Characterization of morphogenetic functions of transposable bacteriophages Mu, D108 and D3112

By

Ian Russell Siboo

Department of Microbiology and Immunology

McGill University

A thesis submitted to the

Faculty of Graduate Studies and Research

in partial fulfillment of the requirements

of the degree of Doctor of Philosophy

Abstract

Morphogenesis of the dsDNA phages is a highly regulated process that utilizes multicomponent pathways to assemble precursors of heads and tails. These mechanisms have been extensively investigated in many phage systems but have yet to be well characterized for the transposable phages. To investigate morphogenesis in these systems, we have cloned and sequenced 4.5 kb the of the DNA of phage Mu immediately downstream of the first late promoter P_{lys}; expressed the associated proteins; examined phage development in thin section; and refined our existing *in vitro* packaging assay Previous amber mutant analysis had located the lytic and maturation genes of Mu in the P_{lys} transcript. Sequence data were derived from this region and were analyzed with various computer algorithms, to characterize all potential open reading frames (ORFs), and to assign function to some ORFs based on homology to previously characterized genes. Sequencing of *lys*, *D* and *E* amber mutants was used to assign these genes to specific ORFs. Cloning and expression of the region proximal to the P_{lys} promoter resulted in the production of protein extracts that specifically recognize the phage-encoded packaging signal, characteristic of phage maturation enzymes. Additionally, the growth characteristics of strains expressing these late gene constructs show that lysis of the cultures is induced as a result of the presence of one or more phage-specific proteins expressed from this region. Our studies indicate a deviation from existing phage lysis paradigms. Thin section electron micrographs show that the development of the structures from induced prophages occurs relatively late in the lytic cycle. Sucrose gradient fractions highly enriched for proheads are associated with *pac*-specific DNA

recognition functions. A specific DNA substrate can be packaged, *in vitro*, into these active proheads. Packaging of the DNA substrate may be improved by the inclusion of purified, *E. coli* HU or IHF.

Résumé

La morphogénèse des phages à double-brin d'ADN se poursuit selon un procédé strictement regulé qui utilise des cheminements à plusieurs étapes dans le but d'assembler les précurseurs de têtes et de queues des phages. Ces cheminements ont été étudié en détail pour plusieurs types de bactériophages, mais n'ont pas été caractérisé chez les bactériophages qui utilisent la transposition comme méthode de réplication. Pour étudier la morphogénèse dans ces systèmes, nous avons cloné et séquencé 4.5 kb d'ADN du bactériophage Mu situé directement downstream du promoteur tardif P_{lys}; exprimé les protéines associées à cette région d'ADN; examiné le dévelopement de Mu par tranches minces; et raffiné notre essai d'emballement *in vitro*. De précédentes analyses de mutants amber avaient localisé les gènes lytiques et de maturation du Mu dans le transcript de P_{lys}. La séquence de cette région a été déterminée et analysée par divers algorithmes informatiques afin d'identifier tous les cadres de lecture ouverte (CLO) potentiels, et pour assigner des fonctions à ces CLOs par comparaison avec des gènes homologues caractérisés précédemment. Le clonage et l'expression d'une région d'ADN proximale à P_{lys} a résulté en la production d'extraits protéiniques qui étaient capables de reconnaître spécifiquement le signal d'emballement encodé par le phage. Cette abilité est une caractéristique d'enzymes de maturation de phages. De plus, la culture de souche bactériennes exprimant des constructions plasmiques contenant la région tardive démontre que la lyse des cultures est induite par une ou des protéines encodée(s) par cette région. Notre étude suggère que la façon par laquelle Mu lyse son hôte est différente des paradigmes établis précédemment chez d'autres phages. Des micrographes électroniques

de tranches minces démontrent qu'après l'induction du prophage Mu, le développement de particles virales se produit relativement tard durant le cycle lytique. Des fractions protéiniques obtenues à partir d'extraits protéiniques post-induction, séparés par gradient linéaire de sucrose, sont enrichies en précurseurs de têtes capables de se lier spécifiquement au site d'ADN d'emballage de Mu, appelé *pac*. Un subtrat d'ADN hétérologue, mais contenant des sites *pac*, peut aussi être emballé, *in vitro*, par ces structures pré-têtes. La réaction d'emballage peut être améliorée par l'addition de HU et IHF, deux protéines encodées par l'hôte.

Acknowledgements

I would like to thank my supervisor Michael DuBow for his help and support during my studies. Mike introduced me to bacteriophages and gave me an appreciation for science for which I will always be grateful.

Thanks to Dr. H.W. Ackermann and Kathy Hewitt for sharing their expertise about electron microscopy. Without them I would never have had access to a world that is truly amazing.

I am indebted to many of the members of the DuBow lab for their help and support. In particular, I would like to thank Dr. Gina MacIntyre, Madani Thiam, and Caroline Diorio. They helped me maintain my perspective and sanity.

I would like to acknowledge the financial support of the Fonds pour la Formation de Chercheurs et L'Aide à la Recherche and the Department of Microbiology and Immunology.

Finally I have to thank the most important people in my life: Doris Fortin for her wisdom, intelligence and humour; my brother Neil for always being someone I can count on; my parents Russell and Ulla, the best people I know.

Preface to the thesis

I have opted to present the experimental portion of this thesis in the form of manuscripts in accordance with the "Guideline for Thesis Preparation" and the approval of the Department of Microbiology and Immunology. An excerpt of the guidelines follows:

> "Candidates have the option of including, as part of the thesis, the text of one or more papers submitted or to be submitted for publication, or the clearly duplicated text of one or more published papers. These texts must be bound as an integral part of the thesis.
>
> If this option is chosen, connecting texts that provide logical bridges between the different papers are mandatory. The thesis must be written in such a way that it is more than a mere collection of manuscripts; in other words, results of a series of papers must be integrated.
>
> The thesis must still conform to all other requirements of the 'Guidelines for Thesis Preparation'. The thesis must include: A Table of Contents, an abstract in English and French, an introduction which clearly states the rationale and objectives of the study, a review of the literature, a final conclusion and summary, and a thorough bibliography or reference list.
>
> Additional material must be provided where appropriate (e.g. in appendices) and in sufficient detail to allow a clear and precise judgement to be made of the importance and originality of the research reported in the thesis.
> In the case of manuscripts co-authored by the candidate and others, the candidate is required to make an explicit statement in the thesis as to who contributed to

such work and to what extent. Supervisors must attest to the accuracy of such statements at the doctoral oral defense. Since the task of the examiners is made more difficult in these cases, it is in the candidate's interest to make perfectly clear the responsibilities of all the authors of the co-authored papers."

Chapters 2and 3 of this thesis are two original papers which will be submitted shortly for publication. The manuscripts in order of appearance are:

1. Siboo, I. R., F. Sieder, M.M Howe and M.S. DuBow. 1999. In preparation. Characterization of the enzymatic functions encoded within the P_{lys} transcript of bacteriophage Mu.

2. Siboo, I. R. and M.S. DuBow. 1999. In preparation. Characterization of the isolated proheads from transposable phages.

As has been attested by my thesis supervisor, Dr. Michael S. DuBow, I was responsible for the research described in chapters 2 and 3. Felix Sieder constructed pFS100 and its deletion derivatives. Using these constructs, Felix sequenced the 3711bp of the phage Mu genome downstream of P_{lys}. Dr. Howe provided the positions of the amber mutations for the *lys* and *D* genes.

Original Contributions to Knowledge

1. I have identified seven new potential open readings frames encoded within the P_{lys} transcript of Bacteriophage Mu.
2. I have demonstrated that the lysis of bacteriophage Mu is dependent on a protein which regulates the release of the phage endolysin from the bacterial cytoplasm.
3. I have demonstrated that the lysis regulator is not encoded, in the phage Mu, in a manner analogous to that reported for other phage lysis operons.
4. I have shown that there is a protein(s) encoded within the P_{lys} region, containing the above mentioned seven ORFs, that is able to specifically recognize the Mu packaging signal.
5. I have shown, using thin section electron micrographs, that two transposable phages, Mu and D3112, produce morphogenetic precursors late in the lytic cycle, after the initiation of host lysis
6. I have shown by electron microscopy and mobility shift assay that phage proheads can be enriched from crude lysates using sucrose gradients, and that these proheads are associated with a packaging signal recognition function.
7. I have shown that the major protein present within preparations of isolated D3112 proheads has a molecular mass of 37kDa
8. I have refined an existing in vitro packaging assay and shown that isolated Mu proheads are able to package a phage-specific substrate. The packaging reaction is dependent on ATP and spermidine and can be enhanced by two bacterially encoded proteins HU and IHF.

List of Abbreviations

a.a.:	Amino Acid
ATP:	Adenosine Triphosphate
BHI:	Brain Heart Infusion
BLAST:	Basic Local Alignment Sequence Tool
bp:	Basepair
BSA:	Bovine Serum Albumin
CTAB:	Hexadecyltrimethyl ammonium bromide
CTP:	Cytosine Triphosphate
DNA:	Deoxyribonucleic Acid
DTT:	Dithiothreitol
EDTA:	Ethylenediaminetetraacetic acid
gp:	Gene Product
HEPES:	N-(2-Hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid)
HTH:	Helix-Turn-Helix
IHF:	Integration Host Factor
IPTG:	Isopropyl-β-D-Thiogalactopyranoside
kb:	Kilobase
kDa:	Kilodalton
LB:	Luria Broth
ORF:	Open Reading Frame
PCR:	Polymerase Chain Reaction
PMSF:	Phenylmethyl-Sulfonylfluoride
pRNA:	Packaging RNA
PVDF:	Polyvinylidenedifluoride
RBS:	Ribosome Binding Site
RNA:	Ribonucleic Acid
RNAP:	RNA Polymerase
SDS-PAGE:	Sodium Dodecylsulfate Polyacrylamide Gel Electrophoresis

List of Figures and Tables

Chapter 1

Figure 1: Assembly pathway for phage proheads 2
Figure 2: Geometrical characteristics of the phage prohead 4
Figure 3: Assembly pathway of phage portals 13
Table 1: 14
Figure 4: Pathway of scaffold/capsid protein interaction 26
Figure 5: The Mu genome 49

Chapter 2

Figure 1:Partial genetic map of bacteriophage Mu and potential ORFs 79
Figure 2: Positions of P_{ly} proximal amber mutants 81
Figure 3: Growth profiles of *E. coli* Bl21(DE3) strains containing phage constructs 83
Figure 4: Mu-*pac* DNA binding properties of extracts containing phage proteins as determined by mobility shift analysis. 84
Figure 5: Mobility shift assay of the pWM5 extract and individual probes 87

Chapter 3

Figure 1: Electronmicroscopic examination of thin sections from lysogens 105
Figure 2: Protein profiles of sucrose gradient fractions 107
Figure 3: Electronmicroscopic images of phage structures 109
Figure 4: Protein content of D3112 prohead fractions 111
Figure 5: Mobility shift assay of sucrose gradient fractions 112
Figure 6: Characterization of refined *in vitro* packaging reactions 114

Table of Contents

Chapter 1: General Literature Review

Introduction	1
The Capsid Shell	3
Capsid Geometry	3
Shell Proteins	5
Isometric Shell Assembly	6
Sizing of Isometric Capsids	7
Sizing of Prolate Capsids	9
The Portal	11
Structure and Function	12
Terminase Interaction	16
Measurement of Packaged DNA	17
Regulation of Packaging Initiation	18
Role in Prohead Assembly	19
The Scaffold	25
Prohead Assembly-Initiation and Fidelity	27
Scaffold Structure and Protein-Protein Interactions	29
Capsid Sizing	34
An Alternative Scaffold-Sid	35
Sid Scaffold Structures	36
Modified Assembly Mechanisms	39
HK97	39
Crosslinking of the Major Capsid Protein	40
Regulation of Head Assembly	42
Crosslinking Mechanism	43
Capsid Scaffold Function	44
ϕ29 Packaging RNA	45
Bateriophage Mu	47
Overview	47
Lifecycle	48
Mu Transposition	50
Regulation of Transcription	56
Middle Transcripition – Mor	57
Late Transcription – C	59
Mu Late Genes Morphogenesis and Packaging	61

Chapter 2: Characterization of the Enzymatic Functions of the P_{lys} Transcript of Bacteriophage Mu.

Abstract	66
Introduction	67

Table of Contents (continued)

Materials and Methods 72
Media and reagents 72
Bacterial strains and plasmids 72
DNA manipulations 74
Production of crude extracts containing phage proteins 75
Growth curves of strains expressing phage proteins 76
Mobility Shift Assays 76
Results 78
Sequencing of the early morphogenetic functions of phage Mu 78
Protein expression effects on cell growth 80
Mobility Shift analysis of crude extracts from pWM5 and pKN35 83
Discussion 88

Chapter 3: Characterization of Isolated Proheads from Transposable Phages

Abstract 93
Introduction 94
Material and Methods 97
Media, bacterial strains and plasmids 97
Production of crude packaging extracts 97
Separation of morphogenetic precursors from crude packaging extracts 98
Western blotting 99
Mobility shift assays on gradient fractions 99
Refined *in vitro* packaging assay 100
Electron microscopy of growing phage *in vivo* and isolated phage structures 102
Results 104
In vivo development of transposable bacteriophages 104
Isolation of phage products by sucrose gradient centrifugation 106
The composition of prohead gradient fractions 108
Pac-containing DNA binding activity of prohead-containing fractions 110
In vitro packaging 113
Discussion 116

Chapter 4: Conclusions

Conclusions 123

Chapter 5: References

References 128

Appendix 1 145
Appendix 2 147

Chapter One

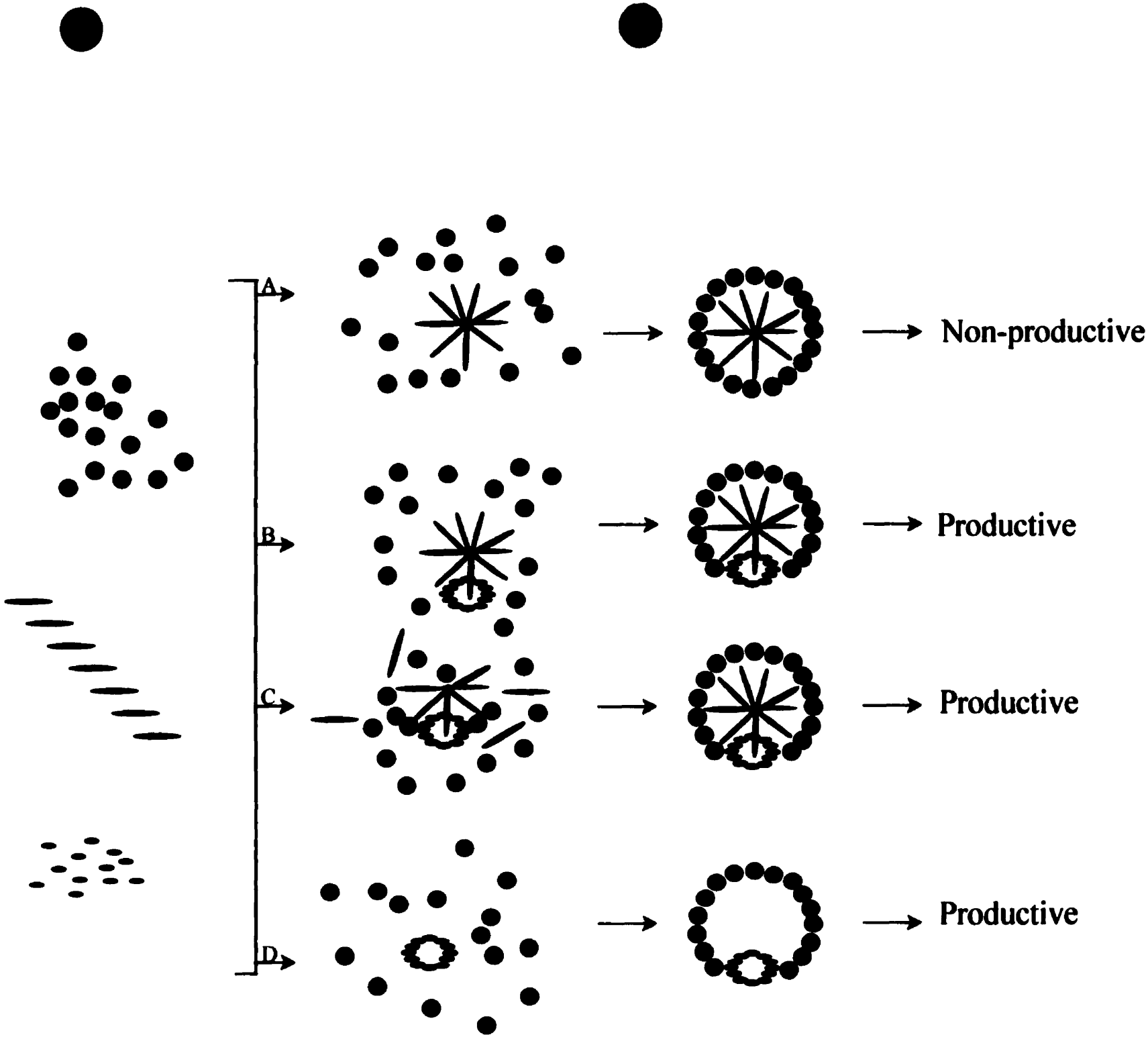
General Literature Review

Introduction

The discovery of mammalian viruses, and perhaps more importantly, bacteriophages, have allowed researchers to explore the intricate mechanisms of various biological processes. Twort (Twort, 1915) and d'Herelle (d'Herelle, 1917) first described an "autolytic principle" of bacteria that was later shown to be the prokaryotic counterpart to the eukaryotic viruses. Similar to animal viruses, the bacteriophages (phages) parasitized their hosts in order to replicate their constituent genomes, and produce progeny phage. The progeny can in turn infect other hosts, thereby maintaining their genetic autonomy. Unlike their eukaryotic counterparts, bacteriophages produce progeny at a relatively fast rate, and are most importantly, only pathogenic in bacteria. As knowledge of phages accumulated, their use in basic biological research grew tremendously. Ultimately, research with phages and their hosts led to a marked comprehension of universal biological processes and to the development of modern molecular biology.

During its lifecycle, a phage goes through distinct phases, all of which have been the focus of intense research. Towards the end of the lytic cycle, constituents of the mature phage are produced in a manner similar to an assembly line. The phage prohead (Figure 1), the precursor to the mature head, and the tail are produced by distinct morphogenetic pathways that are regulated at the transcriptional, translational and structural level (Casjens and Hendrix, 1988). Examination of these processes has allowed investigators to gain an understanding of the assembly of macromolecular structures, the interaction of the individual components and the regulation of these interactions. These paradigms are applicable not only to phages, but other biological systems as well.

Figure 1. Assembly pathways for phage proheads. The assembly of a phage capsid precursor (prohead) occurs upon the proper interaction of its constituent subunits. Despite the necessity for the formation of a structure with a strictly defined geometrical shape, a number of mechanisms had been proposed and demonstrated for the assembly of the prohead. The capsid protein is represented by the circles, the portal by the small ovals and the scaffold by rods. (A) The phage scaffold and capsid proteins may interact to form a closed spherical structure virtually identical to a true prohead. However, this structure contains no portal for the entry/exit of the DNA and is therefore non-functional. (B) Assembly of the prohead may occur through the distinct interaction of precursor proteins. A core may form on which the capsid proteins polymerize forming a functional prohead. (C) The prohead may form through the co-ordinated interaction of all precursor subunits. Such a pathway would also lead to the formation of a functional prohead. (D) Finally, the capsid and portal proteins may assemble in the absence of the scaffold, producing a functional prohead. This assembly mechanism requires that the size and shape determination of the phage head not be carried out by the scaffold protein.


A
B
C
D
Non-productive
Productive
Productive
Productive

Head and tail assembly proceed through a number of distinct steps that produce structures capable of interacting to create progeny phage. Both pathways offer insights into distinct mechanisms of protein oligomerization. However the following review will be devoted to prohead assembly of double-stranded (dsDNA) phages.

The Capsid Shell

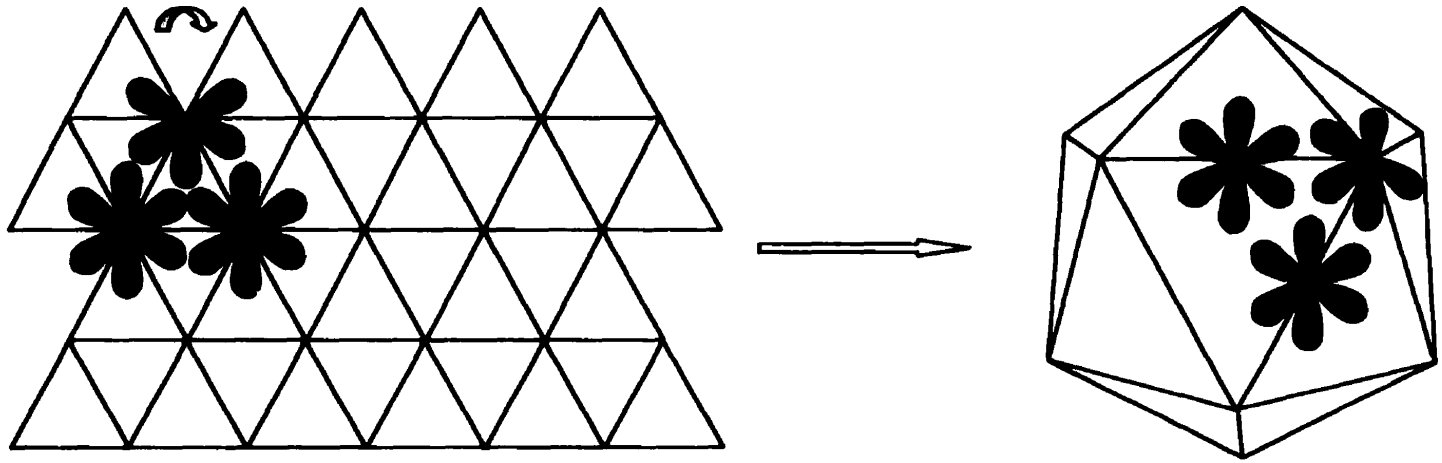
The head or capsid of a tailed dsDNA bacteriophage is composed of at least two proteins. The major capsid protein and the portal protein form a shell in which the genome is housed and to which the tail is attached. The heads of most phages have additional proteins, on or in the head, which provide stability or have distinct roles for packaging/injection of the genome (Casjens and Hendrix, 1988). Additionally, the prohead, a precursor form of the head, has a distinct structure and contains extra components that are not present in the mature capsid. The shape of the capsid is distinct and is modified according to strict geometrical rules, as the size of the head changes.

Capsid Geometry

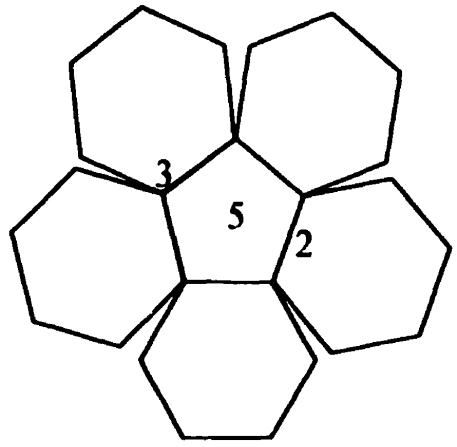
Caspar and Klug (1962) were the first to describe mathematical rules that govern the construction of icosahedral capsids of both viruses and phages. Icosahedra are closed structures that have 5:3:2 cubic symmetry (Figure 2). They contain 12 pentameric vertices and 20 regular faces. The capsid proteins of phages do not take on a pentameric form and therefore there is a minimum of 60 subunits required to construct an

Figure 2. Geometrical characteristics of the prohead. (A) An icosahedron may be formed from a planar lattice. Removing one of the subunits of the hexamer, initiates the formation of a doubly curved surface and creates a pentameric vertex. The closed icosahedron is constructed by creating twelve such pentamers. (B) The symmetry of the basic icosahedron is 5:3:2, which describes the penatmeric vertices, the center of the planar faces and the planar edges.

A



B



icosahedron. Phage capsids are described by multiples of the basic 60-subunit structure, which is identified by a so-called triangulation (T) number. It is not possible, moreover, to place more than sixty subunits, in identical environments, on an icosahedral surface. Caspar and Klug (1962) offered a solution to this problem. They stated that the subunits could be placed in quasi-equivalent environments such that the bonds that held the structure together were of the same type throughout the head, but the individual bonds might be slightly non-symmetrical. The quasi-equivalent environments into which the protein subunits fit support the construction of a large number of capsids of different sizes and therefore can potentially house genomes of any size.

Shell Proteins

The capsids of many dsDNA phages have been studied and a considerable amount of information is known about these structures. While these assemblages are icosahedral, a portal is substituted at one vertex that perturbs its symmetry (Casjens and Hendrix, 1988). The presence of a portal is absolutely necessary for a phage to form a functional prohead (Bazinet and King, 1985) and, consequently, quasi-equivalent interactions by the capsid proteins are essential. The shells themselves are composed of a defined number of a single, or in some cases two, subunits (Casjens and Hendrix, 1988). Additionally, phage heads contain extra proteins, that are added after the maturation of the rounded proheads to the more angular form found in mature phages. These extra proteins may increase the stability of the prohead so that it withstands pressures caused by the packaging of its genome, and environmental stresses that are encountered outside the host (Casjens and

Hendrix, 1988). In phage λ, gpD forms trimers on the shell that is made up of gpE (Casjens and Hendrix, 1988). In phage T4, gp24 forms pentamers at the 5-fold vertices while the rest of the shell is composed of gp23 (Black *et al.* 1994). In addition, two proteins, Hoc and Soc, stabilize the head. Hoc is positioned at the center of the hexameric capsomers, while Soc surrounds the hexamers forming trimeric structures (Black *et al.* 1994). Thus, while the shells themselves may be highly decorated, the phage heads are essentially icosahedral structures that possess one of two forms: isometric or prolate.

Isometric Shell Assembly

The isometric phages whose structures have been studied extensively include λ (Dokland and Murialdo, 1993), HK97 (Conway *et al.* 1995), P2/P4 (Dokland *et al.* 1992), P22 (Prasad *et al.* 1993) and T7 (Casjens and Hendrix, 1988). All of these phages, with the exception of P4, have a triangulation number of 7 and are composed of 60 hexamers, 11 pentamers and a portal. Due to the presence of these distinct geometric structures within the shell lattice, prohead assembly mechanisms have been examined to determine whether or not the monomeric or the oligomeric capsid proteins, are assembly precursors. This process has been examined in HK97 (Xie and Hendrix, 1995), P2/P4 (Marvik *et al.* 1995) and P22 (Prevelige, Jr. *et al.* 1993; Prevelige, Jr. and King, 1993a).

Proheads of HK97 are efficiently assembled from pools of pentamers and hexamers that are produced upon expression of the sole capsid protein, gp5 (Xie and Hendrix, 1995). When present in the normal pentamer to hexamer ratio of 1:6, the oligomers can assemble into head-like structures. Hendrix and Duda (1998) have

reported, however, that mutant forms of gp5 skew monomer oligomerization exclusively to pentamer or hexamer, and proheads are not formed. It is thought, therefore, that there is a natural mechanism that maintains a pool of oligomeric precursors in the proper proportion. A mechanism for the assembly of P2 has been hypothesized from data obtained from cryo-reconstructed images of its parasitizing phage P4 (Bertani and Six, 1988). P4 assembles within a secondary scaffold lattice that is made up of the P4 protein Sid, composed of a defined number of subunits (Marvik *et al.* 1995). It was proposed, that in order for the individual Sid subunits to efficiently and productively assemble a P4 prohead, the Sid protein had to interact with hexamers of the capsid protein. Since the capsid protein present in the P4 shell is in fact a P2 protein, it is theorized that P2 utilizes an oligomeric intermediate for shell assembly (Marvik *et al.* 1995). In the case of P22, kinetic studies have indicated that it is the monomeric form of the protein that is the assembly intermediate (Prevelige, Jr. and King, 1993a). Thus, it appears as though these isometric phages utilize both monomeric and oligomeric substrates for the assembly of their proheads. However, despite the differences in the oligomerization state of the prohead precursor material, all of the assembly pathways lead to the production of a structure of defined size.

Sizing of Isometric Capsids

The size of a prohead is of intrinsic importance, since the internal volume of its shell must be of an appropriate size to incorporate the phage genome in a stable manner. With the exception of phage P4, the size determination function has yet to be definitively

demonstrated in the isometric phages. A recent study, however, indicates that it is partially an inherent property of the capsid protein. Detailed images were reconstructed from cryo-electron micrographs of the P22 prohead that was produced from its major capsid protein. Analysis of the images revealed that structures with triangulation numbers of T=4 and T=7 can be formed (Thuman-Commike *et al.* 1998). A similar change in triangulation number is thought to occur in phage lambda when there are specific point mutations in its major head protein, gpE [referenced in (Thuman-Commike *et al.* 1998)]. Inspection of the P22 pentons and hexons showed that these oligomers were of similar dimension, and that the variations observed in the size of the proheads were due to the placement of the oligomers in different orientations. The rules of quasi-equivalence imply that the number of distinct subunit conformations is equivalent to the triangulation number (Caspar and Klug, 1962). However, due to the conformation of the subunits within the hexon, the T=7 prohead could be formed from the same conformations used in a T=4 structure (Thuman-Commike *et al.* 1998). To resolve this contradiction the authors stated that assembly depended only on interactions of neighbour subunits [referenced in (Thuman-Commike *et al.* 1998)]. Thus, the mechanism by which the T=7 prohead is assembled requires that the interactions leading to a T=4 head be blocked. This role is likely to be filled, in this and other isometric systems, by the scaffold (see below), which was not present in the P22 studies. Despite the need for a phage encoded chaperone that directs the correct sizing of the head, the presence of heads of only two sizes indicate that the capsid protein imposes an intrinsic limitation on that dimension.

A second theory of prohead sizing by capsid intermediates, has been proposed (Hendrix and Duda, 1998). The isometric proheads of P22, λ and HK97 were imaged via

cryo-electron microscopy, and the proheads were shown to contain hexon subunits that exhibited a distinctive skew. This skew is greatest in the HK97 hexons. Unlike P22, HK97 assembles from pre-formed hexons/pentons (Xie and Hendrix, 1995). It was proposed that the HK97 prohead was formed by a stepwise process that included the assembly of hexons, pentons and superpentons (Conway *et al.* 1995). The superpentons have a centrally located penton and five radially arranged hexons. The skewness of the hexons is thought to be an assembly signal that can block further penton-hexon interactions, and may allow hexons to join to other superpentons. The skew present in hexamers is lost upon expansion of the head when protein re-arrangements lead to oligomers with six-fold symmetry (Hendrix and Duda, 1998). This reorientation of the hexamer to change the symmetry is also seen in λ and P22 (Dokland and Murialdo, 1993; Thuman-Commike *et al.* 1996). Although the common hexon skew is seen in these systems, HK97 does not utilize an autonomous scaffold (Xie and Hendrix, 1995). Thus, while the skewing of the hexamers may play a role in the sizing of the HK97 prohead, the skew may be involved in other processes, in isometric phages, which encode distinct scaffolding proteins.

Sizing of Prolate Capsids

Prolate phages, including T4 and ϕ29, are characterized by an elongation of the head parallel to the portal axis. Numerous biophysical studies were devoted to understanding the mechanism of assembly and maturation of the T4 head. Unlike other phages, the T4 prohead is assembled on the inner portion of the bacterial cell membrane

(Black *et al.* 1994). During this process, the two major capsid proteins (gp23 and gp24) assemble around a core consisting of a portal (gp20), scaffold (gp22) and a number of extra internal proteins (gp21, gp67, gp68, IPI, IPII, and IPIII). The size/form determining function of this prolate structure seems to be determined by a number of proteins (Keller *et al.* 1986; Keller *et al.* 1988). Amber mutant analyses of two distinct gp67 variants revealed that a large proportion of the prohead structures that were produced were in fact isometric. In addition, isometric structures were observed from infections with mutant forms of gp68 and gp22 (Keller *et al.* 1988). Thus, the prohead seems to be sized by two core proteins (gp 67 and 68) that are not themselves scaffold proteins (Black *et al.* 1994). However, multi-tailed phages were also observed in these mutant infections (Keller *et al.* 1988). The addition of extra tail structures is indicative of a role for the tail attachment site in this mechanism. The portal structure is absolutely essential for the formation of the prohead since the assembly is known to be dependent on the formation of the ring in the inner membrane of the bacterium. (Black *et al.* 1994). Therefore, direct *in vitro* analysis of the size determining function is not possible. Similar analyses, however, was performed with ϕ29, another prolate phage (Guo *et al.* 1991).

Phage ϕ29 does not assemble its prohead on the inner membrane of the bacterium (Salas, 1988) and thus simplified, *in vitro* assembly experiments are possible. When prohead-like structures of ϕ29 were produced from scaffold and capsid protein, a mixture of products of varying sizes was observed (Guo *et al.* 1991). When the portal, however, was added to the reaction, the products were uniformly of normal dimension. Thus, the size of the ϕ29 capsids seemed to be determined by the portal structure. ϕ29 proheads are lengthened along a five-fold axis, therefore the portal may be required to initiate the

correct placement of scaffold proteins to ensure a suitable pitch for the assembly of the prohead. This pitch may guarantee that the structure lengthens in the proper dimension.

The first images of a prolate phage reconstructed from cryo-electron micrographs were recently published (Tao *et al.* 1998). Analysis of the reconstructed images of ϕ29 has shown that the portal⁻ isometric form of the phage prohead is a T=3 structure. The wild type phage is an extended form of this structure, but it has an extra ring of hexamers that is located equatorially. Furthermore, analysis of these images had shown that, unlike the previously reconstructed images of true isometric phages, the hexamers of the ϕ29 prohead had no skew (Tao *et al.* 1998). The skewing of the hexamers in the isometrics, as mentioned above, has been hypothesised to play a role in prohead sizing. Since this function seems to be provided by the portal in ϕ29, the skew may not be necessary. In addition, the isometrics all increase the volume of their heads during DNA packing by expansion of the prohead. The ϕ29 prohead undergoes no such expansion (Tao *et al.* 1998) and therefore a differential conformation for the hexamers in the prohead and mature capsid may be unnecessary. Thus, distinct differences in the conformation of the capsid structures, viewed in detailed images, may reveal more information regarding the functional relationships of the constituents of the shell. While the prohead shell itself is a dynamic structure capable of impressive conformational rearrangements, a large body of work has also been devoted to the other distinct sub-components that are essential for capsid assembly.

The Portal

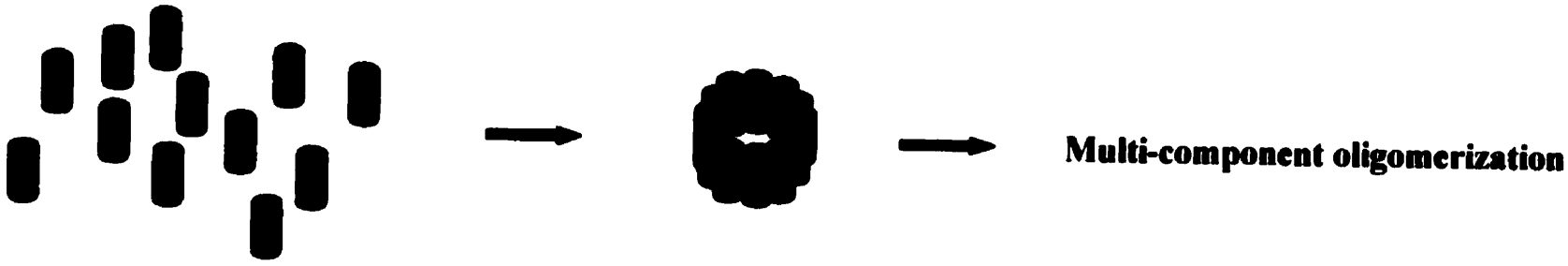
The bacteriophage capsid, the housing for a phage genome, is in its final form a closed structure. As with any closed form that must allow the entry or exit of material contained within it, the bacteriophage head has incorporated a proteinaceous "doorway" which permits the mobilization of its genome into and out of its head. This doorway is called a portal or connector (Fig. 3). The portal acts throughout the morphogenetic process, playing a role in head assembly and packaging, and is also instrumental in the completion of the phage structure with the addition of a pre-formed tail. In fact, tails examined from disrupted phages showed a disk shaped protrusion located on the head-proximal end (Tsui and Hendrix, 1980; Coombs and Eiserling, 1977). This protrusion is the portal. The strong association between the portal and the tail subsequent to phage completion is essential for the proper assembly of the phage. The attachment of the tail is essential for the interaction between the phage and host, and the subsequent injection of DNA. The portal structure of dsDNA bacteriophages has been extensively investigated and the many functions that it executes described (for review see (Bazinet and King, 1985)). Yet, the precise structure and functional mechanism of the portal are still somewhat unclear.

Structure and Function

The portal itself is made up of closed ring, containing a 2-5 nm centrally located hole, of either 12 or 13 subunits polymerized from a unique phage encoded protein (Bazinet and King, 1985; Dube *et al.* 1993). Many of the portal genes have been identified and cloned, and are listed in Table 1. The portal occupies one of the 12 vertices of the icosahedral phage head, by replacing one of the pentamers (Bazinet and King,

Figure 3 Assembly pathways of phage portals. The phage portal is a unique prohead structure located at one of the five fold symmetrical vertices of the head. The ring structure creates an opening that is utilized for DNA packaging, DNA injection and tail attachment. During prohead assembly, the portal protein oligomerizes to form an annular structure composed of 12 or 13 subunits. The ring is then assembled into the prohead as part of a nucleation or co-oligomerization mechanism. The actual assembly of the portal occurs by one of two mechanisms. (A) In most phage systems, the portal subunits are produced in the cytoplasm, then oligomerize upon reaching critical concentrations. Subsequent to the formation of the ring, the complete structure is incorporated into the assembly pathway. (B) Phage T4 utilizes a unique mechanism to assemble the portal. Subsequent to the production of the protein subunits, the portal monomers associate with the cytoplasmic face of the bacterial inner membrane. At the membrane, the portal forms and initiates assembly of the prohead core which precedes the formation of the prohead itself.

A



B

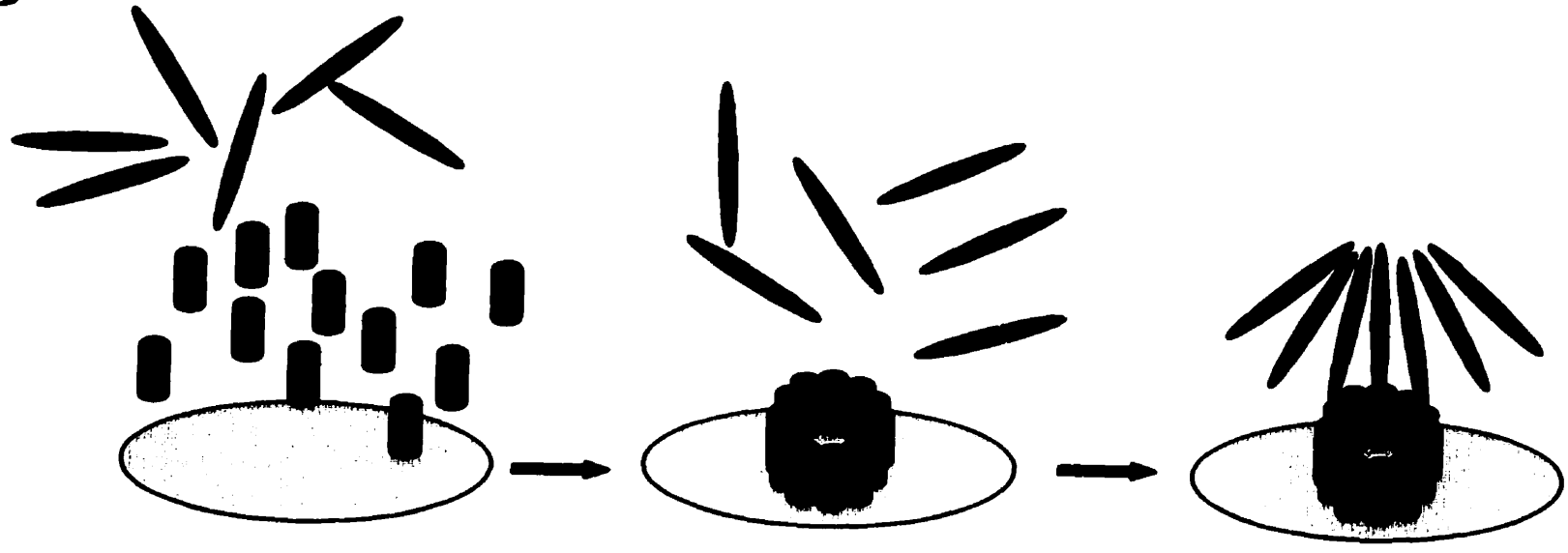


Table 1

Phage	Portal Gene	Mol. Wt. (kDa Per Subunit)	References
P22	gp1	94	Bazinet *et. al.*, 1988
SPP1	gp6	68	Becker *et. al.*, 1997
ϕ29	gp10	36.4	Ibanez *et. al.*, 1984
T4	gp20	65	Driedonks *et. al.*, 1981
Lambda	gpB/B*	62/56	Tsui and Hendrix, 1980
P2	gpQ/h6	39/36	Marvick *et. al.*, 1994
T3	gp8	69	Nakasu *et. al.*, 1985
T7	gp8	62	Dunn and Studier, 1983
Mu	gpH/H*	65/56	Grimaud, 1996

1985). Thus, while the overall icosahedral symmetry of the head can be maintained, the replacement of a true pentameric vertex leads to a local symmetry mismatch. Although the portal symmetry diverges from that of the capsid vertices, it does match the 6-fold symmetry of the tails (Casjens and Hendrix, 1988). Hendrix (1978) proposed that this mismatch was a useful component of phage architecture. The mismatch at the vertex decreases the stability of interactions between the portal structure and the adjacent capsid proteins. The portal is then free to rotate within the prohead. This rotational freedom is thought to be important for the packaging mechanism whereby a freely rotating portal, interacting with the helically arranged phosphate backbone of the phage DNA, may act to "screw" the DNA into the prohead.

The recent advances in electron microscopy and image processing have allowed investigators to examine the assembly of macromolecules. The portals of a number of phages have been examined and their rotational symmetry defined. Published data indicate that the structure of the portal is a homododecamer. Other evidence, however, had indicated that the portal of ϕ29, SPP1 and T7 could have a 12- or 13-fold rotational symmetry (Dube *et al.* 1993; Tsuprun *et al.* 1994; Cerritelli and Studier, 1996a). Higher resolution microscopic data continued to show that the ϕ29 portal was a dodecamer (Muller *et al.* 1997), and this controversy was resolved with recent models derived from crystallographic data that revealed a 12-fold symmetry of the ϕ29 portal (Guasch *et al.* 1998). A 13-fold symmetry, however, was reported for the portal of phage SPP1 (Dube *et al.* 1993). The Hendrix model of portal assisted DNA entry into the capsid was reinterpreted in light of this new data of the symmetry of the SPP1 portal. The rotation of the 13-fold symmetrical structure, rather than behaving as an active pump, could serve to

alleviate the stresses created by the interaction of the protein subunits and the DNA (Dube *et al.* 1993). The direct role of the portal in the entry of phage DNA into the head remains unresolved, as structural data indicate a distinct right-handed vorocity, or twist, in the ϕ29 connector (Carazo *et al.* 1985; Muller *et al.* 1997), and a toroidal shape of the P2/P4 portal (Rishovd *et al.* 1998). This screw-like feature of the portal indicates that it can impose a rotational orientation on the incoming DNA and therefore may play a direct role in packaging.

Terminase Interaction

While the portal may have an unconfirmed active role in packaging, it is known to be associated with the packaging process. Studies with phage portal mutants demonstrated that these mutants were unable to package DNA (Becker *et al.* 1977; Bowden and Modrich, 1985; Bazinet *et al.* 1988). The portal is the entry point of DNA into the head and is consequently central to the mechanism by which the DNA is packaged. Packaging by the dsDNA phages involves a heterodimeric enzyme complex (terminase or pacase) that is able to recognize, cleave, and insert the replicated concatomers of phage DNA into the pre-formed prohead (Black, 1988; Fujisawa and Morita, 1997). The oligomeric terminase/DNA complex and the portal form a high order structure that is capable of inserting the phage genome, in an ATP-dependent manner, into the phage head. The subunit interaction, between the enzyme complex and the portal ring, has been shown to occur through the C-terminal portion of the large subunit of the terminase (Frackman *et al.* 1984; Morita *et al.* 1995a). This interaction occurs primarily

after the initiation of packaging that is signalled by the recognition and cleavage of the phage DNA (Black, 1988). However, the packaging enzyme of phage T3 functions differently. The T3 terminase is made up of products encoded by genes 18 and 19 of the phage (Hamada *et al.* 1986). While the small subunit (gp18) binds to and recognizes the phage genome, the larger subunit is bound to the portal structure of the phage (Fujisawa *et al.* 1991).

Measurement of Packaged DNA

Packaging of phage genomes occurs either through the encapsidation of unit length genomes (e.g. Lambda: (Catalano *et al.* 1995)) or through a headful mechanism whereby greater than unit lengths of genomes are packaged (e.g. P1, P22 (Coren *et al.* 1995; Adams *et al.* 1983)) (Black, 1988). Isolation of portal-mutants has implicated the ring in determining the length of DNA packaged into the head (Casjens *et al.* 1992; Tavares *et al.* 1995). Casjens and co-workers (1992) showed that mutations in the P22 portal gene (gp1) were able to increase the size of DNA packaged into the pre-formed phage head. This effect was mapped to individual point mutations within the gp1 coding region. In addition, it was shown that progeny virions produced from mixed wild type/gp1 mutant infections packaged DNA of an intermediate length. This suggested that portals which were produced were the result of co-assembly of mutant and wild type subunits and that the information or mechanistic effect for genome sizing was derived from multiple subunits. Mutants of SPP1 that were defective in genome sizing were also mapped to the portal gene (gp6). In this case, the mutations resulted in the packaging of

DNA of smaller than unit genome size (Tavares *et al.* 1995). These naturally occurring mutations were all the result of changes from the wild type amino acid residue to a lysine (E_{251}-K; N_{365}-K; E_{424}-K). These effects therefore may give clues to the validity of the various hypotheses concerning the direct role of the portal in DNA packaging. Structural information that definitively demonstrates the position of these various mutations may indicate whether or not these amino acids are available for direct interaction with the incoming DNA or if they are sequestered away from the inner channel. If indeed these amino acids are unable to interact with the phage genome, the role of structural transformation during packaging may be further investigated as a mechanism for transducing signals from the filling capsid to the packaging enzymes.

Regulation of Packaging Initiation

An important secondary signal for packaging may also be present as an integral part of the portal structure. Rishovd *et al.* (1994) characterized the portal structure for the P2/P4 system. They noted that the portal structure, in the mature phage, was made up of a processed form of the Q gene product. While the full-length form of the protein is incorporated during prohead assembly, it is a processed form of the protein that is present after the phage genome has been packaged into the precursor structure. The processed form of the gpQ is designated h6 that is produced after the removal of the first 26 N-terminal amino acids from the protein (Rishovd *et al.* 1994). It was proposed, by the authors, that due to the basic nature of the N-terminal peptide, it might serve to bind DNA. The peptide would then be cleaved away to end the packaging process. Further

studies have indicated that there is a direct relationship between packaging and cleavage. Rishovd *et al.* (1998) showed that in the absence of a large terminase subunit (gpP), not only was DNA packaging inhibited, but also proheads with uncleaved gpQ accumulated. The portal subunit gpQ may therefore be required to initiate packaging by aiding the assembly of a portal/terminase/DNA structure by binding of DNA at the N-terminal region of gpQ. Completion of the assembly process would then proceed only when the basic peptide was removed from the portal monomers. A similar structural transition has been reported for the portal of λ in which the portal protein, gpB, is processed to form gpB*. Unlike the P2/P4 system, this cleavage reaction occurs prior to the packaging of the phage genome. Moreover, rather than processing all of the subunits, approximately three-quarters of the subunits are processed. It is therefore unlikely that the processing of the portal regulates the assembly of the packaging machinery of the λ prohead in a manner analogous to P2/P4.

Role in Prohead Assembly

As an integral, yet distinct, component of the phage head, the portal has often been implicated in initiation and regulation of head assembly. Its role as an initiator of head assembly has been studied in the T4 system. The T4 portal is encoded by gene 20 (Driedonks *et al.* 1981). Initiation of head assembly has been defined clearly in this system because of the location of the prohead assembly machinery of T4 on the inner membrane of *E. coli* (Michaud *et al.* 1989). The portal protein initiates prohead assembly upon oligomerization at the inner membrane. This process, rather than being completely

autonomous and spontaneous, is mediated by another phage encoded protein gp40. This accessory protein acts as a chaperone. Studies have shown that at a decreased temperature, when membrane fluidity is decreased, gp20 mediated prohead assembly can occur in the absence of gp40 (Hsiao and Black, 1978). However, under normal growth temperatures, gp40 is thought to independently insert into the plasma membrane in order to facilitate the localization and oligomerization of the T4 portal proteins that are unable to carry out these processes in the absence of gp40 (Michaud *et al.* 1989). Subsequent mutant analysis of the gp20-gp40 pathway hinted at a traditional chaperone function for gp40. Mutants in gp40 are temperature sensitive and are phenotypically similar to gp20am mutants. Second site suppressor mutants in gp20 termed 40bypass had been isolated to further characterize the portal, polymerisation/prehead initiation pathway (Yap and Rao, 1996). Interestingly the 40bypass mutants are upstream, non-coding changes of gp20. The mutations affect translational efficiency, leading to marked increases in the production of the portal protein. Yap and Rao (1996) speculated that due to the circumstantial nature of the evidence that led to the membrane-based theory of gp40-gp20 interaction the conclusions might be questionable. Instead, they suggest that, due to the nature of the bypass that they observe, the effect of gp40 may be related more to correct gp20 folding. The suppression of the gp40 mutation is due to an increase in the effective pools of gp20 rather than compensation for the loss of gp40. Recently, it was shown that T4 gp31 functionally replaced the groES portion of the *E. coli* encoded chaperonin complex during productive assembly of capsid (gp23) subunits (van der Vies *et al.* 1994). It has been reported that similar to gp31, gp40 acts to ensure proper folding of the portal monomers (Kato and Baschong, 1997). Complexes formed by gp20 during infections in

which membrane association cannot occur, reveal a complex that contains the portal, whiskers and neck of the phage. This complex was similar to that seen in the fully formed phage, and was shown to be efficient in the reconstitution of proheads and naked cores *in vitro* (Kato and Baschong, 1997). The absence of gp40 in these complexes, as well as the similarity in size between gp40 and gp31, lead Kato and Baschong to conclude that gp40 might substitute for groES in a manner analogous to gp31. Therefore, the decreased requirement for gp40 at lower temperatures, which had been speculated to correlate with decreased membrane fluidity and therefore increased oligomer stability (Hsiao and Black, 1978) might in fact represent a more favourable temperature for proper folding of the monomers. This would lessen the requirement for chaperone mediated protein folding.

In contrast to T4, all other well-studied phage systems assemble their portals, and therefore proheads, cytoplasmically. The portal proteins in other systems, such as λ and ϕ29, require a traditional chaperone, GroESL, in order to assemble properly (Tsui and Hendrix, 1980; Tsuprun *et al.* 1995). In addition to this conventional chaperone, it was speculated that P22 utilized nucleic acid in order to regulate polymerisation of its portal subunits. The portal protein (gp1) of P22 was reported to bind to the upstream portion of late messenger RNA [referenced in (Prevelige, Jr. and King, 1993a)]. This is believed to regulate the translation of gp8 and also serves as a site that facilitates monomer oligomerization. In addition to this unique method of facilitating portal oligomerization, the scaffolding and major head proteins of P22 are able to interact, and form a prohead of correct dimensions in the absence of a functional portal structure (Bazinet and King, 1988). This structure, however, cannot package DNA and therefore is unable form functional phages because of the absence of an entry point for the replicated phage

genome. The existence of this structure also brings into question the role of the portal as an initiator of prohead assembly. In order for the portal to be incorporated into the prohead at a single position, inclusion of the structure as part of an initiator complex can constrain the position of the connector early during polymerization. This would also allow the ring to regulate the rate of prohead synthesis in relation to the rate of portal synthesis, and orient the major head proteins into specific conformations that are required to begin their positioning in the various quasi-equivalent orientations. However, the spontaneous formation of P22 proheads of appropriate dimensions, in the absence of a portal structure, indicates that prohead formation can occur independently of this type of initiator event. While these closed, portal-deficient structures may form, the natural order of oligomerization may involve the portal participating early in the process of prohead formation.

Studies with λ have shown that the absence of a connector results in the formation of structures similar to so-called *petit* lambda structures (Hendrix and Casjens, 1975). Similar to P22 structures that are assembled in the absence of the portal, the aberrant lambda proheads are incapable of packaging DNA. Murialdo and Becker (1977; 1978) have demonstrated that a functional portal is required early in the lambda prohead assembly process to initiate proper oligomerization of the lambda prohead. Experiments with extracts from *E. coli,* infected with various mutants of λ, demonstrated that only a combination of a major capsid mutant extract (E^-) and a scaffold mutant extract ($Nu3^-$) could effectively donate active precursors for the correct assembly of a λ prohead (Murialdo and Becker, 1977). In these experiments, both extracts contained active portals that were available early in the assembly process. Additionally, it was shown that the λ

portal was unable to functionally associate with a preassembled shell that contained the major head protein, but could associate with slowly sedimenting (unassembled monomers or small oligomers) forms of gpE. This indicates that the portal must associate with uncommitted molecules of gpE to be incorporated into the prohead (Bazinet and King, 1985).

The ϕ29 prohead has an elongated shape because the mature phage head is not isometric, but is a prolate structure. Studies of prohead formation of ϕ29, in the absence of its portal structure, revealed that this phage was capable of forming proheads composed solely of scaffolds and capsid protein. The structures that formed were a mixture of isometric and prolate shells (Guo *et al.* 1991). Therefore, while structures similar to those reported for P22 (Bazinet and King, 1988) could be formed, the improper shape of the ϕ29 structures indicated, that the incorporation of the portal, early in the formation of the prohead, was an essential part of this assembly pathway.

In the P2/P4 system, it had been reported that the portal was incorporated exclusively into the prohead early in the polymerization process (Rishovd *et al.* 1994). In the absence of a functional portal, the capsid and scaffold proteins of P2 are capable of forming an isometric shell of proper dimension (Marvik *et al.* 1994b). The closed structures are not limited to shells resembling P2. Some structures also resemble the smaller P4 prohead. However, the formation of these smaller structures requires the presence of the phage-encoded protein Sid (Marvik *et al.* 1994b).

In contrast to the proposed assembly pathways for the proheads of T4, λ, ϕ29 and P2/P4, Cerritelli and Studier (1996b) proposed a different role for the connector in the T7 system. Expression of cloned genes that encoded the various prohead proteins revealed

that structures similar in dimension and shape to a normal prohead, again, could be formed from scaffold and capsid protein alone. Thus, T7 follows a pathway that seems to be common to the dsDNA phages. In addition to the presence of these aberrant proheads, two other structures were identified. Scaffold and capsid proteins interacted to form large tubular polyheads and structures which Cerritelli and Studier (1996b) named "pac-men". They believed that the pac-men and polyheads resulted from the inability of the prohead assembly pathway to complete the assembly of the precursor head structure through the addition of the portal ring. This implies that the ring is a terminator rather than an initiator structure. The pac-men were found to have the mass of a structure that lacked a pentamer and the five surrounding hexamers. Pac-men, therefore, are believed to be produced when the scaffold and capsid proteins are nucleated in the absence of a portal ring, producing this aberrant structure. The polyheads, however, are thought to result from uncontrolled polymerization nucleated by the pac-men. In the absence of the portal, pac-men form and remain uncompleted when the scaffold/capsid protein ratio differs from that normally found in the wildtype prohead. The unbound capsid proteins on the growing edge of the structure can then begin to bind monomeric capsid protein to initiate a cascade of monomer binding that leads to the formation of a tubular polyhead. Cerritelli and Studier (1996b) conclude that the presence of these structures is indicative of the role of the portal as a terminator in the formation of the T7 prohead.

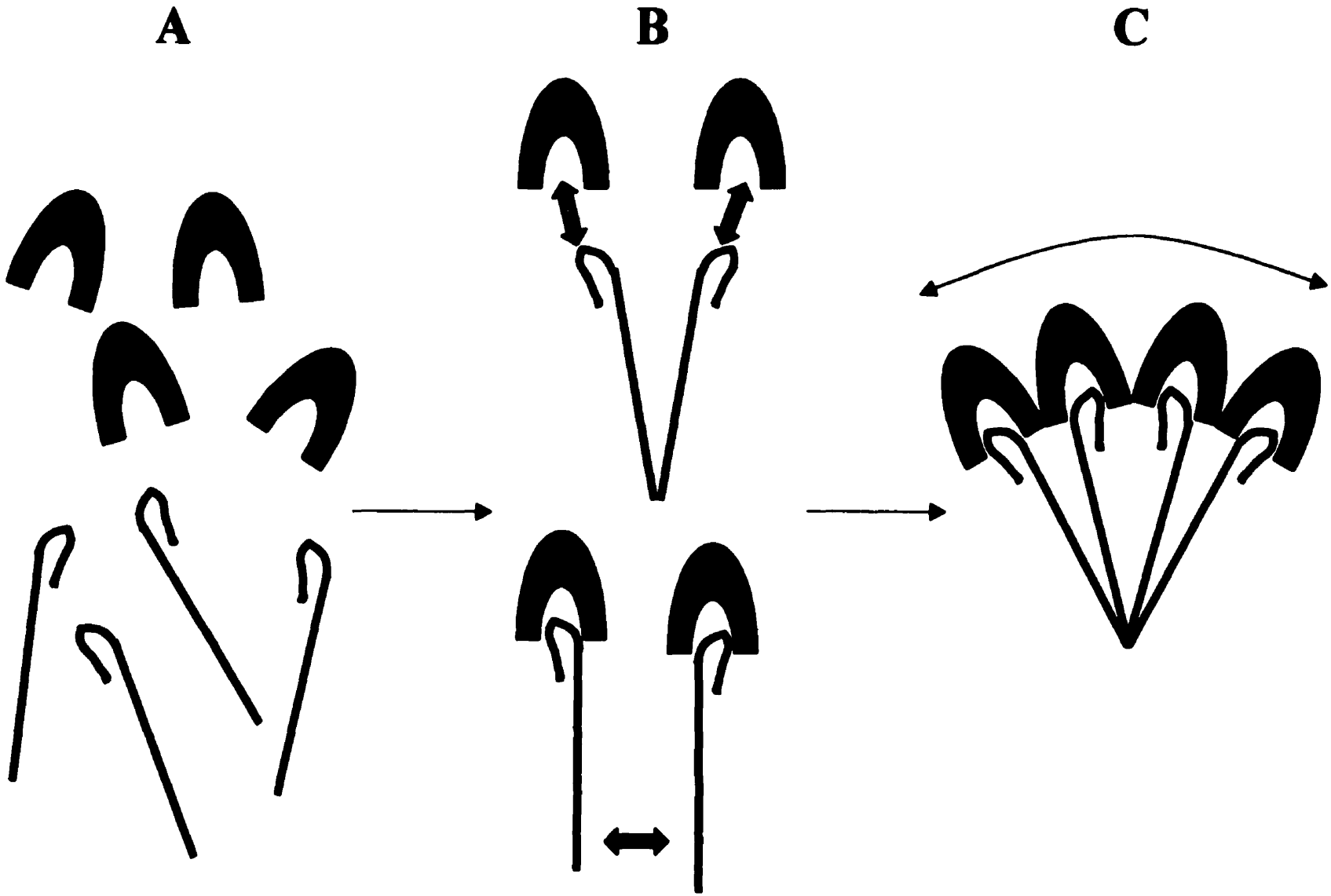
The study of portal mutants has identified a number of structures that helps to recognize the role of this oligomer in the assembly of a prohead. While the role of the portal as an initiator of the assembly process remains unresolved, examination of the structures produced in its absence hints that it plays a central role in this process. With the

exception of the studies of Cerritelli and Studier (1996b), the consistent discovery of closed structures, in the absence of the portal, may be a consequence of the strict geometrical form of the prohead as well as the limited number of orientations that may be held by the capsid protein. ϕ29 portal mutants illustrate this by the presence of isometric particles in the absence of the ring. The portal is absent and therefore is unavailable to orient the capsid monomers in a manner that constrain them in quasi-equivalent environments that favour the formation of a prolate structure. Furthermore, studies on various phage systems seem to indicate that the portal is involved early in the prohead assembly process.

The Scaffold

The prohead of dsDNA bacteriophages is a complex structure that can efficiently assemble under normal conditions but requires factors to promote and regulate this assembly process. In addition to the portal ring, which is present in the fully formed phage and which may be required to initiate the assembly process, the prohead contains a scaffold which seems to be absolutely required for the formation of this precursor assemblage (Casjens and Hendrix, 1988) (Fig. 4). In contrast to the portal, the scaffold is only transiently present within the prohead and must be removed prior to, or concomitant with, the insertion of the phage genome (Casjens and Hendrix, 1988).

Figure 4. Pathway of scaffold/capsid protein interaction. The scaffold protein has been shown, in most phage systems, to be essential for the efficient assembly of the prohead. The scaffold interacts with the major capsid protein as well as many of the minor constituents of the precursor structure. But, the scaffold is not a component of the mature head and is therefore removed from the prohead subsequent to its assembly. The scaffold appears to perform a number of functions during prohead assembly. (A) The scaffold and major capsid proteins are produced as soluble monomers. (B) During prohead assembly, the two proteins begin a co-ordinated oligomerization process. The scaffold proteins of various phages have been shown to exist as monomers, dimers, trimers and /or tetramers. The major capsid proteins exist as monomer, dimers, pentamers and/or hexamers. These assembly intermediates interact to initiate the assembly of the shell. The scaffold is thought to enable the interaction of capsid subunits in order to overcome energetic barriers or to ensure positioning of the subunits within proper quasi-equivalent environments. (C) The specific interaction of the scaffold and the capsid subunits may serve to define an angle of curvature for the prohead. Fixing the curvature through this interaction defines the size of the spherical prohead.


A
B
C

Prohead Assembly-Initiation and Fidelity

Similar to portal mutants, scaffold mutants are able to form closed prohead-like structures (Earnshaw and King, 1978; Ray and Murialdo, 1975; Rajagopal *et al.* 1993). In addition to the closed structures, the absence of scaffolding protein leads to the formation of various aberrant oligomers of the major capsid protein. The nature of these aberrant structures is different in the various systems. For example, mutations in P22 gp8 leads to the formation of long spirals of polymerized capsid protein (Earnshaw and King, 1978). In the case of λ, the aberrant structures are so-called "monsters", structures with no apparent geometrical symmetry (Mendelson *et al.* 1991). The presence of both prohead-like and non-prohead structures is indicative of the strict geometrical constraints placed on the polymerization process. In addition, the appearance of the prohead-like structures is delayed during P22 gp8$^-$ infection compared to the appearance of these structures during wild type infection. This may be understood in terms of the equilibrium model of self-assembly proposed by Zlotnick (Zlotnick, 1994). The author hypothesizes that critical concentrations of the various substrates of assembly are required prior to productive assembly of a viral precursor. If the system is perturbed through the removal of one of the interacting components, the equilibrium of the system will no longer reflect that of the wild-type pathway. Moreover, not only is the timing of the system regulated by the catalytic or helper function of some of the components, but the fidelity of the system is controlled in this manner as aberrant structures are produced if the equilibrium is perturbed. The role of the scaffold is therefore central to the timing and fidelity of prohead morphogenesis.

In contrast to the above systems, in which the examined products were derived from infections with a phage mutant in the scaffolding gene, it was observed that the expression of the major capsid protein of the phage, on its own, led to the formation of inclusion bodies or aberrant structures (Cerritelli and Studier, 1996b; Lee and Guo, 1995). This may be indicative of the presence of phage-encoded factors, other than the scaffold or the portal, that are required for assembly of a prohead. However, prohead-like shells can be produced when the major capsid and scaffold proteins interact in a cloned system (Lee and Guo, 1995). The requirement for co-expression of these genes is again indicative of the equilibrium that exists between the phage-encoded components that are needed for the assembly of this precursor structure. Despite the formation of small amounts of P22 and λ prohead-like structures, in the absence of scaffold, it nevertheless appears that the scaffold is essential for incorporation of normal levels of major capsid protein into prohead-like structures. The scaffold participates in this assembly reaction, orienting and facilitating the interaction of the capsid proteins.

Evidence for an intricate equilibrium state for the formation of the phage prohead can be garnered from the systems described above. A more traditional, step-wise assembly pathway has been detected for the formation of the T4 prohead through the formation of scaffolding core precursors. In addition to initiating head formation at the bacterial cell membrane, T4 is able to form a complex discrete core composed of a number of proteins; the majority of which is the gp22 scaffold (Kuhn *et al.* 1987; Traub and Maeder, 1984). The naked cores observed were the result of infection by mutant phages. However, cores with varying amounts of major capsid protein have been observed during infections with wildtype phages. This indicates that these structures are

true assembly intermediates (Kuhn *et al.* 1987). The presence of this intermediate, however, may simply be an extension of Zlotnick's equilibrium theory (Zlotnick, 1994) whereby the core forms subsequent to the constituents reaching critical concentrations. However, this assembly process is also regulated via its dependence on portal assembly at the membrane (Michaud *et al.* 1989).

Scaffold Structure and Protein-Protein Interactions

Unlike the assembly of the core at the membrane, phages other than T4 do not assemble scaffolding structures in the absence of capsid protein. Prevelige *et. al.* (1988) found that purified P22 scaffold existed as a monomer or dimer in solution but was unable to form a core-like structure. However when the scaffold monomers and dimers were mixed with purified coat protein, they interacted to form prohead-like structures. These structures contained the two proteins at ratios identical to wild type (Prevelige, Jr. *et al.* 1988). Detailed study of the P22 scaffold (gp8) by sedimentation analysis and Raman spectroscopy has indicated that it is a prolate ellipsoid, composed mainly of alpha helical domains (Fuller and King, 1981; Tuma *et al.* 1996). During DNA packaging and prohead expansion, the scaffold must be removed to make room for the DNA and facilitate prohead expansion. Rather than degrading the scaffold within the prohead, the P22 scaffold subunits are released intact and are utilized in further rounds of prohead assembly (Casjens and Hendrix, 1988). Reconstructed images, of immature proheads and mature heads of P22, have revealed the presence of holes at the six-fold vertices. The holes, which are present in the prohead, are absent in the mature head and are therefore

assumed to be the exit point for the scaffold protein (Prasad *et al.* 1993). The flexibility imparted to the protein as a result of the mainly α-helical structure of the scaffold monomer imparts plasticity to this protein (see above) and this facilitates the exit of the protein from the maturing prohead (Tuma *et al.* 1996).

The defined capsid/scaffolding protein assays, available with the P22 system, and in depth structural analysis of the two proteins, have led to a detailed understanding of scaffold function in this system. Analysis of specific mutations first indicated functional regions within the protein (Greene and King, 1996). Four specific amino acid substitutions have revealed diverse effects on the process of prohead assembly and maturation. C-terminal changes in both S242F and Y214W resulted in the loss of incorporation of minor capsid proteins (Greene and King, 1996). The S242F mutation, however, was unable to incorporate the portal and other minor proteins, while the Y214W mutation was deficient only in binding of the portal (Greene and King, 1996). The mutations L177I and Q149W had a more pronounced effect on prohead structure. These mutations had opposing effects on scaffold release, with Q149W being released more readily. Conversely, the L177I mutant was unable to release the scaffold and therefore matured very poorly (Greene and King, 1996). Each of these mutations reflects the multifunctional nature of the scaffold. Not only is the scaffold required for initiation of assembly through the action of individual prohead components, it is also required for the proper maturation of the prohead. This process may not be intuitively linked to this catalytic protein since it need not be present in the mature structure. However the existence of scaffold mutants that perturb this process is indicative of such a function. The mutations themselves are most likely to disrupt normal scaffold function rather than

directly affect the capsid shell. It is known that the P22 scaffold is recycled during prohead assembly (Bazinet and King, 1985) and distinct holes exist within the immature prohead (Thuman-Commike *et al.* 1996); these holes that may serve as exit points for the protein. The maturation mutants, therefore, may disrupt the protein, in such a way, that it is unable to revert to a monomeric state capable of exiting through the capsid holes. The mutations may instead affect conformational signals, which are intended to initiate scaffold release when packaging commences. (Greene and King, 1996).

A signal may be essential for initiation of scaffold release due to the mechanism by which the scaffold forms the core during prohead assembly. Studies by Parker *et. al.* (1997) on the oligomerization state of the scaffold protein in solution have shown that the P22 scaffold is present as monomers, dimers, and tetramers (most likely dimers of dimers). Previously, another phage scaffold protein, Nu3 of bacteriophage lambda, was also shown to form oligomers (dimer and trimers) in solution (Ziegelhoffer *et al.* 1992). It is also known that the association between the scaffolding monomers is fairly weak (Parker *et al.* 1997). This weak association may be essential for the subsequent release of the protein. This is especially important since it has been suggested that it is dimers of the scaffold, interacting with coat pentamers that are the active species in prohead assembly (Prevelige, Jr. *et al.* 1993). The importance of the scaffold dimer was further noted (Parker *et al.* 1997) when a specific mutant form of the protein was used. An R74C/L177I double mutant is able to form stable dimers through the formation of a disulphide bridge. These disulphide-linked dimers catalyze prohead formation more efficiently than the wildtype. In a reducing environment, however, the double mutant is less efficient (Parker *et al.* 1997). A mutant form of the protein, with a deletion of the first 140 amino acids

(mutant 141-303) can dimerize but unable to form tetramers (Tuma *et al.* 1998), yet is still able to form procapsid like structures. However, removal of the C-terminal 11 amino acids (mutant 141-292) from the protein completely abrogated the ability of the scaffold fragment to function in prohead assembly (Parker *et al.* 1998). Examination of the two deletion mutants revealed that while both were able to form dimers, the 141-303 dimers were slightly more stable than the 141-292 fragment (Tuma *et al.* 1998). This stability may be attributed to the presence of the 11 C-terminal amino acids in the assembly competent fragment. While these residues themselves were not essential for overall dimerization, removal of these terminal residues led to a loss of α-helicity in approximately 20 amino acids and a concomitant increase in conformational disorder for approximately 10 more amino acids (Tuma *et al.* 1998). The stability of the wildtype scaffold had been determined previously (Parker *et al.* 1997) and found to be greater than that of either of the mutant scaffolds. Denaturation and deuterium exchange studies have revealed that the wildtype monomer could interact through their N and C-terminal helices to produced stabilized intramolecular forms (Tuma *et al.* 1998). While the wild type form of the protein generates more stable interactions, the 141-303 fragment is able to produce prohead structures albeit with decreased fidelity. The assembly competent fragment and the capsid protein form proheads at a ratio similar to that of the wildtype scaffold/capsid protein structures (Parker *et al.* 1998). Since the assembly competent fragment is approximately half the size of the wildtype, twice the number of the mutant fragments is present in these structures (Parker *et al.* 1998). It had been hypothesized previously that one of the roles of the scaffold was to exclude host cytoplasmic material from the growing shell (Earnshaw and Casjens, 1980). Recently, this function was demonstrated

when the contents of proheads, formed in the absence of a scaffold, were examined (Thuman-Commike *et al.* 1998). The 141-303 fragment can pack within the prohead, in numbers, that maintain the scaffold-capsid ratio of the wild type. This indicates that the incorporation of this protein is also dependent on the interior volume of the structure and not simply the availability of binding sites. However, the decreased efficiency of prohead formation, at increased concentrations, may reflect the multifunctional nature of the N-terminal fragment of the scaffold. This portion of the protein may serve to stabilize dimer interactions, increase the fidelity of structure formation, and serve to occupy a greater volume in order to exclude extraneous cytoplasmic contents.

The importance of the dimer or dimer-dimer protein oligomers suggested that the scaffold might be acting as an "entropy sink" in which the binding energy of scaffold dimerization facilitated coat protein interaction within the polymerizing shell (Parker *et al.* 1997; Parker *et al.* 1998). The 11 C-terminal amino acids of the scaffold can form a highly charged helix that may interact with the phage coat proteins. This may polymerize the capsid proteins into the growing shell by dimerization of the scaffold subunits. The binding of the scaffold to the capsid through such electrostatic interactions was demonstrated by Parker and Prevelige (1998). The importance of electrostatic interactions was determined by carrying out prohead assembly experiments in high ionic strength buffers. A high concentration of salt completely inhibits proper assembly of the prohead from a mixture of scaffold and capsid proteins. Additionally, scaffold re-entry into proheads, from which the scaffold had been specifically removed, was inhibited in the high ionic environment. Finally, induced expansion of proheads was also inhibited in the presence of high ionic strength buffers. These data indicated that an electrostatic force

played a role in the assembly and maturation of the prohead. Parker and Prevelige (1998) hypothesize that the positively charged C-terminal helix-loop-helix of the scaffold interacts, electrostatically, with the capsid protein that carries an overall negative charge. The coupling of charge neutralization and dimerization by the scaffold drives the association of the capsid proteins. In addition, the scaffold may orient the capsid monomers in order that they associate in the proper quasi-equivalent environments.

Capsid Sizing

The orientation of capsid monomers by the scaffold emphasizes the potential of this protein to determine the size of the prohead. While the size determining function of the portal has been demonstrated for the assembly of a prolate phage, the portal does not seem to have a similar role in the proper sizing of isometric capsids. Closed, capsid-like structures of P22 are produced in the absence of the portal (Bazinet and King, 1988). However, in the absence of the scaffold protein, capsid-like structures of two sizes, T=7 and T=4 are formed (Thuman-Commike *et al.* 1998). Images reconstructed from cryo-electron micrographs of large and small capsids were used to examine the effect of the absence of the scaffold on prohead assembly. While the overall structure of the hexon and penton capsid subunits is the same in the two structures, there are distinct differences in binding interactions and relative subunit positioning (Thuman-Commike *et al.* 1998). It was hypothesized that the scaffold acted as an assembly chaperone that limited the binding interactions of the capsid monomers in order to exclude interactions that favoured the formation of the smaller capsids. In addition, the scaffold might establish a specific

binding angle for the capsid subunits that would, in turn, create a specific curvature for the growing icosahedral shell (Thuman-Commike *et al.* 1998). Once such a specific curvature was fixed, the size of the structure would be determined. The structures that are formed in the absence of the scaffold also lack a portal, since it is through an interaction with the scaffold that the portal is incorporated into the prohead (Greene and King, 1996). Therefore, a possible role for the portal in size determination cannot be discounted.

An Alternative Scaffold-Sid

A direct sizing mechanism, mediated by a scaffold structure, has been demonstrated in the P2/P4 system (Marvik *et al.* 1994b). Unlike other phages, P2 and P4 are related; P4 depends on P2 for the production of many of its morphogenetic components (Bertani and Six, 1988). While P2 forms a T=7 capsid to house its genome, the P4 genome is smaller and therefore requires a capsid, T=4, of smaller dimensions (Dokland *et al.* 1992). Rather than relying on the chance production of a T=4 capsid that is shown in the P22 system to be a rare event, the parasitic P4 phage encodes a specific protein that directs the assembly of a prohead of correct size (Marvik *et al.* 1994b).

The assembly of the P2 prohead is accomplished in a manner similar to that of most other bacteriophages. The precursor structure is formed from a small number of morphogenetic proteins including capsid (gpN) (Bertani and Six, 1988), portal (gpQ), and scaffold (gpO) proteins (Rishovd *et al.* 1994). However, unlike P22, the scaffold is apparently degraded, and one of the products of the processing, a peptide designated h7, remains associated with the mature virion (Rishovd *et al.* 1994). A similar mechanism is

employed by the parasitic P4 phage. P4 encodes a protein, Sid (Bertani and Six, 1988) that is able to utilize the P2 derived morphogenetic proteins and, along with the normal scaffold, re-orient the polymerization of the major capsid protein to produce a T=4 prohead (Marvik *et al.* 1994b). In addition to Sid, differences in capsid composition had been noted that were related to the processing of the major capsid protein. These differences were shown to be a direct result of the sizing process (Rishovd and Lindqvist, 1992; Marvik *et al.* 1994a). Prior to this direct evidence for the role of the Sid protein as a sizing regulator/scaffold, other studies had implied this protein played an important role in the formation of the P4 capsid. Mutations in the *sid* gene resulted in a loss of specific P4 capsid formation. Additionally, P2 mutants, designated *sir,* had been isolated in the *N* gene that made the gene product refractory to Sid (Bertani and Six, 1988). These data are indicative of a direct role for Sid in the formation of the smaller P4 capsids, but the exact function of the protein remains unresolved. Evidence has been collected recently that has defined the role of Sid as an alternate scaffold that is utilized specifically to assemble P4 proheads. It was shown that a P4 encoded protein directly controlled the assembly of T=4 sized capsids when P4 was used to infect an *E. coli* strain that expressed the P2 capsid and scaffold proteins (Marvik *et al.* 1994b). Expression of the P2 scaffold and major capsid proteins, from cloned genes, produced prohead-like structures of a size similar to that of wild-type structures. Infection of strains expressing the P2 scaffold and capsid protein by P4, however, resulted in a dramatic decrease in the production of P2 prohead-like structures and a concomitant increase in P4 prohead-like structures.

Sid Scaffold Structures

Image reconstructions of structures produced from strains infected with P4 and co-expressing gpN revealed icosahedrally arranged particles within a dodecahedral lattice (Marvik *et al.* 1995). The protein composition of these structures included Sid and the major capsid protein. The Sid protein-lattice was arranged to form an external cage around the prohead-like structure along the geometrical axes of the icosahedron. Sid monomers seem to interact with each other at the three-fold axes, while the scaffold–capsomer interactions occur along the two-fold domain of the hexamers. The different interactions seem to be mediated by different ends of the molecule, which appears as a multi-domain rod. Unlike other scaffolds that need not take on any specific icosahedral symmetry, the reconstructions of the Sid containing structures indicate a specific geometrical form. Consequently, the Sid lattice must be composed of a defined multiple of individual subunits in order to maintain the symmetry. The authors favoured an arrangement, based on densitometric and mass difference measurements, of a lattice of 120 monomers. This number is geometrically favourable since icosahedrally arranged structures are formed from 60 subunits or multiples thereof. The authors also discussed a possible mechanism by which the Sid protein might act to define a T=4 prohead. They speculated that the Sid protein necessarily interacted with pre-formed hexamers, since the Sid-gpN contacts were limited in the final structure; this therefore seemed to preclude interactions between the scaffold and individual monomers. The interaction of Sid with the capsid hexamer may influence the topology of the hexamer and the interactions of entire complex may be limited due to the orientation of Sid-Sid contacts at the three-fold

axes. Consequently, the formation of the T=4 structure may be regulated by two different protein-protein interactions that provide some redundancy to the process.

The above data strongly suggest a direct role of the Sid protein in mediating P4 T=4 prohead formation. However, the structures produced in these experiments were made in the presence of other P4 derived proteins. Direct analysis of the effect of the Sid protein was studied with a cloned version of the gene. Expression of the Sid protein, from the cloned gene, could not only complement a P4 *sid* mutant, but could directly mediate the formation of smaller proheads in the absence of other P4 proteins (Nilssen *et al.* 1996a). Interestingly, the burst size of the P2 helper was only reduced 3-4 fold as compared to a 90-fold reduction in the presence of a P2::P4 hybrid. Therefore, while Sid may act on its own to generate T=4 proheads, the reaction is potentiated by another P4 encoded protein. Based on studies to define the nature of this helper protein, the authors reported that it was not encoded by any of the previously characterized P4 genes. The authors, however, proposed two hypotheses for the function of this as yet uncharacterized protein. It might be required to ensure proper stoichiometric levels of the Sid protein through stabilisation of the *sid* transcript, stimulation of *sid* translation or stabilisation of the Sid protein. Alternatively, the auxiliary protein could serve to stabilize the Sid-Sid or Sid-gpN interactions. While such a function may be necessary to ensure efficient interference with P2 helper formation, the structural and molecular data demonstrate the function of Sid as a unique scaffold. Analysis of mutations in the protein has revealed that, as suggested by the reconstruction data (Marvik *et al.* 1995), Sid can be divided into three segments demarcated by the N and C termini and the middle of the protein. Furthermore, computer modelling based on the sequence of Sid, has suggested that the

protein is highly α-helical in nature and contains large hydrophilic areas (Nilssen *et al.* 1996b). The elongated rod structure proposed for the Sid protein (Marvik *et al.* 1995) is very similar to that of P22 gp8 (Tuma *et al.* 1996). This structural similarity may be indicative of a general functional configuration of these structural proteins. While the flexibility of gp8 may be of paramount importance for the release and recycling of this scaffold (Bazinet and King, 1985), this property may also be necessary to allow these proteins to undergo subtle rearrangements as they position the capsid proteins during the formation of a prohead.

Modified Assembly Mechanisms

The paradigm developed to describe the assembly and maturation of tailed, dsDNA phages is, as with any model, sufficiently descriptive. As these processes are investigated in greater detail, refinements and differences are discovered that may adhere to the general mechanism while introducing significant changes to the process. Phages that demonstrate such differences include HK97 and ϕ29.

HK97

HK97 is a lambdoid phage. It is able to recombine, at low frequency, with phage λ, with which it shares a similar host range and immunity function, to produce hybrid progeny (Dhillon *et al.* 1980). In addition, it belongs to the same morphological class, possesses a flexible, non-contractile tail and an isometric head (Dhillon *et al.* 1980).

Although this phage shares extensive homology with the better-characterized phage λ, it utilizes a novel mechanism to enhance the stability of its head. HK97 is able to perform extensive cross-linking of its major head protein subunits. Preliminary experiments, by Popa *et al.* (1991), revealed the protein components of the HK97 head and the cross-linked nature of the major head protein. Separation of the phage proteins by SDS-PAGE showed that the head was made up of a 45 kDa portal protein along with the major head protein, which took the form of discretely migrating bands of increasing molecular weight. A large proportion of the head protein was unable to enter the gel. All attempts to dissociate this material were unsuccessful. But, partial proteolysis of isolated bands showed similar patterns that indicated that the bands were derived from a common precursor. Amber mutant analysis and partial proteolysis were used to demonstrate that a 42 kDa protein, seen in gels, was the monomeric form of the high molecular weight protein species. Edman degradation analysis of proteolytic fragments revealed that the crosslink occurred through the side chain of an internal lysine residue. Despite the identification of the cross-linked lysine residue, the second residue involved in the cross-link was not found.

Crosslinking of the Major Capsid Protein

Additional characterization of HK97 was obtained from experiments that utilized structures derived from cloned morphogenetic genes (Duda *et al.* 1995b; Duda *et al.* 1995a). Production of prohead components, from cloned genes, led to the identification of the major capsid protein, with a molecular mass of 31 kDa (Duda *et al.* 1995b).

Furthermore, it was shown that the 31 kDa protein, designated the major capsid protein, was derived from the 42 kDa protein mentioned previously. The recognition of this processed form of the major capsid protein facilitated the interpretation of data pertaining to peptide fragments of the cross-linked protein. Ultimately, the mature cross-linked form of the major capsid protein was used as a substrate for regulated proteolysis and mass spectroscopy. The spectroscopy data confirmed that K_{169} was one of the cross-linking partners. In addition, the second cross-linking partner was identified as N_{356} (Duda *et al.* 1995a). The authors suggested that the cross-linking partners came from two adjacent monomers that were present within individual hexamers and pentamers. The reaction, therefore, served to impart greater stability on the individual subassembly structures that were directly involved in the formation of the prohead (Xie and Hendrix, 1995). However, the discovery of intra-monomer subunit bonds did not account for the presence of bands that might represent oligomers of molecular weights greater than pentamers and/or hexamers, on SDS-PAGE gels. Consequently, two hypotheses were developed to explain the presence of these high molecular weight oligomers (Duda, 1998). The cross-linked pentamers and hexamers could possess other sites that might cross-link via secondary bonds that stabilized the maturing prohead. Alternatively, the individual monomers could be arranged to form catenated oligomers, likened to "chainmail" by the authors. This catenation might also enable the phage to further strengthen its head. Despite intensive investigation, secondary cross-links have not been discovered (Duda, 1998). While the lack of evidence for the presence of secondary cross-links tends to argue against the first of these hypotheses, it does not address the possibility of catenation of the pentamers and hexamers. To understand the mechanism by which the high molecular

weight oligomers form, Duda (Duda, 1998) performed partial proteolysis experiments. It was argued that a limited degree of proteolysis of the oligomers and separation the resulting fragments might reveal the nature of oligomer formation. If the assemblages were formed from secondary crosslinks, the resulting fragments would have a range of molecular masses. Conversely, if the oligomers were the result of catenation of hexamers and pentamers, the proteolytic fragments would be linearized forms of these two subassemblies in addition to their breakdown products. Upon resolution of partially digested cross-linked proheads, it was observed that the products of the reaction migrated as hexamers or pentamers. Further proteolysis resulted in the appearance of bands that migrated as circular hexamers and pentamers; these were available for resolution only after release from all potential catenation partners. In addition, tetramers, trimers, dimers, and monomers were present. These data clearly indicate that individual pentamers and hexamers are catenated during prohead maturation.

Regulation of Head Assembly

This catenation/crosslinking event is the final step during a head maturation pathway that is characterized by a number of distinct steps. The prohead formed from individual pentamers and hexamers was designated Prohead I. It is an immature round structure in which the individual subunits are all 42 kDa in size (Duda *et al.* 1995b). Subsequent to the formation of Prohead I, the phage-encoded protease, is incorporated into the prohead during assembly of the structure. The enzyme becomes active and cleaves the amino terminal 102 amino acids from the major capsid protein, which is

converted to the 31 kDa mature form. This structure is Prohead II. Upon initiation of DNA packaging, the HK97 Prohead II expands to produce Head I (Hendrix and Duda, 1998). Following expansion and subunit rearrangement, the cross-linking reaction is initiated and all subunits within the head are bonded to each other forming Head II (Duda *et al.* 1995a). These distinctly separable chain of events indicate another possible way by which the phage is able to control the substantial enzymatic events that make all transitions irreversible. The cross-linking reaction must occur after the subassemblies have oriented themselves in such a manner that once cross-linked, the entire structure will be catenated. Therefore, the dependence of cross-linking on structural rearrangements may be due to the mechanism of the reaction that is described below.

Crosslinking Mechanism

Cross-linking reactions, similar to those seen with HK97 major capsid protein have been described in other systems. In those cases, the linkage was made with a glutamine and catalyzed by a separate transglutaminase (Duda *et al.* 1995a). The HK97 capsid cross-link is, however, thought to be autocatalytic and is controlled by the positioning of the reactant lysine and asparagine residues subsequent to Head I formation (Duda *et al.* 1995a). Temporal regulation of head formation is mediated by distinct structural characteristics of a protein. This seems to be a common theme in the various dsDNA phages, as demonstrated by the mechanism described above as well as changes in proheads mediated by portals and scaffolds in various other phage systems. However, HK97 lacks a distinct scaffold protein.

Capsid Scaffold Function

As described above, the scaffold is the only other phage protein, along with the major capsid protein, that is essential for the efficient assembly of a prohead-like structure. Experiments using purified proteins, however, demonstrate that the major capsid protein alone can efficiently assemble the head in the HK97 system (Xie and Hendrix, 1995). The capsid protein, in its immature form, contains 102 amino acids at its amino terminus, referred to as the delta domain. This domain is proteolytically removed during the maturation of the prohead (Duda *et al.* 1995a). It has been hypothesized that this domain acts in a manner analogous to the autonomous scaffolds that are produced by other phages (Duda *et al.* 1995b; Xie and Hendrix, 1995; Hendrix and Duda, 1998). Cryo-reconstructions of prohead-like structures, with and without the delta domain, indicate that this region of the protein forms a complex overlapping network immediately underneath the shell (Conway *et al.* 1995). Furthermore, computer analysis of the sequence of the region also indicated it might play a role as a scaffold. It is highly probable that the delta domain can form α-helical coiled coils and can maintain an extended conformation reminiscent of the predicted scaffold structures. The HK97 capsid protein, therefore, may have evolved a "built-in" scaffold. However, it has been noted (Hendrix and Duda, 1998) that the scaffold genes of most phages are located immediately upstream of the capsid gene. Consequently it has been theorized that the HK97 capsid protein is the product of gene fusion between a scaffold gene and head gene. Thus while the assembly mechanism of the HK97 prohead I appears to deviate from the mechanism

of most dsDNA phages, the essential scaffolding function appears to be conserved in this system as well.

ϕ29 packaging RNA

Unlike the modifications to common prohead sub-components observed in the HK97 system, ϕ29 has incorporated a unique structural component essential for phage assembly. ϕ29 incorporates a unique packaging ribonucleic acid (pRNA) into its prohead in association with the portal complex (Guo *et al.* 1987a). However, the pRNA is not present in the mature phage (Wichitwechkarn *et al.* 1989). The pRNA was first identified when it was shown that defined, *in vitro* packaging reactions could be abolished by the presence of ribonucleases (Guo *et al.* 1987b). The pRNA, isolated from proheads and used in these *in vitro* experiments, was 120 nucleotides in length and was shown, by hybridization, to be transcribed from a region within the extreme left-end of the phage. Upon further investigation, it was determined that the isolation process resulted in the truncation of the pRNA, and that the molecule was in fact 174 nucleotides in length (Wichitwechkarn *et al.* 1989). Despite the difference in length between the two molecules, the 120 nucleotide pRNA is active, implying that this molecule contains all information necessary for activity. The number of pRNA molecules, associated with the ϕ29 connector, was determined by dot blot quantification of labelled pRNA from defined numbers of proheads. Based on these measurements, it has been determined that there are 5-6 copies of the molecule per prohead (Wichitwechkarn *et al.* 1989). Further investigation has determined that there are in fact 6 copies of pRNA on the prohead

(referenced in (Chen and Guo, 1997)) and that the molecules are linked to each other through Watson-Crick base pairing of complementary single stranded loops (referenced in (Hendrix, 1998)). The copy number of the molecule is geometrically agreeable because of the dodecahedral nature of the portal (Guasch *et al.* 1998). This suggests a role for the pRNA in the symmetry mismatch machine that is hypothesized to be active in DNA translocation during packaging (Hendrix, 1978). It was shown, *in vitro*, that the addition of mutant pRNA could severely impair packaging, while *in vivo* packaging was completely blocked (Trottier *et al.* 1996). In fact, a single mutant pRNA on a prohead can completely block packaging despite the presence of five other wild type molecules (Trottier *et al.* 1996). The use of partially active pRNAs allowed Chen and Guo (1997) to investigate the mechanism of action of these molecules but they did not address the function that the nucleic acid itself provided. Based on their results, they determined that the pRNAs acted in a sequential manner. All molecules contributed equally to the reaction and consequently a hypothesis was developed that was similar to the mismatch hypothesis. The pRNAs move through a defined step-wise rotation acquiring different energetic states that force interactions with the capsid walls or vertices. Hendrix discussed this hypothesis in a recent review (1998). He was the first to propose the symmetry mismatch mechanism in which the portal participates in phage packaging. In this recent article, he reviews his theory in light of new ideas regarding the shape of proteins that translocate along DNA, and the pRNA of ϕ29. He states that if the packaging complex, in association with the capsid, "walks down" the incoming DNA, the portal may rotate to relieve twists which develop. The pRNA could then perform two functions. It may be an integral part of the packaging machinery, or it may play a structural role providing a

bearing surface that modulates rotation. As yet, these theories remain unproven and the role of the pRNA undefined.

While much is understood about the general mechanism by which dsDNA phages build their morphogenetic precursors, models of the individual phage assembly pathways may reveal biophysical refinements; these may have evolved to complement the individual peculiarities of each virus. An icosahedral prohead provides a stable structure that houses the phage genome. The assembly of these structures, however, is varied in these different systems and therefore may provide a framework for studying complex structural assembly pathways.

Bacteriophage Mu

Overview

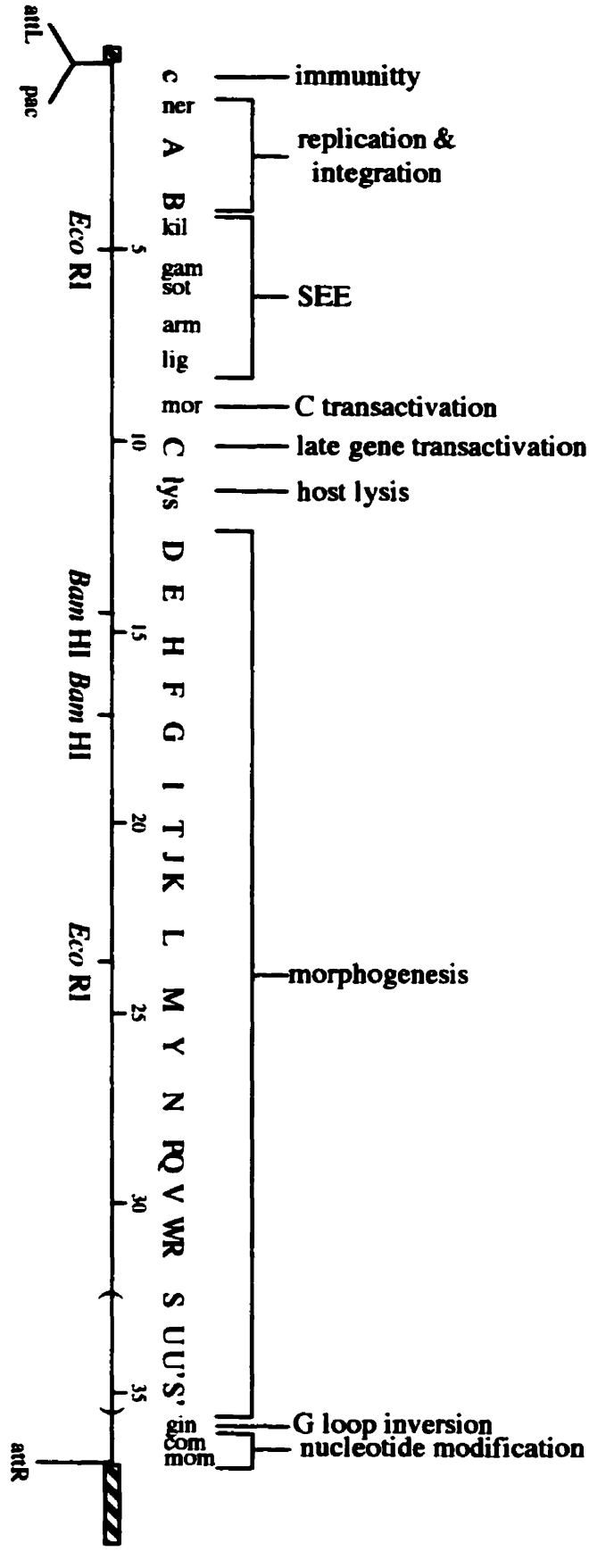
The discovery of Mu, the first of the transposable bacteriophages of *E. coli,* was the result of an investigation by Larry Taylor (1963), during screens of lysogenized bacteria which had become auxotrophs upon phage infection. Subsequent to the discovery of Mu, another transposable coliphage, D108 (Hull *et al.* 1978), was found. The two coliphages, Mu and D108, are homologous in approximately 90% of their DNA sequence, but diverge in their immunity region and at their right ends. A *Pseudomonas* transposable phage, D3112, has been shown to have homology to Mu and D108 with respect to function and position of genes, however it shows little sequence homology to

the coliphages. The Mu genome has recently been completely sequenced. A physical/genetic map of the genome can be seen in figure 5.

Lifecycle

The mutations observed in the lysogenized *E. coli* strains are the result of phage function. The integration of the phage genome is the first event in the phage life cycle following the injection of its DNA (Pato, 1989). The insertion of the phage genome by a random transpositional integration event (Allet, 1978) and subsequent lysogenization of the host can inactivate the chromosomal genes into which the phage genome has inserted (Martuscelli *et al.* 1971). At some time, prior to or after the initial integration event, a molecular decision is made to select one of the two life cycles the phage will undertake (Harshey, 1988). The switch is controlled by phage-encoded proteins that bind to *cis*-acting DNA sequences in the immunity region. This control site is located at approximately 1 kb from the left end of the phage. The immunity region contains two divergent promoters (P_e and P_c) overlapped by operator sites (Krause and Higgins, 1984), an IHF binding site (van Rijn *et al.* 1991), and an activator site for transposition (IAS) (Pato, 1989). The phage encoded repressor (c) and negative early regulator (Ner) proteins bind to the operator sites within the immunity region to control transcription from P_c and P_e. In contrast to all other Mu genes, the repressor, c, is transcribed from right to left (van Meeteren *et al.* 1980). The binding of the repressor mediates lysogeny by repressing transcription of the early genes from P_e whereas Ner relieves c repression and regulates early function transcription (Harshey, 1988). The early functions of the phage are

Figure 5. The Mu genome. The solid line represents the Mu genome and the hatched boxes represent attached host chromosomal DNA. The genome is divided into 5kb intervals, marked above the line. SEE is the Semi Essential Early region.


attL
pac
c
ner
A
B
kil
gam
sot
arm
lig
mor
C
lys
D
E
H
F
G
I
T
J
K
L
M
Y
N
RQ
V
WR
S
UU'S'
gin
com
mom
attR
Eco RI
Bam HI
Bam HI
Eco RI
5
10
15
20
25
30
35
immunitty
replication &
integration
SEE
C transactivation
late gene transactivation
host lysis
morphogenesis
G loop inversion
nucleotide modification

transcribed as a single unit; *ner* being the first gene transcribed. The *ner* gene is followed by the phage transposase (*A*) and transposase-associated (*B*) genes. Immediately downstream of the replicative functions are the semi-essential early genes *kil*, *gam*, *sot*, and *arm* (Paolozzi and Symonds, 1987). These proteins play a variety of roles in lytic development, including phage genome protection and host killing.

The transcription of the Mu genome is also temporally regulated. The middle promoter and its activator, which transcribe the C gene product of the phage, are located at the end of this semi-essential region (Mathee and Howe, 1990). The C protein is the late gene transactivator of the phage and is absolutely essential for the transcription of the four late promoters of the phage (P_{lys}, P_I, P_P, and P_{mom}) (Harshey, 1988). The promoters P_{lys}, P_I, and P_P, transcribe the morphogenetic machinery of the phage and are required to complete the lytic life cycle (Margolin *et al.* 1989). P_{mom} transcribes a gene whose product is responsible for modifying some adenine bases within the phage sequence (Harshey, 1988).

Mu Transposition

The transposition of the phage genome into the host chromosome not only precedes stable lysogenization of the host, but is also the replicative mechanism of the phage during the lytic cycle (Ljungquist and Bukhari, 1977). These phages may also be considered to be transposons, since the genomes are flanked, at all times, by heterologous host DNA. Studies have shown that the ends of Mu are flanked by varying lengths of host DNA (Bukhari and Taylor, 1975; Daniell *et al.* 1975). It was later shown that the left end

host sequences attached to the Mu genome were of a defined range of lengths characterized by differences of 11 bp (George and Bukhari, 1981). Recognition of the transposable nature of the Mu led investigators to define the mechanisms of phage integration and replication.

Transfections of *E. coli* with phenol extracted Mu DNA, in strains other than *rec*BC mutants, proved to be inefficient. This indicated that the free termini of the phage genome were sensitive to exonucleolytic degradation. However, DNA isolated by freeze thaw treatment was efficient in establishing phage infection via transfection (Chase and Benzinger, 1982). Inspection of the isolated DNA revealed a 65kDa protein that was associated with the ends of the phage genome. This protein mediated the formation of a non-covalently closed circular DNA structure. The 65kDa protein was shown, subsequently, to have a mass of 64kDa and was derived from the phage encoded N gene (Gloor and Chaconas, 1986). The N gene is part of the tail gene cluster in the morphogenetic region of the Mu genome (Harshey, 1988). The protein is injected into the host along with the phage DNA. The actual function of the N protein has yet to be proven definitively. The protein may serve in the formation a nucleoprotein complex in which the ends of the phage are oriented in such a manner that they may integrate efficiently into the host chromosome; or, the protein may be utilized by the phage to protect the free termini of the genome from exonucleolytic attack. Finally, the protein may suppress replication of the phage genome so that the primary integration event of the phage DNA occurs via conservative transposition. Indeed, all these functions may be required and are not mutually exclusive.

As previously mentioned, one of the potential functions of the N protein during Mu infection is the suppression of replication of the injected DNA. The phage genome integrates into the host chromosome during a lytic or lysogenic infection and consequently a specific mechanism for this integration must exist. Transposition may proceed either by a replicative mechanism, utilizing the host replication functions to copy the phage genome into a new site within the host chromosome, or it may proceed via a conservative "cut and paste" mechanism (Craig, 1996). During this latter process, the host replicative functions repair gaps that are created during the insertion process. As part of both processes, there is a duplication of the host sequence at the target integration site. This is a hallmark of transposition in all systems except, IS91 and Tn911 (Craig, 1996). Genetically-based investigations, undertaken to determine which mechanism was used by Mu during this primary integration event, were not definitive. However, experiments that examined the methylation state of phage DNA from a *dam*$^+$ source, after integration in a *dam*$^-$ host, indicated that the initial transposition event was conservative (Harshey, 1984). Following this conservative event, all transposition of the phage genome occurs via a replicative mechanism (Craig, 1996).

Due to the ease with which Mu can be manipulated, transposition of the phage has been studied in great detail and is well understood. The basic mechanism that is accepted to describe both non-replicative and replicative Mu transposition is the Shapiro model (Pato, 1989). A double stranded nick, 5 bp apart, is introduced into the target DNA. Free 3' ends from the nicked donor are ligated to the 5' ends on the target. During conservative replication, the nicking of the donor occurs on both strands. This frees 5' phosphate groups that are nucleolytically attacked by the free 3' hydroxyl groups present on the

target. The gapped regions, present as a result of the mechanism of target cleavage, are repaired by host repair functions, which result in a 5 bp duplication at the target. During replicative transposition, each donor end is cleaved only on one strand. The 3' hydroxyl groups, from the phage genome, are ligated to the free 5' ends on the donor, creating a branched intermediate. The free 3' hydroxyl groups on the target act to prime replication by the host machinery, which is able to copy through the branched structure, copying the phage genome. While *E. coli* replication proteins copy the phage genome, transposition is mediated by phage-encoded proteins (Craig, 1996).

The Mu transposase is encoded by the *A* gene of the phage (Pato, 1989). The transposase protein has a molecular mass of 75kDa and can be divided into three globular domains (Rice and Mizuuchi, 1995). The N-terminal domain contains the site-specific DNA binding activity that recognizes the terminal attachment (*att*) sites of the phage as well as the internal activation sequence (IAS) (Leung *et al.* 1989). The central domain has non-specific DNA binding activity because it contains a helix-turn-helix domain. In addition this region contains a D,D(35)E catalytic domain commonly found in other transposases (IS4 family) (Craig, 1996). The C-terminal domain contains residues that are important for protein-protein interaction. These inter-protein contacts mediate specificity and reaction kinetics due to interaction with Mu B. Furthermore the C-terminal sequence mediates disassembly of the transposase complex when it interacts with the ClpXP protease (Levchenko *et al.* 1997).

Transposase functions, as part of a regulated nucleoprotein complex (transpososome), to ensure legitimate transposition (Mizuuchi *et al.* 1995). The protein is monomeric in solution [referenced in (Rice and Mizuuchi, 1995)] and binds as a

monomer to the transposase binding sites. These sites are located at the ends of the phage DNA genome and define the phage/host junction (Craig, 1996). The left *att* site is composed of three distinct binding sites. Two of these sites (L2 & L3) are separated from L1, the leftmost site, by approximately 100 bp. The right *att* site is one contiguous sequence of three binding sites that are located at the rightmost end of the genome (Craigie *et al.* 1984). Binding of the A monomers to the individual sites is reversible (Harshey, 1988). In order for the transpososome to form, the binding must be stabilized and a higher order nucleoprotein complex assembled (Mizuuchi *et al.* 1995). Stabilization of the transposase/DNA interaction occurs via interaction with *cis* and *trans* activating factors.

Approximately 1kb from the left end of the Mu genome is the immunity region for the phage, an IHF binding site and the IAS (Harshey, 1988). The operator sites, IHF site and IAS all overlap and yet are bound in distinct manners. The IAS was shown to interact with gpA for stabilization of the binding of the protein to the left end *att* sites (Mizuuchi *et al.* 1995). This stabilization was shown to occur *in vitro* with superhelical substrates. When IHF was present in these reactions, however, the required degree of superhelicity was reduced to near physiological levels. IHF is known to bind to and bend DNA (Rice, 1997) and it is believed that the bending of the phage genome near the IAS facilitates binding of the transposase to this site. In addition to IHF, another host encoded histone-like protein, HU, is required for transposase stabilization. *In vitro* experiments, with isolated transpososomes, revealed that HU was present within the complex and bound to the DNA near the *att* sites (Lavoie and Chaconas, 1993). The binding of HU to the DNA, with transposase at the *att* sites, was shown to bend DNA. Again, this bending may

facilitate the initiation of stable transpososome formation (Craig, 1996). A complex DNA structural rearrangement, involving gyrase, is believed to restructure the phage genome in such a way that the various proteins and *cis*-acting factors may interact. A strong gyrase binding site is present at approximately 18.1kb from the left end of the phage genome (Pato *et al.* 1990). The action of gyrase at this site increases the local superhelicity of the genome. This, it is believed, brings the phage ends together in order for the transoposasome to form (Pato and Banerjee, 1996). The location of the site, midway between the two *att* sites, makes the mechanism credible for the formation of such a DNA structure. In addition to protein and nucleic acid cofactors, the initiation of this structural rearrangement requires the presence of divalent cations. While Ca^{++} and Mg^{++} can be utilized for complex formation, only Mg^{++} is used by the transposase for catalytic activity (Craig, 1996). As the transposasome forms, the transposase undergoes structural rearrangement; the protein forms a tightly bound tetrameric complex, the stable synaptic complex (SSC) (Lavoie *et al.* 1991), with the three endmost sites on the phage. The SSC is primed to initiate transposition upon interaction with the final transactivating protein, the phage encoded B protein (Craig, 1996).

Although the B protein is not essential for transposition, it stimulates the reaction 100-fold and mediates transpositional immunity. The SSC recognizes a target site via interaction of the transposase tetramer with the B protein bound to a potential insertion site on the host DNA. The B protein has ATPase activity and binds DNA in its ATP bound form. Interaction of B with A stimulates ATP turnover and the release of B from the DNA (Craig, 1996). Target selection by the B protein is random. However, a preference for a highly repeated sequence (N-Py-G/C-Pu-N) has been shown (Craig,

1996). While random target selection is important for the highly processive Mu replication mechanism, transposition into or very near to a previously integrated genome may have deleterious effects. Transpositional immunity prevents such events and is mediated by the B protein (Adzuma and Mizuuchi, 1988).

The SSC-target-MuB complex, in the presence of Mg^{++} (Craig, 1996) can be converted to a cleaved donor complex (CDC) when the phage ends are nicked by the transposase. This event can be stimulated by the B protein even without its binding to DNA (Surette and Chaconas, 1991). The CDC can initiate strand transfer to the target site via a transesterification mechanism to form the strand transfer complex (STC) (Craig, 1996). The transesterifiction mechanism is energetically neutral and therefore while transposition requires ATP, for B function, it is not required as an energy source The STC is the classical Shapiro intermediate which is primed for replication by host mechanisms, due to the release of 3'-OH groups on opposing target strands. After the joining of the donor and target strands, replication is completed by the host DNA polymerase (Craig, 1996).

Regulation of Phage Transcription

Transcription of regulatory and replicative functions, during the lytic cycle of phage Mu, is the first in a series of temporally regulated transcription events typical of dsDNA phages. Temporal regulation is employed, at requisite times, to produce various phage specific gene products during the replicative cycle. Transcription of the early functions terminates within the SEE region 9.2kb from the left end of the phage (Marrs

and Howe, 1990). Transcriptional regulators that are produced during this phase of transcription include the Ner protein, which acts as a repressor to control early transcription (Harshey, 1988), and a protein designated Mor (Mathee and Howe, 1990). This protein, which is produced from a gene within the SEE region, regulates the transcription of the middle operon of the phage (Mathee and Howe, 1990) in which the late gene transactivator (C) is produced (Margolin and Howe, 1986). Thus, by staggering the transcription of transactivators, the phage is able to temporally regulate the transcription of its gene products.

Middle Transcription-Mor

The *E. coli* RNA polymerase holoenzyme (RNAP) is required for transcription of the middle promoter. *In vitro* transcription experiments have demonstrated that when the DNA recognition subunit, σ^{70}, is specifically removed from the reaction, by antibodies directed against the recognition subunit, transcription from the P_m promoter is lost (Mathee and Howe, 1993). Furthermore, it was shown that in a reconstituted system, that consisted of the RNAP core and purified Mor protein, transcription was dependent on the addition of σ^{70} to the reaction. Therefore, the Mor protein itself does not direct transcription by acting as a replacement for the normal bacterial σ^{70} RNAP subunit (Mathee and Howe, 1993). Due to the absence of any leader sequences that could act in antitermination and the similarity between Mor and the Mu late gene transactivator *C*, it was hypothesized that Mor acted as an activator (Mathee and Howe, 1993). The capacity to act as an activator was further examined when the ability of Mor to bind to DNA was

demonstrated with a DNA fragment from the region upstream of the *C* gene (Kahmeyer-Gabbe and Howe, 1996). It was shown previously that the promoter of the *C* gene differed from the canonical sequence of *E. coli* σ^{70} promoters (Stoddard and Howe, 1990). While a recognisable –10 region was present, there was no homology to a consensus –35 region. Thus, the Mor protein may function as an activator by removing the requirement for a consensus –35 region. In fact, with the use of DNase protection assays, it was demonstrated that the Mor protein could bind a region of the C promoter that overlapped the sequence at –35 (Kahmeyer-Gabbe and Howe, 1996). Kahmeyer-Gabbe and Howe (1996) also showed that there were potential binding sites for the *E. coli* IHF protein, Mu A and c. The presence of the repressor binding sites is important for regulation of the function of Mor. It is believed that the phage genome has three regions that are transcribed during the lysogenic state; these include the repressor, the *gin* invertase, and the *gem* region that overlaps *mor*. The binding of c to its sites, within the *C* promoter, serves to block any aberrant transcription initiated in the lysogen by any Mor protein produced during lysogeny. This secondary safeguard complements a primary requirement for transcription of P_m that inhibits production of C in the lysogen. Mor is not only required for transcription of the C protein but, *in vivo,* the *C* gene is only transcribed from a replicating template (Marrs and Howe, 1990). In the absence of these inhibitory factors, Mor activated transcription at P_m through interaction with the C-terminal region of the α subunit and σ^{70}subunit of RNAP. Both mutational studies and yeast "trap assays" were used to determine that the transactivator protein interacted directly with the polymerase (Artsimovitch *et al.* 1996b). In addition to the interaction with RNAP, Mor seems to affect transcription through changes in the conformation of DNA in the P_m

region. Circular permutation assays revealed that the binding of Mor to P_m did not induce any gross bends in the DNA (Artsimovitch and Howe, 1996). Nevertheless, distortions of the promoter immediately downstream of the Mor region were observed when a variety of DNase protection assays were carried out in the presence of both Mor and RNAP (Artsimovitch *et al.* 1996a). The distortion may be necessary to facilitate the interaction between the proteins or to nucleate open complex formation during the initiation of transcription. Ultimately, transcription from P_m produces the late gene transcriptional activator C.

Late Transcription-C

Transcription, from the late promoters (P_{lys}, P_l, P_P and P_{mom}) of Mu is absolutely dependent on the late gene transactivator C (Stoddard and Howe, 1989; Harshey, 1988). The C protein is homologous to the Mor transactivator, but the proteins may not functionally substitute for each other (Mathee and Howe, 1990). In a manner analogous to Mor, C is able to up-regulate the transcription of the Mu late genes. The sequence of *C* led investigators originally to surmise that the protein might function as an alternate sigma factor (Margolin and Howe, 1986). More detailed binding studies, however, indicated that C was likely to be an activator protein was capable of binding to specific regions of the Mu late promoters. Similar to P_m, the late promoters lacked –35 regions that were recognized by *E. coli* RNAP (Bolker *et al.* 1989). Studies by Margolin and Howe (1990) with antibody depletion of σ^{70} and defined *in vitro* transcription assays, confirmed that the transcription of the late promoters required the RNAP holoenzyme.

These studies indicate that C is indeed an activator. The mechanism by which C activates transcription may therefore occur either by an interaction with RNAP or by a conformational change in the promoter DNA structure. The activation of P_m was shown to occur when the Mor protein interacted with both the α and σ subunits of RNAP (Artsimovitch *et al.* 1996b). Despite the homology between Mor and C, similar RNAP-C interactions do not appear to be involved in the activation of late transcription. Deletion studies by Sun *et. al.*(1998) indicated that *in vitro* transcription could proceed from a late promoter, P_{mom}, in the absence of an interaction between C and RNAP. This is in direct contrast to the mechanism of P_m activation by Mor. Other studies have indicated, that similar to the P_m activator, C is able to change the conformation of the DNA within late promoters (Ramesh and Nagaraja, 1996). It had been recognized that the C protein was required to recruit the RNAP from a non-productive interaction in the upstream region of the P_{mom} promoter to a position that could produce a proper open complex formation [referenced in (Ramesh and Nagaraja, 1996)]. Ramesh and Nagaraja (1996) utilized a number of footprinting techniques to discover perturbations in the P_{mom} promoter that might facilitate the correct binding of RNAP. The footprints revealed a number of physical changes to the twist of the promoter that could lessen the improper spacing found in P_{mom}. This allows RNAP to bind correctly and to initiate transcription of the late genes.

Mu Late Genes Morphogenesis and Packaging

The transcripts, produced from the C dependent promoters P_{lys}, P_I and P_P, encode all of the morphogenetic genes of phage Mu (Howe, 1987). While the morphogenesis and packaging of many of the tailed dsDNA phages are understood in great detail, little is known about these processes in Mu, or its related phages. The head of phage Mu is approximately 54 nm in diameter, while the non-flexible tail, to which tail fibers are attached, is approximately 100nm in length (Harshey, 1988). Few studies have been carried out, to identify all of the genes that encode morphogenetic proteins. However, characterization and functional analysis of proteins produced from minicells infected with amber mutant phages led to the identification of some of the late proteins (Giphart-Gassler *et al.* 1981). Furthermore, functional characterization of phages with amber mutations in individual late genes has been carried out by electron microscopic examination (Grundy and Howe, 1985).

The late proteins of Mu were divided into two groups. One group of proteins produces heads and the other group of proteins produces tails. However, there are at least two additional proteins that are encoded within the late region that are not strictly required for head or tail synthesis but rather function in lysis (*lys*) and packaging (Harshey, 1988). The analysis of proteins produced in minicells has identified the major capsid proteins to be encoded by *T* (33kDa) and *F* (54kDa) while *H* (64kDa) was a minor component of the capsid (Giphart-Gassler *et al.* 1981). The tail is made up of a number of proteins, the tube, sheath, baseplate and fibers. The proteins recognized as part of the tail included *L* (55 kDa), *Y* (12.5 kDa), *S* (56kDa), *U* (20.5kDa) and *N* (60kDa) (Giphart-

Gassler *et al.* 1981). The major proteins isolated from purified Mu phage included *T, L* and *Y*. The T protein is the major capsid protein (Shore and Howe, 1982). The L protein is the tail sheath while Y is a component of the tail of unknown function (Harshey, 1988). The data gave some indication of the content and function of the proteins in Mu phage. Direct examination of the structures produced during amber mutant infection, however, provide more information on the function for each of the individual proteins.

An exhaustive study (Grundy and Howe, 1985) examined such amber mutant structures and assigned more definitive functions to the proteins of protein function. Mutations in *lys, F, G, S, U* and in some cases *H* had no effect on the phage structure. The product of the *lys* gene aids in the disruption of the host cell and therefore has no effect on the structure of the phage. The products of the *S* and *U* are components of the tail fiber (Grundy and Howe, 1984). Thus, while mutations in these genes are expected to affect a component of the phage, the changes are such that they do not perturb the overall structure of the particle. While the F protein was found to be a major component of the head (Giphart-Gassler *et al.* 1981), its function was undefined and therefore no conclusions could be drawn about its function during morphogenesis. The effects of different *H* mutants varied and consequently, no definitive function could be attributed to this protein (Grundy and Howe, 1985).

Mutants in *L, M, N, P, Q, R, Y, V,* and *W* have resulted in the accumulation of normal prohead structures; mutations in *K* have produced heads and abnormally long tails (Grundy and Howe, 1985); *L* is known to encode the tail sheath and *K* functions in tail length determination. But, a specific role in tail assembly could not be assigned to any of the other genes. Extracts made from cultures infected with phages with mutations in *T*

and *I* contained only tails. The T gene is known to encode the major head protein (Shore and Howe, 1982) while the I gene is believed to encode the scaffold (Grundy and Howe, 1985). Extracts made from cultures with phages with mutations in the *D* and the *E* genes contain fully formed proheads and tails. This suggests that they are involved in DNA maturation and packaging.

Recently, a more detailed study of Mu head assembly was undertaken. This study demonstrated that *H* encoded the portal protein, while *I* encoded the scaffold (Grimaud, 1996). The use of a number of phages, with various amber mutations in the *H* gene, showed that the protein could assemble into an oligomeric structure. This structure has a sedimentation coefficient of 25S. The portal protein is apparently processed from 64kDa to 50kDa after assembly of the capsid precursor. The protease that cleaves the portal protein was not definitively identified. However, it was speculated that the proteolytic function might be associated with *I* since promoter proximal amber mutants of this gene produced proheads that accumulated with unprocessed H subunits. However, scaffold function was also ascribed to the I protein by Grimaud (1996). Therefore, the resulting loss of proteolytic function may be due to the inability of the phage to incorporate a protein such as a protease. Such a function had been demonstrated for the scaffold protein in P22 (Greene and King, 1996). Furthermore, other phage systems require the presence of a scaffold for the efficient formation of proheads (Thuman-Commike *et al.* 1998).

Following the formation of the packaging precursors, the phage genome must be matured and packaged into the prohead shell. The Mu genome is packaged by a headful mechanism that is initiated at the left end of the phage and proceeds past the right end, thereby incorporating 1-3 kb of host sequence (Bukhari and Taylor, 1975; Daniell *et al.*

1975). The left end is also flanked by host derived DNA, of which only 46-144bp are bound to the genome (George and Bukhari, 1981). For the phage to initiate the encapsidation of the genome and its flanking sequences, a phage-encoded enzyme must specifically recognize the integrated chromosome. Studies by Harel *et. al.* (1990) have demonstrated that a specific sequence at the left end of the phage genome is responsible for packaging specificity. A region, from base pairs 32 to 54, with respect to the left end of the genome, mediated phage-dependent transduction of a plasmid-borne, antibiotic resistance marker. This region, designated the *pac* site, shows partial dyad symmetry that may be important for enzyme recognition or function. The final phage-directed function necessary for successful completion of the lytic cycle is disruption of the host cell. The lytic functions encoded by phages have been reviewed by Young (1992). Lysis of the host is accomplished by a temporally regulated pathway which consists of a membrane bound regulator and the lytic enzyme. Studies, of lysis by Mu, have yet to demonstrate the presence of a regulatory function in Mu (Young, 1992). Despite the intense research that is devoted to many aspects of the Mu life cycle, little is understood about its mechanism of morphogenesis and packaging. Although many of the assembly precursors of the phage have been identified, thus far the individual components of the prohead and tail have yet to be localized within these structures. Therefore, unlike most other well-studied dsDNA phages, no defined, *in vitro* assembly system for the prohead, or tail exists. Additionally, the packaging enzyme of the phage has yet to be cloned and expressed; this hinders the development of a defined *in vitro* maturation and packaging system. While such systems are available for other phages, the unique mechanisms of Mu replication and encapsidation necessitate that a defined packaging system be developed. Since the *pac*

site is located between the att sites L1 and L2/3 (Harel *et al.* 1990), the packaging machinery must assemble at the same site as the transpososome. A defined assay could, therefore, enable a biochemical diagnosis of the unique competition that might occur at this site.

Chapter Two

Characterization of the Enzymatic Functions encoded within the P_{lys} Transcript of Bacteriophage Mu

Abstract

The double stranded DNA of bacteriophage Mu must be matured and packaged from host DNA late during the lytic cycle in order to produce progeny virions. While control of late gene transcription is becoming well understood, little is known about the phage morphogenetic process. To investigate the latter, we have cloned and sequenced 4 kb of the phage DNA immediately downstream of the first late promoter P_{lys}. Previous amber mutant analysis had located the lytic (*lys*) and DNA maturation genes of the phage in this region of the phage genome. Sequence data of these amber mutants and wild type phage DNA, derived from this region, were analyzed and 7 potential open reading frames (ORFs) found. Cloning and expression of the promoter proximal region of the P_{lys} portion of the phage genome enabled production of cell-free protein extracts that specifically recognize the phage encoded packaging sequence (*pac*), characteristic of phage maturation enzymes. Additionally, the growth characteristics of strains expressing these late gene constructs show that lysis of the cultures is induced as a result of the expression of one or more phage-specific proteins expressed from this region. Deletion mutant and computer analyses of this region, however, suggest a deviation from the phage lysis paradigm best characterized by that of bacteriophage lambda. We therefore propose that Mu lysis may occur by a distinct mechanism, while the packaging enzyme is encoded by gene(s) that share a number of characteristics with previously characterized phage DNA maturation enzymes.

Introduction

The ~37kb double-stranded DNA (dsDNA) genomes of temperate coliphages Mu and D108 replicate via transposition to yield a dispersed concatomer of bacteriophage genomes randomly integrated in the host chromosome. To complete the replicative/lytic cycle, the phage genome must be recognized, cleaved, packaged and the host cell lysed. The recognition/cleavage reaction mechanism has been shown in other phage systems, such as λ, T4, SPP1, P2 (Gold *et al.* 1983; Rao and Black, 1988; Chai *et al.* 1995; Bowden and Modrich, 1985), to be catalyzed by an enzyme complex composed of two phage-encoded proteins, one small subunit which recognizes the specific phage sequence, plus one large subunit which catalyzes the DNA cleavage (Black, 1988). The phage encoded lytic systems of a number of phages, including λ, 21, T4, and P22, have been characterized and been shown to be composed of a multi protein timing/effector system (Young, 1992).

The best-studied maturation system(s) is the terminase complex of bacteriophage λ. Lambda terminase is composed of two subunits of 21 kDa and 74 kDa that are encoded by the phage genes, Nu1 and A, respectively (Gold and Becker, 1983). The genes encoding the two terminase subunits are located immediately downstream of the left end *cos* site. They are transcribed as a single unit, and are arranged in overlapping open reading frames (Johnson *et al.* 1991). The smaller subunit, Nu1, recognizes and binds to the *cos*B region of the *cos* site (Shinder and Gold, 1988), and positions the complex in such a manner that the larger subunit A can specifically cleave the *cos*N region (Feiss *et*

al. 1983) to generate a 12 bp 5' extension. In order for the terminase complex to correctly orient the enzyme to nick at *cos*N, Nu1 must bind ATP (Becker and Gold, 1988). The A subunit of the enzyme can then cleave the *cos*N site to produce the staggered nicks, but must hydrolyze ATP in order to separate the strands, at *cos*, prior to encapsidation (Higgins *et al.* 1988).

The heterodimeric terminase paradigm is used by many other phage systems. The terminase genes of T4, gp16 and gp17, have been described (Valle *et al.* 1996) and possess most of the characteristics of the λ terminase complex. In addition to these coliphages, phages such as SPP1 of *Bacillus subtilis* (Chai *et al.* 1997), P22 of *Salmonella typhimurium* (Poteete and Botstein, 1979), and the Lactococcal phage bIL41 (Parreira *et al.* 1996) all use a multifunctional heterodimeric enzyme complex for the maturation and packaging of their replicated genomes.

Phage-encoded lysis systems have yet to be defined as clearly as the maturation systems. The endolysins of the phage systems mentioned above are soluble proteins that are not exported from the cytoplasm (Young, 1992). Phages have acquired proteins that enable them to control the lysis of their host and allow the endolysins access to the peptidoglycan in the periplasm. These latter proteins are termed holins, the best characterized of which is the S protein of bacteriophage λ.

There are two mechanisms by which the S protein is able to control the release of the λ endolysin: the opposing action of two forms of the protein, and the timed opening of the hole that is formed by the oligomerized protein. However, both of these functions are interconnected. The so-called dual start motif of S translation allows for the production of two proteins that differ by only 2 amino acids at the N-terminus (Blasi *et al.* 1989). The

longer of the two polypeptides contains, in addition to an extra methionine, a lysine residue at position 2, which imparts a slight positive charge to the N-terminus (Young, 1992). Oligomerization of the two forms of the protein in the membrane does not immediately create a hole in the cytoplasmic membrane. Instead, it was shown that the energized state of the membrane impeded the structural conformation that allowed the passage of the endolysin (Blasi *et al.* 1990). Early lysis can be induced by the addition of an energy poison, such as CN^-, which causes the membranes to lose their functionally energized state (Reader and Siminovitch, 1971). While the host is able to maintain its membrane potential, the oligomers remain "closed". Insertion of the holin in the membrane, however, causes a "leakiness" which eventually induces the proton gradient to collapse. Once the collapse occurs and membrane potential is lost, the oligomers can undergo a conformational change creating a channel through which the endolysin may pass (Blasi *et al.* 1990).

The homologous systems, in other phages, have permitted the categorization of various holins according to a number of structural criteria including size and start motif (Young and Blasi, 1995). While the holins, within any group, may show striking sequence homology, it is the structural homologies which are functionally most important. Complementation studies have shown that holins that are completely unrelated, on a sequence level, can be functionally interchangeable.

Phages Mu and D108 maintain their genomes as transposons at all times, with host sequences covalently linked to both the left and right ends of the phage genome (Bukhari *et al.* 1976). Consequently, they must employ a maturation mechanism that not only recognizes the phage genome, but also maintains the genome as an integrated

species during and after encapsidation. The *pac* site of Mu and D108 has been defined as a region of partial dyad symmetry that is located from bp 32-54 with respect to the left end of the phage DNA (Groenen and van de Putte, 1985; Harel *et al.* 1990). The Mu *Pac*ase cleaves in 10 bp intervals beginning approximately 50 bp upstream of the left end of the phage genome, within the adjacent host sequences, to initiate the maturation/packaging process (George and Bukhari, 1981). The phage genome is packaged into the prohead along with 1-3 kb of host chromosomal DNA from the region downstream of its right end, by a "head-full" packaging process (Howe, 1987). The phage must also have sufficient time in order to complete the morphogenetic process. This would indicate that Mu and D108 should employ an endolysin/holin system similar to that of other complex phages in order to properly regulate the lysis of the host (Young, 1992)

Amber mutant analysis of the late region has loosely defined a number of complementation groups that encode structural as well as enzymatic functions associated with the morphogenetic pathway (Grundy and Howe, 1985). Based on the amber mutant analysis, the genes encoding the two subunits of the Mu *Pac*ase were suggested to be *D* and *E*. These two genes lie downstream of the *lys* gene, in the first (P_{Lys}) of four late transcripts and are produced as a result of the Mu-directed, temporally regulated expression of the phage late genes (Howe, 1987). We have sequenced a portion of Mu DNA, encoding this first transcript containing these putative genes, and found a number of open reading frames that might encode various proteins, including an endolysin and the *Pac*ase. In addition, sequencing of a number of amber mutant strains was undertaken to more precisely define the ORFs and assign them to specific genes. A functional analysis

of proteins produced from this region showed an activity was encoded that conferred specific binding for a DNA probe that contains the Mu *pac* site. This indicates that this region encodes at least the *pac* recognition subunit of the Mu *Pac*ase and is suggestive of the gene order being lysis, followed by DNA maturation following the P_{lys} late promoter, a situation seen in other phages such as λ.

Material and Methods

Media and reagents

Luria-Bertani (LB) broth was used for the growth of all bacterial strains with the exception of *in vivo* labelling experiments, when M9 medium was used (Miller, 1992). Media was supplemented, where necessary with the following final concentrations of antibiotics: ampicillin (50 μg/ml), kanamycin (50 μg/ml), and chloramphenicol (10 μg/ml). Radioisotopes L-[^{35}S]Met (15.15 μCi/μl), [α-^{32}P]ATP and [α-^{32}P]CTP (3000 mCi/ml) were purchased from Amersham Inc.

Bacterial Strains and plasmids

E. coli strain JM109 [e14$^-$(McrA$^-$), Δ(*lac-proAB*), *thi, gyrA96* (Nal^r), *endA1, hsdR17* (r_k^- m_k^-), *relA1, supE44, recA1*/ F′ *traD36, lacI^q* Δ(*lacZ*)M15, *proA$^+$B$^+$*; (Yanisch-Perron *et al.* 1985)] was used to harbor plasmids and from which crude lysates were prepared. *E. coli* BL21 (DE3) (F$^-$ *ompT, gal, lon, dcm, hsdS$_B$*(r_B^- m_B^-), DE3; (Studier *et al.* 1990)) was used for the overproduction of phage specific proteins from the T7 promoter. Strains MH2522, MH2524, MH2528 MH2529 (Engler *et al.* 1980) and plasmids pWM5, pKN35 (Marrs and Howe, 1983) and pLC3 (Zha *et al.* 1994) have been described previously, and were the kind gift of Dr. M. M. Howe (U of Tennessee). Plasmid pFS100, which contained approximately 3kb of the P_{lys} region of phage Mu, was constructed by

inserting an approximately 4.5 kb *Hpa*I-*Bam*HI fragment of the phage genome into plasmid pUC119 (Vieira and Messing, 1987) which had been previously hydrolyzed with *Bam*HI and *Hinc*II. Nested deletions of pFS100, used for sequencing, were created using the Exo-Size Deletion Kit (New England Biolabs) according to the manufacturer's instructions. The pFS plasmids were sequenced by the dideoxy-chain termination method (Sanger *et al.* 1977) using the Sequenase version 2.0 Kit (United States Biochemical Corp.). Additional sequencing was performed on Mu DNA downstream of the *Bam*HI site at position 14331 using primers designed based on the sequence obtained from the pFS plasmid series sequence data. Plasmid pRS101 was constructed as follows: a 14.5 Kb *Bam*H I fragment was isolated from purified Mu DNA. The *Bam*HI site was made flush with T4 DNA polymerase (GibcoBRL) as per the manufacturer's instructions and the fragment hydrolyzed with *Bgl*II. The approximately 4.1 Kb blunt-*Bgl*II fragment was then ligated into pET29b (Novagen) that had been previously hydrolyzed with *Bgl*II plus *Eco*RV. Plasmid pRS102 was derived from pRS101. The pRS101 construct was hydrolyzed with *Eco*RI plus *Mfe*I, generating two fragments of approximately 6.3 and 3.1kb in size. The 6.3 kb fragment, containing approximately 900bp of cloned phage DNA, was isolated and recircularized through the complimentary ends generated by the two restriction endonucleases. The pRS300 series of plasmids were created to sequence *E* amber mutations. Oligonucleotide primers PPEU (5'-AACGCCGTCAGTGAAGAACT-3') and PPED (5'-GCTCGCATTTCATCGTCAA-3') were used to amplify the putative *E* gene from chromosomal DNA preparations. The PCR products were isolated from the reaction mixture using a Promega Wizard PCR prep kit according to the manufacturer's instructions. The isolated DNA fragment was then ligated into pBluesript KS-

(Stratagene) that had been previously hydrolyzed with *Eco*RV. The ligation products were sequenced with an ABI Prism sequencer (Applied Biosciences), at the Biomedical Resource Centre (BRC, UCSF), using primers PPEU or PPED. Additionally, the T7 promoter proximal region of pRS101 was resequenced, using a T7 promoter primer, at the BRC.

DNA manipulations

Chromosomal DNA was isolated by a CTAB-extraction, "mini-prep" procedure (Ausubel *et al.* 2000). Plasmid DNA was isolated from CsCl-ethidium bromide gradients for large-scale recovery, or by an alkaline lysis "mini-prep" procedure (Sambrook *et al.* 1989) for small scale or analytical preparations.

All restriction enzymes, T4 DNA polymerase, T4 DNA Ligase, and the Klenow fragment of DNA Pol I were purchased from New England Biolabs, Pharmacia Canada, or GibcoBRL and used according to the manufacturer's instructions. Large DNA fragments (>1Kb) were isolated from 0.7% agarose gels with the Geneclean kit (BIO101) as outlined by the manufacturer. Small DNA fragments were isolated from 8% polyacrylamide gels in 1xTBE (90mM Tris-borate, 2mM EDTA; pH 8.0) (Sambrook *et al.* 1989) by the "crush and soak" method (Maxam and Gilbert, 1980). DNA fragments that were to be made flush were incubated with T4 DNA polymerase, in buffers supplied by the manufacturer, at 37°C for 30 minutes prior to inactivation at 65°C for 20 minutes. The Klenow fragment of DNA polymerase I, in Klenow buffer (7mM Tris-HCl [pH 7.5], 7 mM $MgCl_2$, 50 mM NaCl, 1 mM dithiothreitol (DTT)), was used to fill the ends of

fragments with recessed 3' termini and labelled with 40 μCi of the appropriate [α-^{32}P]-dNTP at 20°C for 60 minutes prior to inactivation at 65°C for 20 minutes. Unincorporated nucleotides were separated from labelled DNA by 2 successive ethanol precipitations, and the labelled DNA resuspended in TE (20mM Tris-HCl, 2mM EDTA; pH7.5).

Production of crude extracts containing phage proteins

E. coli JM109, harbouring pLC3, was made competent via calcium treatment (Sambrook *et al.* 1989) and transformed with either pWM5 or pKN35. The cells were grown overnight (O/N) with the appropriate antibiotics (Chloramphenicol [Cam] and Ampicillin [Amp]) to establish growth of transformed cells within the culture. The O/N cultures were diluted 1:100 and grown to an A_{550} of approximately 0.5. IPTG (isopropyl-β-D-thiogalactopyranoside) was then added to a final concentration of 1 mM to induce the expression of the Mu specific late gene transactivating protein (C) and the concomitant expression of the Mu morphogenetic proteins from pWM5 and pKN35 using the P_{lys} promoter. The cultures were incubated for 90 minutes prior to harvesting the cells by centrifugation (4000 X g, 4°C, 15 minutes). The pelleted cells were washed by resuspension in buffer A (10 mM Tris-HCl [pH 7.5]; 10 mM $MgCl_2$; 10 mM β-mercatoethanol), repelleted as described previously and weighed. For each gram of cells, 2.5 grams of levigated alumina was added, and the mixture ground to a gummy paste at 4°C with a prechilled mortar and pestle. The proteins were extracted from the paste with 1.5 ml of buffer A supplemented with 467 mg/ml of phenylmethyl-sulfonylfluoride (PMSF) and 1 M NaCl. The alumina and cell debris was separated from the fluid extracts

by centrifugation (4000 X g, 4°C, 10 minutes). The protein extracts were further clarified by centrifugation (80,000 X g, 4°C, 2.5 hours). The clarified extract was removed and dialyzed at 4°C for 24 hours against 3 changes of storage buffer B (25 mM Tris-HCl [pH 7.5]; 1 mM EDTA; 25 mM NaCl; 10 mM β-mercaptoethanol; 5% [v/v] glycerol) After dialysis, the extracts were aliquoted and stored at -70°C. The protein concentration was determined by the Lowry method (Lowry *et al.* 1951).

Growth curves of strains expressing phage proteins.

Cultures of *E. coli* BL21(DE3) transformed with either pRS101, pRS102, or pET29b were grown overnight in LB with the appropriate antibiotics. The overnight cultures were diluted 1:20 in LB with appropriate antibiotic selection and grown to midlog phase (A_{550}=0.45), at which point T7 specific transcription of the cloned genes was induced by the addition of IPTG to a final concentration of 0.4mM. The growth of the cultures was monitored spectrophotometrically at 550nm. Disruption of inner membrane integrity was initiated by the addition of 0.01 volumes of $CHCl_3$ to the cultures of interest.

Mobility shift assays

Mobility shift assays were performed essentially as described in Tolias and DuBow (1989). The specific 206 bp *Bsa*H1 Mu *pac* site-containing fragment, from plasmid pE8 (Harel *et al.* 1990), and a non-specific 154 bp *Hin*F1 pBR322 fragment were

radioactively labelled. Reactions were carried out in 20 µl volumes. Ten µg of crude protein extract was combined with 5000 cpm of each of the labelled specific and non-specific probes (except in the control lanes), along with 2µl of 10X binding buffer (20 mM Tris-HCl, pH 7.5, 100 µg/ml bovine serum albumin, 0.02% [v/v] Nonidet P-40) and varying concentrations of sonicated calf thymus DNA. The reactions were incubated (37°C, 10 minutes) and then subjected to electrophoresis through an 8% polyacrylamide gel in 1xTBE until the unbound probe had migrated to the bottom of the gel. The gels were then dried and exposed to Kodak XAR-5 X-ray film, under Dupont Cronex intensifying screens, at -70°C, after which the film was developed.

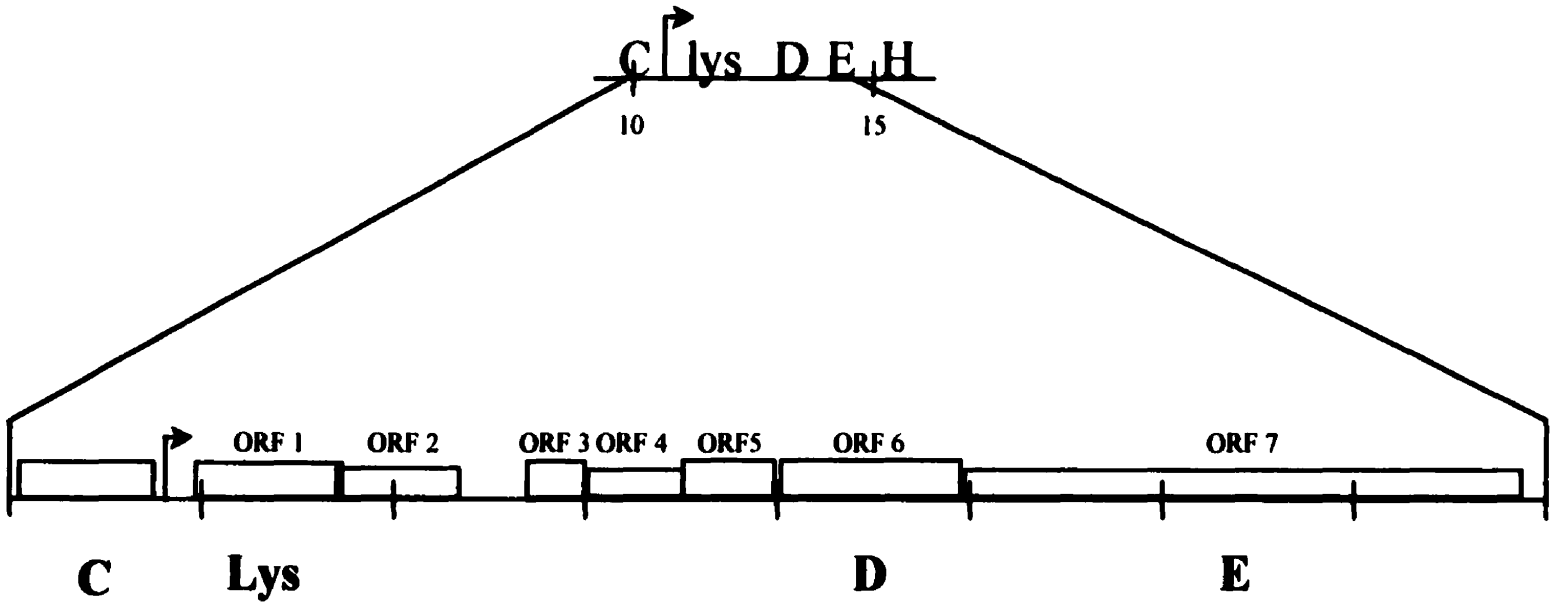
Results

Sequencing of the early morphogenetic functions of phage Mu

Previous amber mutant analysis had indicated that the genes coding for the maturation enzyme of phage Mu were located in the promoter proximal region of the P_{lys} transcript (Grundy and Howe, 1985). In order to investigate the role of genes encoded in this region, sequencing of approximately 4.5 kb of the phage genome downstream of the P_{lys} promoter was undertaken. This sequence has been deposited in the NCBI Genbank and may be accessed through accession number AF007792. The complete Mu sequence has also been determined by Dr. Roger Hendrix and coworkers and deposited in Genebank under accession number AF083977.

Computer analysis of the sequence was used to define the potential open reading frames (ORFs) within the region. Using GeneMarkHMM (Lukashin and Borodovsky, 1998) with default conditions, 7 potential ORFs, shown in Figure 1, spaced through the early morphogenetic region, were found. Examination of the upstream region of the ORFs revealed polypurine tracts, characteristic of ribosome binding sites (RBS), positioned favourably for recognition and translation initiation (Chen *et al.* 1994). The calculated molecular weights for the 7 ORFs are: 18.9 kDa, 14.3 kDa, 7.3 kDa, 11.3kDa, 10.9 kDa, 21.6 kDa, and 62.6 kDa (Figure 1, 1-7 respectively)

Figure 1 Partial genetic map of bacteriophage Mu and potential ORFs. (Top) Map of the P_{lys} proximal region including known genes and those defined through amber mutant analysis. (Lower) Expanded view of the P_{lys} proximal region including ORFS (1-7) defined through computer analysis of the nucleotide sequence of the region. Positions of known or inferred genes are indicated in bold below the schematic map. Vertical lines delineate 600 bp regions from 9900 to 14700 with respect to the left end of the phage. Genebank accession number AF007792. Bent arrows represent the P_{lys} promoter.


C Lys D E H
10
15
ORF 1
ORF 2
ORF 3
ORF 4
ORF5
ORF 6
ORF 7
C
Lys
D
E

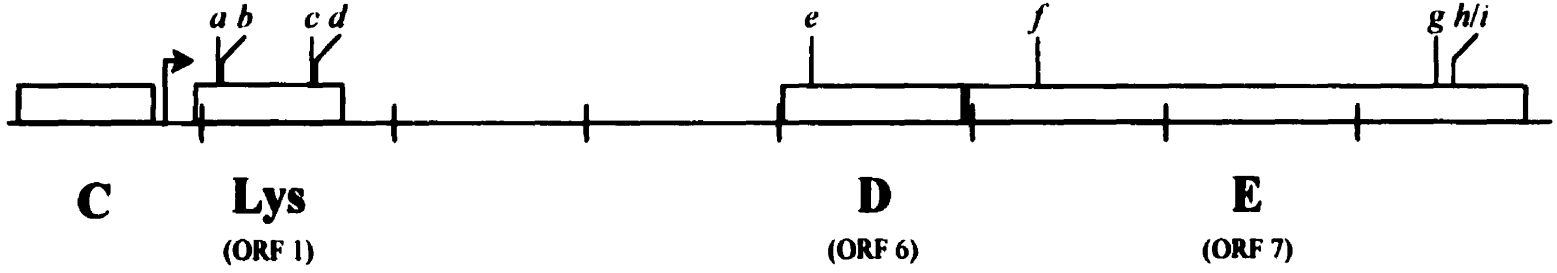
Sequence comparison of the ORFs using the BLASTP algorithm (Altschul *et al.* 1997-searches performed Sept. 1998) indicated that the first ORF of the P_{lys} transcript (Fig. 1,1) encoded a protein that was homologous to endolysins from both bacteriophage and bacterial sources. All other ORFs showed no significant sequence homology to any known proteins in the databases. The final two putative Mu ORFS (Fig. 1; 6 and 7) are arranged with the smaller of the two genes preceding the larger. Protein sequence analysis indicates that the small ORF may take on a highly compact structure due to the presence to a number of helical regions separated by short turns.

Sequence analysis of amber mutations within the P_{lys} proximal region allowed us to identify three of the open reading frames displayed in Fig 2. The mutations indicate that the lysis gene of Mu is encoded by the first ORF downstream of P_{lys} while genes D and E are encoded by the ORFs numbered 6 and 7 in figure 2.

Protein expression effects on cell growth.

To examine which of the ORFs from this region may be expressed into proteins, plasmid pRS101 was constructed. Plasmids pRS101 and pET29b were inserted, individually, into *E. coli* BL21(DE3) (Studier *et al.* 1990) by chemical transformation. Figure 3a shows that cultures of both strains grew normally into log phase prior to induction of T7 RNA polymerase directed protein expression. After a lag of 30 minutes, during which the culture grew normally, the strain containing pRS101 began to lyse rapidly, with the culture essentially completely clearing 60 minutes post induction. The

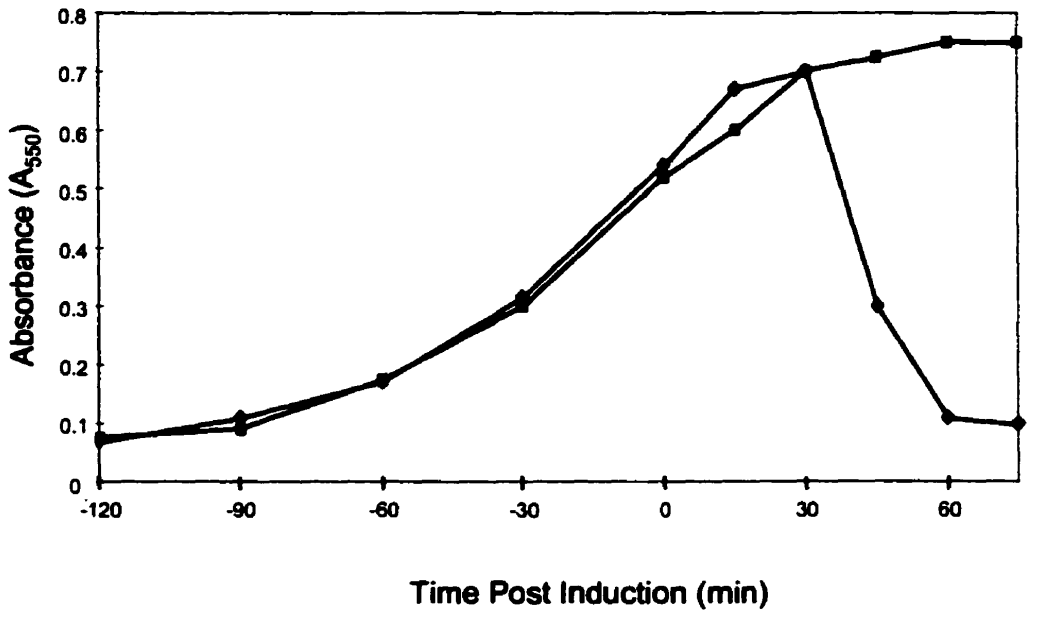
Figure 2. Positions of P_{lys} proximal amber mutants. The positions of the Mu amber mutations are indicated above the boxes representing *lys* (ORF 1), *D* (ORF 6) and *E* (ORF 7) on thepartial genetic map. The table below the diagram lists the specific amber mutation tested, the position of the mutation in the Mu genome and the codon that has been changed.



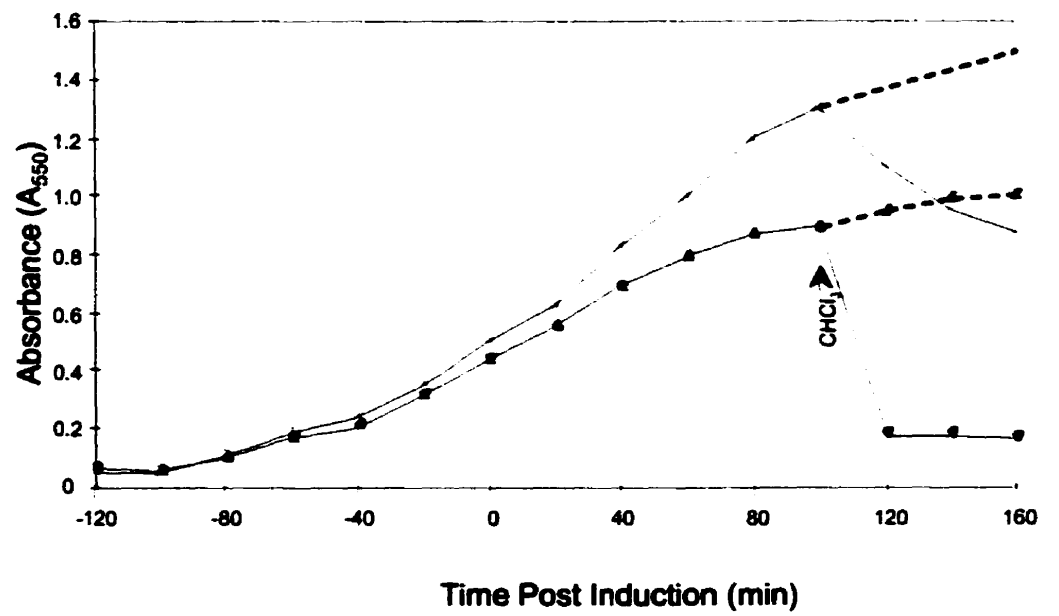
	Mutation	Position	Change
a	*lys* 1030	10555	cag - tag
b	*lys* 7325	10570	cag - tag
c	*lys* 7297	10882	cag - tag
d	*lys* 1025	10895	tgg - tag
e	*D* 7330	12415	cag - tag
f	*E* 7344	13113	tgg - tag
g	*E7367*	14288	cag - tag
h	*E* 7291	14339	cag - tag
i	*E* 7310	14339	cag - tag

Figure 3 Growth profiles of *E. coli* BL21(DE3) strains containing phage constructs. (A) Cells harbouring pET29b (-■-) or pRS101 (-◆-) were induced with 0.4mM IPTG at time zero and followed by monitoring A_{550}. (B) Cells containing the pET29b (-▲-), pRS102 (-X-) [no $CHCl_3$ treatment] or pET29b (-■-), pRS102 (-◆-) [$CHCl_3$ treatment (0.01 volumes) at time indicated by the vertical arrow] were induced, as above, at time zero and followed by monitoring the A_{550}. strain containing the parent pET29b plasmid continued to grow normally after the addition of the inducer.

A


Absorbance (A_{550})
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-120
-90
-60
-30
0
30
60
Time Post Induction (min)

B


Absorbance (A_{550})
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0
-120
-80
-40
0
40
80
120
160
$CHCl_3$
Time Post Induction (min)

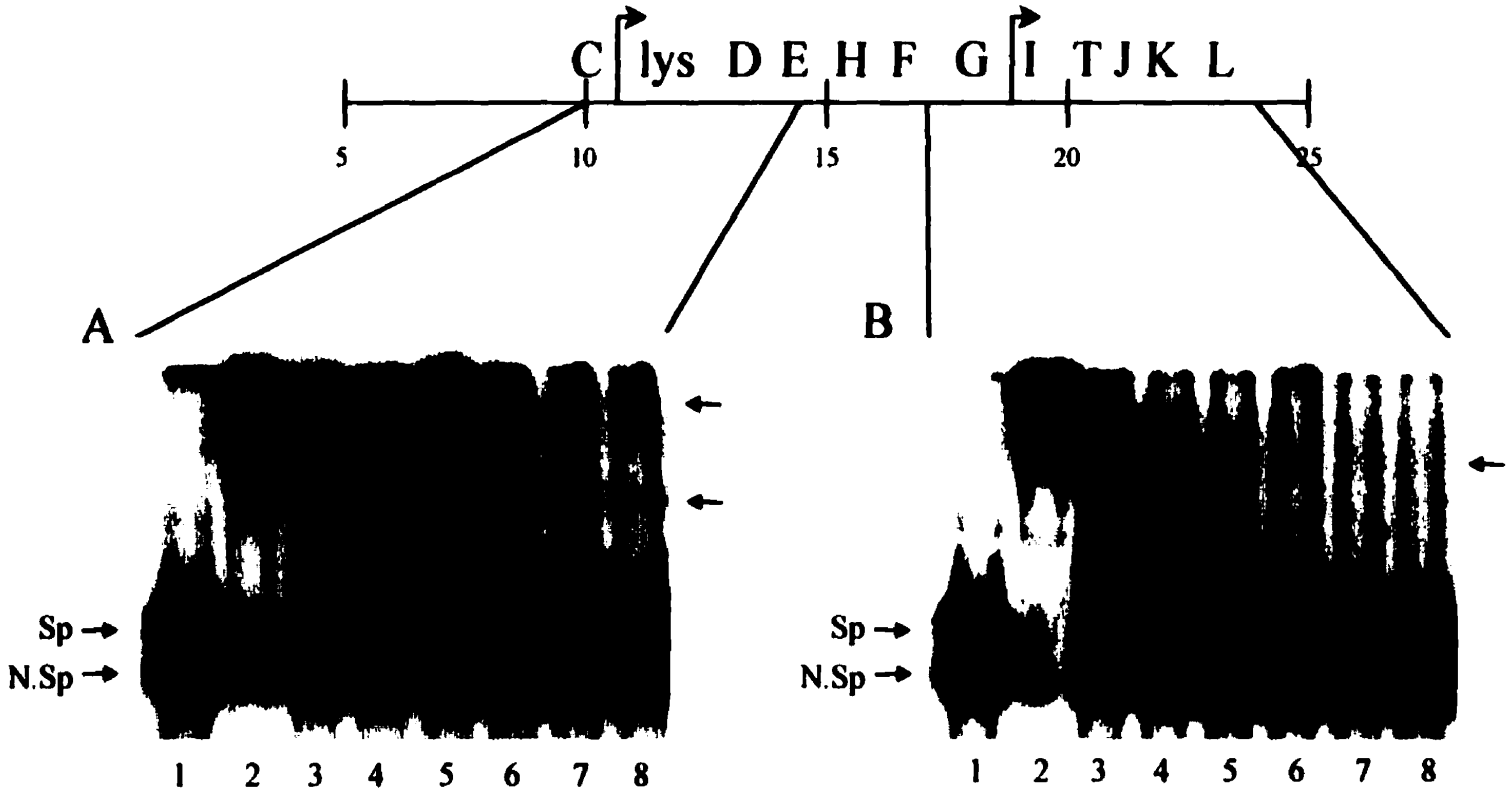
Induction of protein expression from the pRS102 construct allowed us to determine whether or not the lytic effects observed during the expression of pRS101 were due to ORF 1. Figure 3b shows the growth of two BL21(DE3) strains transformed with pRS102 or the parent vector, pET29b. Both cultures grew normally, into log phase, prior to the induction of protein expression from the T7 specific promoter. Subsequent to induction of protein expression, both cultures continued to grow normally. In contrast to the effects of protein expression from a strain transformed with pRS101, the strain containing pRS102 showed no signs of lytic activity during and after the period in which pRS101 had lysed completely.

Cultures, expressing proteins from either pRS102 or pET29b, were grown and split, and one portion of each treated with chloroform. The effects of membrane disruption on these cultures are seen in Figure 3b. Treatment of the culture that contained BL21(DE3) transformed with pET29b produced a decrease in the A_{550} of the culture from 1.3 to 1.0 within 20 minutes. In contrast the strain transformed with the pRS102 construct, upon treatment with chloroform decreased markedly in A_{550} from 0.9 to 0.19 during the same 20-minute incubation period.

Mobility shift analysis of crude extracts from pWM5 and pKN35.

Mobility shift experiments were performed to determine if any of the Mu late genes cloned in pRS101 might encode the DNA-binding component of the maturation function (*pac*ase) of phage Mu. Figure 4 shows autoradiograms of the reactions using

Figure 4 Mu-*pac* DNA binding properties of extracts containing phage proteins as determined by gel mobility shift assay. (Top) A partial genetic map of bacteriophage Mu containing morphogenetic genes present in the first two late transcripts (P_{lys} and P_l) in addition to the late gene transcriptional activator (C). (Lower) Autoradiograms of protein-DNA complexes, prepared from strains containing pWM5 (A) or pKN35 (B), separated on native polyacrylamide gels. Lane 1 specific (Mu *pac*) and non-specific probes without protein. Lanes 2-8 probes with 10 μg/reaction of phage protein containing extracts and 0, 0.05, 0.075, 0.1, 0.2, 0.4, 0.6 μg/reaction of unlabelled, competitor calf thymus DNA respectively. Arrows to the right of the autoradiograms indicate the position of retarded complexes within the gels. Positions of the unbound probes are indicated to the left of the autoradiograms. Angled lines from the genetic map to the autoradiograms delineate the portions of the Mu genome present with the pWM5 and pKN35 constructs. Sp=Specific; N.Sp=Non-Specific


C lys D E H F G I T J K L
5
10
15
20
25
A
B
Sp
N.Sp
Sp
N.Sp
1 2 3 4 5 6 7 8
1 2 3 4 5 6 7 8

extracts of crude proteins produced from portions of the P_{lys} and P_I transcripts of phage Mu. The reactions were carried out in the presence of increasing concentrations of non-labelled, competitor calf-thymus DNA to detect activity, from the crude extract, that indicated proteins bound to the Mu *pac* site. The extract produced from the strain that contained the P_I specific genes revealed a shift of both the specific and non-specific labelled fragments in the absence of any non-labelled competitor DNA (Figure 4 B, lane 2). The subsequent reactions, performed in the presence of competitor DNA, showed disruption of any specific complex formation at low calf thymus DNA concentrations (Figure 4 B, lanes 3-8). The reactions with the lowest concentrations of competitor DNA (Figure 4 B lanes 3 and 4), 0.05 and 0.075μg respectively, showed a smeared band indicative of non-specific interactions which readily dissociate under these conditions (Lane *et al.* 1999). At the higher concentrations of competitor DNA (Fig.4 B lanes 5-8), 0.1 to 0.6μg, no specific protein-DNA complexes are detected.

Reactions performed with an extract from the strain containing the proximal P_{lys} specific genes are shown in Figure 4 (A). Similar to the reaction with the P_I specific extract, the reaction of the P_{lys} extract with the labelled probes, in the absence of competitor DNA, resulted in a strong shift of both probes (Figure 4 A lane 2). In the presence of increasing concentrations of competitor calf-thymus DNA, the presence of two differentially-retarded bands, dependent on distinct competitor concentrations, were observed. In the presence of low concentrations of competitor DNA, the autoradiogram revealed a strongly retarded complex (Figure 4 A lanes 3-5). As the concentration of competitor DNA increases above 0.1μg per sample, the strongly retarded complex begins to diminish and is replaced by a faster migrating complex of lower intensity (Figure 4 A

lanes 5 and 6). This lower complex persists in the presence of concentrations of competitor DNA up to 0.6μg per reaction (Figure 4 A, lanes 7 and 8).

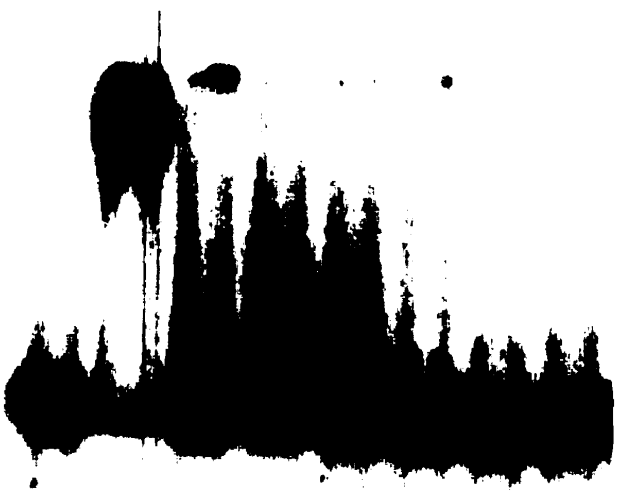
Despite the use of competitor DNA to disrupt non-specific interactions, the complexes revealed in the P_{lys} specific extract (Figure 4 A) cannot be definitively assigned to either the specific or non-specific probes due to the large amounts of free probe present in the reactions. In order to define which of the probes is present in the complexes, the reactions were carried out with each probe individually. The autoradiograms of these reactions are shown in Figure 5. The complexes produced, with the non-specific probe, showed that proteins within the extract bind this DNA fragment in the absence of competitor DNA (Figure 5 B lane 2). However, this complex is not observed when 0.05 to 0.6μg of unlabelled competitor DNA is included in the reaction (Figure 5 B lanes 3-8). In contrast, reactions performed with the specific fragment indicated that the complexes that had formed previously in the presence of both probes were also formed with only the specific probe. The pattern of protein-DNA complex formation, in the presence of competitor DNA, paralleled that seen in reactions with both the specific and non-specific probes.

Figure 5 Mobility shift assay of the pWM5 extract and individual probes. Lane 1 is the specific (A) and non-specific (B) probes without added protein. Lanes 2-8 specific (A) and non-specific (B) probes with 10μg/reaction of the pWM5 extract and 0, 0.05, 0.075, 0.1, 0.2, 0.4 and 0.6 μg/reaction of unlabelled, competitor calf thymus DNA respectively. Arrows to the right of the autoradiograms indicate the positions of retarded complexes while the position of the unbound probes are indicated on the left.

A

Specific →

B



Non-specific →

Discussion

Transcriptional control of the Mu morphogenetic region has been extensively investigated, but relatively little is known about the proteins encoded by this DNA segment. Here we report the sequencing a portion of the late region of the phage along with an initial characterization of some of the functions encoded therein.

The P_{lys} transcript of Mu is the first of the late transcripts. To date, the only gene within this transcript to be cloned and assigned a function is the portal gene (*H*) (Grimaud, 1996). Additionally, amber mutant analysis has assigned lytic (*lys*) and packaging (*D/E*) gene functions to the promoter proximal region of this polycistronic transcript (Grundy and Howe, 1985). We have sequenced 4349 bp of wild-type Mu phage DNA that overlaps sequences from the left end of the genome and extends past the first *Bam*H1 site of the first late transcript. Computer analysis of this region with the GeneMark HMM and BLASTP algorithms reveals the presence of seven potential ORFs with high coding probability and varying degrees of homology to known genes. These ORFs have also been identified by Hendrix and included as part of the whole Mu genome sequence deposited in Genbank. The first of these putative genes has significant homology to a number of phage and bacterial lytic enzymes. In addition, sequence analysis of *lys* amber mutants indicated that the amber stop codons were present within the first ORF downstream of P_{lys}. The next four ORFs displayed no significant homology to any genes of known function. However, they are highly homologous to ORFs present within a cryptic Mu-like prophage of *Haemophilius influenzae* (Fleischmann *et al.* 1995). The final two ORFs again show no significant homology to known genes other than the

reported *H. influenzae* ORFs. However, these two share general similarities with genes that encode the maturation enzyme of other phages.

The maturation enzymes of dsDNA phages, with the exception of P2/P4 (Linderoth *et al.* 1991), are encoded by two adjacent genes. The small subunit-encoding gene precedes the large subunit-encoding gene. Sequence homology was not found between the previously characterized maturation genes and the Mu ORFs. This is generally true for this group of enzymes. Searches of databases using sequence data from previously characterized maturation enzymes reveals homology only within a phage group. In contrast to the differences in sequence, the molecular weights and functions of the two subunits are similar for all phages. The two Mu ORFs have the potential to encode proteins that have molecular weights of 21.6 kDa and 62.6 kDa, within the range reported for the two subunits of the maturation enzymes (Black, 1988). Additionally, database searches with amino acid sequences derived from the nucleotide sequence data indicated that the large subunit Mu ORF has a potential P-loop type ATP binding site. This P-loop type binding center was shown to be the high affinity binding site for ATP in both phages λ and χ DNA maturation enzymes (Hwang and Feiss, 1996; Hwang *et al.* 1996). Finally, sequence analysis of *D* and *E* amber mutants revealed that *D* amber stop codons were within the ORF encoding the 21.6 kDa protein, while the *E* amber stop codons were within the ORF encoding the 62.6 kDa protein. The *E* amber mutant data rules out the possibility that the ORF encoding the 62.6 kDa protein is the *H* gene as stated in Genebank file AF083977. The evidence derived from the sequence data warranted examination of the proteins produced from this region to determine whether or not a *pac*ase-specific function (binding to the *pac* site) could be detected.

Mobility shift assays were performed with crude extracts derived from strains containing plasmids encoding promoter-proximal genes from the P_{lys}, and as a control, P_I promoters. The extracts derived from the pKN35-transformed strain contained a protein(s) that was able to bind to *pac*-specific DNA probes. This binding, however, was not observed even in the presence of low levels of non-specific competitor DNA. In contrast, protein(s) within the extract, produced from the promoter proximal region of the P_{lys} transcript, formed complexes with the *pac*-specific probe that displayed resistance to detection by competitor DNA. Two distinctly migrating complexes were formed at discrete competitor DNA concentrations; we were unable to detect cleavage of *pac*-containing DNA *in vitro* (not shown).

Terminase binding studies in λ have shown that a higher order nucleoprotein complex is formed during initiation of encapsidation (Catalano *et al.* 1995). Thus, the enzyme is not a simple heterodimer but rather a heterooligomeric structure. In addition to the binding of subunit oligomers, accessory proteins such as IHF have been shown to be present within the nucleoprotein complex formed upon initiation of genome maturation (Feiss *et al.* 1988). During the genome maturation process of Mu, the initial left-end cleavage event occurs within the host sequences adjacent to the integrated phage genome (George and Bukhari, 1981). However, the phage *pac* site is located distal to the cleavage sites and within the phage genome itself (Harel *et al.* 1990). The physical separation of the binding and cleavage sites suggests that there must be a mechanism by which these two sites are brought into close proximity. Since the *pac* site is located at the extreme left end of the phage genome, between the transposase binding sites L1 and L2/3, the formation of a maturation complex may occur when the *pac*ase displaces the majority of

the transposasome proteins and reorients the DNA in conjunction with previously bound HU protein. The complexes produced by the proteins encoded by the P_{lys} transcript appeared to confirm the amber mutant data that indicated that the Mu *pac*ase was encoded by this region (Grundy and Howe, 1985). The two ORFs (6 and 7), which have similarity to previously described maturation enzymes (Figure 1, 6&7), both have H-T-H domains, as defined by computer algorithms, which may bind DNA.

The distinct and rapid clearing of the induced cultures upon expression of the P_{lys} proximal region indicates that a lysis function is being produced. Examination of other phage lytic systems suggests that there should be two proteins produced (Young, 1992), an endolysin and holin, if a phage-directed timing system is encoded by Mu. Previously characterized phage lytic systems have been shown to be encoded by genes arranged in a well-conserved manner. The dual-function holin gene immediately precedes the gene encoding the endolysin(s) (Young, 1992). ORF 1 shows homology, by sequence analysis, to endolysins. To further define the functions of the proteins that are necessary for lysis of the host cell, plasmid pRS102 was created. This construct is a deletion derivative of pRS101 in which all but ORF1 has been removed. The growth of this strain, subsequent to induction of protein expression, is similar to that of the parent vector, indicating that any lytic protein that may be expressed is unable to interact with the substrate peptidoglycan present in the periplasm. However, mild disruption of the inner membrane of the bacteria by chloroform, which mimics the effect of a holin, results in the rapid lysis of the cells containing the pRS102 construct (Fig 3b). This suggests that the deleted plasmid no longer contains a gene encoding a holin or holin-like protein, and that ORF1 may encode a protein with endolytic activity. These results suggest that the phage

holin/endolysin paradigm developed from the λ system may not be directly applicable to Mu.

Further investigation into the function of the various ORFs encoded in this region is necessary to identify a protein that enables the phage lytic enzyme access to its periplasmic substrate, as well as purification of the *pac*ase enzyme required for the unusual maturation reaction of this transposable phage's DNA.

The preceeding study allowed us to identify the region that encodes the Mu maturation enzyme, *pac*ase. Furthermore we were able to produce protein extracts containing phage-encoded proteins that have *pac*ase-like activity. We therefore wished to purify other morphogenetic precursors in order to begin refining our *in vitro* packaging assay. The isolation of active proheads from two transposable phages, Mu and D3112, are described in the following chapter.

Chapter Three

Characterization of Isolated Proheads from Transposable Phages

Abstract

The transposable phages of *Escherichia coli* (Mu and D108) and *Pseudomonas aeruginosa* (D3112) have a unique mode of DNA replication and maturation while maintaining the general morphogenesis pathway of other double stranded DNA, tailed phages. Thin section electron micrographs of bacterial cells during the Mu and D3112 lytic cycles show that the development of completed phage particle structures occurs late in the lytic cycle, concomitant with the initiation of host cell lysis. Fractions that are highly enriched for phage proheads were isolated on linear sucrose gradients from crude extracts prepared from thermoinduced Mu *c*ts62 lysogens late in the lytic cycle. These prohead fractions were shown to contain a function that can bind to Mu *pac*-specific DNA using mobility shift assays. A specific Mu DNA substrate can be packaged into these active prohead fractions in *in vitro* packaging reactions that are dependent on ATP and spermidine. The packaging of the DNA substrate is stimulated by the inclusion of the host cell histone-like proteins IHF and HU. The refined packaging system has allowed us to propose a mechanism that controls the switch from replicative transposition to DNA packaging in the transposable phages.

Introduction

Mu and D108 are two related temperate phages of *Escherichia coli* that replicate their 37kb linear double-stranded DNA (dsDNA) genomes by replicative DNA transposition. These two phages are the only known coliphages that use such a mechanism for genome amplification and lysogenization/prophage integration. However, a large number of transposable phages have been isolated from *Pseudomonas aeruginosa* (Pato, 1989). The best studied of the Pseudomonas transposable phages is D3112 (Bidnenko *et al.* 1989). The genetic organization of the transposable coliphages and D3112 is strikingly similar. However, there is little homology between the two phage groups (Bidnenko *et al.* 1989; Pato, 1989). In contrast, Mu and D108 share over 90% homology, at both the DNA and protein level (Gill *et al.* 1981; DuBow and Bukhari, 1981). Despite the extensive homology between these two phages, they differ in their regulatory regions and, therefore, cannot confer immunity to superinfection by the other phage

The transposable phages have been extensively utilized as a model system for the study of DNA transposition, and a great deal is understood about their replicative functions (Pato, 1989; Craig, 1996). In addition, the regulatory regions which control their lytic/lysogenic switch and early gene expression, present at the left end of the phages, have also been extensively investigated (van de Putte *et al.* 1981; Van Leerdam *et al.* 1982; Tolias and DuBow, 1985; Levin and DuBow, 1987; van Rijn *et al.* 1989; Kukolj and DuBow, 1992). Despite the depth of knowledge gained about the early and replicative

functions of these phages, little is known about the genetics and biochemistry of their morphogenesis.

Structural studies of several dsDNA phages have revealed the strict geometric nature of their assembly intermediates (Casjens and Hendrix, 1988). Early studies utilized thin sections of host bacteria that contained the growing phages and their precursors (Murialdo and Becker, 1977; Earnshaw and King, 1978; Black *et al.* 1994; Lee and Guo, 1995). More recently, cryo-electron microscopy and image reconstruction techniques have enabled researchers to enhance the visualization of protein-protein and protein-nucleic acid interactions during phage morphogenesis (Dokland *et al.* 1993; Dokland and Murialdo, 1993; Marvik *et al.* 1995; Thuman-Commike *et al.* 1998; Hendrix and Duda, 1998). Phage precursors (either in crude extracts or purified) have been utilized to develop *in vitro* packaging assays to elucidate the biochemistry of many packaging reactions. This type of assay has been used in the ϕ29 system to describe the role of a unique RNA molecule that is absolutely required for productive packaging of the phage genome (Trottier *et al.* 1996). In addition, similar assays have been used in the T3 system to describe a unique apportioning of the phage maturation enzyme subunits between soluble and head-bound subunits (Morita *et al.* 1995b). These assays are tools that are essential for the elucidation of the packaging pathways of the various phages.

Many morphogenetic precursors of Mu and D3112 have been identified and partially characterized. Studies have described several of the precursor proteins and structures that are responsible for the assembly of these phages (Grundy and Howe, 1985; Giphart-Gassler *et al.* 1981; Grimaud, 1996; DuBow and Bukhari, 1981; Bidnenko *et al.* 1989). However, despite electron microscopic examination of these structures, detailed

biochemical characterization of the precursors has not been carried out. The most recent studies of the morphogenesis of these phages have described the development of the Mu prohead, the packaging of the coliphages by crude extracts of infected cells late in the lytic cycle, and the structures of the D3112 precursors. Bidnenko *et al.* (1989) demonstrated that the head precursors of D3112 were very similar to those of Mu and D108. However, no other morphogenetic studies have been carried out with D3112. The assembly of the Mu prohead was recently examined by Grimaud (Grimaud, 1996). It was reported that, subsequent to the assembly of a prohead, the protein monomers that formed the phage portal are proteolytically cleaved. This proteolytic maturation of a head component has been previously demonstrated in such phages as HK97, P2/P4 and λ (Hendrix and Duda, 1998; Marvik *et al.* 1994a; Bazinet and King, 1985). An *in vitro* packaging system for the transposable coliphages has been developed (Burns *et al.* 1990).

In this paper, we describe the characterization of the morphogenesis of the transposable phages with electron microscopy and *in vitro* packaging techniques. Thin section studies of developing phages allowed us to assess the development of precursor structures. Furthermore, we refined the existing *in vitro* packaging assay with sucrose gradient purified proheads, and also used these gradients to determine the size of the D3112 major capsid protein. The refinement of the *in vitro* packaging assay has allowed us to examine some of the requirements for effective packaging of an exogenous DNA substrate, including ATP, spermidine and host derived histone-like proteins.

Material and Methods

Media, bacterial strains and plasmids

Luria broth (LB) was used for the growth of *E. coli* strains while BHI medium (Difco) was used for *P. aeruginosa*. *E. coli* LF999 (Burns *et al.* 1990), a Mu *cts*62 lysogen of HB101 (*hsdS*20, *recA*13, *ara*-14, *proA*2, *lacY*1, *galK*2, *rpsL*, *xyl*-5, *mtl*-1, *supE*44), was used for the production of phage packaging extracts. HM8305 (F', *pro-lacZ* 8305::Mu *cts*62/Δ *pro-lac*, *his*, *met*, *rpsL*, Mu^R) (DuBow and Bukhari, 1981) was used for thin section studies. Strain PAS429, a D3112 *cts*15 lysogen of the *P. aeruginosa* wild type strain PAO1 was used for the electron microscopic examination of D3112 development and prohead isolation. Plasmid pSZ53 (Tc^R, pSC101::D108*cts*10) (Szatmari *et al.* 1986) was the DNA substrate utilized for *in vitro* packaging.

Production of crude packaging extracts

Extracts were prepared essentially as described by Burns *et al.* (1990). An overnight culture of strain LF999 was diluted 1:50 in LB and grown to an A_{550}=0.5. The temperature of incubation was then shifted to 42°C (T=0) to inactivate the phage repressor and to induce the lytic cycle of the prophage. The phage lytic cycle was arrested prior to lysis of the culture (T=50 min) by rapidly chilling the culture in a dry ice-ethanol bath. The chilled culture was subjected to centrifugation (3300 X g, 4°C, 15 minutes), and the pelleted cells were resuspended in 1.5 ml of resuspension buffer (50 mM HEPES-

NaOH [pH 8.0], 10% [w/v] sucrose) and the mixture frozen at -70°C. The cell suspension was thawed and 0.1 volumes of lysis buffer (6 mM HEPES-NaOH [pH 7.0], 18 mM $MgCl_2$, 30 mM β-mercatoethanol) was added and the mixture incubated on ice for 10 minutes. The cells were sonicated 3 X 10 seconds, with 1 minute intervals between sonication bursts, and the mixtures were kept on ice. The sonicated mixture was subjected to centrifugation (32,500 X g, 4°C, 20 minutes) to pellet the cell debris. The supernatent fluid was removed, aliquoted and stored frozen at -70°C.

Separation of morphogenetic precursors from crude packaging extracts

A 5 ml, 10-30% continuous sucrose gradient (in 50 mM HEPES-NaOH [pH 8.0], 2 mM $MgCl_2$) was prepared with a Radian model 117 gradient maker (Biocomp, New Brunswick, Canada) according to the manufacturers instructions. An aliquot (300 µl) of the crude packaging extract was layered onto the gradient, which was then subjected to centrifugation at 200,000 X g, 4°C, 45 minutes in a Beckman SW50.1 rotor. The gradient was fractionated into eighteen, 300 µl aliquots by bottom puncture. In order to visualize the protein components within each fraction, 10 µl samples from the aliquots were boiled in sodium dodecylsulphate (SDS)-polyacrylamide gel electrophoresis (PAGE) sample buffer and separated on a discontinuous 15% SDS-PAGE matrix by electrophoresis (Laemmli, 1970). The gels were stained with Coomassie Brilliant blue R-250, and washed with a destaining solution (10% [v/v] acetic acid, 40 % [v/v] methanol, 50% deIH_2O) in order to visualize the proteins.

Western Blotting

Crude packaging extracts (Burns *et al.* 1990) were separated on sucrose gradients, the gradients fractionated and samples of the fractions subjected to discontinuous electrophoresis through a 15% SDS-PAGE matrix as described above. After electrophoresis, the gels were washed, for 15 minutes, in transfer buffer (12.5 mM Trizma base, 96 mM glycine, 20% [v/v] methanol, pH 8.2). Immobilon-P membranes (polyvinylidenedifluoride [PVDF]; Millipore) were wetted in pure methanol and subsequently washed in the transfer buffer. Transfer of the electrophoretically separated proteins onto the PVDF membranes was performed in a Model TE50 Transphor appartus (Hoeffer Scientific Instruments) at 200mA for 12-16 hrs. The primary antibody used to perform the western blot was a polyclonal rabbit anti-Mu antibody (DuBow and Bukhari, 1981) (dilution 1 in 250) generated against CsCl purified Mu phage particles. The western blot was completed with a Bio-Rad Immuno-Blot kit as described by the manufacturer.

Mobility shift assays on gradient fractions

Proteins present in the gradient fractions, prepared as described above, were tested for their ability to bind to Mu *pac*-specific DNA with mobility shift assays performed essentially as described in Tolias and DuBow (1989). Ten µl of each individual fraction were incubated with 5000 cpm of either the 206 bp, specific *pac*-containing pE8 *Bsa*H1

fragment (Harel *et al.* 1990) or the 154 bp, non-specific pBR322 *Hin*F1 fragment labelled with [α-^{32}P]ATP and [α-^{32}P]CTP (3000 mCi/ml, Amersham Inc), respectively. The reactions also contained 2 µl of 10X binding buffer (20 mM Tris-HCl, pH 7.5, 100 µg/ml bovine serum albumin, 0.02% [v/v] Nonidet P-40) and were made up to a final reaction volume of 20 µl with deH_2O. The reactions were incubated (10 minutes, 37°C) and then separated by PAGE, on an 8% gel in 1xTBE (90mM Tris-borate, 2mM EDTA; pH 8.0), until the unbound probe had migrated to the bottom of the gel. The gels were then dried and exposed to Kodak XAR-5 X-ray film at -70°C under Dupont Cronex intensifying screens prior to development.

Refined *in vitro* packaging assay

Crude packaging extracts were prepared, separated on 10-30% continuous sucrose gradients and fractionated as previously described. Fractions detected to have affinity for the Mu *pac* site (by mobility shift assay) were pooled and used for *in vitro* packaging reactions, essentially as previously described (Burns *et al.* 1990). Twenty five µl of the pooled fractions were combined with spermidine (6 mM), ATP (1 mM), 1.25 µg of plasmid pSZ53, and packaging buffer (13mM HEPES-NaOH [pH 7.5], 9 mM $MgCl_2$, 15 mM β-mercaptoethanol, 2% (w/v) sucrose), and incubated at 37°C for 1 hour. In addition to the basic reaction listed above, other reactions were performed with one of the following: soluble proteins present at the top of the sucrose gradients, purified Polyoma large T antigen (kindly provided by Yu Cai Peng and Dr. N. Acheson, McGill University), bovine serum albumin (BSA), HU, and IHF. Purified IHF and HU were the

kind gift of Dr. G. Chaconas, University of Western Ontario, London Ontario. After incubation, the reaction volume was increased from 50 to 100 µl by the addition of 40 µl of delH_2O and 10 µl of 10X DNaseI buffer (500 mM Tris-HCl [pH 7.5], 100 mM $MgCl_2$, 20 mM $CaCl_2$, 500 µg/ml BSA). DnaseI (75 units), *Eco*RV (5 units), and *Bgl*II (5 units) were then added to the mixture and the reactions incubated at 37°C for 1 hour. The reactions were stopped by the addition of EDTA $[250mM]_{final}$ and SDS $[0.5\%]_{final}$. The samples were extracted twice with equal volumes of phenol:choloroform:isoamyl alcohol (25:24:1) at 65°C. The extracted samples were brought to a volume of 150 µl with "crush and soak" buffer (Maxam and Gilbert, 1980). The mixture was extracted once with chloroform isoamyl alcohol (24:1) and then precipitated with ice-cold absolute ethanol. The precipitated DNA was subjected to centrifugation (16000 X g, 4°C, 30 minutes) then washed with 70% ethanol. The pellets were dried in a Savant speed vac and resuspended in 25 µl 1X TE. DNase I resistant DNA was detected by PCR with 15 pmols of each of the PCR primers nerR (5'-GAACTAACCGCCAAGATTGG-3') and nerL (5'-GACGGCCAAATATCTTGCGG-3'), complementary to the *ner* gene of phage D108. These were combined with 400 µM dNTPs, 25 µl of the extracted DNA plus 1 unit of *Taq* polymerase in PCR buffer (10 mM Tris-HCl [pH 9.0], 50 mM KCl, 0.1% Triton X-100). The products of the PCR reaction were separated by PAGE on an 8% polyacrylamide gel in 1xTBE (Sambrook *et al.* 1989). After electrophoresis, the gels were stained with ethidium bromide, destained and visualized by UV illumination. Relative intensity of the stained DNA bands from scanned images was assessed using ImageQuant (Molecular Dynamics).

Electron microscopy of growing phage *in vivo* and isolated phage structures

Overnight cultures of *E. coli* HM8305 and *P. aeruginosa* PAS429 were grown in LB and BHI, respectively, and then diluted 1:20 in fresh medium. The diluted cultures were incubated at 32°C until the A_{550} reached 0.4. At this point, the incubation temperature was increased to 42°C to induce the lytic cycle of the prophages, and the incubation continued. One ml aliquots of the lysogens, to be examined in thin section, were removed at times=0, 30, and 45 minutes post induction. The samples were subjected to centrifugation (16000 X g, 4°C, 2.5 min) and the cell pellets resuspended and incubated in fixative solution (2.5% (w/v) glutaraldehyde, 0.1 M NaCacodylate) for 1 hour at 4°C. The resuspended and fixed lysogens were diluted 1:2 with 1.5% agarose (in 0.1M NaCacodylate) and the mixture was allowed to gel. The solid samples were washed (3X15 minutes) with 0.1M NaCacodylate. The washed samples were osmicated (1% (w/v) OsO_4 in 0.1M NaCacodylate) for 2 hours at 4°C. The samples were then subjected to an alcohol dehydration series (1X10 minutes in 30, 50, 70, 90, 95% EtOH, 3X15 minutes in 100% EtOH) followed by a 10 minute treatment with propylene oxide. Subsequent to the dehydration series, the samples were infiltrated for 2hours with a 1:1 solution of propylene oxide and Epon embedding medium. The samples were then treated, overnight, with a 1:2 solution of propylene oxide and Epon. The diluted Epon was then replaced with pure Epon and the samples incubated for 2 hours. The samples were then transferred to Beem capules (SPI supplies, Toronto, Canada), labelled, covered with fresh Epon and allowed to harden. Sections of 700-800 Å were cut from the embedded samples with a Reichert-Jung Ultra Cut E microtome equipped with a diamond

knife, and applied to copper grids. The sections were stained with uranyl acetate (Watson, 1958) and lead citrate (Venable and Coggeshall, 1965) and examined in a Philips 410 transmission electron microscope.

Samples of isolated phage precursors from sucrose gradients were examined by negative staining rather than thin section. The samples were separated from crude extracts on 10-30% sucrose gradients as described. Gradient fractions were dialyzed against 2% ammonium acetate and then applied to formvar coated copper grids. Ten µl of 2% phosphotungstic acid was added to the sample droplets on the grids. The mixtures were allowed to incubate for 30 seconds, after which the fluid was removed from the sample by blotting the edge of the grid. The grids were allowed to dry completely, after which they were examined in a Philips 410 transmission microscope.

Results

In vivo development of transposable bacteriophages

Thin sections of thermoinduced lysogens of either Mucts62 or D3112*cts*15 were examined at specific times during lytic development of the phages. Electron micrographs of thin-sectioned bacteria grown to mid-log phase immediately upon induction of the prophage are shown in Figure 1. The *P. aeruginosa* lysogen PAS429 (Figure 1A, a) has an intact membrane and displays no obvious perturbations within the cytoplasm. The Mu lysogen (Figure 1A, b) also shows no gross disruptions; the cell appears to be dividing and the nucleoid is distinct. By thirty minutes post-induction (Figure 1B, a & b) the inner cell membrane begins to separate from the peptidoglycan and outer membrane, though no morphogenetic precursors could be detected.

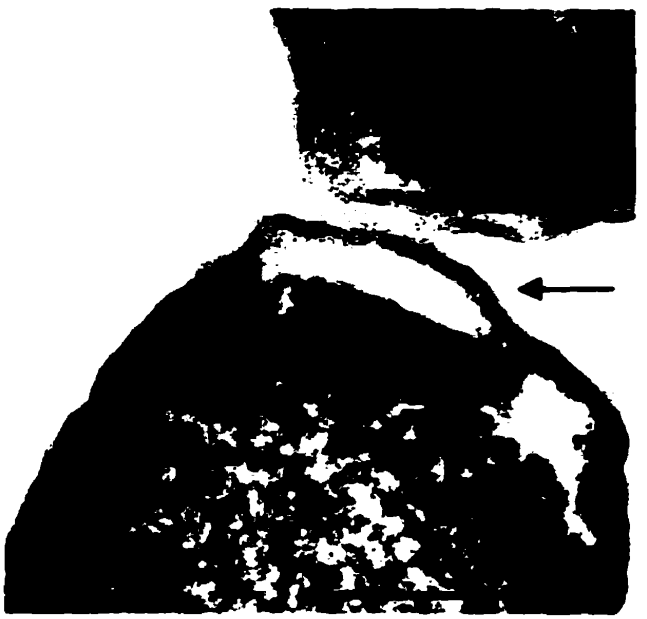
By 45 minutes post-induction, proheads were prominent. Portions of two proheads are visible in the section of *P. aeruginosa* (Figure 1C, a). In addition, electron dense granules were present in a number of the *P. aeruginosa* cells viewed in thin section. The visible proheads in the *E. coli* section contain an intense, electron dense columnar core (Figure 1C, b). Small neck structures have been shown to be present in isolated proheads (Grundy and Howe, 1985). However, such an electron dense structure has not been reported. The Mu prohead, viewed in thin section, is approximately 54 nm in diameter.

Figure 1 Electronmicroscopic examination of thin sections from lysogens. (A) Thin section of lysogenic bacteria prior to induction. *P. aeruginosa* PAS429 (a) and *E. coli* HM8305 (b). Bar=500nm. (B) Thin section of induced lysogens at 30 minutes subsequent to thermoinduction of the prophages from PAS429 (a) and HM8305 (b) are shown. The arrows show regions of cell wall disruption. Bar=250nm. (C) Thin section of thermoinduced lysogens at 45 minutes post-induction. The D3112 structures (a) are marked by the arrows. Mu prohead (b) and a dense columnar structure present within the prohead. Bar= 250nm.

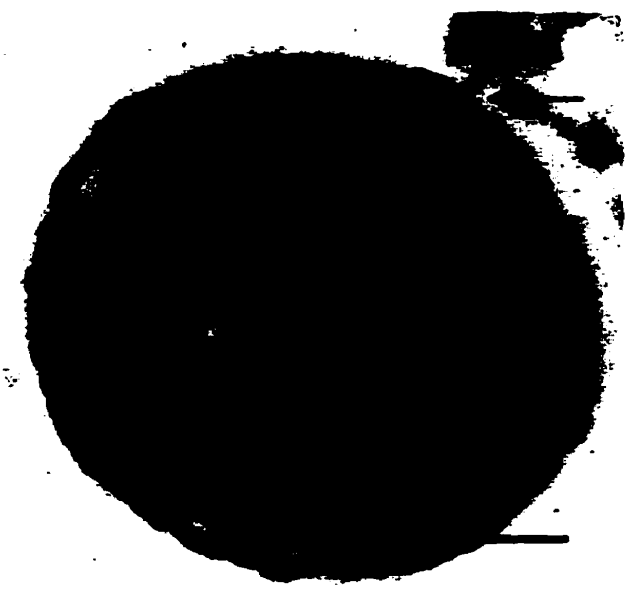
A a

b

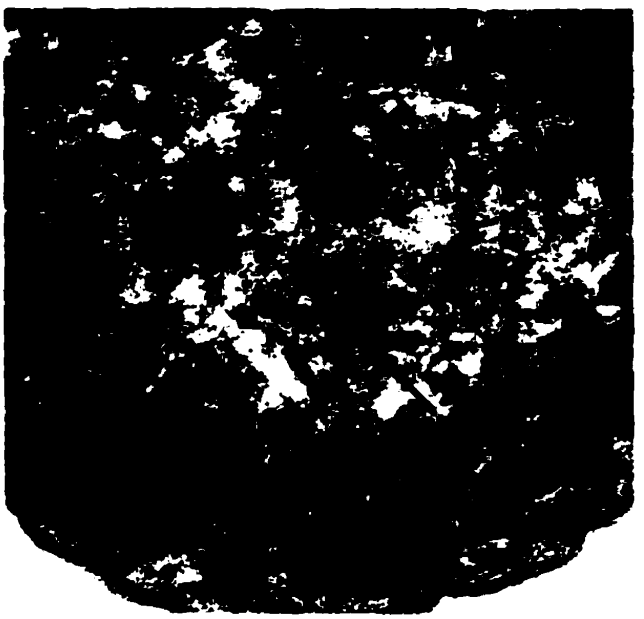
B a



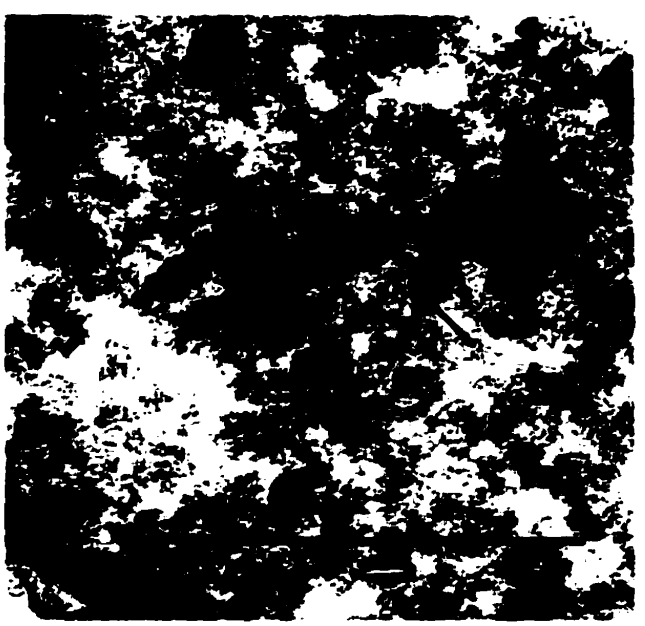
b



C a



b

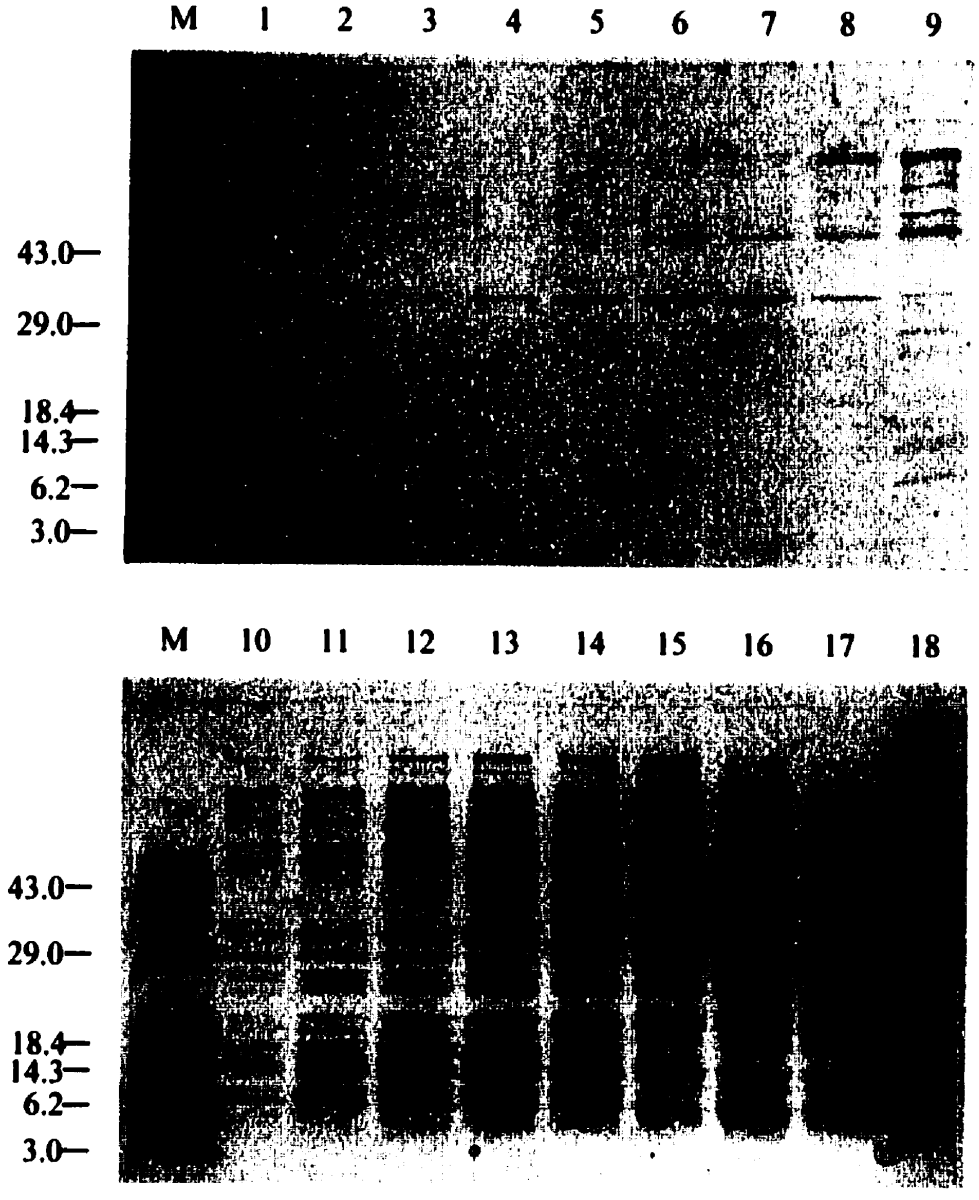


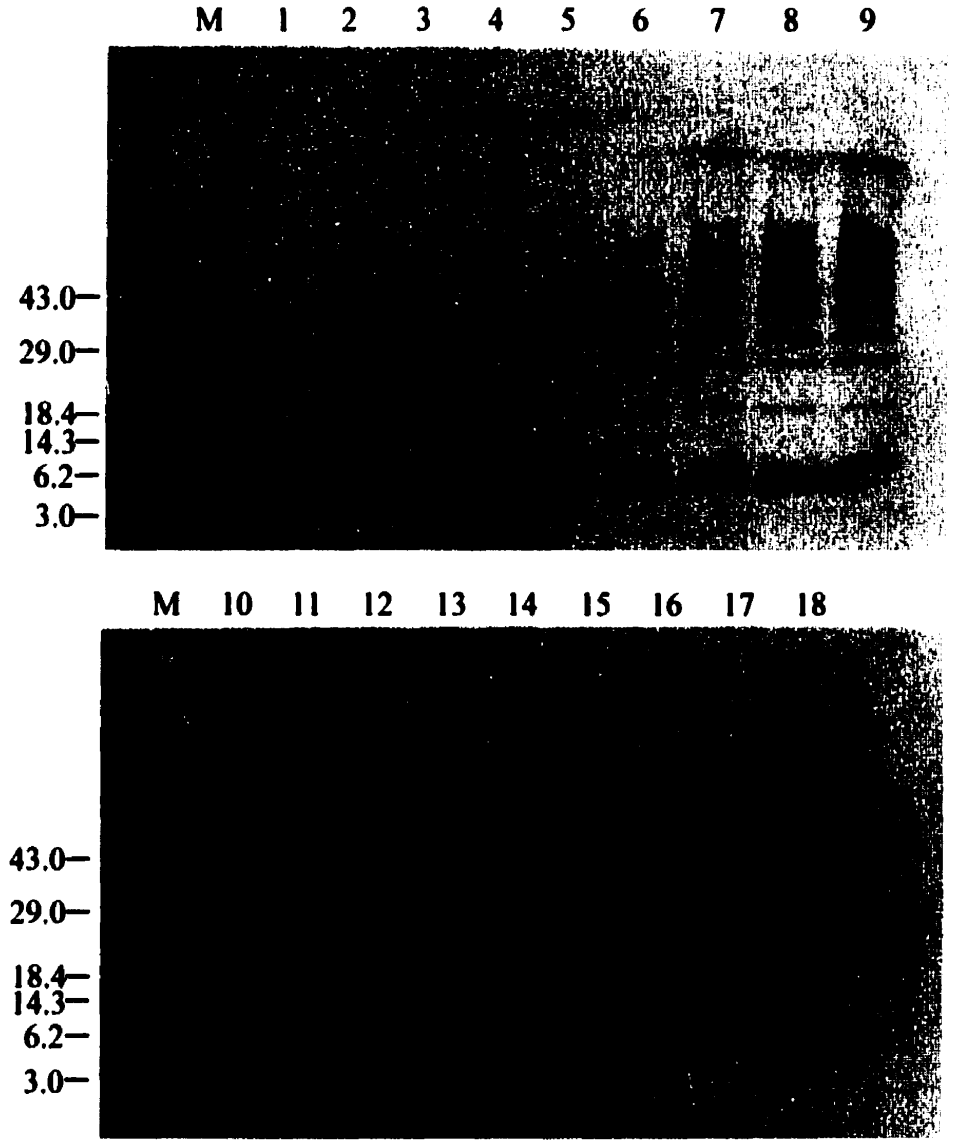
Isolation of phage products by sucrose gradient centrifugation

It was previously reported (Burns *et al.* 1990) that crude lysates produced from induced Mu lysogens could be utilized for *in vitro* packaging of exogenously supplied DNA substrates. Figure 2 shows the proteins from fractions collected from a 10-30% continuous sucrose gradient subsequent to centrifugation of crude lysates, prepared from thermo-induced Mu lysogens late in the lytic cycle, competent for *in vitro* Mu DNA maturation and encapsidation. The gradient was separated into fractions, numbered 1-18, from bottom to top. The bulk of the proteins are soluble and of low sedimentation coefficients, and consequently do not migrate far into the gradient (Figure 2A lane 18). Lanes 5 to 7 contain a single, prominent protein band whose molecular mass of 33kDa corresponds to that of the major capsid protein of phage Mu (Howe, 1987). Within these fractions, there were no other strongly staining protein bands with molecular masses equivalent to those reported for the morphogenetic proteins of Mu. While the 33 kDa band is predominant within fractions 5-7, varying amounts are also present in other fractions higher in the gradient. The proteins present in the various gradient fractions were studied by western blotting using antisera prepared against purified Mu phage particles (DuBow and Bukhari, 1981).

Immunoblots revealed a prominent 33 kDa band present in fractions 5-7 (Figure 2B). In addition to the major head protein, two other protein bands, of approximately 40 and 45 kDa, were strongly stained. A band corresponding to a protein of approximately 50 kDa is also present in fractions 5-7 (Figure 2B).

Figure 2. (A) Protein profile of sucrose gradients fractions. The proteins present in a packaging lysate were separated on a sucrose gradient as described in Materials and Methods. The protein content of the individual fractions was assessed by separation using SDS-PAGE. The fractions are numbered 1 through 18, representing the samples from the bottom to the top of the gradient. (B) Phage proteins separated by gradient fractionation and detected by western blotting. The proteins separated by SDS PAGE were transferred to membranes and probed with Mu antisera as described in Materials and Methods. The samples numbered 1-18 represent the fractions from the bottom to the top of the gradient.


M 1 2 3 4 5 6 7 8 9
43.0
29.0
18.4
14.3
6.2
3.0
M 10 11 12 13 14 15 16 17 18
43.0
29.0
18.4
14.3
6.2
3.0


M 1 2 3 4 5 6 7 8 9
43.0
29.0
18.4
14.3
6.2
3.0
M 10 11 12 13 14 15 16 17 18
43.0
29.0
18.4
14.3
6.2
3.0

Other minor protein bands in fractions 5-7, detected by western blotting, have molecular weights that corresponded to values reported for some of the Mu virion proteins. A band of approximately 64 kDa is present. A faint band, migrating with an approximate molecular weight of 55 kDa, is present within fractions 6 and 7 but undetectable in fraction 5.

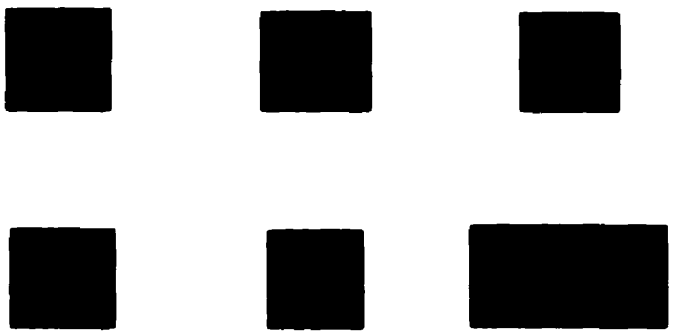
The composition of prohead gradient fractions

In order to determine if gradient fractions 5-7 contained Mu morphogenetic precursors, their constituents were examined by electron microscopy. In addition, D3112 lysates were also prepared and separated on continuous 10-30% sucrose gradients as described for Mu, and Fractions 5-7 of the D3112 gradients were also prepared for electron microscopy.

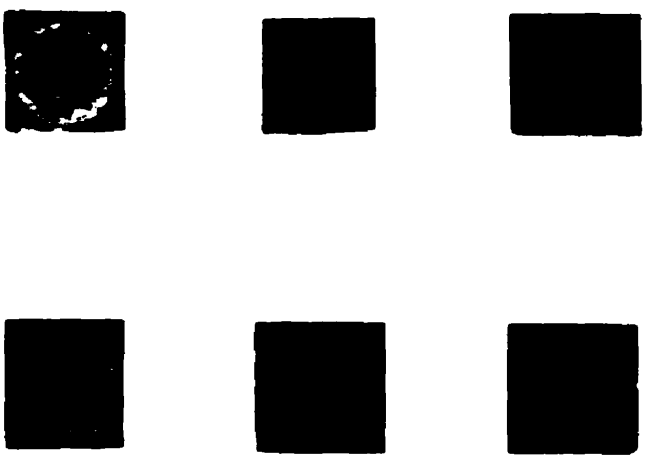
Examination of the pooled fractions revealed the presence of prohead structures for both Mu and D3112. The Mu proheads appeared to be of two types (Figure 3). One subset of proheads was smaller (diameter=51±1.6nm) with thick walls, while the second subset was larger (diameter=54.2±2.7nm) and had slightly thinner walls. Both structures exhibited a thickening at a unique position in their wall. The thickening was seen as a round structure when the prohead was suitably oriented within the field. Gradient fractions 5-7 would thus seem to contain unexpanded and expanded Mu phage proheads.

Figure 3 Electron microscopic images of phage structure. Gradient fractions that contain phage proheads, as determined by molecular mass and reactivity to anti-Mu antisera, were examined by electron microscopy. In addition to Mu structures, D3112 lysates were prepared and examined as per Materials and Methods. Representative structures from the micrographs were utilized to construct the gallery. Bar=100nm

D3112



Mu



Unlike Mu, the D3112 proheads (Figure 3) were of a single diameter (55.6±3.1nm). There is a distinct thickening within the D3112 prohead shell that is similar to that of Mu. A number of proheads were also found to interact with each other through the portal ring (Figure 3). The SDS gel profiles of the D3112 gradient fractions demonstrated that the predominant protein within these fractions is the major capsid protein. The protein components of the pooled fractions of the D3112 lysate also reveal a single predominant band (Figure.4), with an apparent molecular weight of 37 kDa.

Pac-containing DNA binding activity of prohead-containing fractions

Fractions 5-7 were studied to determine if they are able to specifically bind to the Mu *pac* site, a property of phage maturation enzymes, using a gel mobility shift assay. Autoradiograms of the reaction products from the mobility shift assay are shown in Figure 5. Protein-DNA complexes are visible when the gradient fractions were incubated with a *pac*-specific probe (Figure 5A). The strongest shifts occur with fractions 4 to 6. Protein-DNA complexes are present in the fractions above and below fractions 4 to 6, but the bands have a diminished intensity. Reactions performed with a non-specific probe are shown in the lower panel (B) of Figure 5. No significant protein-DNA complexes are formed between the constituents of the gradient fractions and non-*pac* site containing DNA.

Figure 4. Protein content of D3112 prohead fractions. The gradient fractions shown to contain D3112 prohead structure were separated on SDS-15%-PAGE and stained according to Materials and Methods.

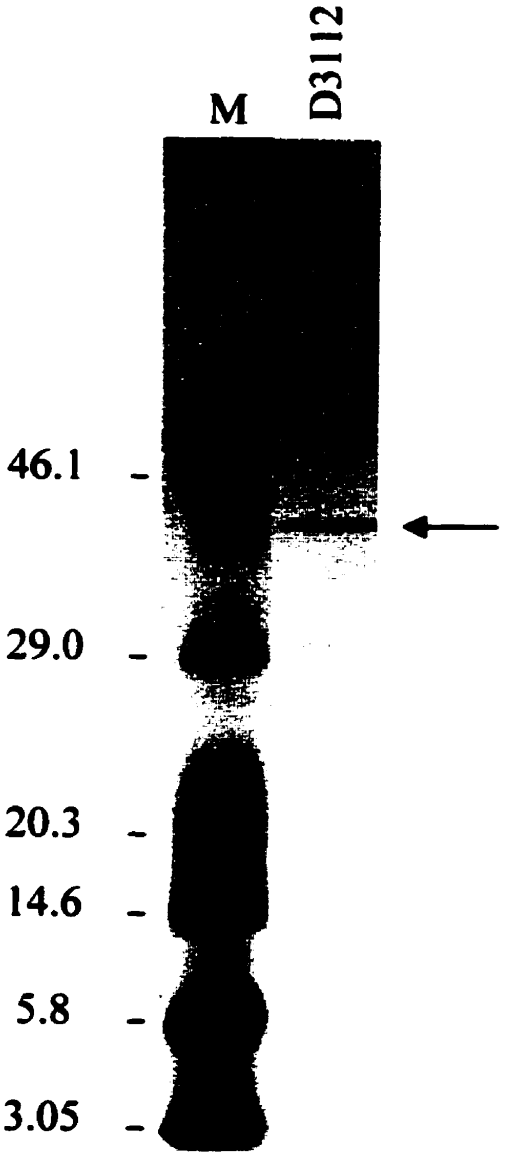
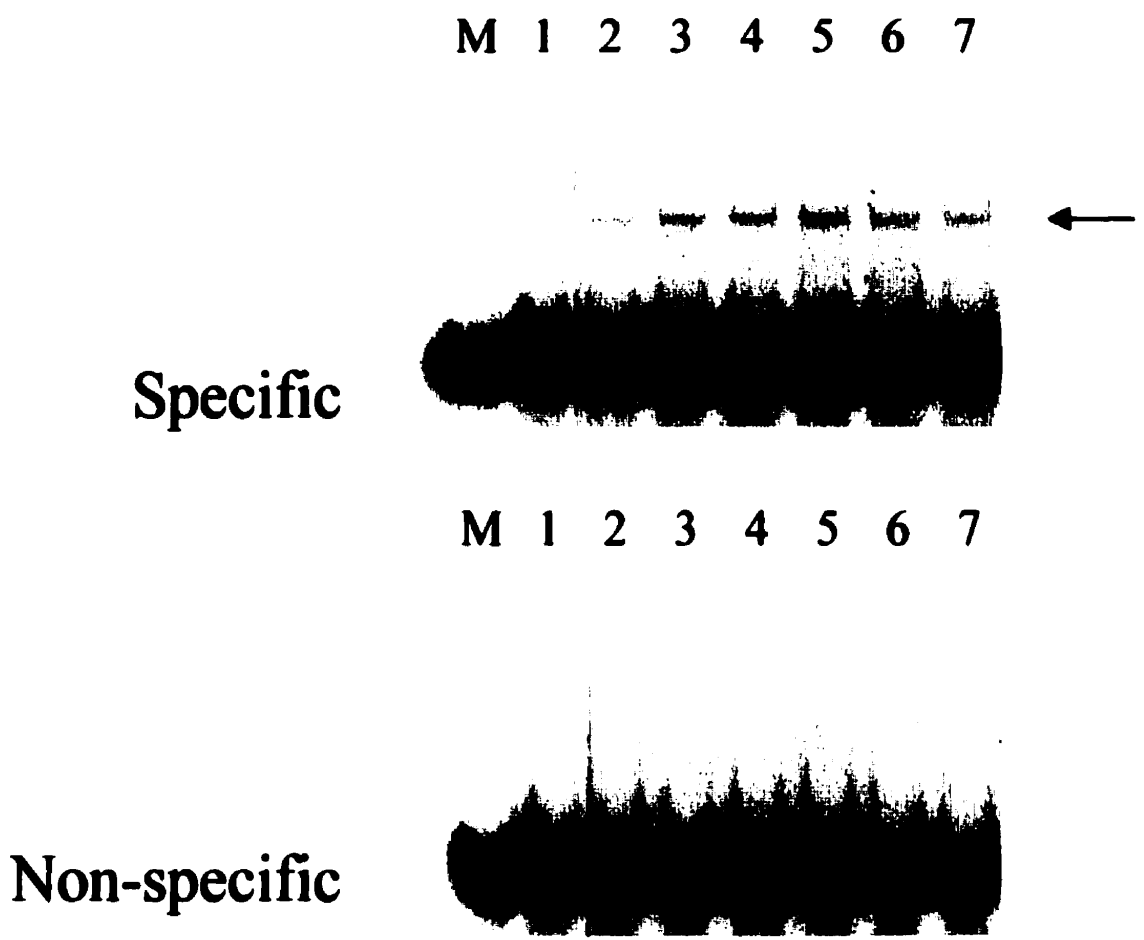

M
D3112
46.1
29.0
20.3
14.6
5.8
3.05

Figure 5. Mobility Shift Assay of sucrose gradient fractions. Lane M contains the free probe while lanes 1-7 represent the corresponding gradient fractions in the presence of a *pac*-containing (A) or control (B) fragment

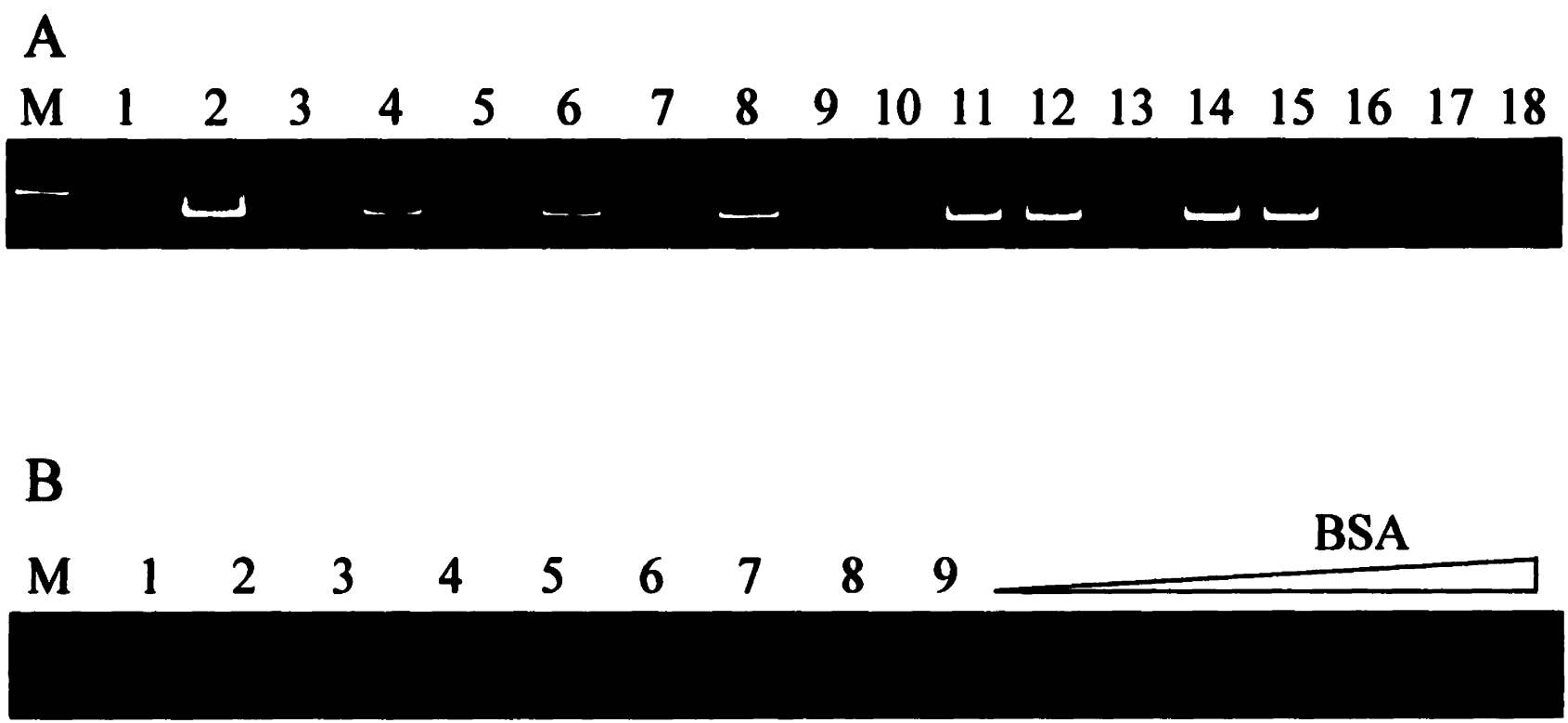

M 1 2 3 4 5 6 7
Specific
M 1 2 3 4 5 6 7
Non-specific

In vitro packaging

The ability of the prohead containing fractions to specifically bind to Mu *pac* DNA is suggestive of the presence of the phage DNA maturation enzyme. *In vitro* DNA encapsidation assays (Burns *et al.* 1990) were performed with aliquots of the pooled fractions which exhibited the strongest DNA binding ability to assess the capacity of these fractions to package exogenously supplied D108 DNA as Mu and D108 can encapsidate each other's genomes *in vivo* and *in vitro*. The absence of purified tail components in these fractions necessitated that the packaging of the DNA substrate be monitored in a manner other than by plaque formation. The assay chosen utilizes PCR to detect the D108 *ner* gene from a pSC101::D108*cts*10 plasmid substrate, which can be packaged by Mu morphogenetic proteins, after digestion of the reaction mixture with nucleases to degrade non-encapsidated DNA. Protection from nuclease digestion is indicative of packaging of the DNA into the purified proheads.

The packaging reactions were carried out under various conditions to assess the requirements for DNA encapsidation (Figure. 6). The DNA substrate was found to be protected from degradation after incubation with the crude lysate; a mixture of fractions 5-7 plus an aliquot of material from the top of the gradient; and by the material in fractions 5-7 alone (Figure 6A lanes 4, 6, 8). In the absence of an exogenously added DNA substrate, no PCR product is generated (Figure. 6A lanes 5, 7, 9). *E. coli* histone-like proteins IHF and HU were added to the packaging reactions to assess their effects on the assay, as supercoiled DNA is the optimal substrate for packaging. An approximately 40% increase in the amount of PCR product in reactions containing the histone-like proteins was observed based on intensity of the stained PCR products. (Figure 6A, lanes

Figure 6. Characterization of refined *in vitro* packaging reactions. In addition to the basic reaction, the assay was performed in the presence or absence of various potential activators or inhibitors as indicated. (A) lane M 100 bp ladder, Lane 1 no DNA, Lane 2 pSZ53 only, Lane 3 pSZ53 digested, Lane 4 packaging lysate, Lane 5 packaging lysate without pSZ53, Lane 6 prohead fractions + fraction 18, Lane 7 prohead fractions +fraction 18 without pSZ53, Lane 8 prohead fraction, Lane 9 prohead fraction without pSZ53, Lane 10 no proheads, HU (100nM), Lane 11 prohead fraction + HU (50nM), Lane 12 prohead fraction + HU (100nM), Lane 13 no proheads, IHF (10 nM), Lane 14 prohead fraction + IHF (5nM), Lane 15 prohead fraction + IHF (10nM), Lane 16 prohead fraction without ATP, Lane 17 prohead fraction + HU (50nM) without ATP, Lane 18 prohead fraction + IHF (5nM) without ATP. (B) Lane M 100bp ladder, Lane 1 no DNA, Lane 2 pSZ53, Lane 3 pSZ53 digested, Lane 4 lysate, Lane 5 prohead fraction, Lane 6 prohead fraction without spermidine, Lane 7 prohead fraction + Rnase A, Lane 8 prohead fraction + LTAg (5nM), Lane 9 prohead fraction + LTAg (10 nM), lanes 10-15 prohead fraction + BSA (50, 100, 200, 300, 400, 500 nM)


A
M 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
B
M 1 2 3 4 5 6 7 8 9
BSA

11, 12, 14, 15). The packaging reactions were performed with a purified, non-prokaryotic DNA binding protein, Polyoma large T antigen, as well bovine serum albumin (BSA) as controls. Neither large T antigen (Figure 6B, lanes 8, 9) nor BSA (Figure 6B, lanes 10-15) was able to increase the amount of DNaseI resistant material available for PCR amplification. In the absence of ATP, the quantity of PCR product produced in the reactions which contain the isolated proheads and those supplemented with IHF and HU is decreased to background levels (Figure 6A, lanes 16-18).

Discussion

In this study, we have begun to characterize the morphogenetic steps of DNA encapsidation of the transposable phages Mu and D3112. Growth *in vivo*, as well as *in vitro* maturation and encapsidation of the phage genome, were undertaken to add insights into the mechanism(s) by which the phages are able to form mature capsids.

Thin section electron micrographs of induced lysogens allowed us to identify precursor phage structures. The host cells contain no structures resembling phage precursors at the time of induction of the replicative cycle (Figure 1A, a & b). Moreover, the cells themselves show no gross deformations and appear to be growing normally. The bacterial cell membrane and nucleoid are intact (Valkenburg *et al.* 1985), and the cells are dividing (Figure 1A). During the lytic cycle, phage precursors are produced, the genome is packaged, and the cell is lysed. In thin section, the host cells show signs of phage development at 30 minutes post-induction. No phage-specific structures are visible, but the host cell begins to break down. The nucleoid is no longer as distinct and cell wall integrity begins to be compromised.

At 45 minutes post-infection, phage structures are visible within the host cells. Typically, phages build their head and tail structures by independent assembly pathways (Casjens and Hendrix, 1988). Head structures are visible in both cells. Unlike isolated proheads, which have diameters similar to the mature head, the diameter of the head structures in thin sections may appear smaller than that of the mature phage due to the plane of the section. The proheads visible in D3112-infected *P. aeruginosa* (Figure 1C, a) appear smaller than expected for these dsDNA phages, however they do have a

characteristic spherical shape. The electron dense structure above the proheads does not resemble a phage product and is most likely a phosphate granule (Takade *et al.* 1991). Similar to the Mu proheads, the D3112 structures are not associated with the inner membrane and therefore are likely to be assembled free in the cytoplasm. The Mu prohead shown in Figure 1C (b) has a diameter similar to that of the mature phage and most likely has been sectioned equatorially. In addition to the characteristic shell, an electron dense structure is visible as a columnar core within the prohead. The structure of the phage DNA within the head has yet to be defined. However, a model based on the spiral packaging of phage DNA has been proposed (Black *et al.* 1985). The visible electron dense core (Figure 1C, b) may therefore represent partially packaged DNA spooling around the portal axis of the prohead. Despite the apparent absence of tail structures at this point in the phage lytic cycle, proheads have been assembled and are beginning to package DNA. Generally, these events are initiated randomly and are part of a process that spans a relatively long period during the late phase of the phage lytic cycle (Kellenberger, 1990).

It has been previously shown that a crude lysate prepared from an induced Mucts62 lysogen is able to package an exogenously supplied D108 DNA substrate (Burns *et al.* 1990). The Mu lysate used in those studies was prepared 50 minutes post-induction of a 60 minute lytic cycle. This period is within the phase of the lytic cycle in which head precursors are available for packaging, as is evident in the sectioned images (Figure 1C, b). Preparation of a defined packaging reaction requires that a purified DNA substrate, purified proheads and the maturation enzyme (*pac*ase) must be available for the reaction. Sucrose gradient fractionation was used to separate the proheads from the bulk of the proteins within the packaging lysate. Analysis of the protein content of the fractions, by

SDS-PAGE (Figure 2A), indicates that three fractions, 5 to 7 (F5-7), in the dense bottom portion of the gradient contain predominantly one protein, with a molecular mass identical to that of the major capsid protein of Mu. These fractions appear to be enriched in Mu proheads and serve as a source of proheads for a defined packaging assay. Western blots of the gradient fractions (Figure 2B) reveal that there are, in addition to the major capsid protein, several other proteins that are recognized by anti-Mu phage particle antisera.

Previous studies have defined a number of "late" Mu proteins that are either structural or involved in the assembly of the phage (Grimaud, 1996; Giphart-Gassler *et al.* 1981; Grundy and Howe, 1985; DuBow and Bukhari, 1981). The 50 kDa portal protein (Grimaud, 1996) is an integral component of the prohead and appears to be present in fractions 5-7. In addition to the 33kDa major capsid protein, two proteins of 40.3 and 45.2 kDa also strongly react with anti-Mu phage particle antiserum in fractions 5 to 7 (Figure 2B). A number of tail-associated genes (*R*, *W* and *P*) have been reported to encode proteins with molecular weights in the range of 40-50 kDa (Giphart-Gassler *et al.* 1981). However, only two phage proteins in the 40-50 kDa range have been shown to be present in the mature virion (DuBow and Bukhari, 1981). The major tail protein was reported to have a molecular mass of either 52 or 55kDa (Grimaud, 1996). Proteins with these molecular weights are not visible in the Coomassie stained gel. However, a very faintly staining band of 55 kDa is visible on the western blot. This band may represent the *F* gene product, a head related protein with a molecular mass of 54 kDa. Electron microscopy showed that fractions 5-7 of the gradient contained a large number of prohead structures (Figure 3), but tails were not visibly evident.

During the process of packaging and maturation, the prohead of most phages undergoes a number of structural changes (Casjens and Hendrix, 1988). The changes associated with packaging and maturation are not simply the gross structural transitions that occur when the prohead assumes the form of a mature phage head. A T4 structure can be isolated which is larger than a naïve prohead (Jardine and Coombs, 1998). This structure, called an empty large particle, is unable to package DNA and is likely to have been formed as a result of an aborted packaging event. In the HK97 system, Prohead II is slightly larger than Prohead I. The former is formed by a proteolytic cleavage event that is required to produce the mature form of the major capsid protein (Hendrix and Duda, 1998). Thus, the two forms of the Mu prohead (Figure 3) may represent essential intermediates in the assembly pathway or two distinct pools of morphogenetic precursors that have different packaging potential. Unfortunately, these structures could not be separated by gradient centrifugation and therefore their ability to package DNA could not be tested individually. Interestingly, D3112 proheads prepared in a similar manner all appeared to be of a uniform size (Figure 3) and consisted of a predominant major capsid protein of 37kDa (Figure 4).

Phage proheads are an essential structural precursor for the packaging of the replicated genome. The packaging of the DNA, however, also requires a phage encoded maturation enzyme (terminase) (Black, 1988). The Mu maturation enzyme (*pac*ase) is thought to be encoded by genes *D* and *E*. In most of the systems that have been studied, the maturation enzyme and the proheads interact after the formation of a specific enzyme-DNA complex (Black, 1988). In at least two of the dsDNA phages, this packaging mechanism has been modified and includes molecules whose functions are normally

associated with free maturation enzymes on the prohead. Phages ϕ29 and T3 use RNA and the large terminase subunit, respectively, to apportion the packaging functions between the prohead and another distinct protein (Guo *et al.* 1987b; Morita *et al.* 1995a). We have shown that the prohead-containing sucrose gradient fractions from Mu lysates are able to specifically bind to the Mu packaging signal (Figure 5). It appears therefore that protein(s) capable of specifically interacting with the Mu packaging signal co-migrate with the phage proheads in the sucrose gradient. The SDS-PAGE protein profiles of fractions 5-7 do not show any other proteins, in significant amounts, other than the major capsid protein (Figure 2A). Moreover, the packaging enzyme of dsDNA phages has not been found to be a component of the mature virion, and antisera that are produced against the mature phage are found not to contain antibodies directed against the enzyme subunits. The western blots, therefore, are not able to reveal the presence or absence of these two proteins.

When used in a modified Mu packaging assay, the Mu prohead fractions 5-7 are able to package a D108 substrate, as monitored by protection from DNaseI degradation and PCR (Figure 6). The dense fractions, when included in the standard reaction, were only able to package the DNA substrate in the presence of ATP. The energy source for packaging used by phage maturation enzymes has previously been shown to be ATP; for every molecule of ATP hydrolysed, two bases are packaged (Black, 1988). When ATP is omitted from the Mu packaging reaction, there is a complete loss of packaging, which implies that the DNA protection monitored in the assay is ATP-dependent (Figure 6A, lanes16-18). Additionally, the omission of spermidine, a polyanion used to neutralize the charge of the nucleic acid substrate, resulted in a loss of packaging (Figure 6B, lane 6).

The packaged phage DNA is confined to a small volume, resulting in a pseudo-crystalline state and therefore, the strong positive charge present on the sugar-phosphate backbone of the DNA must be neutralized. The inability of the reaction precursors to package the DNA substrate in the absence of this molecule reflects the need for a cellular constituent, normally present *in vivo*, to aid in the condensation of the DNA. While ATP and spermidine are essential for the packaging of the phage genome, Mu does not appear to require an RNA molecule to package its genome (Figure 6B lane 7).

The addition of HU or IHF to a packaging reaction with prohead fractions led to an increase in the amount of PCR product after DnaseI treatment of the packaging reactions (Figure 6A, lanes 11, 12, 14, 15). When proteins unrelated to phage packaging were included in the reaction, no specific increase in DNA protection was observed. A stimulatory effect by IHF in a crude Mu packaging reaction has been previously reported (Burns *et al.* 1990). Moreover, both IHF and HU are important for the proper assembly of nucleoprotein complexes during the lytic growth of Mu (Pato, 1989; Higgins *et al.* 1989; Lavoie and Chaconas, 1994; Craig, 1996).

The maturation of the Mu genome prior to packaging requires that the phage DNA is specifically recognized and then cleaved so that it may be encapsidated. It is known that the genome is recognized by the packaging machinery at the *pac* site, located near the extreme left end of the phage genome (Harel *et al.* 1990). Following the recognition of the genome, a cleavage event occurs which results in the retention of approximately 50-150 basepairs of host DNA upstream of the left end of the integrated phage chromosome (George and Bukhari, 1981). Thus, subsequent to *pac*ase binding to its recognition site, a conformational change in the nucleoprotein complex must occur. This may position the

*pac*ase in such a manner that it cleaves the host DNA adjacent to the phage genome. While the phage enzyme itself may induce such a conformation, the inclusion of host histone-like proteins, which are able to strongly bend DNA, may facilitate the positioning of the *pac*ase. The presence of histone-like proteins at or near the Mu *pac* site has been previously demonstrated (Lavoie and Chaconas, 1994; Lavoie and Chaconas, 1993; Craig, 1996). Both IHF and HU play a role in assembling the Mu transposasome during replication and they are known to have binding sites at which protein-DNA complexes can assemble. The Mu transposasome contains Mu A, B and *E. coli* HU proteins. After transposition, the complex is disassembled through the proteolytic removal of transposase (Levchenko *et al.* 1995). The removal of this protein is likely essential for the freeing of the *pac* site, which can apparently be blocked by transposase (Burns *et al.* 1990). Moreover, the structure induced by IHF during the assembly of the transposasome may be maintained subsequent to the removal of the transposase and be available for the maturation complex. Thus, binding of these histone-like proteins may also facilitate the assembly of an ordered nucleoprotein complex that is necessary for the productive cleavage of the Mu genome.

The partially purified packaging assay described here has enabled us to demonstrate the active packaging potential of gradient-isolated proheads. Purification of both the phage *pac*ase and tail structures, however, is necessary for further refinement of the *in vitro* packaging assay. Finally, the assembly of the packaging complex itself can be monitored in order to define the switch that controls the temporal regulation of these unique replicative and packaging functions.

Chapter Four

Conclusions

Conclusions

Bacteriophage morphogenesis is a complex process that results in the assembly of a particle defined by strict geometrical dimensions. The assembly process utilizes multi-step pathways to form these structures. In addition to building structural precursors, the replicated phage genome must be recognized and matured by a phage encoded enzyme complex. Finally, the complete phage virion is released from the host as a result of a regulated lytic process. Bacteriophages Mu, D108 and D3112 are three related phages that replicate their genomes using a well-characterized transposition mechanism. The latter part of their lifecycle, during which the genome is matured and packaged into morphogenetic precursors, has not been well characterized.

The lytic and packaging functions of bacteriophage Mu have been localized, by amber mutant analysis, to the region of the Mu genome immediately downstream of the P_{lys} late promoter (Grundy and Howe, 1985). The lytic mechanism, by which Mu lyses its host, has yet to be characterized, however, it was reported that no holin-like function could be detected during phage infection of *E.coli* (Young, 1992). Cloning and sequencing of approximately 4.5 kb of the Mu genome downstream of P_{lys} revealed that seven potential ORFs were encoded in this region. Characterization of the ORFs, using computer-based homology algorithms, indicated that the first ORF, immediately downstream of P_{lys} had homology to previously characterized endolytic enzymes. Sequencing of Mu *lys* amber mutants confirmed that this first ORF encodes the phage endolysin. The remaining ORFs were found to be homologous to uncharacterized ORFs from a cryptic Mu-like phage of *H. influenzae*. The maturation enzymes of a number of

phages have been identified previously and have been shown to contain similar protein motifs and are encoded by genes that are arranged in a similar manner (Black, 1989). The genes encoding the small and large enzyme subunits are encoded by adjacent ORFs. The small subunit contains a number of α-helices and is believed to bind DNA using an H-T-H motif. The large subunit is thought contain the catalytic domain of the heteroligomer enzyme. Both subunits bind ATP. Two of the identified Mu ORFs have a number of characteristics identical to the maturation enzymes of other phages. ORFs 6 and 7 have the potential to encode proteins with sizes similar to previously described maturation enzyme subunits. The small ORF, 6, precedes 7. ORF 6 has the potential to form a structure containing a number of α-helices. The larger ORF has potential H-T-H and ATP-binding P loop domains. Finally, amber mutant analysis revealed that ORFs 6 and 7 encode the D and E genes of Mu, which are thought to encode the *pac*ase.

The plasmid pRS101 was constructed by cloning 4.5kb, from the Mu genome containing the 7 identified ORFS, into an expression vector to identify proteins produced from this region. Initiation of protein production from this construct resulted in rapid lysis of the bacterial cultures. Phage lytic systems are encoded by an operon in which a lysis regulator, or holin, precedes the lytic enzyme (Young, 1992). The holin controls the release of the endolysin from the cytoplasm to temporally regulate the lysis of the bacterium. Deletion of the holin or the endolysin abolishes lysis by the phage. However, it has been shown that the function of the phage holin can be mimicked by mild chloroform treatment of the bacteria (Reader and Siminovitch, 1971). To identify genes that may encode lytic enzymes ORFs 2-7 were removed from pRS101 to create pRS102. Induction of protein production from pRS102 did not result in lysis of the culture. When the

induced pRS102 culture weas treated with chloroform the bacteria were readily lysed. These studies indicate that the Mu lysis system encodes a regulator-effector system similar to other phages but that this system does not conform to the arrangement previously described for these proteins.

The maturation enzyme of Mu is thought to be encoded within the 4.5 kb region downstream of P_{lys}. Plasmid pWM5 contains this portion of the phage genome but induction of protein production from this construct does not result in rapid lysis of the culture. Crude protein extracts produced from strains containing pWM5 were used to localize the production of the Mu *pac*ase. Mobility shift assays, using a *pac*- specific probe, were used to demonstrate the presence of protein(s) that could recognize the packaging signal in pWM5 and control lysates. The pWM5 culture contained protein (s) that are able to specifically recognize the Mu *pac*ase binding site. Two complexes of differing mobility were revealed. Both complexes were visible in the presence of low concentrations of competitor DNA. The higher complex, however, was disrupted at higher concentrations of competitor DNA. The complexes might represent oligomerization of a single protein on the DNA or assembly of a heterooligomeric protein-DNA structure. Further investigation of the nature of the complexes will be facilitated by the use of purified *pac*ase.

An *in vitro* packaging assay has been developed for use with the transposable phages (Burns *et al.* 1990). The assay uses crude lysates produced from thermoinduced lysogenic cultures. Detailed examination of the mechanism of packaging will be facilitated by the refinement of this assay. The use of purified phage components will enable more rigorous biochemical examination of this process. To refine this assay an

examination of phage development and purification of structural precursors was carried out. Electron micrographs of thin sectioned, thermoinduced, bacterial cultures lysogenic for Mu and D3112 showed development of the phage head precursor. The bacteria showed evidence that phage-induced lysis was occurring at 30 minutes after induction of phage development, prior to the appearance of morphogenetic precursors. By 45 minutes post-induction phage proheads are visible. Crude extracts produced from the induced lysogenic cultures subsequent to the appearance of proheads were separated on continuous sucrose gradients. Protein profiles of the gradient using SDS-PAGE and western blotting revealed that three fractions, in the dense portion of the gradient, were highly enriched for proheads. Electron micrographs of the material within these dense fractions show the presence of proheads. However no phage tails are present. Separation of proheads, in this manner, indicate that the major head protein of D3112 has a mass of 37kDa.

Mobility shift analysis of these fractions, again using a *pac*-specific DNA probe revealed that a *pac* recognition function was associated with the prohead-rich fractions. When used in *in vitro* packaging assays, the prohead rich fragment was able to protect the phage-specific DNA substrate from DNaseI degradation and restriction enzyme digestion. This protection is indicative of packaging of the substrate into the phage prohead rendering it inaccessible to nuclease treatment. Packaging by phage maturation enzymes is dependent on ATP for use as en energy source (Black, 1988). The *in vitro* packaging reactions were dependent on ATP and Spermidine. Spermidine is thought to facilitate the packaging of the DNA by neutralizing the high charge concentration that results from the packing of the DNA into a paracrystalline state. In addition, packaging was improved by

the presence of bacterially derived HU or IHF. These two proteins participate in many phage processes, aiding in the formation of higher order of nucleoprotein complexes (Craig, 1996; Pato, 1989). These two proteins may therefore help to orient the phage enzyme so that it may productively cleave the replicated phage genome. Further purification of phage precursors, and refinement of the *in vitro* assay, will provide the tools necessary to examine and define the mechanism that mediates the switch from replication to packaging and morphogenesis of the Mu, D108 and D3112.

Chapter Five

References

References

Adams, M. B., Hayden, M., and Casjens, S.(1983). On the sequential packaging of bacteriophage P22 DNA. *J. Virol.* **46**, 673-677.

Adzuma, K. and Mizuuchi, K.(1988). Target immunity of Mu transposition reflects a differential distribution of Mu B protein. *Cell* **53**, 257-266.

Allet, B.(1978). Nucleotide sequences at the ends of bacteriophage Mu DNA. *Nature* **274**, 553-558.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J.(1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389-3402.

Artsimovitch, I., Kahmeyer-Gabbe, M., and Howe, M. M.(1996a). Distortion in the spacer region of Pm during activation of middle transcription of phage Mu. *Proc. Natl. Acad. Sci. USA* **93**, 9408-9413.

Artsimovitch, I., Murakami, K., Ishihama, A., and Howe, M. M.(1996b). Transcription activation by the bacteriophage Mu Mor protein requires the C-terminal regions of both alpha and sigma70 subunits of Escherichia coli RNA polymerase. *J. Biol. Chem.* **271**, 32343-32348.

Artsimovitch, I. and Howe, M. M.(1996). Transcription activation by the bacteriophage Mu Mor protein: analysis of promoter mutations in Pm identifies a new region required for promoter function. *Nucleic Acids Res.* **24**, 450-457.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K.(2000). *Current protocols in molecular biology*, John Wiley and Sons, Boston.

Bazinet, C., Benbasat, J., King, J., Carazo, J. M., and Carrascosa, J. L.(1988). Purification and organization of the gene 1 portal protein required for phage P22 DNA packaging. *Biochemistry* **27**, 1849-1856.

Bazinet, C. and King, J.(1985). The DNA translocating vertex of dsDNA bacteriophage. *Ann. Rev. Microbiol.* **39**, 109-129.

Bazinet, C. and King, J.(1988). Initiation of P22 procapsid assembly in vivo. *J. Mol. Biol.* **202**, 77-86.

Becker, A., Murialdo, H., and Gold, M.(1977). Studies on an in vitro system for the packaging and maturation of phage lambda DNA. *Virology* **78**, 277-290.

Becker, A. and Gold, M.(1988). Prediction of an ATP reactive center in the small subunit, gpNu1, of the phage lambda terminase enzyme. *J. Mol. Biol.* **199**, 219-222.

Becker, B., de la Fuente, N., Gassel, M., Gunther, D., Tavares, P., Lurz, R., Trautner, T. A., and Alonso, J. C.(1997). Head morphogenesis genes of the Bacillus subtilis bacteriophage SPP1. *J. Mol. Biol.* **268**, 822-839.

Bertani, L. E. and Six, E. W.(1988). *The Bacteriophages* (Calendar, R. Ed.) Plenum Press, New York and London. 73-143.

Bidnenko, E. M., Akhverdyan, V. Z., Khrenova, E. A., Yanenko, A. S., and Krylov, V. N.(1989). Genetic control of morphogenesis of *Pseudomonas aeruginosa* transposable phage D3112. *Genetika* **25**, 2126-2137.

Black, L. W., Newcomb, W. W., Boring, J. W., and Brown, J. C.(1985). Ion etching bacteriophage T4: support for a spiral-fold model of packaged DNA. *Proc. Natl. Acad. Sci. USA* **82**, 7960-7964.

Black, L. W.(1988). *The Bacteriophages* vol. 2 (Calendar, R. Ed.) Plenum Publishing Corp. New York. 321-373.

Black, L. W.(1989). DNA packaging in dsDNA bacteriophages. *Ann. Rev. Microbiol.* **43**, 267-292.

Black, L. W., Showe, M. K., and Steven, A. C.(1994). *Molecular biology of bacteriophage T4* (Karam, J. D., Drake, J. W., Kreuzer, K. N., Mosig, G., Hall, D. H., Eiserling, F. A., Black, L. W., Spicer, E. K., Kutter, E., Carlson, K., and Miller, E. S. Eds.) American Society for Microbiology, Washington D.C. 218-258.

Blasi, U., Nam, K., Hartz, D., Gold, L., and Young, R.(1989). Dual translational initiation sites control function of the lambda S gene. *EMBO J* **8**, 3501-3510.

Blasi, U., Chang, C. Y., Zagotta, M. T., Nam, K. B., and Young, R.(1990). The lethal lambda S gene encodes its own inhibitor. *EMBO J* **9**, 981-989.

Bolker, M., Wulczyn, F. G., and Kahmann, R.(1989). Role of bacteriophage Mu C protein in activation of the mom gene promoter. *J. Bact.* **171**, 2019-2027.

Bowden, D. W. and Modrich, P.(1985). In vitro maturation of circular bacteriophage P2 DNA. Purification of ter components and characterization of the reaction. *J Biol Chem* **260**, 6999-7007.

Bukhari, A. I., Froshauer, S., and Botchan, M.(1976). Ends of bacteriophage mu DNA. *Nature* **264**, 580-583.

Bukhari, A. I. and Taylor, A. L.(1975). Influence of insertions on packaging of host sequences covalently linked to bacteriophage Mu DNA. *Proc. Natl. Acad. Sci, USA* **72**, 4399-4403.

Burns, C. M., Chan, H. L., and DuBow, M. S.(1990). In vitro maturation and encapsidation of the DNA of transposable Mu-like phage D108. *Proc. Natl. Acad. Sci, USA* **87**, 6092-6096.

Carazo, J. M., Santisteban, A., and Carrascosa, J. L.(1985). Three-dimensional reconstruction of bacteriophage phi 29 neck particles at 2 X 2 nm resolution. *J. Mol. Biol.* **183**, 79-88.

Casjens, S., Wyckoff, E., Hayden, M., Sampson, L., Eppler, K., Randall, S., Moreno, E. T., and Serwer, P.(1992). Bacteriophage P22 portal protein is part of the gauge that regulates packing density of intravirion DNA. *J. Mol. Biol.* **224**, 1055-1074.

Casjens, S. and Hendrix, R.(1988). *The Bacteriophages* (Calendar, R. Ed.) Plenum Press, New York and London. 15-91.

Caspar, D. L. D. and Klug, A.(1962). Physical principles in the construction of regular viruses. *Cold Spring Harbor Symposia on Quantitative Biology* **27**, 1-24.

Catalano, C. E., Cue, D., and Feiss, M.(1995). Virus DNA packaging: the strategy used by phage lambda. *Mol. Microbiol.* **16**, 1075-1086.

Cerritelli, M. E. and Studier, F. W.(1996a). Purification and characterization of T7 head-tail connectors expressed from the cloned gene. *J Mol Biol* **258**, 299-307.

Cerritelli, M. E. and Studier, F. W.(1996b). Assembly of T7 capsids from independently expressed and purified head protein and scaffolding protein. *J. Mol. Biol.* **258**, 286-298.

Chai, S., Lurz, R., and Alonso, J. C.(1995). The small subunit of the terminase enzyme of Bacillus subtilis bacteriophage SPP1 forms a specialized nucleoprotein complex with the packaging initiation region. *J Mol Biol* **252**, 386-398.

Chai, S., Szepan, U., and Alonso, J. C.(1997). Bacillus subtilis bacteriophage SPP1 terminase has a dual activity: it is required for the packaging initiation and represses its own synthesis. *Gene* **184**, 251-256.

Chase, C. D. and Benzinger, R. H.(1982). Transfection of *Escheichia coli* spherolplasts with a bacteriophage Mu DNA-protein complex. *J. Virol.* **42**, 176-185.

Chen, C. and Guo, P.(1997). Sequential action of six virus-encoded DNA-packaging RNAs during phage phi29 genomic DNA translocation. *J. Virol.* **71**, 3864-3871.

Chen, H., Bjerknes, M., Kumar, R., and Jay, E.(1994). Determination of the optimal aligned spacing between the Shine-Dalgarno sequence and the translation initiation codon of *Escherichia coli* mRNAs. *Nucleic Acids Res* **22**, 4953-4957.

Conway, J. F., Duda, R. L., Cheng, N., Hendrix, R. W., and Steven, A. C.(1995). Proteolytic and conformational control of virus capsid maturation: the bacteriophage HK97 system. *J. Mol. Biol.* **253**, 86-99.

Coombs, D. H. and Eiserling, F. A.(1977). Studies on the structure, protein composition and aseembly of the neck of bacteriophage T4. *J. Mol. Biol.* **116**, 375-405.

Coren, J. S., Pierce, J. C., and Sternberg, N.(1995). Headful packaging revisited: the packaging of more than one DNA molecule into a bacteriophage P1 head. *J. Mol. Biol.* **249**, 176-184.

Craig, N. L.(1996). *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. (F.C.Neidhardt, R.Curtiss III, J.L.Ingraham, E.C.C.Lin, K.B.Low, B.Magasamik, W.S.Reznikoff, M.Riley, M.Schaechter, and H.G.Umberger, Eds.) 2nd edition. American Society for Microbiology Press, Washington D.C. 2339-2362.

Craigie, R., Mizuuchi, M., and Mizuuchi, K.(1984). Site-specific recognition of the bacteriophage Mu ends by the Mu A protein. *Cell* **39**, 387-394.

d'Herelle, F.(1917). Sur un microbe invisible antagoniste des bacilles dysenteriques. *Compt. Rend. Acad. Sci.* **165**, 373-375.

Daniell, E., Kohne, D. E., and Abelson, J.(1975). Characterization of the inhomogeneous DNA in virions of bacteriophage Mu by DNA reannealing kinetics. *J. Virol.* **15**, 739-743.

Dhillon, E. K., Dhillon, T. S., Lai, A. N., and Linn, S.(1980). Host range, immunity and antigenic properties of lambdoid coliphage HK97. *J Gen Virol* **50**, 217-220.

Dokland, T., Lindqvist, B. H., and Fuller, S. D.(1992). Image reconstruction from cryo-electron micrographs reveals the morphopoietic mechanism in the P2-P4 bacteriophage system. *EMBO Journal* **11**, 839-846.

Dokland, T., Isaksen, M. L., Fuller, S. D., and Lindqvist, B. H.(1993). Capsid localization of the bacteriophage P4 Psu protein. *Virology* **194**, 682-687.

Dokland, T. and Murialdo, H.(1993). Structural transitions during maturation of bacteriophage lambda capsids. *J. Mol. Biol.* **233**, 682-694.

Driedonks, R. A., Engel, A., tenHeggeler, B., and van Driel, R.(1981). Gene 20 product of bacteriophage T4 its purification and structure. *J Mol Biol* **15**, 641-662.

Dube, P., Tavares, P., Lurz, R., and van Heel, M.(1993). The portal protein of bacteriophage SPP1: a DNA pump with 13-fold symmetry. *EMBO Journal* **12**, 1303-1309.

DuBow, M. S. and Bukhari, A. I.(1981). *Bacteriophage Assembly* (DuBow, M. S. Ed.) Alan R. Liss Inc. New York. 47-67.

Duda, R. L., Hempel, J., Michel, H., Shabanowitz, J., Hunt, D., and Hendrix, R. W.(1995a). Structural transitions during bacteriophage HK97 head assembly. *J. Mol. Biol.* **247**, 618-635.

Duda, R. L., Martincic, K., and Hendrix, R. W.(1995b). Genetic basis of bacteriophage HK97 prohead assembly. *J. Mol. Biol.* **247**, 636-647.

Duda, R. L.(1998). Protein chainmail: Catenated protein in viral capsids. *Cell* **94**, 55-60.

Dunn, J. J. and Studier, F. W.(1983). Complete nucleotide sequence of bacteriophage T7 DNA and the location of T7 genetic elements. *J Mol Biol* **166**, 477-535.

Earnshaw, W. and King, J.(1978). Structure of phage P22 coat protein aggregates formed in the absence of the scaffolding protein. *J. Mol. Biol.* **126**, 721-747.

Earnshaw, W. C. and Casjens, S. R.(1980). DNA packaging by the double-stranded DNA bacteriophages. *Cell* **21**, 319-331.

Engler, J. A., Forgie, R. A., and Howe, M. M.(1980). Restriction endonuclease *Bam* HI cleaves bacteriophage Mu DNA within cistrons E and F. *Gene* **10**, 79-83.

Feiss, M., Kobayashi, I., and Widner, W.(1983). Separate sites for binding and nicking of bacteriophage lambda DNA by terminase. *Proc. Natl. Acad. Sci, USA* **80**, 955-959.

Feiss, M., Fogarty, S., and Christiansen, S.(1988). Bacteriophage lambda DNA packaging: a mutant terminase that is independent of integration host factor. *Mol. Gen. Genet.* **212**, 142-148.

Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J., Dougherty, B. A., Merrick, J. M., McKenney, K., Sutton, G. G., FitzHugh, W., Fields, C. A., Gocayne, J. D., Scott, J. D., Shirley, R., Liu, L. I., Glodek, A., Kelley, J. M., Weidman, J. F., Phillips, C. A., Spriggs, T., Hedblom, E., Cotton, M. D., Utterback, T., Hanna, M. C., Nguyen, D. T., Saudek, D. M., Brandon, R. C., Fine, L. D., Fritchman, J. L., Fuhrmann, J. L., Geoghagen, N. S., Gnehm, C. L., McDonald, L. A., Small, K. V., Fraser, C. M., Smith, H. O., and Venter, J. C.(1995). Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. *Science* **269**, 496-512.

Frackman, S., Siegele, D. A., and Feiss, M.(1984). A functional domain of bacteriophage lambda terminase for prohead binding. *J. Mol. Biol.* **180**, 283-300.

Fujisawa, H., Shibata, H., and Kato, H.(1991). Analysis of interactions among factors involved in the bacteriophage T3 DNA packaging reaction in a defined *in vitro* system. *Virology* **185**, 788-794.

Fujisawa, H. and Morita, M.(1997). Phage DNA packaging. *Genes Cells* **2**, 537-545.

Fuller, M. T. and King, J.(1981). Purification of the coat and scaffolding proteins from procapsids of bacteriophage P22. *Virology* **112**, 529-547.

George, M. and Bukhari, A. I.(1981). Heterogeneous host DNA attached to the left end of mature bacteriophage Mu DNA. *Nature* **292**, 175-176.

Gill, G. S., Hull, R. C., and Curtiss, R. ,3d(1981). Mutator bacteriophage D108 and its DNA: an electron microscopic characterization. *J. Virol.* **37**, 420-430.

Giphart-Gassler, M., Wijffelman, C., and Reeve, J.(1981). Structural polypeptides and products of late genes of bacteriophage Mu:characterization and functional aspects. *J Mol Biol* **145**, 139-163.

Gloor, G. and Chaconas, G.(1986). The bacteriophage Mu N gene encodes the 64-kDa virion protein which is injected with, and circularizes, infecting Mu DNA. *J. Biol. Chem.* **261**, 16682-16688.

Gold, M., Becker, A., and Parris, W.(1983). The bacteriophage lambda terminase enzyme. *Methods in Enzymology* **100**, 183-191.

Gold, M. and Becker, A.(1983). The bacteriophage lambda terminase. Partial purification and preliminary characterization of properties. *J. Biol. Chem.* **258**, 14619-14625.

Greene, B. and King, J.(1996). Scaffolding mutants identifying domains required for P22 procapsid assembly and maturation. *Virology* **225**, 82-96.

Grimaud, R.(1996). Bacteriophage Mu head assembly. *Virology* **217**, 200-210.

Groenen, M. A. and van de Putte, P.(1985). Mapping of a site for packaging of bacteriophage Mu DNA. *Virology* **144**, 520-522.

Grundy, F. J. and Howe, M. M.(1984). Involvement of the invertible G segment in bacteriophage mu tail fiber biosynthesis. *Virology* **134**, 296-317.

Grundy, F. J. and Howe, M. M.(1985). Morphogenetic structures present in lysates of amber mutants of bacteriophage Mu. *Virology* **143**, 485-504.

Guasch, A., Pous, J., Parraga, A., Valepuesta, J. M., Carrascosa, J. L., and Coll, M.(1998). Crystallographic analysis reveals the 12-fold symmetry of the bacteriophage phi29 connector particle. *J Mol Biol* **281**, 219-225.

Guo, P. X., Bailey, S., Bodley, J. W., and Anderson, D.(1987a). Characterization of the small RNA of the bacteriophage phi 29 DNA packaging machine. *Nucleic Acids Res.* **15**, 7081-7090.

Guo, P. X., Erickson, S., and Anderson, D.(1987b). A small viral RNA is required for in vitro packaging of bacteriophage phi 29 DNA. *Science* **236**, 690-694.

Guo, P. X., Erickson, S., Xu, W., Olson, N., Baker, T. S., and Anderson, D.(1991). Regulation of the phage phi 29 prohead shape and size by the portal vertex. *Virology* **183**, 366-373.

Hamada, K., Fujisawa, H., and Minagawa, T.(1986). Overproduction and purification of the products of bacteriophage T3 genes 18 and 19, two genes involved in DNA packaging. *Virology* **151**, 110-118.

Harel, J., Duplessis, L., Kahn, J. S., and DuBow, M. S.(1990). The cis-acting DNA sequences required in vivo for bacteriophage Mu helper-mediated transposition and packaging. *Arch. Microbiol.* **154**, 67-72.

Harshey, R. M.(1984). Transposition without duplication of infecting bacteriophage Mu DNA. *Nature* **311**, 580-581.

Harshey, R. M.(1988). *The Bacteriophages.* (Calendar, R. Ed.) Plenum Press, New York and London. 193-234.

Hendrix, R. W.(1978). Symmetry mismatch and DNA packaging in large bacteriophages. *Proc. Natl. Acad. Sci. USA* **75**, 4779-4783.

Hendrix, R. W.(1998). Bacteriophage DNA packaging: RNA gears in a DNA transport machine. *Cell* **94**, 147-150.

Hendrix, R. W. and Casjens, S. R.(1975). Assembly of bacteriophage lambda heads: protein processing and its genetic control in petit lambda assembly. *J. Mol. Biol.* **91**, 187-199.

Hendrix, R. W. and Duda, R. L.(1998). Bacteriophage HK97 head assembly: a protein ballet. *Adv. Virus Res.* **50**, 235-288.

Higgins, N. P., Collier, D. A., Kilpatrick, M. W., and Krause, H. M.(1989). Supercoiling and integration host factor change the DNA conformation and alter the flow of convergent transcription in phage Mu. *J. Biol. Chem.* **264**, 3035-3042.

Higgins, R. R., Lucko, H. J., and Becker, A.(1988). Mechanism of cos DNA cleavage by bacteriophage lambda terminase: multiple roles of ATP. *Cell* **54**, 765-775.

Howe, M. M.(1987). *Phage Mu* (Symonds, N., Toussaint, A., van de Putte, P., and Howe, M. M. Eds.) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York. 63-74.

Hsiao, C. L. and Black, L. W.(1978). Head morphogenesis of bacteriophage T4. II. The role of gene 40 in initiating prehead assembly. *Virology* **91**, 15-25.

Hull, R. A., Gill, G. S., and Curtiss, R. ,3d(1978). Genetic characterization of Mu-like bacteriophage D108. *J. Virol.* **27**, 513-518.

Hwang, Y., Catalano, C. E., and Feiss, M.(1996). Kinetic and mutational dissection of the two ATPase activities of terminase, the DNA packaging enzyme of bacteriophage Chi. *Biochemistry* **35**, 2796-2803.

Hwang, Y. and Feiss, M.(1996). Mutations affecting the high affinity ATPase center of gpA, the large subunit of bacteriophage lambda terminase, inactivate the endonuclease activity of terminase. *J. Mol. Biol.* **261**, 524-535.

Ibanez, C., Garcia, J. A., Carrascosa, J. L., and Salas, M.(1984). Overproduction and purification of the connector protein of Bacillus subtilis phage phi 29. *Nucleic Acids Res.* **12**, 2351-2365.

Jardine, P. J. and Coombs, D. H.(1998). Capsid expansion follows the initiation of DNA packaging in bacteriophage T4. *J Mol Biol* **284**, 661-672.

Johnson, G., Widner, W., Xin, W. N., and Feiss, M.(1991). Interference with phage lambda development by the small subunit of the phage 21 terminase, gp1. *J. Bact.* **173**, 2733-2738.

Kahmeyer-Gabbe, M. and Howe, M. M.(1996). Regulatory factors acting at the bacteriophage Mu middle promoter. *J. Bact.* **178**, 1585-1592.

Kato, H. and Baschong, C.(1997). Isolation of a gp20-complexand its role in *in vitro* assembly of both prohead and core of bacteriophage T4. *Virology* **227**, 400-408.

Kellenberger, E.(1990). Form determination of the heads of bacteriophages. *Eur. J. Biochem.* **190**, 233-248.

Keller, B., Maeder, M., Becker-Laburte, C., Kellenberger, E., and Bickle, T. A.(1986). Amber mutants in gene 67 of phage T4. Effects on formation and shape determination of the head. *J. Mol. Biol.* **190**, 83-95.

Keller, B., Dubochet, J., Adrian, M., Maeder, M., Wurtz, M., and Kellenberger, E.(1988). Length and shape variants of the bacteriophage T4 head: mutations in the scaffolding core genes 68 and 22. *J. Virol.* **62**, 2960-2969.

Krause, H. M. and Higgins, N. P.(1984). On the mu repressor and early DNA intermediates of transposition. *Cold Spring Harbor Symposia on Quantitative Biology* **49**, 827-834.

Kuhn, A., Keller, B., Maeder, M., and Traub, F.(1987). Prohead core of bacteriophage T4 can act as an intermediate in the T4 head assembly pathway. *J. Virol.* **61**, 113-118.

Kukolj, G. and DuBow, M. S.(1992). Integration host factor activates the Ner-repressed early promoter of transposable Mu-like phage D108. *J. Biol. Chem.* **267**, 17827-17835.

Laemmli, U. K.(1970). Cleavage of structural proteins during assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

Lane, D., Prentki, P., and Chandler, M.(1999). Use of gel retardation to analyze protein-nucleic acid interactions. *Microbiol. Rev.* **56**, 509-528.

Lavoie, B. D., Chan, B. S., Allison, R. G., and Chaconas, G.(1991). Structural aspects of a higher order nucleoprotein complex: induction of an altered DNA structure at the Mu-host junction of the Mu type 1 transpososome. *EMBO J.* **10**, 3051-3059.

Lavoie, B. D. and Chaconas, G.(1993). Site-specific HU binding in the Mu transpososome: conversion of a sequence-independent DNA-binding protein into a chemical nuclease. *Genes Dev.* **7**, 2510-2519.

Lavoie, B. D. and Chaconas, G.(1994). A second high affinity HU binding site in the phage Mu transpososome. *J. Biol. Chem.* **269**, 15571-15576.

Lee, C. S. and Guo, P.(1995). Sequential interactions of structural proteins in phage phi 29 procapsid assembly. *J. Virol.* **69**, 5024-5032.

Leung, P. C., Teplow, D. B., and Harshey, R. M.(1989). Interaction of distinct domains in Mu transposase with Mu DNA ends and an internal transpositional enhancer. *Nature* **338**, 656-658.

Levchenko, I., Luo, L., and Baker, T. A.(1995). Disassembly of the Mu transposase tetramer by the ClpX chaperone. *Genes Dev.* **9**, 2399-2408.

Levchenko, I., Yamauchi, M., and Baker, T. A.(1997). ClpX and MuB interact with overlapping regions of Mu transposase: implications for control of the transposition pathway. *Genes Dev.* **11**, 1561-1572.

Levin, D. B. and DuBow, M. S.(1987). Cloning and localization of the repressor gene (c) of the Mu-like transposable phage D108. *FEBS Letters* **222**, 199-203.

Linderoth, N. A., Ziermann, R., Haggard-Ljungquist, E., Christie, G. E., and Calendar, R.(1991). Nucleotide sequence of the DNA packaging and capsid synthesis genes of bacteriophage P2. *Nucleic Acids Res.* **19**, 7207-7214.

Ljungquist, E. and Bukhari, A. I.(1977). State of prophage Mu DNA upon induction. *Proc. Natl. Acad. Sci. USA* **74**, 3143-3147.

Lowry, O. H., Roebrough, N. J., Farr, A. L., and Randall, R. J.(1951). Protein measurement with the Folin phenol reagent. *J. Biol. Chem.* **193**, 265-275.

Lukashin, A. and Borodovsky, M.(1998). GeneMark.hmm:new solutions for gene finding. *Nucleic Acids Res* **26**, 1107-1115.

Margolin, W., Rao, G., and Howe, M. M.(1989). Bacteriophage Mu late promoters: four late transcripts initiate near a conserved sequence. *J. Bact.* **171**, 2003-2018.

Margolin, W. and Howe, M. M.(1986). Localization and DNA sequence analysis of the C gene of bacteriophage Mu, the positive regulator of Mu late transcription. *Nucleic Acids Res.* **14**, 4881-4897.

Margolin, W. and Howe, M. M.(1990). Activation of the bacteriophage Mu lys promoter by Mu C protein requires the sigma 70 subunit of Escherichia coli RNA polymerase. *J. Bact.* **172**, 1424-1429.

Marrs, C. F. and Howe, M. M.(1983). AvaII and BglI restriction maps of bacteriophage Mu. *Virology* **126**, 563-575.

Marrs, C. F. and Howe, M. M.(1990). Kinetics and regulation of transcription of bacteriophage Mu. *Virology* **174**, 192-203.

Martuscelli, J., Taylor, A. L., Cummings, D. J., Chapman, V. A., DeLong, S. S., and Canedo, L.(1971). Electron microscopic evidence for linear insertion of bacteriophage MU-1 in lysogenic bacteria. *J. Virol.* **8**, 551-563.

Marvik, O. J., Jacobsen, E., Dokland, T., and Lindqvist, B. H.(1994a). Bacteriophage P2 and P4 morphogenesis: assembly precedes proteolytic processing of the capsid proteins. *Virology* **205**, 51-65.

Marvik, O. J., Sharma, P., Dokland, T., and Lindqvist, B. H.(1994b). Bacteriophage P2 and P4 assembly: alternative scaffolding proteins regulate capsid size. *Virology* **200**, 702-714.

Marvik, O. J., Dokland, T., Nokling, R. H., Jacobsen, E., Larsen, T., and Lindqvist, B. H.(1995). The capsid size-determining protein Sid forms an external scaffold on phage P4 procapsids. *J. Mol. Biol.* **251**, 59-75.

Mathee, K. and Howe, M. M.(1990). Identification of a positive regulator of the Mu middle operon. *J. Bact.* **172**, 6641-6650.

Mathee, K. and Howe, M. M.(1993). Bacteriophage Mu Mor protein requires sigma 70 to activate the Mu middle promoter. *J. Bact.* **175**, 5314-5323.

Maxam, A. M. and Gilbert, W.(1980). Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol* **65**, 499-560.

Mendelson, I., Gottesman, M., and Oppenheim, A. B.(1991). HU and integration host factor function as auxiliary proteins in cleavage of phage lambda cohesive ends by terminase. *J Bacteriol* **173**, 1670-1676.

Michaud, G., Zachary, A., Rao, V. B., and Black, L. W.(1989). Membrane-associated assembly of a phage T4 DNA entrance vertex structure studied with expression vectors. *J. Mol. Biol.* **209**, 667-681.

Miller, J. H.(1992). A short course in bacterial genetics. A laboratory manual and handbook for *Escherichia coli* and related bacteria, Cold Spring Harbor Laboratory Press, New York.

Mizuuchi, M., Baker, T. A., and Mizuuchi, K.(1995). Assembly of phage Mu transpososomes: cooperative transitions assisted by protein and DNA scaffolds. *Cell* **83**, 375-385.

Morita, M., Tasaka, M., and Fujisawa, H.(1995a). Structural and functional domains of the large subunit of the bacteriophage T3 DNA packaging enzyme: importance of the C-terminal region in prohead binding. *J. Mol. Biol.* **245**, 635-644.

Morita, M., Tasaka, M., and Fujisawa, H.(1995b). Structural and functional domains of the large subunit of the bacteriophage T3 DNA packaging enzyme: importance of the C-terminal region in prohead binding. *J Mol Biol* **245**, 635-644.

Muller, D. J., Engel, A., Carrascosa, J. L., and Velez, M.(1997). The bacteriophage phi29 head-tail connector imaged at high resolution with the atomic force microscope in buffer solution. *EMBO J.* **16**, 2547-2553.

Murialdo, H. and Becker, A.(1977). Assembly of biologically active proheads of bacteriophage lambda in vitro. *Proc. Natl. Acad. Sci, USA* **74**, 906-910.

Murialdo, H. and Becker, A.(1978). A genetic analysis of bacteriophage lambda prohead assembly in vitro. *J. Mol. Biol.* **125**, 57-74.

Nakasu, S., Fujisawa, H., and Minagawa, T.(1985). Purification of characterization of gene 8 product of bacteriophage T3. *Virology* **143**, 422-434.

Nilssen, O., Fossdal, C. G., Johansen, B. V., and Lindqvist, B. H.(1996a). Bacteriophage P4 capsid-size determination and its relationship to P2 helper interference. *Virology* **219**, 443-452.

Nilssen, O., Six, E. W., Sunshine, M. G., and Lindqvist, B. H.(1996b). Mutational analysis of the bacteriophage P4 capsid-size-determining gene. *Virology* **219**, 432-442.

Paolozzi, L. and Symonds, N.(1987). *Phage Mu.* (Symonds, N., Toussaint, A., van de Putte, P., and Howe, M. M. Eds.) Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. 53-62.

Parker, M. H., Stafford, W. F.,3rd, and Prevelige, P. E.,Jr.(1997). Bacteriophage P22 scaffolding protein forms oligomers in solution. *J. Mol. Biol.* **268**, 655-665.

Parker, M. H., Casjens, S., and Prevelige, P. E.,Jr.(1998). Functional domains of bacteriophage P22 scaffolding protein. *J Mol Biol* **281**, 69-79.

Parker, M. H. and Prevelige, P. E.,Jr.(1998). Electrostatic interactions drive scaffolding/coat protein binding and procapsid maturation in bacteriophage P22. *Virology* **250**, 337-349.

Parreira, R., Valyasevi, R., Lerayer, A. L., Erlich, S. D., and Chopin, M. C.(1996). Gene organization and transcription of a late-expressed region of a Lactococcus lactis phage. *J Bacteriol* **178**, 6158-6165.

Pato, M. L.(1989). *Mobile DNA* (Berg, D. E. and Howe, M. M. Eds.) American Society for Microbiology, Washington, D.C. 23-52.

Pato, M. L., Howe, M. M., and Higgins, N. P.(1990). A DNA gyrase-binding site at the center of the bacteriophage Mu genome is required for efficient replicative transposition. *Proc. Natl. Acad. Sci, USA* **87**, 8716-8720.

Pato, M. L. and Banerjee, M.(1996). The Mu strong gyrase-binding site promotes efficient synapsis of the prophage termini. *Mol. Microbiol.* **22**, 283-292.

Popa, M. P., McKelvey, T. A., Hempel, J., and Hendrix, R. W.(1991). Bacteriophage HK97 structure: wholesale covalent cross-linking between the major head shell subunits. *J. Virol.* **65**, 3227-3237.

Poteete, A. and Botstein, D.(1979). Purification and properties of proteins essential to DNA encapsulation by phage P22. *Virology* **95**, 565-573.

Prasad, B. V., Prevelige, P. E., Marietta, E., Chen, R. O., Thomas, D., King, J., and Chiu, W.(1993). Three-dimensional transformation of capsids associated with genome packaging in a bacterial virus. *J. Mol. Biol.* **231**, 65-74.

Prevelige, P. E.,Jr., Thomas, D., and King, J.(1988). Scaffolding protein regulates the polymerization of P22 coat subunits into icosahedral shells in vitro. *J. Mol. Biol.* **202**, 743-757.

Prevelige, P. E.,Jr., Thomas, D., and King, J.(1993). Nucleation and growth phases in the polymerization of coat and scaffolding subunits into icosahedral procapsid shells. *Biophys. J.* **64**, 824-835.

Prevelige, P. E.,Jr. and King, J.(1993a). Assembly of bacteriophage P22: A model for ds-DNA virus assembly. *Prog. Med. Virol.* **40**, 206-221.

Rajagopal, B. S., Reilly, B. E., and Anderson, D. L.(1993). Bacillus subtilis mutants defective in bacteriophage phi 29 head assembly. *J. Bact.* **175**, 2357-2362.

Ramesh, V. and Nagaraja, V.(1996). Sequence-specific binding of phage Mu C protein:Footprinting analysis reveals altered DNA conformation upon protein binding. *J. Mol. Biol.* **260**, 22-33.

Rao, V. B. and Black, L. W.(1988). Cloning, overexpression and purification of the terminase proteins gp16 and gp17 of bacteriophage T4. Construction of a defined in-vitro DNA packaging system using purified terminase proteins. *J. Mol. Biol.* **200**, 475-488.

Ray, P. and Murialdo, H.(1975). The role of gene *Nu3* in bacteriophage lambda head Morphogenesis. *Virology* **64**, 247-263.

Reader, R. W. and Siminovitch, L.(1971). Lysis defective mutants of bacteriophage lambda: on the role of the S function in lysis. *Virology* **43**, 623-637.

Rice, P. and Mizuuchi, K.(1995). Structure of the bacteriophage Mu transposase core: a common structural motif for DNA transposition and retroviral integration. *Cell* **82**, 209-220.

Rice, P. A.(1997). Making DNA do a U-turn: IHF and related proteins. *Curr. Opin. Struct. Biol.* **7**, 86-93.

Rishovd, S., Marvik, O. J., Jacobsen, E., and Lindqvist, B. H.(1994). Bacteriophage P2 and P4 morphogenesis: identification and characterization of the portal protein. *Virology* **200**, 744-751.

Rishovd, S., Holzenburg, A., Johansen, B. V., and Lindqvist, B. H.(1998). Bacteriophage P2 and P4 morphogenesis: structure and function of the connector. *Virology* **245**, 11-17.

Rishovd, S. and Lindqvist, B.(1992). Bacteriophage P2 and P4 morphogenesis: protein processing and capsid size determination. *Virology* **187**, 548-554.

Salas, M.(1988). *The Bacteriophages* (Calendar, R. Ed.) Plenum Press, New York, London. 169-191.

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

Sanger, F., Nicklen, S., and Couldson, A. R.(1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74**, 5463-5467.

Shinder, G. and Gold, M.(1988). The Nul subunit of bacteriophage lambda terminase binds to specific sites in cos DNA. *J. Virol.* **62**, 387-392.

Shore, S. H. and Howe, M. M.(1982). Bacteriophage Mu T mutants are defective in synthesis of the major head polypeptide. *Virology* **120**, 264-268.

Stoddard, S. F. and Howe, M. M.(1989). Localization and regulation of bacteriophage Mu promoters. *J. Bact.* **171**, 3440-3448.

Stoddard, S. F. and Howe, M. M.(1990). Characterization of the C operon transcript of bacteriophage Mu. *J. Bact.* **172**, 361-371.

Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W.(1990). Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* **185**, 60-89.

Sun, W., Hattman, S., Fujita, N., and Ishihama, A.(1998). Activation of bacteriophage Mu mom transcription by C protein does not require specific interaction with the carboxyl-terminal region of the alpha or sigma 70 subunit of Escherichia coli RNA polymerase. *J Bacteriol* **180**, 3257-3259.

Surette, M. G. and Chaconas, G.(1991). Stimulation of the Mu DNA strand cleavage and intramolecular strand transfer reactions by the Mu B protein is independent of stable binding of the Mu B protein to DNA. *J. Biol. Chem.* **266**, 17306-17313.

Szatmari, G. B., Kahn, J. S., and DuBow, M. S.(1986). Orientation and sequence analysis of right ends and target sites of bacteriophage mu and D108 insertions in the plasmid pSC101. *Gene* **41**, 315-319.

Takade, A., Umeda, A., Misumi, T., Sawae, Y., and Amako, K.(1991). Accumulation of phosphate-containing granules in the nucleoid area of *Pseudomonas aeruginosa*.. *Microbiol. Immunol.* **35**, 367-374.

Tao, Y., Olson, N. H., Xu, W., Anderson, D. L., Rossmann, M. G., and Baker, T. S.(1998). Assembly of a tailed bacterial virus and its genome release studied in three dimensions. *Cell* **95**, 431-437.

Tavares, P., Droge, A., Lurz, R., Graeber, I., Orlova, E., Dube, P., and van Heel, M.(1995). The SPP1 connection. *FEMS Microbiol. Rev.* **17**, 47-56.

Taylor, A. L.(1963). Bacteriophage induced mutations in Escherichia coli. *Proc. Natl. Acad. Sci. USA* **50**, 1043-1051.

Thuman-Commike, P. A., Greene, B., Jakana, J., Prasad, B. V., King, J., Prevelige, P. E.,Jr., and Chiu, W.(1996). Three-dimensional structure of scaffolding-containing phage p22 procapsids by electron cryo-microscopy. *J. Mol. Biol.* **260**, 85-98.

Thuman-Commike, P. A., Greene, B., Malinski, J. A., King, J., and Chiu, W.(1998). Role of the scaffolding protein in P22 procapsid size determination suggested by T = 4 and T = 7 procapsid structures. *Biophys. J.* **74**, 559-568.

Tolias, P. P. and DuBow, M. S.(1985). The cloning and characterization of the bacteriophage D108 regulatory DNA-binding protein ner. *EMBO J.* **4**, 3031-3037.

Tolias, P. P. and DuBow, M. S.(1989). A subsequence-specific DNA-binding domain resides in the 13 kDa amino terminus of the bacteriophage Mu transposase protein. *J. Mol. Recogn.* **1**, 172-178.

Traub, F. and Maeder, M.(1984). Formation of the prohead core of bacteriophage T4 in vivo. *J. Virol.* **49**, 892-901.

Trottier, M., Zhang, C., and Guo, P.(1996). Complete inhibition of virion assembly in vivo with mutant procapsid RNA essential for phage phi 29 DNA packaging. *J. Virol.* **70**, 55-61.

Tsui, L. and Hendrix, R. W.(1980). Head-tail connector of bacteriophage lambda. *J. Mol. Biol.* **142**, 419-438.

Tsuprun, V., Anderson, D., and Egelman, E. H.(1994). The bacteriophage phi 29 head-tail connector shows 13-fold symmetry in both hexagonally packed arrays and as single particles. *Biophys J* **66**, 2139-2150.

Tsuprun, V., Rajagopal, B. S., and Anderson, D.(1995). Electron microscopy of Bacillus subtilis GroESL chaperonin and interaction with the bacteriophage phi 29 head-tail connector. *J. Struct. Biol.* **115**, 258-266.

Tuma, R., Prevelige, P. E.,Jr., and Thomas, G. J.,Jr.(1996). Structural transitions in the scaffolding and coat proteins of P22 virus during assembly and disassembly. *Biochemistry* **35**, 4619-4627.

Tuma, R., Parker, M. H., Weigele, P., Sampson, L., Sun, Y., Krishna, N. R., Casjens, S., Thomas, G. J.,Jr., and Prevelige, P. E.,Jr.(1998). A helical coat protein recognition domain of the bacteriophage P22 scaffolding protein. *J Mol Biol* **281**, 81-94.

Twort, F. W.(1915). An investigation on the nature of ultramicroscopic viruses. *Lancet* **189** 1241-1243.

Valkenburg, J. A. C., Woldringh, C. L., Brakenhoff, G. J., van der Voort, H. T. M., and Nanninga, N.(1985). Confocal scanning light microscopy of the *Escherichia coli* nucleoid: Comparison with phase-contrast and electron microscope images. *J Bacteriol* **161**, 478-483.

Valle, M., Valpuesta, J. M., Carrascosa, J. L., Tamayo, J., and Garcia, R.(1996). The interaction of DNA with bacteriophage phi 29 connector: a study by AFM and TEM. *J. Struct. Biol.* **116**, 390-398.

van de Putte, P., Giphart-Gassler, M., Goosen, N., Goosen, T., and Van Leerdam, E.(1981). Regulation of integration and replication functions of bacteriophage Mu. *Cold Spring Harbor Symposia on Quantitative Biology* **45 Pt 1**, 347-353.

van der Vies, S. M., Gatenby, A. A., and Georgopoulos, C.(1994). Bacteriophage T4 encodes a co-chaperonin that can substitute for Escherichia coli GroES in protein folding. *Nature* **368**, 654-656.

Van Leerdam, E., Karreman, C., and van de Putte, P.(1982). Ner, a cro-like function of bacteriophage Mu. *Virology* **123**, 19-28.

van Meeteren, R., Giphart-Gassler, M., and van de Putte, P.(1980). Transcription of bacteriophage Mu. II. Transcription of the repressor gene. *Mol. Gen. Genet.* **179**, 185-189.

van Rijn, P. A., Goosen, N., Turk, S. C., and van de Putte, P.(1989). Regulation of phage Mu repressor transcription by IHF depends on the level of the early transcription. *Nucleic Acids Res.* **17**, 10203-10212.

van Rijn, P. A., van de Putte, P., and Goosen, N.(1991). Analysis of the IHF binding site in the regulatory region of bacteriophage Mu. *Nucleic Acids Res.* **19**, 2825-2834.

Venable, J. H. and Coggeshall, R.(1965). A simplified lead citrate stain for use in electron microscopy. *J Cell Biol.* **25**, 407-408.

Vieira, J. and Messing, J.(1987). Production of single-stranded plasmid DNA. *Methods Enzymol* **153**, 3-11.

Watson, M. L.(1958). Staining of tissue sections for electron microscopy with heavy metals. *J. Biophys. Biochem. Cytol.* **4**, 475-478.

Wichitwechkarn, J., Bailey, S., Bodley, J. W., and Anderson, D.(1989). Prohead RNA of bacteriophage phi 29: size, stoichiometry and biological activity. *Nucleic Acids Res.* **17**, 3459-3468.

Xie, Z. and Hendrix, R. W.(1995). Assembly in vitro of bacteriophage HK97 proheads. *J. Mol. Biol.* **253**, 74-85.

Yanisch-Perron, C., Vieira, J., and Messing, J.(1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**, 103-119.

Yap, N. L. and Rao, V. B.(1996). Novel mutants in the 5' upstream region of the portal protein gene 20 overcome a gp40-dependent prohead assembly block in bacteriophage T4. *J. Mol. Biol.* **263**, 539-550.

Young, R.(1992). Bacteriophage lysis: mechanism and regulation. *Microbiol. Rev.* **56**, 430-481.

Young, R. and Blasi, U.(1995). Holins: form and function in bacteriophage lysis. *FEMS Microbiol. Rev.* **17**, 191-205.

Zha, J., Zhao, Z., and Howe, M. M.(1994). Identification and characterization of the terminators of the lys and P transcripts of bacteriophage Mu. *J. Bact.* **176**, 1111-1120.

Ziegelhoffer, T., Yau, P., Chandrasekhar, G. N., Kochan, J., Georgopoulos, C., and Murialdo, H.(1992). The purification and properties of the scaffolding protein of bacteriophage lambda. *J. Biol. Chem.* **267**, 455-461.

Zlotnick, A.(1994). To build a virus capsid. An equilibrium model of the self assembly of polyhedral protein complexes. *J. Mol. Biol.* **241**, 59-67.

Appendix 1: Sequence of Mu late morphogenetic region described in chapter 2

The attached sequence represents the data derived by Felix Seider and Ian Siboo. The sequence is numbered based on the complete Mu sequence (Genbank file AF083977). The first nucleotide, 10300, is the first position of the *Bgl*II site. The region from the *Bgl*II site at bp10300 to the *Bam*HI site at 14330 was cloned into pET29b to create pRS101. The amino acid sequences of the proteins corresponding to the potential ORFs 1-7 are included on the attached file. In addition these sequences are numbered to the left.

GeneMark.hmm predicted 7 ORFs in this region when default conditions were used. For ORF1 (lys), the second methionine codon was chosen as the start site since its position relative to the RBS was more favorable. Of the methionine codons that had a probability of acting as an initiator for ORF 7 (*E*), the triplet starting at position 12899 was positioned most favorably with respect to an RBS. Therefore this position was chosen as the start site.

Analyses of the ORFs with programs such as DAS (Csezo, 1997) and TMHMM (Sonnhammer, 1998), which recognize potential transmembrane domains, indicated that ORF 4 had a strong probability of forming a transmembrane domain between amino acids 3 and 25. This type of domain is essential to phage holins. In addition a potential ORF (not recognize by Genemark) with a dual methione start and a transmembrane domain is encoded from 11226-11567. Consequently ORF4 or the dual met start ORF may encode the Mu holin. However, additional functional studies are required to confirm this activity.

M. Cserzo, E. Wallin, I. Simon, G. von Heijne and A. Elofsson. 1997 Prediction of transmembrane alpha-helices in procariotic membrane proteins: the Dense Alignment

Surface method. *Prot. Eng.* **10**: 673-676.

E. L.L. Sonnhammer, G. von Heijne, and A. Krogh. 1998 A hidden Markov model for predicting transmembrane helices in protein sequences. p 175-182. In *Proc. of Sixth Int. Conf. on Intelligent Systems for Molecular Biology*. J. Glasgow, T. Littlejohn, F. Major, R. Lathrop, D. Sankoff, and C. Sensen (ed). AAAI Press, Menlo Park, CA.

Appendix 2: Dimensions of Mu and D3112.

	Proheads (unexpanded)	Proheads (expanded)	Heads	Tails*
Mu	51 nm	54.2 nm	58 nm***	100 nm***
D3112	55.6 nm	-	64 nm**	200 nm**

*The tails of Mu are straight and contractile while those of D3112 are curved.
** These dimension were measured from electronmicrographs published in :
E.M. Bidenko, V.Z. Akhverdyan, E.A Khrenova, A.S Yanenko and V.N. Krylov. Genetic control of morphogenesis of *Pseudomonas aeruginosa* transposable phage D3112. *Genetika* **25**:2126-2137.
*** F.J. Grundy and M.M. Howe. Morphogenetic structures present in lysates of amber mutants of bacteriphage Mu. *Virology* **143**:285-504.