EXPRESSION, PURIFICATION AND CHARACTERIZATION OF BACTERIOPHAGE LAMBDA TAIL TIP PROTEINS

by

Xiaoxian Dai

B.S., Zhejiang University, 2003

Submitted to the Graduate Faculty of

Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2009

UNIVERSITY OF PITTSBURGH

FACULTY OF ARTS AND SCIENCES

This dissertation was presented

by

Xiaoxian Dai

It was defended on

June 17, 2009

and approved by

Peter B. Berget, PhD, Associate Professor

Jeffrey L. Brodsky, PhD, Professor

Graham F. Hatfull, PhD, Eberly Family Professor

James M. Pipas, PhD, Professor

Dissertation Director: Roger W. Hendrix, PhD, Professor

EXPRESSION, PURIFICATION AND CHARACTERIZATION OF BACTERIOPHAGE LAMBDA TAIL TIP PROTEINS

Xiaoxian Dai, PhD

University of Pittsburgh, 2009

Bacteriophage tails, despite differences in their morphology, all play a key role in host recognition and DNA injection. It is widely believed that the tail, especially the baseplate / tail tip, has to undergo conformational change and protein rearrangement during infection. This change has been observed in both the long, contractile tail of T4 and the short tail of T7. In contrast, little is known about this aspect of the long, non-contractile tail of bacteriophage λ. Four proteins are involved in the λ tail tip assembly. I present evidence that gpJ, gpI, gpL and gpK are components of the tail tip complex. My results also suggest that there may be about three copies of gpI and three copies of gpL involved in the λ tail assembly. In addition, I have successfully purified gpL, which contains eight cysteines. My results show that when the conserved cysteine at position 173, 182 or 205 is mutated to serine, the mutant protein is defective in tail assembly. However, the C212S mutant accumulates a small amount of tail. Further analysis of this mutant indicates that C212 may have roles in both tail assembly and DNA injection.

gpK is required for λ tail assembly, but is not detected in the mature virion. Two different amber mutations were introduced into gene *K*. Neither of these mutants is able to complement *in vivo*.

However, the short amber fragment is unable to assemble λ tail whereas phage-like particles with little infectivity accumulate in the long amber fragment lysate. The results indicate that the function of gene *K* can be bypassed to some extent in the Kam768 mutant, but not in the Kam6 mutant.

TABLE OF CONTENTS

PREFACE……...xiii

1.0 INTRODUCTION……1

1.1 THE TAIL OF BACTERIOPHAGES……...……..……2

1.1.1 Tail structure……..…….2

1.1.2 Structural movements and DNA injection……..……8

1.1.3 Cysteine residues in the bacteriophage tail proteins……..……...12

1.2 THE BACTERIOPHAGE LAMBDA……...……13

1.2.1 Life cycle……..…….……...13

1.2.2 Lambda tail assembly pathway……..……..…….17

1.2.3 Function of the individual tail proteins…….20

1.2.4 Morphogenetic genes and λ tail assembly……..……..……23

1.2.5 The tail tip of phage λ……..……...……...26

1.3 THESIS PLAN……..…….……26

2.0 MATERIALS AND METHODS……..……...…….29

2.1 BACTERIAL STRAINS……..…….……29

2.2 MEDIA AND BUFFERS…….……32

2.2.1 Media…….……...32

2.2.2 Buffers……..…….33

2.3 METHODS……..…...……37

2.3.1 DNA manipulation……..……..37

2.3.2 Polymerase chain reaction……..……..45

2.3.3 Phage preparations……..……..48

2.3.4 Protein manipulation……..……...56

2.3.5 Complementation……..……...….61

2.3.6 Separation of λ tail / tail tip assembly intermediates……..……..62

2.4 SEPARATION OF λ PHAGE BY CESIUM CHLORIDE STEP GRADIENT……...63

3.0 EXPRESSION AND PURIFICATION OF λ TAIL TIP PROTEINS……..……64

3.1 INTRODUCTION…….……....64

3.2 ESTABLISHMENT OF THE IDENTITIES OF gpI, gpL AND gpK……...…65

3.2.1 Construction of amber mutation / deletion in the gene of interest……..…..65

3.2.2 The amber / deletion mutant has little biological activity *in vivo* ……..…..65

3.2.3 Identification of gpI, gpL and gpK …….……67

3.3 EXPRESSION OF THE λ TAIL TIP COMPLEX…….……..…….71

3.3.1 Addition of 6XHis tag to gpJ……..……..71

3.3.2 Effects of the polyhistidine tag on the biological activity of λ gpJ……..76

3.3.3 Expression of the tail tip complex…….…...76

3.4 PURIFICATION OF λ TAIL TIP PROTEINS…….……....82

3.4.1 Construction of the recombinant proteins……..……...82

3.4.2 Determination of the biological activities of the recombinant proteins……....84

3.4.3 Expression of λ gpI……..….84

3.4.4 Expression and purification of λ gpL…….…..95

3.4.5 Expression and purification of λ gpK……..……...105

3.4.6 Co-expression of the λ tail tip proteins……...110

3.5 DISCUSSION…….……...……..111

3.5.1 Comparisons of the different purification systems……..……...111

3.5.2 The solubility of the λ tail tip complex……..…….111

3.5.3 Expression and purification of λ tail tip proteins……..……..112

4.0 DETERMINATION OF THE COMPOSITION AND STOICHIOMETRY OF THE LAMBDA TAIL TIP COMPLEX……..……113

4.1 INTRODUCTION……..…….……113

4.2 PRECIPITATION OF λ TAIL TIP ASSEMBLY INTERMEDIATES BY *E. COLI* CELLS……...…….....114

4.3 CHARACTERIZATION OF λ TAIL TIP BY SUCROSE VELOCITY GRADIENT...119

4.3.1 Construction of biologically active T7-tagged gpL……..……..120

4.3.2 Separation of λ tail tip complex by sucrose velocity gradient…..……..…120

4.4 SEPARATION OF λ TAIL BY GLYCEROL GRADIENT…….……..……138

4.5 PURIFICATION OF λ PHAGE BY CESIUM CHLORIDE STEP GRADIENT……..142

4.6 DISCUSSION……...…….…..146

4.6.1 gpK may be loosely associated with the complex…….…….146

4.6.2 Estimation of the stoichiometries of the tail tip proteins…….…...147

4.6.3 DnaK and the biologically active assembly intermediates……..…..149

5.0 CHARACTERIZATION OF THE CYSTEINE RESIDUES IN λ gpL…….151

5.1 INTRODUCTION…….……..……151

5.2 DETERMINATION OF THE ROLE OF THE CYSTEINE RESIDUES IN λ gpL… ..154

5.2.1 Construction of mutation in gene *L*……..……...154

5.2.2 Determination of the biological activity of the mutant proteins *in vivo*…….……154

5.2.3 Investigation of the conserved cysteine residues in λ gpL…….……158

5.3 FURTHER CHARACTERIZATION OF C212 OF λ gpL…….……...….171

5.4 DISCUSSION……...…….……..175

6.0 CHARACTERIZATION OF λ gpK…….……...……..181

6.1 INTRODUCTION……..…….……181

6.2 CONSTRUCTION AND CHARACTERIZATION OF AMBER MUTATIONS IN λ GENE *K*……..……..182

6.2.1 Construction of amber mutations in λ gene *K*…….……...182

6.2.2 Determination of the biological activity of the amber fragments *in vivo*……..….183

6.2.3 Characterization of the amber mutants *in vitro*……..…...……..183

6.3 DISCUSSION……...…….……..188

7.0 DISCUSSION……...……192

7.1 CHANGES ON ATTACHMENT TO THE HOST CELL……..……192

7.2 DnaK MAY BE REQUIRED FOR THE TAIL TIP COMPLEX TO BE BIOLOGICALLY ACTIVE…….……...192

7.3 gpK MAY BE A LYSOZYME……..……...……...193

7.4 COMPOSITION AND STOICHIOMETRY OF THE TAIL TIP COMPLEX……...194

BIBLIOGRAPHY……...……195

LIST OF TABLES

Table 2.1 Plasmids used in this study……..……..38

Table 2.2 Primers used in this study……..……51

Table 3.1 Description of plasmids used in the spot complementation assay……..……...66

Table 3.2 Phages used in the spot complementation assay……..……..68

Table 3.3 Predicted molecular weights of gpI, gpL and gpK…….....……...72

Table 3.4 Description of plasmids used in the spot complementation assay……..…...75

Table 3.5 Description of the plasmids used in the spot complementation assay……..….85

Table 4.1 Stoichiometries of λ tail tip proteins in the samples precipitated by *E. coli* cells…...117

Table 4.2 Plasmids used in the spot complementation assay……..…….121

Table 4.3 Quantification of λ tail proteins from tails purified by glycerol gradient……...141

Table 4.4 Quantification of lambda tail proteins from purified phage…….…...145

Table 4.5 Summary of the copy numbers of tail proteins from different preparations……..….148

Table 5.1 Description of plasmids and positions of mutation…….....155

LIST OF FIGURES

Figure 1.1 Schematic diagram of five bacteriophages……..……...……3

Figure 1.2 Protein rearrangements during the hexagonal-to-star change of the T4 baseplate…..…9

Figure 1.3 The morphology of bacteriophage lambda……..…….14

Figure 1.4 The main steps of the life cycle of bacteriophage lambda…….……..15

Figure 1.5 The tail assembly pathway of bacteriophage lambda……..……..…18

Figure 1.6 The morphogenetic genes of bacteriophage λ……..……24

Figure 2.1 Overview of the recombinant PCR……..…….46

Figure 2.2 Strategy for QuikChange site-directed mutagenesis…….……...49

Figure 3.1 The mutant protein has little *in vivo* complementation activity …...…….…..69

Figure 3.2 Establishment of the identities of gpI, gpL and gpK……..……..73

Figure 3.3 The C-terminal part of λ gpJ…….……...77

Figure 3.4 Spot complementation assay for the polyhistidine tagged gpJ……..……...79

Figure 3.5 Expression of λ tail tip complex……..…….80

Figure 3.6 Determination of the biological activity of 6XHis tagged proteins…….……86

Figure 3.7 Determination of the biological activity of the MBP fusion proteins…….….88

Figure 3.8 Determination of the biological activity of the GST fusion proteins……...90

Figure 3.9 Time-course expression of MBP and MBP-I……..…….92

Figure 3.10 Time-course expression of GST-I in Rosseta^{TM}2……..…….94

Figure 3.11 MBP-I is subject to degradation after being synthesized …….…….....96

Figure 3.12 Purification of 6XHis tagged gpL by nickel affinity chromatography……..99

Figure 3.13 Purification of gpL by MBP fusion…….....…100

Figure 3.14 Purification of gpL by GST fusion……..…….103

Figure 3.15 The effect of temperature during expression on the solubility of MBP-K….……..106

Figure 3.16 Purification of gpK by MBP fusion……..……108

Figure 4.1 Precipitation of λ tail tip assembly intermediates by *E. coli* cells……..……115

Figure 4.2 The T7-tagged gpL is biologically active…….....…….122

Figure 4.3 Schematic diagram of the *in vitro* complementation assay……..……..124

Figure 4.4 The majority of the *in vitro* complementation activity appears in the fifth fraction from the sucrose gradient…….…….…….127

Figure 4.5 Most of the tail proteins are present in the top fractions from the gradient and are not able to complement *in vitro*……..……..…….….129

Figure 4.6 The T7-tagged gpI is biologically active……..……..……132

Figure 4.7 The fraction with most of the *in vitro* complementation activity does not overlap with the fraction containing most of the tail tip proteins…….....134

Figure 4.8 Band xx3 is DnaK..……..……...136

Figure 4.9 Separation of lambda tail by glycerol gradient……..…….....139

Figure 4.10 Purification of lambda phage by cesium chloride step gradient……..….143

Figure 5.1 Multi-sequence alignments of phage λ gpL and its homologs……..….152

Figure 5.2 Determination of the biological activity of the mutant proteins *in vivo*…….156

Figure 5.3 Spot complementation assay to examine the activity of the mutants…….159

Figure 5.4 Effect of single amino acid substitution on lambda tail assembly……..…162

Figure 5.5 The mutant tail is as efficient as the wildtype tail in attachment to λ head…..……167

Figure 5.6 The mutant tail can attach to host cells as efficiently as the wildtype tail…….169

Figure 5.7 Spot complementation test to determine the biological activity of C212A and C212H mutants…...……...……..……..…….172

Figure 5.8 Separation of λ tail by glycerol gradient…….……...173

Figure 5.9 Precipitation of λ tail assembly intermediates by *E. coli* cells……..…….176

Figure 5.10 Summary of the cysteine mutants in λ gpL……..……178

Figure 6.1 Determination of the biological activity of the amber fragments *in vivo*……...184

Figure 6.2 Separation of λ tail by glycerol gradient…….……...186

Figure 6.3 Separation of lambda phage by cesium chloride step gradient…...…….…..189

PREFACE

I would like to thank Dr. Roger Hendrix for constructive suggestions and guidance during my graduate career in his laboratory.

I wish to thank my thesis committee members, Dr. Peter Berget, Dr. Jeffrey Brodsky, Dr. Graham Hatfull and Dr. James Pipas for their insightful suggestions.

I want to thank all the members in the laboratory, especially Dr. Robert Duda and Brian Firek who gave me lots of helpful advice for my research.

Finally, I would like to express my special thanks to my parents and my brothers. Without their support, I would not have gone so far.

1.0 INTRODUCTION

Bacteriophages are the most abundant life form on the planet (Wommack and Colwell, 2000; Hendrix, 2003). They are widely distributed in the world, including soils, hot springs, deep seas and so on. Due to their relatively simple structures, phages have been used as model systems to learn how proteins interact with each other to assemble macromolecular structures.

The bacteriophage lambda has been proven to be a good model system to study the regulation of the assembly of complex structures. It has an icosahedral head and a long, flexible, non-contractile tail (Kellenberger, 1961; Kemp et al., 1968). In tailed phages, the tail carries out an important role in host recognition and DNA injection. It is likely that the tail, especially the tail tip / baseplate, needs to undergo conformational change and protein rearrangement in order to make channels for the translocation of viral DNA into the cytoplasm. This change has been reported in the long, contractile tail of phage T4 and the short tail of phage T7 (Simon and Anderson, 1967; Crowther et al., 1977; Molineux, 2001; Kemp et al., 2005). In terms of the flexible, non-contractile tail of bacteriophage lambda, the tail assembly pathway has been established for more than thirty years (Katsura and Kuhl, 1975). However, little is known about the function and positions of the tail proteins, neither do we know much about the molecular machine of DNA injection. This thesis is aimed to learn some themes of the lambda tail tip, which may help understand the general mechanism of lambda DNA injection.

1.1 THE TAIL OF BACTERIOPHAGES

The tailed bacteriophages are generally believed to be the most diversified of all virus groups. Most of them have an icosahedral head and a tail that may vary in structure among different phages. Based upon the morphology of the tail, bacteriophages can be divided into three groups: (1) myoviridae: a family of phage with long, non-flexible, contractile tails like T4; (2) siphoviridae: a family of phage with long, flexible, non-contractile tails like λ; (3) podoviridae: a family of phage with short, non-contractile tails like T7 (Figure1.1) (Maniloff and Ackermann, 1998). The difference in the structure of the tail among different phage groups may reflect the different mechanisms they adopted to inject viral DNA during infection.

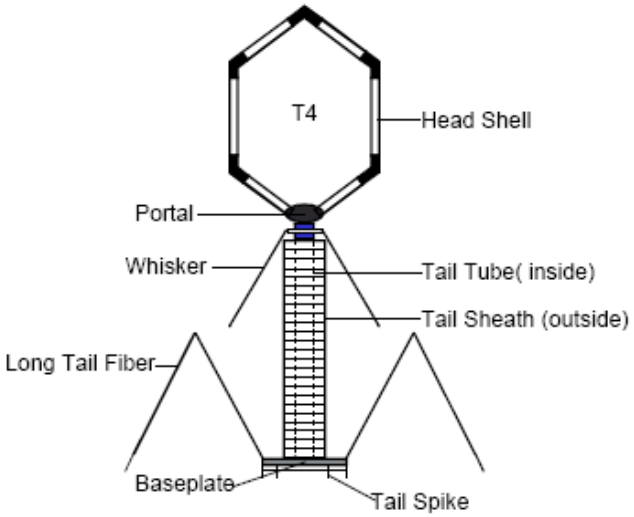
1.1.1 Tail structure

Although most of the tailed bacteriophages have a similar morphology of the head, the morphology of the tail differs greatly among different groups of bacteriophages (Casjens and Hendrix, 1988) (Figure 1.1). In spite of the difference in the morphology of the tail, all the bacteriophage tails carry out a key role in phage adsorption to the host cell and DNA injection during infection. Therefore, it is not surprising for bacteriophages to share some common features of the tail, even though the same substructure may be different in morphology. The common features of the bacteriophage tail will be listed below.

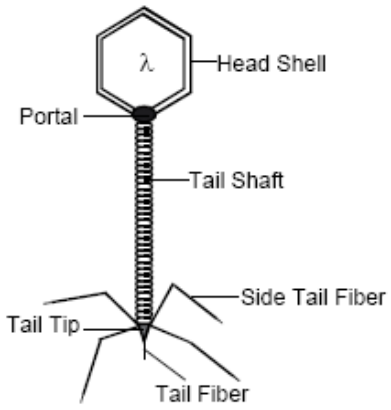
Tail fiber: The tail fiber is located at the distal end of the tail. It is the most important substructure in the process of attachment of phage to the host cell. The structure of tail fibers varies among different types of phages. The bacteriophage T4 has six tail fibers that are

Figure 1.1 Schematic diagram of five bacteriophages (Casjens and Hendrix, 1988; Xu, 2001).

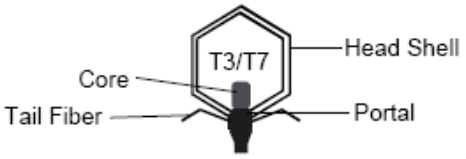
Although the morphology of the head is quite similar among tailed bacteriophages, the morphology of the tail is dramatically different. The myoviridae phage T4, for instance, possesses a long, rigid, contractile tail whereas the siphoviridae phage λ has a long, flexible, non-contractile tail. Phages T3/T7, P22 and ϕ29 have short tails and belong to the podoviridae family.



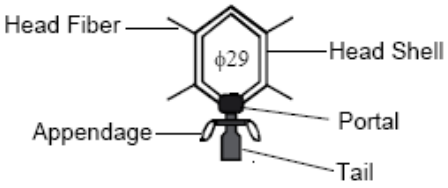
(myoviridae)



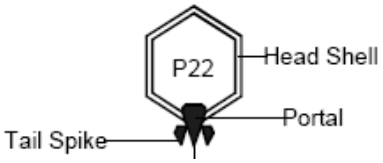
(siphoviridae)



podoviridae



podoviridae



podoviridae

responsible for binding to the host cells (Kellenberger et al., 1965; Simon and Anderson, 1967; Simon and Anderson, 1967). The well-studied bacteriophage λ has a single, thin tail fiber which interacts with the LamB host receptor (Randall-Hazelbauer and Schwartz, 1973; Szmelcman and Hofnung, 1975) and determines the host specificity during infection (Dove, 1966; Mount et al., 1968; Buchwald and Siminovitch, 1969). In the short-tailed bacteriophages, the tail fiber may be also called spikes, thick fibers, appendages or tail proteins.

Tail tip / Baseplate: The tail tip (λ like) or baseplate (T4 like) is where the tail fiber is bound. It is a multi-protein structure in the tail with various shapes and complexity in different phages and plays an important role in host recognition and DNA injection. It may undergo dramatic conformational change and protein rearrangement in order to allow the passage of phage DNA into the cell. The T4 baseplate is a well-characterized multiprotein machine which carries out an important role during infection. It has a dome-shaped structure with a central spike formed by the gp5-gp27 (gp for **g**ene **p**roduct) complex (Kanamaru et al., 2002; Kostyuchenko et al., 2003). The positions and stoichiometry of individual proteins have been determined by fitting X-ray structures into the cryo-EM density (Rossmann et al., 2001; Rossmann et al., 2004).

In terms of bacteriophage λ, four proteins are known to be required for the tail tip formation. The order of interactions among these proteins was determined by *in vitro* complementation assays (Katsura and Kuhl, 1975; Katsura, 1976). However, little is known about the structure and composition of the tail tip, probably due to the small size of the tail tip and the difficulty in the expression and purification of these proteins as will be shown in the following chapters.

The short-tailed bacteriophages like P22 and T7 have a single, central structure surrounded by a baseplate or collar. The shape of this structure appears greatly variable among them. In addition to the function of the baseplate proteins, they may also behave like head and tail completion proteins.

Side tail fiber (stf): In addition to the tail fiber, many phages also have several long, L or Z shaped side tail fibers attached to the tail tip / baseplate. They help the adsorption of the bacteriophages to the host cell surface and increase host specificity during infection. In some phages, such as T4, the long fibers are absolutely required infection, possibly because they trigger essential conformational rearrangement of the baseplate (Makhov et al., 1993). In other phages, including λ and T5, the phage is still infective in the absence of the fibers, although the kinetics of adsorption to the cell is slower (Heller and Braun, 1979; Hendrix and Duda, 1992). In both T4 and λ, the assembly of the side tail fiber requires a chaperone (Tfa in λ and gp38 in T4). The chaperone is absent in the mature structure of the side tail fiber of phage T4, but it is present in the final structure of the side tail fiber of phage λ (Hendrix and Duda, 1992). It has been reported that the λ Tfa can functionally substitute for the T4 gp38 *in vivo* (Montag and Henning, 1987). The four side tail fibers have been missing in the common laboratory stain of bacteriophage λ (λ PaPa) since 1950's due to a single base deletion in the gene encoding side tail fiber. Compared to λ PaPa, Ur-λ, the original isolate with side tail fibers adsorbs to host cells more quickly (Hendrix and Duda, 1992).

Tail shaft: Bacteriophages with long tails have a structure called the tail shaft formed on the tail tip / baseplate, which is likely to be a passive conduit for phage DNA due to the lack of evidence

that it actively participates in propelling phage genome into the cell (Casjens and Hendrix, 1988). There are generally two types of tail shaft: rigid, contractile ones and flexible, non-contractile ones. In the myoviridae family of T4, the tail shaft consists of an internal cylindrical tube assembled from gp19, and an outer, contractile sheath which is composed of 138 molecules of gp18 arranged into 23 hexameric rings (King, 1968; Amos and Klug, 1975; Moody and Makowski, 1981; Leiman et al., 2004). It is the tail sheath that contracts dramatically during DNA injection. The tail shaft of the siphoviridae phage λ contains 32 stacked rings of six gpV subunits each (Buchwald et al., 1970; Casjens and Hendrix, 1974; Katsura, 1983). Despite the difference in the morphology of the tail shaft of phage T4 and λ, both of them use the tape measure protein to determine the tail length. The detailed mechanism about how the tape measure protein determines the tail length is not well understood. It is interesting that the tails of phages which are even not very closely related can have an identical appearance except for the different tail lengths (King, 1968; Youderian, 1978; King, 1980; Katsura and Hendrix, 1984). The long tails may allow phages to search for suitable receptors in the surrounding environment more efficiently (Schwartz, 1976).

Tail completion protein: At the head-proximal end of the tail are the tail completion proteins which stop the polymerization of the major tail protein and probably create a site for the attachment of head. In bacteriophage λ, the terminator protein gpU forms a hexameric ring which caps the tail shaft when it reaches the correct length, and the tail is activated by the addition of gpZ (Casjens and Hendrix, 1974; Katsura, 1976; Edmonds et al., 2007). In bacteriophage T4, two proteins, gjp3 and gp15, are required for the completion of tail assembly. Each of them forms a hexameric ring, which are required for the termination of the tail tube

elongation (gp3) and formation of a connector which enables the tail to bind to the head (gp15) (King, 1968; Vianelli et al., 2000; Zhao et al., 2003). In terms of the short- tailed phages, the tail is assembled directly on the head and it is not useful to distinguish between the head completion proteins and the tail completion proteins.

1.1.2 Structural movements and DNA injection

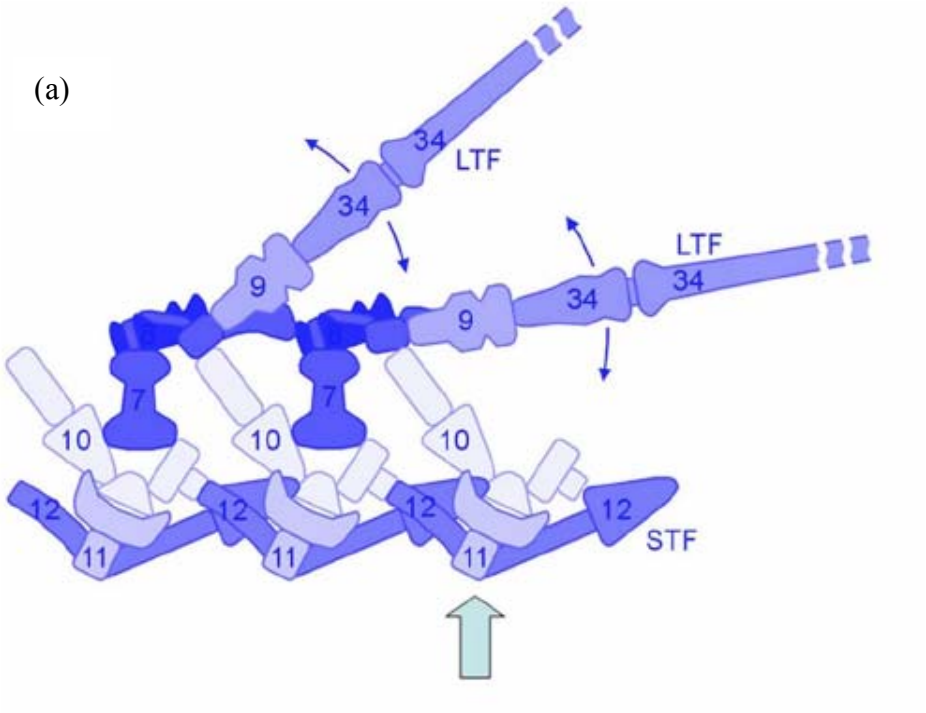
In spite of the dramatic differences in the morphology, all the tails of bacteriophages play a key role in host recognition and DNA injection. The course of infection begins with the attachment of phage to the host cell surface with the help of the tail fiber, followed by an intricate process to inject phage genome into the cytoplasm. The injection of DNA is generally accompanied by dramatic conformational change and protein rearrangement.

In bacteriophage T4, the baseplate has a hexagonal shape in the mature virion. Upon attachment to the host cells, the T4 baseplate undergoes extensive protein rearrangement which results in a star shape (Simon and Anderson, 1967; Crowther et al., 1977). For instance, upon the completion of the hexagonal-to-star conformational transition, gp11, a component of the wedge, rotates by approximately 100° and interacts with the long instead of the short tail fiber (Leiman et al., 2004; Arisaka, 2005) (Figure 1.2). The hexagonal-to-star conformational transition triggers sheath contraction, which is due to the substantial rearrangement of the domains within the sheath protein (Moody, 1973; Leiman et al., 2004). The sheath contraction leads to the extension of the tail tube beneath the baseplate (Leiman et al., 2004). With the help of the tail lysozyme, gp5, the tail tube penetrates through the cell envelope (Kanamaru et al., 2002), followed by the release of the phage DNA into the cytoplasm through the tube (Simon and Anderson, 1967).

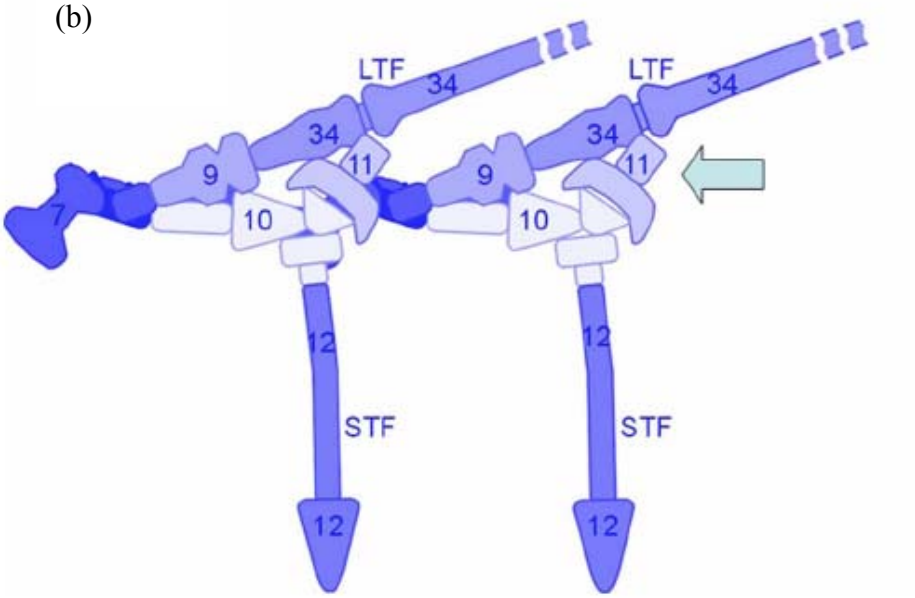
Figure 1.2 Protein rearrangements during the hexagonal-to-star change of the T4 baseplate (Arisaka, 2005).

In the mature virion, gp11 (as indicated by arrows in (a) and (b)) is bound to the kink region of the short tail fiber (STF) gp12. After completion of the conformational change, it moves to interact with the long tail fiber (LTF) gp34. Instead of being accommodated beneath the baseplate in the mature virion, the short tail fibers extrude downward with their tips binding tightly to the cell surface after the conformational change. The numbers in the figure indicate the corresponding gene products.

(a)


34
LTF
34
LTF
9
34
34
9
7
7
10
10
10
12
12
12
12
11
11
11
STF

(b)


LTF
34
LTF
34
34
34
11
11
9
9
7
10
10
12
12
STF
STF
12
12

The podoviridae phage T7 has a short, non-contractile tail. Unlike most other tailed bacteriophages whose tail assembly is independent of head assembly, the stubby tail of T7 forms directly on the DNA-filled head (Studier, 1972; Roeder and Sadowski, 1977; Matsuo-Kato et al., 1981). The 23-nm-long tail is too short to span the *E. coli* cell envelope. In contrast to the well-studied contractile tail of T4, five virion proteins of T7 (three internal core proteins: gp14, gp15 and gp16, the head protein gp6.7, and the tail protein gp7.3) are ejected into the cytoplasm at the initiation of infection, which functionally endows T7 with an extensible tail and allows the translocation of phage genome into the cell (Molineux, 2001; Kemp et al., 2005; Serwer et al., 2008). The ejected virion proteins need to undergo dramatic conformational change at the initiation of infection (Moak and Molineux, 2000; Molineux, 2001; Kemp et al., 2005). The three internal core proteins, gp14, gp15 and gp16, must disaggregate from the structure and almost completely unfold in order to pass through the head-tail connector (Moak and Molineux, 2000; Molineux, 2001; Kemp et al., 2005). The internal core protein of T7, gp16, has transglycosylase activity and facilitates infection (Moak and Molineux, 2000). The details of how the proteins change their conformation in order to be ejected and how they interact with each other to form the extended tail remain unknown.

Little is known about how the siphoviridae phage λ translocates its viral DNA into the host cell. This is probably due to the lack of detailed information about the structures of the tail and the individual proteins. It has been inferred from previous experiments that the tail fiber protein, gpJ, probably moves during infection (Roessner and Ihler, 1984) and gpH* may exit the tail and enter into the cell with the phage genome during DNA injection (Roa, 1981).

1.1.3 Cysteine residues in the bacteriophage tail proteins

In prokaryotes, many proteins contain cysteine residues, some of which are essential for protein folding (Gilbert, 1990). The cytoplasm of bacteria is widely believed to be unfavorable for disulfide bond formation due to its reducing environment. In spite of this, there has been evidence that some cysteine residues in the bacteriophage tail proteins carry out important roles in protein formation or tail assembly which may involve the transient formation of disulfide bonds.

In the baseplate of bacteriophage T4, each of the six wedge complexes is composed of seven proteins, including three copies of gp10 (Zhao et al., 2000). Gp10 has five cysteines, three of which are involved in disulfide bond formation (Zhao et al., 2000). Two molecules of gp10 in the complex participate in a disulfide bond formation while the remaining one does not (Zhao et al., 2000). Although the detailed information about how this asymmetric structure is formed by gp10 and what the role of the disulfide bond is in tail assembly remains an open question, it may be safe to say that some or all of the five cysteines in gp10 may have a key role in the wedge formation.

The *Salmonella* phage P22 has a short tail. The tailspike protein of P22 has 8 cysteines among its 666 amino acids (Sauer et al., 1982). No disulfide bonds have been found in the native state of the tailspike protein (Sargent et al., 1988; Steinbacher et al., 1994). However, there is a transient interchain disulfide bond formation in the protrimer intermediates which is required for the correct folding and assembly of the tailspike (Sather and King, 1994; Robinson and King, 1997).

The λ tail assembly is controlled by at least eleven proteins (Campbell, 1961; Parkinson, 1968), some of which contain several cysteine residues in their protein sequences. For example, gpJ has 11 cysteines among its 1132 amino acids whereas gpL contains 8 cystienes among its 232 amino acids. Some of the cysteine residues are highly conserved among the homologs and may carry out important roles in protein formation, tail assembly or DNA injection. The cysteine residues in gpL will be investigated in this study.

1.2 THE BACTERIOPHAGE LAMBDA

The bacteriophage λ has an icosaheral head with 55nm in diameter and a long, non-contractile tail that is 150 nm long excluding the 23-nm tail fiber (Kellenberger, 1961; Kemp et al., 1968; Katsura, 1983) (Figure 1.3). The morphogenesis of λ is controlled by a number of genes located in the left arm of the phage chromosome (Campbell, 1961; Weigle, 1966; Parkinson, 1968). Due to the large number of mutants available and the vast accumulation of knowledge about this phage, the λ tail has been recognized as a good model system for the study of biological assemblies. The following will give a summary of the λ growth cycle and some knowledge concerning the λ tail.

1.2.1 Life cycle

λ is a temperate phage and can follow two different life cycles (Figure 1.4). At the beginning of the life cycle, λ binds to its host receptor, LamB, by its tail fiber and forms a reversible complex (Randall-Hazelbauer and Schwartz, 1973; Szmelcman and Hofnung, 1975). Subsequently, an

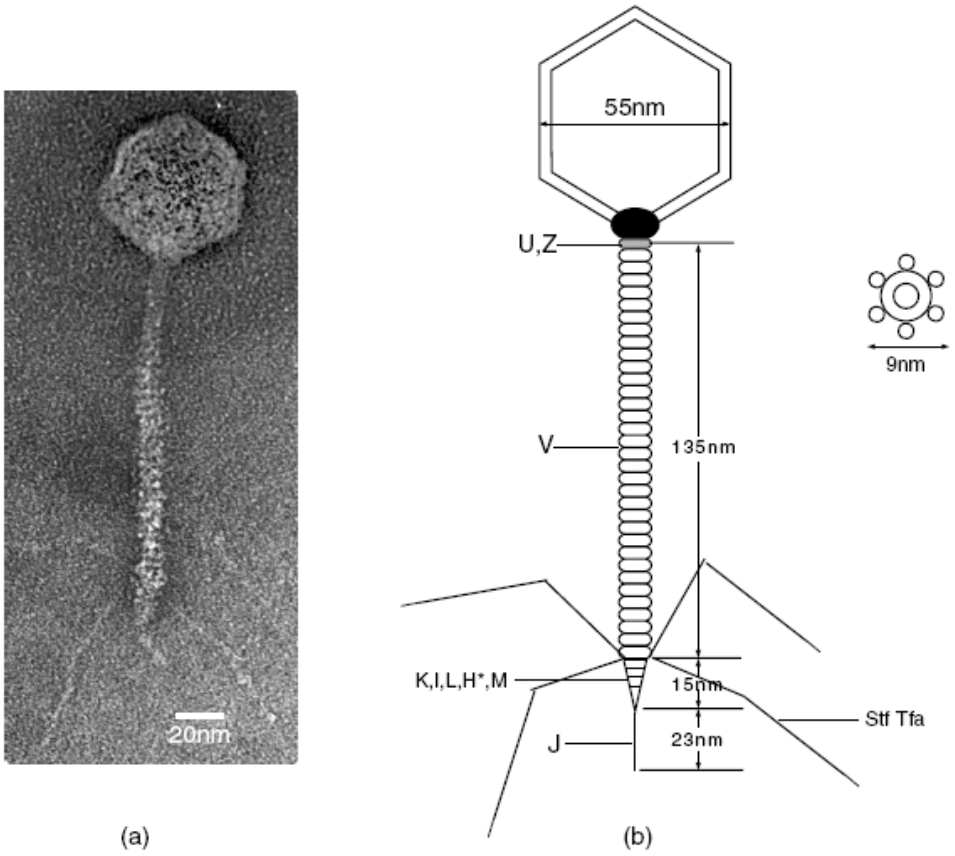


Figure 1.3 The morphology of bacteriophage lambda. (a). Negatively stained bacteriophage Ur-λ (Courtesy of Duda, R.L.). (b). Schematic diagram of the bacteriophage λ and a disk at end view (Katsura, 1983; Xu, 2001). The bacteriophage λ has an icosahedral head 55 nm in diameter and a long, non-contractile tail. The tail is composed of the 135-nm-long tail shaft and 38-nm-long tail tip. The tail shaft is comprised of 32 hexameric rings.

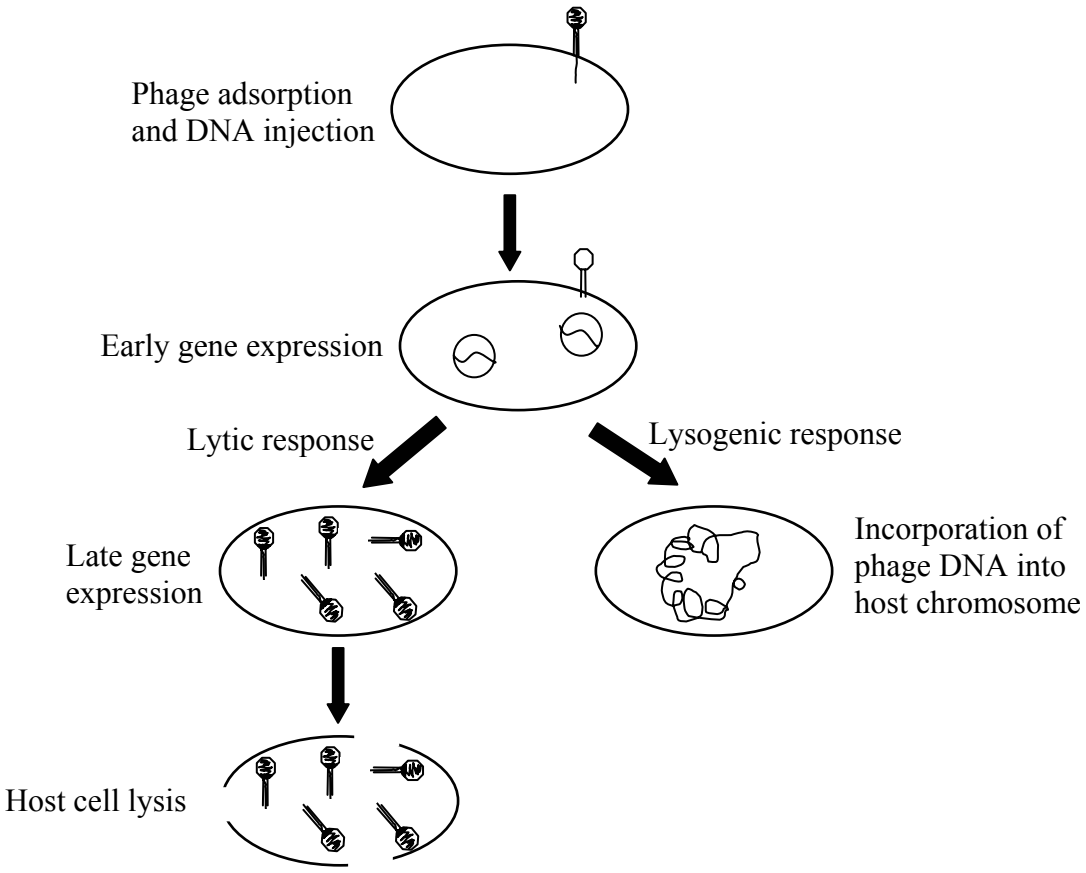


Figure 1.4 The main steps of the life cycle of bacteriophage lambda (Redrawn from (Arber, 1983) with modifications). λ is a temperate phage and can follow two different life cycles. In the lytic cycle, λ multiplies inside the bacteria. Progeny phages are released after cell lysis. In the lysogenic cycle, the phage genome is incorporated into the host cell chromosome and replicated in a passive way.

inactivation step makes the phage-LamB complex irreversible, followed by a 'delay time' before triggering the injection of the linear double-stranded phage DNA into the cytoplasm (Roa and Scandella, 1976). This pause can be shortened by increasing the temperature, with a maximum at 37℃(Lof et al., 2007). When the temperature is at 4℃, phage DNA will not be injected at all (Novick and Baldeschwieler, 1988; Lof et al., 2007). The general mechanism of DNA injection is not well understood. However, it is possible hat the tail undergoes conformational change and protein rearrangement in order to inject the viral DNA.

Early genes, which are mainly responsible for the regulation of phage development, are expressed a little after DNA injection. Subsequently, either of the following two pathways may occur: (1) In the lytic cycle, the phage multiplies inside the bacteria. A phage-encoded enzyme called transglycosylase then lyses the host cell, leading to the liberation of about one hundred progeny phages which are synthesized within 50 minutes at 37℃; (2) In the lysogenic cycle, the phage DNA is incorporated into the host cell chromosome at a characteristic site and replicated in a passive way (Hershey, 1983). Almost all of the phage genes are repressed in the lysogenic cycle and therefore no progeny phages are produced. The phage DNA at this stage (prophage) can enter the lytic cycle through induction, a process in which the phage genes are derepressed and phage DNA is excised from the host chromosome. The lysogenic (prophage-containing) cells can be induced at high frequency by exposure to DNA-damaging agents such as ultraviolet light. The signal which allows prophage to undergo this transition "spontaneously" at low frequency *in vivo* is not clear. The early genes determine which of the above pathways the phage will take (Herskowitz and Hagen, 1980).

The late genes, including the phage morphogenetic genes and lysis genes, are expressed in the lytic cycle a little while after infection or induction until cell lysis. The morphogenetic proteins of λ are transcribed from the same operon and the functional half-life of the mRNA for each gene is not significantly different, but the morphogenetic proteins are synthesized in widely different molar amounts (Murialdo and Siminovitch, 1972; Ray and Pearson, 1974; Sampson et al., 1988). This is probably due to the various rates of the initiation of translation (Sampson et al., 1988). Lambda head and tail assemble independently from the morphogenetic proteins, which subsequently join together to form a biologically active phage (Weigle, 1966; Weigle, 1968). The products of lysis genes cause host cell lysis, which leads to the release of progeny phages.

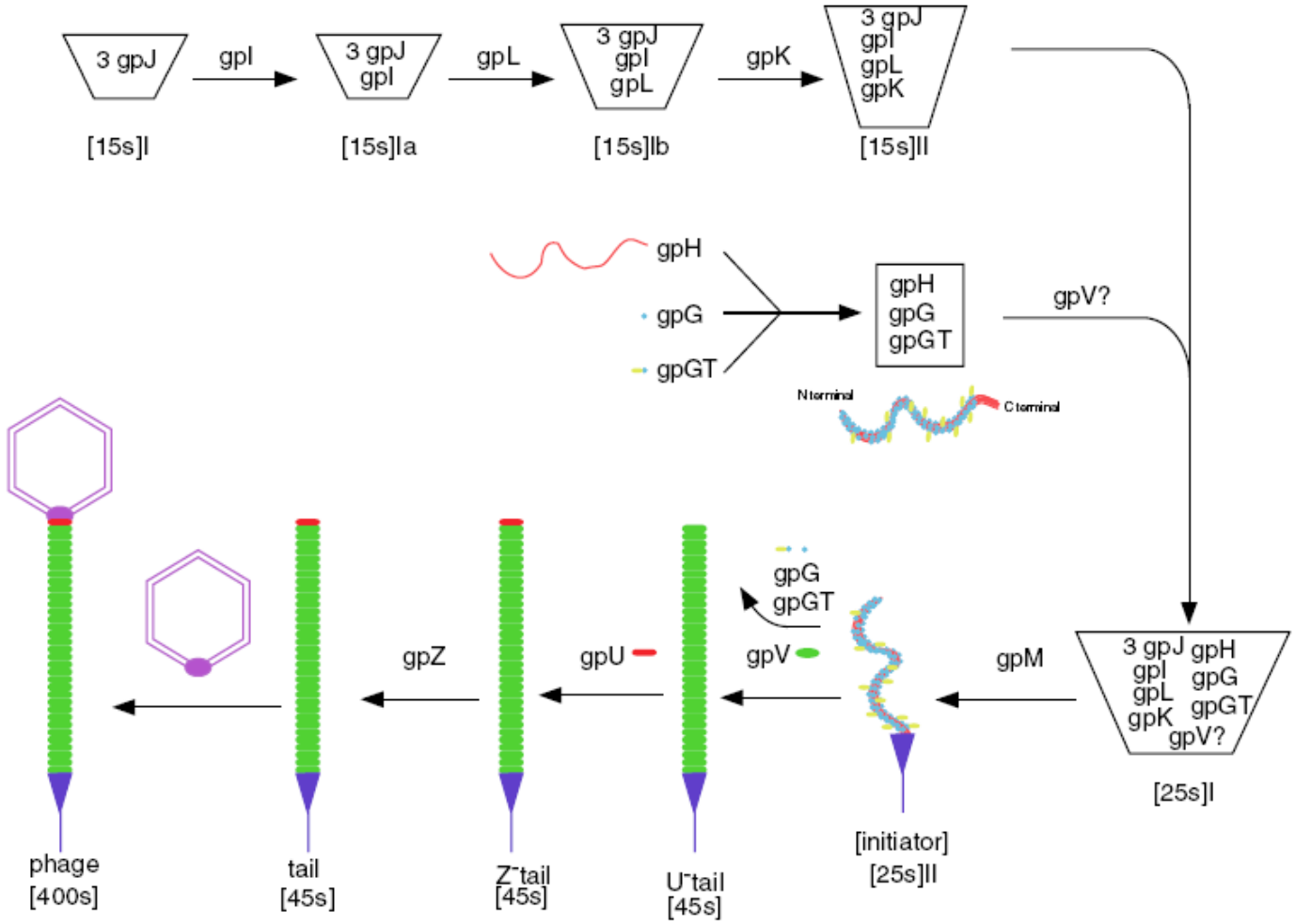
1.2.2 Lambda tail assembly pathway

The assembly of λ tail starts from the tail fiber, located at the distal end of the tail, and proceeds to the proximal end of the tail. At least eleven genes are involved in the tail assembly pathway (Campbell, 1961; Parkinson, 1968). The assembly pathway was established by a series of *in vitro* complementation assays in 1970s (Katsura and Kuhl, 1975; Katsura, 1976) and was modified recently based on new data (Xu, 2001) (Figure 1.5). The λ tail assembly pathway can be divided into three steps and the following will give a summary of each step.

Formation of the initiator: The formation of the initiator begins with three copies of gpJ to form a [15S]I particle, followed by the sequential addition of gpI, gpL and gpK to form a [15S]II particle (Katsura and Kuhl, 1975; Katsura, 1976). Before joining the [15S]II complex to form a complex, gpH and its chaperones, gpG and gpGT, form a subassembly (Xu, 2001). Finally, gpM is added to stabilize the 25S complex (Katsura and Kuhl, 1975).

Figure 1.5 The tail assembly pathway of bacteriophage lambda (Katsura and Kuhl, 1975; Xu, 2001).

The primary assembly pathway has been worked out by a series of *in vitro* complementation assays (Katsura and Kuhl, 1975). A small modification has been made by Xu (2001) who showed that gpG and gpGT are chaperones for gpH, and the three proteins form a subassembly before joining the 15S complex. It is not clear, however, whether or not a few molecules of the major tail protein, gpV, might be involved in the formation of the initiator.


3 gpJ
gpI
3 gpJ gpI
gpL
3 gpJ gpI gpL
gpK
3 gpJ gpI gpL gpK
[15s]I
[15s]Ia
[15s]Ib
[15s]II
gpH
gpG
gpGT
gpH gpG gpGT
gpV?
Nterminal
Cterminal
3 gpJ gpH gpI gpG gpL gpGT gpK gpV?
[25s]I
gpM
gpG gpGT
gpV
gpU
gpZ
[initiator] [25s]II
U⁻tail [45s]
Z⁻tail [45s]
tail [45s]
phage [400s]

Polymerization of the major tail protein: After formation of the initiator, the major tail protein gpV begins to polymerize upon the initiator. In the absence of the initiator or of something that imitates the initiator, the major tail protein exists in a monomer-dimer equilibrium and can not assemble into tail-related structures (Katsura and Tsugita, 1977).

Termination and maturation: The polymerization of the major tail protein gpV pauses for a while when the tail reaches the correct length which is determined by the tape measure protein gpH (Katsura and Hendrix, 1984). At this point, gpU, the terminator protein which can form hexameric rings, joins tail assembly and stops polymerization of gpV. In the absence of gpU, gpV continues to polymerize which results in polytails (Mount et al., 1968; Katsura, 1976). At about the same time, a peptide of about 10 kDa is cleaved off from the tape measure protein gpH (Murialdo and Siminovitch, 1972; Tsui and Hendrix, 1983). At this stage, the tail looks normal, but it is defective. After addition of gpZ, the tail is activated and is ready to join the head to form a biologically active phage (Katsura and Kuhl, 1975).

The picture of λ tail assembly pathway has been known for more than thirty years. However, it leaves many questions unanswered. For instance, the function and exact positions of most proteins in the structure remain unknown. Besides, we know little about the molecular machine of DNA injection.

1.2.3 Function of the individual tail proteins

The λ tail assembly is under the control of at least eleven phage proteins. Thanks to the various mutants available and the *in vitro* complementation assay, the order of interactions among these

proteins has been established (Katsura and Kuhl, 1975). However, little is known about the structure and function of most of the tail proteins. The following is a summary of the available information for λ tail proteins.

gpJ: gpJ, the tail fiber protein of phage λ, is the first protein involved in the tail assembly. Three copies of gpJ comprise the single, thin tail fiber, which is located at the distal end of the tail (Casjens and Hendrix, 1974). The size of the fiber only accounts for a minority of the mass of the three molecules of gpJ, so most of the mass of the gpJ subunits must be in the conical tail tip (Roessner and Ihler, 1987). It interacts with the host cell receptor and determines host specificity during infection (Dove, 1966; Mount et al., 1968; Buchwald and Siminovitch, 1969). It has been reported that the C-terminal part of gpJ is responsible for the interaction with the LamB host receptor (Randall-Hazelbauer and Schwartz, 1973; Szmelcman and Hofnung, 1975; Wang et al., 2000).

gpG and gpGT: gpG and gpGT are related by a programmed -1 translational frameshift (Levin et al., 1993). They are chaperones for the tape measure protein, gpH (Xu, 2001). Although both proteins interact with gpH, only the T part of gpGT has interactions with the major tail protein, gpV (Xu, 2001). Neither gpG nor gpGT is present in the mature tail (Levin et al., 1993).

gpH: The tape measure protein, gpH, is the one that determines the tail length of phage λ (Katsura and Hendrix, 1984) and may be involved in DNA injection (Scandella and Arber, 1976; Katsura, 1983). Shorter-tailed phage particles can be obtained by in-frame deletion of gpH (Katsura and Hendrix, 1984). About six copies of gpH are present in each phage (Casjens and

Hendrix, 1974). The 90-kDa gpH is processed to a 78-kDa product, gpH*, during tail assembly (Hendrix and Casjens, 1974). It is gpH*, but not gpH, that is present in the mature virion.

gpV: gpV is the major tail protein which forms the tail shaft. About one-third of the major tail protein near the C-terminus can be removed and phage particles with the short version of gpV can maintain its total shape and infectivity (Katsura, 1981). The removable C-terminal portion of gpV has been identified as an Ig-like domain (Fraser et al., 2006). Such domains are found at the C-terminus of some but not all major structural proteins of tailed phages. Under certain conditions, gpV may polymerize into polytails (a structure of undertermined tail length with tail tip) or polytubes (a structure of undertermned tail length without the tail tip) (Katsura, 1976).

gpU: gpU, the terminator protein, is located at the head-proximal end of the tail and serves as an interface for attachment of the head to the tail (Katsura and Tsugita, 1977). In the absence of Mg(II), gpU exists as a globular monomer, but it changes to ring-like hexamer upon treatment with 20 mM Mg(II) (Katsura and Tsugita, 1977). The NMR structure of the monomeric form reveals that gpU is a single domain protein possessing a mixed α/β motif and several dynamic loops at the periphery (Edmonds et al., 2007).

gpI, gpL, gpK, gpM and gpZ: The function of these proteins remains unknown. We do know, however, what part of the assembly pathway they act in and we know the Z^- phenotype. The Z^--lysate contains phage-like particles with little infectivity (Katsura and Kuhl, 1975). gpI, gpL and gpK will be investigated in this study.

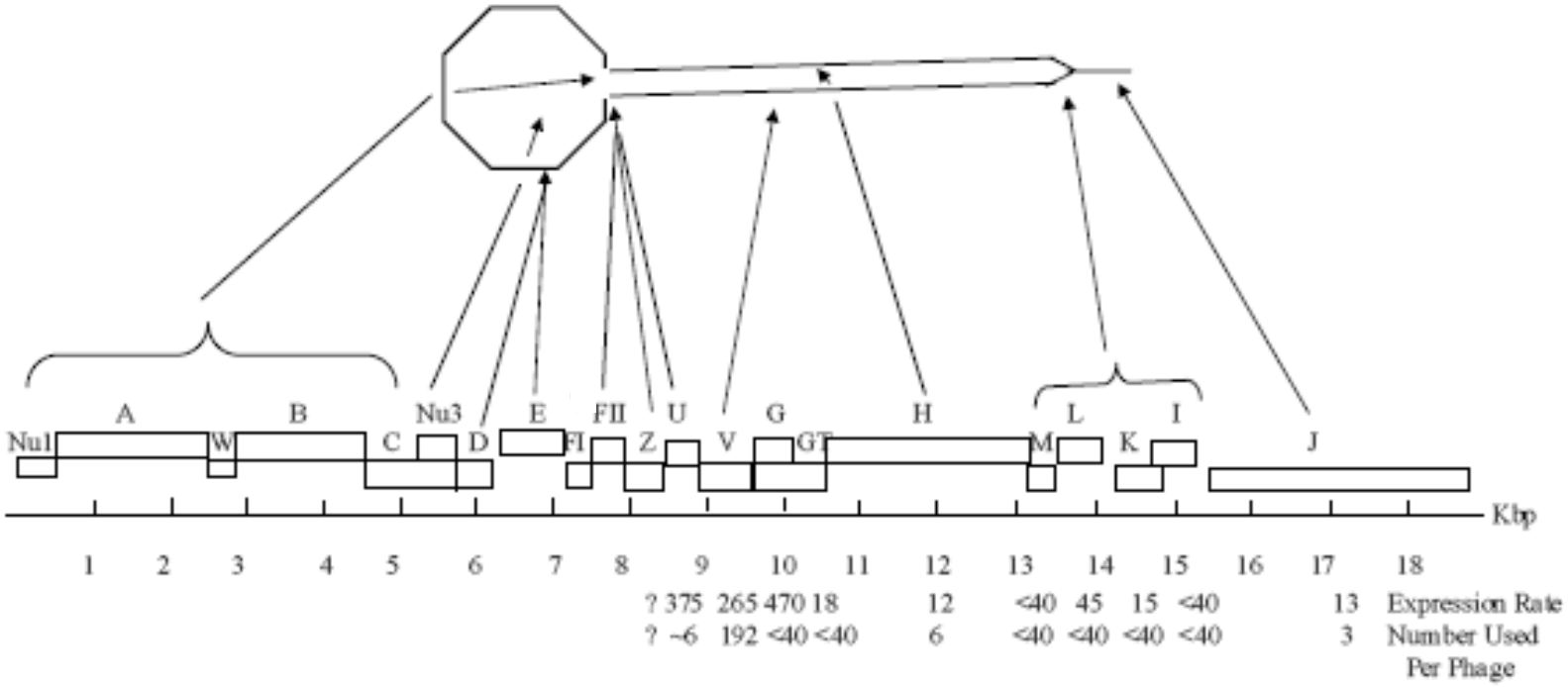
1.2.4 Morphogenetic genes and λ tail assembly

Gene order: Despite the fact that bacteriophages show dramatic differences in morphology, it is interesting to note that the gene order of bacteriophages in their genomes seems to be conserved among different families, particularly for the genes encoding structural components of the virions (Casjens and Henrix, 1988). It is not clear if this conservation of gene order will help us understand the details of the phage assembly process. But, there may be evolutionarily advantages since this linkage may allow gene modules to be evolutionally exchanged among phages and minimizes the formation of defective hybrids (Stahl and Murray, 1966; Casjens and Hendrix, 1974; Echols and Murialdo, 1978; Campbell, 1983). λ is a good example for the study of gene order. Figure 1.6 shows that the head and tail genes of λ are arranged in two non-overlapping, adjacent clusters. In contrast to the order in which the gene products are added to the tail during assembly, transcription of the tail genes of λ proceeds in the reverse order.

Relative expression rates of the morphogenetic genes: It has been known in bacteriophages that the relative expression rates of the morphogenetic genes generally reflect the amounts of the corresponding gene products needed to assemble phage particles. In λ, the structural genes lie within a single operon and the functional half-life of the mRNA for each gene differs less than two-fold, but the morphogenetic proteins are synthesized in widely different molar amounts (Murialdo and Siminovitch, 1972; Ray and Pearson, 1974; Sampson et al., 1988) (Figure 1.6). This ratio among structural proteins may be important in controlling the tail assembly. Deviation from the normal protein ratios may cause problems in assembly. One instance is the λ tail terminator protein gpU, which is made in approximately 50-fold over the amount found in the

Figure 1.6 The morphogenetic genes of bacteriophage λ (Redrawn from (Casjens and Hendrix, 1988) with modifications).

The physical size and positions of all the morphogenetic genes of bacteriophage λ are shown (with a scale in kbp). All the morphogenetic genes lie within a single operon. Normal transcription is from left to right. The schematic diagram of λ is shown above, and the positions of the structural proteins in the virion are indicated by arrows. gpFI is involved in DNA packaging and is not found in the mature virion. gpG and gpGT are chaperones for gpH and they are not present in the mature structure of the virion. The relative expression rate of tail genes and the current best estimated number of molecules of each tail protein required to construct one particle are shown below.


A
B
Nu3
E
FII
U
G
H
L
I
Nu1
W
C
D
FI
Z
V
GT
M
K
J
Kbp
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
? 375 265 470 18 12 <40 45 15 <40 13 Expression Rate
? ~6 192 <40 <40 6 <40 <40 <40 <40 3 Number Used Per Phage

mature virion. In the absence of gpU, gpV continues to polymerize which results in long tails with little infectivity. Thus, it may be essential to have a large amount of gpU in order to make sure gpU terminates the tail in time and no abnormal assembly happens. Another example is gpG and gpGT of λ. They are chaperones for gpH and are related by a programmed frameshift. It has been shown that the ratio between these two proteins is important for normal tail assembly (Xu, 2001). Changes in the ratio of the two proteins may cause trouble in tail assembly (Xu, 2001).

1.2.5 The tail tip of phage λ

The λ tail tip is composed of a conical part and a short, central tail fiber. The conical part may be composed of three or four disks (Katsura, 1983) while the tail fiber is composed of three copies of gpJ. The assembly of λ tail tip is controlled by at least four proteins (gpJ, gpI, gpL and gpK). The order of the interactions between those proteins has been established for many years. However, little is known about its structure probably due to the small size of the tail tip, neither do we know much about the function and stoichiometry of the tail tip proteins. This study is aimed to understand some themes of the lambda tail tip.

1.3 THESIS PLAN

This study is composed of four parts:

(1) The first part deals with the expression and purification of λ tail tip proteins. I first tried to label gpJ with a polyhistidine tag to facilitate the purification of the tail tip complex. However,

most of the complex is insoluble, which makes the purification impractical. Then, I tried to purify the individual tail tip proteins. Each target protein was labeled with three different tags. Several purification methods were used to purify the tail tip proteins. Protein gpL has presented the fewest difficulties of expression and solubility, and its purification has been the most successful.

(2) The second part is about determination of the composition and stoichiometry of the λ tail tip complex. Several different techniques were used to separate and analyze the tail tip complex. The stoichiometry of the lambda tail tip complex was estimated by densitometry.

(3) The third part is about the characterization of the cysteine residues in λ tail tip protein, gpL. According to the multisequence alignment, four of the cysteine residues (cysteine residues at positions 173, 182, 205 and 212) are highly conserved among lambda gpL homologs while the other four (cysteine residues at positions 12, 39, 140 and 208) are not. Each of the individual cysteine residues was mutated to serine or other residues. The results from the *in vivo* complementation assay show that the single amino acid substitutions at the nonconserved cysteines have little effect on the biological activity of the protein while the single amino acid substitutions at the conserved cysteine residues abolish the *in vivo* activity of the protein. This implies that the conserved cysteines may be important in protein folding or tail assembly or for functioning of the assembled tail. A series of assays were carried out to further investigate the role of the conserved cysteine residues of gpL.

(4) The last part is concerned with λ tail tip protein, gpK. gpK is a mystery because it is required for tail assembly, but it has not been found in the mature virion. Two different amber mutations were constructed into gene *K*. One is at the beginning of the gene while the other is near the end of the gene. The *in vivo* complementation assay indicates that both of them have little complementation activity. These mutants were further characterized by *in vitro* assays.

2.0 MATERIALS AND METHODS

2.1 BACTERIAL STRAINS

(1) BL21(DE3)-Δtail

F^- *dcm ompT hsdS*$_B$(r_B^- m_B^-) *gal* λ(DE3)-Δtail. This strain contains λ prophage, DE3, with the tail portion of DE3 being knocked out by deletion. It was used as the host strain for most protein expression experiments and spot complementation assays.

Source: Hendrix lab collection; constructed by J. Xu (Xu, 2001).

(2) BL21 Star^{TM}(DE3)

F^- *dcm ompT hsdS*$_B$(r_B^- m_B^-) *gal rne131* λ(DE3). This strain carries a mutation in the gene encoding RNaseE (*rne131*) and was used as the host strain for pLC3 based plasmids.

Source: Invitrogen.

(3) Rosetta^{TM} 2

F^- *dcm ompT hsdS*$_B$(r_B^- m_B^-) *gal* pRARE2 (Cam^R). This strain provides tRNAs for 7 rare codons (AGA, AGG, AUA, CUA, GGA, CCC, and CGG) and enhances the expression of proteins which contain codons rarely used in *E. coli*.

Source: Novagen.

(4) XL1-Blue

recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F' *proAB lacI*^{q}*Z*△*M15* Tn10 (Tet^{r})] was used as the host strain for the nicked dsDNA from site-directed mutagenesis.

Source: Stratagene.

(5) DH10B

F^{-} *mcrA* Δ(*mrr-hsdRMS-mcrBC*) 80d*lacZ*ΔM15 Δ*lacX74 deoR recA1 araD139* Δ(*ara leu*)7697 *galU galK rpsL endA1 nupG*.

Source: Hendrix lab collection.

(6) Ymel

supF (suppressor strain) was used in the plaque assay and the phage plate stock experiments. Besides, it was also used in the experiment designed to precipitate lambda tail and its assembly intermediates.

Source: Hendrix lab collection.

(7) C600

supE (suppressor strain) was used to select Sam7 revertants.

Source: Hendrix lab collection.

(8) CR63

F+, *LAM*-, *serU60*(AS), *lamB63* was used as the negative control for the experiment using *E. coli* cells to precipitate lambda tail and its assembly intermediates.

Source: Hendrix lab collection.

(9) Lysogens for making phage stocks

594 (λSam7cI857)

C600 (λJam27Sam7cI857)

C600 (λIam838Sam7cI857)

C600 (λLam756Sam7cI857)

C600 (λKam768Sam7cI857)

Source: Hendrix lab collection.

(10) Lysogen used for λ head preparation

594 (λVam750Sam7cI857)

Source: Hendrix lab collection.

(11) Lysogens for making concentrated tail-defective lysates

594 (λJam442Sam7cI857)

594 (λIam838Sam7cI857)

594 (λLam756Sam7cI857)

594 (λKam768Sam7cI857)

Source: Hendrix lab collection.

2.2 MEDIA AND BUFFERS

2.2.1 Media

(1) Luria Broth (LB)

1% (w/v) Bacto tryptone, 0.5% (w/v) Bacto yeast extract and 0.5% NaCl in ddH_2O. Antibiotics were added at 50 μg/ml for ampicillin, 25 μg/ml for chloramphenicol and 25 μg/ml for kanamycin when needed. 0.4% (w/v) maltose or glucose was added to induce or inhibit phage receptor expression in the host when necessary.

(2) LB agar

Same ingredients as LB plus 1.5% (w/v) Bacto agar.

(3) Soft agar

Same ingredients as LB plus 0.7% (w/v) Bacto agar.

(4) 10X M9A medium (use 1X) (Katsura and Kuhl, 1975)

10.54% $Na_2HPO_4 \cdot 7H_2O$, 3% KH_2PO4, 0.5% NaCl, 1% NH_4Cl, 4% glucose, 0.25% $MgSO_4 \cdot 7H_2O$, 0.015% $CaCl_2 \cdot 2H_2O$ and 10% Casamino acids (Difco) in ddH_2O.

(5) M9-MC medium

1X M9 salts, 0.4% (w/v) glucose, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$ and 1% MC medium (same ingredients as Difco methionine assay medium except that no cysteine was added in the MC

medium) in ddH_2O. Appropriate antibiotics were added to 50 μg/ml for ampicillin, 25 μg/ml for chloramphenicol and 25 μg/ml for kanamycin when needed.

(6) 10X M9 salts (use 1X)

13.3% $Na_2HPO_4 \cdot 7H_2O$, 3% KH_2PO4, 0.5% NaCl, 1% NH_4Cl in ddH_2O.

(7) KG

1X M9 salts, 5 mM $MgSO_4$, 1.5% casamino acids, 0.2% glucose in ddH_2O.

2.2.2 Buffers

(1) λ dil

10 mM Tris-HCl pH 7.4, 10 mM $MgCl_2$

(2) TKG (Duda et al., 1995)

20 mM Tris-HCl pH 7.5, 100 mM potassium glutamate

(3) Lysis buffer

50 mM Tris-HCl pH8.0, 5 mM EDTA

(4) Buffer A

0.2 M Tris-HCl pH 8.8, 40 μg/ml DNase I

(5) T1/10E

10 mM Tris-HCl pH 8.0, 0.1 mM EDTA

(6) Head buffer

50 mM Tris-HCl pH 8.0, 10 mM $MgSO_4$, 20 mM putrescene

(7) Gel electrophoresis buffer

50XTAE

1 M Tris-acetate, 50 mM EDTA pH 8.0

4X Lower buffer for SDS-PAGE

1.5 M Tris-HCl pH 8.8, 0.4% SDS

4X Upper buffer for SDS-PAGE

0.5 M Tris-HCl pH 6.8, 0.4% SDS

10X Running buffer for SDS-PAGE

0.25 M Tris base, 2.5 M glycine, 1% SDS

4X Sample buffer for SDS-PAGE

0.25 M Tris-HCl pH 6.8, 40% glycerol, 20% β-mercaptoethanol, 8% SDS

(8) Nickel affinity column buffers (Novagen with modifications)

1X Charging Buffer

50 mM $NiSO_4$

1X Binding Buffer

20 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl pH 7.9

1X Wash Buffer

80 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl pH 7.9

1X Elute Buffer

1 M imidazole, 0.5 M NaCl, 20 mM Tris-HCl pH 7.9

1X Strip Buffer

100 mM EDTA, 0.5 M NaCl, 20 mM Tris-HCl pH 7.9

(9) Amylose affinity column buffers (BioLabs)

1X Column buffer

20 mM Tris-HCl pH 7.4, 200 mM NaCl, 1 mM EDTA, 1 mM azide

1X Elution buffer

1X Column buffer plus 10 mM maltose

20X TEV buffer (use 1X)

1 M Tris-HCl pH 8.0, 10 mM EDTA

(10) Glutathione affinity column buffers (GE Healthcare)

Buffer B

50 mM Tris-HCl pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.1% Triton-X100, 1 mM DTT

1X PBS

140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4

1X Elution buffer

50 mM Tris-HCl pH 8.0, 10 mM reduced Glutathione

1X Cleavage buffer for PreScission Protease

50 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, pH 7.0 at 25℃, chill before use

(11) Hydroxyapatite column buffers

1X Wash buffer

20 mM sodium phosphate buffer pH 7.2, 200 mM NaCl

1X Elution buffer

0.5 M sodium phosphate buffer pH 7.2

(12) Co-immunoprecipitation buffer

Co-IP buffer

50 mM Tris-HCl pH 8.0, 10 mM $MgCl_2$, 100 mM NaCl

2.3 METHODS

2.3.1 DNA manipulation

(1) Plasmid preparation

The plasmids were purified using QIAGEN® plasmid miniprep kit with modifications. DH10B carrying the target plasmid was grown in LB with appropriate antibiotics at 37℃ overnight. About 3 ml of the bacteria cells were pelleted and resuspended in 250 μl Buffer P1. After addition of 250 μl Buffer P2 to lyse the cells, 350 μl Buffer P3 was added and mixed thoroughly by inverting the tubes 4-6 times. The mixture was centrifuged for 10 minutes at 13,000 RPM. The supernatant was applied to the QIAprep spin column by pipetting. After centrifugation at 13,000 RPM for 1 minute, the sample in the flow-through was discarded. 500 μl Buffer PB was added to the column followed by centrifugation at 13,000 RPM for 1 minute. The column was washed by 750 μl Buffer PE and centrifuged for 1 minute. The flow-through was discarded and the sample was centrifuged again at 13,000 RPM for 2-4 minutes to get rid of residual PE. Finally, 50-100 μl T1/10E buffer was added and the sample was held at room temperature for about 5 minutes before it was centrifuged at 13,000 RPM for 1 minute. Table 2.1 lists all the plasmids and construction methods used in this study.

Table 2.1 Plasmids used in this study

Name	Description	Source*
pGEX-6P-1	This vector contains an ampicillin resistance gene, an internal *lacI*^{q} gene and a *tac* promoter. It can be used to construct N-terminal GST fusion protein.	3
pGEX-6P-1-J	PCR fragment of J1R and J4F was digested by BamHI and BsrGI. The resulting fragment was inserted into the BamHI-BsrGI sites of pGEX-6P-1-Ltail.	1
pGEX-6P-1-I	PCR fragment of I5F and I6R was digested by BamHI and EcoRI. The resulting fragment was inserted into the BamHI-EcoRI sites of pGEX-6P-1.	1
pGEX-6P-1-K	PCR fragment of K5F and K6R was digested by BamHI and EcoRI The resulting fragment was inserted into the BamHI-EcoRI sites of pGEX-6P-1.	1
pGEX-6P-1-L	PCR fragment of L9F and L10R was digested by BamHI and EcoRI. The resulting fragment was inserted into the BamHI-EcoRI sites of pGEX-6P-1.	1
pGEX-6P-1-JI	PCR fragment of J1R and I5F was digested by BamHI and BsrGI. The resulting fragment was inserted into the BamHI-BsrGI sites of pGEX-6P-Ltail.	1
pGEX-6P-1-IK	PCR fragment of K5F and I6R was digested by BamHI and EcoRI. The resulting fragment was inserted into the BamHI-EcoRI sites of pGEX-6P-1.	1
pGEX-6P-1-IKL	PCR fragment of I6R and L9F was digested by BamHI and EcoRI. The resulting fragment was inserted into the BamHI-EcoRI sites of pGEX-6P-1 vector	1
pGEX-6P-1-Ltail (pGLtail)	The RsrII-BamHI fragment of pGEX-6P-1-L was inserted into the RsrII-BamHI sites of pGEX-6P-1-Mtail are digested by RsrII and BamHI.	1
pGEX-6P-1-Iter	PCR fragment using I5F and I7R as primers was digested by BamHI and EcoRI. The resulting fragment was cloned into the BamHI-EcoRI sites of pGEX-6P-1. A terminator was introduced after the stop codon of gene *I*.	1
pGEX-6p-1-L1S (pGL1)	QuikChange site-directed mutagenesis using L11F and L11R as primers and pGEX-6P-1-L as template DNA, was performed to construct C12S mutation in gene *L*.	1
pGEX-6p-1-L2S (pGL2)	QuikChange site-directed mutagenesis using L12F and L12R as primers and pGEX-6P-1-L as template DNA, was performed to construct C39S mutation in gene *L*.	1

pGEX-6p-1-L3S (pGL3)	QuikChange site-directed mutagenesis using L13F and L13R as primers and pGEX-6P-1-L as template DNA, was performed to construct C140S mutation in gene *L*.	1
pGEX-6p-1-L4S (pGL4)	QuikChange site-directed mutagenesis using L14F and L14R as primers and pGEX-6P-1-L as template DNA, was performed to construct C173S mutation in gene *L*.	1
pGEX-6p-1-L5S (pGL5)	QuikChange site-directed mutagenesis using L15F and L15R as primers and pGEX-6P-1-L as template DNA, was performed to construct C182S mutation in gene *L*.	1
pGEX-6p-1-L6S (pGL6)	QuikChange site-directed mutagenesis using L17F and L17R as primers and pGEX-6P-1-L as template DNA, was performed to construct C205S mutation in gene *L*.	1
pGEX-6p-1-L7S (pGL7)	QuikChange site-directed mutagenesis using L18F and L18R as primers and pGEX-6P-1-L as template DNA, was performed to construct C208S mutation in gene *L*.	1
pGEX-6p-1-L8S (pGL8)	QuikChange site-directed mutagenesis using L19F and L19R as primers and pGEX-6P-1-L as template DNA, was performed to construct C212S mutation in gene *L*.	1
pGEX-6p-1-L3M (pGL3M)	QuikChange site-directed mutagenesis using L20F and L20R as primers and pGEX-6P-1-L as template DNA, was performed to construct C140M mutation in gene *L*.	1
pGEX-6p-1-L12 (pGL12)	QuikChange site-directed mutagenesis using L11F and L11R as primers and pGL2 as template DNA, was performed to construct C12S mutation in gene *L*.	1
pGEX-6p-1-L127 (pGL127)	QuikChange site-directed mutagenesis using L18F and L18R as primers and pGL12 as template DNA, was performed to construct C208S mutation in gene *L*.	1
pGEX-6p-1-L4C (pGL4C)	QuikChange site-directed mutagenesis using L20F and L20R as primers and pGL127 as template DNA, was performed to construct C140M mutation in gene *L*.	1
pET21+	This vector contains an ampicillin resistance gene, a *lacI* gene and a T7 ϕ10 promoter adjacent to a Lac operator.	4
pET21-Ltail (pLtail)	pEtail4 was digested by EarI, filled in by Klenow fragment and then digested by XhoI. The resulting fragment containing λ tail genes *L-J* was cloned into HindIII filled in, XhoI digestion of pET21+.	2
pIam	PCR fragment using J1R / L1F as outer primers and IamF / IamR as inside primers was digested by BsrGI and RsrII. The resulting fragment was inserted into the BsrGI-RsrII sites of pLtail.	1

pKam6	PCR fragment using J1R / L1F as outer primers and KamF / KamR as inside primers was digested by BsrGI and RsrII. The resulting fragment was inserted into the BsrGI-RsrII sites of pLtail.	1
pLdel	pLtail was digested by RsrII and SacI. The resulting fragment containing λ tail genes *K-J* was filled in by Klenow and ligated.	1
pLtail-m1	PCR fragment using XD1F / XD2R as outer primers and XD17F / XD18R as inside primers was digested by BstBI and AvaI. The resulting fragment was inserted into the BstBI-AvaI sites of pLtail.	1
pLtail-m2	PCR fragment using XD1F / XD2R as outer primers and XD15F / XD16R as inside primers was digested by BstBI and AvaI. The resulting fragment was inserted into the BstBI-AvaI sites of pLtail.	1
pLtail-m3	PCR fragment using XD1F / XD2R as outer primers and XD13F / XD14R as inside primers was digested by BstBI and AvaI. The resulting fragment was inserted into the BstBI-AvaI sites of pLtail.	1
pLtail-m4	PCR fragment using XD9F / XD10R as outer primers and XD19F / XD20R as inside primers was digested by BstBI and AgeI. The resulting fragment was inserted into the BstBI-AgeI sites of pLtail.	1
pGtail	The SacI-NsiI fragment of pT7-5-GT was inserted into the SacI-NsiI sites of pETail4.	2
pGtail-LT7	PCR fragment using L2R / H1F as outside primers and L21F / L21R as inside primers was digested by StuI and RsrII. The resulting fragment was cloned into the StuI-RsrII sites of pGtail.	1
pGtail-IT7	PCR fragment using L1F / J1R as outside primers and I8F / I9R as inside primers was digested by BsrGI and RsrII. The resulting fragment was cloned into the BsrGI-RsrII sites of pGtail.	1
pGtail-4S	PCR fragment using H1F / L26R as outside primers and L25F / L25R as inside primers was digested by StuI and RsrII. The resulting fragment was inserted into the StuI-RsrII sites of pGtail.	1
pGtail-5S	PCR fragment using L23R and H1F as primers was digested by StuI and RsrII. The resulting fragment was inserted into the StuI-RsrII sites of pGtail.	1

pGtail-6S	PCR fragment using L1F / J1R as outside primers and L24F / L24R as inside primers was digested by BsrGI and RsrII. The resulting fragment was inserted into the BsrGI-RsrII sites of pGtail.	1
pGtail-8A	The BsrGI-RsrII fragment of pEtail4-8A was inserted into the BsrGI-RsrII sites of pGtail.	1
pGtail-8H	The BsrGI-RsrII fragment of pEtail4-8H was inserted into the BsrGI-RsrII sites of pGtail.	1
pGtail-8S	PCR fragment using L1F / J1R as outside primers and L19F / L19R as inside primers was digested by BsrGI and RsrII. The resulting fragment was inserted into the BsrGI-RsrII sites of pGtail.	1
pEtail4	All 11 genes involved in λ tail assembly were cloned into pET21+ vector, resulting in pEtail.	2
pEtail4-Kam6	The BrsGI-RsrII fragment of pKam6 was inserted into the BsrGI-RsrII sites of pEtail4.	1
pEtail4-Kam768	PCR fragment using J1R / L1F as outer primers and K768F / K768R as inside primers was digested by BsrGI and RsrII. The resulting fragment was inserted into the BsrGI-RsrII sites of pEtail4.	1
pEtail4-Lam	PCR fragment using L26R / H1F as outside primers and L27F / L27R as inside primers was digested by StuI and RsrII. The resulting fragment was inserted into the StuI-RsrII sites of pEtail4.	1
pEtail4-4S	The BsrGI-StuI fragment of pGtail-4S was inserted into the BsrGI-StuI sites of pEtail4.	1
pEtail4-5S	The BsrGI-StuI fragment of pGtail-5S was inserted into the BsrGI-StuI sites of pEtail4.	1
pEtail4-6S	The BsrGI-RsrII fragment of pGtail-6S was inserted into the BsrGI-RsrII sites of pEtail4.	1
pEtail4-7S	PCR fragment using L1F / J1R as outside primers and L17F / L17R as inside primers was digested by BsrGI and RsrII. The resulting fragment was inserted into the BsrGI-RsrII sites of pEtail4.	1
pEtail4-8S	The BsrGI-RsrII fragment of pGtail-8S was inserted into the BsrGI-RsrII sites of pEtail4.	1
pEtail4-8A	PCR fragment using L1F / J1R as outside primers and L28F / L28R as inside primers was digested by BsrGI and RsrII. The resulting fragment was inserted into the BsrGI-RsrII sites of pEtail4.	1

pEtail4-8H	PCR fragment using L1F / J1R as outside primers and L29F / L29R as inside primers was digested by BsrGI and RsrII. The resulting fragment was inserted into the BsrGI-RsrII sites of pEtail4.	1
pLC3	This vector contains a kanamycin resistance gene, a *lacI* gene and a T7 promoter. It can be used to construct N-terminal MBP fusion protein.	5
pLC3-I	PCR fragment using I3F and I4R as primers was digested by NdeI and BamHI. The resulting fragment was inserted into the NdeI-BamHI sites of pLC3.	1
pLC3-K	PCR fragment using K3F and K4R as primers was digested by NdeI and BamHI. The resulting fragment was inserted into the NdeI-BamHI sites of pLC3.	1
pLC3-L	PCR fragment using L7F and L8R as primers was digested by NdeI and BamHI. The resulting fragment was inserted into the NdeI-BamHI sites of pLC3.	1
pTQ30	This vector contains an ampicillin resistance gene and a T7 promoter. It can be used to construct 6XHis tag to the N-terminus of the target proteins.	2
pTQ30-J	PCR fragment using J3F and J1R as primers was digested by BsrGI and SacI. The resulting fragment was inserted into the BsrGI-SacI sites of pTQ30-G-his-tail.	1
pTQ30-I	PCR fragment using I1F and I2R as primers was digested by HindIII and SacI. The resulting fragment was inserted into the HindIII-SacI sites of pTQ30.	1
pTQ30-K	PCR fragment using K1F and K2R as primers was digested by HindIII and SacI. The resulting fragment was inserted into the HindIII -SacI sites of pTQ30.	1
pTQ30-L	PCR fragment using L5F and L4R as primers was digested by HindIII and SacI. The resulting fragment was inserted into the HindIII-SacI sites of pTQ30.	1
pTQ30-JI	PCR fragment using J1R and I1F as primers was digested by BsrGI and SacI. The resulting fragment was inserted into the BsrGI-SacI sites of pTQ30-G-his-tail.	1
pTQ30-Ltail	PCR fragment using L5F and L2R as primers was digested by RsrII and SacI. The resulting fragment was inserted into the RsrII-SacI sites of pTQ30-G-his-tail.	1

*Source:

1. Constructed for this study
2. Constructed by Jun Xu (Xu, 2001)
3. GE Healthcare
4. Novagen
5. From Dr. Graham Hatfull lab (University of Pittsburgh)

(2) Restriction digestion

The digestion reaction was carried out as described in the enzyme manufactures (New England BioLab).

(3) Ligation

The ligation reactions were carried out in a 10 μl volume, including 1 U ligase, 1X ligation buffer, appropriate amounts of vectors and inserts, and ddH_2O if needed. The reactions were incubated at 14℃ overnight.

(4) Transformation

1 μl ligation reaction or 1 μl of 1/100-1/10 diluted purified plasmid was mixed with 40 μl competent cells. After electroporation on a BioRAD GenePulser machine, the sample was resuspended in 1 ml LB, incubated at 37℃ for 40-60 minutes and plated on LB agar with appropriate antibiotics.

(5) DNA gel electrophoresis

DNA gel electrophoresis was performed as described with 0.8-1.2% agarose in 1X TAE buffer (Sambrook, 1989). The DNA was visualized by ethidium bromide staining.

(6) DNA purification

The PCR products were purified following the protocol described in QIAprep miniprep kit (QIAGEN). To purify DNA fragments from digestion reactions, the samples were first separated

on an agarose gel. Bands of interest were excised from the gel and the DNA fragments were purified by pressure-extrusion method (Gubin and Kincaid, 1998).

(7) Sequencing

The DNA samples were prepared as described and sent to GENEWIZ for sequencing. The results were analyzed by Sequencher.

2.3.2 Polymerase chain reaction (PCR)

(1) Regular PCR

PCR reactions were performed in a 50 μl volume, including 1XPFU buffer, 0.5 μM of each primer, 0.2 mM of each dNTP, 5 μl of the template (for plasmid templates, the plasmid stocks were diluted 50-100 fold), and 2 units of PFU polymerase. The cycle was performed at 94℃ for 50 seconds, 49℃ (the annealing temperature was adjusted in each experiment) for 50 seconds and 72℃ for 2.5 minutes for 20 cycles on a 9600 or 9700 GeneAmp PCR machine (Perkin-Elmer).

(2) Recombinant PCR

Four primers (two "outside" and two "inside" primers) were needed for the recombinant PCR. Figure 2.1 illustrates the strategy of the recombinant PCR. The products from the first-round PCR were purified by QIAprep miniprep kit. The purified products were used as the template in the second-round PCR.

Figure 2.1 Overview of the recombinant PCR ((Russell, 1990)with modifications).

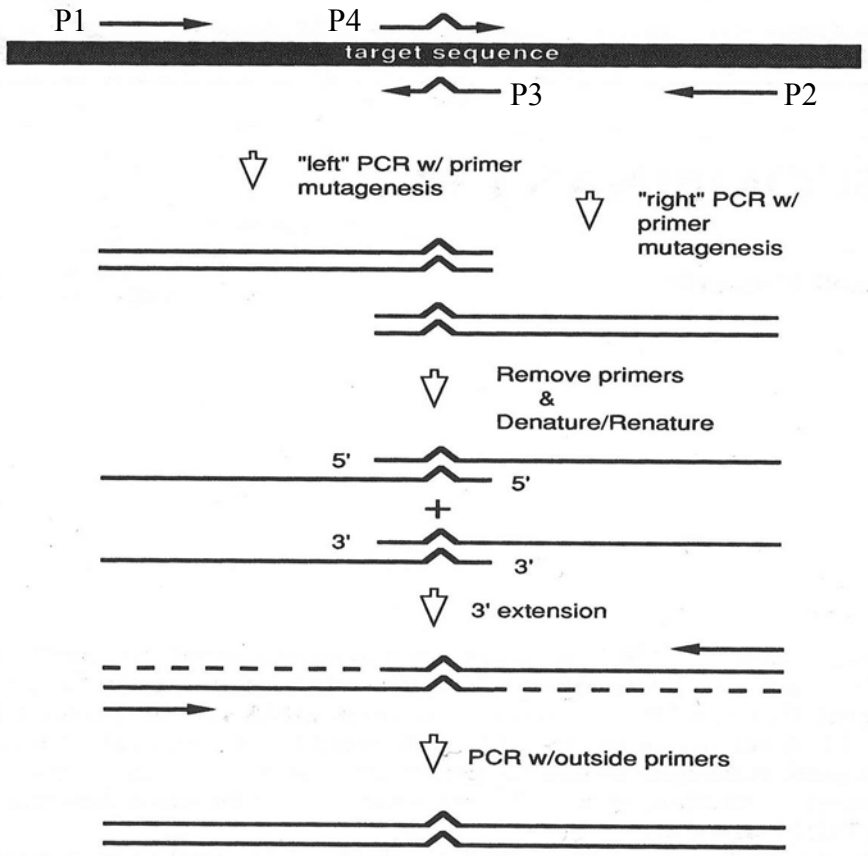
The recombinant PCR is a two-step PCR, which requires two "outside" primers (P1 / P2) and two "inside" primers (P3 / P4). It combines the two separate PCR products with overlapping sequence into one longer product. This technique can be used to make a base substitution, deletion or a small insertion.

(a) The basic steps for making a base substitution by recombinant PCR is shown. Both of the inside primers, P3 and P4, contain the same mutation. Two separate PCR reactions are performed in the first-round PCR, using P1 / P3, and P2 / P4 as primers, respectively. The first-round PCR products are purified by QIAprep miniprep kit. The overlapping sequence in the first-round PCR products can be denatured and allowed to reanneal together to form a whole fragment with the target mutation. This long fragment can be amplified by P1 and P2 in the second-round PCR.

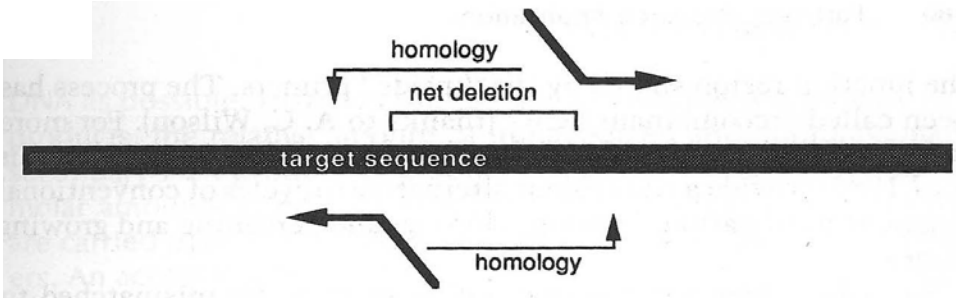
(b) It shows inside primers for the creation of deletion. The remaining steps are the same as those for making a base substitution by recombinant PCR.

(c) It shows inside primers for the creation of small insertion. The remaining steps are the same as those for making a base substitution by recombinant PCR.

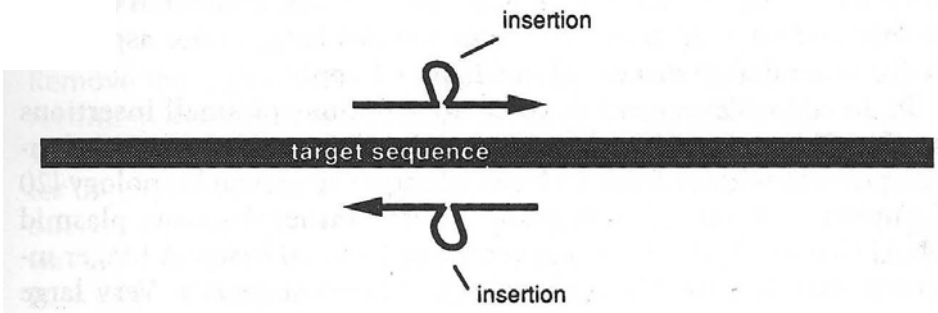
(a)


P1
P4
target sequence
P3
P2
"left" PCR w/ primer mutagenesis
"right" PCR w/ primer mutagenesis
Remove primers & Denature/Renature
5'
5'
3'
3'
3' extension
PCR w/outside primers

(b)


homology
net deletion
target sequence
homology

(c)


insertion
target sequence
insertion

(3) QuikChange site-directed mutagenesis

The site-directed mutagenesis was performed following the QuikChange site-directed mutagenesis kit (Stratagene). The basic procedure for this experiment is shown in Figure 2.2.

(4) Primers

All the primers used in this study are listed in Table 2.2.

2.3.3 Phage preparations

(1) Heat induction of lysogens

The lysogens were grown at 30℃ overnight in LB medium. The overnight culture was 1:100 diluted into LB medium and grown at 37℃with vigorous shaking until it reached a density of about $3X10^8$ cell/ml. The culture was then heat-induced at 43℃ for 18 minutes in a water bath. Subsequently, the culture was transferred to 37℃ for 3 hours. Cells were chilled and harvested by centrifugation at 8,000 RPM for 5 minutes. The pellet was resuspended in 1 ml λ dil buffer. Several drops of chloroform were added to lyse the cells. The pellet was removed by centrifugation at 8,000 RPM for 10 minutes. The supernatant was kept at 4℃. This method applies to prepare λSam7cI857 phage stock as well as some mutant phages stocks which are prepared from the induction of the appropriate C600 lysogens.

(2) Revertant selection

Phages with Sam7 mutation can grow on Ymel, but not on C600 or other *E. coli* strains (such as BL21(DE3)-Δtail) bearing the appropriate plasmid. In order to get rid of the Sam7 mutation,

Figure 2.2 Strategy for QuikChange site-directed mutagenesis (adapted from Stratagene's QuikChange site-directed mutagenesis kit: instruction manual).

The QuikChange site-directed mutagenesis kit can be used to generate point mutations, switch amino acids, delete or insert single or multiple amino acids in double-stranded plasmids. The basic procedure requires a supercoiled dsDNA plasmid with an insert of interest and two primers with the desired mutation. The plasmid is amplified by the oligonucleotide primers which anneal to the same sequence on opposite strands of the plasmid. This results in the nicked, circular strands. Subsequently, the parental DNA template is digested with DpnI and the circular, nicked dsDNA is transformed into XL1-Blue competent cells which can repair the nicks in the mutated plasmid. The mutated plasmid is extract as described above.

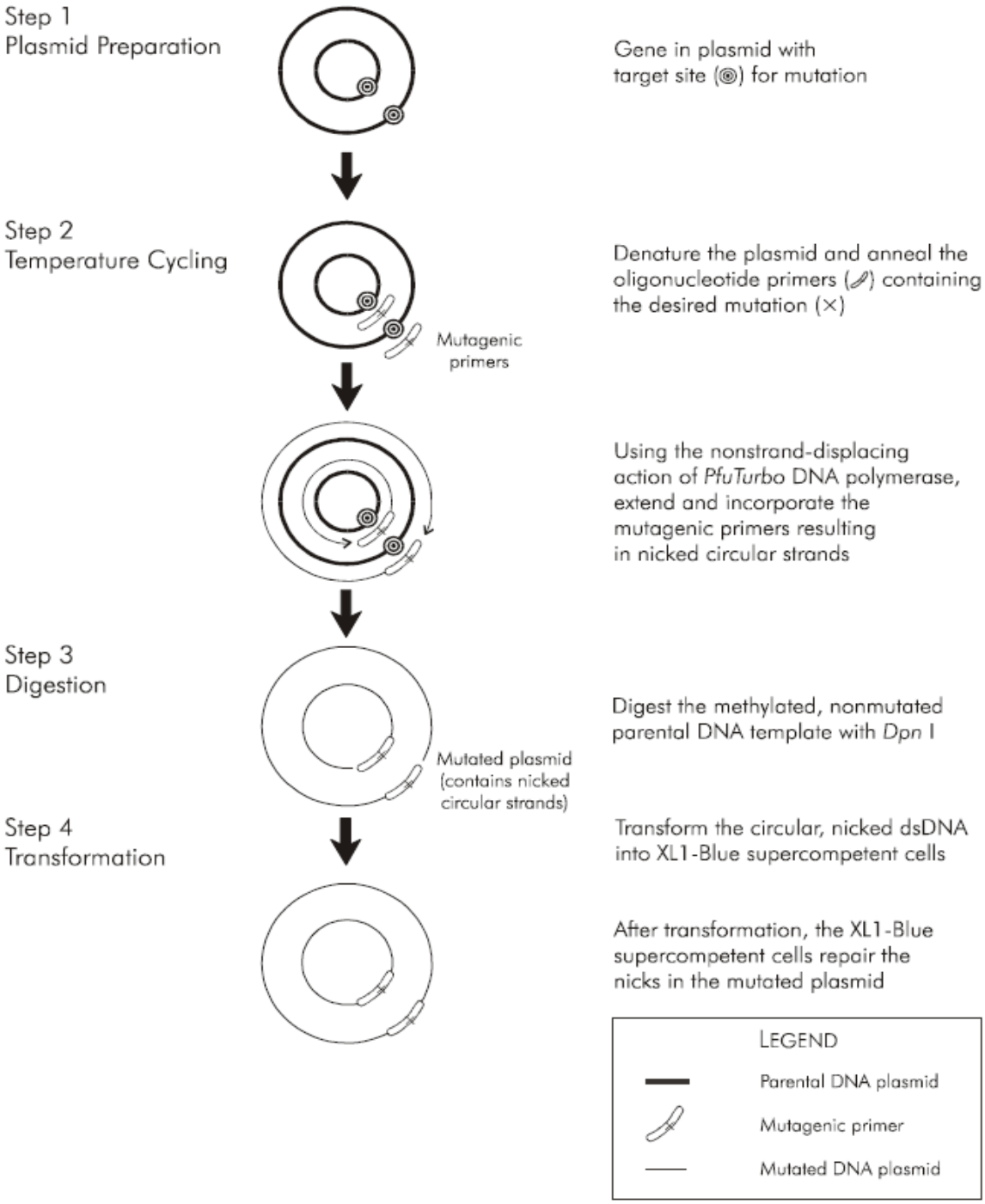

Step 1
Plasmid Preparation
Gene in plasmid with
target site (◎) for mutation
Step 2
Temperature Cycling
Mutagenic
primers
Denature the plasmid and anneal the
oligonucleotide primers (⁄) containing
the desired mutation (×)
Using the nonstrand-displacing
action of PfuTurbo DNA polymerase,
extend and incorporate the
mutagenic primers resulting
in nicked circular strands
Step 3
Digestion
Mutated plasmid
(contains nicked
circular strands)
Digest the methylated, nonmutated
parental DNA template with Dpn I
Step 4
Transformation
Transform the circular, nicked dsDNA
into XL1-Blue supercompetent cells
After transformation, the XL1-Blue
supercompetent cells repair the
nicks in the mutated plasmid
LEGEND
Parental DNA plasmid
Mutagenic primer
Mutated DNA plasmid

Table 2.2 Primers used in this study

Primer Name	Sequence
J1R	ccgttacgttgtatctgaac
J3F	gagctcatgggtaaaggaagcagtaa
J4F	ggatccatgggtaaaggaagcagt
CD1F	cggtggcagcagccaactcag
XD2R	cgaatcccatctcggcaagg
XD9F	ttggtctgctcaattttgac
XD10R	gaagggcttcgccatgtacc
XD13F	gtgatggtgatggtgatgcgggtcaataaatgcgatacga
XD14R	catcaccatcaccatcacccgatgtttgtggcgcaggg
XD15F	gtgatggtgatggtgatgcacactgctttcacgctggcg
XD16R	catcaccatcaccatcaccgtactgtcaccgtgaccga
XD17F	gtgatggtgatggtgatgtccgcgaaacgtcagcgga
XD18R	catcaccatcaccatcacggcaggacaacgtattcgatg
XD19F	gtgatggtgatggtgatggacttttttccagcagctccttg
XD20R	catcaccatcaccatcacgagctgacggaggataacgc
IamF	gtggcgtattctagattgtcc
IamR	ggacaatctagaatacgccac
I1F	gagctcatggcagcgacgcacaca
I2R	aagctttcagcgaccaatcaccaca
I3F	catatggcagcgacgcacacactcc
I4R	ggatcctcagcgaccaatcaccaca
I5F	ggatccatggcagcgacgcacaca
I6R	gaattctcagcgaccaatcaccac
I7R	gaattctcctcacgctccataaat
I8F	atggctagcatgactggtggacagcaaatgggttgatgcaaaatgttttatgtg
I9R	acccatttgctgtccaccagtcatgctagccatgcgaccaatcaccacaacctg
L1F	tttccgggacgtatcatgct
L2R	tgctgcatttatccttcgtg

L4R	aagcttttactgcgaaagtttgtt
L5F	gagctcatgcaggatatccggcag
L7F	catatgcaggatatccggcaggaa
L8R	ggatccttactgcgaaagtttgtt
L9F	ggatccatgcaggatatccggcag
L10R	gaattcttactgcgaaagtttgtt
L11F	atgaaagcacacgtgcggagcagtc
L11R	gcacgtgtgctttcattcagtgtttc
L12F	tttctctaatgagcagaacgaaaaagg
L12R	tctgctcattagagaaaaaataacg
L13F	tgagcagagcagcgaattgagcgcgg
L13R	ttcgctgctctgctcaatgcgccag
L14F	tggcaaacaccagcacctggacctatc
L14R	ggtgctggtgtttgccagcatgatac
L15F	acgagagcggttatagcggcccggctg
L15R	taaccgctctcgtcaccgcgatagg
L17F	ataaaagcagcaaatgcttgagcgg
L17R	tgctgcttttatccttcgtgatatc
L18F	cagcaaaagcttgagcggttgtaag
L18R	gctcaagcttttgctgcatttatcc
L19F	agcggttcaaagttccgcaataacg
L19R	aactttgaaccgctcaagcatttgc
L20F	gagcagatgagcgaactgagcgcgg
L20R	gttcgctcatctgctcgatgcgccag
L21F	atggctagcatgactggtggacagcaaatgggtcaggatatccggcaggaaaca
L21R	acccatttgctgtccaccagtcatgctagccatcatcagttcaccacctgttc
L23R	gacagccggaccgctataaccgctctcgtc
L24F	acgaaggataaaagcagcaaatgc
L24R	gcatttgctgcttttatccttcgt
L25F	atgctggccaacaccagcacctggaccta

L25R	taggtccaggtgctggtgttggccagcat
L26R	gtttgttaatggaaaggaagcc
L27F	ggaagatatgtagactctggtcgg
L27R	ccgaccagagtctacatatcttcc
L28F	atgcctgagcggtgctaagttcc
L28R	ggaacttagcaccgctcaggcat
L29F	ctgagcggtcataagttccgca
L29R	tgcggaacttatgaccgctcag
KAMF	gaagactagttgcaggcagaa
KAMR	ttctgcctgcaactagtcttc
K768F	acaaatggtagcgacgcaca
K768R	tgtgcgtcgctaccatttgt
K1F	gagctcatgtcgccggaagactgg
K2R	aagctttcacacgaaggtcgatgc
K3F	catatgtcgccggaagactggctg
K4R	ggatcctcacacgaaggtcgatgc
K5F	ggatccatgtcgccggaagactgg
K6R	gaattctcacacgaaggtcgatgc
GST1F	gctggcaagccacgtttgg
GST1F	gctggcaagccacgtttgg
H1F	cgctgaataacgtcatgtc

mutant phages with Sam7 mutation were plated on C600 for selection of revertants. About 10^6-10^7 phages were mixed with 100 μl overnight culture of C600. The mixture was incubated at 37℃ for 15 minutes before they were plated on LB agar. After overnight incubation at 37℃, single plaques that can grow on C600 were picked and used for making phage plate stock. Phages made from the single plaques do not contain Sam7 mutation and therefore can be used in the spot complementation assay.

(3) Plate stock

Phages which do not exist as lysogens, such as the revertant phages, can be grown by plate stock method. A single plaque was picked from a fresh plate, resuspended in 1 ml λ dil buffer and incubated at 37℃ for about 1 hour. After this, 0.1 ml of the phage solution was mixed with 3 drops of freshly grown Ymel cells (in LB / Mal) and the mixture was incubated at 37℃ for 15 minutes. The mixture was then mixed with 4 ml of diluted soft agar (equal volume of soft agar and KG), poured onto a fresh tryptone agar plate, and incubated at 37℃ right-side up to confluent lysis (about 6-8 hours). The top soft agar was scraped into a sterile centrifuge tube and several drops of chloroform were added. The mixture was homogenized by vortexing and the pellet was removed by centrifugation at 8,000 RPM for 10 minutes.

(4) Plaque assay

The phage stock was made series dilution. 100 μl of the diluted phage was mixed with 100 μl of overnight bacteria cells and incubated at 37℃ for 15 minutes to allow phage adsorption. Subsequentely, the solution was mixed with 3-4 ml of melted soft agar and poured on LB agar plates. The plates were incubated at 37℃ overnight.

(5) Preparation of concentrated tail-defective lysates

The overnight culture of appropriate 594 lysogen was 10^4-10^5 diluted into M9A medium and grown at 28℃ overnight. The next morning, the growth temperature was shifted to 30℃ until they reached about $3X10^8$ cell/ml. The culture was then transferred to a 43℃ water bath for 18 min for heat induction. After 3 hours' growth at 37℃, the culture was chilled and harvested by centrifugation at 4000 g for 10 minutes. The pellet was resuspended in 1/250-1/300 of the culture volume in 0.2 M Tris-HCl (pH 8.8).

(6) λ head preparation

The head lysate was prepared from 594(λVam750Sam7cI857) lysogen in the same way as making conecentrated tail-defective lysates except that cells were grown in LB medium instead of M9A medium. After cells were harvested, they were resuspended in head buffer. Several drops of chloroform were added to lyse the cells. After lysis, 20 μg/ml DNase I was added and the mixture was incubated in a 24℃ water bath for about 5 minutes to promote DNase action. The pellet was removed by centrifugation at 8,000 RPM for 15 minutes. The supernatant was loaded onto the top of 1.5 ml of 25% sucrose and centrifuged in Ti80 at 30,000 RPM for 30 minutes at 4℃. After centrifugation, appropriate amount of head buffer was added and the solution was incubated overnight at 4℃ to allow the head to be dissolved. The undissolved pellet was removed by low speed centrifugation and the supernatant was used for the head-tail joining reaction.

2.3.4 Protein manipulation

(1) Protein expression

Cells with the target plasmid were grown in LB with appropriate antibiotics (50 μg/ml Amp and 25 μg/ml Cam for BL21(DE3)-Δtail pLysS with pET21+ / pTQ30 / pGEX-6p-1 based plasmids, 25 μg/ml Kan and 25 μg/ml Cam for BL21 Star^{TM}(DE3) pLysS with pLC3 based plasmids). The next morning, the overnight culture was inoculated into LB medium with appropriate antibiotics by 1:500-1:1000 dilution. The culture was grown at 37℃ until they reached a density of about 4X10^{8} cell/ml. IPTG (1 mM for pET21+ / pGEX-6p-1 / pLC3 based plasmids and 0.4 mM for pTQ30 based plasmids) was added to induce protein expression at 37℃ for about 2.5 hours. Alternatively, cells can grow at 30℃ until they reached a density of about 4X10^{8} cell/ml. They were induced with 0.4 mM IPTG overnight at 28℃. Cells were chilled, harvested and resuspended in an appropriate buffer. They were lysed by sonication and the pellet was removed by centrifugation at 10,000 RPM (for lysates used for amylose or glutathione affinity chromatography) or 13,000 RPM (for lysates used for nickel affinity chromatography) for 20 minutes at 4℃.

(2) Preparation of ^{35}S-Met labeled proteins

A single colony was picked from a fresh plate and resuspended in 1 ml of M9-MC medium. This solution was then made 1/100 dilution into a larger volume of M9-MC medium with appropriate antibiotics in a 25 ml flask and the culture was left standing overnight at 30℃. The next morning, the culture was transferred to a 37℃ water bath with shaking until it reached a density of about

$3X10^8$ cell/ml. The temperature was shifted to 30℃ and appropritate amount of IPTG (1 mM for pET21+ / pLC3 / pGEX-6p-1 based plasmids or 0.4 mM for pTQ30 based plasmids) was added to induce protein expression for 30 minutes. After this, 200 μg/ml rifampicin was added to shut down host gene expression for 20 minutes. ^{35}S-Met was added to 5-10 μCi/ml and the culture was grown at 30℃ for 1 hour. The cells were chilled and harvested by centrifugation at 8,000 RPM for 10 minutes. The pellet was resuspended in 1/10 culture volume of cold lysis buffer. Triton X-100 was added to a final concentration of 0.1% and the mixture was warmed to 24℃ for about 5 minutes to promote lysozyme action. Subsequently, $MgSO_4$ and DNase I were added to a final concentration of 7.5 mM and 20 μg/ml, respectively. The mixture was incubated at 24℃ for about 5 minutes to promote DNase activity. The supernatant was separated from the pellet by centrifugation at 8,000 RPM for 15 minutes at 4℃.

For radiolabeled samples used for the sucrose velocity gradient followed by the *in vitro* complementation assay, the cells were harvested by centrifugation at 4,000 g for 10 minutes at 4℃ and resuspended in 1/50-1/100 of the culture volume of Buffer A. Cells were lysed by freezing in the liquid nitrogen for 10 seconds and thawing at room temperature for about 15 minutes. The pellet was removed by centrifugation at 6,000 g for 15 minutes at 4℃. The supernatant was further analyzed by a 10-25% sucrose velocity gradient.

(3) Pulse-chase experiment

Before addition of ^{35}S-Met, the procedure for this experiment is the same as that of the preparation of ^{35}S-Met labeled proteins as described above except that the temperature was kept

at 37℃ after induction. After addition of rifampicin for 20 minutes, 2 μg/ml of ^{35}S-Met was added for 30 seconds and then 1000-fold excess of cold methionine was added. Samples were taken at 0 (right before addition of cold methionine), 2, 5, 10, 15, 20, 30, 40, 50, and 60 minutes after addition of cold methionine. Each sample was TCA precipitated and resuspended in 40 μl of 1X SDS sample buffer. The samples were boiled for 2.5 minutes before they were loaded onto SDS-PAGE.

(4) TCA precipitation of proteins

Cold TCA was added to the protein sample with a ratio of about 4:1 (TCA:protein) The mixture was vortexed immediately, incubated on ice for 10-30 minutes and then centrifuged at 12,000 RPM for 15 minutes at 4℃. The supernatant was aspirated and 0.5-1 ml of ice cold acetone was added to each pellet to get rid of residual TCA. The sample was incubated at -20℃ for more than 30 minutes before it was centrifuged at 12,000 RPM for 10 minutes at 4℃. The supernatant was aspirated and the pellet was placed in a desiccator for 15 minutes. The pellet was resuspended in 1X SDS sample buffer and boiled for 2.5 minutes.

(5) Nickel affinity chromatography

This experiment was carried out as described in Novagen protocols with buffers modified. In brief, the column was first charged with 4 bed volumes of 1X changing buffer and then equilibrated with 3 bed volumes of 1X binding buffer. Subsequently, the sample was loaded onto the column. This was followed by 10 bed volumes of 1X binding buffer, 8 bed volumes of 1X washing buffer and 3 bed volumes of 1X elute buffer.

(6) Amylose affinity chromatography

This experiment was carried out as described in the instruction manual for pMAL^{TM} protein fusion and purification system (New England BioLabs) with buffers modified. Brief, the column was equilibrated with 8 bed volumes of 1X column buffer. The sample was first diluted to reduce the protein to about 2.5 mg/ml before it was loaded onto the column. After this, the column was first washed with 12 bed volumes of 1X column buffer and then eluted with 3 bed volumes of 1X elution buffer. The eluted protein was dialyzed in 1X TEV buffer and then used for AcTEV protease digestion following the instruction (Invitrogen).

(7) Glutathione affinity chromatography

This experiment was performed following the instruction manual for Glutathione Sepharose 4B (Amersham Biosciences) with buffers modified. Briefly, the column was first equilibrated with ice cold 1XPBS. After the sample was loaded, the column was washed with 10 bed volumes of ice cold 1XPBS and then eluted with 5 bed volumes of 1X elution buffer. The eluted protein was dialyzed in 1X PreScission cleavage buffer and used for PreScission Protease cleavage following the instructions (GE Healthcare).

(8) Hydroxyapatite chromatography

The hydroxyapatite chromatography was carried out as described in the instruction manual for pMAL^{TM} protein fusion and purification system (New England BioLabs). Briefly, the column was first equilibrated with 1X wash buffer and then the sample was loaded onto the column. The column was washed with 10 bed volumes of 1X wash buffer and then eluted with 3 bed volumes of 1X elution buffer. The elution fractions were assayed for protein A_{280}.

(9) Ammonium sulfate precipitation

Appropriate amount of solid $(NH_4)_2SO_4$ was slowly added to the sample, which was kept stirring on an ice box, until the desired concentration. After this, the solution was kept stirring for about 1 hour to fully equilibrate. The solution was then centrifuged at 13,000 RPM for 20 minutes at 4℃ to pellet out protein. More solid ammonium sulfate was added to the supernatant to make the next desired concentration, and the stirring and centrifugation steps were repeated. Each pellet was resuspended in the same volume of appropriate buffer.

(10) Co-immunoprecipitation (Co-IP)

The Co-IP reaction was performed in a 200 μl reaction with 1 μl of 1:5 diluted T7-tag antibody and the appropriate amount of the sample. The mixture was incubated on a rotating wheel for 2 hours at 4℃. Protein G Sepharose 4 Fast Flow slurry was washed 3 times with Co-IP buffer before it was added to the mixture. After overnight incubation with protein G-coupled beads on a rotating wheel, each sample was washed 3 times with Co-IP buffer, resuspended in 20 μl of 1X SDS sample buffer and boiled for 5 minutes.

(11) SDS polyacrylamide gel electrophoresis

The experiment was performed as described (Laemmli, 1970). The SDS gels were run at 150 V constant voltage until the blue dye ran out of the gel. Gels were stained with coomassie blue.

(12) Densitometry

For SDS gels with radioactive samples, the gels were dried before they were exposed to Kodak BioMax MR films. For quantification purpose, the gels were exposed to the films for several

different times and the experiments were generally repeated 1-3 times. After development of the films, they were scanned and analyzed by Kodak 1D image analysis software (version 3.5) following the instruction. The net intensity of the bands of interest was correlated to the exposure time. A curve was obtained by plotting the net intensity of the target bands against the exposure time. If a number was significantly deviated from the curve, it was thought to be inaccurate and was not used for further analysis. For the estimation of the stoichiometries of the tail tip proteins, the copy numbers of the tail tip proteins were corrected for the number of sulfur atoms in the proteins and they were normalized to either 3 copies of gpJ or 6 copies of gpU.

2.3.5 Complementation

(1) Head-tail joining reaction

The head-tail joining reaction was carried out as described (Weigle, 1966). 10 μl of tail and 10 μl of head were mixed with 80 μl of λ dil buffer. The mixture was incubated at room temperature for 30-60 minutes before they were titered for phage production.

(2) *In vitro* complementation assay

The *In vitro* complementation assay was performed as described (Katsura and Kuhl, 1975). 15 μl of each fraction from the sucrose gradient was mixed with 10 μl of concentrated tail-defective lysates. The mixtures were then frozen in liquid nitrogen for 10 seconds and then incubated at room temperature overnight. Subsequently, 0.1 ml of λ dil buffer containing 10 μg/ml DNase I was used to suspend the mixtures. The mixtures were diluted by λ dil buffer and titered for phage production.

(3) *In vivo* spot complementation assay

Cells containing the target plasmid were grown in LB with appropriate antibiotics at 37℃ overnight. 150 μl of the overnight cultures were plated on LB agar plates. 4 μl of 10-fold diluted phages (10^4-10^8 PFU/ml) were spotted on the lawn of bacteria. The plates were incubated at 37℃ overnight.

2.3.6 Separation of λ tail / tail tip assembly intermediates

(1) Sucrose velocity gradient

The sucrose velocity gradient was performed as described (Katsura and Kuhl, 1975). The sample was loaded onto a 10-25% linear sucrose gradient in 0.2 M Tris-HCl buffer (pH 7.8) and centrifuged in an SW60Ti rotor at 58,000 RPM for 3.5 hours at 4℃. After this, the centrifuge tubes were punctured at the bottom and the gradients were collected into 12-13 fractions.

(2) Precipitation of λ tail / tail tip assembly intermediates by *E. coli* cells

The experiment was performed as described in (Tsui and Hendrix, 1983). Briefly, the radiolabeled tail lysate was incubated with excess amounts of *E. coli* cells at 4℃ for 1 hour or overnight to allow the binding of gpJ to LamB receptors. Subsequently, the mixture was washed 3 times with λ dil buffer by centrifugation at 6,000 RPM for 10 minutes to remove the unbound proteins. The bound proteins were resuspended in 1X SDS sample buffer and boiled for 2.5 minutes.

(3) Glycerol gradient to separate λ tail

The radiolabeled tail lysate was loaded onto a 10-30% glycerol gradient in TKG buffer and centrifuged in an SW41 rotor at 40,000 RPM for 3.5 hours at 4℃. After centrifugation, a whole was punctured at the bottom of the gradient tube and the gradient was collected into 12-13 fractions.

2.4 SEPARATION OF λ PHAGE BY CESIUM CHLORIDE STEP GRADIENT

The phage solution was layered on 4.5 ml of the cesium chloride gradient, including 1 ml of 10% glycerol in λ dil buffer, 2 ml of 1.4 g/cm^3 CsCl and 1.5 ml of 1.6 g/cm^3 CsCl. After centrifugation in an SW41 rotor at 30,000 RPM for 90 minutes at 20℃, a whole was punctured at the bottom of the tube and about 1/3 from the bottom of the gradient was collected.

3.0 EXPRESSION AND PURIFICATION OF λ TAIL TIP PROTEINS

3.1 INTRODUCTION

The tail tip of phage λ, located at the distal end of the tail, is composed of a thin, central tail fiber and a conical part (Katsura, 1983). It recognizes host receptors and may undergo conformational change and protein rearrangement in order to allow the passage of the phage genome into the cytoplasm. The general mechanism of DNA injection remains an open question.

The assembly of the tail tip begins with the formation of a 15S complex by 3 copies of gpJ and this is followed by the sequential addition of gpI, gpL and gpK. Although the order of the interactions between these proteins has been known for many years, little is know about the function and positions of the tail tip proteins in the structure, neither do we know about the composition and stoichiometry of the tail tip complex. Besides, the tail is generally believed to be a 3-fold symmetric structure. Therefore, it will be of interest to know how the 4 side tail fibers are bound to the 3-fold symmetric tail.

In order to answer these questions, it is essential to study the structure of the λ tail tip. In this chapter, I will describe attempts to express and purify the λ tail tip complex and the individual

tail tip proteins. A long term goal of this work, as has been done for the T4 baseplate, is to determine the composition and stoichiometry of the λ tail tip by fitting X-ray structures of the individual tail tip proteins into the cryo-EM density of the tail tip complex.

3.2 ESTABLISHMENT OF THE IDENTITIES OF gpI, gpL AND gpK

At lease four gene products are involved in lambda tail tip assembly. Except for gpJ, the identities of the remaining three tail tip proteins are not well established. In order to identify these proteins, amber mutation or deletion is used to delete most of the target gene product.

3.2.1 Construction of amber mutation / deletion in the gene of interest

The pET21+ vector (Novagen) has a *lacI* gene and has a T7 promoter adjacent to a *lac* operator. This vector provides tighter control of gene expression under non-induced conditions. The four genes encoding λ tail tip proteins were cloned into this vector, resulting in plasmid pLtail (Xu, 2001). An amber mutation (in gene *I* or gene *K*) or deletion (in gene *L*) was introduced to delete most of the target gene product. Table 3.1 is a description of the plasmids used in this section.

3.2.2 The amber / deletion mutant has little biological activity *in vivo*

After construction of each mutation, I wanted to ask if the amber mutation / deletion abolishes the ability of the target gene product to complement *in vivo*. This can be done by the spot

Table 3.1 Description of plasmids used in the spot complementation assay.

The pLtail plasmid encodes all four proteins required for lambda tail tip assembly. The pIam, pLdel and pKam6 plasmids are derivatives of pLtail. The amber mutations were introduced by recombinant PCR whereas the deletion was made by restriction digestion of pLtail plasmid (see MATERIALS AND METHODS).

Plasmid	**Vector used**	**Tail Proteins**	**Description**
pET21+	pET21+	None	Empty vector
pLtail	pET21+	gpJ, gpI, gpL, gpK	λ tail genes *J*, *I*, *K* and *L* were cloned into the pET21+ vector.
pIam	pET21+	gpJ, gpL, gpK	An amber mutation was introduced around the middle of gene *I* resulting in a 102-amino-acid amber fragment.
pLdel	pET21+	gpJ, gpI, gpK	A deletion was made by digesting pLtail with SacI and RsrII. 716 base pairs, including the first 559 base pairs of gene *L*, were deleted from pLtail.
pKam6	pET21+	gpJ, gpI, gpL	An amber mutation was made in the sixth codon of gene *K,* resulting in a 5-amino-acid amber fragment.

complementation assay. In this experiment, decreasing concentrations of phage are spotted on a lawn of cells containing the target plasmid. The phage usually carries an amber mutation in one of the tail genes and therefore is unable to form plaques. If proteins expressed from the plasmid can complement the defect, a clear spot will form. The better the complementation, the fewer phages that are required to form a clear spot.

Tables 3.1 and 3.2 list the plasmids and phages used in the experiment. Each plasmid was spotted with λIam838cI857, λLam756cI857 and λKam768cI857. Figure 3.1 shows that no complementation activity is found between pIam and λIam838cI857, and, pLdel and λLam756cI857. Phage λKam768cI857 does make plaques on the strain carrying pKam6 at about 0.1% efficiency. Since pKam6 encodes only the first 5 amino acids of gpK, it is most likely that this low level of phage growth is due to recombination between Kam6 in the plasmid and Kam768 in the superinfecting phage at low frequency to give a wildtype sequence instead of complementation. This hypothesis is suggested by the fact that the level of growth is insensitive to induction of expression from the plasmid. Therefore, the amber mutation / deletion in the gene of interest abolishes the biological activity of the target gene product. In addition, the result also shows that the amber mutation or deletion has little effect on the growth of the other two amber phages tested, which suggests that the amber mutation / deletion has little polar effect on the other two genes examined.

3.2.3 Identification of gpI, gpL and gpK

The tail tip proteins expressed from the target plasmids were radiolabeled by ^{35}S-Met and the crude lysates were prepared for SDS-PAGE followed by autoradiography (see MATERIALS

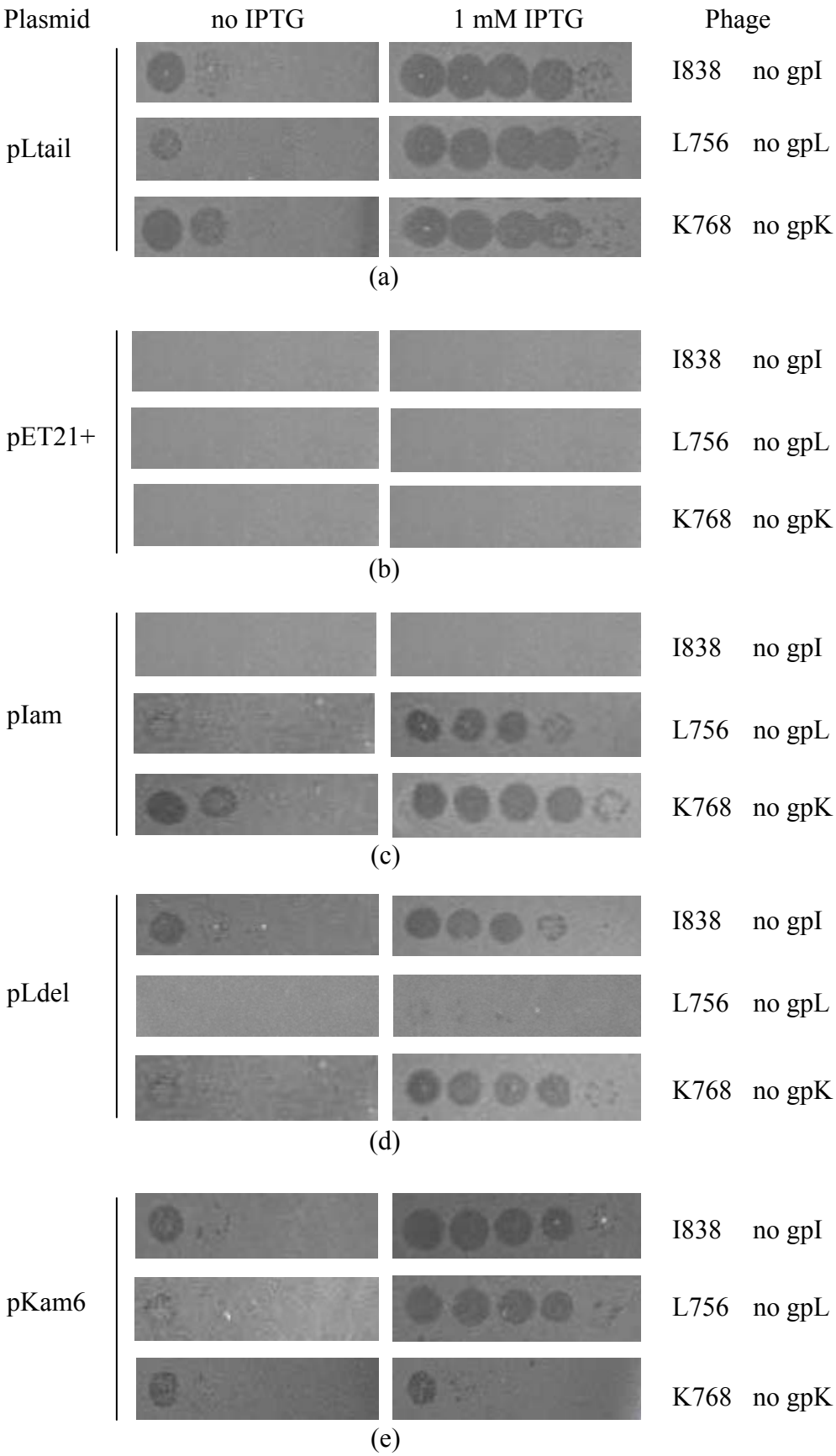
Table 3.2 Phages used in the spot complementation assay.

Phage (abbreviation)	Defects
λIam838cI857 (I838)	No gpI, cI temperature sensitive mutant
λLam756cI857 (L756)	No gpL, cI temperature sensitive mutant
λKam768cI857 (K768)	No gpK, cI temperature sensitive mutant

Note: cI857 is a temperature sensitive mutation in gene *cI*. It prevents lysogenic growth of λ phage at 37℃.

Figure 3.1 The mutant protein has little *in vivo* complementation activity.

BL21(DE3) is an *E. coli* strain containing a λ prophage, DE3. In order to minimize the effect of the tail proteins expressed from the prophage and to prevent recombination between plasmid and prophage, the tail genes of DE3 were knocked out, resulting in BL21(DE3)-Δtail (Xu, 2001). This new strain was used as the host for the target plasmids in this assay. λIam838cI857 (I838), λLam756cI857 (L756) and λKam768cI857 (K768) were used as the test phages. Phages were spotted onto each row with 10-fold decrease in the number from left to right with the left-most spot having about $4X10^5$ phages. The first column indicates the plasmid used. The second and third column indicates the concentration of isopropyl-β-D-thiogalactopyranoside (IPTG) added to the top soft agar. The last column is about the phage used and its defect.


Plasmid
no IPTG
1 mM IPTG
Phage
pLtail
I838 no gpI
L756 no gpL
K768 no gpK
(a)
pET21+
I838 no gpI
L756 no gpL
K768 no gpK
(b)
pIam
I838 no gpI
L756 no gpL
K768 no gpK
(c)
pLdel
I838 no gpI
L756 no gpL
K768 no gpK
(d)
pKam6
I838 no gpI
L756 no gpL
K768 no gpK
(e)

AND METHODS). Table 3.3 shows the predicted molecular weights of the target proteins and Figure 3.2 shows the identities of these proteins on SDS-PAGE. GpI band is absent from the sample containing pIam plasmid (lane 2) whereas gpL or gpK band is absent from the sample containing pLdel or pKam6 (lanes 3-4). The position where gpI appears on SDS-PAGE is consistent with the predicted molecular weight. In terms of gpL and gpK, the positions where they migrate on SDS-PAGE are a little different the positions expected from the calculated molecular weights.

3.3 EXPRESSION OF THE λ TAIL TIP COMPLEX

3.3.1 Addition of 6XHis tag to gpJ

The pLtail plasmid encodes all four proteins involved in lambda tail tip assembly. I first tried to purify the complex by nickel affinity chromatography. This can be done by labeling gpJ with a polyhistidine tag since gpJ is the first protein involved in λ tail tip assembly and anything associated with it will be co-purified. Early studies in LamB mutants tightly blocking phage adsorption revealed that amino acid substitutions at the C-terminal part of gpJ are responsible for the extended host range (Hofnung et al., 1976; Werts et al., 1994). These results imply that the C-terminal part of gpJ may be exposed and directly interact with *E. coli* LamB. Therefore, I tried to insert a 6X or 8XHis tag to four different positions of the C-terminal part of gpJ (Table 3.4). To minimize the disruption of helices and strands, each polyhistidine tag was inserted into the position where it is predicted to be coils. I also tried not to disrupt the sites which are found to be responsible for extended host range. The detailed information about the predicted secondary

Table 3.3 Predicted molecular weights of gpI, gpL and gpK.

The protein sequences used for calculation are based on the DNA sequences of the target genes. The predicted molecular weights of the tail tip proteins are obtained from the following website: http://www.encorbio.com/protocols/Prot-MW.htm.

Protein	Predicted molecular weight (kDa)
gpI	23.1
gpL	25.7
gpK	23.0

Figure 3.2 Establishment of the identities of gpI, gpL and gpK.

The pLysS plasmid produces T7 lysozyme which reduces basal level expression of the genes of interest. Each target plasmid was co-transformed with pLysS into strain BL21(DE3)-Δtail. Cells were grown in M9-MC medium with antibiotics at 37℃. When they reached a density of about $3X10^8$ cell/ml, IPTG (1 mM) was added to induce protein expression. After 30 minutes, rifampicin (200 μg/ml) was added to shut down host strain protein expression. About 20 minutes after the addition of rifampicin, 100 μl sample was taken from each culture and 0.25 μl ^{35}S-Met (10 mCi/ml) was added to each sample for 15 minutes at 37℃. Subsequently, samples were chilled and centrifuged at 8000 RPM for 5 minutes at 4℃. The pellet was resuspended in 40 μl 1XSDS sample buffer. Each sample was boiled for 2.5 minutes before they were loaded onto 15% SDS-PAGE. Proteins were visualized by autoradiography. The positions of the low molecular mass markers are indicated on the left.

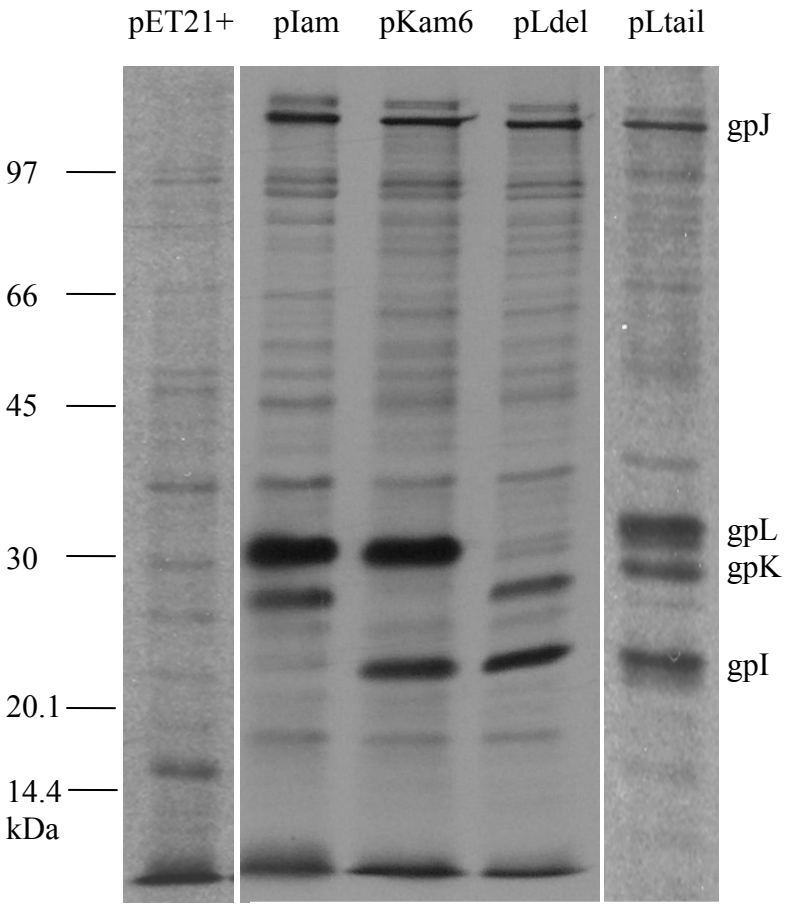

pET21+
pIam
pKam6
pLdel
pLtail
97
66
45
30
20.1
14.4
kDa
gpJ
gpL
gpK
gpI

Table 3.4 Description of plasmids used in the spot complementation assay.

All plasmids list below are derived from pET21+ vector. The polyhistidine tag was introduced into gpJ by recombinant PCR.

Plasmid	Tail proteins	Description
pET21+	None	Empty vector
pLtail	gpJ, gpI, gpL, gpK	Encodes wildtype tail tip proteins
pLtail-m1	gpJ-m1, gpI, gpL, gpK	Differs from pLtail in that it encodes gpJ with an 8XHis tag inserted between 1043^{th} and 1044^{th} residues of wildtype gpJ.
pLtail-m2	gpJ-m2, gpI, gpL, gpK	Differs from pLtail in that it encodes gpJ with a 6XHis tag inserted between 1013^{th} and 1014^{th} residues of wildtype gpJ.
pLtail-m3	gpJ-m3, gpI, gpL, gpK	Differs from pLtail in that it encodes gpJ with a 6XHis tag inserted between 907^{th} and 908^{th} residues of wildtype gpJ.
pLtail-m4	gpJ-m4, gpI, gpL, gpK	Differs from pLtail in that it encodes gpJ with a 6XHis tag inserted between 835^{th} and 836^{th} residues of wildtype gpJ.

structure of the C-terminal part of gpJ, sites responsible for extended host range and the positions where the polyhistidine tags are inserted is shown in Figure 3.3.

3.3.2 Effects of the polyhistidine tag on the biological activity of λ gpJ

After construction of each plasmid, a spot complementation assay was performed to determine the biological activity of the recombinant protein. Each target plasmid was tested for its ability to complement λJam27cI857 which carries an amber mutation in gene *J*. Figure 3.4 shows that in the absence of inducer, the recombinant proteins can complement almost as well as the wildtype protein. However, when lactose is present, only gpJ-m3 can complement almost as well as the wildtype protein. Decreased complementation ability is observed in the remaining three mutants. These results suggest that gpJ-m1, gpJ-m2 and gpJ-m4 may become precipitated at the presence of the inducer. It may also be that these three mutant proteins are able to form heterotrimers with the amber fragment of gpJ in λJam27cI857 in the absence of the inducer and the heterotrimers may become homotrimers when inducer is present.

3.3.3 Expression of the tail tip complex

A time-course expression of the tail tip proteins was carried out and the result is shown in Figure 3.5. Except for gpJ, it is very hard to see the remaining three tail proteins on SDS-PAGE by coomassie blue stain. The fact that the majority of gpJ is present in the pellet suggests that most of the tail tip proteins are insoluble. Different expression temperatures (37℃, 30℃ and 26℃) and different buffer conditions (such as detergent concentrations) have been tried with little improvement in the solubility of gpJ. In addition to the polyhistidine tag, I also tried to clone all

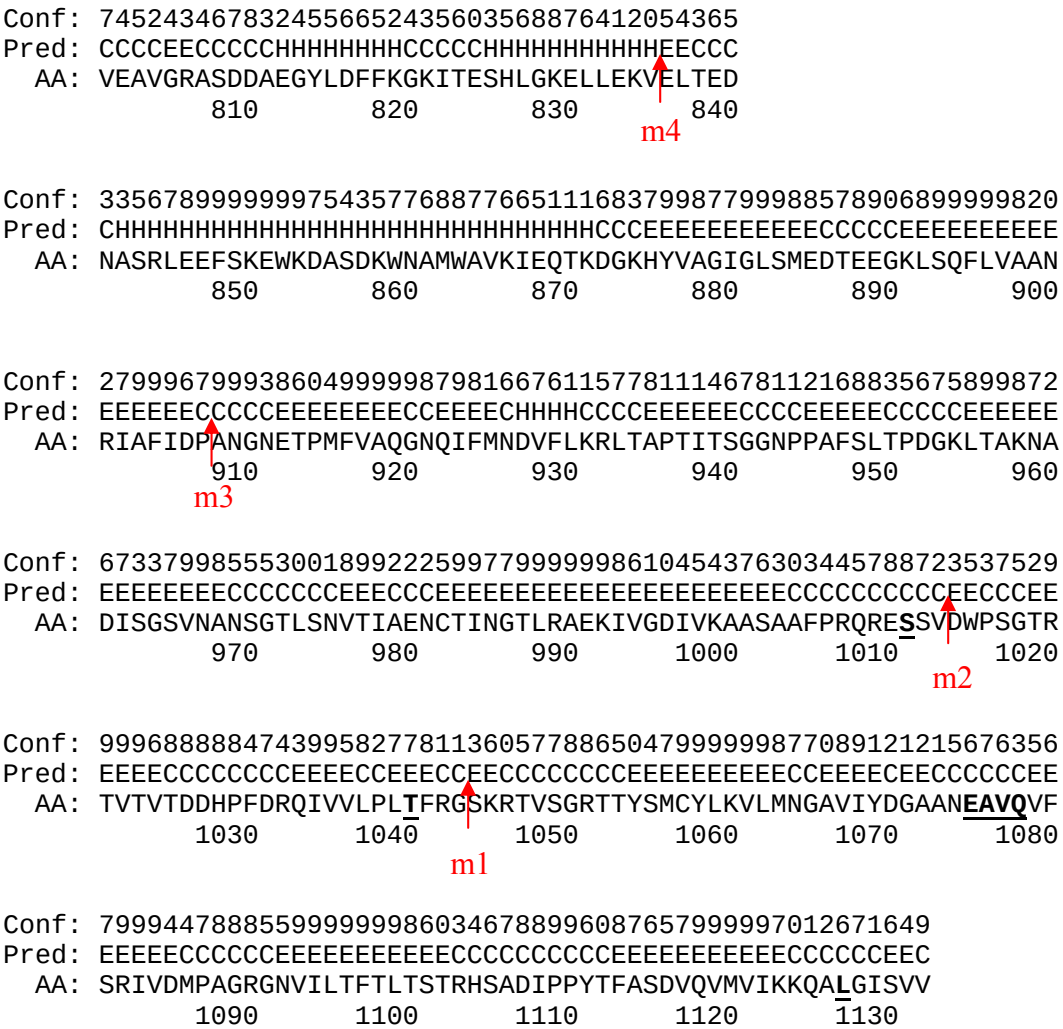
Figure 3.3 The C-terminal part of λ gpJ.

λ gene *J* encodes a 1132 amino acid protein. The last 332 residues are shown here. The secondary structure is predicted by PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/). The residues which are found to be responsible for extended host range are bolded and underlined. Red arrows indicate the positions where the polyhistidine tags were inserted.

Conf: Confidence (0=low, 9=high)

Pred: Predicted secondary structure (H=helix, E=strand, C=coil)

AA: Target sequence


Conf: 74524346783245566524356035688764120543 65
Pred: CCCCEECCCCCHHHHHHHHCCCCCHHHHHHHHHHHEECCC
AA: VEAVGRASDDAEGYLDFFKGKITESHLGKELLEKVELTED
810 820 830 840
m4
Conf: 335678999999975435776887766511168379987799988578906899999820
Pred: CHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHCCCEEEEEEEEEEECCCCCEEEEEEEEE
AA: NASRLEEFSKEWKDASDKWNAMWAVKIEQTKDGKHYVAGIGLSMEDTEEGKLSQFLVAAN
850 860 870 880 890 900
Conf: 279996799938604999998798166761157811146781121688356758998 72
Pred: EEEEEECCCCCEEEEEEEECCEEEECHHHHCCCCEEEEEECCCCEEEEECCCCCEEEEEE
AA: RIAFIDPANGNETPMFVAQGNQIFMNDVFLKRLTAPTITSGGNPPAFSLTPDGKLTAKNA
910 920 930 940 950 960
m3
Conf: 673379985553001899222599779999998610454376303445788723537529
Pred: EEEEEEEECCCCCCCEEECCCEEEEEEEEEEEEEEEEEEEECCCCCCCCCCEECCCEE
AA: DISGSVNANSGTLSNVTIAENCTINGTLRAEKIVGDIVKAASAAFPRQRESSVDWPSGTR
970 980 990 1000 1010 1020
m2
Conf: 99968888847439958277811360577886504799999987708912121567635 6
Pred: EEEECCCCCCCCEEEECCEEECCEECCCCCCCCEEEEEEEEEECCEEEECEECCCCCCEE
AA: TVTVTDDHPFDRQIVVLPLTFRGSKRTVSGRTTYSMCYLKVLMNGAVIYDGAANEAVQVF
1030 1040 1050 1060 1070 1080
m1
Conf: 799944788855999999986034678899608765799999701267164 9
Pred: EEEEECCCCCCEEEEEEEEEECCCCCCCCCCEEEEEEEEEEECCCCCCEEC
AA: SRIVDMPAGRGNVILTFTLTSTRHSADIPPYTFASDVQVMVIKKQALGISVV
1090 1100 1110 1120 1130

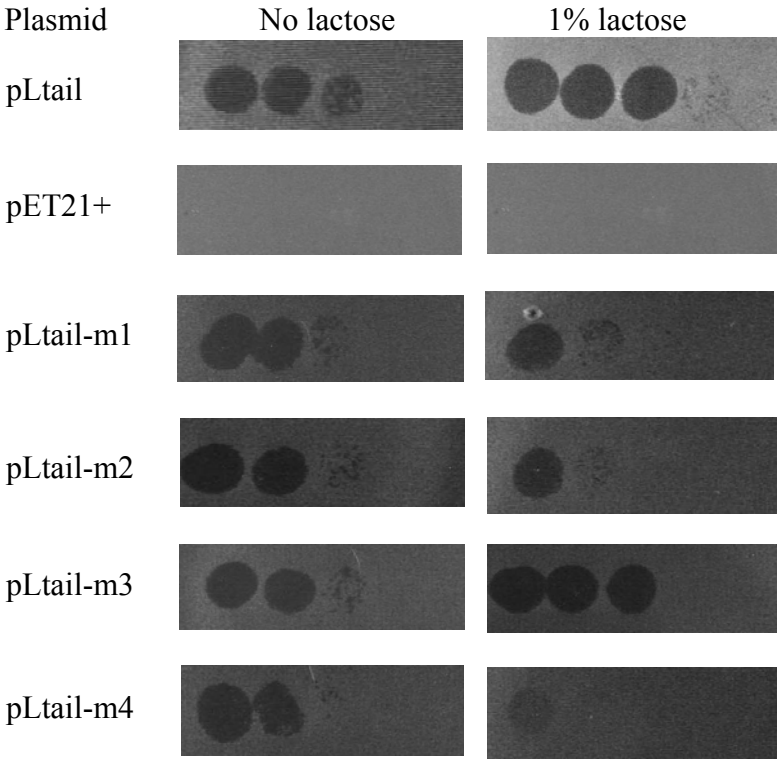
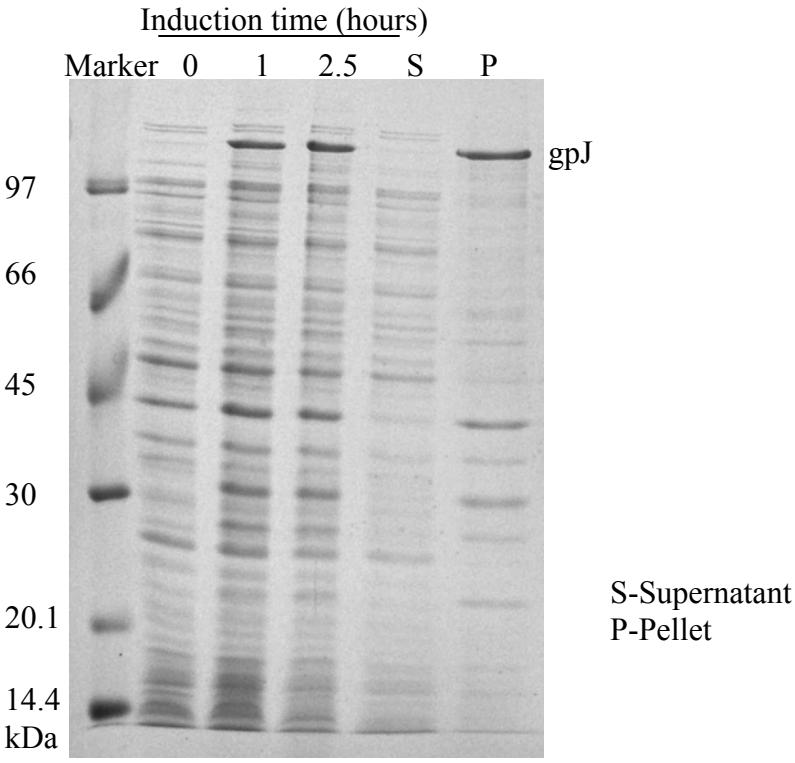


Figure 3.4 Spot complementation assay for the polyhistidine tagged gpJ. BL21(DE3)-Δtail is the host strain for the plasmids of interest. Each plasmid was tested for its ability to complement λJam27cI857, which encodes a defective amber fragment of gpJ. The experiment was carried out as described previously. No lactose or 1% lactose is present in the top soft agar.

Figure 3.5 Expression of λ tail tip complex.

BL21(DE3)-△tail with the target plasmid was grown in LB with antibiotics at 37℃ until it reached about $4X10^8$ cell/ml. IPTG (1 mM) was added to induce protein expression. 100 μl sample was taken at 0 minute, 1 hour and 2.5 hours after induction. Subsequently, the culture was harvested and resuspend in 1Xbinding buffer. Cells were lysed by sonication and the pellet was removed from the supernatant by centrifugation at 13,000 RPM 20 minutes at 4℃. The samples were prepared for 15% SDS-PAGE and visualized by coomassie blue stain. This experiment was done for all the mutants, but the results look very similar. Therefore, only the result for pLtail-m3 is shown here. Except for gpJ, the yield of the remaining three tail tip proteins is too low to be detected on SDS-PAGE by coomassie blue stain.


Induction time (hours)
Marker
0
1
2.5
S
P
gpJ
97
66
45
30
20.1
14.4
kDa
S-Supernatant
P-Pellet

four tail tip genes into the pGEX-6p-1 vector which adds a GST tag at the N-terminus of gpL. Although the yield of the tail tip complex are high (based on the amounts of gpJ), most of them are insoluble (data not shown).

3.4 PURIFICATION OF λ TAIL TIP PROTEINS

As described above, I had some trouble in obtaining a soluble form of the tail tip complex. Subsequently, I tried to express and purify each of the individual tail tip proteins. gpJ will not be described below since the addition of 6XHis tag, MBP, or GST to the C-terminus of gpJ results in an inactive protein (data not shown). Besides, although the recombinant gpJ is produced in large amounts, the majority of the recombinant protein is insoluble (data not shown). The expression and purification of the remaining three tail tip proteins will be described below.

3.4.1 Construction of the recombinant proteins

Three different tags were used to help purify the tail tip proteins. Below is a description of the vectors used for the construction of the recombinant proteins.

pTQ30: The EcoRI-XbaI fragment of pQE30 (QIAGEN) was cloned into the EcoRI- XbaI sites of pT7-5, resulting in pTQ30. Proteins expressed from this vector are under the control of the T7 promoter. In my experiments, the target gene was inserted into the HindIII-SacI sites of pTQ30. BL21(DE3)-Δtail is used as the host strain. A 16-amino-acid sequence (MRGSHHHHHHGSACEL), including a 6XHis tag, is added to the N-terminus of the target

protein. The main advantage of this system is that the short polyhistidine tag can minimize the effect of the foreign sequence on the activity of the target protein.

pLC3: The PCR fragment encoding N-terminal 6XHis tagged maltose binding protein (MBP) was inserted into the NcoI-NdeI sites of pET28b vector. The resulting construct (pLC3) can be used for the construction of proteins with MBP at the N-terminus. The MBP fusion protein can be cleaved by AcTEV protease, which recognizes the seven-amino-acid sequence ENLYFQG and cleaves between glutamine (Q) and glycine (G) with high specificity. The 29-kDa protease contains a polyhistidine tag at the N-terminus and therefore can be removed by affinity chromatography. In the experiments, the target gene was cloned into the BamHI-NdeI sites of pLC3. After AcTEV protease digestion, an additional 5-amino-acid sequence (GDITH) remains at the N-terminus of the target protein.

pGEX-6p-1: The pGEX-6p-1 vector (GE healthcare) contains an internal *lacI*^{q} gene and a *tac* promoter. This vector can be used for the construction of N-terminal glutathione S-transferase (GST) fusion proteins. In the experiments, the gene of interest was cloned into the BamHI-EcoRI sites of the vector. GST can be cleaved from the target protein by PreScission Protease. The protease recognizes a subset of sequences, including the core amino acid sequence LFQGP, and cleaves between glutamine (Q) and glycine (G) residues with high specificity. The 46-kDa PreScission Protease can be easily removed from the cleavage reaction by affinity chromatography since it is fused to GST. The main advantage of using PreScission Protease is that it allows low-temperature cleavage (5℃) of fusion proteins, which minimizes the

degradation of the target proteins. After cleavage, a 5-amino-acid sequence (GPLGS) remains at the N-terminus of the target protein.

3.4.2 Determination of the biological activities of the recombinant proteins

Each recombinant protein was tested for its ability to complement *in vivo*. The spot complementation assay was carried out as described previously. Table 3.5 is a description of the plasmids used and Figures 3.6-3.8 show the results. The amounts of IPTG added to the top soft agar in each experiment was adjusted to obtain the best result. The results (Figures 3.6-3.8) show that the addition of 6XHis tag, MBP or GST to the N-terminus of gpI or gpL has little effect on the biological activity of the target protein. However, the addition of any of the three tags to the N-terminus of gpK abolishes its ability to complement *in vivo*.

3.4.3 Expression of λ gpI

gpI is a protein with 223 amino acids. A 6XHis tag, MBP or GST was added to the N-terminus of gpI. Unfortunately, the yield for each recombinant protein was too low to be detected on SDS-PAGE by coomassie blue stain. An example of the time-course expression of MBP-I is shown in Figure 3.9. Although MBP is expressed in large amounts, the MBP-I fusion protein is not detectable on SDS-PAGE by coomassie blue stain. In order to improve the yield of gpI, several approaches were tried.

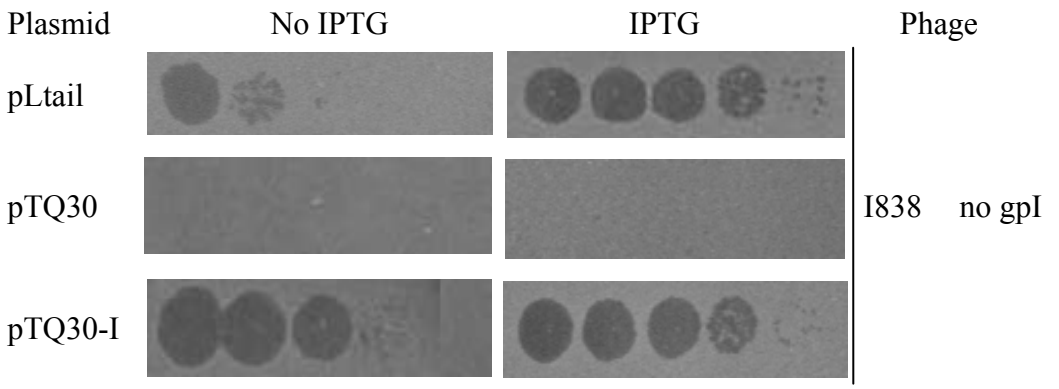
Table 3.5 Description of the plasmids used in the spot complementation assay.

Each tag (6XHis tag, MBP or GST) was added to the N-terminus of the target proteins.

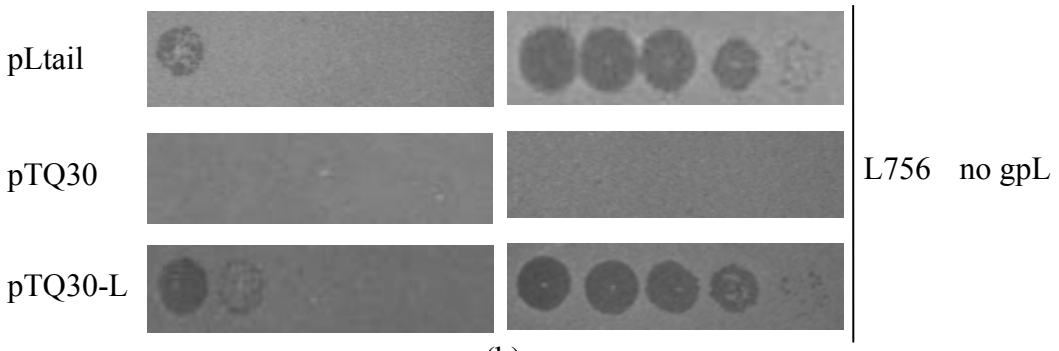
Plasmid	Vector	Tail proteins
pLtail	pET21+	Wildtype gpJ, gpI, gpL, gpK
pTQ30	pTQ30	None
pTQ30-I	pTQ30	6XHis tagged gpI
pTQ30-L	pTQ30	6XHis tagged gpL
pTQ30-K	pTQ30	6XHis tagged gpK
pLC3	pLC3	None
pLC3-I	pLC3	MBP-I
pLC3-L	pLC3	MBP-L
pLC3-K	pLC3	MBP-K
pGEX-6p-1	pGEX-6p-1	None
pGEX-6p-1-I	pGEX-6p-1	GST-I
pGEX-6p-1-L	pGEX-6p-1	GST-L
pGEX-6p-1-K	pGEX-6p-1	GST-K

Figure 3.6 Determination of the biological activity of 6XHis tagged proteins.

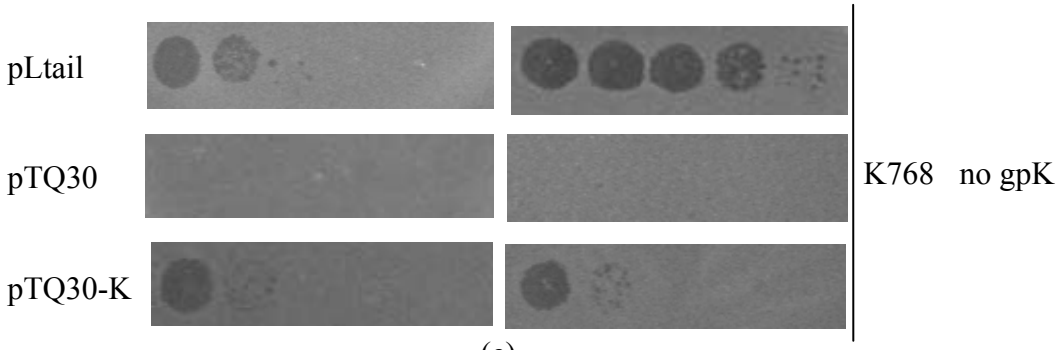
BL21(DE3)-△tail was used as the host strain. The spot complementation assay was performed as described previously. No IPTG (second column) or 0.01 mM IPTG (third column, except for pLtail which contains 1 mM IPTG) is present in the top soft agar.


Plasmid
No IPTG
IPTG
Phage
pLtail
pTQ30
pTQ30-I
I838
no gpI

(a)


pLtail
pTQ30
pTQ30-L
L756
no gpL

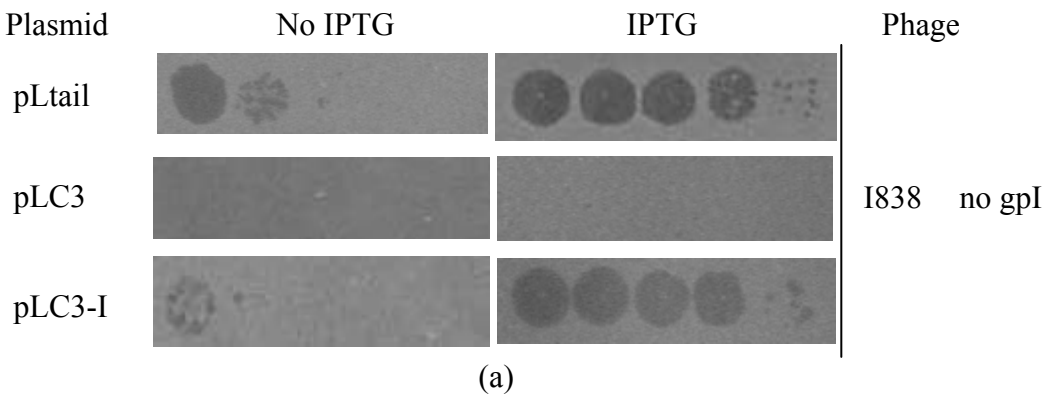
(b)


pLtail
pTQ30
pTQ30-K
K768
no gpK

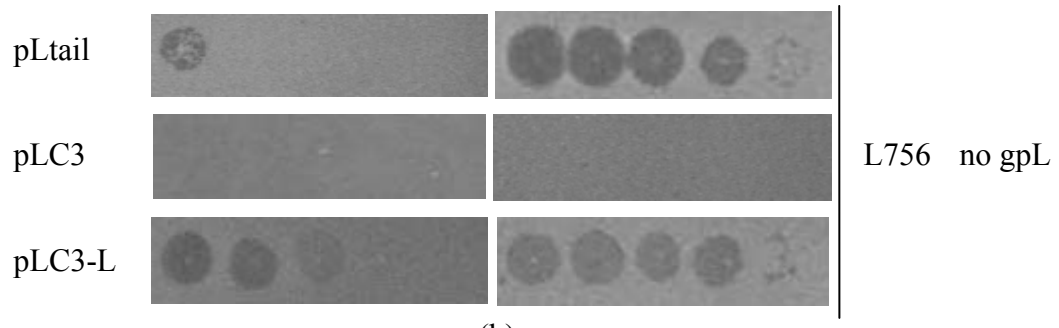
(c)

Figure 3.7 Determination of the biological activity of the MBP fusion proteins.

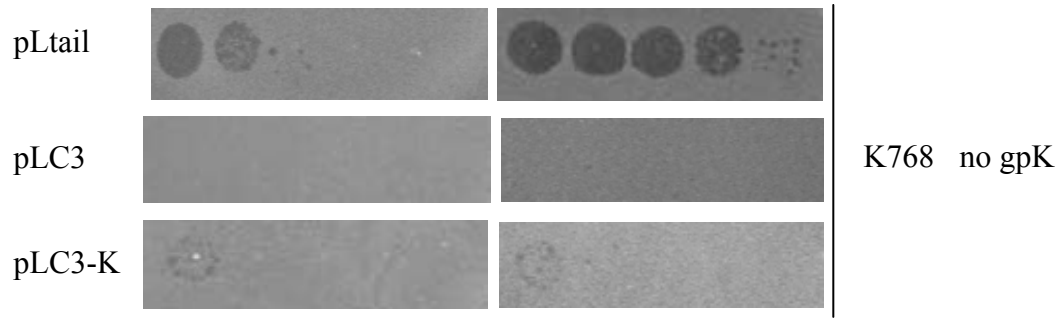
BL21 Star^{TM}(DE3) (Invitrogen) is an *E. coli* strain designed for improving protein yield in a T7 promoter-based expression system. Transcription from the T7 promoter is not coupled to translation in *E. coli* since the amounts of mRNA synthesized by T7 RNA polymerase are much more than those by *E. coli* RNA polymerases. This causes the unprotected mRNA to be easily degraded by endogenous RNases. This strain contains a mutation in the gene encoding RNaseE (*rne131*), one of the major sources for this mRNA degradation, and therefore improves the stability of mRNA transcripts and increases protein expression yield from T7 promoter-based vectors. This strain was used as the host for pLC3 based plasmids. The experiment was performed as described previously. No IPTG (second column) or 0.2 mM IPTG (third column, except for pLtail which contains 1 mM IPTG) was added to the top soft agar.


Plasmid
No IPTG
IPTG
Phage
pLtail
pLC3
pLC3-I
I838
no gpI

(a)


pLtail
pLC3
pLC3-L
L756
no gpL

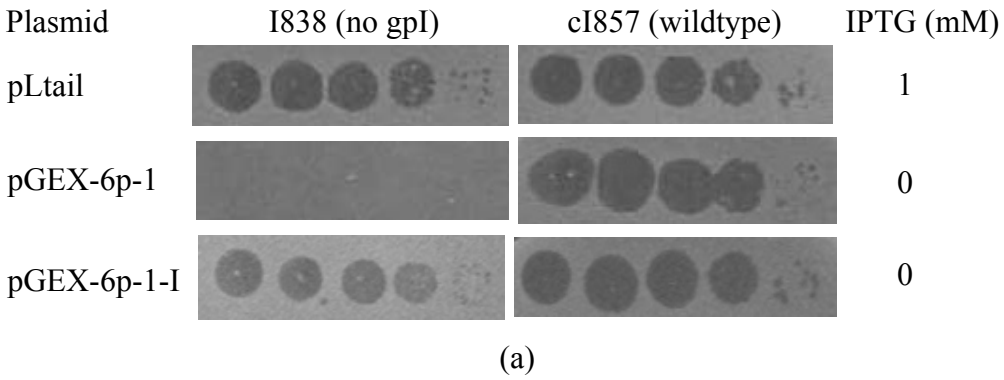
(b)


pLtail
pLC3
pLC3-K
K768
no gpK

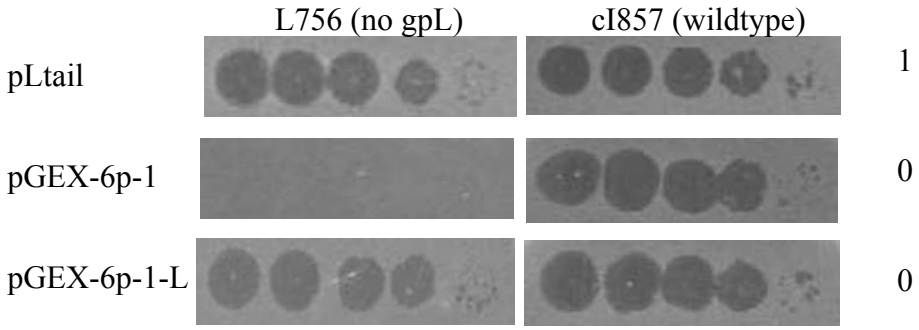
(c)

Figure 3.8 Determination of the biological activity of the GST fusion proteins.

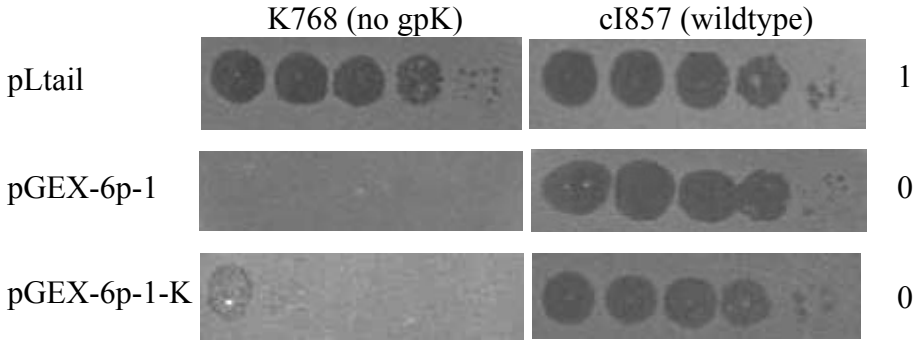
BL21(DE3)-Δtail was used as the host strain. Except for pLtail which contains 1 mM IPTG in the top soft agar, no inducer was added to the remaining plasmids since the recombinant proteins complement best in the absence of inducer. As an alternative, each target plasmid was also tested for its ability to complement λcI857 (cI857), a type of phage which differs from the wildtype λ phage in that it carries a *ts* mutation in gene *cI*. This phage behaves like the wildtype phage in the complementation assay.


Plasmid
I838 (no gpI)
cI857 (wildtype)
IPTG (mM)
pLtail
1
pGEX-6p-1
0
pGEX-6p-1-I
0

(a)


L756 (no gpL)
cI857 (wildtype)
pLtail
1
pGEX-6p-1
0
pGEX-6p-1-L
0

(b)


K768 (no gpK)
cI857 (wildtype)
pLtail
1
pGEX-6p-1
0
pGEX-6p-1-K
0

(c)

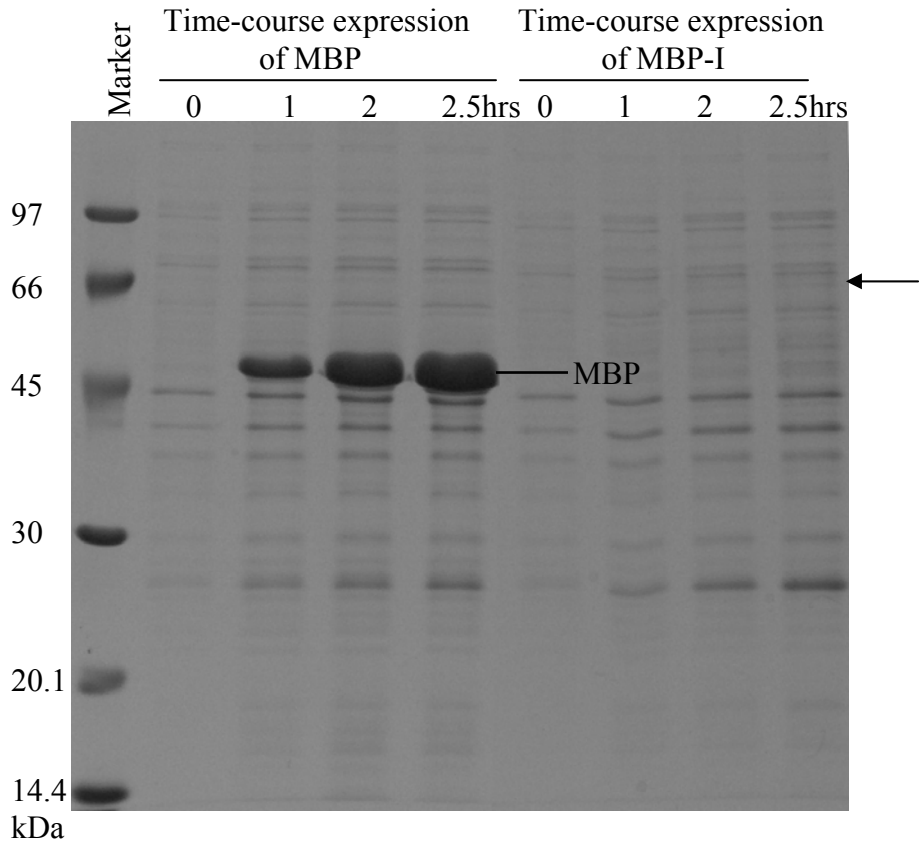


Figure 3.9 Time-course expression of MBP and MBP-I. BL21Star^{TM}(DE3) with the appropriate plasmid was used for protein expression. Cells was grown at 37℃ until they reached about 4X10^{8} cell/ml. IPTG was added to a final concentration of 1 mM to induce protein expression at 37℃. 100 µl sample was taken at 0 minute, 1 hour, 2 hours and 2.5 hours after induction. Subsequently, the samples were prepared for 15% SDS-PAGE and visualized by coomassie blue stain. Arrow indicates the position where the MBP-I fusion protein is supposed to be.

(1) Toxicity of the protein

After induction, cells with plasmid encoding recombinant gpI do not grow as well as cells with the empty vector. This suggests that gpI may be toxic to *E. coli* cells. In order to provide more stringent control of gene expression, pLysE is used instead of pLysS. Cells with pLysE produce higher amounts of T7 lysozyme than those with pLysS and therefore provide tighter control. Time-course expression of the recombinant gpI was conducted. However, there is little improvement in the yield of gpI (data not shown since the result looks the same as the one using pLysS).

(2) Rare codons

Codons that correspond to rare tRNAs are the least used codons in *E. coli*. An excess of any of these codons may cause problems during translation and thus a reduction in quantity of the protein produced. One solution is to use a host strain which contains a plasmid with the appropriate tRNA. Gene *I* contains several rare codons, such as AGA, CGA and AUA. I wanted to examine if the low expression level of gpI is due to the presence of the rare codons. If the poor yield of gpI is due to the presence of rare codons, an improvement in the yield of the protein will be obtained when the rare tRNAs are supplied. Otherwise, no improvement will be observed.

Rosetta^{TM} 2 (Novagen) is a derivative of BL21, which supplies rare tRNAs for seven codons, AUA, AGG, AGA, CUA, CCC, GGA, and CGG. The target plasmid was transformed into this host strain. Compared to the yield of gpI using BL21(DE3)-Δtail as the host strain, there is no improvement in the yield of this protein when expressed in strain Rosetta^{TM}2 (Figure 3.10). This suggests that the low expression level of gpI is not due to the presence of rare codons.

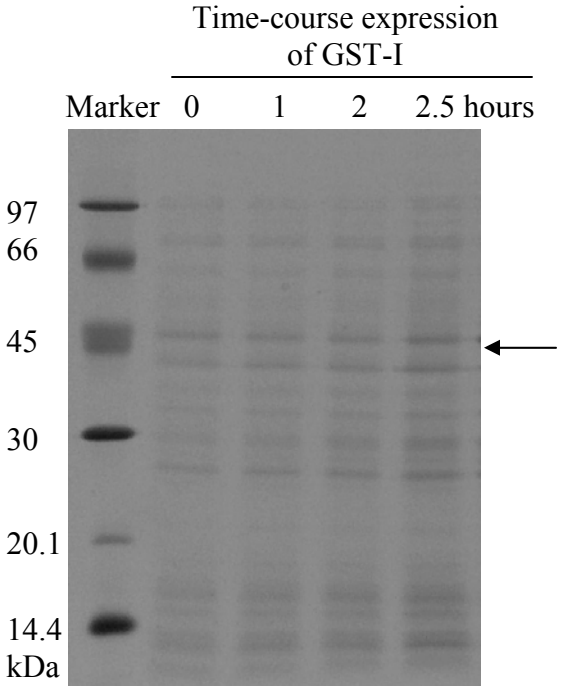


Figure 3.10 Time-course expression of GST-I in Rosseta^{TM}2. The time-course expression of 6XHis-tagged gpI, MBP-I and GST-I was performed, but only the result for GST-I fusion protein is shown here. BL21(DE3)-Δtail with the appropriate plasmid was used for protein expression. The experiment was carried out as described previously (Figure 3.9). Instead of adding 1 mM IPTG, 0.4 mM IPTG was added to induce protein expression. The samples were loaded to 15% SDS-PAGE and visualized by coomassie blue stain. Arrow indicates the position where the GST-I fusion protein is supposed to be.

(3) Stability of the mRNA

One possible reason for the low yield of gpI is that the mRNA is not stable. This is because theoretically MBP or GST has its own translational start site and should be translated efficiently even if gpI is not translated well. However, in my experiments, I could not detect any MBP or GST on SDS-PAGE by coomassie blue stain when they are fused to gpI. When I looked at the original DNA sequence of gene *I*, I noticed that there is a sequence after the terminator of gene *I* which forms a secondary structure. It is likely that this secondary structure is required for the stability of mRNA. In order to find out if this is true, this sequence was added after the terminator of gene *I* in the pGEX-6p-1-I construct. However, I did not observe any change in the yield of the protein (data not shown). So far, I still do not know how to improve the yield of gpI.

(4) Stability of the protein

It is possible that the low yield of gpI is due to the instability of the protein. A pulse-chase experiment (see MATERIAL AND METHODS) was performed to investigate the stability of MBP-I. Figure 3.11 shows that over 50% of radiolabeled MBP-I is degraded within 10 minutes and over 90% of the protein is degraded 40 minutes after being synthesized. This indicates that the MBP-I fusion protein is not stable and subject to being degraded. Attempts to stabilize gpI by co-expressing it with gpJ are desrcribed below in section 3.4.6.

3.4.4 Expression and purification of λ gpL

(1) Expression and purification of gpL by 6XHis tag

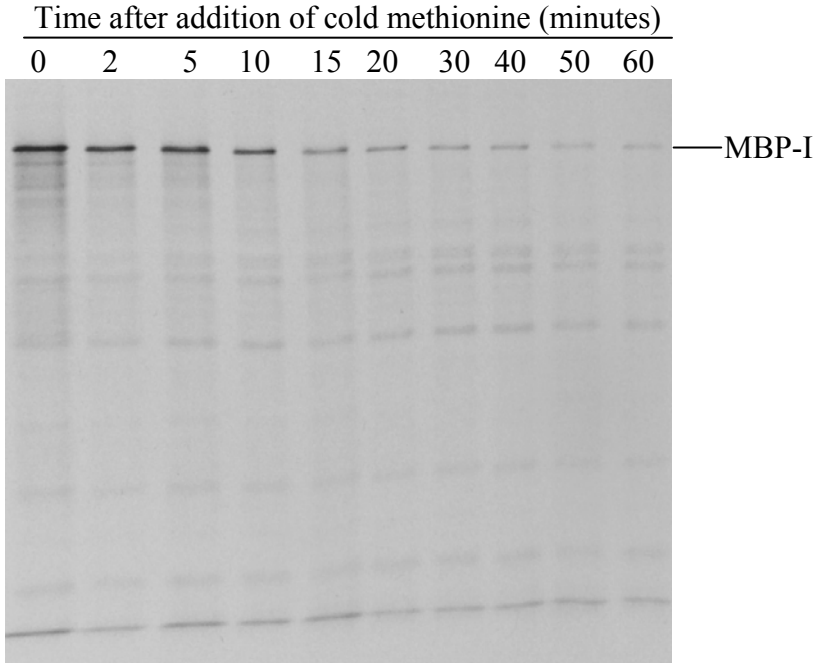
The wildtype gpL contains 232 amino acids. At first, I labeled it with a 6XHis tag by cloning gene *L* into the pTQ30 vector. The polyhistidine tagged gpL is expressed in large amounts and

Figure 3.11 MBP-I is subject to degradation after being synthesized.

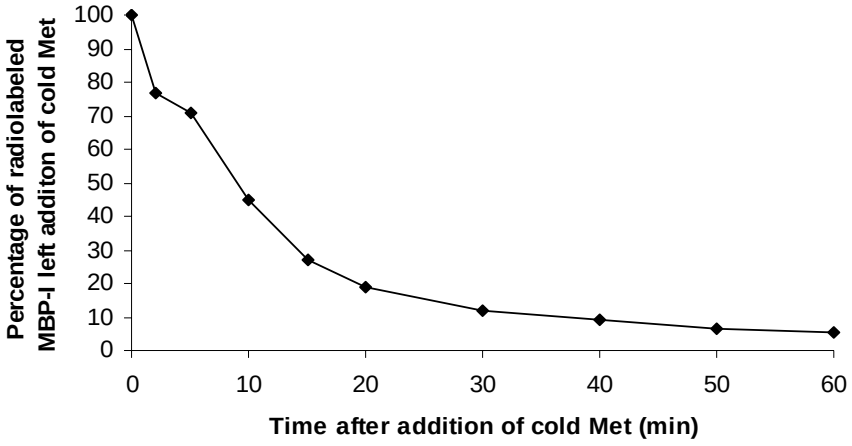
BL21Star^{TM}(DE3)pLysS with pLC3-I was grown in M9-MC medium at 37℃ until it reached about 3X10^{8} cell/ml. IPTG (1 mM) was added to induce protein expression. About 30 minutes later, rifampicin (200 μg/ml) was added to shut down host strain protein expression. After 20 minutes, 1 μl (10 mCi/ml) ^{35}S-Met was added for 30 seconds. Subsequently, about 1000-fold excess of cold methionine was added. Samples were taken at 0, 2, 5, 10, 15, 20, 30, 40, 50 and 60 minutes after the addition of cold methionine. The samples were concentrated by TCA precipitation before they were prepared for 15% SDS PAGE and visualized by autoradiography.

(a) Lanes 1-10: samples taken at 0, 2, 5, 10, 15, 20, 30, 40, 50 and 60 minutes after the addition of cold methionine.

(b) The net intensity of MBP-I band was quantified. The percentage of radiolabeled MBP-I was plotted against time after addition of cold methionine. The percentage of the radiolabeled fusion protein right before the addition of cold methionine (0 minute) is set to 100%.


Time after addition of cold methionine (minutes)
0 2 5 10 15 20 30 40 50 60
MBP-I

(a)


Percentage of radiolabeled MBP-I left additon of cold Met
100
90
80
70
60
50
40
30
20
10
0
0 10 20 30 40 50 60
Time after addition of cold Met (min)

(b)

most of the recombinant protein is soluble (Figure 3.12 lanes 12-13). The 6XHis tagged gpL was first purified by nickel affinity chromatography (see MATERIALS AND METHODS) and the result is shown in Figure 3.12. Under the experimental conditions, most of the 6XHis-tagged gpL is present in elution fractions. There are small amounts of some other contaminant proteins which co-elute with the recombinant protein. Several different buffer conditions were tried with only a very slight improvement in the purity of the target protein (data not shown). Several other purification methods, such as ion-exchange chromatography, were tried to further purify the polyhistidine tagged gpL, but there is only a very slight improvement in the purity of the protein (data not shown). In addition, I noticed that the nickel column turns to grey after loading the sample containing 6XHis tagged gpL. The elution fraction which has the deepest color contains the highest concentration of gpL. This suggests that the cysteine residues in gpL may react with the nickel ion. Therefore, I decided to switch to MBP fusion to purify gpL.

(2) Purification of gpL by MBP fusion

MBP-L can complement as well as the wildtype gpL *in vivo*. This fusion protein was expressed at 37℃ and the purification result is shown in Figure 3.13. Panel (a) shows the result for the purification of MBP-L by amylose affinity chromatography. Most of the fusion protein is present in the elution fractions. Although the fusion protein is not very pure after amylose column purification, I decided to try to do the cleavage reaction. The purified MBP-L was incubated with AcTEV Protease to cleave MBP from gpL. The result shows that the cleavage reaction is not complete (Panel b lane 1). When the digestion mixture was loaded onto the hydroxyapatite chromatography to remove maltose, the amount of cleaved gpL in the elution fraction is greatly decreased (Panel b lane 2). The elution fraction from the hydroxyapatite chromatography was

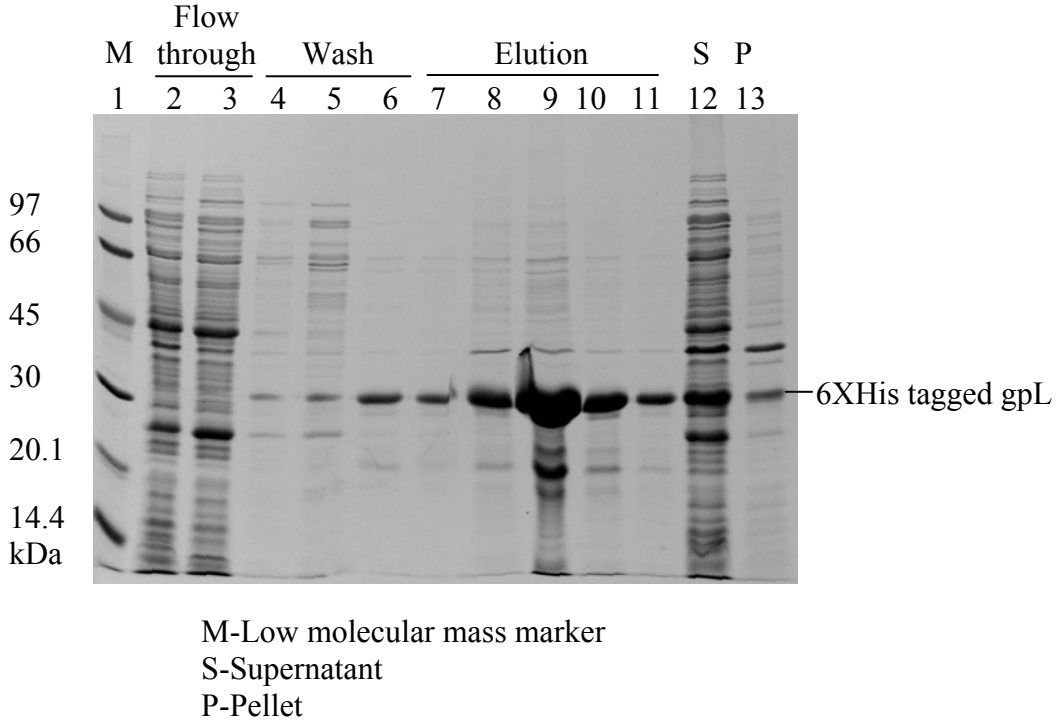
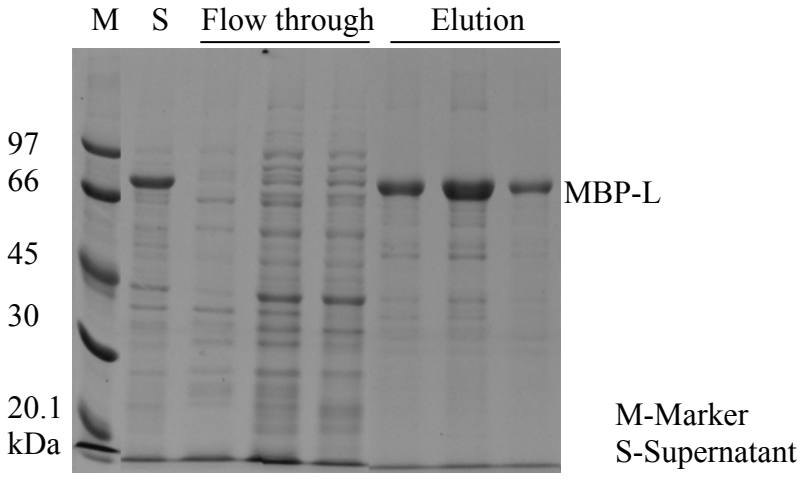


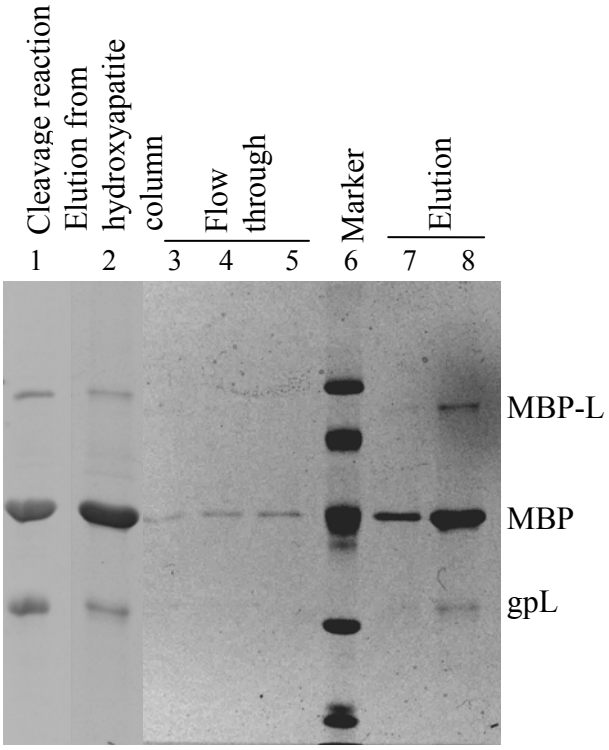
Figure 3.12 Purification of 6XHis tagged gpL by nickel affinity chromatography. BL21(DE3)-Δtail pLysS with pTQ30-L was grown in LB with antibiotics at 37℃ until it reached about $4X10^8$ cell/ml. IPTG (0.4 mM) was added to induce protein expression for 2.5 hours. Subsequently, cells were chilled, harvested and resuspended in 1Xbinding buffer (20 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl pH 7.9). Cells were lysed by sonication and the supernatant was separated from the pellet by centrifugation at 13,000 RPM for 20 minutes at 4℃. The supernatant was loaded onto a nickel column for purification (see MATERIALS AND METHODS). The concentrations of imidazole used in the binding and wash buffers are 20 mM and 80 mM, respectively. No reductant was added in the buffers used in this experiment.

Figure 3.13 Purification of gpL by MBP fusion.

BL21Star^{TM}(DE3) pLysS with pLC3-L was grown in LB with 0.2% (W/V) glucose and antibiotics at 37℃ until it reached about 4X10^{8} cell/ml. IPTG (1 mM) was added to induce protein expression. About 2.5 hours later, cells were chilled on ice, harvested and resuspended in the column buffer (20 mM Tris-HCl pH 7.4, 200 mM NaCl, 1 mM EDTA, 1 mM DTT). Cells were lysed by sonication and the pellet was removed by centrifugation at 10,000 RPM for 20 minutes at 4℃. The supernatant (Panel a lane 2) was used for amylose affinity chromatography as described in the MATERIALS AND METHODS. Since the wildtype gpL has 8 cysteines, all the buffers used contains 1 mM dithiothreitol (DTT) to prevent the formation of disulfide bonds.


M
S
Flow through
Elution
97
66
45
30
20.1
kDa
MBP-L
M-Marker
S-Supernatant

(a)


Cleavage reaction
Elution from hydroxyapatite column
Flow through
Marker
Elution
1
2
3
4
5
6
7
8
MBP-L
MBP
gpL

(b)

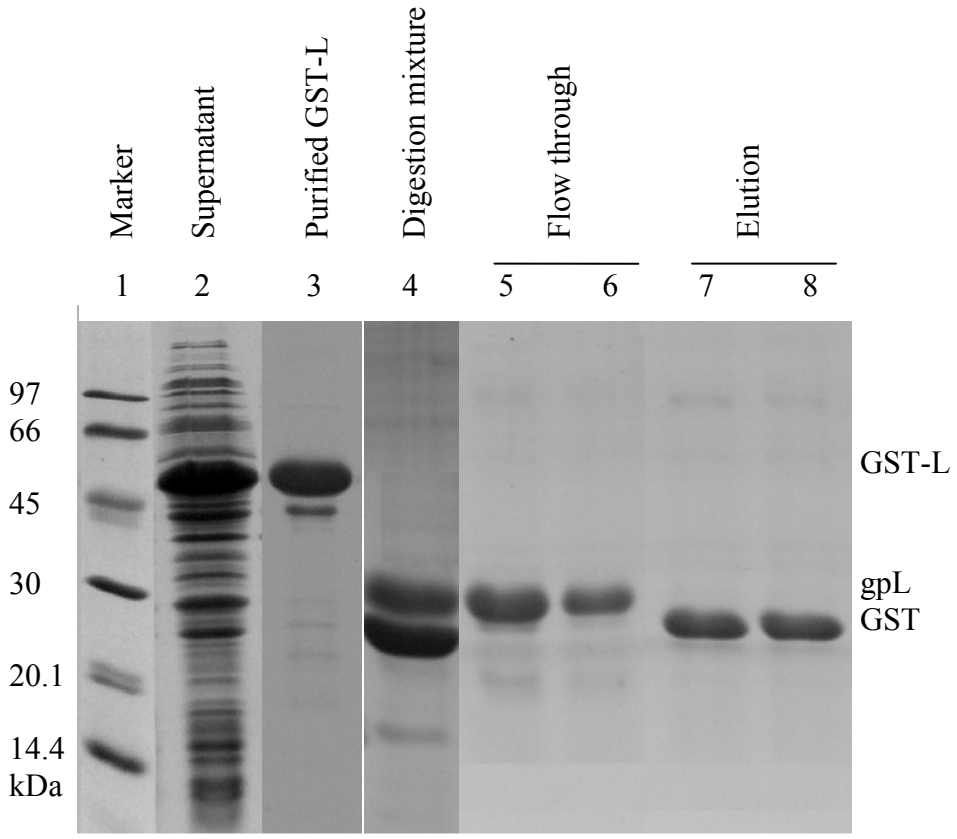
reloaded onto the amylose column to separate gpL from MBP and uncleaved MBP-L. However, the majority of gpL co-elutes with MBP and MBP-L (Panel b). Several different purification methods, such as ion exchange chromatography and gel filtration, were tried (data not shown). But, most gpL is either precipitated during purification or co-elutes with MBP and MBP-L.

(3) Purification of gpL by GST fusion

Due to the difficulty of the removal of MBP and MBP-L from gpL, I tried to purify gpL by GST fusion and the result is shown in Figure 3.14. GST-L is expressed in large amounts in a soluble form at 37℃ (lane 2). Early attempts to purify GST-L suggest that the bacterial Hsp70, DnaK, co-elutes with it (data not shown). In order to dissociate the complex, the crude lysate was first incubated with ATP and Mg^{2+} at 37℃ for 20 minutes. After removal of pellet by centrifugation, the supernatant was loaded onto the Glutathione Sepharose chromatography for purification. About 60-70 mg of the fusion protein can be purified from 1 liter culture. The result shows that there are several minor proteins co-purified with GST-L (lane 3). The minor bands disappear after digestion by PreScission Protease (lane 4), suggesting that they may be degradation products of GST-L. The digestion mixture was reloaded onto the Glutathione Sepharose chromatography. Lanes 5-8 shows that the cleaved gpL is present in the flow through fractions while GST is in the elution fractions. 15-20 mg of cleaved gpL can be obtained from 60-70 mg of the purified fusion protein. This protein has been sent to our collaborators for structural studies.

Figure 3.14 Purification of gpL by GST fusion.

BL21(DE3)-△tail is used as the host strain. Cells containing the target plasmid were grown at 37℃ until they reached about $4X10^8$ cell/ml. IPTG (1 mM) was added to induce protein expression at 37℃ for 2.5 hours (or cells can be induced with 0.4 mM IPTG at 28℃ overnight). After this, the culture was chilled on ice, harvested and resuspended in Buffer 1 (50 mM Tris-HCl pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.1% Triton X-100, 1mM DTT). Cells were broken down by sonication and the lysate was incubated with ATP and $Mg2^+$ before it was centrifuged to remove the pellet. The supernatant was used for Glutathione Sepharose affinity chromatography as described in MATERIALS AND METHODS. The 1X wash buffer used in this experiment contains 0.5% Triton X-100. However, the buffers used for the Glutathione Sepharose chromatography to separate gpL from the cleavage mixture do not contain any detergent. All the buffers used contain 1 mM DTT.


Marker
Supernatant
Purified GST-L
Digestion mixture
Flow through
Elution
1
2
3
4
5
6
7
8
97
66
45
30
20.1
14.4
kDa
GST-L
gpL
GST

3.4.5 Expression and purification of λ gpK

(1) Expression of the recombinant protein

The gene *K* of phage lambda encodes a 199-amino-acid sequence. When it is labeled with a 6XHis tag at the N-terminus, the yield of this protein is too low to be used for purification (data not shown). The GST-K fusion protein is expressed in large amounts. However, the majority of the protein is insoluble. Different expression temperatures (37℃, 30℃ and 26℃) and buffer conditions (such as presence or absence of detergents) have been tried with no improvement in the solubility of the protein (data not shown).

In terms of the MBP fusion protein, most of MBP-K is insoluble when it is induced at 37℃ with 1 mM IPTG (Figure 3.15). However, when the growth temperature is lowered, more fusion protein becomes soluble. When the temperature is reduced to 26℃, most of the fusion protein made is soluble. The purification of MBP-K fusion will be described below.

(2) Purification of gpK by MBP fusion

The fusion protein was expressed as described above. Figure 3.16 shows the purification result. 70-85 mg of the fusion protein can be purified from 1 liter culture. Several other proteins co-elute with MBP-K (Panel a). In order to further purify the fusion protein, ammonium sulfate precipitation was performed. The result shows that most of the fusion protein precipitates between 60-70% saturated ammonium sulfate (Panel b). Subsequently, the purified MBP-K fusion protein was digested by AcTEV Protease. The digestion reaction is not complete (Panel c lane 2). When this mixture was loaded onto the hydroxyapatite chromatography to remove

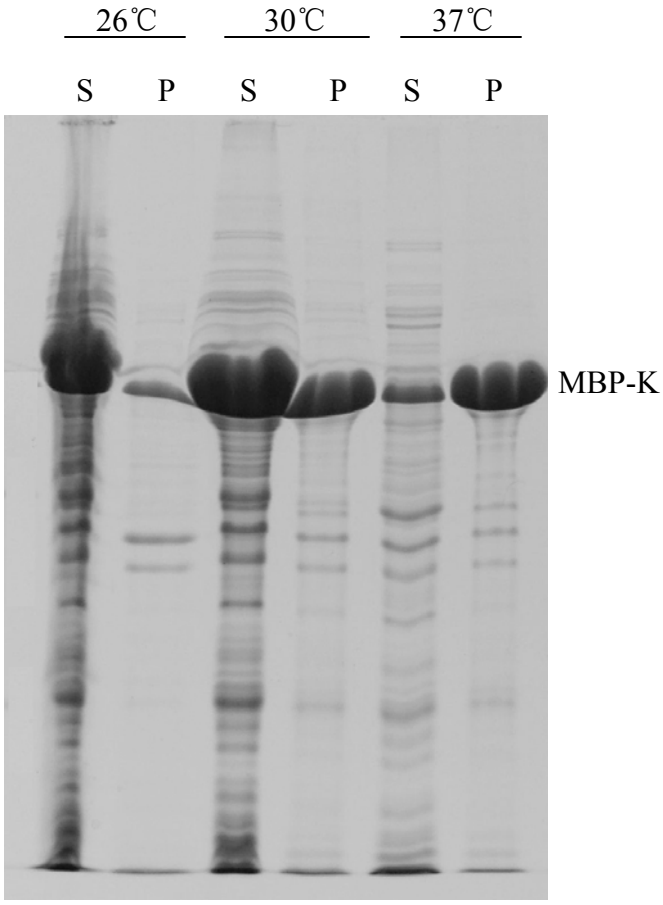
Figure 3.15 The effect of temperature during expression on the solubility of MBP-K.

BL21 Star^{TM}(DE3) was used as the host strain. The culture was grown at different temperatures until they reached about 4X10^{8} cell/ml. IPTG was added to induce protein expression. Subsequently, cells were chilled on ice, harvested and resuspended in the column buffer. After lysis by sonication, the pellet was removed from the supernatant by centrifugation at 10,000 RPM for 20 minutes at 4℃. The pellet was resuspended in the same volume of the column buffer as the supernatant. Finally, the supernatant and pellet was prepared for 15% SDS-PAGE and visualized by coomassie blue stain.

Lanes 1-2: The culture was grown at 30℃ before induction. At the time of adding 0.4 mM IPTG, the temperature was switched to 26℃ overnight.

Lanes 3-4: The culture was grown at 30℃. When cells reached the right density, 1 mM IPTG was added to induce protein expression for 3.5 hours at 30℃.

Lanes 5-6: The culture was grown at 37℃ and induced with 1 mM IPTG for 2.5 hours at the same temperature.



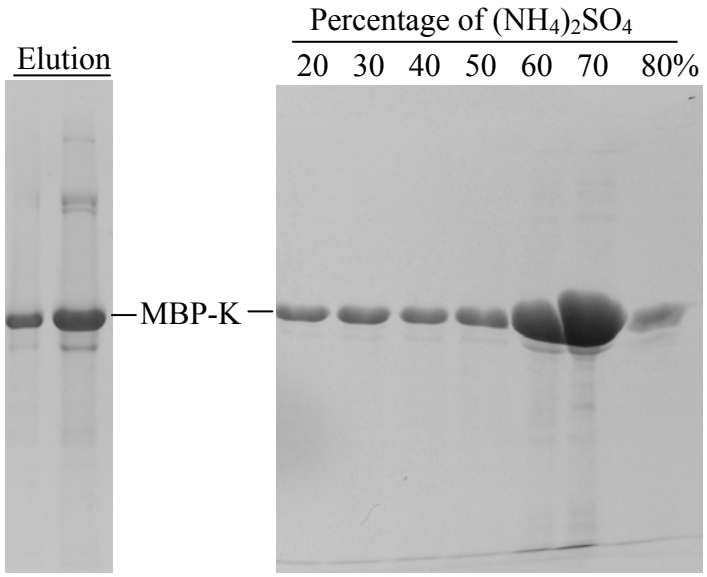
S-Supernatant
P-Pellet

Figure 3.16 Purification of gpK by MBP fusion.

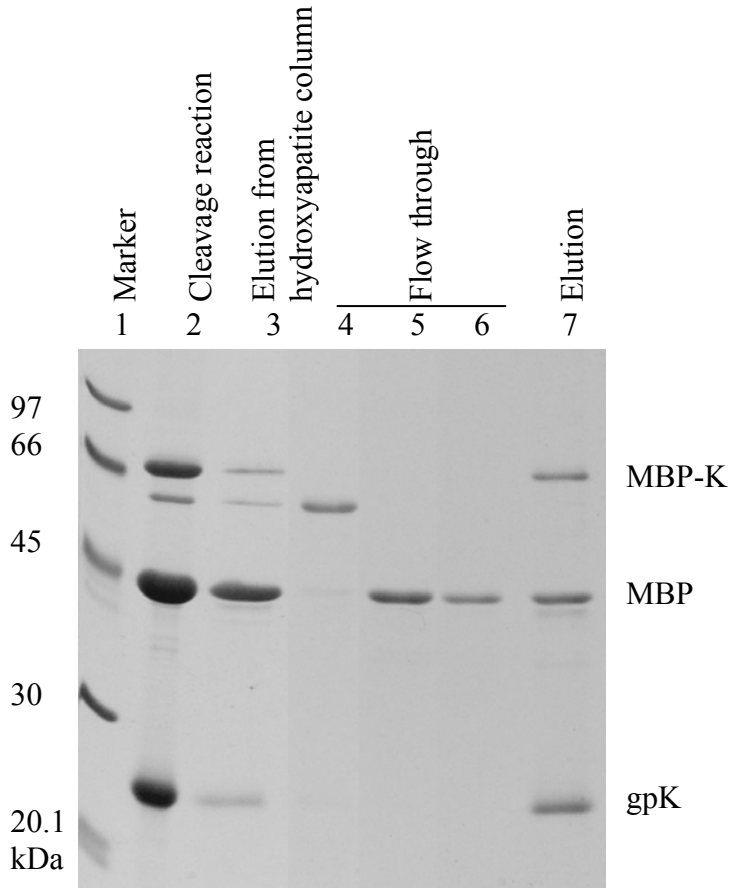
(a) The supernatant from the culture induced at 26℃ was used for amylose affinity chromatography purification (see MATERIALS AND METHODS) and the elution fractions are shown.

(b) Ammonium sulfate precipitation was performed to further purify MBP-K fusion protein. Proteins that were pelleted by ammonium sulfate at different concentrations are shown.

(c) The purified protein was cleaved by AcTEV Protease and the result is shown in lane 2. After this, the cleavage mixture was loaded onto the hydroxyapatite chromatography to remove maltose. The elution fraction (lane 3) was reloaded onto the amylose affinity chromatography to purify gpK and the result is shown in lanes 4-7. No DTT is present in the buffers used for purification.


Elution
Percentage of $(NH_4)_2SO_4$
20 30 40 50 60 70 80%
MBP-K

(a) (b)


Marker
Cleavage reaction
Elution from hydroxyapatite column
Flow through
Elution
1 2 3 4 5 6 7
97
66
45
30
20.1 kDa
MBP-K
MBP
gpK

(c)

maltose, 80-90% of gpK is lost due to precipitation (Panel c lane 3). The elution fraction was reloaded onto the amylose affinity chromatography to purify gpK from the other components. However, like cleaved gpL from MBP-L, the cleaved gpK co-elutes with MBP and uncleaved MBP-K. Several other purification methods, such as ion-exchange chromatography, DEAE-Sepharose chromatography and gel filtration, have been tried. But, the majority of gpK is either lost possibly due to its precipitation on the top of the column or co-eluted with MBP and MBP-K possibly due to the incompelete cleavage of the fusion protein.

3.4.6 Co-expression of the λ tail tip proteins

As shown above, I have had some trouble in the solubility of gpJ and in the expression of gpI. It is possible that some tail proteins may need to be co-expressed in order to obtain a stable or soluble form of the protein. In order to examine if this is true, I cloned the target genes into the high expression plasmid pGEX-6p-1 and tried four different co-expression experiments: (1) gpJ and GST-I, (2) gpJ, gpI and GST-K, (3) gpI and GST-K, (4) gpI, gpK and GST-L. Although gpJ alone is expressed in large amounts in an insoluble form, no tail proteins are observed in the SDS-PAGE by coomassie blue stain when gpJ is co-expressed with GST-I (data not shown). When gpJ is co-expressed with gpI and GST-K, most of gpJ is insoluble (data not shown). In terms of the remaining two constructs (gpI and GST-K or gpI, gpK and GST-L), only the protein with the tag has a good yield and can be detected on SDS-PAGE by coomassie blue stain. The other tail protein(s) in the construct can not be detected in the crude lysate or purified complex (data not shown).

3.5 DISCUSSION

3.5.1 Comparisons of the different purification systems

Three different tags were used in the above experiments to facilitate protein purification. The main advantage of polyhistidine tag is that it is very small and therefore can minimize the effect of additional sequence on the activity of the target protein. However, the yield of some proteins is low and this may not be improved by simply adding a polyhistidine tag. MBP or GST fusion can sometimes solve this problem since they have their own translational start site. Besides, the MBP or GST fusion can sometimes help improve the solubility and stability of the target protein. The main disadvantage for these two systems is that the tag is large and therefore they may impair the activity of the target protein.

3.5.2 The solubility of the λ tail tip complex

As shown above, the majority of the tail tip complex is insoluble, which makes the purification impractical. The early studies for the *in vitro* complementation assays between different tail-defective lysates suggest that the tail tip complex must be soluble to some extent. This disagreement can be explained by two possible reasons. First, the conditions for sample preparation are different. For samples prepared for purification, the lysates are centrifuged much harder than those used for *in vitro* complementation assays. Besides, the buffer conditions are also different. Second, it is possible that only a small fraction of the tail tip proteins is soluble and this amount is enough for the *in vitro* complementation assay. The results from Chapter 4 also imply that the amounts of the tail tip proteins required for the *in vitro* complementation

assay is very small. If this is true, it may be that lambda tail assembly initiates at the cell membrane and then continues to assemble in the cytoplasm.

3.5.3 Expression and purification of λ tail tip proteins

In my experiments to date, gpJ and gpI have presented the most intractable problems for purification due to their insolubility and for gpI, instability. GpL has been the easiest to work with, and this has lead to a successful purification. gpK has been intermediate in these regards, showing good purification as the MBP fusion protein, but poor recovery after cleavage from MBP. GpK would seem to be an alternative target for future attempts to carry forward the purification to tail tip proteins.

4.0 DETERMINATION OF THE COMPOSITION AND STOICHIOMETRY OF THE LAMBDA TAIL TIP COMPLEX

4.1 INTRODUCTION

The assembly of λ tail tip is controlled by at least four proteins. The order of the interactions among these proteins has been known for many years (Figure 1.5). However, little is known about the composition, stoichiometry and detailed structure of the tail tip. There are several different ways of addressing these questions. A fundamental approach is to isolate the tail tip complex, separate it into its components by denaturing gel electrophoresis, and quantify the relative amounts of the protein components by densitometry. The copy numbers of gpJ, gpH and gpV have been determined in a similar way (Casjens and Hendrix, 1974). The composition can be related to the structure of the complex if X-ray structures of individual proteins can be fit into the cryo-EM density of the tail tip. The arrangement of subunits in the T4 baseplate has been determined by this method (Rossmann et al., 2001; Rossmann et al., 2004). In Chapter 3, I described attempts to purify the tail tip complex and the individual tail tip proteins for structural studies, but the success of this work to date has been only the purification of gpL.

In this chapter, I tried to determine the composition and stoichiometry of the tail tip complex which were prepared by several different methods.

4.2 PRECIPITATION OF λ TAIL TIP ASSEMBLY INTERMEDIATES BY *E. COLI* CELLS

One crude way of examining the compositions and stoichiometries of the tail tip assembly intermediates is to precipitate them from *E. coli* cells followed by SDS-PAGE, autoradiography and densitometry. This method depends on the specific interaction between the λ gpJ tail fiber protein and the λ receptor, LamB, on the surface of *E. coli* (Tsui and Hendrix, 1983). The pIam, pLdel, pKam6 and pLtail plasmids were previously described in Table 3.1. The tail tip proteins expressed from these plasmids were radiolabeled (see MATERIALS AND METHODS). After lysis, the lysate was incubated with *E. coli* cells to allow the binding of the tail fiber to the LamB receptor. Subsequently, cells were washed three times with λ dil to remove the unbound components and the samples were prepared for SDS-PAGE followed by autoradiography and densitometry. A strength of this approach is that all tail tip assembly intermediates are expected to be precipitated, because gpJ is the first protein in the assembly pathway. In this experiment, *E. coli* cells will precipitate any complexes containing gpJ, including assembled structures as well as any partially assembled structures. Therefore, the stoichiometries of the proteins may be distorted toward over-representation of components that add early in the assembly pathway.

Figure 4.1 is the result of this experiment and Table 4.1 shows the copy numbers of the tail tip proteins in the complex quantified from Figure 4.1. The result (lane 2) indicates that in the absence of gpI, a very small amount of gpL can assemble onto gpJ. In contrast, the absence of gpL does not seem to affect the assembly of gpI (lane 4). These results agree with the *in vitro* complementation data, which indicate that gpI precedes gpL in the assembly pathway. In the

Figure 4.1 Precipitation of λ tail tip assembly intermediates by *E. coli* cells.

The pLtail plasmid encodes all four proteins required for lambda tail tip assembly. The pIam, pKam6 and pLdel are derivatives of pLtail and carry an amber mutation in gene *I* (pIam) or gene *K* (pKam6), or a deletion in gene *L* (pLdel). The detailed information about these plasmids was described in Table 3.1. The plasmids were co-transformed with pLysS into BL21(DE3)-△tail. Tail proteins expressed from the plasmids were radiolabeled and lysed (see MATERIALS AND METHODS). The samples were incubated with *E. coli* cells for 1 hour or overnight at 4℃ before they were washed with λ dil to remove the unbound components. The bound proteins were prepared for 15% SDS-PAGE. The gel was dried and visualized by autoradiography.

Plasmids used in this experiment:

Lane 1: pET21+ (empty vector)

Lane 2: pIam (I^-, encodes gpJ, gpL and gpK)

Lane 3: pKam6 (K^-, encodes gpJ, gpI and gpL)

Lane 4: pLdel (L^-, encodes gpJ, gpL and gpK)

Lane 5: pLtail (complete, encodes gpJ, gpI, gpL and gpK)

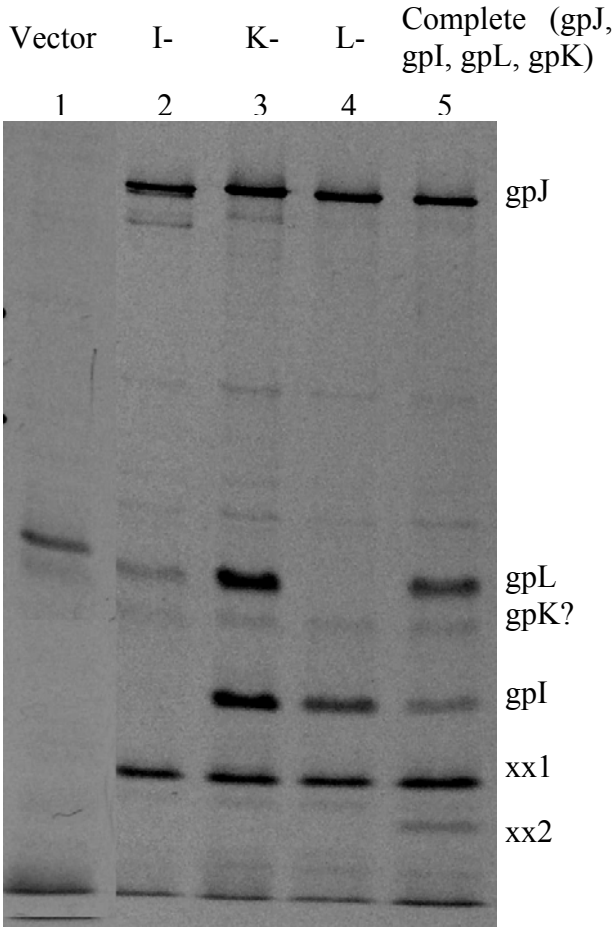

Vector
I-
K-
L-
Complete (gpJ, gpI, gpL, gpK)
1
2
3
4
5
gpJ
gpL
gpK?
gpI
xx1
xx2

Table 4.1 Stoichiometries of λ tail tip proteins in the samples precipitated by *E. coli* cells.

The experiment shown in Figure 4.1 was repeated twice. The SDS-PAGE was exposed to films for different times (For example: 1 hour, 2 hours, 3 hours and 6 hours). The films were then scanned and analyzed as described in MATERIALS AND METHODS. The copy numbers of gpI, gpL and gpK were corrected for the number of sulfur atoms in the proteins and normalized to 3 copies of gpJ. The copy numbers were rounded to the nearest integers.

	pIam	pKam6	pLdel	pLtail
gpJ	**3**	**3**	**3**	**3**
gpI		3-5	3-4	2-3
gpL	~1	2-3		2-4
gpK				<0.6*

* The band of gpK examined is the gpK-like band shown in Figure 4.1. The intensity of the gpK-like band is obtained by subtraction of the net intensity of the gpK-like band in lane 3 from that of lane 5.

absence of gpK, the other three proteins appear to be present in amounts similar to the situation when it is present. This again agrees with the *in vitro* complementation data, which say that gpK acts after gpJ, gpI and gpL. There is a very faint band which migrates about the same position as gpK in lanes 2-5 (Figure 4.1). The fact that the band is also present in the lysate in which there is an amber mutation early in gene *K* shows that this band is not gpK. This band could possibly obscure small amounts of gpK in other lysates. Therefore, it is not clear whether or not gpK is present in the tail tip assembly intermediates from this experiment, but it is clearly not present in amounts comparable to the other tail proteins. In fact, the amount of material in the gpK position indicating the contaminant band amounts to less than one molecule per structure. The absence of a gpK band in the K^-, L^- and I^- lanes of the gel in Figure 4.1 is not unexpected, in the case of the K^- lane because of the K amber mutation and in the cases of the L^- and I^- lanes because I and L precede K in the assembly pathways. Its absence in the pLtail plasmid lane is unexpected. Possible explanation include: 1) gpK is not efficiently assembled into the complex in these conditions, 2) it is assembled but dissociates in the *in vitro* conditions, or 3) it is assembled but is proteolytically processed or degraded. An unknown band, xx2, which has a mobility corresponding to about 14 kDa, is only detected in the sample from pLtail plasmid. It is not clear if it is a processed or degradation product of one of the tail proteins. The intensity of the gpI band in this sample is reduced relative to the other samples where it is present, suggesting that band xx2 may be a fragment of gpI. Alternatively, it could be a fragment of gpK.

A striking feature of the gel in Figure 4.1 is the band labeled xx1 with a mobility corresponding to about 16 kDa. This is apparently not one of the phage-encoding proteins, or a processed form of one of them, because it is present in the lanes with I, K and L mutations as well as in the intact

pLtail lane. It is also not derived from the experimental plasmid since it is absent in the empty vector lane. It may be a cellular protein that associates stably with gpJ. If we assume an average sulfur content for this protein (the sulfur content in the protein corresponding to band xx1 is the same as that of λ gpJ), it is present in about 8-12 copies for each 3 copies of gpJ. It has not been possible to attempt further identification of this protein because it is not available in biochemical scale quantity.

4.3 CHARACTERIZATION OF λ TAIL TIP BY SUCROSE VELOCITY GRADIENT

There are two main drawbacks to the technique described in the previous section. First, *E. coli* cells will precipitate assembled structures as well as any partially assembled structures. Thus, the stoichiometries of the tail proteins may be distorted. Second, there is no way to determine the biological activity of the tail tip proteins. To solve this, sucrose gradient was used to separate lambda tail tip complex followed by *in vitro* complementation assay to examine the biological activity of the tail proteins in each fraction. The fractions from the gradient were further assayed by co-immunoprecipitation (Co-IP) followed by SDS-PAGE, autoradiography and densitometry to determine the compositions and stoichiometries of the fractions from the gradient. So, this experiment is able to determine if the tail tip proteins in the complex have any biological activity.

4.3.1 Construction of biologically active T7-tagged gpL

The pGtail plasmid encodes the first 8 proteins involved in lambda tail assembly. This plasmid was chosen instead of pLtail because the crude lysate from pGtail shows much better *in vitro* complementation activity than that of pLtail.

A T7 tag was added to the N-terminus of gpL and the new plasmid was called pGtail-LT7. The spot complementation assay was performed to determine the ability of the recombinant protein to complement λcI857 and λLam756cI857. Table 4.2 lists the plasmids used in this assay. Figure 4.2 shows that the recombinant protein can complement as well as the wild type protein, suggesting that the addition of T7 tag has little effect on the biological activity of gpL *in vivo*.

4.3.2 Separation of λ tail tip complex by sucrose velocity gradient

λ tail proteins expressed from pGtail-LT7 plasmid were radiolabeled by ^{35}S-Met and prepared for 10-25% sucrose gradient (see MATERIALS AND METHODS). After centrifugation in an SW60 rotor at 58,000 RPM for 3.5 hours at 4℃, the centrifuge tube was punctured at the bottom and the sucrose gradient collected into 12 fractions. Each fraction was examined by *in vitro* complementation assay and co-immunoprecipitation.

(1) *In vitro* complementation assay

This assay is aimed to determine the biological activity of the tail tip proteins in the fractions from the gradient. Figure 4.3 is an illustration of the experimental design. The lysogen which is not able to make tails due to an amber mutation in one of the tail genes is used to make

Table 4.2 Plasmids used in the spot complementation assay.

Plasmid	**Vector used**	**Protein expressed**	**Description**
pET21+	pET21+	None	Empty vector
pGtail	pET21+	gpJ, gpI, gpL, gpK, gpM, gpG, gpGT, gpH	Wildtype genes J, I, L, K, M, G, T and H were introduced into pET21+ vector
pGtail-LT7	pET21+	gpJ, gpI, T7-tagged gpL, gpK, gpM, gpG, gpGT, gpH	Differs from pGtail in that it encodes gpL with a T7 tag at the N-terminus

Note: The sequence of T7 tag added to gpL is MASMTGGQQMG.

Figure 4.2 The T7-tagged gpL is biologically active.

The target plasmids were co-transformed with pLysS into strain BL21(DE3)-Δtail. They were spotted with λcI857 (cI857) and λLam756cI857 (L756). The first column is the plasmids being tested. No IPTG (second column) or 1 mM IPTG (third column) is present in the top soft agar. The fourth column shows the type of phage spotted and the defect in each phage. Each row was spotted with 10-fold decrease in number of phages from left to right with the left-most spot having about $4X10^5$ phages.

(a) pET21+ is an empty vector and produces no tail proteins.

(b) pGtail encodes the first 8 proteins involved in lambda tail assembly.

(c) pGtail-LT7 differs from pGtail in that it encodes a T7-tagged gpL.

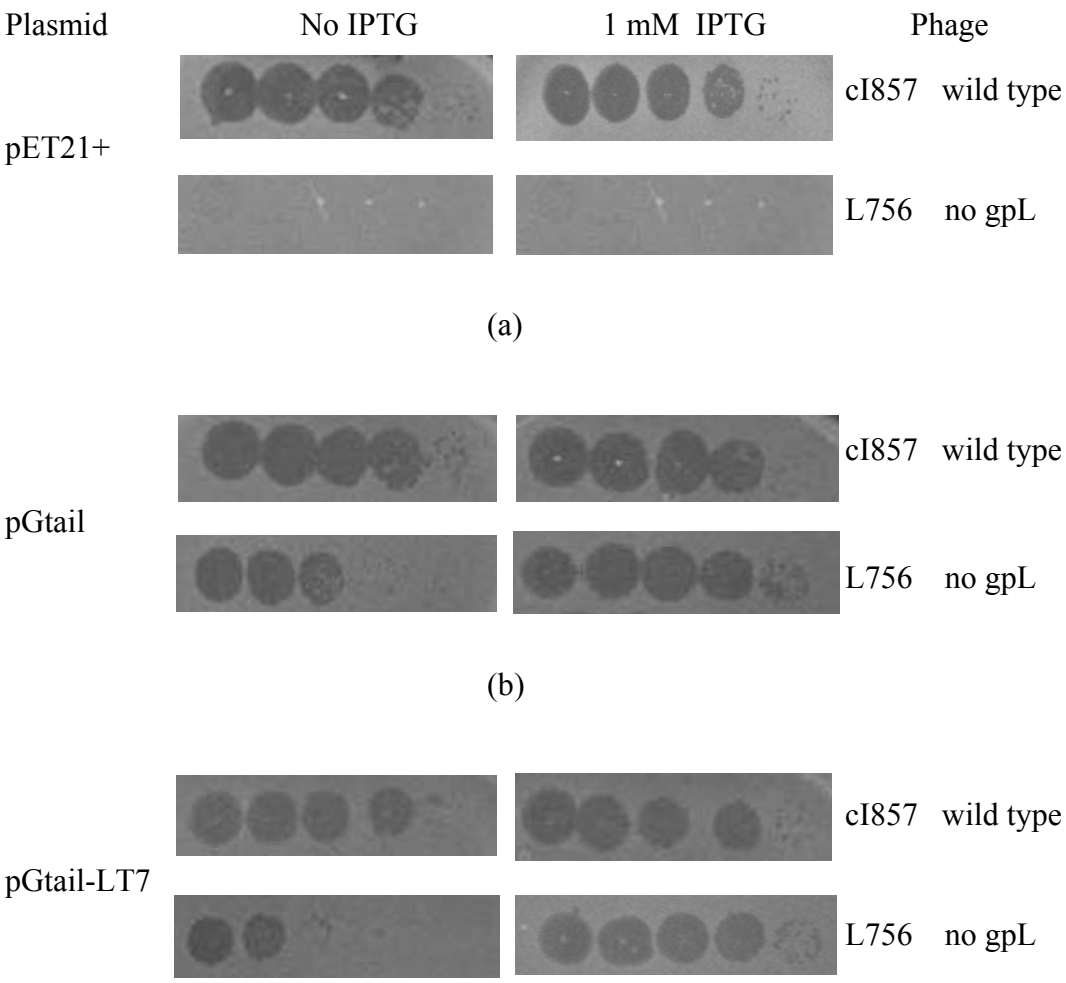
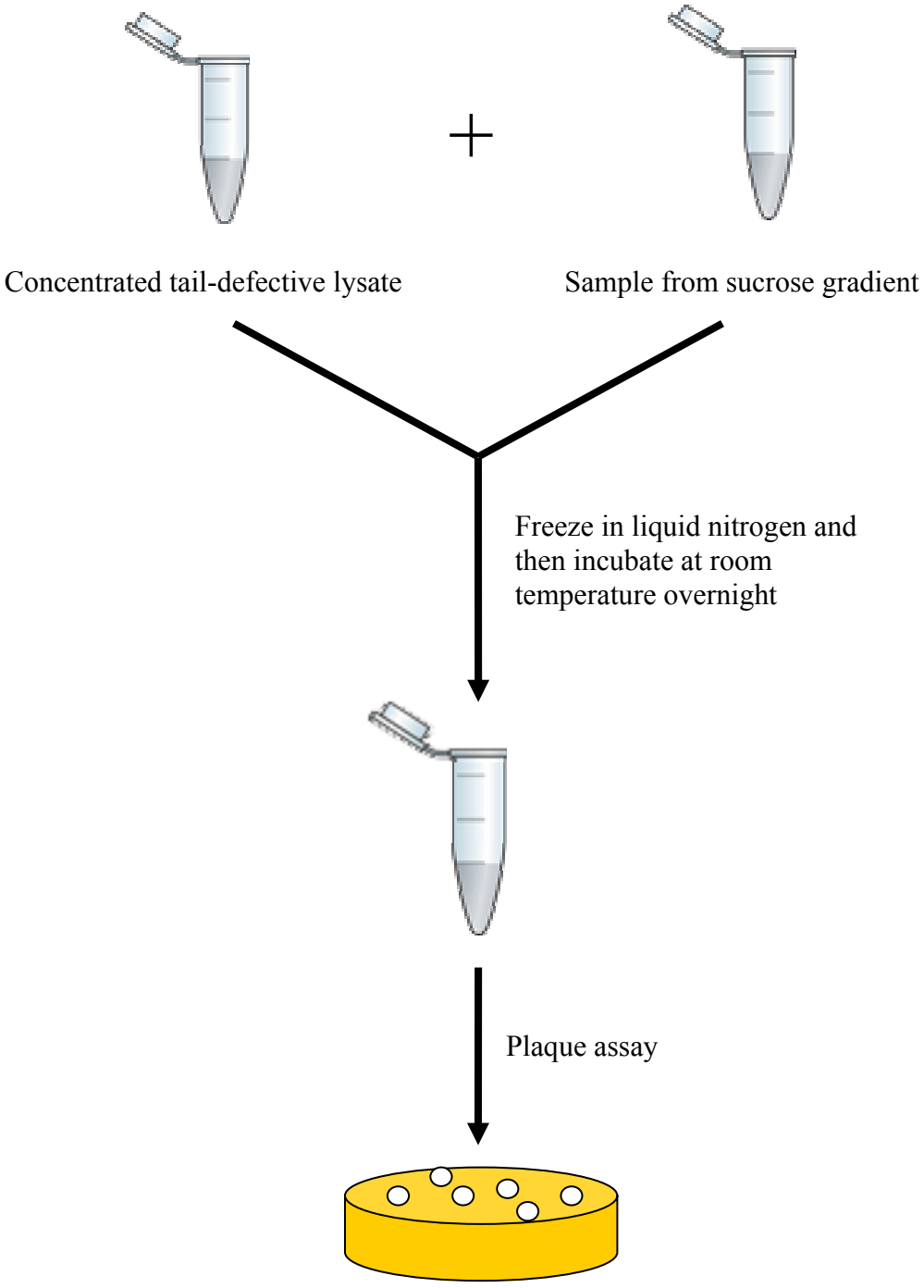

Plasmid
No IPTG
1 mM IPTG
Phage
pET21+
cI857 wild type
L756 no gpL
(a)
pGtail
cI857 wild type
L756 no gpL
(b)
pGtail-LT7
cI857 wild type
L756 no gpL
(c)

Figure 4.3 Schematic diagram of the *in vitro* complementation assay.

A lysogen which contains lambda phage with an amber mutation in one of the tail genes is used to make concentrated tail-defective lysate. A small sample is taken from each of the sucrose gradient fractions and added to the tail-defective lysate. The mixture is frozen in liquid nitrogen for 10 seconds and then sits at room temperature overnight. Subsequently, the mixture is diluted by λ dil containing 10 μg/ml DNase to stop the reaction. Plaque assay is performed to determine the phage titer in the mixture. If tail proteins in the fractions from the gradient are able to complement the tail-defective lysate, an increase in the phage titer over the background titer will be observed. The complementation efficiency (C.E.) is expressed as the ratio of complementation phage titer to the background phage titer. The complementation efficiency is concentration dependent. In order to get better complementation efficiency, the tail proteins in both samples have to be very concentrated. In addition, the pH of the samples should be kept above 7 due to the poor *in vitro* complementation activities for most λ tail precursors at pH below 7.


Concentrated tail-defective lysate
Sample from sucrose gradient
Freeze in liquid nitrogen and then incubate at room temperature overnight
Plaque assay

concentrated tail-defective lysate. The lysate is then mixed with the fractions collected from the gradient. The mixture is frozen in liquid nitrogen for a short time to lyse the cells and then incubated at room temperature overnight. Subsequently, λ dil with 10 μg/ml DNase is added to stop the reaction and the mixture is used for plaque assay to determine the phage titer in the reaction. If the tail tip proteins in the fraction are able to complement the tail-defective lysate, an increase in the phage titer over the background phage titer will be observed.

Figure 4.4 shows that the fifth fraction has the highest phage titer and the phage titer for most of the other fractions is close to the background phage titer. Therefore, the majority of the complementation activity is present in the fifth fraction.

(2) Co-immunoprecipitation

This assay is to determine the composition of the material that co-precipitates with gpL- presumably including tail assembly complexes- from each of the fractions from the sucrose gradient. Each fraction from the sucrose gradient was prepared for Co-IP by anti-T7 antibody (see MATERIALS AND METHODS). After this, the samples were prepared for SDS-PAGE and visualized by autoradiography. Figure 4.5 shows that the majority of the tail tip proteins appear near the top of the gradient. Based on the data in Figures 4.4-4.5, it can be concluded that the tail tip proteins in the fraction which contains most of the *in vitro* complementation activity only constitutes a very small fraction of the tail tip proteins in the lysate. Most of the tail tip proteins are near the top of the gradient and are not biologically active in this complementation assay. Besides, the net intensity of the tail tip protein bands in the fraction which has most of the *in vitro* complementation activity is similar to that of the surrounding fractions. This may indicate

Figure 4.4 The majority of the *in vitro* complementation activity appears in the fifth fraction from the sucrose gradient.

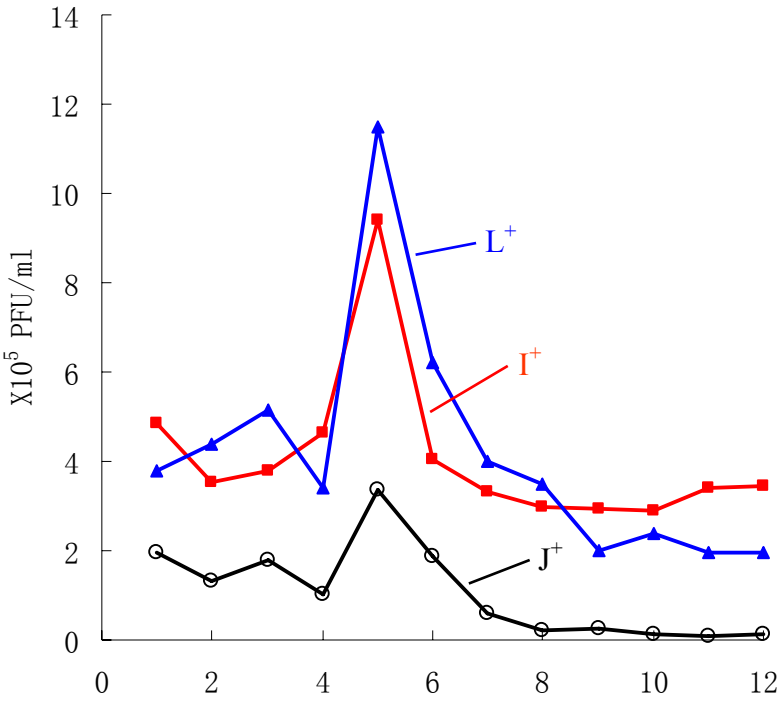
Lysogens 594(λJam442cI857), 594(λIam838cI857), 594(λLam756cI857) and 594(λKam768cI857) were used to make tail-defective lysates. Fractions from sucrose gradient were complemented with the tail-defective lysates. The phage titer in each reaction was examined by plaque assay. The phage titer was plotted against the fraction number. If a fraction is able to complement J^--lysate, it will be said that it has J^+ activity and similarly for I^+ or L^+ activity. Due to the high background phage titer in the K^--lysate (about 10^8 PFU/ml in the concentrated lysate), it is hard to tell if there is an increase in the phage titer in the reaction. So, the complementation activity for K^--lysate is not shown here.

Background phage titer:

J^--lysate: $8X10^3$ PFU/ml

I^--lysate: $4.6X10^5$ PFU/ml

L^--lysate: $3.5X10^5$ PFU/ml



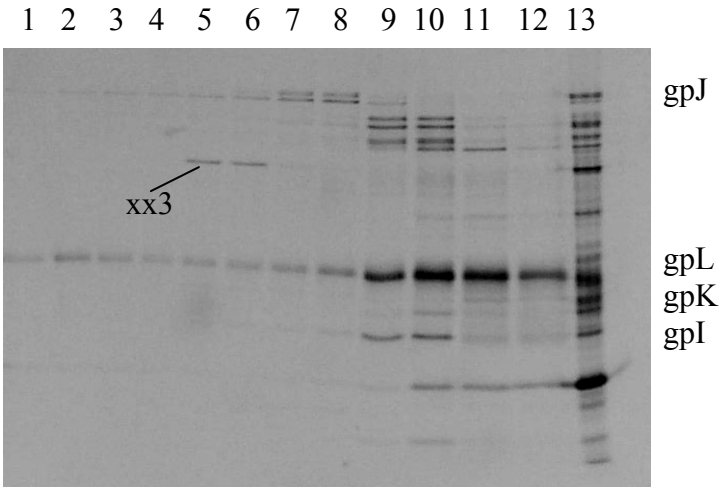
Fractions 1-12: Bottom to top of the gradient

Figure 4.5 Most of the tail proteins are present in the top fractions from the gradient and are not able to complement *in vitro*.

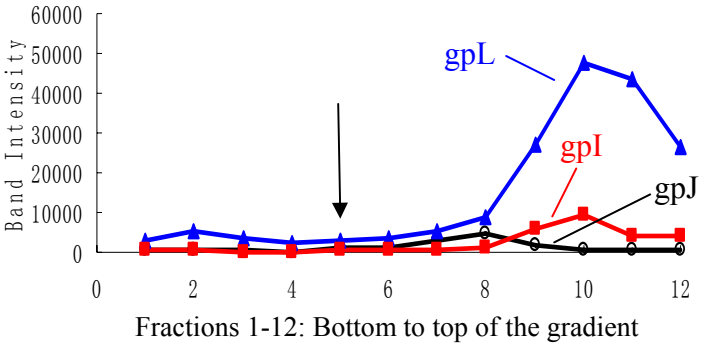
200 μl of each fraction from the gradient was incubated with diluted anti-T7 antibody on a rotating wheel at 4℃ for 2 hours before the addition of the beads. After overnight incubation with the beads, the mixture was washed 3 times with Co-IP buffer and resuspended in 40 μl 1X SDS sample buffer. The sample was boiled and loaded to 15% SDS-PAGE.

(a) Lanes 1-12 are fractions 1-12 from the bottom to the top of the gradient. Each fraction was prepared by Co-IP using anti-T7 antibody. Lane 13 is the sample used for the gradient. There is a very faint of gpK band present in the fifth and sixth fraction. The gpK band is more intensive in the fractions without Co-IP (data not shown).

(b) The film from Panel (a) was scanned and the net intensity of the bands of interest was measured. The net intensity of the bands was plotted against the fraction number. Arrow indicates the fraction that contains most of the *in vitro* complementation activity.


1 2 3 4 5 6 7 8 9 10 11 12 13
xx3
gpJ
gpL
gpK
gpI

(a)


Band Intensity
60000
50000
40000
30000
20000
10000
0
0 2 4 6 8 10 12
gpL
gpI
gpJ
Fractions 1-12: Bottom to top of the gradient

(b)

that only very small amounts of the tail proteins are involved in the *in vitro* complementation assay; this may explain the notorious inefficiency of the *in vitro* complementation assay for the early steps of assembly. In fact, when I tried to examine the copy numbers of the tail tip proteins in the fraction which has *in vitro* complementation activity (fifth fraction), the bands corresponding to gpK and gpI are too weak to be quantified.

In addition to the pGtail-LT7 plasmid, a T7 tag was added to the C-terminus of gpI in the pGtail plasmid, resulting in pGtail-IT7. The recombinant protein complements as well as the wild type protein in the *in vivo* complementation assay (Figure 4.6). The experiments for pGtail-LT7 as described above were done for the pGtail-IT7 plasmid. The result is comparable to that of the pGtail-LT7 plasmid (Figure 4.7).

Band xx3, which migrates about 70 kDa on SDS-PAGE, seems to co-sediment with the tail tip complex, and it is co-precipitated with gpL and with gpI in the experiments shown in Figures 4.5 and 4.7. This experiment was repeated and fractions 4-6 were prepared for Co-IP using either anti-DnaK or anti-groEL antibody. Figure 4.8 shows that band xx3 is present in the Co-IP samples using anti-DnaK antibody, not in those using anti-groEL antibody. This suggests that band xx3 is DnaK. gpK seems to be co-precipitated with DnaK. A very faint gpJ band can be found on the original film in lanes 4-9, but it is too faint to be seen in this figure. The presence of gpL or gpI is uncertain due to overloading of samples in lanes 10-11. Taken together, these results agree that DnaK is a component of a complex of tail tip proteins that co-sediments with the *in vitro* complementation activity of the tail tip proteins. This leads to a suggestion that DnaK has a role in tail tip assembly.

Figure 4.6 The T7-tagged gpI is biologically active.

The plasmids of interest were co-transformed with pLysS into BL21(DE3)-Δtail. Each plasmid was tested for its ability to complement λcI857 (cI857) and λIam838cI857 (I838). No IPTG or 1 mM IPTG is present in the top soft agar. The complementation assay was performed as described (see MATERIALS AND METHODS).

(a) pET21+ is an empty vector and produces no tail proteins.

(b) pGtail encodes the first 8 proteins involved in lambda tail assembly.

(c) pGtail-IT7 differs from pGtail in that it encodes gpI with a T7 tag at its C-terminus.

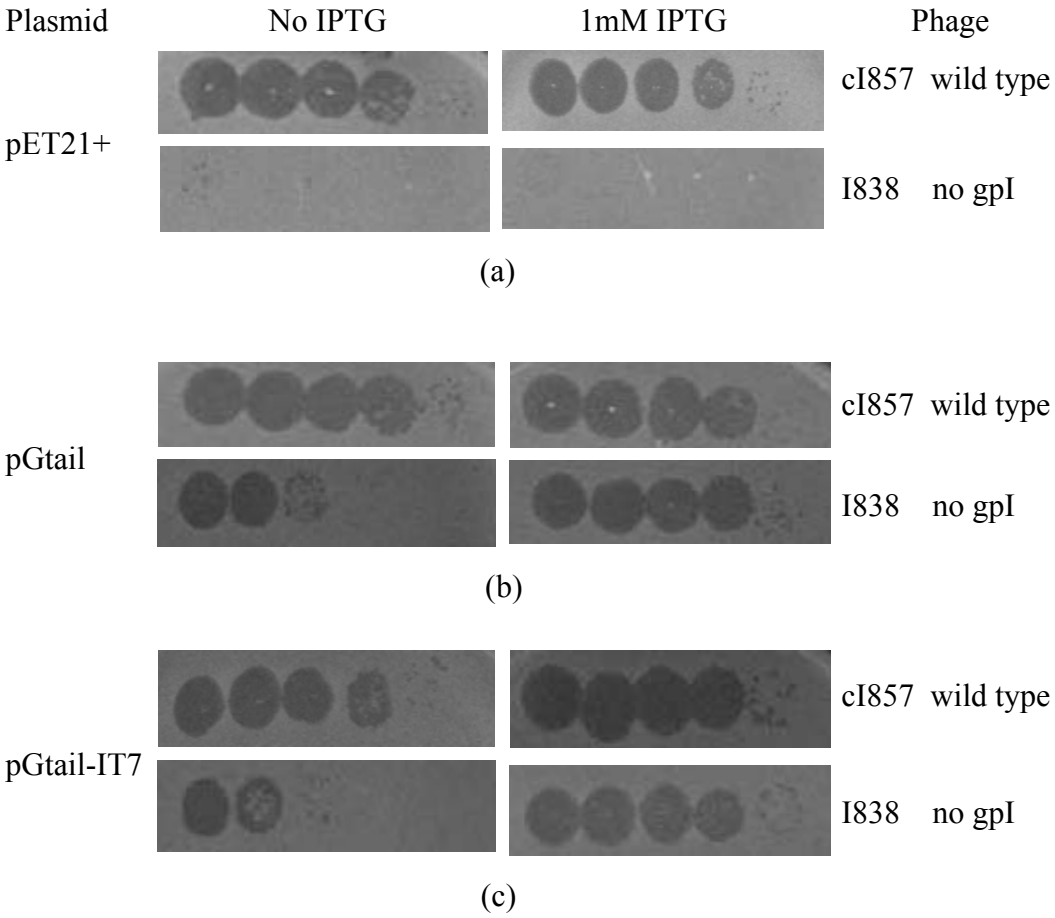

Plasmid
No IPTG
1mM IPTG
Phage
pET21+
cI857 wild type
I838 no gpI
(a)
pGtail
cI857 wild type
I838 no gpI
(b)
pGtail-IT7
cI857 wild type
I838 no gpI
(c)

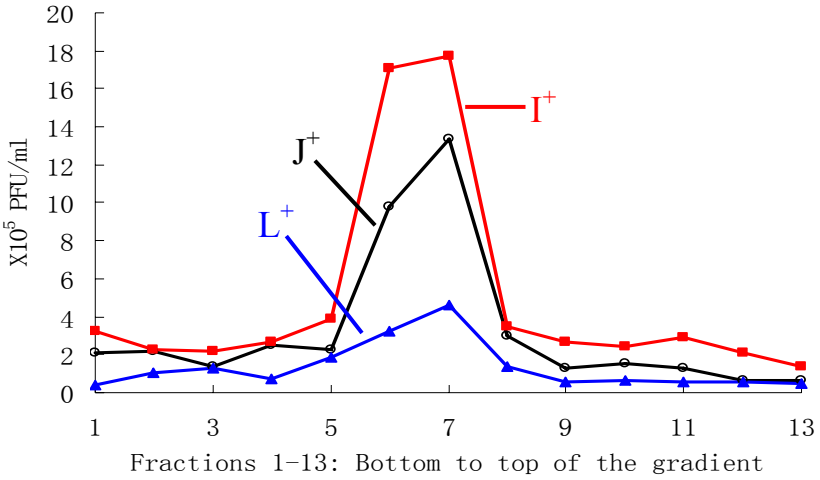
Figure 4.7 The fraction with most of the *in vitro* complementation activity does not overlap with the fraction containing most of the tail tip proteins.

The experimental procedure, including sucrose gradient, *in vitro* complementation assay and Co-IP, for the pGtail-IT7 lysate is the same as that for the pGtail-LT7. About 13 fractions were collected from the gradient.

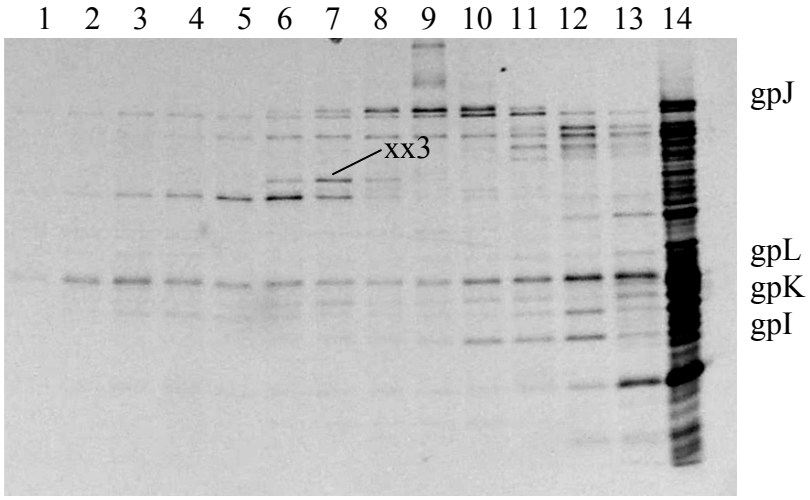
(a) Each fraction from the gradient was assayed for its ability to complement J^--, I^--, L^-- and K^--lysates. The phage titer was plotted against the fraction number. The complementation activity of K^--lysate is not shown here due to the high background phage titer in this lysate. Background phage titer: J^--lysate: $2.5X10^4$ PFU/ml; I^--lysate: $2.3X10^5$ PFU/ml; L^--lysate: $1.0X10^5$ PFU/ml.

(b) Lanes 1-13 are samples from the bottom to the top of the gradient. Each fraction was prepared by Co-IP using anti-T7 antibody. Lane 14 is the sample used for the gradient.

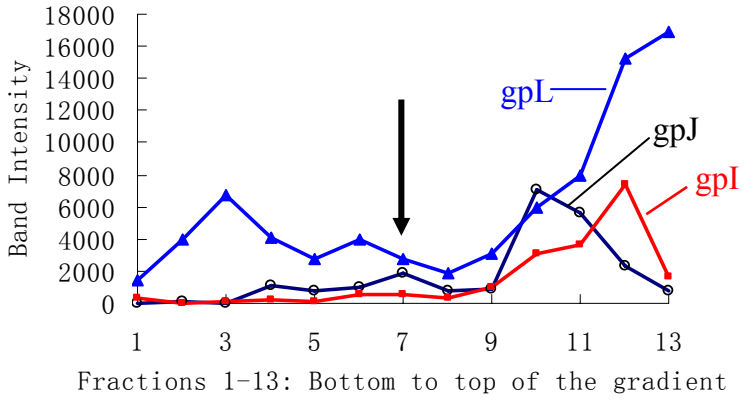
(c) The film from Panel (b) was scanned and bands of interests were quantified. The net intensity of the bands was plotted against the fraction number. Arrow shows the fraction with most *in vitro* complementation activity.


20
18
16
14
12
10
8
6
4
2
0
X10^5 PFU/ml
I+
J+
L+
1
3
5
7
9
11
13
Fractions 1-13: Bottom to top of the gradient

(a)


1 2 3 4 5 6 7 8 9 10 11 12 13 14
xx3
gpJ
gpL
gpK
gpI

(b)


18000
16000
14000
12000
10000
8000
6000
4000
2000
0
Band Intensity
gpL
gpJ
gpI
1
3
5
7
9
11
13
Fractions 1-13: Bottom to top of the gradient

(c)

Figure 4.8 Band xx3 is DnaK.

Tail proteins expressed from plasmid pGtail-LT7 were radiolabeled and prepared for sucrose gradient as described previously. After centrifugation, the gradient was collected into 12 fractions. Fractions 4-6 were used for co-immunoprecipitation using anti-DnaK or anti-groEL antibody. Subsequently, the samples were prepared for 15% SDS-PAGE and visualized by autoradiography.

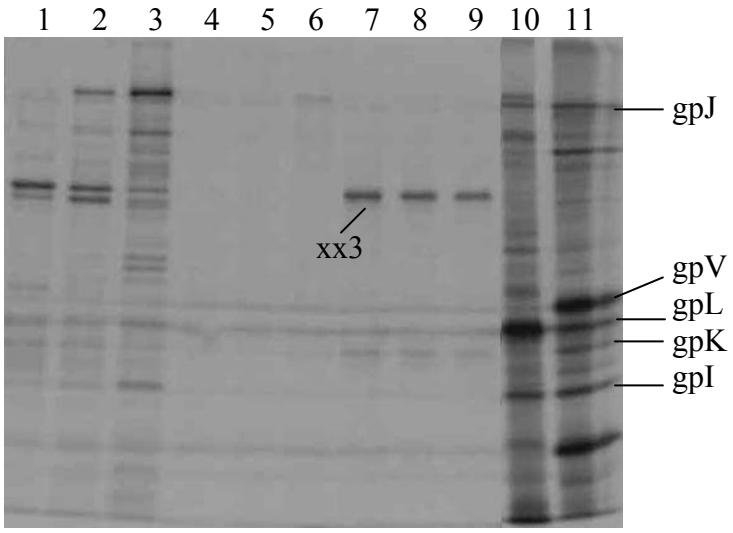
Lanes 1-3: Fractions 4-6 from the sucrose gradient without Co-IP.

Lanes 4-6: Fractions 4-6 from the sucrose gradient prepared by Co-IP using anti-groEL antibody.

Lanes 7-9: Fractions 4-6 from the sucrose gradient prepared by Co-IP using anti-DnaK antibody.

Lane 10: Lysate prepared from pKam6 (contains λ tail genes *J, I* and *L*) was used as marker.

Lane 11: Lysate prepared from pT7-5-Vtail (contains λ tail genes *V-J*) was used as as marker.


1 2 3 4 5 6 7 8 9 10 11
xx3
gpJ
gpV
gpL
gpK
gpI

4.4 SEPARATION OF λ TAIL BY GLYCEROL GRADIENT

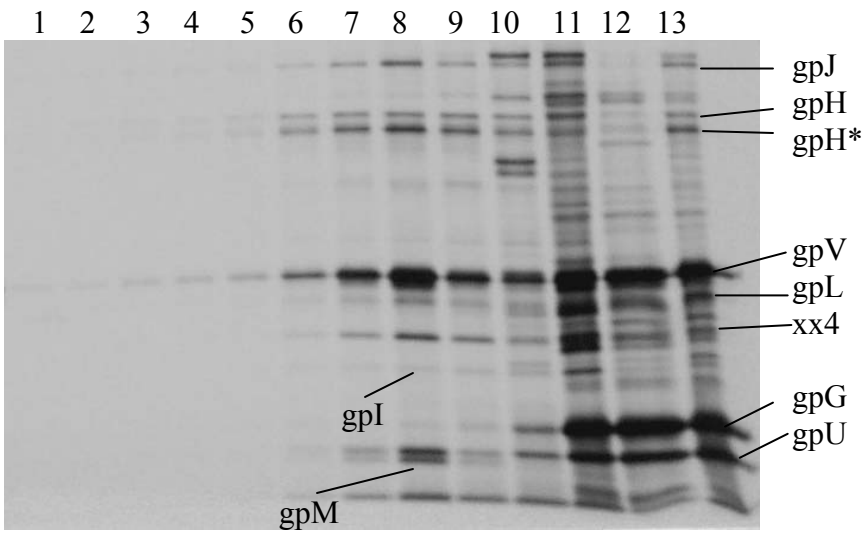
The stoichiometries of the tail tip proteins determined in Section 6.2 are variable and subject to uncertainties about whether the complexes being assayed accurately reflect the composition of either the tail assembly intermediates or the mature tail. Therefore, I decided to determine the stoichiometries of the tail tip proteins in the purified tail. The pEtail4 plasmid encodes all 11 proteins required for lambda tail assembly. Tail proteins expressed from the plasmid were radiolabeled and prepared for velocity sedimentation in a 10-30% glycerol gradient as described in MATERIALS AND METHODS. After centrifugation, the tube was punctured at the bottom and about 12 fractions were collected. TCA precipitation was carried out to concentrate the sample. After TCA precipitation, each fraction was prepared for SDS-PAGE followed by autoradiography and densitometry. Figure 4.9 shows that there is a peak at the eighth fraction for tail proteins which are present in the mature structure. The amount of gpG increases from the bottom to the top of the gradient because gpG acts as a chaperone and is not present in the mature tail. Therefore, the majority of the tail is present in the eighth fraction. Table 4.3 is the quantification from the eighth fraction. The copies of gpV are a little lower than what has been reported probably due to the dramatic difference in the intensity of the two bands, gpJ and gpV. The copies of gpL are consistent with what has been shown in Table 4.1. The band intensity of gpI and xx4, an unknown band with a mobility corresponding to about 25 kDa, seems to vary greatly from individual experiments. In one extreme example, the intensity of the xx4 band is almost 1/3-1/2 as intensive as gpV in the glycerol-gradient purified tail prepared from pEtail4-7S (This plasmid differs from pEtail4 in that it carries a non-conserved cysteine to serine mutation in gpL. The detailed information about the plasmid is described in Chapter 5), and it is very hard

Figure 4.9 Separation of lambda tail by glycerol gradient.

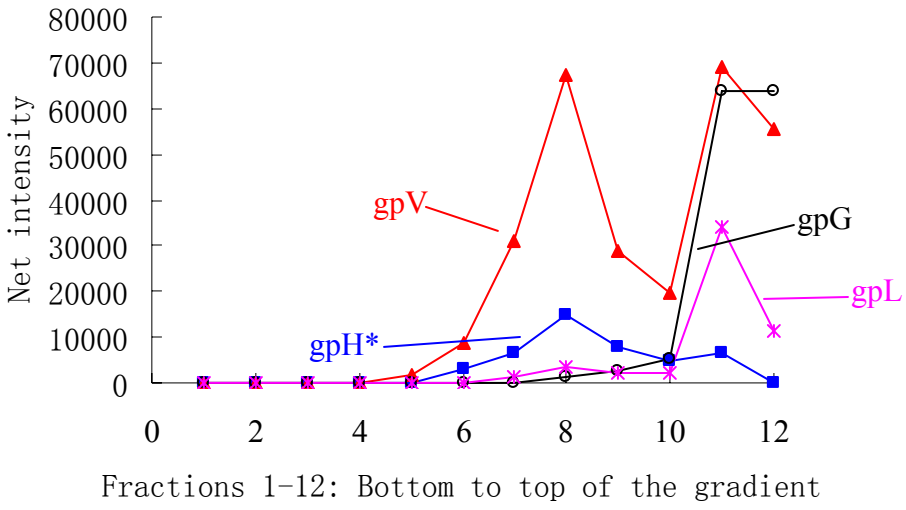
The pEtail4 plasmid encodes all 11 proteins required for lambda tail assembly. Tail proteins expressed from this plasmid were labeled by ^{35}S-Met and prepared for 10-30% glycerol gradient. After centrifugation at 40,000 RPM for 3.5 hours at 4℃, the gradient tube was punctured at the bottom and fractions collected. The fractions were then prepared for 15% SDS-PAGE and visualized by autoradiography. In this experiment, proteins / protein complexes will go to different positions of the gradient according to their sedimentation coefficient. Most unassembled proteins will remain in the top of the gradient.

(a) Lanes 1-12 are fractions 1-12 from the bottom to the top of the gradient. The amount of the sample loaded onto lane 12 is 10 times less than those of lanes 1-11. Lane 13 is the sample used for the gradient

(b) The net intensity of some tail protein bands was quantified. The net intensity of the bands of interest was plotted against the fraction number.


1 2 3 4 5 6 7 8 9 10 11 12 13
gpJ
gpH
gpH*
gpV
gpL
xx4
gpG
gpU
gpI
gpM

(a)


Net intensity
80000
70000
60000
50000
40000
30000
20000
10000
0
0 2 4 6 8 10 12
gpV
gpG
gpL
gpH*
Fractions 1-12: Bottom to top of the gradient

(b)

Table 4.3 Quantification of λ tail proteins from tails purified by glycerol gradient.

The experiment was repeated twice. The fraction which contains most of the tail was loaded to 22.5% SDS-PAGE and several different lengths of exposure were made. The bands were visualized by autoradiography and examined by densitometry. The copy numbers of the tail proteins were normalized to 3 copies of gpJ.

	Copies of λ tail proteins
gpJ	**3**
gpI	1-3
gpL	3-4
gpH	1-3
gpH*	5-8
gpV	153-177

to detect gpL and gpI bands in this case. It is not clear whether xx4 is a processed or degradation product of one of the tail proteins.

4.5 PURIFICATION OF λ PHAGE BY CESIUM CHLORIDE STEP GRADIENT

gpH, the unprocessed tape measure protein, is not present in the mature tail. But a small amount of gpH is present in the eighth fraction (Figure 4.9). This suggests that the purified tail is a mixture of the mature tail and partially assembled tail. In order to solve this problem, I tried to determine the stoichiometries of the tail tip proteins from purified lambda phage. In this experiment, the tail proteins expressed from pEtail4 were radiolabeled and the cells were lysed as described in MATERIALS AND METHODS. The 594(λVam750Sam7cI857) lysogen produces λ head without tail attached. This lysogen was used for head preparation (see MATERIALS AND METHODS). Excess amount of the purified head were mixed with the radiolabeled tail for the head-tail joining reaction (see MATERIALS AND METHODS). After this, the phage mixture was loaded onto a cesium chloride step gradient to separate lambda phage virions. The tube was punctured at the bottom and about one third of the gradient from the bottom was collected after centrifugation. The samples were prepared for SDS-PAGE followed by autoradiography and densitometry. Each fraction from the gradient was also examined by plaque assay to determine the phage titer. The result is shown in Figure 4.10 and Table 4.4. Most of the phage is present in the fourth fraction. Fractions 8-13 contain most of the other components from the phage mixture, including a small amount of λ phage, tail assembly intermediates, unassembled phage proteins and so on. With elongated exposure time, several faint bands can be observed (Panel b). gpL behaves a little strangely. The band looks a little

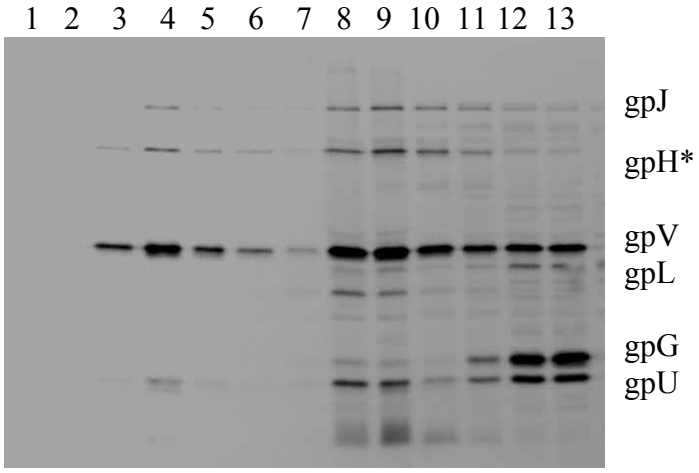
Figure 4.10 Purification of lambda phage by cesium chloride step gradient.

λ tail proteins expressed from the pEtail4 plasmid were radiolabeled by ^{35}S-Met. The 594(λVam750Sam7cI857) lysogen was used asa donor of non-radioactive heads. The head and tail were prepared for the head-tail joining reaction as described in MATERIALS AND METHODS. Subsequently, the phage mixture was loaded onto the top of two layers of cesium chloride solutions with different density (1.4 g/cm^3 and 1.6 g/cm^3). After centrifugation, the tube was punctured at the bottom and about 1/3 from the bottom of the gradient tube was collected.

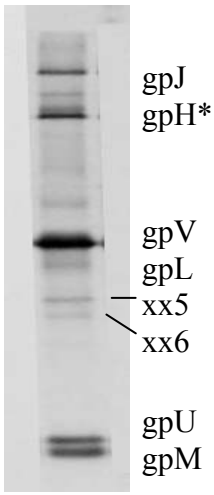
(a) Each fraction from the gradient was prepared for 15% SDS-PAGE and visualized by autoradiography. Lanes 1-13: Fractions 1-13 collected from the bottom to about 1/3 of the cesium chloride gradient tube. The film was exposed to the dried SDS gel for 1 day.

(b) The fourth fraction was loaded to 20% SDS-PAGE and the film was exposed to the SDS gel for 14 days in order to see the faint bands.

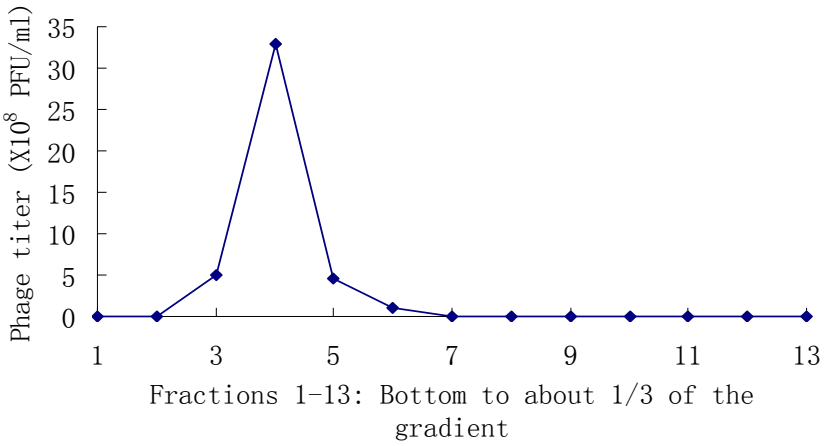
(c) Each fraction from the gradient tube was examined by plaque assay to determine the phage titer. The base-10 logarithm of the phage titer in each fraction was plotted against the fraction number.


1 2 3 4 5 6 7 8 9 10 11 12 13
gpJ
gpH*
gpV
gpL
gpG
gpU

(a)


gpJ
gpH*
gpV
gpL
xx5
xx6
gpU
gpM

(b)


Phage titer (X10^8 PFU/ml)
35
30
25
20
15
10
5
0
1
3
5
7
9
11
13
Fractions 1-13: Bottom to about 1/3 of the gradient

(c)

Table 4.4 Quantification of lambda tail proteins from purified phage.

The fourth fraction from the gradient was loaded onto 22.5% SDS-PAGE. The film was exposed to the gel for different times and visualized by autoradiography. The bands of interest were quantified. The copy numbers of gpH* and gpV were normalized to 3 copies of gpJ. These numbers are close to the previous studies. The copy numbers of gpL and gpM were normalized to 6 copies of gpU. The gpL band is very weak and looks a little fuzzy. Therefore, the quantification for this protein may not be accurate.

	Copies of λ tail proteins in purified phage
gpJ	**3**
gpI	
gpL	~1
gpM	8
gpH*	6
gpV	178
gpU	**6**

fuzzy possibly due to the presence of 8 cysteine residues. No band is found in the position corresponding to gpI. But, a weak band, xx6, which migrates a little slower than gpI is found. Another band (xx5) which appears between gpI and gpL, is probably the same one as xx4 found in Figure 4.9. So far, it is not clear what these two bands are.

4.6 DISCUSSION

4.6.1 gpK may be loosely associated with the complex

gpK presents a conundrum because it is required for lambda tail assembly, but it is not found in the mature structure. The presence of gpK-like band in the sample prepared from the pKam6 plasmid indicates that it may be a background band which migrates about the same position as lambda gpK (Figure 4.1). This suggests that gpK may be loosely associated with the complex. There are two experiments which support this hypothesis. In one experiment, 6XHis tag was added to the C-terminus of gpI encoded from pLtail. The modified gpI can complement as well as the wildtype protein in the *in vivo* complementation assay. Radiolabeled proteins from this plasmid were loaded onto a nickel column. When the immidazole concentration in the wash buffer was increased from 90 mM to 120 mM, there is a significant decrease in the amount of gpK found in the elution fractions. The fact that K^+ activity can be found in late assembly intermediates such as V^--lysate and U^--lysate (Katsura and Kuhl, 1975) also supports the idea that it may be loosely associated with the early assembly intermediates.

4.6.2 Estimation of the stoichiometries of the tail tip proteins

Sections 4.2-4.5 show the examination of lambda tail tip proteins (gpI, gpL and gpK) from different preparations. There are advantages and disadvantages for each method. The strength of the technique using *E. coli* cells to precipitate lambda tail tip assembly intermediates is that gpJ and anything associated with it will be precipitated. This may also be a drawback because the stoichiometries of the tail tip assembly intermediates will be distorted due to the over-representation of proteins added early in the assembly pathway. Another drawback to this technique is that the activity of the assembly intermediates can not be determined. In order to solve this problem, sucrose gradient centrifugation followed by *in vitro* complementation and Co-IP were performed. The main advantage of this method is that the biological activity of the tail assembly intermediates can be examined by an *in vitro* complementation assay. Like the technique using *E. coli* cells to precipitate assembly intermediates, the stoichiometries of the tail tip proteins are distorted in this method. This is because the tail tip and the tail tip assembly intermediates have the same sedimentation coefficient and therefore they will sediment to the same position in the gradient. This disadvantage also applies to the method using glycerol gradient to purify lambda tail. The only method that does not have this problem is the one using cesium chloride step gradient to purify lambda phage. Phages are separated by density not sedimentation as in a sucrose or glycerol gradient. Desipte the caveats listed above, the approximate consistency among the different methods for the amounts measured for most of the proteins suggests that these numbers are likely to be relatively accurate.

Table 4.5 is a summary of the copy numbers of the tail proteins from Tables 4.1, 4.3 and 4.4. The lambda tail is thought to be three-fold symmetric and the copy numbers of the tail proteins are

Table 4.5 Summary of the copy numbers of tail proteins from different preparations.

The following table is a summary of Tables 4.1, 4.3 and 4.4. The copies of other proteins were normalized to 3 copies of gpJ expect for the copies of gpL and gpM in purified phage, which were normalized to 6 copies of gpU.

	pIam	pKam6	pLdel	pLtail	λ tail	λ phage
gpJ	**3**	**3**	**3**	**3**	**3**	**3**
gpI		3-5	3-4	2-3	1-3	
gpL	~1	2-3		2-4	3-4	~1
gpK				<0.6		
gpH					1-3	
gpH*					5-8	6
gpV					153-177	178
gpU						**6**
gpM						8

therefore thought likely to be a multiple of 3 or 6. If this is true, the best estimation of the stoichiometries of the tail proteins from my studies is that there are probably 3 copies of gpL and 3 copies of gpI involved in each tail assembly pathway. The copy number of gpK required for the assembly remains unknown due to the fact that it was not detected in the various structures examined (except in the case of DnaK, see below) possibly because it may be loosely associated with the early assembly intermediates. The absence of gpK in the mature virion is consistent with previous studies. So far, it is not clear how many copies of gpL are present in the mature virion. The gpL band appears a little fuzzy on SDS-PAGE. The same phenomenon has been observed in other sample preparations. In most cases, this seems to happen after the sample has been kept in 4℃ or -20℃ for several days. But this is not the case for the gpL in the mature virion. The gpL band looks a little fuzzy even when the sample is fresh. This may be an artifact of gel electrophoresis due to the presence of 8 cysteine residues in gpL.

In addition to gpL, the copy number of gpI in the mature phage remains an open question. We know that gpI is required for lambda tail assembly and that it is present in the tail which is not attached to the head. No gpI-like band is found in the mature virion. This is not consistent with previous studies. One possibility is that the gpI band is not well established in the previous experiment. The band observed in the previous studies which is thought to be gpI may be the band which migrates a little slower than gpI as I found in the cesium chloride purified phage.

4.6.3 DnaK and the biologically active assembly intermediates

Attempts to determine the protein composition and stoichiometry for the biologically active assembly intermediates were thwarted by the fact that the position of the sucrose gradients with

in vitro complementation activity had almost no detectable amounts of the tail proteins after precipitation with antibodies that bring down gpI or gpL. This was unexpected because in principle these intermediates are the same as the ones that were precipitated with *E. coli* cells, which had substantial amounts of gpJ, gpI and gpL. Possibly the assembly complexes are unstable in the sucrose gradient, which could also explain the long-established fact that the level of *in vitro* complementation from such gradients is extremely low.

A potentially important result from these experiments is that a band identified as DnaK was co-precipiated by antibodies against immunotags on gpI and gpL. The DnaK detected in this way co-migrated on the sucrose gradients with the biological activity of the tail proteins. Although pre-immune serum was not available for this commercial antibody for a negative control, the DnaK band was not precipitated by the groEL antibody, suggested that the DnaK was indeed being precipitated as a result of forming a complex with the immunotagged tail proteins. If this is correct, it suggests that DnaK, which has well-established roles in protein folding and assembly / disassembly, has a role in production of lambda tails. It is interesting that one protein that co-precipitates with anti-DnaK antibodies has the mobility of gpK, which has not been seen in any of the other attempts to visualize tail assembly complexes.

A puzzling feature of these data is that DnaK is only co-precipitated by the tail tip complex separated by sucrose velocity gradient, but not by *E. coli* cells. One possible interpretation is that DnaK is released after attachment of the tail tip complex to *E. coli* LamB receptors.

5.0 CHARACTERIZATION OF THE CYSTEINE RESIDUES IN λ gpL

5.1 INTRODUCTION

Many proteins contain cysteine residues, some of which may play a key role in the protein folding pathway. In spite of the fact that the reducing environment of the cytoplasm of bacteria is generally believed to be unfavorable for disulfide bond formation, there has been some evidence indicating that some cysteines in bacteriophages carry out important roles in protein formation or phage assembly. The tailspike protein of *Salmonella* phage P22, for example, has 8 cysteine residues among its 666 amino acids. Although no disulfide bonds have been found in the native state of the tailspike protein, there is a transient interchain disulfide bond formation in the protrimer intermediates which is required for the correct folding and assembly of the tailspike (Sauer et al., 1982; Sargent et al., 1988; Sather and King, 1994; Steinbacher et al., 1994; Robinson and King, 1997). Cysteines can have other roles in proteins as well. A common example is coordinating metal ions, as in zinc fingers (Berg, 1990).

The λ tail tip protein, gpL, has 8 cysteines among its 232 amino acids. According to multi-sequence alignments, 4 of them are perfectly conserved among λ gpL homologs (Figure 5.1). The conserved cysteine residues may be essential for correct protein folding or tail assembly or for functioning of the assembled tail. In order to characterize the cysteine residues, I used

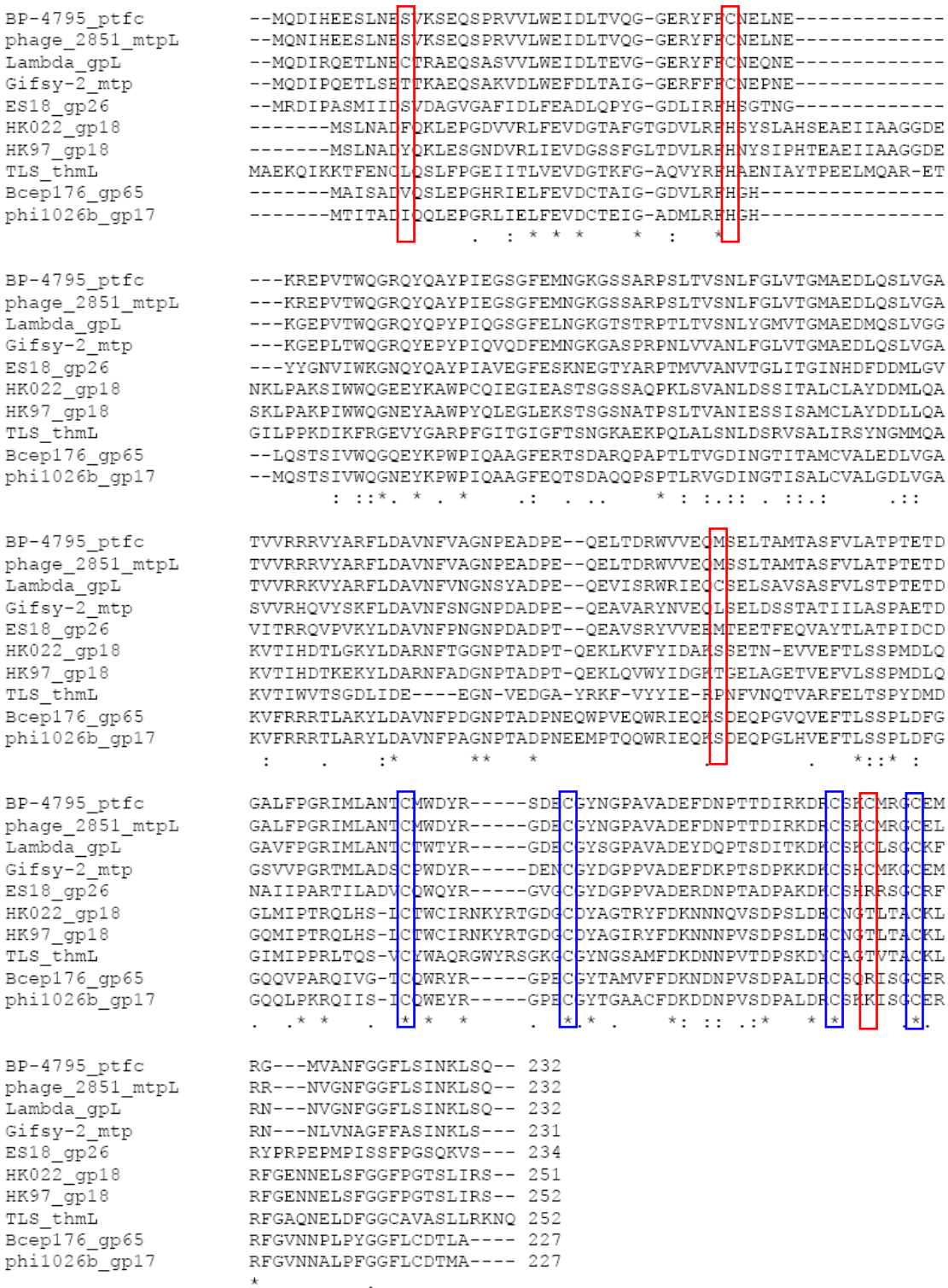
Figure 5.1 Multi-sequence alignments of phage λ gpL and its homologs.

Lambda tail tip protein, gpL, contains 8 cysteines among its 232 amino acids. ClustalW2 was used to produce multiple sequence alignments among homologous sequences identified by a BlastP search. Each protein in the examined group contains 5-8 cysteines. The result shows that the cysteine residues at positions 173, 182, 205 and 212 are perfectly conserved among lambda gpL and the examined group of its homologs while the cysteine residues at positions 12, 39, 140 and 208 are non-conserved. The non-conserved cysteine residues are highlighted by red rectangles whereas the conserved cysteine residues are highlighted by blue rectangles. On the left side of the figure, the word before the underline shows the name of phage whereas the word after the underline is the λ gpL homolog in this phage. For instance, ES18_gp26 means that in phage ES18, gp26 is a homolog of λ gpL.

ptfc: putative tail fiber component

mtpL: minor tail protein L

mtp: minor tail protein


BP-4795_ptfc --MQDIHEESLNESVKSEQSPRVVLWEIDLTVQG-GERYFFCNELNE----------
phage_2851_mtpL --MQNIHEESLNESVKSEQSPRVVLWEIDLTVQG-GERYFFCNELNE----------
Lambda_gpL --MQDIRQETLNECTRAEQSASVVLWEIDLTEVG-GERYFFCNEQNE----------
Gifsy-2_mtp --MQDIPQETLSETTKAEQSAKVDLWEFDLTAIG-GERFFFCNEPNE----------
ES18_gp26 --MRDIPASMIIDSVDAGVGAFIDLFEADLQPYG-GDLIRFHSGTNG----------
HK022_gp18 -------MSLNADFQKLEPGDVVRLFEVDGTAFGTGDVLRFHSYSLAHSEAEIIAAGGDE
HK97_gp18 -------MSLNADYQKLESGNDVRLIEVDGSSFGLTDVLRFHNYSIPHTEAEIIAAGGDE
TLS_thmL MAEKQIKKTFENCLQSLFPGEIITLVEVDGTKFG-AQVYRFHAENIAYTPEELMQAR-ET
Bcep176_gp65 -------MAISADVQSLEPGHRIELFEVDCTAIG-GDVLRFHGH----------
phi1026b_gp17 -------MTITADIQQLEPGRLIELFEVDCTEIG-ADMLRFHGH----------
BP-4795_ptfc ---KREPVTWQGRQYQAYPIEGSGFEMNGKGSSARPSLTVSNLFGLVTGMAEDLQSLVGA
phage_2851_mtpL ---KREPVTWQGRQYQAYPIEGSGFEMNGKGSSARPSLTVSNLFGLVTGMAEDLQSLVGA
Lambda_gpL ---KGEPVTWQGRQYQPYPIQGSGFELNGKGTSTRPTLTVSNLYGMVTGMAEDMQSLVGG
Gifsy-2_mtp ---KGEPLTWQGRQYEPYPIQVQDFEMNGKGASPRPNLVVANLFGLVTGMAEDLQSLVGA
ES18_gp26 ---YYGNVIWKGNQYQAYPIAVEGFESKNEGTYARPTMVVANVTGLITGINHDFDDMLGV
HK022_gp18 NKLPAKSIWWQGEEYKAWPCQIEGIEASTSGSSAQPKLSVANLDSSITALCLAYDDMLQA
HK97_gp18 SKLPAKPIWWQGNEYAAWPYQLEGLEKSTSGSNATPSLTVANIESSISAMCLAYDDLLQA
TLS_thmL GILPPKDIKFRGEVYGARPFGITGIGFTSNGKAEKPQLALSNLDSRVSALIRSYNGMMQA
Bcep176_gp65 --LQSTSIVWQGQEYKPWPIQAAGFERTSDARQPAPTLTVGDINGTITAMCVALEDLVGA
phi1026b_gp17 --MQSTSIVWQGNEYKPWPIQAAGFEQTSDAQQPSPTLRVGDINGTISALCVALGDLVGA
BP-4795_ptfc TVVRRRVYARFLDAVNFVAGNPEADPE--QELTDRWVVEQMSELTAMTASFVLATPTETD
phage_2851_mtpL TVVRRRVYARFLDAVNFVAGNPEADPE--QELTDRWVVEQMSSLTAMTASFVLATPTETD
Lambda_gpL TVVRRKVYARFLDAVNFVNGNSYADPE--QEVISRWRIEQCSELSAVSASFVLSTPTETD
Gifsy-2_mtp SVVRHQVYSKFLDAVNFSNGNPDADPE--QEAVARYNVEQLSELDSSTATIILASPAETD
ES18_gp26 VITRRQVPVKYLDAVNFPNGNPDADPT--QEAVSRYVVEEMTEETFEQVAYTLATPIDCD
HK022_gp18 KVTIHDTLGKYLDARNFTGGNPTADPT-QEKLKVFYIDAKSSETN-EVVEFTLSSPMDLQ
HK97_gp18 KVTIHDTKEKYLDARNFADGNPTADPT-QEKLQVWYIDGKTGELAGETVEFVLSSPMDLQ
TLS_thmL KVTIWVTSGDLIDE----EGN-VEDGA-YRKF-VYYIE-RPNFVNQTVARFELTSPYDMD
Bcep176_gp65 KVFRRRTLAKYLDAVNFPDGNPTADPNEQWPVEQWRIEQKSDEQPGVQVEFTLSSPLDFG
phi1026b_gp17 KVFRRRTLARYLDAVNFPAGNPTADPNEEMPTQQWRIEQKSDEQPGLHVEFTLSSPLDFG
BP-4795_ptfc GALFPGRIMLANTCMWDYR-----SDECGYNGPAVADEFDNPTTDIRKDRCSKCMRGCEM
phage_2851_mtpL GALFPGRIMLANTCMWDYR-----GDECGYNGPAVADEFDNPTTDIRKDRCSKCMRGCEL
Lambda_gpL GAVFPGRIMLANTCTWTYR-----GDECGYSGPAVADEYDQPTSDITKDRCSKCLSGCKF
Gifsy-2_mtp GSVVPGRTMLADSCPWDYR-----DENCGYDGPPVADEFDKPTSDPKKDRCSKCMKGCEM
ES18_gp26 NAIIPARTILADVCQWQYR-----GVGCGYDGPPVADERDNPTADPAKDRCSKRRSGCRF
HK022_gp18 GLMIPTRQLHS-LCTWCIRNKYRTGDGCDYAGTRYFDKNNNQVSDPSLDECNGTLTACKL
HK97_gp18 GQMIPTRQLHS-LCTWCIRNKYRTGDGCDYAGIRYFDKNNNPVSDPSLDECNGTLTACKL
TLS_thmL GIMIPPRLTQS-VCYWAQRGWYRSGKGCGYNGSAMFDKDNNPVTDPSKDYCAGTVTACKL
Bcep176_gp65 GQQVPARQIVG-TCQWRYR-----GPECGYTAMVFFDKNDNPVSDPALDRCSQRISGCER
phi1026b_gp17 GQQLPKRQIIS-ICQWEYR-----GPECGYTGAACFDKDDNPVSDPALDRCSKKISGCER
BP-4795_ptfc RG---MVANFGGFLSINKLSQ-- 232
phage_2851_mtpL RR---NVGNFGGFLSINKLSQ-- 232
Lambda_gpL RN---NVGNFGGFLSINKLSQ-- 232
Gifsy-2_mtp RN---NLVNAGFFASINKLS--- 231
ES18_gp26 RYPRPEPMPISSFPGSQKVS--- 234
HK022_gp18 RFGENNELSFGGFPGTSLIRS-- 251
HK97_gp18 RFGENNELSFGGFPGTSLIRS-- 252
TLS_thmL RFGAQNELDFGGCAVASLLRKNQ 252
Bcep176_gp65 RFGVNNPLPYGGFLCDTLA---- 227
phi1026b_gp17 RFGVNNALPFGGFLCDTMA---- 227

site-directed mutagenesis to mutate individual cysteine residues and then characterized the mutant proteins by a series of assays.

5.2 DETERMINATION OF THE ROLE OF THE CYSTEINE RESIDUES IN λ gpL

5.2.1 Construction of mutation in gene *L*

The pGEX-6p-1-L plasmid encodes gpL with a GST fusion at the N-terminus. This recombinant protein functions as well as the wildtype protein *in vivo*, as judged by the complementation experiments shown in Chapter 3. QuikChange site-directed mutagenesis was used to mutate each of the individual cysteines to serine (see MATERIALS AND METHODS). Table 5.1 is a brief description of the plasmids and the positions of mutations introduced.

5.2.2 Determination of the biological activity of the mutant proteins *in vivo*

Each target plasmid was co-transformed with pLysS into strain BL21(DE3)-Δtail. Spot complementation assay was conducted to examine if the mutation has any effect on the activity of gpL. Figure 5.2 shows that the C12S, C39S and C208S mutants can complement as well as GST-L. Single amino acid substitution at position 140, 173, 182, 205 or 212 abolishes the ability of the mutant protein to complement *in vivo*. According to the multi-sequence alignments, C140 is not conserved among lambda gpL homologs. Since some of its homologs, which show high sequence similarity with lambda gpL, have methionine at the position corresponding to lambda gpL C140, site-directed mutagenesis was performed to construct the C140M mutation. The result

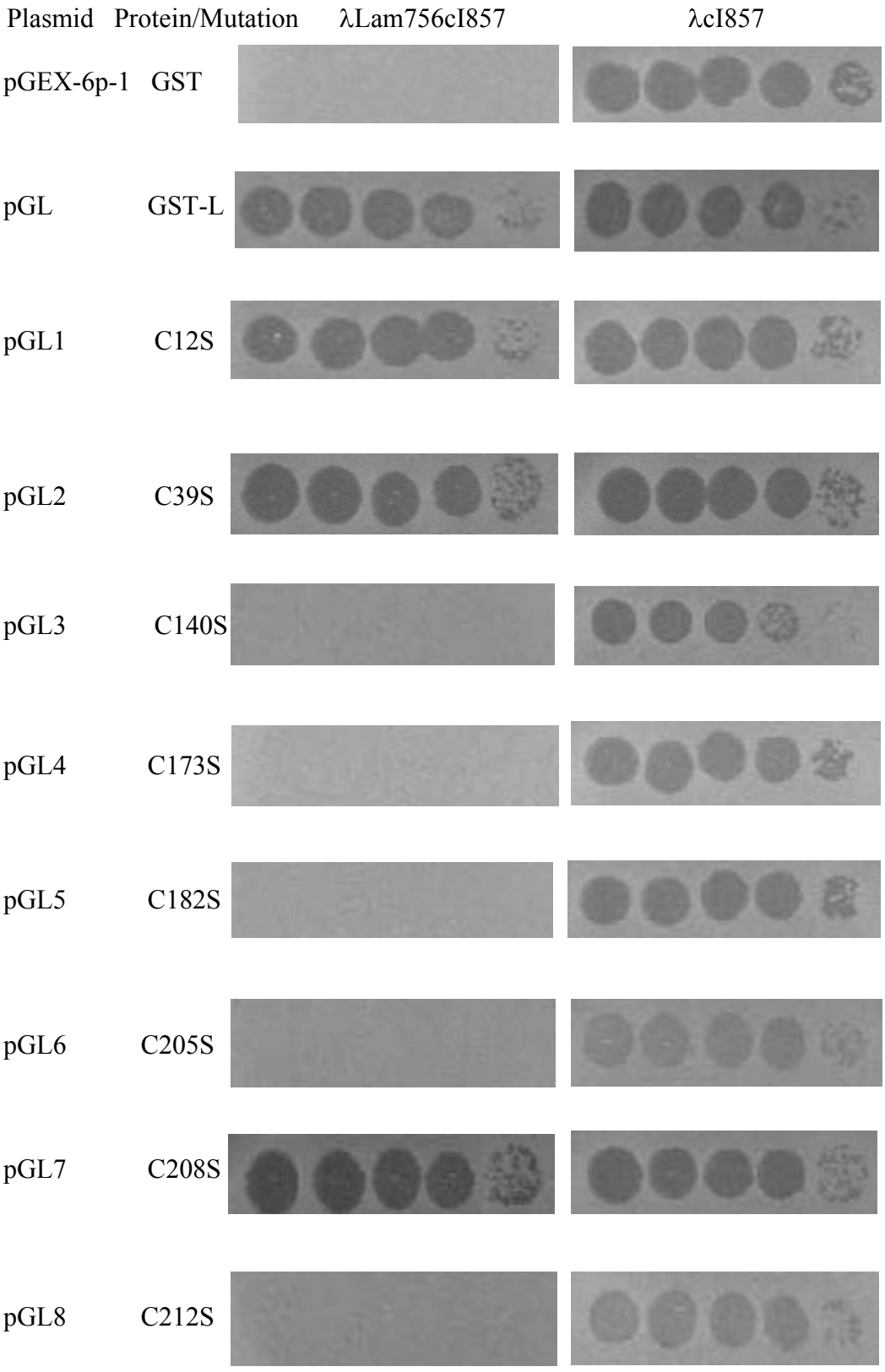
Table 5.1 Description of plasmids and positions of mutation.

All plasmids listed below use pGEX-6p-1 vector which adds a GST fusion at the N-terminus of the target protein. The numbering of the cysteine residues is based on their positions in the wildtype gpL, not GST-L.

Plasmid	Protein
pGEX-6p-1-L (pGL)	GST-L
pGEX-6p-1-L1S (pGL1)	GST-L with C12S mutation
pGEX-6p-1-L2S (pGL2)	GST-L with C39S mutation
pGEX-6p-1-L3S (pGL3)	GST-L with C140S mutation
pGEX-6p-1-L4S (pGL4)	GST-L with C173S mutation
pGEX-6p-1-L5S (pGL5)	GST-L with C182S mutation
pGEX-6p-1-L6S (pGL6)	GST-L with C205S mutation
pGEX-6p-1-L7S (pGL7)	GST-L with C208S mutation
pGEX-6p-1-L8S (pGL8)	GST-L with C212S mutation

Figure 5.2 Determination of the biological activity of the mutant proteins *in vivo*.

Each target plasmid was co-transformed with pLysS into strain BL21(DE3)-△tail. The spot complementation assay was conducted as described previously. Each plasmid was spotted with λcI857 and λLam756cI857. The first column shows the plasmid tested and the second column is the type of the protein or the position of mutation. The third and fourth columns show the type of phage used.


Plasmid
Protein/Mutation
λLam756cI857
λcI857
pGEX-6p-1
GST
pGL
GST-L
pGL1
C12S
pGL2
C39S
pGL3
C140S
pGL4
C173S
pGL5
C182S
pGL6
C205S
pGL7
C208S
pGL8
C212S

from spot complementation assay shows that the C140M mutation has little effect on the biological activity of gpL (Figure 5.3 (a)-(c)).

In addition to the single amino acid substitutions, I also constructed a triple mutant with cysteine residues at positions 12, 39 and 208 changed to serines, and a quadruple mutant which only contains the four conserved cysteines. The triple mutant behaves nearly as well as the wildtype protein while there is a ~10-fold decrease in the activity of the quadruple mutant (Figure 5.3 (d)-(e)). As has been shown in the previous chapters, cysteines in gpL may cause problems during purification and it may contribute to the abnormal behavior of gpL on SDS-PAGE. The triple mutant, which contains only 5 cysteine residues and can function almost as well as the wildtype protein *in vivo*, may be a good candidate for structural studies.

5.2.3 Investigation of the conserved cysteine residues in λ gpL

The previous section shows that the single amino acid substitutions of the conserved cystiene residues result in the inability of gpL to complement *in vivo*. This suggests that these cysteine residues are likely to be important for protein folding or tail assembly. There are four different steps at which the cysteines may play a role: (1) tail assembly, (2) tail attachment to head, (3) phage adsorption to host cells, and (4) DNA injection. In order to find out which step is defective in these mutants, the target mutation was introduced into the pEtail4 plasmid and this is followed by a series of assays to investigate the role of the conserved cysteine residues in gpL.

Figure 5.3 Spot complementation assay to examine the activity of the mutants.

The spot complementation assay was conducted as described previously. The first and second columns show the plasmid tested and the protein encoded by the plasmid or the position of mutation, respectively. The third and fourth columns show the type of phage used.

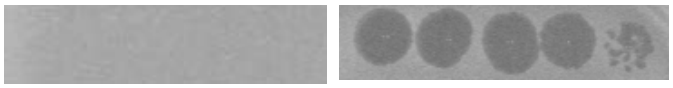
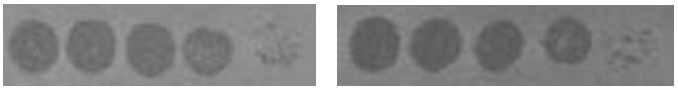
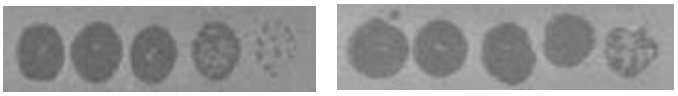
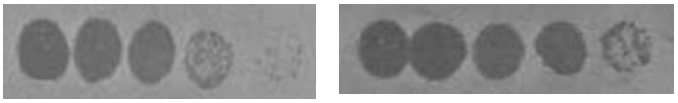
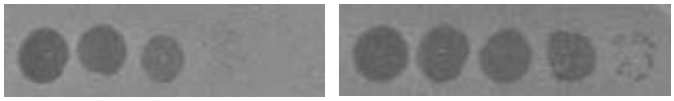
(a) pGRX-6p-1: empty vector.

(b) pGEX-6p-1-L (pGL): encodes λ gpL with a GST fusion at the N-terminus.

(c) pGEX-6p-1-L3M (pGL3M): encodes GST-L with C140M mutation.

(d) pGEX-6p-1-L127 (pGL127): encodes GST-L with C12S, C39S and C208S mutations.

(e) pGEX-6p-1-L4C (pGL4C): encodes GST-L with the four non-conserved cysteines mutated (C12S, C39S, C140M and C208S).

Plasmid	Protein/Mutation	λLam756cI857	λcI857
pGEX-6p-1	GST		
		(a)	
pGL	GST-L		
		(b)	
pGL3M	C140M		
		(c)	
pGL127	C12S, C39S, C208S		
		(d)	
pGL4C	C12S, C39S, C140M, C208S		
		(e)	

(1) Tail assembly

One possibility for the inability of the mutants to complement *in vivo* is that they are unable to assemble lambda tail. In order to find out if the conserved cysteine residues are essential for lambda tail assembly, the individual cysteine to serine mutations were introduced into the pEtail4 plasmid which encodes all the 11 proteins required for lambda tail assembly. Lambda tail proteins expressed from the plasmid were radiolabeled and prepared for 10-30% glycerol gradient (see MATERIALS AND METHODS). About 12 fractions were collected from each gradient tube after centrifugation. Figure 5.4 shows that the single amino acid substitution at position 173, 182 or 205 disables lambda tail assembly. The C212S mutant, however, is able to assemble lambda tails. The gpV band in the fraction which contains most of the purified tail from both the wildtype and C212S mutant samples was quantified in the same way as described in the previous chapter. Based on the net intensity of the gpV band, the amount of the tail in the C212S mutant sample is about 4.2±1.0% of that of the wildtype sample.

(2) Tail attachment to head

Figure 5.4 shows that a small amount of tail can accumulate in the C212S mutant, but it does not tell whether or not the mutant tail is able to attach to lambda head or host cells. In order to investigate the ability of the mutant tail to attach to lambda head, radiolabeled tail proteins expressed from pEtail or pEtail4-8S were incubated with non-radiolabeled, purified head prepared from 594(λVam750S7cI857) lysogen to allow the binding of tail to head. This was followed by cesium chloride step gradient to separate phage from other components in the lysate (see MATERIALS AND METHODS). The gradient tube was fractionated after centrifugation. If the mutant tail is able to bind to the head, mutant phage will be found in about the same fraction

Figure 5.4 Effect of single amino acid substitution on lambda tail assembly.

The conserved cysteine to serine mutation was introduced into pEtail4. The plasmid was co-transformed with pLysS into strain BL21(DE3)-Δtail. λ tail proteins expressed from the plasmid were labeled by ^{35}S-Met and prepared for 10-30% glycerol gradient followed by fractionation (see MATERIALS AND METHODS). Each fraction was prepared for 15% SDS-PAGE followed by autoradiography. The bands representing tail proteins are labeled on the right. For Panels (a)-(g), lanes 1-12 are fractions from the bottom to the top of the gradient, and lane 13 is the sample used for the gradient. Red rectangles indicate the fraction which contains most of the tail.

(a) pEtail4-Lam: an amber mutation was introduced to gene *L*, resulting in an amber fragment with 95 amino acids.

(b) pEtail4: encodes all 11 proteins required for lambda tail assembly.

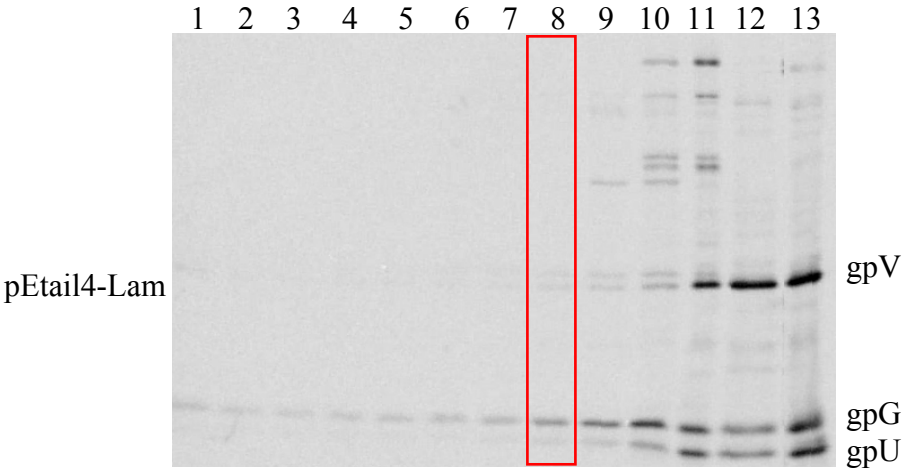
(c) pEtail4-7S: the non-conserved cysteine at 208 in gpL was mutated to serine.

(d) pEtail4-4S: encodes gpL with C173S mutation.

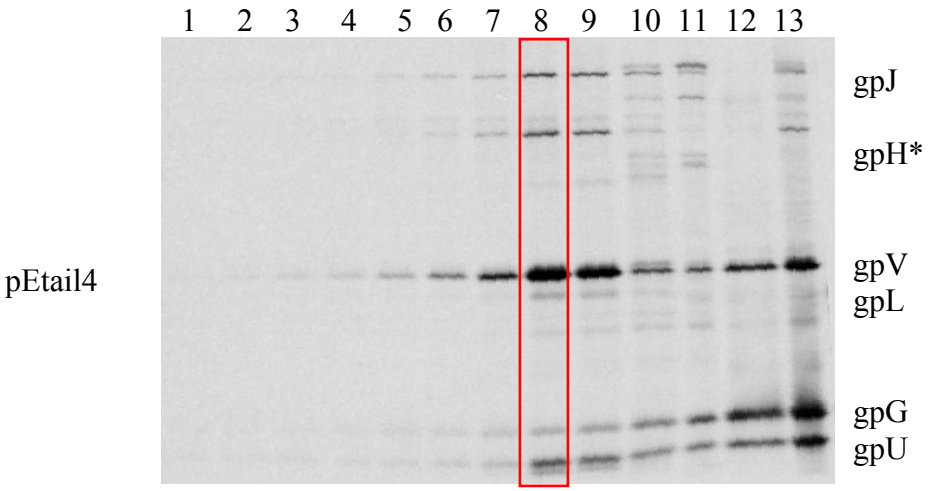
(e) pEtail4-5S: encodes gpL with C182S mutation.

(f) pEtail4-6S: encodes gpL with C205S mutation.

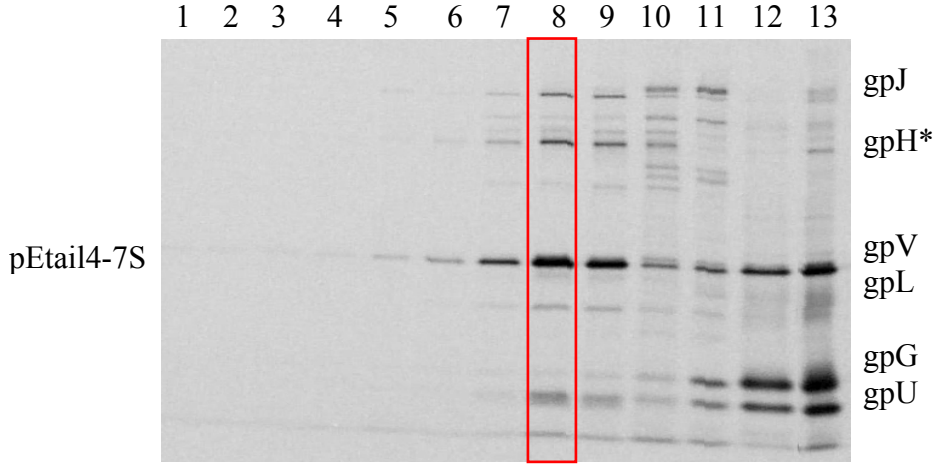
(g) pEtail4-8S: encodes gpL with C212S mutation.


1 2 3 4 5 6 7 8 9 10 11 12 13
pEtail4-Lam
gpV
gpG
gpU

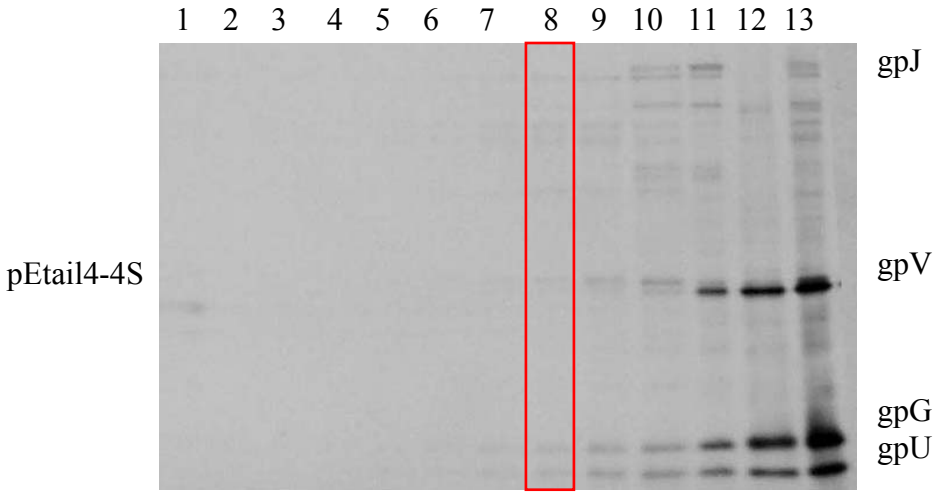
(a)


1 2 3 4 5 6 7 8 9 10 11 12 13
pEtail4
gpJ
gpH*
gpV
gpL
gpG
gpU

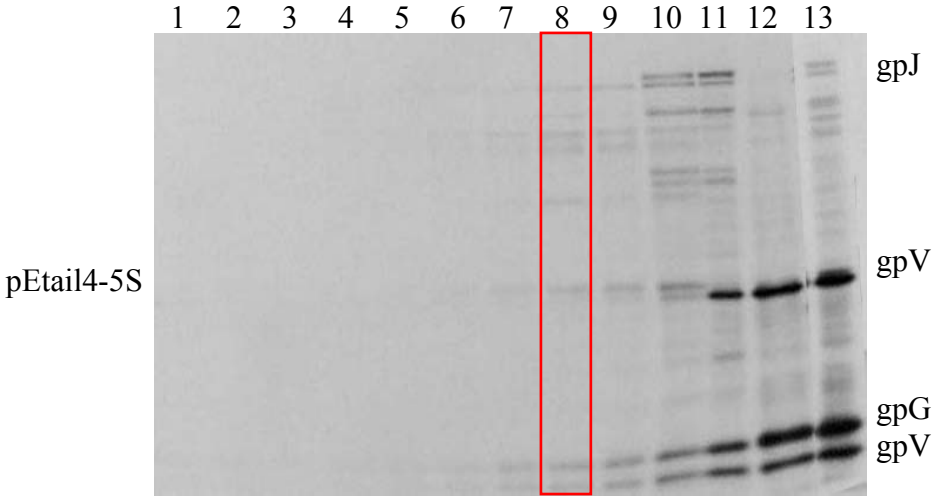
(b)


1 2 3 4 5 6 7 8 9 10 11 12 13
pEtail4-7S
gpJ
gpH*
gpV
gpL
gpG
gpU

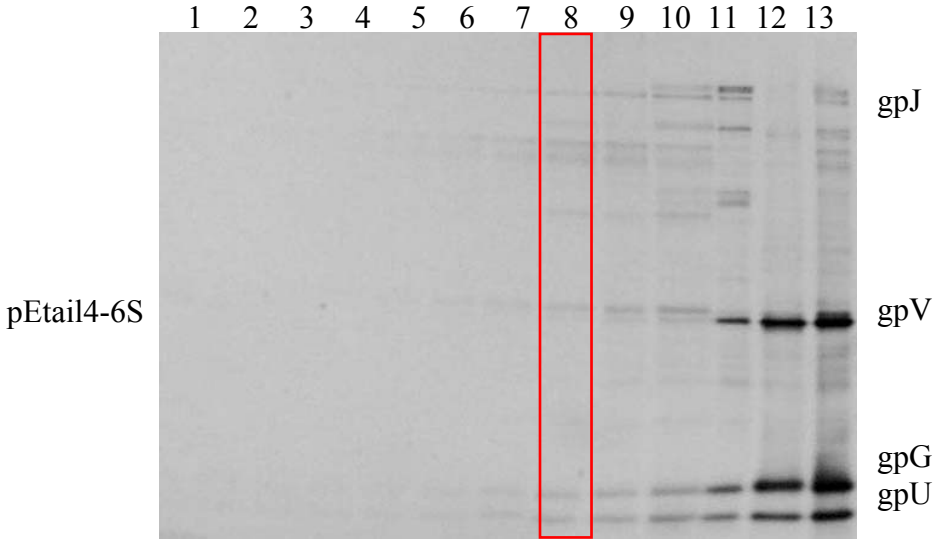
(c)


1 2 3 4 5 6 7 8 9 10 11 12 13
pEtail4-4S
gpJ
gpV
gpG
gpU

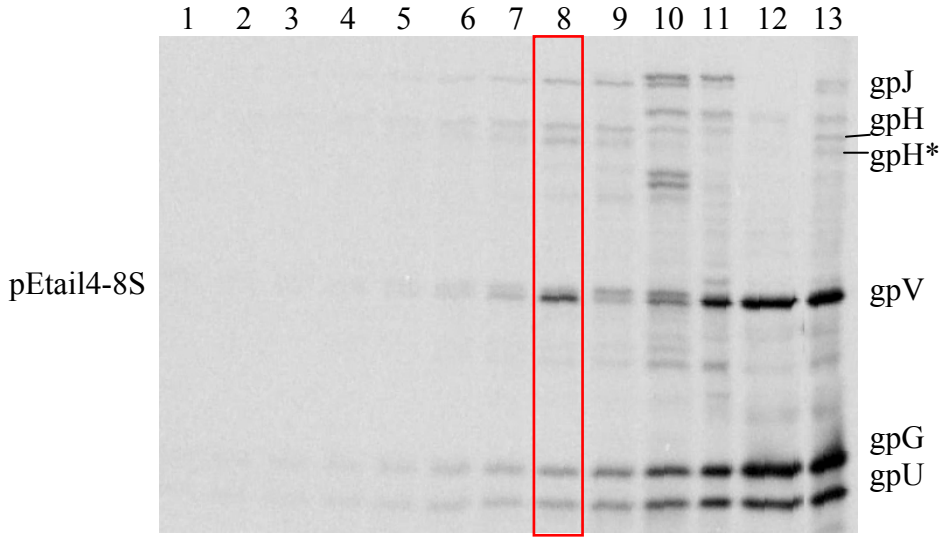
(d)


1 2 3 4 5 6 7 8 9 10 11 12 13
pEtail4-5S
gpJ
gpV
gpG
gpV

(e)


1 2 3 4 5 6 7 8 9 10 11 12 13
pEtail4-6S
gpJ
gpV
gpG
gpU

(f)


1 2 3 4 5 6 7 8 9 10 11 12 13
pEtail4-8S
gpJ
gpH
gpH*
gpV
gpG
gpU

(g)

as the wildtype phage. Figure 5.5 shows that a small amount of mutant phage was found in about the same position as the wildtype phage. By using the same quantification method as described above, it was found that there is about 5.0±1.2% of the mutant phage compared to the wildtype phage in the fifth fraction. This suggests that the mutant tail can attach to head as efficiently as the wildtype tail.

In addition, I used the plaque assay to determine the viable phage titer in the fifth fraction from both the wildtype and mutant sample. The result shows that the viable phage titer in the mutant sample is only about 0.08±0.05% to that of the wildtype phage. Thus the specific infectivity of the phage particles with mutant tails is ~1.6% of that of the wildtype particles. This indicates that the mutant phage is probably defective in phage attachment to cell envelope or DNA injection.

(3) Phage adsorption to host cells

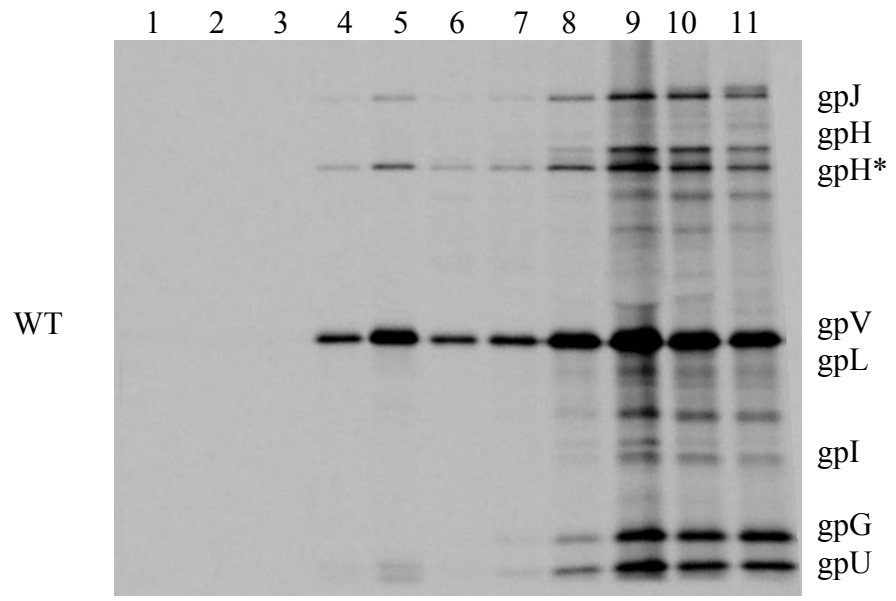
In order to find out if the C212S mutant phage is able to attach to *E. coli* cells, I used *E. coli* cells to precipitate lambda tail. If the mutant tail is able to bind to LamB receptor, it will be precipitated by *E. coli* cells. Lambda tail proteins expressed from the pEtail4 or pEtail4-8S plasmid were radiolabeled and prepared for incubation with *E. coli* cells as described previously. Figure 5.6 shows that the mutant tail can attach to host cells. Based on the net intensity of the gpV band in the wildtype and C212S mutant samples, the amount of mutant tail bound to bacteria is about 4.3±0.6% to that of the wildtype tail. This suggests that the mutant tail can adsorb to bacteria as efficiently as the wildtype tail. Taken together, the results from the above experiments imply that the C212S mutant is probably defective in DNA injection.

Figure 5.5 The mutant tail is as efficient as the wildtype tail in attachment to λ head.

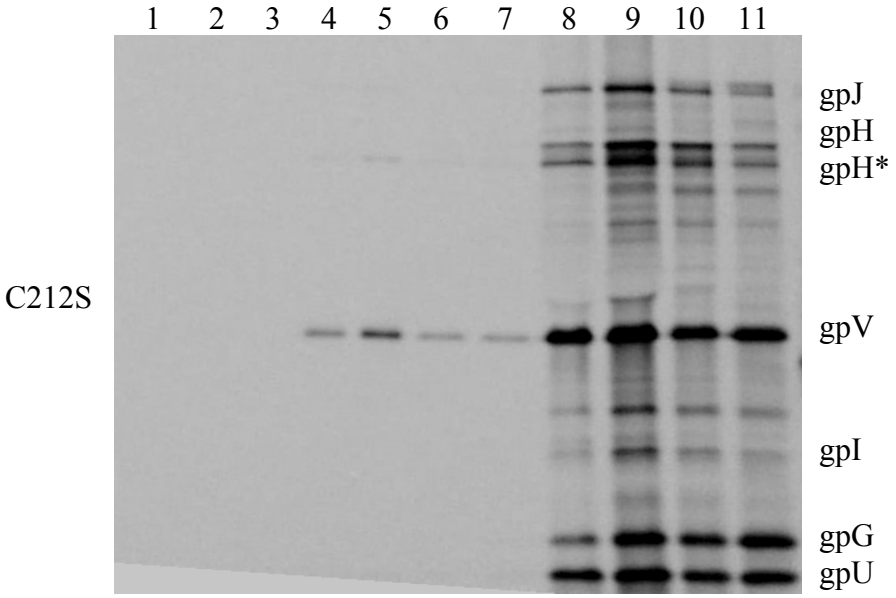
λ tail proteins expressed from pEtail4 or pEtail4-8S were radiolabeled by ^{35}S-Met and lysed as described (see MATERIALS AND METHODS). The 594(λVam750Sam7cI857) lysogen was used for head preparation. The radiolabeled tail lysate was incubated with non-radiolabeled head to allow the binding of tail to head. Subsequently, the mixture was loaded onto cesium chloride step gradient to purify phage. About 11 fractions were collected from the bottom of the tube. Each fraction was prepared for 15% SDS-PAGE followed by autoradiography. In Panels (a) and (b), lanes 1-11 are fractions from the bottom to about 1/3 of the gradient tube. The bands representing tail proteins are labeled on the right.

(a) The pEtail4 plasmid which encodes all 11 proteins required for lambda tail assembly is used for tail preparation.

(b) The pEtail4-8S plasmid differs in pEtail4 in that it encodes gpL with C212S mutation.


1 2 3 4 5 6 7 8 9 10 11
WT
gpJ
gpH
gpH*
gpV
gpL
gpI
gpG
gpU

(a)


1 2 3 4 5 6 7 8 9 10 11
C212S
gpJ
gpH
gpH*
gpV
gpI
gpG
gpU

(b)

Figure 5.6 The mutant tail can attach to host cells as efficiently as the wildtype tail.

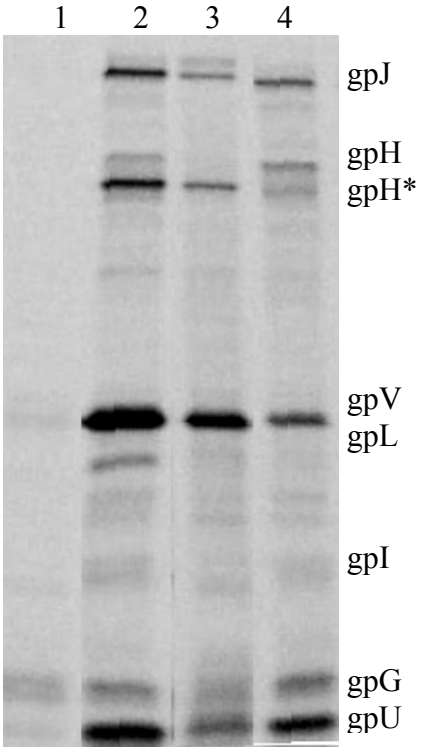
Tail proteins expressed from the target plasmids were radiolabeled and lysed as described previously. The radiolabeled tail lysate was incubated with *E. coli* cells to allow the binding of tail fiber to *E. coli* cells. Subsequently, the mixture was washed and prepared for 15% SDS-PAGE. The bands were visualized by autoradiography. Bands representing lambda tail proteins are labeled on the right.

Lane 1: pEtail4-Lam: An amber mutation was introduced to gene *L*.

Lane 2: pEtail4: encodes all 11 proteins required for lambda tail assembly.

Lane 3: pEtail4-7S: differs in pEtail4 in that it encodes gpL with the nonconserved cysteine residue at position 208 substituted by serine.

Lane 4: pEtail4-8S: differs in pEtail4 in that it encodes gpL with C212S mutation. The weak gpH* band is not a feature of this mutant since this band can be seen clearly when I repeated the experiment.


1
2
3
4
gpJ
gpH
gpH*
gpV
gpL
gpI
gpG
gpU

5.3 FURTHER CHARACTERIZATION OF C212 OF λ gpL

In addition to my results, the results from Dr. Davidson's lab at the University of Toronto indicate that gpL binds to a metal, possibly iron (personal communication). It is suggested that the conserved cysteine residues in gpL may form an iron-sulfur cluster, which may be important for the correct folding of gpL or for the formation of the tail tip. If this is true, it is possible that substitution of C212 by histidine may not completely abolish tail assembly since histidines can sometimes form zinc rings. In order to further characterize C212, two more mutants, C212A and C212H, were constructed into pEtail4 by recombinant PCR. After construction of plasmids, a spot complementation assay was performed to determine the biological activity of the mutant. Figure 5.7 shows that λLam756cI857 can grow on any of the mutants tested at ~0.1% efficiency. Since the mutant protein encoded by pEtail4-6S is defective in tail assembly, it is likely that this level of phage growth is due to recombination between two alleles to give a wildtype sequence.

The C212A and C212H mutants were then examined by glycerol gradient for their ability to assemble λ tails as described previously. Figure 5.8 shows that the majority of the wildtype tail is found in the eighth and ninth fractions whereas no tail-like structures are present in these fractions in either mutant. Therefore, neither of the mutants is able to assemble tails. I also tried to find out if there is any difference in the early tail assembly intermediates for the C212A, C212H and C212S mutants. Each mutation was introduced to the pGtail plasmid, which encodes the first eight proteins involved in lambda tail assembly. *E. coli* cells were used to precipitate lambda tail assembly intermediates. Surprisingly, the C212A and C212H mutants behave like the wildtype protein while there is a significant decrease in the amount of the tail tip proteins

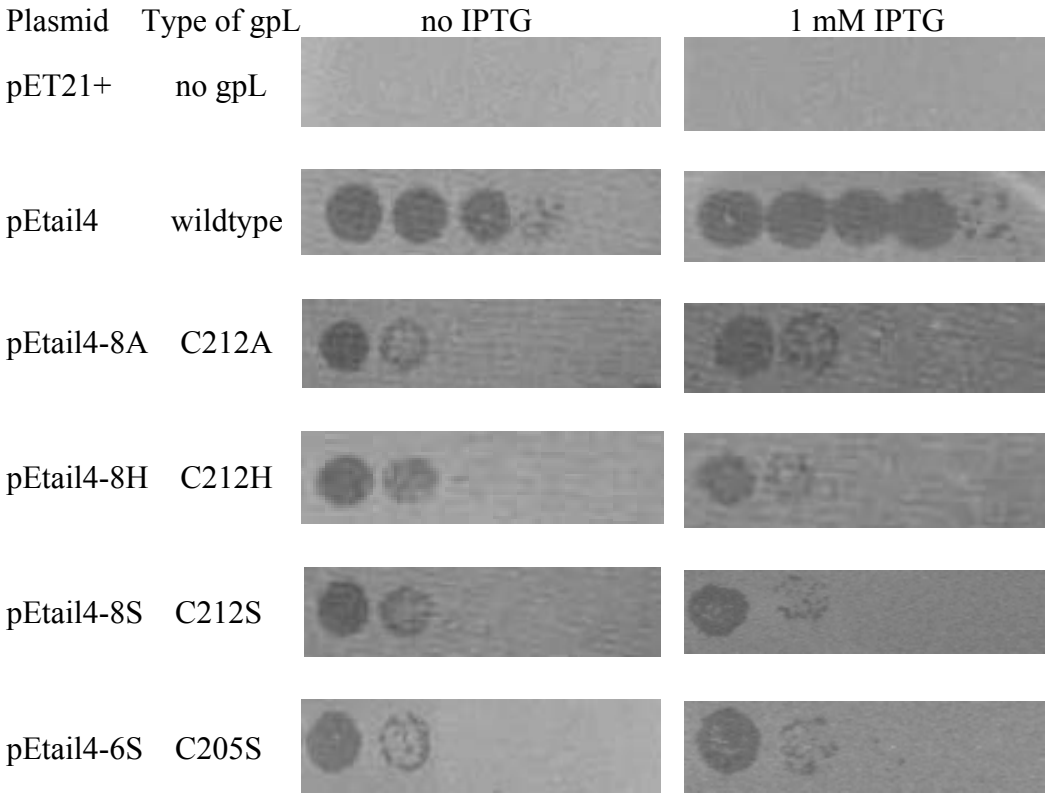
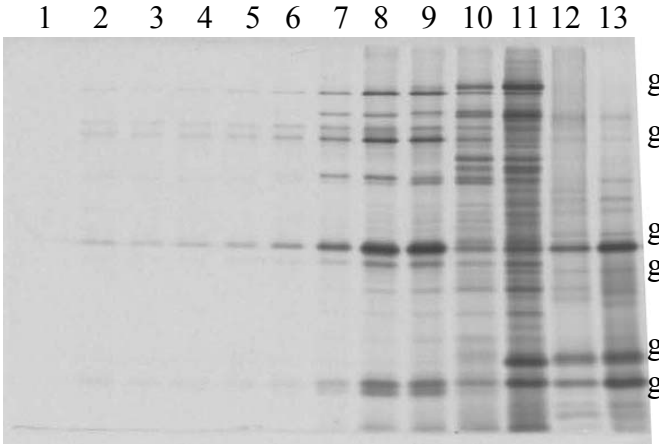


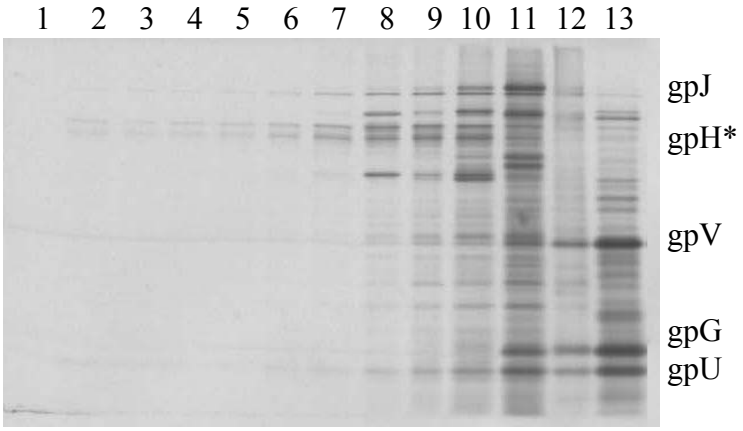
Figure 5.7 Spot complementation test to determine the biological activity of C212A and C212H mutants. Each plasmid was co-transformed with pLysS into BL21(DE3)-△tail. The spot complementation experiment was performed as described previously. Each target plasmid was tested for its ability to complement λLam756cI857 (L756). No IPTG or 1 mM IPTG was present in the top soft agar.

Figure 5.8 Separation of λ tail by glycerol gradient.

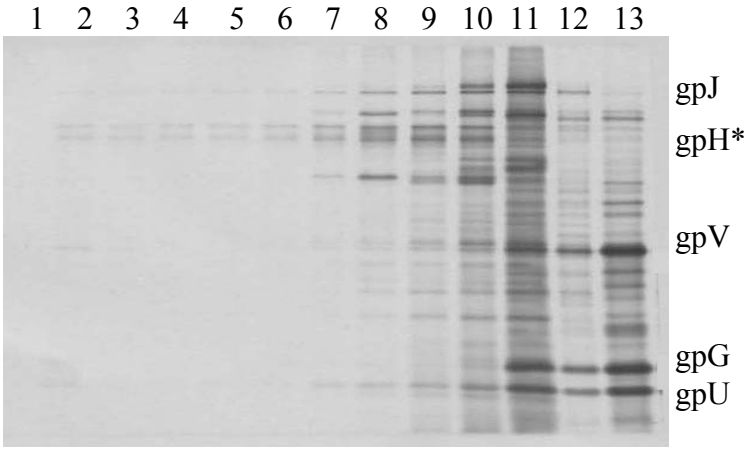
The radiolabeled tail lysates were prepared and loaded onto 10-30% glycerol gradient as described previously. After centrifugation, each gradient was punctured at the bottom and about 13 fractions were collected. Each fraction was prepared for 15% SDS-PAGE followed by autoradiography. In Panels (a)-(c), lanes 1-13 are fractions 1-13 from the bottom to the top of the gradient.



(a) pEtail4



(b) pEtail4-8A



(c) pEtail4-8H

precipitated by *E. coli* cells in the C212S mutant (Figure 5.9). Since the C212S mutant lysate can accumulate a small amount of tail as shown in the previous sections, the result suggests that in the absence of the other tail proteins, the tail assembly intermediates produced by pGtail with C212S mutation may be not as stable as that from the wildtype protein. In addition, the result suggests that the C212A and C212H mutants may be defective in some late steps in the tail assembly pathway. So far, it remains an open question for why the conserved cysteine to serine substitution causes problems in the tail assembly. To answer this question, it is probably essential to study the structure of gpL.

5.4 DISCUSSION

Sections 5.2-5.3 describe the investigation of the cysteine residues in the lambda tail protein, gpL. A summary of the results for the cysteine mutants is shown in Figure 5.10. The result from the complementation assay shows that the single amino acid substitution at position 12, 39, 140 or 208 has little effect on the biological activity of the protein whereas the single amino acid substitution at position 173, 182, 205 or 212 results in the inability of the protein to complement *in vivo*. This result is consistent with the multi-sequence alignments. In addition, the result also suggests that the conserved cysteine residues may be important for protein folding or tail assembly or for the functioning of the assembled tail or metal binding.

In order to further characterize the conserved cysteine residues, the individual mutation was introduced into pEtail4. The radiolabeled tail lysate was first examined by glycerol gradient. Figure 5.4 shows that the single amino acid substitution at position 173, 182 or 205 completely

Figure 5.9 Precipitation of λ tail assembly intermediates by *E. coli* cells.

The tail proteins expressed from the target plasmid were radiolabeled and mixed with *E. coli* cells as described previously. Each sample was incubated with Ymel (an *E. coli* K12 strain with LamB receptor) and CR63 (an *E. coli* strain which carries a point mutation in gene lamb and therefore is resistant to lambda phage). Lambda tail tip proteins are labeled on the right.

Lanes 1-2: pGtail: encodes the first eight proteins involved in lambda tail assembly. The lysate was precipitated by Ymel (lane 1) or CR63 (lane 2).

Lanes 3-4: pGtail-8A: differs in pGtail in that it encodes gpL with C212A mutation. The lysate was precipitated by Ymel (lane 3) or CR63 (lane 4).

Lanes 5-6: pGtail-8H: differs in pGtail in that it encodes gpL with C212H mutation. The lysate was precipitated by Ymel (lane 5) or CR63 (lane 6).

Lanes 7-8: pGtail-8S: differs in pGtail in that it encodes gpL with C212S mutation. The lysate was precipitated by Ymel (lane 7) or CR63 (lane 8).

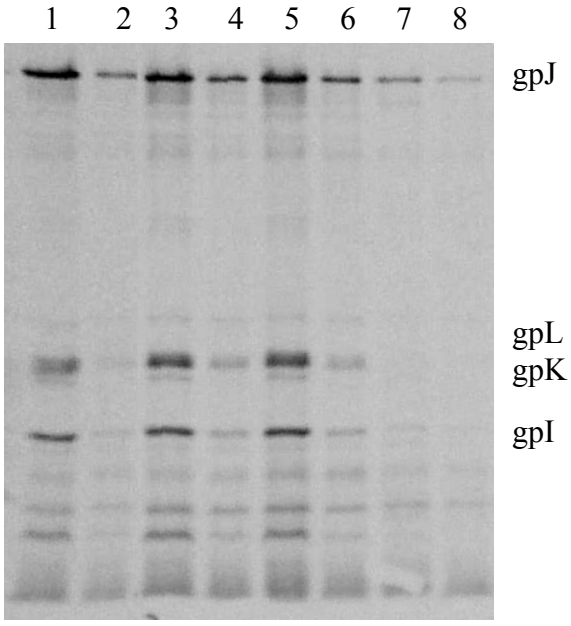

1 2 3 4 5 6 7 8
gpJ
gpL
gpK
gpI

Figure 5.10 Summary of the cysteine mutants in λ gpL.

This is a summary of the results for the cysteine mutants in lambda gpL from Sections 5.2-5.3.

(a) The positions of the cysteine residues in protein gpL are shown. Whether or not the cysteine residues are conserved is based on the multi-sequence alignments shown in Figure 5.1.

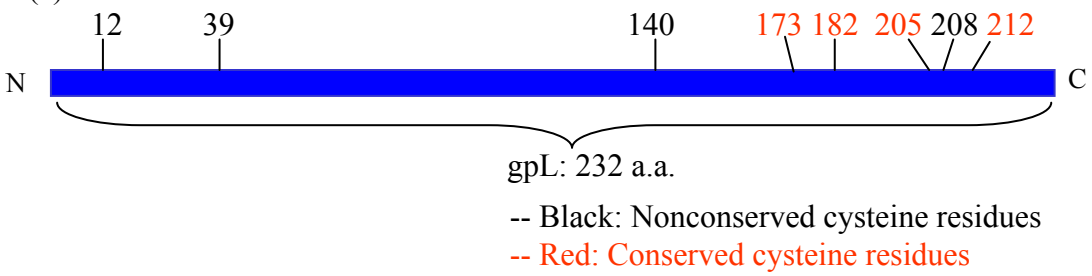
(b) A summary of the phonotypes for the cysteine mutants is shown. The triple mutant contains C12S, C39S and C208S mutations whereas the quadruple mutant contains C12S, C39S, 140M and C208S mutations.

+: the number of complementation phage titers is at least 1000 fold over the background phage

-: the complementation phage titer is about the same as the backround phage titer

N/A: the mutant has not been examined

(a)



(b)

Mutants	Phenotypes	
	In vivo activity	*In vitro*
C12S	+	N/A
C39S	+	N/A
C140S	-	N/A
C140M	+	N/A
C173S	-	Defective in the tail assembly
C182S	-	Defective in the tail assembly
C205S	-	Defective in the tail assembly
C208S	+	Makes tails
C212S	-	Makes tails inefficiently; defective in DNA injection
C212A	-	Defective in the tail assembly
C212H	-	Defective in the tail assembly
Triple	+	N/A
Quadruple	+	N/A

abolishes lambda tail assembly. A small amount of tail can accumulate in the C212S mutant lysate. Further characterization of this mutant indicates that the mutant tail is as efficient as the wildtype tail in binding to head and attachment to host cells. Taken together, these results show that all of the mutants that have a change in a conserved cysteine are defective in tail assembly, and that the small number of tails successfully assembled by the C212S mutant are defective in a step of infection following adsorption to the cell. These experiments therefore identify two different steps in tail assembly and function where gpL has an essential role.

6.0 CHARACTERIZATION OF λ gpK

6.1 INTRODUCTION

The lambda gene *K* encodes a 199 amino acid protein. Its gene product, gpK, is required for tail assembly, but it is not detected in the mature virion. In previous studies, a phage with an amber mutation in gene *K* (λKam768Sam7cI857) was isolated (Parkinson, 1968). The lysate made from this mutant phage can accumulate some phage-like particles which sediment as fast as normal phages (Katsura and Kuhl, 1975). These K^- phages are not infective by themselves, but they can complement all other tail-defective lysates except the K^--lysate (Katsura and Kuhl, 1975). These results suggest that gpK can be bypassed to some extent (Katsura and Kuhl, 1975). However, it is also possible that the amber fragment of gpK is sufficient to rescue the formation of a small fraction of phage-like particles.

In order to further characterize gpK, the gene *K* in λKam768Sam7cI857 phage was sequenced to identify the position of the amber mutation. The result shows that this mutant phage encodes a 167-amino-acid amber fragment. Another amber mutation, Kam6, was introduced early in gene *K*. These two amber mutants were further characterized by *in vivo* and *in vitro* assays.

6.2 CONSTRUCTION AND CHARACTERIZATION OF AMBER MUTATIONS IN λ GENE *K*

6.2.1 Construction of amber mutations in λ gene *K*

The pEtail4 plasmid encodes all 11 proteins required for lambda tail assembly. Two different amber mutations were constructed into this plasmid, resulting in pEtail4-Kam6 and pEtail4-Kam768. These two plasmids differ from pEtail4 in that they carry an amber mutation early in gene *K* (pEtail4-Kam6) or late in gene *K* (pEtail4-Kam768).

An amber mutation in the sixth codon of gene *K* was constructed by recombinant PCR in plasmid pLtail, to give plasmid pKam6 (Table 3.1). Plasmid pEtail4-Kam6 was constructed by inserting the BsrGI-RsrII fragment of pKam6 into the BsrGI-RsrII sites of pEtail4.

The amber mutation in pEtail4-Kam768 is the same amber mutation in gene *K* as in the previously isolated mutant phage λKam768Sam7cI857. PCR was performed to amplify gene *K* in this mutant phage. Subsequently, the PCR product was sequenced to identify the position of the amber mutation. The sequence result shows that the introduction of Kam768 mutation should result in a 167-amino-acid amber fragment (data not shown). Subsequently, recombinant PCR was performed to introduce the Kam768 mutation into pEtail4 to construct the pEtail4-Kam768 plasmid (see MATERIALS AND METHODS).

6.2.2 Determination of the biological activity of the amber fragments *in vivo*

After construction of plasmids, a spot complementation assay was carried out to determine the ability of the amber fragment to complement λKam768cI857 *in vivo*. The experiment was performed as described previously and the result is shown in Figure 6.1. The wildtype phage, λcI857, grows equally well on all the strains tested, demonstrating that neither *K* amber fragment interferes with phage growth at the low level of expression tested. No complementation activity is found between pEtail4-Kam768 and λKam768cI857 as expected since they contain the same amber mutation in gene *K*. Phage λKam768cI857 does make plaques on the strain carrying pEtail4-Kam6 at about 1% efficiency. I ascribe this not to complementation but to recombination between Kam6 and Kam768, as mentioned in Chapter 3.

6.2.3 Characterization of the amber mutants *in vitro*

(1) Glycerol gradient

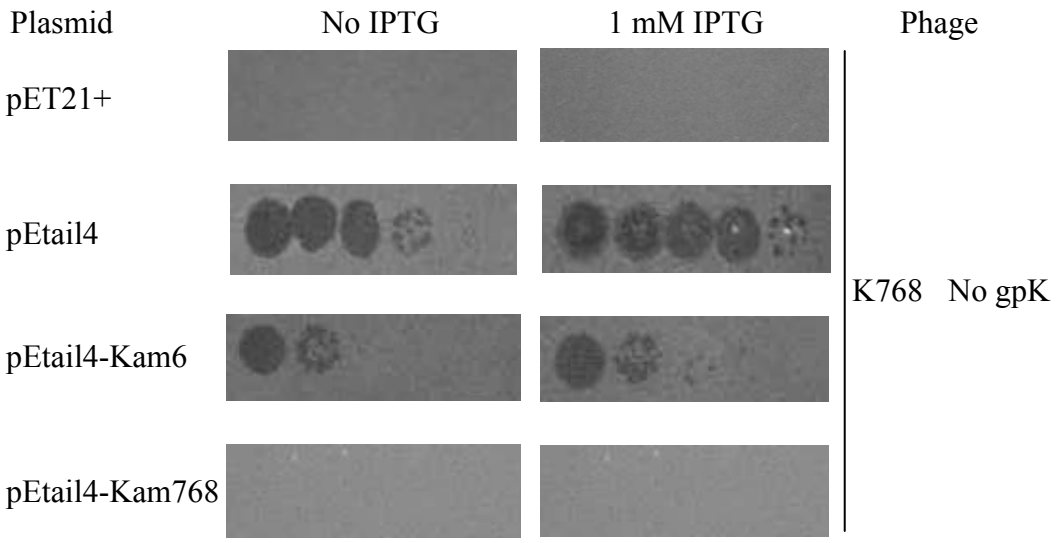
The *in vivo* complementation assay shows that both of the amber fragments have little biological activity. However, it does not tell which step is affected by the mutation. In order to further characterize these amber mutants, glycerol gradient was performed to examine if there is any tail formation in the lysate containing the amber fragment. The experiment was carried out as described previously. Figure 6.2 shows that no tail is found in the lysate made from pEtail4-Kam6 (Panel b). However, a small amount of tail-like structure is accumulated in the lysate made from pEtail4-Kam768 (Panel c). The comparison of the band intensity of gpV reveals that the assembly efficiency of the mutant protein is about 3.1±0.9% to that of the wildtype protein.

Figure 6.1 Determination of the biological activity of the amber fragments *in vivo*.

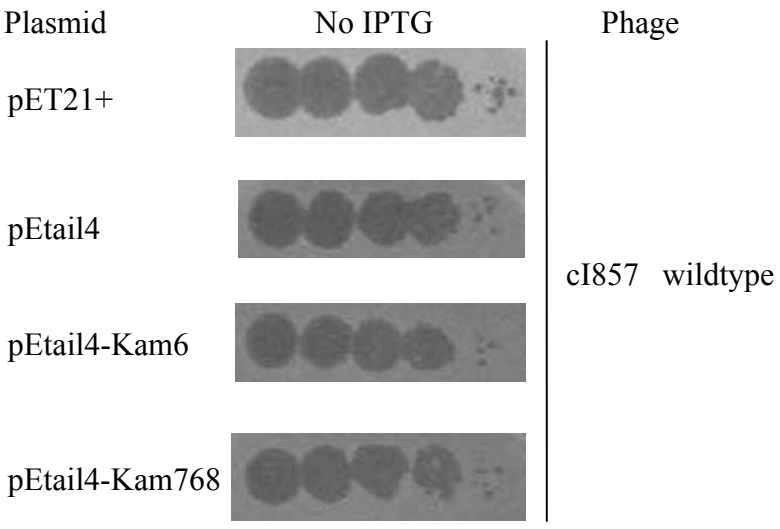
Each target plasmid was co-transformed with pLysS into BL21(DE3)-Δtail. The spot complementation experiment was carried out as described previously. The plasmids of interest were tested for their ability to complement λKam768cI857 (K768). The wildtype phage, λcI857 (cI857), was used as a positive control. No IPTG or 1 mM IPTG was added to the top soft agar. The presence or absence of IPTG makes no difference for the growth of λcI857 on the strains carrying the target plasmids. Therefore, only the result without IPTG in the top soft agar is shown here.

(a) Each target plasmid was tested for its ability to complement λKam768cI857 (K768).

(b) The strains with the target plasmids were spotted with λcI857 (cI857).


Plasmid
No IPTG
1 mM IPTG
Phage
pET21+
pEtail4
pEtail4-Kam6
pEtail4-Kam768
K768 No gpK

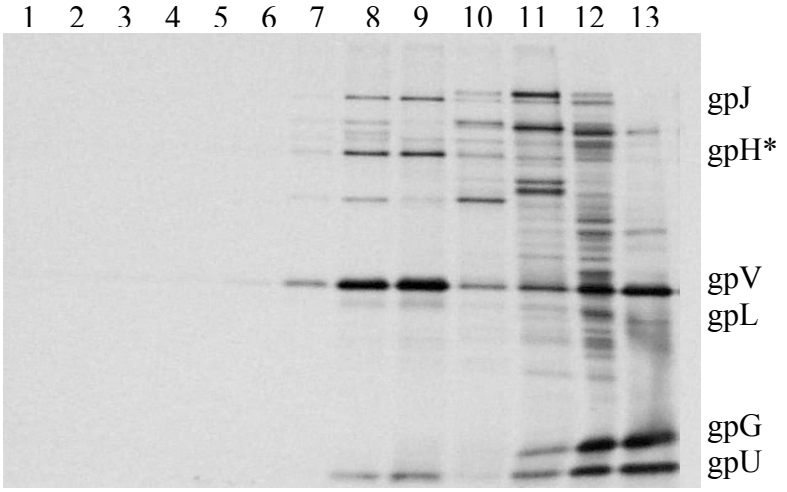
(a)


Plasmid
No IPTG
Phage
pET21+
pEtail4
pEtail4-Kam6
pEtail4-Kam768
cI857 wildtype

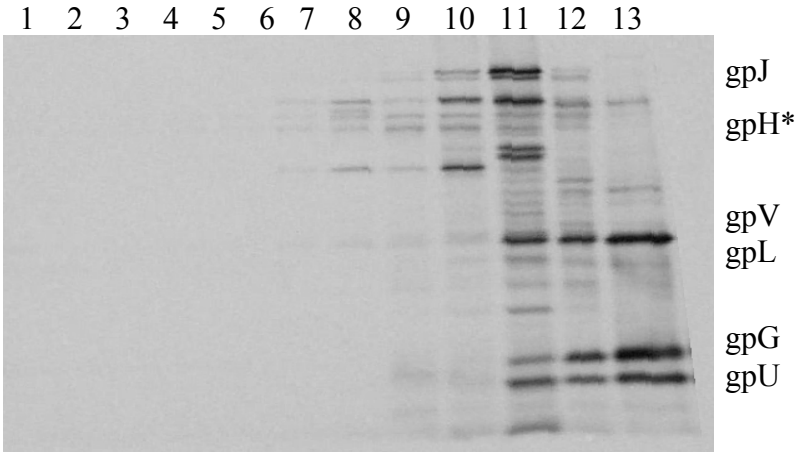
(b)

Figure 6.2 Separation of λ tail by glycerol gradient.

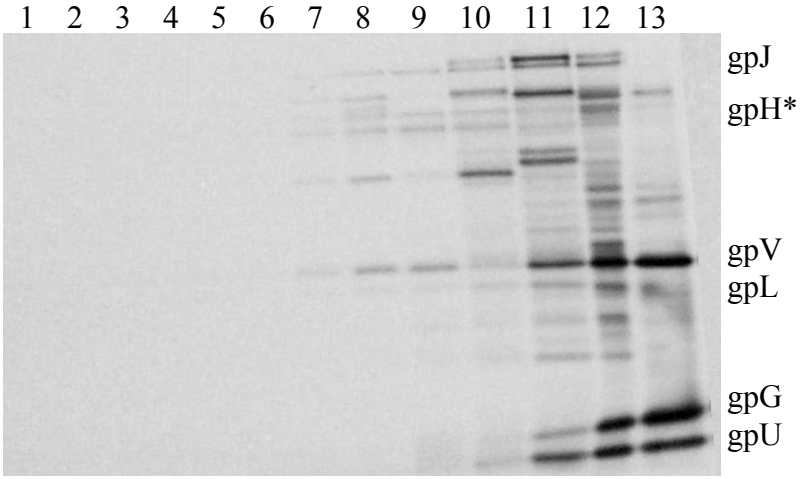
BL21(DE3)-△tail was used as the host strain. The tail proteins expressed from the target plasmid were radiolabeled by ^{35}S-Met and prepared for 10-30% glycerol gradient as described in MATERIALS AND METHODS. After centrifugation, each gradient tube was collected into about 12 fractions. Each fraction was prepared for 15% SDS-PAGE followed by autoradiography. Lanes 1-12 in Panels (a)-(c) are fractions 1-12 from the bottom to the top of the gradient. Lane 13 is the sample used for the gradient.



(a) pEtail4



(b) pEtail4-Kam6



(c) pEtail4-Kam768

(2) Cesium chloride step gradient

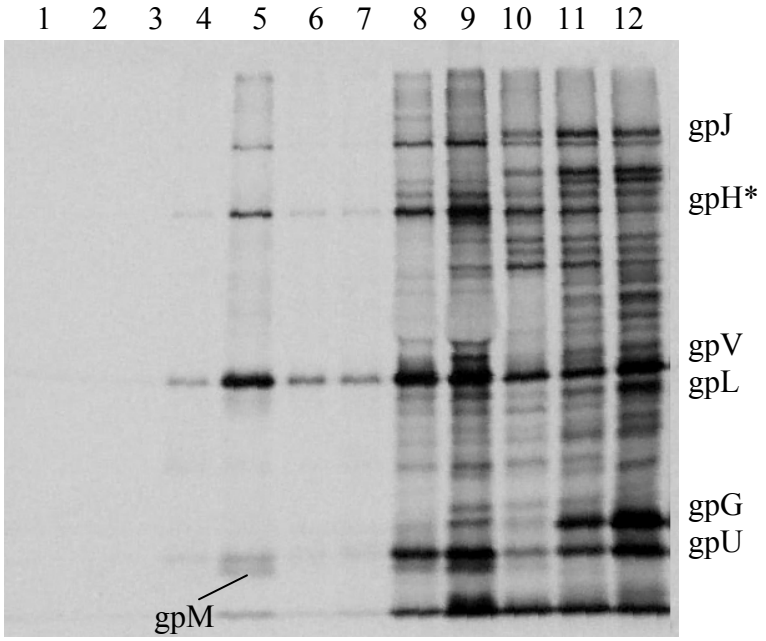
Next, I want to examine if the tail made from pEtail4-Kam768 is able to attach to lambda head. The experiment was carried out as described previously and the result is shown in Figure 6.3. The mutant tail is able to attach to head. Based on the net band intensity of gpV, the amount of the mutant phage is approximately 3.6 ±1.1% of the wildtype phage. This suggests that the mutant tail can attach to the head as efficiently as for the wildtype phage.

6.3 DISCUSSION

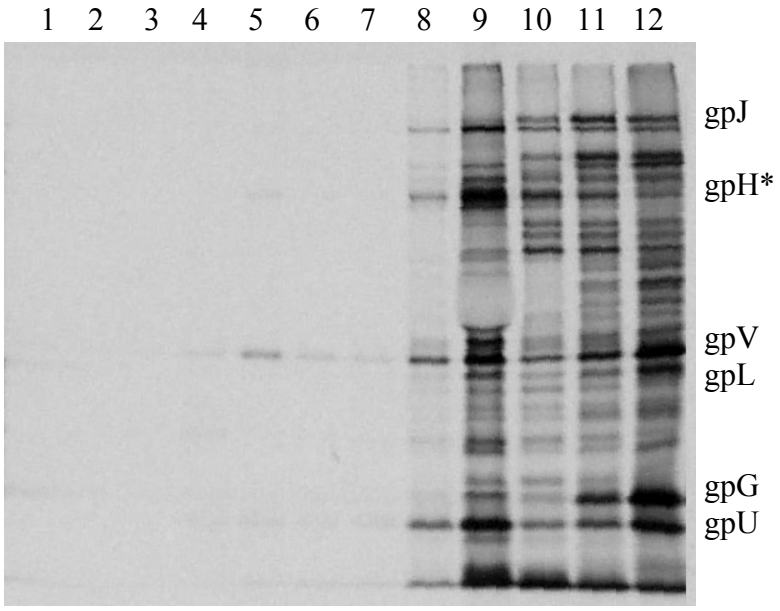
The function of gpK is a mystery because it is required for lambda tail assembly, but it is not detected in the mature tail structure. One possible interpretation of this observation is that it may play a chaperone-like role. Based on the result from PSI-BLAST, the protein sequence of gpK matches many cell wall-associated hydrolases and NLP/P60 family proteins, a family of cell-wall peptidases. All characterized members of this family have been shown to either hydrolyze D-glutamyl-meso-diaminopimelate or *N*-acetylmuramate-L-alanine linkages (Anantharaman and Aravind, 2003). Many phages contain a protein with lysozyme activity (Moak and Molineux, 2004). Gp5 of bacteriophage T4, for example, has lysozyme activity and helps the tail tube penetrate through the cell envelope during infection (Kanamaru et al., 2002). It is therefore tempting to speculate that gpK could be a lysozyme in the λ tail. Arguing against this possibility is the failure to detect a band corresponding to gpK in SDS gels of λ virions. However, I cannot rule out the possibility that gpK is proteolytically processed during tail maturation and the fragment or fragments that remain in the virion have not been detected.

Figure 6.3 Separation of lambda phage by cesium chloride step gradient.

594(λVam750Sam7cI857) was used for head preparation (see MATERIALS AND METHODS). Lambda tail proteins expressed from the target plasmid were radiolabeled by ^{35}S-Met as described in MATERIALS AND METHODS. The non-radiolabeled head was incubated with radiolabeled tail to allow the binding of head to tail (see MATERIALS AND METHODS). Subsequently, the mixture was loaded onto a cesium chloride step gradient. After centrifugation, the tube was punctured at the bottom and about 12 fractions were collected. Each sample was prepared for 15% SDS-PAGE and visualized by autoradiography. In Panels (a) and (b), lanes 1-12 were fractions 1-12 from the bottom to about 1/3 of the gradient tube.



(a) pEtail4



(b) pEtail4-Kam768

My experiments suggest that a small amount of phage-like particles can accumulate in the Kam768 lysate. This result is consistent with previous studies about this mutant (Katsura and Kuhl, 1975). In addition, my result shows that the Kam6 is unable to assemble lambda tail. Taken together, these results argue that gpK is required for assembly of tails but that the 167-amino-acid amber fragment of Kam768 does support assembly of tail-like structures of approximately normal size and composition, which are able to attach to heads (Katsura and Kuhl, 1975). However, these tails are not functional unless they are complemented with a lysate capable of supplying wildtype gpK. Thus, it appears that gpK has at least two discernable functions. The first is to assist tail assembly. This function can be carried out, albeit inefficiently, by a gpK lacking the last 32 amino acids. The second function, which requires that the rest of the gpK sequence be present, makes the assembled virion capable of infecting. This could mean that gpK is present in the mature virion, acting, for example, as a lysozyme to make a hole in the cell wall for DNA injection. Alternatively, it could mean that gpK modifies the propertities of the tail during assembly to make it competent to support infection, and may or may not itself be present in the mature tail.

7.0 DISCUSSION

7.1 CHANGES ON ATTACHMENT TO THE HOST CELL

The result of the experiment using *E. coli* cells to precipitate lambda tail tip assembly intermediates (Figure 4.1) shows that a mystery band, xx2, is only present in the pLtail lane. Therefore, it is likely that band xx2 is a processed or degradation product of one of the tail tip proteins. The presence of this band may be explained by the possibility of the change in the tail tip complex upon attachment to the host cell. That is upon attachment to the host cell, the tail tip may undergo certain change which may result in one of the tail tip proteins to be proteolytically degraded or processed and this change may only happen in the tail tip complex, but not in the tail tip assembly intermediates. Since the relative amount of gpI in the pLtail lane seems less than that in the pLdel or pKam6 lane, it is possible that band xx2 may be a processed or degradation product of gpI.

7.2 DnaK MAY BE REQUIRED FOR THE TAIL TIP COMPLEX TO BE BIOLOGICALLY ACTIVE

The results from the separation of lambda tail tip complex by the sucrose velocity gradient followed by *in vitro* complementation assay and Co-IP (Figures 4.4, 4.5 and 4.7) show that the

fraction which has most *in vitro* complementation activity contains very small amounts of the lambda tail tip proteins. In contrast, the majority of the tail tip proteins are present in the top fractions from the gradient and they are not biologically active. One possible explanation is that DnaK may be required for the tail tip complex to be biologically active, since DnaK is present in the fraction which has most *in vitro* complementation activity, but not in the top fractions. Arguing against this possibility is the absence of DnaK in the tail tip complex precipitated by *E. coli* cells (Figure 4.1). However, this can be explained by that DnaK may be released upon attachment of the tail fiber to the cell envelope.

7.3 gpK MAY BE A LYSOZYME

Based on the result from PSI-BLAST, the sequence of gpK matches many cell wall-associated hydrolases and the NLP/P60 family proteins. This leads to the assumption that gpK may be a lysozyme. The results from my experiments about Kam6 and Kam768, suggest that the function of gene *K* can not be bypassed. The results about Kam768 mutation from my experiments are consistent from previous studies (Katsura and Kuhl, 1975). Besides, previous studies also show that the Kam768 mutant phage has little infectivity. However, when it is complemented by other tail-defective lysates, the mutant phage will become infectious. Taken together, these data suggest that gpK may be a lysozyme and may help facilitate penetration of the cell wall during infection. The fact that the gpK band is not detected in the mature tail purified by glycerol gradient and in the phage purified by cesium chloride step gradient may argue the possibility of gpK to be a lysozyme. However, in both purified mature tail and lambda phage, there is a band which migrates a little faster than gpK (band xx4 in Figure 4.9 and band xx5 in Figure 4.10). It is

possible that gpK may need to be processed during tail assembly and this band may be a processed form of gpK. If this is true, it may not be surprising since the lambda tape measure protein, gpH, has been shown to be processed during assembly in order to form a biologically active tail (Murialdo and Siminovitch, 1972; Tsui and Hendrix, 1983).

7.4 COMPOSITION AND STOICHIOMETRY OF THE TAIL TIP COMPLEX

The result from precipitation of the tail tip complex by *E. coli* cells shows that gpJ, gpI and gpL are present in the tail tip complex (Figure 4.1). The presence of gpK in the complex is not clear from this experiment. However, the fact that gpK is co-precipitated with DnaK in a sucrose velocity gradient indicates that gpK is present in the tail tip complex (Figure 4.8). The presence of gpK in the tail tip complex is consistent with previous *in vitro* complementation data which show that gpK activity is found in a complex late in the tail assembly pathway (Katsura and Kuhl, 1975). One possible explanation for the absence of gpK in the tail tip complex precipitated by *E. coli* cells is that upon attachment to the host cell, the tail tip complex may undergo some change which may result in gpK to be proteolytically degraded or processed. Another possibility is that gpK may be loosely associated under the experimental conditions.

The quantification results from several different methods suggest that there may be three copies of gpI and three copies of gpL involved in the tail assembly pathway. The copy numbers of gpK involved in the tail assembly pathway is unknown since gpK is not detectable in most experiments.

BIBLIOGRAPHY

Amos, L. A. and A. Klug (1975). "Three-dimensional image reconstructions of the contractile tail of T4 bacteriophage." J Mol Biol **99**(1): 51-64.

Anantharaman, V. and L. Aravind (2003). "Evolutionary history, structural features and biochemical diversity of the NlpC/P60 superfamily of enzymes." Genome Biol **4**(2): R11.

Arber, W. (1983). "A beginner's guide to lambda biology". Lambda II. R. W. Hendrix, Roberts, J.W., Stahl, F.W. and Weisburg, R.A., eds. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory**:** 381-394.

Arisaka, F. (2005). "Assembly and infection process of bacteriophage T4." Chaos **15**(4): 047502.

Berg, J. M. (1990). "Zinc fingers and other metal-binding domains. Elements for interactions between macromolecules." J Biol Chem **265**(12): 6513-6.

Buchwald, M. and L. Siminovitch (1969). "Production of serum-blocking material by mutants of the left arm of the lambda chromosome." Virology **38**(1): 1-7.

Buchwald, M., P. Steed-Glaister, et al. (1970). "The morphogenesis of bacteriophage lambda. I. Purification and characterization of lambda heads and lambda tails." Virology **42**(2): 375-89.

Campbell, A. (1961). "Sensitive mutants of bacteriophage lambda." Virology **14**: 22-32.

Campbell, A., Botstein, D. (1983). "Evolution of the lambdoid phage". Lambda II. R. W. Hendrix, Roberts, J.W., Stahl, F.W. and Weisburg, R.A., eds. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory**:** 365-380.

Casjens, S. and R. Hendrix (1974). "Comments on the arrangement of the morphogenetic genes of bacteriophage lambda." J Mol Biol **90**(1): 20-5.

Casjens, S. R. and R. W. Hendrix (1974). "Locations and amounts of major structural proteins in bacteriophage lambda." J Mol Biol **88**(2): 535-45.

Casjens, S. R., Hendrix, R. W. (1988). "Control mechanisms in dsDNA bacteriophage assembly". The Bacteriophages. R. Calendar, ed. New York, NY, Plenum Press. **1:** 15-91.

Crowther, R. A., E. V. Lenk, et al. (1977). "Molecular reorganization in the hexagon to star transition of the baseplate of bacteriophage T4." J Mol Biol **116**(3): 489-523.

Dove, W. F. (1966). "Action of the lambda chromosome. I. Control of functions late in bacteriophage development." J Mol Biol **19**(1): 187-201.

Duda, R. L., J. Hempel, et al. (1995). "Structural transitions during bacteriophage HK97 head assembly." J Mol Biol **247**(4): 618-35.

Echols, H. and H. Murialdo (1978). "Genetic map of bacteriophage lambda." Microbiol Rev **42**(3): 577-91.

Edmonds, L., A. Liu, et al. (2007). "The NMR structure of the gpU tail-terminator protein from bacteriophage lambda: identification of sites contributing to Mg(II)-mediated oligomerization and biological function." J Mol Biol **365**(1): 175-86.

Fraser, J. S., Z. Yu, et al. (2006). "Ig-like domains on bacteriophages: a tale of promiscuity and deceit." J Mol Biol **359**(2): 496-507.

Gilbert, H. F. (1990). "Molecular and cellular aspects of thiol-disulfide exchange." Adv Enzymol Relat Areas Mol Biol **63**: 69-172.

Gubin, A. N. and R. L. Kincaid (1998). "A pressure-extrusion method for DNA extraction from agarose gels." Anal Biochem **258**(1): 150-2.

Heller, K. and V. Braun (1979). "Accelerated adsorption of bacteriophage T5 to Escherichia coli F, resulting from reversible tail fiber-lipopolysaccharide binding." J Bacteriol **139**(1): 32-8.

Hendrix, R. W. (2003). "Bacteriophage genomics." Curr Opin Microbiol **6**(5): 506-11.

Hendrix, R. W. and S. R. Casjens (1974). "Protein cleavage in bacteriophage lambda tail assembly." Virology **61**(1): 156-9.

Hendrix, R. W. and R. L. Duda (1992). "Bacteriophage lambda PaPa: not the mother of all lambda phages." Science **258**(5085): 1145-8.

Hershey, A. D., Dover, W. (1983). "Introduction to Lambda". Lambda II. R. W. Hendrix, Roberts, J.W., Stahl, F.W. and Weisburg, R.A., eds. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory**:** 3-11.

Herskowitz, I. and D. Hagen (1980). "The lysis-lysogeny decision of phage lambda: explicit programming and responsiveness." Annu Rev Genet **14**: 399-445.

Hofnung, M., A. Jezierska, et al. (1976). "lamB mutations in E. coli K12: growth of lambda host range mutants and effect of nonsense suppressors." Mol Gen Genet **145**(2): 207-13.

Kanamaru, S., P. G. Leiman, et al. (2002). "Structure of the cell-puncturing device of bacteriophage T4." Nature **415**(6871): 553-7.

Katsura, I. (1976). "Morphogenesis of bacteriophage lambda tail. Polymorphism in the assembly of the major tail protein." J Mol Biol **107**(3): 307-26.

Katsura, I. (1981). "Structure and function of the major tail protein of bacteriophage lambda. Mutants having small major tail protein molecules in their virion." J Mol Biol **146**(4): 493-512.

Katsura, I. (1983). "Tail assembly and injection". Lambda II. R. W. Hendrix, Roberts, J.W., Stahl, F.W. and Weisburg, R.A., eds. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory**:** 331-346.

Katsura, I. and R. W. Hendrix (1984). "Length determination in bacteriophage lambda tails." Cell **39**(3 Pt 2): 691-8.

Katsura, I. and P. W. Kuhl (1975). "Morphogenesis of the tail of bacteriophage lambda. III. Morphogenetic pathway." J Mol Biol **91**(3): 257-73.

Katsura, I. and A. Tsugita (1977). "Purification and characterization of the major protein and the terminator protein of the bacteriophage lambda tail." Virology **76**(1): 129-45.

Kellenberger, E. (1961). "Vegetative bacteriophage and the maturation of the virus particles." Adv Virus Res **8**: 1-61.

Kellenberger, E., A. Bolle, et al. (1965). "Functions And Properties Related To The Tail Fibers Of Bacteriophage T4." Virology **26**: 419-40.

Kemp, C. L., A. F. Howatson, et al. (1968). "Electron microscopy studies of mutants of lambada bacteriophage. I. General description and quantitation of viral products." Virology **36**(3): 490-502.

Kemp, P., L. R. Garcia, et al. (2005). "Changes in bacteriophage T7 virion structure at the initiation of infection." Virology **340**(2): 307-17.

King, J. (1968). "Assembly of the tail of bacteriophage T4." J Mol Biol **32**(2): 231-62.

King, J. (1980). "Regulation of structural protein interactions as revealed in phage morphogenesis". Biological Regulation and Development. R. Goldberger, ed. New York, NY, Plebum Press**:** 101-132.

Kostyuchenko, V. A., P. G. Leiman, et al. (2003). "Three-dimensional structure of bacteriophage T4 baseplate." Nat Struct Biol **10**(9): 688-93.

Laemmli, U. K. (1970). "Cleavage of structural proteins during the assembly of the head of bacteriophage T4." Nature **227**(5259): 680-5.

Leiman, P. G., P. R. Chipman, et al. (2004). "Three-dimensional rearrangement of proteins in the tail of bacteriophage T4 on infection of its host." Cell **118**(4): 419-29.

Levin, M. E., R. W. Hendrix, et al. (1993). "A programmed translational frameshift is required for the synthesis of a bacteriophage lambda tail assembly protein." J Mol Biol **234**(1): 124-39.

Lof, D., K. Schillen, et al. (2007). "Forces controlling the rate of DNA ejection from phage lambda." J Mol Biol **368**(1): 55-65.

Makhov, A. M., B. L. Trus, et al. (1993). "The short tail-fiber of bacteriophage T4: molecular structure and a mechanism for its conformational transition." Virology **194**(1): 117-27.

Maniloff, J. and H. W. Ackermann (1998). "Taxonomy of bacterial viruses: establishment of tailed virus genera and the order Caudovirales." Arch Virol **143**(10): 2051-63.

Matsuo-Kato, H., H. Fujisawa, et al. (1981). "Structure and assembly of bacteriophage T3 tails." Virology **109**(1): 157-64.

Moak, M. and I. J. Molineux (2000). "Role of the Gp16 lytic transglycosylase motif in bacteriophage T7 virions at the initiation of infection." Mol Microbiol **37**(2): 345-55.

Moak, M. and I. J. Molineux (2004). "Peptidoglycan hydrolytic activities associated with bacteriophage virions." Mol Microbiol **51**(4): 1169-83.

Molineux, I. J. (2001). "No syringes please, ejection of phage T7 DNA from the virion is enzyme driven." Mol Microbiol **40**(1): 1-8.

Montag, D. and U. Henning (1987). "An open reading frame in the Escherichia coli bacteriophage lambda genome encodes a protein that functions in assembly of the long tail fibers of bacteriophage T4." J Bacteriol **169**(12): 5884-6.

Moody, M. F. (1973). "Sheath of bacteriophage T4. 3. Contraction mechanism deduced from partially contracted sheaths." J Mol Biol **80**(4): 613-35.

Moody, M. F. and L. Makowski (1981). "X-ray diffraction study of tail-tubes from bacteriophage T2L." J Mol Biol **150**(2): 217-44.

Mount, D. W., A. W. Harris, et al. (1968). "Mutations in bacteriophage lambda affecting particle morphogenesis." Virology **35**(1): 134-49.

Murialdo, H. and L. Siminovitch (1972). "The morphogenesis of bacteriophage lambda. IV. Identification of gene products and control of the expression of the morphogenetic information." Virology **48**(3): 785-823.

Novick, S. L. and J. D. Baldeschwieler (1988). "Fluorescence measurement of the kinetics of DNA injection by bacteriophage lambda into liposomes." Biochemistry **27**(20): 7919-24.

Parkinson, J. S. (1968). "Genetics of the left arm of the chromosome of bacteriophage lambda." Genetics **59**(3): 311-25.

Randall-Hazelbauer, L. and M. Schwartz (1973). "Isolation of the bacteriophage lambda receptor from Escherichia coli." J Bacteriol **116**(3): 1436-46.

Ray, P. N. and M. L. Pearson (1974). "Evidence for post-transcriptional control of the morphogenetic genes of bacteriophage lambda." J Mol Biol **85**(1): 163-75.

Roa, M. (1981). "Receptor-triggered ejection of DNa and protein in phage lambda." FEMS Microbiol. Lett. **11**: 257-262.

Roa, M. and D. Scandella (1976). "Multiple steps during the interaction between coliphage lambda and its receptor protein in vitro." Virology **72**(1): 182-94.

Robinson, A. S. and J. King (1997). "Disulphide-bonded intermediate on the folding and assembly pathway of a non-disulphide bonded protein." Nat Struct Biol **4**(6): 450-5.

Roeder, G. S. and P. D. Sadowski (1977). "Bacteriophage T7 morphogenesis: phage-related particles in cells infected with wild-type and mutant T7 phage." Virology **76**(1): 263-85.

Roessner, C. A. and G. M. Ihler (1984). "Proteinase sensitivity of bacteriophage lambda tail proteins gpJ and pH in complexes with the lambda receptor." J Bacteriol **157**(1): 165-70.

Roessner, C. A. and G. M. Ihler (1987). "Sequence of amino acids in lamB responsible for spontaneous ejection of bacteriophage lambda DNA." J Mol Biol **195**(4): 963-6.

Rossmann, M. G., R. Bernal, et al. (2001). "Combining electron microscopic with x-ray crystallographic structures." J Struct Biol **136**(3): 190-200.

Rossmann, M. G., V. V. Mesyanzhinov, et al. (2004). "The bacteriophage T4 DNA injection machine." Curr Opin Struct Biol **14**(2): 171-80.

Russell, H. (1990). "Recombinant PCR". PCR protocols: a guide to methods and applications. M. A. Innis, Gelfand, D. H., and Sninsky, J. J., eds. San Diego, CA, Academic Press**:** 177-183.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual, 2nd ed. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press.

Sampson, L. L., R. W. Hendrix, et al. (1988). "Translation initiation controls the relative rates of expression of the bacteriophage lambda late genes." Proc Natl Acad Sci U S A **85**(15): 5439-43.

Sargent, D., J. M. Benevides, et al. (1988). "Secondary structure and thermostability of the phage P22 tailspike. XX. Analysis by Raman spectroscopy of the wild-type protein and a temperature-sensitive folding mutant." J Mol Biol **199**(3): 491-502.

Sather, S. K. and J. King (1994). "Intracellular trapping of a cytoplasmic folding intermediate of the phage P22 tailspike using iodoacetamide." J Biol Chem **269**(41): 25268-76.

Sauer, R. T., W. Krovatin, et al. (1982). "Phage P22 tail protein: gene and amino acid sequence." Biochemistry **21**(23): 5811-5.

Scandella, D. and W. Arber (1976). "Phage lambda DNA injection into Escherichia coli pel-mutants is restored by mutations in phage genes V or H." Virology **69**(1): 206-15.

Schwartz, M. (1976). "The adsorption of coliphage lambda to its host: effect of variations in the surface density of receptor and in phage-receptor affinity." J Mol Biol **103**(3): 521-36.

Serwer, P., E. T. Wright, et al. (2008). "Evidence for bacteriophage T7 tail extension during DNA injection." BMC Res Notes **1**: 36.

Simon, L. D. and T. F. Anderson (1967). "The infection of Escherichia coli by T2 and T4 bacteriophages as seen in the electron microscope. I. Attachment and penetration." Virology **32**(2): 279-97.

Simon, L. D. and T. F. Anderson (1967). "The infection of Escherichia coli by T2 and T4 bacteriophages as seen in the electron microscope. II. Structure and function of the baseplate." Virology **32**(2): 298-305.

Stahl, F. W. and N. E. Murray (1966). "The evolution of gene clusters and genetic circularity in microorganisms." Genetics **53**(3): 569-76.

Steinbacher, S., R. Seckler, et al. (1994). "Crystal structure of P22 tailspike protein: interdigitated subunits in a thermostable trimer." Science **265**(5170): 383-6.

Studier, F. W. (1972). "Bacteriophage T7." Science **176**(33): 367-76.

Szmelcman, S. and M. Hofnung (1975). "Maltose transport in Escherichia coli K-12: involvement of the bacteriophage lambda receptor." J Bacteriol **124**(1): 112-8.

Tsui, L. C. and R. W. Hendrix (1983). "Proteolytic processing of phage lambda tail protein gpH: timing of the cleavage." Virology **125**(2): 257-64.

Vianelli, A., G. R. Wang, et al. (2000). "Bacteriophage T4 self-assembly: localization of gp3 and its role in determining tail length." J Bacteriol **182**(3): 680-8.

Wang, J., M. Hofnung, et al. (2000). "The C-terminal portion of the tail fiber protein of bacteriophage lambda is responsible for binding to LamB, its receptor at the surface of Escherichia coli K-12." J Bacteriol **182**(2): 508-12.

Weigle, J. (1966). "Assembly of phage lambda in vitro." Proc Natl Acad Sci U S A **55**(6): 1462-6.

Weigle, J. (1968). "Studies on head-tail union in bacteriophage lambda." J Mol Biol **33**(2): 483-9.

Werts, C., V. Michel, et al. (1994). "Adsorption of bacteriophage lambda on the LamB protein of Escherichia coli K-12: point mutations in gene J of lambda responsible for extended host range." J Bacteriol **176**(4): 941-7.

Wommack, K. E. and R. R. Colwell (2000). "Virioplankton: viruses in aquatic ecosystems." Microbiol Mol Biol Rev **64**(1): 69-114.

Xu, J. (2001). "A conserved frameshift strategy in dsDNA long tailed bacteriophages", University of Pittsburgh. **Ph.D. thesis**.

Youderian, P. (1978). "Genetic analysis of the length of the tails of lambdoid bacteriophages", Massachusetts Institute of Technology. **Ph.D. thesis**.

Zhao, L., S. Kanamaru, et al. (2003). "P15 and P3, the tail completion proteins of bacteriophage T4, both form hexameric rings." J Bacteriol **185**(5): 1693-700.

Zhao, L., S. Takeda, et al. (2000). "Stoichiometry and inter-subunit interaction of the wedge initiation complex, gp10-gp11, of bacteriophage T4." Biochim Biophys Acta **1479**(1-2): 286-92.