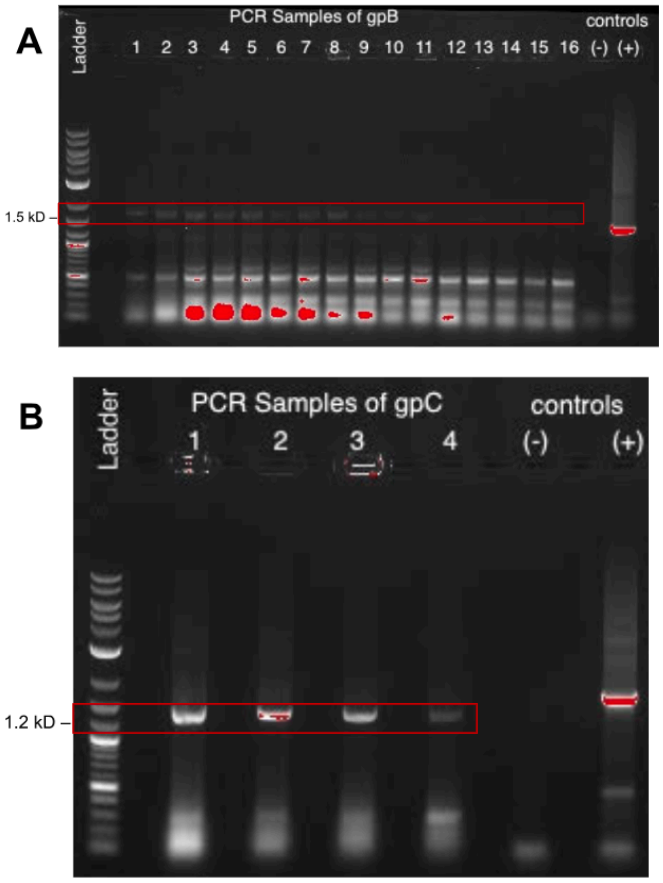


Figure 6: Gel images to conclude that the plasmids of correct size for gpB and gpC were made. (A) There was a large volume of plasmids of gpB made (16 colonies tested), hence the need for multiple wells. The plasmid for gpB is known to have 1602 base pairs. According to the 1kb DNA ladder used, bands appeared around 1600 bp. (B) The plasmid for gpC is known to have 1320 base pairs. The bands on the gel are located around 1300 bp, according to the ladder. These bands were more concentrated than those of gpB. The size of the bands of both gpB and gpC indicate accuracy in the desired plasmid size.

To confirm the success of the cloning, samples 2, 3, 4, 5, and 8 of gpB, and all four samples of gpC in **Figure 6** were sequenced via Eurofins Genomics LLC. The data received was analyzed using Snapgene. The plasmid map showed the plasmid to have both the protein and the GST tag (**Figure 7a and 7b**). However, upon analysis of each of the sequences, only gpC showed to have the insert cloned in frame in the plasmid, meaning that gpC was successfully cloned while gpB was not. While the plasmid of gpB had successfully taken up the insert so that both the sequence for the B portal protein and the GST affinity tag, the sequence was not in frame and was off by one base pair.

In Vitro Assembly of Bacteriophage Lambda Procapsids

As a result of the sequence analysis, only gpC was used in the ratio of combinations, for *in vitro* assembly, with already purified gpE, gpD, and gpNu3. At this point, the GST affinity tags have been cleaved from the individual proteins via the ThermoScientific GST Spin Column Protocol and the HRV 3C protease. The volumes of each protein sample added to the combination vials were calculated based on construct ratios of gpE:gpD:gpNu3:gpC and the sample concentrations as a result of the protein isolation and purification. The calculations were put into a table, as shown in **Figure 8a**.

To test the uniformity of procapsid assembly and formation, each of the combination samples were analyzed using SEC and compared to data of combinations only containing ratios of gpE:gpD:gpNu3 (Appendix A) and procapsid controls (Appendix B). The SEC data of the five different combinations is shown as a function of intensity over time (**Figures 8b-8f**). In comparing the results of the SEC data from previous research, it is evident that there is a shift of the first peak in the most recent data, especially in the results for Sample 4. The new results have the first peak coming off a little after 20 minutes while previous results had the first peak at 40 minutes.

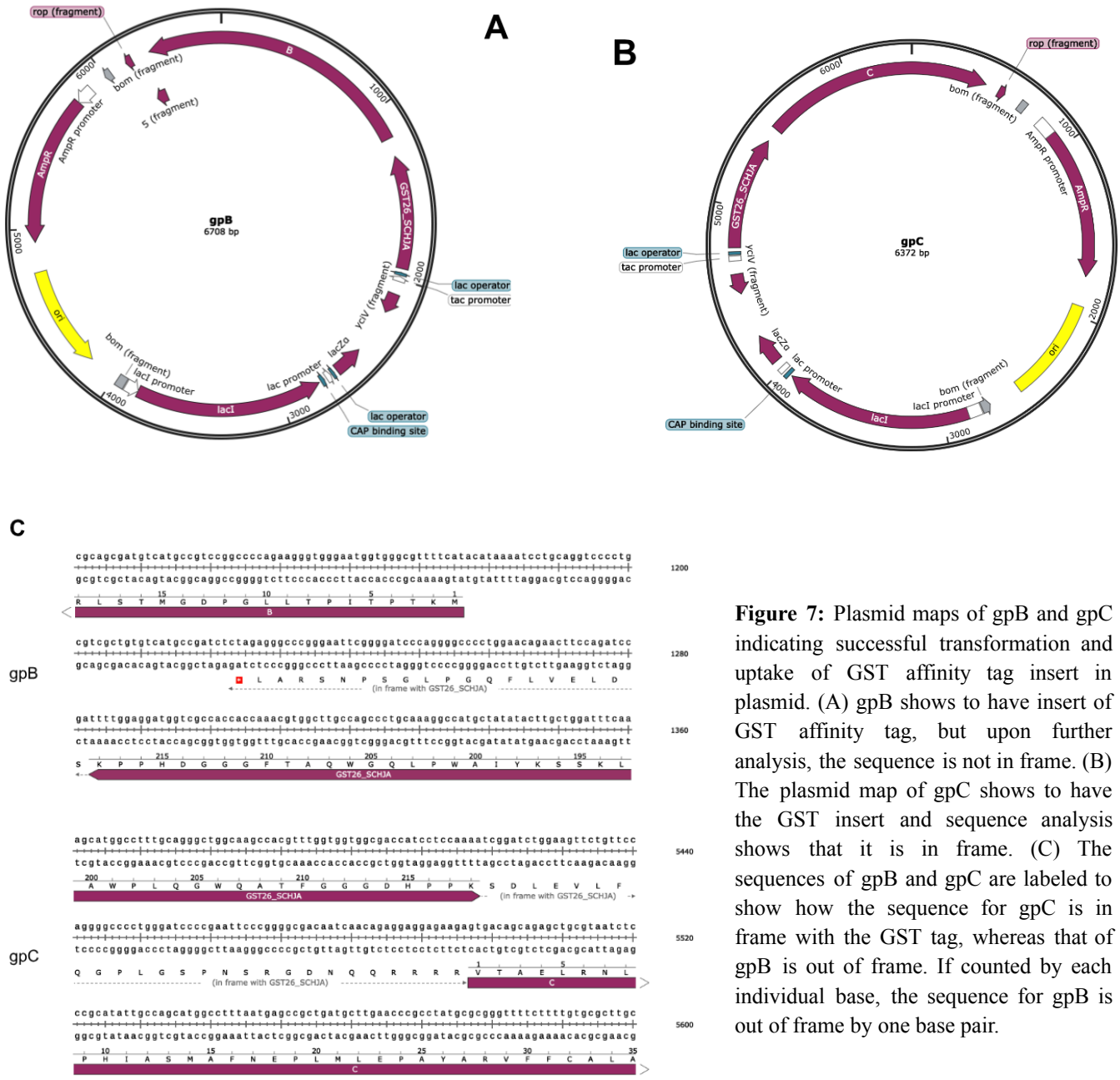


Figure 7: Plasmid maps of gpB and gpC indicating successful transformation and uptake of GST affinity tag insert in plasmid. (A) gpB shows to have insert of GST affinity tag, but upon further analysis, the sequence is not in frame. (B) The plasmid map of gpC shows to have the GST insert and sequence analysis shows that it is in frame. (C) The sequences of gpB and gpC are labeled to show how the sequence for gpC is in frame with the GST tag, whereas that of gpB is out of frame. If counted by each individual base, the sequence for gpB is out of frame by one base pair.

A

Sample #	Ratio (gpE:gpD:gpNu3:gpC)	gpE (µL)	gpD (µL)	gpNu3 (µL)	gpC (µL)	Total volume of protein	Tris Buffer up to 1000 µL	Total sample volume
1	1:1:0.3:0	100.000	73.244	52.560	0	225.804	774.196	1000.000
2	1:1:0.3:0.3	100.000	73.244	52.560	71.41304348	297.217	702.783	1000.000
3	25:0:4:0	100.000	0	28.032	0	128.032	871.968	1000.000
4	25:0:4:4	100.000	0	28.032	38.08695652	166.119	833.881	1000.000
5	12:0:4:4	100.000	0	58.4	79.34782609	237.748	762.252	1000.000

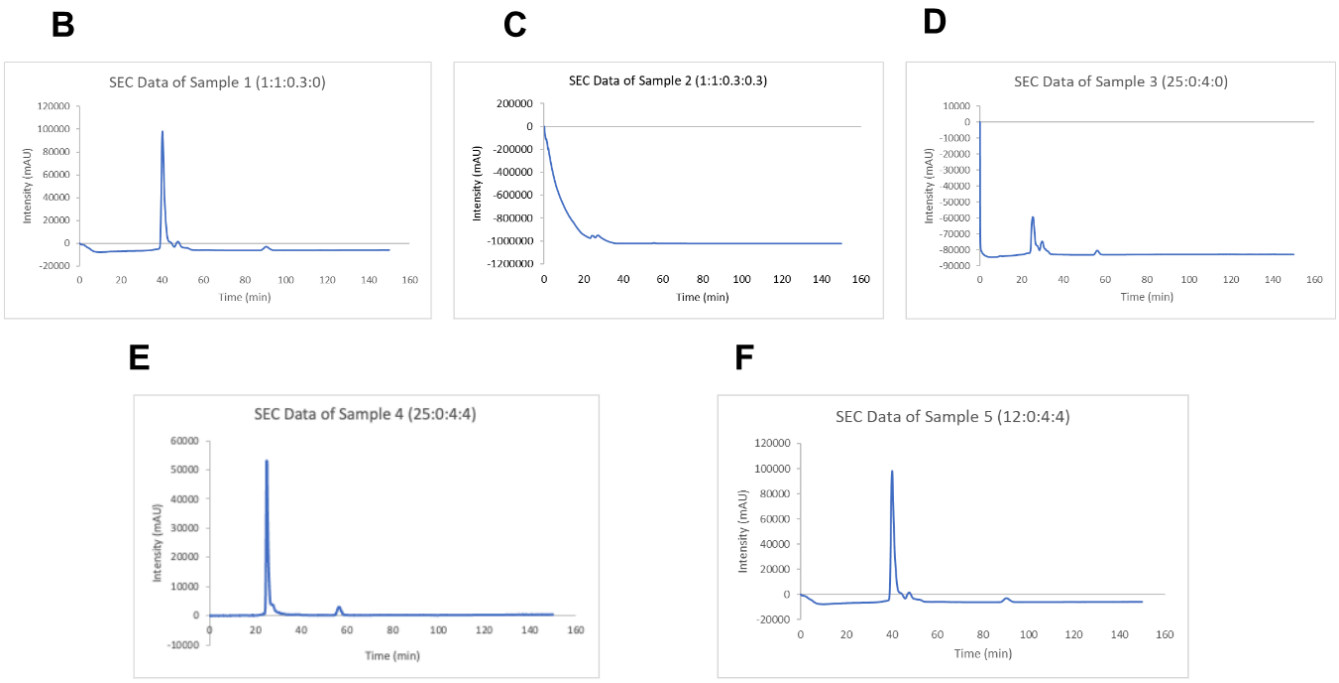


Figure 8: The protein combinations of various ratios are made *in vitro*. (A) Three different ratios of gpE:gpD:gpNu3:gpC were numbered 1-3, combined, and analyzed. The purified proteins were combined based on the concentrations and construct ratios. (B) Using size exclusion chromatography, the incubated samples were analyzed to determine the possibility of intact procapsids. Sample 1 does have a clear peak, but the aggregate comes off at 40 minutes. It did not have gpC present in the combination. (C) Sample 2 did not have a clear peak. Sample 2 had all four proteins present. (D) Sample 3 does not include gpD, and has a small peak at around 25 minutes. (E) Having similar results, Sample 4 did not have gpD and has a sharper peak around 20 minutes, indicating structure uniformity. (F) Sample 5, also lacking gpD, showed a similar peak, but higher. However, the peak was at 40 minutes.

Discussion

Because the bacteriophage lambda procapsid has a unique and purposeful protein structure, it has the potential to be used as vehicles for transporting drugs to specific sites of disease. The fully formed icosahedral shell utilizes the five major proteins, gpE, gpD, gpNu3, gpB, and gpC. Thus, it was necessary to clone the plasmids of gpB and gpC to have the GST affinity tag so the proteins can be produced, isolated, and used for the full structure. Once cloning was completed successfully, then the proteins were produced, purified, and isolated for use in combination and assembly. The combinations of differing ratios of the proteins were done *in vitro* to determine the effectiveness of procapsid formation with the presence of the

new proteins. It was found through sequencing that only gpC was successfully cloned to have the GST affinity tag in frame, so as a result, the ratios of proteins were of gpE:gpD:gpNu3:gpC. In comparison to results from previous research (Appendix A), the presence of gpC in the protein combinations influences a greater structure uniformity. This will be explained through the following insights concerning the *in vitro* assembly of the procapsid.

Sequencing of gpB shows need for site-directed mutagenesis

Upon obtaining the sequencing results of the gpB cloning samples, it was found that although the gene was successfully inserted into the plasmid, as shown in **Figure 7a**, the actual sequence of gpB within the context of the plasmid was determined to be off-set by one base pair. As a result, any gpB produced by the bacterial cells would not be viable for the *in vitro* assembly of bacteriophage procapsids. This single base pair difference in the sequence would potentially result in drastic changes in gene expression, and will not produce the proper protein needed for procapsid formation.

The presence of the single base pair error in the cloning process necessitates site-directed mutagenesis to repair the mistake. Site specific mutagenesis in the plasmid would allow for the insertion/deletion of the missing/extra base pair currently present in the plasmid's sequence. Using a pair of complementary primers containing the desired mutation, the base-pair shift can be corrected, thus allowing the modified plasmid to produce gpB viable for procapsid assembly.

Successful cloning of gpC determined by sequencing allows for its addition in the combination of proteins for procapsid production

PCR gel electrophoresis of the plasmids for gpB and gpC post-cloning indicated that the DNA contained the correct number of base pairs for the respective protein. Knowing that the plasmids are around the right size, and there were no other significant bands created other than the desired band for each protein, sequencing can be completed to analyze the success of the cloning. Sequencing of gpC indicated that the insert was successfully positioned into the vector and in frame. Now the plasmid of gpC contains the sequence for the GST affinity tag, by which the protein can be isolated and purified.

With the plasmid now with the ability to produce gpC with a GST affinity tag, the protein was successfully produced and isolated. The cleavage of the protein from the GST tag using HRV 3C protease completed the purification of the protein and gpC can now be incorporated into the combinations of proteins for assembly. As a result, further research can now produce the gpC protein and isolate the protein through GST tag affinity chromatography and then use gpC in the combination ratios.

The function of the protease gpC is to cleave the scaffolding protein gpNu3 from the structure.

In vitro assembly of capsid proteins with gpC present yield SEC results similar to those of full, intact procapsids

Previous research has found that intact procapsids made *in vivo* come off of the SEC column at around 20 minutes (Appendix A). This data is used as a reference of comparison. The combination of proteins with ratio 1:1:1 of gpE:gpD:gpNu3 has been imaged using

Cryo-EM (**Figure 4**), which revealed tubular aggregates. However, it has also been shown in previous research that a sharper peak indicating a uniform structure occurs when gpE, gpD, and gpNu3 are combined in a 1:1:0.3 ratio (Appendix B). This combination has not been imaged through Cryo-EM and therefore it cannot be determined how the structure of the procapsids is being affected. The SEC data for this sample shows a peak at around 40 minutes, which had been the most uniform peak obtained in previous research. For this project, Sample 1 was used to replicate these results with the same ratio. The SEC data obtained for this sample is much cleaner and sharper than in previous research, reaching about 1.0E5 mAU in intensity, but still comes off late around 40 minutes (**Figure 8b**).

Sample 2 was included as a means of testing all four proteins obtained: gpE, gpD, gpNu3 and gpC. Since the function of gpC is to be a protease that cleaves out gpNu3, the same protein amount of gpC was added to the combination. For this particular sample, gpNu3 was added first, then gpE, gpD, and lastly, gpC. Although all four proteins were present and utilized, there was no significant peak in the SEC results (**Figure 8c**), but only a significant decline from the start. The combinations possibly formed large aggregates and came off of the column early due to its size. The order in which the proteins were added could have been a factor in how they aggregated. Another factor is incubation time.

The ratio of Sample 3 comes from literature[6] stating that between gpE and gpNu3, a ratio of around 25:4 is present in a procapsid shell made *in vivo*. A small peak at around 20-30 minutes is present (**Figure 8d**), but with a negative intensity value, meaning there could have been an error or difference in the refractive index during the SEC run.

Like the previous sample, Sample 4 did not have gpD present, but had the addition of gpC for a ratio of gpE:gpD:gpNu3:gpC to be 25:0:4:4. As mentioned previously, the equal amount of gpC protein as gpNu3 protein was added since the role of gpC is to degrade the scaffolding protein once gpE has formed around it. The SEC run of this sample resulted in the sharpest and cleanest peak yet, coming off the column at 20 minutes with an intensity of about 5.25E4 mAU (**Figure 8e**). Due to the sharpness of the peak and the time in which the peak shows, there is a high possibility of high structure uniformity of the procapsids. However, this cannot be fully determined until imaging of the sample by Cryo-EM. Since gpD was not added, this could indicate that this protein needs to be added to the sample later on in the order of proteins, perhaps after gpC has cleaved gpNu3. The function of gpD, a decoration protein, is to stabilize the structure of the procapsid, especially the gpE capsid proteins. Since the procapsid would eventually be broken down to administer the drug, the presence of gpD is not completely necessary. The other proteins will still come together. However, the ability to stabilize the structure can still be an important facet, thus, the protein could be added after the cleaving of gpNu3.

Sample 5 has a 12:0:4:4 ratio of gpE:gpD:gpNu3:gpC, which was not based on any literature or previous knowledge, but was a construct ratio that was yet to be tried. While the peak obtained for this sample had a higher intensity value of about 1.0E5 mAU in comparison to the previous sample, the peak occurred at 40 minutes (**Figure 8f**). This indicates high structure uniformity, but is still

off in size to intact procapsids due to its elution time.

It is evident in the final two combination samples that the presence of gpC contributes greatly to the aggregation and size of the procapsids being made. It cannot be determined if the proteins are coming together in the right orientation until they are imaged through Cryo-EM. However, based on the results of the SEC data, the uniformity of the structure based on the clean, sharp peaks greatly increases with the presence of gpC. It is also important to note that gpD was not present in these final two combination samples. When it was present in Sample 2, the uniformity of the structure seems to decrease based on the SEC data. While gpD is important in procapsid stability, it is possible it operates best after gpNu3 has been degraded by gpC. A future step would be to test what happens when gpD is added after incubating with gpNu3, as this has not been determined yet.

Conclusion and Future Directions

In sum, this study shows the evident importance of gpC in the *in vitro* assembly of the bacteriophage lambda procapsid. The samples in which gpC was present and gpD was not present had the cleanest peaks. However, the combination ratio that appears to have the highest structure uniformity is the ratio of gpE:gpD:gpNu3:gpC of 25:0:4:4. In determining the actual uniformity of the procapsids, future directions for this sample would be to image the procapsids with Cryo-EM.

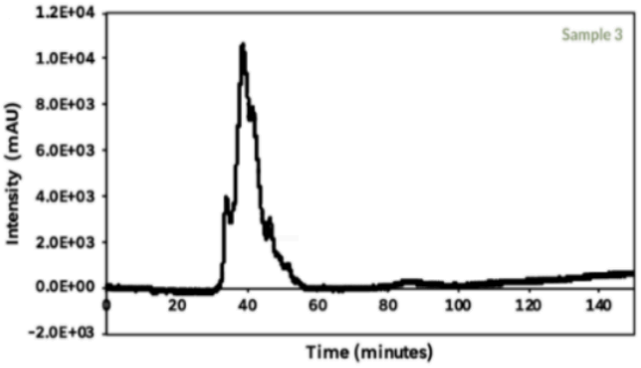
Since the presence of gpC significantly improves the uniformity of the structure, potential avenues for the next steps in this research would be to test the presence of gpB and gpD on the assembly of the procapsid.

For gpD, the presence and amount of this protein greatly influences the assembly of the structure. Future research can be done in determining the effect of the presence of gpD on the 25:0:4:4 ratio mentioned earlier. Various amounts of gpD can be added to the combination after the degradation of gpNu3 by gpC.

Finally, the incorporation of gpB in the combination ratio would be the next potential avenue to explore. In order for this to be done, site-directed mutagenesis of the gpB cloned in this study would be a required step to complete. This portal protein is present in the final icosahedral structure of procapsids and functions for the passage of molecules. While it can be utilized as the means of encapsulation and administration of drugs, its presence in the *in vitro* assembly could play an important role in the stability of the structure and its formation.

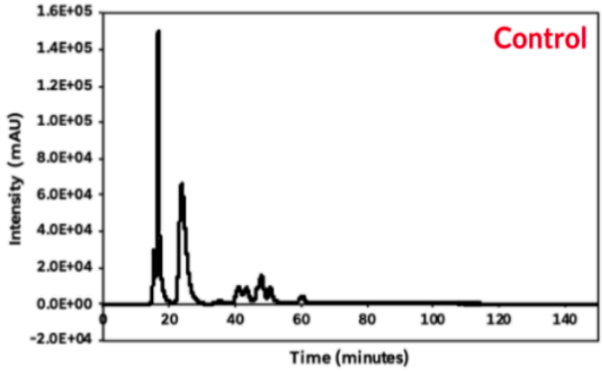
Appendix

A



Sample #	Ratio (gpE:gpD:gpNu3)	gpE (µL)	gpD (µL)	gpNu3 (µL)	Total volume of protein	Tris Buffer up to 1500 µL	Total sample volume
3	1:1:0.3	330.000	220.000	87.429	637.429	862.571	1500.000

B



Appendix: Size Exclusion Chromatography data of previous combinations and procapsid controls made in the 2023 Honors Research Project by Katie Biscocho. Combination samples with various ratios of gpE:gpD:gpNu3 were run through the SEC column, with Sample 3 showing the most uniform peak. (A) Sample 3 with a peak at 40 minutes of sample combination ratio of 1:1:0.3 of gpE:gpD:gpNu3. (B) Procapsid control samples were made *in vivo* and run through the SEC column to use as comparison for the *in vitro* combination samples. A sharp peak at 20-30 indicates the reference peak for intact procapsids.

References

[1]Yusuf, Azeez et al. "Nanoparticles as Drug Delivery Systems: A Review of the Implication of Nanoparticles' Physicochemical Properties on Responses in Biological Systems." Polymers vol. 15, 23 Mar. 2023, doi:10.3390/polym15071596

[2]Anselmo, A.C. and S. Mitragotri. 2016. Nanoparticles in the Clinic. Bioeng Transl Med. 1:10-29. doi: 10.1002/btm2.10003.

[3]Chung, Y.H, Cai H., & Steinmetz, N.F. (2020). Viral nanoparticles for drug delivery, imaging, immunotherapy, and theranostic applications. *Advanced Drug Delivery Reviews*, 156, 214-235. https://doi.org/10.1016/j.addr.2020.06.024

[4]Kasman L.M, Porter L.D. 2022. Bacteriophages. StatPearls. Treasure Island (FL): StatPearls Publishing; 2024,Jan. https://www.ncbi.nlm.nih.gov/books/NBK493185/

[5]Lander, G.C., A. Evilevitch, M. Jeembaeva, C.S. Potter, B. Carragher, and J.E. Johnson. 2008. Bacteriophage Lambda Stabilization by Auxiliary Protein gpD: Timing, Location, and Mechanism of Attachment Determined by Cryo-EM. Cell Press. 16(9):1399-1406. doi.org/10.1016/j.str.2008.05.016

[6]Medina, E. M., Andrews, B. T., Nakatani, E., & Catalano, C. E. (2011). The bacteriophage lambda gpNu3 scaffolding protein is an intrinsically disordered and biologically functional procapsid assembly catalyst. *Journal of molecular biology*, *412*(4), 723–736. https://doi.org/10.1016/j.jmb.2011.07.045

[7]Medina, E., Wieczorek, D., Medina, E. M., Yang, Q., Feiss, M., & Catalano, C. E. (2010). Assembly and Maturation of the Bacteriophage Lambda Procapsid: gpC Is the Viral Protease. Journal of Molecular Biology, 401(5), 813–830. https://doi.org/10.1016/J.JMB.2010.06.060

[8]Dedeo, C. L., Cingolani, G., & Teschke, C. M. (2019, September 29). *Portal protein: The orchestrator of capsid assembly for the dsdna tailed bacteriophages and herpesviruses*. Annual review of virology. https://doi.org/10.1146/annurev-virology-092818-015819