Construction and characterization of a hybrid bacteriophage gene delivery platform

by

Shirley Wong

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Pharmacy

Waterloo, Ontario, Canada, 2022

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner DR. ZEINAB HOSSEINIDOUST
Assistant Professor
Department of Chemical Engineering
McMaster University

Supervisor DR. RODERICK SLAVCEV
Associate Professor
School of Pharmacy
University of Waterloo

Internal Member DR. JONATHAN BLAY
Professor
School of Pharmacy
University of Waterloo

Internal Member DR. TREVOR CHARLES
Professor
Department of Biology
University of Waterloo

Internal-external Member DR. MARC AUCOIN
Professor
Department of Chemical Engineering
University of Waterloo

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

The work described in this thesis was carried out solely by Shirley Wong under the supervision of Dr. Roderick Slavcev, with the following exceptions:

- in Chapter 3, Dr. Salma Jimenez contributed to experimental design and agarose gel images in Figures 3.7 and 3.8;

- in Chapter 5, Dr. Salma Jimenez contributed to experimental design and transfection data in Figures 5.2 and 5.8 to 5.10; and

- in Chapter 6, Dr. Salma Jimenez contributed to experimental design and agarose gel images in Figures 6.5 and 6.9.

Abstract

Bacteriophages, viruses that specifically infect bacteria, may be considered as structurally simplistic protein-based vehicles for ferrying nucleic acid cargo. Since phages cannot infect nor replicate within human cells, they can be used for non-viral gene delivery. The filamentous phage M13 is a single-stranded DNA phage that co-exists with its host after infection. M13 possesses several attractive characteristics for gene delivery: 1) high titres of phage progeny, 2) capsid proteins amenable to display of foreign peptides to modify tropism, 3) a simple, well-characterized genome that is easy to genetically modify, and 4) a flexible, theoretically limitless transgene cargo capacity. Although non-viral vectors may be safer than vectors derived from infection-competent human viruses, they overall suffer from a lack of efficacy due to the absence of innate pathways for mammalian cell uptake and gene expression. Bacterial sequences in conventional gene transfer plasmids and in phage genomes often consist of elements only necessary for amplification in a prokaryotic background, and as such, are superfluous in the mammalian cell. These problematic elements include antibiotic resistance genes, which can disseminate antibiotic resistance, and CpG motifs, which are inflammatory in animals and can lead to transgene silencing.

We examined how M13 could be improved as a gene delivery vector by removing its phage backbone in conjunction with phage display of a cell-specific targeting ligand. As a proof of principle, a mammalian transgene cassette was flanked by isolated initiation and termination elements of the phage origin of replication. Phage proteins provided in trans by a separate helper phage would replicate only the cassette, without any bacterial or phage backbone. Helper phage rescue of "miniphagemids" from these split origins was equal to, if not greater than, isogenic "full phagemids" arising from intact origins. However, helper phage efficiency appears to be highly dependent on host background, as striking differences in phagemid:helper composition were observed between different host bacterial strains.

To examine their capacity for gene delivery, we produced mini and full phagemid particles encoding the luciferase reporter transgene cassette and displaying a cell-specific targeting ligand, the epidermal growth factor (EGF). Across cell lines expressing the EGF receptor (EGFR), the display of EGF was the single greatest factor affecting transgene expression, underlining the necessity of targeting receptor-mediated endocytosis for efficient phage uptake. Importantly, smaller vector size was associated with improved transgene expression both for purified single-stranded phagemid

DNA and phagemid virion particles. The display of cell-specific targeting ligands improves cellular internalization while the miniaturized phagemid vector improves intracellular localization and evades immune-induced transgene silencing. As single-stranded DNA transfects poorly in mammalian cells, we also set out to construct a double-stranded phage gene transfer vector. We were able to demonstrate helper phage-mediated rescue of double-stranded phagemid DNA. Production of a double-stranded phagemid was highly inefficient, but we identified several key areas for further investigation and improvement.

Overall, we characterized the production and function of a novel set of targeted transgene delivery vectors based on the filamentous phage M13. Its capacity for the display of a cell-targeting ligand in combination with the straightforward miniaturization of its genetic cargo showcases the flexibility of M13 as a gene delivery tool with immense therapeutic potential. Altogether, these advancements can further increase the efficiency of M13 phage-mediated gene transfer.

Acknowledgements

My deepest gratitude goes out to my family. Throughout my graduate studies, they have encouraged me after every misstep and celebrated every milestone. I am immensely grateful for their unconditional and loving support.

Finally, this thesis could not exist in its current form without the help of Dr. Dmitri Iouchtchenko, to whom I am eternally grateful.

Dedication

To my beloved mother, Sio Hong Wong

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AAV	adenoviral associated virus
Ad5	adenovirus serotype 5
AGE	agarose gel electrophoresis
AP	activator protein
BCA	bicinchoninic acid
BSA	bovine serum albumin
C_q	threshold cycle number
CCC	circular covalently closed
CD	cluster of differentiation
CELiD	closed-ended linear duplex
CFU	colony forming units
CMV	cytomegalovirus
CpG	cytosine-guanine dinucleotide
CpG-DNA	DNA containing unmethylated CpG motifs
CPP	cell-penetrating peptide
CTL	cytotoxic T lymphocyte
DAPI	$4',6$-diamidino-2-phenylindole
DC	dendritic cell
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
ds	double-stranded
DSB	double-stranded DNA break
DTS	DNA nuclear targeting sequence

EGF	epidermal growth factor
EGFR	epidermal growth factor receptor
ELISA	enzyme-linked immunosorbent assay
f1 *ori*	f1 functional origin
FBS	fetal bovine serum
F-factor	fertility factor
Ff phage	F-specific filamentous phage
GFP	green fluorescent protein
HAP	hydroxyapatite
HBEGF	heparin-binding EGF-like growth factor
HIV	human immunodeficiency virus
HMW	high molecular weight
HRP	horseradish peroxidase
IFN	interferon
IHF	integration host factor
IL	interleukin
IR	intergenic region
IRF	interferon regulatory factor
ITR	inverted terminal repeat
LB	Luria–Bertani
LCC	linear covalently closed
lipoplex	complex of cationic lipid and nucleic acid
LPS	lipopolysaccharide
LRR	leucine-rich repeat
mAb	monoclonal antibody
MAPK	mitogen-activated protein kinase
MAPKK	MAPK kinase
MAPKKK	MAPKK kinase
MBN	mung bean nuclease
MHC	major histocompatibility complex
MIDGE	minimalistic immunologically defined gene expression
MiLV	micro-linear vector

MOI	multiplicity of infection
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
MyD	myeloid differentiation primary-response protein
NF	nuclear factor
NK	natural killer
NLS	nuclear localization signal
NSCLC	non-small cell lung cancer
ori	origin of replication
OTC	orinithine transcarbamylase
PAMP	pathogen-associated molecular pattern
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
pDC	plasmacytoid dendritic cell
PEG	polyethylene glycol
PEI	polyethylenimine
PFU	plaque forming units
phage	bacteriophage
PRR	pattern recognition receptor
PS	packaging signal
PTB	phosphotyrosine binding
qPCR	quantitative PCR
RES	reticuloendothelial system
RF	replicative factor
RFP	red fluorescent protein
RFx	recombinant RF
RGD	Arg-Gly-Asp
RNA	ribonucleic acid
RTK	receptor tyrosine kinase
SAS	sense-antisense
SD	standard deviation
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SEM	standard error of the mean

Shc2	Src homology 2
SIV	simian immunodeficiency virus
SRE	c-fos serum response element
ss	single-stranded
SSB	single-stranded binding
TAK	TGF-β-activated kinase
TGF-α	transforming growth factor-α
Th	T helper lymphocyte
TIR	Toll/interleukin-1 receptor homology
TLR	Toll-like receptor
TN	Tris-NaCl
TNF-α	tumour necrosis factor α
UV	ultraviolet
VITT	vaccine-induced immune thrombotic thrombocytopenia
X-SCID	X-linked severe combined immunodeficiency

Chapter 1

Introduction

1.1 Gene therapy

Gene therapy strives to treat disease through the delivery, transcription, and translation of exogenous genes for therapeutic effect. As the nature of the effect is dictated by the choice of transgene, gene therapy has innumerable applications. To date, over 3000 indications have been addressed by gene therapy clinical trials, the majority of which ($\geq$65%) were cancer diseases [1]. Additionally, over 3000 gene therapy clinical trials have been reported to date, 130 of which were initiated in 2020 alone.

A myriad of diverse gene delivery vector types are under investigation or currently in use. Out of 22 gene therapies that have been approved worldwide as of 2020, 7 were viral, 7 were ribonucleic acid (RNA) (including antisense RNA), 2 were deoxyribonucleic acid (DNA), and 6 were recombinant cell therapy [2].

Efficacious gene transfer depends on the exogenous transgene bypassing both inter- and intracellular barriers to penetrate into a cell's nucleus where gene expression can occur (Figure 1.1). The success of any gene therapy is wholly dependent on the ability of the delivery vehicle, known as the vector, to navigate these barriers. Depending on the route of administration, a vector first requires transport across the skin, the respiratory system, or the digestive system. The most common route is needle injection, either into dermal or muscular tissue or directly into the bloodstream [3, 4]. Other forms of delivery, including nasal sprays or oral capsules are under rapid development. In the blood, a vector must avoid clearance by the reticuloendothelial system (RES) and other immune cells, or aggregation with serum proteins and other blood components. To impart therapeutic effect, the transgene must bypass the plasma membrane and enter a cell; this is generally mediated by endocytic, phagocytic or pinocytic mechanisms which would force the vector into the endolysomal compartment. After endosomal escape, some gene therapies may be effective in the cytoplasm, but nuclear translocation is generally necessary for transcription of the transgene into functional products.
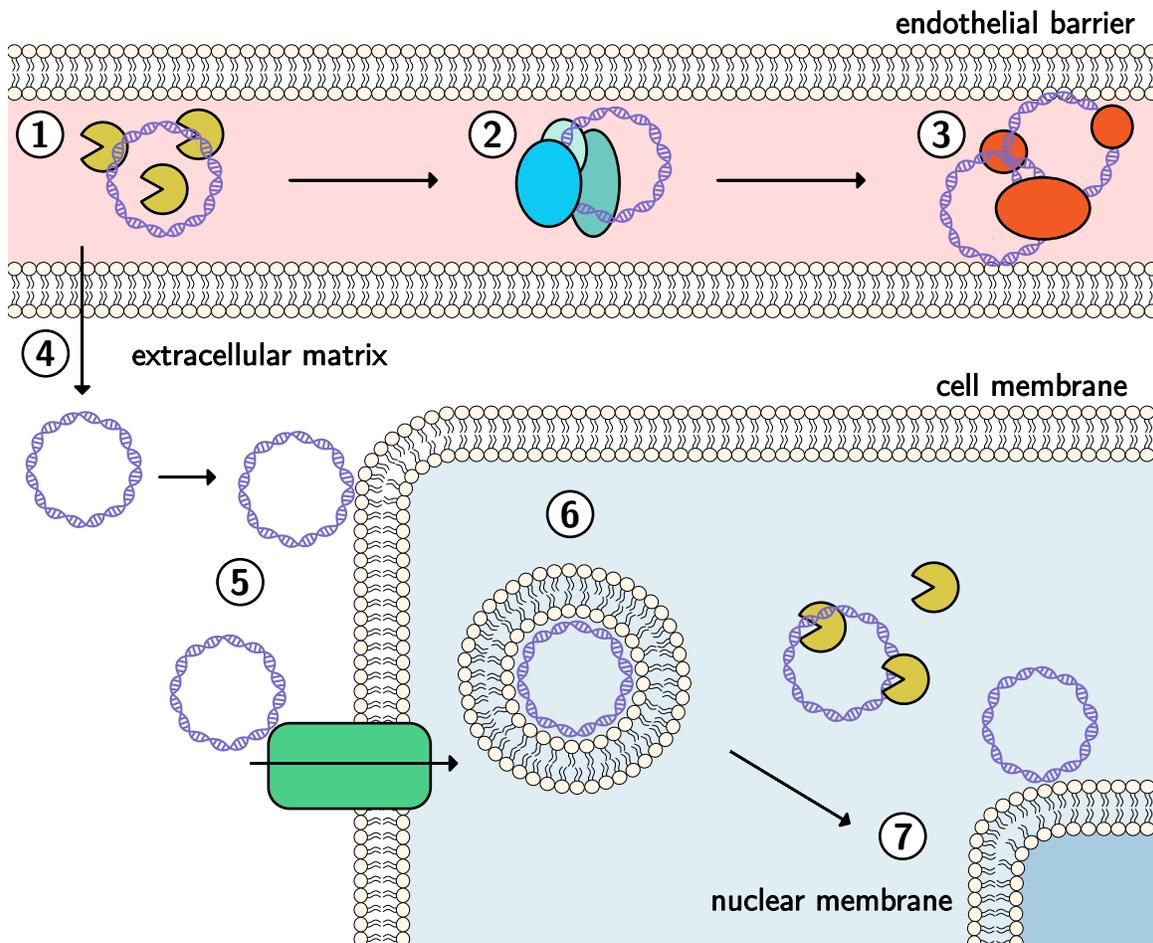
Figure 1.1: Barriers to gene delivery. Once inside the body, gene delivery vectors must bypass several barriers to achieve gene transfer. 1) Serum nucleases that can break down DNA, 2) clearance by the reticuloendothelial system, 3) aggregation by serum proteins, 4) extravasation across the endothelial membrane, 5) uptake by target cells, 6) translocation across cell to nucleus, including escape from the endosomal pathway, 7) import across the nuclear membrane.

1.1.1 Viral gene therapy

Most gene transfer vectors that reach clinical trial are recombinant human viruses (Figure 1.2), such as adenovirus (542 trials to date) and retrovirus (528 trials). In particular, the adenoviral associated virus (AAV) has received increasingly more attention as a vector (263 trials) due to its lack of pathogenicity [5]. AAV infection can only be established in the presence of a helper virus (either adenovirus or herpesvirus), so it has emerged as a highly safe and efficacious viral vector [6]. Many different RNA and DNA-encoding viruses have been investigated for gene transfer overall [7]. Their natural tropism for human cells and high transduction efficiency make viruses highly capable as gene delivery vectors. In general, viral gene transfer vectors are typically rendered deficient of replicative genes for safety [8]; in their place, a transgene is inserted. In some indications, retaining the ability to replicate can be more useful, for example, in oncolytic viral therapy [9]. Even so, a number of dangerous disadvantages have plagued development of viral vectors since their first use decades ago.
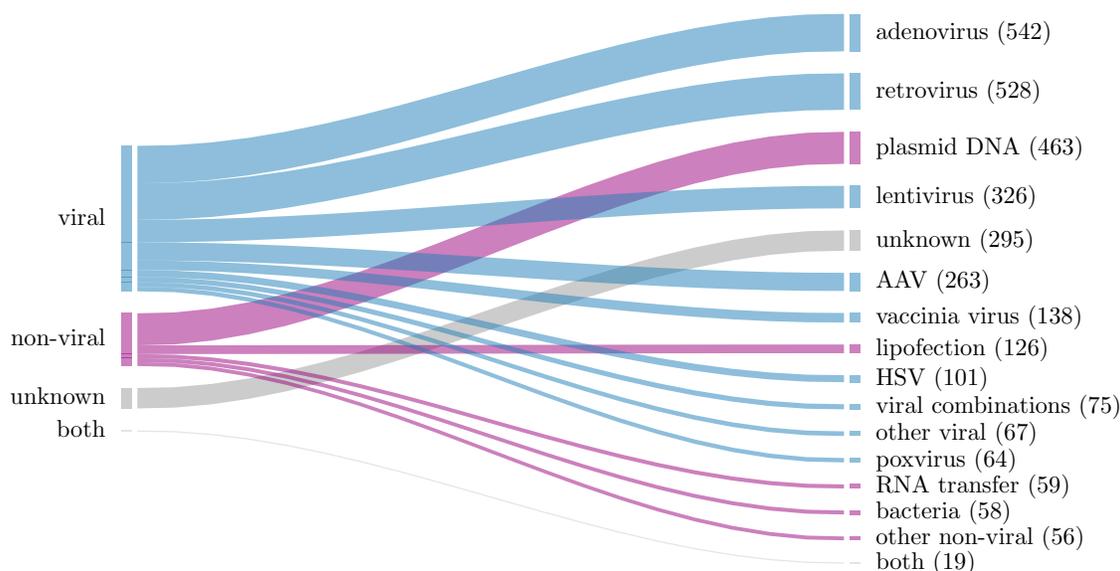
Figure 1.2: Types of vectors in gene therapy clinical trials. The majority of vectors that reach clinical trial are viral, making up 75% of all represented vectors. Out of 3180 total trials, 19 pair a viral vector with either DNA or RNA ("both"). AAV: adeno-associated virus, HSV: herpes simplex virus. Data is from Gene Therapy Clinical Trials Worldwide, November 2021 (https://www.abedia.com/wiley) [1].

In retroviral vector production, two laboratories have reported spontaneous recombination events between the producer cell line and the retroviral vector: this is a dangerous safety concern as it enables the uncontrolled production of replication-competent retrovirus [10–12]. In gene therapy clinical trials for X-linked severe combined immunodeficiency (X-SCID) [13, 14], retroviral vector integration into the genome led to insertional mutagenesis and, consequently, onset of leukemia-like disorders. This manifested in four of nine treated patients in one trial [15, 16] and one patient

in another trial [17]. In another clinical trial for orinithine transcarbamylase (OTC) deficiency, severe immunotoxicity from the adenoviral vector resulted in the tragic death of 18-year-old Jesse Gelsinger [18]. His death prompted serious questions regarding the state of the field [18–20].

More recently, onset of thrombosis with thrombocytopenia and anti-platelet antibodies has been linked to two adenoviral COVID-19 vaccines. This condition, termed vaccine-induced immune thrombotic thrombocytopenia (VITT), was hypothesized to arise from adenovirus-mediated immune interactions [21]. Also recently, a clinical trial evaluating the delivery of the myotubularin-1 gene to treat X-linked myotubular myopathy using AAV was halted due to the deaths of four trial patients [22]. This has prompted concerning questions regarding the previously well-accepted safety profile [23] of AAV for gene transfer [24, 25].

Due to their high prevalence in nature, pre-existing immunity against many of the common viral vector serotypes has developed across the globe [26, 27]. Pre-existing immunity to adenovirus serotype 5 (Ad5) was associated with increased risk of human immunodeficiency virus (HIV) acquisition in men after vaccination with an Ad5-vectored HIV vaccine [28, 29]. Even aside from immunotoxic effects, immune clearance of the vector can dramatically inhibit gene transfer efficacy and prohibit repeated administration of a viral vector [30, 31]. Furthermore, genetic cargo capacity is dictated by the relatively inflexible capsid size of the virus, potentially limiting what genes a viral vector may transfer. In general, high immunogenicity, risk of insertional mutagenesis, as well as the possibility of recombination into replication-competent viruses are dire safety concerns for viral vectors, while pre-existing immunity and cargo capacity are limiting efficacy concerns [32].

1.1.2 Non-viral gene therapy

Alternatively, there are gene transfer vectors not based on human viruses: non-viral vectors. Non-viral vectors encompass all other types of vectors not originating from human viruses. Broadly speaking, a non-viral vector is a combination of nucleic acid material (DNA or RNA) that may have some sort of protective vehicle to bypass barriers to transfection (Figure 1.1). Both nucleic acids are used in gene therapy strategies, but the focus of this thesis will be on the transfer of DNA, although the transfer of RNA is also an area of immense therapeutic potential.

The DNA molecule is a polymer of two complementary chains of nucleotides coiled together in a helical conformation [33]. The nucleotide is the monomeric base unit; each nucleotide contains one of four nitrogenous bases (A: adenine, G: guanine, T: thymine, or C: cytosine) covalently linked to a deoxyribose saccharide unit and a phosphate group. As such, each DNA strand is comprised of a sequence of nucleotides covalently linked through a sugar-phosphate backbone. Hydrogen bonding between corresponding base pairs (A with T, C with G) of two complementary DNA strands forms the double-stranded (ds) DNA helix. The genetic sequence thus defines the set of instructions for downstream RNA and protein production.

The negative charge from the phosphate backbone of DNA can promote interactions with serum proteins, inviting destruction by serum nucleases and/or RES clearance [34]. A large charged macromolecule like DNA cannot easily extravasate across the blood vessel [35] nor cross the phospholipid bilayer of the cell (Figure 1.1). Spontaneous uptake of DNA by cells has been observed in vitro, albeit at very low frequency [36]. Uptake seems primarily mediated by endocytic mechanisms as confocal microscopy has shown exogenous DNA within the endosomal compartment. DNA exhibits a half-life between 50 and 90 min once in the cytoplasm, as measured in HeLa and COS cells [37]. Diffusion does not appear to play a large role in the cytoplasmic movement of internalized DNA; in fact, DNA longer than 2000 bp shows little to no diffusion at all [38]. Nuclear translocation in the cytoplasm is primarily mediated by active transport along the microtubule network [39]. Inclusion of protein nuclear localization signals (NLSs) [40] or nucleotide DNA nuclear targeting sequences (DTSs) [41, 42] is associated with vastly improved nuclear uptake, mediated by existing active nuclear import mechanisms [39, 43].

Non-viral methods of gene delivery can be divided into two major categories: physical methods that force exogenous DNA into cells, or chemical methods that encapsulate the material to assist its endocytic uptake and endosomal escape [4, 32, 44]. Physical methods such as needle injection, jet-injection, or electroporation can deposit DNA directly into the cytoplasm of cells, bypassing several barriers to entry altogether [4]. The most common non-viral gene transfer strategy to date has been direct injection of plasmid DNA in 463 clinical trials (Figure 1.2). Methods such as electroporation have been shown to enhance uptake over simple injection [45]. However, physical methods of gene transfer tend to be restricted in use due to tissue accessibility and the capacity for tissue damage; as such, they are generally not suitable for systemic gene delivery [4].

In contrast, chemical gene delivery carriers encompass a broad range of cationic polymers that can negate the negative charge of the phosphate backbone of DNA and mediate interactions with the cell membrane [4]. A common popular reagent is the cationic lipid, which typically consists of a hydrophobic domain linked to a positively charged head group. Cationic lipids are typically formulations of a mixture of cationic and neutral lipids that can encapsulate negatively charged DNA. Upon mixing, cationic lipids condense nucleic acids into complexes of cationic lipid and nucleic acid (lipoplexes), which can then fuse with the cell membranes to facilitate cellular uptake and mediate endosomal escape. Gene transfer using such reagents is commonly referred to as "lipofection", which is now the gold standard for transfection. Lipid nanoparticles comprised of a mixture of cationic and neutral lipids have also shown great promise as DNA carriers. Similarly, cationic polymers can form "polyplexes" with DNA for enhanced cell uptake and processing. Non-viral gene transfection has been additionally demonstrated using a rapidly growing array of different cationic polymers and other materials such as gold, carbon, and chitosan [46].

For the most part, non-viral gene delivery is episomal and has not generally been associated with insertional mutagenetic events [47, 48]. Non-viral gene delivery methods are also not as immunogenic as viral vectors and therefore boast a much stronger safety profile [4]. Neither

anti-nuclear nor anti-DNA antibodies have been observed after repeated administrations of plasmid DNA in vivo [49]. Without the limitations of a capsid shell, non-viral gene transgene vectors also do not have any genetic capacity limits. Despite these safety advantages, the main hurdle facing non-viral gene vector development is low transfection efficiency, especially in comparison to viral vectors [4, 44]. Viruses have evolved to efficiently enter a cell, escape the endosomal compartment, and enter the nucleus; but non-viral gene transfer strategies must incorporate solutions to each obstacle de novo.

1.1.3 Gene transfer of plasmid DNA

Non-viral gene therapies encompasses many forms of DNA, but the most common and most well-studied DNA vector is purified bacterial plasmid DNA. The plasmid is a circular covalently closed (CCC) dsDNA molecule typically amplified in bacteria. It is comprised of two major components: the expression cassette encoding the gene of interest and regulatory elements, and the prokaryotic backbone necessary for amplification and maintenance of that plasmid (Figure 1.3). The backbone includes elements such as a bacterial origin of replication (*ori*) and an antibiotic resistance selective marker [34]. Potential horizontal gene transfer of the *ori* and antibiotic resistance marker risks dissemination of replicative plasmid and antibiotic resistance in the environment [50, 51]. Regulatory bodies discourage the use of such selection markers in plasmid therapeutics and recommend non-antibiotic methods of selection for clinical use [50].
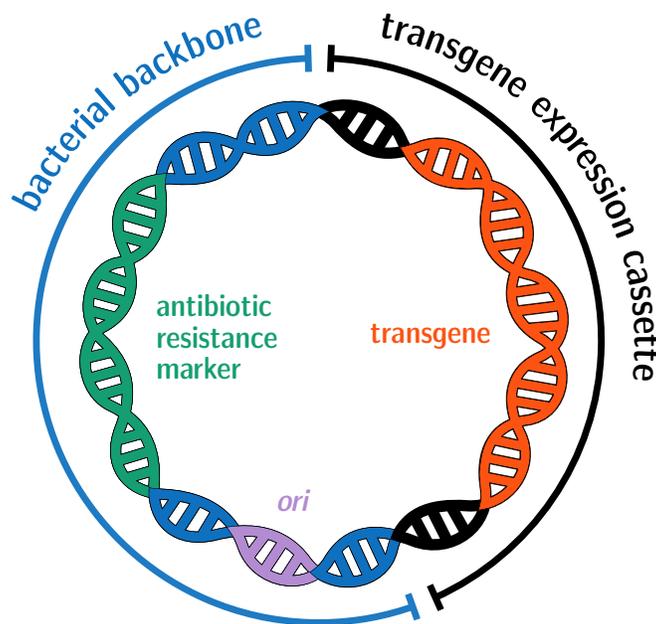
Figure 1.3: Schematic of a gene transfer plasmid. A plasmid is a closed circular double-stranded DNA vector. For expression in mammalian cells, an expression cassette (black) typically contains the transgene (orange) with all necessary regulatory elements. It must also have a prokaryotic backbone (blue) that contains a bacterial origin of replication (lavender) and an antibiotic resistance marker (green) for amplification in bacterial cells.

Despite their rarity [47, 48], recombination events between plasmid and chromosomal DNA can still occur [45]. In such an event, vector integration could still lead to insertional mutagenesis and, potentially, oncogenesis. Using a quantitative gel purification assay [52], the rate of plasmid integration into the chromosome was measured by several groups. Briefly, DNA isolated from treated tissues is subjected to pulsed-field gel electrophoresis to separate high molecular weight (HMW) genomic DNA from free episomal vector [47]. After multiple rounds of purification, any remaining associated vector DNA is assumed to have integrated; the HMW DNA is then probed by quantitative PCR (qPCR) to quantify vector presence. Using this method, the amount of genome-associated plasmid DNA from simple muscular injection was observed to be between 1 and 30 copies/µg DNA [45, 47, 48]. From these results, Ledwith et al. (2000) calculated a "worst-case" gene mutation frequency on the order of 10^{-9} mutations per gene, which is well below their estimated basal frequency of gene-inactivating mutations: 2×10^{-6} [47, 53].

Wang et al. (2004) observed that supplementing injection with electroporation greatly increased the gene transfer rate (up to 38-fold); however, it also increased the amount of genome-associated vector to 700–980 copies/µg DNA, which was more than 40-fold greater than injection without electroporation [45]. Even the higher corresponding gene mutation frequency still fell below the estimated spontaneous frequency of gene-inactivating mutations, suggesting that vector integration is not as high a concern for plasmid-based gene transfer. However, it is expected that gene transfer rates will improve alongside new advances in non-viral technology; as such, plasmid integration rates may consequently continue to climb and may become a problem in the future.

While anti-DNA antibody production is not associated with DNA gene transfer, it has been well-documented that foreign DNA can induce inflammatory responses [54–56]. What differentiates mammalian genomes from bacterial is the frequency of the cytosine-guanine dinucleotide (CpG). CpG motifs occur at an expected frequency assuming random base utilization in bacterial genomes (1 in 16 bases), but far more rarely in vertebrate genomes (1 in 50-60 bases) [57–59]. Furthermore, the cytosine in these CG pairs is almost always methylated in vertebrates [59], but tends to be unmethylated in bacteria. As a characteristic of "non-self" DNA, CpG motifs are a pathogen-associated molecular pattern (PAMP), one of many motifs recognized by pattern recognition receptors (PRRs) of the innate immune system [60–62].

The Toll-like receptor (TLR) 9 mediates innate immunity against DNA containing unmethylated CpG motifs (CpG-DNA). As a member of the TLR family of membrane glycoproteins, TLR9 is characterized by its leucine-rich repeat (LRR) [63] and Toll/interleukin-1 receptor homology (TIR) domains [64]. Highly abundant in plasmacytoid dendritic cells (pDCs) and B cells, TLR9 is both necessary and sufficient for CpG-mediated immunostimulation [65, 66]. CpG immunostimulation strongly resembles a T helper lymphocyte (Th)-mediated immune response, specifically that of Th1 [36, 67, 68]. Recognition of CpG-DNA by TLR9 stimulates the TLR/interleukin (IL)-1R signalling pathway, ultimately activating both nuclear factor (NF)-κB and mitogen-activated protein kinase (MAPK) pathways [36, 55].

Upon binding to CpG-DNA in the endosomal compartment [69], TLR9 recruits TIR-domain-adaptors such as the myeloid differentiation primary-response protein (MyD) 88 at its cytoplasmic TIR domain [61, 70, 71]. Through the MyD88-dependent signalling pathway [71, 72], the MAPKK kinase (MAPKKK) TGF-β-activated kinase (TAK) 1 and interferon regulatory factors (IRFs) 7 are phosphorylated and activated [73–76]. TAK1-mediated phosphorylation of IκB and MAPK kinase (MAPKK) 6, in turn, activate NF-κB and MAPKs respectively [62, 70]. The MAPK signal cascade activates activator protein (AP) 1. Transcription factors NF-κB, AP1, and IRF7 mediate the inflammatory response against CpG motifs through upregulating expression of pro-inflammatory genes [36, 77]. Overall, the immune response against CpG-DNA is characterized by stimulation of immune cells, including B cells, dendritic cells (DCs), natural killer (NK) cells, and Th cells [36, 78]; and production of cytokines, such as IL-6, IL-12, tumour necrosis factor α (TNF-α), interferon (IFN)-α, IFN-β, and IFN-γ.

CpG-DNA stimulates B cell proliferation [79] and increases the expression of cell surface receptors such as the major histocompatibility complex (MHC) class II. B cells were also demonstrated to have increased production of IL-6, IL-12, and IgM [80]. DC differentiation and maturation is also stimulated, as evidenced by increased cell surface expression of MHC II and expression of a cluster of differentiation (CD) associated with mature DCs, CD83 [81]. Specifically in pDCs, activated IRF7 induces expression of type 1 IFNs [82]. CpG motifs are thought to indirectly activate NK and Th1 release of release of IFN-γ through cytokine production from other cells [36]. As a result of CpG immunostimulation, transgene expression may be silenced, prohibiting long-term expression in DNA gene therapies [83, 84]. Removal or separation of the bacterial backbone from the transgene has been shown to avoid transgene silencing [58, 85, 86]. In some contexts, however, CpG-immunostimulation may be useful and even desirable. For example, CpG oligodeoxyribonucleotides have shown great promise as adjuvants in vaccine applications [87, 88].

Overall, the conventional plasmid imparts low gene expression efficiency due to its prokaryotic backbone and has the potential for insertional mutagenesis due to its closed circular structure. Alternatives to the conventional plasmid structure should and have been investigated.

1.1.4 DNA minivectors

Gene transfer with vectors devoid of the bacterial backbone, or DNA "minivectors" [50, 89], presents a better alternative to plasmids for non-viral transgene delivery. Minicircle vectors are produced through directed excision of a eukaryotic transcription unit from a parent plasmid using site-specific recombinases [90, 91]. This produces a minimized CCC dsDNA molecule containing only the gene expression cassette. Early reports of minicircle production were generated using λ integrase [92, 93]. Since then, minicircles have also been produced with P1 Cre recombinase [94], ParA resolvase [95, 96], or ϕC31-integrase [97, 98]. To generate minicircles, a gene expression cassette flanked by recombinase target sites is encoded on a parent plasmid carrying bacterial amplification elements. Induction of recombination activity at the target sites produces multiple

species: minicircles, miniplasmids containing backbone elements, minivector intermediates, and residual parent plasmid. A purification step is generally required to separate minicircle DNA from parent plasmid and other byproducts.

Ministring vectors are similarly produced through recombination, but generate dsDNA linear covalently closed (LCC) vectors instead of circular [99]. Linear minivectors such as the ministring may avoid the risk of insertional mutagenesis. Integration of a linear vector into a chromosome results in a double-stranded DNA break (DSB) that activates cell death pathways in both prokaryotic [99] and eukaryotic systems [100], thus removing cells with integrated plasmids from the proliferating cell population. Nafissi et al. (2014) demonstrated much reduced integration frequencies with LCC minivectors in comparison to CCC vectors [100]. Another LCC minivector, closed-ended linear duplex (CELiD) DNA, incorporates inverted terminal repeat (ITR) elements from AAV and can therefore be replicated using the AAV Rep protein in insect cells [101]. A number of other notable linear gene delivery minivectors have also been produced through in vitro methods, such as a minimalistic immunologically defined gene expression (MIDGE) vector [102], micro-linear vector (MiLV) [103], and DoggyBone [104].

While minivector technology is very promising, its unique nature has led to some bottlenecks in its manufacture [91]. Choice of recombinase has had great impact on minivector production. For minicircles, ParA resolvase [95] has shown the highest recombination efficiency (99.5%), while λ integrase [92] and P1 Cre recombinase have the lowest (60%) [105]. For ministrings, Tel protelomerase may have greater recombination efficiency than TelN [99]. Removal of contaminating plasmid and miniplasmid byproducts from the minivector product has also been challenging as they all share the same structural and chemical properties. The inclusion of endonuclease target sites on the plasmid backbone for digestion of plasmid and miniplasmid products in vivo prior to extraction greatly improved minicircle purity [98, 106]. Affinity chromatography has been investigated as an attractive option for purification of minicircles and ministrings due to its high efficiency, cost-efficiency, and overall simplicity of usage [91, 107, 108].

Minivectors overall confer greater, more durable gene expression than their isogenic plasmid counterparts [86, 97, 100, 109]. Elimination of bacterial elements like the antibiotic resistance marker, the *ori* and CpG motifs improves safety and prevents immunostimulatory transgene silencing [97]. The smaller size afforded by losing the backbone (≥ 2 kb of DNA [110]) also improves transfection and general utility. Shorter DNA can also better resist shear forces for better synergy with many physical methods of gene delivery. For example, minivectors less than 1200 bp were better able to survive nebulization for aerosolized gene delivery [111]. More minivector molecules are present per unit mass than plasmid molecules, increasing bioavailability and permitting high dosing [110]. Their shorter length is also correlated with greater vector mobility [38]; short DNA can more easily diffuse into the perinuclear space, increasing nuclear uptake. For already lengthy transgene cassettes, minivector delivery may be necessary as the additional bulk of a prokaryotic backbone could prohibit its nuclear translocation and subsequent gene expression.

1.2 Bacteriophage for gene delivery

Safety concerns exist with viral vectors, as do efficacy concerns with non-viral vectors. As might be expected, there does not yet exist a "perfect" gene delivery vector. Therefore, explorations into alternative vehicles for gene transfer are important as they may lead to greater innovation within the gene delivery space. Existing biological systems are a reservoir of untapped gene transfer mechanisms [112]. From mammalian systems, modified whole cells or exosomes have been investigated for their capacity for gene transfer [113]. Prokaryotic systems also hold much promise; microorganisms within the commensal microbiota already symbiotically coexist within humans. "Bactofection", the use of bacteria for transgene delivery, has also been successfully demonstrated with a number of bacterial strains [112].

Of particular interest from the prokaryotic world is the bacteriophage (phage), a virus that specifically infects bacteria [114]. The most abundant entity on Earth, bacteriophages are prevalent across the globe and within the human microbiome, where they number at least 10^{12} viral particles per gram feces [115, 116]. Without any intrinsic tropism for eukaryotic cells and an inability to replicate within them, the phage is considered a non-viral gene delivery vector despite being a virus itself. In fact, a phage particle may be considered as a simplistic proteinaceous coat around a nucleic acid core when applied to the human body: a promising foundation for a gene transfer vector. As they are relatively easy to genetically modify, they present an appealing option for non-viral gene delivery.

Several key characteristics further increase the appeal of phages as biotechnology products. Phage are generally inexpensive to manufacture in large quantities as they replicate in bacterial culture, which has been highly amenable to scale-up. They exhibit good stability in the face of a number of environmental factors, including pH and temperature [117–120]. As they are amenable to desiccation, phage vectors may also be inexpensive to store long-term [121]. Taken together, the genetic flexibility, robust stability, and amenability to scale-up make the bacteriophage a promising candidate for gene transfer applications.

1.2.1 Filamentous bacteriophage

The filamentous phages (*Inoviridae*) are long non-lytic viruses that primarily infect Gram-negative bacteria [122]. These phages do not lyse their hosts; after infection, progeny continuously bud out from the infected host over the course of the infected bacterium's lifetime. Although they are not lysed, the burden of phage production slows the growth of infected *E. coli* cells in comparison to uninfected cells; hence, parasitized *E. coli* can be visually identified on a lawn of infected bacteria as turbid plaques, zones of reduced growth. The most well-characterized of the filamentous phages are the F-specific filamentous phages (Ff phages), three almost identical filamentous phages that specifically infect *E. coli*: f1, fd, and M13. First discovered in sewage samples in the 1960s [123–125], studies on the Ff phages have since contributed to numerous advances in molecular biology [126].

These phages are the basis of common modern cloning vectors [127–130] and of phage display, which is the decoration of phage capsid proteins with foreign peptides [131]. The focus in this thesis will be phage M13, but the three Ff phages share 98% sequence identity and have been studied interchangeably [122, 132, 133].

Phage M13 is comprised of a circular closed single-stranded (ss) genome (6407 nucleotides [134]) encoding 11 gene products, 5 of which are capsid proteins (Figure 1.4). The genes are numbered I-XI and denoted by the prefix "g" (ex: gIII), while their products are denoted by the prefix "p" (ex: pIII). Of note, gX is an in-frame coding region contained entirely within gII. Similarly, gXI is an in-frame coding region contained entirely in gI.
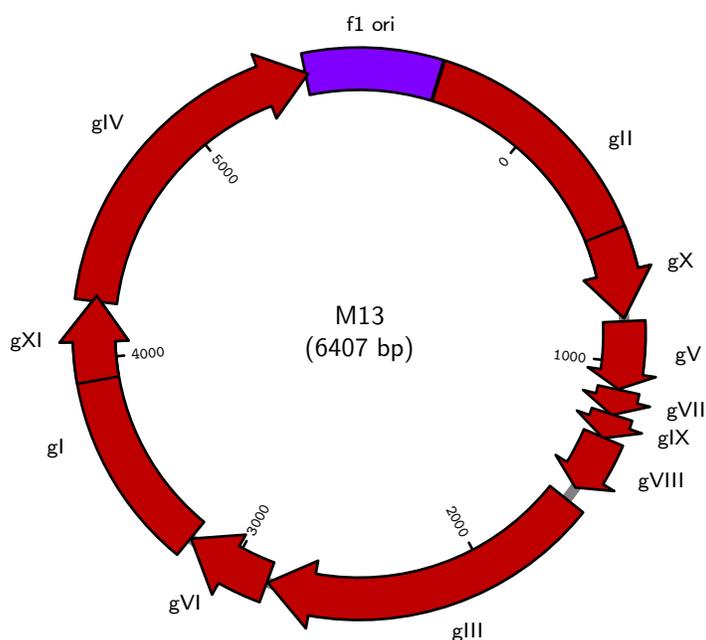
Figure 1.4: Genome map of bacteriophage M13. Bacteriophage M13 encodes 11 genes on its genome (6407 nt). Genes gX and gIX are contained entirely within and in-frame of gII and gI, respectively. The non-coding intergenic region between gIV and gII contains signalling structures for phage replication and packaging.

Proteins pII and pX are involved in the amplification of the replicative factor (RF), the circular dsDNA episomal form of the phage genome. The replicative protein pII (410 a.a.) is necessary and sufficient for rolling circle amplification of the filamentous phage genome [135]. Identical to the last 111 residues of pII, pX (111 a.a.) also regulates replication. The DNA binding protein pV (87 a.a.) sequesters amplified single-stranded viral genome molecules for phage assembly [136]. Proteins pI (348 a.a.), pXI (108 a.a.) and pIV (405 a.a.) mediate assembly and export of progeny phage, along with the capsid proteins (pIII, pVI, pVII, pVIII, and pIX) that coat the extruding phage genome as it exits the cell [137].

Structure of bacteriophage M13

The M13 phage particle is long and rod-like in shape (Figure 1.5). Approximately 2700 copies of the major coat protein pVIII (50 a.a.) wrap around the ssDNA core to form the vast majority of the tubular protein coat [138]. Each pVIII unit is comprised of a hydrophilic N-terminal domain oriented externally in the virion particle and a positively charged C-terminus facing inwards in the virion particle to interact with the packaged phage DNA [132, 139, 140]. These domains are connected by a hydrophobic connecting domain. Copies of pVIII are arranged in a right-handed helix held together by interactions between the hydrophobic domains of overlapping copies of pVIII [141–143]. At one end of the tube is a blunt cap comprised of 3–5 copies each of minor coat proteins pVII (32 a.a.) and pIX (32 a.a.) [144]. Both pVII and pIX are quite small and hydrophobic [145] with charged surface-exposed N-termini. The other end is a protruding cap of 5 copies each of pIII (406 a.a.) and pVI (112 a.a.) [144]. The pIII protein mediates infectivity, while pVI stabilizes and anchors pIII to pVIII and the rest of the protein coat [146, 147]. Comprised of three functional domains, pIII is anchored to pVI via a C-terminal domain (CT), while two N-terminal domains (N1 and N2) mediate attach to host cell receptors.

The filament-like phage particle measures approximately 7 nm in diameter while its length is essentially dictated by the size of its genome; for wildtype M13, this is approximately 900 nm [148, 149]. Virion length is simply proportional to genome length so no theoretical limit exists on the amount of DNA an M13 phage particle can package. In practice, the physical limits encompass a large range: from microphage variants about 50 nm long carrying as little as 221 nt of DNA [150, 151] to larger highly unstable "polyphage" particles almost 20 μm in length packaging approximately ten phage genome copies or over 60 knt of DNA [152].
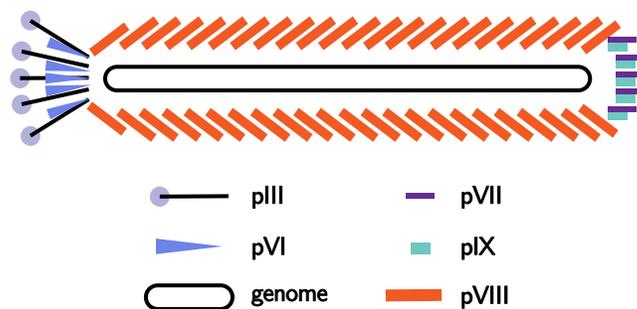
Figure 1.5: Structure of bacteriophage M13. The circular closed single-stranded genome (6407 nt) is surrounded by approximately 2000 copies of pVIII. At one end, pVII and pIX pentamers form a blunt cap. At the other end, a protruding cap is formed by pIII and pVI pentamers.

1.2.2   Lifecycle of bacteriophage M13

Adsorption and entry.   The Ff phages primarily infect F$^+$ *E. coli*, *E. coli* that carry the fertility factor (F-factor). This is an episome that encodes genes for conjugation, which is a form of horizontal gene transfer through a pilus. The pilus is a tubular protein structure that directly connects two bacterial cells through which transfer of genetic material can occur [153]. M13 infection begins with phage attachment to an extended F-pilus [154], specifically mediated by N2 domain binding to the pilus tip [155, 156]. Upon pilus retraction (Figure 1.6), phage particles are brought into close contact with the host cell surface [154]. Binding of N2 to the pilus activates an unfolding reaction, exposing the N1 domain in pIII which contains binding sites for the cell surface receptor complex, TolA-C [157, 158]. Initial attachment to the pilus is not absolutely required for successful infection: M13 phage uptake has been demonstrated in F$^-$ bacteria in the presence of Ca^{2+} ions or polyethylene glycol (PEG), albeit at a lower frequency [159]. On the other hand, Tol proteins are necessary: *tolQ*, *tolR*, or *tolA* mutants are resistant to filamentous phage infection [160, 161]. Conversely, a mutant lacking its pIII pilus-specific N2 domain can actually infect F$^-$ *E. coli* better than wild-type [158]; this is attributed to the exposure of N1 in pIII, which is normally only released after pilus-N2 binding.

The Tol proteins (TolA, TolQ, TolR) are membrane proteins that associate with a number of other membrane proteins, including Pal and OmpF [147, 162], to mediate cell functions including membrane assembly [163] and macromolecule transport [164, 165]. These proteins are also implicated in the uptake of group A colicins, plasmid-borne toxins produced by other competing bacteria [166]. The Tol proteins localize within adhesion zones, which are contact sites between the inner and outer membranes where they are physically closer. Tol proteins thereby bridge the inner and outer membranes across the periplasmic space and the peptidoglycan [167]. The N1 domain of pIII interacts specifically with TolA [132]. TolA (421 a.a.) is comprised of three domains: TolAI: the N-terminal domain that anchors in the inner membrane and interacts with TolQ and TolR; TolAII: a domain spanning the periplasm; and TolAIII: a C-terminal domain that can interact with the outer membrane [162, 165].

As the phage nears the cell surface, the N1 domain of pIII binds TolAIII. Coat disassembly begins as the phage breaches the cell envelope. The pIII and pVIII coat proteins from the infecting phage particle have been observed to insert into the inner membrane [168–170]. In studies with biotinylated M13 phage, virions were visualized penetrating the cell in two steps. First they entered the outer membrane and migrated across the periplasmic space, an event characterized by the swelling of the periplasm [171]. Next, virions moved past the inner membrane as coat proteins were deposited. Although the precise mechanism of how the infecting phage deposits its genome in the cytoplasm remains unknown, it is believed to function similarly to Tol-mediated uptake of colicins [172]. TolR and TolQ function together as a motor, using the proton motive force to drive TolA-mediated active transport across the cell envelope [163, 173–175]. The process from initial pilus attachment to cell penetration can occur within 20 s at 30 °C [176].
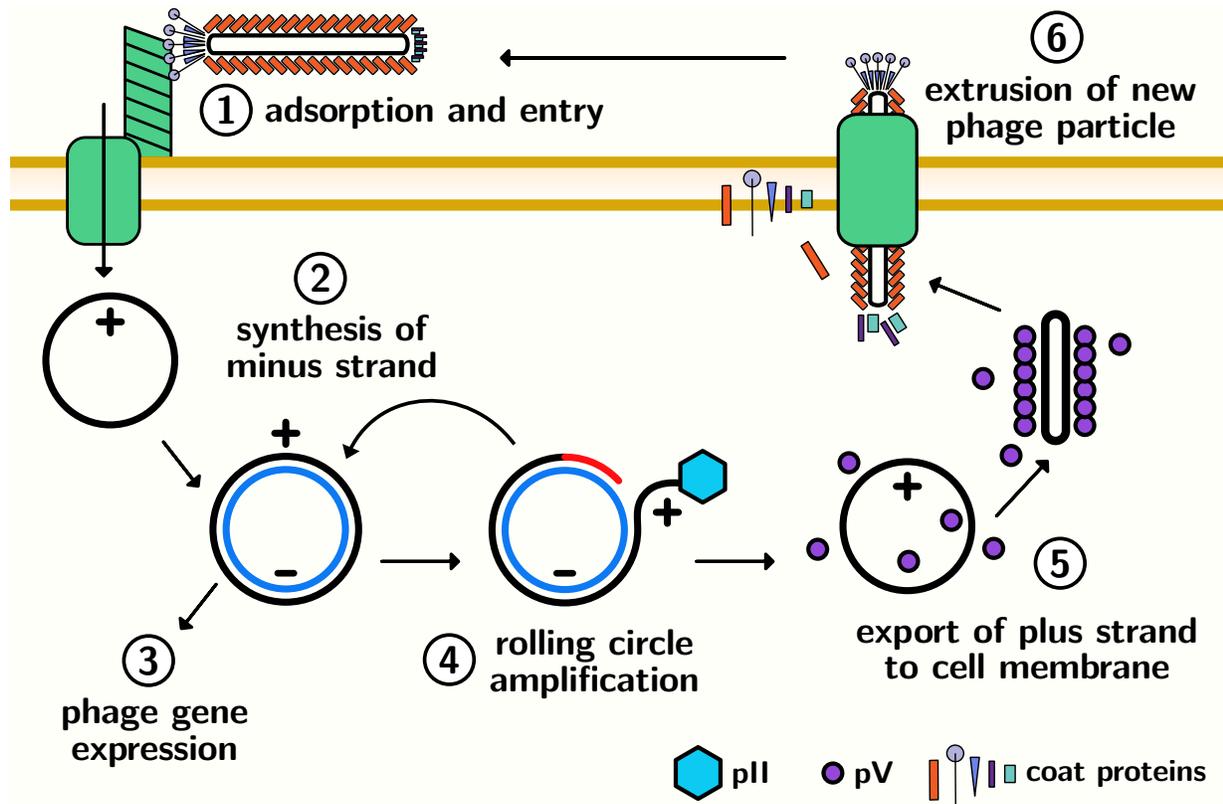
Figure 1.6: Lifecycle of the filamentous phage M13. The lifecycle of phage M13 can be broadly broken down into six steps: 1) M13 binds to the extended pilus via its minor coat protein, pIII. Retraction of the pilus brings the phage in contact with Tol proteins, through which it can enter the cell. 2) Upon cell entry, the plus ssDNA genome is converted into a dsDNA replicative form (RF) as host machinery synthesize the complementary minus strand using it as a template. 3) Phage genes are expressed. 4) The RF is replicated through a rolling circle mechanism of amplification, initiated by binding of phage protein pII. 5) Once pV levels increase and outcompete replication proteins, they bind a free viral plus strand and prepare it for export. 6) The phage particle is extruded as coat proteins assemble around the exported DNA molecule at the membrane.

Minus strand synthesis.   Before phage replication can proceed, the single-stranded viral plus strand must be converted to a double-stranded episome, the RF (Figure 1.6). The complementary minus strand is synthesized by host proteins using the plus strand as a template [177]. Both minus strand synthesis and phage replication are initiated within the intergenic region (IR), located between positions 5769-5909 in the M13 plus strand (Figure 1.4), according to the sequence first described by Beck and Zink [134]. Also known as the f1 functional origin (f1 *ori*), it contains several stretches of palindromic sequences that form secondary structures recognized by both phage and host replicative machinery. The structure of the f1 *ori* has been extensively characterized and five hairpin loop structures have been identified [178, 179] as *cis*-acting sites for phage protein function (Figure 1.7): A) immediately downstream of gIV is the packaging signal (PS) that mediates phage assembly and export, B-C) two hairpins flank the minus strand origin where host polymerases begin minus strand synthesis, D) another hairpin contains the pII nicking site as well as signals for replication termination, and E) immediately after is another hairpin containing pII signals for replication initiation.
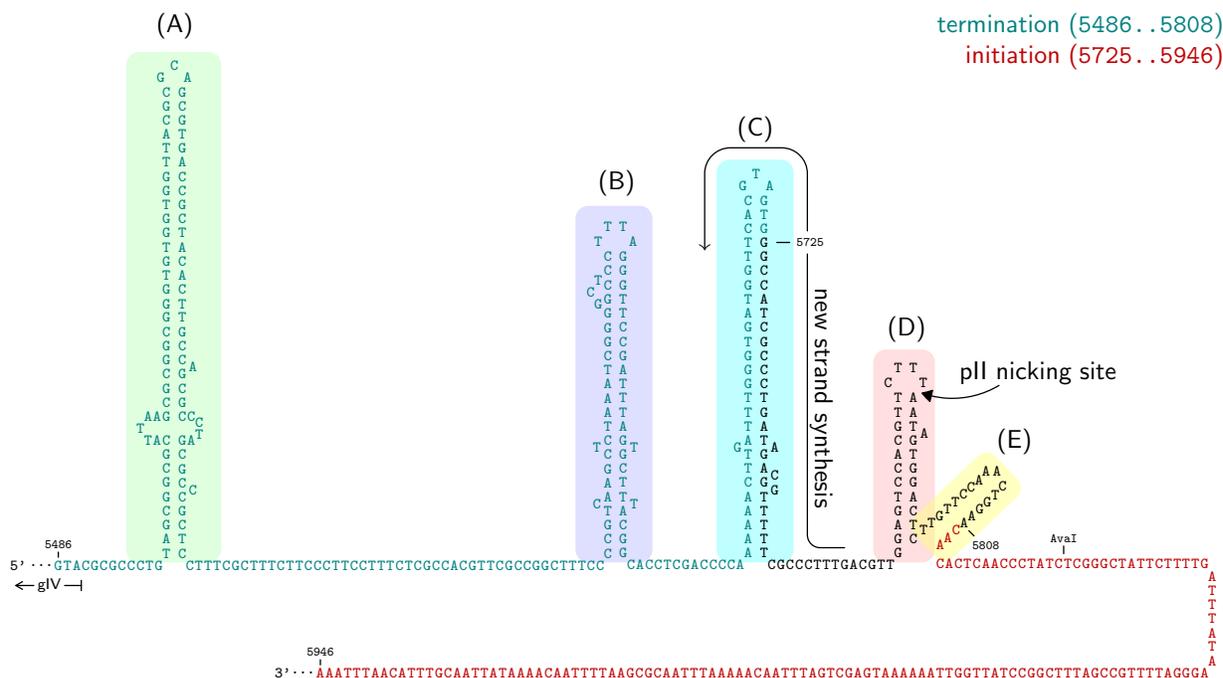
Figure 1.7: Secondary structure of the f1 functional origin on the viral plus strand. Hairpin structures are formed by palindromic sequences, with several of particular importance. From 5′ to 3′, beginning immediately downstream of gIV: A) packaging signal, B) binding site for RNA polymerase, C) binding site of RNA primer synthesis, D) location of pII nicking site, E) binding site for pII. As indicated by the black arrow, pII nicks at a site within the overlap of the termination and initiation domains to initiate new strand synthesis, which proceeds as indicated by the hook arrow. Sequence and secondary structure domains from [178–180].

To synthesize the minus strand, *E. coli* RNA polymerase binds to the minus strand origin in the f1 *ori* (Figure 1.7C). It generates a 20 nt RNA primer complementary to a stretch of nucleotides (positions 5736 to 5717 on the plus strand) of the minus strand [179, 181]. Extension from this

primer by *E. coli* DNA polymerase III results in a relaxed, closed double-stranded molecule (RF type IV) that is acted upon by DNA gyrase to generate a supercoiled RF molecule (RF type I) [182]. After establishment of the supercoiled dsDNA episome, the newly synthesized minus strand of the RF serves as a template for phage gene expression (Figure 1.6). Expression of the replication protein pII is required before further phage replication can occur.

Expression of phage M13 genes. Mapping of the phage genome has revealed at least eight promoter sites [183, 184] and approximately three terminator sites [185–187], two of which are dependent on the transcriptional terminator Rho. Activity of six promoters (preceding genes II, III, IV, V, X, and VIII) have been positively confirmed in vivo while two promoters (preceding gI and VI) have been identified in vitro, although not clearly demonstrated in vivo [184].

The M13 genome can be divided into two separate transcriptional domains that exhibit vastly differing levels of transcription activity [186, 188]. These two domains are separated by the IR on one side and a Rho-independent transcription terminator immediately downstream of gVIII (Figure 1.4), known as the "central terminator" [185], on the other side. Eleven polycistronic messenger RNA transcripts are generated between the two domains; they range in size between 370–2000 nt and in half-life between $\leq$2–10 min [187–190]. No transcript spans across either the central terminator nor the IR [186]. The host endoribonuclease RNase E is implicated in post-transcriptional processing of these phage transcripts [189, 191].

Within the highly transcribed region beginning downstream of the IR are gII, gX, gV, gVII, and gVIII [187]. Of the six transcripts arising from this domain, nearly all encode gVIII, and the majority also encode gV [186]. Furthermore, transcripts containing gV and gVIII have been observed to be more stable than all other transcripts [189, 192]. As might be expected, these genes encode the proteins required in the greatest abundance for phage production. In contrast, only one transcript encodes the entirety of gII [189, 190], making it the least abundant of the highly transcribed transcripts [186, 189].

The second less active domain starts immediately downstream of the central terminator and ends at the IR [187]. It encompasses the other five genes: gIII, gVI, gI, gXI, and gIV. Both gIII and gVI are always on the same transcript, which follows as their gene products act in a stoichiometric fashion. The inclusion of gI and gXI in a gIII – gXI transcript mostly appears to be a result of RNA polymerase read-through of a weak Rho-dependent terminator signal within gI [187, 193]. A separate promoter preceding gIV independently drives its expression [187], although a rare longer transcript encoding all genes in this domain (gIII–gIV) has been observed [186, 187]. Transcription termination of this second domain appears to be mediated by a Rho-dependent terminator within the IR, although RNA polymerase can read through this terminator. Rho-dependent termination generally produces RNA transcripts with 3′ ends within the PS (Figure 1.7A) of the IR in vivo [194]. However, transcripts terminating at hairpins C, D, or E have been observed (Figure 1.7). It is believed that the strong secondary structure within the f1 *ori* most likely interferes with RNA

polymerase processivity, leading to transcription cessation independent of Rho. Overall, differential expression of the 11 viral genes arises from a complex interplay of disparate transcriptional patterns of otherwise very simple transcription products.

During gene expression, both pIII and pVIII are actually expressed as precursor molecules with N-terminal signal peptides that are cleaved upon translocation and insertion into the membrane [169, 195–197]. In fact, all the coat proteins and assembly proteins localize to the inner and outer membranes after gene expression [168, 198, 199]. Localization of these morphogenetic proteins to the membranes is dependent on the host *E. coli* Sec translocation system [200], whereas translocation of pVIII and pIII occur in a SecA-independent manner and instead depend on YidC proteins. Although none of pVI, pVII, or pIX have a signal sequence, these minor coat proteins are small and hydrophobic and it is likely their long hydrophobic domains are what enable their membrane insertion [145]. The accumulation of phage coat proteins – particularly the numerous copies of major coat protein pVIII – in the infected cell membrane alters its phospholipid composition and reduces its capacity for pili formation [195].

Replication. M13 replication (Figure 1.6D) proceeds through rolling circle amplification of the RF to generate more double-stranded RF. This process is initiated by phage protein pII and carried out by host replicative machinery. The f1 *ori* is both necessary and sufficient for f1 phage replication as it contains the binding site for pII.

A nicking-closing enzyme with topoisomerase I-like activity [178, 182], pII initiates replication in two steps specifically at the f1 *ori* of the plus strand of a supercoiled RFI substrate [178]. First, pII binds to a core sequence approximately 15 bp long just downstream of its nicking site in the RF (Figure 1.7D-E) [201]. The *E. coli* integration host factor (IHF) binds to AT-rich replication enhancer domains approximately 50 and 100 nucleotides downstream [202]. Although not necessary, phage replication proceeds poorly in the absence of IHF. Interestingly, compensatory mutations in pII can overcome IHF-dependency. It is hypothesized that IHF activity within the f1 *ori* forms supportive secondary structures, analogous to its role in the replication of another *E. coli* phage, bacteriophage λ [202]. Binding of pII precedes RF DNA duplex melting, facilitating pII-nicking of a liberated single-stranded viral plus strand [203].

Next, replication is initiated when pII cleaves between the thymine and adenine of its nicking site [182] within the signalling hairpin loops [178]. The supercoiled RFI is therefore converted into a relaxed, nicked circular intermediate (RF type II) [204]. After nicking, pII remains covalently bound to the released 5′ hydroxyl end of the nick [205]. The released 3′-hydroxyl end of the released plus strand is elongated by host DNA polymerase III and accessory proteins to synthesize a new plus strand using the minus strand as a template [206]. Replication proceeds through a rolling circle mechanism. Elongation by DNA polymerase III proceeds unidirectionally as host Rep helicase and DNA binding protein I displace the original plus strand [204, 206–208]. Each round of replication terminates with pII nicking-closing activity again, in this case, re-circularizing

17

the displaced single plus strand of DNA and releasing it (Figure 1.7D). The displaced plus strand is again converted to RF by host polymerases, while DNA gyrase supercoils the newly synthesized, relaxed CCC RFIV into RFI [182].

Replication of phage DNA builds a pool of RFs, which can be detected as early as five minutes after infection [209]. As phage expression proceeds, levels of the DNA binding protein pV accumulate in the cytoplasm, where they compete with replication proteins to bind available ssDNA (Figure 1.6). Regulation of phage DNA synthesis also appears to be partially mediated by pX. Overexpression of pX leads to the complete cessation of phage DNA synthesis [210], while overexpression of pII leads to conversion of all phage ssDNA to RF dsDNA. The DNA binding protein pV also regulates phage replication; it has been observed that pV represses translation of pII and pX [211, 212] and, at high concentrations, inhibits DNA polymerase I and II [213].

Newly synthesized plus strands are sequestered for phage progeny export instead of further replication approximately 11 min after the start of infection, evidenced by both a detectable pool of cytoplasmic ssDNA and viable phage progeny at this point [209]. Levels of pV at this point reach approximately 500 molecules per cell, and continue to increase to approximately 2×10^5 molecules roughly 20 min after infection [212]. It is the ratio of pV to DNA in the cytoplasm that determines whether replicative intermediates will produce further double- or single-stranded DNA [214]. The pV phage protein is absolutely required for synthesis of progeny phage [215]. In the absence of functional pV, small dsDNA molecules approximately a quarter of the size of the wildtype f1 genome accumulate in the cell. The production of this species is most likely due to rising levels of single-stranded intermediates not stabilized by pV binding [216].

The binding protein pV [217] functions as a homodimer [217, 218] and was long thought to non-specifically and cooperatively bind to available ssDNA nucleotides [214, 219, 220]. It binds at a ratio of approximately 4 nucleotides per protein monomer [221], or approximately 8 pV dimers per helical turn of bound DNA [221, 222]. Later studies demonstrated pV does exhibit some selective binding affinity; for example, it binds better to deoxyribonucleotides over ribonucleotides [223]. Dimers assemble onto DNA cooperatively, mediated through inter-dimer electrostatic interactions [224]. The pV dimer also exhibits sequence specificity: the repression of phage replication by pV occurs through pV sequence-specific binding of a motif present in the gII mRNA leader sequence, thereby preventing its translation [225, 226]. This 16-mer sequence (5′-GUUUUUGGGGCUUUUC-3′) [227, 228] is bound by pV both as DNA or RNA [225, 226]. Oliver et al. (1999, 2000) discovered that this sequence can form a tetraplex structure, where single-stranded regions emerge from central guanine quartet [225, 226]. Two pV dimers are thought to bind to two of the DNA strands emerging from one side of the G-quartet, with two dimers symmetrically binding to the other two strands on the other side. Such structures may serve as nucleation sites for pV-DNA binding, whereupon subsequent pV binding continues through cooperative dimer-dimer interactions. Similar stem-loop structures are present in the mRNA leader sequences of other genes also potentially translationally regulated by pV [228] (gI, gII, gIII, gV, and gX). More recently, other 58-mer DNA

motifs were also found to be preferentially bound by pV [228]: these form stem-loop structures with large loops of unpaired nucleotides "enclosed" by one or more hairpins. Like the G-quartet motif, this stem-loop structure is thought to be the nucleation site for initial pV-DNA binding, before cooperative assembly of pV units saturate the molecule [228]. Similar structures were prevalent throughout the fd genome [228], further supporting the nucleation hypothesis.

Upon pV binding, the viral genome collapses into a rod-like form, leaving the PS (Figure 1.7A) in the f1 *ori* exposed [229]. The nucleoprotein complex forms a left-handed superhelical molecule with a pitch of 1.9 Å [221]. The absence of pV binding at the f1 *ori* may result from weaker affinity for the duplex secondary structures present in the region [122, 223]. This newly displaced plus strand is now sequestered and not converted to an RF for further replication.

Assembly and export. Coat proteins are localized to the membrane prior to assembly [196, 199, 230, 231]. In addition to the capsid proteins localized in the inner membrane, three morphogenetic phage proteins that participate in the assembly and export of filamentous phage progeny are also localized here: pI and pIX are localized in the inner membrane, and pIV, a secretin, is localized in the outer membrane (Figure 1.6) [232]. Together, pI, pXI, and pIV form a transmembrane channel, through which phage progeny DNA is ultimately exported [233].

Phage assembly is dependent on the presence of a functional PS, which is recognized by pI [234, 235]. In the absence of this *cis*-acting signal in the phage genome, assembly proceeds at approximately 0.1% the normal rate [234] although mutations in gI can compensate for a missing PS. A key characteristic of the PS is the hairpin loop, a double-stranded DNA structure left uncoated by pV and therefore available for assembly machinery to bind [223, 235]. Part of the PS is conserved between Ff phage and a more distant filamentous phage, Ike [234]: this palindromic sequence can form a tight hairpin, and might also be bound in a sequence-specific manner.

Through its large N-terminal cytoplasmic domain, pI binds the PS signal of the phage nucleoprotein complex [236]. At the inner membrane, pI complexes with pXI, which itself comprised of the C-terminal 108 residues of pI [200, 237]. While pI has been shown to mediate interactions between other substrates, the role of pXI appears to be more structural [238]. Both pI and pXI interact with a pIV multimeric complex via short C-terminal periplasmic domains [239, 240]. Like other secretins, pIV forms the bulk of a multimeric channel in the outer membrane [240]. Comprised of approximately fifteen identical subunits, the large aqueous channel interacts with inner membrane proteins pI and pXI through N-terminal domains that project into the periplasmic space [239, 240].

Together, the pI-pXI-pIV complex forms a preinitiation complex to phage assembly (Figure 1.8). Through their interactions in across the periplasm, the complex bridges across the inner and outer membranes [241] in the host cell as an adhesion zone. Between 150–300 such preinitiation complexes are synthesized in an infected cell [152, 241].
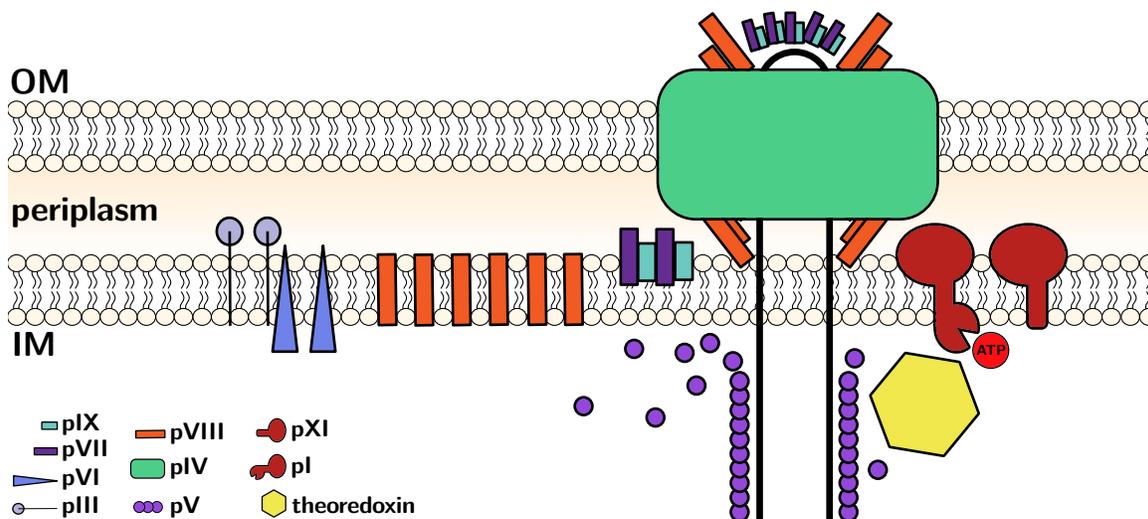
Figure 1.8: Assembly and export of progeny filamentous phage. Phage proteins pI, pXI, and pIV form a transmembrane channel across the outer and inner membranes of the host cell. Upon binding the packaging signal of the pV-coated viral plus strand, pI displaces pV proteins. The association of the pI-pXI-pIV complex with minor capsid proteins pVII and pIX and host factor theoredoxin initiates coating by pVIII units and phage extrusion out of the cell.

The cytoplasmic host factor thioredoxin is also required [242]. Rather than its canonical role as a reductase, theoredoxin appears to function as a structural component for two different *E. coli* phages: it is part of the M13 phage assembly complex, and also part of the T7 phage polymerase complex [243–245]. The association of phage DNA, thioredoxin, and minor coat proteins pVII and pIX at the preinitiation complex forms the initiation complex [239], where phage assembly begins (Figure 1.8). The host factor theoredoxin may cooperate with pI in displacing pV for pVIII during the phage assembly process, which is analogous to its role in T7 DNA polymerase processivity [244–246]. Upon binding to the phage DNA, pI mediates interactions between the DNA, pVII, and pIX, and also displaces units of pV. Phage assembly requires both ATP and the proton motive force [239, 247]; this may be mediated by pI ATPase activity [240].

Proteins pVII and pIX also bind to the PS to form the blunt end cap [199]. Since pIX has more basic residues than pVII, interactions between the capping proteins and phage DNA are thought to be mediated more by pIX; mutations in pIX are more also likely to compensate for a missing PS [138, 234]. Approximately five copies each of pVII and pIX and pVIII assemble as part of the initiation complex in the inner membrane [138]. An assembled phage extrudes from the cell surface blunt end first [248]; in the final virion, the phage genome is oriented such that the PS aligns with the pVII-pIX cap. Coating by membrane pVIII units proceeds simultaneously as the phage molecule is translocated through the pIV pore. The major coat protein undergoes major conformational change as it moves out of the membrane into the elongating phage particle [138, 249]. However, the first five copies of pVIII in complex with pVII and pIX in the membrane most likely retain their membrane conformation [138].

When the end of the DNA is reached, terminal cap proteins pIII and pVI bind last and terminate the assembly [152]. Prior to capping, pIII and pVI form a pre-termination complex in the membrane that docks to the extruding phage filament through recognition by either the assembly proteins or the incomplete pVIII tube [250]. Rakonjac et al. (1999) proposed that binding of this pre-termination complex to the extruding phage filament induces conformational changes in the pIII CT domain that releases the capped phage particle from the membrane. At this terminus, interactions between the cap proteins pIII and pVI again necessitate a different conformation for some units of pVIII [138].

If no pIII or pVI proteins are available, another unit length of DNA is added and the elongation process continues. Although the PS is required to initiate phage assembly, elongation of additional DNA does not require another PS. In a normal phage population, approximately 5% of progeny phage particles have been observed to encapsulate multiple genome molecules [152, 234]. Phage production is very proficient: up to 1000 progeny phage per cell can be released approximately an hour after infection [251].

Phage progeny cannot be produced without the major coat protein pVIII, although the coat protein itself can sustain a number of mutations without impacting coat incorporation or phage viability [139]. In contrast, progeny assembly and extrusion can carry on in spite of deficiencies in any of the minor coat proteins, albeit rather inefficiently. In the absence of pIII, release of progeny phage from the cell is very inefficient: only approximately 2.5×10^{-7} infective particles are released per cell compared to 700 per cell in the presence of pIII [152, 252]. Instead, longer noninfective polyphage particles equivalent in size to ten "normal" length phage particles are extruded and mostly remain associated with the cell membrane rather than dissociate into the medium [152]. The release of viable phage progeny into the environment is correlated to the levels of pIII in the infected cell [152]. Similarly, deficiencies in pVI [152] and in the other minor capsid proteins (pVII, or pIX) [199, 253] also produce low levels of long polyphage particles. Without pIII or pVI, capping of the completed phage particle cannot occur so extruded filaments mostly remain tethered to the cell membrane. Production of normal length phage particles is dependent on sufficient cellular concentrations of functional minor coat proteins.

1.2.3 Applications of bacteriophage M13

Since their discovery, Ff phages have contributed to numerous advances in molecular biology [254]. The simplistic biochemical and physical structure of the phage make it attractive as a nanomaterial [255], with support for chemical conjugation of functional moieties [140, 256]. Development of a number of cloning vehicles and techniques were prompted from investigations of the phage [127]. Arguably its most well-known and important contribution is its role in the development of phage display [257].

Bacteriophage M13 peptide display

As the initial model for phage display by George Smith in 1985 [131], the filamentous phage has cemented itself as an integral workhorse in molecular biology. A diverse range of peptide moieties has been displayed on filamentous phage, with innumerable applications: peptides [131], antibody fragments and/or whole antibody domains [258–260], cDNA libraries [261, 262], scaffolding proteins [263], and whole enzymes [257, 264]. Phage display libraries can also be constructed for biopanning, a screening and enrichment process for the discovery of novel peptide moieties.

While all the coat proteins of M13 have shown some capacity for peptide display [257, 264], the two most common sites of display are pIII and pVIII. M13 phage display can be homogeneous, in which every copy of the capsid protein contains the translational fusion, or heterogeneous if both wild-type and fusion capsid proteins are incorporated into the phage particle. The display peptide sequence must be incorporated in-frame into one of the coat protein genes. Smith (1993) developed a classification system for display vectors according to the arrangement of the coat protein genes [265] (Figure 1.9). As only a single type of coat protein form is used in an homogeneous display system, it can be represented using a single Arabic numeral [257, 265]: for example, Type 3 vectors display gIII::peptide fusions, while Type 8 vectors display pVIII::peptide fusions. Homogeneous phage display can be less stable if the foreign peptide is large enough to disrupt phage structure [249]. Hence, heterogeneous display is particularly important for the display of larger proteins ($\geq$100 residues).

Another common approach is the usage of a phagemid, which is a plasmid that carries an f1 *ori* for M13-mediated replication and assembly, to encode a copy of the fusion peptide for display. Separately, a helper phage then encodes all other phage genes, including the wild-type capsid proteins. For homogeneous display, the helper phage must completely lack the gene for the coat protein of choice so that only phagemid-encoded fusion coat proteins are available [252, 266]. For heterogeneous display, helper phage-derived wild-type coat proteins compete with the phagemid-derived fusion coat protein to generate mosaic progeny virion particles. Smith denotes this dual vector system as a sum to represent gene separation on different genomes [257, 265]: for example, a Type 3 + 3 system is constructed from one wild-type pIII from the helper and one fusion pIII from the phagemid (Figure 1.9). The phagemid DNA is preferentially packaged over helper phage DNA [267]. Isolated phage particles will therefore encode the gene corresponding to the displayed ligand. It is this simple yet powerful linkage of genotype (as encoded on the packaged phagemid) to phenotype (the function derived from the displayed peptide) that enables rapid screening of vast genetic libraries for ligands of interest [268–270].

Alternatively, a second copy of the fusion gene may be inserted alongside the wild-type copy [257, 258, 271]. This also leads to assembly of mosaic progeny phage, but both genes are encoded on one vector. Since a single genome bears both copies of the gene, this system is represented by two Arabic numerals indicating the gene [257, 265]. For example, a Type 33 vector encodes both a wild-type gIII and gIII::peptide fusion on one genome.
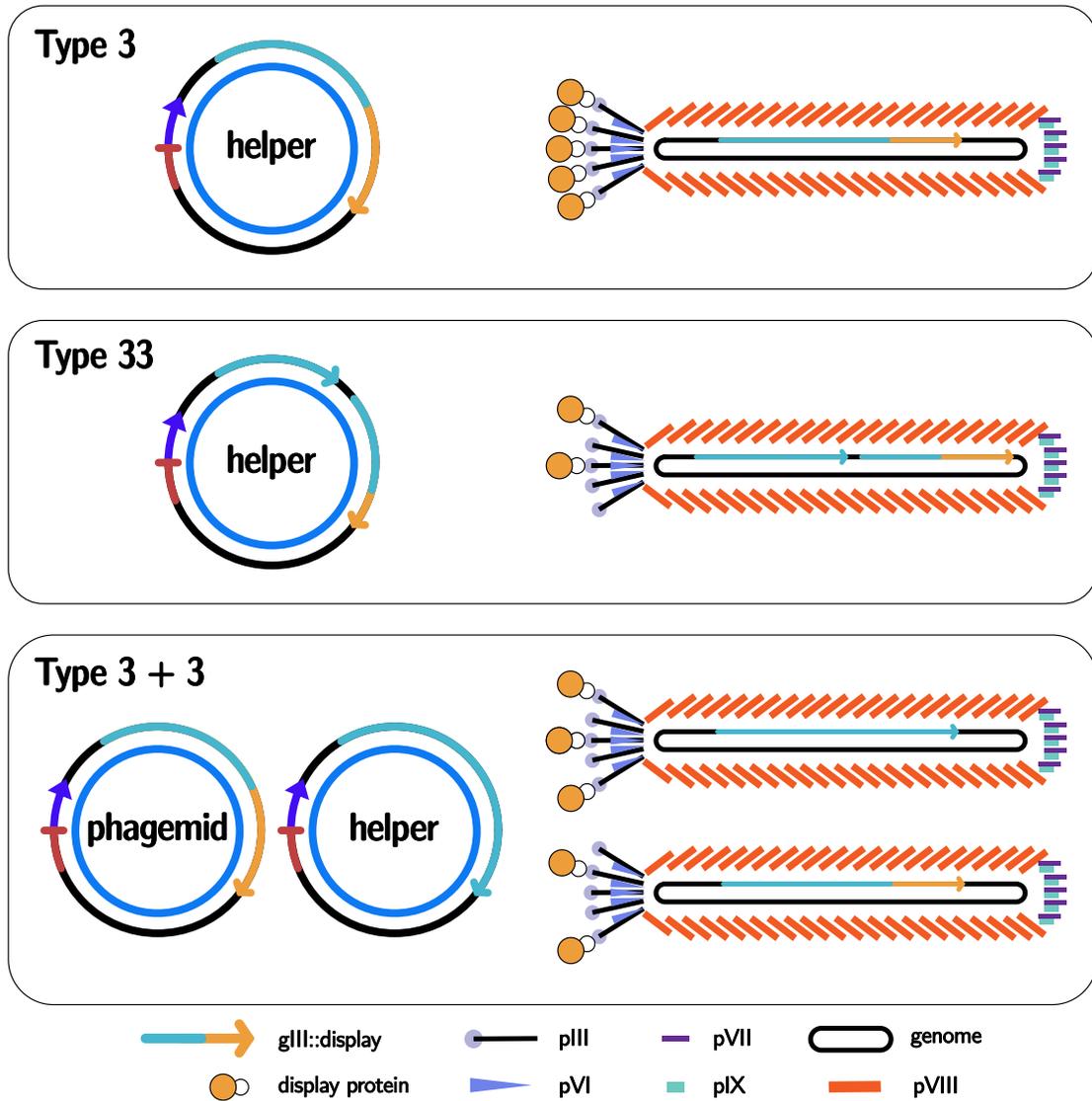
Figure 1.9: Phage display systems. This schematic shows three M13 display systems using pIII as a model. In Type 3 display, only one copy of gIII is available, leading to homogeneous display of the fusion peptide. In Type 33, two cis copies of gIII are available, leading to heterogeneous display, but encapsulation. In Type 3+3, two trans copies of gIII are available, again leading to heterogeneous display.

Display on pIII. The display of exogenous peptides through in-frame translation fusion was first demonstrated on the minor coat protein pIII in filamentous phage f1 [131]. As with all the coat proteins, pIII inserts into the inner membrane of *E. coli* with its C-terminus exposed in the cytoplasm and its N-terminus facing outwards towards the periplasm [197, 272]. Its localization to the inner membrane is mediated by the signal peptide on its N-terminus; N-terminal fusions must therefore be inserted downstream of this signal peptide. Upon cleavage of the signal peptide, the fused peptide will then be exposed to the periplasm. Display has been demonstrated in multiple domains of pIII, including at both termini [257]. The highly exposed amino terminus has substantial capacity for the incorporation of a broad range of peptides and proteins [249], but is not an ideal site for the incorporation of cytoplasmic proteins [273–275]. Display of proteins on the C-terminal end of pIII has been investigated as it is exposed in the cytoplasm [275, 276]. Interestingly, modifications to the pIII localization signal and introduction of other bacterial localization tags have enabled N-terminal fusion of otherwise incompatible proteins [274].

Overall there are only five copies of pIII in each virion and all of them share the same physical space (Figure 1.5). Type 3 display, that is, pentavalent pIII display is best suited for shorter peptide sequences [257]. However, phage assembly may not be possible with pentavalent display of larger proteins as this would lead to steric hindrance. Instead, Type 33 or Type 3+3 systems can facilitate pIII monovalent display of larger proteins [257, 265, 274, 277]. Most concerningly, inserts in pIII can interfere with the infection process [131] since pIII mediates phage adsorption to host cell receptors. However, insertions of very small peptides may not have any effect on infectivity. Hence, heterogeneous display can additionally retain phage infectivity.

Despite some of its limitations, pIII remains a very common site for display due to the ease of fusion introduction and the large existing body of pIII display vectors [249, 278]. Display on pIII has been combined with display on other coat proteins as well, including pVIII [279–282] and pIX or pVII [282–284]. Dual display enables functional complementation between two displayed moieties. Display of fibrinogen-binding or streptavidin-binding peptides on pVIII ("capture-peptide") facilitated phage capture in combination with the display of the actual library peptide of interest ("assay-peptide") on a different coat protein (pIII or pVIII) [281]. Similarly, another Type 3 + 88 system enabled pVIII-mediated immobilization or visualization of phage particles in combination with pIII display of targeting peptides [280]. Fusion of His-tags to pIX allowed immobilization of phage particles, where pIII display could bind peptides of interest [283, 284]. Løset et al. (2011) developed helper phages with one of several fusion tags (FLAG-, His-, or AviTag) fused to the N-terminus of pVII [282] for phage immobilization in tandem with with a phagemid encoding a display peptide of interest. This system therefore enables heterogeneous display on either pIII (Type 3 + 7) or pVIII (Type 8 + 7) in combination with homogeneous pVII display.

Display on pVI. As part of the same protruding cap of M13, pVI anchors pIII to the rest of the protein coat and is also present in five copies. However, display on pVI has not been as thoroughly studied as pIII [285–288]. Only C-terminal display has been explored as it is exposed to the surface, while its hydrophobic N-terminal is buried in the virion particle [138]. The exposed C-terminal end of pVI has more potential for the display of cDNA libraries as cDNA stop codons would not prematurely end pVIII expression and phage assembly [262, 287]. Jespers et al. (1995) constructed a plasmid and phage combination (Type 6 + 6) for pVI monovalent display of a cDNA library [285], but observed lower display levels than with pIII. Other groups also demonstrated that pVI could display functional proteins from cDNA libraries [287, 288]. Additionally, Jespers et al. observed that pVI may confer improved display of certain proteins with C-terminal biological activity over pIII display [285]. Thus far, pVI display remains more limited in use in comparison to the other coat proteins.

Display on pVIII. Use of the pVIII coat protein as a site for display is highly attractive [289–291] since it is present in more than 2700 copies per virion. As with pIII, display on both the N-terminus [290] and C-terminus [292, 293] have been demonstrated. Although the pVIII C-terminus is buried within the virion coat, it can tolerate some degree of fusion, albeit less efficiently than the surface-exposed N-terminus [293]. However, because of the dense packing of the small 55 a.a protein, it can only accommodate homogeneous display of sequences approximately 6–10 residues long [290]. On the other hand, heterogeneous pVIII display of even very large proteins does not appear to disrupt phage assembly so long as some wild-type pVIII proteins are available [277, 294]. It is believed that the five or so pVIII units at each end of the virion particle are less tolerant of insertions due to the differences in their conformation. These pVIII units must stabilize connections between the pVIII coat with the minor proteins; hence, they are thought to be more structurally integral than the rest of the core pVIII coat. Overall, only a small selection of short epitopes in pVIII appear to be absolutely necessary for coat assembly and phage function [290, 292, 295]. Four epitopes have been positively identified: three small hydrophobic epitopes (near both termini and the protein core) and one positively-charged epitope near the C-terminus [139, 148, 292]. The hydrophobic epitopes mediate packing of pVIII units to assemble the virion coat, while the charged epitope mediates pVIII-DNA binding. So long as these epitopes are unmodified, pVIII appears highly tolerant of large changes in sequence and structure [139, 292, 294, 296].

Expanding even further upon heterogeneous pVIII display, Malik et al. (1997) produced "double hybrid" display phage by combining three copies of gVIII in a Type 8 + 88 system: two copies in fusion with different display peptides and one wild-type [297]. This can lead to the production of highly heterogeneous phage coats, which can support a diverse range of applications.

Display on pVII and pIX. The last two minor coat proteins pVII and pIX form the blunt end cap of M13; both proteins are fairly small and hydrophobic, but have surface-exposed N-termini as potential sites for display. N-terminal display on pVII and pIX have been explored through several

systems [285, 298–302]. C-terminal display is not thought possible as these ends of both proteins are buried in the virion particle. Similarly with pIII and pVI, only up to five copies of each protein are available which can limit the applicability of the displaying virion. Display on pVII or pIX has thus far seen usage in the construction of biopanning libraries where such a limitation does not apply. Gao et al. (1999) demonstrated a combinatorial "7/9+7/9" display of antibody heavy- and light-chain variable regions respectively [303]. The close proximity of the two fusion coat proteins enabled heterodimeric formation of a functional antibody Fv region. Antibody fragment libraries could also be constructed using only pIX (Type 9 + 9) [304] or only pVII (Type 7) [300, 302] display. Both pVII and pIX tolerate fusions of proteins up to 50 kDa [302].

To date, all coat proteins of M13 have demonstrated the capacity for display of a broad, diverse range of proteins. In general, the coat proteins insert into *E. coli* inner membranes such that their N-termini face the periplasm, while their C-termini are cytoplasmic. This does place a constraint on the conformations of the fusion proteins, although a number of methods to overcome this constraint have since been demonstrated [138, 274]. For monovalent or otherwise low valency displays, minor coat proteins pIII, pIX, and pVI can all tolerate even fairly large N-terminal fusions. The display of cDNA libraries may be better suited on pVI as it can accommodate C-terminal fusions, although pIII also has this capacity. For higher valency displays, the major coat protein pVIII has exhibited great flexibility, although the displayed peptide must either be very small or otherwise the display be heterogeneous. Since its inception, phage display has played a large role in both drug and druggable target discovery [257, 278, 305], and can functionalize the displaying virion towards almost any application.

M13 cloning vectors

Single-stranded DNA production. Aside from phage display, filamentous phages have also been instrumental as cloning vehicles and have precipitated a number of inventions that remain staples of research laboratories around the world to this day [126]. As the phage genome itself is single-stranded, filamentous phage M13 was useful as workhorse for the production of ssDNA [267]. Notably, a number of commonplace cloning vectors were developed based on work with filamentous M13 [306, 307]. Messing and his colleagues incorporated multiple endonuclease recognition sites (a polylinker) within *lacZα*. The inclusion of this into the IR of M13 gave rise to cloning vectors (the M13mp vectors, including notable examples M13mp18 and M13mp19) and enabled blue-white screening for M13-mediated propagation [308].

Development of universal primers for M13mp vectors enabled their use in DNA sequencing [309, 310]. Messing pioneered the technique of shotgun sequencing [311, 312]: multiple fragments were cloned in M13mp vectors, which could then be sequenced in parallel using an M13 universal primer. These sequences were stitched later together in silico. This was demonstrated using the genome of the cauliflower mosaic virus [126]. Later, the pUC series of plasmids (including notable examples pUC18 and pUC19) were developed containing the polylinker from M13mp7 and both a high copy

origin of replication (pMB1) and an f1 *ori* [129, 307]. The inclusion of the pMB1 *ori* enabled highly efficient production of dsDNA [313]. Infection with phage then enabled the efficient production of ssDNA from the same vector through the f1 *ori*.

To improve the propagation of M13-derived vectors, Messing and his colleagues also generated a series of *E. coli* strains that are now routinely employed for the maintenance of phage and supercoiled plasmids [308]. Among them, JM109 and JM110 are recombination-deficient (*recA*), making them ideal hosts for plasmid amplification while avoiding recombination-induced mutations.

Phagemids. As the f1 *ori* is necessary and sufficient for both pII-mediated replication and pI-mediated extrusion, it can be integrated into any plasmid vector for phage-mediated replication of the vector independent of its plasmid origin; such plasmids are known as phagemids [314]. Helper phages are needed to amplify phagemids as they provide necessary phage proteins in trans.

Any of the Ff phages or their derivatives can be used as a helper phage; however, the f1 *ori* present on the phage genome naturally competes with the origin on the phagemid for replication and assembly. Hence, phagemid lysates will contain both helper phage and phagemid [315]. Enea and Zinder (1982) generated "interference-resistant" f1 phage mutants with gII mutations that enabled superior replication; derivatives of these mutants have been useful as helper phages [316]. Vieira et al. (1987) constructed M13KO7, now a common helper phage, for the production of single-stranded DNA from phagemids [267]. M13KO7 (Figure 1.10) is a derivative of phage M13 with the origin of replication from p15a and a kanamycin resistance gene from Tn903 inserted into the AvaI site (5825) of the IR. The insertion of these elements interrupts the structure of the f1 *ori*. When a phagemid vector is present, phage replicative proteins will replicate the intact phagemid f1 *ori* instead. Meanwhile, the plasmid *ori* enables replication independent of pII.
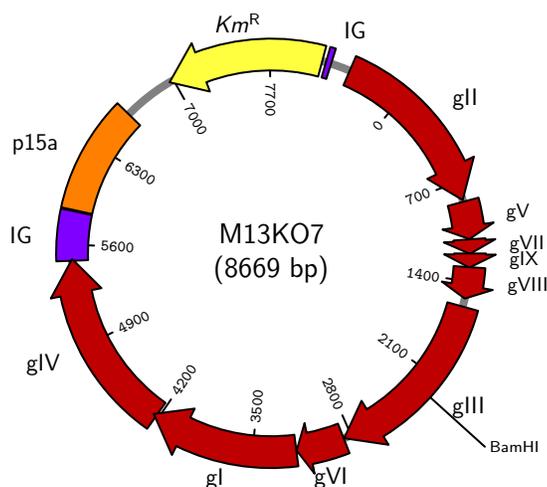
Figure 1.10: Map of helper phage M13KO7. The helper phage M13KO7 is almost identical to the wild-type M13 phage, with one notable difference. Its IR is interrupted by a plasmid origin of resistance and a kanamycin resistance marker.

Similarly, the IR in M13KE (based on M13mp19) is interrupted by the insertion of *lacZα*, again leading to deficient self-replication in the presence of a wild-type uninterrupted f1 *ori* [271]. Additionally, M13KE has two endonuclease recognition sites integrated within the N-terminal region of gIII, making it very easy to clone N-terminal gIII fusions. Other display-specific helper phages have been developed: for example, two gIII-deficient helpers, R408d3 [252], and CT [317], were developed for the propagation of pIII display libraries. As these phages are unable to contribute wild-type pIII, all progeny phage must exhibit homogeneous multivalent display of a pIII fusion library peptide.

Other efforts have resulted in the production of helper plasmids without any f1 *ori* sequences at all [318]. Helper plasmids by Chasteen et al. (2006) carried a plasmid origin of replication (p15a) similarly to M13KO7. The production of phage progeny carrying phagemid DNA was demonstrated successfully using this system while avoiding a large degree of helper plasmid self-packaging.

Interestingly, genome integration of phage genes has been attempted [318, 319] but without success. Likely, single copy gene expression is insufficient for successful phage replication. Filamentous virion production results from a complex interplay that is highly dependent on stoichiometric balance between phage products. This is perhaps best exemplified by the transition from phage amplification as dsDNA RFs to assembly and extrusion of ssDNA phage, which is governed by the accumulation of pV. Single-stranded DNA will be converted to double-stranded RF by pII if pV levels are inadequate or if pII levels are too high [210]. On the other hand, a high pV:RF ratio may halt conversion to RF such that phage replication can no longer progress and the cell is cured of M13 infection. Overall, it appears that phage copy number is paramount to good phage production. In particular the phagemid:helper ratio is dependent on phagemid DNA outcompeting the helper for assembly, but not at the expense of helper-driven phage gene expression. Hence, development of helper phage deficient in self-replication and/or self-packaging is crucial for phagemid propagation and maintenance.

1.2.4   Mammalian gene transfer by M13

M13 does not have any natural tropism for human cells and, as such, functions as proteinaceous carrier of genetic material. Display of peptides on the capsid surface enable targeting and uptake by human tissues. Although already well-established as a platform for discovery of cell-selective peptides in the 1990s, it was unclear if phages could serve as gene delivery vectors [320].

M13 biodistribution varies widely depending on the displayed organ-specific ligand; indeed, phage display libraries have been instrumental in the discovery of such targeting ligands [270, 321–324]. Investigations of untargeted versus targeted phage in mice demonstrated rapid (4 min) phage distribution in the blood compartment of tissues of administration, and tissue extravasation by 24 h [325, 326]. Clearance by the RES appears to be the major factor impacting its biodistribution. A majority of phages can be cleared by the RES within 90 min of administration; however, cell uptake and tissue extravasation could occur in as little as 30 min [326]. Shortly after administration,

phages were shown to accumulate mostly in the liver [325, 326], but clear over time. Targeted and untargeted filamentous phage are more likely found in the lungs, spleen, and kidneys. Intriguingly, M13 targeted to the brain has even demonstrated the ability to cross the blood-brain barrier [321, 327]. This capability was hypothesized to be, at least partially, due to its rod-like shape able to bypass cell barriers better than globular vectors of similar mass.

Phage uptake by cells through macropinocytosis and phagocytosis is a readily common phenomenon [328, 329] even without the display of cell-specific targeting ligands. The rate of phage uptake appears to be both tissue and phage species-dependent [328]. Phages internalized via macropinocytosis have been to shown to transcytose across epithelial layers, as demonstrated with T4 [330]. Transcytosis of a phage of *Salmonella enterica* serovar Typhimurium [331] has also been demonstrated in models in vitro.

However, the display of a cell-targeting ligand can also facilitate the phage's uptake by exploiting receptor-mediated endocytosis. Internalization of ligand-displaying phages by targeted cells was shown to occur within 30 min in mice [326]. Uptake of ligand-displaying phage is specific to the ligand-receptor interaction and confers cell-specificity to the displaying phage particle [326, 332, 333]. Non-targeted M13 were observed to internalize primarily via clathrin-coated endocytosis and macropinocytosis in epithelial cells such as HeLa in vitro, but via caveolae-mediated endocytosis in an endothelial cell line like HDMEC [334]. The phage is also internalized through the mechanism conferred by the displaying peptide. Phage displaying peptides containing the Arg-Gly-Asp (RGD) motif, for example, are targeted to cells that over-express αv-integrins and robustly internalized via clathrin-coated receptor-mediated endocytosis [289, 335]. In contrast, phage displaying cell-penetrating peptides (CPPs) such as the TAT peptide from HIV-1 could internalize via clathrin- or caveolae-mediated endocytosis [336]. The displayed peptide also determines the intracellular route internalized phage follow. For example, tracking of internalized TAT-displaying phages showed they were routed for lysosomal degradation within 2 h of administration. In contrast, display of a different CPP (3D8 VL) facilitated escape into the cytosol.

The flexibility enabled by phage display suggests that M13 — and perhaps other viruses — can be modified for targeted gene delivery into mammalian cells. As they are naturally predators of bacteria, bacteriophages do not have any intrinsic mechanisms for optimally transducing eukaryotic, much less mammalian, cells [337, 338]. Internalized filamentous phage have been tracked to the endosomal compartment where they are eventually trafficked to lysosomes [334]. In vitro experiments have shown that while 100% of cells internalized phage displaying a targeting ligand, only up to 10% expressed the transgene [339–341]. Display of a targeting ligand improved phage internalization 100-fold over non-displaying phage, but only increased gene expression threefold. This suggests other barriers still prevent optimal gene transfer by phages. Larocca et al. (1999) demonstrated that phage gene delivery improved upon the addition of chloroquine, an endosomal disrupter, which highlights this important barrier to phage-mediated gene transfer [333]. Phage escape from the endosomal compartment is therefore necessary for gene delivery.

In the case of filamentous phages, although they are readily internalized, they do have a major disadvantage compared to other viruses and gene transfer vectors: they encode single-stranded DNA. In general, dsDNA performs better for mammalian gene transfer [342]. Gene expression does not start until approximately 48 h post-transfection, after endosomal escape, nuclear import, and conversion to dsDNA. Self-complementary sequences have been introduced in M13 to improve mammalian M13 gene delivery [343]. Alternatively, phage gene delivery supplemented with DNA-damaging agents also improve gene transfer, likely through accelerated second strand synthesis initiated through DNA repair mechanisms [344].

Despite its shortcomings, however, cell-targeted M13-mediated gene delivery has been documented [333, 344–350], with transfection efficiencies up to 40%. Filamentous phage M13 presents a novel method for gene transfer with much flexibility and potential from its large capacity for capsid decoration and ease of genetic manipulation.

1.2.5 Immunity against bacteriophages

In general, bacteriophage are highly immunogenic particles foreign to animal hosts and therefore cleared rapidly in the regardless of the presence of their bacterial hosts [351, 352]. The major route of phage clearance in the body is the RES [353, 354]. B cells have also been implicated in phage clearance, but this may be specific to certain phage species [355]. More specifically, filamentous phages may be cleared by the RES within 90 min after administration [326]. The rapid mammalian removal of phage is likely mediated through the recognition of PAMP motifs in coat proteins. As such, alterations of coat proteins have led to the isolation of phage particles capable of RES evasion and prolonged systemic circulation [356, 357].

Although phage modulation of both the innate and adaptive immune responses has been well-documented [354, 358], much still remains unclear. Within the body, phages primarily reside alongside their natural hosts; as such, immunomodulatory effects from "natural" phages are primarily exerted from within the gastrointestinal tract [358]. Phages may have potent potential for immunomodulation [359] through the maintenance of gut homeostasis in the intestinal microbiome [360, 361]. Translocation of phage outside of the gut into the bloodstream ("phagemia") has been observed on numerous occasions [352, 362, 363]; but their circulation does not appear to be harmful. Within the gut, the virome exerts antimicrobial effects against host bacteria, driving competition between bacterial species and expanding ecological niches [364, 365]. Thusly, predator-prey dynamics shape and support diversity in the commensal microbiota, which in turn modulates their interactions with gut immune cells [366, 367]. It is believed that the gut virome may modulate gut immunity indirectly through this regulation of the gut microbiome [365], particularly since the virome may be dominated by nonlytic phages rather than lytic [368, 369]. However, the adherence of lytic virions to gut mucosal surfaces has been demonstrated to protect against bacterial invasion [370] and, in this way, directly contribute to gut innate and adaptive immunity [367].

Some phages like T4 are able to interact with immune cells through motifs present in their capsid proteins [360, 371]. Both anti- and pro-inflammatory effects have been attributed to phages, although these reports have been conflicting [358, 372, 373]. As phages amplify within bacterial hosts, phage preparations run a risk of bacterial debris carryover, which are highly immunostimulatory in their own right. In particular, lipopolysaccharide (LPS) has well-documented immunostimulatory effects [374, 375]; as such its presence may alter any immunomodulatory effects of bacteriophage preparations [359]. This can be an advantage in vaccine or cancer immunotherapy applications, however; for example, the use of LPS was used to enhance phage-mediated tumour destruction [376]. Conflictingly, phages of *E. coli* and *Staphylococcus aureus* have also been shown to inhibit LPS-mediated inflammation [377, 378]. Van Belleghem et al. (2019) demonstrated upregulation of both anti-inflammatory and pro-inflammatory cytokines from administration of highly purified phages both with and without LPS [373]. Another major mechanism by which phage appear to reduce inflammation is through the clearance of infecting bacteria in the host and therefore removal of the immune response against bacteria [358, 359].

Immunity against M13. Adaptive immunity against M13 is primarily mediated through pVIII, which comprises most of the phage surface, and the partially exposed pIII [351]. Antibody responses are restricted to the exposed N-terminal twelve residues of pVIII [379] and the N-terminal domains of pIII [380, 381]. Removal of pIII N-terminal regions can decrease the immunogenicity of circulating M13 because the number of available phage-specific immunodominant epitopes would be reduced [380]. Filamentous phages have been demonstrated to activate Th cells, including Th1, Th2, and Th17 [382, 383], and generate cytotoxic T lymphocyte (CTL) responses through MHC I and II presentation [384, 385].

Immune activation against M13 appears to be dependent on TLR signalling (specifically through MyD88) [376, 386, 387], leading to upregulation of pro-inflammatory cytokines such as TNF-α, IL-6, and IL-1 [376]. Murine studies have showed strong IgG antibody responses against M13, also dependent on MyD88 [388]. Indeed, induction of immune responses against displayed antigenic peptides was found to be abolished in the absence of TLR9 expression [386]. TLR9 signalling activates inflammation pathways in response to intracellular bacterial or viral DNA through the recognition of CpG motifs [65]. Activation of TLR9 signalling by M13 most likely occurs through recognition of CpG motifs present in the phage genome [387, 389]. This has been supported with the demonstration of co-localization of TLR9 receptors with M13 particles in late endosomes [387]. As such, innate immunity against M13 appears to be primarily induced after cell uptake.

M13 display of antigenic peptides has been heavily explored both to elucidate mechanisms of the immune system and for the development of novel vaccine strategies [351, 380]. The display of antigenic peptides can evoke strong adaptive immune responses without the need for adjuvanting, including T cell responses and antibody production against the displayed peptide [380, 390–394]. Specifically, the display of Th peptides derived from HIV or human cytomegalovirus (CMV) were able to elicit CD4$^+$ T cell responses [394, 395]. Display of epitopes from *Candida albicans* also

induced potent Th-mediated immune responses [382, 383, 396, 397]. Cytotoxic T cell responses are also induced through the display of immunogenic peptides on M13 [386, 395, 398]. Strong antibody responses against typically weakly antigenic peptides can be elicited using M13 as the display platform [399]. As a highly immunogenic particle able to activate both innate and adaptive immune responses, M13 can be an excellent model antigen [381].

Overall, however, phages have been repeatedly demonstrated to be safe in clinical trials despite their immunomodulatory properties [400–403]. Our understanding of the phage immune response is still an evolving process, but it is clear that phages hold immense promise as modulators of biological processes.

Chapter 2

Project hypotheses and objectives

2.1   Project rationale

Bacteriophage-mediated gene delivery holds much potential. Phage particles are simplistic protein particles that protect DNA cargo easily functionalized by phage display of cell-specific ligands to target uptake into specific tissues. Presence of the f1 functional origin (f1 *ori*) is sufficient to direct M13-mediated replication of any vector, which is referred to as a phagemid [314]. Thus, a mammalian transgene vector can be easily propagated and encapsulated into M13 virion particles. However, a number of intracellular barriers still need to be overcome for successful gene transfer by phages. It is now well-known that bacterial DNA can detrimentally reduce transgene expression both due to cytosine-guanine dinucleotide (CpG)-mediated inflammatory silencing and excess DNA bulk impacting nuclear import [38, 83]. Therefore, its presence on a phagemid vector may very well contribute to the lack of transgene expression efficiency by filamentous bacteriophages.

In the f1 *ori*, only the loops near the pII nicking site and the domain downstream of it are necessary for the initiation of replication (Figure 2.1). Only the loop containing the pII nicking site and the domain immediately upstream are necessary for the termination of replication. So long as the pII nicking site is maintained, these domains can function more or less independently of each other. Therefore, separation of the initiation and termination signals can enable replication of only DNA between the two replication signals [208], producing smaller phage variants [150].

Dotto et al. (1984) previously produced chimeric phagemids with two phage origins in the same orientation [178]. Replication began at the initiation domain of one f1 *ori* and terminated at the termination domain of the next, resulting in the amplification of the DNA between the two signals. Later, Short et al. (1988) demonstrated pII-mediated propagation of a sequence of their choice when flanked with these domains [319]. They further demonstrated that these domains were sufficient for excision and propagation of an intact phagemid from the λ genome.

```
        5487
        RsaI
5'      GTACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGC


        GCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCT


        TTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCT


        AAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCG
                                                5725
                                                Sau96I
        ACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCC
                                5767
                                HinfI
        TGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTT*AATAG

        5789
        HinfI
        TGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATT


        CTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAAT


        GAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGTT
                5947
                DraI
        TACAATTTAAA  3'
```

Figure 2.1: Sequences of the terminator and initiator domains of the f1 origin. The nicking site is indicated (*) and is found in the overlapping region. Reprinted from ref. [319] with permission.

We propose the use of these separated f1 *ori* domains to flank a mammalian transgene expression cassette, isolating it from the bacterial backbone. By doing so, we hope to direct phage replicative and assembly machinery to produce phagemid gene delivery vectors lacking bacterial backbone. In other words, we propose the construction of "miniphagemid" vectors. Another major constraint of phage M13 is its single-stranded (ss) nature [342–344], necessitating the extra step of ssDNA synthesis prior to gene expression in mammalian host cells. Single-stranded viral vectors have demonstrated excellent gene transfer efficacy, which demonstrates the potential for single-stranded phagemid vectors for gene delivery [404]. The introduction of dsDNA regions within phagemid vectors was shown to improve gene expression of filamentous DNA in mammalian cells [343]. If the entirety of a double-stranded transgene cassette could be assembled into a filamentous phage virion particle, we postulate that it could lead to vastly improved phage-delivered gene expression. These projects, as depicted in Figure 2.2, will contribute to the advancement of the filamentous phage M13 as a mammalian gene delivery vehicle.
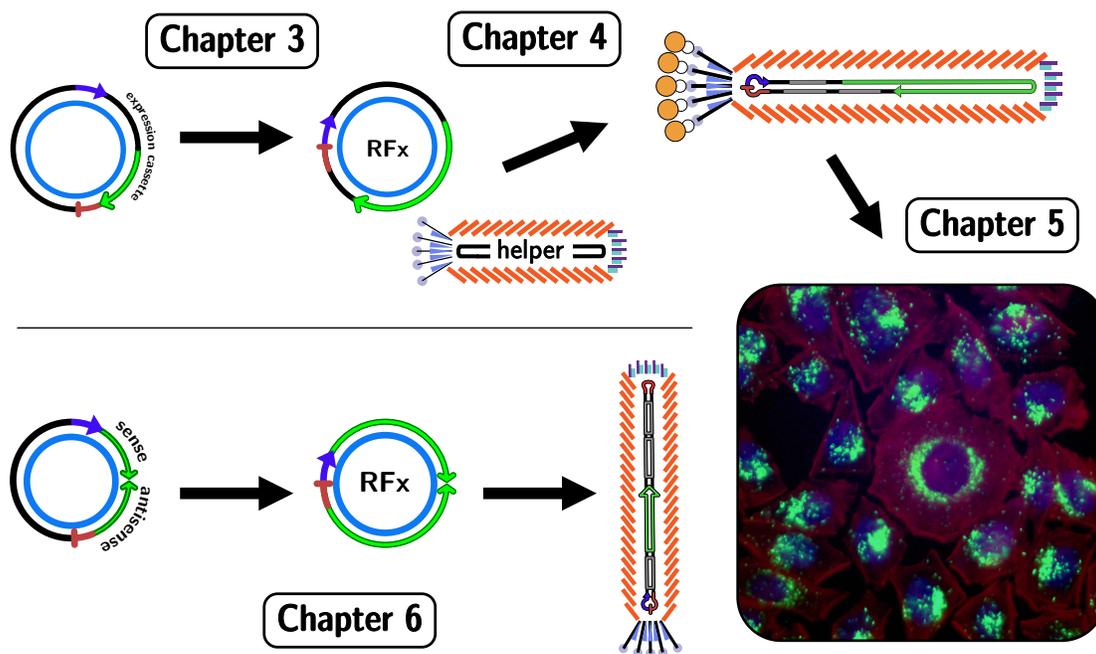
Figure 2.2: Overview of the project aims in this thesis. As indicated, each chapter implements a step in the construction and application of a set of novel hybrid targeted M13 gene transfer vectors. In the top branch, we first aim to construct a system for the production of a single-stranded M13-mediated DNA minivector (Chapter 3). This minivector can be encapsulated into a tissue-targeted phage virion (Chapter 4), which we evaluate for gene delivery in mammalian cells (Chapter 5). In the bottom branch, we further expand the platform to demonstrate production of a double-stranded DNA minivector (Chapter 6).

2.1.1 Hypotheses

A single-stranded DNA miniphagemid: Flanking a transgene expression cassette with separated initiation and termination regions of the filamentous replication origin will enable production of a miniaturized phagemid ("miniphagemid") without any bacterial backbone, packaged into a filamentous phage virion (Figure 2.3). Helper phage proteins will package a vector with a split origin with less efficiency than the wild-type origin. However, a miniphagemid will transfect more efficiently into mammalian cells than a "full" phagemid that retains the entirety of the bacterial backbone. These hypotheses pertain to the top branch of Figure 2.2.

A double-stranded DNA miniphagemid: Encoding both the sense $(+)$ and antisense $(-)$ sequences of a transgene cassette between separated f1 *ori* domains will enable production of a double-stranded miniphagemid (Figure 2.4). Not only will this miniphagemid vector not retain any bacterial backbone, the complementary sequences will bind and form a linear covalently-closed double-stranded vector, packaged into a filamentous phage virion. These hypotheses pertain to the bottom branch of Figure 2.2.
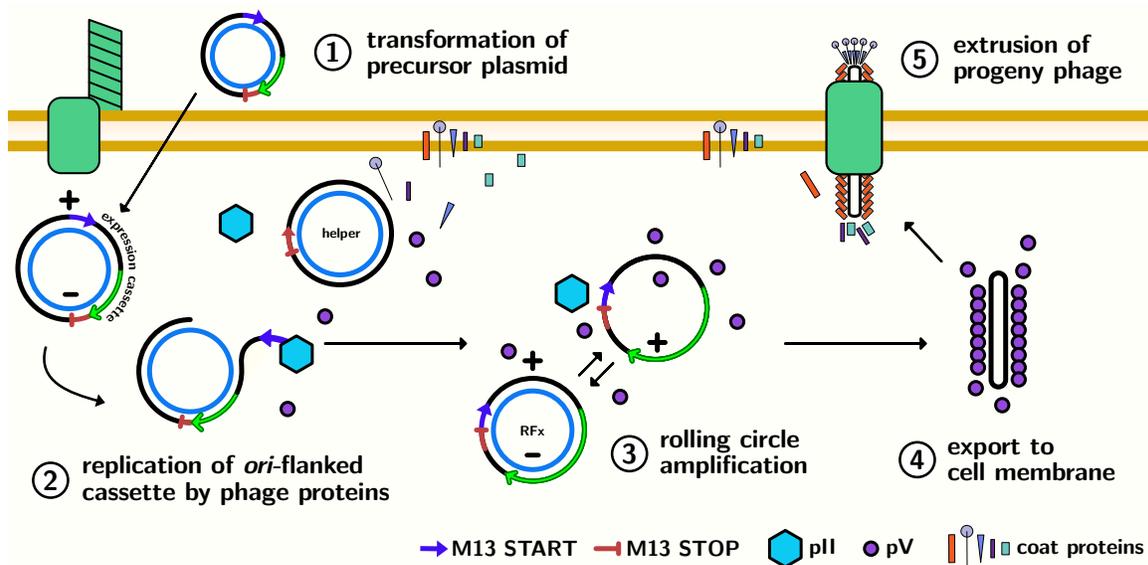
Figure 2.3: Proposed mechanism of single-stranded miniphagemid production. Based on existing knowledge of the lifecycle of M13 and how it interacts with its genomic signalling structures, here is a schematic depicting how single-stranded miniphagemid production could occur: 1) A precursor plasmid encoding a transgene cassette flanked by the separated f1 *ori* domains will be introduced into a susceptible host cell for replication, alongside a helper phage. Thus, two DNA species are present in the cell. 2) Phage gene expression will occur, enabling pII-mediated replication of the transgene cassette. 3) A third DNA species will be introduced, the replicative factor containing only the transgene cassette and the reconstituted f1 *ori* ("RFx"). 4) Phage replication continues as pV sequesters single-stranded miniphagemid DNA to the cell membrane. 5) Progeny phage carrying the miniphagemid are extruded from the cell.
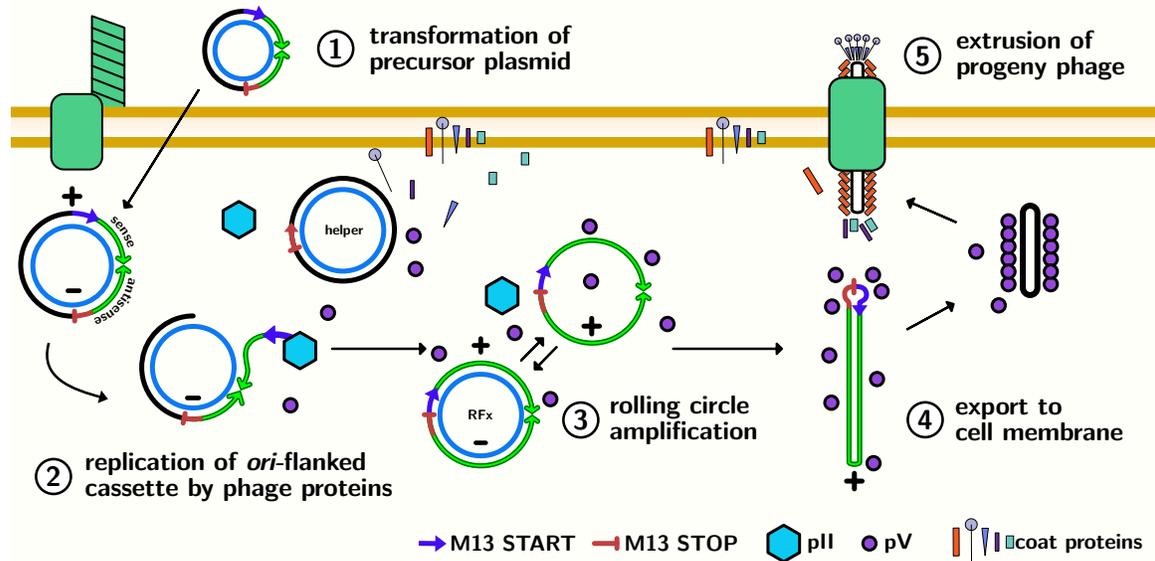
Figure 2.4: Proposed mechanism of double-stranded miniphagemid production. Based on existing knowledge of the lifecycle of M13 and how it interacts with its genomic signalling structures, here is a schematic depicting how double-stranded miniphagemid production could occur: 1) A precursor plasmid encoding a both complementary sequences of a transgene cassette flanked by the separated f1 *ori* domains will be introduced into a susceptible host cell for replication, alongside a helper phage. Thus, two DNA species are present in the cell. 2) Phage gene expression will occur, enabling pII-mediated replication of the entire sense-antisense cassette. 3) A third DNA species will be introduced, the replicative factor containing only the sense-antisense region and the reconstituted f1 *ori* ("RFx"). 4) Phage replication continues as pV sequesters the plus strand to the cell membrane, during which it may self-anneal to form a linear covalently closed (LCC) double-stranded miniphagemid. 5) Progeny phage carrying the miniphagemid are extruded from the cell.

2.1.2   Project aims

The overall aim of this thesis is to characterize the production and efficiency of miniaturized filamentous M13 phagemids for gene delivery in mammalian cells.

To evaluate the hypotheses, the project is broken down into the following objectives.

- Chapter 3: Construction and characterization of miniaturized phagemids from precursor plasmids encoding mammalian gene cassettes flanked by separated domains of the f1 *ori*.

  - Precursor plasmids with wild-type f1 *ori* and split f1 *ori* will be contrasted in their ability to produce phagemid virion particles.

- Chapter 4: Characterization and construction of helper phages.

  - As a proof of principle, a cell-specific display peptide (the epidermal growth factor) will be incorporated into the helper phage to generate targeted virion particles in a Type 3 display system.
  - A novel helper phage without a packaging signal will be assessed for its ability to package phagemid particles without packaging its own genome.

- Chapter 5: Characterization of the ability of miniaturized phagemids to deliver gene cassettes to mammalian cell lines in vitro.

  - Gene expression as a result of phage transfection will be assessed by luciferase activity across different cell lines: kidney (HEK-293T), cervical (HeLa), colon (HT-29), and lung (A549, MRC-5).
  - Phage displaying a targeting ligand (epidermal growth factor) will be compared to phage not displaying any targeting ligand.
  - The addition of a cationic lipid transfection reagent will be assessed for its capacity to improve phage transfection.
  - The cytotoxicity of the phage lysates will be assessed in a representative cell line (HeLa).

- Chapter 6: Construction and characterization of precursor vector with both sense (+) and antisense (−) sequences of a gene expression cassette for phage-mediated production of a linear dsDNA miniphagemid.

  - Since the proposed vector encodes large inverted repeats, its construction will be carried out in an *E. coli* strain deficient in recombination.
  - The precursor plasmid will be characterized in its ability to produce double-stranded phagemid virion particles.

Chapter 3

M13-mediated assembly of a single-stranded DNA minivector phagemid

3.1 Introduction

Viruses such as the filamentous bacteriophage M13 exist as inert virions outside their host cell, *E. coli*. In order to propagate, M13 needs to infect a bacterial host, replicate within the cell, and finally assemble phage progeny that then extrude from the cell. Its simplistic genome encodes all eleven proteins necessary for these activities, which are controlled by signalling structures within a short non-coding intergenic region (IR). A phagemid is a plasmid that contains the only the sequences within the IR, also known as the f1 functional origin (f1 *ori*). A conventional plasmid with both a plasmid origin and a phage origin, the phagemid can be replicated by filamentous phage machinery independent of its plasmid origin [314, 315]. However, these elements provide no benefit for mammalian gene transfer applications. The prokaryotic backbone which is necessary for amplification and maintenance in a bacterial host is rich in unmethylated cytosine-guanine dinucleotide (CpG) motifs that are known to inhibit transgene expression in mammalian cells [83]. It furthermore contains an antibiotic resistance marker that can disseminate antibiotic resistance into the environment [50]. In this chapter, we aim to produce miniaturized phagemids, which lack the superfluous and potentially immunostimulatory genetic material for improved gene transfer in mammalian cells.

3.1.1 Replication of filamentous phage M13

Upon infection, the M13 genome is converted to a dsDNA episomal replicative factor (RF), whereupon phage gene expression occurs. Both replication of filamentous phage DNA and assembly of phage virion particles are directed by key signalling structures within the non-coding IR of the

phage genome, located between gIV and gII. It has long been established which of the hairpin
loops within this f1 *ori* are necessary for replication and assembly [208]. The domain containing
cis-acting elements to initiate replication begins with hairpin loops C-E (Figure 3.1) that comprise
the site of action for the phage replication protein, pII [319]. The domain further encompasses the
region directly upstream of gII where the host integration host factor (IHF) is hypothesized to
bind and participate in phage replication [202]. The main factors governing M13 phage replication
are the interactions between pII and host replicative machinery with f1 *ori* structures. Replication
is initiated by pII-mediated cleavage at the hairpin loop [182, 205], and continues unidirectionally
along the genome, terminating when it reaches the pII cleavage site again [206]. Single-stranded
episomes are later sequestered for assembly into progeny phage and extruded out of the cell over
the course of the bacterium's lifetime.

Figure 3.1: The M13 functional origin contains binding regions for phage and host machinery.
From left to right, immediately downstream of gIV: A) packaging signal, B) RNA polymerase
recognition signal, C) site for primer synthesis, D-E) site for pII recognition and cleavage.
E. coli integration host factor binds within enhancer regions downstream of loops D and
E (red). The AvaI site is indicated. The termination-specific domain is indicated in teal,
initiation-specific in red, and regions participating in both in black.

The introduction of the f1 *ori* fragment into any vector backbone confers upon it the ability
to be packaged into phage, which has been exploited for the production of single-stranded (ss)
DNA [267, 308]. Such a vector is referred to as a "phagemid". Since the phagemid itself does
not encode any phage proteins, a helper phage is necessary to provide proteins in trans. The
helper phage M13KO7 has a defective origin of replication, but it is not completely deficient in
self-replication and assembly [267]. Insertion of a plasmid origin and a kanamycin resistance marker
in the AvaI site separates the hairpin loops from the binding site for host *E. coli* IHF (Figure 3.1).
Phage replication generally proceeds poorly without IHF activity [202]; hence wild-type f1 *ori* are
replicated preferentially over the M13KO7 genome.

Filamentous virion production results from a complex interplay that is highly dependent on stoichiometric balance between phage products. This is perhaps best exemplified by the transition from phage amplification as double-stranded RFs to assembly and extrusion of single-stranded phage. This transition is governed by accumulation of pV [214, 219, 220]. If there is an insufficient pV:RF ratio, single-stranded intermediates are converted to double-stranded RF before they can be sequestered by pV [214]. On the other hand, if there is too much pV:RF, no ssDNA conversion to dsDNA RF may take place so phage replication ceases and the cell is cured of M13 infection. Thus, phage copy number is paramount to good phage production. In the case of phagemids, the phagemid:helper ratio depends on the phagemid outcompeting the helper for assembly, without losing helper-provided phage gene expression.

Dotto et al. (1981) demonstrated that M13-mediated replication of a vector could initiate and terminate from two distant cis f1 *ori* regions on the same vector [208], resulting in a chimeric phage progeny population. Phage progeny all contained only one origin each, and contained only the sequences between the flanking origins. Short et al. (1988) leveraged this ability of the M13 replication system by isolating the sequences necessary for initiation and termination, then flanking a desired sequence for amplification within the separated domains [319]. They demonstrated that the f1 *ori* could be reconstituted by rescuing a phagemid vector that was inserted into the λ genome.

3.1.2 Rationale and hypothesis

In this chapter, we describe the construction of a precursor vector encoding a mammalian transgene cassette flanked by the separated signalling domains. Our aim was to generate DNA minivectors, deficient in bacterial or phage genetic sequences, that can be used for gene delivery in mammalian cells. As the transgene is encoded between the separated domains of the f1 *ori*, pII-mediated replication was hypothesized to excise and re-circularize the transgene cassette, thereby separating it from the plasmid backbone (Figure 3.2). Here, we evaluated if minivectors containing only the transgene cassette can be assembled from such precursors using the helper phage M13KO7.

We evaluated the production of mini or "full" phagemids arising from precursor plasmids carrying a split or wild-type f1 *ori* and quantify the resultant phage progeny. We hypothesized that the wild-type f1 *ori* present on the backbone of our wild-type phagemid controls would be preferentially replicated and packaged over the helper phage itself. The split origin in the miniphagemid precursors separates the sequences known only to participate in the initiation of replication from those known only to participate in the termination thereof. As such, it was hypothesized that production efficiency of miniphagemid progeny would be less efficient due to the extra requirement for forming a recombinant RF (RFx) prior to phage assembly. Although we expect less efficient replication of the split f1 *ori* than from the wild-type *ori*, the high titres normally expected from phage M13 purification may overcome this disadvantage.
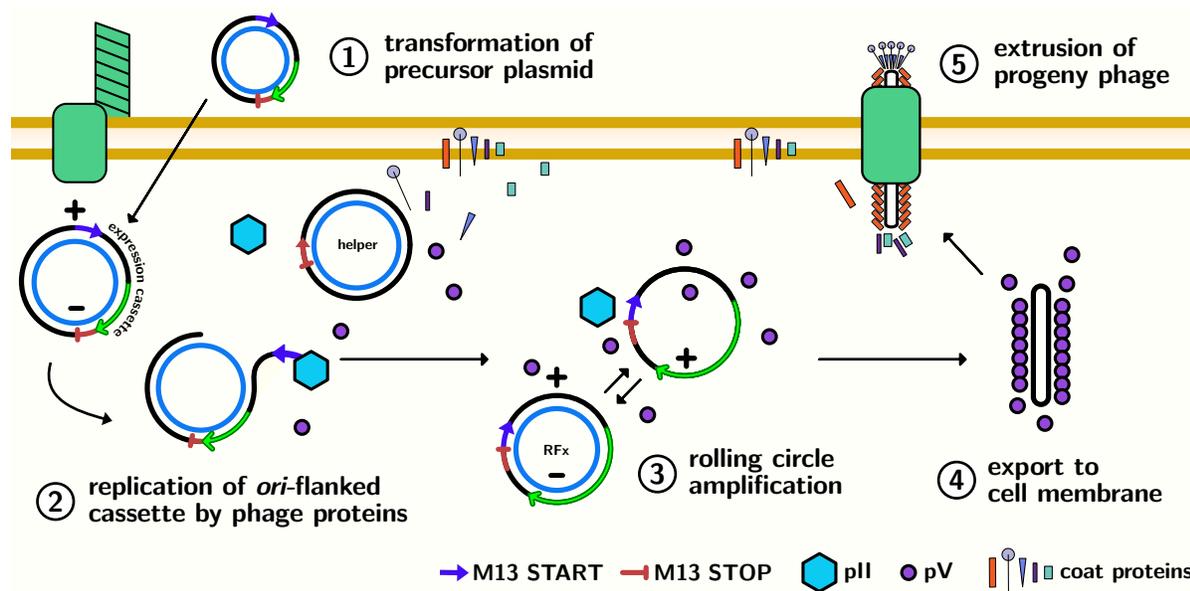
Figure 3.2: Production and purification of miniphagemid DNA. The schematic depicts an overview of miniphagemid production within the host cell. 1) The host cell is transformed by the precursor plasmid encoding a region of interest between the separated M13 START and STOP domains. 2) After infection with a helper phage, phage protein expression can occur. 3) Rolling circle amplification generates a recombinant replicative factor (RFx) from the plasmid that reconstitutes the f1 *ori* and loses the plasmid backbone. 4) The phage ssDNA binding protein pV sequester single-stranded minivector and prevents further amplification. 5) Assembly proteins extrude progeny phage particles encapsulating the minivector.

3.2   Materials and methods

Strains and vectors.   *E. coli* K-12 strains were used in the generation of all constructs. *E. coli* JM109 was the host for plasmid amplification and purification unless otherwise specified. Bacterial strains are listed in Table 3.1, plasmids in Table 3.2, and phages in Table 3.3.

Table 3.1: Bacterial strains used in the characterization of miniphagemids

Strain	Genotype	Source
JM109	F′ traD36 proAB$^+$ lacI^q lacZΔM15/ Δ(lac-proAB) endA1 glnV44 thi-1 e14$^-$ recA1 gyrA96 relA1 hsdR17	NEB
XL1-Blue	F′ traD36 proAB$^+$ lacI^q lacZΔM15 Tn10/ lac endA1 glnV44 thi-1 recA1 gyrA96 relA1 hsdR17	Agilent Technologies
ER2738	F′ zzf::Tn10(TcR) proAB$^+$ lacI^q lacZΔM15/ Δ(lac-proAB) thi-1 glnV44, Δ(hsdS-mcrB)5, fhuA2	NEB

NEB: New England BioLabs

Table 3.2: Plasmids used in the characterization of miniphagemids

Plasmid	Genotype	Source
pGL2-SS-CMV-GFP-SS	pGL2-Promoter (GN: X65326.2) *cmv-gfp* replaced *SV40-luc*, Ap^R	gift, Mediphage Bioceuticals Inc.
pGL3-CMV	pGL3-Basic (GN: U47295.2) *cmv* inserted in BglII-HindIII, Ap^R	gift, Dr. N. Oviedo
pBluescript II KS+	wild-type f1 ori, pUC ori, Ap^R	[405]
pBR322	Ap^R, Tc^R	[406]
pSW9	pBluescript II KS+, *cmv-gfp* inserted in BamHI-EcoRI, ApR	This study
pSW10	pBluescript II KS+, *cmv-luc* inserted in KpnI, ApR	This study
pM13ori	split f1 ori inserted in pUC57 (Genbank Accession No. Y14837.1), ApR	This study
pM13ori2	pM13ori, BsaI cut site replaced with SphI, ApR	This study
pM13ori2.cmvgfp	pM13ori2, *cmv-gfp* inserted in EcoRI-PacI	This study
pM13ori2.cmvluc	pM13ori2, *cmv-luc* inserted in EcoRI-KpnI	This study
pM13ori2.cmvgfp-tet	pM13ori2.cmvgfp, Tc^{R}from pBR322 inserted into KpnI	This study
pSW9-tet	pSW9, Tc^{R}from pBR322 inserted into SmaI	This study

GN: GenBank Accession No.

Table 3.3: Phages used in the characterization of miniphagemids

Phage	Genotype	Source
M13	wild-type	CGSC #13587
M13KO7	M13, p15a *ori*, Tn903 (Km^R)	NEB
M13KE	M13, KpnI and EagI target sites in gIII	NEB
full-(gfp)	*cmv-gfp*, from precursor pSW9, Ap^R	This study
mini-(gfp)	*cmv-gfp*, from precursor pM13ori2.cmvgfp	This study
full-(luc)	from precursor pSW10, Ap^R	This study
mini-(luc)	from precursor pM13ori2.cmvluc, Ap^R	This study
mini-(gfp-tet)	from precursor pM13ori2.cmvgfp-tet, Tc^R	This study
full-(gfp-tet)	from precursor pSW9-tet, Tc^R	This study

CGSC: Coli Genetics Stock Center   NEB: New England BioLabs

Media and buffers.   All media and buffers were prepared according to recipes in Appendix A, unless otherwise specified. Bacterial strains were cultured in Luria–Bertani (LB) liquid medium, while plaque assays were carried out in top agar. Phage lysates were purified and stored in Tris-NaCl (TN) buffer. Where indicated, phage lysates were concentrated with polyethylene glycol (PEG).

Design and construction of a precursor vector for miniphagemid production.   The precursor vector backbone pM13ori was designed based on the regions identified by Short et al. [319]. A polylinker was designed in Snapgene 5.3 (`https://www.snapgene.com/`) and flanked by the separated regions (M13 START upstream and M13 STOP downstream) (Figure 3.3). The entire cassette was synthesized and subcloned into pUC57 to construct pM13ori (Genscript Inc, Piscataway, USA). The BsaI site present within the ampicillin resistance marker (*bla*) was replaced with SphI through insertional mutagenesis with primers BsaI-SphI-F and BsaI-SphI-R to construct pM13ori2 (Table 3.4). This was to accommodate potential BsaI-dependent Golden Gate assembly [407, 408] in the future.

Table 3.4: Primers for vector construction

Primer	Template	Sequence (5′ – 3′)[1]
BsaI-SphI-F	pM13ori	<u>GCATG</u>CCACGCTCACCGGCTC
BsaI-SphI-R	pM13ori	GCGGTATCATTGCAGCACTG
EcoRI-luc-F	*cmv-luc*	<u>GAATTC</u>CCAGATATACGCGTTGAC
KpnI-luc-R	*cmv-luc*	<u>GGTACC</u>CTGCAGATCCTTATCGATTTTACCAC
KpnI-tet-F	pBR322 (4325..1835)	<u>GGTACC</u>AAATAGGCGTATCACGAGGC
KpnI-tet-R	pBR322 (4325..1835)	<u>GGTACC</u>GTTACTGATGATGAACATGCC

[1] Underlined nucleotides indicate primer overhang with endonuclease target site as indicated in primer name.
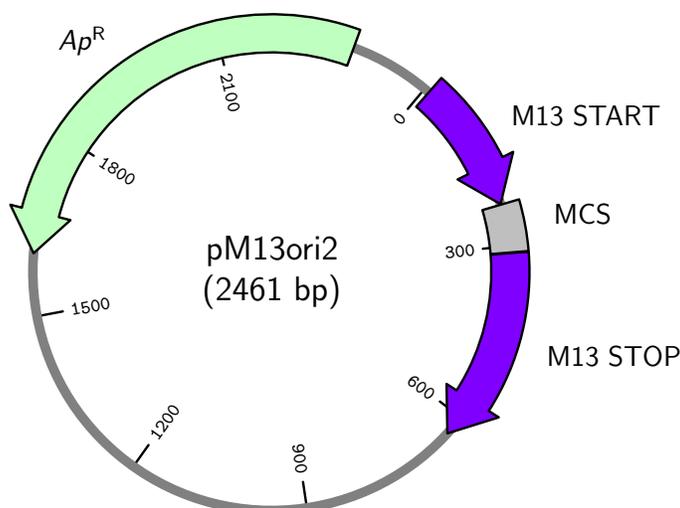
Figure 3.3: The regions required for initiation (M13 START) and termination (M13 STOP) were inserted into a pUC57 backbone, separated by a polylinker, to generate pM13ori2.

Construction of a precursor vector for mini or full phagemid production. To construct precursor vectors for miniphagemid production, reporter gene cassettes (green fluorescent protein (GFP) and luciferase) were subcloned into the polylinker of pM13ori2. These reporters serve as stuffer sequences of different lengths (approximately 2 kb versus 2.5 kb) in this proof of principle, and can laterbe used to evaluate phage-mediated gene transfer. To construct precursor vectors for "full" phagemid production, reporter gene cassettes were subcloned into the same wild-type f1 *ori* backbone: pBluescript II KS+.

The expression cassette *cmv-gfp* from pGL2-SS-CMV-GFP-SS was subcloned directly into the EcoRI-PacI sites (Genscript Inc) to generate pM13ori2.cmvgfp (Figure 3.4A). The expression cassette *cmv-gfp* was then liberated from pM13ori2.cmvgfp by digestion with EcoRI and BamHI (New England BioLabs, Ipswich, USA). The phagemid pBluescript II KS+ was similarly digested. Purified digested fragments were ligated with T4 DNA ligase (New England BioLabs), then selected on LB agar supplemented with ampicillin to generate pSW9 (Figure 3.5A).

Separately, the expression cassette *cmv-luc* was amplified by polymerase chain reaction (PCR) from the source plasmid pGL3-CMV using primers EcoRI-CMVluc-F and KpnI-CMVluc-R (Table 3.4), which added EcoRI and KpnI target sites on either side of the amplified fragment. The vector backbone, pM13ori2, was digested with EcoRI and KpnI for ligation with the similarly digested purified PCR product to generate pM13ori.cmvluc (Figure 3.4B). The EcoRI-*cmv-luc*-KpnI PCR product from above was also used for subcloning into pBluescript II KS to generate pSW10 (*cmv-luc*) (Figure 3.5B).

45

(A) gfp   **(B)** luc   **(C)** gfp-tet
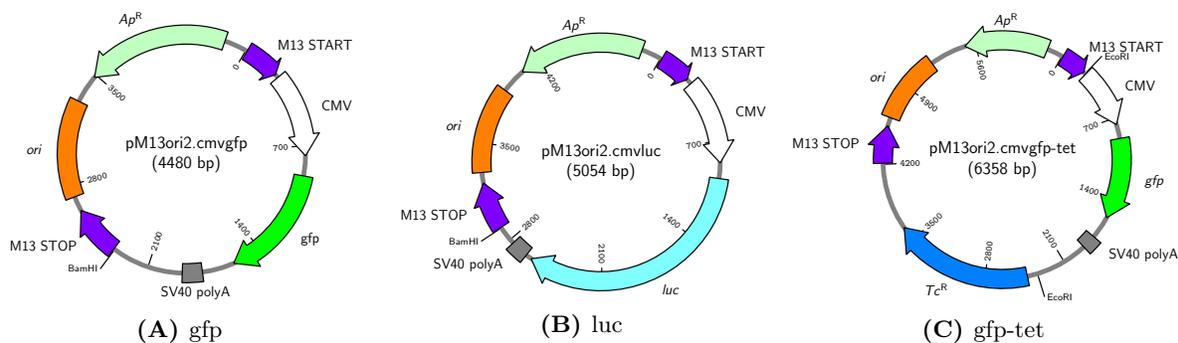
Figure 3.4: Maps of miniphagemid precursors. Reporter genes *gfp* (A) and *luc* (B) under the control of the universal promoter, CMV, were inserted between M13 START and M13 STOP in pM13ori2. A Tc$^{\mathrm{R}}$ fragment from pBR322 was inserted downstream of *cmv-gfp* to generate the third construct (C).

(A) pSW9   **(B)** pSW10   **(C)** gfp-tet
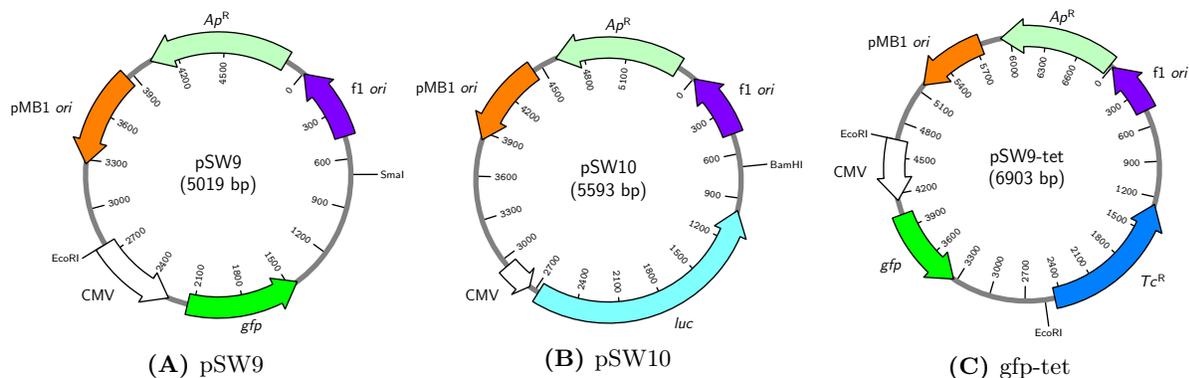
Figure 3.5: Maps of phagemid controls. Reporter genes *gfp* (A) and *luc* (B) under the control of the universal promoter, CMV, were inserted into the polylinker of the phagemid backbone pBluescript II KS+, which itself has an f1 *ori*. A Tc$^{\mathrm{R}}$ fragment from pBR322 was inserted downstream of *cmv-gfp* to generate the third construct (C).

46

Construction of longer phagemid expression cassettes.   To examine the impact of length on miniphagemid production, a 1884 bp fragment from pBR322 (containing the *tet* gene for Tc^R) was inserted into precursor plasmids encoding *cmv-gfp*. The fragment was amplified by PCR from pBR322 using primers KpnI-tet-F and KpnI-tet-R (Table 3.4), which added KpnI target sites. For insertion into the miniphagemid precursor vector pM13ori2.cmvgfp, both the vector and the purified PCR product were digested with KpnI. The blunt PCR product was also inserted into SmaI cut site of pSW9. Respective insert-vector pairs were ligated and selected on LB agar supplemented with ampicillin and tetracycline. The purified final constructs were pM13ori2.cmvgfp-tet (*cmv-gfp, tet*), and pSW9-tet (*cmv-gfp, tet*). Both Tc^R constructs are approximately 2 kb larger than their non-Tc^R counterparts (Figures 3.4 and 3.5).

Phage amplification.   Figure 3.6 depicts the overall workflow for production and purification of phage, phagemids, and RF DNA. A fresh colony of the host carrying the phagemid precursor plasmid was first inoculated into rich media and incubated at 37 °C with aeration until slightly turbid ($0.01 < A_{600} < 0.4$). The culture was then infected with helper phage M13KO7 and returned to the same conditions for an additional 1–1.5 h. It was then supplemented with kanamycin (70 µg/mL) and returned to the same conditions overnight. The next day, the culture was centrifuged ($8000 \times g$, 10 min) to separate the bacterial pellet containing RF from the supernatant containing the phage lysate.
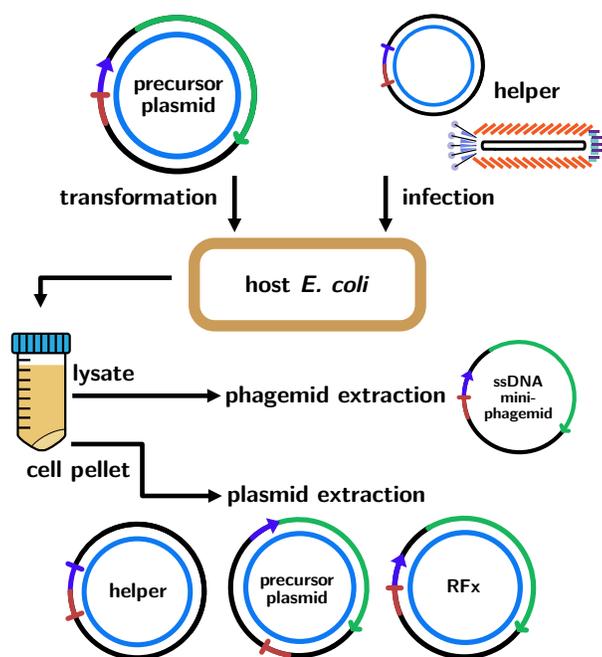
Figure 3.6: Workflow of miniphagemid production and purification. The schematic depicts the overall workflow for the production and purification of a phagemid lysate and phagemid RF DNA from a precursor plasmid and helper phage.

Purification of phage lysate.   After removal of the cell pellet, the phage lysate was further purified through a 0.45 μm filter to remove residual bacterial debris. Filtered lysate was concentrated through precipitation with PEG: 1/5 of the lysate volume was added in PEG, and the mixture was incubated at 4 °C for at least 2 h. The lysate was then centrifuged at $12\,000 \times g$ for 15 min at 4 °C to separate the PEGylated phage (pellet) from media (supernatant). The pellet was then re-suspended in a smaller volume of ice-cold TN buffer and a second PEG precipitation was carried out to further concentrate the phage. PEG-precipitated phage suspensions were treated with DNase I (Promega, Madison, USA) to remove any extraneous phage or bacterial DNA in the sample [409, 410]. The concentrated phage lysate was stored at 4 °C.

Purification of the double-stranded RF and single-stranded miniphagemid DNA. Episomal RF dsDNA were extracted from the pellet with the Monarch Plasmid Miniprep Kit (New England Biolabs). Phagemid DNA were extracted from the phage lysate through phenol-chloroform extraction. Phenol was added to PEG-concentrated phage lysate (1:1, v/v) and mixed by vortexing. After centrifugation for 5 min at 4 °C, the top aqueous layer was extracted. This layer was then extracted again with an equivalent volume of phenol:chloroform twice, and chloroform once. Finally, the phagemid ssDNA was precipitated overnight with 100% ethanol at −80 °C. The precipitated DNA was washed with 70% ethanol, dried, and re-suspended in RNase and DNase-free water.

Quantification of phagemid and RF DNA.   Extracted DNA was analyzed on a NanoDrop 2000 to determine concentration and purity. Extracted RFs were linearized through digestion with BamHI and visualized via agarose gel electrophoresis (AGE), while phagemid ssDNA was visualized directly via AGE without digestion.

Quantification of phage lysates.   The populations of helper and recombinant phagemids within each phage suspension were quantified through SYBR Green quantitative PCR (qPCR). Calibration curves for quantifying phage were constructed using external standards: M13KE RF (7222 bp; New England BioLabs) for helper phage, pGL2-SS-CMV-GFP-SS (5257 bp; Mediphage Bioceuticals) for *gfp*-encoding target phage, and pGL3-CMV (5678 bp) for *luc*-encoding target phage. Primers specific to a region within gV of M13KO7 [409] were used to amplify helper phage. Primers targeting a region within *gfp* or *luc* were designed to amplify *gfp*- or *luc*-encoding phage (Table 3.5). To generate each calibration curve, 10-fold serial dilutions of each external control were prepared (10^{-1} to 10^{-8}) as template for the PCR reaction.

To prepare the phage particles for PCR, the lysates were denatured by heat for 100 °C for 15 min to isolate phage DNA [409]. Ten-fold serial dilutions of each cleared lysate were prepared as template for the PCR reaction. As the phage lysates were of unknown concentration, dilution series of each were amplified through qPCR. For PEGylated lysates, dilutions in the range 10^{-3}–10^{-5} were examined, while for non-PEGylated lysates, 10^{-1}–10^{-3}.

Table 3.5: Primers for qPCR to quantify phagemid and helper phage

Primer	Amplicon	Sequence $(5' - 3')$
g5-F	gV	CACCGTTCATCTGTCCTCTTT
g5-R	gV	CGACCTGCTCCATGTTACTTA
gfp-F	gfp	CAAGATGAAGAGCACCAAAGG
gfp-R	gfp	CGAAGTGGTAGAAGCCGTAG
luc-F	luc	GCGCGGAGGAGTTGTGTT
luc-R	luc	TCTGATTTTTCTTGCGTCGAGTT

Each 10 µL PCR reaction was prepared using 5 µL of PowerUp SYBR Green Mix (Thermo Fisher Scientific, Waltham, USA), 1 µL each of 500 nM primer (forward and reverse), 2 µL of template, and 1 µL of dH$_2$O. PCR cycling conditions were as follows: 50 °C for 2 min, 95 °C for 2 min, followed by 40 cycles at 95 °C for 15 s and 60 °C for 1 min. Next, the melt curve was set for 1 cycle at 95 °C for 15 s, 60 °C for 1 min, and 95 °C for 15 s. PCR reactions were run in triplicate on the StepOne Plus Real-Time PCR system (Applied Biosystems, Waltham, USA). The quantification cycle or threshold cycle number (C$_q$) for each reaction was recorded and used in subsequent analysis.

Conversion from the mass of dsDNA to the number of genome copies is given by

$$gc = \left(\frac{mass}{size} \times 607.4 + 157.9\right) \times (6.02 \times 10^{23}), \tag{3.1}$$

where gc is the concentration of phage genome copies in genome copies/µL, $mass$ is the mass of the dsDNA standard in g/µL, and $size$ is its length in bp [409]. Phage concentrations are estimated from their respective calibration curve. First, the C$_q$ values from each exogenous control are plotted against the log of the known concentrations for each concentration in the dilution series. Linear regression produces a familiar equation of the form

$$C_q = mx + b \tag{3.2}$$

where C$_q$ is the measured threshold cycle number, m is the slope, x is the base-10 log of the concentration in gc/µL and b is the x-intercept. PCR amplification efficiency is given by

$$E = 10^{-1/m} - 1. \tag{3.3}$$

Subsequently, virion concentration V can be estimated simply by

$$V = 10^{(C_q - b)/m} \times 2 \tag{3.4}$$

where multiplication by 2 adjusts for the estimation of single-stranded products (gc/µL) from dsDNA standards. Only C$_q$ measurements within the bounds of the calibration curve were used to calculate phage concentration.

Quantification of viable phage through plaque and colony assay.   A plaque assay was conducted to quantify viable phage. Plaque and colony assays were all using ER2738, a fast-growing F^+ strain that is recommended for the production of filamentous phage (New England BioLabs). First, the appropriate dilution to generate countable (10–300) plaques was determined through a spot plate. Briefly, $200\,\mu$L of early log-phase *E. coli* ER2738 was mixed with $3\,$mL top agar supplemented with $5\,$mM $MgSO_4$, then poured on a pre-warmed LB agar plate. Separately, phage lysates were serially diluted in TN buffer, then $5\,\mu$L of each phage lysate dilution were spotted onto the top agar. Plates were incubated overnight at $37\,°$C and analyzed the next day. The lowest dilution that generated distinct plaques on the bacterial lawn was selected to generate full plates for plaque assay. For the full plaque assay, the same procedure was followed except that the entire volume of diluted phage was added directly to the volume of bacterial culture prior to the addition of top agar and pouring. The titre was determined from the mean of three counts and expressed as plaque forming units (PFU) per millilitre.

A colony assay was conducted to quantify the ability of infectious helper phage to confer antibiotic resistance. As with the plaque assay, the appropriate dilution to generate countable (10–300) colonies was determined through a spot plate. First, $100\,\mu$L aliquots of serially diluted phage lysates were prepared in TN buffer. $100\,\mu$L of early log-phase *E. coli* ER2738 was added to each aliquot, and incubated at $37\,°$C for $0.5\,$h. $5\,\mu$L of each aliquot was spotted on LB agar plates supplemented with kanamycin. Dried plates were incubated overnight at $37\,°$C and analyzed the next day. For full plates, the same procedure was followed except that the entire dilution volume $(200\,\mu$L) was spread on an LB agar plate supplemented with kanamycin. The titre was determined from the mean of three counts and expressed as colony forming units (CFU) per millilitre.

Prediction of phagemid secondary structure.   To further ascertain differences between the phagemid constructs, rudimentary secondary structure prediction was performed. Centroid secondary structures of miniphagemids, phagemids, and phages were predicted using Vienna RNAfold 2.5.0 [411, 412], using available DNA thermodynamic parameters [412, 413].

Statistical analysis.   All statistical analyses were performed using Python (with the packages NumPy [414], Pandas [415, 416], SciPy [417], scikit-bio [418], and statsmodels [419]). Values are reported as means of n independent experiments with uncertainty reported as the standard error of the mean (SEM). Statistical hypothesis tests were evaluated using one-way ANOVA, followed by the Tukey range test for multiple comparisons. Values of $p < 0.05$ were considered statistically significant; in some cases, values of $p < 0.1$ are noted and discussed.

The phagemid fraction was determined as the concentration of target phagemid divided by the total virion concentration, expressed as percentages. As compositional data [420], they have a fixed constant sum constraint (100%). In order not to violate this constraint, the data were transformed using an isometric log ratio transformation before performing statistical analyses. Data were then transformed back to percentages for reporting.

3.3 Results

The initiation and termination regions of the f1 *ori*, as identified by Short et al. [319], were synthesized and inserted into a pUC57 backbone to create the base precursor plasmid, pM13ori2. Transgene cassettes expressing different reporter genes were subcloned into the polylinker of pM13ori2: *cmv-gfp* (2.1 kb) to generate the minivector phage particle ("miniphagemid") mini-(gfp), and *cmv-luc* (2.6 kb) to generate mini-(luc). They were also cloned into the phagemid vector, pBluescript II KS+, which was used as our wild-type f1 *ori* control, to generate "full" phagemid particles full-(gfp) and full-(luc), respectively.

3.3.1 Production of single-stranded DNA minivector phagemids

We investigated the ability of helper phage M13KO7 to rescue a new recombinant replicative form (RFx) from a split origin precursor plasmid (Figure 3.6). Purified RF dsDNA was linearized with BamHI and visualized via AGE (Figure 3.7). RF DNA isolated from cells infected with wild-type M13 or helper phage M13KO7 alone resolved to their respective genome sizes: 6.4 kb and 8.7 kb. Two expected RF DNA species were present in helper-infected cells transformed by pBluescript II KS+: the helper phage genome and the phagemid (3 kb). In cells transformed by the miniphagemid precursor, pM13ori2.cmvgfp, a single band corresponding to the expected plasmid size (4.5 kb) was observed after BamHI digestion. An additional DNA species was present only in the presence of helper phage and was the expected size for the RFx molecule (2.6 kb).
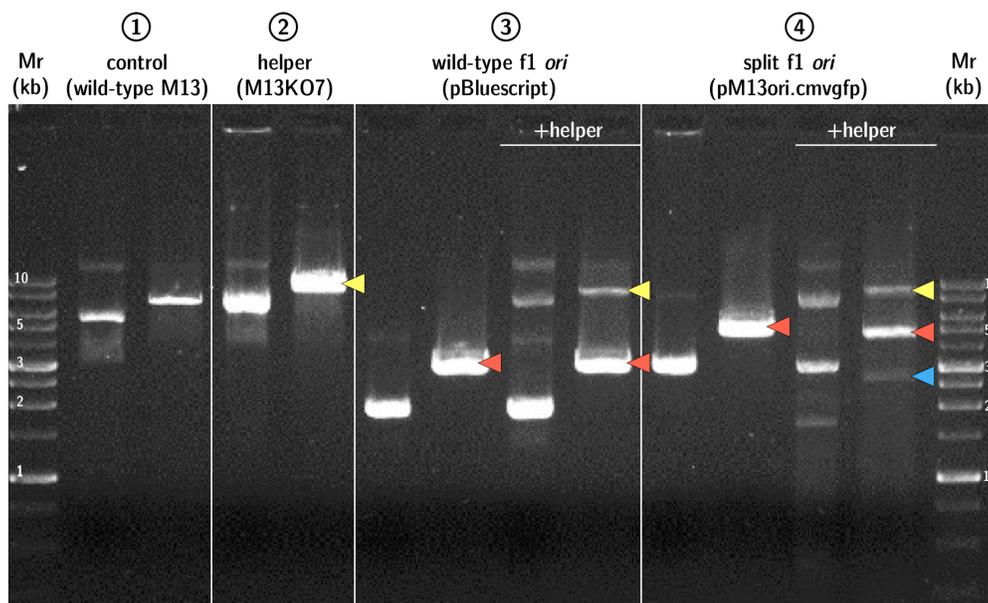
Figure 3.7: Helper phage M13KO7 produces recombinant RFx species from miniphagemid precursor. Double-stranded RFs were linearized with BamHI and visualized on AGE. From left to right (uncut, followed by BamHI-cut): 1) wild-type M13 only (6.4 kb), 2) helper M13KO7 alone (8.7 kb), 3) pBluescript II KS+ (3 kb) alone or with helper, 4) pM13ori.cmvgfp (4.5 kb) alone or with helper. Arrows indicate bands of interest: yellow (helper genome), red (plasmid), blue (recombinant RF: RFx). Mr: 1 kb ladder (FroggaBio).

We next examined purified phagemid ssDNA via AGE without digestion (Figure 3.8), noting that ssDNA runs at approximately half the expected size of the respective dsDNA. For wild-type M13 (6.0 knt) and helper phage alone (8.7 knt), a single ssDNA band was observed in each lane. For M13KO7-assembled pBluescript II KS+, a single band was observed that corresponded to the expected phagemid size (3 knt).

Alongside pM13ori2.cmvgfp, we also constructed an intermediate precursor with a 1 kb *gfp* fragment stuffer sequence. Two distinct phage species were purified as ssDNA: the larger band corresponded to the expected size of helper phage M13KO7, while the smaller band corresponded to the expected size of the miniphagemid. There were three bands from mini-(gfp) that were the expected sizes for the helper genome, the precursor plasmid (4.5 knt), and a third band corresponding to the expected size of recombinant miniphagemid. As we were able to observe miniphagemid production from pM13ori2.cmvgfp, we discontinued experiments with the stuffer precursor in favour of the full *cmv-gfp* cassette.

(A) Yield of ssDNA from phagemid lysates

(B) Visualization of ssDNA from phagemid lysates
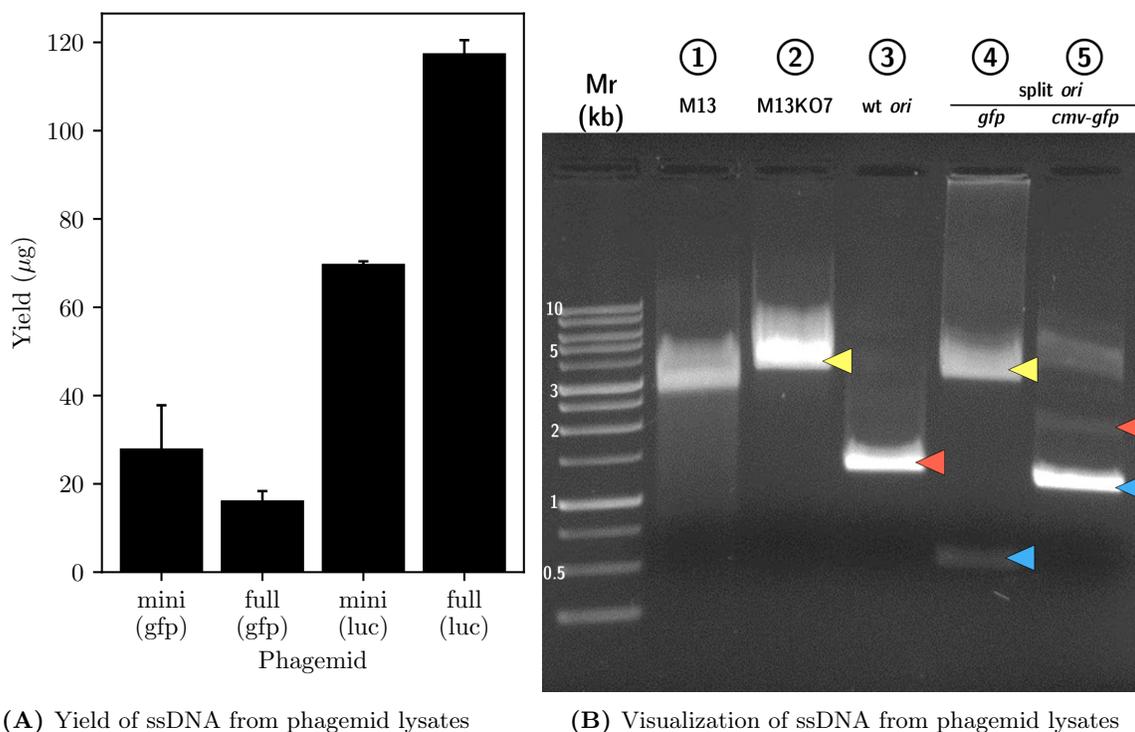
Figure 3.8: Single-stranded miniphagemid DNA isolated from phage lysates. A) Purified ssDNA was quantified on Nanodrop and compared across different phagemids. B) The ssDNA was also visualized on AGE, alongside a dsDNA ladder. The ssDNA runs at approximately half its expected dsDNA size. From left to right, lysates were prepared from: 1) wild-type M13 (6.4 knt), 2) M13KO7 alone (8.7 knt), 3) pBluescript II KS+ (3 knt), 4) pM13ori2 with an approximately 1 kb fragment of *gfp* inserted, 5) pM13ori2.cmvgfp (plasmid: 4.5 knt, miniphagemid: 2.2 knt). Arrows indicate bands of interest: yellow (helper genome), red (plasmid or full phagemid), blue (recombinant miniaturized phagemid). Mr: 1 kb ladder (FroggaBio).

The yield of single-stranded DNA from the phage lysate was comparable across both wild-type and split *ori* (Figure 3.8). No significant difference was observed in the yield of ssDNA arising from a split f1 *ori* compared to an "intact" one for *cmv-gfp*, although more full-(luc) ssDNA was produced than mini-(luc). Intriguingly, the ssDNA yield appeared greater for the *luc*-encoding phagemids than *gfp*-encoding, under the same growth and extraction conditions. However, assessment of yield based on ssDNA mass biases towards larger molecules. Conversion from mass to number of molecules shows that the ssDNA yields are actually comparable: 2.116×10^{13} for mini-(gfp), 1.077×10^{13} for full-(gfp), 4.322×10^{13} for mini-(luc) and 4.088×10^{13} for full-(luc).

3.3.2 Production efficiency of mini and full phagemids

We next quantified the helper phage and target phagemid particles resulting from rescue of a split origin versus intact. DNA quantification approaches such as qPCR are highly accurate methods for enumerating viral genomes, which is equivalent to counting the number of virion particles since every virion encapsulates one phage genome [421–423]. To estimate phage concentration, calibration curves were generated from a dilution series of each exogenous plasmid control: M13KE (gV), pGL2-SS-cmv-gfp-SS (*gfp*), and pGL3-cmv (*luc*). Representative calibration curves are presented in Figure 3.9. Lines of best fit were estimated using linear regression; values of R^2 for each fit were above 0.99. PCR amplification efficiency was calculated using Equation (3.3); in general, PCR efficiencies within 0.9–1.1 were considered acceptable. Representative amplification plots, melt curves, and multicomponent plots for the three amplicons are in Appendix C.
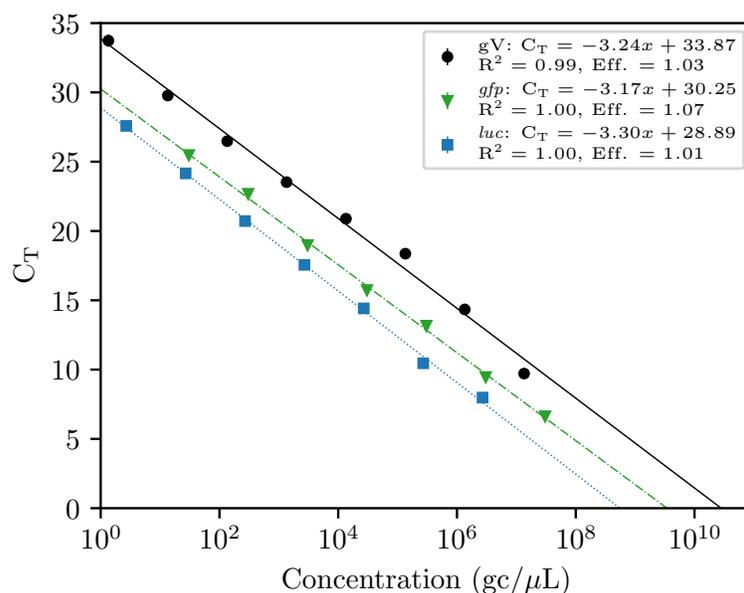
Figure 3.9: Calibration curves to estimate phage concentration. Representative calibration curves are shown for each amplicon, as measured from qPCR of exogenous dsDNA controls: M13KE (gV), pGl2-ss-cmv-gfp (*gfp*), pGL3-cmv (*luc*). The equation for each line of best fit, the R^2 value, and PCR efficiency are reported. Data points represent the mean of 3 technical replicates.

We further contrasted concentrations determined through qPCR against those determined through plaque assay and colony assay (Figure 3.10). Since M13KO7 encodes a selectable marker, it is possible to quantify M13KO7-infected cells that gain resistance to kanamycin. We compared the titre of plaque or colony forming units as estimated against the concentrations of total and helper phage only as estimated by qPCR. For clarity, the quantification by plaque or colony assay is reported as the phage titre, with units PFU or CFU per volume, as it refers to the number of viable or infectious phage. In contrast, quantification by qPCR is reported with units of genome copies (gc) per volume.
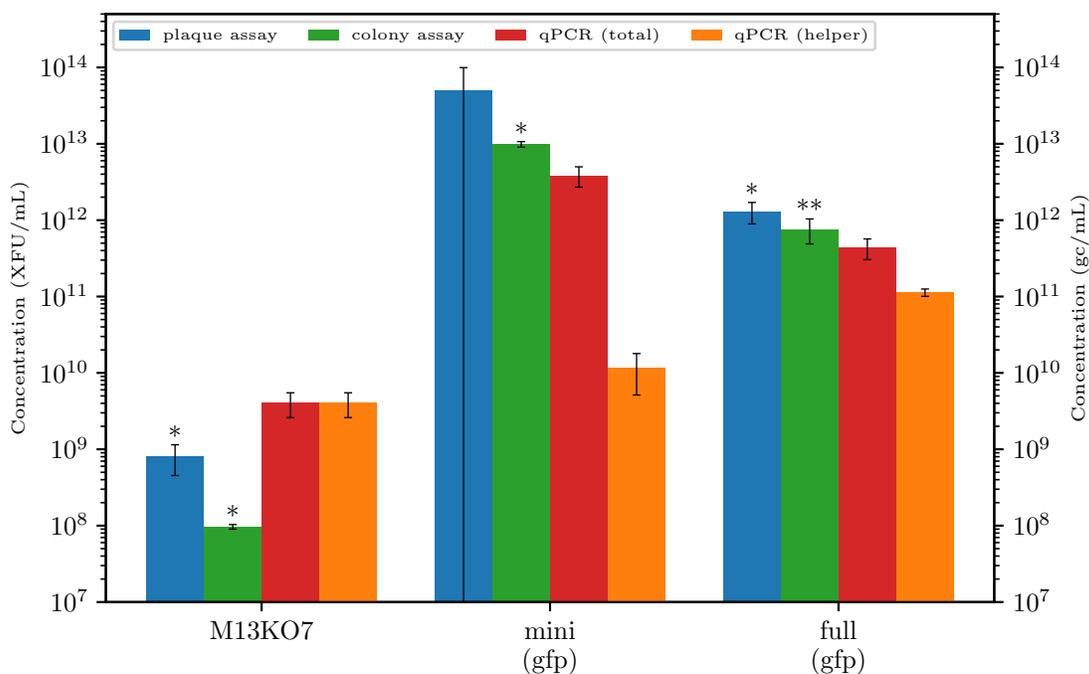
Figure 3.10: Plaque titre estimated by plaque or colony assay compared to qPCR. The titre of viable phage was assessed by plaque or colony assay in XFU/mL where "X" represents "plaque" (P) or "colony" (C), respectively. This was compared to the total estimated phage concentration or helper phage concentration only (genome copies/mL) as estimated by qPCR. Error bars represent SEM$\times 2$, $n = 3$. The $*$ and $**$ above the bars indicate differences at significance levels $p < 0.05$ and $p < 0.1$, respectively, relative to helper phage concentration per lysate.

As a control, we performed both assays and qPCR on a commercial stock of M13KO7 with a known titre of 10^8 PFU/mL. Since no phagemid was present, total virion and helper concentrations were identical. For our control, M13KO7, the plaque assay estimated 8×10^8 PFU/mL, the colony assay estimated 9.7×10^7 PFU/mL, and qPCR estimated 4.0×10^9 gc/mL. The qPCR-estimated concentration over-estimated the phage titre compared to the manufacturer's reported titre and both the plaque and colony assays ($p < 0.05$). Interestingly, the plaque assay also over-estimated the phage titre compared to the manufacturer's reported titre.

For the heterogeneous lysates comprised of both phagemid and helper, we observed the opposite behaviour. The plaque and colony assay-estimated titres were not significantly different from each other, but were both greater than the qPCR-estimated virion concentration. The qPCR-estimated helper concentration of the mini-(gfp) lysate was statistically significantly lower than the titre as estimated by colony assay ($p < 0.05$). In contrast, the qPCR-estimated helper concentration of the full-(gfp) lysate was statistically significantly lower than the titre as estimated by plaque assay ($p < 0.05$). It was also lower than the colony assay-estimated titre, but this was above the significance level ($0.05 < p < 0.1$).

Total virion concentration (including both helper and target phagemid) are reported in Table 3.6. Both helper and target phagemid species were highly abundant across all PEGylated samples, with no phage species present at concentrations below 10^8 gc/µL. No significant difference in overall yield was observed.

Table 3.6: Comparison of phage yield from mini and full phagemids

	Total virion concentration ($\times 10^8$ gc/µL)[1]	
Cassette	mini	full
cmv-gfp	38.49 ± 19.82	4.36 ± 1.32
cmv-luc	45.51 ± 13.72	25.87 ± 16.80

[1] Uncertainty reported as SEM$\times$2, $n = 3$.

Figure 3.11 summarizes the estimated virion concentration for each phagemid and helper phage per lysate in the top panel. In the bottom panel, the fraction that the target phagemid comprises of the overall lysate is shown as a percentage for each sample. The composition of phage lysates were observed to differ depending on the encoded cassette and whether the origin was split or intact. The phagemid fraction decreased between the mini and full lysates for each respective transgene: mini-(gfp) comprised 99.7% of the lysate while full-(gfp) comprised 73.2%. Similarly, mini-(luc) comprised 42.4% while full-(luc) comprised only 5.8%. Differences in composition were all statistically significant ($p < 0.05$), except for the difference between full-(gfp) and mini-(luc) ($0.05 < p < 0.1$).

M13KO7 appeared to preferentially package itself over the *cmv-luc* phagemids, but packaged *cmv-gfp* encoding phagemids over itself. A potential factor influencing the target phagemid:helper ratio may be cassette length or phagemid length. The *cmv-luc* transgene cassette is longer than the *cmv-gfp* cassette and, of course, the full phagemids are longer than their mini counterparts. The shortest phagemid, mini-(gfp) (2.5 knt), was rescued with near complete efficiency; helper phage comprised less than 1% of the final lysate. Similarly, a dramatic decrease in packaging efficiency was observed between the longest phagemid, full-(luc) (5.9 knt) and its mini-counterpart (mini-(luc) at 3.1 knt). However, while full-(gfp) (5 knt) is longer than mini-(luc), full-(gfp) was packaged with comparable or better efficiency ($0.05 < p < 0.1$).
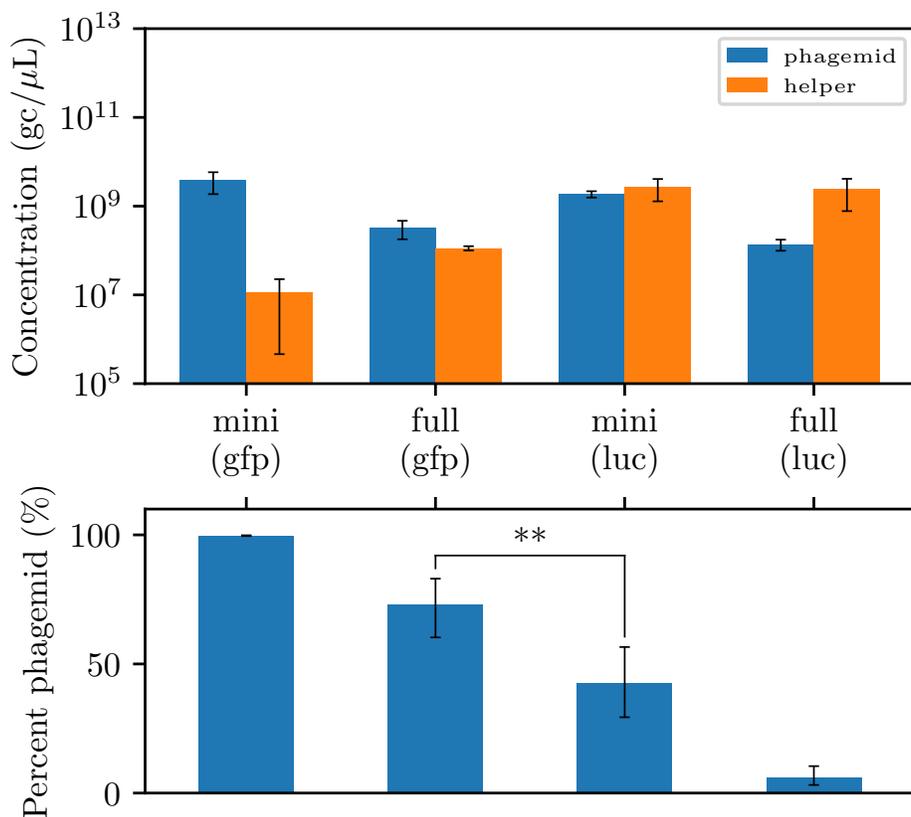
Figure 3.11: Composition of phage lysates encoding either *cmv-gfp* or *cmv-luc*. Above, the concentration (gc/µL) of each phage species (target phagemid or helper) in each lysate is presented. Below, the proportion of the target phagemid is shown as a percentage of the total phage population. Error bars represent SEM×2, $n = 3$. The ∗∗ above the bars indicates a difference at significance level $p < 0.1$.

3.3.3 Production efficiency of a longer miniphagemid precursor

We sought to investigate the potential link between phagemid production efficiency and phagemid length further by constructing a third set of phagemids (Figure 3.12). The insertion of a 2.5 kb fragment from pBR322 (which encodes *tet*) downstream of *cmv-gfp* in pM13ori2.cmvgfp generated pM13ori2.cmvgfp-tet (Figure 3.4). The same fragment was also inserted into the wild-type *ori* phagemid, pSW9, to generate pSW9-tet, as a full phagemid control (Figure 3.5). The length of this final *cmv-gfp-tet* cassette was approximately 3.9 kb. The predicted length of the recombinant gfp-tet miniphagemid would be approximately 4.4 knt, which is longer than the luc miniphagemid (3.1 knt). The length of full-(gfp-tet) would be approximately 6.9 knt, the longest of all the phagemid constructs in this work. Susceptible *E. coli* were transformed by a gfp-tet miniphagemid precursor or wild-type *ori* control, then infected with helper phage M13KO7 for production of RF DNA and phage. Of note, cultures of 10 mL were used in the preparation of RF, ssDNA, and phage, in contrast to 0.5 L previously (Figure 3.11).
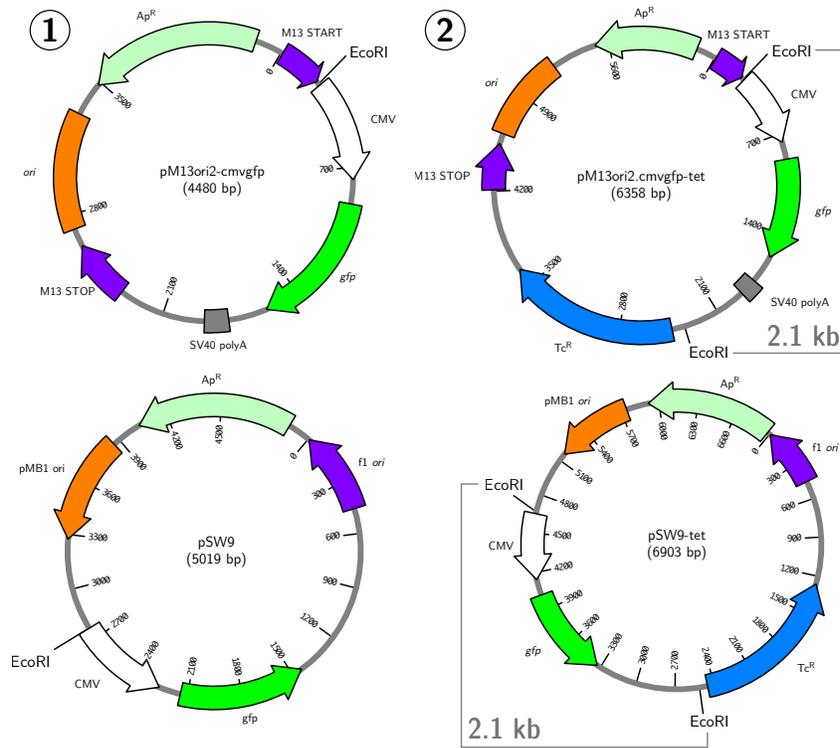
(A) Plasmid schematics
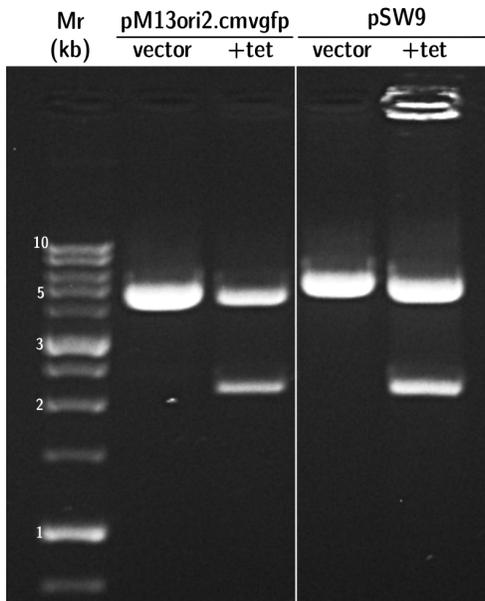
(B) EcoRI digest

Figure 3.12: A fragment from pBR322 is inserted into the *gfp* phagemid vectors. A-1) One
EcoRI cut site precedes the *cmv-gfp* cassette on backbone vectors, pM13ori2.cmvgfp and
pSW9. A-2) Insertion of the pBR322 fragment downstream of this cassette introduces
a second EcoRI cut site. B) EcoRI digestion of the new constructs (+tet) releases the
2.1 kb *cmv-gfp* fragment, leaving another 4.2 kb or 4.8 kb, respectively. Mr: 1 kb ladder
(FroggaBio).

Table 3.7: Comparison of phage yield across different phagemid cassettes

	Total virion concentration ($\times 10^8$ gc/µL)[1]	
Cassette	mini	full
cmv-gfp	5062.13 ± 6562.92	3372.98 ± 1276.59
cmv-luc	0.71 ± 0.68	207.71 ± 413.69
cmv-gfp-tet	860.52 ± 902.79	305.38 ± 313.02

[1] Uncertainty reported as SEM$\times 2$, $n = 3$.

Overall virion concentrations may be found in Table 3.7; intriguingly, this set exhibits far more variability than yields under earlier culture conditions (Table 3.6). The total virion concentration of full-(gfp) significantly increased during these culture conditions compared to the previous conditions, while the total concentration of mini-(luc) significantly decreased ($p < 0.05$). Overall yield was not found to differ significantly between lysates in this cohort.
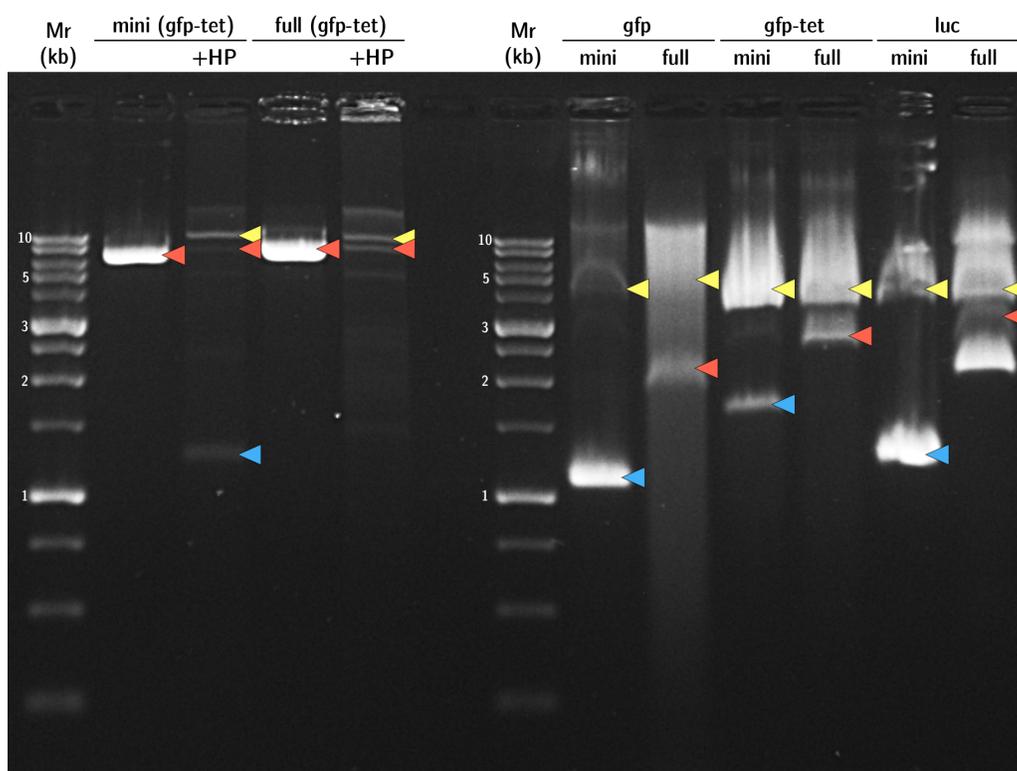
Figure 3.13: M13KO7 produces miniphagemids from *cmv-gfp-tet* precursors. On the left, double-stranded RFs isolated from bacterial pellets transformed with a (gfp-tet) precursor plasmid are linearized with PacI. On the right, single-stranded phagemid DNA was isolated from (gfp), (gfp-tet), and (luc) phage lysates. From left to right: 1) pM13ori2.cmvgfp-tet (6.4 kb) with/without helper M13KO7 2) pSW9-tet (6.9 kb) with/without helper M13KO7; 3) mini-(gfp), full-(gfp), 4) mini-(gfp-tet), full-(gfp-tet), 5) mini-(luc), full-(luc). Arrows indicate bands of interest: yellow (helper genome), red (plasmid or corresponding full phagemid), blue (recombinant RF: RFx, or miniphagemid). Mr: 1 kb ladder (FroggaBio).

Examination of PacI-linearized RF DNA on AGE revealed faint bands approximately around the sizes of the expected RFx, although bands of other sizes are also visible (Figure 3.13). Production of the miniphagemid mini-(gfp-tet) and its wild-type *ori* counterpart full-(gfp-tet) was observed by visualizing purified phagemid ssDNA alongside that of the mini and full phagemids for the other two cassettes, *cmv-gfp* and *cmv-luc*. Single-stranded DNA was observe to run at approximately half the size expected for its dsDNA counterpart, when referring to a double-stranded ladder. Helper phage genome was visible across all lanes for purified ssDNA, indicated by yellow arrows in Figure 3.13. Distinctive bands corresponding to the miniphagemids and full phagemids were also observed, marked by blue arrows. However, full-(luc) showed multiple bands, indicating some recombinant secondary species present in the lysate.

Quantification of (gfp), (luc), and (gfp-tet) phagemid lysates also revealed high ratios of mini-(gfp) and full-(gfp) phagemid to helper phage and much lower ratios of mini-(luc) and full-(luc), which is consistent with Figure 3.11 (Figure 3.14). The mini-(gfp) (96.9%) and full-(gfp) (93.0%) phagemid fractions were higher than both their (gfp-tet) and (luc) counterparts ($p < 0.05$). Under these culture conditions, no significant difference was observed between the mini and full phagemid fractions for *cmv-gfp* ($p > 0.05$). The mini-(gfp-tet) phagemid comprised 39.2% of its lysate and the full-(gfp-tet) comprised 41.5%. The *cmv-luc* phagemids again comprised the smallest fractions: 16.7% for mini-(luc) and 16.1% for full-(luc). Although the length of the *cmv-gfp-tet* cassette was much longer than *cmv-luc*, no significant difference was observed between their lysate compositions ($p > 0.05$).
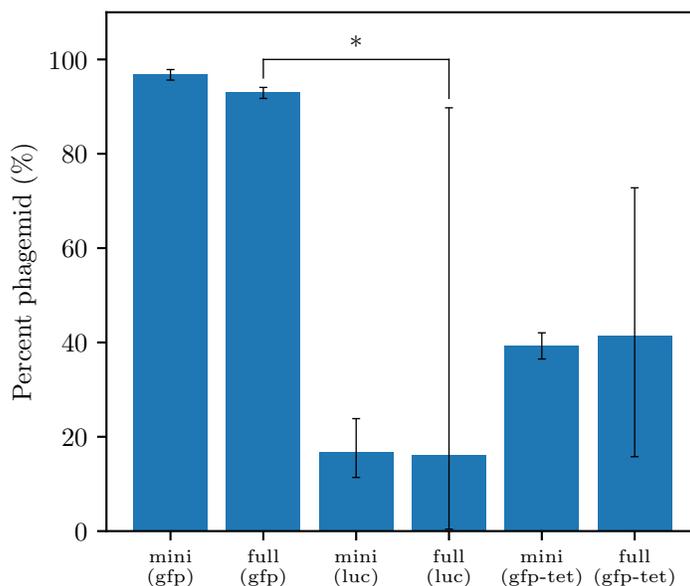
Figure 3.14: Composition of phage lysates encoding one of *cmv-gfp*, *cmv-luc*, or *cmv-gfp-tet*. The proportion of the target phagemid is shown as a percentage of the total phage population. Error bars represent SEM×2, $n = 3$. The ∗ above the bars indicates a difference at significance level $p < 0.05$.

3.3.4 Secondary structure analysis of miniphagemid DNA

Figure 3.15 compares the effect of either the cassette itself or the entire phagemid length to the resulting fraction of target phagemid in the lysate. A slight trend emerged when examining the effect of the length of the gene cassette only on the resulting phagemid fraction than phagemid length; no trend was apparent when comparing total phagemid length to phagemid fraction. Interestingly, a larger degree of spread was noticeable for larger phagemids. However, the *length* of the transgene cassette is not likely the key influencing factor as both *cmv-gfp-tet* phagemids are packaged with a comparable phagemid:helper ratio as the *cmv-luc* phagemids. Instead, the differences could be attributed to the *composition* of the transgene cassette.
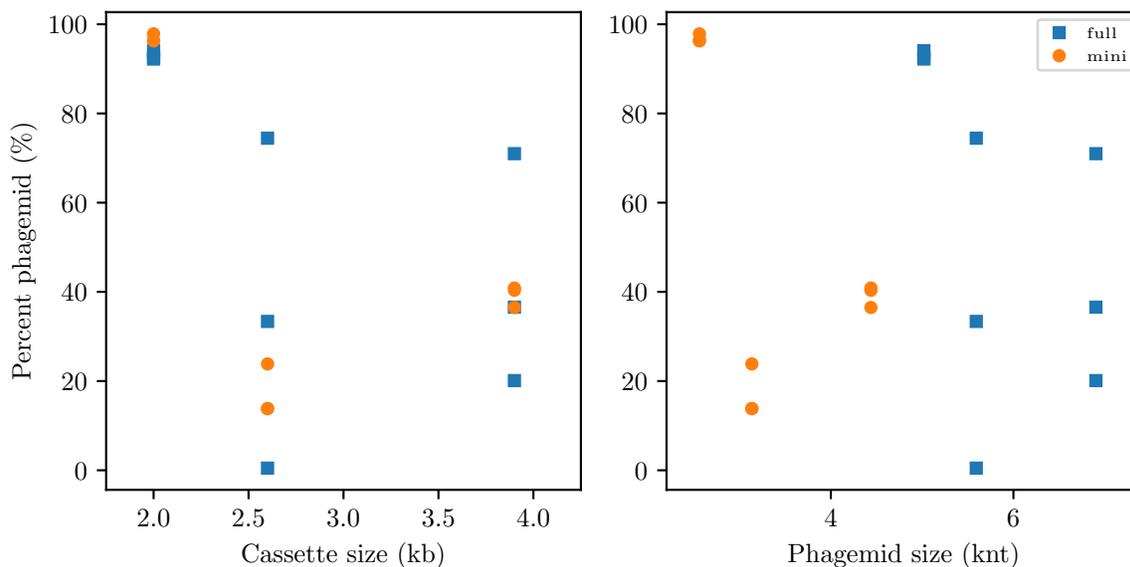
Figure 3.15: Target phagemid proportion over cassette or phagemid size. The proportion of target phagemid in the total lysate is examined over the size of the transgene cassette (left) and the phagemid (right). Three replicates per lysate are represented individually.

We examined the sequences of the transgenes and phagemids themselves to assess potential influences on phagemid replication and assembly. The differences in cassette length have already been described: *cmv-gfp* = 2.1 kb, *cmv-luc* = 2.6 kb, and *cmv-gfp-tet* = 3.9 kb. Miniphagemids share the same precursor backbone (pM13ori2) while their wild-type f1 *ori* counterparts share a pBluescript II KS+ backbone. The plasmid backbones have 49% and 50% GC content, respectively. The cassettes share many elements, including the same *cmv* promoter, polyA sequence, and SV40 enhancer elements. GC content per cassette was found to be 54%, 46%, 56%, respectively. The constructs do not exhibit major differences in base composition.

Next, we examined potential secondary structures of the single-stranded miniphagemids (Figure 3.16) in comparison to the genomes of wild-type M13 and helper phage M13KO7 (Figure 3.17). Secondary structures were predicted using RNAfold [411, 412]; the ensemble centroid secondary structures [424] are depicted. Colours represent base-pair probability; the colours of the unpaired

nucleotides indicate the probability of being unpaired. The length of the M13KO7 genome exceeded the maximum limits accepted by the RNAfold webserver, so this structure was generated locally without a base-pairing probability colour key. Sequences and dot-bracket notation of the centroid structures may be found in Appendix D. Unpaired regions are predicted with medium or low probability across all three miniphagemid structures. A large region of unpaired nucleotides is predicted in wild-type M13; smaller unpaired regions are predicted in M13KO7. Similarly, several large regions of unpaired nucleotides predicted in both mini-(gfp) and mini-(gfp-tet); these regions were identified as part of the gene sequences for *gfp* and *tet*. Small unpaired regions were observed in mini-(luc); these regions were identified as part of *luc*.

(A) mini-(gfp)

(B) mini-(luc)

(C) mini-(gfp-tet)
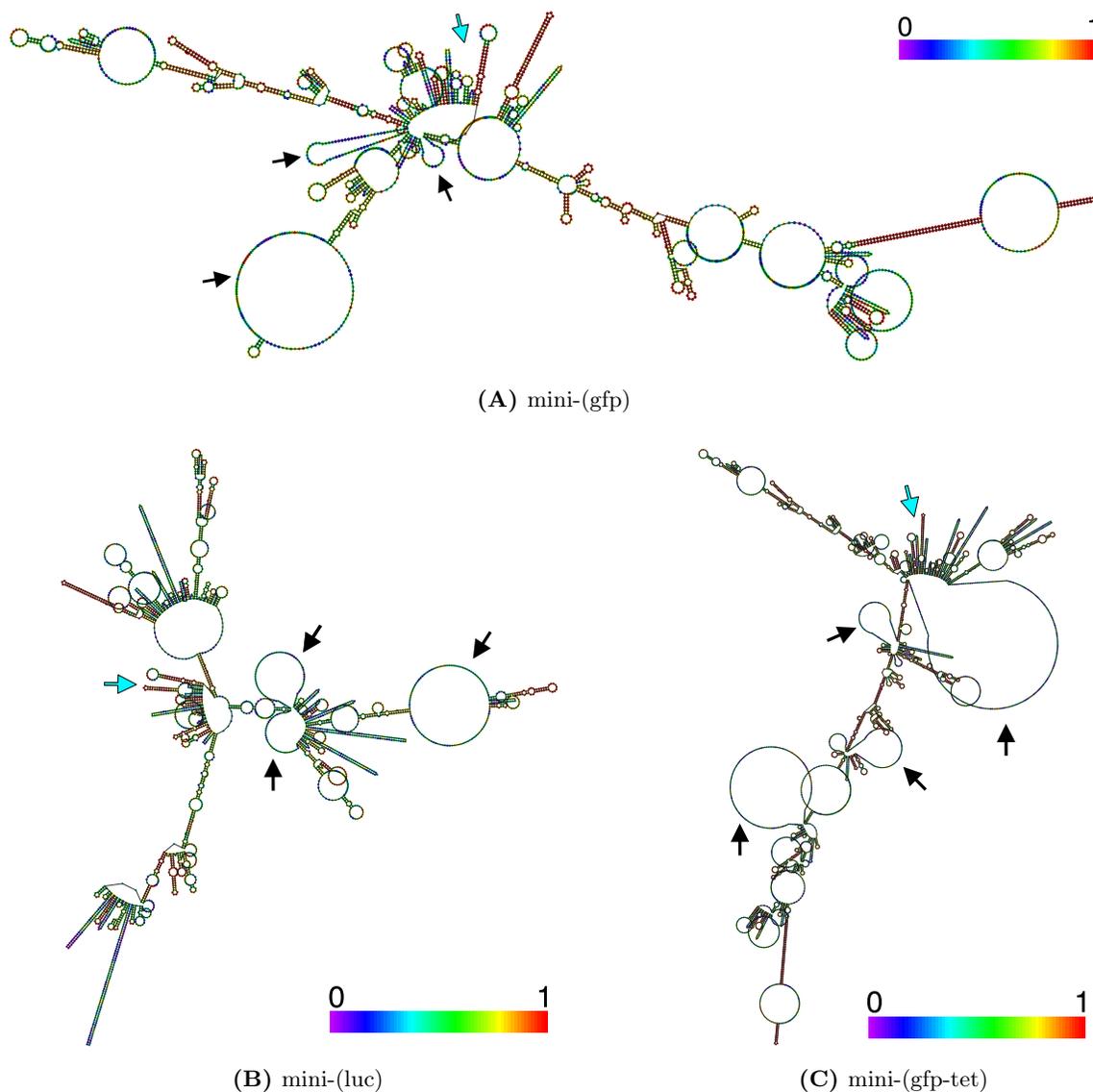
Figure 3.16: Predicted secondary structures of miniphagemids. The predicted centroid structure of the recombinant single-stranded phagemid molecules were produced using RNAfold [411]. Black arrows indicate some larger unpaired regions, while cyan arrows indicate the f1 *ori*. Colours denote base pairing probabilities.
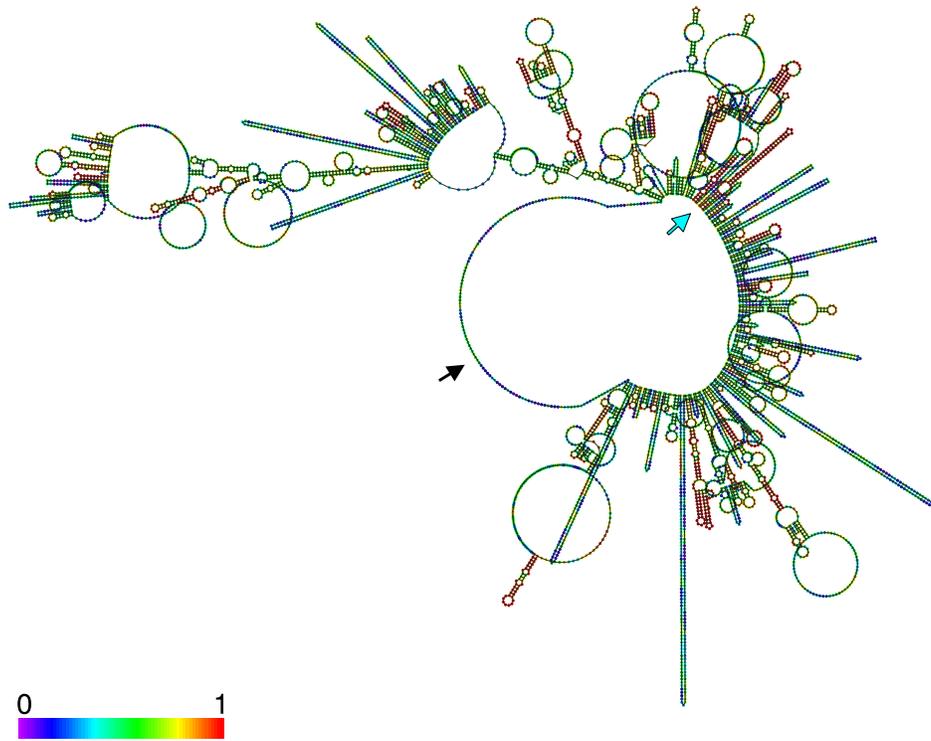
(A) M13
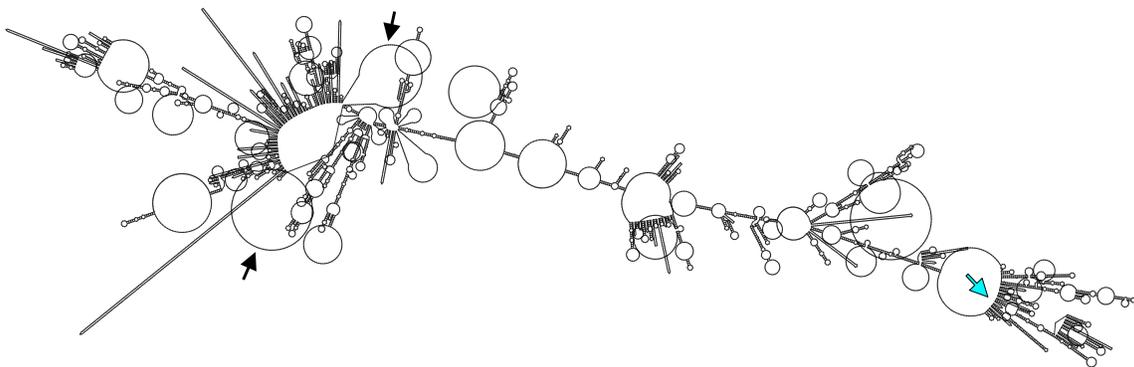
(B) M13KO7

Figure 3.17: Predicted secondary structures of M13 and M13KO7. The predicted centroid structures of wildtype M13 (B) and helper phage M13KO7 (B) were produced using RNAfold [411]. Black arrows indicate some larger unpaired regions, while cyan arrows indicate the f1 *ori*. Colours denote base pairing probabilities.

3.4   Discussion

In this chapter, we demonstrated that helper phage M13KO7 was able to act upon a split f1 *ori* present on a double-stranded plasmid precursor to generate a recombinant RFx molecule that is assembled into a miniphagemid phage particle. Through examination of multiple plasmid precursors carrying different transgene cassettes, we characterized the rescue efficiency of M13KO7 for these phagemids.

3.4.1   Miniphagemids can be produced from a split origin of replication

The visualization of the RFx molecule in M13KO7-infected cells carrying pM13ori2.cmvgfp demonstrated the ability of a helper phage to rescue phagemids from a split origin, which is consistent with observations reported in literature [208, 319]. Phage lysates were pre-treated with DNase prior to ssDNA extraction as free DNA, whether from lysed bacteria or malformed phage particles, contaminated agarose gels with smears. Upon examination of ssDNA through AGE, bands corresponding to the helper phage genome were observed in the lanes with our split origin phagemids, but not pBluescript II KS+. Our results initially suggested that the helper phage M13KO7 replicates the wild-type f1 *ori* preferentially over the split f1 *ori*. Quantification of the ssDNA from phage lysates showed that yield was not impacted by whether the phagemid carried a split or intact f1 *ori* (Figure 3.8) and instead, depended more on the choice of transgene cassette.

Visualization of ssDNA purified from phage lysates revealed some bands that could correspond to the precursor plasmid (Figures 3.8 and 3.13). Free DNA can be present in phage preparations due to cell lysis or malformed phage particles [409]. Phage preparations from which single-stranded DNA were extracted were not treated with DNase, which may explain the additional DNA species visible through AGE [409, 410].

Additionally, although we refer to our separated origin as "split", all structures necessary for phage amplification are present. Consequently, an intact packaging signal is also present on the phagemid precursor plasmid so they are available for assembly into progeny phage as well. Therefore, some encapsulated plasmid may be present in the lysate. However, we expect this fraction to be negligible as double-stranded DNA molecules are much less likely to interact with pV, a key controller in the M13 lifecycle, or the assembly proteins. Likewise, no bands corresponding to double-stranded pBluescript II KS+, wild-type M13, nor M13KO7 have been observed, although all will have a functional packaging signal (PS) (Figure 3.8).

3.4.2   Quantitative PCR is a robust method for phage quantification

Nucleic acid methods of quantification are commonly used for the detection and enumeration of many viral species in both research and healthcare diagnostics [409, 421, 425–427]. Quantitative PCR is a very common type of nucleic acid quantification, where the accumulation of amplicon can

be measured in real time by the use of DNA-intercalating fluorescent dyes such as SYBR Green, or probes such as TaqMan. Methods such as qPCR quantify based on the number of genome copies and would therefore include free DNA, degraded particles, or otherwise non-infective virions [428]. This is in contrast to quantification methods such as the plaque assay [429], which quantifies phage based their infectivity.

For enumeration of bacteriophage M13, SYBR Green was reported to be more sensitive than TaqMan [409]; since the viral abundance in our samples was unknown, we cautiously proceeded with SYBR Green. From our results, we saw that both helper and target phagemids were highly abundant; even the least abundant phage species was found to number at least 10^6 genome copies per microlitre (Table 3.6). In future similar studies, TaqMan PCR may prove preferable to SYBR as multiplexing would allow the simultaneous quantification of multiple phage species per run.

We wished to assess the accuracy of the qPCR method to quantify M13 lysates by comparing to plaque and colony assays (Figure 3.10). The plaque assay quantifies the number of viable M13 in a sample by their ability to infect susceptible host *E. coli* and generate zones of clearing, where infected *E. coli* grow more slowly. These turbid plaques indicate the number of plaque forming units in a sample. The colony assay operates in a similar manner: infected *E. coli* are only able to grow on LB agar supplemented with kanamycin if infected with M13KO7, which encodes a Km^R marker. Kanamycin-resistant colonies therefore indicate the number of colony forming units that that have taken up helper phage specifically.

Phage titres and qPCR virion concentrations can only be equal if one plaque or one colony forming unit arises from one infection event by a single virion particle. Hence, infection of a cell by a phagemid particle alone was not expected to result in plaque formation nor colony formation. A phage particle encapsulating a phagemid retains all factors necessary for cell entry. However, the phagemid does not contain any phage genes necessary for replication nor the Km^R marker so it would be unable to replicate on its own. Helper phage must also infect the same cell in order for phage gene or Km^R expression to occur. As such, plaques should arise from cells that were infected by both phagemid and helper as well as cells infected *only* by helper. Similarly, Km^R colonies should arise from cells infected by both phagemid and helper as well as those only infected by helper. We therefore expected that phage titres would be closer to qPCR concentrations of helper phage only, rather than the total virion concentrations.

High variability has been documented in infectivity-based assays, which arises from inherent differences in bacterial growth parameters, and variability between plate pouring and spreading techniques, among other factors [422, 423, 430]. Overall, there was generally good agreement between the plaque and colony assays in this study. The mean plaque assay titre of our control, M13KO7, was approximately 8-fold greater than the manufacturer's reported titre while the qPCR estimated concentration was about 40-fold or a magnitude greater. Relative to the plaque assay, the qPCR estimate was approximately 5-fold greater. Altogether, the fold discrepancies between the three methods are consistent with discrepancies reported in literature [409, 428].

The phagemid lysates were pre-treated with DNase prior to enumeration in order to reduce contaminating free bacterial DNA or other DNA not encapsulated within intact phage particles [409, 410]. For our heterogeneous phage populations, the infectivity-based assays overall resulted in higher estimated phage titres than qPCR. This was surprising, as the presence of residual non-infective virions was expected to result in higher phage concentration estimations through qPCR compared to assays based on infectivity. In the case of full-(gfp), the phage titre as estimated by plaque assay is approximately 10-fold greater than the helper concentration as estimated by qPCR, while it is almost 2000-fold greater for mini-(gfp). This was inconsistent with the results comparing titres of M13KO7 alone [409, 428]. Helper phage concentrations may be under-estimated in qPCR when in the presence of phagemid. Alternatively, the colony and plaque assays may over-estimate helper titre when in the presence of phagemid. The high variability of infectivity assays does preclude any conclusions without further investigation; it would be interesting to investigate both possibilities in future studies.

Both the plaque and colony assays are laborious procedures requiring 1–2 days before analysis [409, 422, 423]. In comparison, qPCR can differentiate between multiple phage species in a single experiment and requires only 2–6 h, but it is generally more expensive and requires more specialized equipment. Neither the plaque assay nor the colony assay can differentiate between different phage species within the same sample. Other validated methods for phage enumeration have been reviewed [422, 423], but many of these also require specialized equipment (ex: flow cytometry, NanoSight), are also laborious and time-consuming (ex: transmission electron microscopy), and not all can differentiate between heterogeneous phage populations. On the whole, the qPCR method for enumeration is highly sensitive, reproducible, and faster than the standard infectivity assays. For phages such as the miniphagemid vectors generated here, qPCR is furthermore a superior quantification method in the absence of a screening or selection marker as multiple viral species can be identified from the same lysate.

3.4.3 Miniphagemids are preferentially assembled over full phagemids

High phage yields were associated with both *cmv-luc* lysates initially (Table 3.6), although there was large variability between two culturing conditions (Tables 3.6 and 3.7). The major differences in these two sets of data may arise from the manner of preparation: they differed in culture volume (0.5 L versus 10 mL) and the former lysates were precipitated with PEG. M13 production can vary considerably depending on inherent variability of the host cell and culturing environment, and changes in culturing volume and culture vessel can impact yield [431]. Future phagemid production will also need to take this into account.

We hypothesized that the intact f1 *ori* on our wild-type phagemid controls (pSW9, pSW10, pSW9-tet) would be preferentially replicated and packaged over the helper phage genome. The split origin (pM13ori2.cmvgfp, pM13ori2.cmvluc, pM13ori2.cmvgfp-tet) was hypothesized to be less efficiently processed as the intact f1 *ori* first needs to be formed. Indeed, significant differences

were observed in the production of target phagemids by the helper phage M13KO7. Consistent with observations by Short et al. [319], however, the split *ori* was actually not observed to impact overall yields nor compromise phagemid production. In fact, full phagemids comprised a smaller fraction of lysates than their mini-counterparts (Figures 3.11 and 3.14). We observed that the helper phage preferentially assembled itself over a *cmv-luc*-encoding phagemid, regardless if the origin of the target phagemid was split or wild-type Figure 3.11. It was initially hypothesized that this was due to the increased overall phagemid length, as previous studies have observed a decrease in phage titre for longer phagemids [315]. Poor conversion of large phagemids (5–10 kb) from intracellular RF to extracellular phage-encapsulated DNA has been observed. To investigate this, a longer transgene cassette was constructed through the introduction of a pBR322 fragment into the *cmv-gfp* cassette, leading to a final cassette size of 3.9 kb and final phagemid size of 4.4 knt. If phagemid length did affect production efficiency, mini-(gfp-tet) and full-(gfp-tet) (6.9 knt) would be more poorly rescued than mini-(luc) (3.1 knt) and full-(luc) (5.9 knt).

The mini and full-(gfp) phagemid fractions were much higher than their (luc) and (gfp-tet) counterparts. No significant difference in lysate composition was observed between the (luc) and (gfp-tet) phagemids despite their large differences in phagemid length. Notably, greater variance in composition was observed for both full-(luc) and full-(gfp-tet) under the culturing conditions in this study. When examining the effect of cassette or the phagemid length on phage lysate composition, the cassette length may exert a greater effect than overall phagemid length although the effect is not strong (Figure 3.15). However, only three cassettes were examined. Hence, it may not be the *length* of the cassette, but rather the *composition* of transgene cassette. We did not find any significant differences in the overall base composition of the three transgene cassettes: as they all share the *cmv* promoter and other non-coding elements, the differences in resultant rescue efficiency may arise from the sequences of the coding regions (*gfp, luc* and *gfp, tet*). Indeed, examination of the predicted secondary structures of the three miniphagemids (gfp, luc, and gfp-tet) led to a new hypothesis.

Phage production is governed by interactions with three main phage proteins during key steps during the infection lifecycle: 1) pII interaction with the f1 *ori*, 2) pV coating of the ssDNA molecule, and 3) pI interaction with the packaging signal. A key step between M13 replication and phage particle assembly is pV coating of free ssDNA M13 intermediates. As pV preferentially binds single-stranded DNA, the formation of double-stranded domains may impact its ability to coat the molecule. We therefore examined possible secondary structure formation by the single-stranded miniphagemids. We chose to examine the entire miniphagemid structure itself instead of the transgenes in isolation in order to assess potential interactions between the transgenes and the phagemid backbones.

Prediction of nucleic secondary structure is more robust for RNA [432] than DNA, although many common and publicly available tools for RNA structure prediction also accept DNA sequences [411, 433] with some caveats. Firstly, only secondary structure of single-stranded nucleic acids can be

predicted. Secondly, structure prediction is less reliable and more computationally expensive for longer sequences. The ViennaRNA webserver accepts RNA (and DNA) sequences up to 7500 nt for some analyses [411, 434]. Predicted structures should be corroborated with empirical evidence and cannot be considered definitive. Nevertheless, the differences between the three predicted structures may offer some insight into how M13 phage proteins may interact with our different miniphagemids.

The centroid secondary structures were selected for comparison between the three miniphagemids for (gfp), (luc), and (gfp-tet). These structures were compared to the predicted centroid structure for wild-type M13 and M13KO7. Briefly, the centroid secondary structure is one that minimizes the base pairing distance over the entire Boltzmann ensemble of possible structures; it describes the "central tendency" of the ensemble [424]. The Boltzmann ensemble encapsulates all possible secondary structures, each one weighted by its probability [435]. The base pairing distance is a measure of how many base pairs differ between one predicted structure and the next; by minimizing this value, the ensemble centroid best represents all sampled structures [424, 436]. Because the ensemble centroid represents the entire ensemble, it is not equivalent to any single structure; thus, the centroid structure may not truly represent the molecular structure in reality.

Visual inspection of the predicted secondary structure indicates that the largest differences between each of the miniphagemids appear to be the unpaired regions. The centroid structure for wild-type M13 predicted a large unpaired region (Figure 3.17); some smaller unpaired regions were also in the centroid structure of M13KO7. Since there is a preference for pV binding to ssDNA, the presence of unpaired regions is likely necessary for efficient coating and sequestering after amplification. It is also well accepted that shorter phagemid sequences are rescued better than longer sequences [315].

Given that 1) predictive algorithms like RNAfold are optimized for folding RNA, 2) these algorithms are not as accurate as comparative sequence analysis with experimentally-confirmed structures, and 3) the centroid structure describes the central tendency of all possible structures and therefore may not accurately describe a real existing structure, we offer the following hypothesis. There is not likely not a stringent requirement for single-strandedness so long as f1 *ori* signals are accessible. As the ssDNA intermediates of the phagemid and M13KO7 compete for pV coating, *cmv-luc* phagemids may fare poorly because they are both long and tend towards more secondary structure in comparison to the helper phage. Single-stranded phagemids not maintained by pV coating will be converted back to duplex forms. The *cmv-gfp* and *cmv-gfp-tet* phagemids fare better, especially the *cmv-gfp* phagemid because of its short length and sufficient exposure of single-stranded DNA. We therefore postulate this is why they are more likely to be exported for assembly and extrusion as progeny phage. This aligns with the observations from Levinson et al. who also postulated the same bottleneck for large phagemids [315]: either insufficient phagemid ssDNA are stabilized for assembly or the phagemid ssDNA cannot compete with helper phage ssDNA for assembly resources.

Other groups have observed that regions within gIV upstream of the PS (roughly 86 bp of the tail end of gIV) may include signals that enhance genome packaging [315, 437], although other investigations of the f1 *ori* do not include this region [179, 180]. Hence, inclusion of this region in gIV in phagemids may improve their assembly. The helper phage M13KO7 was selected for its optimal production of phage progeny (exhibited through high ssDNA production) and high phagemid:helper packaging ratio [267]. Intriguingly, for longer phagemid particles, Levinson et al. found that using wild-type M13 as the helper led to the best phagemid:helper ratio [315]. Short et al. found the IR1 helper phage superior to M13KO7 for phagemid rescue from lambda ZAP, but concluded that the high efficiency of the split origin lent itself to the use of any helper [319]. The IR1 helper phage differs in a mutation within pII, which could enhance rolling circle amplification of split origin precursors. A combination of different helpers in tandem with different phagemids depending on phagemid length or other characteristics may reveal unpredicted synergies.

3.4.4   Summary

In this chapter, we demonstrated the replication and assembly of a miniphagemid from a precursor plasmid encoding a transgene of interest flanked by separated regions of the M13 f1 *ori*. We observed that assembly of a recombinant miniphagemid from a split f1 *ori* was superior to assembly of the same transgene cassette in a backbone containing a wild-type f1 *ori*. Overall, the high titres of M13 do compensate for low production of any of the phagemids under study. For example, although full-(luc) comprises only a small fraction of a PEGylated phage lysate, well over 10^8 virions are present.

For a given transgene, a shorter phagemid sequence may be assembled more efficiently; however, this effect appears to be transgene-dependent. The specific transgene sequence may interfere with phage replication and packaging to a greater extent than phagemid length. We propose that greater tendency towards secondary structures within a phagemid sequence may lead to an inability to efficiently convert from double-stranded RF to a single-stranded form, which precludes phage particle assembly and extrusion.

Chapter 4

Construction and characterization of helper phages derived from M13KO7

4.1 Introduction

In Chapter 3, we showed the successful assembly of miniphagemids, which do not contain any prokaryotic backbone aside from the short f1 functional origin (f1 *ori*). Phagemid replication and assembly into filamentous phage particles can only occur if phage proteins are present from another source, generally referred to as the helper. This helper phage must provide the full complement of phage genes required for replication, assembly, and extrusion [267, 315, 316]. In order to use virion-encapsulated miniphagemid for mammalian gene transfer, the incorporation of targeting peptides can facilitate cellular uptake in a tissue-specific manner. In this chapter, we construct a helper phage displaying a cell-specific targeting ligand, which would allow the production of phage particles targeting mammalian cells. As we have demonstrated how the phagemid can be optimized for mammalian gene delivery production, we now further investigate how the helper phage can be modified to improve phagemid rescue.

4.1.1 Phage display

Peptide fusion into the helper phage and/or phagemid allows for a diverse combination of peptide display, which is useful for extending phage function and peptide library screening [257]. In tandem with gene insertions in the phagemid, peptide display can be paired with gene transfer for a myriad of gene transfer applications. Present in five copies in the capsid, the coat protein pIII is the most common site for peptide display. Display on pIII can be accomplished in number of ways (Figure 4.1): if the only source of pIII is from the helper, this results in homogeneous, multivalent pIII display or a Type 3 display. If the phagemid encodes a competing gIII, a pool of chimeric pIII molecules compete for assembly into the phage particle, leading to heterogeneous display (Type 3+3). Translational fusion of peptides to both the N- and C-termini of pIII has

been demonstrated [249, 257, 276]. The coat protein pIII mediates infectivity while it is part of the phage capsid structure [122]. After infection and coat disassembly, pIII units are localized to the cell membrane where they participate in phage assembly by capping off extruding phage particles. Protein fusions to pIII, particularly to its N-terminus, must therefore avoid disruption of its functions outside and within the cell. Despite these limitations, pIII remains a popular site for peptide display, due to the accessibility of the protein to the environment and its ability to accommodate large fusions [249, 257].
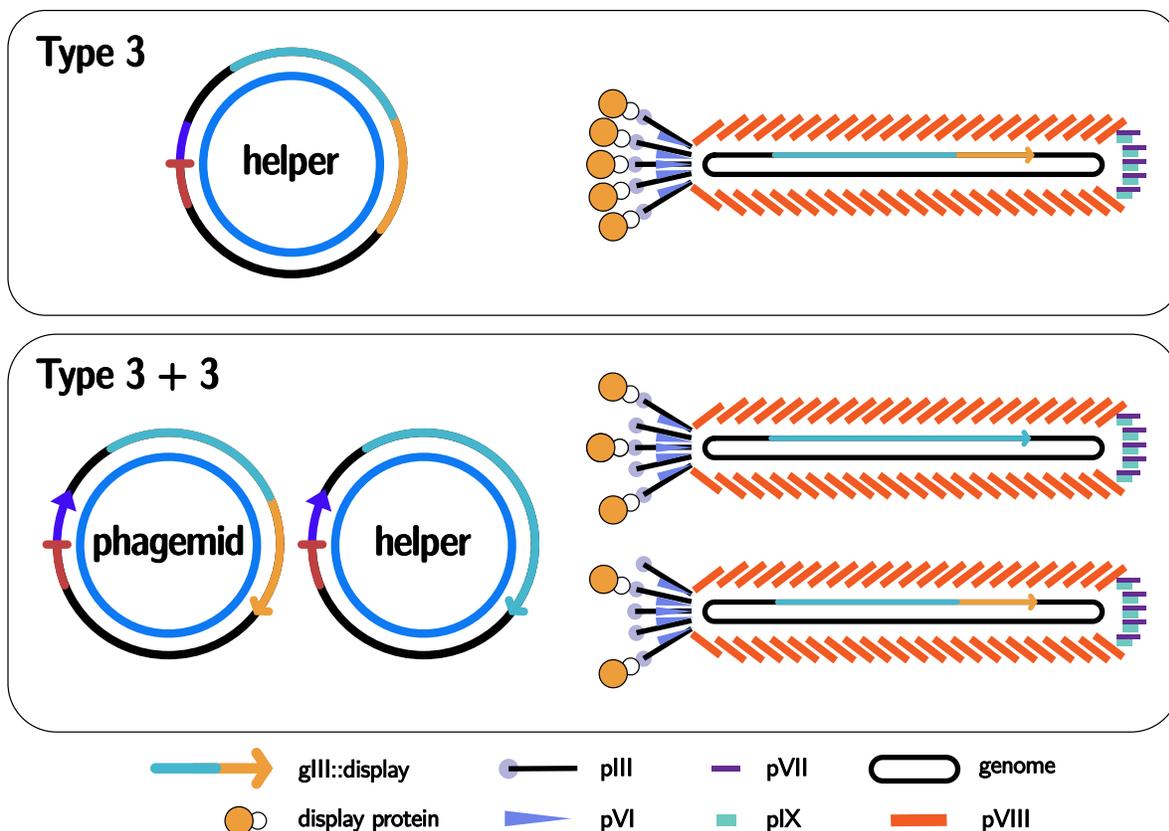
Figure 4.1: Display systems of M13. This schematic shows two M13 display systems using pIII as a model. In Type 3 display, only one copy of gIII is available, leading to homogeneous display of the fusion peptide. In Type 3+3, two trans copies of gIII are available, instead leading to heterogeneous display.

4.1.2 Helper phage manipulation

While helper phage is necessary to supply phage structural, replicative, and assembly proteins, both a helper phage and phagemid contain the full f1 *ori*. As such, both helper and phagemid are substrates for the phage replicative protein, pII, which initiates and terminates rolling circle amplification during the phage lifecycle [178]. Interference by the helper generally reduces the overall yield of phagemid particles [267, 315] and can overtake phage progeny production. Work on reducing interference from the helper has proven fruitful, giving rise to an array of different

helper phages [316], including the helper phage M13KO7 [267]. The insertion of the transposon Tn903 into the AvaI site of M13mp1 interrupts the f1 *ori*, creating an origin that is less efficiently replicated by pII. To counteract this, Tn903 carries the plasmid p15a origin of replication, enabling M13KO7 to replicate without complete dependence on pII. As a helper, M13KO7 provides phage proteins necessary for high amplification and production of phagemid progeny while contributing less of its own progeny.

The f1 origin of another helper phage, M13KE, is interrupted by the insertion of *lacZα*, preferentially amplifying a phagemid with a wild-type f1 *ori* if one is present. Moreover, M13KE easily facilitates N-terminal fusions to pIII through the incorporation of two endonuclease target sites within the N-terminus of gIII [438]. The polylinker within lacZα also accommodates direct insertion of genetic sequences. M13KE is therefore suitable as either a Type 3 (one copy of gIII) or Type 3+3 (two trans copies of gIII) display system in tandem with a phagemid (Figure 4.1). It can also support a Type 33 display (two cis copies of gIII).

4.1.3 *Escherichia coli* host strains for the propagation of M13

A number of common laboratory *E. coli* strains have been frequently used in the propagation of M13 and production of single-stranded (ss) DNA [439, 440]. They include strains from the Messing laboratory [308, 312] (for example: JM109), XL1-Blue [439, 441], TG1 [439], and NM522 [442].

Messing and colleagues constructed a number of *E. coli* strains for the propagation of M13 cloning vectors [308, 312]. Of them, JM109 is a recombination-deficient (*recA*) strain commonly used for plasmid purification. XL1-Blue is another common *recA* host for M13 [261, 319, 443] and often used in applications requiring stable propagation of plasmids. They both have very similar published genotypes (Table 4.1); the biggest difference is a Tc^R marker through Tn10 insertion in the F′ episome in XL1-Blue. Both carry the *gyrA96* mutation, where *gyrA* encodes for the GyrA subunit of DNA gyrase. DNA gyrase has been implicated in homology-dependent, RecA-independent deletions; the *gyrA96* mutation gives resistance to nalidixic acid and stabilizes the cloning of repetitive DNA [444] without adversely impact supercoiling of plasmid DNA [445].

In contrast, NM522 and TG1 are both recombination-competent strains often used for the propagation of filamentous phage. NM522 is *hsd*Δ5 derivative of 71-18, a *(lac-pro)* strain of *E. coli* [446], commonly used for the production of single-stranded DNA. TG1 is commercially available from a number of companies (Lucigen, Agilent, Antibody Design Laboratories, GoldBio) for the generation of phage display libraries [447]. It is also closely related to CSH50, a fairly well-known strain which used to be distributed by the Cold Spring Harbor Laboratory to research laboratories [439, 448] and was frequently used in M13 studies.

4.1.4 Rationale and hypothesis

In Chapter 3, we characterized the ability of helper phage M13KO7 to rescue miniaturized phagemids from precursor plasmids with a split origin in comparison to full phagemids with wild-type origins. Here, we constructed a derivative of M13KO7 for Type 3 display of a mammalian ligand, the epidermal growth factor (EGF), for use in mammalian cell targeting. When used as a cell-specific targeting ligand, EGF recognition by its cognate receptor has been shown to enable phage internalization into mammalian cells, which can then enable phage-mediated gene transfer in a tissue-specific manner [347]. Mono- and multivalent display of EGF has previously been demonstrated on M13 [347, 449–451], although EGF-displaying helper rescue of phagemid has not been thoroughly investigated. The effect of EGF on phage infectivity has also not been deeply investigated. However, if phage are able to infect, then pIII::EGF fusions are not expected to adversely impact replication. We hypothesized that display of EGF on the helper would not impact phagemid:helper ratios for either full or mini phagemids.

Additionally, we describe the construction of another derivative of M13KO7 that deletes the packaging signal (PS) directly downstream of gIV. Deletion of the PS would prohibit phage genome interaction with phage assembly complexes, thereby preventing extrusion of the helper; the loss of PS has been linked to reduced phage particle production [452]. We hypothesized that the absence of a PS on the helper should lead to preferential extrusion of phagemid progeny over the helper and improve the phagemid:helper ratio.

Strain selection. Based on recommendations from Green and Sambrook [439], we selected six common or commercially available laboratory strains to assess phagemid rescue by M13KO7 and M13SW8. 1) JM109 [308] and 2) XL1-Blue [441] are often used for plasmid and M13 propagation. 3) ER2738 [442, 453] is a derivative of NM522 with a Tc^R marker present on its F′ episome [442] and commercially sold as a host for M13 phage display (New England BioLabs, Lucigen, Amid Biosciences). 4) The NEB Turbo strain from New England BioLabs is notable for its rapid growth rate and good plasmid yield; its genotype bears resemblance to TG1, which was also recognized for its rapid growth [447]. 5) Stbl4 is derived from Stbl2 [454], which itself is a derivative of JM109; both strains were developed for applications with large repeats of DNA or large plasmids that are normally unstable in *E. coli* strains [455]. 6) DH5α [456] was included as our F$^-$ control.

DH5α, and Stbl4 are both recombination-deficient, like JM109 and XL1-Blue. Conversely, ER2738 and NEB Turbo are both fast-growing *recA*$^+$ strains. Specifically, New England BioLabs reports NEB Turbo colony growth in as little as 6 h and plasmid preparation after 4 h [457]. Similarly to XL1-Blue, both ER2738 and Stbl4 are resistant to tetracycline by way of a Tn10 insertion in the F′ episome. As the helper phages, phagemids, and miniphagemid replicative factor (RF) species all share the same f1 *ori* (∼500 bp), recombination events may occur; therefore, Stbl4 was investigated for potentially improved plasmid/phagemid maintenance.

4.2 Materials and methods

Strains and vectors. *E. coli* K-12 strains were used in the generation of all phage and plasmid constructs. *E. coli* JM109 was the host for plasmid amplification and purification unless otherwise noted. Bacterial strains are listed in Table 4.1, plasmids in Table 4.2, and phages in Table 4.3.

Media and buffers. All media and buffers were prepared according to recipes in Appendix A, unless otherwise specified. Bacterial strains were cultured in Luria–Bertani (LB) liquid medium, while plaque assays were carried out in top agar. Phage lysates were purified and stored in Tris-NaCl (TN) buffer and concentrated with polyethylene glycol (PEG), as indicated.

Construction of an M13KO7 derivative capable of display on pIII. M13KE has endonuclease target sites in gIII that simplify N-terminal peptide fusions [438], while M13KO7 contains the plasmid p15a *ori* and a Km^R marker [267]. The gIII from M13KE was inserted into the helper phage M13KO7 using Gibson assembly in order to generate a helper phage that could easily take on N-terminal pIII fusions while retaining the p15a *ori*. In M13KE, gIII was amplified using primers, gIII-F and gIII-R (Table 4.4). The M13KO7 backbone was amplified with primers extending outward from its gIII region: M13KO7-F and M13KO7-R. PCR products were assembled using NEBuilder HiFi DNA Assembly Master Mix (New England BioLabs) then transformed, and selected on LB agar supplemented with 25 μg kanamycin. The purified final construct, M13SW7 (Figure 4.2A), was transformed into an F^+ *E. coli* strain for further characterization.

Table 4.1: Bacterial strains used in the construction of helper phages

Strain	Genotype	Source
JM109	F′ traD36 proAB$^+$ lacI^q lacZΔM15/ Δ(lac-proAB) endA1 glnV44 thi-1 e14$^-$ recA1 gyrA96 relA1 hsdR17	NEB
DH5α	F$^-$ ϕ80 Δ(lacZYA-argF)U169 endA1 thi-1 recA1 gyrA96 relA1 hsdR17(r$_K^-$, m$_K^+$) phoA supE44	CGSC #12384
XL1-Blue	F′ traD36 proAB$^+$ lacI^q lacZΔM15 Tn10/ lac endA1 glnV44 thi-1 recA1 gyrA96 relA1 hsdR17	Agilent Technologies
Stbl4	F′ proAB$^+$ lacIq lacZΔM15 Tn10/ endA1 glnV44 thi-1 recA1 gyrA96 relA1 Δ(lac-proAB) mcrA Δ(mcrBC-hsdRMS-mrr) λ$^-$ gal	Invitrogen
ER2738	F′ zzf::Tn10(TcR) proAB$^+$ lacI^q lacZΔM15/ Δ(lac-proAB) thi-1 glnV44, Δ(hsdS-mcrB)5, fhuA2	NEB
NEB Turbo	F′ traD36 proAB$^+$ lacI^q lacZΔM15/ Δ(lac-proAB) glnV44 thi-1 galE15 galK16 R(zgb-210::Tn10)TetS endA1 fhuA2 Δ(mcrB-hsdSM)5(r$_K^-$, m$_K^-$)	NEB

NEB: New England BioLabs CGSC: Coli Genetics Stock Center

Table 4.2: Plasmids used in the construction of helper phages

Plasmid	Genotype	Source
pGL2-SS-CMV-GFP-SS	pGL2-Promoter (GN : X65326.2) *cmv-gfp* replaced *SV40-luc*, Ap^R	gift, Mediphage Bioceuticals Inc.
pGL3-CMV	pGL3-Basic (GN: U47295.2) *cmv* inserted in BglII-HindIII, Ap^R	gift, Dr. N. Oviedo
pBluescript II KS+	wild-type f1 ori, pUC ori, Ap^R	[405]
pPL451-gpD::egf	*egf* in translation fusion to gpD from λ in pPL451 (GN: AB248919.1)	[458]
pSW9	pBluescript II KS+, *cmv-gfp* inserted in BamHI-EcoRI, ApR	Chapter 3
pSW10	pBluescript II KS+, *cmv-luc* inserted in KpnI, ApR	Chapter 3
pM13ori2.cmvgfp	pM13ori2, *cmv-gfp* inserted in EcoRI-PacI	Chapter 3
pM13ori2.cmvluc	pM13ori2, *cmv-luc* inserted in EcoRI-KpnI	Chapter 3

GN: GenBank Accession No.

Table 4.3: Phages used in the construction of helper phages

Phage	Genotype	Source
M13	wild-type	CGSC #13587
M13KO7	M13, p15a *ori*, Tn903 (Km^R)	NEB
M13KE	M13, KpnI and EagI target sites in gIII	NEB
M13SW7	M13KO7, gIII from M13KE	This study
M13SW7-EGF	M13SW7, *egf* from pPL451-gpD::egf	This study
M13SW8	M13KO7, PS removed, Km^R	This study
full-(gfp)	*cmv-gfp*, from precursor pSW9, Ap^R	Chapter 3
full$_{egf}$-(gfp)	*cmv-gfp*, from precursor pSW9, pIII::EGF, Ap^R	This study
mini-(gfp)	*cmv-gfp*, from precursor pM13ori2.cmvgfp	Chapter 3
mini$_{egf}$-(gfp)	*cmv-gfp*, from precursor pM13ori2.cmvgfp, pIII::EGF	This study
full-(luc)	from precursor pSW10, Ap^R	Chapter 3
full$_{egf}$-(luc)	from precursor pSW10, pIII::EGF, Ap^R	This study
mini-(luc)	from precursor pM13ori2.cmvluc, Ap^R	Chapter 3
mini$_{egf}$-(luc)	from precursor pM13ori2.cmvluc, pIII::EGF, Ap^R	This study

CGSC: Coli Genetics Stock Center   NEB: New England BioLabs

(A) M13SW7

(B) M13SW7-EGF
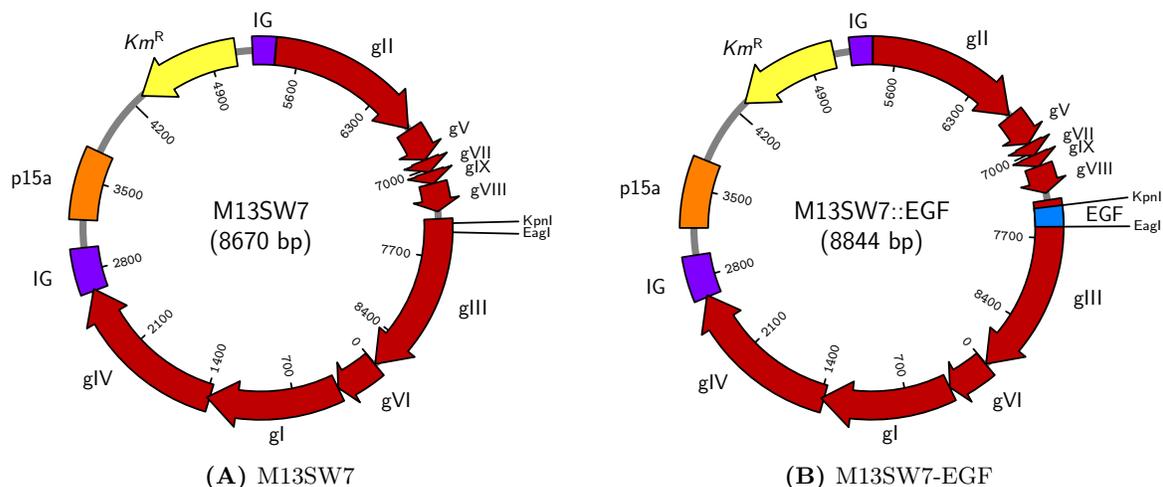
Figure 4.2: Maps of modified display helper phages. The endonuclease sites for insertion of *egf::GGGS* (KpnI and EagI) are indicated.

Table 4.4: Primers for phage construction

Primer	Amplicon	Sequence (5′ – 3′)[1]
gIII-F	gIII	TTTTTTTGGAGATTTTCAACGTGAAAAAATTATTATTCGCA ATTCC
gIII-R	gIII	CCCAAAAGAACTGGCATGATTTAAGACTCCTTATTACGCAG TATG
M13KO7-F	M13KO7	GTTGAAAATCTCCAAAAAAAAGGC
M13KO7-R	M13KO7	TCATGCCAGTTCTTTTGGG
KpnI-egf-F	*egf*	<u>*GGTACCTTTCTATTCT*</u>CACTCTAATAGTGACTCTGAATGTC CC[2]
EagI-egf-R	*egf*	<u>CGGCCGA*AGAACCACCACC*</u>GCGCAGTTCCCACCAC[3]
gIV-R	gIV	<u>*GGCT*</u>CGCTACAGGGCGCGTAC[4]
Tn903-F	Tn903	<u>*ATAATGGC*TAATCCCACAGGCCGCCAGTTCCGCTGGCGGC ATTTT</u>GCCAGGAAGATACTTAACAGGG[4]

[1] Underlined nucleotides indicate primer overhang. [2] Italicized nucleotides indicate gIII leader sequence.
[3] Italicized nucleotides indicate the GGGS linker. [4] Italicized nucleotides indicate spacer.

Construction of EGF-displaying phages. EGF was inserted in the N-terminus of pIII after the peptide leader sequence, via a GGGS (Gly-Gly-Gly-Ser) linker, commonly used to join components of fusion proteins [459]. Primers were designed to amplify *egf* (Table 4.4) from pPL451-gpD-EGF [458, 460]. In the forward primer (KpnI-egf-F), the gIII leader sequence and KpnI target site were part of the 5′ primer overhang. In the reverse primer (EagI-egf-R), the linker and the EagI target site were part of the 3′ overhang. The amplified PCR fragment and the M13SW7 genome were digested with KpnI and EagI (New England BioLabs). Purified digested fragments were ligated with T4 DNA ligase (New England BioLabs), transformed, then selected on LB agar supplemented with kanamycin to generate M13SW7-EGF (Figure 4.2B).

Phage amplification and purification. First, a fresh colony was inoculated in 2X YT supplemented with $MgSO_4$, and incubated at 37 °C with agitation until slightly turbid (0.05 < A_{600} < 0.4). Helper phage was then added to a titre of approximately 10^8 PFU/mL and returned to the same conditions for another 1.5 h. The culture was then supplemented with kanamycin (70 µg/mL) and returned to the same conditions overnight. The next day, the culture was centrifuged (8000 × g, 10 min) to separate the bacterial pellet from the supernatant.

The supernatant containing the phage lysate was further purified through a 0.45 µm filter to remove residual bacterial debris. Filtered lysate was concentrated through precipitation with PEG: 1/5 of the lysate volume was added in PEG, and the mixture was incubated at 4 °C for at least 2 h. The lysate was then centrifuged at 12 000 × g for 15 min at 4 °C to separate the PEGylated phage (pellet) from media (supernatant). The pellet was then re-suspended in a smaller volume of ice-cold TN buffer and a second PEG precipitation was carried out to further concentrate the phage. PEG-precipitated phage suspensions were treated with DNase I (Promega, Madison, USA) to remove extraneous phage or bacterial DNA [409, 410], then stored at 4 °C.

Purification of the double-stranded RF and single-stranded miniphagemid DNA. Episomal double-stranded (ds) RF were extracted from a cell pellet with the Monarch Plasmid Miniprep Kit (New England BioLabs). Phagemid DNA were extracted from the phage lysate through phenol-chloroform extraction. Phenol was added to PEG-concentrated phage lysate (1:1, v/v) and mixed by vortexing. After centrifugation for 5 min at 4 °C, the top aqueous layer was extracted. This layer was extracted again with an equivalent volume of phenol:chloroform twice, and chloroform once. Finally, the phagemid ssDNA was precipitated overnight with 100% ethanol at −80 °C. The precipitated DNA was washed with 70% ethanol, dried, and re-suspended in RNase and DNase-free water.

Quantification of phagemid and RF DNA. DNA was analyzed using the NanoDrop 2000 to determine concentration and purity. Extracted RFs were linearized with BamHI and visualized via agarose gel electrophoresis (AGE), while phagemid ssDNA was visualized directly without digestion.

Dot blot of EGF-displaying helper phage. To qualitatively confirm EGF display, a dot blot was performed with M13SW7-EGF and purified recombinant EGF (Thermo Fisher Scientific). Serial dilutions of recombinant EGF and M13SW7-EGF were spotted onto nitrocellulose membrane and left to dry. The membrane was then blocked with 4% bovine serum albumin (BSA) buffer. Anti-EGF monoclonal mouse antibody (Thermo Fisher Scientific) was applied, followed by a secondary horseradish peroxidase (HRP)-conjugated anti-mouse antibody. The blot was visualized with a colorimetric HRP substrate (Thermo Fisher Scientific).

ELISA of EGF displaying helper phage. To further confirm EGF display, a semi-quantiative enzyme-linked immunosorbent assay (ELISA) was performed, also with M13SW7-EGF and purified recombinant EGF. 96-well plates were coated with capture anti-EGF monoclonal mouse antibody and blocked with 4% BSA buffer. Serial dilutions of recombinant EGF and M13SW7-EGF were applied to the coated wells. Detection anti-EGF and anti-M13 antibodies were applied. Following the addition of streptavadin-HRP and HRP substrate (Thermo Fisher Scientific), absorbances were measured at 450 nm.

Construction of a self-packaging deficient helper phage. The PS immediately downstream of gIV was deleted in M13KO7. Since the PS functions as the gIV transcriptional terminator, a short terminator sequence was incorporated in its stead [461]. The well-characterized Rho-independent *trp* terminator [462, 463] was selected based on its short length as it was easily integrated into a primer overhang. The genome of M13KO7 was amplified using primers starting at the 3′ end of gIV, gIV-R, and beginning within the Tn903 fragment, Tn903-F (Table 4.4). The amplified fragment was purified and ligated with T4 DNA ligase (New England BioLabs), transformed into DH5α, then selected on LB agar supplemented with kanamycin. The final construct, M13SW8 (Figure 4.3), was transformed into an F$^+$ *E. coli* strain for further characterization.
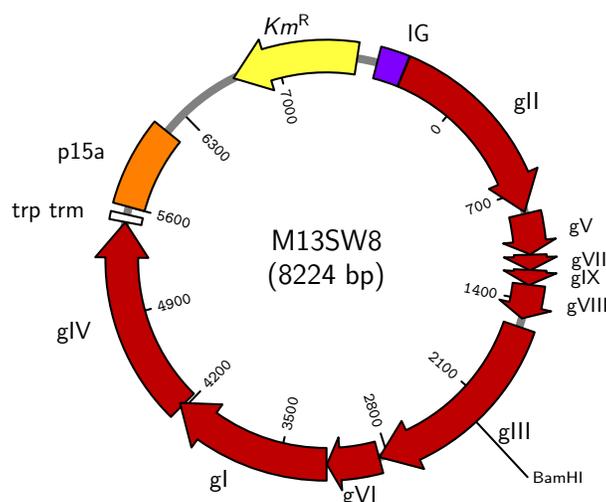
Figure 4.3: Map of modified non-packaging helper phage. Helper phage M13SW8 has the *trp* terminator instead of its original terminator and PS downstream of gIV.

Assessment of helper phage-assembled phagemid viability through infectivity assays.
Both infectivity assays were conducted using ER2738, a fast-growing F$^+$ strain that is recommended
for the production of filamentous phage (New England BioLabs). To quantify the degree of helper
self-packaging, plaque assays were performed comparing the infectivity of M13SW8 and M13KO7.
First, the appropriate dilution to generate countable (10–300) plaques was determined through
a spot plate. Briefly, 200 µL of early log-phase *E. coli* ER2738 was mixed with 3 mL top agar
supplemented with 5 mM MgSO$_4$, then poured on a pre-warmed LB agar plate. Separately, phage
lysates were serially diluted in TN buffer, then 5 µL of each phage lysate dilution were spotted
onto the top agar. Plates were incubated overnight at 37 °C and analyzed the next day. The lowest
dilution that generated distinct plaques on the bacterial lawn was selected to generate full plates
for plaque assay. For the full plaque assay, the same procedure was followed except that the entire
volume of diluted phage was added directly to the volume of bacterial culture prior to the addition
of top agar. The titre was expressed as plaque forming units (PFU) per millilitre.

A colony assay was conducted to quantify the ability of infectious helper phage to confer antibiotic
resistance. Cells were first transformed by either the helper M13KO7 or M13SW8. They were
then also transformed with pBluescript II KS+, which carries the f1 origin and confers Ap^R. A
culture of cells with both the helper phage and target phagemid vector was grown overnight in
2X YT supplemented with MgSO$_4$, kanamycin, and ampicillin. Phage lysates were purified from
these cultures the next day and concentrated through PEG precipitation. These phage lysates
were incubated with early log-phase *E. coli* ER2738 (A$_{600}$ $\sim$ 0.4) and spotted on LB agar plates
supplemented with ampicillin, kanamycin, or both. For full plates, the same procedure was followed
except that the entire dilution volume (200 µL) was spread on an LB agar plate supplemented
with ampicillin, kanamycin, or both. The titre was expressed as colony forming units (CFU) per
millilitre.

Comparison of helper phage activity across *E. coli* strains. A range of *E. coli* host
backgrounds were examined for their effect on helper phage productivity: JM109, XL1-Blue,
STLB4, NEB Turbo, ER2738, and DH5α. Strains were transformed with one of two kanamycin-
resistant helper phages: M13KO7 (PS present), or M13SW8 (PS deleted). Lysogens were then
transformed with one of two phagemid precursor plasmids: pSW9 (intact *ori*), or pM13ori2.cmvgfp
(split *ori*). To produce phagemid particles, strains harbouring both helper and precursor plasmid
were inoculated into 2X YT liquid medium supplemented with kanamycin and ampicillin. Cultures
were incubated overnight at 37 °C with agitation.

Quantification of phage species. The populations of helper and recombinant phagemids
within each phage suspension were quantified through SYBR Green quantitative PCR (qPCR).
Calibration curves for quantifying phage were constructed using external standards: M13KE RF
(7222 bp; New England BioLabs) for helper phage, pGL2-SS-CMV-GFP-SS (5257 bp; Mediphage
Bioceuticals) for *gfp*-encoding target phage, and pGL3-CMV (5678 bp) for *luc*-encoding target

phage. Primers specific to a region within gV of M13KO7 [409] were used to amplify helper phage. Primers targeting a region within *gfp* or *luc* were designed to amplify *gfp*- or *luc*-encoding phage (Table 4.5). To generate each calibration curve, 10-fold serial dilutions of each external control were prepared (10^{-1} to 10^{-8}) as template for the polymerase chain reaction (PCR) reaction. To prepare the phage particles for PCR, the lysates were denatured by heat for $100\,^\circ$C for $15\,$min to isolate phage DNA [409]. Ten-fold serial dilutions of each cleared lysate were prepared as template for the PCR reaction. As the phage lysates were of unknown concentration, dilution series of each were amplified through qPCR. For PEGylated lysates, dilutions in the range 10^{-3}–10^{-5} were examined, while for non-PEGylated lysates, 10^{-1}–10^{-3}.

Each $10\,$µL PCR reaction was prepared using $5\,$µL of PowerUp SYBR Green Mix (Thermo Fisher Scientific, Waltham, USA), $1\,$µL each of 500 nM primer (forward and reverse), $2\,$µL of template, and $1\,$µL of dH$_2$O. PCR cycling conditions were as follows: $50\,^\circ$C for $2\,$min, $95\,^\circ$C for $2\,$min, followed by 40 cycles at $95\,^\circ$C for $15\,$s and $60\,^\circ$C for $1\,$min. Next, the melt curve was set for 1 cycle at $95\,^\circ$C for $15\,$s, $60\,^\circ$C for $1\,$min, and $95\,^\circ$C for $15\,$s. PCR reactions were run in triplicate on the StepOne Plus Real-Time PCR system (Applied Biosystems, Waltham, USA). The quantification cycle or threshold cycle number (C_q) for each reaction was recorded and used in subsequent analysis.

Table 4.5: Primers for qPCR to quantify phagemid and helper phage

Primer	Amplicon	Sequence ($5' - 3'$)
g5-F	gV	CACCGTTCATCTGTCCTCTTT
g5-R	gV	CGACCTGCTCCATGTTACTTA
gfp-F	gfp	CAAGATGAAGAGCACCAAAGG
gfp-R	gfp	CGAAGTGGTAGAAGCCGTAG
luc-F	luc	GCGCGGAGGAGTTGTGTT
luc-R	luc	TCTGATTTTTCTTGCGTCGAGTT

Conversion from the mass of dsDNA to the number of genome copies is given by

$$gc = \left(\frac{mass}{size} \times 607.4 + 157.9\right) \times (6.02 \times 10^{23}), \tag{4.1}$$

where gc is the concentration of phage genome copies in genome copies/µL, $mass$ is the mass of the dsDNA standard in g/µL, and $size$ is its length in bp [409]. Phage concentrations are estimated from their respective calibration curve. First, the C_q values from each exogenous control are plotted against the log of the known concentrations for each concentration in the dilution series. Linear regression produces a familiar equation of the form

$$C_q = mx + b \tag{4.2}$$

where C_q is the measured threshold cycle number, m is the slope, x is the base-10 log of the concentration in gc/µL and b is the x-intercept.

PCR amplification efficiency is given by

$$E = 10^{-1/m} - 1.$$ (4.3)

Subsequently, virion concentration V can be estimated simply by

$$\mathrm{V} = 10^{(C_q - b)/m} \times 2$$ (4.4)

where multiplication by 2 adjusts for the estimation of single-stranded products (gc/µL) from dsDNA standards. Only C_q measurements within the bounds of the calibration curve were used to calculate phage concentration.

Statistical analysis.   All statistical analyses were performed using Python (with the packages NumPy [414], Pandas [415, 416], SciPy [417], scikit-bio [418], and statsmodels [419]). Values were reported as means of n independent experiments with uncertainty reported as the standard error of the mean (SEM) or standard deviation (SD), as indicated. Statistical hypothesis tests were evaluated using one-way ANOVA, followed by the Tukey range test for multiple comparisons. Values of $p < 0.05$ were considered statistically significant; in some cases, values of $p < 0.1$ are noted and discussed.

Phagemid fractions were determined as the concentration of target phagemid divided by the total virion concentration, expressed as percentages. As compositional data [420], they have a fixed constant sum constraint (100%). In order not to violate this constraint, the data were transformed using an isometric log ratio transformation before performing statistical analyses. Data were then transformed back to percentages for reporting.

4.3   Results

First, M13SW7 was constructed by replacing the gIII in M13KO7 with the display-accommodating gIII from M13KE. The targeting ligand EGF was then inserted in-frame and the ability of this ligand-displaying helper (M13SW7-EGF) to replicate and assemble mini and full phagemids encoding *cmv-gfp* was investigated. Next, a derivative of M13KO7 deficient of its packaging signal was constructed (M13SW8). The ability of this helper phage to rescue phagemid was characterized across multiple *E. coli* strain backgrounds.

4.3.1   Construction of a helper phage vector displaying EGF

Construction of M13SW7 proceeded through Gibson assembly of M13KE-derived gIII (1.4 kb) and the M13KO7 backbone without gIII (7.3 kb); fragments were amplified from their respective genomes (Figure 4.4A). Two endonuclease target sites (KpnI and EagI) are present in M13KE-derived gIII. KpnI does not cut M13KO7; instead, M13KO7 contains a unique BstBI cut site

within its the p15a *ori*, which is not present in M13KE. To verify successful Gibson assembly, putative M13SW7 clones were digested with KpnI-BstBI and visualized via AGE (Figure 4.4B). In their respective lanes, BstBI-linearized M13KO7 and KpnI-linearized M13KE ran at their expected genome sizes: 8.7 kb and 7.2 kb. M13SW7, having both cut sites, was expected to show two bands: 4.9 kb and 3.8 kb. Several putative clones exhibited this pattern (Figure 4.4), confirming the successful construction of M13SW7.

(A) Products for assembly   **(B)** KpnI-BstBI digestion of putative M13SW7
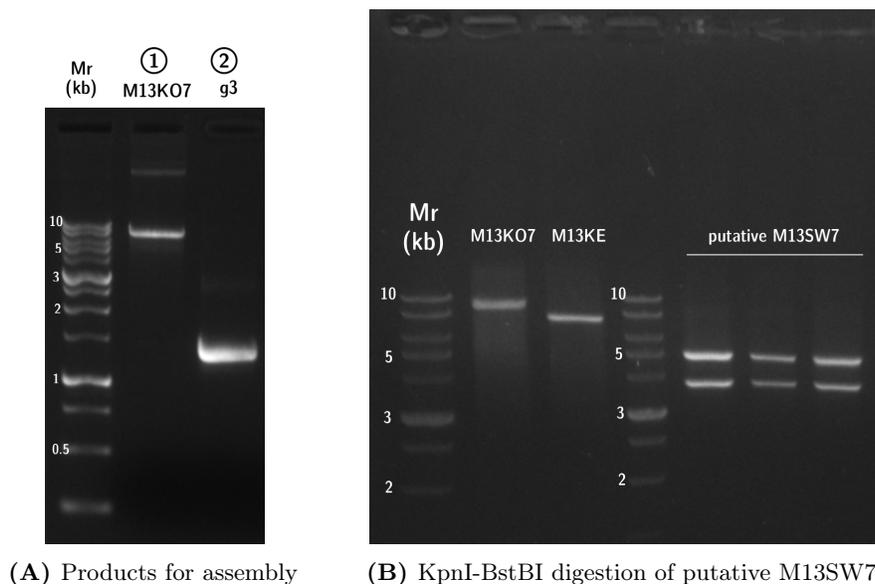
Figure 4.4: Construction of a recombinant helper phage from M13KO7 and M13KE. A) Lanes show PCR products for Gibson assembly: M13KO7 backbone (7.3 kb) and M13KE gIII (1.4 kb). B) Lanes show products digested with KpnI and BstBI. From left to right: M13KO7 (8.7 kb) and M13KE (7.2 kb), 3 putative M13SW7 clones (4.9 kb, 3.8 kb). Mr: 1 kb, FroggaBio.

Next, the targeting ligand EGF was amplified by PCR to include a GGGS linker (*egf::GGGS*) for insertion into the KpnI-EagI sites of M13SW7. Phage tolerance of the fusion was assessed by plaque assay of PEGylated lysates (Table 4.6). M13SW7-EGF titres were comparable to those of its parents, M13KO7 and M13KE, as well as wild-type M13. Thus, the EGF fusion did not impede phage infectivity.

Results summarized in Figure 4.5 confirm the genetic insertion of *egf::GGGS* between KpnI and EagI, and display of EGF on the M13SW7 virion. Successful insertion of the fusion was verified through PCR (Figure 4.5B). Without insertion, the expected amplicon size was 54 bp; if inserted, the expected amplicon size was 228 bp, which was confirmed through PCR and visualized via AGE. Display of the peptide was also verified through immunoassays: both through an ELISA (Figure 4.5A) and dot blot (Figure 4.5C). Signal was detected for M13SW7-EGF and not M13SW7, confirming the presence of the EGF ligand.

Table 4.6: Phage titre of the EGF-displaying helper phage

Phage	Titre ($\times 10^{13}$ PFU/mL)[1]
M13	0.31 ± 0.06
M13KO7	1.90 ± 0.12
M13KE	0.12 ± 0.02
M13SW7	3.43 ± 1.73
M1SW7-EGF	2.75 ± 1.04

[1] Uncertainty reported as SEM$\times 2$, $n = 3$.

(A) ELISA of EGF-displaying helper phage

(B) PCR of insertions in gIII

(C) Dot blot capturing EGF display on helper phage
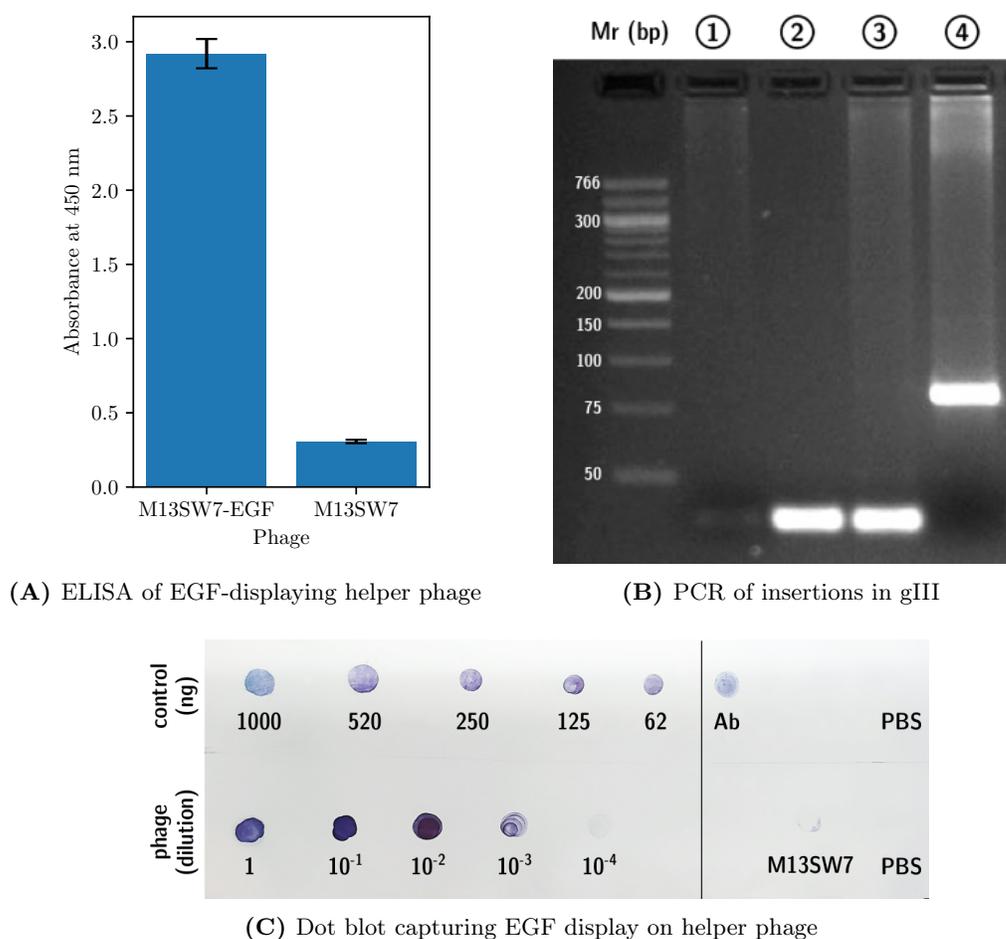
Figure 4.5: EGF display is confirmed on recombinant helper phage. A) ELISA of the putative EGF-displaying helper is compared to non-displaying helper. Error bars represent SD, $n = 3$. B) The gIII KpnI-EagI region was amplified by PCR and visualized via AGE. From left to right: 1) M13KO7 (negative control, no amplification), 2) M13KE (54 bp), 3) M13SW7 (54 bp), 4) M13SW7-EGF (228 bp), Mr: Low Molecular Weight Ladder (New England BioLabs). C) A dot blot comparing dilutions of recombinant human EGF (control, top) to dilutions of M13SW7-EGF (bottom). As a positive control, the antibody was spotted directly on the membrane, alongside M13SW7 (negative control, no display) and PBS.

4.3.2 Characterization of EGF-displaying helper performance

We next examined the ability of M13SW7-EGF to assemble phagemid from split and wild-type origins. We compared the overall yield as well (Table 4.7) as the phagemid:helper ratios between M13SW7-EGF and M13KO7 (Figure 4.6). Yield of full$_{egf}$-(luc) was highest overall; in fact, yields from lysates prepared using M13SW7-EGF tended to be greater than those prepared with M13KO7, but not always to a significant degree. As observed previously (Chapter 3), the *cmv-gfp* encoding phagemids were produced at a higher phagemid:helper ratio than *cmv-luc*. However, no significant difference in phagemid compositions was observed between the use of M13KO7 or M13SW7-EGF as the helper. Both mini-(gfp) and mini$_{egf}$-(gfp) comprised the largest fractions of their lysates, followed by full-(gfp) and full$_{egf}$-(gfp). Similarly, both mini-(luc) and mini$_{egf}$-(luc) were also rescued to a greater extent better than full-(luc) and full$_{egf}$-(luc), which comprised the smallest fractions of their lysates. Overall, the display of EGF did not affect the helper phage's ability to rescue mini or full phagemids.
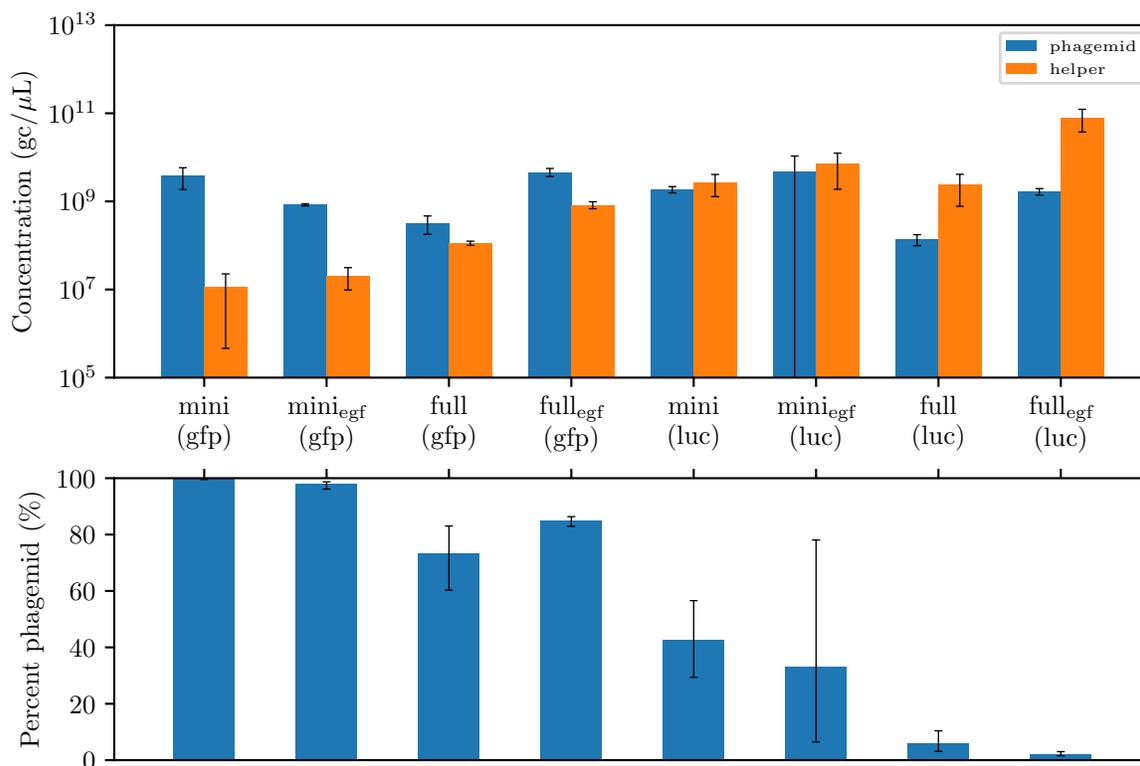
Figure 4.6: Composition of EGF-displaying phage lysates. At the top, the concentration (gc/μL) of each phage species (target phagemid or helper) in each lysate is presented. Below, the phagemid fraction of each phage species is shown as a percentage of the total phage population. Error bars represent SEM×2, $n = 3$.

Table 4.7: Comparison of phage yield from the EGF-displaying helper

	Total phage concentration ($\times 10^8$ gc/µL)[1]			
	mini		full	
Cassette	$-$EGF	$+$EGF	$-$EGF	$+$EGF
cmv-gfp	38.49 ± 19.82	8.56 ± 0.33	4.36 ± 1.32	54.56 ± 10.84
cmv-luc	45.51 ± 13.72	118.81 ± 39.42	25.87 ± 16.80	819.78 ± 430.10

[1] Uncertainty reported as SEM$\times$2, $n = 3$.

4.3.3 Construction of a helper phage vector without a packaging signal

Another helper phage was derived from M13KO7; here, the PS downstream of gIV was removed and replaced with a Rho-independent terminator to generate M13SW8. PCR amplification of the region between gIV and the p15a *ori* would result in a 386 bp (Figure 4.7). In M13SW8, deletion of the region should result in amplification of a 169 bp band. Bands at the expected sizes were observed, confirming the construction of M13SW8.
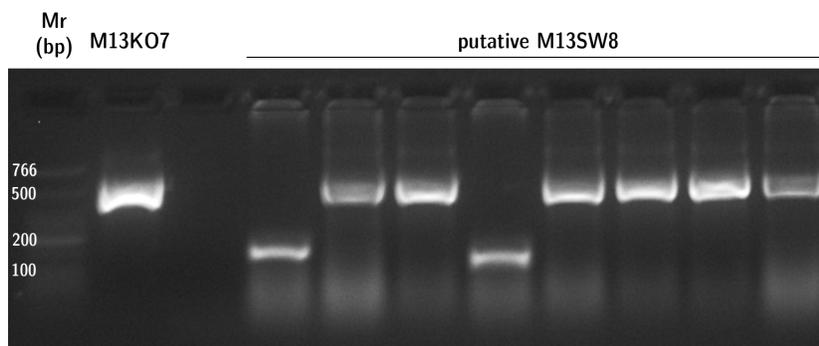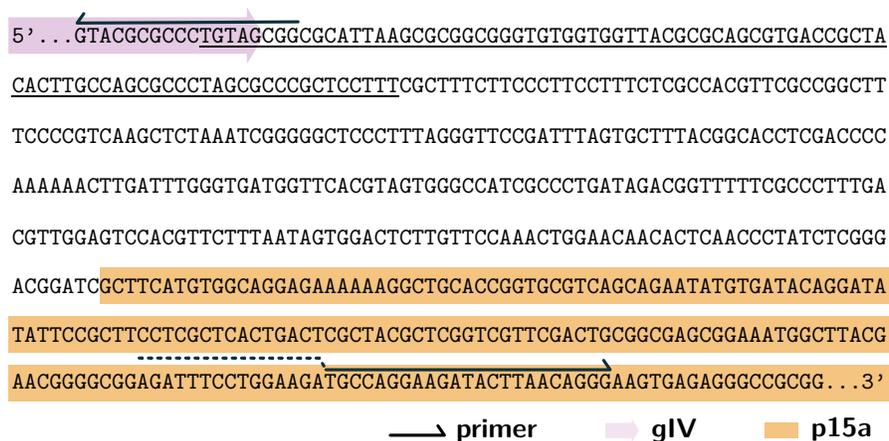
Figure 4.7: The PS is deleted in M13KO7 to create M13SW8. On top, the region between gIV and the p15a *ori* in M13KO7 is shown. Nucleotides identified as part of the PS loop are underlined. Primers for mutagenesis and relevant features are depicted. Below, PCR products of this region are visualized via AGE. From left to right: M13KO7 (386 bp), putative M13SW8 clones (169 bp). Mr: Low Molecular Weight ladder (New England BioLabs).

M13KO7 and M13SW8 phage were propagated in ER2738 alone and in presence of pBluescript II KS+. Figure 4.8A shows plaque formation by spots of M13KO7 lysate on a lawn of ER2738, but no plaques even at low dilutions for M13SW8. Notably, attempts to extract ssDNA from M13SW8 lysates were unsuccessful. The absence of plaque formation and ssDNA in the lysate suggest no M13SW8 phage production.

Next, lysates of pBluescript II KS+ phagemid assembled by M13KO7 or M13SW8 were incubated with susceptible ER2738. The infected cultures were spotted onto LB agar supplemented with ampicillin (Figure 4.8B) or ampicillin and kanamycin (Figure 4.8C). Ampicillin resistance arose by cell uptake of Ap^R pBluescript phagemid particles; single colonies could not be discerned even at high dilutions (10^{-8}), indicating a high phagemid titre from both helpers. Kanamycin resistance also emerged by cell uptake of Km^R M13KO7 or M13SW8 helper phage particles. However, no $Ap^R Km^R$ colonies were visible past the 10^{-5} dilution spot for the M13SW8-packaged lysate. In comparison, single colonies could not be discerned for the M13KO7-packaged lysate, indicating much higher concentration of M13KO7 in the phagemid lysate compared to M13SW8.
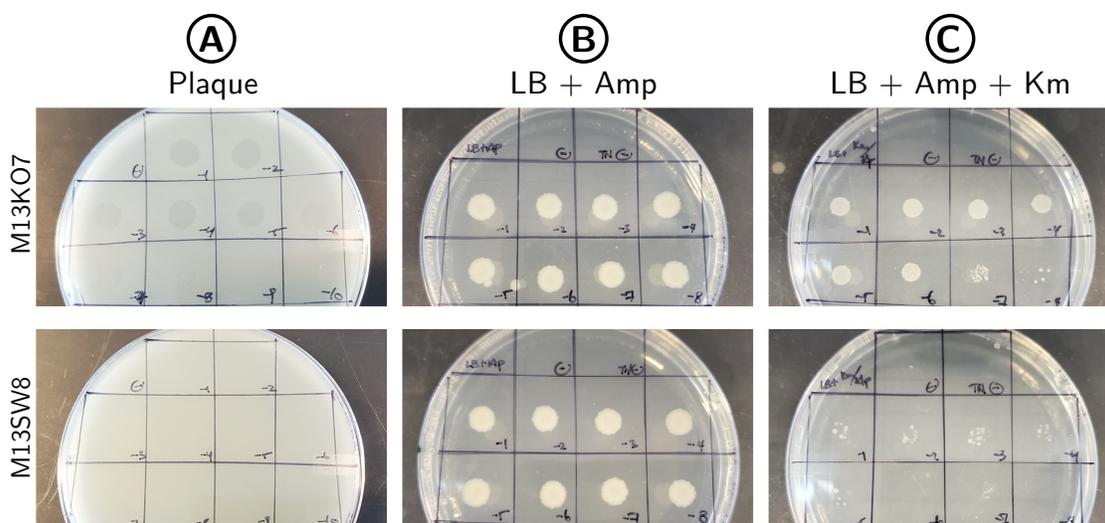
Figure 4.8: Helper phage M13SW8 is deficient in self-packaging. The top row shows the helper phage M13KO7, while the bottom row shows M13SW8. A) Helper phage lysates were spotted onto a lawn of susceptible *E. coli* ER2378 (Km^S and Ap^S). B) After incubation with the helper phage, infected ER2738 were spotted onto LB agar supplemented with kanamycin. C) After incubation with helper phage-packaged pBluescript II KS+ (Ap^R), infected ER2738 were spotted onto LB agar supplemented with kanamycin and ampicillin.

Titres determined from helper phage or helper-rescued pBluescript lysates show similar results (Table 4.8). While the lysate generated from M13SW8-rescue of pBluescript exhibited a tenfold higher Ap^R colony titre than M13KO7-rescued lysate (3.00×10^{14} versus 4.37×10^{13} CFU/mL), the Km^R colony titre was much lower. Indeed, the Km^R colony titre (helper phage) is seven orders of magnitude lower than that of the Ap^R (phagemid) in the M13SW8-rescued lysate. In comparison, the Ap^R and Km^R colony titres are much more similar when M13KO7 is the helper.

Table 4.8: Phage titres from a helper deficient of its packaging signal

	M13KO7[1]	M13SW8[1]
Phage only		
PA	$3.16 \times 10^{13} \pm 0.42 \times 10^{13}$	–
CA (Km)	$2.61 \times 10^{14} \pm 0.32 \times 10^{14}$	–
pBluescript II KS+		
CA (Ap)	$4.37 \times 10^{13} \pm 0.58 \times 10^{13}$	$3.00 \times 10^{14} \pm 1.15 \times 10^{14}$
CA (Km)	$4.43 \times 10^{13} \pm 0.29 \times 10^{13}$	$3.93 \times 10^{7} \pm 0.24 \times 10^{7}$
CA (Ap/Km)	$4.53 \times 10^{13} \pm 0.68 \times 10^{13}$	$3.94 \times 10^{7} \pm 0.32 \times 10^{7}$

[1] Uncertainty reported as SEM×2, $n = 3$.   PA: Plaque assay (PFU/mL).
CA: Colony assay (CFU/mL), antibiotic indicated in parentheses.

4.3.4 Rescue efficiency of a helper phage deficient in self-packaging

We next compared the efficiency of M13KO7 and M13SW8 in the production of mini-(gfp) and full-(gfp) across six strains that represent common laboratory strains for plasmid or M13 propagation [439, 440]. The five F^{+} strains were JM109, XL1-Blue, Stbl4, NEB Turbo, and ER2738; an F^{-} control, DH5α, was also compared. Each strain was transformed by the mini phagemid precursor pM13ori2.cmvgfp or the full phagemid plasmid (pSW9) and either M13KO7 or M13SW8 to act as the helper. Phagemid fractions for each lysate are reported in Figure 4.9 while total phage concentrations are given in Table 4.9
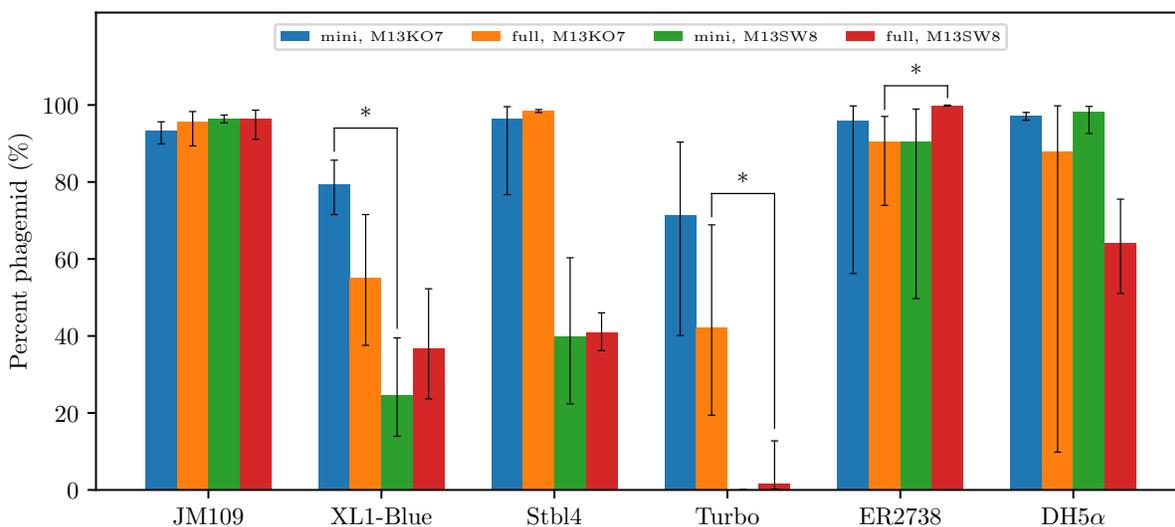
Figure 4.9: Composition of lysates packaged by PS-deficient helper phage. The proportion of packaged target phagemid (either mini-(gfp) or full-(gfp)) is shown as a percentage of the total phage population, as produced by either M13KO7 or M13SW8 across 6 different *E. coli* strains. Error bars represent SEM×2, $n = 3$. The ∗ above the bars indicates a difference at significance level $p < 0.05$.

Total phage titres rescued by M13KO7 did not differ significantly across strains; however a higher total phage titre was observed in M13SW8 rescue of full-(gfp) in JM109 when compared to all other strains ($p < 0.05$). Both M13KO7 and M13SW8 were able to rescue phagemid in DH5α, despite the lack of an F$'$ episome. In NEB Turbo, M13SW8-rescue of mini-(gfp) was detected, but the phagemid comprised only 0.04% of the lysate, with a concentration of 1.35×10^4 gc/µL. In contrast, the helper phage comprised the vast majority of the lysate at 4.19×10^7 gc/µL. The best-performing strains were JM109 (mini and full phagemids comprising 93.3% and 96.7% of their respective lysates) and ER2738 (90.5%–99.9%).

Table 4.9: Comparison of phage yield by two helper phages across different strains

| | Total phage concentration ($\times$ 10^8 gc/µL)[1] | | | |
| | M13KO7 | | M13SW8 | |
Strain	full	mini	full	mini
JM109	5.24 ± 6.92	2.33 ± 0.94	8.33 ± 5.88	5.08 ± 0.42
XL1-Blue	6.61 ± 5.38	2.32 ± 0.72	2.78 ± 1.34	1.29 ± 1.10
Stbl4	11.77 ± 7.24	13.93 ± 10.56	8.96 ± 6.68	0.80 ± 0.80
Turbo	1.10 ± 0.68	1.01 ± 0.08	0.42 ± 0.20	0.78 ± 0.76
ER2738	6.85 ± 5.18	1.38 ± 0.70	3.85 ± 4.24	0.69 ± 0.36
DH5α	14.72 ± 16.36	17.47 ± 33.16	55.38 ± 94.10	1.10 ± 1.16

[1] Uncertainty reported as SEM$\times$2, $n = 3$.

M13KO7 performed better than M13SW8 in XL1-Blue (55.0%–79.0% versus 37.4%–25.7%), Stbl4 (92.5%–98.5% versus 40.7%–41.1%) and NEB Turbo (43.6%–68.0% versus 0.04%–3.78%). Within JM109, ER2738, and DH5α, M13KO7 and M13SW8 performed more similarly. Interestingly, M13SW8 performed better than M13KO7 in ER2738 for one phagemid: the proportion of full-(gfp) was greater than when rescued with M13KO7 (99.9% versus 90.6%). Differences were statistically significant ($p < 0.05$).

Purified RF DNA linearized by BamHI was also visualized via AGE (Figure 4.10). In JM109 specifically, the mini-(gfp) RFx band as processed by M13SW8 was much brighter than that of M13KO7. This was correlated with an increase in the phagemid:helper ratio, although the difference was not statistically significant ($0.05 < p < 0.1$). In DH5α, pM13ori2.cmvgfp appears to be completely lost, processed by M13KO7 to RFx. This was correlated with a very high phagemid:helper ratio (97.2%). In comparison with RF DNA isolated from JM109, purified RF DNA from other strains also shows bands around the 1–2 kb range; these bands appear similar in size to single-stranded full-(gfp) or mini-(gfp) species (Chapter 3). Presence of these bands is weak in the two strains associated with higher phagemid:helper ratios, JM109 and Stbl4, but very strong in the worst-performing strains, XL1-Blue and NEB Turbo. In XL1-Blue and NEB Turbo, bands around 4 kb are also visible in the M13KO7-mini-(gfp) and M13KO7-full-(gfp) lanes, respectively.

Across all strains, M13SW8 rescue of full-(gfp) also appears to produce another extraneous 3 kb band. This band is not present when the helper phage is M13KO7 in JM109, but occurs in all other strains. It also does not appear to manifest from the rescue of miniphagemids by M13SW8, only during the rescue of full phagemids.
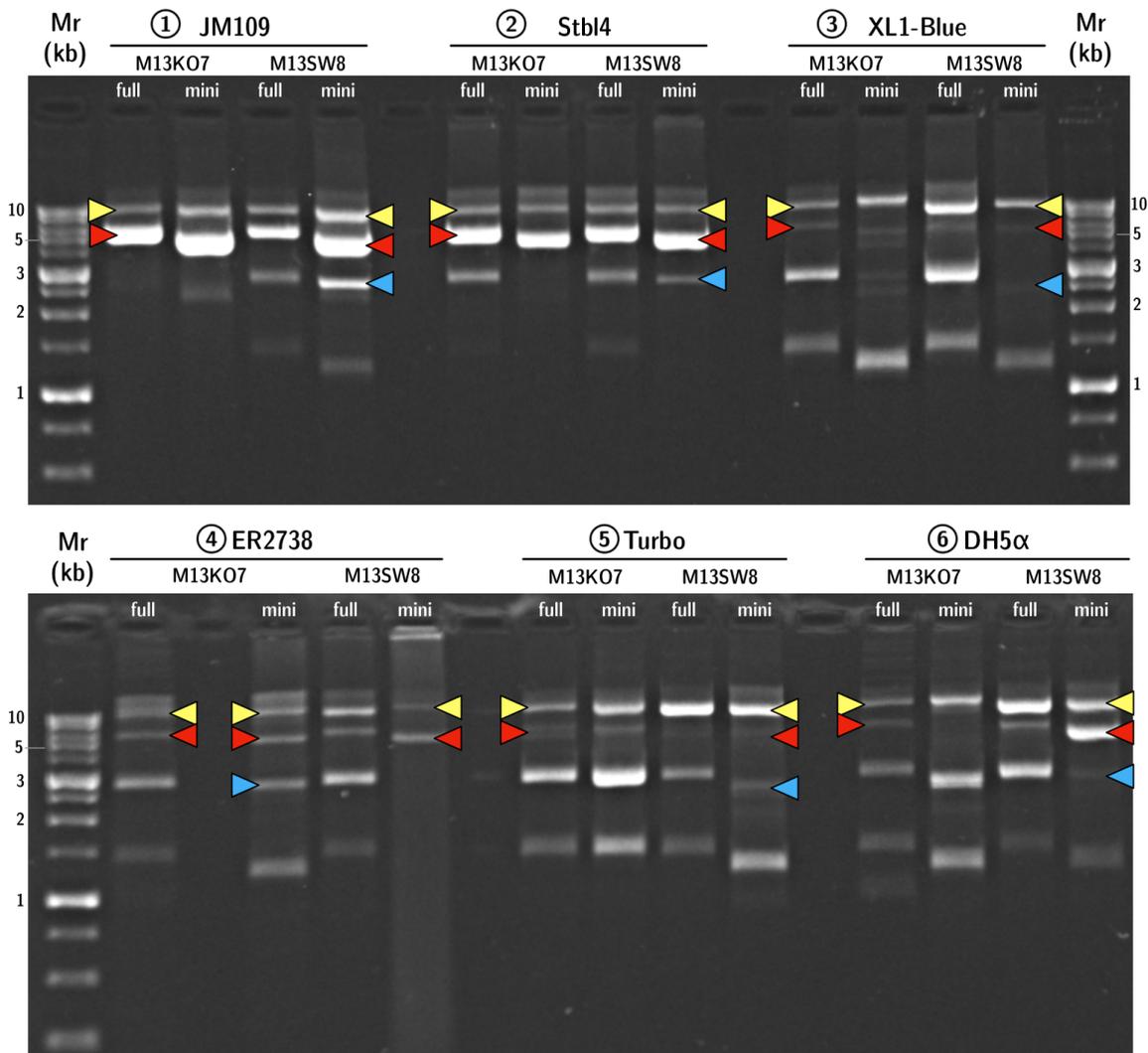
Figure 4.10: PS-deficient helper phage produces recombinant RF (RFx) across 6 *E. coli* strains. BamHI-digested RF DNA was extracted from cells carrying either pSW9 (full) or pM13ori2.cmvgfp (mini) with helpers M13KO7 or M13SW8. Top: 1) JM109, 2) Stbl4, and 3) XL1-Blue. Bottom: 4) ER2738, 5) NEB Turbo, and 6) DH5α. Arrows indicate bands of interest: yellow (helper genome), red (plasmid or corresponding full phagemid), blue (recombinant RF: RFx). Mr: 1 kb ladder (FroggaBio).

4.4 Discussion

4.4.1 Display of EGF does not impact phage infectivity nor rescue of mini or full phagemids

EGF has been successfully displayed on pIII, in both N-terminal [347, 449–451] and C-terminal fusions [347]. Consistent with this body of work, we observed no impact from M13 pIII display of EGF on phage viability (Table 4.6), nor ligand recognition by its cognate antibody (Figure 4.5). In fact, the M13SW7 and M13SW7-EGF titres were among the highest of the group under the current study conditions.

The displayed peptide totaled 57 residues in length when including the linker. In general, pIII can tolerate pentavalent display of peptides up to approximately 50 residues long. Display of longer peptides generally decrease display valency; large proteins including antibody fragments are only possible through monovalent display on pIII [257]. Thus, the inclusion of a flexible linker helps to avoid steric hindrance; the GGGS linker used here is the most common form of flexible linker used to join components of fusion proteins [459].

The display of EGF was not expected to impact phagemid rescue as pIII has no role in the replication phase of the phage lifecycle. Quantification of phage lysates by qPCR showed higher titres with the EGF-displaying helper than the non-displaying, although this was only significant for full$_{egf}$-(luc). In general, M13SW7-EGF performed as well as M13KO7, its non-displaying parent, in the production of both mini and full phagemids. Lysate compositions were consistent with those obtained with the non-displaying parent: again, efficiency of phagemid production is dependent on the phagemid-encoded transgene. Overall, these results demonstrate the viability of a pIII display helper system for the packaging of phagemids [347, 451].

4.4.2 Removal of the packaging signal greatly reduces helper phage self-packaging

The self-packaging deficient helper phage M13SW8 was constructed by deleting a ∼500 bp region containing the filamentous PS between gIV and the p15a *ori* in M13KO7. Termination of gIV gene expression is dependent on a Rho-dependent terminator near or within the PS signal [194]; however, Rho-independent termination of the gIV transcript has also been documented [187, 194]. It is believed that the strong secondary structure within the f1 *ori* can halt transcription independent of Rho. Replacement of the PS with a *trp* terminator did not appear to impact phagemid rescue, as M13SW8-rescued pBluescript lysates were viable and able to confer ampicillin resistance to susceptible Ap^S cells (Figure 4.8).

Levinson et al. (1984) postulated that a low phagemid rescue from a helper phage resulted from ssDNA helper intermediates outcompeting phagemid intermediates for pV processing and assembly [315]. Therefore, we hypothesized that removal of the helper phage PS would decouple helper phage replication from assembly and extrusion. We would expect to see accumulation

of ssDNA helper intermediates within the cell, which would convert back to duplex RF, and a reduction in helper phage virions in the final lysate. Accumulated helper RF molecules could serve as more templates for phage gene expression, which may further improve phagemid yield.

Spot plates of M13SW8 on ER2738 showed no visible plaque formation even at low dilution (Figure 4.8). However, kanamycin resistance was still transferred to cells after incubation with a pBluescript lysate using M13SW8 as the helper. Still, titres of Km^R phage were much lower when compared to cells infected by the M13KO7-mediated pBluescript lysate (Table 4.8). This indicates that loss of the PS does reduce helper self-packaging to a large extent. Despite the absence of plaque formation and the lack of ssDNA from the phage lysate, M13SW8 infection still occurs in the presence of phagemid. Together, this suggests that M13SW8 phage extrusion only occurs when a functional PS is present. Single-stranded M13SW8 molecules may simply be assembled into extruding pBluescript particles. Although the PS is required to initiate assembly, phage particle elongation with additional DNA molecules does not appear to have this requirement [152]. In wild-type phage populations, approximately 5% of phage progeny encapsulate multiple genomes [152, 234]. Based on the titres of Ap^R and Km^R phage from M13SW8-mediated rescue of pBluescript, M13SW8 comprises less than 1% of the lysate.

Hence, we postulate that virion production of the PS-deficient helper phage may manifest primarily as multi-length virion particles co-extruded with PS-proficient phagemids. Moreover, the accumulation of unextruded phage genome within the cell may serve as more template for greater phage gene expression.

4.4.3   Choice of host background greatly impacts helper phage efficiency

We next contrasted M13SW8 and M13KO7 helper activity on *cmv-gfp* encoding phagemids across an array of different host backgrounds. At minimum, F-specific filamentous phages (Ff phages) have the capacity to infect any Gram-negative bacteria that express cell-surface TolA [159]. However, for more robust filamentous infection, F^+ *E. coli* capable of pilus formation are necessary. Overall, M13 phage replication places an enormous metabolic burden on the infected host; in addition, the presence of the precursor phagemid further increases the metabolic strain. Both the phagemid and phagemid precursors under investigation have the high copy number ColE1 origin (300–500 copies/cell) [464] while the p15a *ori* present on the helper phages M13KO7 and M13SW8 is medium copy (10 copies/cell). In wild-type M13 infection, the maximal RF copy number has been observed to reach approximately 30–50 molecules per cell [465]. As high levels of episomal DNA must be continually maintained, a suitable host should be both proficient in the propagation of large plasmids and be susceptible to Ff phage infection. Additionally, both the helper and phagemid have plasmid origins so phage re-transduction or re-infection may not be necessary for phagemid production. We sought to investigate this using DH5α, an F^- strain, as the absence of the fertility factor and lack of pili formation should prohibit infection or re-infection by M13 progeny during production.

From a head-to-head comparison, the host background more strongly impacted phagemid production than the choice of helper phage (Figure 4.9). Total phage yield was also strain-dependent (Table 4.9). Note that cells were transformed by helper phage RF DNA in this study instead of being directly infected as in Chapter 3. This was because M13SW8 virions could not be isolated for infection. However, no difference in helper phage activity was observed between these two methods of DNA delivery.

Overall, JM109 was the best-performing strain, with few to no byproducts in its RF extraction and a high phagemid:helper ratio. While the ER2738-M13SW8 combination was better at rescue of full-(gfp), its ability to rescue mini-(gfp) was highly variable. ER2738 is a $recA^+$ strain and examination of RF products (Figure 4.10) shows multiple DNA species in the cell. This suggests that phagemid processing may activate host recombination pathways, producing other DNA species during infection. Accordingly, $recA^-$ hosts should be preferred over $recA^+$.

However, inactivation of $recA$ has been found to be insufficient to completely eliminate undesired recombination. Another $recA$ host, Stbl4 has been recommended for the propagation of unstable inserts or large repetitive sequences [466]. Such sequences are difficult to maintain in *E. coli* as the potential for genetic rearrangement between repetitive regions are homology-dependent and RecA-independent [455]. Stbl4 is a derivative of Stbl2, which itself is a derivative of JM109/J5, a JM109 clone found to stably maintain the simian immunodeficiency virus (SIV) genome [454]. Stbl2 was constructed by introducing mutations in JM109/J5 to enable cloning of methylated DNA. As of this writing, the published genotypes of Stbl4 (Appendix B) and Stbl2 [454] do not reveal any mutations that explain their prowess in the maintenance of unstable recombination-prone plasmids. However, the entire genome sequence for Stbl4 was very recently published [466], which may soon shed more light on these strains.

M13KO7-mediated rescue of both the mini and full phagemids were comparable to that of JM109; however, M13SW8 performed much more poorly. M13SW8 performance was similarly poor in the other $recA$ strain, XL1-Blue, but better in the $recA$ F$^-$ strain, DH5α. It is possible that RecA-independent homologous recombination [455] may be impacting phage replication when intracellular levels of helper genome are high due to the lack of phage extrusion, although it is unclear why this occurs in Stbl4 and XL1-Blue but not JM109. We postulate some potential reasons below.

Plasmid nicking may adversely impact phagemid rescue. In contrast to JM109, the $recA$ strain XL1-Blue performed poorly despite their genotypic similarities. Both XL1-Blue and NEB Turbo were poor hosts for miniphagemid processing by both M13KO7 and M13SW8, which may be related to plasmid nicking by endonuclease A (*endA*). While all six strains under investigation carry the *endA* mutation, it has previously been observed that nicking and subsequent plasmid degradation can still occur in some *endA* backgrounds [467]. A head-to-head comparison of multiple *E. coli* strains found that supercoiled plasmid DNA was mostly retained in some

endA strains like JM107 and DH5α (99% retention), but was lost in others like NEB Turbo and XL1-Blue (77% and 82% retention) [445]. It may be possible that, despite being genotypically absent of endonuclease A, some *endA* strains still exhibit plasmid nicking activity through other compensatory mutations.

Nicking of supercoiled duplex RF DNA releases DNA gyrase-imposed torsion, leading to a circular relaxed form. Unsupercoiled phage genome is not a template for phage gene expression [204]. This may negatively impact some phage proteins such as pVIII and pV, which are normally required in very large amounts. Random nicks may interfere with DNA polymerase processivity during rolling circle amplification. If the nick is in the minus strand, it would impede DNA polymerase processivity; if the nick is on the positive strand, no ssDNA intermediate would be formed. In either case, no daughter RF would be produced. On the whole, uncontrolled random nicking of RFs may impede RF amplification and phage gene expression. Because host backgrounds which are not truly deficient of endonuclease A appear to be poor hosts for replication of plasmid DNA, this might explain the poor performance of XL1-Blue and NEB Turbo.

Accumulation of M13SW8 RF may reduce phage gene expression. A 3 kb band also appears in RF DNA isolated from helper-full-(gfp) combinations (Figure 4.10). Curiously, this is the size of pBluescript II KS+, the backbone of pSW9. It may be possible that deletion of the *cmv-gfp* cassette is occurring in the presence of M13SW8, but this is unlikely in recombination-deficient hosts such as Stbl4, JM109, XL1-Blue, and DH5α. Alternatively, accumulation of M13SW8 due to its lack of PS may lead to pII-mediated generation of smaller DNA intermediates independently of host recombination pathways.

Spontaneous formation of "miniphages" has been previously observed with infection of filamentous phages at high multiplicity of infection (MOI) [235, 468]. These miniphages primarily consist of the intergenic region between gIV and gII: in M13SW8, this region is approximately 2.7 kb long. Although we predicted that helper phage accumulation could increase the level of phage gene products, miniphage conversion would instead reduce the pool of available templates. Since miniphages lack all elements of the phage genome aside from the f1 *ori*, they cannot be a template for phage gene expression. Instead they would compete for phage resources alongside phagemid molecules. Fortunately, observations of miniphages show that they are generally outcompeted by their full phage counterparts [235]. However, prolonged rates of conversion from intact helper genome to miniphage could eventually suppress gene expression and reduce overall phagemid processing. Strains with potentially strong miniphage presence (XL1-Blue, NEB Turbo, and DH5α) correspondingly show reduced phagemid:helper ratio.

AGE analysis also showed multiple bands not consistent with the expected RF sizes, particularly in XL1-Blue and NEB Turbo. The smaller 1–2 kb bands correspond to the expected sizes of single-stranded mini-(gfp) and full-(gfp). Bands of approximately 4 kb can also be seen, which are most likely the helper phage ssDNA intermediates. The strains with some of the highest

phagemid:helper lysate ratios, JM109 and Stbl4, show very little ssDNA presence in purified RF. Accumulation of phagemid ssDNA intermediates occurs when the helper phage is M13SW8 across all strains, and also occurs strongly when the helper is M13KO7 in XL1-Blue and NEB Turbo. The presence of these intermediates suggests that pII-mediated phagemid amplification and conversion to RFx occurs readily with a PS-deficient helper. High ssDNA retention in the cell suggests these molecules are not efficiently assembled into phage progeny after replication nor converted into more dsDNA RFs. These results imply that the rate of assembly and extrusion is inadequate for the amount of ssDNA intermediates. Efficient phage assembly requires a substantial number of pVIII to coat the extruding phage particle. Low levels of pVIII may be due to insufficient supercoiled helper phage template, whether from helper phage genome conversion to miniphages or conversion to circular relaxed forms of RF.

Earlier, we postulated that M13SW8 presence in the phage lysate may be primarily driven by co-encapsulation with extruding phagemid. This may be true in optimal host backgrounds such as JM109. In some strains, M13SW8 comprises the majority of the phage lysate (specifically, XL1-Blue, NEB Turbo, and Stbl4), which appears inconsistent with this hypothesis. The accumulation of large amounts of M13SW8 product may drive recombination events between phagemid and helper, possibly reconstituting the PS in the helper at the expense of the phagemid. Furthermore, these three strains were also under heavier metabolic burden due to the presence of antibiotic resistance markers on their F′ episomes [445]. With the phagemids (Ap^R) and the helper phages (Km^R), the burden of protein production may reduce overall fitness in these strains. It has been suggested that metabolic burden is generally more due to increased protein production rather than DNA (specifically plasmid) maintenance [445, 469].

On the whole, choice of host strain background may be the main driving factor of optimal miniphagemid production over choice of helper phage. The recA, gyrA, endA host, JM109, was the best overall host for phagemid production. Mutations such as recA and gyrA reduce occurrences of homology-dependent recombination events. These mutations are likely more important for propagation of larger phagemids (>10 kb). We observed that although the F′ episome may not be strictly required, its presence may help improve phage yield.

Other *E. coli* strains may also be attractive as potential hosts. The strains under evaluation were all derivatives of *E. coli* K12; investigation of strains derived from other wild-type strains may reveal other genetic factors that can influence M13 phagemid rescue. HB101 [470], for example, is a cross between *E. coli* K12 and B. Another interesting strain to investigate may be Stbl3, a recA derivative of HB101 [471] with similar properties to Stbl4. A further example is Mach1, a fast-growing strain similar to NEB Turbo, which is derived from *E. coli* W [472]. Yau et al. also reported runaway plasmid replication (up to 2708 copies of a 5.8 kb plasmid) without any decrease in growth rate in *E. coli* strains TG1, HB101, and MG1655 [445]. All three of these strains may be of particular interest to increase overall phage yield.

4.5   Summary

In this chapter, we modified M13KO7 to display a mammalian cell-specific ligand, EGF. The resulting phage, M13SW7-EGF, maintained comparable infectivity to its parent phages and was able to process split origin precursor plasmids into miniphagemids as efficiently as M13KO7.

We further modified M13KO7 to delete its PS to reduce helper phage contamination in the phagemid lysates and examined the capacity of both helper phages across six *E. coli* hosts. A *recA* background may be beneficial for plasmid maintenance, but is insufficient to prevent formation of other DNA species, as we saw with the production of an unidentified 3 kb band. Nevertheless, *recA* strains performed as well as, if not better than, $recA^+$ strains, indicating that their slower growth rate did not impact phage production. Poor helper performance was associated with high recovery of ssDNA intermediates from the cell pellet, which may be a result of insufficient assembly and extrusion. This may be related to breakthrough plasmid nicking in *endA* backgrounds, thereby reducing levels of supercoiled ds RF DNA.

Taking our results together, we found that the PS-deficient helper phage does improve the phagemid:helper ratio. The effect is more pronounced in a *recA*, *endA* strain such as JM109. Overall, JM109 and M13SW8 appear to be the best strain-helper combination for miniphagemid production. In some cells, we postulate that accumulation of M13SW8 genomes may actually reduce phagemid production. This may arise due to helper phage genome conversion into miniphage elements, which are comprised only of the phage origin and do not contribute to phage gene expression. While this did not appear to adversely impact the phagemid:helper ratio in JM109, other strains fared more poorly.

In essence, RF and RFx production do not appear to be the limiting factor; rather, the rate of ssDNA intermediate uptake by phage assembly complexes may be the bottleneck. Disruptions to phage gene expression may negatively affect phagemid assembly if cellular levels of pV and coat protein pVIII cannot keep up with the amount of ssDNA strand production. This is consistent with our results from Chapter 3, and underscores the complexity of the interplay between phage gene expression and the transition from RF amplification to phage assembly.

Chapter 5

Transfection of a targeted miniphagemid particle in mammalian cells

5.1 Introduction

Chapters 3 and 4 resulted in the construction of miniphagemid virions which display epidermal growth factor (EGF). This ligand binds to the epidermal growth factor receptor (EGFR) overexpressed in many cancers and represents a promising target for cancer gene therapy [473–476]. The display of EGF enables our phage to specifically target cells expressing its cognate receptor for cell uptake and gene transfer applications. In this chapter, we quantify the ability of these phage particles to infect mammalian cells with and without the EGF receptor.

5.1.1 Phage gene delivery

Bacteriophages do not possess intrinsic tropism for mammalian cells; thus, they can function as simplistic proteinaceous carriers of genetic material for gene transfer applications. However, this is a double-edged sword. Although there is no danger of unregulated phage replication in mammalian cells, phages also have no means by which to transduce cells. In order to successfully transfer exogenous genetic material, a gene transfer vector must overcome a number of major cellular barriers: 1) bypassing the plasma membrane to enter the cell, 2) navigation across the cytoplasm, and 3) bypassing the nuclear envelope to enter the nucleus.

The most common gene transfer vectors are modified mammalian viruses [2, 7], which exhibit natural tropism for human cells and have evolved mechanisms to bypass these cellular barriers to gene transfer. Hence, they exhibit high transduction efficiency, although a number of safety concerns limit their usage. Specifically, high immunogenicity, risk of insertional mutagenesis, and the possibility of recombination into replication-competent viruses are some serious safety concerns that have been raised regarding viral vectors [10, 14, 19, 32].

Non-viral methods of gene transfer incorporate other means of bypassing cellular barriers for gene transfer independent of mammalian viruses. Traditionally, in vitro non-viral transfection of cell cultures has been accomplished primarily using chemical transfection reagents that mediate transport across cell membranes [4, 477]. In general, these chemical reagents tend to be positively-charged lipids or polymers that interact with the negatively-charged phosphate backbone of nucleic acids [4, 46]. However, these transfection reagents can be prohibitively expensive and cannot target to specific tissues; functionalization requires additional covalent linkage [477].

As phages cannot enter nor replicate within mammalian cells, they may be considered as non-viral gene transfer vectors. To overcome their lack of tropism, display of a cell-targeting ligand can facilitate cellular uptake by exploiting receptor-mediated endocytosis. Through the targeting of cell-surface receptors, internalization of ligand-displaying phages can occur in a cell-specific manner [326, 333, 347]. While uptake of non-targeted phage can occur [329, 334, 477], ligand-directed endocytosis appears to be more robust and can confer tissue specificity [339, 347].

The choice of ligand determines the mechanism of internalization. Display of a cell-specific ligand generally exploits clathrin-coated receptor-mediated endocytosis for cell entry [329]. However, the ligand-displaying phage then becomes trapped in the endosomal compartment; without any means of escape, it will eventually be routed to lysosomes for degradation. Additionally, a DNA molecule encapsulated by phage proteins needs to be exposed and available for transcription. Interestingly, the harsh conditions of the endolysosomal compartment, such as low pH, may aid in the separation of the phage protein coat from the DNA. As such, endosomal entry may, in fact, be necessary for phage-mediated gene transfer. Endosomal escape is a known bottleneck for non-viral gene transfer strategies [478, 479]. Traditional polymeric chemical transfection reagents can mediate endosomal escape via the proton sponge effect [478, 480]: within the acidic environment of the endosome, increasing protonation of the amine groups in cationic polymers such as polyethylenimine (PEI) results in an influx of counter-ions. The consequent osmotic swelling is thought to lead to rupture of the endosome and release of polymer-DNA complexes into the cytoplasm.

Upon endosomal escape, the DNA molecule still cannot confer any downstream effect without the transcription machinery within the nucleus; hence, it must be able to bypass the nuclear envelope. A number of strategies have evolved for increased nuclear import, such as the inclusion of protein or nucleotide localization tags [40, 42]. DNA minivectors, which are reduced in size compared to their plasmid equivalents due to the removal of bacterial backbone elements, have shown improved nuclear localization from increased nuclear diffusion [38].

Here, we evaluate M13 miniphagemid particles in the delivery of mammalian transgenes based on the following three factors: 1) the display of a cell-specific ligand to improve internalization, 2) complexation with a cationic polymer to improve endosomal escape, and 3) the absence of the prokaryotic backbone. As a proof of principle, we constructed phagemid particles displaying the EGF as the cell-specific ligand and encoding mammalian transgene cassettes that express the reporters luciferase (*cmv-luc*) and green fluorescent protein (GFP) (*cmv-gfp*) in Chapter 4.

5.1.2 EGFR as a model for receptor-mediated cell internalization of phage

EGF binds to the epidermal growth factor receptor Erb1 (EGFR), a member of the ErbB/HER family [481]. The ErbB/HER family are among the most well-studied in the receptor tyrosine kinase (RTK) superfamily of cell surface signal receptors [482, 483]. Aside from EGFR, members also include ErbB2 (Her2/Neu), ErbB3 (Her3), and ErbB4 (Her4) [483]. All are transmembrane receptors (170–180 kDa) with similar structure: an extracellular cysteine-rich domain for ligand binding, a transmembrane domain, and a cytoplasmic tyrosine kinase domain [481]. The ErbB/HER family of receptors play important roles in the regulation of cellular homeostasis, proliferation, and differentiation [476, 481, 482]. Overexpression of EGFR and other ErbB/HER family members has been implicated in many cancers, including breast, lung, colon, pancreatic, prostate, liver, and kidney [473, 474]. As EGFR signalling modulates tumour growth, the receptor is a promising candidate for anti-cancer therapies [475, 476]. Tyrosine kinase inhibitors and anti-EGFR monoclonal antibodys (mAbs) have been approved for the treatment of non-small cell lung cancer (NSCLC), colorectal cancer, and head and neck cancers [475, 484, 485].

The most well-known ligand of EGFR, EGF, is a polypeptide chain 53 amino acids long, derived from cleavage of its precursor, prepro-EGF [486–488]. However, EGFR is also known to bind six other ligands: transforming growth factor-α (TGF-α), heparin-binding EGF-like growth factor (HBEGF), amphiregulin, betacellulin, epiregulin, and epigen [488]. All are members of the group I EGF family [487] and share a cysteine-rich "EGF motif" that forms three structural integral disulfide bonds [488].

The endocytosis of EGFR has been thoroughly investigated and characterized; now, it serves as a model for most other RTKs [489]. Upon ligand recognition, EGFR dimerizes, which activates the cytosolic kinase domain of each unit to phosphorylate C-terminal tyrosine residues on the dimer partner [490]. Autophosphorylation of the dimer opens docking sites for proteins with Src homology 2 (Shc2) or phosphotyrosine binding (PTB) domains, which then activate downstream signal transduction pathways involved in cell proliferation, including mitogen-activated protein kinase (MAPK) [481, 491, 492]. Specifically, phosphorylation proceeds rapidly and in a specific order, beginning at the tyrosine in position 1173 (Tyr1173), then 1148, 1068, 992, 845, and 1045 [493, 494]. Dimerization can occur between different members of the ErbB/HER family; indeed, the varied combinations of receptor dimers diversify the downstream signalling pathways activated upon ligand binding. For example, Erb2 does not have a canonical ligand and only activates upon forming heterodimers with other ErbB/HER family members [483].

Endocytosis of the bound EGF-EGFR complex appears to be activated by dimerization and has been shown to occur independently of EGFR autophosphorylation [495]. The ligand-dimer complex is internalized via clathrin-dependent or clathrin-independent endocytic pathways. Endocytosed EGFR molecules may be recycled to the cell surface or directed to lysosomes for degradation, depending on the receptor-ligand combination [496, 497]. EGF-EGFR complexes are primarily internalized via clathrin-coated endocytosis at low physiological concentrations of EGF [498, 499]. At

high concentrations, EGF-bound receptor complexes are also internalized in a clathrin-independent manner, whereby downstream ubiquitination mediates lysosomal targeting of endocytosed EGFR dimers [489, 500]. In this way, EGFR protein levels are downregulated in a negative feedback loop.

5.1.3   Rationale and hypothesis

Previously, we constructed miniaturized phagemids (Chapter 3) expressing reporter genes *gfp* or *luc* under control of the strong universal promoter *cmv* in conjunction with multivalent EGF display (Chapter 4). Here, we assess these hybrid targeted phage particles for their capacity for mammalian gene transfer. Previous investigations into phage transfection found an increase in gene transfer with the addition of chemical cationic transfection reagents [477]. Here, we compare the delivery of phage virions alone or in combination with the commercial cationic polymer, TurboFect (Thermo Fisher Scientific, Waltham, USA). We examine the ability of these miniphagemids and their full phagemid counterparts to transfect mammalian cells using the well-characterized EGF and its receptor as a model to assess phage uptake and transgene expression.

We hypothesized that the display of EGF would be necessary for phage uptake and transgene expression in EGFR$^+$ cell lines. To test this, we selected an array of epithelial cell lines in which EGFR expression has been demonstrated: the cervical carcinoma epithelial cell line HeLa, the colorectal carcinoma epithelial cell line HT-29, and the lung carcinoma epithelial cell line A549. We also selected a non-epithelial cell line with endogenous EGFR expression, MRC-5 (lung fibroblast), and an EGFR$^-$ cell line, HEK293T, to compare the efficacy of gene transfer across different tissues. We furthermore hypothesized that the miniphagemids would confer increased gene expression as compared to their isogenic full phagemid counterparts. Their smaller size is expected to confer improved cytoplasmic diffusion, while their lack of prokaryotic backbone reduces cytosine-guanine dinucleotide (CpG)-mediated gene silencing.

5.2   Materials and methods

Strains and vectors.   *E. coli* K-12 strains were used in the generation of all phage and plasmid constructs. *E. coli* JM109 was the host for plasmid amplification and purification unless otherwise noted. All bacterial strains and mammalian cell lines are listed in Table 5.1, plasmids in Table 5.2, and phages in Table 5.3.

Maintenance of mammalian cell lines.   Mammalian cell lines were maintained in tissue culture plates (Thermo Scientific) at 37 °C in a humidified atmosphere with 10% CO_2. All cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) (Thermo Scientific) supplemented with 10% heat-inactivated fetal bovine serum (FBS) and 1% penicillin/streptomycin.

Table 5.1: Bacterial strains and mammalian cell lines used for transfection

Strain	Genotype	Source
JM109	F′ traD36 proAB⁺ lacI�q lacZΔM15/ Δ(lac-proAB) endA1 glnV44 thi-1 e14⁻ recA1 gyrA96 relA1 hsdR17	NEB

Cell line	Description	Source
HEK-293T	embryonic kidney epithelial	gift, Dr. M. Aucoin
HeLa	uterus, cervix adenocarcinoma	gift, Serenity Bioworks
MRC-5	lung fibroblast	ATCC CCL-171
HT-29	colon epithelial adenocarcinoma	gift, Dr. J. Blay
A549	lung epithelial carcinoma	ATCC CCL-185

NEB: New England BioLabs   ATCC: American Type Culture Collection

Table 5.2: Plasmids used for transfection

Plasmid	Genotype	Source
pGL2-SS-CMV-GFP-SS	pGL2-Promoter (GN: X65326.2) *cmv-gfp* replaced *SV40-luc*, Ap^R	gift, Mediphage Bioceuticals Inc.
pGL3-CMV	pGL3-Basic (GN: U47295.2) *cmv* inserted in BglII-HindIII, Ap^R	gift, Dr. N. Oviedo
pSW9	pBluescript II KS+, *cmv-gfp* inserted in BamHI-EcoRI, ApR	Chapter 3
pSW10	pBluescript II KS+, *cmv-luc* inserted in KpnI, ApR	Chapter 3
pM13ori2	pM13ori, BsaI cut site replaced with SphI, ApR	Chapter 3
pM13ori2.cmvgfp	pM13ori2, *cmv-gfp* inserted in EcoRI-PacI	Chapter 3
pM13ori2.cmvluc	pM13ori2, *cmv-luc* inserted in EcoRI-KpnI	Chapter 3

GN: GenBank Accession No.

Table 5.3: Phages used for transfection

Phage	Genotype	Source
M13KO7	M13, p15a *ori*, Tn903 (Km^R)	NEB
M13SW7-EGF	M13SW7, *egf* from pPL451-gpD::egf	Chapter 4
full-(gfp)	*cmv-gfp*, from precursor pSW9, Ap^R	Chapter 3
full$_{egf}$-(gfp)	*cmv-gfp*, from precursor pSW9, pIII::EGF, Ap^R	Chapter 4
mini-(gfp)	*cmv-gfp*, from precursor pM13ori2.cmvgfp	Chapter 3
mini$_{egf}$-(gfp)	*cmv-gfp*, from precursor pM13ori2.cmvgfp, pIII::EGF	Chapter 4
full-(luc)	from precursor pSW10, Ap^R	Chapter 3
full$_{egf}$-(luc)	from precursor pSW10, pIII::EGF, Ap^R	Chapter 4
mini-(luc)	from precursor pM13ori2.cmvluc	Chapter 3
mini$_{egf}$-(luc)	from precursor pM13ori2.cmvluc, pIII::EGF	Chapter 4

NEB: New England BioLabs

Purification of DNA vectors and phage. Plasmids and double-stranded (ds) replicative factors (RFs) were extracted and purified using the EZNA Endo-Free maxi-prep kit (Omega BioTek, Norcross, USA). Phage and single-stranded (ss) phagemid DNA lysates as quantified in Chapters 3 and 4 were used for here for transfection.

Optimization of transfection of double-stranded plasmid DNA and single-stranded phagemid DNA. The gene transfer capacity of the phagemid ssDNA was measured in comparison to plasmid dsDNA. Gene transfer was assessed in both the HEK293T (EGFR$^-$) and HeLa (EGFR$^+$) cell lines as these were the representative cell lines for later EGF$^+$ phage transfections. HEK293T cells were seeded in 24-well plates (Thermo Fisher Scientific) at 1×10^5 cells/mL, while HeLa cells were seeded at 5×10^4 cells/mL. The following day, 1 µg of dsDNA or 2 µg of ssDNA was transfected into each well after complexing with commercial cationic polymer transfection reagent, TurboFect (Thermo Fisher Scientific). The empty vector, pM13ori2, and helper phage M13KO7 alone, were also transfected as negative controls. Efficiency of gene transfer was assessed via the Luciferase Assay System (Promega, Madison, USA) and reported as raw luminescence.

Optimization of transfection of single-stranded phagemid DNA and phage particles. The capacity for gene transfer was assessed for both purified ssDNA and phage-encapsulated phage over time in the HeLa cell line. Cells were seeded at 5×10^4 cells/mL in 24-well plates. Phage particles were added to a final concentration of 5×10^7 virions/mL: mini-(luc), mini$_{egf}$-(luc), full-(luc), and full$_{egf}$-(luc). Phages were added alone or complexed with TurboFect. For the purified phagemid DNA, pGL3-CMV and precursor plasmids, pM13ori2.cmvluc and pSW10, were also transfected as positive controls. For the phagemid particles, helper phage alone was also transfected as a negative control. Gene expression was assessed at 24, 48, 72, and 96 h via the Luciferase Assay System (Promega) and reported as raw luminescence.

Assessment of cell viability after exposure to phage particles.   HeLa cells were seeded in 96-well plates (Thermo Fisher Scientific) at 5×10^4 cells/mL. The following day, EGF-displaying (M13SW7-EGF) and non-displaying (M13KO7) phage were transfected into each well at concentrations between 5×10^5 and 5×10^{10} virions/mL. As a negative control, dimethyl sulfoxide (DMSO) was added at a concentration of 12%. Cell viability was assayed using 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) after 24 and 96 h. Conversion of MTT to formazan was allowed to proceed for 1 h at 37 °C, at which point 100 µL of DMSO was added to dissolve the formazan. The absorbance of each well was taken at 492 nm on a Varioskan LUX multimode plate reader (Thermo Fisher Scientific). Cell viability was reported as a percentage of the difference between the absorbances of each sample (A_{sample}) and negative control (A_{negative}) relative to the absorbance of the untreated control (A_{NTC}):

$$\text{viability} = \frac{A_{\text{sample}} - A_{\text{negative}}}{A_{\text{NTC}}} \times 100\%. \tag{5.1}$$

Localization of transfected phage particles.   Phage particles displaying EGF (M13SW7-EGF) or not displaying EGF (M13KO7) were purified and concentrated with polyethylene glycol (PEG) in phosphate-buffered saline (PBS), then labeled with Alexa Fluor 488 (Thermo Fisher Scientific). HeLa cells were seeded in 24-well plates at 5×10^4 cells/mL. The following day, labeled phage particles and mini-(gfp) were transfected at 5×10^7 virions/mL. Wells were imaged 1, 6, 24, 48, 72, and 96 h after transfection. To image, cells were fixed with 4% para-formaldehyde and permeabilized with 0.1% Triton-X. Nuclei were stained with 4′,6-diamidino-2-phenylindole (DAPI) and actin was stained with rhodamine phalloidin. Fixed cells were imaged on the EVOS FL Auto Imaging System (Thermo Fisher Scientific) at 40X magnification using the DAPI (nuclei), red fluorescent protein (RFP) (actin), and GFP (phage or expressed GFP) channels.

Transfection of phage in four EGFR$^+$ cell lines.   Cell lines were seeded at the following cell densities in 24-well plates: 5×10^4 cells/mL (HeLa), 1×10^5 cells/mL (HEK293T, MRC-5, HT-29). Phage particles were added to a final concentration of 5×10^7 virions/mL. Helper phage alone (M13KO7) was transfected as a negative control. To assess the influence of a cationic polymer transfection carrier, phage particles were also complexed with 2 µL of TurboFect. Transfection was assessed by fluorescence microscopy on the EVOS FL Auto Imaging System (Thermo Fisher Scientific) or quantified through luminescence (Luciferase Reporter Assay System; Promega) after 96 h. Luminescence was normalized against whole protein content, which was estimated via a bicinchoninic acid (BCA) assay (Thermo Fisher Scientific). The efficiency of gene transfer was reported as luminescence per 100 µg of whole protein content.

Semi-quantification of EGFR across cell lines.   Presence of the EGF receptor across the cell lines HEK293T, HeLa, HT-29, MRC-5, and A549 was visualized via Western blot. Cell lines were propagated in 6-well plates until 80% confluency, at which point they were lysed. Denatured

proteins were separated by size via sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), transferred to nitrocellulose membranes, and probed by an anti-EGFR rabbit mAb (D38B1; Cell Signaling Technology, Danvers, USA). β-actin was used as a loading control (sc-47778; Santa Cruz Biotechnology, Dallas, USA). Secondary antibodies conjugated to horseradish peroxidase (HRP) (Pierce Goat H+L anti-rabbit mAb, Pierce Goat H+L anti-mouse mAb, Thermo Fisher Scientific) were used for visualization with the chemiluminescent Pierce ECL Western Blotting Substrate (Thermo Fisher Scientific). Images were captured using the iBright FL1500 Imaging System (Thermo Fisher Scientific) and analyzed using ImageJ.

Phosphorylation EGFR in HeLa.   We examined if EGF-displaying phage particles were able to activate EGFR. Activation of EGFR by EGF binding, as determined by phosphorylation of the tyrosine residue at 1173, was visualized via Western blot. HeLa cells were treated with EGF$^-$ Helper phage M13KO7, recombinant EGF (Thermo Fisher Scientific), or EGF$^+$ M13SW7-EGF for 5 min prior to cell lysis. In addition, HeLa cells were also pre-treated with EGFR inhibitor gefitinib (Cell Signaling Technologies) 2 h prior to addition of EGF and M13SW7-EGF. The Western blot was performed using antibodies as outlined above, including the use of β-actin as a loading control. Lysates were also probed for phosphorylated EGFR using a phospho-specific EGFR rabbit mAb specifically against Tyr1173 (53A5, Cell Signaling Technology).

Statistical analysis.   All statistical analyses were performed using Python (with the packages NumPy [414], Pandas [415, 416], SciPy [417], scikit-bio [418], and statsmodels [419]). Values were reported as means of n independent experiments with uncertainty reported as standard deviation (SD) or standard error of the mean (SEM), as indicated. Statistical hypothesis tests were evaluated using one-way ANOVA, followed by the Tukey range test for multiple comparisons. Values of $p < 0.05$ were considered statistically significant.

5.3   Results

To assess if miniaturization of the phagemid cassette could improve gene transfer, purified mini and full phagemid ssDNA and encapsulated phage particles were administered across multiple cell lines. Phage particles displaying EGF were evaluated in contrast to phage without any display for its effect on cellular uptake and subsequent gene transfer. Complexation with a commercial cationic polymer transfection reagent was also evaluated.

5.3.1   Transfection of purified single-stranded miniphagemid DNA

To determine the optimal length of time for the assessment of ssDNA transfection, purified plasmid dsDNA and phagemid ssDNA were transfected over period of 96 h in a HeLa cell line (Figure 5.1). As expected, purified plasmid dsDNA transfected well. As a negative control, the empty backbone vector pM13ori2 was also transfected; similarly, pGL3-CMV (the source of the *cmv-luc* cassette)

was transfected as a positive control. No background luminescence was observed after transfection with pM13ori2. Gene expression was also rapid: luciferase activity peaked approximately 24 to 48 h post-transfection. No significant difference was observed between any of the three plasmid vectors. For purified ssDNA, the onset of gene expression was delayed and not as potent. Luciferase activity peaked around 72 to 96 h; overall, gene expression as measured by luciferase activity was approximately 100-fold lower for purified ssDNA compared to each plasmid counterparts. Notably, transfection of purified mini-(luc) ssDNA was correlated with increased luciferase activity over purified full-(luc) ssDNA.
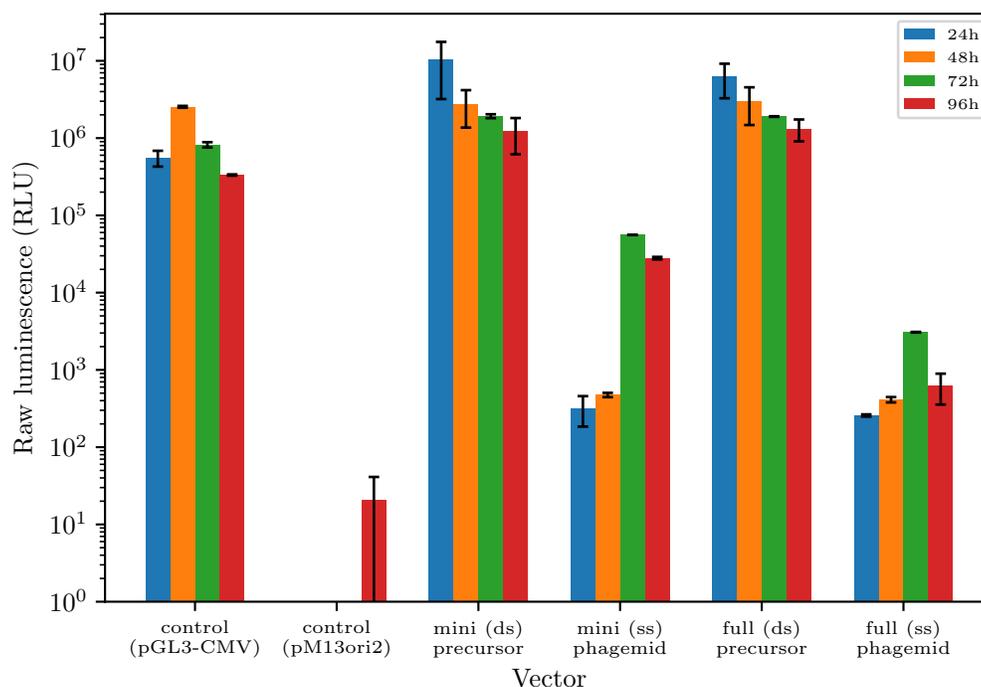
Figure 5.1: Gene expression of precursor vectors and purified phagemid DNA over 96 h. HeLa cells were treated with 1 µg of dsDNA purified plasmid DNA or 2 µg of ssDNA purified phagemid DNA, respectively. DNA was complexed with TurboFect in all cases. Raw luminescence was reported 24, 48, 72, and 96 h after transfection. Error bars represent SD, $n = 3$.

Next, purified ssDNA from mini-(luc) and full-(luc) lysates were transfected into HEK293T as well as HeLa confirm ssDNA expression at 96 h in another cell line (Figure 5.2). As expected, no significant differences were observed between the plasmid vectors; furthermore, all plasmid vectors performed comparably well across both HEK293T and HeLa. Purified dsDNA transfects significantly better than ssDNA across both cell lines. Still, purified mini-(luc) ssDNA resulted in higher gene expression than purified full-(luc) ssDNA, resulting in a ten-fold increase in expression in HEK293T. Overall, ssDNA does appear to be a viable material for gene transfer, albeit at lower efficiency than ds plasmid DNA. From these combined results, 96 h was chosen as the optimal duration for sufficient ssDNA gene expression for analysis.
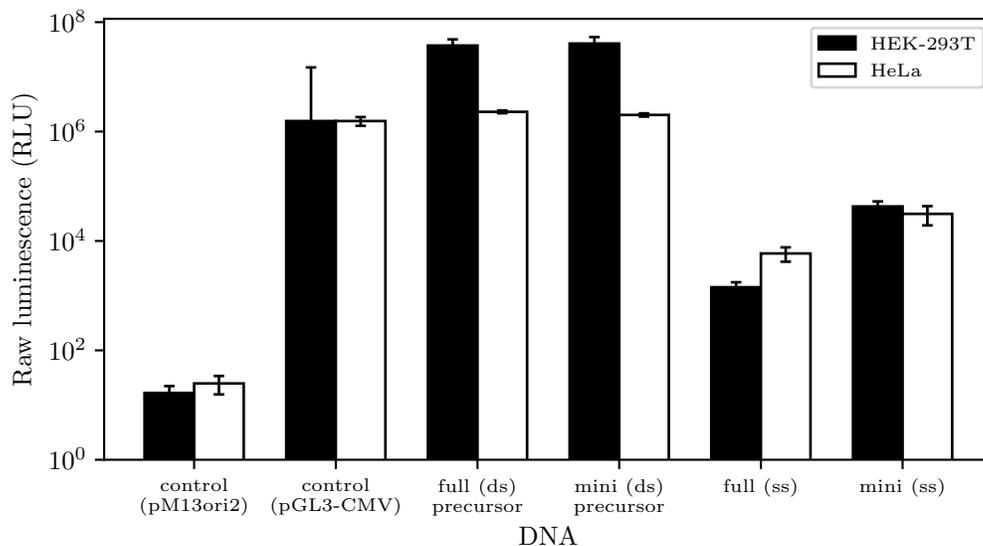
Figure 5.2: Gene expression of precursor vectors and purified phagemid DNA in HeLa and HEK293T after 96 h. Cell lines HEK293T and HeLa were treated with $1\,\mu g$ of dsDNA purified plasmid DNA or $2\,\mu g$ ssDNA purified phagemid DNA, respectively. DNA was complexed with TurboFect in all cases. Error bars represent SD, $n = 3$.

5.3.2 Optimization of phage particle transfection

We next examined luciferase activity after treatment with phagemid particles in the HeLa cell line over time to see if gene expression followed the same patterns as purified ssDNA (Figure 5.3). Gene expression from transfection of mini-(luc), mini$_{egf}$-(luc), full-(luc), and full$_{egf}$-(luc) was measured based on luciferase activity, alongside helper phage alone as a negative control.

Peak luciferase activity was approximately ten-fold lower than the levels observed for the purified phage ssDNA. This is likely due to the differences in number of molecules transfected. Transfection of $2\,\mu g$ of mini-(luc) ssDNA resulted in approximately 1×10^9 phagemid DNA copies available per millilitre, roughly two orders of magnitude higher than the 5×10^7 phage transfected per millilitre here. Due to practical constraints on phage volume, this concentration was selected to minimize the volumes of required lysate. Quantification of lysates showed that full-(luc) and full$_{egf}$-(luc) lysates in particular were mostly comprised of helper phage (Chapter 4), which we were unable to separate away from phagemids at this time. Hence, in order to transfect equal amounts of *cmv-luc* encoding phage without transfecting exorbitant volumes of full phagemids, this amount was chosen as a compromise.

We expected similar trends when transfecting phage particles as with purified DNA. Interestingly, negligible luciferase activity was observed from transfection of the full phagemid without EGF display; it was comparable to that of the negative control (helper only). In contrast, transfection of the mini-(luc) phagemid did result in some luciferase gene expression. For both phagemids, display of EGF dramatically increased gene expression, which is a strong indicator of the necessity

of receptor-mediated cell internalization for phage-mediated gene transfer. Overall, we observed increasing luciferase activity until the end of the analysis period at 96 h. Based on these results, 96 h was selected as a good timeframe to evaluate gene expression in other cell lines.
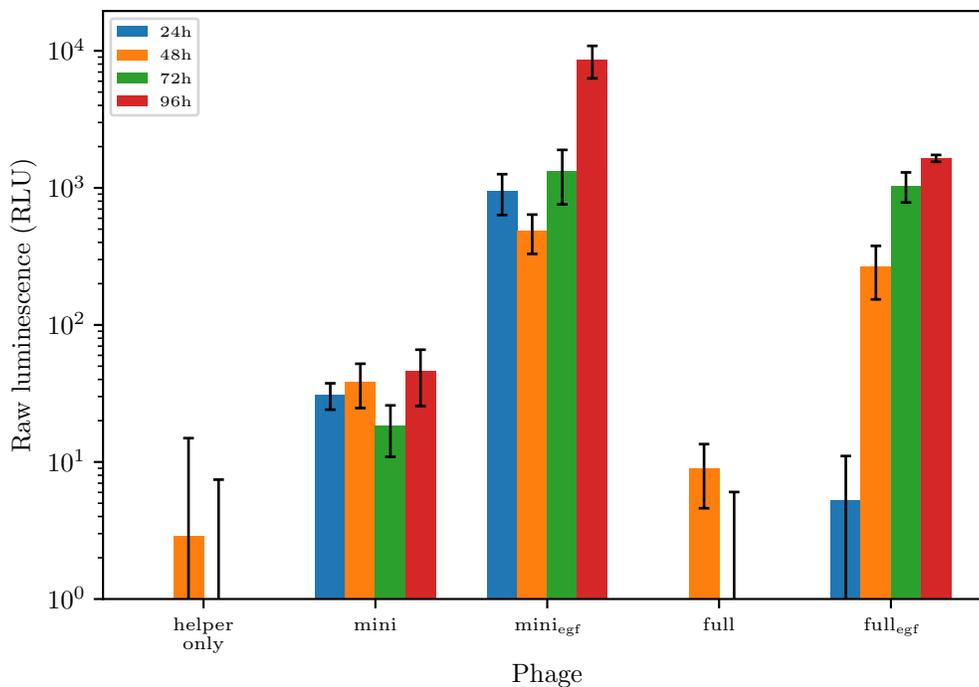
Figure 5.3: Phagemid gene expression in HeLa over 96 hours. HeLa cells were treated with 5×10^7 phage/mL. Raw luminescence was reported 24, 48, 72, and 96 h after transfection. Error bars represent SD, $n = 3$.

Alexa Fluor 488-tagged M13SW7-EGF (EGF$^+$) or M13KO7 (EGF$^-$) were visualized at 1, 6, 24, 48, 72, and 96 h post-treatment in HeLa cells (Figure 5.4). The rightmost column also shows fluorescence from GFP expression over the same timeframe after administration of mini$_{egf}$-(gfp). This phagemid was chosen as its *cmv-luc* counterpart previously showed the greatest luciferase expression of all the phagemids (Figure 5.3). EGF-displaying phage were abundantly localized to the cells within an hour of administration in a ligand-dependent manner, as the same cannot be said for M13KO7. After 6 h, EGF$^+$ phage accumulated around the nucleus, indicating successful cell uptake and cytoplasmic translocation.

Considering the strong perinuclear localization, it is unclear if these particles have escaped the endosomal compartment, as cytoplasmic phage particles are expected to diffuse throughout the cytosol [336]. Alexa Fluor 488-tagged M13SW7-EGF were detectable even after 96 h, although it is unclear if the phage particles are still intact. Consistent with the results from our luciferase assays, expression of GFP from mini$_{egf}$-(gfp) was not detectable until approximately 72 h after transfection. In contrast, few M13KO7 particles were detected even by 96 h. This speaks to the strong influence of the cell-targeting ligand in both cellular localization and internalization.
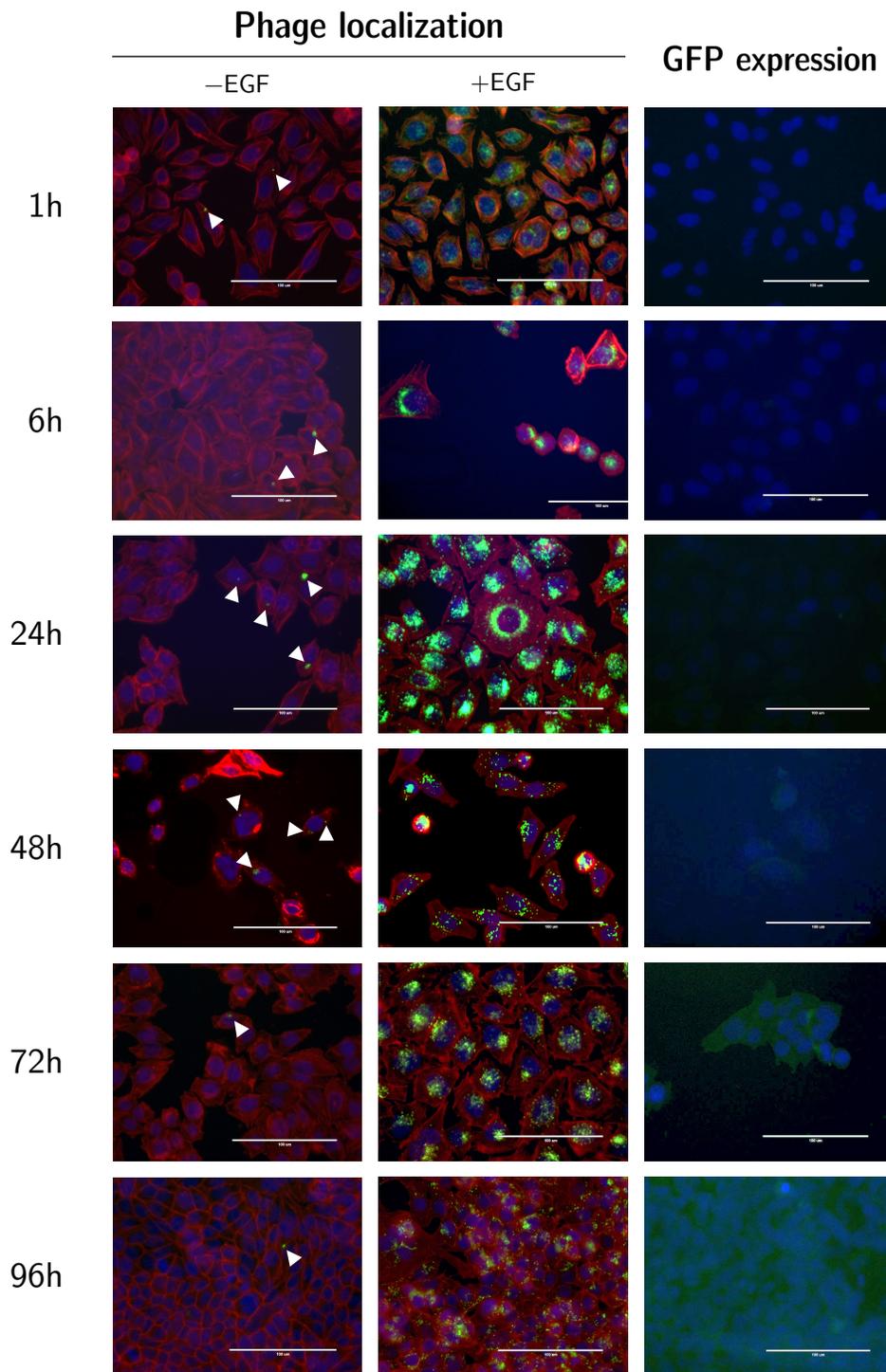
Figure 5.4: Localization of EGF-displaying phage particles in HeLa cells. EGF$^+$ (M13SW7-EGF) and EGF$^-$ phage particles (M13KO7) were tagged with Alexa Fluor 488 (green). Phage were delivered at 5×10^7 phage/mL, then fixed and stained at the indicated timepoints. Cell nuclei were stained with DAPI (blue) while cell cytoskeletons were stained with rhodamine phalloidin (red). Expression of GFP as encoded by mini$_{egf}$-(gfp) is also shown on the right (green). Fluorescent non-displaying phage are indicated with white arrows. Scale bars represent $100\,\mu m$.
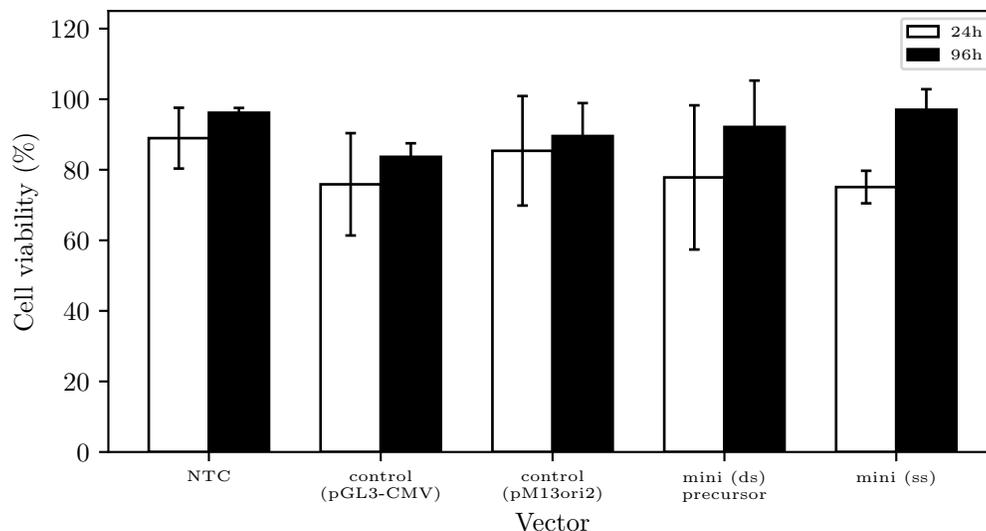
(A) Cell viability after DNA transfection
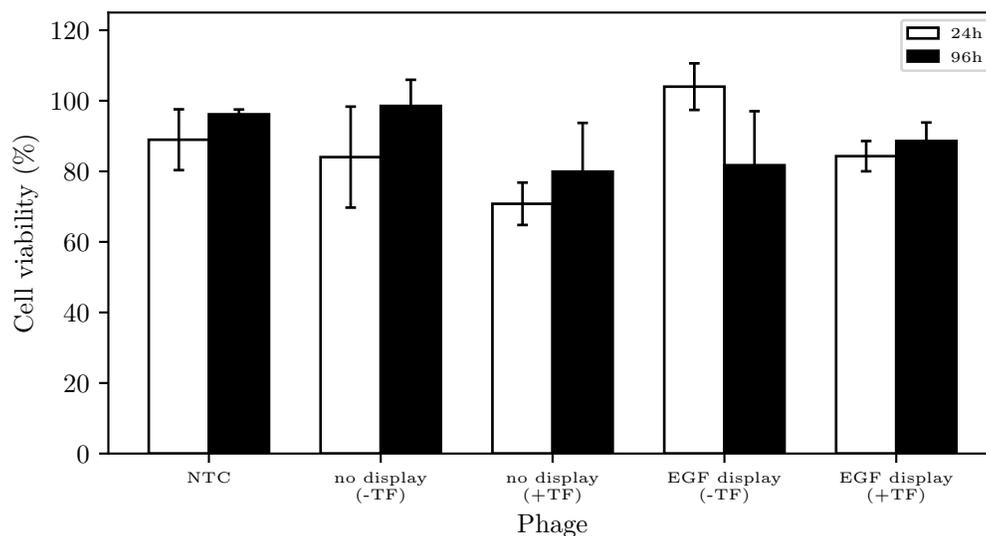
(B) Cell viability after phage transfection

Figure 5.5: Cell viability after vector administration. Purified DNA (A) and phage (B) were administered to HeLa cells. Cell viability assessed through the MTT assay 24 and 96 h post-transfection. A) The plasmid precursor pM13ori2.cmvluc (pDNA) and the corresponding purified ssDNA of mini-(luc) (DNA) were transfected alongside the source plasmid pGL3 and the empty vector pM13ori2 as controls. B) Helper phage M13KO7 (no display) and M13SW7-EGF (EGF display) were transfected alone or with TurboFect. Purified dsDNA was transfected at 1 μg/mL, ssDNA at 2 μg/mL, phage particles at 5×10^7 virions/mL. NTC: no treatment control, TF: TurboFect. Error bars represent SD, $n = 3$.
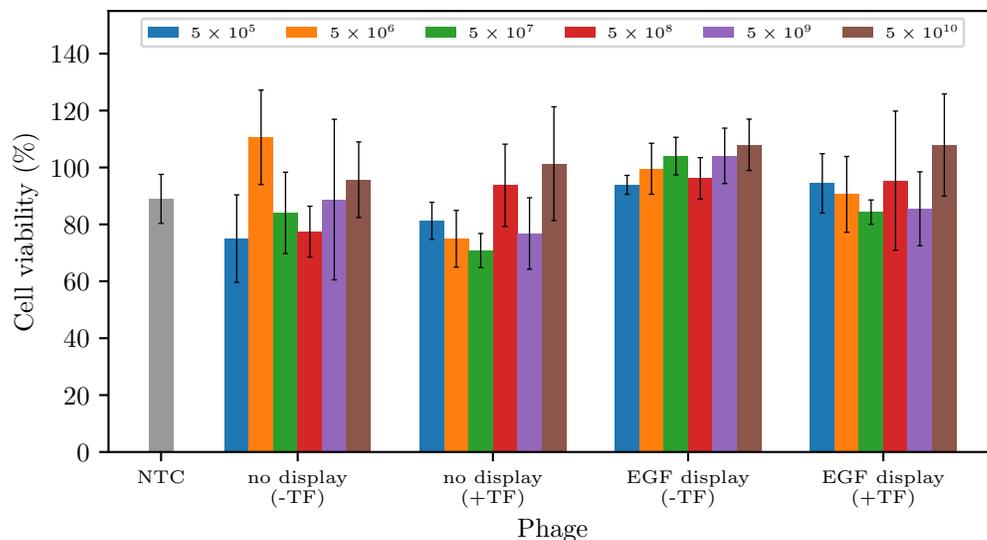
(A) Cell viability after increasing doses of phage at 24 h
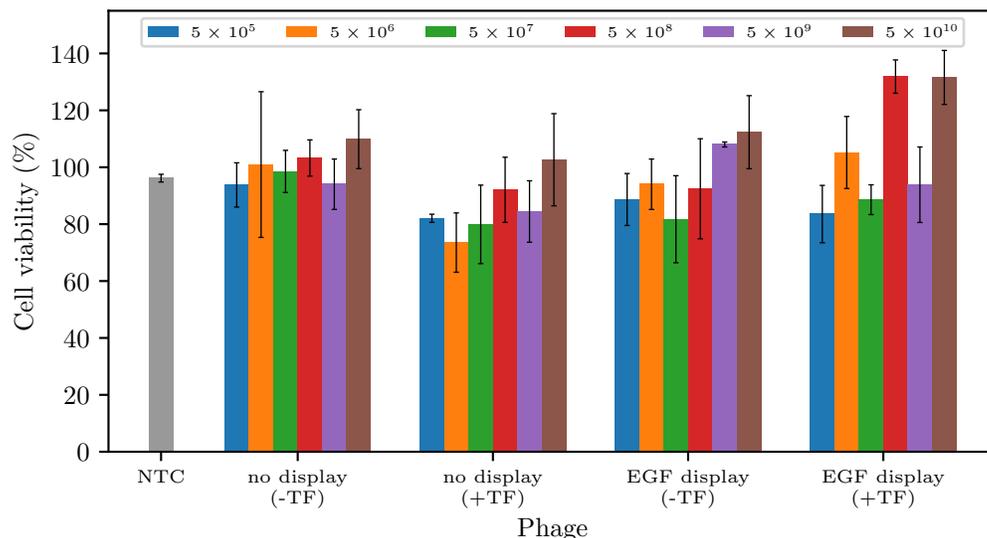
(B) Cell viability after increasing doses of phage at 96 h

Figure 5.6: Increasing phage dose does not decrease cell viability. The non-displaying helper phage M13KO7 and EGF-displaying helper M13SW7-EGF were transfected alone or with TurboFect. Phage were administered to HeLa cells at increasing concentrations between 5×10^5 to 5×10^{10} virions/mL. Cell viability was assessed via the MTT assay at 24 h (A) and 96 h (B) after transfection. NTC: no treatment control, TF: TurboFect. Error bars represent SD, $n = 3$.

We next examined if phage particles were cytotoxic in HeLa cells. Cell viability was assessed using the MTT assay both 24 h and 96 h after treatment with purified DNA or phage particles (Figure 5.5). Phages both displaying EGF (M13SW7-EGF) or not displaying EGF (M13KO7) were assessed across a range of concentrations, between 5×10^5 and 5×10^{10} phage/mL (Figure 5.6). As might be expected, no significant difference was observed in cell viability for all transfections

with purified DNA from the untreated control. Although chemical transfection reagents such as TurboFect can be cytotoxic at high concentrations, the amounts used in this study were insufficient to adversely impact cell viability.

Neither the EGF$^+$ nor the EGF$^-$ phage decreased cell viability, whether or not in combination with TurboFect. Indeed, there was generally no significant difference in cell viability observed from increasing doses of EGF$^-$ or EGF$^+$ phage 24 h after treatment. However, there was a trend towards increased cell viability ($>$100%) with EGF$^+$ phage, particularly between 5×10^8 and 5×10^{10} phage/mL 96 h after treatment. Intriguingly, for these doses of EGF-displaying phage, a statistically significant increase in cell viability was observed as compared to the untreated control both without and with TurboFect ($p < 0.05$).

Overall, however, cell viability was not found to be adversely impacted by phage administration even at high concentrations of filamentous phage M13, regardless of EGF display. As 5×10^7 phage/mL was associated with cell viability on par with the untreated control, this concentration was used for later transfection experiments.

5.3.3   Transfection efficiency of miniphagemid phage particles

The capacity of mini and full phagemids for gene transfer was compared across four cell lines known to express EGFR: HeLa, HT-29, MRC5, and A549. Transfection was also performed in an EGFR$^-$ cell line, HEK293T. EGFR expression was verified via Western blot (Figure 5.7) and relative levels of EGFR expression were reported using β-actin as the loading control. The EGFR$^+$ cell lines appeared to have very similar levels of EGFR; differences were not significant. EGFR expression in HEK293T was negligible, as expected.

(A) Western blot of EGFR expression

(B) EGFR expression relative to β-actin
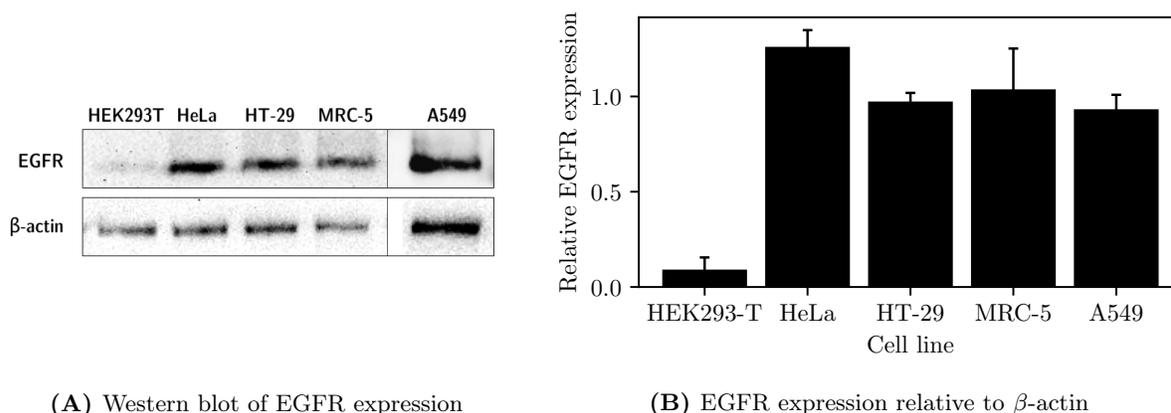
Figure 5.7: EGFR expression across cell lines as measured by Western blot. A) Cell lines HeLa, HT-29, MRC-5, and A549 express EGFR, while HEK293T is negative for EGFR expression. β-actin was used as the loading control. The blot was repeated 3 times and a representative blot is shown here. B) Relative to β-actin, the levels of EGFR expression are reported per cell line. Error bars represent SD, $n = 3$.

Gene expression was reported as luciferase activity per $100\,\mu g$ whole protein content $96\,h$ post-transfection (Figure 5.8). The effects of the display of EGF and complexation with TurboFect were assessed for both miniaturized and non-miniaturized phagemid particles. Fold differences in luciferase activity from transfection of the mini and full phagemids generally showed an increase in gene expression as a result of phagemid miniaturization (Table 5.4). More dramatically, fold differences in luminescence between EGF-displaying and non-displaying counterparts generally showed much larger increase in gene expression from the display of EGF (Table 5.5).

In general, luciferase activity was highest in HEK293T and HeLa. The biggest factor contributing to increased gene expression was the display of the cell-specific ligand in the EGFR$^+$ cell lines. The miniaturized phagemid also conferred increased luciferase activity as compared to its full counterpart. This increase in gene expression was statistically significant in combination with the display of EGF *and* when complexed with TurboFect in cell lines HeLa, HT-29, and A549 ($p < 0.05$).

Table 5.4: Fold difference in luciferase expression between mini and full phagemids

| | Fold difference in gene expression (mini/full)[1] | | | |
| | EGF$^-$ | | EGF$^+$ | |
Cell line	$-$TurboFect	$+$TurboFect	$-$TurboFect	$+$TurboFect
HEK293T	5.22 ± 5.63	4.21 ± 3.79	2.03 ± 2.59	1.78 ± 1.79
HeLa	∞^2	0.51 ± 0.12	1.98 ± 1.92	1.48 ± 0.37
HT-29	∞	7.44 ± 7.63	1.74 ± 1.40	2.81 ± 0.67
MRC-5	0.52 ± 1.04	∞	0.43 ± 0.33	3.51 ± 1.73
A549	∞	∞	3.20 ± 0.05	2.08 ± 0.18

[1] Uncertainty reported as SEM$\times 2$, $n = 3$.　　[2] Luminescence was below threshold for EGF$^-$ full phagemid.

Table 5.5: Fold difference in luciferase expression between phage based on EGF display

| | Fold difference in gene expression (EGF$^+$/EGF$^-$)[1] | | | |
| | mini | | full | |
Cell line	$-$TurboFect	$+$TurboFect	$-$TurboFect	$+$TurboFect
HEK293T	11.73 ± 11.07	7.17 ± 4.24	43.94 ± 59.44	17.65 ± 10.54
HeLa	25.59 ± 33.01	704.94 ± 351.36	∞^2	261.37 ± 182.45
HT-29	∞	∞	∞	∞
MRC-5	∞	5.73 ± 2.79	4.14 ± 0.47	∞
A549	∞	9.96 ± 3.34	∞	∞

[1] Uncertainty reported as SEM$\times 2$, $n = 3$.　　[2] Luminescence was below threshold for EGF$^-$ phagemid.
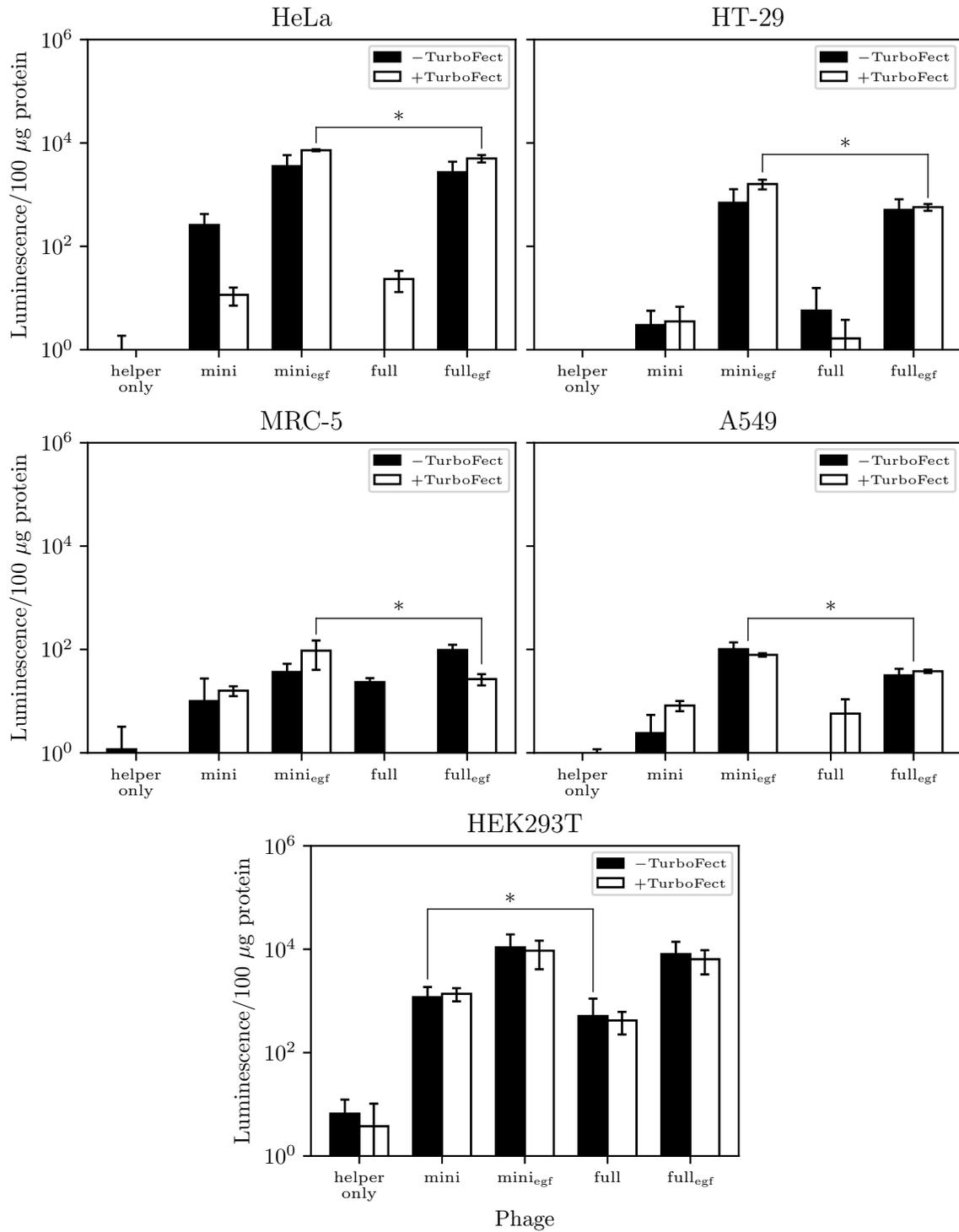
Figure 5.8: Luciferase gene expression of phage-delivered transgenes. Cell lines HeLa, HT-29, MRC-5, A549, and HEK293T were treated with miniphagemids encoding *cmv-luc*. The display of the cell-specific ligand EGF was compared to phagemids without any display. Gene expression is reported per $100\,\mu g$ of whole protein content. The impact of the commercial cationic polymer reagent, TurboFect, was also assessed. Error bars represent SD, $n = 3$. The $*$ above the bars indicates a difference at significance level $p < 0.05$.

In the case of the EGFR$^-$ control, HEK293T, mini-(luc) was correlated with increased luciferase activity over full-(luc) when complexed with TurboFect ($p < 0.05$). However, the increase in gene expression conferred by miniaturization of the phagemid was not statistically significant in the absence of TurboFect in HEK293T. Interestingly, no statistically significant difference in gene expression was observed between the mini and full phagemids when they displayed EGF in HEK293T. Display of the EGF was associated with an increase in gene expression for full$_{egf}$-(luc) over full-(luc) when complexed with TurboFect ($p < 0.05$), but this effect was not significant in the absence of the cationic polymer. The display of EGF also did not lead to a significant increase in luciferase activity for mini$_{egf}$-(luc) over mini-(luc), interestingly. Moreover, the improvement in gene expression conferred by EGF display when targeting this EGFR$^-$ cell line was only approximately twofold.

In contrast, display of a cell-targeting ligand was, as expected, much more impactful on resultant gene expression if the cell line expressed the targeted receptor. Gene expression was maximally 700-fold greater in HeLa, and over 100-fold greater in HT-29 when the miniphagemid displayed EGF. Often, luciferase activity was below the threshold for detection after administration of full phagemids that did not display EGF.

Miniphagemids performed better than their full counterparts only if the phagemid also displayed EGF. The addition of TurboFect tended towards positive enhancement of gene transfer, but it alone did not contribute a statistically significant difference in gene expression. Notably, variability in gene expression was reduced when phages were complexed with TurboFect prior to treatment. Overall, mini$_{egf}$-(luc) complexed with TurboFect tended to perform better than full$_{egf}$-(luc) with TurboFect across all cell lines ($p < 0.05$). More specifically, in A549, the difference was significant regardless of the addition of TurboFect.

Additionally, cells transfected by mini$_{egf}$-(gfp) or full$_{egf}$-(gfp) were visualized by GFP fluorescence 96 h after transfection (Figure 5.9). Similarly, Figure 5.10 shows cells transfected by EGF$^-$ phage. Cells were all over 100% confluent at this stage. As expected, few cells exhibited any fluorescence when transfected by phage without the targeting ligand. While some GFP fluorescence was detectable across all cell lines, it was generally weak. Although consistent with our observations of ssDNA-mediated luciferase expression (purified or phage-encapsulated), these results show that phage-mediated transfection efficiency (number of cells transfected) may be poor overall although the efficiency of gene expression per cell can be improved using miniaturized phagemids.
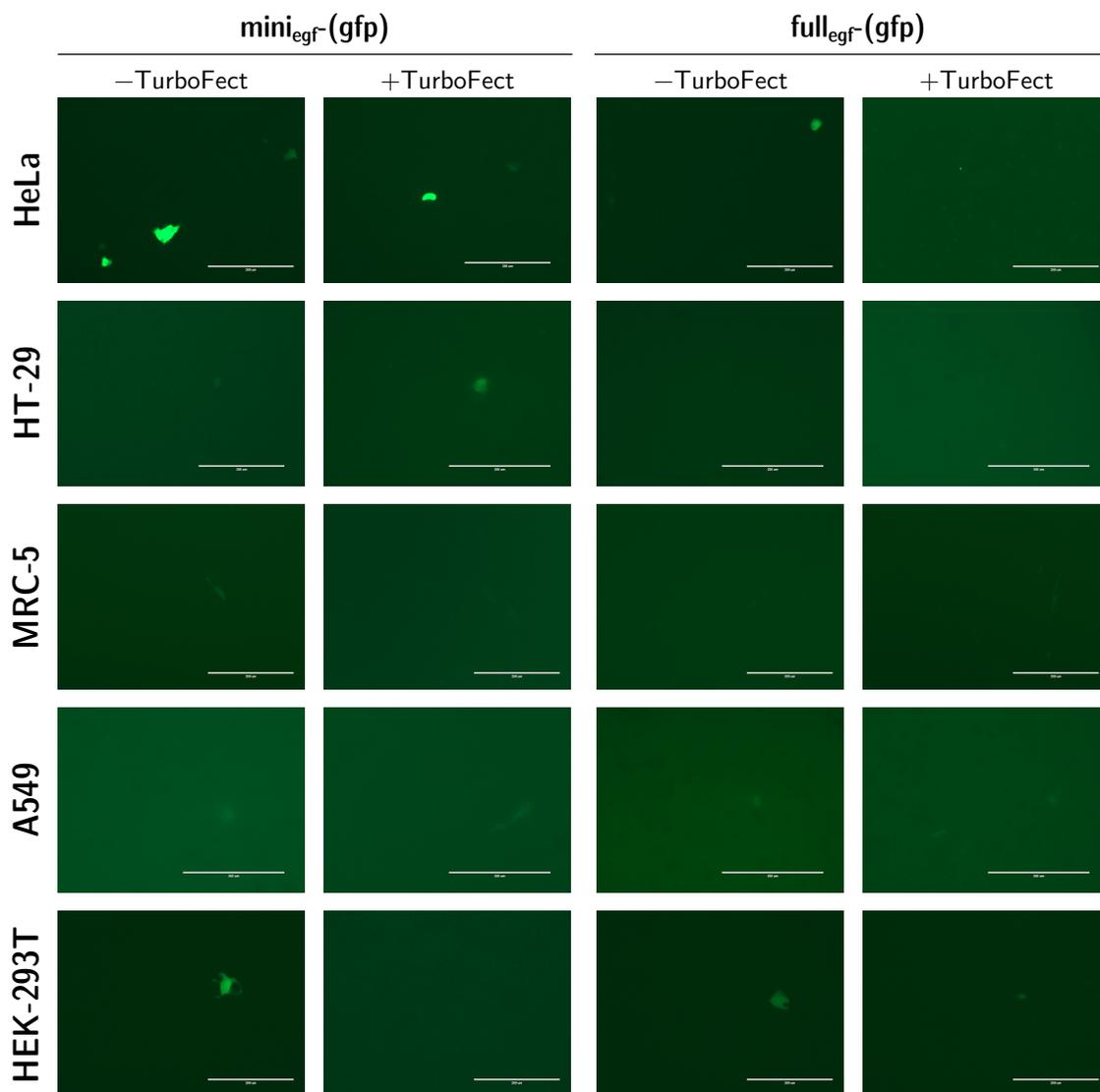
Figure 5.9: GFP gene expression of EGF-displaying phage-delivered transgenes. Cell lines HeLa, HT-29, MRC-5, A549, HEK293T were treated with EGF-displaying mini or full phagemids encoding *cmv-gfp*. Transfections were compared with or without the transfection reagent TurboFect. Scale bars represent $200\,\mu m$.
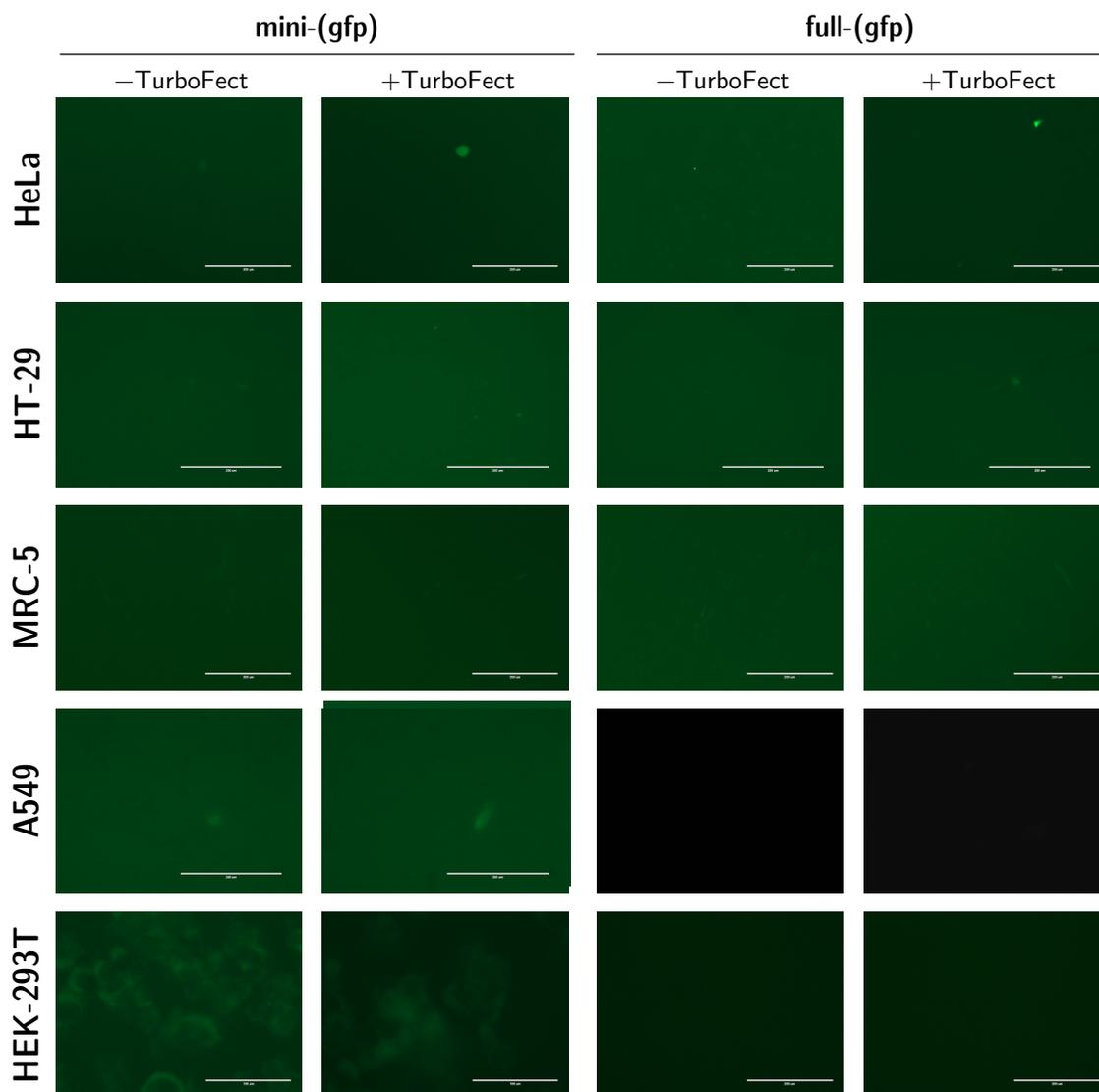
Figure 5.10: GFP gene expression of phage-delivered transgenes. Cell lines HeLa, HT-29, MRC-5, A549, HEK293T were treated with mini or full phagemids without any targeting ligands encoding *cmv-gfp*. Transfections were compared with or without the transfection reagent TurboFect. Scale bars represent 200 µm.

5.3.4 Effect of the target receptor on phage-mediated transfection

Based on the results from the MTT assay (Figure 5.6), we next asked if the filamentous phage particle displaying EGF was able to lead to receptor activation. Activation of the EGF receptor stimulates signal transduction pathways involved in cell proliferation, which may be related to the apparent increase in cell viability previously observed.

Canonically, ligand-EGFR is necessary for receptor dimers to autophosphorylate prior to internalization via clathrin-coated endocytosis. Figure 5.11 shows a Western blot of the HeLa cell lysates after treatment with EGF^- or EGF^+ phage in comparison with treatment with purified recombinant EGF. Phosphorylation of the tyrosine at position 1173 occurs within minutes of EGF binding and is one of the necessary steps for downstream signal transduction. EGF-mediated activation of EGFR is also prevented when in presence of the RTK inhibitor, gefitinib [494]. As seen in Figure 5.11, phosphorylated EGFR is detectable after treatment with EGF but not if cells were pre-treated with gefitinib. In contrast, no phosphorylation is detected when treated with EGF^+ phage only in absence of gefitinib. Based on these results, EGF-displaying phage are likely internalized through EGFR-mediated endocytosis, but this does not appear to lead to EGFR phosphorylation.
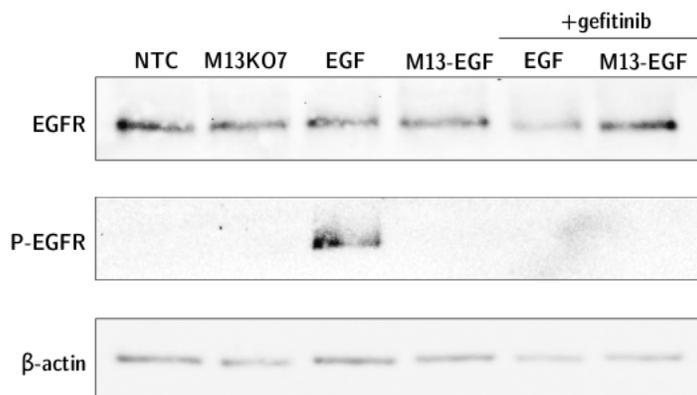
Figure 5.11: EGF-displaying M13 does not activate EGFR as measured by Western blot.
From left to right, untreated HeLa cell extracts were compared to cells treated with M13KO7 (no display), purified recombinant EGF (100 ng/µL) or M13SW7-EGF (EGF display) for 5 min prior to analysis. Additionally, cells were also pre-treated with the EGFR inhibitor, gefitinib (10 µM), prior to treatment with EGF or M13SW7-EGF. From top to bottom, cell lysates were probed for presence of EGFR or phosphorylated EGFR (Tyr1173). β-actin was used as the loading control. Phage were added at 5×10^{10} virions/mL. NTC: no treatment control, M13-EGF: M13SW7-EGF, P-EGFR: Tyr1173.

5.4 Discussion

5.4.1 Display of a cell-specific ligand determines phage cellular uptake

The use of EGF as a cell-targeting ligand has been shown to improve the internalization of filamentous phage in previous studies and consequently, improved gene transfer, both in vitro and in vivo [333, 347, 450, 451]. Using the EGFR$^+$ cell line, HeLa, we observed cellular localization of phages within 1 h of administration (Figure 5.4). Within 1 h, phages were likely already internalized although we did not differentiate between surface-adhered or internalized phage. Upon ligand binding, EGF-EGFR complexes can internalize within 15–20 min [501]. Others have observed phage internalization within 10–60 min [328, 336, 358, 502]. Our results appear consistent with these previous findings.

In phage transcytosis studies, Nguyen et al. (2017) observed significant cytoplasmic translocation by phage T4 within 2 h [330]. Juxtanuclear localization was evident by 6 h in several images, as fluorescently-tagged phages particles were clearly seen surrounding the DAPI-stained nucleus but not within it. Upon ligand binding, internalized EGFR complexes are routed to lysosomes for degradation; therefore, EGFR-bound phage may be prone to accumulate within juxtanuclear lysosomes, which positions them perfectly for subsequent escape and nuclear transport. Expression of GFP was detectable by fluorescence microscopy after 72 h, which is similar to results from delivery of purified ssDNA. This suggests that the additional requirement of DNA separation from the filamentous phage coat may not contribute as much of a hindrance to phage-mediated gene transfer as previously thought. Indeed, the low luminal pH of juxtanuclear lysosomes can contribute to phage coat shedding [503]. From these results, we assert that the two obstacles, lysoendosomal escape and nuclear penetration, are the key bottlenecks to successful phage-mediated gene transfer.

To assess the utility of EGF as the targeting ligand in this proof of principle, we evaluated target receptor expression across the cell lines HeLa, HT-29, MRC-5, A549, and HEK293T (Figure 5.7). As expected, we did not see endogenous expression of EGFR in HEK293T [504–506]. Relative to HeLa, we found that EGFR levels in A549 and HT-29 were similar when normalized to the loading control, β-actin. Elsewhere, HeLa, A549, and MRC-5 have all been reported to express moderate levels of EGFR [505, 507–509]. More specifically, HeLa and A549 have been reported to have similar endogenous levels of EGFR; Zhang et al. (2015) measured EGFR density in HeLa and A549 to to be approximately 270 and 142 receptors per μm^2, respectively [505]. Although both we and others [508, 509] show endogenous EGFR expression in MRC-5, this has been contested by other studies [510]. Chowdhury et al. (2014) reported MRC-5 to be in the same expression level group as HeLa and A549 [509], which we similarly observed here. HT-29 has been reported to have moderate to high levels of EGFR expression amongst colorectal cancer cell lines [511, 512], although conflicting observations have been reported [513–516]. Our results show that HT-29 exhibits similar levels of EGFR in comparison to HeLa, MRC-5, and A549.

While all the EGFR$^+$ cell lines exhibited similar EGFR expression to HeLa, treatment with mini$_{egf}$ or full$_{egf}$ phagemids did not result in comparable levels of *luc* expression (Figure 5.8). Although display of an EGFR-targeting ligand appeared necessary for gene transfer in many of the EGFR$^+$ cell lines, it alone appears insufficient for high levels of transgene expression. The rate of phage internalization is very much tissue-dependent; for example, Bichet et al. (2021) identified slow uptake of different bacteriophages in several cell lines, including HT-29 [328].

On the other hand, robust luciferase expression was observed in HEK293T, which does not express detectable levels of EGFR (Figure 5.8). In this cell line, although greater gene expression was observed with the EGF-displaying phage, the difference in gene transfer imparted by EGF$^+$ versus EGF$^-$ phagemids was not statistically significant. Regardless of EGF display, the miniaturized phagemid was associated with greater luciferase activity: up to a 5-fold increase in luciferase activity over its full counterpart (Table 5.4). In the absence of cell uptake mediated by ligand-receptor interactions, filamentous phage have been observed to enter via caveolae-mediated endocytosis [334, 517]. Larger phage particles must enter via phagocytosis and macropinocytosis [329]. In the absence of a target receptor, filamentous phage uptake in HEK293T is likely clathrin-independent, but the specific endocytic mechanism remains unclear. The high rate of phage-mediated gene transfer into HEK293T independent of a receptor-targeted ligand presents an exciting avenue for further investigation. Overall, our results here are generally consistent with previous reports of EGF-mediated improved phage internalization and subsequent gene transfer [450, 451, 518].

5.4.2 TurboFect may improve phage-mediated gene transfer

The combination of biological gene transfer vectors with so-called traditional chemical reagents has been investigated on many occasions. We observed that even with the display of a cell-specific ligand, luciferase activity was greater when complexed with a cationic polymer. This is consistent with the observations of Donnelly et al. (2015), as they also found that the cationic polymer transfection reagent PEI improved filamentous phage-mediated gene transfer [477].

In contrast to the work by Donnelly et al. [477], however, we did not observe a large increase in phage-mediated gene transfer when complexed together with a cationic polymer. Of note, we used a different cationic polymer, the commercial reagent TurboFect, not PEI. Additionally, the amount of TurboFect used in this study was low, but greater amounts were likely to lead to greater cytotoxicity. Donnelly et al. proposed that the complex of cationic polymer and phage were able to better penetrate cell membranes, which would suggest these complexes bypass receptor-mediated endocytic uptake altogether. Another possibility is that the polymer-phage complex can better escape the endosomal compartment, perhaps through the proton sponge effect. The increase in gene transfer due to TurboFect was only evident in conjunction with the display of the cell-specific ligand in this study. TurboFect enhanced EGF-dependent gene transfer but did not appear to permit EGF-independent gene transfer, thereby supporting the latter hypothesis. However, the use of such reagents in vivo is highly undesirable as they can be toxic and expensive to scale up.

Notably, all phage lysates in these transfection studies were purified and concentrated with PEG, which is the most common approach to filamentous phage purification. Residual PEG is known to remain in the phage lysate [519, 520], which actually may be an advantage for gene transfer. The addition of PEG to chemical non-viral gene delivery vehicles is associated with improved cellular uptake and intracellular translocation [521–523]. PEG has also been suggested to improve systemic delivery of plasmid DNA [524, 525], as it can protect circulating DNA complexes from clearance by blood serum components.

Investigation into combinations of other cationic polymer reagents with phage may more prove fruitful; hence, it may be important to remove PEG as a confounding factor. It would be elucidating to examine phage-mediated gene transfer with cationic polymers in the absence of PEG. A number of different strategies for purifying filamentous phage have been investigated [520], including ultracentrifugation, size-exclusion chromatography, and others. More recently, filamentous phage have been further purified from PEG-precipitated lysates based on the isoelectric point of the phage [430, 519, 520].

5.4.3   The single-stranded miniphagemid improves gene transfer

Both purified and phage-encapsulated miniphagemid DNA were correlated with greater luciferase activity over their full phagemid counterparts. However, this increase in luciferase expression as conferred by the miniaturized *luc* phagemid was generally only evident when the phage particles also displayed EGF in EGFR-positive cell lines. This is to be expected as the phage must first be taken up through receptor-ligand interactions in order for any benefits from improved cytoplasmic trafficking can manifest.

Another benefit conferred by the smaller size of the phagemid particle may be more efficient internalization upon ligand-receptor binding. Clathrin-mediated endocytosis has been previously shown to accommodate filamentous phage particles up to 900 nm in length [336], although clathrin-coated vesicles are typically on the scale of 200 nm [526]. In general, particles smaller than 500 nm are more likely to be taken up through clathrin or caveolae-mediated endocytosis, while larger particles are likely internalized through other mechanisms such as phagocytosis or macropinocytosis [527, 528]. Although long, their flexible rod-like structure enables filamentous phages to be compacted; as such, they are more readily taken up alongside receptor-mediated clathrin or caveolae-mediated endocytosis in contrast to other larger, more globular proteins and phages. This effect is therefore likely enhanced by the short nature of the miniphagemid particles since filamentous phage length is determined largely by the length of the encapsulated DNA molecule. The lengths of $mini_{egf}$-(luc) and $mini_{egf}$-(gfp) are less than half that of the wild-type M13 genome (2.5–3 knt versus 6.4 knt) so the corresponding phagemid particles would be 400 nm or shorter.

While luciferase activity peaked around 72 h for purified ssDNA, luciferase activity from the delivery of $mini_{egf}$-(luc) was still increasing at 96 h post-transfection. Investigations into long-term miniphagemid-mediated gene transfer in future time course studies could be very interesting.

Furthermore, these phagemids displayed EGF on pIII, which is present in five copies on one end of the phage particle. Display of targeting ligands on the much more highly abundant pVIII may improve the likelihood of ligand-receptor binding and further improve cell internalization and subsequent gene transfer. This has been demonstrated in other phage-mediated gene transfer applications using a different cell-specific ligand, RGD (arginine-glycine-aspartate) [529]. However, pVIII cannot accommodate homogeneous display of the entire EGF peptide (53 residues); thus either shorter EGFR-targeted peptides or the use of Type 88 or Type 8+8 display system for heterogeneous display must be incorporated.

5.4.4 Internalization of EGF-displaying M13 does not activate EGFR

Despite the use of EGFR as an internalization target in multiple studies, activation of EGFR-mediated signal transduction by EGF-displaying phage has not been deeply investigated. M13 is not believed to prohibit EGFR dimerization, but it is unclear if it interferes with receptor phosphorylation events. Larger molecules such as the monoclonal antibody cetuximab [530] have been shown to internalize with EGFR as antibody-receptor complexes without triggering the downstream signalling pathway [531]. Internalization of bacteriophage λ displaying EGF was demonstrated in HT-29 cells [458], also independent of signal transduction. In fact, EGFR-targeted λ phage were shown to reduce cell proliferation, which was believed to be due to antagonistic receptor binding by the EGF-displaying λ phage. It is suggested that cetuximab-EGFR internalization is mediated through macropinocytosis [531], which is also the likely mechanism of uptake for a large phage particle such as λ. However, it is unclear if this may also be the primary mechanism by which the rod-like filamentous phage M13 may be internalized.

A number of studies have demonstrated that ligand-induced EGF receptor dimerization is the primary mechanism by which ligand-EGFR endocytosis occurs and that it occurs independently of EGFR autophosphorylation [495, 532–534]. Although the cell viability results from our MTT assay indicated a possibility for EGFR-induced cell proliferation in cells treated with EGF$^+$ phage (M13SW7-EGF), we did not observe autophosphorylation of EGFR after phage treatment. EGFR phosphorylation is known to occur rapidly after ligand-binding; hence, we examined EGFR phosphorylation 5 min after treatment (Figure 5.11). No phosphorylation was observed as a result of EGF$^+$ phage treatment, although phosphorylated EGFR was detected after purified EGF treatment. We also examined cell lysates after treatment with EGF or M13SW7-EGF for 1 h, but did not observe any phosphorylation after either treatment. Since receptor-mediated phage internalization is known to occur within 15 to 60 min, it is possible that phage-displayed EGF requires longer than 5 min (but less than 1 h) to activate EGFR.

Interestingly, the use of filamentous phage with pIII display of EGF to perturb EGFR signalling pathways has been reported [449]. Souriau et al. (1997) observed induction of c-fos serum response element (SRE)-mediated transcription, a late event in the EGFR-stimulated pathway, triggered by treatment of EGFR$^+$ cells with EGF-displaying phage. It has been demonstrated elsewhere that

SRE-mediated transcriptional activation can occur even in cells with kinase-defective EGFR [535]. This suggests alternative mechanisms to stimulate EGFR-mediated signal transduction separately from the receptor's kinase activity, which warrants further investigation in future studies with EGF-displaying phage. Additionally, nonactivated EGF-EGFR complexes within intracellular endosomes can still autophosphorylate and transduce signals intracellularly [536–538]. Endosomal EGFR signalling was demonstrated by Wang et al. (2002) by treating cells with EGF in the presence of a tyrosine kinase inhibitor and a EGFR recycling inhibitor, then removing both inhibitors after EGF-EGFR internalization [537]. Thusly, downstream cell proliferation pathways could be activated in the absence of cell surface EGFR phosphorylation. Overall, we showed that EGF displayed on filamentous phage did not activate EGFR by way of phosphorylation of the key tyrosine residue 1173, although this does not rule out endosomal EGFR signalling nor other kinase-independent signal transduction.

5.4.5   Summary

Internalization of EGF-displaying phage was observed to lead to juxtanuclear localization, likely via the lysoendosomal pathway. This was associated with vastly improved gene transfer in EGFR$^+$ cell lines, particularly in HeLa, over phage that did not display a ligand. Although it is known that endocytosis of the EGF receptor is generally precipitated by autophosphorylation and receptor activation, it remains unclear if EGF-displaying filamentous phage leads to EGFR activation.

Overall, the display of a cell-specific targeting ligand greatly enhanced gene transfer due to ligand-receptor endocytosis enabling the bypass of a significant obstacle to gene transfer: the plasma membrane. Furthermore, complexation of phage particles with a cationic polymer also further enhances gene transfer; we postulate this effect likely manifests by improving phage endosomal escape. Future developments in bacteriophage gene delivery must address the need for endosomal escape in order to increase intracellular phagemid bioavailability. Finally, reduction in phagemid size through removal of the bacterial backbone was found to improve phage-mediated gene transfer, likely by increasing cytoplasmic diffusion towards the nucleus.

Chapter 6

M13-mediated assembly of a linear covalently-closed double-stranded DNA minivector

6.1 Introduction

In Chapters 3 and 5, we characterized the production of a single-stranded (ss) DNA minivector phagemid ("miniphagemid") encoding a mammalian transgene cassette and demonstrated their capacity for gene transfer in mammalian cells. Although the filamentous bacteriophage M13 has been successfully used for gene transfer both in vitro and in vivo [326, 332, 347], one major disadvantage to its usage in gene delivery is its single-stranded nature [344]. In general, ssDNA transfects more poorly compared to double-stranded (ds) DNA [338, 342]. Here, we propose to construct a double-stranded miniaturized phagemid for gene delivery to mammalian cells. Using the split origin system for minivector construction we previously characterized (Chapter 3), our aim here is to produce a dsDNA miniphagemid that only encodes a mammalian transgene, without any prokaryotic backbone.

6.1.1 Transfection of double-stranded DNA

The efficiency of gene transfer even by cell-targeted filamentous phages is around 10% [342, 344, 518], which is lower than other double-stranded phages. Similarly to ssDNA eukaryotic viruses such as the adenoviral associated virus (AAV), the M13 ssDNA molecule must be converted to a duplex form prior to transcription, a rate-limiting step that is highly inefficient in mammalian cells [344, 539]. To overcome this disadvantage, the incorporation of some duplex structures has improved M13-mediated gene transfer [339, 343]. The integration of inverted terminal repeats (ITRs) from AAV was found to improve M13-mediated gene transfer [339, 340]; in AAV, these repeats form hairpins that are thought to induce synthesis of the second strand. Upon infection, AAV is

converted to a dsDNA episome induced by the hairpin secondary structure of its ITRs. Indeed, phage transfection was dramatically superior in the presence of a genotoxic drug like camptothecin, which increased transfection efficiency from 10% to 30–45% [344]. DNA damaging treatments including camptothecin, ultraviolet (UV) irradiation, and heat shock have been associated with increased transduction of both AAV [540] and M13 [344, 541]; it is postulated that the improved gene transfer arises from increased second strand synthesis through the activation of DNA repair mechanisms. Although it remains unclear if the introduction of AAV ITR elements induces the formation of dsDNA M13 vectors in mammalian cells, Hajitou et al. (2006) showed the chimeric phage vectors had increased stability, possibly through the formation of concatemers, and improved gene expression [339].

Phagemid vectors with self-complementary regions encompassing only the promoter as well as the entire transgene cassette were associated with superior gene transfer, without any observable decrease in phage titre [343]. Prieto and Sànchez (2007) observed that even a partial duplex region encompassing only the promoter was sufficient to improve gene expression. They postulated that the double-stranded sequence upstream of the gene cassette could function as a primer to stimulate polymerization of the downstream cassette or that binding of RNA polymerase to the dsDNA promoter may be sufficient to induce transcription-coupled DNA repair, again, stimulating polymerization of the downstream DNA sequence [343]. However, in these applications of phage-mediated gene transfer, a prokaryotic backbone was still present in the delivered DNA molecule.

6.1.2 DNA minivectors

It has been well-documented that foreign DNA, such as viral or bacterial DNA, will induce inflammatory responses [54–56]. Prokaryotic DNA is rich in unmethylated cytosine-guanine dinucleotides (CpGs) [59], which are a type of pathogen-associated molecular pattern (PAMP) and are recognized by the innate immune system [60–62]. Abundant in the endosomal compartments of plasmacytoid dendritic cells (pDCs) and B cells, the Toll-like receptor (TLR) 9 mediates innate immunity against cell-internalized DNA containing unmethylated CpG motifs (CpG-DNA) [65, 66]. Recognition of CpG-DNA by TLR9 stimulates the TLR/interleukin (IL)-1R signalling pathway, activates downstream inflammatory pathways [36, 55]: notably, nuclear factor (NF)-κB and mitogen-activated protein kinase (MAPK)). The TLR9 signalling pathway is dependent on myeloid differentiation primary-response protein (MyD) 88 [71, 72] to activate the MAPK signal cascade and NF-κB-mediated transcription of pro-inflammatory genes [36, 77]. CpG immunostimulation strongly resembles a T helper lymphocyte (Th)-mediated immune response, specifically that of Th1 [36, 67, 68]. It is characterized both by immune cell stimulation and cytokine production [36, 78]. CpG immunostimulation is thought to be a primary factor in the silencing of transgene expression, which can prohibit long-term expression of exogenous DNA [83, 84]. Vectors absent in prokaryotic backbones ("minivectors") have shown improved transfection efficiencies and prolonged gene expression [58, 85, 86].

Double-stranded DNA minivectors [50, 89] lack the bacterial backbone present in conventional plasmids. Minivectors are minimized double-stranded circular covalently closed (CCC) or linear covalently closed (LCC) molecules containing only the gene expression cassette. As minivectors are also much smaller than their isogenic plasmid counterparts, they exhibit improved cytoplasmic diffusion [38], which can further improve gene transfer from improved nuclear translocation.

Circular minivectors (minicircles) are generated by using recombinases to excise the minivector from a precursor plasmid; minicircle recombinase systems have included λ integrase [92, 93], P1 Cre recombinase [94], ParA resolvase [95, 96], or ϕC31-integrase [97, 98]. Recombination activity at target sites on the precursor produces multiple species: minicircles, miniplasmids containing backbone elements, minivector intermediates, and residual parent plasmid. A purification step is required to separate minicircle DNA from residual parent plasmid and other byproducts. Similarly, dsDNA linear ministring vectors are produced using different recombinase enzymes, but generate LCC vectors instead of circular ones [99]. A number of LCC minivectors have also been produced in vitro [102–104, 542, 543], generally through capping of open linear dsDNA DNA molecules. Linear minivectors offer a superior safety and transfection efficiency profile over their isogenic plasmid and even their minicircle counterparts [100]. Recombination between a linear vector and a chromosome results in double-stranded DNA break (DSB) that activates cell death pathways in both prokaryotic [99] and eukaryotic systems [100], thus safely removing these cells from the proliferating cell population.

While we demonstrated the utility of minimization of gene transfer phagemid DNA (Chapter 5), these efforts were precipitated on CCC vectors of a single-stranded nature. Nevertheless, we were able to observe an increase in gene transfer correlated with the miniaturization of the phagemid vector. To investigate a dsDNA phagemid assembled by filamentous phage, we must first construct a phagemid vector containing both complementary sequences of a transgene cassette on the same DNA strand; in other words, we must first construct a plasmid with a large inverted repeat.

6.1.3 Cloning of inverted repeats in *Escherichia coli*

The success of molecular cloning in prokaryotic systems can be impaired by the activation of host pathways involved in DNA replication or repair. Such pathways include 1) restriction systems that cleave foreign DNA; 2) endo- and exonuclease degradation of DNA; 3) homologous or non-homologous recombination; and 4) DNA repair mechanisms.

Indeed, *E. coli* strains most commonly employed for the propagation of plasmid DNA are deficient in restriction systems. Notably, suppression of EcoKI restriction (*hsdRMS*) [544] and MrcA, MrcBC, and Mrr (*mrr* and *mcr*) [545] permit the introduction of foreign DNA. Mutation in *endA* improves the quality of DNA plasmid extractions [546, 547]. Mutations in homologous recombination and DNA repair pathways improve the stability of such vectors, and further allow propagation of other unstable DNA species.

In *E. coli*, the major mediator of homologous recombination is RecA (*recA*), which is involved in the RecBCD, RecE, and RecF pathways of recombination [548, 549]. RecA catalyzes recombination through the pairing of homologous strands of DNA and subsequent strand exchange.

RecBCD mediates the primary pathway for recombination in *E. coli*; it operates in concert with RecA, DNA polymerase I, single-stranded binding (SSB) proteins, and DNA gyrase, among other Rec proteins [549, 550]. RecBCD is essential for the repair of DSB and is also a major player in recombination processes involving free dsDNA ends: for example, conjugation, transduction, and transformation [549]. Upon binding to a blunt dsDNA end, RecBCD exhibits helicase activity to unwind the DNA. Through recognition of a χ motif (5′-GCTGGTGG-3′), its nuclease activity modifies the end to produce a 3′-ssDNA overhang [549, 551] upon which RecA can be loaded for the catalysis of homologous alignment and strand exchange [552].

The RecF pathway is implicated in the DNA repair of lesions arising from UV damage or stalled replication forks; it may also participate in recombination processes of dsDNA ends in the absence of RecBCD [553, 554]. In place of RecBCD, RecQ unwinds double-stranded DNA to enable RecJ exonuclease activity in order to produce the 3′-ssDNA overhang for RecA loading [555, 556]. RecE-mediated recombination is activated in *recB*, *recC*, *sbcA* mutants [557]. Acting on dsDNA as a substrate, RecE exonuclease activity produces 3′ overhangs [558, 559], which enables RecT-mediated homologous DNA strand exchange and annealing [558, 560].

As RecA is the primary recombinase across these pathways [548, 549], *recA* is the primary target for recombination silencing. Knockout of the RecBCD pathway by mutations in *recB* and/or *recCD* can drastically reduce homologous recombination events in *E. coli*. RecF-mediated recombination can be ablated by mutations in *recA*, although *recF* and *recJ* are also notable targets. Importantly, DNA gyrase has been implicated in plasmidic recombination [444], so mutations in *gyrA* are important to prevent undesired plasmid sequence modification.

Large unstable episomal DNA molecules have given rise to other peculiarities. Repetitive sequences in DNA are not well tolerated in either eukaryotes or prokaryotes and can form deleterious secondary structures [561, 562]. Leach (1996) defined a DNA palindrome as a region of double-stranded DNA with twofold rotational symmetry; self-complementarity in each strand permits intra-strand base-pairing giving rise to hairpin or cruciform structures [562]. The formation of hairpins from inverted repeats stalls replication in bacterial, yeast, and mammalian cells alike [561, 563], which can lead to genomic rearrangements and inviability. Propagation of inverted repeats larger than 200–300 bp in rec^+ hosts is limited and subject to deletion or inviability [562]. Plasmids with such regions are prone to deletion or rearrangements [444, 455], as their palindromic nature activates homologous recombination pathways [548]. The use of a central asymmetric region that lacks self-complementarity to interrupt a DNA palindrome was found to improve stability in *E. coli* [564]. Warren et al. (1985) observed that the addition of a central asymmetric spacer with a minimum length of 50 bp restored viability of otherwise nonviable DNA palindromes. A spacer greater than 150 bp was necessary for stable propagation.

For the stable propagation and manipulation of large stretches of repetitive DNA, *E. coli* strains have been developed with combinations of mutations in the following pathways: homologous recombination (*recA*, *recB*, *recJ*), UV repair (*umuC*, *uvrC*), and restriction (*hsd*, *mcr*, *mrr*) [454, 565–568]. Mutations in genes involved in the RecBCD homologous recombination pathway (*recA*, *BC*, *recD*, *sbcB*) improved the stability of repetitive DNA [566]. In particular, *sbcC* and *sbcB* were implicated in the stability of DNA palindromes in the range of 500–1600 bp [565, 569]. It has been shown that SbcCD causes lethal DSBs during replication of DNA palindromes [570]. A quadruple combination of mutations in *recB*, *recC*, *sbcC*, *recJ*, and/or *recN* prevented deletion of unstable mammalian DNA sequences [567]. The SURE *E. coli* strains (SURE and SURE2) are deficient in recombination (*recB*, *recJ*, *sbcC*), UV repair (*uvrC*, *umuC*), and restriction (*mrcAB*, *mrr*, *hsdRMS*). Similarly, the Stbl *E. coli* strains (Stbl2 and Stbl4) are deficient in recombination (*recA*) and restriction. A known disadvantage of using recombination-deficient *E. coli* strains is their slower growth activity; however, we and others [445] have observed that some *rec⁻* strains can perform comparably to *rec⁺* strains.

6.1.4 Rationale and hypothesis

Here, we construct and characterize a sense-antisense (SAS) precursor vector that contains both the sense and antisense sequences of a transgene cassette approximately 2.0 kb in length, flanked by the separated domains of the f1 functional origin (f1 *ori*). We hypothesized that the insertion of a large inverted repeat could be stabilized through the use of a minimized spacer region (50 bp) in the *E. coli* strain Stbl4. Propagation of large unstable repetitive DNA such as lentiviral and retroviral sequences has been demonstrated in Stbl *E. coli* strains [454, 471]. Hence, it appears to be a good strain for the cloning of a large inverted repeat. We hypothesized that during amplification, the complementary regions of the released plus (+) strand can re-anneal to form a dsDNA LCC minivector phagemid that can be assembled into progeny virions.

6.2 Materials and methods

Strains and vectors. *E. coli* K-12 strains were used in the generation of all phage and plasmid constructs. *E. coli* JM109 was the host for plasmid amplification and purification unless otherwise noted. Phage propagation was carried out using materials and methods described Chapter 3. All bacterial strains and mammalian cell lines are listed in Table 6.1, plasmids in Table 6.2, and phages in Table 6.3.

Maintenance of mammalian cell lines. Mammalian cell lines were maintained in tissue culture plates (Thermo Scientific, Waltham, USA) at 37 °C in a humidified atmosphere with 10% CO_2. All cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) (Thermo Scientific) supplemented with 10% heat-inactivated fetal bovine serum (FBS) and 1% penicillin/streptomycin.

Table 6.1: Bacterial and mammalian cell lines used to assess double-stranded miniphagemid function

Strain	Genotype	Source
JM109	F$'$ traD36 proAB$^+$ lacI^q lacZΔM15/ Δ(lac-proAB) endA1 glnV44 thi-1 e14$^-$ recA1 gyrA96 relA1 hsdR17	NEB
Stbl4	F$'$ proAB$^+$ lacIq lacZΔM15 Tn10/ endA1 glnV44 thi-1 recA1 gyrA96 relA1 Δ(lac-proAB) mcrA Δ(mcrBC-hsdRMS-mrr) λ^- gal	Invitrogen

Cell line	Description	Source
HeLa	uterus, cervix adenocarcinoma	gift, Serenity Bioworks

NEB: New England BioLabs

Table 6.2: Plasmids used in the construction of a double-stranded miniphagemid

Plasmid	Genotype	Source
pGL2-SS-CMV-GFP-SS	pGL2-Promoter (GN: X65326.2) *cmv-gfp* replaced *SV40-luc*, Ap^R	gift, Mediphage Bioceuticals Inc.
pM13ori2.cmvgfp	pM13ori2, *cmv-gfp* inserted in EcoRI-PacI	Chapter 3
pJET1.2/blunt	GN: EF694056.1	Thermo Fisher Scientific
pJET1.2-cmvgfp	pJET1.2, *cmv-gfp* inserted into EcoRV site, Ap^R	This study
pM13ori2.cmvgfp.SAS (L)	pM13ori2.cmvgfp, second *cmv-gfp* inserted in HindIII-BamHI	This study
pM13ori2.cmvgfp.SAS (NL)	pM13ori2.cmvgfp, second *cmv-gfp* inserted in HindIII-BamHI	This study
pM13ori2.cmvgfp.SAS (GG)	pM13ori2.cmvgfp, second *cmv-gfp* inserted via Golden Gate assembly	This study

GN: GenBank Accession No.

Table 6.3: Phages used in the construction of a double-stranded miniphagemid

Phage	Genotype	Source
M13KO7	M13, p15a *ori*, Tn903 (Km^R)	NEB
mini-(gfp)	*cmv-gfp*, from precursor pM13ori2.cmvgfp	Chapter 3
ds-mini-(gfp)	*cmv-gfp* – *cmv-gfp*, from precursor pM13ori2.cmvgfp.SAS (L)	Chapter 3

NEB: New England BioLabs

Construction of a precursor vector with both sense and antisense transgene cassettes using Golden Gate Assembly.   Two approaches were taken to construct a SAS precursor vector encoding inverted repeats of the mammalian transgene cassette, *cmv-gfp* (Figure 6.1). The first was through BsaI-dependent Golden Gate Assembly and incorporated a small central region absent of self-complementarity (50 bp). The "sense" plasmid pM13ori2.cmvgfp was used both as the template and target destination vector. A BsaI recognition site within the Ap^R marker of the pM13ori2 backbone was previously mutagenized in Chapter 3. Both target destination vector and the "antisense" *cmv-gfp* cassette were amplified from pM13ori2.cmvgfp (Table 6.4) with BsaI recognition sites in the primer overhangs. Digestion by BsaI and ligation by T4 ligase proceeded within the same reaction using the Golden Gate Assembly kit (New England BioLabs, Ipswich, USA) to generate the final construct pM13ori2.cmvgfp.SAS (GG), which is depicted in Figure 6.2B. The ligation was transformed into Stbl4 and selected on Luria–Bertani (LB) agar supplemented with ampicillin.

Table 6.4: Primers for Golden Gate assembly of a sense-antisense precursor

Primer	Sequence (5′ – 3′)[1]
cmvgfp-AS-F	GTCGAA*GGTCTC*CAACCTCTTAATTAAGGAATGTGTGTC
cmvgfp-AS-R	GTCGAA*GGTCTC*CACGAATTCCGTTACATAACTTACGG
M13ori-F	GGCTAC*GGTCTC*TTCGTCGTACGCGCCCTGTAGCG
M13ori-R	GGCTAC*GGTCTC*TGGTTGGAATGTGTGTCAGTTAGGGTGTG

[1] Underlined nucleotides indicate primer overhang. Italicized nucleotides indicate BsaI recognition site.

Construction of a precursor vector with both sense and antisense transgene cassettes using traditional endonuclease digestion.   A second approach used traditional restriction and ligation, and incorporated a larger central region absent of self-complementarity by way of the transit cloning vector pJET1.2 from the CloneJET PCR Cloning Kit (Thermo Fisher Scientific). The *cmv-gfp* cassette, including the PacI site, from pM13ori2.cmvgfp was first inserted into pJET1.2. The cassette was then excised from pJET1.2 using endonucleases BamHI and HindIII, which included a 264 bp region of pJET1.2, referred to as the linker. Digestion of the destination vector pM13ori2.cmvgfp (containing the "sense" cassette) was performed using the same enzymes, both of which are located near the M13 STOP. The "antisense" fragment was then ligated using T4 ligase into the digested pM13ori2.cmvgfp to construct the SAS precursor with linker: pM13ori2.cmvgfp.SAS (L). The ligation was transformed into Stbl4 and selected on LB agar supplemented with ampicillin. Additionally, the linker was later excised by digestion with PacI to construct a precursor variant without a linker: pM13ori2.cmvgfp.SAS (NL). Both constructs are depicted in Figure 6.2.
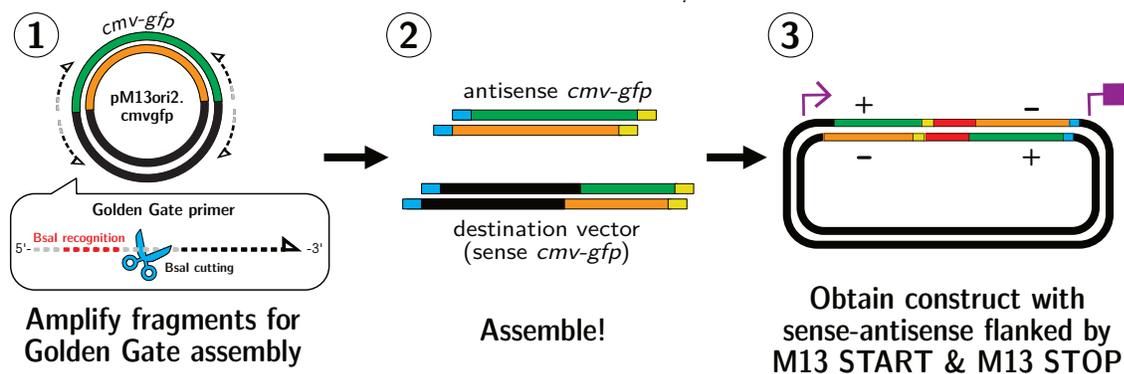
(A) Golden Gate strategy for construction of the SAS precursor vector
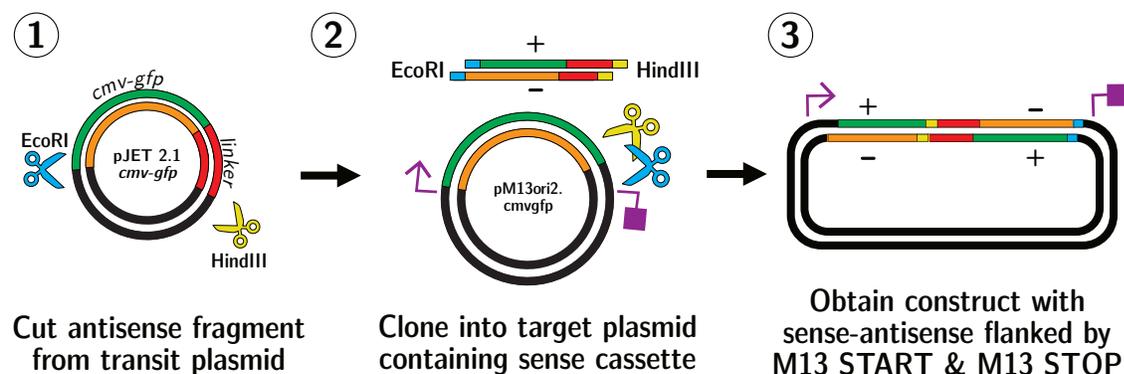
(B) Traditional cloning strategy for construction of the SAS precursor vector

Figure 6.1: Two approaches to the construction of the sense-antisense precursor. A) To assemble the SAS precursor via Golden Gate, primers with BsaI cut sites within the primer overhang were used to amplify the antisense fragment and the destination vector (containing the sense sequence). Assembly proceeded through thermocycling to activate BsaI and T4 ligase activity in order to construct the final plasmid. B) To assemble the SAS precursor via the traditional restriction and ligation approach, the *cmv-gfp* cassette was first amplified by PCR and inserted into a transit vector (pJET 2.1) to incorporate a HindIII cut site downstream of the cassette. The destination vector (containing the sense sequence) was then digested with EcoRI and HindIII as was the pJET1.2-cmv-gfp transit vector. Ligation with T4 ligase would produce the final construct with a linker region from pJET1.2.

(A) sense vector

(B) Golden Gate Assembly, short linker

(C) BamHI-HindIII digestion, large linker

(D) BamHI-HindIII digestion, no linker
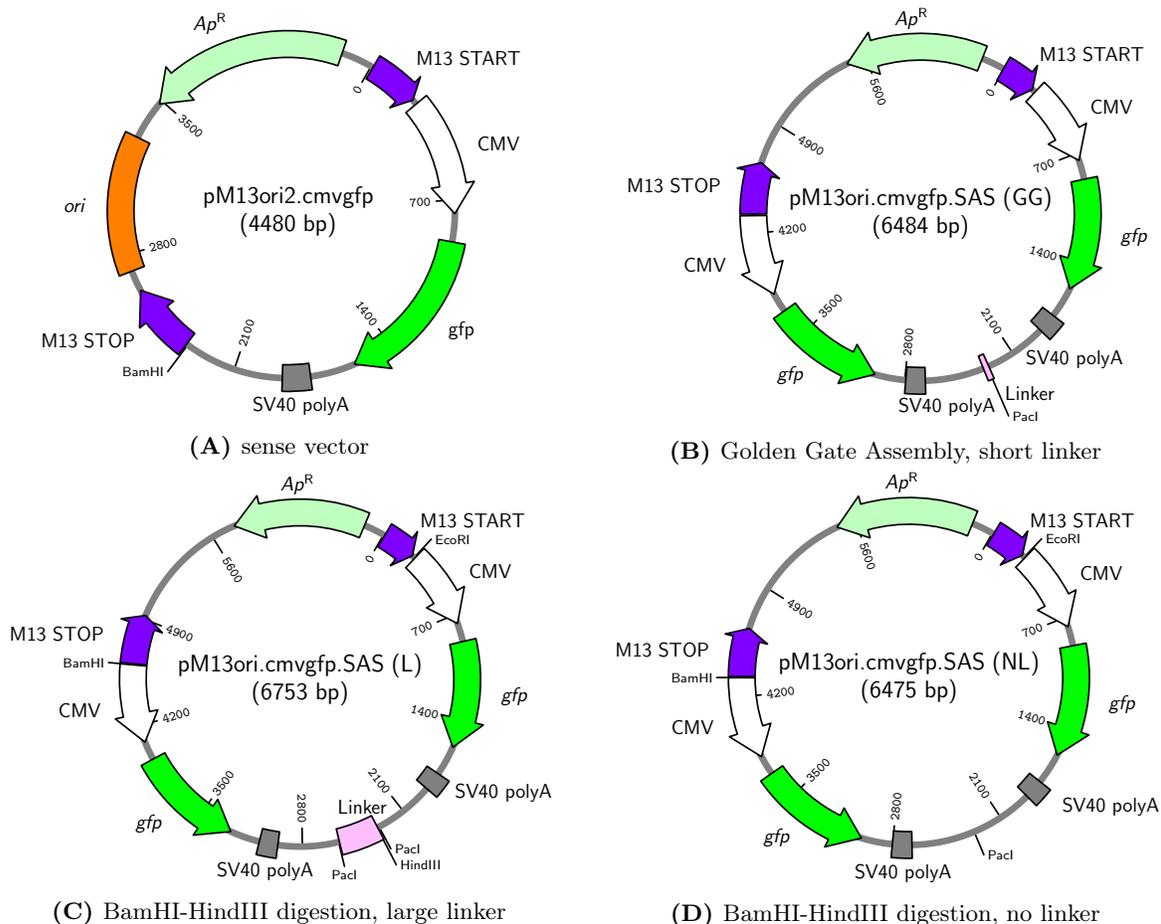
Figure 6.2: Maps of sense-antisense miniphagemid precursors. Predicted maps of precursor plasmids for the production dsDNA miniphagemids. An inverted repeat of the *cmv-gfp* cassette is inserted upstream of the M13 STOP. Both (B) and (C) show the same predicted construct assembled through different methods, and (D) is the same construct as (C) with the linker (pink) excised. GG: Golden Gate Assembly, L: pJET1.2 fragment (linker), NL: No linker.

Characterization of the putative sense-antisense precursor plasmid.   Putative clones were subjected to endonuclease digestion to verify construction. For the plasmid constructed through Golden Gate Assembly, these enzymes were: ApaI, EcoRI, MfeI, MluI, PacI (New England BioLabs). For the plasmid constructed through traditional restriction and ligation, these enzymes were: BamHI, EcoRI, HindIII (New England BioLabs). Digested products were visualized via agarose gel electrophoresis (AGE).

Production and purification of phage.   A fresh *E. coli* colony carrying the phagemid precursor plasmid was inoculated into rich media and incubated at $37\,°C$ with aeration until slightly turbid $(0.01 < A_{600} < 0.4)$. The culture was then infected with helper phage M13KO7 to a titre of $1 \times 10^8\,\text{PFU/mL}$ and returned to the same conditions for an additional 1–1.5 h. It was supplemented with kanamycin $(70\,\mu g/mL)$ and returned to the same conditions overnight. The next day, the culture was centrifuged $(8000 \times g)$ to separate the bacterial pellet (replicative factor (RF) DNA) from the supernatant (phage lysate).

After removal of the pellet, the phage lysate was purified through a $0.45\,\mu m$ filter to remove residual bacterial debris. Phage were precipitated with polyethylene glycol (PEG) and re-suspended in Tris-NaCl (TN) buffer. Filtered lysate was concentrated through precipitation with PEG: 1/5 of the lysate volume was added in PEG, and the mixture was incubated at $4\,°C$ for at least 2 h. It was then centrifuged at $12\,000 \times g$ for 15 min at $4\,°C$ to separate the PEGylated phage (pellet) from media (supernatant). The pellet was then re-suspended in a smaller volume of ice-cold TN buffer and a second PEG precipitation was carried out. PEG-precipitated phage suspensions were treated with DNase I (Promega, Madison, USA) to remove any extraneous phage or bacterial DNA in the sample [409, 410]. The concentrated phage lysate was stored at $4\,°C$.

RF dsDNA were extracted from the pellet with the Monarch Plasmid Miniprep Kit (New England Biolabs). Phagemid ssDNA were extracted from the phage lysate through phenol-chloroform extraction. Briefly, phenol was added to the PEGylated lysate (1:1, v/v) and mixed by vortexing. After centrifugation for 5 min at $4\,°C$, the top aqueous layer was extracted. This layer was extracted again with an equivalent volume of phenol:chloroform twice, and chloroform once. Finally, the ssDNA was precipitated overnight with 100% ethanol at $-80\,°C$. The precipitated DNA was washed with 70% ethanol, dried, and re-suspended in RNase- and DNase-free water.

Quantification of phagemid and RF DNA.   Extracted DNA was analyzed on a NanoDrop 2000 to determine concentration and purity. Extracted RFs were linearized with BamHI and visualized via AGE, while phagemid ssDNA was visualized directly via AGE without digestion.

Mung bean nuclease assay.   Mung bean nuclease preferentially and nonspecifically digests single-stranded DNA. We used it to confirm the form of DNA in phage lysates arising from the SAS precursor vector. Purified DNA was subjected to mung bean nuclease (1 or 10 units) for 0.5 or 2 h (New England BioLabs). Digested samples were then visualized via AGE.

Optimization of SAS phage production and RF processing.   Key parameters were investigated for their influence on the processing of the SAS precursor plasmid into a recombinant RF (RFx). Phage production was performed as described above, except for modifications to the parameters under investigation. The parameters were: optical density at time of infection with helper phage, the multiplicity of infection (MOI) of helper phage, presence of phagemid antibiotic (ampicillin) and/or helper antibiotic (kanamycin), and the duration of infection by helper phage. As a baseline, cultures were infected with helper phage M13KO7 at $A_{600} = 0.02$ to a final concentration of 1×10^8 PFU/mL for 16 h before phage purification or RF extraction. Plasmids and RFs were extracted and purified from cell pellets using the Monarch Plasmid Miniprep Kit (New England Biolabs). RFx DNA was analyzed via AGE.

Separation of DNA species.   The phage lysate arising from helper phage action on the SAS precursor vector contained both ssDNA helper phage genome and double-stranded LCC miniphagemid. Helper phage genome was separated from the dsDNA SAS minivector in the phage lysate using hydroxyapatite (HAP) chromatography (Bio-Rad), according to the method described by Fadrosh et al. [571, 572]. Briefly, the DNA mixture was incubated with 0.12 M phosphate buffer. The mixture was passed through hydrated HAP (Bio-Gel HTP hydroxyapatite) in an open glass Econo-column (Bio-Rad) at 60 °C. DNA was eluted from the column in fractions of increasing phosphate concentration: 1) 0.12 M (ssDNA), 2) 0.2 M, and 3) 0.4 M and 1.0 M (dsDNA). Phosphate buffer recipes can be found in Appendix A. Eluted fractions were extracted using phenol:chloroform (1:1, v/v), followed by chloroform (1:1, v/v), and desalted using Amicon Ultra Centrifugal Filter Devices (30,000 MWCO; Millipore, Billerica MA). Finally, nucleic acids were precipitated overnight with 100% ethanol at −80 °C. The precipitated DNA was washed with 70% ethanol, dried, and re-suspended in RNase and DNase-free water.

Transfection of phage and purified DNA into HeLa.   HeLa cells were seeded at 5×10^4 cells/mL in a 24-well plate. Purified double-stranded plasmid DNA was complexed with TurboFect (Thermo Fisher Scientific) for 15 min prior to transfection. Transfection was visualized by fluorescence microscopy on the EVOS FL Auto Imaging System (Thermo Fisher Scientific).

6.3   Results

A plasmid containing both the complementary sequences of a mammalian expression cassette (*cmv-gfp*) flanked by separated domains of the f1 *ori* was constructed. Two molecular cloning approaches were investigated, resulting in the successful construction of the vector in *E. coli* Stbl4. Helper phage M13KO7-mediated rescue of a double-stranded miniphagemid was then demonstrated and investigated. The resulting lysate was subjected to HAP chromatography to isolate the LCC double-stranded miniphagemid for downstream applications.

6.3.1 Construction of a sense-antisense miniphagemid precursor

The construction of a plasmid with a large interrupted palindrome was evaluated with two sizes of spacers. Through Golden Gate assembly, a linker approximately 50 bp long was incorporated, while through traditional restriction and ligation, a 264 bp linker was incorporated.

Although assembly of products was viable, clones were unstable and prone to deletion or recombination. As seen in Figures 6.3 and 6.4, a substantial deletion event was observed. This is evidenced by differences in the observed sizes of restriction digestion products compared to their expected sizes when visualized via AGE. Table 6.5 summarizes all the expected DNA fragment sizes resulting from digestion of the destination/source plasmid and the predicted sequence of the anticipated pM13ori2.cmvgfp.SAS (GG) construct. EcoRI digestion of pM13ori2.cmvgfp.SAS (GG) was expected to result in two fragments: one 2.4 kb containing the plasmid backbone and one 4.1 kb containing both sense and antisense cassettes. While the 2.4 kb fragment was identified, the other fragment was noticeably smaller than 4 kb, suggesting a deletion of at least 100 bp. Notably EcoRI-PacI digestion was expected to result in one fragment approximately 2.4 kb in size (identified in Figure 6.3B-2) and two fragments approximately 2 kb in size. Instead, one approximately 4 kb fragment was identified. Similarly, in the digestion with ApaI-PacI, one fragment in the range of 1–1.5 kb instead of 2×820 bp was identified. Based on the results in Figure 6.3, the linker PacI site did not appear to be present in the putative SAS clone.

Using endonucleases MluI, ApaI, and MfeI, we sought to narrow down where the deletion occurred. Two recognition sites for each of these nucleases should be present in the SAS precursor; MluI cuts upstream of *gfp*, ApaI cuts at the tail end of *gfp*, while MfeI cuts just upstream of the polyA sequence (Figure 6.4A). Digestion with MluI showed the expected fragment size spanning the plasmid backbone (3.6 kb), but the fragment expected to contain *gfp* and the linker was smaller than predicted (Figure 6.4B-2). Similarly, the 1.9 kb ApaI-ScaI fragment containing the "sense" *cmv-gfp* region was present after digestion of both the source vector and the putative SAS precursor (Figure 6.4B-1).

However, the predicted 1.6 kb band corresponding to the linker region between ApaI sites was reduced in size; the deletion event occurred downstream of the *gfp* sequence. Finally, digestion with MfeI again showed the fragment containing the backbone was as predicted, while the 1 kb fragment expected to contain the linker ran just above the 0.5 kb ladder. Thus, one or multiple deletions totalling approximately 200–500 bp occurred near (or likely at) the centre of the palindrome, also resulting in the deletion of the PacI site.
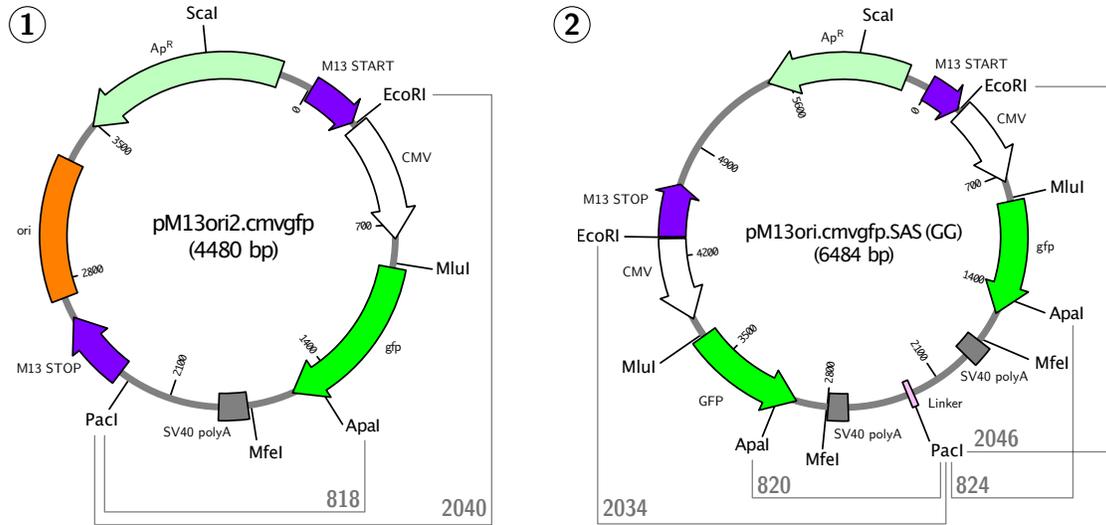
(A) Indicated regions correspond to predicted fragment sizes (bp)
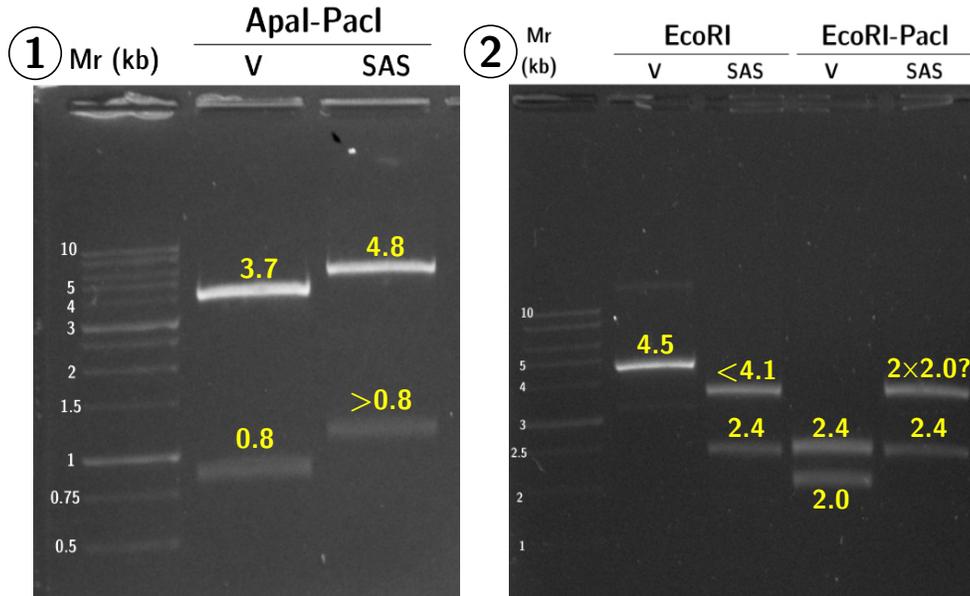
(B) Endonuclease digestion products

Figure 6.3: The PacI site may be deleted during Golden Gate assembly of the SAS precursor.
A) Schematics illustrating expected digestion products of interest from 1) the sense precursor, pM13ori2.cmvgfp (V), and 2) the SAS precursor (SAS). Fragment sizes are listed in bp. B) DNA fragments visualized via AGE as a result of digestion with: 1) ApaI and PacI, 2) EcoRI and PacI. Some SAS digestion products are smaller or larger than expected, as indicated in yellow. Mr: 1 kb ladder (FroggaBio).
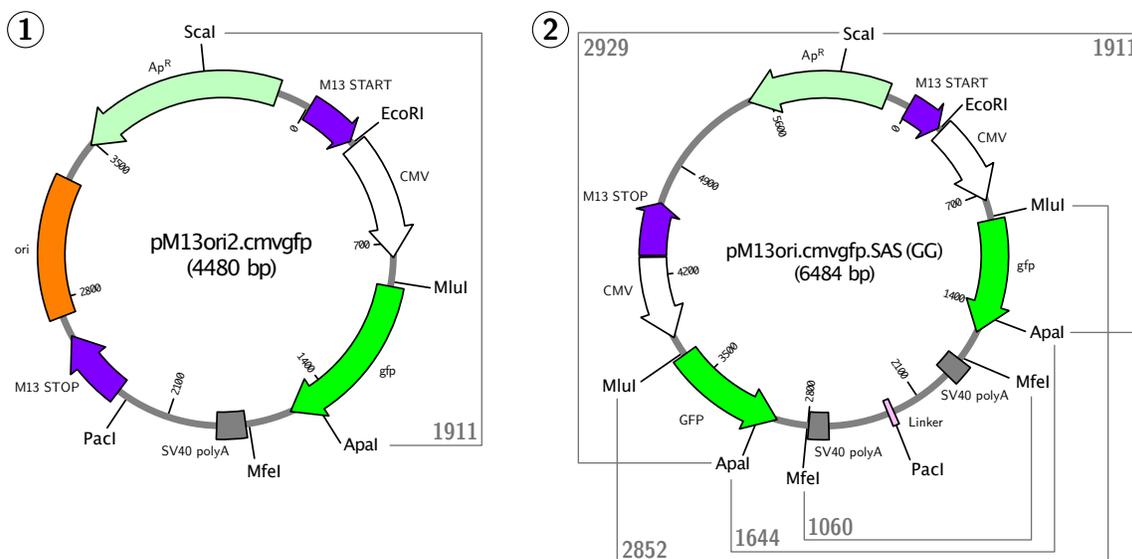
(A) Indicated regions correspond to predicted fragment sizes (bp)
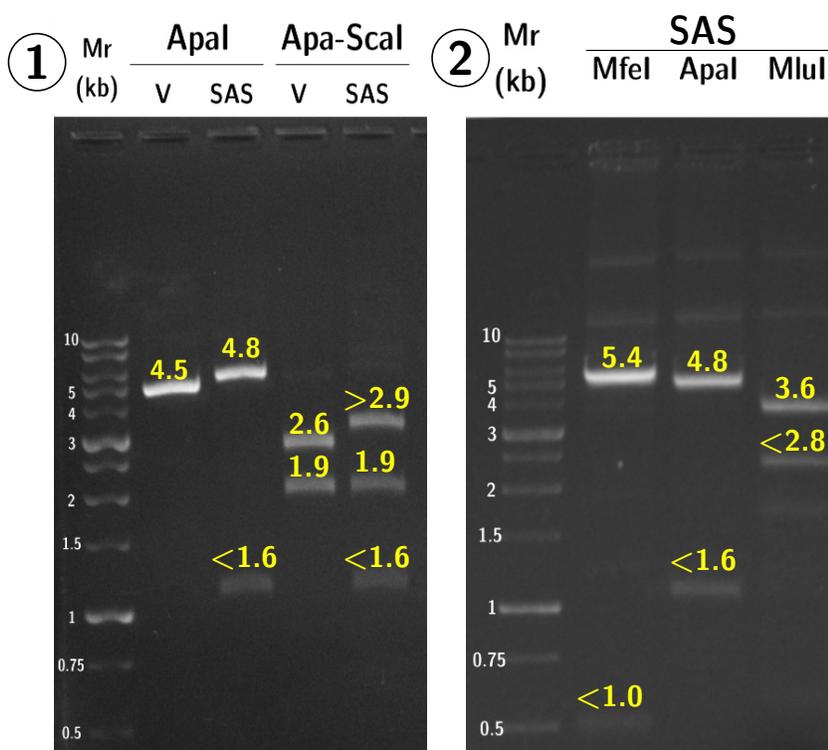
(B) Endonuclease digestion products

Figure 6.4: Digestion with endonucleases narrows down missing region in SAS precursor. A) Schematics illustrating expected digestion products of interest from 1) the sense precursor, pM13ori2.cmvgfp (V), and 2) the SAS precursor (SAS). Fragment sizes are listed in bp. B) DNA fragments visualized via AGE as a result of digestion with: 1) ApaI and ScaI, 2) MfeI, ApaI, or MluI. Some SAS digestion products are smaller or larger than expected, as indicated in yellow. Mr: 1 kb ladder (FroggaBio).

134

Table 6.5: Expected fragment sizes from digestion of sense-antisense precursor (GG)

Endonuclease(s)	Sense vector fragments (bp)[1]	SAS vector fragments (bp)[2]
PacI	4480	6484
EcoRI	4480	4080, 2404
ApaI-PacI	3662, 818	4840, 824, 820
EcoRI-PacI	2440, 2040	2404, 2046, 2034
ScaI	4480	6484
ApaI	4480	4840, 1644
ApaI-ScaI	2569, 1911	2929, 1911, 1644
MluI	4480	3632, 2852
MfeI	4480	5424, 1060

[1] Predicted from sequence for pM13ori2.cmvgfp.   [2] Predicted from sequence for pM13ori2.cmvgfp.SAS (GG).

The second approach to construct the SAS precursor vector was by way of a larger linker from pJET1.2. By amplifying the *cmv-gfp* cassette using a forward primer containing a BamHI site in the overhang and the HindIII site in pJET1.2, the "antisense" *cmv-gfp* could be inserted downstream of its "sense" complement in a reverse orientation using the unique HindIII-BamHI cut sites present upstream of the M13 STOP (Figure 6.1). Putative pM13ori2.cmvgfp.SAS (L) clones were analyzed by digestion with EcoRI, HindIII, and BamHI to assess if the antisense-linker fragment inserted stably (Figure 6.5). The bands corresponded to the predicted sizes as summarized in Table 6.6; in particular, the BamHI-HindIII fragment containing the antisense insert was successfully excised back out of the final clone. Thus, construction of the sense-antisense precursor plasmid, pM13ori2.cmvgfp.SAS (L) was confirmed.

Table 6.6: Expected fragment sizes from digestion of sense-antisense precursor (L)

Endonuclease(s)	SAS vector fragments (bp)[1]
EcoRI-HindIII	4703, 2050
HindIII-BamHI	4451, 2302
EcoRI-BamHI	4352, 2401

[1] Predicted from sequence for pM13ori2.cmvgfp.SAS (L).

Next, we assessed if removal of the linker was viable. We obtained clones of pM13ori.cmvgfp.SAS (NL) after PacI digestion, ligation, and selection in Stbl4; however, transformation of Stbl4 with purified (NL) plasmid was unstable. Figure 6.6 shows six random clones subjected to digestion by EcoRI, BamHI, HindIII, or PacI; all endonucleases were expected to cut once, except for HindIII whose recognition site was in the removed linker. EcoRI consistently cut once, but multiple bands were present in two clones when digested by BamHI. Instead, faint bands approximately 0.75–1 kb can also be seen after digestion with EcoRI or BamHI, indicating the prevalence of other DNA species. Furthermore, the PacI site appeared to have been lost in all clones.
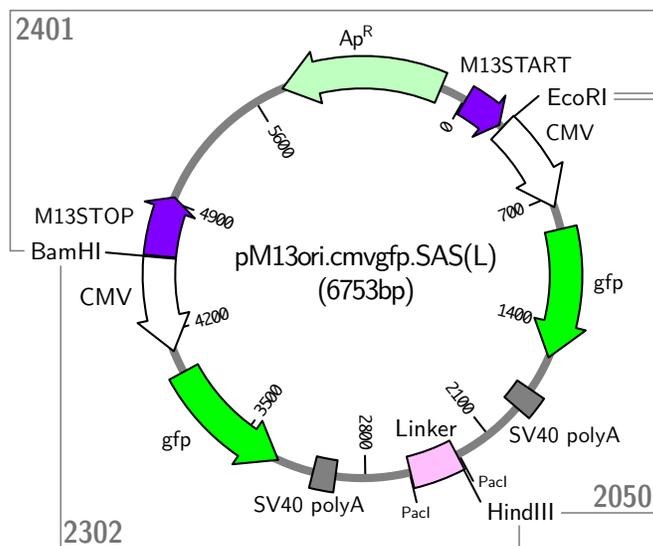
(A) Indicated regions correspond to predicted fragment sizes (bp)
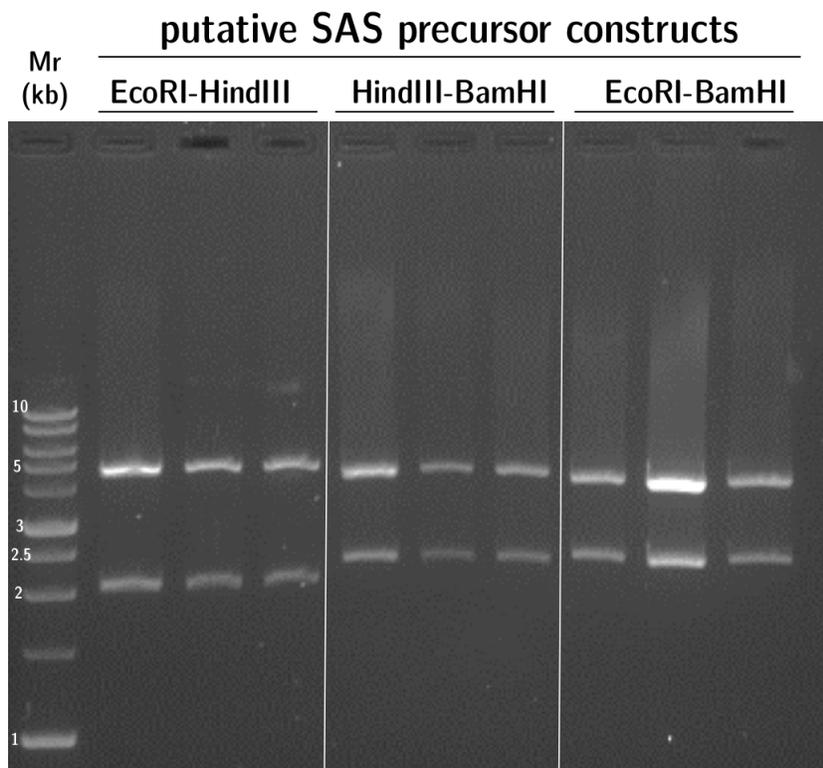
(B) Endonuclease digestion products

Figure 6.5: Construction of a SAS precursor plasmid. Putative clones were assessed by endonuclease digestion. A) Schematics illustrating expected digestion products of interest from the SAS precursor. Fragment sizes are listed in bp. B) DNA fragments visualized via AGE as a result of digestion with: EcoRI-HindIII, HindIII-BamHI, or EcoRI-BamHI. The products are the predicted size for pM13ori2.cmvgfp.SAS (L), confirming its construction. Mr: 1 kb ladder (FroggaBio).

Although the linker was removed from the (L) precursor plasmid in a separate step, this (NL) plasmid was unable to propagate stably even in Stbl4. Overall, a plasmid containing a large inverted repeat of a *cmv-gfp* cassette was constructed in Stbl4 using an intermediate spacer that linked the two repeats together. Construction of the same vector without a linker via Golden Gate Assembly was unsuccessful, likely due to the unstable conformation of two very close inverted repeats. Notably, attempts to repeat the construction in strains such as JM109 were unsuccessful; however, transformation with the constructed pM13ori2.cmvgfp.SAS (L) was viable and stable.
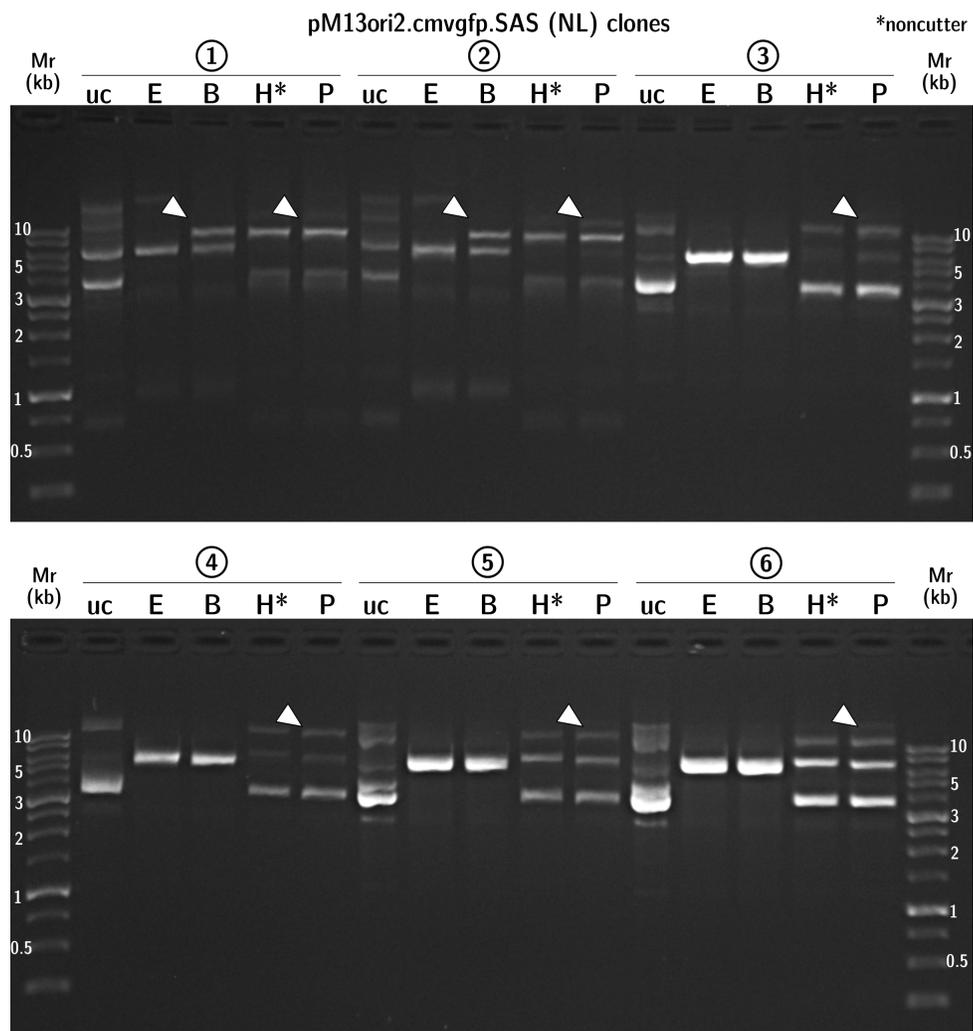
Figure 6.6: Removal of the central spacer destabilizes the sense-antisense plasmid. Stbl4 was transformed with purified pM13ori2.cmvgfp.SAS (NL); six random clones were miniprepped and subjected to digestion with enzymes cutting once (EcoRI, BamHI, PacI); the noncutter HindIII, was also used. From left to right for each clone, uc: uncut; E: EcoRI; B: BamHI; H: HindIII; P: PacI. A single band at 6.8 kb was expected in lanes E, B, and P; undigested plasmid was expected in lanes uc and H. White arrows indicate lanes where multiple bands were unexpectedly observed. Mr: 1 kb ladder (FroggaBio).

6.3.2   Production of a recombinant RF from the sense-antisense precursor

We next infected cells carrying the pM13ori2.cmvgfp.SAS (L) precursor plasmid with helper phage M13KO7 to evaluate if the sense-antisense cassette could be rescued by phage proteins. M13KO7-mediated rescue and assembly of the CCC single-stranded *cmv-gfp* miniphagemid was compared to that of the putative LCC dsDNA miniphagemid (Figure 6.7). The predicted ds-mini-(gfp) miniphagemid should be approximately 2.5 kb, which was indeed present in ssDNA purified from phage lysates. Visualization via AGE, however, showed that a majority of the ds-mini-(gfp) lysate was comprised of helper phage genome. In comparison, helper phage genome was not always detectable via AGE in the mini-(gfp) lysate.

The RFx was predicted to be approximately 4.5 kb (Figure 6.8). Both the RFx and ds-mini-(gfp) products were present in the extracted dsDNA RF and the ssDNA phage lysate, respectively (Figure 6.8B). From these results, M13KO7 did appear able to rescue a double-stranded mini-(gfp) via the formation of the predicted RFx (Figure 6.8A), but tended to preferentially assembled its own genome into phage progeny. This was expected as phage proteins exert optimal activity on ssDNA.
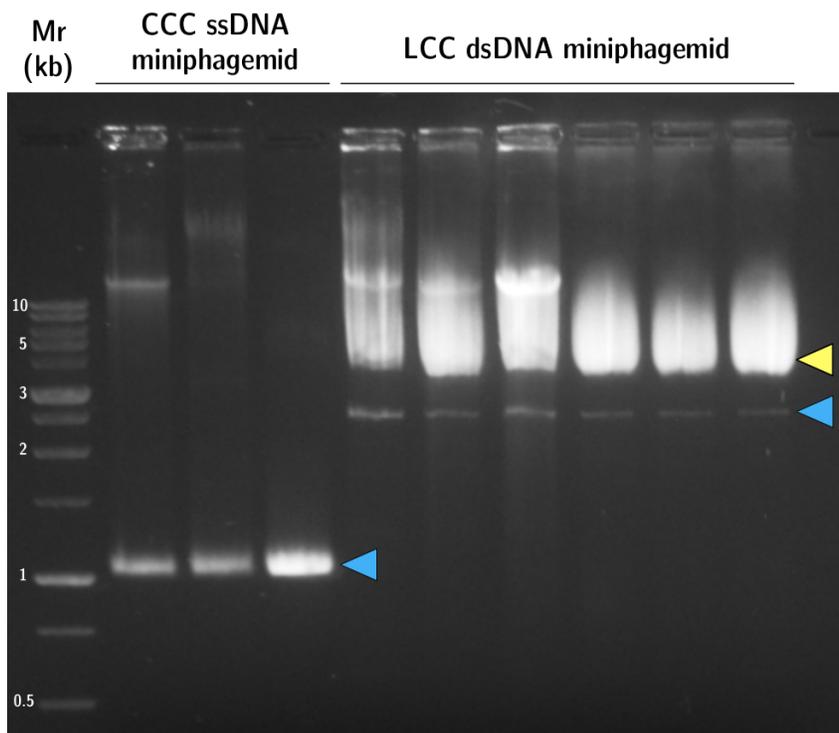
Figure 6.7: Helper phage M13KO7 can package a double-stranded miniphagemid. On the left, CCC purified single-stranded miniphagemid was separated from helper phage genome via AGE. On the right, putative LCC dsDNA miniphagemid was separated from helper phage genome. A faint band can be seen in each lysate that corresponds to the expected size of the double-stranded *cmv-gfp* miniphagemid. Bands are indicated as helper genome (yellow) and ds-mini-(gfp) (blue). Mr: 1 kb ladder (FroggaBio).
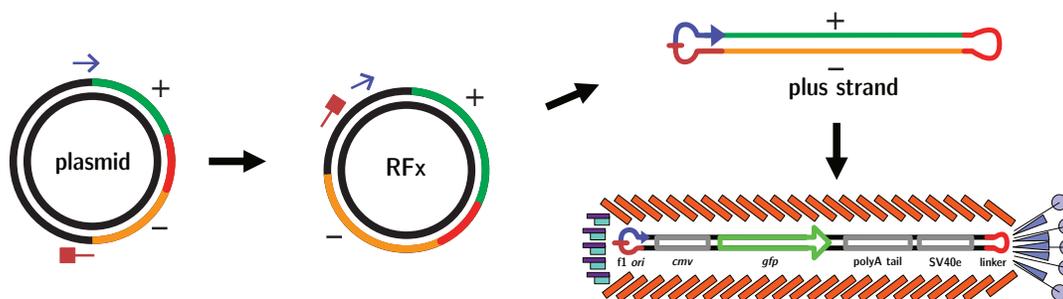
(A) DNA species from helper rescue of the double-stranded LCC miniphagemid
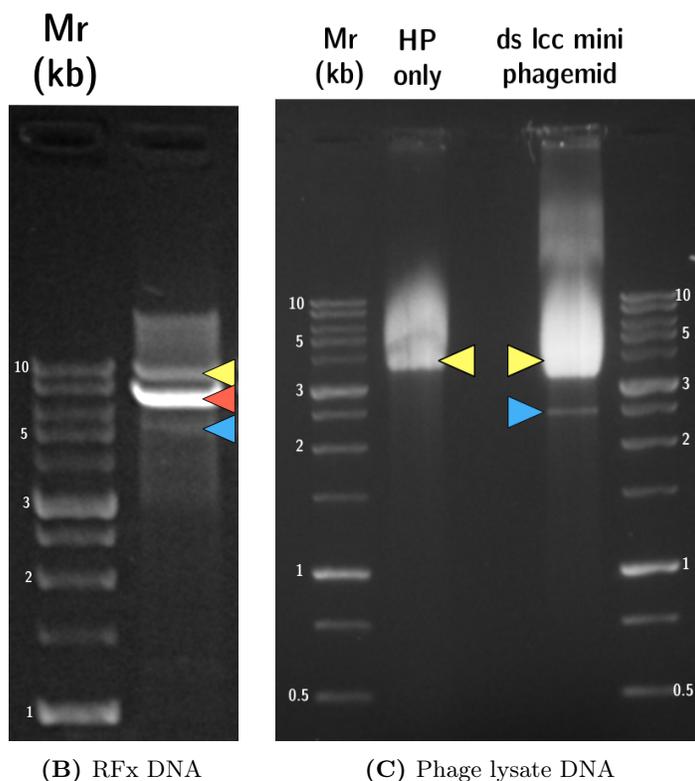
(B) RFx DNA

(C) Phage lysate DNA

Figure 6.8: The SAS RFx is necessary for double-stranded miniphagemid production. A) The schematic illustrates the species expected from helper phage rescue of a double-stranded LCC miniphagemid from the SAS precursor vector: 1) Helper phage rescue from the precursor plasmid forms 2) a double-stranded RFx with both the sense-antisense sequences. It undergoes rolling circle amplification to form 3) an ssDNA species where the sense-antisense regions anneal to form a dsDNA molecule for assembly into 4) a double-stranded LCC miniphagemid. Double-stranded RF products (B) and single-stranded purified phagemid DNA (C) were visualized via AGE. Bands are indicated as helper genome (yellow), precursor plasmid (red), and RFx or ds-mini-(gfp) (blue). Mr: 1 kb ladder (FroggaBio).

To verify if the 4.5 kb band in the phage lysate was indeed double-stranded in nature, it was treated with mung bean nuclease (MBN), which preferentially cuts ssDNA. We expected that helper phage genome and mini-(gfp) would be degraded after MBN treatment.

Like plasmid DNA, both ds-mini-(gfp) (L) and (NL) were resistant to MBN treatment (Figure 6.9). Even with 1 unit of MBN, all helper phage genome was degraded after 30 min, while the ds-mini-(gfp) remained almost completely intact. Importantly, prolonged treatment with MBN even degrades plasmid dsDNA, so it was expected that band intensity of the ds-mini-(gfp) would also decrease. With prolonged or increased MBN treatment, the band size corresponding to the dsDNA miniphagemid decreased. This is suspected to be due to the MBN-mediated removal of ssDNA regions (the f1 *ori* and the linker). Thus, we have demonstrated that the split origin of the sense-antisense precursor plasmid was a viable substrate for helper phage rescue of a double-stranded miniphagemid.
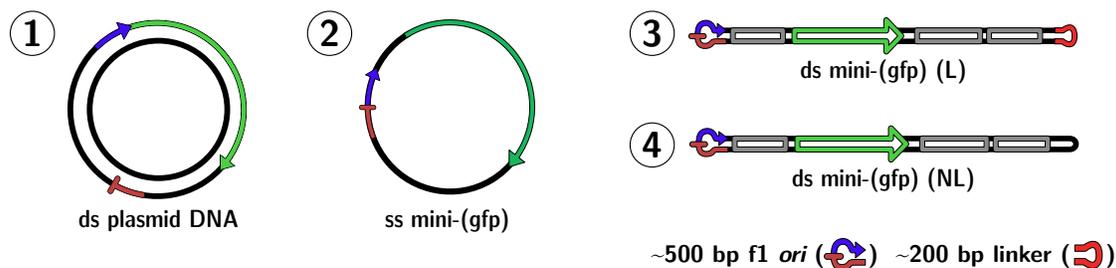
(A) Schematic drawings of the products treated with MBN
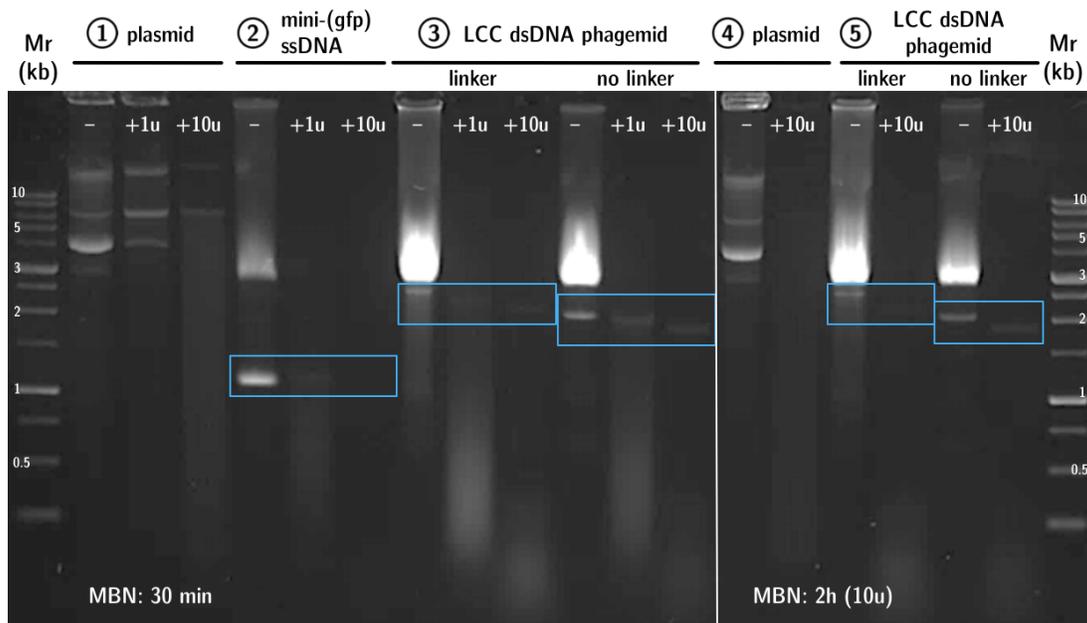
(B) AGE of phage lysates treated with MBN

Figure 6.9: The SAS phage product is resistant to mung bean nuclease. Phagemid and plasmid constructs (A) were subjected to digestion with MBN for 30 min or 2 h and visualized via AGE (B). Lysates were treated for 30 min with 1 or 10 units of MBN: 1) CCC plasmid dsDNA, 2) CCC mini-(gfp) ssDNA phagemid, 3) putative LCC ds-mini-(gfp) phagemid (with or without a linker). The blue boxes enclose the bands expected to be phagemid DNA. Lysates were treated for 2 h with 10 units of MBN: 4) plasmid DNA, 5) ds-mini-(gfp) phagemid (with or without a linker). Mr: 1 kb ladder (FroggaBio).

6.3.3   Optimization of the production of a recombinant sense-antisense RF

Although helper phage rescue of a sense-antisense RFx was demonstrated, filamentous production of a dsDNA progeny particle was not efficient. Thus, we examined if any parameters in the production method for ds-mini-(gfp) could be modified to improve it (Figure 6.10). For this study, we focused on the conversion of precursor sense-antisense plasmid to the recombinant RF, which is one of the rate-limiting steps to phagemid rescue. Optimizations were carried out in JM109 as RFx production was not detectable in Stbl4. Four parameters for phage production were evaluated for their influence on production of the dsDNA miniphagemid: length of time from infection to phage harvest, antibiotic selection for the precursor plasmid (ampicillin) and the helper phage (kanamycin), optical density of the culture at time of infection, and MOI.
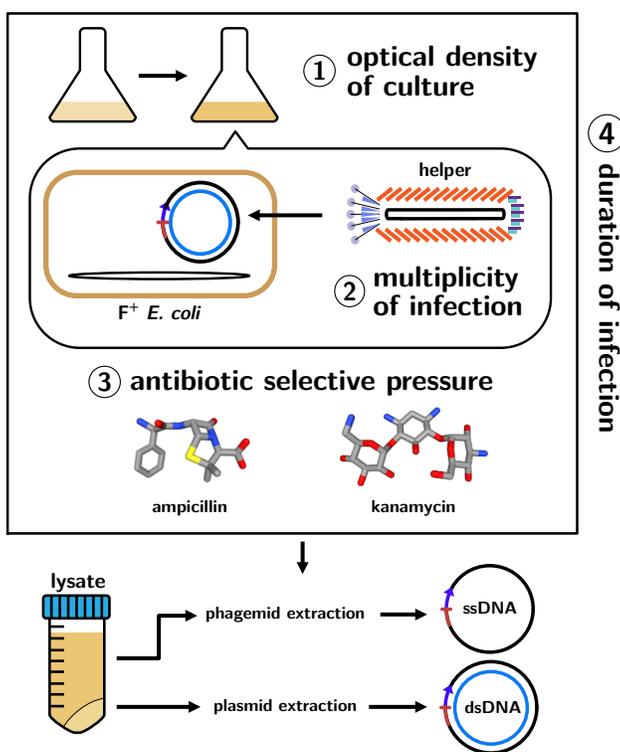
Figure 6.10: Key parameters optimized in the production of a sense-antisense miniphagemid. Parameters are indicated in this schematic of the overall workflow: 1) optical density of the culture at time of infection, 2) multiplicity of infection (MOI), 3) the degree of antibiotic selective pressure for the phagemid precursor or the helper phage, 4) duration of infection. Molecular structures of kanamycin [573] and ampicillin [574] are reproduced from PubChem [575].

The optimal duration for infection appears to be 16–18 h, as we were unable to observe appreciable amounts of RFx or ssDNA via AGE both when infection was stopped before 16 h and when it proceeded beyond 18 h. RFx was detected in most lysates if infection was allowed to proceed for 16–18 h, but was never observed outside this range. This contrasts with the production of "typical" CCC wild-type phage, which can be obtained after as little as 6 h [576].

We next examined if we could optimize M13KO7 rescue of the sense-antisense cassette from the precursor in order to generate the RFx. First, we examined if a higher culture optical density, an indicator of the bacterial growth stage, had an effect on the RFx production. Purified RF DNA was harvested from cultures carrying the SAS precursor plasmid infected by helper phage at optical densities from $A_{600} = 0.022$ to 0.094 (Figure 6.11), and $A_{600} = 0.562$ (Figure 6.12). Both undigested and BamHI-linearized DNA were visualized via AGE. We also examined RFx production with or without antibiotic selective pressure (kanamycin) for the helper phage in Figure 6.11. Notably, multiple extraneous bands of different sizes were detected; for example, a band slightly larger than 1 kb can be seen both in the uncut and cut lanes. These bands did not correspond to any expected DNA species and, notably, were not produced in the absence of helper phage. This may be indicative of recombination events stemming from replication from f1 *ori*, not the plasmid origin of replication (*ori*).
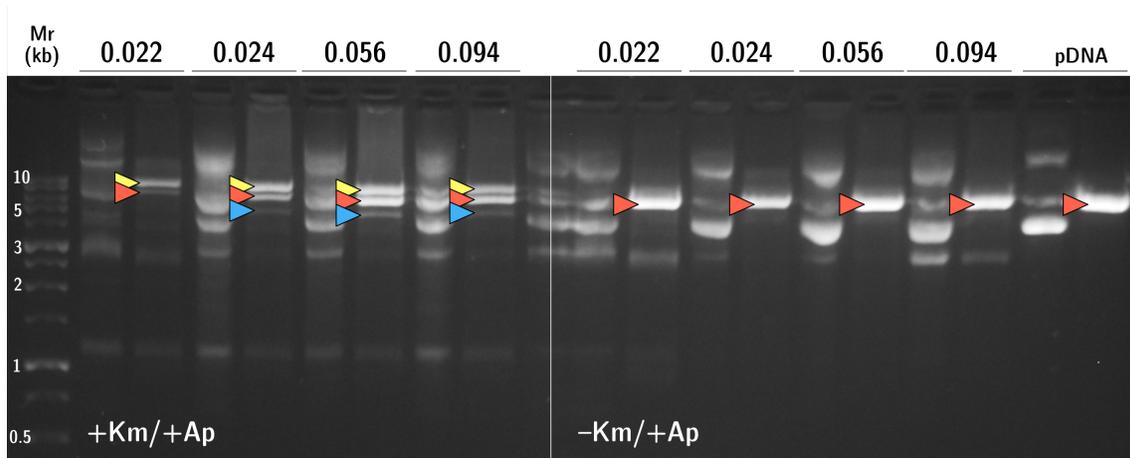
Figure 6.11: Effect of optical density and helper antibiotic on SAS RFx conversion. Each sample was infected by M13KO7 at the optical density indicated above the wells. Purified RF DNA was run in pairs (undigested, BamHI-digested). Lysates on the left side were prepared from cultures supplemented both kanamycin and ampicillin. Lysates on the right side were not treated with kanamycin after addition of M13KO7. Arrows indicate bands of interest: yellow (helper genome), red (plasmid), blue (RFx). Mr: 1 kb ladder (FroggaBio).

The brightest RFx band was associated with infection of culture at $A_{600} = 0.056$, when supplemented with kanamycin; neither RFx production nor the presence of the helper phage were observed in absence of kanamycin at these optical densities. This suggests that antibiotic selection is required to maintain helper phage activity. However, Figure 6.12 showed strong bands correlating to the M13KO7 genome even in absence of kanamycin when cultures were infected at $A_{600} = 0.562$. Here, we observed that both ampicillin and kanamycin appeared dispensable when cells were in late log phase upon infection. Unidentified DNA species were also present in both undigested and digested samples, including the >1 kb band.
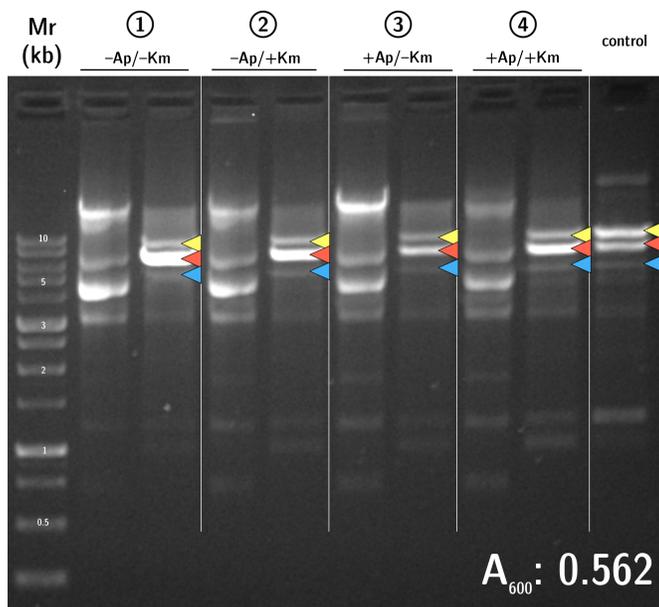
Figure 6.12: Effect of phagemid and helper antibiotic on SAS RFx conversion. All samples were infected when $A_{600} = 0.562$. Samples were loaded in pairs (undigested, BamHI-digested). From left to right, media after infection contained 1) no antibiotic; 2) kanamycin only; 3) ampicillin only; 4) both ampicillin and kanamycin. Arrows indicate bands of interest: yellow (helper genome), red (plasmid), blue (RFx). Mr: 1 kb ladder (FroggaBio).
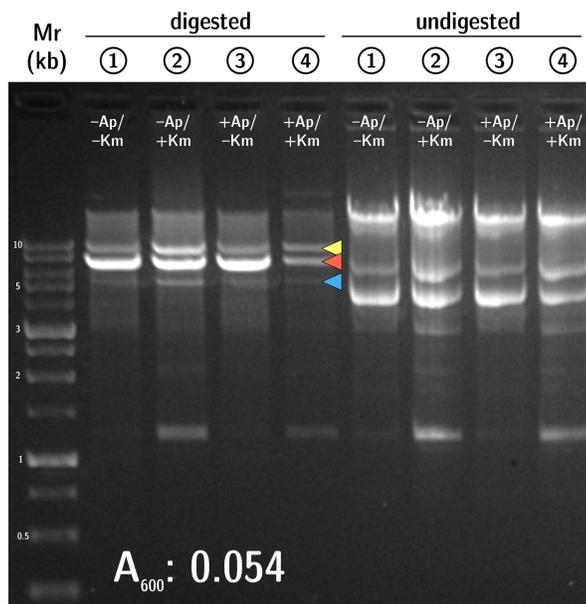
Figure 6.13: Effect of antibiotic on SAS RFx conversion. All samples were infected when $A_{600} = 0.054$. Samples on the left were digested with BamHI while the same undigested samples are on the left. From left to right, media after infection contained 1) no antibiotic; 2) kanamycin only; 3) ampicillin only; 4) both ampicillin and kanamycin. Arrows indicate bands of interest: yellow (helper genome), red (plasmid), blue (RFx). Mr: 1 kb ladder (FroggaBio).

143

Initial results (Figure 6.11) indicated that an optical density of approximately $A_{600} = 0.05$ appeared to be better for RFx production. The impact of selective pressure against the phagemid precursor vector was therefore evaluated at a low optical density ($A_{600} = 0.054$) in contrast to a high optical density($A_{600} = 0.562$) in Figures 6.12 and 6.13. Although helper phage genome was present even in the absence of kanamycin, bands correlating to linearized RFx were distinctly fainter in samples not supplemented with kanamycin; thus, maintenance of helper phage with antibiotic selection appeared to be required for good processing of the sense-antisense precursor. On the other hand, removal of ampicillin from the media did not appear to negatively impact RFx production; in fact, the brightest RFx band (Figure 6.13) was from a culture supplemented with kanamycin but devoid of ampicillin.

Next, we evaluated higher optical densities in Figure 6.14. Cultures were infected with M13KO7 at $A_{600} = 0.505$ or 0.605 to a final titre of 1×10^7 or 1×10^8 PFU/mL. Production of the RFx appeared robust even when $A_{600} = 0.605$, when cells may start entering stationary phase. The increase in helper phage concentration also appeared to be correlated with improved RFx production, particularly in the presence of kanamycin. However, we could not determine an optimal optical density; RFx production was similar both around $A_{600} = 0.05$ and $A_{600} = 0.6$.
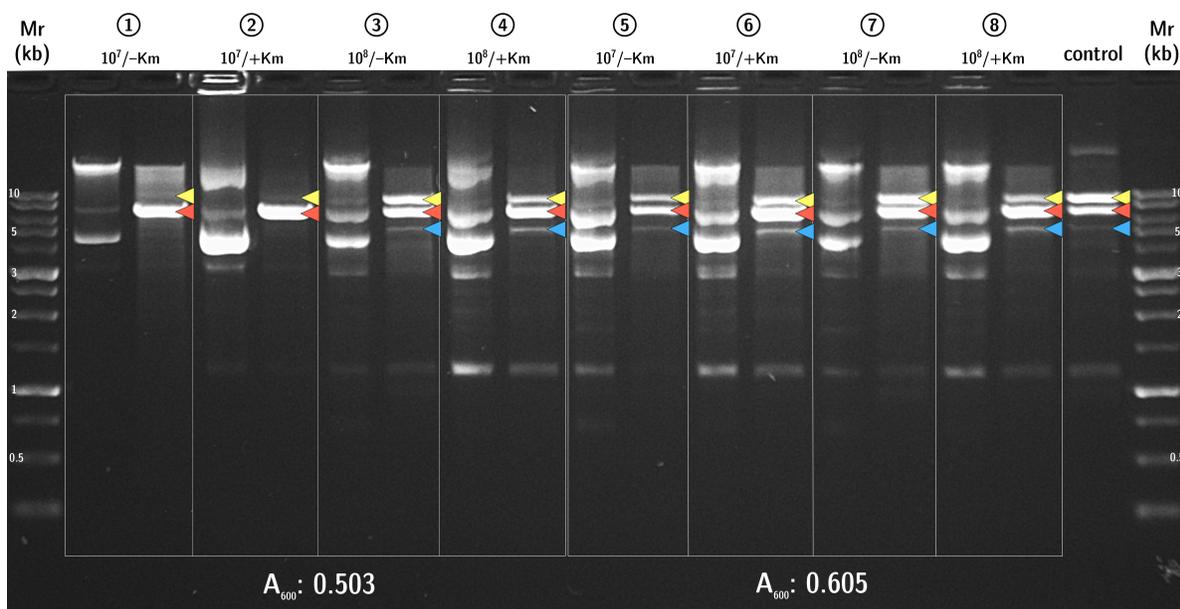
Figure 6.14: Effect of antibiotic, optical density, and MOI on SAS RFx conversion. On the left, samples were infected when $A_{600} = 0.503$; on the right, when $A_{600} = 0.605$. Samples were loaded in pairs (undigested, BamHI-digested). When the culture reached $A_{600} = 0.503$, helper phage M13KO7 was added at a concentration of 1×10^7 PFU/mL without kanamycin (1) and with (2); and at a concentration of 1×10^8 PFU/mL without kanamycin (3) and with (4). When the culture reached $A_{600} = 0.605$, helper phage M13KO7 was added at a concentration of 1×10^7 PFU/mL without kanamycin (5) and with (6); and at a concentration of 1×10^8 PFU/mL without kanamycin (7) and with (8). Arrows indicate bands of interest: yellow (helper genome), red (plasmid), blue (RFx). Mr: 1 kb ladder (FroggaBio).

Hence, we asked if it was the MOI that was defining factor. The MOI is defined as the number of phage that infect per cell. Given the same phage concentration, a lower optical density (where cell mass is low) results in a higher MOI; likewise, a higher optical density results in a lower MOI. However, no pattern emerged from examining the MOI across Figures 6.11 and 6.14. Specifically, we observed superior RFx production at A = 0.05, 0.5, and 0.6 when helper phage added at 1×10^8 PFU/mL in the presence of kanamycin. This corresponds to MOIs of 430, 60, and 111, respectively; it does not appear that MOI is a major impact on RFx production, although the absolute number of helper phage molecules may be. Aside from the duration of infection, we observed that antibiotic selective pressure against the helper phage improved precursor-RFx conversion, while phagemid antibiotic selective pressure after infection appeared to be non-essential.

6.3.4   Hydroxyapatite chromatography separation of DNA species

Given the high degree of helper phage in the lysate, the dsDNA phagemid needed to be purified prior to downstream applications. HAP chromatography has emerged as a robust method to separate different forms of DNA [572, 577, 578]. A mock chromatography run was first performed using a known mixture ssDNA and dsDNA to determine the efficacy of recovery (Figure 6.15).

For this mock run, 5 μg each of purified ssDNA M13KO7 genome and dsDNA pM13ori2 were combined. Fractions were eluted with 0.12 M, 0.2 M, 0.4 M, and 1.0 M phosphate buffer. The 0.4 M and 1.0 M fractions were combined in accordance with Fadrosh et al. (2011) [571]. DNA purified from each fraction was visualized via AGE; DNA purity and concentration were assessed by NanoDrop. The A260/280 ratios were in the range 1.87–1.92, which is within boundaries for high-purity DNA [579]. On the other hand, A260/230 ratios were in the range 1.70–2.26, which suggests possible contaminants that absorb strongly at 230 nm. In the 0.12 M fraction, approximately 68% (3.4 μg) of ssDNA was recovered. No DNA was detectable in the 0.2 M fraction. In the combined 0.4 M/1.0 M fraction, only 10% (approximately 0.5 μg) of the dsDNA was recovered. Overall, despite the low recovery, separation of dsDNA species from ssDNA species via HAP chromatography was confirmed in this study.

As the mock run was able to separate dsDNA from ssDNA, we followed by investigating the utility of HAP chromatography in separating phage genomic ssDNA from ds-mini-(gfp) DNA in consolidated phage lysates. SAS miniphagemid lysates were combined into two volumes for separation via HAP chromatography. Fractions were eluted across four fractions of increasing phosphate concentrations as detailed above [571], and visualized via AGE (Figure 6.16). The A260/280 ratios ranged from 1.80 to 1.97, again an acceptable ratio for pure DNA [579]. The A260/230 ratios ranged from 1.56 to 2.49; the low end is indicative of the presence of contaminants that absorb strongly at 230 nm.
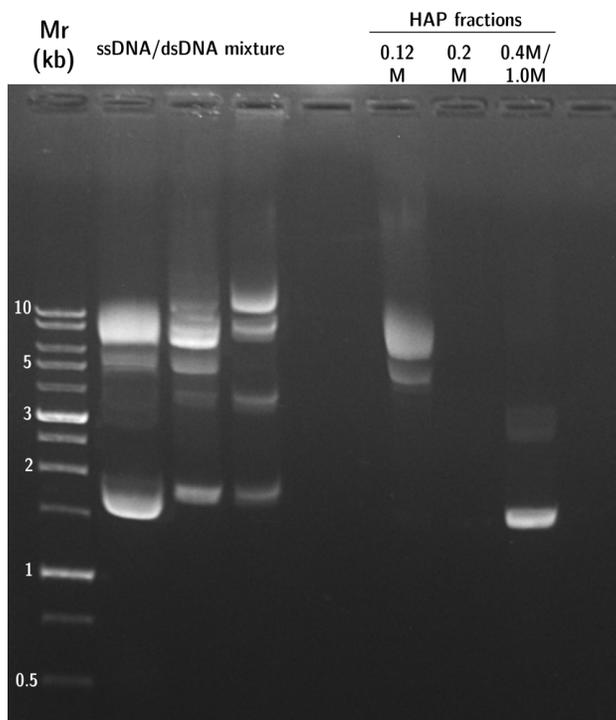
Figure 6.15: Hydroxyapatite chromatography separates DNA forms. On the left, undigested mixtures of ssDNA (M13KO7 genome) and dsDNA (pM13ori2 plasmid) are run together, unseparated. After performing HAP chromatography, the 0.12 M fraction contained the ssDNA M13KO7 genome, the 0.2 M fraction would contain RNA if there were any, and the combined 0.4 M/1.0 M fraction contained supercoiled dsDNA pM13ori2. Mr: 1 kb ladder (FroggaBio).

Approximately 307.9 µg and 129.2 µg of ssDNA was recovered for each 0.12 M fraction. However, DNA was present in the 0.2 M fraction as well: in total, 59.8 µg and 18.9 µg were recovered per each 0.2 M fraction. In the combined 0.4 M/1.0 M fractions, 5.2 µg and 12.9 µg of DNA was recovered per fraction.

Upon visualization via AGE, the DNA recovered in the 0.2 M fractions appears to be primarily ssDNA. The large amount of ssDNA per sample likely overwhelmed column capacity and was not completely collected during elution with 0.12 M phosphate buffer. However, practically all ssDNA was eluted with 0.12 M and 0.2 M phosphate buffer, with little to none found in the 0.4 M/1.0 M fractions. Additionally, a smaller band (approximately 2 kb) was visible alongside the expected 2.4 kb dsDNA minivector band. Some smearing was visualized It is likely that this DNA species was always present in the lysate, but at such low abundance that it was not detectable until after concentration through HAP chromatography.
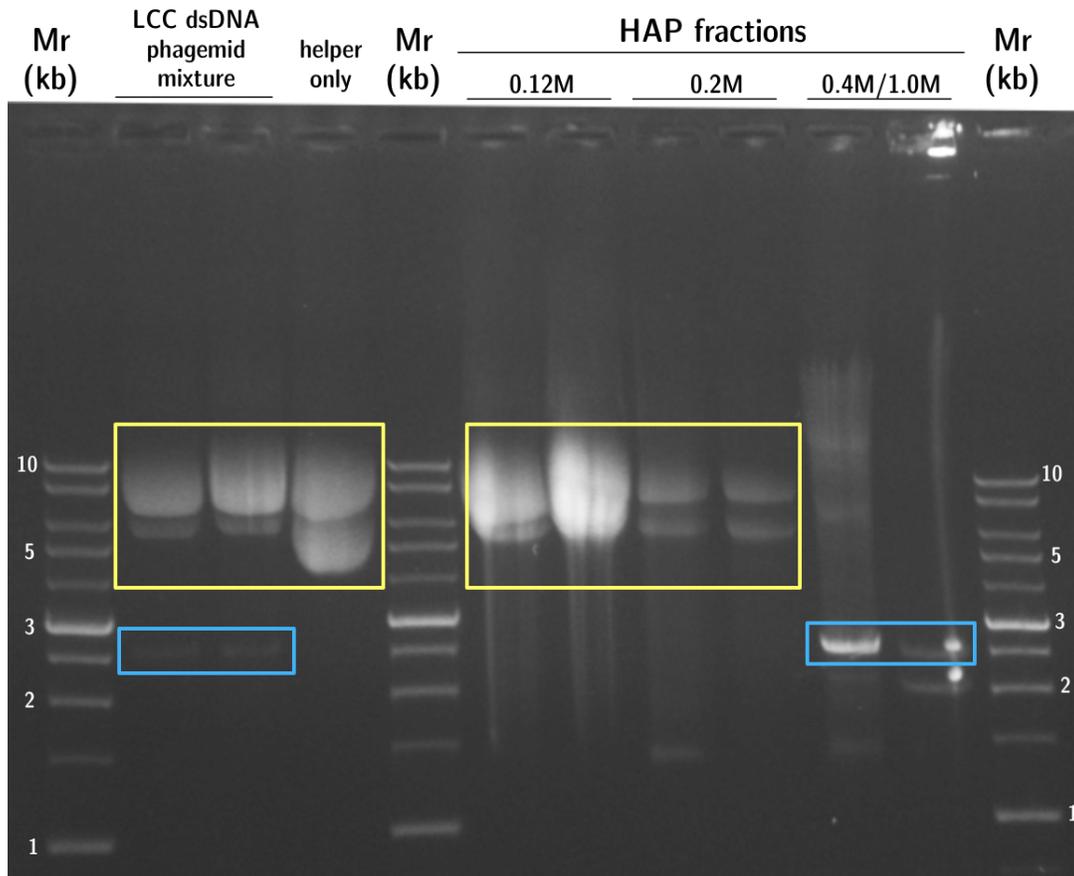
Figure 6.16: Separation of the LCC dsDNA miniphagemid from ssDNA M13KO7 genome using hydroxyapatite chromatography. A mixture of helper genome and LCC dsDNA miniphagemid was subjected to HAP chromatography. The majority of ssDNA M13KO7 genome was recovered in the 0.12 M phosphate fraction, while the ds-mini-(gfp) miniphagemid was recovered in the 0.4 and 1.0 M fractions. Due to the overwhelmingly large amounts of ssDNA, some of it was also present in the 0.2 M fraction. Yellow boxes indicate helper phage ssDNA while blue boxes indicate ds-mini-(gfp) DNA. Mr: 1 kb ladder (FroggaBio).

6.3.5   Transfection of HeLa using SAS

Lastly, we applied the HAP-purified ds-mini-(gfp) phagemid DNA to HeLa cells to assess their capacity for transfection. Gene expression over 24 to 96 h was visualized by fluorescence of green fluorescent protein (GFP) (Figure 6.17).
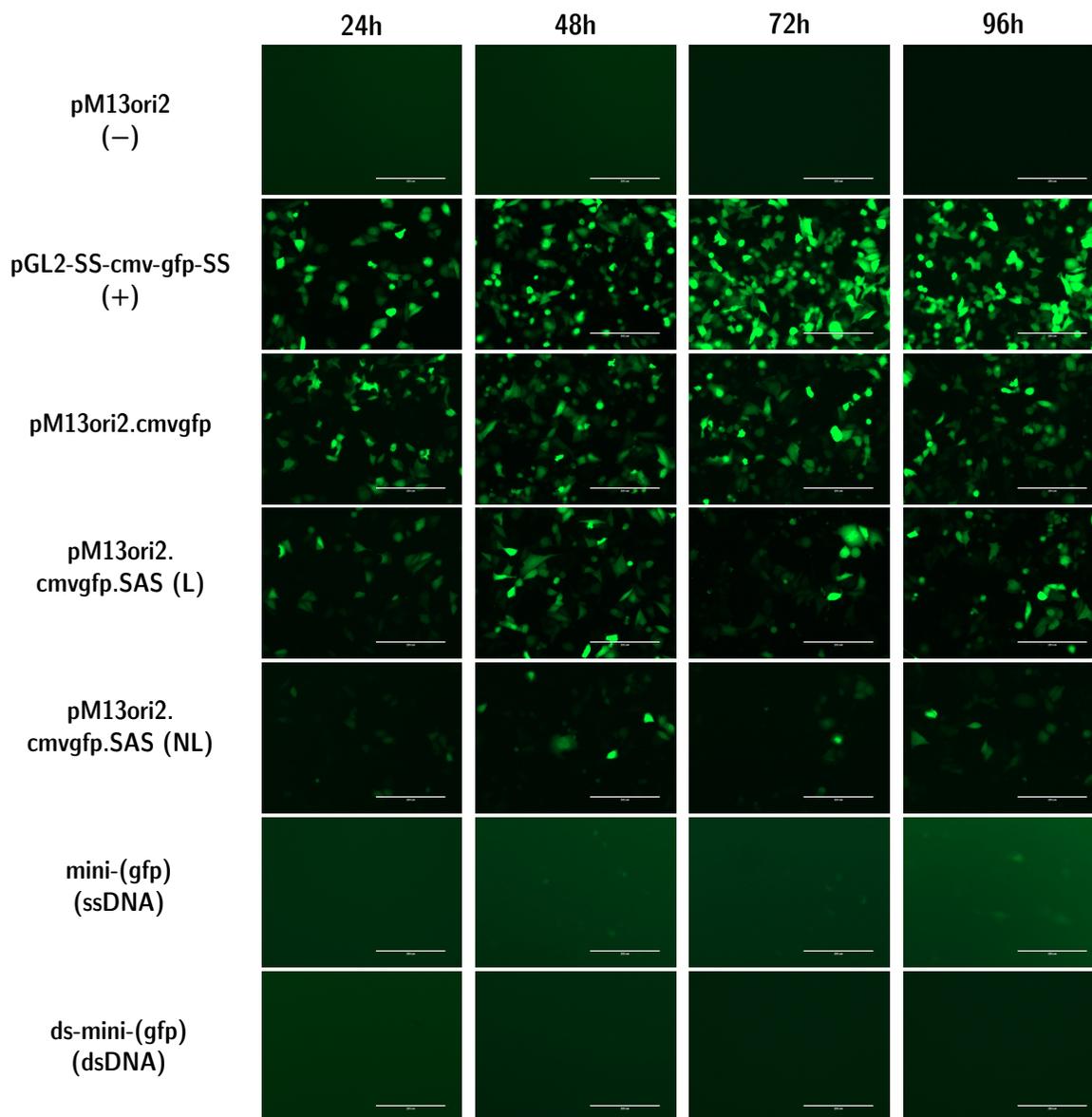
Figure 6.17: GFP fluorescence was not detected from the LCC dsDNA miniphagemid. Transfection of purified DNA was performed in HeLa cells. Gene expression was tracked via visualization of GFP fluorescence at 24, 48, 72, and 96 h after transfection (left to right). From top to bottom, transfection was performed with: 1) the empty vector pM13ori2 as a negative control, 2) the source plasmid pGL2-SS-cmv-gfp-SS as a positive control, 3) the precursor plasmid for mini-(gfp), 4-5) the SAS precursor plasmid (L) and (NL), 6) mini-(gfp) ssDNA, and 7) SAS (L) miniphagemid DNA. Scale bars represent 200 μm.

The ds-mini-(gfp) phagemid DNA was transfected alongside the empty vector, pM13ori2 as a negative control and the source plasmid pGL2-SS-cmv-gfp-SS as a positive control. GFP expression was then compared to transfection of the single-stranded mini-(gfp), and the SAS precursor plasmids with and without a linker. As might be expected, strong GFP fluorescence was observed for the plasmid vectors; however, both SAS precursor plasmids exhibited lower fluorescent intensity per cell compared to pGL2-SS-CMV-GFP-SS and pM13ori2.cmvgfp. This is likely due to the presence of the inverted repeats, as pM13ori2.cmvgfp.SAS (NL) additionally performed worse than pM13ori2.cmvgfp.SAS (L). No GFP fluorescence was detected over the 96 h period from ds-mini-(gfp), indicating it was unsuccessful as a gene transfer vector. In comparison, fluorescence was detected from administration of ssDNA mini-(gfp) after approximately 72 h.

6.4 Discussion

6.4.1 A 4 kb DNA palindrome with a large central spacer is stable in Stbl4

We constructed a plasmid with a large interrupted DNA palindrome approximately 4 kb in size in Stbl4, an *E. coli* strain designed for the cloning of unstable inserts [454, 466, 471]. We attempted to insert a 2 kb inverted repeat with a 50 bp or a 267 bp central spacer region and found that the larger linker was necessary for successful cloning of the inverted repeat in the Stbl4 host background. Interruption of the palindrome with a central non-palindromic spacer appears to be necessary for viable propagation, particularly with larger repeats [564]. Warren and Green (1985) postulated that this spacer could be theoretically as small as 20 bp, but experimentally observed a minimum requirement of 50 bp for viability and 150 bp for stability [564]. Plasmid-borne DNA palindromes have been reported in host backgrounds highly deficient of recombination, including Stbl2, and a number of combined *rec* and *sbc* mutants [562, 567, 569]. Consistent with these observations, we were unable to clone the same inverted repeat with the large linker in JM109 nor with a short linker in Stbl4. Our attempts with a short linker led to a high frequency of deletion events of around 500 bp near the central region of the palindrome (Figures 6.3 and 6.4).

In contrast, deletion of the linker was viable and relatively more stable in Stbl4. We were able to obtain pM13ori2.cmvgfp.SAS (NL) after PacI-mediated excision of the linker, although we found that the PacI site was nonfunctional after ligation. Propagation of this perfect palindrome was unstable and resulted in a high frequency of plasmid splicing in Stbl4, as evidenced by the presence of small unidentified DNA species visible via AGE (Figure 6.6).

The stable propagation of such a large perfect palindrome has not yet been reported to our knowledge. Plasmids with perfect palindromes 2.2 kb long (2×1.1 kb) or interrupted palindromes 500 bp long have been propagated in *sbcC*-deficient *E. coli* strains, but still with rearrangements or deletion frequencies up to 50% [562, 580]. The dsDNA phagemid constructed by Prieto and Sànchez separated two 1.7 kb repeats by a 180 bp central spacer in the phagemid backbone, pBluescript II SK− [343]. Palindrome inviability appears partially dependent on SbcCD cleavage of palindromic

secondary structure, leading to the formation of DSBs [581–583]. Replication of DNA palindromes is susceptible to mutations from misalignment of the palindromic arms [584], which can induce replication slippage. Replication-slipped mispairings can lead to deletions near the centre of symmetry in a SbcCD-independent fashion [583].

Overall, the Stbl4 host background in conjunction with the large linker were sufficient to permit the cloning and propagation of the large interrupted palindrome. Hosts with a combination of *rec* and *sbc* mutations comprise another promising avenue for investigation. In particular, knockout of *sbcCD* [567] and *recF*, which are implicated in the DNA palindrome instability and plasmidic DNA recombination [554], may improve cloning. Importantly, we also found that although the construction required a Stbl4 background, we were able to stably transform and propagate the final construct pM13ori2.cmvgfp.SAS (L) into another *recA* strain like JM109.

6.4.2 A double-stranded miniphagemid vector can be assembled by M13

The helper phage M13KO7 can act upon the split origin of the sense-antisense precursor plasmid to rescue a double-stranded miniphagemid. Both the production of the RFx (Figure 6.8) and the ds-mini-(gfp) (Figure 6.7) were confirmed by AGE visualization of DNA bands at the predicted sizes. The duplex nature of the ds-mini-(gfp) was further confirmed by its resistance to MBN digestion (Figure 6.9). The decrease in size of the ds-mini-(gfp) after extensive MBN digestion is attributed to loss of single-stranded loop regions at the covalent ends of the molecule.

Filamentous phage-mediated DNA replication is initiated by pII binding at the f1 *ori*, then proceeds unidirectionally through rolling circle amplification. The plus strand is displaced by a newly synthesized strand as DNA polymerase proceeds along the minus strand as a template until the termination signal is reached [135]. Normally, the displaced strand is coated by SSB proteins until a new complementary minus strand is synthesized or the plus strand is sequestered from the replication process by pV coating. Since each single strand has a large region of self-complementarity, hairpin and cruciform structure formation can prohibit these activities. It is intriguing that the ds-mini-(gfp) molecule can be extruded through the assembly machinery at all; this suggests both the pII binding region and the packaging signal (PS) are still accessible in the f1 *ori* on a RFx with inverted repeats. Although pV is known to bind nonspecifically to ssDNA, it also has some sequence specificity; in particular, it is known to bind to a 16-mer nucleotide sequence (5′-GUUUUUGGGGCUUUUC-3′) [227, 228] both as an RNA or DNA sequence [225, 226]. Notably, this sequence forms a structure resembling a stem loop. More recently, 58-mer DNA motifs constrained by hairpins were also found to be preferentially bound by pV [228]: again, large loops of unpaired nucleotides are "enclosed" by one or more hairpin loops in a stem-loop structure. It is believed that such stem-loop structures can serve as nucleation sites for pV-DNA [225, 226, 228], whereby saturation of the DNA molecule occurs through cooperative dimer-dimer interactions between pV units. Given that the sense-antisense is contains extraordinary amounts of secondary structure, it would not be surprising if pV nucleation sites exist throughout the single-stranded

RFx intermediate. Importantly, formation of the PS hairpins should be occurring in the SAS RFx in order to facilitate ds-mini-(gfp) phagemid assembly and extrusion. However, we cannot discount the possibility of PS-deficient dsDNA miniphagemid simply assembling alongside helper genomes into multi-genome progeny.

Production of the sense-antisense RFx is strongly correlated to strength of helper phage presence.   We examined if the replication and assembly process could be perturbed by modifications to the phage production workflow outside of genetic engineering (Figures 6.11 to 6.14). As a baseline, we used the following parameters as a reference: M13KO7 helper phage were added to a final concentration of 1×10^8 PFU/mL to a culture of susceptible cells carrying the target phagemid and supplemented with ampicillin once the culture reached $A_{600} = 0.02$. The culture would then be supplemented with kanamycin, and infection proceeded for 16 to 18 h.

Overall, we observed that production of the ds-mini-(gfp) phagemid was more sensitive to perturbations than production of the single-stranded mini-(gfp) (Chapter 3). The duration of infection was limited to 16–18 h, for example. Between the parameters of optical density at time of infection, the MOI, and antibiotic selective pressure, the latter was the most important.

In complex media such as LB, *E. coli* are thought to be in steady-state growth, commonly referred to as the log phase, until approximately $A_{600} = 0.6 - 1.0$. It has even been reported that steady-state growth can stop much earlier, around $A_{600} = 0.3$ [585], after which the growth rate and cell mass gradually decrease. Therefore, we expected that infection by helper phage at low optical density should be superior; pili, the adsorption sites for M13, are lost as cells enter stationary phase [439]. The start of stationary phase, as indicated by the rise of intracellular levels of stationary phase sigma factor (σ^S), begins around $A_{600} = 0.5$ [585, 586]. However, we observed robust RFx conversion even when cells were infected between $A_{600} = 0.5$ and 0.6, which is indicative of sufficient pilus formation to support M13 infection even at this stage of growth. More importantly, the presence of kanamycin was most strongly correlated with good RFx production (Figures 6.11 to 6.14). A strong helper phage presence appears to be necessary to overcome the inefficiency of processing the split origin around a DNA palindrome.

Although construction of the sense-antisense vector was only possible in Stbl4, not JM109, we found that the reverse was true for phagemid rescue. In our previous work with Stbl4 (Chapter 4), phagemid rescue was highly efficient with both a wild-type and split f1 origin. However, RFx production and ds-mini-(gfp) was not detectable in Stbl4 during this study; hence, we moved the process to JM109. Phage-mediated replication of this phagemid particularly is adversely affected in Stbl4 but the reason is unclear. Interestingly, Leach (1996) observed that stability of inverted repeats in M13 was particularly poor [562]. Furthermore, palindrome-bearing plasmids are more susceptible to rearrangements or deletion if present in high copy number [562]. Modifying the plasmid origin, the helper phage origin, and/or the host strain background may lead to a more suitable combination for double-stranded filamentous phagemid production.

Our work here only scratches the surface of the highly complex interplay between phage proteins and dsDNA substrates. Taking our results together, we have identified several areas for further study. Firstly, strain background may have an even greater impact on the rescue of the double-stranded miniphagemid than it did on single-stranded miniphagemids. Hence, investigation into different strain backgrounds may prove fruitful. Secondly, robust helper phage presence appears strongly correlated with good RFx production. Therefore, increasing MOI (i.e. increasing phage concentrations) appears to be a promising line of inquiry. The strength of the antibiotic selective pressure for the helper phage is another parameter that can be investigated more thoroughly. Thirdly, the phagemid plasmid origin is high copy number, which may be contributing to palindrome instability. The use of lower copy plasmids may be beneficial to increase palindrome stability, although it is unclear what the impact on phagemid rescue would be. For helper phage rescue of palindromic phagemids, switching the high copy origin (pMB1) and medium copy origin (p15a) of the SAS phagemid and helper phage may prove useful. In addition, we observed in earlier studies that culture conditions can drastically alter resulting phage yield (Chapter 3). The results here were based on $10\,\text{mL}$ cultures; thus, it will become important to investigate how scaling up may affect yield. In investigations of large-scale production of M13, growth medium and culture pH were two important parameters that affected host cell health (and consequently, cell phage output) [431, 587]. Presence of antibiotic reduced yield from the additional metabolic burden; in these small-scale studies, we found that antibiotic selective pressure to maintain helper phage levels important. Indeed, maintenance of plasmid, helper phage, and RFx subspecies likely all contribute to substantial metabolic burden on host cells. Thus, the need to further balance SAS precursor-RFx conversion adds another layer of complexity to scale-up, which will need to be addressed for future applications of this technology.

6.4.3   HAP chromatography can purify ds-mini-(gfp) DNA

Purification of double-stranded products from ssDNA helper phage can be reliably accomplished using hydroxyapatite [571, 572]. HAP has long been used to fractionate on the basis of DNA composition [577, 588]. An important caveat is the amount of helper phage, as the large degree of helper phage contamination in our lysates overwhelmed the capacity of the hydroxyapatite. Robust column optimization is important, which may be difficult to achieve if sample composition is not known ahead of time. In all our samples of recovered dsDNA, we observed a reduced A260/230 ratio, indicative of residual contaminants that absorb strongly at $230\,\text{nm}$. This is most likely phenol, as phenol:chloroform extraction was necessary to purify the fractionated DNA. Phenol contamination was not apparent in fractions of purified ssDNA. Likely, the higher salt content in the (up to $1.0\,\text{M}$ phosphate buffer) dsDNA fractions impeded efficient purification. The recovery of dsDNA was also much lower than that of ssDNA, which may also be related. On the whole, we were able to use HAP chromatography to fully separate undesired ssDNA helper phage genome.

Transfection of the ds-mini-(gfp) was inconclusive.   While we were able to demonstrate the production of a double-stranded DNA vector using filamentous phage M13, we were unable to observe gene expression from its transfection in cells in vitro. We hypothesize this was most likely related to the low A260/230 ratio as potential contaminants such as phenol will impair transfection.

Another explanation for the lack of gene expression with the ds-mini-(gfp) may be imperfect annealing of the two complementary sense-antisense sequences. Replication of inverted repeats is prone to error, where replication slippage can lead to nucleotide mismatch [584]. Although the size and duplex nature of the ds-mini-(gfp) molecule were as confirmed, debilitating point mutations may have accrued during its replication. Transfection of both the interrupted palindromic *cmv-gfp* (L) and perfect palindromic (NL) plasmids showed qualitatively less fluorescence: GFP fluorescence was observed in fewer cells and the fluorescence was also less intense per cell. Gene expression from the (NL) plasmid was also worse in comparison to the (L) plasmid, which can be attributed to the greater instability of the perfect palindrome. This is in line with previous observations of the instability of inverted repeats in mammalian and bacterial cells [561].

6.4.4   Summary

Overall, we demonstrated that a phagemid vector carrying a large DNA palindrome could be constructed if a sufficiently large central spacer region was present. We found that a central spacer of 267 bp was superior to a 50 bp spacer for this size of palindrome. Although filamentous phage M13 is a ssDNA phage, we were able to demonstrate helper phage-rescue of a dsDNA phagemid DNA, confirmed by its resistance to MBN degradation. Furthermore, we were able to isolate the dsDNA phagemid from helper phage genome in the resulting lysate using HAP. This lays the groundwork for a novel phage-based gene transfer platform.

Chapter 7

Conclusions and future directions

7.1 Conclusions

The genetic simplicity and flexibility of the filamentous bacteriophage (phage) M13 make it a promising platform for mammalian cell targeting and gene transfer. Its capacity for phage display simplifies tissue targeting by facilitating the display of any cell-specific ligand. As phages are generally inexpensive to produce in large quantities, they are highly amenable to large-scale manufacture as biopharmaceuticals. In this work, we constructed and characterized a novel phage gene delivery vector platform and demonstrated its application in vitro.

In Chapter 3, we showed that a single-stranded (ss) miniaturized phagemid devoid of bacterial or phage backbone could be produced when a mammalian gene expression cassette was flanked by separated domains of the f1 functional origin (f1 *ori*). Despite the split origin on the precursor plasmid, M13KO7 was able to produce a recombinant RF (RFx) absent of the plasmid backbone and assemble the minivector phagemid into progeny virions. Miniphagemid yields arising from a split origin precursor plasmid were comparable, and even superior, to yields from phagemids with intact wild-type origins. The efficiency of phagemid rescue by helper phage M13KO7 is known to be dependent on length [315], but we further demonstrated that the sequence composition plays a far more influential role.

Quantification of the phagemid and helper phage in the same lysate showed that miniphagemids encoding *cmv-gfp* were more efficiently packaged than isogenic non-miniaturized phagemids ("full" phagemids) with a wild-type f1 *ori*. For other genetic cassettes, the phagemid:helper ratio was far lower, although uncorrelated with total phagemid length. While phagemid rescue could be very efficient for some transgenes, the final lysate still contained some degree of helper phage contamination, which we were unable to separate from the target phagemids. Additionally, phagemid:helper ratio was only evaluated across three mammalian transgene cassettes, limiting the scope of the study. We postulated that secondary structure within the genetic cassette could be the driving factor behind a preferential helper:phagemid ratio.

Miniphagemid production was further characterized across different *E. coli* host backgrounds in Chapter 4. In a head-to-head comparison of six *E. coli* strains, JM109 and Stbl4 appeared to be the best strains for helper phage M13KO7-mediated rescue of miniphagemid virions, with miniphagemid lysate compositions over 95%. Additionally, robust mini and full phagemid assembly in DH5α showed that the F-factor was non-essential if the helper phage could be maintained via a plasmid origin. Based on our results in these six strains, strong miniphagemid rescue appeared to be correlated with hosts that had mutations in *endA*, *recA*, and *gyrA*.

Furthermore, the packaging signal (PS) was eliminated in M13KO7 in order to reduce extrusion of the helper genome into the phage lysate. Phage without the PS (M13SW8) were unable to form plaques on susceptible *E. coli*, but were still able to rescue phagemid with efficiency on par to its parent phage. The removal of the PS dramatically reduced the number of infective helper phage virions in the phage lysate, leading to a much improved phagemid:helper ratio. Based on the results in this chapter, we postulated that the small level of PS-deficient helper phage genomes still in the phage lysate may be co-extruded alongside phagemid molecules, or that recombination events between the helper phage genome and phagemid in the cell could reconstitute a wild-type PS. Analysis of replicative factor (RF) DNA also indicated that other recombination events may be occurring in cells transformed by PS-deficient helper phage. The presence of an unexpected DNA species in the RF DNA of many M13SW8-rescued lysates was suggestive of "miniphage" formation; these are recombination products believed to arise from phage genome rearrangement whereby all genetic elements except for the phage origin are lost [589]. We postulated that the high accumulation of unextruded helper phage genome could lead to miniphage conversion. The propensity of accumulated helper phage to form miniphage remains a challenge for this system.

Additionally, a helper phage was constructed to display the cell-specific targeting ligand epidermal growth factor (EGF). This helper phage was used to assemble EGF$^+$ mini and full phagemids encoding *cmv-gfp* or *cmv-luc*. In Chapter 5, these hybrid targeted phages were applied to cells expressing the EGF receptor (EGFR) in order to assess their capacity for gene transfer. Internalization of EGF-displaying phages in was confirmed in the EGFR$^+$ HeLa cell line; display of the cell-specific ligand dramatically improved gene transfer of phagemids. The EGF ligand was not observed to activate phosphorylation of the EGF receptor prior to cellular internalization, although this does not preclude later signal transduction, nor transduction independent of phosphorylated-epidermal growth factor receptor (EGFR) [534, 537, 538].

Display of receptor-targeting peptides can therefore confer mammalian cell tropism to phage particles; this overcomes one of the biggest obstacles to the use of phage as a genetic therapeutic in an animal model. One caveat, however, is that the display of a receptor-targeted ligand alone appears insufficient to ensure highly efficient gene transfer. Treatment of other cell lines with similar levels of EGFR did not result in as high gene expression, although EGF display did improve expression over non-displaying phage per cell line. Other cell-specific factors can impede

the rate of phage uptake, the rate of nuclear penetration, the rate of gene transcription, or all of the above [328]. Furthermore, the addition of a commercial cationic polymer transfection reagent further improved phage-mediated gene transfer, which we postulated was related to improved endosomal escape. While receptor targeting was the single largest factor influencing successful gene transfer, miniaturization of EGF-displaying phagemids was also associated with an increase in gene expression. Intriguingly, gene expression was demonstrated from both EGF^+ and EGF^- miniphagemids in the $EGFR^-$ cell line, HEK293T; this shows that cell internalization of filamentous phage can occur independently of ligand-receptor endocytosis, although the mechanism is unclear. Based on the results in this chapter, the combination of a cationic polymer, cell-specific targeting ligand, and miniaturization of the transgene-encoding phagemid was found to improve phage-mediated gene transfer.

Finally, we sought to further advance the utility of M13 gene delivery by assembling a double-stranded (ds) phagemid minivector in Chapter 6. Building upon the miniphagemid design from Chapter 3, we constructed a sense-antisense (SAS) precursor plasmid in which a repeat of the *cmv-gfp* gene cassette was inserted in reverse, such that both the sense and antisense sequences were present on one DNA strand. Rescue of the phagemid from this SAS precursor demonstrated the assembly and production of a dsDNA filamentous phagemid virion. While Stbl4 was a suitable host for the production of single-stranded miniphagemid particles, we were unable to produce a double-stranded equivalent. Consistent with our hypothesis in Chapter 3, an RF molecule with direct inverted repeats is a poor substrate for the phage replication protein pII. A major challenge with this system will be the construction and maintenance of a stable palindrome precursor plasmid.

Overall, the incorporation of two separated domains of the f1 *ori* is a straightforward improvement over a single wild-type origin and results in high titres of miniphagemid gene transfer vectors. In support of this platform, we constructed a backbone vector, pM13ori, which contains a polylinker to accommodate cloning of other transgenes. The optimizations presented this work here establish how filamentous phage gene delivery vectors can better overcome cellular gene transfer barriers, including: bypassing the plasma membrane through receptor-targeted endocytosis, efficiently navigating the cytosol through improved endosomal escape, and accessing nuclear transcriptional machinery with the miniaturized phagemid structure. These explorations expand the utility of bacteriophage M13 as a platform for the production of genetic therapies.

7.2 Future directions

A number of new hypotheses and new lines of inquiry have arisen from the results compiled in the preceding chapters. These questions can form the basis of new research directions, which will further advance the technology characterized in this work.

7.2.1 Improving the efficiency of miniphagemid production

In Chapter 3, we investigated the utility of a split origin precursor plasmid for the generation of miniaturized phagemid vectors encoding mammalian transgenes for gene transfer applications. We hypothesized that good helper phage rescue of miniphagemid from a precursor plasmid was dependent on secondary structure formation in the encoded sequence. As such, it would be useful to examine a larger array of sequences where secondary structures, if present, have been characterized. We found that phagemids encoding *cmv-gfp* were assembled with great efficiency but the addition of a *tet* fragment from pBR322 reduced M13KO7-mediated rescue drastically. Hence, it would be beneficial to examine these elements in isolation: for example a comparison of miniphagemids encoding portions of each genetic cassette. Importantly, the *cmv-luc* cassette was assembled with very poor efficiency, although it shares a great number of features with the *cmv-gfp* and *cmv-gfp-tet* cassettes. Quantification of M13KO7-rescued phagemids encoding isolated portions of the *cmv-luc* gene cassette could determine the problematic sequences that interfere with phage replication. The hypothesized secondary structure should also be verified experimentally: DNA cruciform structures can be visualized using atomic force microscopy, for example [590].

Furthermore, we showed the production of miniaturized phagemids as a proof of principle using the helper phage M13KO7 across 6 *E. coli* host backgrounds, all of which were derived from *E. coli* K12. Based on the results of ch:helper, other *E. coli* strains are also promising candidate hosts. In particular, we suggested a *recA* host derived from *E. coli* B, Stbl3. Notably, a study evaluating strains for their capacity to propagate supercoiled plasmid DNA identified a phenomenon of runaway plasmid replication in TG1, HB101, and MG1655 [445]. Plasmid copy number was much higher than expected in these strains, which could be very beneficial from a production standpoint, and may contribute to a much higher phagemid yield. A number of other helper phages and helper plasmid vectors have been constructed by other groups [252, 316, 591], including helper plasmids completely devoid of the f1 *ori* [318]. The optimal host strain/helper phage combination has yet to be identified.

7.2.2 Hybrid phage display applications

Functionalization of the phagemid particle via phage display has been applied by many groups to extend the applicability of filamentous phage-mediated mammalian targeting. We demonstrated pIII display of an approximately 50 a.a. peptide for cell-specific targeting, which did not impact infectivity. However, all the coat proteins of M13 have the capacity for display; as no coat protein is involved in replication, it is not likely that fusion coat proteins would have any impact on miniphagemid processing. It would be useful to explore combinatorial display; for example, a 388 or 38 + 8 system, whereby the helper phage has fusion pIII and pVIII for multifunctional display. Importantly, homogeneous pVIII display can only support very short peptides; for the display of larger peptides, a second copy of wild-type gVIII should be present either on the phagemid or in the helper phage for heterogeneous display.

Indeed, pIII display of EGF only presents five copies of the ligand per phage. A natural extension of the work presented in this thesis would, therefore, be the generation of Type 8 or 88 helper phages displaying many more copies of EGF to see this further improves gene transfer. As the peptide is too large for homogeneous pVIII display, either shorter EGFR-targeting peptides [482] or heterogeneous pVIII display must be used. Additionally, other cell-specific ligands should be tested to better characterize cell-specific uptake of phage. Towards this goal, we have begun work on the display of other peptides on M13SW7. In particular, we have successfully displayed a different targeting peptide from the lung Ace2 receptor-targeting spike protein of the coronavirus SARS-CoV-2. To further functionalize the hybrid miniphagemid particle, a combinatorial display of a peptide for cell-targeting, another peptide for intracellular trafficking, and/or a peptide for endosomal escape may prove superior.

We further investigated the pIII display of an 11-residue highly cationic peptide that may improve endosomal escape [592]. However, this fuses poorly in the N-terminal region of pIII and drastically impedes infectivity. This peptide may be best suited for C-terminal pIII display, as it is exposed to the cytoplasm whereas the N-terminus is exposed to the periplasm. Altogether, we envision a multifunctional chimeric M13 phage particle encapsulating a miniphagemid gene delivery vector, not only functionalized by the display of a cell-specific targeting ligand on pVIII, but also by the display of other peptides to improve cell and nuclear penetration on other coat proteins.

7.2.3   Improved characterization of phage-mediated gene delivery

Our observations of miniphagemid gene transfer indicate that filamentous phage M13 internalization is dependent on EGF-EGFR endocytosis, which is primarily mediated by clathrin. Many bacteriophages, including filamentous phages, can enter mammalian cells by both clathrin-dependent and clathrin-independent pathways. Based on the results in Chapter 5, we propose alternative studies to better understand the intracellular localization of phage over time. Internalized phage can be quantified by lysing cells after phage uptake both through infectivity and DNA quantitation assays, such as quantitative PCR (qPCR). Importantly, this can determine the proportion of phage particles that remain intact within the cell. Additionally, visualization through confocal microscopy, for example, can precisely show the intracellular location of phage particles. The use of endocytic disruptive agents [527] such as inhibitors of clathrin (chlorpromazine), caveolae (methyl-β-cyclodextrin), and endosome acidification (chloroquine) can also help shed further light which pathway is used.

Here, a cell-specific targeted miniaturized phagemid particle complexed with a cationic polymer transfection reagent appeared to be the best combination for phage-mediated gene transfer. To better characterize the dynamics of the polymer-phage complexes, future studies can determine particle size, through dynamic light scattering techniques or through visualization via scanning or transmission electron microscopy. As advances in nanoparticle development have demonstrated superior cell penetration and cell trafficking [593], novel phage-nanoparticle formulations may be

of interest. Since the phagemids under study were concentrated using polyethylene glycol (PEG), separation of phage particles from PEG [430, 520] may be necessary to fully characterize the interactions between phages and co-transfection agents. Given that M13 phagemid DNA is single-stranded, the use of DNA damaging agents can drastically improve transfection efficiency [344]. A combination of camptothecin (or other DNA damaging agent), cationic polymer, and cell-targeted miniphagemid may therefore prove to be highly synergistic [344, 541].

Further characterization of phage-mediated EGFR activation. Although we suspect filamentous phage displaying EGF activates the EGFR downstream signalling pathway, the results were inconclusive. We were unable to observe autophosphorylation from the treatment of EGFR$^+$ cells with EGF$^+$ phage, but it has been reported that cell-surface phosphorylation of EGFR is not the only means by which signal transduction can occur [534, 537, 538]. Thus, future work should assess if gene transcription or other late events in EGFR-stimulated pathways, for example, can be activated by EGF-displaying filamentous phage or EGF-displaying miniphagemids.

7.2.4 Improved characterization of double-stranded miniphagemid vectors

Construction of the SAS precursor plasmid and M13KO7-mediated rescue of the resultant dsDNA phagemid are both highly inefficient processes. Notably, *sbcCD* mutations have been implicated in impaired propagation of DNA palindromes [565, 567, 569, 594, 595] and *recF* in plasmidic recombination. Hence, future SAS vector construction may benefit from the use of a strain such as SURE or SURE2 [568]. A relatively large central spacer was necessary for the cloning of a direct inverted repeat; however it would be fruitful to see if an *sbcCD* and/or *recF* host can support the construction of a large perfect palindrome or and/or one with a smaller spacer.

We identified several areas of further study with regard to the production workflow in Chapter 6: increased multiplicity of infection (MOI), assessment of a range of kanamycin concentrations, and construction of a lower copy SAS precursor plasmid. As helper phage maintenance was vital to robust production of the RFx, construction of a higher copy helper plasmid with a low or medium copy SAS precursor plasmid would be interesting to explore. In addition, different complex media formulations may also be viable. We did not observe major differences between preliminary tests in Luria–Bertani (LB) versus 2×YT medium. However, other rich media including Terrific Broth [596], Super Optimal Broth [597], and other plasmid production supporting media [598] may prove fruitful [598, 599]. Along these lines, the use of a defined medium would help elucidate particular metabolites vital to phagemid production. For example, Mg^{2+} ions are important for DNA replication so supplementing cultures with Mg^{2+} improves phage production [600]. Furthermore, elucidation of the entire metabolome could lead to better culture conditions.

Genetic manipulations can also improve helper phage action on the split origin. Firstly, we previously postulated that genome copies of the M13SW8 helper phage we constructed in Chapter 4 accumulate intracellularly due to the removal of the PS. Although we further postulated that this

also led to spontaneous miniphage conversion of the helper, it would still be interesting to see how this helper and others perform in the rescue of the dsDNA phagemid. We also recommended modifying the plasmid origins of the SAS precursor and the helper phage in order to assess phagemid rescue with a high copy helper phage (ex: pMB1) and low copy phagemid precursor (ex: p15a). This could saturate the cell with helper phage genome copies and reduce the number of palindromes available for recombination.

Given that pV overall binds poorly to dsDNA [214, 219, 220], another avenue for improved extrusion could be to increase the number of available pV proteins. This may be accomplished by increasing the number of copies of gV in the helper phage genome, or by incorporating a stronger promoter. Additionally, incorporation of multiple PS, while contributing to the overall number of repetitive DNA sequences, may help the extrusion of dsDNA progeny phage.

Also, while we demonstrated the successful isolation of dsDNA miniphagemid from ssDNA helper phage via hydroxyapatite (HAP) chromatography, this process can be optimized for improved recovery of the double-stranded miniphagemid. Given the considerable amount of helper phage genome in extracted ssDNA lysates, we propose the calibration of HAP against a range of increasing ssDNA concentrations in order to determine to maximal capacity per column. The discovery of other DNA bands upon agarose gel electrophoresis (AGE) visualization of the dsDNA HAP fraction suggested extrusion of other DNA species. Thus, more thorough genetic characterization of products from helper phage rescue of the dsDNA phagemid may prove informative. For example, the resultant helper phage genomes, putative dsDNA phagemid, and the other byproducts we observed via AGE should be sequenced to determine if point mutations did occur. Visualization of products through techniques such as atomic force microscopy, including the SAS RFx, may better demonstrate if inhibitory DNA cruciform structures are indeed forming. Secondary structure may also be inferred from non-denaturing gel electrophoresis.

Letters of Copyright Permission

Copyright permissions were obtained for the reproduction of the image in Figure 2.1:

License Number	5191000332052	
License date	Nov 16, 2021	

✅ Licensed Content

Licensed Content Publisher	Oxford University Press
Licensed Content Publication	Nucleic Acids Research
Licensed Content Title	λ ZAP: a bacteriophage λ expression vector with *in vivo* excision properties
Licensed Content Author	Short, Jay M.; Fernandez, Joseph M.
Licensed Content Date	Aug 11, 1988
Licensed Content Volume	16
Licensed Content Issue	15

📋 Order Details

Type of Use	Thesis/Dissertation
Requestor type	Educational Institution/Non-commercial/ Not for-profit
Format	Print and electronic
Portion	Figure/table
Number of figures/tables	1
Will you be translating?	No

📄 About Your Work

Title	Construction and characterization of a hybrid phage gene delivery platform
Institution name	University of Waterloo
Expected presentation date	Mar 2022

📁 Additional Data

Portions	Figure 3, p. 7590

📍 Requestor Location

Requestor Location	Shirley Wong 200 University Ave W Waterloo, ON N2L 3G1 Canada Attn: Shirley Wong

🧾 Tax Details

Publisher Tax ID	GB125506730

💲 Price

Total	0.00 CAD

References

(1) Gene Therapy Clinical Trials Worldwide `https://www.abedia.com/wiley/`.

(2) Ma, C.-C.; Wang, Z.-L.; Xu, T.; He, Z.-Y.; Wei, Y.-Q. The approved gene therapy drugs worldwide: from 1998 to 2019. *Biotechnology Advances* **2020**, *40*, 107502.

(3) Pushparajah, D.; Jimenez, S.; Wong, S.; Alattas, H.; Nafissi, N.; Slavcev, R. A. Advances in gene-based vaccine platforms to address the COVID-19 pandemic. *Advanced Drug Delivery Reviews* **2021**, *170*, 113–141.

(4) Sum, C. H.; Shortall, S. M.; Wong, S.; Wettig, S. D. In *Nanomedicine: Gene Delivery, Imaging and Evaluation Systems*, Slavcev, R. A., Wettig, S., Zeng, Z., Eds.; Experientia Supplementum; Springer International Publishing: Cham, 2018, pp 3–68.

(5) Kotterman, M. A.; Schaffer, D. V. Engineering adeno-associated viruses for clinical gene therapy. *Nature Reviews Genetics* **2014**, *15*, 445–51.

(6) Daya, S.; Berns, K. I. Gene therapy using adeno-associated virus vectors. *Clinical Microbiology Reviews* **2008**, *21*, 583–593.

(7) Ginn, S. L.; Amaya, A. K.; Alexander, I. E.; Edelstein, M.; Abedi, M. R. Gene therapy clinical trials worldwide to 2017: An update. *Journal of Gene Medicine* **2018**, *20*, 1–16.

(8) Giacca, M.; Zacchigna, S. Virus-mediated gene delivery for human gene therapy. *Journal of Controlled Release* **2012**, *161*, 377–388.

(9) Zeyaullah, M.; Patro, M.; Ahmad, I.; Ibraheem, K.; Sultan, P.; Nehal, M.; Ali, A. Oncolytic viruses in the treatment of cancer: A review of current strategies. *Pathology & Oncology Research* **2012**, *18*, 771–781.

(10) Chong, H.; Vile, R. G. Replication-competent retrovirus produced by a 'split-function' third generation amphotropic packaging cell line. *Gene Therapy* **1996**, *3*, 624–629.

(11) Chong, H.; Starkey, W.; Vile, R. G. A replication-competent retrovirus arising from a split-function packaging cell line was generated by recombination events between the vector, one of the packaging constructs, and endogenous retroviral sequences. *Journal of Virology* **1998**, *72*, 2663–2670.

(12) Garrett, E.; Miller, A. R.; Goldman, J. M.; Apperley, J. F.; Melo, J. V. Characterization of recombination events leading to the production of an ecotropic replication-competent retrovirus in a GP+envAM12-derived producer cell line. *Virology* **2000**, *266*, 170–179.

(13) Cavazzana-Calvo, M.; Hacein-bey, S. Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. *Science* **2000**, *288*, 669–672.

(14) Hacein-Bey Abina, S.; von Kalle, C.; Schmidt, M.; Le Deist, F.; Wulffraat, N.; McIntyre, E.; Radford, I.; Villeval, J.-L.; Fraser, C. C.; Cavazzana-Calvo, M.; Fischer, A. A serious adverse event after successful gene therapy for x-linked severe combined immunodeficiency. *New England Journal of Medicine* **2003**, *348*, 255–266.

(15) Hacein-Bey-Abina, S. et al. Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. *The Journal of Clinical Investigation* **2008**, *118*, 3132–3142.

(16) Hacein-Bey-Abina, S. et al. Efficacy of gene therapy for X-linked severe combined immunodeficiency. *The New England Journal of Medicine* **2010**, *363*, 355–364.

(17) Howe, S. J. et al. Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. *The Journal of Clinical Investigation* **2008**, *118*.

(18) Wilson, J. M. Lessons learned from the gene therapy trial for ornithine transcarbamylase deficiency. *Molecular Genetics and Metabolism* **2009**, *96*, 151–157.

(19) Jenks, S. Gene therapy death – "Everyone has to share in the guilt". *Journal of the National Cancer Institute* **2000**, *92*, 98–100.

(20) Verma, I. M. A tumultuous year for gene therapy. *Molecular Therapy* **2000**, *2*, 415–416.

(21) Klok, F. A.; Pai, M.; Huisman, M. V.; Makris, M. Vaccine-induced immune thrombotic thrombocytopenia. *The Lancet Haematology* **2021**, *9*, e73–e80.

(22) Philippidis, A. Fourth boy dies in clinical trial of Astellas' AT132. *Human Gene Therapy* **2021**, *32*, 1008–1010.

(23) Flotte, T. R. Birth of a new therapeutic platform: 47 years of adeno-associated virus biology from virus discovery to licensed gene therapy. *Molecular Therapy* **2013**, *21*, 1976–1981.

(24) Srivastava, A. AAV Vectors: Are they safe? *Human Gene Therapy* **2020**, *31*, 697–699.

(25) Maurya, S.; Sarangi, P.; Jayandharan, G. R. Safety of adeno-associated virus-based vector-mediated gene therapy—impact of vector dose. *Cancer Gene Therapy* **2022**, 1–2.

(26) Fausther-Bovendo, H.; Kobinger, G. P. Pre-existing immunity against Ad vectors. *Human Vaccines & Immunotherapeutics* **2014**, *10*, 2875–2884.

(27) Calcedo, R.; Vandenberghe, L. H.; Gao, G.; Lin, J.; Wilson, J. M. Worldwide epidemiology of neutralizing antibodies to adeno-associated viruses. *The Journal of Infectious Diseases* **2009**, *199*, 381–390.

(28) Duerr, A.; Huang, Y.; Buchbinder, S.; Coombs, R. W.; Sanchez, J.; del Rio, C.; Casapia, M.; Santiago, S.; Gilbert, P.; Corey, L.; Robertson, M. N.; Step/HVTN 504 Study Team Extended follow-up confirms early vaccine-enhanced risk of HIV acquisition and demonstrates waning effect over time among participants in a randomized trial of recombinant adenovirus HIV vaccine (Step Study). *The Journal of Infectious Diseases* **2012**, *206*, 258–266.

(29) Buchbinder, S. P.; McElrath, M. J.; Dieffenbach, C.; Corey, L. Use of adenovirus type-5 vectored vaccines: a cautionary tale. *The Lancet* **2020**, *396*, e68–e69.

(30) Manning, W. C.; Zhou, S.; Bland, M. P.; Escobedo, J. A.; Dwarki, V. Transient immunosuppression allows transgene expression following readministration of adeno-associated viral vectors. *Human Gene Therapy* **1998**, *9*, 477–485.

(31) Moskalenko, M.; Chen, L.; van Roey, M.; Donahue, B. A.; Snyder, R. O.; McArthur, J. G.; Patel, S. D. Epitope mapping of human anti-adeno-associated virus type 2 neutralizing antibodies: implications for gene therapy and virus structure. *Journal of Virology* **2000**, *74*, 1761–1766.

(32) Louise, C. In Philip, H.; Marlene, R. *Transplantation Immunology*; Humana Press: New Jersey, 2006, pp 201–226.

(33) Watson, J. D.; Crick, F. H. C. Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid. *Nature* **1953**, *171*, 737–738.

(34) Sum, C. H.; Wettig, S.; Slavcev, R. Impact of DNA vector topology on non-viral gene therapeutic safety and efficacy. *Current Gene Therapy* **2014**, *14*, 309–329.

(35) Egawa, G.; Nakamizo, S.; Natsuaki, Y.; Doi, H.; Miyachi, Y.; Kabashima, K. Intravital analysis of vascular permeability in mice using two-photon microscopy. *Scientific Reports* **2013**, *3*, 1932.

(36) Krieg, A. M.; Hartmann, G.; Yi, A.-K. In *Immunobiology of Bacterial CpG-DNA*, Wagner, H., Ed.; Current Topics in Microbiology and Immunology; Springer: Berlin, Heidelberg, 2000, pp 1–21.

(37) Lechardeur, D.; Sohn, K. J.; Haardt, M.; Joshi, P. B.; Monck, M.; Graham, R. W.; Beatty, B.; Squire, J.; O'Brodovich, H.; Lukacs, G. L. Metabolic instability of plasmid DNA in the cytosol: a potential barrier to gene transfer. *Gene Therapy* **1999**, *6*, 482–497.

(38) Lukacs, G. L.; Haggie, P.; Seksek, O.; Lechardeur, D.; Freedman, N.; Verkman, A. S. Size-dependent DNA mobility in cytoplasm and nucleus. *The Journal of Biological Chemistry* **2000**, *275*, 1625–1629.

(39) Bai, H.; Lester, G. M. S.; Petishnok, L. C.; Dean, D. A. Cytoplasmic transport and nuclear import of plasmid DNA. *Bioscience Reports* **2017**, *37*, BSR20160616.

(40) Lu, J.; Wu, T.; Zhang, B.; Liu, S.; Song, W.; Qiao, J.; Ruan, H. Types of nuclear localization signals and mechanisms of protein import into the nucleus. *Cell Communication and Signaling* **2021**, *19*, 1–10.

(41) Dean, D. A.; Dean, B. S.; Muller, S.; Smith, L. C. Sequence requirements for plasmid nuclear import. *Experimental Cell Research* **1999**, *253*, 713–722.

(42) Van Gaal, E. V. B.; Oosting, R. S.; van Eijk, R.; Bakowska, M.; Feyen, D.; Kok, R. J.; Hennink, W. E.; Crommelin, D. J. A.; Mastrobattista, E. DNA nuclear targeting sequences for non-viral gene delivery. *Pharmaceutical Research* **2011**, *28*, 1707–1722.

(43) Lam, A. P.; Dean, D. A. Progress and prospects: nuclear import of nonviral vectors. *Gene Therapy* **2010**, *17*, 439–47.

(44) Yin, H.; Kanasty, R. L.; Eltoukhy, A. A.; Vegas, A. J.; Dorkin, J. R.; Anderson, D. G. Non-viral vectors for gene-based therapy. *Nature Reviews Genetics* **2014**, *15*, 541–555.

(45) Wang, Z. et al. Detection of integration of plasmid DNA into host genomic DNA following intramuscular injection and electroporation. *Gene Therapy* **2004**, *11*, 711–721.

(46) Mintzer, M. A.; Simanek, E. E. Nonviral vectors for gene delivery. *Chemical Reviews* **2009**, *109*, 259–302.

(47) Ledwith, B. J.; Manam, S.; Troilo, P. J.; Barnum, A. B.; Pauley, C. J.; Ii, T. G. G.; Harper, L. B.; Beare, C. M.; Bagdon, W. J.; Nichols, W. W. Plasmid DNA vaccines: Investigation of integration into host cellular dna following intramuscular injection in mice. *Intervirology* **2000**, *43*, 258–272.

(48) Manam, S.; Ledwith, B. J.; Barnum, A. B.; Troilo, P. J.; Pauley, C. J.; Harper, L. B.; Griffiths, T. G.; Niu, Z.; Denisova, L.; Follmer, T. T.; Pacchione, S. J.; Wang, Z.; Beare, C. M.; Bagdon, W. J.; Nichols, W. W. Plasmid DNA vaccines: Tissue distribution and effects of DNA sequence, adjuvants and delivery method on integration into host DNA. *Intervirology* **2000**, *43*, 273–281.

(49) Jiao, S.; Williams, P.; Berg, R. K.; Hodgeman, B. A.; Liu, L.; Repetto, G.; Wolff, J. A. Direct gene transfer into nonhuman primate myofibers in vivo. *Human Gene Therapy* **1992**, *3*, 21–33.

(50) Vandermeulen, G.; Marie, C.; Scherman, D.; Préat, V. New generation of plasmid backbones devoid of antibiotic resistance marker for gene therapy trials. *Molecular Therapy* **2011**, *19*, 1942–1949.

(51) Hardee, C.; Arévalo-Soliz, L.; Hornstein, B.; Zechiedrich, L. Advances in non-viral DNA vectors for gene therapy. *Genes* **2017**, *8*, 65.

(52) Ledwith, B. J.; Manam, S.; Troilo, P. J.; Barnum, A. B.; Pauley, C. J.; Griffiths, T. G.; Harper, L. B.; Schock, H. B.; Zhang, H.; Faris, J. E.; Way, P. A.; Beare, C. M.; Bagdon, W. J.; Nichols, W. W. Plasmid DNA vaccines: Assay for integration into host genomic DNA. *Developments in Biologicals* **2000**, *104*, 33–43.

(53) Cole, J.; Skopek, T. R. Working paper no. 3 Somatic mutant frequency, mutation rates and mutational spectra in the human population in vivo. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **1994**, *304*, 33–105.

(54) Yamamoto, S.; Yamamoto, T.; Tokunaga, T. The discovery of immunostimulatory DNA sequence. *Springer Seminars in Immunopathology* **2000**, *22*, 11–19.

(55) Häcker, G.; Redecke, V.; Häcker, H. Activation of the immune system by bacterial CpG-DNA. *Immunology* **2002**, *105*, 245–251.

(56) Krieg, A. M. CpG Motifs in Bacterial DNA and Their Immune Effects. *Annual Review of Immunology* **2002**, *20*, 709–760.

(57) Krieg, A. M.; Wu, T.; Weeratna, R.; Efler, S. M.; Love-Homan, L.; Yang, L.; Yi, A.-K.; Short, D.; Davis, H. L. Sequence motifs in adenoviral DNA block immune activation by stimulatory CpG motifs. *Proceedings of the National Academy of Sciences of the United States of America* **1998**, *95*, 12631–12636.

(58) Chen, Z.-Y.; Yant, S. R.; He, C.-Y.; Meuse, L.; Shen, S.; Kay, M. A. Linear DNAs concatemerize in vivo and result in sustained transgene expression in mouse liver. *Molecular Therapy* **2001**, *3*, 403–410.

(59) Bird, A. P. DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Research* **1980**, *8*, 1499–1504.

(60) Kawasaki, T.; Kawai, T. In *International Review of Cell and Molecular Biology*, Vanpouille-Box, C., Galluzzi, L., Eds.; Nucleic Acid Sensing and Immunity, Part A, Vol. 344; Academic Press: 2019, pp 1–30.

(61) Kumagai, Y.; Takeuchi, O.; Akira, S. TLR9 as a key receptor for the recognition of DNA. *Advanced Drug Delivery Reviews* **2008**, *60*, 795–804.

(62) Akira, S.; Uematsu, S.; Takeuchi, O. Pathogen recognition and innate immunity. *Cell* **2006**, *124*, 783–801.

(63) Bell, J. K.; Mullen, G. E. D.; Leifer, C. A.; Mazzoni, A.; Davies, D. R.; Segal, D. M. Leucine-rich repeats and pathogen recognition in Toll-like receptors. *Trends in Immunology* **2003**, *24*, 528–533.

(64) Bowie, A.; O'Neill, L. A. J. The interleukin-1 receptor/Toll-like receptor superfamily: signal generators for pro-inflammatory interleukins and microbial products. *Journal of Leukocyte Biology* **2000**, *67*, 508–514.

(65) Bauer, S.; Kirschning, C. J.; Häcker, H.; Redecke, V.; Hausmann, S.; Akira, S.; Wagner, H.; Lipford, G. B. Human TLR9 confers responsiveness to bacterial DNA via species-specific CpG motif recognition. *Proceedings of the National Academy of Sciences of the United States of America* **2001**, *98*, 9237–9242.

(66) Hemmi, H.; Takeuchi, O.; Kawai, T.; Kaisho, T.; Sato, S.; Sanjo, H.; Matsumoto, M.; Hoshino, K.; Wagner, H.; Takeda, K.; Akira, S. A Toll-like receptor recognizes bacterial DNA. *Nature* **2000**, *408*, 740–745.

(67) Chu, R. S.; Targoni, O. S.; Krieg, A. M.; Lehmann, P. V.; Harding, C. V. CpG oligodeoxynucleotides act as adjuvants that switch on T Helper 1 (Th1) immunity. *Journal of Experimental Medicine* **1997**, *186*, 1623–1631.

(68) Carson, D. A.; Raz, E. Oligonucleotide adjuvants for T Helper 1 (Th1)–specific vaccination. *Journal of Experimental Medicine* **1997**, *186*, 1621–1622.

(69) Latz, E.; Schoenemeyer, A.; Visintin, A.; Fitzgerald, K. A.; Monks, B. G.; Knetter, C. F.; Lien, E.; Nilsen, N. J.; Espevik, T.; Golenbock, D. T. TLR9 signals after translocating from the ER to CpG DNA in the lysosome. *Nature Immunology* **2004**, *5*, 190–198.

(70) Akira, S.; Takeda, K. Toll-like receptor signalling. *Nature Reviews Immunology* **2004**, *4*, 499–511.

(71) Häcker, H.; Vabulas, R. M.; Takeuchi, O.; Hoshino, K.; Akira, S.; Wagner, H. IMmune cell activation by bacterial CpG-DNA through myeloid differentiation marker 88 and tumor necrosis factor receptor–associated factor (Traf)6. *Journal of Experimental Medicine* **2000**, *192*, 595–600.

(72) Takeda, K.; Kaisho, T.; Akira, S. Toll-like receptors. *Annual Review of Immunology* **2003**, *21*, 335–376.

(73) Deng, L.; Wang, C.; Spencer, E.; Yang, L.; Braun, A.; You, J.; Slaughter, C.; Pickart, C.; Chen, Z. J. Activation of the IκB kinase complex by TRAF6 requires a dimeric ubiquitin-conjugating enzyme complex and a unique polyubiquitin chain. *Cell* **2000**, *103*, 351–361.

(74) Wang, C.; Deng, L.; Hong, M.; Akkaraju, G. R.; Inoue, J.-i.; Chen, Z. J. TAK1 is a ubiquitin-dependent kinase of MKK and IKK. *Nature* **2001**, *412*, 346–351.

(75) Li, S.; Strelow, A.; Fontana, E. J.; Wesche, H. IRAK-4: A novel member of the IRAK family with the properties of an IRAK-kinase. *Proceedings of the National Academy of Sciences of the United States of America* **2002**, *99*, 5567–5572.

(76) Takaoka, A.; Yanai, H.; Kondo, S.; Duncan, G.; Negishi, H.; Mizutani, T.; Kano, S.-i.; Honda, K.; Ohba, Y.; Mak, T. W.; Taniguchi, T. Integral role of IRF-5 in the gene induction programme activated by Toll-like receptors. *Nature* **2005**, *434*, 243–249.

(77) Sen, R.; Smale, S. T. Selectivity of the NF-κB Response. *Cold Spring Harbor Perspectives in Biology* **2010**, *2*, a000257.

(78) Krieg, A. M. The role of CpG motifs in innate immunity. *Current Opinion in Immunology* **2000**, *12*, 35–43.

(79) Krieg, A. M.; Yi, A. K.; Matson, S.; Waldschmidt, T. J.; Bishop, G. A.; Teasdale, R.; Koretzky, G. A.; Klinman, D. M. CpG motifs in bacterial DNA trigger direct B-cell activation. *Nature* **1995**, *374*, 546–549.

(80) Klinman, D. M.; Ylt, A.-k.; Beaucaget, S. L.; Conover, J.; Kriegt, A. M. CpG motifs present in bacteria DNA rapidly induce lymphocytes to secrete interleukin 6, interleukin 12, and interferon gamma. *Proceedings of the National Academy of Sciences of the United States of America* **1996**, *93*, 2879–83.

(81) Hartmann, G.; Weiner, G. J.; Krieg, A. M. CpG DNA: A potent signal for growth, activation, and maturation of human dendritic cells. *Proceedings of the National Academy of Sciences of the United States of America* **1999**, *96*, 9305–9310.

(82) Uematsu, S.; Sato, S.; Yamamoto, M.; Hirotani, T.; Kato, H.; Takeshita, F.; Matsuda, M.; Coban, C.; Ishii, K. J.; Kawai, T.; Takeuchi, O.; Akira, S. Interleukin-1 receptor-associated kinase-1 plays an essential role for Toll-like receptor (TLR)7- and TLR9-mediated interferon-α induction. *Journal of Experimental Medicine* **2005**, *201*, 915–923.

(83) Hodges, B. L.; Taylor, K. M.; Joseph, M. F.; Bourgeois, S. A.; Scheule, R. K. Long-term transgene expression from plasmid DNA gene therapy vectors is negatively affected by CpG dinucleotides. *Molecular Therapy* **2004**, *10*, 269–278.

(84) Chen, Z.-Y.; Riu, E.; He, C.-Y.; Xu, H.; Kay, M. A. Silencing of episomal transgene expression in liver by plasmid bacterial backbone DNA is independent of CpG methylation. *Molecular Therapy* **2008**, *16*, 548–556.

(85) Chen, Y.; Lenert, P.; Weeratna, R.; McCluskie, M.; Wu, T.; Davis, H. L.; Krieg, A. M. Identification of methylated CpG motifs as inhibitors of the immune stimulatory CpG motifs. *Gene Therapy* **2001**, *8*, 1024–1032.

(86) Hyde, S. C. et al. CpG-free plasmids confer reduced inflammation and sustained pulmonary gene expression. *Nature Biotechnology* **2008**, *26*, 549–551.

(87) Lipford, G. B.; Bauer, M.; Blank, C.; Reiter, R.; Wagner, H.; Heeg, K. CpG-containing synthetic oligonucleotides promote B and cytotoxic T cell responses to protein antigen: A new class of vaccine adjuvants. *European Journal of Immunology* **1997**, *27*, 2340–2344.

(88) Yamamoto, S. et al. Discovery of immunostimulatory CpG-DNA and its application to tuberculosis vaccine development. *Japanese Journal of Infectious Diseases* **2002**, *55*, 37–44.

(89) Grund, M.; Schleef, M. In *Minicircle and Miniplasmid DNA Vectors*, Schleef, M., Ed.; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2013, pp 1–6.

(90) Gaspar, V.; de Melo-Diogo, D.; Costa, E.; Moreira, A.; Queiroz, J.; Pichon, C.; Correia, I.; Sousa, F. Minicircle DNA vectors for gene therapy: advances and applications. *Expert Opinion on Biological Therapy* **2015**, *15*, 353–379.

(91) Almeida, A. M.; Eusébio, D.; Queiroz, J. A.; Sousa, F.; Sousa, A. The use of size-exclusion chromatography in the isolation of supercoiled minicircle DNA from Escherichia coli lysate. *Journal of Chromatography A* **2020**, *1609*, 460444.

(92) Darquet, A. M.; Cameron, B.; Wils, P.; Scherman, D.; Crouzet, J. A new DNA vehicle for nonviral gene delivery: supercoiled minicircle. *Gene Therapy* **1997**, *4*, 1341–1349.

(93) Darquet, A. M.; Rangara, R.; Kreiss, P.; Schwartz, B.; Naimi, S.; Delaère, P.; Crouzet, J.; Scherman, D. Minicircle: an improved DNA molecule for in vitro and in vivo gene transfer. *Gene Therapy* **1999**, *6*, 209–218.

(94) Bigger, B. W.; Tolmachov, O. E.; Collombet, J. M.; Fragkos, M.; Palaszewski, I.; Coutelle, C. An araC-controlled bacterial cre expression system to produce DNA minicircle vectors for nuclear and mitochondrial gene therapy. *Journal of Biological Chemistry* **2001**, *276*, 23018–23027.

(95) Jechlinger, W.; Tabrizi, C. A.; Lubitz, W.; Mayrhofer, P. Minicircle DNA immobilized in bacterial ghosts: In vivo production of safe non-viral DNA delivery vehicles. *Microbial Physiology* **2004**, *8*, 222–231.

(96) Simcikova, M.; Prather, K. L. J.; Prazeres, D. M. F.; Monteiro, G. A. On the dual effect of glucose during production of pBAD/AraC-based minicircles. *Vaccine* **2014**, *32*, 2843–2846.

(97) Chen, Z.-Y. Minicircle DNA vectors devoid of bacterial DNA result in persistent and high-level transgene expression in vivo. *Molecular Therapy* **2003**, *8*, 495–500.

(98) Chen, Z.-Y.; He, C.-Y.; Kay, M. A. Improved production and purification of minicircle DNA vector free of plasmid bacterial sequences and capable of persistent transgene expression in vivo. *Human Gene Therapy* **2005**, *16*, 126–131.

(99) Nafissi, N.; Slavcev, R. Construction and characterization of an in-vivo linear covalently closed DNA vector production system. *Microbial Cell Factories* **2012**, *11*, 154.

(100) Nafissi, N.; Alqawlaq, S.; Lee, E. A.; Foldvari, M.; Spagnuolo, P. A.; Slavcev, R. DNA ministrings: Highly safe and effective gene delivery vectors. *Molecular Therapy Nucleic Acids* **2014**, *3*, e165.

(101) Li, L.; Dimitriadis, E. K.; Yang, Y.; Li, J.; Yuan, Z.; Qiao, C.; Beley, C.; Smith, R. H.; Garcia, L.; Kotin, R. M. Production and characterization of novel recombinant adeno-associated virus replicative-form genomes: A eukaryotic source of dna for gene transfer. *PLOS ONE* **2013**, *8*, e69879.

(102) Schakowski, F.; Gorschlüter, M.; Junghans, C.; Schroff, M.; Buttgereit, P.; Ziske, C.; Schöttker, B.; König-Merediz, S. A.; Sauerbruch, T.; Wittig, B.; Schmidt-Wolf, I. G. H. A novel minimal-size vector (MIDGE) improves transgene expression in colon carcinoma cells and avoids transfection of undesired DNA. *Molecular Therapy* **2001**, *3*, 793–800.

(103) Wang, H.-S.; Chen, Z.-J.; Zhang, G.; Ou, X.-L.; Yang, X.-L.; Wong, C. K. C.; Giesy, J. P.; Du, J.; Chen, S.-Y. A novel micro-linear vector for in vitro and in vivo gene delivery and its application for EBV positive tumors. *PLOS ONE* **2012**, *7*, e47159.

(104) Walters, A. A.; Kinnear, E.; Shattock, R. J.; McDonald, J. U.; Caproni, L. J.; Porter, N.; Tregoning, J. S. Comparative analysis of enzymatically produced novel linear DNA constructs with plasmids for use as DNA vaccines. *Gene Therapy* **2014**, *21*, 645–652.

(105) Almeida, A. M.; Queiroz, J. A.; Sousa, F.; Sousa, Â. Minicircle DNA: THe future for DNA-based vectors? *Trends in Biotechnology* **2020**, *38*, 1047–1051.

(106) Kay, M. A.; He, C.-Y.; Chen, Z.-Y. A robust system for production of minicircle DNA vectors. *Nature Biotechnology* **2010**, *28*, 1287–1289.

(107) Sum, C. H.; Chong, J. Y.; Wettig, S.; Slavcev, R. Separation and purification of linear covalently closed deoxyribonucleic acid by Q-anion exchange membrane chromatography. *Journal of Chromatography A* **2014**, *1339*, 214–218.

(108) Diamantino, T.; Pereira, P.; Queiroz, J. A.; Sousa, Â.; Sousa, F. Minicircle DNA purification using a CIM® DEAE-1 monolithic support. *Journal of Separation Science* **2016**, *39*, 3544–3549.

(109) Gracey Maniar, L. E.; Maniar, J. M.; Chen, Z.-Y.; Lu, J.; Fire, A. Z.; Kay, M. A. Minicircle DNA vectors achieve sustained expression reflected by active chromatin and transcriptional level. *Molecular Therapy* **2013**, *21*, 131–138.

(110) Arévalo-Soliz, L. M.; Hardee, C. L.; Fogg, J. M.; Corman, N. R.; Noorbakhsh, C.; Zechiedrich, L. Improving therapeutic potential of non-viral minimized DNA vectors. *Cell and Gene Therapy Insights* **2020**, *6*, 1489–1505.

(111) Catanese, D. J.; Fogg, J. M.; Schrock, D. E.; Gilbert, B. E.; Zechiedrich, L.; Ii, D. E. S. Supercoiled Minivector DNA resists shear forces associated with gene therapy delivery. *Gene Therapy* **2012**, *19*, 94–100.

(112) Seow, Y.; Wood, M. J. Biological gene delivery vehicles: beyond viral vectors. *Molecular Therapy* **2009**, *17*, 767–777.

(113) Alvarez-Erviti, L.; Seow, Y.; Yin, H.; Betts, C.; Lakhal, S.; Wood, M. J. A. Delivery of siRNA to the mouse brain by systemic injection of targeted exosomes. *Nature Biotechnology* **2011**, *29*, 341–345.

(114) Bakhshinejad, B.; Sadeghizadeh, M.; Karimi, M. Bacteriophages as vehicles for gene delivery into mammalian cells: prospects and problems. *Expert Opinion on Drug Delivery* **2014**, *11*, 110.

(115) Hoyles, L.; McCartney, A. L.; Neve, H.; Gibson, G. R.; Sanderson, J. D.; Heller, K. J.; van Sinderen, D. Characterization of virus-like particles associated with the human faecal and caecal microbiota. *Research in Microbiology* **2014**, *165*, 803–812.

(116) Castro-Mejía, J. L.; Muhammed, M. K.; Kot, W.; Neve, H.; Franz, C. M. A. P.; Hansen, L. H.; Vogensen, F. K.; Nielsen, D. S. Optimizing protocols for extraction of bacteriophages prior to metagenomic analyses of phage communities in the human gut. *Microbiome* **2015**, *3*, 64.

(117) Jepson, C. D.; March, J. B. Bacteriophage lambda is a highly stable DNA vaccine delivery vehicle. *Vaccine* **2004**, *22*, 2413–2419.

(118) Jończyk, E.; Kłak, M.; Międzybrodzki, R.; Górski, A. The influence of external factors on bacteriophages—review. *Folia Microbiologica* **2011**.

(119) Ackermann, H.-W.; Tremblay, D.; Moineau, S. Long-term bacteriophage preservation. *WFCC Newsletter* **2004**, *38*, 35–40.

(120) Olofsson, L.; Ankarloo, J.; Andersson, P.-O.; Nicholls, I. A. Filamentous bacteriophage stability in non-aqueous media. *Chemistry & Biology* **2001**, *8*, 661–671.

(121) Branston, S. D.; Stanley, E. C.; Ward, J. M.; Keshavarz-Moore, E. Determination of the survival of bacteriophage M13 from chemical and physical challenges to assist in its sustainable bioprocessing. *Biotechnology and Bioprocess Engineering* **2013**, *18*, 560–566.

(122) Russel, M.; Model, P. In *The Bacteriophages*, 2nd edition; Oxford University Press: 2006, pp 146–160.

(123) Zinder, N.; Valentine, R.; Roger, M.; Stoeckenius, W. f1, a rod-shaped male-specific bacteriophage that contains DNA. *Virology* **1963**, *20*, 638–640.

(124) Loeb, T. Isolation of a bacteriophage specific for the F+ and Hfr mating types of Escherichia coli K-12. *Science* **1960**, *131*, 932–933.

(125) Marvin, D. A.; Hoffmann-Berling, H. Physical and chemical properties of two new small bacteriophages. *Nature* **1963**, *197*, 517–518.

(126) Messing, J. In *Methods in Molecular Biology*, Griffin, H., Griffin, A., Eds.; Humana Press Inc: Totowa, NJ, 1993; Vol. 23; Chapter 2, pp 9–22.

(127) Messing, J. Cloning in M13 phage or how to use biology at its best. *Gene* **1991**, *100*, 3–12.

(128) Dente, L.; Cesareni, G.; Cortese, R. PEMBL: A new family of single stranded plasmids. *Nucleic Acids Research* **1983**, *11*, 1645–1655.

(129) Vieira, J.; Messing, J. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* **1982**, *19*, 259–268.

(130) Zinder, N. D.; Boeke, J. D. The filamentous phage (Ff) as vectors for recombinant DNA— a review. *Gene* **1982**, *19*, 1–10.

(131) Smith, G. P. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **1985**, *228*, 1315–1317.

(132) Marvin, D. Filamentous phage structure, infection and assembly. *Current Opinion in Structural Biology* **1998**, *8*, 150–158.

(133) Model, P.; Russel, M. In *The Bacteriophages*, Calendar, R., Ed., 2nd ed.; Oxford University Press: New York, 1988; Chapter 12, pp 146–160.

(134) Beck, E.; Zink, B. Nucleotide sequence and genome organisation of filamentous bacteriophages fl and fd. *Gene* **1981**, *16*, 35–58.

(135) Dotto, G. P.; Dotto, G. P.; Horiuchi, K.; Zinder, N. D. Initiation and termination of phage f1 plus-strand synthesis. *Proceedings of the National Academy of Sciences of the United States of America* **1982**, *79*, 7122–7126.

(136) Chen, W. P.; Cheng, C. M.; Wang, A. H.; Kuo, T. T. Single-stranded DNA binding protein from bacteriophage cf: Characterization, gene localization and protein-ssDNA complex. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* **1996**, *1309*, 147–155.

(137) *The Bacteriophages*, 2nd ed; Calendar, R., Ed.; Oxford University Press: Oxford; New York, 2006; 746 pp.

(138) Makowski, L. Terminating a macromolecular helix. Structural model for the minor proteins of bacteriophage M13. *Journal of Molecular Biology* **1992**, *228*, 885–892.

(139) Roth, T. A.; Weiss, G. A.; Eigenbrot, C.; Sidhu, S. S. A minimized M13 coat protein defines the requirements for assembly into the bacteriophage particle. *Journal of Molecular Biology* **2002**, *322*, 357–367.

(140) Chung, W.-J.; Lee, D.-Y.; Yoo, S. Y. Chemical modulation of M13 bacteriophage and its functional opportunities for nanomedicine. *International Journal of Nanomedicine* **2014**, *9*, 5825–5836.

(141) Hunter, G. J.; Rowitch, D. H.; Perham, R. N. Interactions between DNA and coat protein in the structure and assembly of filamentous bacteriophage fd. *Nature* **1987**, *327*, 252–254.

(142) Greenwood, J.; Hunter, G. J.; Perham, R. N. Regulation of filamentous bacteriophage length by modification of electrostatic interactions between coat protein and DNA. *Journal of Molecular Biology* **1991**, *217*, 223–227.

(143) Rowitch, D. H.; Hunter, G. J.; Perham, R. N. Variable electrostatic interaction between DNA and coat protein in filamentous bacteriophage assembly. *Journal of Molecular Biology* **1988**, *204*, 663–674.

(144) Simons, G.; Konings, R. N.; Schoenmakers, J. G. Genes VI, VII, and IX of phage M13 code for minor capsid proteins of the virion. *Proceedings of the National Academy of Sciences of the United States of America* **1981**, *78*, 4194–4198.

(145) Simons, G.; Konings, R. N.; Schoenmakers, J. G. Identification of two new capsid proteins in bacteriophage M13. *FEBS Letters* **1979**, *106*, 8–12.

(146) Lubkowski, J.; Hennecke, F.; Plückthun, A.; Wlodawer, A. Filamentous phage infection: crystal structure of g3p in complex with its coreceptor, the C-terminal domain of TolA. *Structure* **1999**, *7*, 711–722.

(147) Karlsson, F.; Borrebaeck, C. A. K.; Nilsson, N.; Malmborg-Hager, A.-C. The mechanism of bacterial infection by filamentous phages involves molecular interactions between TolA and phage protein 3 domains. *Journal of Bacteriology* **2003**, *185*, 2628–2634.

(148) Marvin, D.; Hale, R. D.; Nave, C.; Citterich, M. H. Molecular models and structural comparisons of native and mutant class I filamentous bacteriophages Ff (fd, f1, M13), If1 and IKe. *Journal of Molecular Biology* **1994**, *235*, 260–286.

(149) Russel, M.; Linderoth, N. A.; Sali, A. Filamentous phage assembly: variation on a protein export theme. *Gene* **1997**, *192*, 23–32.

(150) Specthrie, L.; Bullitt, E.; Horiuchi, K.; Model, P.; Russel, M.; Makowski, L. Construction of a microphage variant of filamentous bacteriophage. *Journal of Molecular Biology* **1992**, *228*, 720–724.

(151) Sattar, S.; Bennett, N. J.; Wen, W. X.; Guthrie, J. M.; Blackwell, L. F.; Conway, J. F.; Rakonjac, J. Ff-nano, short functionalized nanorods derived from Ff (f1, fd, or M13) filamentous bacteriophage. *Frontiers in Microbiology* **2015**, *6*, 316.

(152) Rakonjac, J.; Model, P. Roles of pIII in filamentous phage assembly. *Journal of Molecular Biology* **1998**, *282*, 25–41.

(153) Silverman, P. M. Towards a structural biology of bacterial conjugation. *Molecular Microbiology* **1997**, *23*, 423–429.

(154) Jacobson, A. B. Role of F pili in the penetration of bacteriophage f1. *Journal of Virology* **1972**, *10*, 835–843.

(155) Deng, L.-W.; Malik, P.; Perham, R. N. Interaction of the globular domains of pIII protein of filamentous bacteriophgae fd with the F-pilus of Escherichia coli. *Virology* **1999**, *277*, 271–277.

(156) Deng, L.-W.; Perham, R. N. Delineating the site of interaction on the pIII protein of filamentous bacteriophage fd with the F-pilus of Escherichia coli. *Journal of Molecular Biology* **2002**, *319*, 603–614.

(157) Eckert, B.; Schmid, F. X. A conformational unfolding reaction activates phage fd for the infection of Escherichia coli. *Journal of Molecular Biology* **2007**, *373*, 452–461.

(158) Lorenz, S. H.; Schmid, F. X. Reprogramming the infection mechanism of a filamentous phage. *Molecular Microbiology* **2011**, *80*, 827–834.

(159) Russel, M.; Whirlow, H.; Sun, T.-P.; Webster, R. E. Low-frequency infection of F- bacteria by transducing particles of filamentous bacteriophages. *Journal of Bacteriology* **1988**, *170*, 5312–5316.

(160) Sun, T.-p.; Webster, R. E. Fii, a bacterial locus required for filamentous phage infection and its relation to colicin-tolerant tolA and tolB. *Journal of Bacteriology* **1986**.

(161) Sun, T.-p.; Webster, R. E. Nucleotide sequence of a gene cluster involved in entry of E colicins and single-stranded DNA of infecting filamentous bacteriophages into Escherichia coli. *Journal of Bacteriology* **1987**.

(162) Derouiche, R.; Bénédetti, H.; Lazzaroni, J.-C.; Lazdunski, C.; Lloubès, R. Protein complex within Escherichia coli inner membrane: TolA N-terminal domain interacts with TolQ and TolR proteins. *Journal of Biological Chemistry* **1995**, *270*, 11078–11084.

(163) Lloubès, R.; Cascales, E.; Walburger, A.; Bouveret, E.; Lazdunski, C.; Bernadac, A.; Journet, L. The Tol-Pal proteins of the Escherichia coli cell envelope: an energized system required for outer membrane integrity? *Research in Microbiology* **2001**, *152*, 523–529.

(164) Webster, R. E. The tol gene products and the import of macromolecules into Escherichia coli. *Molecular Microbiology* **1991**, *5*, 1005–1011.

(165) Levengood-Freyermuth, S. K.; Click, E. M.; Webster, R. E. Role of the carboxyl-terminal domain of TolA in protein import and integrity of the outer membrane. *Journal of Bacteriology* **1993**, *175*, 222–228.

(166) Cascales, E.; Buchanan, S. K.; Duché, D.; Kleanthous, C.; Lloubès, R.; Postle, K.; Riley, M.; Slatin, S.; Cavard, D. Colicin biology. *Microbiology and Molecular Biology Reviews* **2007**, *71*, 158–229.

(167) Guihard, G.; Boulanger, P.; Bénédetti, H.; Lloubés, R.; Besnard, M.; Letellier, L. Colicin A and the Tol proteins involved in its translocation are preferentially located in the contact sites between the inner and outer membranes of Escherichia coli cells. *Journal of Biological Chemistry* **1994**, *269*, 5874–5880.

(168) Click, E. M.; Webster, R. E. The TolQRA proteins are required for membrane insertion of the major capsid protein of the filamentous phage f1 during infection. *Journal of Bacteriology* **1998**, *180*, 1723–1728.

(169) Davis, N. G.; Boeke, J. D.; Model, P. Fine structure of a membrane anchor domain. *Journal of Molecular Biology* **1985**, *181*, 111–121.

(170) Riechmann, L.; Holliger, P. The C-terminal domain of TolA is the coreceptor for filamentous phage infection of e. coli. *Cell* **1997**, *90*, 351–360.

(171) Nakamura, M.; Tsumoto, K.; Kumagai, I.; Ishimura, K. A morphologic study of filamentous phage infection of Escherichia coli using biotinylated phages. *FEBS Letters* **2003**, *536*, 167–172.

(172) Duché, D.; Houot, L. Similarities and differences between colicin and filamentous phage uptake by bacterial cells. *EcoSal Plus* **2019**, *8*.

(173) Cascales, E.; Gavioli, M.; Sturgis, J. N.; Lloubes, R. Proton motive force drives the interaction of the inner membrane TolA and outer membrane Pal proteins in Escherichia coli. *Molecular Microbiology* **2000**, *38*, 904–915.

(174) Cascales, E.; Lloubès, R.; Sturgis, J. N. The TolQ-TolR proteins energize TolA and share homologies with the flagellar motor proteins MotA-MotB. *Molecular Microbiology* **2001**, *42*, 795–807.

(175) Germon, P.; Ray, M. C.; Vianney, A.; Lazzaroni, J. C. Energy-dependent conformational change in the TolA protein of Escherichia coli involves its N-terminal domain, TolQ, and TolR. *Journal of Bacteriology* **2001**, *183*, 4110–4114.

(176) Tzagoloff, H.; Pratt, D. The initial steps in infection with coliphage M13. *Virology* **1964**, *24*, 372–380.

(177) Geider, K.; Kornberg, A. Conversion of the M13 viral single strand to the double-stranded replicative forms by purified proteins. *Journal of Biological Chemistry* **1974**, *249*, 3999–4005.

(178) Dotto, G. P.; Horiuchi, K.; Zinder, N. D. The functional origin of bacteriophage f1 DNA replication. Its signals and domains. *Journal of Molecular Biology* **1984**, *172*, 507–521.

(179) Horiuchi, K. Initiation mechanisms in replication of filamentous phage DNA. *Genes to Cells* **1997**, *2*, 425–432.

(180) Short, J. M.; Sorge, J. A. In *Recombinant DNA Methodology II*, Wu, R., Ed.; Selected Methods in Enzymology; Academic Press: Boston, 1995, pp 185–198.

(181) Higashitani, N.; Higashitani, A.; Horiuchi, K. Nucleotide sequence of the primer RNA for DNA replication of filamentous bacteriophages. *Journal of Virology* **1993**, *67*, 2175–2181.

(182) Meyer, T. F.; Geider, K. Bacteriophage fd gene II-protein. I. Purification, involvement in RF replication, and the expression of gene II. *Journal of Biological Chemistry* **1979**, *254*, 12636–12641.

(183) Edens, L.; van Wezenbeek, P. M.; Konings, R. N.; Schoenmakers, J. G. Mapping of promoter sites on the genome of bacteriophage M13. *European Journal of Biochemistry* **1976**, *70*, 577–587.

(184) Edens, L.; Konings, R. N.; Schoenmakers, J. G. Transcription of bacteriophage M13 DNA: existence of promoters directly preceding genes III, VI, and I. *Journal of Virology* **1978**, *28*, 835–842.

(185) Edens, L.; Konings, R. N.; Schoenmakers, J. G. Physical mapping of the central terminator for transcription on the bacteriophage M13 genome. *Nucleic Acids Research* **1975**, *2*, 1811–1820.

(186) La Farina, M. L.; Model, P. Transcription in bacteriophage f1-infected Escherichia coli: Messenger populations in the infected cell. *Journal of Molecular Biology* **1983**, *164*, 377–393.

(187) Smits, M. A.; Jansen, J.; Konings, R. N.; Schoenmakers, J. G. Initiation and termination signals for transcription in bacteriophage M13. *Nucleic Acids Research* **1984**, *12*, 4071–4081.

(188) Smits, M. A.; Schoenmakers, J. G.; Konings, R. N. Expression of bacteriophage M13 DNA in vivo. *European Journal of Biochemistry* **1980**, *112*, 309–321.

(189) Goodrich, A. F.; Steege, D. A. Roles of polyadenylation and nucleolytic cleavage in the filamentous phage mRNA processing and decay pathways in Escherichia coli. *RNA* **1999**, *5*, 972–985.

(190) Smeal, S. W.; Schmitt, M. A.; Pereira, R. R.; Prasad, A.; Fisk, J. D. Simulation of the M13 life cycle II: Investigation of the control mechanisms of M13 infection and establishment of the carrier state. *Virology* **2017**, *500*, 275–284.

(191) Stump, M.; Steege, D. Functional analysis of filamentous phage f1 mRNA processing sites. *RNA* **1996**, *2*, 1286–1294.

(192) La Farina, M.; Model, P. Transcription in bacteriophage F1-Infected Escherichia coli I. Translation of the RNA in Vitro. *Virology* **1978**, *86*, 368–375.

(193)   Horabin, J. I.; Webster, R. E. Morphogenesis of f1 filamentous bacteriophage. Increased expression of gene I inhibits bacterial growth. *Journal of Molecular Biology* **1986**, *188*, 403–413.

(194)   Moses, P. B.; Model, P. A rho-dependent transcription termination signal in bacteriophage f1. *Journal of Molecular Biology* **1984**, *172*, 1–22.

(195)   Boeke, J. D.; Model, P.; Zinder, N. D. Effects of bacteriophage f1 gene III protein on the host cell membrane. *Molecular Genetics and Genomics* **1982**, *186*, 185–192.

(196)   Boeke, J. D.; Model, P. A prokaryotic membrane anchor sequence: carboxyl terminus of bacteriophage f1 gene III protein retains it in the membrane. *Proceedings of the National Academy of Sciences of the United States of America* **1982**, *79*, 5200–5204.

(197)   Wickner, W.; Ito, K.; Mandel, G.; Bates, M.; Nokelainen, M.; Zwizinski, C. The three lives of M13 coat protein: a virion capsid, an integral membrane protein, and a soluble cytoplasmic proprotein. *Annals of the New York Academy of Sciences* **1980**, *343*, 384–390.

(198)   Dunker, A. K.; Ensign, L. D.; Arnold, G. E.; Roberts, L. M. A model for fd phage penetration and assembly. *FEBS Letters* **1991**, *292*, 271–274.

(199)   Endemann, H.; Model, P. Location of filamentous phage minor coat proteins in phage and in infected cells. *Journal of Molecular Biology* **1995**, *250*, 496–506.

(200)   Rapoza, M. P.; Webster, R. E. The filamentous bacteriophage assembly proteins require the bacterial SecA protein for correct localization to the membrane. *Journal of Bacteriology* **1993**, *175*, 1856–1859.

(201)   Greenstein, D.; Horiuchi, K. Interaction between the replication origin and the initiator protein of the filamentous phage f1: Binding occurs in two steps. *Journal of Molecular Biology* **1987**, *197*, 157–174.

(202)   Greenstein, D.; Zinder, N. D.; Horiuchi, K. Integration host factor interacts with the DNA replication enhancer of filamentous phage f1. *Proceedings of the National Academy of Sciences of the United States of America* **1988**, *85*, 6262–6266.

(203)   Higashitani, A.; Greenstein, D.; Hirokawa, H.; Asano, S.; Horiuchi, K. Multiple DNA conformational changes induced by an initiator protein precede the nicking reaction in a rolling circle replication origin. *Journal of Molecular Biology* **1994**, *237*, 388–400.

(204)   Geider, K.; Bäumel, I.; Meyer, T. F. Intermediate stages in enzymatic replication of bacteriophage fd duplex DNA. *The Journal of Biological Chemistry* **1982**, *257*, 6488–6493.

(205)   Asano, S.; Higashitani, A.; Horiuchi, K. Filamentous phage replication initiator protein gpII forms a covalent complex with the 5' end of the nick it introduced. *Nucleic Acids Research* **1999**, *27*, 1882–1889.

(206)   Meyer, T. F.; Geider, K. Enzymatic synthesis of bacteriophage fd viral DNA. *Nature* **1982**, *296*, 828–832.

(207)   Horiuchi, K.; Horiuchi, K. Origin of DNA replication of bacteriophage f1 as the signal for termination. *Proceedings of the National Academy of Sciences of the United States of America* **1980**, *77*, 5226–5229.

(208)   Dotto, G. P.; Horiuchi, K. Replication of a plasmid containing two origins of bacteriophage f1. *Journal of Molecular Biology* **1981**, *153*, 169–176.

(209)   Pratt, D.; Erdahl, W. S. Genetic control of bacteriophage M13 DNA synthesis. *Journal of Molecular Biology* **1968**, *37*, 181–200.

(210)   Fulford, W.; Model, P. Regulation of bacteriophage f1 DNA replication: I. New functions for genes II and X. *Journal of Molecular Biology* **1988**, *203*, 49–62.

(211)   Model, P.; Mcgill, C.; Mazur, B. J.; Fulford, W. The replication of bacteriophage f1: Gene V protein regulates the synthesis of gene II protein. *Cell* **1982**, *29*, 329–335.

(212) Yen, T.; Webster, R. E. Translational control of bacteriophage f1 gene II and gene X proteins by gene V protein. *Cell* **1982**, *29*, 337–345.

(213) Fulford, W.; Model, P. Specificity of translational regulation by two DNA-binding proteins. *Journal of Molecular Biology* **1984**, *173*, 211–226.

(214) Mazur, B. J.; Model, P. Regulation of coliphage f1 single-stranded DNA synthesis by a DNA-binding protein. *Journal of Molecular Biology* **1973**, *78*, 285–300.

(215) Timmis, K.; Marvin, D. Filamentous bacterial viruses. XVI. Inherent temperature sensitivity of gene 5 protein and its involvement in abortive infection. *Virology* **1974**, *59*, 293–300.

(216) Grandis, A. S.; Webster, R. E. A new species of small covalently closed f1 DNA in Escherichia coli infected with an amber mutant of bacteriophage f1. *Virology* **1973**, *55*, 14–19.

(217) Cuypers, T.; Ouderaa, F. J. V. D.; de Jong, W. W. The amino acid sequence of gene 5 protein of bacteriophage M13. *Biochemical and Biophysical Research Communications* **1974**, *59*, 557–563.

(218) Pretorius, H. T.; Klein, M.; Day, L. A. Gene V protein of fd bacteriophage. Dimer formation and the role of tyrosyl groups in DNA binding. *Journal of Biological Chemistry* **1975**, *250*, 9262–9269.

(219) Pratt, D.; Laws, P.; Griffith, J. D. Complex of bacteriophage M13 single-stranded DNA and gene 5 protein. *Journal of Molecular Biology* **1974**, *82*, 425–439.

(220) Oey, J. L.; Knippers, R. Properties of the isolated gene 5 protein of bacteriophage fd. *Journal of Molecular Biology* **1972**, *68*, 125–138.

(221) Torbet, J.; Gray, D. M.; Gray, C. W.; Marvin, D.; Siegrist, H. Structure of the fd DNA–gene 5 protein complex in solution. A neutron small-angle scattering study. *Journal of Molecular Biology* **1981**, *146*, 305–320.

(222) Gray, C. W. Three-dimensional structure of complexes of single-stranded DNA-binding proteins with DNA. IKe and fd gene 5 proteins form left-handed helices with single-stranded DNA. *Journal of Molecular Biology* **1989**, *208*, 57–64.

(223) Bulsink, H.; Harmsen, B.; Hilbers, C. W. Specificity of the binding of bacteriophage M13 encoded gene-5 protein to DNA and RNA studied by means of fluorescence titrations. *Journal of Biomolecular Structure & Dynamics* **1985**, *3*, 227–247.

(224) Guan, Y.; Zhang, H.; Konings, R. N.; Hilbers, C. W.; Terwilliger, T. C.; Wang, A. H.-J. Crystal structures of Y41H and Y41F mutants of gene v protein from ff phage suggest possible protein-protein interactions in the GVP-ssDNA complex. *Biochemistry* **1994**, *33*, 7768–7778.

(225) Oliver, A. W.; Kneale, G. G. Structural characterization of DNA and RNA sequences recognized by the gene 5 protein of bacteriophage fd. *Biochemical Journal* **1999**, *339*, 525–531.

(226) Oliver, A. W.; Bogdarina, I.; Schroeder, E.; Taylor, I. A.; Kneale, G. G. Preferential binding of fd gene 5 protein to tetraplex nucleic acid structures. *Journal of Molecular Biology* **2000**, *301*, 575–584.

(227) Michel, B.; Michel, B.; Zinder, N. D.; Zinder, N. D. Translational repression in bacteriophage f1: characterization of the gene V protein target on the gene II mRNA. *Proceedings of the National Academy of Sciences of the United States of America* **1989**, *86*, 4002–4006.

(228) Wen, J.-D.; Gray, D. M. Ff gene 5 single-stranded DNA-binding protein assembles on nucleotides constrained by a DNA hairpin. *Biochemistry* **2004**, *43*, 2622–2634.

(229) Grant, R. A.; Webster, R. E. The bacteriophage f1 morphogenetic signal and the gene V protein/phage single-stranded DNA complex. *Virology* **1984**, *133*, 329–340.

(230) Webster, R. E.; Cashman, J. S. Abortive infection of Escherichia coli with the bacteriophage f1: Cytoplasmic membrane proteins and the f1 DNA-gene 5 protein complex. *Virology* **1973**, *55*, 20–38.

(231) Houbiers, M. C.; Hemminga, M. A. Protein-lipid interactions of bacteriophage M13 gene 9 minor coat protein. *Molecular Membrane Biology* **2004**, *21*, 351–359.

(232) Russel, M.; Kaźmierczak, B. Analysis of the structure and subcellular location of filamentous phage pIV. *Journal of Bacteriology* **1993**, *175*, 3998–4007.

(233) Marciano, D. K.; Russel, M.; Simon, S. M. An aqueous channel for filamentous phage export. *Science* **1999**, *284*, 1516–1519.

(234) Russel, M.; Model, P. Genetic analysis of the filamentous bacteriophage packaging signal and of the proteins that interact with it. *Journal of Virology* **1989**, *63*, 3284–3295.

(235) Dotto, G. P.; Zinder, N. D. The morphogenetic signal of bacteriophage f1. *Virology* **1983**, *130*, 252–256.

(236) Russel, M. Phage assembly: a paradigm for bacterial virulence factor export? *Science* **1994**, *265*, 612–614.

(237) Rapoza, M. P.; Webster, R. E. The products of gene I and the overlapping in-frame gene XI are required for filamentous phage assembly. *Journal of Molecular Biology* **1995**, *248*, 627–638.

(238) Haigh, N. G.; Webster, R. E. The pI and pXI assembly proteins serve separate and essential roles in filamentous phage assembly. *Journal of Molecular Biology* **1999**, *293*, 1017–1027.

(239) Feng, J.-n.; Model, P.; Russel, M. A trans-envelope protein complex needed for filamentous phage assembly and export. *Molecular Microbiology* **1999**, *34*, 745–755.

(240) Conners, R.; McLaren, M.; Łapińska, U.; Sanders, K.; Stone, M. R. L.; Blaskovich, M. A. T.; Pagliara, S.; Daum, B.; Rakonjac, J.; Gold, V. A. M. CryoEM structure of the outer membrane secretin channel pIV from the f1 filamentous bacteriophage. *Nature Communications* **2021**, *12*, 1–14.

(241) Lopez, J.; Webster, R. E. Assembly site of bacteriophage f1 corresponds to adhesion zones between the inner and outer membranes of the host cell. *Journal of Bacteriology* **1985**, *163*, 1270–1274.

(242) Russel, M.; Model, P. A bacterial gene, fip, required for filamentous bacteriophage fl assembly. *Journal of Bacteriology* **1983**, *154*, 1064–1076.

(243) Russel, M.; Model, P. Thioredoxin is required for filamentous phage assembly. *Proceedings of the National Academy of Sciences of the United States of America* **1985**, *82*, 29–33.

(244) Russel, M.; Model, P. The role of thioredoxin in filamentous phage assembly. Construction, isolation, and characterization of mutant thioredoxins. *Journal of Biological Chemistry* **1986**, *261*, 14997–15005.

(245) Marvin, D.; Symmons, M. F.; Straus, S. K. Structure and assembly of filamentous bacteriophages. *Progress in Biophysics & Molecular Biology* **2014**, *114*, 80–122.

(246) Huber, H. E.; Russel, M.; Model, P.; Richardson, C. C. Interaction of mutant thioredoxins of Escherichia coli with the gene 5 protein of phage T7. The redox capacity of thioredoxin is not required for stimulation of DNA polymerase activity. *Journal of Biological Chemistry* **1986**, *261*, 15006–15012.

(247) Feng, J.-n.; Russel, M.; Model, P. A permeabilized cell system that assembles filamentous bacteriophage. *Proceedings of the National Academy of Sciences of the United States of America* **1997**, *94*, 4068–4073.

(248) Lopez, J. A.; Webster, R. E. Morphogenesis of filamentous bacteriophage f1: Orientation of extrusion and production of polyphage. *Virology* **1983**, *127*, 177–193.

(249) Makowski, L. Phage display: structure, assembly and engineering of filamentous bacteriophage M13. *Current Opinion in Structural Biology* **1994**, *4*, 225–230.

(250) Rakonjac, J.; Feng, J.-n.; Model, P. Filamentous phage are released from the bacterial membrane by a two-step mechanism involving a short C-terminal fragment of pIII. *Journal of Molecular Biology* **1999**, *289*, 1253–1265.

(251) Russel, M. Filamentous phage assembly. *Molecular Microbiology* **1991**, *5*, 1607–1613.

(252) Rakonjac, J.; Jovanovic, G.; Model, P. Filamentous phage infection-mediated gene expression: construction and propagation of the gIII deletion mutant helper phage R408d3. *Gene* **1997**, 99–103.

(253) Horiuchi, K.; Vovis, G. F.; Model, P. The filamentous phage genome: Genes, physical structure, and protein products. *Cold Spring Harbor Monograph Archive* **1978**, 113–137.

(254) Henry, K. A.; Arbabi-Ghahroudi, M.; Scott, J. K. Beyond phage display: non-traditional applications of the filamentous bacteriophage as a vaccine carrier, therapeutic biologic, and bioconjugation scaffold. *Frontiers in Microbiology* **2015**, *6*, 755.

(255) Jin, H.-E.; Lee, S.-W. In *Virus-Derived Nanoparticles for Advanced Technologies*, Wege, C., Lomonossoff, G. P., Eds.; Methods in Molecular Biology, Vol. 1776; Springer New York: New York, NY, 2018, pp 487–502.

(256) Bernard, J. M. L.; Francis, M. B. Chemical strategies for the covalent modification of filamentous phage. *Frontiers in Microbiology* **2014**, *5*, 734.

(257) Smith, G. P.; Petrenko, V. A. Phage display. *Chemical Reviews* **1997**, *97*, 391–410.

(258) McCafferty, J.; Griffiths, A. D.; Winter, G.; Chiswell, D. J. Phage antibodies: filamentous phage displaying antibody variable domains. *Nature* **1990**, *348*, 552–554.

(259) Winter, G.; Griffiths, A. D.; Hawkins, R. E.; Hoogenboom, H. R. Making antibodies by phage display technology. *Annual Review of Immunology* **1994**, *12*, 433–455.

(260) Ledsgaard, L.; Kilstrup, M.; Karatt-Vellatt, A.; McCafferty, J.; Laustsen, A. H. Basics of antibody phage display technology. *Toxins* **2018**, *10*, 236.

(261) Crameri, R.; Jaussi, R.; Menz, G.; Blaser, K. Display of Expression Products of cDNA Libraries on Phage Surfaces. *European Journal of Biochemistry* **1994**, *226*, 53–58.

(262) Georgieva, Y.; Konthur, Z. Design and Screening of M13 Phage Display cDNA Libraries. *Molecules* **2011**, *16*, 1667–1681.

(263) Zhao, N.; Schmitt, M. A.; Fisk, J. D. Phage display selection of tight specific binding variants from a hyperthermostable Sso7d scaffold protein library. *The FEBS Journal* **2016**, *283*, 1351–1367.

(264) Sidhu, S. S. Engineering M13 for phage display. *Biomolecular Engineering* **2001**, *18*, 57–63.

(265) Smith, G. P. Preface: Surface display and peptide libraries. *Gene* **1993**, *128*, 1–2.

(266) Rondot, S.; Koch, J.; Breitling, F.; Dübel, S. A helper phage to improve single-chain antibody presentation in phage display. *Nature Biotechnology* **2001**, *19*, 75–78.

(267) Vieira, J.; Messing, J. In *Methods in Enzymology*; Elsevier: 1987; Vol. 153, pp 3–11.

(268) Scott, J. K.; Smith, G. P. Searching for peptide ligands with an epitope library. *Science* **1990**, *249*, 386–390.

(269) Kang, A. S.; Barbas, C. F.; Janda, K. D.; Benkovic, S. J.; Lerner, R. A. Linkage of recognition and replication functions by assembling combinatorial antibody Fab libraries along phage surfaces. *Proceedings of the National Academy of Sciences of the United States of America* **1991**, *88*, 4363–4366.

(270) Gustafson, H. H.; Olshefsky, A.; Sellers, D. L.; Sylvestre, M.; Pun, S. H. Current state of in vivo panning technologies: designing specificity and affinity into the future of drug targeting. *Advanced Drug Delivery Reviews* **2018**, *130*, 39–49.

(271) Tjhung, K. F.; Deiss, F.; Tran, J.; Chou, Y.; Derda, R. Intra-domain phage display (ID-PhD) of peptides and protein mini-domains censored from canonical pIII phage display. *Frontiers in Microbiology* **2015**, *6*, 340.

(272) Wickner, W. Asymmetric orientation of phage M13 coat protein in Escherichia coli cytoplasmic membranes and in synthetic lipid vesicles. *Proceedings of the National Academy of Sciences of the United States of America* **1976**, *73*, 1159–1163.

(273) Strandberg, L.; Enfors, S. O. Factors influencing inclusion body formation in the production of a fused protein in Escherichia coli. *Applied and Environmental Microbiology* **1991**, *57*, 1669–1674.

(274) Vispo, N. S.; Callejo, M.; Ojalvo, A. G.; Santos, A.; Chinea, G.; Gavilondo, J. V.; Araña, M. J. Displaying human interleukin-2 on the surface of bacteriophage. *Immunotechnology* **1997**, *3*, 185–193.

(275) Velappan, N.; Fisher, H. E.; Pesavento, E.; Chasteen, L.; D'Angelo, S.; Kiss, C.; Longmire, M.; Pavlik, P.; Bradbury, A. R. M. A comprehensive analysis of filamentous phage display vectors for cytoplasmic proteins: an analysis with different fluorescent proteins. *Nucleic Acids Research* **2010**, *38*, e22.

(276) Fuh, G.; Sidhu, S. S. Efficient phage display of polypeptides fused to the carboxy-terminus of the M13 gene-3 minor coat protein. *FEBS Letters* **2000**, *480*, 231–234.

(277) Kretzschmar, T.; Geiser, M. Evaluation of antibodies fused to minor coat protein III and major coat protein VIII of bacteriophage M13. *Gene* **1995**, *155*, 61–65.

(278) Kehoe, J. W.; Kay, B. K. Filamentous phage display in the new millennium. *Chemical Reviews* **2005**, *105*, 4056–4072.

(279) Light, J.; Lerner, R. A. Phophabs: Antibody-phage-alkaline phosphatase conjugates for one step ELISA's without immunization. *Bioorganic & Medicinal Chemistry Letters* **1992**, *2*, 1073–1078.

(280) Chen, L.; Zurita, A. J.; Ardelt, P. U.; Giordano, R. J.; Arap, W.; Pasqualini, R. Design and validation of a bifunctional ligand display system for receptor targeting. *Chemistry & Biology* **2004**, *11*, 1081–1091.

(281) Bonnycastle, L. L. C.; Brown, K. L.; Tang, J.; Scott, J. K. Assaying phage-borne peptides by phage capture on fibrinogen or streptavidin. *Biological Chemistry* **1997**, *378*, 509–516.

(282) Løset, G. Å.; Bogen, B.; Sandlie, I. Expanding the versatility of phage display I: Efficient display of peptide-tags on protein vii of the filamentous phage. *PLOS ONE* **2011**, *6*, e14702.

(283) Khalil, A. S.; Ferrer, J. M.; Brau, R. R.; Kottmann, S. T.; Noren, C. J.; Lang, M. J.; Belcher, A. M. Single M13 bacteriophage tethering and stretching. *Proceedings of the National Academy of Sciences of the United States of America* **2007**, *104*, 4892–4897.

(284) Alsteens, D.; Pesavento, E.; Cheuvart, G.; Dupres, V.; Trabelsi, H.; Soumillion, P.; Dufrêne, Y. F. Controlled manipulation of bacteriophages using single-virus force spectroscopy. *ACS Nano* **2009**, *3*, 3063–3068.

(285) Jespers, L. S.; Messens, J.; Keyser, A. D.; Eeckhout, D.; Ilse Van Den Brande; Gansemans, Y. G. J.; Lauwereys, M.; Vlasuk, G. P.; Stanssens, P. Surface expression and ligand-based selection of cDNAs fused to filamentous phage gene VI. *Nature Biotechnology* **1995**, *13*, 378–382.

(286) Jespers, L. S.; De Keyser, A.; Stanssens, P. E. λZLG6: A phage lambda vector for high-efficiency cloning and surface expression of cDNA libraries on filamentous phage. *Gene* **1996**, *173*, 179–181.

(287) Hufton, S. E.; Moerkerk, P. T.; Meulemans, E. V.; de Bruïne, A. P.; Arends, J.-W.; Hoogenboom, H. R. Phage display of cDNA repertoires: the pVI display system and its applications for the selection of immunogenic ligands. *Journal of Immunological Methods* **1999**, *231*, 39–51.

(288) Fransen, M.; Veldhoven, P. P. V.; Subramani, S. Identification of peroxisomal proteins by using M13 phage protein VI phage display: molecular evidence that mammalian peroxisomes contain a 2,4-dienoyl-CoA reductase. *Biochemical Journal* **1999**, *340*, 561–568.

(289) Hart, S. L.; Knight, A. M.; Harbottle, R. P.; Mistry, A.; Hunger, H. D.; Cutler, D. F.; Williamson, R.; Coutelle, C. Cell binding and internalization by filamentous phage displaying a cyclic Arg-Gly-Asp-containing peptide. *Journal of Biological Chemistry* **1994**, *269*, 12468–12474.

(290) Iannolo, G.; Minenkova, O.; Petruzzelli, R.; Cesareni, G. Modifying Filamentous Phage Capsid: Limits in the Size of the Major Capsid Protein. *Journal of Molecular Biology* **1995**, *248*, 835–844.

(291) Merzlyak, A.; Indrakanti, S.; Lee, S.-w. Genetically engineered nanofiber-like viruses for tissue regenerating materials. *Nano Letters* **2009**, *9*, 846–852.

(292) Held, H. A.; Sidhu, S. S. Comprehensive mutational analysis of the M13 major coat protein: improved scaffolds for C-terminal phage display. *Journal of Molecular Biology* **2004**, *340*, 587–597.

(293) Fuh, G.; Pisabarro, M. T.; Li, Y.; Quan, C.; Lasky, L. A.; Sidhu, S. S. Analysis of PDZ domain-ligand interactions using carboxyl-terminal phage display. *Journal of Biological Chemistry* **2000**, *275*, 21486–21491.

(294) Sidhu, S. S.; Weiss, G. A.; Wells, J. A. High copy display of large proteins on phage for functional selections. *Journal of Molecular Biology* **2000**, *296*, 487–495.

(295) Weiss, G. A.; Wells, J. A.; Sidhu, S. S. Mutational analysis of the major coat protein of M13 identifies residues that control protein display. *Protein Science* **2000**, *9*, 647–654.

(296) Sidhu, S. S.; Feld, B. K.; Weiss, G. A. In *Protein Engineering Protocols*, Arndt, K. M., Müller, K. M., Eds.; Methods in Molecular Biology™; Humana Press: Totowa, NJ, 2007, pp 205–219.

(297) Malik, P.; Perham, R. N. Simultaneous display of different peptides on the surface of filamentous bacteriophage. *Nucleic Acids Research* **1997**, *25*, 915–916.

(298) Høydahl, L. S.; Nilssen, N. R.; Gunnarsen, K. S.; du Pré, M. F.; Iversen, R.; Roos, N.; Chen, X.; Michaelsen, T. E.; Sollid, L. M.; Sandlie, I.; Løset, G. Å. Multivalent pIX phage display selects for distinct and improved antibody properties. *Scientific Reports* **2016**, *6*, 1–13.

(299) Løset, G. Å.; Sandlie, I. Next generation phage display by use of pVII and pIX as display scaffolds. *Methods* **2012**, *58*, 40–46.

(300) Kwaśnikowski, P.; Kristensen, P.; Markiewicz, W. T. Multivalent display system on filamentous bacteriophage pVII minor coat protein. *Journal of Immunological Methods* **2005**, *307*, 135–143.

(301) Tornetta, M.; Baker, S.; Whitaker, B.; Lu, J.; Chen, Q.; Pisors, E.; Shi, L.; Luo, J.; Sweet, R. W.; Tsui, P. Antibody Fab display and selection through fusion to the pIX coat protein of filamentous phage. *Journal of Immunological Methods* **2010**, *360*, 39–46.

(302) Shi, L.; Wheeler, J. C.; Sweet, R. W.; Lu, J.; Luo, J.; Tornetta, M.; Whitaker, B.; Reddy, R.; Brittingham, R.; Borozdina, L.; Chen, Q.; Amegadzie, B.; Knight, D. M.; Almagro, J. C.; Tsui, P. De novo selection of high-affinity antibodies from synthetic Fab libraries displayed on phage as pIX fusion proteins. *Journal of Molecular Biology* **2010**, *397*, 385–396.

(303) Gao, C.; Mao, S.; Lo, C.-H. L.; Wirsching, P.; Lerner, R. A.; Janda, K. D. Making artificial antibodies: A format for phage display of combinatorial heterodimeric arrays. *Proceedings of the National Academy of Sciences of the United States of America* **1999**, *96*, 6025–6030.

(304) Gao, C.; Mao, S.; Kaufmann, G. F.; Wirsching, P.; Lerner, R. A.; Janda, K. D. A method for the generation of combinatorial antibody libraries using pIX phage display. *Proceedings of the National Academy of Sciences of the United States of America* **2002**, *99*, 12612–12616.

(305) Bratkovič, T. Progress in phage display: evolution of the technique and its applications. *Cellular and Molecular Life Sciences* **2010**, *67*, 749–767.

(306) Messing, J. New M13 vectors for cloning. *Methods in Enzymology* **1983**, *101*, 20–78.

(307) Norrander, J.; Kempe, T.; Messing, J. Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. *Gene* **1983**, *26*, 101–106.

(308) Yanisch-Perron, C.; Vieira, J.; Messing, J. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **1985**, *33*, 103–119.

(309) Heidecker, G.; Messing, J.; Gronenborn, B. A versatile primer for DNA sequencing in the M13mp2 cloning system. *Gene* **1980**, *10*, 69–73.

(310) Anderson, S.; Gait, M. J.; Mayol, L.; Young, I. G. A short primer for sequencing DNA cloned in the single-stranded phage vector M13mp2. *Nucleic Acids Research* **1980**, *8*, 1731–1743.

(311) Sanger, F.; Coulson, A.; Barrell, B.; Barrell, B. G.; Smith, A.; Roe, B. A. Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. *Journal of Molecular Biology* **1980**, *143*, 161–178.

(312) Messing, J.; Crea, R.; Seeburg, P. H. A system for shotgun DNA sequencing. *Nucleic Acids Research* **1981**, *9*, 309.

(313) Greenstein, D.; Brent, R. Introduction to vectors derived from filamentous phages. *Current Protocols in Molecular Biology* **2001**, *Chapter 1*, Unit1.14.

(314) Cleary, J. M.; Ray, D. S. Replication of the plasmid pBR322 under the control of a cloned replication origin from the single-stranded DNA phage M13. *Proceedings of the National Academy of Sciences of the United States of America* **1980**, *77*, 4638–4642.

(315) Levinson, A.; Silver, D.; Seed, B. Minimal size plasmids containing an M13 origin for production of single-strand transducing particles. *Journal of Molecular and Applied Genetics* **1984**, *2*, 507–517.

(316) Enea, V.; Zinder, N. D. Interference resistant mutants of phage f1. *Virology* **1982**, *122*, 222–226.

(317) Kramer, R. A.; Cox, F.; van der Horst, M.; van den Oudenrijn, S.; Res, P. C. M.; Bia, J.; Logtenberg, T.; de Kruif, J. A novel helper phage that improves phage display selection efficiency by preventing the amplification of phages without recombinant protein. *Nucleic Acids Research* **2003**, *31*, e59.

(318) Chasteen, L.; Ayriss, J.; Pavlik, P.; Bradbury, A. R. M. Eliminating helper phage from phage display. *Nucleic Acids Research* **2006**, *34*, e145.

(319) Short, J. M.; Fernandez, J. M.; Sorge, J. A.; Huse, W. D. Lambda ZAP: a bacteriophage lambda expression vector with in vivo excision properties. *Nucleic Acids Research* **1988**, *16*, 7583–7600.

(320) Russell, S. J. Peptide–displaying phages for targeted gene delivery? *Nature Medicine* **1996**, *2*, 276–277.

(321) Pasqualini, R.; Ruoslahti, E. Organ targeting in vivo using phage display peptide libraries. *Nature* **1996**, *380*, 364–366.

(322) Barry, M.; Dower, W.; Johnston, S. A. Toward cell-tergetting gene therapy vectors: selection of cell-binding peptides from random peptide-presenting phage libraries. *Nature Medicine* **1996**, *2*, 299–305.

(323) Becerril, B.; Poul, M.-A.; Marks, J. D. Toward Selection of Internalizing Antibodies from Phage Libraries. *Biochemical and Biophysical Research Communications* **1999**, *255*, 386–393.

(324) Uppala, A.; Koivunen, E. Targeting of phage display vectors to mammalian cells. *Combinatorial Chemistry & High Throughput Screening* **2000**, *3*, 373–392.

(325) Yip, Y. L.; Hawkins, N. J.; Smith, G.; Ward, R. L. Biodistribution of filamentous phage-Fab in nude mice. *Journal of Immunological Methods* **1999**, *225*, 171–178.

(326) Molenaar, T. J. M.; Michon, I.; de Haas, S. A. M.; van Berkel, T. J. C.; Kuiper, J.; Biessen, E. A. L. Uptake and processing of modified bacteriophage M13 in mice: implications for phage display. *Virology* **2002**, *293*, 182–191.

(327) Frenkel, D.; Solomon, B. Filamentous phage as vector-mediated antibody delivery to the brain. *Proceedings of the National Academy of Sciences of the United States of America* **2002**, *99*, 5675–5679.

(328) Bichet, M. C.; Chin, W. H.; Richards, W.; Lin, Y.-W.; Avellaneda-Franco, L.; Hernandez, C. A.; Oddo, A.; Chernyavskiy, O.; Hilsenstein, V.; Neild, A.; Li, J.; Voelcker, N. H.; Patwa, R.; Barr, J. J. Bacteriophage uptake by mammalian cell layers represents a potential sink that may impact phage therapy. *iScience* **2021**, *24*, 102287.

(329) Miernikiewicz, P.; Dabrowska, K. Endocytosis of Bacteriophages. *Current Opinion in Virology* **2022**, *52*, 229–235.

(330) Nguyen, S.; Baker, K.; Padman, B. S.; Patwa, R.; Dunstan, R. A.; Weston, T. A.; Schlosser, K.; Bailey, B.; Lithgow, T.; Lazarou, M.; Luque, A.; Rohwer, F.; Blumberg, R. S.; Barr, J. J. Bacteriophage transcytosis provides a mechanism to cross epithelial cell layers. *mBio* **2017**, *8*, ed. by Racaniello, V. R., e01874–17.

(331) Otero, J.; García-Rodríguez, A.; Cano-Sarabia, M.; Maspoch, D.; Marcos, R.; Cortés, P.; Llagostera, M. Biodistribution of liposome-encapsulated bacteriophages and their transcytosis during oral phage therapy. *Frontiers in Microbiology* **2019**, *10*, 689.

(332) Larocca, D.; Witte, A.; Johnson, W.; Pierce, G. F.; Baird, A. Targeting bacteriophage to mammalian cell surface receptors for gene delivery. *Human Gene Therapy* **1998**, *9*, 2393–2399.

(333) Larocca, D.; Kassner, P. D.; Witte, A.; Ladner, R. C.; Pierce, G. F.; Baird, A. Gene transfer to mammalian cells using genetically targeted filamentous bacteriophage. *The FASEB Journal* **1999**, *13*, 727–734.

(334) Tian, Y.; Wu, M.; Liu, X.; Liu, Z.; Zhou, Q.; Niu, Z.; Huang, Y. Probing the endocytic pathways of the filamentous bacteriophage in live cells using ratiometric pH fluorescent indicator. *Advanced Healthcare Materials* **2015**, *4*, 413–419.

(335) Pasqualini, R.; Koivunen, E.; Ruoslahti, E. av Integrins as receptors for tutnor targeting by circulating ligands. **1997**, *15*, 5.

(336) Kim, A.; Shin, T. H.; Shin, S. M.; Pham, C. D.; Choi, D. K.; Kwon, M. H.; Kim, Y. S. Cellular internalization mechanism and intracellular trafficking of filamentous M13 phages displaying a cell-penetrating transbody and TAT peptide. *PLOS ONE* **2012**, *7*, e51813.

(337) Di Giovine, M.; Salone, B.; Martina, Y.; Amati, V.; Zambruno, G.; Cundari, E.; Failla, C. M.; Saggio, I. Binding properties, cell delivery, and gene transfer of adenoviral penton base displaying bacteriophage. *Virology* **2001**, *282*, 102–12.

(338) Monaci, P.; Urbanelli, L.; Fontana, L. Phage as gene delivery vectors. *Current Opinion in Molecular Therapeutics* **2001**, *3*, 159–169.

(339) Hajitou, A. et al. A hybrid vector for ligand-directed tumor targeting and molecular imaging. *Cell* **2006**, *125*, 385–398.

(340) Hajitou, A.; Rangel, R.; Trepel, M.; Soghomonyan, S.; Gelovani, J. G.; Alauddin, M. M.; Pasqualini, R.; Arap, W. Design and construction of targeted AAVP vectors for mammalian cell transduction. *Nature Protocols* **2007**, *2*, 523–531.

(341) Stoneham, C. A.; Hollinshead, M.; Hajitou, A. Clathrin-mediated endocytosis and subsequent endo-lysosomal trafficking of adeno-associated virus/phage. *Journal of Biological Chemistry* **2012**, *287*, 35849–35859.

(342) Abedheydari, E.; Khalaj-kondori, M.; Hosseinpour-Faizi, M.-A.; Kosari-nasab, M. A comparative investigation on efficiency of bacteriophage lambda and M13 based vectors for delivering and expression of transgene in eukaryote cells. *Journal of Cell and Molecular Research* **2014**, *6*, 69–75.

(343) Prieto, Y.; Sánchez, O. Self-complementary sequences induce the formation of double-stranded filamentous phages. *Biochimica et Biophysica Acta (BBA) - General Subjects* **2007**, *1770*, 1081–1084.

(344) Burg, M. A.; Jensen-pergakes, K.; Gonzalez, A. M.; Cells, C.; Ravey, P.; Baird, A.; Larocca, D. Enhanced phagemid particle gene transfer in camptothecin-treated carcinoma cells. *Cancer Research* **2002**, *62*, 977–981.

(345) Kassner, P. D.; Burg, M. A.; Baird, A.; Larocca, D. Genetic selection of phage engineered for receptor-mediated gene transfer to mammalian cells. *Biochemical and Biophysical Research Communications* **1999**, *264*, 921–928.

(346) Poul, M.-A.; Marks, J. D. Targeted gene delivery to mammalian cells by filamentous bacteriophage. *Journal of Molecular Biology* **1999**, *288*, 203–211.

(347) Larocca, D.; Jensen-Pergakes, K.; Burg, M. A.; Baird, A. Receptor-targeted gene delivery using multivalent phagemid particles. *Molecular Therapy* **2001**, *3*, 476–484.

(348) Urbanelli, L.; Ronchini, C.; Fontana, L.; Menard, S.; Orlandi, R.; Monaci, P. Targeted gene transduction of mammalian cells expressing the HER2/neu receptor by filamentous phage11Edited by J. Karn. *Journal of Molecular Biology* **2001**, *313*, 965–976.

(349) Mount, J. D.; Samoylova, T. I.; Morrison, N. E.; Cox, N. R.; Baker, H. J.; Petrenko, V. A. Cell targeted phagemid rescued by preselected landscape phage. *Gene* **2004**, *341*, 59–65.

(350) Baird, A. Gene transfer into mammalian cells using targeted filamentous bacteriophage. *Cold Spring Harbor Protocols* **2011**, *2011*, 950.

(351) Van Houten, N. E.; Zwick, M. B.; Menendez, A.; Scott, J. K. Filamentous phage as an immunogenic carrier to elicit focused antibody responses against synthetic peptide. *Molecular Biology* **2007**, *24*, 4188–4200.

(352) Olszowska-Zaremba, N.; Borysowski, J.; Dąbrowska, K.; Górski, A. In *Bacteriophages in health and disease*, Hyman, P., Abedon, S. T., Eds.; CABI: Wallingford, 2012, pp 168–184.

(353) Halpern, B. N. The role and function of the reticulo-endothelial system in immunological processes. *Journal of Pharmacy and Pharmacology* **1959**, *11*, 321–338.

(354) Nicastro, J.; Wong, S.; Slavcev, R. A. In *Bacteriophage Applications - Historical Perspective and Future Potential*; SpringerBriefs in Biochemistry and Molecular Biology; Springer International Publishing: Cham, 2016, pp 69–82.

(355) Srivastava, A. S.; Kaido, T.; Carrier, E. Immunological factors that affect the in vivo fate of T7 phage in the mouse. *Journal of Virological Methods* **2004**, *115*, 99–104.

(356) Merril, C. R.; Biswas, B.; Carlton, R.; Jensen, N. C.; Creed, G. J.; Zullo, S.; Adhya, S. Long-circulating bacteriophage as antibacterial agents. *Proceedings of the National Academy of Sciences of the United States of America* **1996**, *93*, 3188–3192.

(357) Capparelli, R.; Ventimiglia, I.; Roperto, S.; Fenizia, D.; Iannelli, D. Selection of an Escherichia coli O157:H7 bacteriophage for persistence in the circulatory system of mice infected experimentally. *Clinical Microbiology and Infection* **2006**, *12*, 248–253.

(358) Huh, H.; Wong, S.; St. Jean, J.; Slavcev, R. Bacteriophage interactions with mammalian tissue: Therapeutic applications. *Advanced Drug Delivery Reviews* **2019**, *145*, 4–17.

(359) Górski, A.; Dabrowska, K.; Międzybrodzki, R.; Weber-Dabrowska, B.; Łusiak-Szelachowska, M.; Jończyk-Matysiak, E.; Borysowski, J. Phages and immunomodulation. *Future Microbiology* **2017**, *12*, fmb-2017–0049.

(360) Gorski, A.; Dabrowska, K.; Switala-Jele, K.; Nowaczyk, M.; Weber-Dabrowska, B.; Boratynski, J.; Wietrzyk, J.; Opolski, A. New insights into the possible role of bacteriophages in host defense and disease. *Medical Immunology* **2003**, *2*, 2.

(361) Górski, A.; Weber-Dabrowska, B. The potential role of endogenous bacteriophages in controlling invading pathogens. *Cellular and Molecular Life Sciences* **2005**, *62*, 511–519.

(362) Dabrowska, K.; Switała-Jelen, K.; Opolski, A.; Weber-Dabrowska, B.; Gorski, A. Bacteriophage penetration in vertebrates. *Journal of Applied Microbiology* **2005**, *98*, 7–13.

(363) Górski, A.; Wazna, E.; Dąbrowska, B. W.; Dąbrowska, K.; Świtała-Jeleń, K.; Międzybrodzki, R. Bacteriophage translocation. *FEMS Immunology & Medical Microbiology* **2006**, *46*, 313–319.

(364) Hsu, B. B.; Gibson, T. E.; Yeliseyev, V.; Liu, Q.; Lyon, L.; Bry, L.; Silver, P. A.; Gerber, G. K. Dynamic modulation of the gut microbiota and metabolome by bacteriophages in a mouse model. *Cell Host & Microbe* **2019**, *25*, 803–814.e5.

(365) Wahida, A.; Tang, F.; Barr, J. J. Rethinking phage-bacteria-eukaryotic relationships and their influence on human health. *Cell Host & Microbe* **2021**, *29*, 681–688.

(366) Wong, S.; Lam, P.; Nafissi, N.; Denniss, S.; Slavcev, R. Production of double-stranded DNA ministrings. *Journal of Visualized Experiments* **2016**, 3–9.

(367) Barr, J. J.; Youle, M.; Rohwer, F. Innate and acquired bacteriophage-mediated immunity. *Bacteriophage* **2014**, *3*, e25857.

(368) Reyes, A.; Haynes, M.; Hanson, N.; Angly, F. E.; Heath, A. C.; Rohwer, F.; Gordon, J. I. Viruses in the faecal microbiota of monozygotic twins and their mothers. *Nature* **2010**, *466*, 334–338.

(369) Reyes, A.; Semenkovich, N. P.; Whiteson, K.; Rohwer, F.; Gordon, J. I. Going viral: next-generation sequencing applied to phage populations in the human gut. *Nature Reviews Microbiology* **2012**, *10*, 607–17.

(370) Barr, J. J.; Auro, R.; Furlan, M.; Whiteson, K. L.; Erb, M. L.; Pogliano, J.; Stotland, A.; Wolkowicz, R.; Cutting, A. S.; Doran, K. S.; Salamon, P.; Youle, M.; Rohwer, F. Bacteriophage adhering to mucus provide a non–host-derived immunity. *Proceedings of the National Academy of Sciences of the United States of America* **2013**, *110*, 10771–10776.

(371) Dabrowska, K.; Miernikiewicz, P.; Piotrowicz, A.; Hodyra, K.; Owczarek, B.; Lecion, D.; Kaźmierczak, Z.; Letarov, A.; Górski, A. Immunogenicity studies of proteins forming the T4 phage head surface. *Journal of Virology* **2014**, *88*, 12551–12557.

(372) Górski, A.; Międzybrodzki, R.; Weber-Dabrowska, B.; Fortuna, W.; Letkiewicz, S.; Rogóż, P.; Jończyk-Matysiak, E.; Dąbrowska, K.; Majewska, J.; Borysowski, J. Phage therapy: Combating infections with potential for evolving from merely a treatment for complications to targeting diseases. *Frontiers in Microbiology* **2016**, *7*.

(373) Van Belleghem, J. D.; Dabrowska, K.; Vaneechoutte, M.; Barr, J. J.; Bollyky, P. L. Interactions between bacteriophage, bacteria, and the mammalian immune system. *Viruses* **2019**, *11*, 10.

(374) Hoshino, K.; Takeuchi, O.; Kawai, T.; Sanjo, H.; Ogawa, T.; Takeda, Y.; Takeda, K.; Akira, S. Cutting Edge: Toll-like receptor 4 (TLR4)-deficient mice are hyporesponsive to lipopolysaccharide: evidence for TLR4 as the lps gene product. *The Journal of Immunology* **1999**, *162*, 3749–3752.

(375) Alexander, C.; Rietschel, E. T. Bacterial lipopolysaccharides and innate immunity. *Journal of Endotoxin Research* **2001**, *7*, 167–202.

(376) Eriksson, F.; Tsagozis, P.; Lundberg, K.; Parsa, R.; Mangsbo, S. M.; Persson, M. A. A.; Harris, R. A.; Pisa, P. Tumor-specific bacteriophages induce tumor destruction through activation of tumor-associated macrophages. *The Journal of Immunology* **2009**, *182*, 3105–3111.

(377) Miernikiewicz, P.; Klopot, A.; Soluch, R.; Szkuta, P.; Keska, W.; Hodyra-Stefaniak, K.; Konopka, A.; Nowak, M.; Lecion, D.; Kazmierczak, Z.; Majewska, J.; Harhala, M.; Górski, A.; Dabrowska, K. T4 phage tail adhesin gp12 counteracts LPS-induced inflammation in vivo. *Frontiers in Microbiology* **2016**, *7*, 1112.

(378) Zhang, L.; Hou, X.; Sun, L.; He, T.; Wei, R.; Pang, M.; Wang, R. Staphylococcus aureus bacteriophage suppresses LPS-Induced inflammation in MAC-T bovine mammary epithelial cells. *Frontiers in Microbiology* **2018**, *9*, 1614.

(379) Kneissel, S.; Queitsch, I.; Petersen, G.; Behrsing, O.; Micheel, B.; Dübel, S. Epitope structures recognised by antibodies against the major coat protein (g8p) of filamentous bacteriophage fd (Inoviridae). *Journal of Molecular Biology* **1999**, *288*, 21–28.

(380) Van Houten, N. E.; Henry, K. A.; Smith, G. P.; Scott, J. K. Engineering filamentous phage carriers to improve focusing of antibody responses against peptides. *Vaccine* **2010**, *28*, 2174–2185.

(381) Henry, K. A.; Murira, A.; van Houten, N. E.; Scott, J. K. Developing strategies to enhance and focus humoral immune responses using filamentous phage as a model antigen. *Bioengineered Bugs* **2011**, *2*, 275.

(382) Yang, Q.; Wang, L.; Lu, D.-N.; Gao, R.-J.; Song, J.-N.; Hua, P.-Y.; Yuan, D.-w. Prophylactic vaccination with phage-displayed epitope of C. albicans elicits protective immune responses against systemic candidiasis in C57BL/6 mice. *Vaccine* **2005**, *23*, 4088–4096.

(383) Wang, Y.; Su, Q.; Dong, S.; Shi, H.; Gao, X.; Wang, L. Hybrid phage displaying SLAQVKYTSASSI induces protection against Candida albicans challenge in BALB/c mice. *Human Vaccines & Immunotherapeutics* **2014**, *10*, 1057–1063.

(384) Gaubin, M.; Fanutti, C.; Mishal, Z.; Durrbach, A.; De Berardinis, P.; Sartorius, R.; Del Pozzo, G.; Guardiola, J.; Perham, R. N.; Piatier-Tonneau, D. Processing of filamentous bacteriophage virions in antigen-presenting cells targets both HLA class I and class II peptide loading compartments. *DNA and Cell Biology* **2003**, *22*, 11–18.

(385) Wan, Y.; Wu, Y.; Zhou, J.; Zou, L.; Liang, Y.; Zhao, J.; Jia, Z.; Engberg, J.; Bian, J.; Zhou, W. Cross-presentation of phage particle antigen in MHC class II and endoplasmic reticulum marker-positive compartments. *European Journal of Immunology* **2005**, *35*, 2041–2050.

(386) Gomes-Neto, J. F.; Sartorius, R.; Canto, F. B.; Almeida, T. S.; Dias, A. A.; Barbosa, C.-H. D.; Melo, G. A.; Oliveira, A. C.; Aguiar, P.-H. N.; Machado, C. R.; de Matos Guedes, H. L.; Santiago, M. F.; Nóbrega, A.; De Berardinis, P.; Bellio, M. Vaccination with recombinant filamentous fd phages against parasite infection requires TLR9 expression. *Frontiers in Immunology* **2018**, *9*, 1173.

(387) Sartorius, R.; D'Apice, L.; Trovato, M.; Cuccaro, F.; Costa, V.; De Leo, M. G.; Marzullo, V. M.; Biondo, C.; D'Auria, S.; De Matteis, M. A.; Ciccodicola, A.; De Berardinis, P. Antigen delivery by filamentous bacteriophage fd displaying an anti-DEC-205 single-chain variable fragment confers adjuvanticity by triggering a TLR9-mediated immune response. *EMBO Molecular Medicine* **2015**, *7*, 973–988.

(388) Hashiguchi, S.; Yamaguchi, Y.; Takeuchi, O.; Akira, S.; Sugimura, K. Immunological basis of M13 phage vaccine: Regulation under MyD88 and TLR9 signaling. *Biochemical and Biophysical Research Communications* **2010**, *402*, 19–22.

(389) Mori, K.; Kubo, T.; Kibayashi, Y.; Ohkuma, T.; Kaji, A. Anti-vaccinia virus effect of M13 bacteriophage DNA. *Antiviral Research* **1996**, *31*, 79–86.

(390) Greenwood, J.; Anne E. Willis; Richard N. Perham Multiple display of foreign peptides on a filamentous bacteriophage: Peptides from Plasmodium falciparum circumsporozoite protein as antigens. *Journal of Molecular Biology* **1991**, *220*, 821–827.

(391) Willis, A. E.; Perham, R. N.; Wraith, D. C. Immunological properties of foreign peptides in multiple display on a filamentous bacteriophage. *Gene* **1993**, *128*, 79–83.

(392) Minenkova, O. O.; Ilyichev, A. A.; Kishchenko, G. P.; Petrenko, V. A. Design of specific immunogens using filamentous phage as the carrier. *Gene* **1993**, *128*, 85–88.

(393) Di Marzo Veronese, F.; Willis, A. E.; Boyer-Thompson, C.; Appella, E.; Perham, R. N. Structural mimicry and enhanced immunogenicity of peptide epitopes displayed on filamentous bacteriophage: The V3 loop of HIV-1 gp120. *Journal of Molecular Biology* **1994**, *243*, 167–172.

(394) De Berardinis, P.; D'Apice, L.; Prisco, A.; Ombra, M. N.; Barba, P.; Del Pozzo, G.; Petukhov, S.; Malik, P.; Perham, R. N.; Guardiola, J. Recognition of HIV-derived B and T cell epitopes displayed on filamentous phages. *Vaccine* **1999**, *17*, 1434–1441.

(395) Ulivieri, C.; Citro, A.; Ivaldi, F.; Mascolo, D.; Ghittoni, R.; Fanigliulo, D.; Manca, F.; Baldari, C. T.; Li Pira, G.; Del Pozzo, G. Antigenic properties of HCMV peptides displayed by filamentous bacteriophages vs. synthetic peptides. *Immunology Letters* **2008**, *119*, 62–70.

(396) Wang, G.; Sun, M.; Fang, J.; Yang, Q.; Tong, H.; Wang, L. Protective immune responses against systemic candidiasis mediated by phage-displayed specific epitope of Candida albicans heat shock protein 90 in C57BL/6J mice. *Vaccine* **2006**, *24*, 6065–6073.

(397) Shi, H.; Dong, S.; Zhang, X.; Chen, X.; Gao, X.; Wang, L. Phage vaccines displaying YGKDVKDLFDYAQE epitope induce protection against systemic candidiasis in mouse model. *Vaccine* **2018**, *36*, 5717–5724.

(398) Mascolo, D.; Barba, P.; De Berardinis, P.; Di Rosa, F.; Del Pozzo, G. Phage display of a CTL epitope elicits a long-term in vivo cytotoxic response. *FEMS Immunology & Medical Microbiology* **2007**, *50*, 59–66.

(399) Frenkel, D.; Katz, O.; Solomon, B. Immunization against Alzheimer's β-amyloid plaques via EFRH phage administration. *Proceedings of the National Academy of Sciences of the United States of America* **2000**, *97*, 11455–11459.

(400) Bruttin, A.; Brussow, H. Human volunteers receiving Escherichia coli phage T4 orally: A safety test of phage therapy. *Antimicrobial Agents and Chemotherapy* **2005**, *49*, 2874–2878.

(401) Borysowski, J.; Górski, A. Is phage therapy acceptable in the immunocompromised host? *International Journal of Infectious Diseases* **2008**, *12*, 466–471.

(402) Wright, A.; Hawkins, C. H.; Änggård, E.; Harper, D. R. A controlled clinical trial of a therapeutic bacteriophage preparation in chronic otitis due to antibiotic-resistant Pseudomonas aeruginosa; a preliminary report of efficacy. *Clinical Otolaryngology* **2009**, *34*, 349–357.

(403) Merabishvili, M. et al. Quality-controlled small-scale production of a well-defined bacteriophage cocktail for use in human clinical trials. *PLOS ONE* **2009**, *4*, ed. by Ojcius, D. M., e4944.

(404) Flotte, T. R.; Carter, B. J. Adeno-associated virus vectors for gene therapy. *Gene Therapy* **1995**, *2*, 357–362.

(405) Alting-Mees, M. A.; Short, J. M. pBluescript II: gene mapping vectors. *Nucleic Acids Research* **1989**, *17*, 9494.

(406) Bolivar, F.; Rodriguez, R. L.; Betlach, M. C.; Boyer, H. W. Construction and characterization of new cloning vehicles. I. Ampicillin-resistant derivatives of the plasmid pMB9. *Gene* **1977**, *2*, 75–93.

(407) Engler, C.; Kandzia, R.; Marillonnet, S. A one pot, one step, precision cloning method with high throughput capability. *PLOS ONE* **2008**, *3*, e3647.

(408) Scior, A.; Preissler, S.; Koch, M.; Deuerling, E. Directed PCR-free engineering of highly repetitive DNA sequences. *BMC Biotechnology* **2011**, *11*, 1–9.

(409) Peng, X.; Nguyen, A.; Ghosh, D. Quantification of M13 and T7 bacteriophages by TaqMan and SYBR green qPCR. *Journal of Virological Methods* **2018**, *252*, 100–107.

(410) Bichet, M. C.; Patwa, R.; Barr, J. J. Protocols for studying bacteriophage interactions with in vitro epithelial cell layers. *STAR Protocols* **2021**, *2*, 100697.

(411) Gruber, A. R.; Lorenz, R.; Bernhart, S. H.; Neuböck, R.; Hofacker, I. L. The Vienna RNA Websuite. *Nucleic Acids Research* **2008**, *36*, W70–W74.

(412) Lorenz, R.; Bernhart, S. H.; Höner zu Siederdissen, C.; Tafer, H.; Flamm, C.; Stadler, P. F.; Hofacker, I. L. ViennaRNA Package 2.0. *Algorithms for Molecular Biology* **2011**, *6*, 1–14.

(413) SantaLucia, J. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proceedings of the National Academy of Sciences of the United States of America* **1998**, *95*, 1460–1465.

(414) Harris, C. R. et al. Array programming with NumPy. *Nature* **2020**, *585*, 357–362.

(415) McKinney, W. In *Proceedings of the 9th Python in Science Conference*, Python in Science Conference, Austin, Texas, 2010, pp 56–61.

(416) Reback, J. et al. pandas-dev/pandas: Pandas 1.4.1, version v1.4.1, Zenodo, 2022.

(417) Virtanen, P. et al. SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods* **2020**, *17*, 261–272.

(418) The scikit-bio development team. scikit-bio: A Bioinformatics Library for Data Scientists, Students, and Developers, version 0.5.6, 2022.

(419) Seabold, S.; Perktold, J. In *Proceedings of the 9th Python in Science Conference*, Python in Science Conference, Austin, Texas, 2010, pp 92–96.

(420) Aitchison, J. The Statistical Analysis of Compositional Data. *Journal of the Royal Statistical Society: Series B (Methodological)* **1982**, *44*, 139–160.

(421) Mackay, I. M.; Arden, K. E.; Nitsche, A. Real-time PCR in virology. *Nucleic Acids Research* **2002**, *30*, 1292.

(422) Anderson, B.; Rashid, M. H.; Carter, C.; Pasternack, G.; Rajanna, C.; Revazishvili, T.; Dean, T.; Senecal, A.; Sulakvelidze, A. Enumeration of bacteriophage particles. *Bacteriophage* **2011**, *1*, 86–93.

(423) Ács, N.; Gambino, M.; Brøndsted, L. Bacteriophage enumeration and detection methods. *Frontiers in Microbiology* **2020**, *11*, 2662.

(424) Ding, Y.; Chan, C. Y.; Lawrence, C. E. RNA secondary structure prediction by centroids in a Boltzmann weighted ensemble. *RNA* **2005**, *11*, 1157.

(425) Kralik, P.; Ricchi, M. A basic guide to real time PCR in microbial diagnostics: Definitions, parameters, and everything. *Frontiers in Microbiology* **2017**, *8*, 108.

(426) Méndez-Scolari, J. E.; Florentín-Pavía, M. M.; Mujica, M. P.; Rojas, N.; Sotelo, P. H. A qPCR Targeted against the viral replication origin designed to quantify total amount of filamentous phages and phagemids. *Indian Journal of Microbiology* **2019**, *59*, 365–369.

(427) Edelman, D. C.; Barletta, J. Real-time PCR provides improved detection and titer determination of bacteriophage. *BioTechniques* **2003**, *35*, 368–375.

(428) Reitinger, S.; Petriv, O. I.; Mehr, K.; Hansen, C. L.; Withers, S. G. Purification and quantitation of bacteriophage M13 using desalting spin columns and digital PCR. *Journal of Virological Methods* **2012**, *185*, 171–174.

(429) Green, M. R.; Sambrook, J. Plating bacteriophage M13. *Cold Spring Harbor Protocols* **2017**, *2017*, pdb.prot093427.

(430) Dong, D.; Sutaria, S.; Hwangbo, J. Y.; Chen, P. A simple and rapid method to isolate purer M13 phage by isoelectric precipitation. *Applied Microbiology and Biotechnology* **2013**, *97*, 8023–8029.

(431) Warner, C. M.; Barker, N.; Lee, S.-W.; Perkins, E. J. M13 bacteriophage production for large-scale applications. *Bioprocess and Biosystems Engineering* **2014**, *37*, 2067–2072.

(432) Seetin, M. G.; Mathews, D. H. In *Bacterial Regulatory RNA: Methods and Protocols*, Keiler, K. C., Ed.; Methods in Molecular Biology; Humana Press: Totowa, NJ, 2012, pp 99–122.

(433) Hofacker, I. L. Vienna RNA secondary structure server. *Nucleic Acids Research* **2003**, *31*, 3429–3431.

(434) Gruber, A. R.; Bernhart, S. H.; Lorenz, R. The ViennaRNA web services. *Methods in Molecular Biology* **2015**, *1269*, 307–326.

(435) Ding, Y.; Lawrence, C. E. A statistical sampling algorithm for RNA secondary structure prediction. *Nucleic Acids Research* **2003**, *31*, 7280.

(436) Mathews, D. H. Revolutions in RNA Secondary Structure Prediction. *Journal of Molecular Biology* **2006**, *359*, 526–532.

(437) Dotto, G. P.; Enea, V.; Zinderi, N. D.; Zinder, N. D. Functional analysis of bacteriophage f1 intergenic region. *Virology* **1981**, *114*, 463–473.

(438) Noren, K. A.; Noren, C. J. Construction of high-complexity combinatorial phage display peptide libraries. *Methods* **2001**, *23*, 169–178.

(439) Green, M. R.; Sambrook, J. Preparation of double-stranded (replicative form) bacteriophage M13 DNA. *Cold Spring Harbor Protocols* **2017**, *2017*, pdb.prot093443.

(440) Hoseini, S. S.; Sauer, M. G. Molecular cloning using polymerase chain reaction, an educational guide for cellular engineering. *Journal of Biological Engineering* **2015**, *9*, 1–13.

(441) Bullock, W. O.; Fernandez, J. M.; Short, J. M. XL1-Blue—a high-efficiency plasmid transforming recA Escherichia coli strain with β-galactosidase selection. *BioTechniques* **1987**, *5*, 376–381.

(442) Gough, J. A.; Murray, N. E.; Brenner, S. Sequence diversity among related genes for recognition of specific targets in DNA molecules. *Journal of Molecular Biology* **1983**, *166*, 1–19.

(443) Jang, S.; Park, W.; Chung, S.-K.; Jeong, C.; Chung, D. New E. coli cloning vector using a cellulase gene (celA) as a screening marker. *BioTechniques* **2001**, *31*, 1064–1068.

(444)   Saing, K. M.; Orii, H.; Tanaka, Y.; Yanagisawa, K.; Miura, A.; Ikeda, H. Formation of deletion in Escherichia coli between direct repeats located in the long inverted repeats of a cellular slime mold plasmid: Participation of DNA gyrase. *Molecular and General Genetics MGG* **1988**, *214*, 1–5.

(445)   Yau, S. Y.; Keshavarz-Moore, E.; Ward, J. Host strain influences on supercoiled plasmid DNA production in Escherichia coli: Implications for efficient design of large-scale processes. *Biotechnology and Bioengineering* **2008**, *101*, 529–544.

(446)   Messing, J.; Gronenborn, B.; Müller-Hill, B.; Hopschneider, P. H. Filamentous coliphage M13 as a cloning vehicle: insertion of a HindII fragment of the lac regulatory region in M13 replicative form in vitro. *Proceedings of the National Academy of Sciences of the United States of America* **1977**, *74*, 3642–3646.

(447)   Chai, D.; Wang, G.; Fang, L.; Li, H.; Liu, S.; Zhu, H.; Zheng, J. The optimization system for preparation of TG1 competent cells and electrotransformation. *MicrobiologyOpen* **2020**, *9*, e1043.

(448)   Miller, J. H.; Miller, J. B., *Experiments in Molecular Genetics*; Cold Spring Harbor Laboratory: 1972; 494 pp.

(449)   Souriau, C.; Fort, P.; Roux, P.; Hartley, O.; Lefranc, M.-P.; Weill, M. A simple luciferase assay for signal transduction activity detection of epidermal growth factor displayed on phage. *Nucleic Acids Research* **1997**, *25*, 1585–1590.

(450)   Li, Z.; Jiang, H.; Zhang, J.; Gu, J. Cell-targeted phagemid particles preparation using Escherichia coli bearing ligand-pIII encoding helper phage genome. *BioTechniques* **2006**, *41*, 706–707.

(451)   Jiang, H.; Cai, X.-M.; Shi, B.-Z.; Zhang, J.; Li, Z.-H.; Gu, J.-R. Development of efficient RNA interference system using EGF-displaying phagemid particles. *Acta Pharmacologica Sinica* **2008**, *29*, 437–442.

(452)   Geider, K.; Baldes, R. Influence of fd gene 2-protein and the viral replication origin on the compatibility of pfd-plasmids. *Nucleic Acids Research* **1988**, *16*, 6385–6396.

(453)   Woodcock, D.; Crowther, P.; Doherty, J.; Jefferson, S.; DeCruz, E.; Noyer-Weidner, M.; Smith, S.; Michael, M.; Graham, M. Quantitative evaluation of Escherichia coli host strains for tolerance to cytosine methylation in plasmid and phage recombinants. *Nucleic Acids Research* **1989**, *17*, 3469–3478.

(454)   Trinh, T.; Jessee, J.; Bloom, F. R.; Hirsch, V. STBL2: an Escherichia coli strain for the stable propagation of retroviral clones and direct repeat sequences. *Focus* **1994**, *16*, 78–80.

(455)   Bzymek, M.; Lovett, S. T. Instability of repetitive DNA sequences: the role of replication in multiple mechanisms. *Proceedings of the National Academy of Sciences of the United States of America* **2001**, *98*, 8319–8325.

(456)   Grant, S. G.; Jessee, J.; Bloom, F. R.; Hanahan, D. Differential plasmid rescue from transgenic mouse DNAs into Escherichia coli methylation-restriction mutants. *Proceedings of the National Academy of Sciences of the United States of America* **1990**, *87*, 4645–4649.

(457)   Biolabs, N. E. NEB® Turbo Competent E. coli (High Efficiency) New England Biolabs, `https://international.neb.com/products/c2984-neb-turbo-competent-e-coli-high-efficiency#Product%20Information` (accessed 01/22/2022).

(458)   Huh, H. Characterization of bacteriophage λ displaying epidermal growth factor in the uptake, infiltration and formation of HT29 spheroids, University of Waterloo, 2019.

(459)   Chen, X.; Zaro, J.; Shen, W.-C. Fusion protein linkers: Property, design and functionality. *Advanced Drug Delivery Reviews* **2013**, *65*, 1357–1369.

(460)   Slavcev, R. A.; Huh, H.; Blay, J. (Theraphage Inc). EGFR binding moiety-presenting bacteriophages for tumour treatment US Patent, 20210100858A1, 2021.

(461) Chen, Y.-J.; Liu, P.; Nielsen, A. A. K.; Brophy, J. A. N.; Clancy, K.; Peterson, T.; Voigt, C. A. Characterization of 582 natural and synthetic terminators and quantification of their design constraints. *Nature Methods* **2013**, *10*, 659–664.

(462) Wu, A. M.; Platt, T. Transcription termination: Nucleotide sequence at 3′ end of tryptophan operon in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America* **1978**, *75*, 5442–5446.

(463) Rossi, J.; Egan, J.; Hudson, L.; Landy, A. The tyrT locus: Termination and processing of a complex transcript. *Cell* **1981**, *26*, 305–314.

(464) Jahn, M.; Vorpahl, C.; Hübschmann, T.; Harms, H.; Müller, S. Copy number variability of expression plasmids determined by cell sorting and Droplet Digital PCR. *Microbial Cell Factories* **2016**, *15*, 1–12.

(465) Lerner, T. J.; Model, P. The "steady state" of coliphage f1: DNA synthesis late in infection. *Virology* **1981**, *115*, 282–294.

(466) Assad, L.; Matjošaitis, K.; Gross, H. Genome sequence of Escherichia coli Stbl4, a versatile genetic tool for heterologous expression. *Microbiology Resource Announcements* **2021**, *10*, e00823–21.

(467) Taylor, R. G.; Walker, D. C.; McInnes, R. R. E. coli host strains significantly affect the quality of small scale plasmid DNA preparations used for sequencing. *Nucleic Acids Research* **1993**, *21*, 1677.

(468) Geider, K.; Baldes, R.; Bellemann, P.; Metzger, M.; Schwartz, T. Mutual adaptation of bacteriophage fd, pfd plasmids and their host strains. *Microbiological Research* **1995**, *150*, 337–346.

(469) Kay, A.; O'Kennedy, R.; Ward, J.; Keshavarz-Moore, E. Impact of plasmid size on cellular oxygen demand in Escherichia coli. *Biotechnology and Applied Biochemistry* **2003**, *38*, 1–7.

(470) Jeong, H.; Sim, Y. M.; Kim, H. J.; Lee, S. J. Unveiling the hybrid genome structure of Escherichia coli RR1 (HB101 RecA+). *Frontiers in Microbiology* **2017**, *8*.

(471) Al-Allaf, F. A.; Tolmachov, O. E.; Zambetti, L. P.; Tchetchelnitski, V.; Mehmet, H. Remarkable stability of an instability-prone lentiviral vector plasmid in Escherichia coli Stbl3. *3 Biotech* **2013**, *3*, 61–70.

(472) Bloom, F. R.; Schmidt, B. J.; Lin, J.-J. (Invitrogen Corp). Rapid growing microorganisms for biotechnology applications US Patent, 6709852B1, 2004.

(473) Inchley, C. J. The activity of mouse Kupffer cells following intravenous injection of T4 bacteriophage. *Clinical and Experimental Immunology* **1969**, *5*, 173.

(474) Salomon, D. S.; Brandt, R.; Ciardiello, F.; Normanno, N. Epidermal growth factor-related peptides and their receptors in human malignancies. *Critical Reviews in Oncology/Hematology* **1995**, *19*, 183–232.

(475) Seshacharyulu, P.; Ponnusamy, M. P.; Haridas, D.; Jain, M.; Ganti, A.; Batra, S. K. Targeting the EGFR signaling pathway in cancer therapy. *Expert Opinion on Therapeutic Targets* **2012**, *16*, 15.

(476) Sigismund, S.; Avanzato, D.; Lanzetti, L. Emerging functions of the EGFR in cancer. *Molecular Oncology* **2018**, *12*, 3–20.

(477) Donnelly, A.; Yata, T.; Bentayebi, K.; Suwan, K.; Hajitou, A. Bacteriophage mediates efficient gene transfer in combination with conventional transfection reagents. *Viruses* **2015**, *7*, 6476–6489.

(478) Shete, H. K.; Prabhu, R. H.; Patravale, V. B. Endosomal escape: a bottleneck in intracellular delivery. *Journal of Nanoscience and Nanotechnology* **2014**, *14*, 460–474.

(479) Cupic, K. I.; Rennick, J. J.; Johnston, A. P.; Such, G. K. Controlling endosomal escape using nanoparticle composition: current progress and future perspectives. *Nanomedicine* **2019**, *14*, 215–223.

(480) Behr, J.-P. J. The proton sponge: a trick to enter cells the viruses did not exploit. *International Journal for Chemistry* **1997**, *2*, 34–36.

(481) Sabbah, D. A.; Hajjo, R.; Sweidan, K. Review on epidermal growth factor receptor (EGFR) structure, signaling pathways, interactions, and recent updates of EGFR inhibitors. *Current Topics in Medicinal Chemistry* **2020**, *20*, 815–834.

(482) Aloisio, A.; Nisticò, N.; Mimmi, S.; Maisano, D.; Vecchio, E.; Fiume, G.; Iaccino, E.; Quinto, I. Phage-displayed peptides for targeting tyrosine kinase membrane receptors in cancer therapy. *Viruses* **2021**, *13*, 649.

(483) Olayioye, M. A. The ErbB signaling network: receptor heterodimerization in development and cancer. *The EMBO Journal* **2000**, *19*, 3159–3167.

(484) Arteaga, C. Targeting HER1/EGFR: A molecular approach to cancer therapy. *Seminars in Oncology* **2003**, *30*, 3–14.

(485) Mazzarella, L.; Guida, A.; Curigliano, G. Cetuximab for treating non-small cell lung cancer. *Expert Opinion on Biological Therapy* **2018**, *18*, 483–493.

(486) Cohen, S. Isolation of a Mouse Submaxillary Gland Protein Accelerating Incisor Eruption and Eyelid Opening in the New-born Animal. *Journal of Biological Chemistry* **1962**, *237*, 1555–1562.

(487) Zeng, F.; Harris, R. C. Epidermal growth factor, from gene organization to bedside. *Seminars in Cell & Developmental Biology* **2014**, *28*, 2–11.

(488) Singh, B.; Carpenter, G.; Coffey, R. J. EGF receptor ligands: recent advances. *F1000Research* **2016**, *5*.

(489) Sorkin, A.; Goh, L. K. Endocytosis and intracellular trafficking of ErbBs. *Experimental Cell Research* **2009**, *315*, 683–696.

(490) Heldin, C. H. Dimerization of cell surface receptors in signal transduction. *Cell* **1995**, *80*, 213–223.

(491) Shoelson, S. E. SH2 and PTB domain interactions in tyrosine kinase signal transduction. *Current Opinion in Chemical Biology* **1997**, *1*, 227–234.

(492) Wieduwilt, M. J.; Moasser, M. M. The epidermal growth factor receptor family: biology driving targeted therapeutics. *Cellular and Molecular Life Sciences* **2008**, *65*, 1566–1584.

(493) Furdui, C. M.; Lew, E. D.; Schlessinger, J.; Anderson, K. S. Autophosphorylation of FGFR1 kinase is mediated by a sequential and precisely ordered reaction. *Molecular Cell* **2006**, *21*, 711–717.

(494) Kim, Y.; Li, Z.; Apetri, M.; Luo, B.; Settleman, J. E.; Anderson, K. S. Temporal resolution of autophosphorylation for normal and oncogenic forms of EGFR and differential effects of gefitinib. *Biochemistry* **2012**, *51*, 5212–5222.

(495) Wang, Q.; Chen, X.; Wang, Z. Dimerization drives EGFR endocytosis through two sets of compatible endocytic codes. *Journal of Cell Science* **2015**, *128*, 935–950.

(496) Roepstorff, K.; Grandal, M. V.; Henriksen, L.; Knudsen, S. L. J.; Lerdrup, M.; Grøvdal, L.; Willumsen, B. M.; van Deurs, B. Differential effects of EGFR ligands on endocytic sorting of the receptor. *Traffic* **2009**, *10*, 1115–1127.

(497) Alwan, H. A. J.; van Zoelen, E. J. J.; van Leeuwen, J. E. M. Ligand-induced lysosomal epidermal growth factor receptor (EGFR) degradation is preceded by proteasome-dependent EGFR de-ubiquitination. *Journal of Biological Chemistry* **2003**, *278*, 35781–35790.

(498) Sorkina, T.; Huang, F.; Beguinot, L.; Sorkin, A. Effect of tyrosine kinase inhibitors on clathrin-coated pit recruitment and internalization of epidermal growth factor receptor. *Journal of Biological Chemistry* **2002**, *277*, 27433–27441.

(499) Pinilla-Macua, I.; Sorkin, A. Methods to study endocytic trafficking of the EGF receptor. *Methods in Cell Biology* **2015**, *130*, 347–367.

(500) Dikic, I. Mechanisms controlling EGF receptor endocytosis and degradation. *Biochemical Society Transactions* **2003**, *31*, 1178–1181.

(501) Sorkin, A. D.; Teslenko, L. V.; Nikolsky, N. N. The endocytosis of epidermal growth factor in A431 cells: a pH of microenvironment and the dynamics of receptor complex dissociation. *Experimental Cell Research* **1988**, *175*, 192–205.

(502) Ivanenkov, V. V.; Felici, F.; Menon, A. G. Uptake and intracellular fate of phage display vectors in mammalian cells. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1999**, *1448*, 450–462.

(503) Johnson, D. E.; Ostrowski, P.; Jaumouillé, V.; Grinstein, S. The position of lysosomes within the cell determines their luminal pH. *Journal of Cell Biology* **2016**, *212*, 677–692.

(504) Sakai, K.; Arao, T.; Shimoyama, T.; Murofushi, K.; Sekijima, M.; Kaji, N.; Tamura, T.; Saijo, N.; Nishio, K. Dimerization and the signal transduction pathway of a smallin-frame deletion in the epidermal growth factor receptor. *The FASEB Journal* **2006**, *20*, 311–313.

(505) Zhang, F.; Wang, S.; Yin, L.; Yang, Y.; Guan, Y.; Wang, W.; Xu, H.; Tao, N. Quantification of epidermal growth factor receptor expression level and binding kinetics on cell surfaces by surface plasmon resonance imaging. *Analytical Chemistry* **2015**, *87*, 9960.

(506) Brea, M. S.; Díaz, R. G.; Escudero, D. S.; Caldiz, C. I.; Portiansky, E. L.; Morgan, P. E.; Pérez, N. G. Epidermal growth factor receptor silencing blunts the slow force response to myocardial stretch. *Journal of the American Heart Association* **2016**, *5*, e004017.

(507) Sirotnak, F. M.; Zakowski, M. F.; Miller, V. A.; Scher, H. I.; Kris, M. G. Efficacy of cytotoxic agents against human tumor xenografts is markedly enhanced by coadministration of ZD1839 (Iressa), an inhibitor of EGFR tyrosine kinase. *Clinical Cancer Research* **2000**, *6*, 4885–4892.

(508) Ahsan, A.; Ramanand, S. G.; Whitehead, C.; Hiniker, S. M.; Rehemtulla, A.; Pratt, W. B.; Jolly, S.; Gouveia, C.; Truong, K.; Waes, C. V.; Ray, D.; Lawrence, T. S.; Nyati, M. K. Wild-type EGFR is stabilized by direct interaction with HSP90 in cancer cells and tumors. *Neoplasia* **2012**, *14*, 670–IN1.

(509) Chowdhury, S. M.; Mannepalli, P.; Sitharaman, B. Graphene nanoribbons elicit cell specific uptake and delivery via activation of epidermal growth factor receptor enhanced by human papillomavirus E5 protein. *Acta Biomaterialia* **2014**, *10*, 4494–4504.

(510) Zhang, K.; Ma, Y.; Guo, Y.; Sun, T.; Wu, J.; Pangeni, R. P.; Lin, M.; Li, W.; Horne, D.; Raz, D. J. Cetuximab-triptolide conjugate suppresses the growth of EGFR-overexpressing lung cancers through targeting RNA polymerase II. *Molecular Therapy Oncolytics* **2020**, *18*, 304–316.

(511) Keese, M.; Magdeburg, R. J.; Herzog, T.; Hasenberg, T.; Offterdinger, M.; Pepperkok, R.; Sturm, J. W.; Bastiaens, P. I. H. Imaging epidermal growth factor receptor phosphorylation in human colorectal cancer cells and human tissues. *Journal of Biological Chemistry* **2005**, *280*, 27826–27831.

(512) Solmi, R. et al. Displayed correlation between gene expression profiles and submicroscopic alterations in response to cetuximab, gefitinib and EGF in human colon cancer cell lines. *BMC Cancer* **2008**, *8*, 1–22.

(513) Balin-Gauthier, D.; Delord, J.-P.; Rochaix, P.; Mallard, V.; Thomas, F.; Hennebelle, I.; Bugat, R.; Canal, P.; Allal, C. In vivo and in vitro antitumor activity of oxaliplatin in

combination with cetuximab in human colorectal tumor cell lines expressing different level of EGFR. *Cancer Chemotherapy and Pharmacology* **2006**, *57*, 709–718.

(514) Loeffler-Ragg, J.; Skvortsov, S.; Sarg, B.; Skvortsova, I.; Witsch-Baumgartner, M.; Mueller, D.; Lindner, H.; Zwierzina, H. Gefitinib-responsive EGFR-positive colorectal cancers have different proteome profiles from non-responsive cell lines. *European Journal of Cancer* **2005**, *41*, 2338–2346.

(515) Shimizu, M.; Deguchi, A.; Lim, J. T. E.; Moriwaki, H.; Kopelovich, L.; Weinstein, I. B. (-)-Epigallocatechin Gallate and Polyphenon E Inhibit Growth and Activation of the Epidermal Growth Factor Receptor and Human Epidermal Growth Factor Receptor-2 Signaling Pathways in Human Colon Cancer Cells. *Clinical Cancer Research* **2005**, *11*, 2735–2746.

(516) Feng, Y.; Gao, S.; Gao, Y.; Wang, X.; Chen, Z. Anti-EGFR antibody sensitizes colorectal cancer stem-like cells to Fluorouracil-induced apoptosis by affecting autophagy. *Oncotarget* **2016**, *7*, 81402–81409.

(517) Sweere, J. M. et al. Bacteriophage trigger antiviral immunity and prevent clearance of bacterial infection. *Science* **2019**, *363*, eaat9691.

(518) Larocca, D.; Baird, A. Receptor-mediated gene transfer by phage-display vectors: applications in functional genomics and gene therapy. *Drug Discovery Today* **2001**, *6*, 793–801.

(519) Branston, S.; Stanley, E.; Keshavarz-Moore, E.; Ward, J. Precipitation of filamentous bacteriophages for their selective recovery in primary purification. *Biotechnology Progress* **2012**, *28*, 129–136.

(520) Passaretti, P.; Khan, I.; Dafforn, T. R.; Goldberg Oppenheimer, P. Improvements in the production of purified M13 bacteriophage bio-nanoparticle. *Scientific Reports* **2020**, *10*, 1–9.

(521) Mishra, S.; Webster, P.; Davis, M. E. PEGylation significantly affects cellular uptake and intracellular trafficking of non-viral gene delivery particles. *European Journal of Cell Biology* **2004**, *83*, 97–111.

(522) Rejman, J.; Wagenaar, A.; Engberts, J. B. F. N.; Hoekstra, D. Characterization and transfection properties of lipoplexes stabilized with novel exchangeable polyethylene glycol-lipid conjugates. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2004**, *1660*, 41–52.

(523) Suk, J. S.; Xu, Q.; Kim, N.; Hanes, J.; Ensign, L. M. PEGylation as a strategy for improving nanoparticle-based drug and gene delivery. *Advanced Drug Delivery Reviews* **2016**, *99*, 28–51.

(524) Ogris, M.; Brunner, S.; Schüller, S.; Kircheis, R.; Wagner, E. PEGylated DNA/transferrin–PEI complexes: reduced interaction with blood components, extended circulation in blood and potential for systemic gene delivery. *Gene Therapy* **1999**, *6*, 595–605.

(525) Kim, K. P.; Cha, J. D.; Jang, E. H.; Klumpp, J.; Hagens, S.; Hardt, W. D.; Lee, K. Y.; Loessner, M. J. PEGylation of bacteriophages increases blood circulation time and reduces T-helper type 1 immune response. *Microbial Biotechnology* **2008**, *1*, 247–257.

(526) Rejman, J.; Oberle, V.; Zuhorn, I. S.; Hoekstra, D. Size-dependent internalization of particles via the pathways of clathrin- and caveolae-mediated endocytosis. *Biochemical Journal* **2004**, *377*, 159–169.

(527) Rennick, J. J.; Johnston, A. P. R.; Parton, R. G. Key principles and methods for studying the endocytosis of biological and nanoparticle therapeutics. *Nature Nanotechnology* **2021**, *16*, 266–276.

(528) Kumari, S.; Mg, S.; Mayor, S. Endocytosis unplugged: multiple ways to enter the cell. *Cell Research* **2010**, *20*, 256–275.

(529) Choi, D. S.; Jin, H. E.; Yoo, S. Y.; Lee, S. W. Cyclic RGD peptide incorporation on phage major coat proteins for improved internalization by HeLa cells. *Bioconjugate Chemistry* **2014**, *25*, 216–223.

(530) Sunada, H.; Magun, B. E.; Mendelsohn, J.; MacLeod, C. L. Monoclonal antibody against epidermal growth factor receptor is internalized without stimulating receptor phosphorylation. *Proceedings of the National Academy of Sciences of the United States of America* **1986**, *83*, 3825–3829.

(531) Berger, C.; Madshus, I. H.; Stang, E. Cetuximab in combination with anti-human IgG antibodies efficiently down-regulates the EGF receptor by macropinocytosis. *Experimental Cell Research* **2012**, *318*, 2578–2591.

(532) Wang, Q.; Villeneuve, G.; Wang, Z. Control of epidermal growth factor receptor endocytosis by receptor dimerization, rather than receptor kinase activation. *EMBO Reports* **2005**, *6*, 942–948.

(533) Wang, Q.; Zhu, F.; Wang, Z. Identification of EGF receptor C-terminal sequences 1005–1017 and di-leucine motif 1010LL1011 as essential in EGF receptor endocytosis. *Experimental Cell Research* **2007**, *313*, 3349–3363.

(534) Kourouniotis, G.; Wang, Y.; Pennock, S.; Chen, X.; Wang, Z. Non-ligand-induced dimerization is sufficient to initiate the signalling and endocytosis of EGF receptor. *International Journal of Molecular Sciences* **2016**, *17*, 1200.

(535) Eldredge, E. R.; Korf, G. M.; Christensen, T. A.; Connolly, D. C.; Getz, M. J.; Maihle, N. J. Activation of c-fos gene expression by a kinase-deficient epidermal growth factor receptor. *Molecular and Cellular Biology* **1994**, *14*, 7527–7534.

(536) Cohen, S.; Fava, R. A. Internalization of functional epidermal growth factor:receptor/kinase complexes in A-431 cells. *The Journal of Biological Chemistry* **1985**, *260*, 12351–12358.

(537) Wang, Y.; Pennock, S.; Chen, X.; Wang, Z. Endosomal signaling of epidermal growth factor receptor stimulates signal transduction pathways leading to cell survival. *Molecular and Cellular Biology* **2002**, *22*, 7279–7290.

(538) Wang, Y.; Billing, S.; Wang, Z. Activation of endosome-associated inert EGF receptor following internalization. *Methods in Molecular Biology* **2017**, *1652*, 117–126.

(539) Ferrari, F. K.; Samulski, T.; Shenk, T.; Samulski, R. J. Second-strand synthesis is a rate-limiting step for efficient transduction by recombinant adeno-associated virus vectors. *Journal of Virology* **1996**, *70*, 3227–3234.

(540) Alexander, I. E.; Russell, D. W.; Miller, A. D. DNA-damaging agents greatly increase the transduction of nondividing cells by adeno-associated virus vectors. *Journal of Virology* **1994**, *68*, 8282–8287.

(541) Tsafa, E.; Bentayebi, K.; Topanurak, S.; Yata, T.; Przystal, J.; Fongmoon, D.; Hajji, N.; Waramit, S.; Suwan, K.; Hajitou, A. Doxorubicin improves cancer cell targeting by filamentous phage gene delivery vectors. *International Journal of Molecular Sciences* **2020**, *21*, 7867.

(542) Heinrich, J.; Schultz, J.; Bosse, M.; Ziegelin, G.; Lanka, E.; Moelling, K. Linear closed mini DNA generated by the prokaryotic cleaving-joining enzyme TelN is functional in mammalian cells. *Journal of Molecular Medicine* **2002**, *80*, 648–654.

(543) Scott, V. L.; Patel, A.; Villarreal, D. O.; Hensley, S. E.; Ragwan, E.; Yan, J.; Sardesai, N. Y.; Rothwell, P. J.; Extance, J. P.; Caproni, L. J.; Weiner, D. B. Novel synthetic plasmid and Doggybone™ DNA vaccines induce neutralizing antibodies and

provide protection from lethal influenza challenge in mice. *Human Vaccines & Immunotherapeutics* **2015**, *11*, 1972–1982.

(544) Sain, B.; Murray, N. E. The hsd (host specificity) genes of E. coli K 12. *Molecular and General Genetics MGG* **1980**, *180*, 35–46.

(545) Raleigh, E. A.; Murray, N. E.; Revel, H.; Blumenthal, R. M.; Westaway, D.; Reith, A. D.; Rigby, P. W.; Elhai, J.; Hanahan, D. McrA and McrB restriction phenotypes of some E. coli strains and implications for gene cloning. *Nucleic Acids Research* **1988**, *16*, 1563–1575.

(546) Eun, H.-M., *Enzymology Primer for Recombinant DNA Technology*; Elsevier: 1996; 727 pp.

(547) Singer, A.; Eiteman, M. A.; Altman, E. DNA plasmid production in different host strains of Escherichia coli. *Journal of Industrial Microbiology and Biotechnology* **2009**, *36*, 521–530.

(548) Clark, A. J.; Low, K. B. In *The Recombination of Genetic Material*, Low, K. B., Ed.; Academic Press: 1988, pp 155–215.

(549) Kowalczykowski, S. C.; Dixon, D. A.; Eggleston, A. K.; Lauder, S. D.; Rehrauer, W. M. Biochemistry of homologous recombination in Escherichia coli. *Microbiological Reviews* **1994**, *58*, 401–465.

(550) Dillingham, M. S.; Kowalczykowski, S. C. RecBCD Enzyme and the Repair of Double-Stranded DNA Breaks. *Microbiology and Molecular Biology Reviews* **2008**, *72*, 642–671.

(551) Smith, G. R. Homologous recombination in prokaryotes: enzymes and controlling sites. *Genome* **1989**, *31*, 520–527.

(552) Eggleston, A. K.; West, S. C. Recombination initiation: easy as A, B, C, D... chi? *Current Biology* **1997**, *7*, R745–749.

(553) Buljubašić, M.; Hlevnjak, A.; Repar, J.; Đermić, D.; Filić, V.; Weber, I.; Zahradka, K.; Zahradka, D. RecBCD- RecFOR-independent pathway of homologous recombination in Escherichia coli. *DNA Repair* **2019**, *83*, 102670.

(554) Cohen, A.; Laban, A. Plasmidic recombination in Escherichia coli K-12: the role of recF gene function. *Molecular and General Genetics MGG* **1983**, *189*, 471–474.

(555) Lloyd, R. G.; Picksley, S. M.; Prescott, C. Inducible expression of a gene specific to the RecF pathway for recombination in Escherichia coli K12. *Molecular and General Genetics MGG* **1983**, *190*, 162–167.

(556) Lovett, S. T.; Clark, A. J. Genetic analysis of the recJ gene of Escherichia coli K-12. *Journal of Bacteriology* **1984**, *157*, 190–196.

(557) Gillen, J. R.; Willis, D. K.; Clark, A. J. Genetic analysis of the RecE pathway of genetic recombination in Escherichia coli K-12. *Journal of Bacteriology* **1981**, *145*, 521–532.

(558) Kuzminov, A. Recombinational repair of DNA damage in Escherichia coli and bacteriophage lambda. *Microbiology and Molecular Biology Reviews* **1999**, *63*, 751–813, table of contents.

(559) Joseph, J. W.; Kolodner, R. Exonuclease VIII of Escherichia coli. II. Mechanism of action. *The Journal of Biological Chemistry* **1983**, *258*, 10418–10424.

(560) Hall, S. D.; Kolodner, R. D. Homologous pairing and strand exchange promoted by the Escherichia coli RecT protein. *Proceedings of the National Academy of Sciences of the United States of America* **1994**, *91*, 3205–3209.

(561) Voineagu, I.; Narayanan, V.; Lobachev, K. S.; Mirkin, S. M. Replication stalling at unstable inverted repeats: Interplay between DNA hairpins and fork stabilizing proteins. *Proceedings of the National Academy of Sciences of the United States of America* **2008**, *105*, 9936–9941.

(562) Leach, D. R. F. In *Genetic Engineering: Principles and Methods*, Setlow, J. K., Ed.; Genetic Engineering; Springer US: Boston, MA, 1996, pp 1–11.

(563) Branzei, D.; Foiani, M. Leaping forks at inverted repeats. *Genes & Development* **2010**, *24*, 5–9.

(564) Warren, G. J.; Green, R. L. Comparison of physical and genetic properties of palindromic DNA sequences. *Journal of Bacteriology* **1985**, *161*, 1103–1111.

(565) Leach, D. R.; Stahl, F. W. Viability of lambda phages carrying a perfect palindrome in the absence of recombination nucleases. *Nature* **1983–0005**, *305*, 448–451.

(566) Wyman, A. R.; Wertman, K. F.; Barker, D.; Helms, C.; Petri, W. H. Factors which equalize the representation of genome segments in recombinant libraries. *Gene* **1986**, *49*, 263–271.

(567) Ishiura, M.; Hazumi, N.; Koide, T.; Uchida, T.; Okada, Y. A recB recC sbcB recJ host prevents recA-independent deletions in recombinant cosmid DNA propagated in Escherichia coli. *Journal of Bacteriology* **1989**, *171*, 1068–1074.

(568) Greener, A. L. (Stratagene California). Novel cloning host organisms European pat., 0437071A1, 1991.

(569) Chalker, A. F.; Leach, D. R. F.; Lloyd, R. G. Escherichia coli sbcC mutants permit stable propagation of DNA replicons containing a long palindrome. *Gene* **1988**, *71*, 201–205.

(570) Azeroglu, B.; Lincker, F.; White, M. A.; Jain, D.; Leach, D. R. F. A perfect palindrome in the Escherichia coli chromosome forms DNA hairpins on both leading- and lagging-strands. *Nucleic Acids Research* **2014**, *42*, 13206–13213.

(571) Fadrosh, D. W.; Andrews-Pfannkoch, C.; Williamson, S. J. Separation of single-stranded DNA, double-stranded DNA and RNA from an environmental viral community using hydroxyapatite chromatography. *Journal of Visualized Experiments* **2011**, e3146.

(572) Andrews-Pfannkoch, C.; Fadrosh, D. W.; Thorpe, J.; Williamson, S. J. Hydroxyapatite-mediated separation of double-stranded DNA, single-stranded DNA, and RNA genomes from natural viral assemblages. *Applied and Environmental Microbiology* **2010**, *76*, 5039–5045.

(573) National Center for Biotechnology Information PubChem Compound Summary for CID 6032, Kanamycin `https://pubchem.ncbi.nlm.nih.gov/compound/Kanamycin` (accessed 02/02/2022).

(574) National Center for Biotechnology Information PubChem Compound Summary for CID 6249, Ampicillin `https://pubchem.ncbi.nlm.nih.gov/compound/Ampicillin` (accessed 02/02/2022).

(575) Kim, S.; Chen, J.; Cheng, T.; Gindulyte, A.; He, J.; He, S.; Li, Q.; Shoemaker, B. A.; Thiessen, P. A.; Yu, B.; Zaslavsky, L.; Zhang, J.; Bolton, E. E. PubChem in 2021: New data content and improved web interfaces. *Nucleic Acids Research* **2021**, *49*, D1388–D1395.

(576) Green, M. R.; Sambrook, J. Growing bacteriophage M13 in liquid culture. *Cold Spring Harbor Protocols* **2017**, *2017*, pdb.prot093435.

(577) Bernardi, G. Chromatography of nucleic acids on hydroxyapatite. *Nature* **1965**, *206*, 779–783.

(578) Beland, F. A.; Dooley, K. L.; Casciano, D. A. Rapid isolation of carcinogen-bound DNA and RNA by hydroxyapatite chromatography. *Journal of Chromatography A* **1979**, *174*, 177–186.

(579) Koetsier, G.; Cantor, E. A practical guide to analyzing nucleic acid concentration and purity with microvolume spectrophotometers. *New England BioLabs* **2019**, *8*, 1–8.

(580) Malagón, F.; Aguilera, A. Genetic stability and DNA rearrangements associated with a 2×1.1-Kb perfect palindrome in Escherichia coli. *Molecular and General Genetics MGG* **1998**, *259*, 639–644.

(581) Leach, D. R. F.; Okely, E. A.; Pinder, D. J. Repair by recombination of DNA containing a palindromic sequence. *Molecular Microbiology* **1997**, *26*, 597–606.

(582) Connelly, J. C.; Kirkham, L. A.; Leach, D. R. F. The SbcCD nuclease of Escherichia coli is a structural maintenance of chromosomes (SMC) family protein that cleaves hairpin DNA. *Proceedings of the National Academy of Sciences of the United States of America* **1998**, *95*, 7969–7974.

(583) Bzymek, M.; Lovett, S. T. Evidence for two mechanisms of palindrome-stimulated deletion in Escherichia coli: Single-strand annealing and replication slipped mispairing. *Genetics* **2001**, *158*, 527–540.

(584) Lovett, S. T. Encoded errors: Mutations and rearrangements mediated by misalignment at repetitive DNA sequences. *Molecular Microbiology* **2004**, *52*, 1243–1253.

(585) Sezonov, G.; Joseleau-Petit, D.; D'Ari, R. Escherichia coli physiology in Luria-Bertani broth. *Journal of Bacteriology* **2007**, *189*, 8746–8749.

(586) Lange, R.; Hengge-Aronis, R. The cellular concentration of the sigma S subunit of RNA polymerase in Escherichia coli is controlled at the levels of transcription, translation, and protein stability. *Genes & Development* **1994**, *8*, 1600–1612.

(587) Grieco, S.-H. H.; Lee, S.; Dunbar, W. S.; MacGillivray, R. T. A.; Curtis, S. B. Maximizing filamentous phage yield during computer-controlled fermentation. *Bioprocess and Biosystems Engineering* **2009**, *32*, 773–779.

(588) Kanter, P. M.; Schwartz, H. S. A hydroxylapatite batch assay for quantitation of cellular DNA damage. *Analytical Biochemistry* **1979**, *97*, 77–84.

(589) Horiuchi, K. Co-evolution of a filamentous bacteriophage and its defective interfering particles. *Journal of Molecular Biology* **1983**, *169*, 389–407.

(590) Shlyakhtenko, L. S.; Potaman, V. N.; Sinden, R. R.; Lyubchenko, Y. L. Structure and dynamics of supercoil-stabilized DNA cruciforms. *Journal of Molecular Biology* **1998**, *280*, 61–72.

(591) Russel, M.; Kidd, S.; Kelley, M. R. An improved filamentous helper phage for generating single-stranded plasmid DNA. *Gene* **1986**, *45*, 333–338.

(592) Midoux, P.; Kichler, A.; Boutin, V.; Maurizot, J. C.; Monsigny, M. Membrane permeabilization and efficient gene transfer by a peptide containing several histidines. *Bioconjugate Chemistry* **1998**, *9*, 260–267.

(593) Mitchell, M. J.; Billingsley, M. M.; Haley, R. M.; Wechsler, M. E.; Peppas, N. A.; Langer, R. Engineering precision nanoparticles for drug delivery. *Nature Reviews Drug Discovery* **2021**, *20*, 101–124.

(594) Lloyd, R. G.; Buckman, C. Identification and genetic analysis of sbcC mutations in commonly used recBC sbcB strains of Escherichia coli K-12. *Journal of Bacteriology* **1985**, *164*, 836–844.

(595) Gibson, F. P.; Leach, D. R.; Lloyd, R. G. Identification of sbcD mutations as cosuppressors of recBC that allow propagation of DNA palindromes in Escherichia coli K-12. *Journal of Bacteriology* **1992**, *174*, 1222–1228.

(596) Terrific Broth (TB) Medium. *Cold Spring Harbor Protocols* **2015**, *2015*, pdb.rec085894.

(597) SOB Medium. *Cold Spring Harbor Protocols* **2018**, *2018*, pdb.rec102723.

(598) Danquah, M. K.; Forde, G. M. Growth medium selection and its economic impact on plasmid DNA production. *Journal of Biomedicine and Biotechnology* **2007**, *104*, 490–497.

(599) Kram, K. E.; Finkel, S. E. Rich medium composition affects Escherichia coli survival, glycation, and mutation frequency during long-term batch culture. *Applied and Environmental Microbiology* **2015**, *81*, 4442–4450.

(600) Reddy, P.; McKenney, K. Improved method for the production of M13 phage and single-stranded DNA for DNA sequencing. *BioTechniques* **1996**, *20*, 854–860.

Appendix A

Media and buffers

All media and solutions were sterilized by autoclaving, unless otherwise specified.

Media recipes are listed in Table A.1, while solutions are listed in Table A.2. Buffer solutions for hydroxyapatite (HAP) chromatography are listed in Table A.3, as described by Andrews-Pfannkoch (2010) [571, 572]. As needed, media were supplemented with $5\,\text{mM}$ $MgSO_4$ to improve filamentous phage replication and/or with antibiotics for plasmid or phage maintenance. Antibiotics were used in the following concentrations, unless otherwise specified: ampicillin $100\,\text{µg/mL}$, kanamycin $25\,\text{µg/mL}$, tetracycline $10\,\text{µg/mL}$.

Table A.1: Media recipes

Name	Recipe (1 L)
LB broth	10 g tryptone, 5 g yeast extract, 10 g NaCl, in deionized water[*]
2X YT broth	16 g tryptone, 10 g yeast extract, 5 g NaCl in deionized water[*]
LB agar	liquid LB, 14 g agar, divide 25 mL per 100 mm petri dish
top agar	liquid LB, 7 g agar

[*] pH adjusted to 7

Table A.2: Solution recipes

Name	Recipe	pH
Tris-NaCl (TN) buffer	50 mM Tris-Cl, 150 mM NaCl in deionized water	7.5
TBS-T buffer	TN buffer, 0.1% (v/v) Tween 20	7.5
polyethylene glycol (PEG) solution	20% PEG-8000 (w/v), 2.5 M NaCl in deionized water	–
blocking buffer	5% nonfat bovine milk in TBS-T	–

Table A.3: Hydroxyapatite chromatography buffer recipes

Buffer	Recipe	pH
1 M Monosodium Phosphate Stock[1]	119.98 g sodium phosphate monobasic, anhydrous	–
1 M Disodium Phosphate Stock[1]	141.96 g sodium phosphate dibasic	–
1 M Phosphate Buffer Stock Solution	equal volumes of 1 M Mono- & Disodium Phosphate Stock	6.8
0.12M Phosphate Buffer[2]	30 mL 1 M Phosphate Buffer Stock, SDS & EDTA[3]	6.8
0.18M Phosphate Buffer[2]	45 mL 1 M Phosphate Stock, SDS, EDTA[3]	6.8
0.20M Phosphate Buffer[2]	50 mL 1 M Phosphate Stock, SDS, EDTA[3]	6.8
0.40M Phosphate Buffer[2]	100 mL 1 M Phosphate Stock, SDS, EDTA[3]	6.8
1.0M Phosphate Buffer[2]	225 mL 1 M Phosphate Stock, SDS, EDTA[3]	6.8

[1] Final volume: 1 L.   [2] Final volume: 250 mL.   [3] 2.5 mL 10% SDS, 5 mL 0.5 M EDTA.

Appendix B

Strains, plasmids, and phages

Bacterial strains, mammalian cell lines, plasmids, and phages used and/or constructed are summarized in each chapter. Here, all strains (Table B.1), cell lines (Table B.2), phages (Table B.3), and plasmids (Table B.4) across the entire work are summarized as a reference.

Table B.1: Bacterial strains

Strain	Genotype	Source
JM109	F′ traD36 proAB$^+$ lacI^q lacZΔM15/ Δ(lac-proAB) endA1 glnV44 thi-1 e14$^-$ recA1 gyrA96 relA1 hsdR17	NEB
DH5α	F$^-$ ϕ80 Δ(lacZYA-argF)U169 endA1 thi-1 recA1 gyrA96 relA1 hsdR17(r$_K^-$, m$_K^+$) phoA supE44	CGSC #12384
XL1-Blue	F′ traD36 proAB$^+$ lacI^q lacZΔM15 Tn10/ lac endA1 glnV44 thi-1 recA1 gyrA96 relA1 hsdR17	Agilent Technologies
Stbl4	F′ proAB$^+$ lacIq lacZΔM15 Tn10/ endA1 glnV44 thi-1 recA1 gyrA96 relA1 Δ(lac-proAB) mcrA Δ(mcrBC-hsdRMS-mrr) λ$^-$ gal	Invitrogen
ER2738	F′ zzf::Tn10(TcR) proAB$^+$ lacI^q lacZΔM15/ Δ(lac-proAB) thi-1 glnV44, Δ(hsdS-mcrB)5, fhuA2	NEB
NEB Turbo	F′ traD36 proAB$^+$ lacI^q lacZΔM15/ Δ(lac-proAB) glnV44 thi-1 galE15 galK16 R(zgb-210::Tn10)TetS endA1 fhuA2 Δ(mcrB-hsdSM)5(r$_K^-$, m$_K^-$)	NEB

NEB: New England BioLabs   CGSC: Coli Genetics Stock Center

Table B.2: Mammalian cell lines

Cell line	Description	Source
HEK-293T	embryonic kidney epithelial	gift, Dr. M. Aucoin
HeLa	uterus, cervix adenocarcinoma	gift, Serenity Bioworks
MRC-5	lung fibroblast	ATCC CCL-171
HT-29	colon epithelial adenocarcinoma	gift, Dr. J. Blay
A549	lung epithelial carcinoma	ATCC CCL-185

ATCC: American Type Culture Collection

Table B.3: Phages

Phage	Genotype	Source
M13	wild-type	CGSC #13587
M13KO7	M13, p15a *ori*, Tn903 (Km^R)	NEB
M13KE	M13, KpnI and EagI target sites in gIII	NEB
M13SW7	M13KO7, gIII from M13KE	Chapter 4
M13SW7-EGF	M13SW7, *egf* from pPL451-gpD::egf	Chapter 4
M13SW8	M13KO7, PS removed, Km^R	Chapter 4
full-(gfp)	*cmv-gfp*, from precursor pSW9, Ap^R	Chapter 3
full$_{egf}$-(gfp)	*cmv-gfp*, from precursor pSW9, pIII::EGF, Ap^R	Chapter 4
mini-(gfp)	*cmv-gfp*, from precursor pM13ori2.cmvgfp	Chapter 3
mini$_{egf}$-(gfp)	*cmv-gfp*, from precursor pM13ori2.cmvgfp, pIII::EGF	Chapter 4
full-(luc)	from precursor pSW10, Ap^R	Chapter 3
full$_{egf}$-(luc)	from precursor pSW10, pIII::EGF, Ap^R	Chapter 4
mini-(luc)	from precursor pM13ori2.cmvluc	Chapter 3
mini$_{egf}$-(luc)	from precursor pM13ori2.cmvluc, pIII::EGF	Chapter 4
full-(gfp-tet)	from precursor pSW9-tet, Ap^R, Tc^R	Chapter 3
mini-(gfp-tet)	from precursor pM13ori2.cmvgfp-tet, Tc^R	Chapter 3
ds-mini-(gfp)	*cmv-gfp-cmv-gfp*, from precursor pM13ori2.cmvgfp.SAS (L)	Chapter 6

CGSC: Coli Genetics Stock Center   NEB: New England BioLabs

Table B.4: Plasmids

Plasmid	Genotype	Source
pGL2-SS-CMV-GFP-SS	pGL2-Promoter (GN: X65326.2) *cmv-gfp* replaced *SV40-luc*, Ap^R	gift, Mediphage Bioceuticals Inc.
pGL3-CMV	pGL3-Basic (GN: U47295.2) *cmv* inserted in BglII-HindIII, Ap^R	gift, Dr. N. Oviedo
pBluescript II KS+	wild-type f1 ori, pUC ori, Ap^R	[405]
pBR322	Ap^R, Tc^R	[406]
pSW9	pBluescript II KS+, *cmv-gfp* inserted in BamHI-EcoRI, ApR	Chapter 3
pM13ori	split f1 ori inserted in pUC57 (Genbank Accession No. Y14837.1), ApR	Chapter 3
pM13ori2	pM13ori, BsaI cut site replaced with SphI, ApR	Chapter 3
pM13ori2.cmvgfp	pM13ori2, *cmv-gfp* inserted in EcoRI-PacI	Chapter 3
pM13ori2.cmvluc	pM13ori2, *cmv-luc* inserted in EcoRI-KpnI	Chapter 3
pM13ori2.cmvgfp-tet	pM13ori2.cmvgfp, Tc^{R}from pBR322 inserted into KpnI	Chapter 3
pSW9-tet	pSW9, Tc^{R}from pBR322 inserted into SmaI	Chapter 3
pJET1.2/blunt	GN: EF694056.1	Thermo Fisher Scientific
pJET1.2-cmvgfp	pJET1.2, *cmv-gfp* inserted into EcoRV site, Ap^R	Chapter 6
pM13ori2.cmvgfp.SAS (L)	pM13ori2.cmvgfp, second *cmv-gfp* inserted in HindIII-BamHI	Chapter 6
pM13ori2.cmvgfp.SAS (NL)	pM13ori2.cmvgfp, second *cmv-gfp* inserted in HindIII-BamHI	Chapter 6
pM13ori2.cmvgfp.SAS (GG)	pM13ori2.cmvgfp, second *cmv-gfp* inserted via Golden Gate assembly	Chapter 6

GN: GenBank Accession No.

205

Appendix C

Supplementary qPCR data
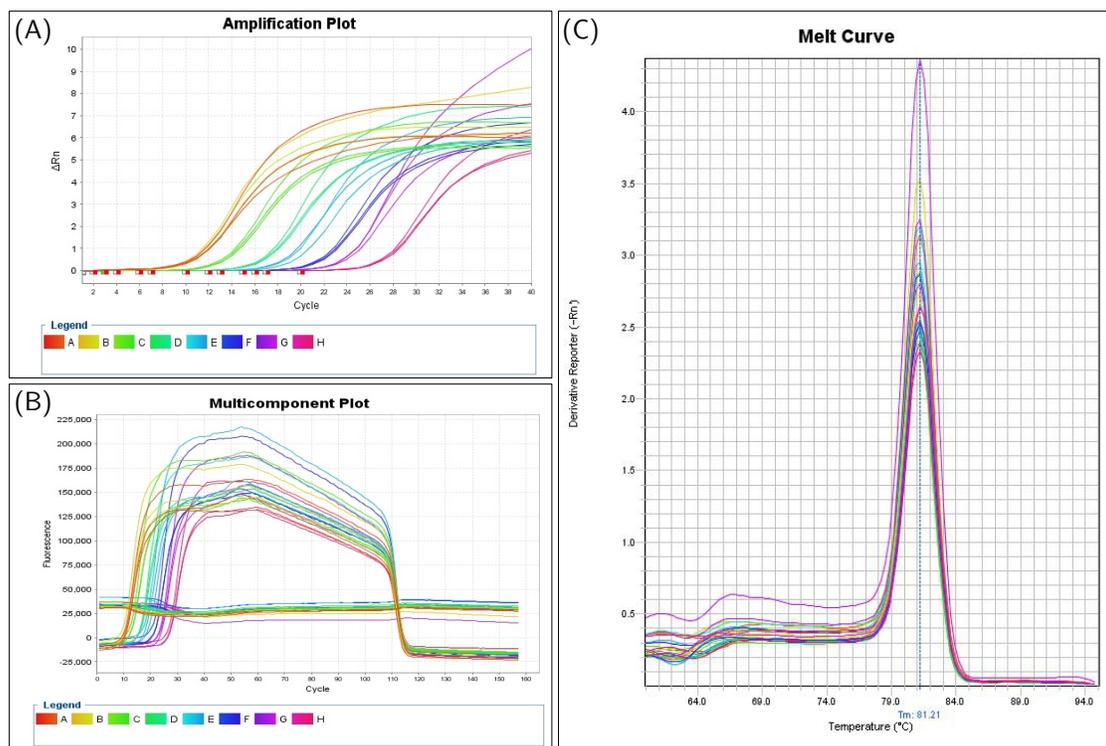
Figure C.1: Amplification plot, multicomponent plot, and melt curve for the gV amplicon.
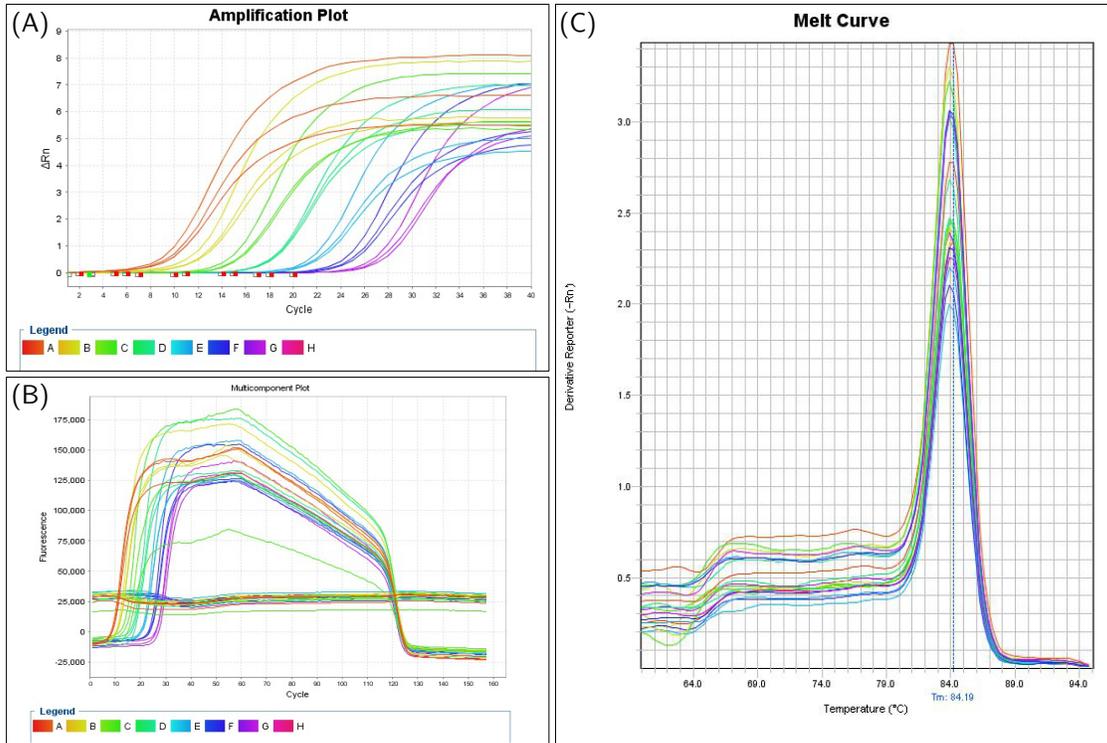
Figure C.2: Amplification plot, multicomponent plot, and melt curve for the *gfp* amplicon.
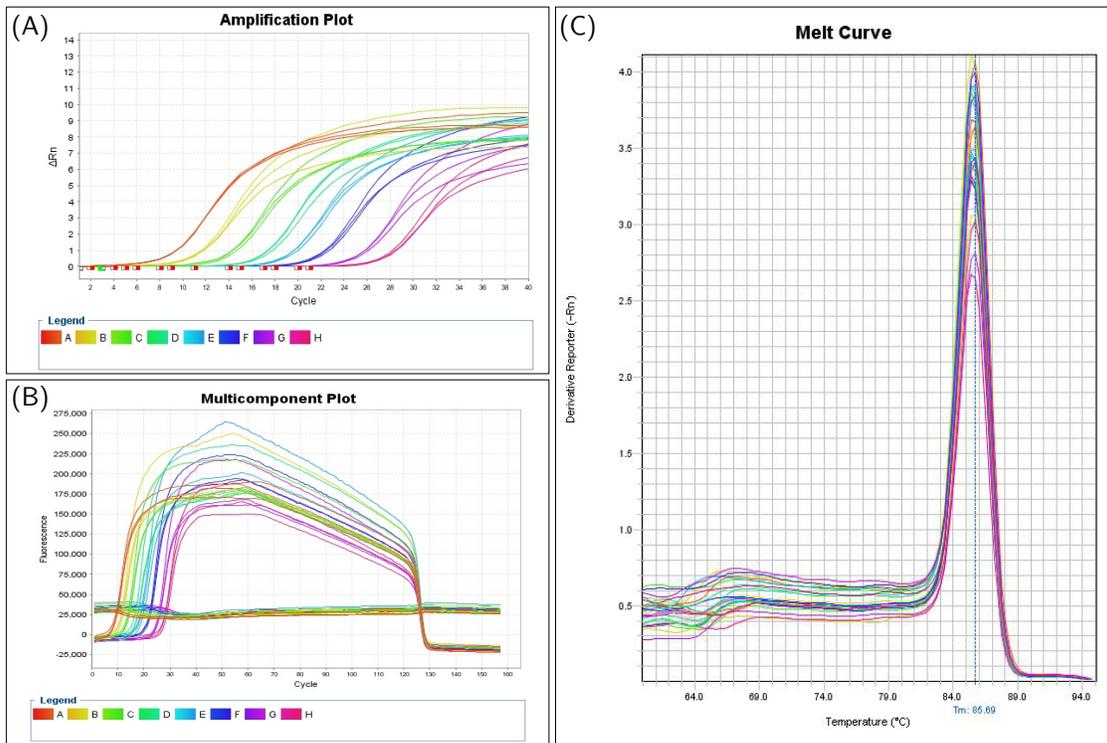
Figure C.3: Amplification plot, multicomponent plot, and melt curve for the *luc* amplicon.

Appendix D

Supplementary RNAfold data

Each miniphagemid sequence is followed by the sequence for secondary structure in dot-bracket notation as output from RNAFold [434]. An unpaired base is denoted with a dot and base pairs are denoted with opening and closing brackets.

Note that although the output sequences are RNA, the input sequences were DNA. RNAFold structure prediction was performed using DNA energy parameters, using the following command:

```
RNAfold -p -d2 --noLP --circ -P dna_parameters.par --noconv < sequence.fa
```

```
>centroid of phagemid mini-(gfp)
GGCCAUCGCCCUGAUAGACGGUUUUUCGCCCUUUGACGUUGGAGUCCACGUUCUUUAAUAGUGGACUCUUGUUCCAAACUGGAACAACACUCAACCC
UAUCUCGGGCUAUUCUUUUGAUUUAUAAGGGAUUUUGCCGAUUUCGGCCUAUUGGUUAAAAAAUGAGCUGAUUUAACAAAAAUUUAACGCGAAUUUU
AACAAAAUAUUAACGUUUACAAUUUAAAGAUAUCUCGAGCUUCGAAUUCCGUUACAUAACUUACGGUAAAUGGCCCGCCUGGCUGACCGCCCAACGA
CCCCCGCCCAUUGACGUCAAUAAUGACGUAUGUUCCCAUAGUAACGCCAAUAGGGACUUUCCAUUGACGUCAAUGGGUGGAGUAUUUACGGUAAACU
GCCCACUUGGCAGUACAUCAAGUGUAUCAUAUGCCAAGUACGCCCCCUAUUGACGUCAAUGACGGUAAAUGGCCCGCCUGGCAUUAUGCCCAGUACA
UGACCUUAUGGGACUUUCCUACUUGGCAGUACAUCUACGUAUUAGUCAUCGCUAUUACCAUGGUGAUGCGGUUUUGGCAGUACAUCAAUGGGCGUGG
AUAGCGGUUUGACUCACGGGGAUUUCCAAGUCUCCACCCCAUUGACGUCAAUGGGAGUUUGUUUUGGCACCAAAAUCAACGGGACUUUCCAAAAUGU
CGUAACAACUCCGCCCCAUUGACGCAAAUGGGCGGUAGGCGUGUACGGUGGGAGGUCUAUAUAAGCAGAGCUGGUUUAGUGAACCGUCAGAUCCGCU
AGCGCUACCGGUCGCCACCCUCGAGUCAUGAACUAGUGCGGCCGCCUGCAGGUGACCAGAUCUGAUAUCACUCGUACGCGUCUCGGCCGGCCACCAU
GGAGAGCGACGAGAGCGGCCUGCCCGCCAUGGAGAUCGAGUGCCGCAUCACCGGCACCCUGAACGGCGUGGAGUUCGAGCUGGUGGGCGGCGGAGAG
GGCACCCCCGAGCAGGGCCGCAUGACCAACAAGAUGAAGAGCACCAAAGGCGCCCUGACCUUCAGCCCCUACCUGCUGAGCCACGUGAUGGGCUACG
GCUUCUACCACUUCGGCACCUACCCCAGCGGCUACGAGAACCCCUUCCUGCACGCCAUCAACAACGGCGGCUACACCAACACCCGCAUCGAGAAGUA
CGAGGACGGCGGCGUGCUGCACGUGAGCUUCAGCUACCGCUACGAGGCCGGCCGCGUGAUCGGCGACUUCAAGGUGAUGGGCACCGGCUUCCCCGAG
GACAGCGUGAUCUUCACCGACAAGAUCAUCCGCAGCAACGCCACCGUGGAGCACCUGCACCCCAUGGGCGAUAACGAUCUGGAUGGCAGCUUCACCC
GCACCUUCAGCCUGCGCGACGGCGGCUACUACAGCUCCGUGGUGGACAGCCACAUGCACUUCAAGAGCGCCAUCCACCCCAGCAUCCUGCAGAACGG
GGGCCCCAUGUUCGCCUUCCGCCGCGUGGAGGAGGAUCACAGCAACACCGAGCUGGGCAUCGUGGAGUACCAGCACGCCUUCAAGACCCCGGAUGCA
GAUGCCGGUGAAGAAAGAGUUUAAACGGUAAAUAUAAAAUUUUUAAGUGUAUAAUGUGUUAAACUACUGAUUCUAAUUGUUUGUGUAUUUUAGAUGG
CCGGCUGAUCAUAAUCAGCCAUACCACAUUUGUAGAGGUUUUACUUGCUUUAAAAAACCUCCCACACCUCCCCCUGAACCUGAAACAUAAAAUGAAU
GCAAUUGUUGUUGUUAACUUGUUUAUUGCAGCUUAUAAUGGUUACAAAUAAAGCAAUAGCAUCACAAAUUUCACAAAUAAAGCAUUUUUUUCACUGC
AUUCUAGUUGUGGUUUGUCCAAACUCAUCAAUGUAUCUUACGAUCGCCCUUCCCAACAGUUGCGCAAUCCCGCCCCUAACUCCGCCCAGUUCCGCCC
AUUCUCCGCCCCAUGGCUGACUAAUUUUUUUUAUUUUAUGCAGAGGCCGAGGCCGCCUCGGCCUCUGAGCUAUUCCAGAAGUAGUGAGGAGGCUUUUU
UGGAGGCCGGAGCGGAGAAUGGGCGGAACUGGGCGGAGUUAGGGGCGGGAUGGGCGGAGUUAGGGGCGGGACUAUGGUUGCUGACUAAUUGAGAUGC
AUGCUUUGCAUACUUCUGCCUGCUGGGGAGCCUGGGGACUUUCCACACCUGGUUGCUGACUAAUUGAGAUGCAUGCUUUGCAUACUUCUGCCUGCUG
GGGAGCCUGGGGACUUUCCACACCCUAACUGACACACAUUCCUUAAUUAAGUCGACAAGCUUGACGUCGGUACCGAGCUCGCCCGGGAUCCGUACGC
GCCCGUAGCGGCGCAUUAAGCGCGGCGGGUGUGGUGGUUACGCGCAGCGUGACCGCUACACUUGCCAGCGCCCUAGCGCCCGCUCCUUUCGCUUUC
UUCCCUUCCUUUCUCGCCACGUUCGCCGGCUUUCCCCGUCAAGCUCUAAAUCGGGGGCUCCCUUUAGGGUUCCGAUUUAGUGCUUUACGGCACCUCG
ACCCCAAAAAACUUGAUUUGGGGUGAUGGUUCACGUAGUG
```

```
(((((((((((.(((....((((((...............(((((((..........))))))).((((((...))))))...........
...................................................................................(..........
............).................................................................................
.((((((((((((((....))))))...............................(((((((((((((((.((.((((.................
...(((((((((((...)))))))....)))))).))))).))..))))))))))))))))..((((.....((.....))....)))).((((.((.(
(...((((.....)))...))))))))......(((((.(((((((((.(((((...)))))))))...))))))))).)))))))))).....((
((..(((((....(((((((....)))))))..(((((((...)))))).((..((((((...))))))..))..(((((((((.....((((..
...(((((....(((.....)))...))).))......))))))))))))...............................))))).))))..)))
....................................................................(((((.(((.((((....(.....)...)))))
)))))))...........................(((((.....)))))...................(((((.(((.............
.......................................)))))))))..............................................
..............................................................................................
.......................................................................((((......(((.(((.(((
(((((......))))))))...(((((.......(((((((....))..(((((...(((.........(((((((.((((...........
......(.((((.((((.....)))))))).)........................))))))))))((((((....((((...(((((......
....(((((((...))))))......((((.......)))).(((....(((....))).....))).......)))))..))))).....(((((
((((..(((((.....................))))))....................(((((((.........((..((((((((...
.))))))))...))......((.((((((((.............))))))).))..........................................
.............((.((((.......))))).).........................)))))))......)))))))))........((...((
.....))..)).......................(((.(((((((..((((.(.((((((((((((((((((((((((((((((((((((((.((..
((((.................((((((((((((...))))))))))))......(((((((.(((.....)))))))))))))))).)))
)))))))))))))))))))))))))))))))))).)))).))).)))).)))...........................(((((.....)))))....
......))))))).))).)))))....)))))......)))))......(((((((((.....))))).)))))))))......))))(((((....
......))))).........................................................(((......((((((((..
...((((((((((((((((((((...))))))))))))))))))))).)))).....)))).))))....)))........................
...............(((((.(((((.((((((((((.((((.....))))))))))))))).)))).)))).............)))))))..))).
)))))))))))((((...)))
```

>centroid of phagemid mini-(luc)
GGCCAUCGCCCUGAUAGACGGUUUUUUCGCCCUUUGACGUUGGAGUCCACGUUCUUUAAUAGUGGACUCUUGUUCCAAACUGGAACAACACUCAACCC
UAUCUCGGGCUAUUCUUUUGAUUUAUAAGGGAUUUUUGCCGAUUUCGGCCUAUUGGUUAAAAAAUGAGCUGAUUUAACAAAAAUUUAACGCGAAUUUU
AACAAAAUAUUAACGUUUACAAUUUAAAGAUAUCUCGAGCUUCGAAUUCCCAGAUAUACGCGUUGACAUUGAUUUAUUGACUAGUUAUUAAUAGUAAU
CAAUUACGGGGUCAUUAGUUCAUAGCCCAUAUAUGGAGUUCCGCGUUACAUAACUUACGGUAAAUGGCCCGCCUGGCUGACCGCCCAACGACCCCCG
CCCAUUGACGUCAAUAAUGACGUAUGUUCCCAUAGUAACGCCAAUAGGGACUUUCCAUUGACGUCAAUGGGUGGACUAUUUACGGUAAACUGCCCAC
UUGGCAGUACAUCAAGUGUAUCAUAUGCCAAGUACGCCCCCUAUUGACGUCAAUGACGGUAAAUGGCCCGCCUGGCAUUAUGCCCAGUACAUGACCU
UAUGGGACUUUCCUACUUGGCAGUACAUCUACGUAUUAGUCAUCGCUAUUACCAUGGUGAUGCGGUUUUGGCAGUACAUCAAUGGGCGUGGAUAGCG
GUUUGACUCACGGGGAUUUCCAAGUCUCCACCCCAUUGACGUCAAUGGGAGUUUGUUUUGGCACCAAAAUCAACGGGACUUUCCAAAAUGUCGUAAC
AACUCCGCCCCAUUGACGCAAAUGGGCGGUAGGCGUGUACGGUGGGAGGUCUAUAUAAGCAGAGCUCUCUGGCUAACUAGAGAACCCACUGCUUACU
GGCUUAUCGAAAUUAAUACGACUCACUAUAGGGAGACCCAAGCUUGGCAUUCCGGUACUGUUGGUAAAGCCACCAUGGAAGACGCCAAAAACAUAAA
GAAAGGCCCGGCGCCAUUCUAUCCGCUGGAAGAUGGAACCGCUGGAGAGCAACUGCAUAAGGCUAUGAAGAGAUACGCCCUGGUUCCUGGAACAAUU
GCUUUUACAGAUGCACAUAUCGAGGUGGACAUCACUUACGCUGAGUACUUCGAAAUGUCCGUUCGGUUGGCAGAAGCUAUGAAACGAUAUGGGCUGA
AUACAAAUCACAGAAUCGUCGUAUGCAGUGAAAACUCUCUUCAAUUCUUUAUGCCGGUGUUGGGCGCGUUAUUUAUCGGAGUUGCAGUUGCGCCCGC
GAACGACAUUUAUAAUGAACGUGAAUUGCUCAACAGUAUGGGCAUUUCGCAGCCUACCGUGGUGUUCGUUUCCAAAAAGGGGUUGCAAAAAAUUUUG
AACGUGCAAAAAAAGCUCCCAAUCAUCCAAAAAAUUAUUAUCAUGGAUUCUAAAACGGAUUACCAGGGAUUUCAGUCGAUGUACACGUUCGUCACAU
CUCAUCUACCUCCCGGUUUUAAUGAAUACGAUUUUGUGCCAGAGUCCUUCGAUAGGGACAAGACAAUUGCACUGAUCAUGAACUCCUCUGGAUCUAC
UGGUCUGCCUAAAGGUGUCGCUCUGCCUCAUAGAACUGCCUGCGUGAGAUUCUCGCAUGCCAGAGAUCCUAUUUUUGGCAAUCAAAUCAUUCCGGAU
ACUGCGAUUUUAAGUGUUGUUCCAUUCCAUCACGGUUUUGGAAUGUUUACUACACUCGGAUAUUUGAUAUGUGGAUUUCGAGUCGUCUUAAUGUAUA
GAUUUGAAGAAGAGCUGUUUCUGAGGAGCCUUCAGGAUUACAAGAUUCAAAGUGCGCUGCUGGUGCCAACCCUAUUCUCCUUCUUCGCCAAAAGCAC
UCUGAUUGACAAAUACGAUUUAUCUAAUUUACACGAAAUUGCUUCUGGUGGCGCUCCCCUCUCUAAAGGAAGUCGGGGAAGCGGUUGCCAAGAGGUUC
CAUCUGCCAGGUAUCAGGCAAGGAUAUGGGCUCACUGAGACUACAUCAGCUAUUCUGAUUACACCCGAGGGGGAUGAUAAACCGGGCGCGGUCGGUA
AAGUUGUUCCAUUUUUUGAAGCGAAGGUUGUGGAUCUGGAUACCGGGAAAACGCUGGGCGUUAAUCAAAGAGGCGAACUGUGUGUGAGAGGUCCUAU
GAUUAUGUCCGGUUAUGUAAACAAUCCGGAAGCGACCAACGCCUUGAUUGACAAGGAUGGAUGGCUACAUUCGGAGACAUAGCUUACUGGGACGAA
GACGAACACUUCUUCAUCGUUGACCGCCUGAAGUCUCUGAUUAAGUACAAAGGCUAUCAGGUGGCUCCCGCUGAAUUGGAAUCCAUCUUGCUCCAAC
ACCCCAACAUCUUCGACGCAGGUGUCGCAGGUCUUCCCGACGAUGACGCCGGUGAACUUCCCGCCGCCGUUGUUGUUUUGGAGCACGGAAAGACGAU
GACGGAAAAAGAGAUCGUGGAUUACGUCGCCAGUCAAGUAACAACCGCGAAAAGUUGCGCGGAGGAGUUGUGUUUGUGGACGAAGUACCGAAAGGU
CUUACCGGAAAACUCGACGCAAGAAAAAUCAGAGAGAUCCUCAUAAAGGCCAAGAAGGGCGGAAAGAUCGCCGUGUAAUUCUAGAGUCGGGGCGGCC
GGCCGCUUCGAGCAGACAUGAUAAGAUACAUUGAUGAGUUUGGACAAACCACAACUAGAAUGCAGUGAAAAAAAUGCUUUAUUUGUGAAAUUUGUGA
UGCUAUUGCUUUAUUUGUAACCAUUAUAAGCUGCAAUAAACAAGUUAACAACAACAAUUGCAUUCAUUUUAUGUUUCAGGUUCAGGGGGAGGUGUGG
GAGGUUUUUUAAAGCAAGUAAAACCUCUACAAAUGUGGUAAAAUCGAUAAGGAUCUGCAGGGUACCGAGCUCGCCCGGGAUCCGUACGCGCCCUGUA
```

```
GCGGCGCAUUAAGCGCGGCGGGUGUGGUGGUUACGCGCAGCGUGACCGCUACACUUGCCAGCGCCCUAGCGCCCGCUCCUUUCGCUUUCUUCCCUUC
CUUUCUCGCCACGUUCGCCGGCUUUCCCCGUCAAGCUCUAAAUCGGGGGCUCCCUUUAGGGUUCCGAUUUAGUGCUUUACGGCACCUCGACCCCAAA
AAACUUGAUUUGGGUGAUGGUUCACGUAGUG
((((((((((.........(((((...............(((((((...........)))))))(((((((...)))))))...........
...............................(((((...))))).......((.....................))............(((((......
..........................(((((.(((...)))........))))).))))))...(((((((.....(((.......))).))))))
))............(((((((((......((.(((.((((.((.(((.....(((((((((((..(((.(((.((((.((((.((....(((((
((((..(((((....))))).(((....)))....................................................................((((
((((((((....)))))....))))))))))..................................((.....))....(((((.((((((((.....)
))(((...)))))))).))))..............................................................................
....................................................((((((....))))).))(((.((............))....
)))..........))))))))).........))).)))).))))....(((((((...(((((((....))))))...))))))...(((
(.(((....)))........)))))(((((..((((.........((.....))....)))).....)))))....))))
))).))))))))).)).))...)))).)).))..))............)))).))...))))))))).....(((((.....((.....
..((((.(((.....................(((((..(((((..(((((...............((((.((........
...))..))))......))))).....(((((.......................................))))))).)))))).))
)))).............(((.((((((.....))))))).)))...............................................
.......................(((((.....))))).......................................................
.............................(((((((.(((((...............((..((((((.(((.((((((((((.......
..............................((((.((...))).))))............((((....((((........
.(((((.((..(((((((.........)))))))...)).)))))))....)))).............................
.........))))))..))))).)))...............))))))..))...................................
..............................(((.((((.............)))).)))...............(.....
...)...............................(((((....))))
.............)))).))))........))........))))))....((((.(((((((.(((.((((((...............(((((....
.........))))).........(((((.....)))))....((.((.((...)).)).))...((((((((((((...((((.......(((
((((...((((((((((((((.((.........)).))))).))))))))...(((..........(((((((.(.....(((......(((
((....)))))..........((((((((........))))).))).)))).)).).))))))))))....)))))))...........))
))......))))).)))))).))).....(((((((.(.....).))))).....)...................(((....)))............
................................................................................(((((...)))).....
..................((.....)).....(((((((...............................(((((.....(.........
..............)....))))...........)))))))...............................((((((((...............
)))))))).................(((.(((((....(((((.................................))))))).(((((((.....(((((((
((((((((((((((((...)))))))))))))))))).))))....)))).))).))))))))..............................)))))))))))))).
.)))))..((((.((((.(((((((((((...........)))))))))))..))))).))))..............))))).......))))))))))
))(((...)))
```

>centroid of phagemid mini-(gfp-tet)
GGCCAUCGCCCUGAUAGACGGUUUUUCGCCCUUUGACGUUGGAGUCCACGUUCUUUAAUAGUGGACUCUUGUUCCAAACUGGAACAACACUCAACCC
UAUCUCGGGCUAUUCUUUUGAUUUAUAAGGGAUUUUGCCGAUUUCGGCCUAUUGGUUAAAAAAUGAGCUGAUUUAACAAAAAUUUAACGCGAAUUUU
AACAAAAUAUUAACGUUUACAAUUUAAAGAUAUCUCGAGCUUCGAAUUCCGUUACAUAACUUACGGUAAAUGGCCCGCCUGGCUGACCGCCCAACGA
CCCCCGCCCAUUGACGUCAAUAAUGACGUAUGUUCCCAUAGUAACGCCAAUAGGGACUUUCCAUUGACGUCAAUGGGUGGAGUAUUUACGGUAAACU
GCCCACUUGGCAGUACAUCAAGUGUAUCAUAUGCCAAGUACGCCCCCUAUUGACGUCAAUGACGGUAAAUGGCCCGCCUGGCAUUAUGCCCAGUACA
UGACCUUAUGGGACUUUCCUACUUGGCAGUACAUCUACGUAUUAGUCAUCGCUAUUACCAUGGUGAUGCGGUUUUGGCAGUACAUCAAUGGGCGUGG
AUAGCGGUUUGACUCACGGGGAUUUCCAAGUCUCCACCCCAUUGACGUCAAUGGGAGUUUGUUUUGGCACCAAAAUCAACGGGACUUUCCAAAAUGU
CGUAACAACUCCGCCCCAUUGACGCAAAUGGGCGGUAGGCGUGUACGGUGGGAGGUCUAUAUAAGCAGAGCUGGUUUAGUGAACCGUCAGAUCCGCU
AGCGCUACCGGUCGCCACCCUCGAGUCAUGAACUAGUGCGGCCGCCUGCAGGUGACCAGAUCUGAUAUCACUCGUACGCGUCUCGGCCGGCCACCAU
GGAGAGCGACGAGAGCGGCCUGCCCGCCAUGGAGAUCGAGUGCCGCAUCACCGGCACCCUGAACGGCGUGGAGUUCGAGCUGGUGGGCGGCGGAGAG
GGCACCCCCGAGCAGGGCCGCAUGACCAACAAGAUGAAGAGCACCAAAGGCGCCCUGACCUUCAGCCCCUACCUGCUGAGCCACGUGAUGGGCUACG
GCUUCUACCACUUCGGCACCUACCCCAGCGGCUACGAGAACCCCUUCCUGCACGCCAUCAACAACGGCGGCUACACCAACACCCGCAUCGAGAAGUA
CGAGGACGGCGGCGUGCUGCACGUGAGCUUCAGCUACCGCUACGAGGCCGGCCGCGUGAUCGGCGACUUCAAGGUGAUGGGCACCGGCUUCCCCGAG
GACAGCGUGAUCUUCACCGACAAGAUCAUCCGCAGCAACGCCACCGUGGAGCACCUGCACCCCAUGGGCGAUAACGAUCUGGAUGGCAGCUUCACCC
GCACCUUCAGCCUGCGCGACGGCGGCUACUACAGCUCCGUGGUGGACAGCCACAUGCACUUCAAGAGCGCCAUCCACCCCAGCAUCCUGCAGAACGG
GGGCCCCAUGUUCGCCUUCCGCCGCGUGGAGGAGGAUCACAGCAACACCGAGCUGGGCAUCGUGGAGUACCAGCACGCCUUCAAGACCCCGGAUGCA
GAUGCCGGUGAAGAAAGAGUUUAAACGGUAAAUAUAAAAUUUUUAAGUGUAUAAUGUGUUAAACUACUGAUUCUAAUUGUUUGUGUAUUUUAGAUGG
CCGGCUGAUCAUAAUCAGCCAUACCACAUUUGUAGAGGUUUUACUUGCUUUAAAAAACCUCCCACACCUCCCCCUGAACCUGAAACAUAAAAUGAAU
GCAAUUGUUGUUGUUAACUUGUUUAUUGCAGCUUAUAAUGGUUACAAAUAAAGCAAUAGCAUCACAAAUUUCACAAAUAAAGCAUUUUUUUCACUGC
AUUCUAGUUGUGGUUUGUCCAAACUCAUCAAUGUAUCUUACGAUCGCCCUUCCCAACAGUUGCGCAAUCCCGCCCCUAACUCCGCCCAGUUCCGCCC
AUUCUCCGCCCCAUGGCUGACUAAUUUUUUUUAUUUAUGCAGAGGCCGAGGCCGCCUCGGCCUCUGAGCUAUUCCAGAAGUAGUGAGGAGGCUUUUU

UGGAGGCCGGAGCGGAGAAUGGGCGGAACUGGGCGGAGUUAGGGGCGGGAUGGGCGGAGUUAGGGGCGGGACUAUGGUUGCUGACUAAUUGAGAUGC
AUGCUUUGCAUACUUCUGCCUGCUGGGGAGCCUGGGGACUUUCCACACCUGGUUGCUGACUAAUUGAGAUGCAUGCUUUGCAUACUUCUGCCUGCUG
GGGAGCCUGGGGACUUUCCACACCUAACUGACACACAUUCCUUAAUUAAGUCGACAAGCUUGACGUCGGUACCAAAUAGGCGUAUCACGAGGCCCU
UUCGUCUUCAAGAAUUCUCAUGUUUGACAGCUUAUCAUCGAUAAGCUUUAAAUGCGGUAGUUUAUCACAGUUAAAUUGCUAACGCAGUCAGGCACCGU
GUAUGAAAUCUAACAAUGCGCUCAUCGUCAUCCUCGGCACCGUCACCCUGGAUGCUGUAGGCAUAGGCUUGGUUAUGCCGGUACUGCCGGGCCUCUU
GCGGGAUAUCGUCCAUUCCGACAGCAUCGCCAGUCACUAUGGCGUGCUGCUAGCGCUAUAUGCGUUGAUGCAAUUUCUAUGCGCACCCGUUCUCGGA
GCACUGUCCGACCGCUUUGGCCGCCGCCCAGUCCUGCUCGCUUCGCUACUUGGAGCCACUAUCGACUACGCGAUCAUGGCGACCACACCCGUCCUGU
GGAUCCUCUACGCCGGACGCAUCGUGGCCGGCAUCACCGGCGCCACAGGUGCGGUUGCUGGCGCCUAUAUCGCCGACAUCACCGAUGGGGAAGAUCG
GGCUCGCCACUUCGGGCUCAUGAGCGCUUGUUUCGGCGUGGGUAUGGUGGCAGGCCCCGUGGCCGGGGGACUGUUGGGCGCCAUCUCCUUGCAUGCA
CCAUUCCUUGCGGCGGCGGUGCUCAACGGCCUCAACCUACUACUGGGGCUGCUUCCUAAUGCAGGAGUCGCAUAAGGGAGAGCGUCGACCGAUGCCCU
UGAGAGCCUUCAACCCAGUCAGCUCCUUCCGGUGGGCGCGGGGCAUGACUAUCGUCGCCGCACUUAUGACUGUCUUCUUUAUCAUGCAACUCGUAGG
ACAGGUGCCGGCAGCGCUCUGGGUCAUUUUCGGCGAGGACCGCUUUCGCUGGAGCGCGACGAUGAUCGGCCUGUCGCUUGCGGUAUUCGGAAUCUUG
CACGCCCUCGCUCAAGCCUUCGUCACUGGUCCCGCCCACCAAACGUUUCGGCGAGAAGCAGGCCAUUAUCGCCGGCAUGGCGGCCGACGCGCUGGGCU
ACGUCUUGCUGGCGUUCGCGACGCGAGGCUGGAUGGCCUUCCCCAUUAUGAUUCUUCUCGCUUCCGGCGGCAUCGGGAUGCCCGCGUUGCAGGCCAU
GCUGUCCAGGCAGGUAGAUGACGACCAUCAGGGACAGCUUCAAGGAUCGCUCGCGGCUCUUACCAGCCUAACUUCGAUCACUGGACCGCUGAUCGUC
ACGGCGAUUUAUGCCGCCUCGGCGAGCACAUGGAACGGGUUGGCAUGGAUUGUAGGCGCCGCCCUAUACCUUGUCUGCCUCCCCGCGUUGCGUCGCG
GUGCAUGGAGCCGGGCCACCUCGACCUGAAUGGAAGCCGGCGGCACCUCGCUAACGGAUUCACCACUCCAAGAAUUGGAGCCAAUCAAUUCUUGCGG
AGAACUGUGAAUGCGCAAACCAACCCUUGGCAGAACAUAUCCAUCGCGUCCGCCAUCUCCAGCAGCCGCACGCGGCGCAUCUCGGGCAGCGUUGGGU
CCUGGCCACGGGUGCGCAUGAUCGUGCUCCUGUCGUUGAGGACCCGGCUAGGCUGGCGGGGUUGCCUUACUGGUUAGCAGAAUGAAUCACCGAUACG
CGAGCGAACGUGAAGCGACUGCUGCUGCAAAACGUCUGCGACCUGAGCAACAACAUGAAUGGUCUUCGGUUUCCGUGUUUCGUAAAGUCUGGAAACG
CGGAAGUCAGCGCCCUGCACCAUUAUGUUCCGGAUCUGCAUCGCAGGAUGCUGCUGGCUACCCUGUGGAACACCUACAUCGUAUUAACGAAGCGCU
GGCAUUGACCCUGAGUGAUUUUUCUCUGGUCCCGCCGCAUCCAUACCGCCAGUUGUUUACCCUCACAACGUUCCAGUAACCGGGCAUGUUCAUCAUC
AGUAACGGUACCGAGCUCGCCCGGGAUCCGUACGCGCCCUGUAGCGGCGCAUUAAGCGCGGCGGGUGUGGUGGUUACGCGCAGCGUGACCGCUACAC
UUGCCAGCGCCCUAGCGCCCGCUCCUUUCGCUUUCUUCCCUUCCUUUCUCGCCACGUUCGCCGGCUUUCCCCGUCAAGCUCUAAAUCGGGGGCUCCC
UUUAGGGUUCCGAUUUAGUGCUUUACGGCACCUCGACCCCAAAAAACUUGAUUUGGGUGAUGGUUCACGUAGUG
((((((((((.....(((((.................(((((((...........)))))))((((((((...)))))))((((.((((..
....(((((.......................((((....))))......(((((..........))))))........(((.........
...........)))...............((((...))))...(((((..........))))....))))))....(((((((.((((.((..(
(((((((((.(((((....)))))).................((((...........)))).((((((((.(((((((..(((((..(((((((.
.((((((((((((.....)))))....))))))))))..)).(((.(((((((((((..(((.........................((.(((...
...)))(((...))).))...................(((((((((((...(....(((((............)))))........).))))
))))..)))))...................))).))))))))))))))))))).(((((....(...((.....))...))))))))...))))
)..))))))).))))))))..)))))))))..))...........))).)))).)))).......))).......))))).))))).)))))).........((((
((.((((.((((..((((((((((((...((((((((...
.........((((.....))))............(((((......))))).....(((((((((..(((((.((((((((((((((((.....
((((.((((.(((.(((((((((..............................(((((.............))))(((.(((((..............
.................(((((((((((...
....((....))......(
((((((.......))))))))...................((.....)).((((......(((....)))(((((((((.(((.((((((((.....
...))))..((((..(((.....))).))))))))).......................)))))))))))((((((..................
.........................((((((..((...(((
((.....(((((((((.......................)))..)))))))..............................(((..(((((((..
..)))))))...))).......((((((((..........)))))))).....................................(((.......
....................)))...))))))..))....((((.((
(.....)))))))).....................................(((....((.(.((((((((((((((((((((((((((((((((((((....
.....................(((((((((((...)))))))))))...))
)))))))))))))))))))))))))))))))))))))).))...))).................)))))).........(((((.....)))))...
.......)))))))).....))))...............................(((((.....)))))....))).))))))))........
....))))))).)))................................
............((((((((...)))))))....(((.(((.((.........)).....)))))))..........................
....))))))))..))).)))))))))))..(((((...(((......
)))(((((.(((((......)))))))))))...(((((.(((((((...))))....)))))))).)))).....(((((.....)))))(((.....
)))))))))))).)).)))))..............((((..((((.....))))..))))..(((((......)))))...))))))))).......
................(((((.((((.((((((((..(((((((((((..(.((((((...............................
..............))))))....)....(((((..(((((...)))).)))))..))))))))......))))))))))........)))))).))))
)..............................(((.(((((.......))))).)))...
.........................(((....)))))))))))).................))))..)))))).)))...))))))))))).))))))....
....
....

```
.........................................................(((((((((((((..((.(((..(((((..............(((((((
.........((((.((.(((((((.....(((((......))))).....))))).)).)).. .))))).))))))(((((......))))).........
.............................................................(((((......)))))..........................
...........................((.((((((..(((((((((((((.......))))))))..)))).....)))))).))..................
...................................))))))))).)))))).))))))))).......(((((((((....(((((.(((........))).
)))))....)))))))))........................................................................................
...........................(((..((((((......((((((((((((.........))))))))))))...........))))))).)))..
.............................................................................(((((..(((........))).)))
)....................(((((.......))))).......................((.....))....................................
..........(((((((((.....(((((((((((((((((((((((...)))))))))))))))))))))))).))))....)))).)))).........
.......................................(((((.((((.((((((((((((............))))))))))).)))).))))......
....................)))))))))))((((...)))))
```

>centroid of phage M13
```
AACGCUACUACUAUUAGUAGAAUUGAUGCCACCUUUUCAGCUCGCGCCCCAAAUGAAAAAUAUAGCUAAACAGGUUAUUGACCAUUUGCGAAAUGUAU
CUAAUGGUCAAACUAAAUCUACUCGUUCGCAGAAUUGGGAAUCAACUGUUACAUGGAAUGAAACUUCCAGACACCGUACUUUAGUUGCAUAUUUAAA
ACAUGUUGAGCUACAGCACCAGAUUCAGCAAUUAAGCUCUAAGCCAUCCGCAAAAAUGACCUCUUAUCAAAAGGAGCAAUUAAAGGUACUCUCUAAU
CCUGACCUGUUGGAGUUUGCUUCCGGUCUGGUUCGCUUUGAAGCUCGAAUUAAAACGCGAUAUUUGAAGUCUUUCGGGCUUCCUCUUAAUCUUUUUG
AUGCAAUCCGCUUUGCUUCUGACUAUAAUAGUCAGGGUAAAGACCUGAUUUUUGAUUUAUGGUCAUUCUCGUUUUCUGAACUGUUUAAAGCAUUUGA
GGGGGAUUCAAUGAAUAUUUAUGACGAUUCCGCAGUAUUGGACGCUAUCCAGUCUAAACAUUUUACUAUUACCCCCUCUGGCAAAACUUCUUUUUGCA
AAAGCCUCUCGCUAUUUUGGUUUUUUAUCGUCGUCUGGUAAACGAGGGUUAUGAUAGUGUUGCUCUUACUAUGCCUCGUAAUUCCUUUUGGCGUUAUG
UAUCUGCAUUAGUUGAAUGUGGUAUUCCUAAAUCUCAACUGAUGAAUCUUUCUACCUGUAAUAAUGUUGUUCCGUUAGUUCGUUUUAUUAACGUAGA
UUUUUCUUCCCAACGUCCUGACUGGUAUAAUGAGCCAGUUCUUAAAAUCGCAUAAGGUAAUUCACAAUGAUUAAAGUUGAAAUUAAACCAUCUCAAG
CCCAAUUUACUACUCGUUCUGGUGUUUCUCGUCAGGGCAAGCCUUAUUCACUGAAUGAGCAGCUUUGUUACGUUGAUUUGGGUAAUGAAUAUCCGGU
UCUUGUCAAGAUUACUCUUGAUGAAGGUCAGCCAGCCUAUGCGCCUGGUCUGUACACCGUUCAUCUGUCCUCUUUCAAAGUUGGUCAGUUCGGUUCC
CUUAUGAUUGACCGUCUGCGCCUCGUUCCGGCUAAGUAACAUGGAGCAGGUCGCGGAUUUCGACACAAUUUAUCAGGCGAUGAUACAAAUCUCCGUU
GUACUUUGUUUCGCGCUUGGUAUAAUCGCUGGGGGUCAAAGAUGAGUGUUUUAGUGUAUUCUUUCGCCUCUUUCGUUUUAGGUUGGUGCCUUCGUAG
UGGCAUUACGUAUUUUACCCGUUUAAUGGAAACUUCCUCAUGAAAAAGUCUUUAGUCCUCAAAGCCUCUGUAGCCGUUGCUACCCUCGUUCCGAUGC
UGUCUUUCGCUGCUGAGGGUGACGAUCCCGCAAAAGCGGCCUUUAACUCCCUGCAAGCCUCAGCGACCGAAUAUAUCGGUUAUGCGUGGGCGAUGGU
UGUUGUCAUUGUCGGCGCAACUAUCGGUAUCAAGCUGUUUAAGAAAUUCACCUCGAAAGCAAGCUGAUAAACCGAUACAAUUAAAGGCUCCUUUUGG
AGCCUUUUUUUUUGGAGAUUUUCAACAUGAAAAAAUUAUUAUUCGCAAUUCCUUUAGUUGUUCCUUUCUAUUCUCACUCCGCUGAAACUGUUGAAAG
UUGUUUAGCAAAACCCCAUACAGAAAAUUCAUUUACUAACGUCUGGAAAGACGACAAAACUUUAGAUCGUUACGCUAACUAUGAGGGUUGUCUGUGG
AAUGCUACAGGCGUUGUAGUUUGUACUGGUGACGAAACUCAGUGUUACGGUACAUGGGUUCCUAUUGGGCUUGCUAUCCCUGAAAAUGAGGGUGGUG
GCUCUGAGGGUGGCGGUUCUGAGGGUGGCGGUUCUGAGGGUGGCGGUACUAAACCUCCUGAGUACGGUGAUACACCUAUUCCGGGCUAUACUUAUAU
CAACCCUCUCGACGGCACUUAUCCGCCUGGUACUGAGCAAAACCCCGCUAAUCCUAAUCCUUCUCUUGAGGAGUCUCAGCCUCUUAAUACUUUCAUG
UUUCAGAAUAAUAGGUUCCGAAAUAGGCAGGGGGCAUUAACUGUUUAUACGGGCACUGUUACUCAAGGCACUGACCCCGUUAAAACUUAUUACCAGU
ACACUCCUGUAUCAUCAAAAGCCAUGUAUGACGCUUACUGGAACGGUAAAUUCAGAGACUGCGCUUUCCAUUCUGGCUUUAAUGAGGAUCCAUUCGU
UUGUGAAUAUCAAGGCCAAUCGUCUGACCUGCCUCAACCUCCUGUCAAUGCUGGCGGCGGCUCUGGUGGUGGUUCUGGUGGCGGCUCUGAGGGUGGU
GGCUCUGAGGGUGGCGGUUCUGAGGGUGGCGGCUCUGAGGGAGGCGGUUCCGGUGGUGGCUCUGGUUCCGGUGAUUUUGAUUAUGAAAAGAUGGCAA
ACGCUAAUAAGGGGGCUAUGACCGAAAAUGCCGAUGAAAACGCGCUACAGUCUGACGCUAAAGGCAAACUUGAUUCUGUCGCUACUGAUUACGGUGC
UGCUAUCGAUGGUUUCAUUGGUGACGUUUCCGGCCUUGCUAAUGGUAAUGGUGCUACUGGUGAUUUUGCUGGCUCUAAUUCCCAAAUGGCUCAAGUC
GGUGACGGUGAUAAUUCACCUUUAAUGAAUAAUUUCCGUCAAUAUUUACCUUCCCUCCCUCAAUCGGUUGAAUGUCGCCCUUUUGUCUUUAGCGCUG
GUAAACCAUAUGAAUUUUCUAUUGAUUGUGACAAAAUAAACUUAUUCCGUGGUGUCUUUGCGUUUCUUUUAUAUGUUGCCACCUUUAUGUAUGUAUU
UUCUACGUUUGCUAACAUACUGCGUAAUAAGGAGUCUUAAUCAUGCCAGUUCUUUUGGGUAUUCCGUUAUUAUUGCGUUUCCUCGGUUUCCUUCUGG
UAACUUUGUUCGGCUAUCUGCUUACUUUUCUUAAAAAGGGCUUCGGUAAGAUAGCUAUUGCUAUUUCAUUGUUUCUUGCUCUUAUUAUUGGGCUUAA
CUCAAUUCUUGUGGGUUAUCUCUCUGAUAUUAGCGCUCAAUUACCCUCUGACUUUGUUCAGGGUGUUCAGUUAAUUCUCCCGUCUAAUGCGCUUCCC
UGUUUUUAUGUUAUUCUCUCUGUAAAGGCUGCUAUUUUCAUUUUUGACGUUAAACAAAAAAUCGUUUCUUAUUUGGAUUGGGAUAAAUAAUAUGGCU
GUUUAUUUUGUAACUGGCAAAUUAGGCUCUGGAAAGACGCUCGUUAGCGUUGGUAAGAUUCAGGAUAAAAUUGUAGCUGGGUGCAAAAUAGCAACUA
AUCUUGAUUUAAGGCUUCAAAACCUCCCGCAAGUCGGGAGGUUCGCUAAAACGCCUCGCGUUCUUAGAAUACCGGAUAAGCCUUCUAUAUCUGAUUU
GCUUGCUAUUGGGCGCGGUAAUGAUUCCUACGAUGAAAAUAAAAACGGCUUGCUUGUUCUCGAUGAGUGCGGUACUUGGUUUAAUACCCGUUCUUGG
AAUGAUAAGGAAAGACAGCCGAUUAUUGAUUGGUUUCUACAUGCUCGUAAAUUAGGAUGGGAUAUUAUUUUUCUUGUUCAGGACUUAUCUAUUGUUG
AUAAACAGGCGCGUUCUGCAUUAGCUGAACAUGUUGUUUAUUGUCGUCGUCUGGACAGAAUUACUUUACCUUUUGUCGGUACUUUAUAUUCUCUUAU
UACUGGCUCGAAAAUGCCUCUGCCUAAAUUACAUGUUGGCGUUGUUAAAUAUGGCGAUUCUCAAUUAAGCCCUACUGUUGAGCGUUGGCUUUAUACU
GGUAAGAAUUUGUAUAACGCAUAUGAUACUAAACAGGCUUUUUCUAGUAAUUAUGAUUCCGGUGUUUAUUCUUAUUUAACGCCUUAUUUAUCACACG
GUCGGUAUUUCAAACCAUUAAAUUUAGGUCAGAAGAUGAAAUUAACUAAAAUAUAUUUGAAAAAGUUUUCUCGCGUUCUUUGUCUUGCGAUUGGAUU
UGCAUCAGCAUUUACAUAUAGUUAUAUAACCCAACCUAAGCCGGAGGUUAAAAAGGUAGUCUCUCAGACCUAUGAUUUUGAUAAAUUCACUAUUGAC
UCUUCUCAGCGUCUUAAUCUAAGCUAUCGCUAUGUUUUCAAGGAUUCUAAGGGAAAAUUAAUUAAUAGCGACGAUUUACAGAAGCAAGGUUAUUCAC
UCACAUAUAUUGAUUUAUGUACUGUUUCCAUUAAAAAAGGUAAUUCAAAUGAAAUUGUUAAAUGUAAUUAAUUUUGUUUUCUUGAUGUUUGUUUCAU
CAUCUUCUUUUGCUCAGGUAAUUGAAAUGAAUAAUUCGCCUCUGCGCGAUUUUGUAACUUGGUAUUCAAAGCAAUCAGGCGAAUCCGUUAUUGUUUC
```

213

```
UCCCGAUGUAAAAGGUACUGUUACUGUAUAUUCAUCUGACGUUAAACCUGAAAAUCUACGCAAUUUCUUUAUUUCUGUUUUACGUGCUAAUAAUUUU
GAUAUGGUUGGUUCAAUUCCUUCCAUAAUUCAGAAGUAUAAUCCAAACAAUCAGGAUUAUAUUGAUGAAUUGCCAUCAUCUGAUAAUCAGGAAUAUG
AUGAUAAUUCCGCUCCUUCUGGGUGGUUUCUUUGUUCCGCAAAAUGAUAAUGUUACUCAAACUUUUAAAAUUAAUAACGUUCGGGCAAAGGAUUUAAU
ACGAGUUGUCGAAUUGUUUGUAAAGUCUAAUACUUCUAAAUCCUCAAAUGUAUUAUCUAUUGACGGCUCUAAUCUAUUAGUUGUUAGUGCACCUAAA
GAUAUUUUAGAUAACCUUCCUCAAUUCCUUUCUACUGUUGAUUUGCCAACUGACCAGAUAUUGAUUGAGGGUUUGAUAUUUGAGGUUCAGCAAGGUG
AUGCUUUAGAUUUUUCAUUUGCUGCUGGCUCUCAGCGUGGCACUGUUGCAGGCGGUGUUAAUACUGACCGCCUCACCUCUGUUUUAUCUUCUGCUGG
UGGUUCGUUCGGUAUUUUUAAUGGCGAUGUUUUUAGGGCUAUCAGUUCGCGCAUUAAAGACUAAUAGCCAUUCAAAAAUAUUGUCUGUGCCACGUAUU
CUUACGCUUUCAGGUCAGAAGGGUUCUAUCUCUGUUGGCCAGAAUGUCCCUUUUAUUACUGGUCGUGUGACUGGUGAAUCUGCCAAUGUAAAUAAUC
CAUUUCAGACGAUUGAGCGUCAAAAUGUAGGUAUUUCCAUGAGCGUUUUUCCUGUUGCAAUGGCUGGCGGUAAUAUUGUUCUGGAUAUUACCAGCAA
GGCCGAUAGUUUUGAGUUCUUCUACUCAGGCAAGUGAUGUUAUUACUAAUCAAAGAAGUAUUGCUACAACGGUUAAUUUGCGUGAUGGACAGACUCUU
UUACUCGGUGGCCUCACUGAUUAUAAAAACACUUCUCAAGAUUCUGGCGUACCGUUCCUGUCUAAAAUCCCUUUAAUCGGCCUCCUGUUUAGCUCCC
GCUCUGAUUCCAACGAGGAAAGCACGUUAUACGUGCUCGUCAAAGCAACCAUAGUACGCGCCCUGUAGCGGCGCAUUAAGCGCGGCGGGUGUGGUGG
UUACGCGCAGCGUGACCGCUACACUUGCCAGCGCCCUAGCGCCCGCUCCUUUCGCUUUCUUCCCUUCCUUUCUCGCCACGUUCGCCGGCUUUCCCCG
UCAAGCUCUAAAUCGGGGGCUCCCUUUAGGGUUCCGAUUUAGUGCUUUACGGCACCUCGACCCCAAAAAACUUGAUUUGGGUGAUGGUUCACGUAGU
GGGCCAUCGCCCUGAUAGACGGUUUUUCGCCCUUUGACGUUGGAGUCCACGUUCUUUAAUAGUGGACUCUUGUUCCAAACUGGAACAACACUCAACC
CUAUCUCGGGCUAUUCUUUUGAUUUAUAAGGGAUUUUGCCGAUUUCGGCCUAUUGGUUAAAAAAUGAGCUGAUUUAACAAAAAUUUAACGCGAAUUU
UAACAAAAUAUUUAACGUUUACAAAUUUAAAUAUUUGCUUAUACAAUCUUCCUGUUUUUGGGGCUUUUCUGAUUAUCAACCGGGGUACAUAUGAUUGAC
AUGCUAGUUUUACGAUUACCGUUCAUCGAUUCUCUUGUUUGCUCCAGACUCUCAGGCAAUGACCUGAUAGCCUUUGUAGACCUCUCAAAAAUAGCUA
CCCUCUCCGGCAUGAAUUUAUCAGCUAGAACGGUUGAAUAUCAUAUUGAUGGUGAUUUGACUGUCUCCGGCCUUUCUCACCCUUUUGAAUCUUUACC
UACACAUUACUCAGGCAUUGCAUUUAAAAUAUAUGAGGGUUCUAAAAAUUUUUAUCCUUGCGUUGAAAUAAAGGCUUCUCCCGCAAAAGUAUUACAG
GGUCAUAAUGUUUUUUGGUACAACCGAUUUAGCUUUAUGCUCUGAGGCUUUAUUGCUUAAUUUUGCUAAUUCUUUGCCUUGCCUGUAUGAUUUAUUGG
AUGUU
....(((((....)))))).........................................(((((((.........(((((.((((.(((.((((((((.............
.))))))))).............(((.((....)).))).............(((((.....)))))..........................))
).))))).)))))(((.(((((((...................(((..((((...(((((((....(((((((((.((((...................
.............))))..))..((.(((.......(((((((((((.......................))))))))))).....(((.....)))
((((......))))..(((((((...)))))))......))).)).)))))))....))))))))...((.......))......(((((.(((((.
.................(((((((.........(((((....))))......................................(((((....))))))...(
((.....)))..........................(((.........(((((.......))))....)))........................
....((((((((((..((....))...)).))))))))...................................................
..............(((((.........)))))...........................................................
.............................................(((((...))))).........................(((((....(((((.(((.(
(((((...)))))))..)))((((((((((.........)))).))).)))))))))))..................................
.......(((((((((.(((((..........))))))))).)))))...........................................
..............................................................((((((((((((((((((.((.(((.......
.......((((((((((...........(((((....((........(((((.((((.....(((.(...))).(((((.......
.................((((...))))...........................(((((....))))..(((((.(((((((........)))))
))).))))).....((((((((((((..((((......................).))).)))..))))))).....(((((((((....)))))))))...
.....................................(((.........)))...................................................
.((.............................(((((...)))))............................................(((((.......))))
.................................((.(((.....))).)).....................))............)))))....))))).))))).
.
(((((.......(((.((.....(((((((((.(((....))).))................................................))))))....
)).))).....)))))...........................................................)).)))))((((.(((.....
..))).))))........))))))......))))))))....))..))))))))))).....))))))............................
.............................................................................(((((....)))))......
...))))))......))))).............)))))).))))).))).))))))).))....))).)))).................................
......................................................................................................
.................((((......(.....(((((((((((((..............(((((....))))..........(((.....))
).(((((((...........................................................(((((.(((.((((
(.......)))))).))).))))))....................................)))))))..))))))))))))).....(((((...(((((...
........................))))..))))).....))))....................................................
.............................................................(((.((....))).)))...................
....(((((...)))))................(((((...)))))...........................................(((((.
....))))..(((((......((((...(.........).))))...............................)))))....................
.............................................................................................
....................................................(((((((((..(((((..(((.((((((((((...(((
((((((....)))))))))))(((((((((((...))))).))))).....))...)))))))).).))).)))).....)))))))))..............
.(((.(((((((.(((((((((......((.(((..........((.(.............)..))...........................(((((..
...)))))......(((.((.......)).)))....................................................))).))..(((((.((....)).
```

```
))))..))).))))))).)))))))..)))...............................................(((((.((((((((
((((..(((..(.(((((.((((((.((((.........((((..((........)).)))))......(((((((.................
.......................................)))))))....))))).))))))).)).)))..........)....))).))))))))
).)))).((((((((.............................((((.........))))....)))).))))).))))))............
...............................................(((...)))................................((
((((((((((.............))))))....)))))))).................................................
.............................................................................................
...((((((((.............((((............))))....))))))))......................................
.......((((.((((..((((............)))).)))).))))................((((...............))))....(
((((.(((((((.......))))))))....(((((..((((((((((...))))....))))))..))))).....))))).............
..............................................................(((...)))...............((.....)
)........((((((((....................................................................(((((((......
...((((......))))................)))))))......)))))))).......((((....)
)))............(((((((.............)))))))........................(((..............
.....................................)))...............................................
.......................(((...))).....................................(((((..((((......)))).)))))...........
.......................((((..(((((((((((.......)))))))).)).....))).......(((((....)))))..........
...................................................((((((...))))))..............
.................................................(((....)))...............................(((((((...))))))))..
.........................((((((((.....(((((((((((((((((((((((...))))))))))))))))))))))).))))....)
))).)))).....................................((((.((((.(((((((((((.............))))))))
))).)))).)))(((....(((((............(((((((((((((...))))))))))))...(((.................
..((((((((...........)))))))))((((((((...))))))))..........)))...............................(((
(....))))......(((.................)))..............((......................))..................
............))))))......((((..(((((((((((...(((.....))))(((.((((.............
.......(((((.....)))))....................)))).))))))))))(((..(((.(((.....))))))))))).)))))))..)))
)......))))................................................(((((.................
...................................................(((...(((((.....((((...))))......)))))..))).
...............................))))))...................
```

>centroid of phage M13KO7
```
AACGCUACUACUAUUAGUAGAAUUGAUGCCACCUUUUUCAGCUCGCGCCCCAAAUGAAAAAUAUAGCUAAACAGGUUAUUGACCAUUUGCGAAAUGUAU
CUAAUGGUCAAACUAAAUCUACUCGUUCGCAGAAUUGGGAAUCAACUGUUACAUGGAAUGAAACUUCCAGACACCGUACUUUAGUUGCAUAUUUAAA
ACAUGUUGAGCUACAGCACCAGAUUCAGCAAUUAAGCUCUAAGCCAUCCGCAAAAAUGACCUCUUAUCAAAAGGAGCAAUUAAAGGUACUCUCUAAU
CCUGACCUGUUGGAGUUUGCUUCCGGUCUGGUUCGCUUUGAAGCUCGAAUUAAAACGCGAUAUUUGAAGUCUUUCGGGCUUCCUCUUAAUCUUUUUG
AUGCAAUCCGCUUUGCUUCUGACUAUAAUAGUCAGGGUAAAGACCUGAUUUUUGAUUUAUGGUCAUUCUCGUUUUCUGAACUGUUUAAAGCAUUUGA
GGGGGAUUCAAUGAAUAUUUAUGACGAUUCCGCAGUAUUGGACGCUAUCCAGUCUAAACAUUUUACUAUUACCCCCUCUGGCAAAACUUCUUUUUGCA
AAAGCCUCUCGCUAUUUUGGUUUUUUAUCGUCGUCUGGUAAACGAGGGUUAUGAUAGUGUUGCUCUUACUAUGCCUCGUAAUUCCUUUUUGGCGUUAUG
UAUCUGCAUUAGUUGAAUGUGGUAUUCCUAAAAUCUCAACUGAUGAAUCUUUCUACCUGUAAUAAUGUUGUUCCGUUAGUUCGUUUUAUUAACGUAGA
UUUUUCUUCCCAACGUCCUGACUGGUAUAAUGAGCCAGUUCUUAAAAAUCGCAUAAGGUAAUUCACAAUGAUUAAAGUUGAAAUUAAACCAUCUCAAG
CCCAAUUUACUACUCGUUCUGGUGUUUCUCGUCAGGGCAAGCCUUAUUCACUGAAUGAGCAGCUUUGUUACGUUGAUUUGGGUAAUGAAUAUCCGGU
UCUUGUCAAGAUUACUCUUGAUGAAGGUCAGCCAGCCUAUGCGCCUGGUCUGUACACCGUUCAUCUGUCCUCUUUCAAAGUUGGUCAGUUCGGUUCC
CUUAUGAUUGACCGUCUGCGCCUCGUUCCGGCUAAGUAACAUGGAGCAGGUCGCGGAUUUCGACACAAUUUAUCAGGCGAUGAUACAAAUCUCCGUU
GUACUUUGUUUCGCGCUUGGUAUAAUCGCUGGGGGUCAAAGAUGAGUGUUUUAGUGUAUUCUUUCGCCUCUUUCGUUUUAGGUUGGUGCCUUCGUAG
UGGCAUUACGUAUUUUACCCGUUUAAUGGAAACUUCCUCAUGAAAAAGUCUUUAGUCCUCAAAGCCUCUGUAGCCGUUGCUACCCUCGUUCCGAUGC
UGUCUUUCGCUGCUGAGGGUGACGAUCCCGCAAAAGCGGCCUUUAACUCCCUGCAAGCCUCAGCGACCGAAUAUAUCGGUUAUGCGUGGGCGAUGGU
UGUUGUCAUUGUCGGCGCAACUAUCGGUAUCAAGCUGUUUAAGAAAUUCACCUCGAAAGCAAGCUGAUAAACCGAUACAAUUAAAGGCUCCUUUUGG
AGCCUUUUUUUUUGGAGAUUUUCAACGUGAAAAAAUUAUUAUUCGCAAUUCCUUUAGUUGUUCCUUUCUAUUCUCACUCCGCUGAAACUGUUGAAAG
UUGUUUAGCAAAACCCCAUACAGAAAAUUCAUUUACUAACGUCUGGAAAGACGACAAAACUUUAGAUCGUUACGCUAACUAUGAGGGUUGUCUGUGG
AAUGCUACAGGCGUUGUAGUUUUGUACUGGUGACGAAACUCAGUGUUACGGUACAUGGGUUCCUAUUGGGCUUGCUAUCCCUGAAAAUGAGGGUGGUG
GCUCUGAGGGUGGCGGUUCUGAGGGUGGCGGUUCUGAGGGAGGCGGUUCCGGUGGUGGCUCUGGUUCCGGUGAUUUUGAUUAUGAAAAGAUGGCAA
ACGCUAAUAAGGGGGCUAUGACCGAAAAUGCCGAUGAAAACGCGCUACAGUCUGACGCUAAAGGCAAACUUGAUUCUGUCGCUACUGAUUACGGUGC
UGCUAUCGAUGGUUUCAUUGGUGACGUUUCCGGCCUUGCUAAUGGUAAUGGUGCUACUGGUGAUUUUGCUGGCUCUAAUUCCCAAAUGGCUCAAGUC
GGUGACGGUGAUAAUUCACCUUUAAUGAAUAAUUUCCGUCAAUAUUUACCUUCCCUCCCUCAAUCGGUUGAAUGUCGCCCUUUUGUCUUUAGCGCUG
```

```
GUAAACCAUAUGAAUUUUCUAUUGAUUGUGACAAAAUAAACUUAUUCCGUGGGUGUCUUUGCGUUUCUUUUAUAUGUUGCCACCUUUAUGUAUGUAUU
UUCUACGUUUGCUAACAUACUGCGUAAUAAGGAGUCUUAAUCAUGCCAGUUCUUUUGGGUAUUCCGUUAUUAUUGCGUUUCCUCGGUUUCCUUCUGG
UAACUUUGUUCGGCUAUCUGCUUACUUUUCUUAAAAAGGGCUUCGGUAAGAUAGCUAUUGCUAUUUCAUUGUUUCUUGCUCUUAUUAUUGGGCUUAA
CUCAAUUCUUGUGGGUUAUCUCUCUGAUAUUAGCGCUCAAUUACCCUCUGACUUUGUUCAGGGUGUUCAGUUAAUUCUCCCGUCUAAUGCGCUUCCC
UGUUUUUAUGUUAUUCUCUCUGUAAAGGCUGCUAUUUUCAUUUUUGACGUUAAACAAAAAAUCGUUUCUUAUUUGGAUUGGGAUAAAUAAUAUGGCU
GUUUAUUUUGUAACUGGCAAAUUAGGCUCUGGAAAGACGCUCGUUAGCGUUGGUAAGAUUCAGGAUAAAAUUGUAGCUGGGUGCAAAAUAGCAACUA
AUCUUGAUUUAAGGCUUCAAAACCUCCCGCAAGUCGGGAGGUUCGCUAAAACGCCUCGCGUUCUUAGAAUACCGGAUAAGCCUUCUAUAUCUGAUUU
GCUUGCUAUUGGGCGCGGUAAUGAUUCCUACGAUGAAAAUAAAAACGGCUUGCUUGUUCUCGAUGAGUGCGGUACUUGGUUUAAUACCCGUUCUUGG
AAUGAUAAGGAAAGACAGCCGAUUAUUGAUUGGUUUCUACAUGCUCGUAAAUUAGGAUGGGAUAUUAUUUUUCUUGUUCAGGACUUAUCUAUUGUUG
AUAAACAGGCGCGUUCUGCAUUAGCUGAACAUGUUGUUUAUUGUCGUCGUCUGGACAGAAUUACUUUACCUUUUGUCGGUACUUUAUAUUCUCUUAU
UACUGGCUCGAAAAUGCCUCUGCCUAAAUUACAUGUUGGCGUUGUUAAAUAUGGCGAUUCUCAAUUAAGCCCUACUGUUGAGCGUUGGCUUUAUACU
GGUAAGAAUUUGUAUAACGCAUAUGAUACUAAACAGGCUUUUUCUAGUAAUUAUGAUUCCGGUGUUUAUUCUUAUUUAACGCCUUAUUUAUCACACG
GUCGGUAUUUCAAACCAUUAAAAUUUAGGUCAGAAGAUGAAAUUAACUAAAAUAUAUUUGAAAAAGUUUUCUCGCGUUCUUUGUCUUGCGAUUGGAUU
UGCAUCAGCAUUUACAUAUAGUUAUAUAACCCAACCUAAGCCGGAGGUUAAAAAGGUAGUCUCUCAGACCUAUGAUUUUGAUAAAUUCACUAUUGAC
UCUUCUCAGCGUCUUAAUCUAAGCUAUCGCUAUGUUUUCAAGGAUUCUAAGGGAAAAUUAAUUAAUAGCGACGAUUUACAGAAGCAAGGUUAUUCAC
UCACAUAUAUUGAUUUAUGUACUGUUUCCAUUAAAAAAGGUAAUUCAAAUGAAAUUGUUAAAUGUAAUUAAUUUUGUUUUCUUGAUGUUUGUUUCAU
CAUCUUCUUUUGCUCAGGUAAUUGAAAUGAAUAAUUCGCCUCUGCGCGAUUUUGUAACUUGGUAUUCAAAGCAAUCAGGCGAAUCCGUUAUUGUUUC
UCCCGAUGUAAAAGGUACUGUUACUGUAUAUUCAUCUGACGUUAAACCUGAAAAUCUACGCAAUUUCUUUAUUUCUGUUUUACGUGCUAAUAAUUUU
GAUAUGGUUGGUUCAAUUCCUUCCAUAAUUCAGAAGUAUAAUCCAAACAAUCAGGAUUAUAUUGAUGAAUUGCCAUCAUCUGAUAAUCAGGAAUAUG
AUGAUAAUUCCGCUCCUUCUGGUGGUUUCUUUGUUCCGCAAAAUGAUAAUGUUACUCAAACUUUUAAAAUUAAUAACGUUCGGGCAAAGGAUUUAAU
ACGAGUUGUCGAAUUGUUUGUAAAGUCUAAUACUUCUAAAUCCUCAAAUGUAUUAUCUAUUGACGGCUCUAAUCUAUUAGUUGUUAGUGCACCUAAA
GAUAUUUUAGAUAACCUUCCUCAAUUCCUUUCUACUGUUGAUUUGCCAACUGACCAGAUAUUGAUUGAGGGUUUGAUAUUUGAGGUUCAGCAAGGUG
AUGCUUUAGAUUUUUCAUUUGCUGCUGGCUCUCAGCGUGGCACUGUUGCAGGCGGUGUUAAUACUGACCGCCUCACCUCUGUUUUAUCUUCUGCUGG
UGGUUCGUUCGGUAUUUUUAAUGGCGAUGUUUUAGGGCUAUCAGUUCGCGCAUUAAAGACUAAUAGCCAUUCAAAAAUAUUGUCUGUGCCACGUAUU
CUUACGCUUUCAGGUCAGAAGGGUUCUAUCUCUGUUGGCCAGAAUGUCCCUUUUAUUACUGGUCGUGUGACUGGUGAAUCUGCCAAUGUAAAUAAUC
CAUUUCAGACGAUUGAGCGUCAAAAUGUAGGUAUUUCCAUGAGCGUUUUUCCUGUUGCAAUGGCUGGCGGUAAUAUUGUUCUGGAUAUUACCAGCAA
GGCCGAUAGUUUGAGUUCUUCUACUCAGGCAAGUGAUGUUAUUACUAAUCAAAGAAGUAUUGCUACAACGGUUAAUUUGCGUGAUGGACAGACUCUU
UUACUCGGUGGCCUCACUGAUUAUAAAAACACUUCUCAAGAUUCUGGCGUACCGUUCCUGUCUAAAAUCCCUUUAAUCGGCCUCCUGUUUAGCUCCC
GCUCUGAUUCCAACGAGGAAAGCACGUUAUACGUGCUCGUCAAAGCAACCAUAGUACGCGCCCUGUAGCGGCGCAUUAAGCGCGGCGGGUGUGGUGG
UUACGCGCAGCGUGACCGCUACACUUGCCAGCGCCCUAGCGCCCGCUCCUUUCGCUUUCUUCCCUUCCUUUCUCGCCACGUUCGCCGGCUUUCCCCG
UCAAGCUCUAAAUCGGGGGCUCCCUUUAGGGUUCCGAUUUAGUGCUUUACGGCACCUCGACCCCAAAAAACUUGAUUUGGGUGAUGGUUCACGUAGU
GGGCCAUCGCCCUGAUAGACGGUUUUUCGCCCUUUGACGUUGGAGUCCACGUUCUUUAAUAGUGGACUCUUGUUCCAAACUGGAACAACACUCAACC
CUAUCUCGGGACGGAUCGCUUCAUGUGGCAGGAGAAAAAAGGCUGCACCGGUGCGUCAGCAGAAUAUGUGAUACAGGAUAUAUUCCGCUUCCUCGCU
CACUGACUCGCUACGCUCGGUCGUUCGACUGCGGCGAGCGGAAAUGGCUUACGAACGGGGCGGAGAUUUCCUGGAAGAUGCCAGGAAGAUACUUAAC
AGGGAAGUGAGAGGGCCGCGGCAAAGCCGUUUUUCCAUAGGCUCCGCCCCCCUGACAAGCAUCACGAAAUCUGACGCUCAAAUCAGUGGUGGCGAAA
CCCGACAGGACUAUAAAGAUACCAGGCGUUUCCCCCUGGCGGCUCCCUCGUGCGCUCUCCUGUUCCUGCCUUUCGGUUUACCGGUGUCAUUCCGCUG
UUAUGGCCGCGUUUGUCUCAUUCCACGCCUGACACUCAGUUCCGGGUAGGCAGUUCGCUCCAAGCUGGACUGUAUGCACGAACCCCCCGUUCAGUCC
GACCGCUGCGCCUUAUCCGGUAACUAUCGUCUUGAGUCCAACCCGGAAAGACAUGCAAAAGCACCACUGGCAGCAGCCACUGGUAAUUGAUUUAGAG
GAGUUAGUCUUGAAGUCAUGCGCCGGUUAAGGCUAAACUGAAAGGACAAGUUUUGGUGACUGCGCUCCUCCAAGCCAGUUACCUCGGUUCAAAGAGU
UGGUAGCUCAGAGAACCUUCGAAAAACCGCCCUGCAAGGCGGUUUUUUCGUUUUCAGAGCAAGAGAUUACGCGCAGACCAAAACGAUCUCAAGAAGA
UCAUCUUAUUAAGGGGUCUGACGCUCAGUGGAACGAAAACUCACGUUAAGGGAUUUUGGUCAUGAGAUUAUCAAAAAGGAUCUUCACCUAGAUCCUU
UUAAAUUAAAAAUGAAGUUUUAAAUCAAUCUAAAGUAUAUAUGAGUAAACUUGGUCUGACAGUUACCAAUGCUUAAUCAGUGAGGCACCUAUCUCAG
CGAUCUGUCUAUUUCGUUCAUCCAUAGUUGCCUGACUCCCCGUCGUGUAGAUAACUACGAUACGGGAGGGCUUACCAUCUGGCCCCAGUGCUGCAAU
GAUACCGCGAGACCCACGCUCACCGGCUCCAGAUUUAUCAGCAAUAAACCAGCCAGCCGAUUCGAGCUCGCCCCGGGGAUCGACCAGUUGGUGAUUU
UGAACUUUUGCUUUGCCACGGAACGGUCUGCGUUGUCGGGAAGAUGCGUGAUCUGAUCCUUCAACUCAGCAAAAGUUCGAUUUAUUCAACAAAGCCG
CCGUCCCGUCAAGUCAGCGUAAUGCUCUGCCAGUGUUACAACCAAUUAACCAAUUCUGAUUAGAAAAACUCAUCGAGCAUCAAAUGAAACUGCAAUU
UAUUCAUAUCAGGAUUAUCAAUACCAUAUUUUUGAAAAAGCCGUUUCUGUAAUGAAGGAGAAAACUCACCGAGGCAGUUCCAUAGGAUGGCAAGAUC
CUGGUAUCGGUCUGCGAUUCCGACUCGUCCAACAUCAAUACAACCUAUUAAUUUCCCCUCGUCAAAAAUAAGGUUAUCAAGUGAGAAAUCACCAUGA
GUGACGACUGAAUCCGGUGAGAAUGGCAAAAGCUUAUGCAUUUCUUUCCAGACUUGUUCAACAGGCCAGCCAUUACGCUCGUCAUCAAAAUCACUCG
CAUCAACCAAACCGUUAUUCAUUCGUGAUUGCGCCUGAGCGAGACGAAAUACGCGAUCGCUGUUAAAAGGACAAUUACAAACAGGAAUCGAAUGCAA
CCGGCGCAGGAACACUGCCAGCGCAUCAACAAUAUUUUCACCUGAAUCAGGAUAUUCUUCUAAUACCUGGAAUGCUGUUUUCCCGGGGAUCGCAGUG
GUGAGUAACCAUGCAUCAUCAGGAGUACGGAUAAAAUGCUUGAUGGUCGGAAGAGGCAUAAAUUCCGUCAGCCAGUUUAGUCUGACCAUCUCAUCUG
UAACAUCAUUGGCAACGCUACCUUUGCCAUGUUUCAGAAACAACUCUGGCGCAUCGGGCUUCCCAUACAAUCGAUAGAUUGUCGCACCUGAUUGCCC
GACAUUAUCGCGAGCCCAUUUAUACCCAUAUAAAUCAGCAUCCAUGUUGGAAUUUAAUCGCGGCCUCGAGCAAGACGUUUCCCGUUGAAUAUGGCUC
AUAACACCCCUUGUAUUACUGUUUAUGUAAGCAGACAGUUUUAUUGUUCAUGAUGAUAUAUUUUUAUCUUGUGCAAUGUAACAUCAGAGAUUUUGAG
ACACAACGUGGCUUUCCCCCCCCCCCCCUGCAGGUCUCGGGCUAUUCUUUUUGAUUUAUAAGGGAUUUUGCCGAUUUCGGCCUAUUGGUUAAAAAAUG
AGCUGAUUUAACAAAAAUUUAACGCGAAUUUUAACAAAAUAUUAACGUUUACAAUUUAAAUAUUUGCUUAUACAAUCUUCCUGUUUUUGGGGCUUUU
CUGAUUAUCAACCGGGGUACAUAUGAUUGACAUGCUAGUUUUACGAUUACCGUUCAUCGAUUCUCUUGUUUGCUCCAGACUCUCAGGCAAUGACCUG
AUAGCCUUUGUAGACCUCUCAAAAAUAGCUACCCUCUCCGGCAUGAAUUUAUCAGCUAGAACGGUUGAAUAUCAUAUUGAUGGUGAUUUGACUGUCU
```

216

```
CCGGCCUUUCUCACCCUUUUGAAUCUUUACCUACACAUUACUCAGGCAUUGCAUUUAAAAUAUAUGAGGGUUCUAAAAAUUUUUAUCCUUGCGUUGA
AAUAAAGGCUUCUCCCGCAAAAGUAUUACAGGGUCAUAAUGUUUUUUGGUACAACCGAUUUAGCUUUAUGCUCUGAGGCUUUAUUGCUUAAUUUUGCU
AAUUCUUUGCCUUGCCUGUAUGAUUUAUUGGAUGUU
....(((((.(....))))))...........................................(((((.((((.(((.((((((((.............
.))))))))).............(((.((.(....)).))).............(((((......)))))....................))
).)))).))))).........................................((((((((...(((((((((.((((................
............)))).))..((.(((.......((((((((((.......................))))))))))......(((.....))
)((((......))))..(((((((...))))))).......)))).))..........)........)))))))....(((((((....)))))))....(((
((((.........................(((((.....))))...................................))))))..(((((....))))))...
(((.....))).....................(((((........))))............................................
......(((((((((..((....))...)).........)).........................................................
...............(((((........)))))..............................................................
..............................................................................(((((....((((.(((
.(((((.....))))))..)))((((((((((........)))).))).)))))))))))..............................................
.........(((((((.(((((.............))))))).))))..............................................
........................................................(((((((((((((((((((.((.(((.....
.........((((((((((.........(((((....((.......(((((.((((.....(((....))).(((((.........
......................((((....)))).................................(((((....)))))..(((((.(((((((........))
)))))).))))...((((((((((..((((.....................).))).))).))))))).....((((((((((...))))))))))..
....................................................................................................
.((.....................................(((((....)))))...........................(((((.......))))
..............................((.(((.....))).)).................))...................)))))....))))).)))).
(((((.......(((.((.....(((((((((.(((.....))).))..........................................))))))....
)).))).....)))))..........................................))...)))))((((.(((.....
..))).))))...........))))))......)))))))))....))..)))))))))))....)))))))..............................
....(((((...........................................................)))))...........(((((....)))))...
..............................((((.(.....................................).))))...............
..............................(.(((.....))).)..............................................
....((((....((((.......(((...........(((((((((((.............(((((...))))...........((....))..
.(((((((.......................................................................(((((((.(((.(((((
......))))).)))).)))))).....................................)))))))..)))))))))))))...........(.....
......................)..................................................................
...............................................))).......(((.((....)).)))......))))....(
(...(((((...)))))).)).)))))....(((((....)))))................................................
...........................................................................................
...........(((((....)))))).(((..(((((.............................(((((((..((((..(((.(((((((((...(((
(((((.....))))))))))((((((((((...)))))).))))......))...))))))..).))).)))).........)))))))........(((((..
.(((.(((((((.((((((((.....(((((((...........(((.(.....................).))).........(((((..........(((((..
...)))))......(((.((......)).)))................................)))))...........)))))))..((((.((....)).
))))..))).))))))).)))))))..))).(((((......................................)))))...)))))........((((((.((((((((
((((..(((..(.(((((.((((((.((((...............(((((..((.......)).)))))...((((((((.(((..((......)).))).)))))))
.....................)))))))).....)))).))))))))).)).)))..........).....)))..)))))))
).)))).(((((((((...................(((((...........))))...))))).))))))).........
.........(((((((...............(((...))).......((
(((((((((((...........))))))))....))))))))
...................((((.((((((..(((((.(((..(((((((.......................((((((((....
..........................(((((((..............................(((((((...(((((.......
......((((.((((..((((................)))).)))).))))...................((((.............))))....
....((((((((.........)))))))).............((((.((((...))))....)))).............((
(.....)))..........................................(((....)))..........((((.....)
)))......(((((((..............................................................((((((...
...(((......)))..................))))))).......)))))))))............((..............(((.(((((...
...((((((...))))((((((.........))))))))).....((((.(.(((((((((((.((((((((((((((.(....((
.................))...).))))))).)))))))...(((.((((((((...))))....
.)))))))..(((...(((((...............................(((((...(((((..(((.....)))).)))))....)))
)).............................................((((((...............(((((((.(
((((.....................................((((((((........(((.(((((((((((((((
............................((((((((((.....................(((((.....)))(((((((.(...)))))
```

```
))............................(((((((.....(((((((((((((((((((((((...))))))))))))))))))))))).))))..
..)))).))))..........(((((((.....(((((((.((((((...(((((((((((.((((.((((((((((............)))))
))))))).))))).))))..........................((((((((((((((...)))))))))))))).(((((..................
.....((((((((..........))))))))(((((((...)))))))).........))))).....................................
....................(((((.........))))))................(((((..((((((.....)))..))))))))).
))))).))..))...))))))).)))).))).(((((((.....))))))).........))))))).................................
....((((...(((((..(((.(((.......))).))).....))))....))))..........................(((((.......))))).
.......................(((((((((((((((((((((.(((((..((.......(((((...........))))........(((
(((((..((................(((((((........(((......))).))))))).........))...))))).))).........)
).))))).)))).(((.....((((.((((....)))).))))..............................))).))))).......))))))))
)))).)).(((((((((..((((((.............................((((.....(....((((((((((((.....)))))
))))))))).....))))........))))))))))))(((((.....))))...(.....)..............................
.....))))))))))).............((((((((((.....)))))))))).............................................
.......(((((........)))))....))))))))))))........)).))))))))))).......)))...).)))))))))((.(((((.
..((((....))))...))))))).......)))))))).............................................................
..................................(((((((((((.............(((((.(((((..((...............
..................................)).))))).))))).....((((((((((((..(((..((((.............
...........................................................................................
.............(((((((..(((((..((...(((((......)))).......)).))))(((((....(((((.........(((((.(
.................)..))))).(((.((((....)))).))).))))).((((((((((((.............................
.........)))))))).)).)))))))))).....)))))))))).(((.(((..................)))..))).................
............)))))..))).......................))))))))))))))).....))))))))))..................(((
(.(((((....))))))))).....................((((.....))))....................................(((((((((
....(((.............)))........)))))))))...........................))))).(((((((((.....))))....))))
.)))))))........((((((....)))))).....))))))))).............((..(((((((.....))))))...))).((((...)))).
..........))).)))))))...............(((((......((......(((((..............)))))...(((((.((.(
((...........))).)).)))))............................................................................
......))......((((....))))............))))))..................................)))))))))))..)))))).))
))))...).))))...........................(((((..((((((((((((...(((.....)))((((.((((...........
............................((((((.....))))).................)))).))))))))))(((..(((.(((.....)))
))))))).))))))))..))))................))))))).............)))))...)))......................
..........................................................................(((...))).
.......................................................................
```