UCLA
UCLA Electronic Theses and Dissertations

Title

Genome and Internal Protein Ejection from DNA Viruses

Permalink

https://escholarship.org/uc/item/9mv7t7tf

Author

Jin, Yan

Publication Date

2015

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Genome and Internal Protein Ejection from

DNA Viruses

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Chemistry

by

Yan Jin

2015

Genome and Internal Protein Ejection from dsDNA viruses

by

Yan Jin

Doctor of Philosophy in Chemistry

University of California, Los Angeles, 2015

Professor William M. Gelbart, Co-chair

Professor Charles M. Knobler, Co-chair

Abstract

Because of the unique simplicity of their life cycles, compared to all other evolving organisms, double-stranded (ds) DNA bacteriophages have served as an extremely

valuable model for elucidating the basic physics and molecular biology of gene replication and expression. In this dissertation, we present experimental work on two phages, lambda and P22, to study the general genome delivery (ejection) mechanisms of dsDNA bacteriophage.

Using lambda phage, we systematically investigated the effects on DNA genome ejection of external osmotic pressure controlled by the concentration of osmolyte (PEG 8000) and the presence of polyvalent cations (tetravalent polyamine spermine, Sp4+). We found that the internal pressure of the capsid decreases from 38 to 17 atm as the [Sp4+] is increased from 0 to 1.5 mM. The existence of Sp4+ can also induce incomplete ejection under zero osmotic pressure when its concentration reaches 0.15 mM or higher; for [Sp4+] below this threshold, the ejection is complete. Further, we observed that the self-attraction induced by Sp4+ affects the configurational dynamics of the encapsidated genome, causing it to get stuck in a broad range of non-equilibrated structures.

In order to further study the DNA ejection mechanism from phage capsids, we have systematically determined how DNA transcription *in vitro* is affected by the presence of different osmolyte and viscogen molecules, so that we can test the transcription-pulling hypothesis of genome delivery in the presence of crowded environments mimicking the cytoplasm of the bacterial cell hosts of phages. We found that at high concentrations of DNA templates, macromolecules can increase the RNA yield due to crowding effects on the initiation step of transcription, while small molecules decrease the yield because of

viscosity effects on the elongation step. Experiments carried out at low concentrations show a decrease in yield for large and small molecules, confirming the dominant effect of viscosity effects in the elongation step.

Having established and quantified the nature of the spontaneous driving force for DNA delivery, we then studied the ejection behavior of internal proteins from another phage, P22, and their functions in an *in vitro* osmotic suppression system controlled by PEG 8000. We found that the Outer Membrane Protein A (OmpA) from *Salmonella*, the natural bacterial host of P22, can significantly enhance the rate of DNA ejection in the presence of the primary receptor, LPS. While the DNA is ejected in the presence of LPS, no ejection of the internal proteins occurs unless OmpA is also present. We also find that their ejection is largely complete before any of the genome is ejected. This finding helps us understand the possible roles that the internal proteins play during infection.

The dissertation of Yan Jin is approved.

Harold Monbouquette

William M. Gelbart, Committee Co-chair

Charles M. Knobler, Committee Co-chair

University of California, Los Angeles

2015

Dedication

This dissertation is dedicated to my parents and sister, Jingde Jin, Aishu Piao, Ying Jin for supporting me with affections and love. And to my advisors, William M. Gelbart and Charles M. Knobler, without whom none of the work would be possible.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

Chapter 1 is a version of "Controlling the Extent of Viral Genome Release by a Combination of Osmotic Stress and Polyvalent Cations" by Yan Jin, submitted for publication.

Chapter 2 is work related to a project in collaboration with the Shimon Weiss Group.

Chapter 3 is a version of "Does Bacteriophage P22 Eject All Its Internal Proteins Before Its Genome?" by Yan Jin, submitted for publication.

There are lots of people who have helped me a lot along the path to a PhD. First, I would like to express my deepest appreciation to my advisors Bill Gelbart and Chuck Knobler, who have always been fully supporting me, giving me generously their time and knowledge. Their incredible consideration, enthusiasm, encouragement and faith in me throughout have been extremely helpful. Without their guidance and patience, this thesis would not be possible.

With my sincere gratitude, I want to thank the following people who have helped me directly on the projects:

Dr. Li Tai Fang, thank you for being so patiently answering my full-of-question emails and providing useful tips for helping me get started with lambda work.

Kristin Parent lab in Michigan State University where I did my P22 radioactive work in Chapter 4: Kristin helped purifying radioactive viruses, gave incredibly helpful guidance throughout the whole project; Natalia Porcek helped purifying OmpA protein; John Dover helped purifying LPS; And Sophia Sdao helped with the ejection rate experiments for P22 hybrids. Without them this project would not be possible. I like working in your lab Kristin!

I am also grateful for everyone in the UCLA virus group:

Dr. Ajay Gopal, with his brilliance he has taught me a lot during conversations covering many different areas from science, technology to entrepreneurship, physical activities… And you have helped me go through my hardest time in life. I will always remember the ice cream you gave me to cheer me up that day.

Dr. Odisse Azizgolshani, my only female labmate who has not only taught me a lot about science but also about life as a woman. Your kindness and elegance will always influence me.

Dr. Aron Yoffe, as my science and English "teacher", thank you for being so patiently correcting my English and explaining a lot of words to me. You have changed my English from "bad" to "not too bad". And the most important thing is: thank you for being so generous every time I played a trick with you.

Dr. Ruben Cadena, Dr. Maurico Comas-Garcia, and Dr. Rees Garmann. As senior group members, you are very good role models as researchers. And thank you for leading me into the world of CCMV virus.

Rich Sportsman, Walter Singaran, Christian Beren, Adam Biddlecome, Devin Brandt, Cho Ki Tam. I love working in such an environment with your company. I learnt a lot from you and you have brought much laughter to the lab. All of you make our lab unique.

I would like to thank all my friends I've met here in UCLA: Ling Peng, Yuxi Zhao, Juyao Dong, Chencai Wang, Junda Zhu... Your company makes my life full of excitement and surprise. Without you, I would not have overcome many difficulties that easily.

With my warmest affection I want to thank my best friends forever in China: my "loft". Thank you for knowing me better than I do. I am glad the distance haven't changed our friendship at all. You are always in my deepest heart.

Finally I would like to include a special note of thanks to express my deepest love to my family: my mother, my father and my sister who are in China:

亲爱的爸爸妈妈臭狗子，感谢一路上有你们的陪伴，你们始终是我在大洋彼岸最牵挂的人，也是因为有你们的支持，我才可以任性的一步一步越飞越远，虽然我知道你们的心中是有多么的不舍与想念。相见的机会总是很少很短暂，原谅我长这么大以来没有做到作为女儿该做的，但你们的平安幸福永远是我最大的追求，答应你们，女儿以后要多多回家。

狗子，很少和你说感谢，但真的是因为有一个这么懂事会照顾家人的你，我才会放心的在外飞翔，否则我真的会义无反顾的放弃这边的生活奔回爸妈身边。辛苦你了，好好照顾自己，我永远年轻貌美如花的狗狗！

Thank you all for your incredible love and support. Words cannot express how I feel.

VITA

2005 B.Sc. Macromolecular Materials and Engineering

     Fudan University, Shanghai, China

CHAPTER 1

Introduction

Bacteriophages are a group of viruses that infect bacteria. Their life cycles have long attracted the attention of biologists and physicists, going back to the 1940s when Delbruck and Lvoff founded their "phage schools" at Caltech and the Pasteur Institute, respectively. Their aim was to elucidate the fundamental processes of DNA replication and gene expression by means of simple viruses and the bacteria they infect; in doing so, they founded the field of "molecular biology". Although it is 62 years ago that the Hershey-Chase experiment was performed (Hershey *et al.,* 1952) showing that when bacteria are infected only the phage genome enters the host cell, with its capsid (protein) remaining outside, many aspects of the process are still under debate. Among them are two major questions: "What is the mechanism of genome ejection?" and "How do the internal proteins that are packaged inside the mature capsid along with DNA get involved in DNA transport through the host cell membranes?"

Bacteriophages were discovered independently by two groups of people in the early 1900's (Twort, 1915; d'Hérelle, 1917) and have since been intensively studied, serving as a tool to understand the physics of biological molecules and the process of life. Like all viruses, bacteriophages consist of genetic material (DNA or RNA, single-stranded or double-stranded) surrounded and protected by a protein capsid. But the phages can differ in shape: icosahedral shell, with or without a tail; or a filamentous structure (Mc Grath *et*

al., 2007). Double-stranded (ds) DNA bacteriophages with tail and icosahedral shell are the main subject of interest here.
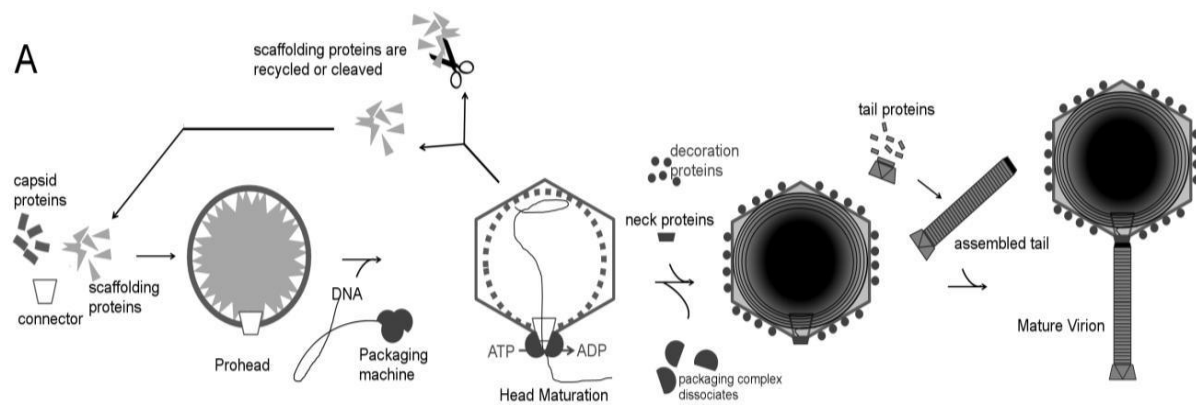
Figure 1-1. Schematic of dsDNA tailed-phage assembly steps: First, scaffold proteins guide the coat protein to form a procapsid; then DNA is packaged, followed by expansion and maturation of the capsid; subsequently, the tail is attached to the head to form a mature phage particle. (Aksyuk, 2011)

Figure 1-1 shows the general pathway of a tailed-phage assembly from a succession of virally-encoded proteins (Aksyuk, 2011): first the scaffold proteins co-assemble with coat proteins to form a double-shelled structure with a portal complex protein located at one vertex. Then a DNA "packaging machine" binds to the portal complex and starts packaging DNA into the capsid. At the same time the capsid expands and the scaffold proteins escape from the capsid. After all the DNA has been packaged, the motor falls off and other tail parts are added on to the portal and a mature virus capsid is formed. Some viruses also package several copies of proteins (internal proteins) that are important in their life cycle inside the capsid during the procapsid formation period.

During the DNA packaging process, the virally-encoded motor needs to consume energy from the hydrolysis of ATP to drive the DNA into the capsid. Work must be expended because DNA is a uniformly-negatively-charged and rigid polymer and the total length ("contour" length) of the genomic DNA that needs to be packaged in one capsid is usually several hundred times longer than the capsid radius. The radius is also smaller than the "persistence" length of the DNA, i.e., the radius into which the DNA can be bent by thermally-available energy. Thus the electrostatic self-repulsion and bending energy of the densely-packed charged DNA make dsDNA viruses extremely highly pressurized (Smith *et al.*, 2001). For instance in bacteriophage lambda, the internal pressure can reach about 40 atm (Evilevitch *et al.*, 2003).
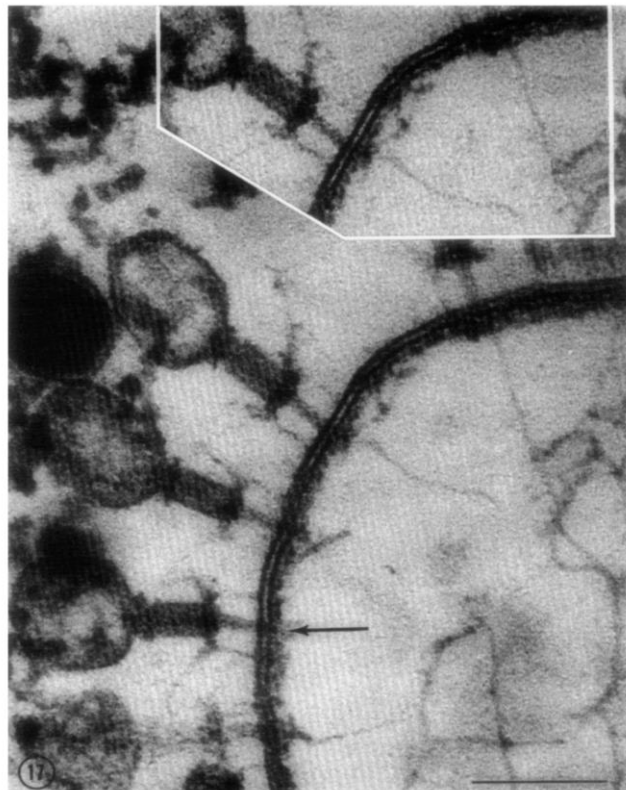
Figure 1-2. Phage T4 infecting E. coli: the tip of the tail as well as the tail spikes bind to receptors on the host surface and then the genome is delivered into the cell, leaving the protein capsid outside (Simon *et al.*, 1967).

Viruses can't "live" without infecting their host cell. From the early Hershey-Chase experiment we already know that a dsDNA virus delivers only its genome into the host and leaves the capsid outside. Figure 1-2 shows T4 viruses infecting E. coli (Simon *et al*., 1967). The phages first diffuse around the host surface until they dock at specific molecules on the cell surface called receptors (often transmembrane proteins), followed by conformational changes in the proteins located in the portal and the tail of the virus which trigger genome transfer into the targeted cell.

Studies done with dsDNA viruses have confirmed that the internal pressure plays a significant role in delivering DNA by helping the genome penetrate the host membrane, working against friction and the inside osmotic pressure of the cell (~3-5 atm due to many proteins and small molecules in the cytoplasm). So it is easy to imagine that anything that can change the pressure difference between the virus capsid and the host will affect the genome delivery process. In Chapter 2, we use bacteriophage lambda as a model system to study the extent of phage genome release by a combination of osmotic stress and polyvalent cations. Previous studies showed the dependence of *in vitro* DNA ejection properties on either ambient salt concentrations or osmotic pressure. But none has systematically investigated how the extent of ejection is controlled by a combination of these conditions. In this work we specifically quantify the role of: the polyvalent cation, spermine, a polyamine with four positive charges, which is permeable to the capsid and can neutralize the negative charges on lambda DNA; and the osmolyte polyethylene glycol (PEG) 8000, a polymer that cannot penetrate into the capsid and thus can affect

only the environmental osmotic pressure. We find that incubation of the phage particles in spermine concentrations as low as 0.15 mM – the threshold for DNA condensation in bulk solution – is sufficient to significantly limit the extent of ejection in the absence of osmolyte; for lower spermine concentrations the ejection is complete in the absence of osmolyte. We also show that any desired extent of DNA ejection can be assured by different combinations of polyvalent cation concentration and osmotic pressure in the host solution. The degree of nonequilibrium dynamics of the packaging and ejection processes can also be investigated at the same time, showing that the presence of polyvalent cation, which induces a self-attraction between neighboring portions of packed DNA, impedes re-organization of the genome, causing it to get stuck in a broad range of non-equilibrated structures.

As the genome comes out of the capsid, the osmotic pressure inside the capsid continues to decrease and finally reaches the pressure of the host cell. Experiments on phages lambda and T5 (Evilevitch *et al*., 2003; Leforestier *et al*., 2008) have shown that at this point there is still around half of the genome remaining inside, but the pressure exerted by it no longer exceeds that of the host. Then, what drives the second half of the genome into the host? One of the proposed mechanisms is that transcription of viral DNA by RNA polymerase can pull the DNA out of the capsid. Transcription of the genome is the first stage of infection in cells. To test this idea *in vitro* we proposed to first use high-molecular-weight osmolyte to stop the DNA ejection half-way and then initiate transcription to see whether the DNA ejected length will change. In order to do this, it is necessary to first know how transcription works in a crowded environment. It also helps us understand the

5

transcription process *in vivo* where the cell contains many biological molecules and organelles. In Chapter 3, we mainly studied the dependence of the transcription process on different types (large molecules and small molecules) of osmolyte/viscogen both at high concentration and at the single-molecule level. When reactant concentrations are high, high-molecular-weight osmolyte can increase the yield of RNA transcription while small molecules inhibit this process, reducing the yield of RNA. Single-molecule measurements allow us to study separately the elongation step of the transcription process, and it is observed that the RNA yield is always *de*creased, independent of the kind of osmolyte involved. These findings make the transcription process in live cells more understandable.
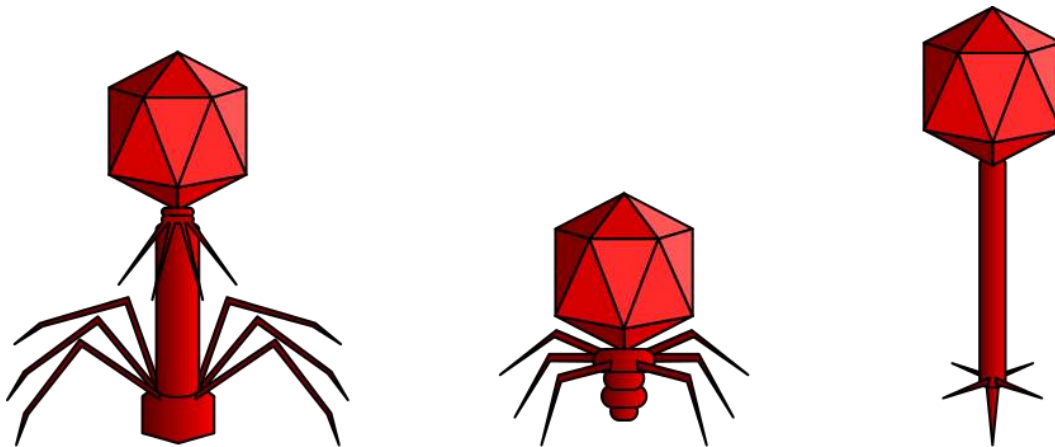
Figure 1-3. Three families of tailed bacteriophage (Mc Grath *et al*., 2007). Left: *Myoviridae*, with long contractile tails; middle: *Podoviridae*, whose tails are very short and usually no longer than one third of those in *Myoviridae* and *Siphoviridae*; right: and *Siphoviridae*, also with long (but non-contractile) tails.

Knowing the driving force for DNA delivery after the virus binds to the outer membrane receptor, we want to further find out what factors are needed to facilitate passage of the viral DNA through both the outer and inner membranes as well as the periplasmic space of the host. Cartoons shown in Figure 1-3 depict three families of the tailed phages (Mc Grath *et al*., 2007). The main difference between them lies in their tail properties, with different tails suggesting that they have different mechanisms of delivering their genome into the cytoplasm. On the left of Figure 1-3 is a sketch of T4, one of the *Myoviridae*. It has a long contractile tail. When it infects bacteria, the tail contracts and facilitates the penetration of DNA into the cell membranes. The phage on the right represents one of the *Siphoviridae,* which has a long flexible tail. Its tail does not contract to lead the genome to cytoplasm. The middle one with the shortest tail is typical of the *Podoviridae*, one of which – P22 – is the focus of Chapter 4. They have short, non-contractile tails, no longer than 17nm. We deal with the question of how they are able to transfer their genomes across the 40-60nm defense barrier of host cell bacterial membranes.

Studies have shown that the internal proteins that are prepackaged (and which do not leave with the scaffolding proteins) are needed to achieve successful infection, but little is known about how and when the internal proteins come out of the virus. We choose to work with P22, one of the *Podoviridae*, and set up an *in vitro* osmotic suppression system with PEG 8000 to study the ejection of the internal proteins by progressively controlling the extent of ejection from the phage.
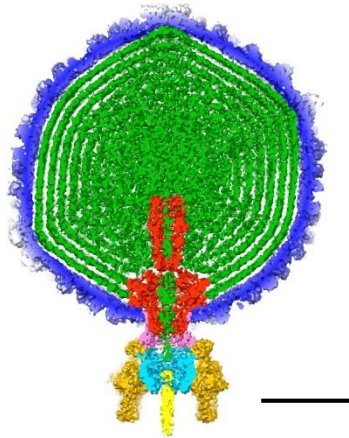
Figure 1-4. P22 virion structure (Tang *et al.*, 2011), segmented to highlight individual components: coat protein (blue); DNA (green); portal protein (red); tail spike protein (dark yellow); tail needle (yellow). Scale bar: 20 nm.

Figure 1-4 is a cryo-reconstruction image of the P22 virion (Tang *et al.*, 2011). Blue color is the capsid made of coat protein. It is filled with viral genomic DNA, which is represented by green. The portal protein (colored red) is located at one of the vertices and extends inside almost to the center the capsid. Connecting with the portal are the tail-spike proteins (labeled with dark yellow) which are functional during the early stage of the infection. In the center below the portal (colored yellow) is the tail-plug protein (needle), which needs to be displaced during infection. Previous studies have shown that lipopolysaccharide (LPS) from the virus' host cell (*Salmonella*) is recognized as a receptor that triggers ejection *in vitro* (Iwashita *et al.*, 1973). LPSs are large molecules consisting of lipid and of a polysaccharide composed of O-antigen, which is a glycan polymer comprising the outermost domain of LPS. LPSs are found in the outer membrane of the bacteria (Raetz *et al.*, 2002).
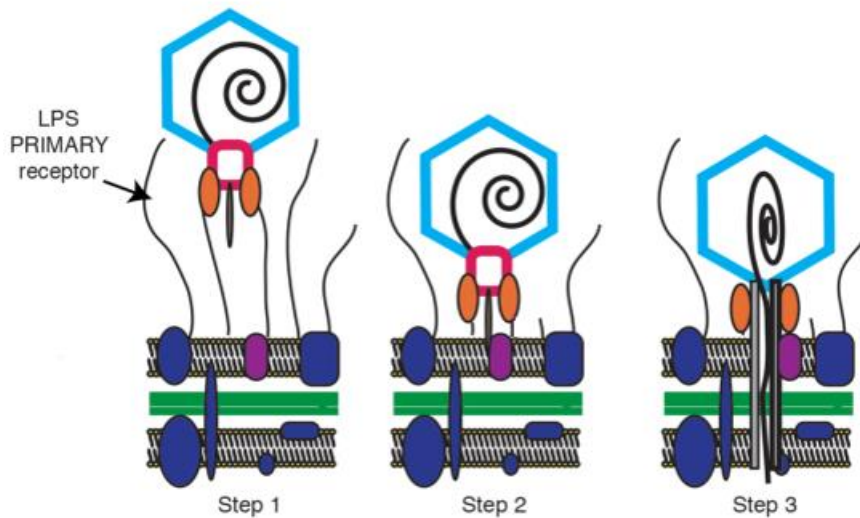
Figure 1-5. P22 infection process (Kristin et al., 2014): A virion utilizes its tail spike proteins to bind to its primary receptor, LPS. By hydrolyzing the LPS, virus is able to move closer to the cell surface and finally deliver their genome into the host cytoplasm. During the genome entry, there has to be some mechanism to facilitate this process. For example, a putative channel is formed through both membranes to transfer DNA into the host.

Figure 1-5 shows the infection process of P22 (Kristin et al., 2014): Infection begins when the tail-spike proteins of P22 interact with LPS from its host *Salmonella* to help position the phage on the bacterial outer membrane. Tail-spike proteins have enzymatic properties that can digest the long O-antigen of the LPS into smaller repeat units and bring the phage closer to the surface of the host. Upon protein structural changes, tail-plug proteins are displaced and finally DNA is released (Andres *et al*., 2010; 2012).

We find from *in vitro* studies of ejection that the Outer Membrane Protein A (OmpA) from *Salmonella* significantly enhances the rate of DNA ejection in the presence of LPS. Further, we observe that while the DNA is ejected in the presence of LPS, no ejection of the internal proteins occurs unless OmpA is also present, and that their ejection is largely

complete before any of the genome is ejected. This finding helps us better understand how P22 infects its host cell and, in particular, about possible roles that the internal proteins play during infection.

CHAPTER 2

Controlling the Extent of Viral Genome Release by a Combination of Osmotic Stress and Polyvalent Cations

Abstract

While several *in vitro* experiments on viral genome release have specially studied the effects of external osmotic pressure and of the presence of polyvalent cations on the ejection of DNA from bacteriophages, few have systematically investigated how the extent of ejection is controlled by a combination of these effects. In this work we quantify the effect of the tetravalent polyamine spermine on the extent of ejection of the DNA from bacteriophage lambda as a function of the osmotic pressure. We found that the pressure required to completely inhibit ejection decreases from 38 to 17 atm as the spermine concentration is increased from 0 to 1.5 mM. Further, incubation of the phage particles in spermine concentrations as low as 0.15 mM - the threshold for DNA condensation in bulk solution - is sufficient to significantly limit the extent of ejection in the absence of osmolyte; for spermine concentrations below this threshold, the ejection is complete. In accord with recent investigations on the packaging of DNA in the presence of a condensing agent,

we observe that the self-attraction induced by the polyvalent cation affects the ordering of the genome, causing it to get stuck in a broad range of non-equilibrated structures.

Many double-stranded (ds) DNA viruses have their stiff, highly-charged, genomes packed to crystalline density in pre-formed rigid shells (capsids), resulting in a high state of stress. The work of packaging is performed by strong virally-encoded protein motors, and the resultant stress provides much of the driving force for genome ejection. Considerable theoretical (Reimer *et al*., 1978; Odijk, 1998; Kindt *et al*., 2001; Tzlil *et al*., 2003; Purohit *et al*., 2005; Forrey *et al*., 2006; Petrov *et al*., 2011) and experimental work (Smith *et al*., 2001; Evilevitch *et al*., 2003; de Frutos *et al*., 2005; Smith *et al*., 2011) has been devoted to calculating and measuring the forces and pressures associated with these processes. It is only recently, however, that attention has been focused on the nonequilibrium dynamics of the packaging process, which involves loading rates as high as 100s of base pairs (bps)/sec. In particular, Smith and co-workers (Berndsen *et al*., 2014; Keller *et al*., 2014) have employed single-molecule techniques to study the heterogeneity of the packaging dynamics and how it varies when the DNA-DNA interactions are changed from repulsive to attractive.

One expects that the genome dynamics associated with ejection from the capsid will also be sensitive to disorder in the packaged DNA and to the nature of the DNA-DNA

interactions. We address these issues here by investigating the separate and combined effects of polyvalent counterions and of osmolytes on the *in vitro* genome ejection properties of bacteriophage λ, a dsDNA virus. As predicted (Tzlil *et al*., 2003), we found that a significant fraction of the genome remains inside the capsid when ejection is triggered in the absence of osmolyte but in the presence of spermine (Sp4+) at concentrations sufficiently high to induce DNA self-attraction (condensation). Comparably limited ejection obtains in the absence of spermine and in the presence of 3-4 atm of osmotic pressure.

Most notably, and consistent with the strong heterogeneity in packaging rates reported by Smith and coworkers (Berndsen *et al*., 2014; Keller *et al*., 2014), we found in our ejection studies that DNA self-attraction induced by polyvalent cations causes the packaged DNA to become stuck in a wide range of non-equilibrium structures. Related effects have been discussed in several simulation studies of phage packaging (Forrey *et al*., 2006; Petrov *et al*., 2011; Ali *et al*., 2006), including the effects of twist (Spakowitz *et al*., 2005) on entanglement. Specifically, we report measurements of the extent and nature of genome ejection for a variety of biologically relevant osmotic pressures and polyvalent cation concentrations, and discuss our results in the context of related work on DNA condensation and of our present understanding of strongly-confined DNA.

Wildtype bacteriophage lambda was purified from infected E. coli c600, and the lambda receptor (LamB) purified from E. coli pop154, both as described earlier. The final titer of phage was around 10^{11} infectious units per ml, determined by plaque assay. Virus

solution was first treated with DNase I (1 unit, New England Biolabs) to remove free DNA due to the breakage of capsids during purification. Each 100 µl of phage sample was then incubated overnight at 4ºC in TM buffer (50mM Tris, pH 7.5, 10mM Magnesium Sulfate) with Sp4+ so that it could diffuse into the capsid and equilibrate with the DNA. PEG 8000, LamB, 1% detergent (oPoE), DNase I buffer and an additional 5 µl of DNase I (2 units) were then added together to the samples, and incubated at room temperature for 15 min to allow the LamB to bind to the phage, triggering ejection. Further incubation for 4 hr at 37 ºC ensured that the ejected DNA was fully digested by DNase I.

The unejected DNA lengths were determined by extracting them from the capsids. To denature the DNase I before extraction, 1 mM EDTA was added and the sample heated for 10 min at 75ºC. After addition of an equal volume (100 µl) of protein lysis buffer (25 mM EDTA, 200 mM Tris, 250 mM NaCl, 1%SDS, pH 7.5) and 1 µl of Protease K (1 unit), the samples were incubated at 65ºC for 1 hr to disrupt the capsids, allowing the DNA inside to be released. Phenol/chloroform extraction was done twice to separate protein from DNA, followed by ethanol precipitation to concentrate the DNA. Finally, the precipitated DNA was redissolved in TE buffer and analyzed by electrophoresis in a 0.3% agarose gel run for 6 hr at 3V/cm and stained with SYBR gold.

Figure 2-1.A is a gel image showing the length of DNA remaining inside the capsid when ejection is triggered in solutions containing no osmolyte, but with Sp4+ at concentrations between 0 and 1.5 mM. Because of a low concentration of the LamB receptor (only one per phage) and the failure of some phages to eject even when bound to receptor, genome

DNA ladder(kbp)	1.5	1	0.5	0.15	0.1	0.05	0	[Sp4+] (mM)

1.A

48.5
20
15
10
8
6
5
4
3
2
1.5
1
0.5

DNA ladder (kbp)	0	0.1 mM	1.5 mM	[Sp4+] (mM)
	7.5	2.5	0	P/atm

1.B

48.5
20
15
10
8
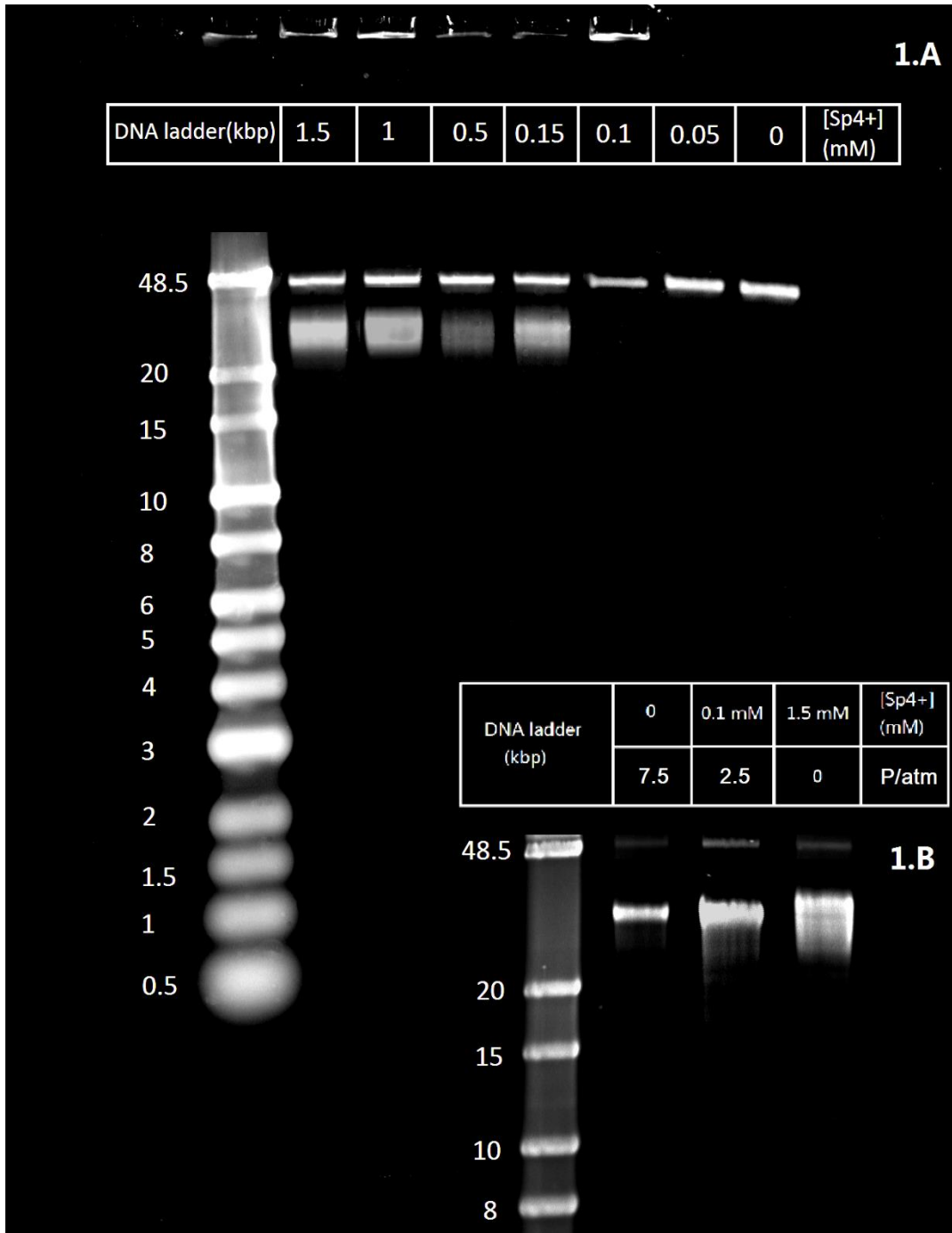
Figure 2-1. 1. A. 0.3% agarose gel of unejected DNA, following ejection from phages triggered by LamB after incubation with Sp4+ at concentrations ranging from 0 to 1.5 mM. 1. B. Same, but with the indicated combinations of PEG and Sp4+ concentration.

ejection is triggered in only about half of the phage particles in any sample. (This estimate is inferred from the relative intensities of the two bands in the gel - see the discussion given immediately below.) Accordingly, all of the lanes show a band corresponding to the full-length (48.5 kbp) DNA that remains unejected and hence protected by the capsid. In the rightmost 3 lanes, for which [Sp4+] < 0.15 mM, we see only this band, because ejection is complete from all opened capsids. For [Sp4+] > 0.15 mM, however, there is another (and distinctly more diffuse – see below) band corresponding to the lengths ($\approx$ 25-33 kbp) of DNA remaining in the capsids following ejection. The presence of a significant length of the genome remaining in the capsid attests to the DNA having been condensed by spermine (Tzlil *et al*., 2003), and the broad range of lengths involved reflects the large non-equilibrium effects associated with the spermine-induced attractions between neighboring portions of packaged duplex. This strong heterogeneity is consistent with the large standard deviation in packaged lengths measured for high spermine concentration by Keller *et al* (2014).

Note that the above 25-33 kbp range of lengths corresponds to 52-68% of the genome remaining in the capsid when ejection is triggered under DNA condensation conditions. Measurement of the UV absorbance of ejected, digested, DNA in the supernatant, following centrifugation, while not providing direct information on the distribution of ejected (and therefore of un-ejected) DNA lengths, does give their average value. In this way Evilevitch (2006) obtained a value of 29 ± 2 kbp for the average length remaining inside the lambda capsid when ejection was triggered in the presence of 1 mM spermine, with the ± 2 kbp uncertainty calculated from propagation of errors. This result, i.e., 60% of the

genome remaining inside in the presence of 1mM spermine and 0 atm osmotic pressure, can be compared with 50% and 90% remaining inside in the case of (1 mM spermine and) 1 atm (Evilevitch *et al.*, 2004) and 3.5 atm (Evilevitch *et al.*, 2008) of osmotic pressure, respectively.

As a control, we checked that spermine-induced condensation of DNA did not result in the DNA running differently than uncondensed DNA in the electrophoresis gel. Samples of lambda DNA incubated in 1 mM Sp4+, and the same DNA incubated in spermine-free buffer solution, ran identically in a 0.3% agarose gel in TE (spermine-free) buffer, thereby confirming our determinations of unejected DNA lengths in the case of super-threshold [Sp4+]. Further, consistent with related studies (Evilevitch *et al.*, 2006), DNA is not protected against DNase I digestion by virtue of it being condensed by spermine: DNA is essentially fully digested into nucleotides by DNase I, after incubation with Sp4+ at concentrations up to 5 mM, under the conditions of our nuclease digestion protocol.

As an additional control we checked that the duration of digestion of ejected DNA did not affect the *amount* of ejection, i.e., the length of DNA remaining in the capsid. In particular, after adding LamB to the phage samples incubated in spermine and/or PEG, we digested for either 3, 4, or 5 hr with DNase I and ran gels as in Figure 2-1.A and 2-1.B for each of the three time durations, and found no differences in the positions and intensities of the bands associated with DNA remaining in the capsids.

Even though the phages have been incubated in spermine before genome release is triggered, the packaged DNA can only begin to feel its self-attraction after a significant amount of ejection has occurred. Up until that point the average interaxial spacing between neighboring duplex portions is smaller than the 2.8-2.9 nm at which the onset of attractive forces occurs (Raspaud *et al.*, 2005). The effective interaction is repulsive and ejection proceeds until the average interaxial spacing becomes this large or, alternately, until the length of DNA remaining inside is sufficiently small that its toroidal condensate has an outer radius equal to the inner radius of the viral capsid (Tzlil *et al.*, 2003; Leforestier *et al.*, 2009). Were the DNA able to relax (equilibrate) during packaging and ejection, the points at which these limits are reached would coincide. But they can vary considerably from phage to phage because the local packaging configurations are highly heterogeneous as a result of large non-equilibrium effects associated with how the genome was packaged in the first place (Berndsen *et al.*, 2014; Keller *et al.*, 2014) and especially because of the onset of strong self-attraction of the packaged DNA as it is ejected following the incubation with spermine.

Measurements were also carried out with mutant bacteriophage lambda b221, whose genome length is 37.8 kbp. In this case, the lengths of DNA remaining in the capsid were - as expected - the same (25-33 kbp) as for the wildtype, because of the ejected DNA being digested by DNase I. And, again, as with the wildtype, the unejected DNA is much more polydisperse than would be expected if the DNA configurations arising in packaging and ejection of the genome were equilibrated for all of the force, pressure, and spermine conditions involved.

This departure from equilibrium was anticipated by Tzlil *et al.*(2003) who, in calculating that 1/4th to 1/3rd of the lambda genome should remain unejected in the presence of self-attraction, pointed out that this prediction should be a lower-bound because of non-equilibrium effects. Computer simulation studies by Petrov and Harvey (2011) also addressed the effect of self-attraction on ejection. In particular, they used molecular dynamics to examine different lengths of DNA - corresponding to 10, 30 and 78% of the full genome - equilibrated in capsids with purely repulsive interactions between neighboring duplex portions. Attractive interactions were then switched on, mimicking the effect of incubation in a sufficiently high concentration of spermine. The shorter (10 and 30% length) chains were able to form toroid-like structures, whereas the trajectories associated with the long (78%) one were associated with pronounced non-equilibrium structures. Similar effects had been reported in earlier phenomenological packaging simulations by Ali, Marenduzzo and Yeomans (2006), and by Forrey and Muthukumar (2006).

In an experiment similar in concept to ours, de Frutos *et al.* (2005) examined the ejection of DNA from T5 phage, for which the full length is 114 kbp, in the presence of sufficient spermine to condense the DNA. Gel electrophoresis analysis of the DNA remaining in the phage capsids showed several broad bands, at approximately 100, 46 and 11 kbp. These widely different lengths are difficult to explain, but are consistent with those found for T5 by Leforestier *et al.* (2008), who measured the lengths of DNA remaining in the capsid following ejection in the presence of low concentrations of PEG; they are also consistent

with the several pauses in ejection reported in kinetic studies of T5 in the absence of PEG (Mangenot *et al.*, 2005).

De Frutos *et al.* did, however, estimate the size of the maximum toroid that could fit inside the capsid, using the empirical value of the ratio between the inner and outer toroidal radii reported earlier from studies by Bloomfield and coworkers (Aracott *et al.*, 1990) of DNA in solution condensed by polyvalent cations into circumferentially-wound/hexagonally-packed toroids with circular cross-sections. This maximum length of unejected DNA is the value that one would measure if the DNA configuration were able to re-equilibrate throughout the ejection process. It is given (in bps) by

$$L_{maxtoroid} = \frac{2.78(R_{inner}^{capsid})^3}{\sqrt{3}h(a_{cond})^2}$$

where R_{inner}^{capsid} is the inner radius of the capsid, $h = 0.34$ nm/bp, and a_{cond} is the interaxial spacing in spermine-condensed DNA. Using $a_{cond} = 2.88$ nm, and 40 nm for the inner radius R_{inner}^{capsid} of the T5 capsid, they obtained an estimate of 36 kbp for the length of the condensed genome equilibrated inside. The same calculation for lambda phage, using the 28-nm internal radius determined by Dokland and Murialdo (1993), leads to an estimate of 12 kbp for the length remaining inside, consistent with the value calculated directly by Tzlil *et al.* using measured DNA bending and self-attraction energies (Rau *et al.*, 1992). The fact that we measure more than twice this length, and a broad distribution of lengths about this average, is a consequence of the large role played by non-equilibrium genome configurations in the case of self-attraction.

It is instructive to compare the ejection studies on lambda and T5 with the series of high-resolution cryo-electron microscopy structural studies by Leforestier and Livolant (2009) in which PEG was used to limit the ejection of DNA from T5 and then 5 mM Sp4+ was added to condense the DNA remaining inside the capsid. By using a range of PEG concentrations they obtain a range of unejected lengths, many of which are short enough for their toroidal condensates to have outer radii smaller than the inner radius of the capsids. The lengths of DNA in the toroids could then be calculated from measurements of their inner and outer radii, again assuming a circular cross-section and hexagonal packing. A stress-free toroid persists until its outer radius becomes as large as the inner radius of the capsid. For longer lengths, the confined condensates are deformed toroids with average interaxial spacings < 2.8 nm. In this way, by using osmolyte to stabilize progressively longer DNA lengths inside the capsid before adding spermine, Leforestier and Livolant were able to conclude that there are about 27 kbp of DNA in the toroid that just fits inside the 40-nm inner radius of the T5 capsid.

Each concentration of PEG in the host (external) solution corresponds to a certain amount of water being sucked out of the phage capsid (from which PEG is excluded). As a result, the water inside is under tension, thereby producing a force resisting the ejection of DNA. As demonstrated by earlier measurements (Evilevitch *et al*., 2003; de Frutos *et al*., 2005; Knobler *et al*., 2009), the higher the PEG concentration the lower the extent of genome ejection. To explore how this mechanism of osmotic suppression competes with the effects of polyvalent-cation-induced DNA self-attraction, we carried out a series of

experiments in which we measure the extent of genome ejection in the presence of both PEG and spermine.

We consider three regimes of spermine concentration, defined with respect to the threshold value (0.15 mM) for mediating a self-attraction: its total absence; a sub-critical value (0.1 mM); and a super-critical value (1.5 mM). We incubate the phage in Sp4+ and PEG, trigger ejection by adding LamB in the presence of DNase I, deactivate the DNase I, extract the unejected DNA, and measure its length in an agarose gel. In this way we determine the fraction of ejected DNA for each specified pair of spermine and PEG concentrations.

Figure 2-2 shows the dependence of fraction ejected vs. osmotic pressure: for sub-threshold [Sp4+] (0.1 mM), see solid squares, which is not high enough to induce an attraction between neighboring DNA duplex portions; and for super-threshold [Sp4+] (1.5 mM), see solid circles. For comparison, we also measured ejection fractions as a function of osmotic pressure for the case of zero spermine (see open circles) and superimpose on those data points a curve (dotted) fit to the measurements reported earlier by Grayson, *et al.* (2006) for lambda ejection in the absence of Sp4+. (The determination of osmotic pressure from PEG w/w% concentration was made using the same calibration employed in our previous studies: see, for example, Evilevitch *et al.*, 2003.) Two of the three curves - those corresponding to no Sp4+ and to non-zero but sub-threshold Sp4+ - are qualitatively similar. In particular, because the DNA self-interaction is purely repulsive in both cases the fraction ejected in the absence of osmotic pressure (i.e., no PEG) is 100%.

The PEG pressure needed to completely suppress ejection is 38 atm for [Sp4+] = 0 and 25 atm for [Sp4+] = 0.1 mM. This is because the presence of a low concentration of polyvalent cation, even though insufficient to mediate a self-attraction, nevertheless contributes to a reduction in the self-repulsion that dominates the capsid pressure. On the other hand, the curve in Figure 2-2 for the DNA fraction ejected for super-critical [Sp4+] (1.5 mM) is qualitatively different. It starts (for no PEG) at a value of about 0.4, corresponding, as discussed earlier, to about 0.6 of the genome remaining in the capsid. Also, because of the much higher spermine concentration, the pressure in the intact phage is as low as 17 atm.

The error bars in Figure 2-2 correspond to the standard deviation in the average DNA length ejected under the corresponding conditions of spermine concentration and osmotic pressure, based on several measurements made for each sample. But the range of lengths observed in each measurement is significantly larger. Recall from the gel in Figure 2-1 that for $\geqslant$ 0.15 mM, and hence for the zero-osmotic-pressure 1.5 mM [Sp4+] sample represented by the left-most filled circle on the lowest curve in Figure 2-2, the lengths of DNA remaining in the capsid range from 25 to 33 kbp. If, on the other hand, an average length of about 28 kbp is made to remain in the capsid because of having imposed a sufficiently high osmotic pressure in the absence of spermine, we expect that the associated range of lengths involved will be smaller because no attractions are operative under these conditions. The same would be true for a combination of a lower osmotic pressure and a sub-threshold spermine concentration, where again there is only a repulsive self-interaction of the DNA. This situation is illustrated in Figure 2-1.B, a gel of
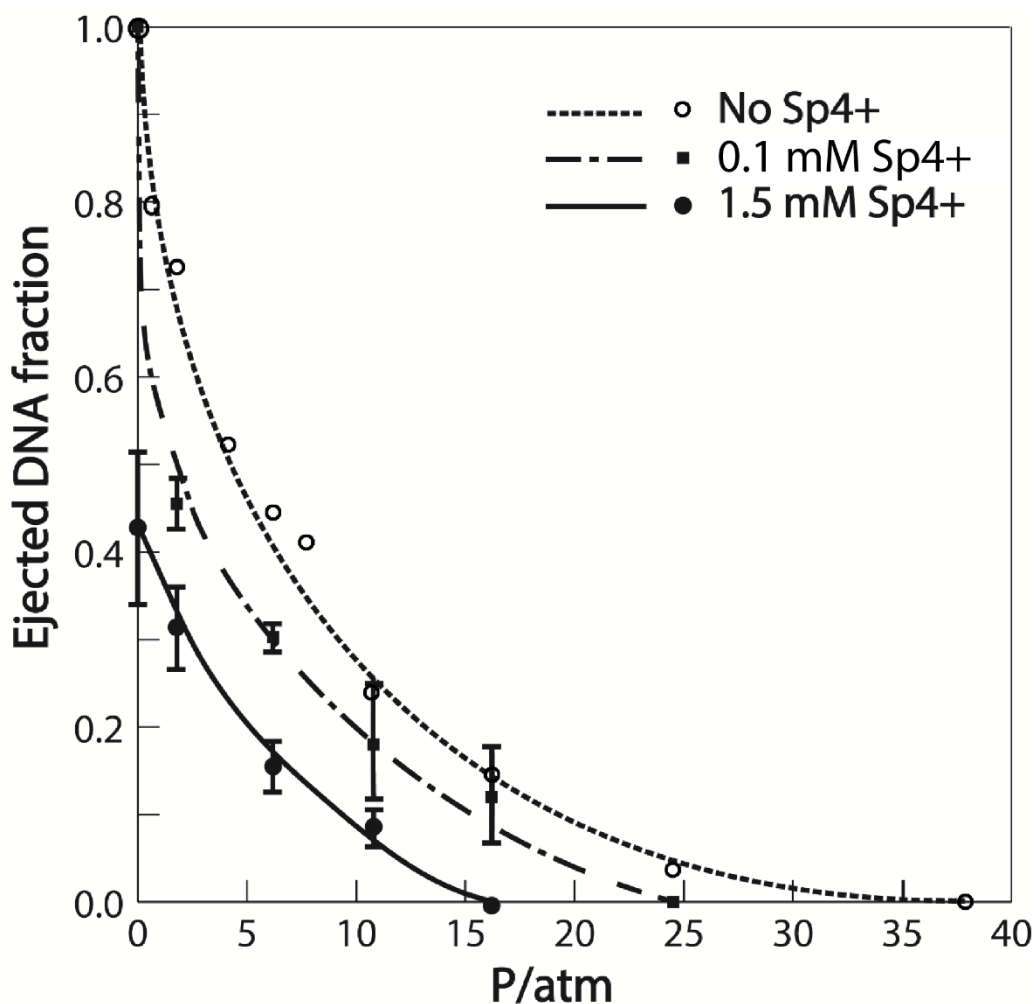
Figure 2-2. The ejected DNA fraction as a function of osmotic pressure (PEG8000 concentrations), for each of three different concentrations of [Sp4+]: open circles, no spermine; filled squares, 0.1 mM spermine; and filled circles, 1.5 mM. In each case the phage are incubated in PEG and/or [Sp4+], ejection is triggered by addition of receptor in the presence of DNase I, the sample is centrifuged to isolate the capsids, and the DNA remaining in the capsids is extracted and run in an agarose gel as described in the discussion of Figure 2-1. The dot-dash and solid curves are fits to our data for sub- and super-critical spermine concentrations, respectively.

extracted, unejected, DNA for these three scenarios. Lane 1 (to the immediate right of the ladder) is for 7.5 atm PEG, no spermine; lane 2 for 2.5 atm PEG, 0.1 mM spermine; and lane 3 for 1.5 mM spermine, no PEG. The average DNA length remaining inside the capsid is 29 kbp in all three cases, with ranges of 2, 4, and 8 kbp for spermine concentrations of 0, 0.1, and 1.5 mM, respectively, as determined from densitometry

traces of the gel. These results suggest, consistent with the situation reported from single-molecule *packaging* measurements (Berndsen *et al*., 2014; Keller *et al*., 2014), that non-equilibrium effects are weaker for more repulsive DNA-DNA interactions. They are also consistent with simulation studies of *ejection* (Marenduzzo, *et al*., 2013) in which reducing the self-repulsion of the confined chain makes it more likely to get stuck ("self-entangle").

It is clear that *in vitro* studies of dsDNA viruses like lambda and T5 have much to teach us about the dependence of genome packaging and ejection on ambient osmotic pressures and salt concentrations, and that future investigations will continue to yield improved physical understanding of these fundamental processes. In particular, the importance of nonequilibrium effects need to be explored further because of the short time scales involved compared to the relaxation times of strongly-confined stiff polyelectrolytes like DNA. These effects show up first in the *packaging* of the genome (Berndsen *et al*., 2014; Keller *et al*., 2014; Petrov *et al*., 2011; Forrey *et al*., 2006, Ali *et al*., 2006; and Spakowitz *et al*., 2005), and then again in its *ejection* (present work; and Marenduzzo *et al*., 2013), and their magnitude increases with weakening of the DNA self-repulsion and with onset of self-attraction.

CHAPTER 3

In vitro DNA Transcription in a Crowded Environment, and Its Role in the Genome Delivery from a Bacterial Virus

Abstract

The bacterial cell cytoplasm is crowded with many large and small molecules which give rise to high osmotic pressure and high viscosity. The vital process of DNA transcription involves a number of proteins and nucleotides and can be affected by the presence of osmolytes and viscogens. In this study, we determined at both "high" concentration and at the single-molecule level how *in vitro* transcription is affected by the presence of various molecules of both high and low molecular weight. Small molecules will decrease the RNA yield mainly because of their effect on viscosity, while macromolecules can increase the yield in bulk by enhancing the binding between RNA polymerase and DNA template, thereby compensating for the viscosity effect. By starting at different transcription stages, we were able to determine that the osmolyte crowding effect that enhances binding plays

a main role during the initiation process, while the elongation step is the one that is affected most by viscosity. On the basis of these findings we design experiments to test the extent to which *in vitro* transcription can complete genome ejection from bacteriophage lambda in the presence of osmolytes and viscogens, and shed light on the possible role of this mechanism *in vivo*.

3.1 Introduction

Previous studies (see Chapter 1) on *in vitro* DNA ejection from bacteriophage lambda inform us about the pressure inside the dsDNA bacteriophage and its function during infection. But pressure alone is not enough to achieve complete genome delivery into the host (Molineux, 2013). The cell cytoplasm is a complicated environment where there are many large and small molecules as well as different cell organelles (Elowitz *et al*., 1999). Accordingly, an osmotic pressure of several atmospheres exists in the intracellular environment. For example, *E. coli* cells have an osmotic pressure of 3-5 atm (Cayley *et al*., 1991). This fact makes the DNA delivery process incomplete if pressure is the only driving force. More explicitly, for bacteriophage lambda, only about half of the genome remains unejected when the pressure inside the capsid drops to the value of that in the bacterial cytoplasm (Evilevitch *et al.,* 2003). To deliver the rest of the genome, there have to be other mechanisms that can provide an ejection – or, pulling – force.

One hypothesis is that transcription of the already delivered DNA genes can pull the remaining portion out of the capsid. Transcription is a process initiated by an enzyme called RNA polymerase (RNAP) that binds to a certain ("promoter") sequence of DNA and then moves along the DNA to produce a complementary strand of RNA. As it moves, it exerts a force against its transcribing direction. In a particularly well-studied example, bacteriophage T7, it has been demonstrated (Molineux *et al.*, 2006) that transcription of an "early" gene, i.e., one delivered by the initial ejection step, is necessary to complete delivery of the genome. Specifically, the phage internal pressure only drives the first several hundred base pairs of genome out of the capsid, but that portion contains the promoter sequence for E. coli RNA polymerase and the rest of the genome is then pulled out by the transcription process with E. coli and virally-encoded (T7) RNA polymerases (Kemp *et al.,* 2004). This raises the question: for other dsDNA bacteriophages, and in particular for lambda, can transcription by *E. coli* RNA polymerase contribute to the delivery of the remainder of the DNA after capsid pressure has fallen to that of the host cell cytoplasm?

To test this idea *in vitro*, we first need to control the extent of DNA ejection from lambda, so that we can then initiate transcription and measure the ejected DNA length difference before and after transcription. Tuning the DNA ejection length can be achieved by setting up an *in vitro* osmotic suppression system with high-molecular-weight PEG 8000 (Evilevitch *et al.,* 2003).

Transcription is a complicated process including the many different steps depicted schematically in Figure 3-1 (Pearson Education Inc., 2012): initiation, in which RNAP diffuses in the solution, reaches the DNA template, moves along it and finds a specific promoter sequence, unwinds that part of the double helix by breaking the hydrogen bonds between complementary nucleotides, and forms a structure called the RNAP-DNA "open complex" (step 1 in Figure 3-1). NTPs are then recruited from the surrounding solution and RNA complementary to the template strand is produced (step 2 and 3 in Figure 3-1). Before producing full-length RNA, many truncated RNAs – referred to as "abortive products" (around 1-10 nucleotides long) – are usually synthesized by DNA "scrunching" while RNAP stays immobile on the promoter. Once RNAP overcomes an energy barrier and leaves the promoter sequence to move downstream of the template, transcription goes into the next step: elongation, in which full-length RNA is produced and freed into solution (Goldman *et al.,* 2009) as RNAP moves along DNA and recruits NTPs for RNA synthesis. The last step is termination, wherein transcription is completed upon RNAP encountering a special (weak-RNAP-binding) sequence and falling off the template. If there is no termination site, transcription will not stop until the RNAP runs off the end of the linearized DNA template.

Most *in vitro* transcription studies to date have been performed in buffered solution with low-concentration DNA template, purified RNAP, and RNA nucleotides, in the absence of any other molecules. But transcription processes are expected to behave differently in high concentrations of viscogen and/or osmolyte, or in the crowded environment of the
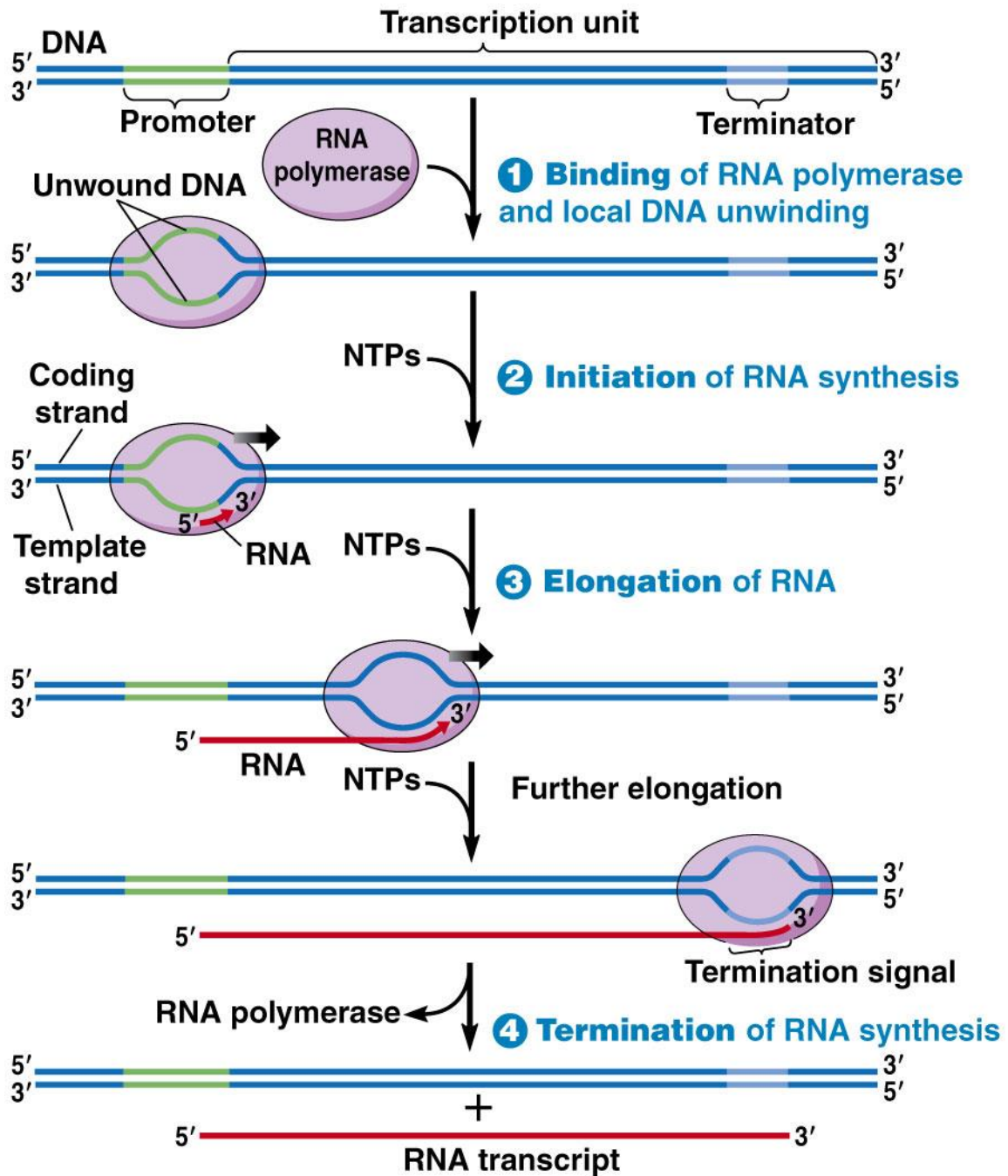
Figure 3-1. Transcription processes (Pearson Education Inc., 2012): binding of the RNA polymerase at a promoter site on DNA to be transcribed, unwinding of a region of the DNA helix and initiation of transcription, and synthesis of an RNA transcript complementary to the template strand of the DNA. The transcript elongates as the polymerase proceeds down the DNA until the polymerase reaches the termination site, falls off and finishes the transcription.

cytoplasm (Demidenko *et al.*, 2011): viscogens will impede the diffusion of the components while osmolytes will produce a crowding effect that alters the effective binding of RNAP to DNA. So in order to better understand these processes it is necessary to study transcription under controlled viscous and high osmotic stress conditions.

We carried out experiments on four viscogens (substances affecting the viscosity of the solution), which also act as osmolytes – PEG8000, PEG 3350 (macromolecules), PEG 400 and glycerol (small molecules) – and with two different RNAPs, T7 RNAP and E.coli RNAP to eliminate any specific chemical effect or specificity of RNAPs. In addition, experiments were also performed in two different concentration regimes: "high" concentration in which RNAPs bind and transcribe multiple times on DNA templates, and "low" concentration where there is only one-time transcription – i.e., once the RNAP finishes transcribing one DNA and falls off, it can no longer find another template during the time of the experiment. Different from the high concentration conditions where we initiate the reaction by mixing RNAP, DNA template and NTPs, the low concentration experiments were started with a pre-made DNA-RNAP open complex (Kim *et al.*, 2011) so that we are able to study each step (initiation vs. elongation) separately during the transcription process and determine the contribution of each factor (mainly viscosity vs. osmotic pressure) to the different steps.

We found at high concentration that the presence of macromolecules can increase the RNA yield in the case of both E.coli and T7 RNAPs, while small molecules have an

opposite effect. These effects are more obvious for higher-molecular-weight *E. coli* RNAP than for the smaller T7 RNAP. The transcription yield enhancement due to the macromolecules can be understood in terms of an increase in the dissociation-association equilibrium constant of binding between template and RNAP. At *low* concentration, on the other hand, by studying two different high-molecular-weight polymers and two low-molecular-weight small molecules, we learned that all of them *de*crease the amount of RNA product, with this effect being most pronounced for the small molecules. In parallel, the low-concentration case was studied with single-molecule FRET measurements carried out by the group of Shimon Weiss, and similar results were found. Viscosity – viscous drag – in this situation is the major reason behind these lower yields of RNA, and is more pronounced for the smaller viscogens because of the competition between those molecules and NTPs at the active transcription site and the resultant slowing-down of NTP incorporation.

3.2 Materials and Methods

Transcription of 200bp DNA by E.coli / T7 RNA polymerase in the presence of PEG 8000/glycerol in high concentrations

Transcription samples with a total volume 20µL were prepared by mixing 4µL of 5x *E. coli* RNAP / 2µL of 10x T7 RNAP transcription buffer (NEB), 1µL of 1.33mg/mL (final concentration ~0.5 µM) 200bp DNA made by PCR reaction containing pL promoter sequence (5'-TTGACATAAATACCACTGGCGGTTATACTGAGCACAGAAGTAGACCA

CGGAACGAGGTTC-3') from bacteriophage lambda/T7 promoter (5'TAATACGACTCAC

TATAGGGAGA-3'), 1 unit (1 μL) of RNase inhibitor and 2 units (2 μL) of *E. coli* / T7 RNA

polymerase in different osmolyte conditions: 0, 2, 3, 4, 8, 12, 16, 20% (w/v) of PEG 8000,

or 0, 5, 10, 20, 30, 40 (v/v) of glycerol, respectively. Samples were incubated at 37º C for

30 min to allow RNA polymerase to bind to the DNA template. Then 1μL of each 75mM

solutions of ATP, UTP, GTP and CTP (with final concentration ~4mM of each NTP) were

added to each sample followed by 1hr incubation at 37º C to allow transcription to occur.

After transcription, 1 unit (1 μL) DNase I (RNase-free, NEB) was added to digest the DNA

template and terminate the transcription. The amount of RNA products was visualized by

running a 1% agarose gel, stained with Ethidium Bromide and quantified by ImageJ. Each

band intensity was converted into relative yield by taking the ratio of the band intensity in

each PEG/ glycerol sample to the one which does not contain any PEG or glycerol.

Transcription of 75bp DNA by E.coli RNA polymerase in the presence of osmolyte/viscogen at low concentrations

Reaction in nanomolar concentration was started with the DNA-RNAP open complex

prepared by the Weiss group (Kim *et al.,* 2011) in KG7 buffer (40mM HEPES, 10mM

MgCl2, 1mM DTT, 100 μg/ml BSA, pH 7.0). The DNA template is 75bp long of which

~40bp is the promoter sequence from E. coli lacCONS+2 with an added 20dA designed

for   FRET   measurement   at   the   end   of   the   template   (5'-

AGGCTTGACACTTTATGCTTCGGCTCGTATAATGTGTGGAATTGTGAGAGCGGAAA

AAAAAAAAAAAAAAAAAAA-3'). These conditions determine that transcription can quickly

go into elongation and go through only one round of RNA synthesis due to the low

concentration. For a 50µl reaction, the RNAP open complex was added to the final concentration 1nM and each rNTP concentration is 500nM (each component is 1000 times lower than the previous high concentration test). Samples were prepared under the same osmolyte conditions as in the high concentration measurements but more types of molecules are used: PEG8000, PEG3350, PEG400 and glycerol. After 15 min incubation, which is long enough for a single round of transcription at room temperature, DNase was added to each sample to digest the template and stop the transcription reaction. OliGreen (Life Technologies O7582), a green-fluorescent nucleic acid quantifying agent, was added to each sample to measure the amount of RNA. Fluorescent signals were obtained with a Tecan Infinite M1000 Plate Reader (excitation wavelength 480 nm; emission wavelength 520 nm).

3.3 Results

At high concentrations, PEG8000 increases the RNA yield while glycerol decreases the RNA yield for both RNA polymerases

Figure 3-2 shows the amounts of RNA product transcribed in solutions with PEG 8000 concentrations corresponding (Ninni *et al.,* 2003) to viscosities ranging from 1 to 17 (relative to the viscosity of the PEG-free solution, for both E. coli RNAP and T7 RNAP). As the relative viscosity due to PEG 8000 is increased from 1 to 9, the relative transcription yield increases 2.5 fold and stabilizes for E. coli RNAP. The change is slightly smaller for T7 RNAP transcription, but also increases monotonically with PEG concentration (viscosity).

Figure 3-2. PEG8000 increases the yield of *in vitro* transcription by E.coli / T7 RNAP. Transcripts were produced in reactions containing E.coli / T7 RNAP, 200bp DNA template with pL/T7 promoter and the different viscosities associated with different concentrations of PEG 8000. Relative efficiencies are calculated from densitometric scans of each RNA band divided by that in the control sample which does not contain any PEG 8000. Viscosity relative to pure water under the same condition is calculated from the literature calibration (Ninni *et al.,* 2003). Error bars are one standard deviation based on at least three times of measurement.

Glycerol is found to have a completely opposite effect: from 0 to 40% (v/v) which corresponds to relative viscosities from 1 to 4 (with respect to pure water) (Segur *et al.,* 1953), transcription yields for both *E. coli* and T7 RNAP *de*crease monotically as the glycerol concentration increases. And, similar to the situation with PEG8000, the effect on RNA yield for E. coli RNAP is larger than for T7 RNAP, probably because of the larger size of the former (450kDa for *E. coli* RNAP and 98kDa for T7 RNAP).
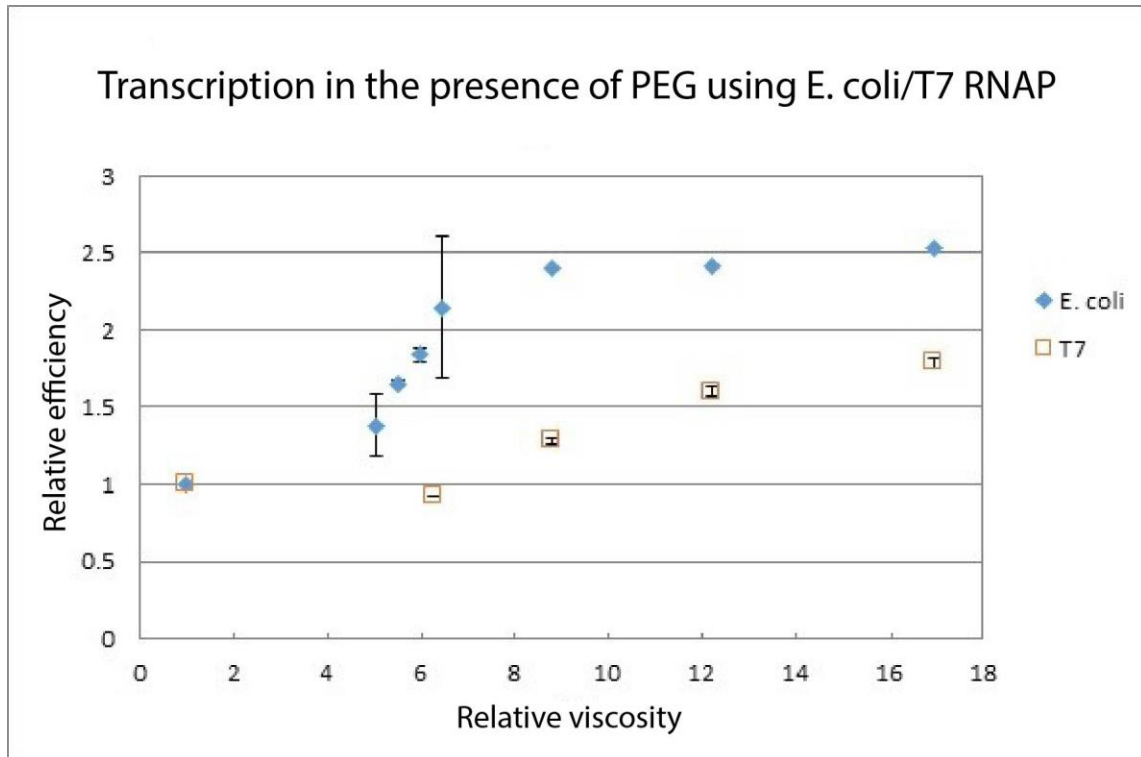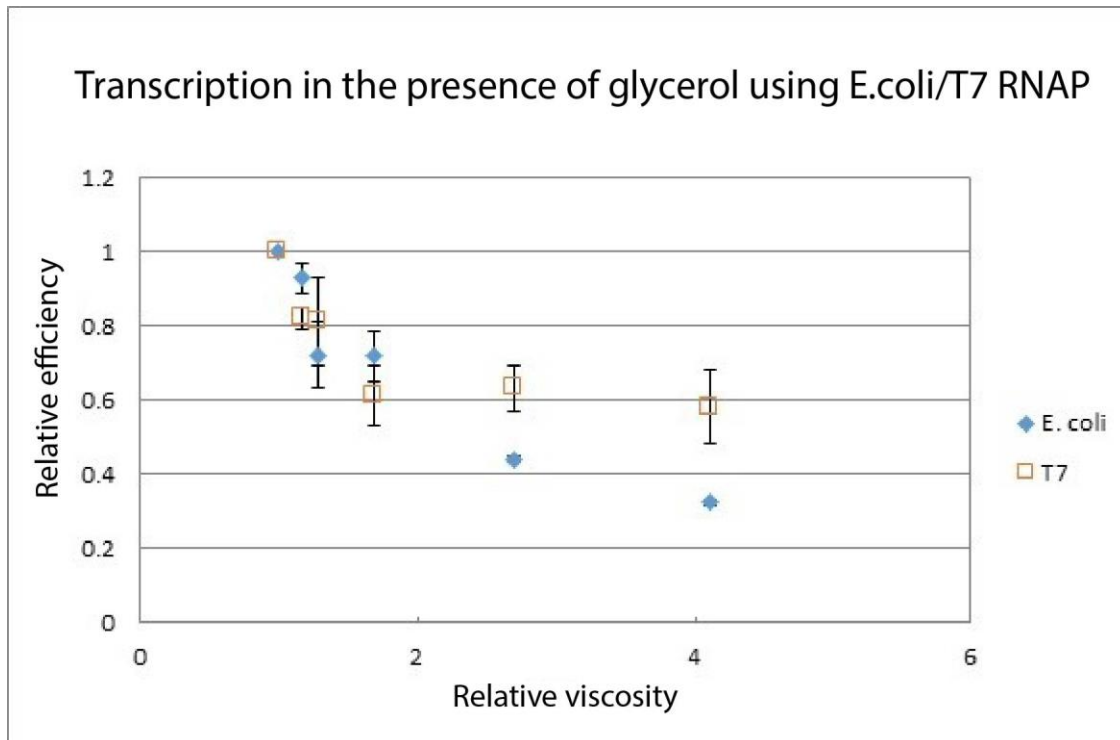
Figure 3-3. Glycerol decreases the yield of *in vitro* transcription by E.coli / T7 RNAP. Transcripts were produced in reactions containing E.coli / T7 RNAP, 200bp DNA template with pL / T7 promoter, and different viscosities (relative to glycerol-free solution) introduced by increasing concentrations of glycerol (Segur *et al.,* 1953). Relative efficiencies are calculated by using densitometric scans of each RNA band divided by that in the control sample which has no glycerol.

At low concentrations, RNA yield by E.coli RNAP decreases for all four different viscogens / osmolytes

In the above high-concentration reaction conditions the overall transcription process includes multiple rounds of initiation and elongation steps as the polymerase successively falls off a template, diffuses around, and binds to the promoter sequence to of a new template to start a new transcription. Thus the relationship between RNA yield and

viscogen / osmolyte concentration arises from the combined effects of viscogen / osmolyte in each of the steps.

Collaboration with the group of Shimon Weiss, who have been using single-molecule Fluorescence Resonance Energy Transfer (FRET) detection of transcription products, made possible the investigation of individual steps in the underlying process. We start with the DNA-RNAP open complex prepared by them and study transcription at low concentrations of all reactants. Thus there is no chance for the rebinding of the RNAP. In this way, we can study the viscogen / osmolyte effect on the NTP incorporation process alone. From the measurement of the fluorescence intensity, we determined the relationships of RNA yield to osmotic pressure/viscosity shown in Figures 3-4 and 3-5. In these studies we used four different, chemically-inert, substances, two of them small molecules and the other two polymers. In Figure 3-4 we have plotted the RNA yield against the osmotic pressure showing that in all cases the yield decreased with increasing osmotic pressure from 0 to 40 atm (Money, 1989). Comparing large molecules with small molecules, we find little difference between them.

But if we plot the RNA yield against viscosity (Gonzllez-Tello *et al*., 1994), we see (Figure 3-5) two different populations: In particular, small molecules including PEG400 and glycerol have a much larger effect on decreasing the RNA yield than do larger molecules like PEG 3350 and PEG 8000.
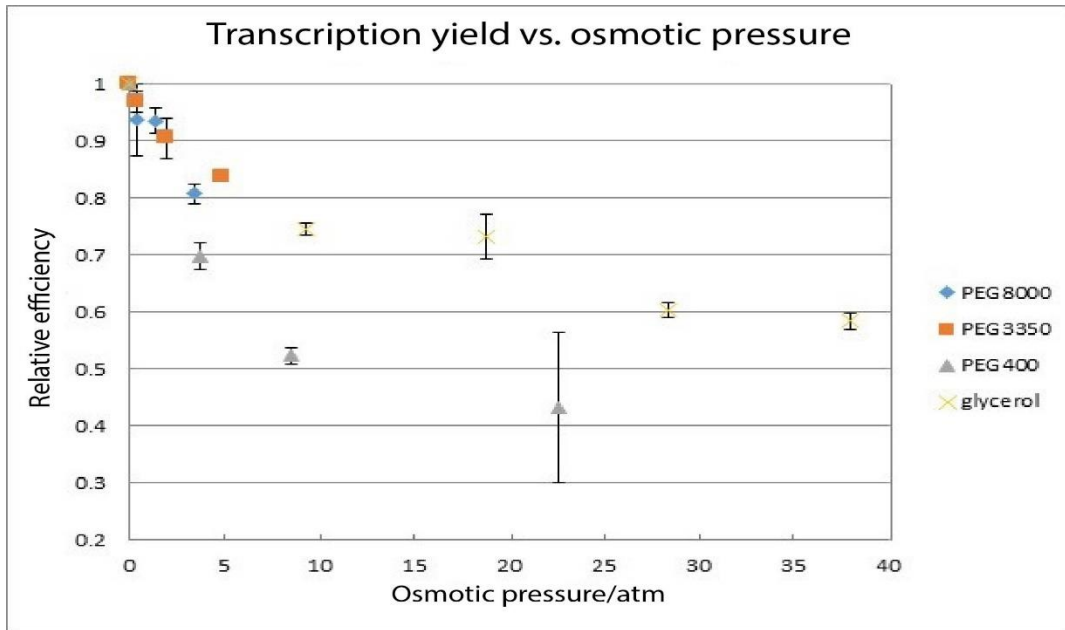
Figure 3-4. Effect of osmotic pressure (Money. 1989) on the yield of RNA transcribed by E. coli RNAP, for low concentrations of DNA. Transcripts were produced in reactions containing the transcription open complex comprised of E.coli RNAP and 75bp DNA template with a lacCONS+2 promoter. Relative efficiencies are calculated as explained in Figures 3-2&3.
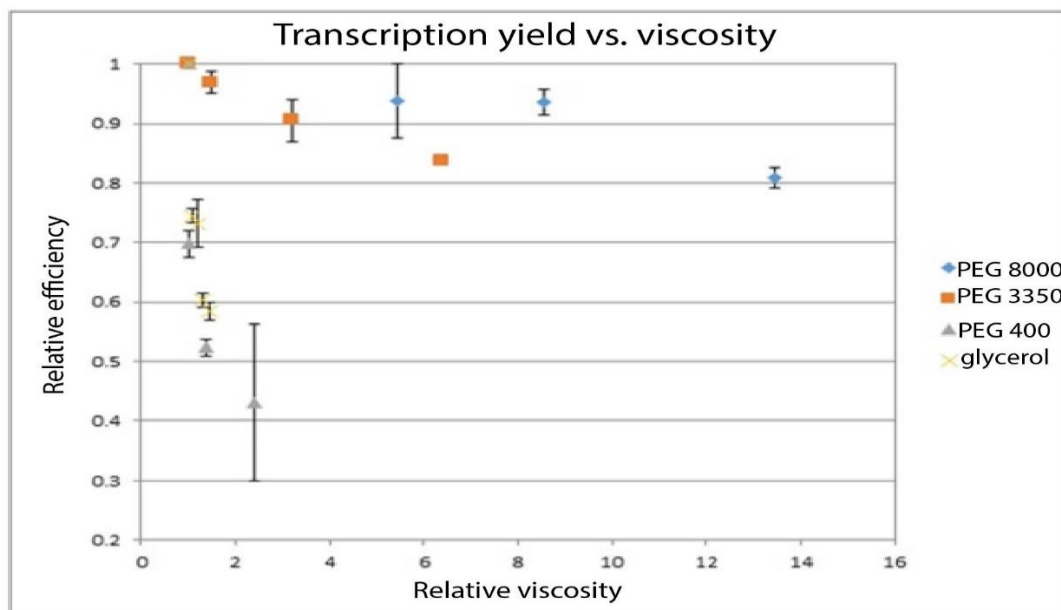
Figure 3-5. Effect of viscosity on RNA yield by E. coli RNAP. Transcripts were produced in low-concentration-DNA solutions containing the transcription open complex comprised of E.coli RNAP and 75bp DNA template with a lacCONS+2 promoter at different relative viscosities introduced by different molecules. Relative viscosity is calculated based on the literature (Gonzllez-Tello *et al.*, 1994).

3.4 Discussion

During the transcription process, large molecules, including the DNA template and the RNA product, need to continuously perform translational and rotational movements; small molecules such as NTPs and pyrophosphate keep diffusing in the solution (Demidenko *et al*., 2011). There are also molecular interactions during initiation and elongation between polymerase and template, along with conformational changes, all of which can be affected by the viscosity and osmotic pressure of the environment. In this study we work with both big molecules and small molecules in order to study the effect of crowding on molecular interactions and of viscosity on diffusion, as well as other possible factors influencing transcription.

Effect of different osmolytes / viscogens on transcription at high conentration

High concentrations allow multiple rounds of binding of both template and polymerases and thus several rounds of transcription. In this case it is not entirely clear why the viscogens / osmolytes glycerol and PEG 8000 appear to have opposite effects on RNA transcription yield. By increasing the viscosity of the solution, both molecules give rise to higher drag forces on all moving particles and thus hinder RNA polymerase from reaching or moving along the DNA template, thereby inhibiting transcription (Demidenko *et al.,* 2011); this is consistent with the monotonic decrease in RNA yield observed in glycerol shown in Figure 3-3. But the situation is more complicated in the case of PEG 8000. In

addition to the viscosity effect, slowing down all motions, there is also a "crowding" effect induced by high-molecular-weight molecules which can increase the yield of RNA. This crowding is associated with an attractive ("depletion") force between the reactants (in this case the DNA template and RNAP) and make the transition from initiation into elongation faster. More explicitly, the conformation after binding excludes less volume to the crowder PEG 8000 and thus alters the dissociation-association equilibrium constant (Morelli *et al.,* 2011), favoring association. This mainly affects the initiation step where two macromolecules need to bind and undergo a conformational change to proceed to transcription. So the increasing yield observed in PEG 8000 at high concentrations (Figure 3-2) suggests that the effect of increasing viscosity (slowing the motion of RNA polymerases) is offset by the increasing strength of association between RNA polymerase and DNA template due to the crowding effect. It also agrees with the fact that transcription is limited by the transition step from individual reactants to an active open complex. So for different types of reactions, introducing polymer can in theory (Minton *et al*., 1989) slow down fast associations (diffusion is the rate-determining step) and accelerate slow associations (association and transition are the rate-determining steps). But specifically for PEG 8000, another factor has been suggested that can affect the reaction: an attractive force between PEG and nonpolar or hydrophobic sidechains on the protein surface that can compensate for the excluded volume contribution to the PEG-protein interaction (Bloustine *et al.,* 2006; Tubio *et al*., 2004; Winzor *et al*., 2006). But clearly in our case this interaction is less strong than the depletion effect.

Effect of large / small molecules under low concentration conditions

Starting with the open complex, we are able to focus on the transition from initiation to elongation and beyond. Four different molecules were studied, and all showed a similar result – decrease of the transcription yield with increasing concentration of osmolyte / viscogen. Because the RNAP is already bound at the outset of reaction, and only a single round of transcription occurs, depletion forces play little role here and viscosity effects dominate. Previous studies had also shown that the elongation step can be largely slowed by high viscosity in an environment that limits the diffusion of NTP and by-product pyrophosphate reagents (Demidenko *et al.*, 2011). But, interestingly, when we plot the transcription yield against viscosity, the results fall into two groups: small molecules have a larger effect in decreasing the yield than do large molecules. It is possible that because of their size, small molecules can compete with free NTPs for occupying the position where NTP gets incorporated into RNA, and thus slow the transcription. Another possibility is that the presence of large molecules may help lower the activation energy between initiation and elongation states, thereby increasing the transition rate and counteracting the effect of the viscosity. But since our method uses a fluorescent dye to test the total intensity and calculate relative yield, we have no idea about the length distribution of our RNA product, i.e., we cannot tell whether the presence of different molecules will affect the accuracy of the RNAP transcription starting point. The reason why we think it might affect the transcription starting point is based on the preliminary results obtained by the Weiss group, where they label the DNA template and RNAP with FRET pairs in a way that allows monitoring of the shape of the open complex with and without crowding molecules. They have shown that macromolecules like PEG 8000 can change the conformation of the open complex. But will this conformational change make

the transition from initiation into elongation easier? Or can it affect the transcription starting point and thus interfere with transcription accuracy? Further studies are needed to answer these questions.

Thought about transcription pulling hypothesis of lambda DNA

To try to answer the question "How is DNA translocated from virus capsid into host cytoplasm?" several different scenarios have been considered: diffusion-based ejection; pressure-based ejection; reversible protein binding; and transcription of early genes. Relevant to the case of T7, which supports the transcription mechanism, it has been shown that transcription by RNAP can generate bout ten picoNewtons of force, which is sufficient to provide a translocation rate of around 50-60bp/s (Grayson *et al*., 2007). Is this mechanism possible for bacteriophage lambda? Can transcription pull the DNA out when pressure no longer generates a net force? To test this idea *in vitro*, we need to stop the ejection about half way (see below) and then initiate transcription followed by the measurement of DNA exposed length. To start and continue transcription, RNAP must bind to a special sequence to start heading in the right direction and surmount termination sites along the template (if there are any), in order to reach the capsid and start pulling out DNA.

Bacteriophage lambda has a 48.5kbp double-stranded DNA genome with 12-base single-stranded overhangs ("cohesive sticky ends") at both ends (Sanger *et al*., 1982). These two single-stranded parts are complementary to each other and will pair up to circularize

the DNA once both ends are available, i.e., once the DNA is fully ejected. Along the DNA there are two promoter sites, pL and pR (L and R indicate the direction of transcription, left and right), which are located about one third of the way from the right end (the end which is always packaged last and ejected first). Once the E. coli RNAP binds to the promoter, transcription will start until the polymerase meets a terminator site and falls off. A brief layout of terminators and promotors on lambda DNA is shown below in Figure 3-6 (Echols *et al.,* 1978).

tJ	*b region*	*tL3*	*tL2*	*tL1*	*N*	*pL*

| 0 | 40 | | 64.4 | 68.9 | 71.1 | 73.5 | 100 |
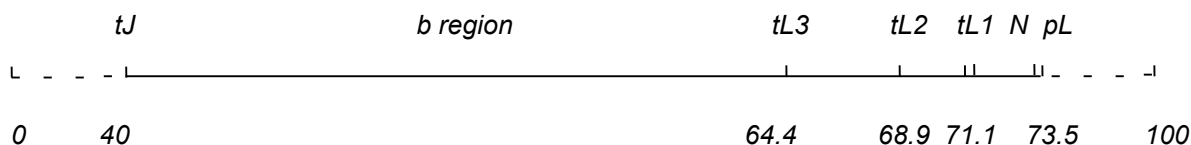
Figure 3-6: Partial genetic map of lambda showing the positions of the promoter and terminators for transcription leftward (towards the capsid). If RNAP can go through all the three terminators shown in the picture, transcription can then proceed all the way to the end of the b region, which is around 40 percent from the left end of the genome (Echols *et al*., 1978).

There is a special gene called N located to the left of the pL promotor and it is the first gene that is transcribed, and by the host cell (E. coli) RNAP. Gene product N, a small protein with molecular weight 14kDa, is an antiterminator that allows transcription to continue through termination sites (Greenblatt *et al*., 1980; Van Glist *et al.,* 1997). To function as an antiterminator, it also needs the DNA region called Nut (Vieu *et al*., 2004) between the pL promoter and the N gene. Nut acts through its messenger (m) RNA. More explicitly, during transcription, N is thought to bind to a small hairpin component of nut

mRNA, known as BoxB, Figure 3-7 so that N can capture the RNAP elongation complex and transform RNAP into the termination-resistant form (Das *et al*., 1992).

NutL site

5'   BOX A   BOX B   3'

73.5 (pL)   73.3(N)

Figure 3-7: Partial RNA transcript of lambda DNA between pL promoter and N gene showing the position of the nutL site. It contains two regions, BOX A (8nt) and BOX B (17nt). The 10nt upstream from BOX A and the 1nt downstream from BOX B are necessary for effective anti-termination; Along with the 7nt spacer between BOX A and BOX B, the minimal length of nutL is 43nt (Das *et al.,* 1992).

But these two factors are not sufficient to transcribe all the genes in lambda. In host cell E. coli, there are several additional protein factors that are important: NusA, B, E, G. These are *E.coli* transcription elongation factors that allow N protein to bind tightly to the elongation complex and lead to efficient suppression of terminators located thousands of base pairs away from the promoter sites.

NusA: During transcription this host-cell protein generally reduces the RNA elongation rate and allows the pause of RNA polymerase to enhance rho-dependent and rho-independent termination (Friedman *et al*., 1974; Mah *et al.,* 1999). Rho is a protein in E.coli involved in transcription termination (Howard *et al.,* 1977). For rho-dependent termination, the rho factor needs to access the RNA 50-100 nt upstream of the terminating site and move to the paused polymerase. In N-assisted anti-termination system, the pause of RNA polymerase at N-recognition sites (nut sites) mediated by NusA facilitates

the action of N protein on RNA polymerase. But when the concentration of N is high enough, and under low salt conditions, N alone in the absence of nut and NusA can suppress the termination. This has been demonstrated by a series of *in vitro* transcription experiments involving different concentrations of N protein and lambda DNA pieces as a template with or without NusA (Mason *et al*., 1992). More specifically, if the concentration of N protein reaches 500nM (versus the normal 50nM in natural system), RNA polymerase can go through a termination site which is several hundred base pairs away from the promoter site.

NusB and NusE: These two proteins form a complex and bind to RNA containing the consensus BOX A sequence. Interaction of NusB with BOX A facilitates the interaction of NusE with RNAP during N-mediated anti-termination. But NusE itself can interact with RNAP independently without NusB or BOX A. This has been shown by *in vitro* experiments (Keppel *et al.,* 1974). By omitting one of the other factors (NusA, NusB and NusG), it is found that NusE can still form a complex with RNA polymerase and that this interaction is affected only when the amount of NusE is low. So, other factors can enhance the interaction between NusE and RNA polymerase (Friedman *et al.,* 1976).

NusG: NusG is required for efficient rho-dependent termination. But it is also a requirement for N-mediated anti-termination, as demonstrated in an experiment in which anti-termination is reduced *in vivo* without NusG. It contributes to the stability of N binding

to the lambda elongation complex, by being part of the complex (Downing *et al*., 1990; Sullivan *et al*., 1992; Li *et al.,* 1992; Zhou *et al*., 2002).

Accordingly, in our system, if we want transcription from the pL promoter to continue all the way to the left end of the b region (around 16k nt away from the promoter), the polymerase has to pass through several terminators such as tL1 (rho-dependent, 71.1% from the left end of lambda DNA), tL2 (rho-independent but a weak terminator, 68.9%) and tL3 (rho-independent, 64.4%) (see Figure 3-6). And the farthest terminator lies thousands of base-pairs away from the promoter. What factors should be in the system in order to ensure that the RNAP travel far enough to pull out the DNA? Ideally, we need to introduce all the factors listed above: N protein, nut sequence, and Nus factors but no rho factors (Stephen *et al*., 1992).

But even if we could succeed in demonstrating that RNAP pulls out the DNA *in vitro,* is this what is happening in the lambda life cycle? Early work by Herskowitz suggests that during the bacteriophage lambda infection process, transcription will not start until the gene circularizes, i.e., the two sticky ends meet and pair up with each other to form a circle (Herskowitz *et al.,* 1973). That means both ends of the DNA need to come into the cytoplasm first. If this is strictly true, transcription would not be involved in pulling the rest of the genome out. But so far no evidence has been found to support the statement that circularization happens first and transcription cannot occur without circularization.

In addition, it is a fact that the structural genes of lambda are located on the left of the DNA that come out last. These are transcribed by the polymerase that binds to the pR promoter and moves to the right side of the gene (Deighan *et al.,* 2007). Only after the DNA ends bind together can those structural genes be reached by the polymerase moving rightward. But will the RNAP binding to the pL promoter keep transcribing the noncoding strand to produce redundant RNA and thus pull all the DNA out? It seems most unlikely. For these reasons, the transcription hypothesis for DNA delivery doesn't appear to apply to bacteriophage lambda.

3.5 Future work

Although the above-described measurements at high and low concentrations, and the single-molecule assays of the Weiss group, allow us to study the effect of different molecules on transcription in considerable detail, there are still many studies needed to fully clarify the mechanisms with which they affect transcription. By changing the pairs of positions of the FRET labeling dyes on the DNA template (upstream, downstream, and the relative positions between two dyes) as well as on the RNAP, we can determine potential different conformational changes of the template and polymerase due to the presence of different osmolyte and viscogen molecules. If we want to focus on the crowding effect brought by large molecules, other alternatives such as dextrans and ficoll rather than PEG should be chosen in order to avoid the potential interaction between PEG and protein. It is possible that the addition of other molecules may affect the shape of the open complex and thus change the transcription rate and accuracy. By analyzing

the length and yield of abortive products, we can also learn whether different molecules

have different effects on the open complex stabilization and about the activation energy

to transit into elongation step. Such findings will help us understand more about the

transcription process in the live cell environment.

CHAPTER 4

Does Bacteriophage P22 Eject All Its Internal Proteins Before Its Genome?

Abstract

Double stranded (ds) DNA bacteriophages are highly pressurized due to the electrostatic self-repulsion and bending of the densely-packed, negatively-charged, DNA. Several studies have shown that this pressure provides a large force driving ejection of the genome from its capsid. In *Podoviridae* there are several ("E") proteins contained in the capsid along with the DNA genome. These internal viral proteins are needed to facilitate passage of the viral DNA through both the outer and inner membranes of the host after the virus binds to the outer membrane. But essentially nothing is known about how and when the internal proteins come out of the virus. To study this question, and to progressively control the extent of ejection from P22, one of the *Podoviridae*, we set up an in vitro osmotic suppression system with high-molecular-weight polyethylene glycol (PEG) mimicking the cell environment. We find that Outer Membrane Protein A (OmpA) from Salmonella significantly enhances the rate of DNA ejection *in vitro* in the presence of lipopolysaccharide (LPS). Further, we observe that while the DNA is ejected in the

presence of LPS, no ejection of the E proteins occurs unless OmpA is also present, and that their ejection is largely complete before any of the genome is ejected. This finding helps us understand more about how P22 infects its host cell and, in particular, about possible roles that the E proteins play during infection.

4.1 Introduction

Tailed double-stranded (ds) DNA bacteriophages need to deliver their genomes into their bacterial hosts to initiate infection. The detailed mechanism of this process is not fully understood for any virus or phage. Generally, proteins located in the phage tail first contact the surface of the bacterium and the phage diffuses along the surface until it finds a specific receptor. Upon binding to the receptor, the phage tail undergoes a series of conformational changes that release the DNA and further transport it from the capsid into the cytoplasm of the host (Poranen *et al.,* 2002). In order to overcome the defense barriers of the bacteria – outer and inner membranes along with the periplasmic space in between – bacteriophages use different strategies based on their tail morphology. Both *Myoviridae* (long, contractile tails) and *Siphoviridae* (long, non-contractile tails) have been well studied (Leiman*,* 2012; Davidson *et al*., 2012). *Podoviridae*, a family of bacteriophages with tails shorter than the width of the periplasm, cannot directly use their tails to penetrate both membranes and there is no generalized mechanism of infection known to date.

P22 is a member of *Podoviridae* that infects *Salmonella enterica*. It is a dsDNA phage with a 43.5 kbp genome that is packaged via a headful mechanism (Casjens *et al*., 1988) into an icosahedral procapsid formed from assembly of the coat, scaffolding, and portal proteins (King *et al*., 1976). There are three internal proteins (called "pilot, "ejection", or "E" proteins) packaged inside P22 procapsids, all incorporated by scaffolding proteins in the early stages of assembly: gp16 (gene 16 product), gp20 and gp7, each with 10-20 copies (Israel, 1977). A short tail "machine" is then connected at one of the five-fold vertices to complete the P22 virion. The majority of these structural proteins have been identified in recent cryoEM investigations (Tang *et al.,* 2011; Lander *et al.,* 2009). Although density in reconstructions of the portal had been ascribed to the E proteins, the most recent asymmetric reconstructions show (Tang *et al.,* 2011) this to be incorrect and the location of the three E proteins within the mature virion remains unknown.

Along with their DNA, most phages eject proteins that have been packaged into the capsid, and this is generally essential for infection (Molineux *et al*., 2013). Although the exact roles that each of these E proteins plays in infection have not been determined, some information is available. Gp16 acts within 10 minutes of infection, and when gp16-deficient phage are used, cells continue to divide normally and do not replicate the P22 genome. Co-infections showed that gp16 can work *in trans;* complementing a gp16-deficient particle, indicating its early function occurs once it is ejected into the host (Hoffman *et al*., 1975a; 1975b). Furthermore, gp16-deficient phage do not induce superinfection exclusion response and gp16 is not part of the replication machinery (Israel *et al.*, 1972). Gp7 and gp20 often co-purify with gp16, and gp7 ejection does not occur

when gp16 is absent (Israel, 1977). Indirect evidence supports a membrane-breaching role because purified gp16 disrupts dye-loaded, lipid vesicles (Perez *et al*., 2009). It has been proposed that the E proteins protect the DNA in the periplasm during infection (Israel, 1977), but function and location post infection has yet to be determined experimentally. Taken together, these data suggest that the E proteins, gp7, gp16, and gp20, may be linked to the efficiency and dynamics of DNA ejection across host membranes.

If the receptor is known and can be solubilized, phage ejection can be triggered *in vitro*, in which case the extent of DNA ejection can be controlled by the presence of an osmolyte in the surrounding buffer solution (Evilevitch *et al.* 2003; Castelnuovo and Evilevitch, 2007). More explicitly, each concentration of high-molecular-weight (8000) PEG in the host ("external") solution corresponds to a certain amount of water being sucked out of the phage capsid (from which PEG is excluded). As a result, the water inside is under tension, thereby producing a force resisting the ejection of DNA. Such osmotic suppression studies done with dsDNA phages like lambda and T5 have shown that the virus capsid is highly pressurized as a result of the electrostatic self-repulsion and bending of the densely-packed, negatively-charged, semi-rigid DNA (Riemer *et al*., 1978). In these *in vitro* experiments, in which pressure is the only driving force for DNA ejection, the length of DNA ejected can be tuned by the osmotic pressure difference between the outside and inside of the capsid (Evilevitch *et al*., 2003). In this study we use osmotic suppression to examine for the first time the ejection of both the DNA and the E proteins from P22, which sheds light on ejection mechanisms in *Podoviridae*.

4.2 Results/Discussion

For P22, the O-antigen portion of lipopolysaccharide (LPS) located on the surface of the host *Salmonella enterica* serovar Typhimurium works as a primary receptor for infection (Iwashita *et al*., 1973); Andres *et al.* (2010) demonstrated that LPS can trigger ejection of DNA from P22 *in vitro*, which makes it possible to study the process with the osmotic suppression technique.

Osmotic suppression experiments show release of DNA from P22 is inhibited at a pressure of 16.8 atm

Using LPS to trigger ejection, we determined the fraction of the DNA remaining unejected in the presence of an osmotic pressure by separating the capsids from the ejected DNA, which was degraded by DNase, and then recovering and analyzing the DNA remaining in the capsids. The results are indicated by the open circles in Figure 4-1. For all samples, capsid "opening" was confirmed by plating experiments in which the number of remaining plaque-forming units (PFUs) after treatment with receptor was counted, as previously described (Parent *et al*., 2014). As evident from the figure, the extent of ejection decreases with increasing PEG concentration until it is completely suppressed at about 17 atm.
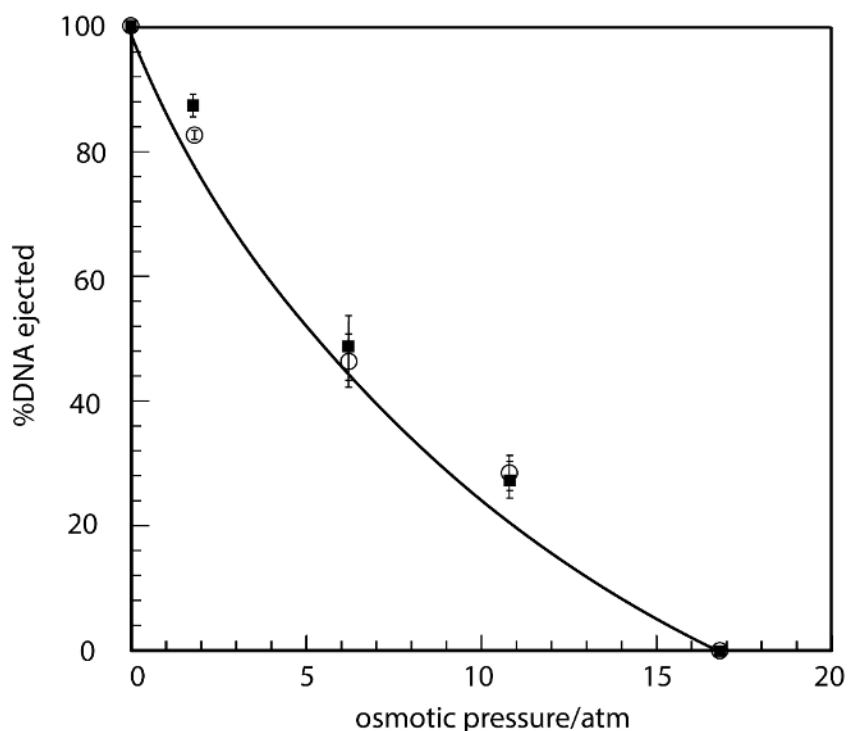
Figure 4-1. Measured DNA ejection percentage from P22 at various osmotic pressures under different receptor conditions: P22+LPS (○); P22+LPS+OmpA (■). Ejection is triggered by addition of receptor (LPS or LPS and OmpA) in the presence of DNase I, the sample is centrifuged to isolate the capsids, and the DNA remaining in the capsids is extracted, run in an agarose gel, and the length is calculated based on a DNA ladder run in the gel at the same time. % DNA ejection is relative to the full length DNA.

E protein ejection does not occur when P22 is treated with LPS alone

Using the equilibrium, partially ejected DNA experiments as described above and shown in Figure 4-1, we centrifuged ^{35}S-labeled samples (see Materials and Methods) to separate capsids from ejected protein and DNA. We then analyzed by SDS-PAGE, autoradiography and densitometry the protein content in the pellet, i.e., those proteins still associated with the capsid/receptor macromolecular complex. Figure 4-2 shows the

results for representative resuspended pellets, for different concentrations of PEG, in which LPS was used to trigger DNA ejection. In all cases, as in Figure 4-1, we assayed a portion of each reaction and counted PFUs before and after treatment with LPS to confirm that the majority of virions were triggered for ejection.
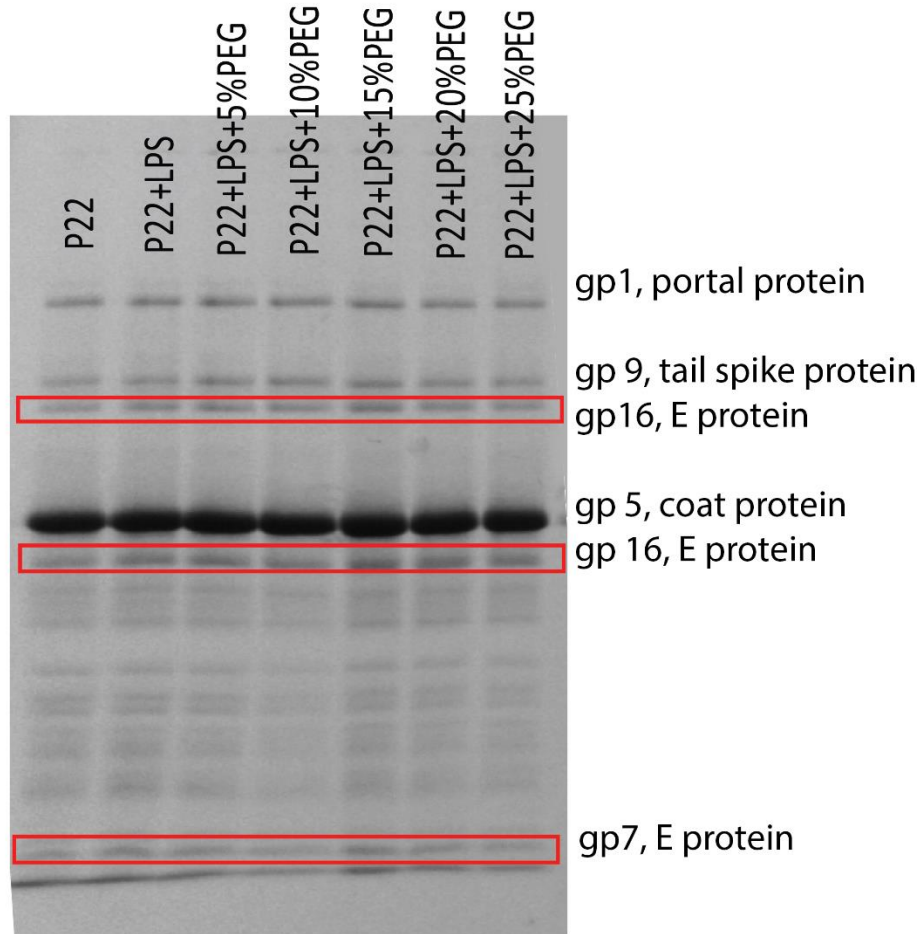

Figure 4-2. SDS-PAGE visualized by autoradiography showing the protein content in the pellet after triggered ejection by LPS under different PEG concentrations. The left-most lane is the control which does not contain any receptor. Red boxes highlight the three E proteins that we are interested in.

If any (or all) of the E proteins had been ejected, they would be present in the supernatant and consequently the bands representing those proteins would become less intense or even disappear, depending on the fraction ejected. As is evident, each lane has the same viral protein pattern with similar band intensities, independent of the presence of LPS and the amount of PEG. The same result was found in 10 trials. Similar results were also found when we performed a trypsin digestion in parallel samples, to digest any protein that was ejected and no longer protected by the capsid (data not shown). Therefore, we can exclude the possibility that protein was ejected but remained bound to the outside of the capsid. We can thus conclude that LPS alone did not trigger E protein ejection *in vitro*.

P22 ejection efficiency increases *in vitro* in the presence of outer membrane protein A (OmpA)

Studies by the Seckler group have shown that while LPS from *Salmonella* can trigger P22 ejection *in vitro*, the kinetics are very slow (~ 5 hours) compared to the phage life cycle (~1 hour) (Andres *et al.*, 2010; 2012). Because of the close relationship between Sf6 and P22 (Casjens *et al.*, 2011), and the fact that both the outer membrane protein OmpA and LPS are required for Sf6 ejection *in vitro* (Parent *et al.*, 2014), we examined if OmpA was able to enhance the rate of ejection in P22. As can be seen in Figure 4-3.A, as early as 10 min after incubation around 75% of the virus in the sample containing both OmpA and LPS has lost infectivity versus only 30% of the sample with LPS alone. Samples with OmpA but not LPS remained essentially unchanged after incubation, showing that OmpA alone does not trigger ejection.

We also performed the same experiment with two P22 hybrids (Leavitt *et al.*, 2013) where

the majority of the phage is comprised of P22 proteins, but the tail needle is from another

phage type. One such hybrid containing an Sf6 tail showed results similar to P22 (Figure

4-3.B). But for a mutant with a T4 "foldon" tail needle (Figure 4-3.C), neither LPS with

OmpA nor LPS alone was able to trigger loss of infectivity. It appears then that for LPS

and OmpA to function as receptors, the tail needle structure must be closely similar to

that of P22, consistent with the earlier suggestion that the P22 tail needle is part of the

trigger that determines DNA ejection (Leavitt, *et al.*, 2013).
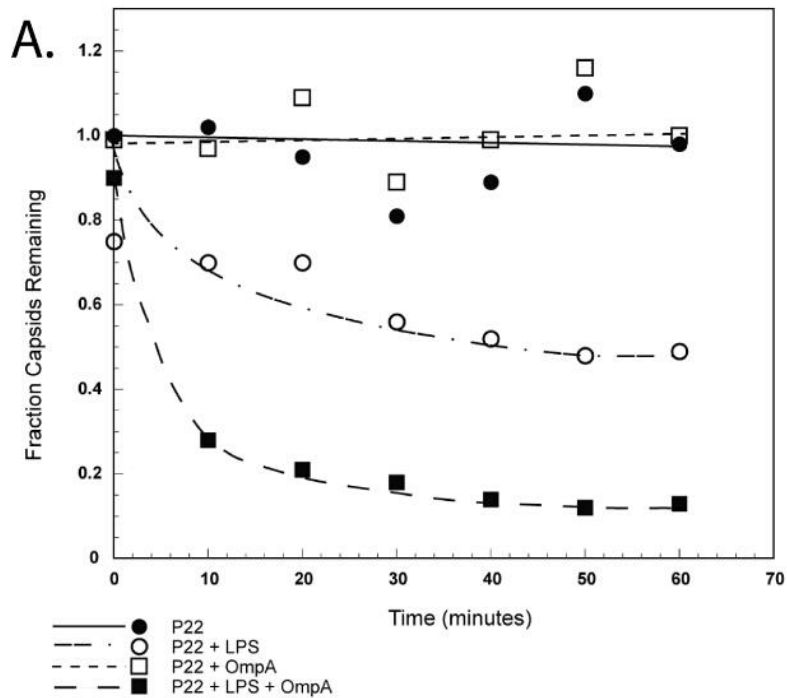
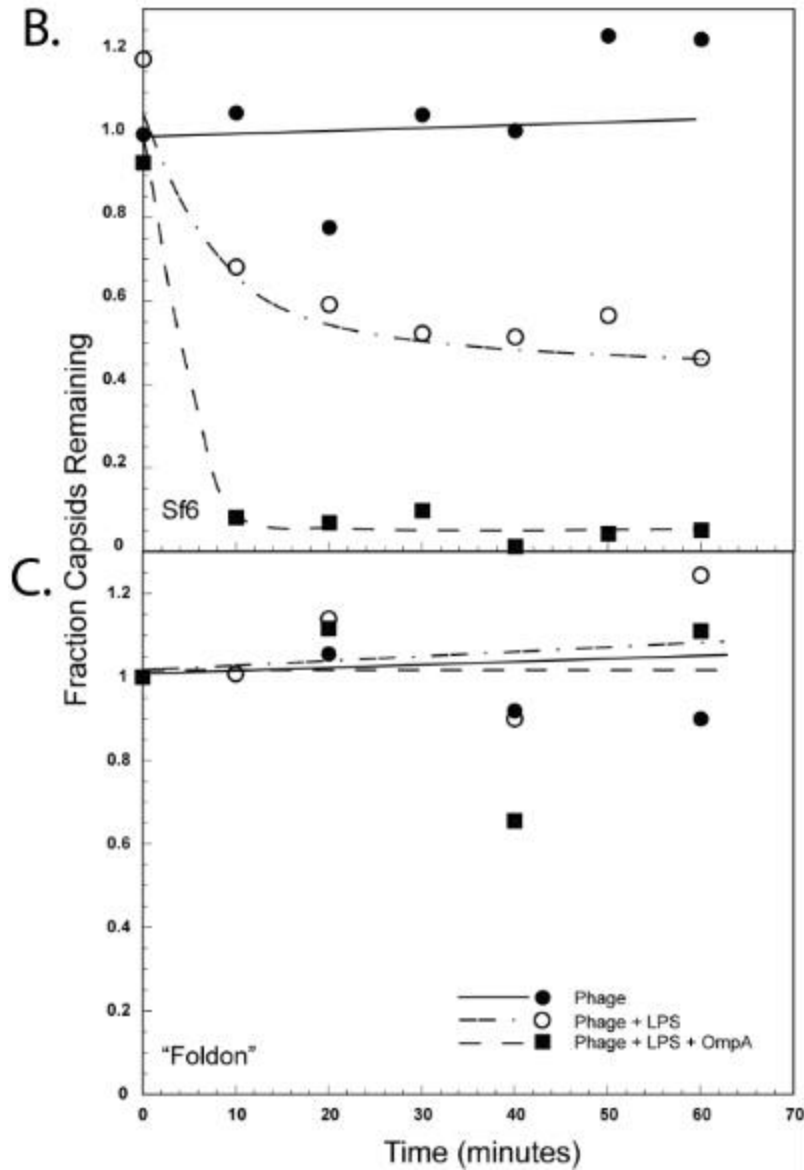Figure 4-3. *In vitro* genome ejection of (A) wildtype P22 and (B) two P22 hybrids. "Fraction capsids remaining" was calculated at each time point as the number of pfu remaining after incubation with LPS, OmpA or LPS and OmpA, divided by the number of pfu when incubated with buffer at t = 0.

P22 E proteins are ejected *in vitro* in the presence of both LPS and an

outer membrane protein

Since OmpA enhanced P22 ejection kinetics and overall efficiency (Figure 4-3.A), we asked if P22 capsids would release the E proteins when both receptors were present. The pressure driving DNA ejection was determined again, but this time in the presence of both LPS and OmpA (Figure 4-1—filled squares). Within the precision of the measurements the data obtained for fast ejection, i.e., with the two receptors, are identical to those for slow ejection, i.e., for LPS alone.

However, when both LPS and OmpA were present, the E protein behavior was different. Figure 4-4 is an autoradiograph showing samples with both receptors at PEG concentrations ranging from 0 to 15%. Comparing lanes with receptors (lane 2-5) and the lane without receptors (lane 1), we can clearly see that bands representing gp16 and gp20 disappear in the presence of the two receptors, independent of [PEG]. A test with trypsin digestion, but without separation by centrifugation, gave consistent results—that the E proteins were rapidly digested when the capsids were triggered "open" with both LPS and OmpA, indicating their release into the solution (gel not shown).

The percentage of ejection of all three E proteins in osmotic pressures ranging from 0 to 38 atm were calculated as follows. For each autoradiograph, there is one lane representing a sample of virus concentration identical to that of the other samples in the gel but lacking receptors. The relative intensity of each E-protein band in each sample is calculated based on the coat protein band in the same sample. The ejection percentage is calculated from the relative intensity (I) in samples with receptors after calibration with
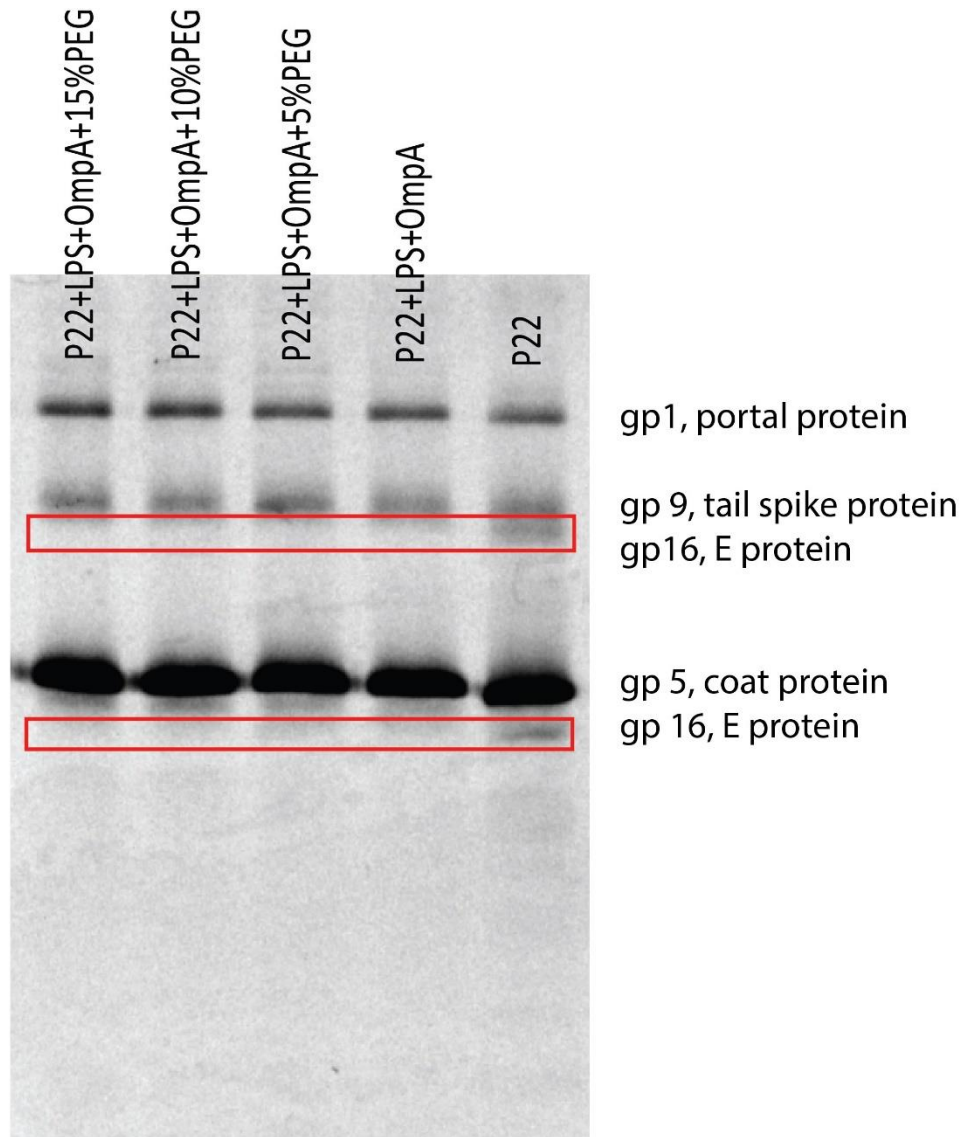
Figure 4-4. SDS-PAGE visualized by autoradiography showing the protein content in the pellet after triggered ejection by LPS and OmpA under different PEG concentrations. The right-most lane is the control which does not contain any receptor. Red boxes highlight the E proteins that we are interested in.

virus ejection efficiency (a; $0 \leqslant a \leqslant 1$) found from the plaque assay, divided by that of pure

virus sample (I_0):

$$f = \frac{I_0 - I}{aI_0} \times 100\%$$

These results are presented in Figure 4-5: 60 – 70% of all three E proteins are ejected from the capsid for osmotic pressures ranging from 0 to 16.8 atm. No E protein ejection can be detected at higher pressures.
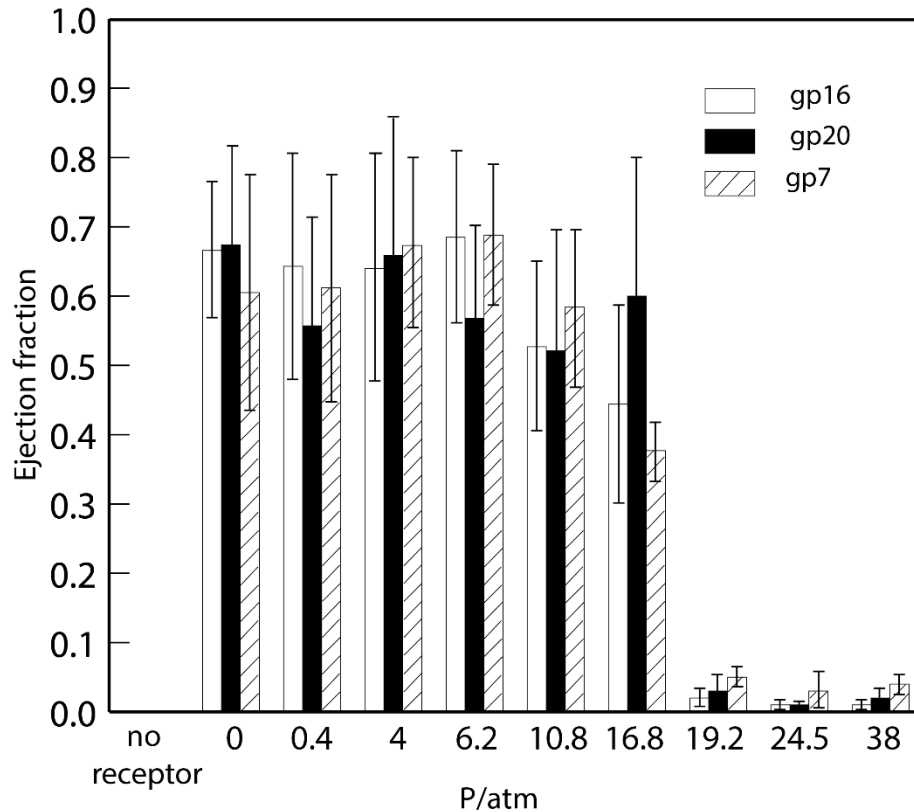
Figure 4-5. *In vitro* E protein ejection fractions under different osmotic pressures. It was calculated by using the relative intensity of E protein band in each sample divided by the relative intensity of that band in an intact capsid.

In contrast to the incomplete ejection of E proteins that we have observed, studies of P22 E protein ejection performed in the early 1970s showed that more than 90% of gp16 and gp20, as well as 70-90% of gp7 were ejected from the capsid during infection. In our system every component such as LPS and OmpA is present as a free molecule in solution while in the cell the LPS and OmpA are confined to the surface of the bacterium and their

relative positions are stationary. It is possible that not all the viruses can interact with an active OmpA protein at the moment they are in contact with LPS. This would cause some viruses to eject only in the presence of LPS and thus leave E proteins inside the capsid. An alternative explanation is that perhaps OmpA is not as functional when purified in detergent micelles as it is *in vivo*. However, our data show very clearly that virtually no E protein release occurs with LPS alone, and we observe near-physiological amounts of protein ejection when OmpA is also present.

If we overlap the DNA partial ejection data (Figure 4-1) with the E protein ejection data (Figure 4-5), we see that at 17 atm DNA ejection is completely suppressed while most of the E proteins are still ejected (Figure 4-6). From this we conclude that in the presence of LPS and OmpA all three E-proteins are ejected prior to DNA ejection. It is only at an osmotic pressure 2 atm higher that the ejection of the E proteins is inhibited. However, we were unable to determine the *order* of E protein ejection. It is possible that all the proteins bind together and come out together, or that they come out in a random sequence, which is not detectable in our bulk population measurements. It is also likely – even if an order exists – that, because all the proteins are ejected before the highly-stressed DNA it is not possible to resolve their order using the osmotic suppression method because the contribution of each protein to the overall pressure is too small.

The force driving ejection of the protein from P22 arises from the confined DNA; it would

be expected to be similar to that associated with the ejection of DNA from λ phage, which

like P22 has a T = 7 capsid and is of closely similar dimensions. Grayson *et al.* (2006)
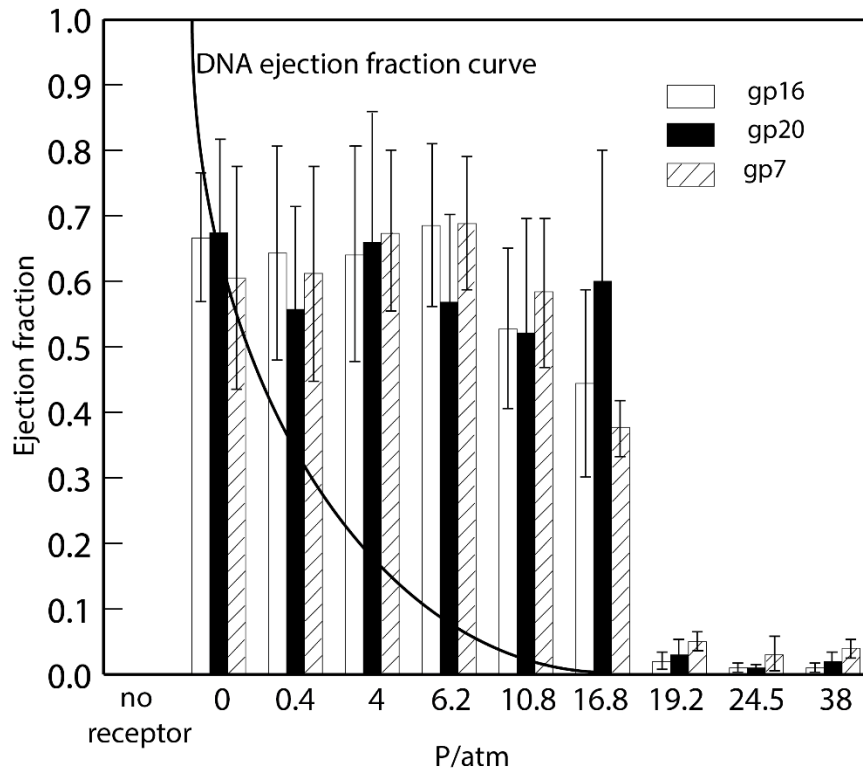
Figure 4-6. Superposition of *in vitro* DNA ejection fraction curve and E protein ejection fraction graph in different osmotic pressures.

carried out osmotic suppression measurements on λ for the 48.5 kb wild-type genome

and a 37.7 kb mutant and found that the ejection was completely inhibited at pressures

of 20-25 and 10-15 atm, respectively. The pressure required to inhibit the E proteins that

we have observed for P22 originates from a 43.5 kb DNA (larger than the 41.7 kb genome

because of head-full packaging) confined to a volume slightly smaller than the capsid

interior because of the presence of the proteins. The fraction of the internal volume

occupied by the E proteins can be estimated from their number and molecular masses. One can estimate their volume by assuming they have a typical protein density of $1.4g/cm^3$. Taking the internal volume to be $1.1 \times 10^5 \; nm^3$ (Patterson *et al.*, 2012) one estimates that the E proteins occupy about 2% of the volume. This decrease in volume available to the DNA is quite small. Nevertheless because of the exponential dependence of osmotic pressure of highly condensed DNA (Rau *et al.*, 1984) it would be expected to increase the ejection pressure by about 10%, comparable to the difference we observe between the pressures that inhibits the ejection of the proteins and the DNA. Finally we note that the turgor pressure in both the periplasm and cytoplasm of *Salmonella* is 3.5 atm (Stock *et al.*, 1977), so that pressure-driven ejection from P22 can be expected to transfer only 60% of the DNA from the phage to the host (see Figure 4-1); another mechanism is necessary to account for the remaining 40%.

4.3 Conclusions

Productive virus infection requires accurate recognition of the host cell to avoid unnecessary genome release. This process is largely controlled by specific interactions between the virus and the receptors from the host. It has been shown that LPS from the host *Salmonella* can trigger slow DNA ejection of P22 *in vitro*. In this work, we have demonstrated for the first time that together with LPS, purified OmpA can dramatically increase the rate and the efficiency of DNA ejection of P22 *in vitro*. OmpA is one of the major outer membrane proteins in bacteria with ~100,000 copies in each cell. Our results suggest that it can serve as a potential secondary receptor during P22 infection. This

finding is similar to the case of Sf6 (another *Podovirus* closely related to P22) for which LPS alone from the host *Shigella* as a primary receptor cannot trigger DNA release by itself; only when a secondary receptor like OmpA or OmpC is present as well can the virus trigger DNA release (Parent *et al*., 2014).

How exactly both receptors interact with P22, induce a possibly different conformational change of the portal complex compared to LPS alone, and promote a more effective DNA ejection and E protein ejection, remains to be determined. After specific binding to the host outer membrane, how viral DNA translocates across both membranes of the host to initiate the infection remains elusive, especially for short-tailed *Podoviridae*.

Compared to *Siphoviridae* and *Myoviridae* which in principle can span the space between the host surface and the cytoplasm with their long tail, *Podoviridae* may require a more complex mechanism to deliver their genome into the cytoplasm. One of the best studied *Podoviridae,* T7, utilizes its internal proteins to further extend its tail across the membranes to facilitate genome delivery (Molineux *et al*., 2013). So far there is no such evidence for P22, and from the EM studies the tail appears unchanged after infection, implying that protein ejection does not extend its tail. Our finding that all E proteins are ejected *before* the DNA, and that all of them need to be present for infectivity, suggests that they play a key role in transferring P22 DNA through the periplasm and inner membrane by either protecting the DNA from periplasmic nuclease digestion or providing an enzymatic activity that can digest the peptidoglycan layer and inner membrane.

Studies of the internal proteins of other viruses have suggested other possible functions for them during infection. For example, studies on N4, another of the *Podoviridae*, showed that the pre-packaged viral RNA polymerase (RNAP) is released before the DNA and that the transcription of the DNA by this RNA polymerase helps pull the rest of the DNA out of the capsid (Casjens, *et al.*, 2012). This function seems unlikely for E proteins in P22, because its genome undergoes headful packaging, which makes the end of the genome in each capsid not necessarily the same (Lander *et al*., 2006).

In this study, no order of E protein ejection has been found. We know that each E protein has 10-20 copies in each capsid, so that in total there are 30-60 copies of proteins ejected before the DNA can exit. The energy for protein ejection must come from the energy that has been stored in DNA during packaging. And, because of the lack of information about the location of the proteins within the capsid their conformations and their specific functions after ejection, it is unclear if they coordinate with each other to come out in a specific order. One possibility for understanding more about these mechanisms is to determine a high-resolution structure during ejection by cryo-EM.

4.4 Materials and Methods

Plaque assay for determining the virus ejection rate

For all experiments, the P22 strain used was a clear plaque mutant that is an obligate lytic strain (Casjens *et al*., 1987). Hybrid P22 with either a Sf6 tail needle knob or a T4 foldon tail were purified from *Salmonella enterica* serovar Typhimurium DB7136 (leuA414am, hisC525am, su^0) LT2 cell lysates. LPS was purified using a kit (BULLDOGBIO), and OmpA was purified as described (Porcek and Parent, 2015). Each type of virus with final concentration of 10^9 pfu/ml was incubated with either LPS (2.5 mg/ml), OmpA (0.2 mg/ml) or both LPS and OmpA, all in the presence of detergent Triton x-100 [1.06% (w/v)]. At each time point (0, 10, 20, 30, 60, 90 min), an aliquot was removed for titering. The percentage of infectious virus remaining was determined using the pfu from each sample divided by that for a sample containing no receptor.

Genome ejection assay

P22 labeled with radioactive ^{35}S was purified as described in (Parent *et al*., 2004). The virus sample (10^{10} pfu/ml) was first treated with DNase I to remove any free DNA. Samples of P22 and the LPS receptors were mixed with PEG 8000 (at concentrations of PEG corresponding to osmotic pressures of either 0, 1.78, 6.2, 10.8 or 16.8 atm) and were incubated overnight at 37°C to ensure complete genome ejection. The ratios between virus and the receptors were the same as those employed in the plaque assays. For samples containing both LPS and OmpA as receptors, the same amounts of P22 and LPS, OmpA, Triton X-100 and the same PEG concentration used in the plaque assays were mixed and incubated at 37°C for 1 hr. In both cases, DNase I (5 units, New England Biolabs) was added to the sample after ejection was triggered, and the mixture was kept at 37°C for 4 hr to digest the DNA ejected from the capsid. Before recovering the

unejected DNA from the capsid, 1mM EDTA was added and the sample heated for 10 min at 75°C to denature the DNase I. After addition of an equal volume of lysis buffer (25mM EDTA, 200mM Tris, 250mM NaCl, 1%SDS, pH 7.5) and 1 unit of Protease K, the samples were incubated at 65°C for 1 hr to disrupt the capsids, allowing the DNA inside to be released. Phenol / chloroform extraction was carried out twice to separate protein from DNA, followed by ethanol precipitation to concentrate the DNA. The remaining DNA, now in the pellet, was redissolved in TE buffer and analyzed by gel electrophoresis. In order to resolve the relatively long DNA, a low-percentage agarose gel (0.3%) was used and the running condition was 6 hr at 3V/cm; SYBR gold was used to stain the gel in the last step.

SDS-PAGE for determining E protein ejection

We used samples prepared as described above for the plaque assays. DNase I was then added followed by another 2hr incubation at 37ºC to completely digest the ejected DNA. An aliquot of each sample was used to determine the efficiency of genome ejection. The remainder of the sample was centrifuged at 27,000g for 1.5 hr to separate the virus capsids from possible ejected proteins and free nucleotides. Resuspended pellets containing capsids were TCA-precipitated and loaded into a 10% SDS-PAGE gel.

Bibliography

Ali, I., Marenduzzo, D., Yeomans, J. (2006). Polymer Packaging and Ejection in Viral Capsids: Shape Matters. *Phys Rev Lett,* 96, 208102.

Andres, D., Hanke, C., Baxa, U., Seul, A., Barbirz, S., Seckler, R. (2010). Tailspike interactions with lipopolysaccharide effect DNA ejection from phage P22 particles in vitro. *J Biol Chem,* 285, 36768-36775.

Andres, D., Roske, Y., Doering, C., Heinemann, U., Seckler, R., Barbirz, S. (2012). Tail morphology controls DNA release in two Salmonella phages with one lipopolysaccharide receptor recognition system. *Mol Microbiol,* 83, 1244-1253.

Arscott, P. G., Li, A.-Z., Bloomfield, V. A. (1990). Condensation of DNA by trivalent cations. 1. Effects of DNA length and topology on the size and shape of condensed particles. *Biopolymers,* 30, 619-630.

Bauer, D. W., Huffman, J. B., Homa, F. L., Evilevitch, A. (2013). Herpes virus genome, the pressure is on. *J Am Chem Soc,* 135, 11216-11221.

Berndsen, Z. T., Keller, N., Grimes, S., Jardine, P. J., Smith, D. E. (2014). Nonequilibrium dynamics and ultraslow relaxation of confined DNA during viral packaging. *Proc Natl Acad Sci U S A,* 111, 8345-8350.

Bloustine, J., Virmani, T., Thurston, G., Fraden, S. (2006). Light Scattering and Phase Behavior of Lysozyme-Poly(Ethylene Glycol) Mixtures. *Phys Rev Lett,* 96, 087803.

Casjens, S., Huang, W.M., Hayden, M., Parr, R. (1987). Initiation of bacteriophage P22 DNA packaging series Analysis of a mutant that alters the DNA target specificity of the packaging apparatus. *J Mol Biol,* 194, 411-422.

Casjens, S., Hayden, M. (1988). Analysis in vivo of the bacteriophage P22 headful nuclease. *J Mol Biol,* 199, 467-474.

Casjens, S. R., Thuman-Commike, P. A. (2011). Evolution of mosaically related tailed bacteriophage genomes seen through the lens of phage P22 virion assembly. *Virol,* 411, 393-415.

Castelnovo, M., Evilevitch, A. (2007). DNA ejection from bacteriophage: towards a general behavior for osmotic-suppression experiments. *Eur Phys J E Soft Matter,* 24, 9-18.

Cayley, S., Lewis, B. A., Guttman, H. J., Record, M. T. Jr. (1991). Characterization of the cytoplasm of Escherichia coli K-12 as a function of external osmolarity Implications for protein-DNA interactions in vivo. *J Mol Biol,* 222, 281-300.

Das, A. (1992). How the phage lambda N gene product suppresses transcription termination communication of RNA polymerase with regulatory proteins mediated by signals in nascent RNA. *J Bacteriol,* 174, 6711-6716.

Davidson, A. R., Cardarelli, L., Pell, L. G., Radford, D. R., Maxwell, K. L. (2012). Long noncontractile tail machines of bacteriophages. *Adv Exp Med Biol, 726*, 115-142.

de Frutos, M., Brasiles, S., Tavares, P., Raspaud, E. (2005). Effect of spermine and DNase on DNA release from bacteriophage T5. *Eur Phys J E Soft Matter,* 17, 429-434.

de Frutos, M., Letellier, L., Raspaud, E. (2005). DNA ejection from bacteriophage T5: analysis of the kinetics and energetics. *Biophys J,* 88, 1364-1370.

Deighan, P., Hochschild, A. (2007). The bacteriophage lambdaQ anti-terminator protein regulates late gene expression as a stable component of the transcription elongation complex. *Mol Microbiol,* 63, 911-920.

Demidenko, A. A., Lee, J., Powers, T. R., Nibert, M. L. (2011). Effects of viscogens on RNA transcription inside reovirus particles. *J Biol Chem,* 286, 29521-29530.

Downing, W. L., Sullivan, S.L., Gottesman, M. E., Dennis, P.P. (1990). Sequence and transcriptional pattern of the essential Escherichia coli secE-nusG operon. *J Bacteriol,* 172, 1621-1627.

Elowitz, M. B., Surette, M. G., Wolf, P., Stock, J. B., Leibler, S. (1999). Protein Mobility in the Cytoplasm of Escherichia coli. *J Bacteriol,* 181, 197-203.

Evilevitch, A. (2006). Effects of Condensing Agent and Nuclease on the Extent of Ejection from Phage λ. *J Phys Chem B,* 110, 22261-22265.

Evilevitch, A., Lavelle, L., Knobler, C. M., Raspaud, E., Gelbart, W. M. (2003). Osmotic pressure inhibition of DNA ejection from phage. *Proc Natl Acad Sci U S A,* 100, 9292-9295.

Evilevitch, A., Fang, L. T., Yoffe, A. M., Castelnovo, M., Rau, D. C., Parsegian, V. A., Gelbart, W. M., Knobler, C. M. (2008). Effects of Salt Concentrations and Bending Energy on the Extent of Ejection of Phage Genomes. *Biophys J,* 94, 1110-1120.

Evilevitch, A., Castelnovo, M., Knobler, C. M., Gelbart, W. M. (2004). Measuring the force ejecting DNA from phage. *J Phys Chem B,* 108, 6838-6843.

Forrey, C., Muthukumar, M. (2006). Langevin dynamics simulations of genome packing in bacteriophage. *Biophys J, 9*1, 25-41.

Friedman, D.I., Baron, L. S. (1974). Genetic characterization of a bacterial locus involved in the activity of the N function of phage lambda. *Virol*, 58, 141-148.

Friedman, D. I., Baumann, M., Baron, L. S. (1976). Cooperative effects of bacterial mutations affecting lambda N gene expression. I. Isolation and characterization of a nusB mutant. *Virol,* 73, 119-127.

Goldman, S. R., Ebright, R. H., Nickels, B. E. (2009). Direct detection of abortive RNA transcripts in vivo. *Science,* 324, 927-928.

Gonzalez-Tello, P., Camacho, F., Blazquez, G. (1994). Density and Viscosity of Concentrated Aqueous Solutions of Polyethylene Glycol. *J Chem Eng Data,* 39, 611-614.

Grayson, P., Evilevitch, A., Inamdar, M. M., Purohit, P. K., Gelbart, W. M., Knobler, C. M., & Phillips, R. (2006). The effect of genome length on ejection forces in bacteriophage lambda. *Virology,* 348, 430-436.

Grayson, P., Molineux, I. J. (2007). Is phage DNA 'injected' into cells--biologists and physicists can agree. *Curr Opin Microbiol,* 10, 401-409.

Greenblatt J., Malnoe P., Li, J. (1980). Purification of the gene N transcription anti-termination protein of bacteriophage lambda. *J Biol Chem*, 255, 1465-1470.

Hayden, S. C. M. (1988). Analysis in vivo of the bacteriophage P22 headful nuclease. *J Mol Biol,* 199, 467-474.

Herskowitz, I. (1973). control of gene expression in bacteriophage lambda. *Annu Rev Genet,* 7, 289-324.

Hoffman, B., Levine, M. (1975a). Bacteriophage P22 virion protein which performs an essential early function I. analysis of 16-ts mutants. *J Virol,* 16, 1536-1546.

Hoffman, B., Levine, M. (1975b). Bacteriophage P22 virion protein which performs an essential early function II. characterization of the gene 16 function. *J Virol,* 16, 1547-1559.

Howard, B. H., de Crombrugghe, B., Rosenberg, M. (1977). Transcription in vitro of bacteriophage lambda 4S RNA studies on termination and rho protein. *Nucleic Acids Res,* 4, 828-842.

Israel, V., Woodworth-Gutai, M., Levine, M. (1972). Inhibitory effect of bacteriophage P22 infection on host cell deoxyribonuclease activity. *J Virol,* 9, 752-757.

Israel, V. (1977). E proteins of bacetriophage P22 I. Identification and Ejection from Wild-Type and Defective Particles. *J Virol,* 23, 91-97.

Iwashita, S., Kanegasaki, S. (1973). Smooth specific phage adsorption endorhamnosidase activity of tail parts of P22. *Biochem Biophys Res Commun,* 55, 403-409.

Keller, N., delToro, D., Grimes, S., Jardine, P. J., Smith, D. E. (2014). Repulsive DNA-DNA Interactions Accelerate Viral DNA Packaging in Phage Phi29. *Phys Rev Lett,* 112, 248101.

Kemp, P., Gupta, M., Molineux, I. J. (2004). Bacteriophage T7 DNA ejection into cells is initiated by an enzyme-like mechanism. *Mol Microbiol,* 53, 1251-1265.

Keppel, F., Georgopoulos, C. P.,Eisen, H. (1974). Host interference with expression of the lambda N gene product. *Biochimie*, 56, 1505–1509.

Kindt, J., Tzlil, S., Ben-Shaul, A., Gelbart, W. M. (2001). DNA packaging and ejection forces in bacteriophage. *Proc Natl Acad Sci U S A,* 98, 13671-13674.

King, J., Botstein, D., Casjens, S., Earnshaw, W., Harrison, S., Lenk, E. (1976). Structure and assembly of the capsid of bacteriophage P22. *Phil Trans R Soc Lond B,* 276, 37-49.

Lander, G. C., Khayat, R., Li, R., Prevelige, P. E., Potter, C. S., Carragher, B., Johnson, J. E. (2009). The P22 tail machine at subnanometer resolution reveals the architecture of an infection conduit. *Structure,* 17, 789-799.

Leavitt, J. C., Gogokhia, L., Gilcrease, E. B., Bhardwaj, A., Cingolani, G., Casjens, S. R. (2013). The tip of the tail needle affects the rate of DNA delivery by bacteriophage P22. *PLoS One,* 8, e70936.

Leforestier, A., Brasiles, S., de Frutos, M., Raspaud, E., Letellier, L., Tavares, P., Livolant, F. (2008). Bacteriophage T5 DNA ejection under pressure. *J Mol Biol,* 384, 730-739.

Leforestier, A., Livolant, F. (2009). Structure of toroidal DNA collapsed inside the phage capsid. *Proc Natl Acad Sci U S A,* 106, 9157-9162.

Leiman, P. G., Shneider, M. M. (2012). Contractile tail machines of bacteriophages. *Adv Exp Med Biol,* 726, 93-114.

Li, J., Horwitz, R., McCracken, S., Greenblatt, J. (1992). NusG, a new Escherichia coli elongation factor involved in transcriptional antitermination by the N protein of phage lambda. *J Biol Chem*, 267, 6012-6019.

Mah, T. F., Li, J., Davidson, A. R., Greenblatt, J. (1999). Functional importance of regions in Escherichia coli elongation factor NusA that interact with RNA polymerase, the bacteriophage lambda N protein and RNA *Mol Microbiol,* 34, 523-537.

Mangenot, S., Hochrein, M., Radler, J., Letellier, L. (2005). Real-time imaging of DNA ejection from single phage particles. *Curr Biol,* 15, 430-435.

Marenduzzo, D., Micheletti, C., Orlandini, E., Sumners, D. W. (2013). Topological friction strongly affects viral DNA ejection. *Proc Natl Acad Sci* U S A, 110, 20081-20086.

Mason, S. W., Joyce Li, J., Greenblatt, J. (1992). Host factor requirements for processive antitermination of transcription and suppression of pausing by N protein of lambda. *J Biol Chem,* 267, 19418-19426.

Minton, A. P.(1989). Holobiochemistry: an integrated approach to the understanding of biochemical mechanism that emerges from the study of proteins and protein associations in volume-occupied solutions. In: Srere, P., Jones, M. E., Mathews, C., eds. Structural and Organizational Aspects of Metabolic Regulation. *Alan R. Liss, Inc., New York*, 291-306.

Molineux, I. J. (2006). Fifty-three years since Hershey and Chase; much ado about pressure but which pressure is it? *Virology,* 344, 221-229.

Molineux, I. J., Panja, D. (2013). Popping the cork: mechanisms of phage genome ejection. *Nat Rev Microbiol,* 11, 194-204.

Money, N. P. (1989). Osmotic pressure of aqueous Polyethylene Glycols. *Plant Physiol,* 91, 766-769.

Morelli, M. J., Allen, R. J., Wolde, P. R. (2011). Effects of macromolecular crowding on genetic networks. *Biophys J,* 101, 2882-2891.

Ninni, L., Burd, B., Fung, W. H., Meirelles, A. J. A. (2003). Kinematic Viscosities of Poly(ethylene glycol) Aqueous Solutions. *J Chem Eng Data,* 48, 324-329.

Odijk, T. (1998). Hexagonally packed DNA within bacteriophage T7 stabilized by curvature stress. *Biophys J,* 75, 1223-1227.

Parent, K. N., Ranaghan, M. J., Teschke, C. M. (2004). A second-site suppressor of a folding defect functions via interactions with a chaperone network to improve folding and assembly in vivo. *Mol Microbiol,* 54, 1036-1050.

Parent, K. N., Erb, M. L., Cardone, G., Nguyen, K., Gilcrease, E. B., Porcek, N. B., Casjens, S. R. (2014). OmpA and OmpC are critical host factors for bacteriophage Sf6 entry in Shigella. *Mol Microbiol,* 92, 47-60.

Patterson, D. P., Prevelige, P. E., Douglas,T. (2012). Nanoreators by Programmed Enzyme Encapsulation Inside the Capsid of the Bacteriophage P22. ACS Nano, 6, 5000-5009.

Perez, G. L., Huynh, B., Slater, M., Maloy, S. (2009). Transport of phage P22 DNA across the cytoplasmic membrane. *J Bacteriol,* 191, 135-140.

Petrov, A. S., Harvey, S. C. (2011). Role of DNA-DNA interactions on the structure and thermodynamics of bacteriophages Lambda and P4. *J Struct Biol,* 174, 137-146.

Plum, G. E., Arscott, P. G., Bloomfield, V. A. (1990). Condensation of DNA by trivalent cations. 2. Effects of cation structure. *Biopolymers,* 30, 631-643.

Poranen, M. M., Daugelavicius, R., & Bamford, D. H. (2002). Common principles in viral entry. *Annu Rev Microbiol,* 56, 521-538.

Purohit, P. K., Inamdar, M. M., Grayson, P. D., Squires, T. M., Kondev, J., Phillips, R. (2005). Forces during bacteriophage DNA packaging and ejection. *Biophys J,* 88, 851-866.

Raspaud, E., Durand, D., Livolant, F. (2005). Interhelical spacing in liquid crystalline spermine and spermidine-DNA precipitates. *Biophys J,* 88, 392-403.

Rau, D. C., Lee, B., Parsegian, V. A. (1984). Measurement of the repulsive force between polyelectrolyte molecules in ionic solution Hydration forces between parallel DNA doubles helices. *Proc Natl Acad Sci U S A,* 81, 2621-2625.

Riemer, S. C., Bloomfield, V.A. (1978). Packaging of DNA in bacteriophage heads some considerations on energetics. *Biopolymers,* 17, 785-794.

Sanger, F., Coulson, A.R., Hong, G.F., Hill, D.F, Petersen, G.B., (1982). Nucleotide sequence of bacteriophage λ DNA*. J Mol, Biol,* 162, 729-773.

Segur, J. B., Oberstar, H. E. (1951). Viscosity of Glycerol and Its Aqueous Solutions. *Ind Eng Chem,* 43, 2117-2120.

Smith, D. E., Tans, S. J., Smith, S. B., Grimes, S., Anderson, D. L., Bustamante, C. (2001). The bacteriophage phi29 portal motor can package DNA against a large internal force. *Nature,* 413, 748-752.

Smith, D. E. (2011). Single-molecule studies of viral DNA packaging. *Curr Opin Virol,* 1, 134-141.

Spakowitz, A. J., Wang, Z. G. (2005). DNA packaging in bacteriophage: is twist important? *Biophys J*, 88, 3912-3923.

Stephen, M. (1992). Host Factor Requirements for Processive Antitermination of Transcription and Suppression of pausing by the N Protein of Bacteriophage lambda*. J Biol Chem*, 267, 19418-19426

Stock, J. B., Rauch, B., Roseman, S. (1977). Periplasmic space in Salmonella typhimurium and Escherichia coli. *J Biol Chem,* 252, 7850-7861.

Sullivan, S. L., Ward, D. F., Gottesman, M. E. (1992). Effect of Escherichia coli nusG function on lambda N-mediated transcription antitermination. *J Bacteriol,* 174, 1339-1344.

Tang, J., Lander, G. C., Olia, A. S., Li, R., Casjens, S., Prevelige, P., Jr., Johnson, J. E. (2011). Peering down the barrel of a bacteriophage portal: the genome packaging and release valve in p22. *Structure,* 19, 496-502.

Tubio, G., Nerli, B., Pico, G. (2004). Relationship between the protein surface hydrophobicity and its partitioning behaviour in aqueous two-phase systems of polyethyleneglycol–dextran. *J Chromatogr B,* 799, 293-301.

Tzlil, S., Kindt, J.T., Gelbart, W.M., Ben-Shaul, A. (2003). Forces and pressures in DNA packaging and release from viral capsids. *Biophys J,* 84, 1616-1627.

Van Gilst, M.R., Rees, W. A., Das, A. and Von Hippel, P. H. (1997). Complexes of N Antitermination Protein of Phage λ with Specific and Nonspecific RNA Target Sites on the Nascent Transcript. *Biochemistry,* 36, 1514-1524.

Vieu, E., Rahmouni, A. R. (2004). Dual role of boxB RNA motif in the mechanisms of termination/antitermination at the lambda tR1 terminator revealed in vivo. *J Mol Biol,* 339, 1077-1087.

Winzor, D. J., Wills, P. R. (2006). Molecular crowding effects of linear polymers in protein solutions. *Biophys Chem,* 119, 186-195.

Zhou, Y., Filter, J. J., Court, D. L., Gottesman, M. E., Friedman, D. I. (2002). Requirement for NusG for Transcription Antitermination In Vivo by the N Protein. J Bacteriol, 184, 3416-3418.