1988

Characterization Of The Integrative Precursor Protein-dna Complex Of Bacteriophage Mu

Gregory Brian Gloor

Follow this and additional works at: https://ir.lib.uwo.ca/digitizedtheses

NOTICE

AVIS

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c C-30

La qualité de cette microforme depend grandement de la qualité de la thèse soumise au microfilmage Nous avons tout fait pour assurer une qualité superieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade

La qualité d'impression de certames pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, tests publiés, etc) ne sont pas microfilmés.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30

Canada

CHARACTERIZATION OF THE INTEGRATIVE PRECURSOR
PROTEIN-DNA COMPLEX OF BACTERIOPHAGE Mu

by

Gregory B. Gloor

Department of Biochemistry

Submitted in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy

Faculty of Graduate Studies
The University of Western Ontario
London, Ontario
February, 1988

ABSTRACT

Bacteriophage Mu integrates and replicates its genome via DNA transposition. Mu transposition during integration is non-replicative (conservative) and generates simple insertions. Transpositon during the lytic cycle is replicative and amplifies the Mu genome by cointegrate production. Mu therefore, must chose between these two pathways. Infecting Mu DNA is found associated with a coinjected, 64 kDa virion protein bound noncovalently to its ends. Characterization of the protein-DNA complex is reported here.

Antiserum was prepared against the virion 64 kDa protein and used to probe an expression library of cloned Mu DNA sequences. The Mu *N* gene was mapped to the overproducing clones by physical and genetic techniques. Partial proteolysis of the N protein produced *in vitro* and the 64 kDa protein isolated from the protein-DNA complex showed them to be identical. The Mu *N* gene was sequenced and a region of homology to many site specific DNA binding proteins was observed.

The transposition end product produced *in vitro* by the N protein-Mu DNA complex isolated from Mu infected cells was examined by neutral and alkaline agarose gel electrophoresis. The protein-DNA complex was found to form an identical strand transferred end product as the mini-Mu control plasmid. The implication of these findings for Mu integration is that *in vivo*, the strand transferred product is probably processed by a second nick following the initial strand transfer reaction of transposition.

ACKNOWLEDGEMENTS

iv

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

°C	degrees centigrade
ϕ-CHCl$_3$	phenol-chloroform-isoamylalcohol 25:24:1
βME	β-mercaptoethanol
ATP	adenosine-5'-triphosphate
ATPγS	adenosine-5'-O-(3-thiotriphosphate)
bp	base pairs
da	dalton
ddNTP	2',3'-dideoxyribonucleotide-5'-triphosphate
dNTP	3'-deoxyribonucleotide-5'-triphosphate
dATP	3'-deoxyadenosine-5'-triphosphate
dCTP	3'-deoxycytidine-6'-triphosphate
dGTP	3'-deoxyguanosine-5'-triphosphate
dTTP	3'-deoxythymidine-5'-triphosphate
DTT	dithiotreitol
EDTA	ethylenediaminetetracetic acid
EGTA	ethyleneglycol-bis-(β-aminoethyl ether)-N,N,N',N'-tetracetic acid
kbp	kilo base pairs
kDa	kilodalton
pfu	plaque forming units
SDS	sodium dodecyl sulfate
TE	10 mM Tris-HCl pH 7.8, 1 mM EDTA
Tris	tris(hydroxymethyl)aminomethane

CHAPTER 1

INTRODUCTION

INTRODUCTION TO Mu

In 1963, Taylor discovered a bacteriophage contaminant in an *Escherichia coli* K12 strain which he received from another investigator. Subsequently, he found that the phage was temperate, like phage lambda. Unlike lambda, however, approximately 2% of the lysogens were found to be auxotrophic for an amino-acid or other growth factor. From this observation, Taylor proposed that the Mu phage DNA was capable of inserting into near random chromosomal locations. Taylor named this phage mutator, or Mu-1, for its ability to mutate the host that it invades.

Since 1963, bacteriophage Mu has been shown, by many investigators, to integrate and replicate its genome by transposition (described in the next section), and is in fact, one of the two *E. coli* phage known to utilize transposition in this way. The other phage, D108, is closely related to Mu in DNA sequence (Mise, 1971, Gill et al., 1981).

Mu DNA extracted from virions is double stranded and approximately 39 kbp in length. A map showing some physical detail of the Mu genome is shown in Fig. 1.1. When the virion DNA is denatured, renatured and examined by electron microscopy, two features of the genome become apparent. First, single-stranded DNA, 1-2 Kb long, is observed at the right end of Mu (Daniell et al. 1973, Hsu and Davidson, 1974), and corresponds to host DNA attached to the ends of the phage (Daniell et al., 1975). This is termed the
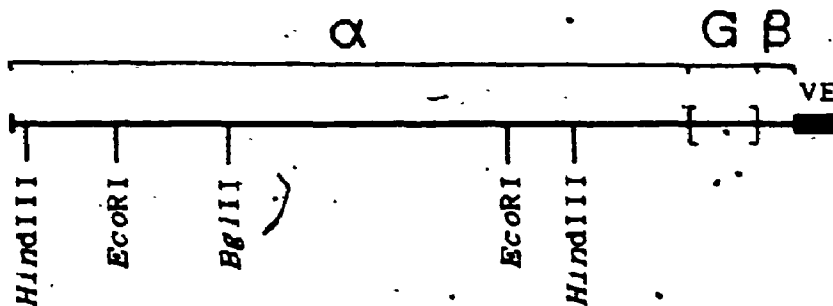
Figure 1.1: PHYSICAL MAP OF Mu DNA

A physical map of the Mu genome is shown along with a few landmark restriction sites. The host sequences at the Mu left end are shown as the thin vertical line, the variable right end (VE) of Mu is represented by the dark box, the invertible G region is enclosed in brackets. The three general regions of the Mu genome (α, G, β) are demarcated above the map. The α region left of the BglII site contains the phage early control genes, the transposition and replication genes and the semi-essential early genes. The remaining α region contains the late control genes and the morphogenesis and structural genes. The invertible G region contains tail spike genes necessary for host recognition. The β region contains a gene that controls the inversion of the G region and genes that control the protective modification of Mu DNA (Howe, 1987). The map is redrawn from Marrs and Howe (1983).

variable end (VE) region of Mu. Second, if the DNA was prepared from Mu lysogens induced to make phage, a denaturation bubble is observed near the variable end region in 50% of the molecules. This denaturation bubble is termed the G region of Mu and is a 3 kbp region that is capable of inverting. Inversion of this sequence is catalyzed by the *gin* gene product and allows the production of alternate phage absorption proteins that broaden the host range (Hsu and Davidson, 1974; Kamp et al., 1978; Kahmann et al., 1984).

Mu DNA is always found covalently associated with the DNA of the last host strain that it infected; there are approximately 56 to 144 base pairs of host DNA at the left end and about 1 500 base pairs at the right end (the variable end region; Bukhari et al., 1976; George and Bukhari, 1981). This results from the fact that Mu DNA is packaged by a headful mechanism starting near the left end of the genome and proceeding rightward until about 38 kbp has been packaged (Bukhari and Taylor, 1975).

Mu presents the investigator with some very important advantages in studying DNA transposition. The first advantage arises because Mu has evolved to utilize transposition very efficiently; transposition of Mu occurs at a frequency which is several orders of magnitude greater than that found in any other system. Second, one can infect a cell, known to have no other copies of Mu, with the phage and follow the fate of the infecting DNA. A third advantage of studying Mu is the convenient *in vitro* Mu transposition system developed by Mizuuchi (1983). Finally, Mu is a

temperate phage and mutants of Mu that can be induced to transpose at will have been isolated (Howe, 1973).

Mu was recognized as a transposing phage immediately upon discovery by Taylor (1963). In order to fully understand the Mu life cycle, one must understand DNA transposition.

INTRODUCTION TO TRANSPOSITION.

Transposition is the movement of a discrete piece of DNA, the transposon, to a new site in the genome. Transposition was first proposed by McClintock in the late 1940's to explain the variety of phenotypes observed in Maize kernel morphology (McClintock, 1950). McClintock observed that these phenotypes correlated with translocations of the maize chromosomes, and she further theorized that the translocations were caused by moveable pieces of DNA, or what are now called transposons. Physical evidence for transposition was not found for about two decades, until insertions in bacterial genes were observed to cause mutations.

There are two end products observed following transposition in bacteria: cointegrates and simple insertions (for a recent review on transposition see Grindley and Reed, 1985). The structure of the two end products is shown in Fig. 12. Cointegrates, also known as replicon fusions, result when the transposon is duplicated during transposition. A new copy of the transposon is generated and the vector portion of the donor molecule is integrated into the target

ve... .target

transposon

simple insertion

recombination
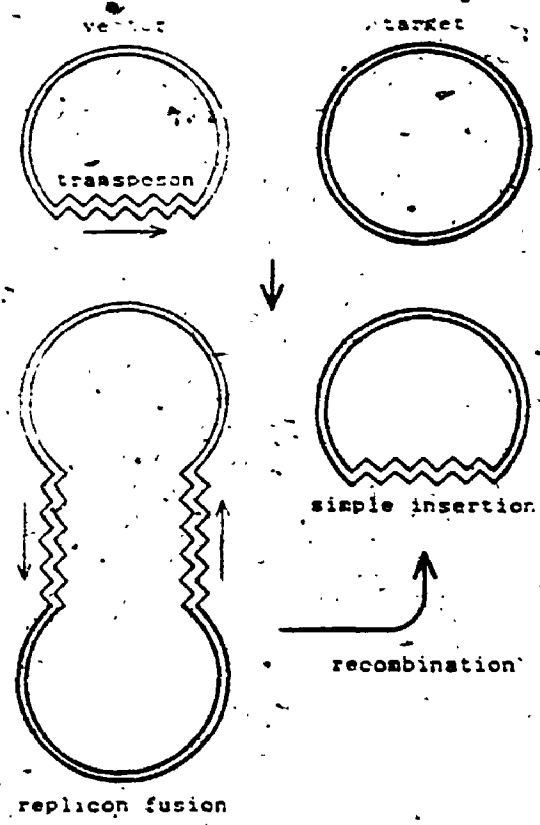
replicon fusion

FIGURE 12: TRANSPOSITION END PRODUCTS

Transposition can generate either of two end products, replicon fusions or simple insertions. Replicon fusions, the end product on the left, consist of the vector and target molecules joined by two copies of the transposon. These two copies are generated during the movement of the transposon and are found as direct repeats in the end product. Another term for replicon fusions is cointegrates, referring to the cointegration of transposon and vector DNA into the target. Simple insertions, the end product shown on the right, are the insertion of the transposon into the target DNA alone. Simple insertions may be generated via site specific recombination of a replicon fusion to regenerate the initial vector/transposon plasmid and the target plasmid containing the transposon. Mechanisms of generating replicon fusions and simple insertions are discussed in the text. The transposon sequences are shown as wavy lines, vector DNA is shown as thin lines and target DNA is shown as thick lines. The small arrows present the transposon's orientation.

site along with both copies of the transposon. Simple insertion end products appear as if the transposon had excised from the donor molecule and then inserted into the target molecule. However, site specific recombination of the cointegrate molecule would also give a simple inserted end product.

In addition to the two potential end products, there are two transposition pathways: replicative and conservative. Replicative transposition involves extensive DNA replication of the transposon, whereas conservative transposition involves a minimal amount of replication.

Movement of a transposon results in the duplication of a short sequence of the target DNA. These short repeats are found flanking the newly inserted transposon and are in a direct orientation. The size of the direct repeats formed is characteristic of each transposon. Mu generates 5 bp duplications of the target DNA during transposition (Allet, 1979).

Several models for replicative transposition have been proposed, but the model that has withstood the scrutiny allowed by in vitro experimentation is the model that was first proposed by Shapiro (1979) and subsequently by Arthur and Sherratt, (1979). Figure 13 shows the model that has come to be known as the Shapiro model of transposition. This model involves the concerted attack of both transposon ends on a target molecule. More specifically, the ends of the transposon are nicked and a double stranded cut is made
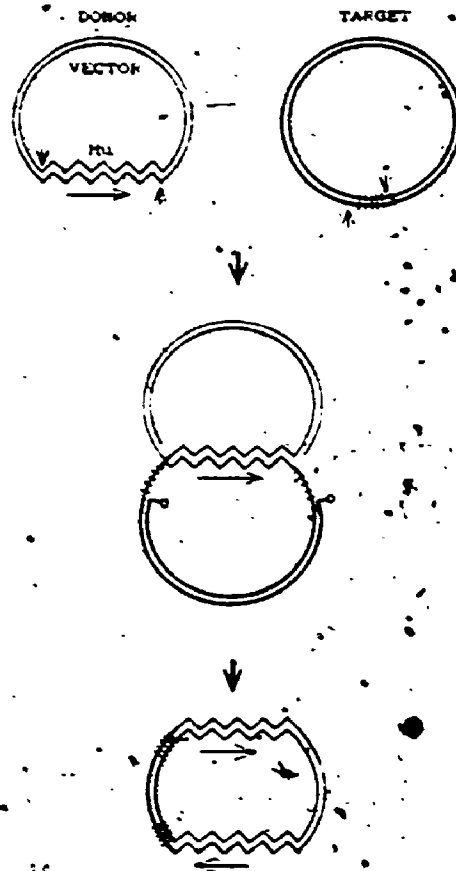
FIGURE 13: THE SHAPIRO MODEL OF TRANSPOSITION

The Shapiro transposition model (1979) predicts the formation of simultaneous nicks at opposite ends of the transposon (wavy lines) on opposite strands. The target molecule (thick lines) is opened by a double stranded cut that produces overhanging ends. The nicks at the ends of the transposon are ligated to the overhanging target sequences, generating the strand transferred structure shown in the middle that resembles a θ. DNA replication proceeds through the transposon stopping once the transposon has been replicated. The end result of this transposition model is known as a cointegrate molecule, where a second copy of the transposon has been generated by DNA replication, and the donor replicon (thin lines) has been joined to the target replicon. The polarity shown, where the 5' phosphate at the end of the target DNA 5 base overhang is joined to the 3' hydroxyl found at the nicked Mu end, was determined by Mizuuchi (1984). The duplication of target DNA found at transposon insertion sites is the result of replication of the small overhanging sequence generated in the target. In the diagram, the nick sites are shown as small arrows, the hydroxyl group in the strand transferred product, from which DNA replication is initiated, is shown as the small open circle. The lines crossing the target DNA represent the base pairs duplicated in the replicated end product. The orientation of the transposon is shown by the arrow parallel to it.

in the target molecule. The double stranded cut ends of the target molecule are then joined to the nicks found at the ends of the transposon. The strand transferred product thus formed resembles a replication fork. The transposon located in this strand transferred product is replicated by host proteins to generate a cointegrate end product. Simple insertions may be formed from the cointegrate molecule by a reciprocal recombination event (Shapiro, 1979).

Simple insertions could also arise by nonreplicative transposition. It has been observed that the Shapiro strand transferred product could also form conservative simple insertions (Ohtsubo et al., 1981; Mizuuchi, 1984; Derbyshire and Grindley, 1986). The strand transferred product shown in Fig. 13 does not have to undergo extensive DNA replication in order to result in an insertion of the transposon into the target DNA. Rather, the transposon could be cut by an endonuclease in the strand opposite to the original nick. This cut could be made either before, or after DNA replication has begun. Following repair of the nick, the end product is the insertion of the transposon into the target DNA. The donor's vector sequences are then degraded since they have been linearized and not resealed. Conservative transposition does not have to proceed via processing of a Shapiro strand transferred product. Double stranded cuts can be made at the ends of the transposon, rather than nicks (Berg, 1977). The double stranded cut transposon is then joined to the double stranded cut target DNA. This model of conservative transposition obviates the necessity of

forming a Shapiro strand transferred product. In this model the donor's vector sequences are presumed to be lost by degradation.

IN VITRO TRANSPOSITION OF Mu

The Shapiro model has been shown to be generally applicable for phage Mu transposition. An in vitro transposition system for Mu involving crude cell extracts has been described by Mizuuchi (1983). Using this system, Craigie and Mizuuchi (1985b) demonstrated biochemically, and Miller and Chaconas (1986) demonstrated by electron microscopy, that the strand transferred product for Mu transposition was that predicted by Shapiro's model.

Bacteriophage Mu requires the products of the A and B early genes to integrate and replicate its DNA (Faelen and Toussaint, 1973; Wijffelman et al., 1974; O'Day et al., 1978). These genes have been cloned and their proteins purified (Craigie and Mizuuchi, 1985a; Chaconas et al., 1985a). The Mu A protein has been shown to bind, to asymmetrically distributed sites, at the Mu ends (Craigie et al., 1984). The A protein binding sites correspond to sites at the ends of Mu required for transposition (Groenen et al., 1985). The B protein is a non-specific DNA binding protein (Chaconas et al., 1985a) that is required for in vitro transposition to proceed (Mizuuchi, 1983).

Craigie et al. (1985) refined the system to allow the strand transfer reaction to proceed in the presence of only three

proteins: the Mu A and B proteins, and the *E. coli* HU protein.
Under some conditions of suboptimal supercoiling, the *E. coli*
protein, IHF is required for efficient strand transfer (Surette and
Chaconas, unpublished).

Mizuuchi (1984) has demonstrated the polarity of the Mu strand
transfer reaction to be the joining of the 5' phosphate on the 5 bp
overhang found at the cut ends of the target DNA to the 3' hydroxyl
group found at the nicked Mu DNA ends. Surette et al., (1987; see
Fig. 1.4) showed that two protein-DNA complexes, the type 1 and 2
transpososomes, are active intermediates in the transposition of
Mu. The type 1 transpososome is the active intermediate in the
strand transfer reaction and consists of the Mu left (L) and right
(R) ends in the donor DNA bound non-covalently to the Mu A protein.
The Hu protein may also be a part of this protein-DNA complex. The
donor DNA molecule is nicked at the ends of the Mu sequences in
this structure. In the type 1 transpososome, the Mu domain of the
plasmid is supercoiled, held together by the proteins, and the
non-Mu domain is relaxed. The type 2 transpososome consists of the
Mu A and B proteins, possibly with HU, bound non-covalently to the
Mu ends in the strand transferred end product.

INTEGRATION OF Mu

The focus of this thesis is the integration of bacteriophage Mu
into the *E. coli* chromosome following infection. Mu DNA in the
virion is linear, double stranded and non-permuted (Allet, 1978).

Mini-Mu donor plasmid — Type 1 Transpososome — Type 2 Transpososome

A. HU, Mg²⁺ Reaction

B. ATP, Mg²⁺ Strand transfer

Target DNA

Topoisomerase — Topoisomerase

Non-Mu donor sequences
Mu sequences
Target sequences

SDS

Target DNA
B. ATP, Mg²⁺ Strand transfer

Type 1 (relaxed)

Type 2 (relaxed)

SDS

SDS

Nicked mini-Mu donor plasmid

Θ structure
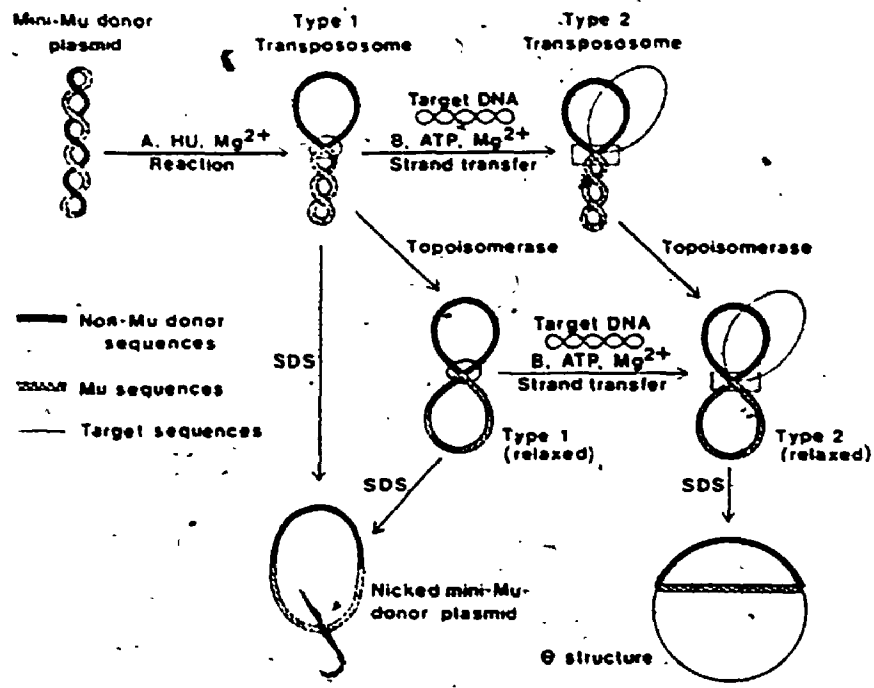
FIGURE 1.4: FORMATION OF TYPE 1 AND 2 TRANSPOSOSOMES

The type 1 transpososome is formed by incubation of the donor molecule with the Mu A and *E. coli* HU proteins in the presence of Mg⁺⁺. These components form a stable protein-DNA complex where the Mu ends in the donor molecule are held together. The Mu and non-Mu domains are topologically distinct in this structure; the Mu domain is supercoiled, while the vector domain is relaxed. Type 2 transpososomes are formed from type 1 transpososomes by the addition of the Mu B protein, target DNA and ATP. The Mu DNA ends are held together with the target DNA ends ligated to them, by another protein-DNA complex. The DNA in the type 2 transpososomes have the structure predicted for the strand transferred product by Shapiro (1979). Supercoiling is only required for the formation of the type 1 transpososome. Formation of the type 2 transpososome can proceed if the type 1 transpososome had been relaxed prior to the addition of the second set of components. This figure is from Surette et al, 1987.

The integration of linear Mu DNA into the chromosome of the host cell by a cointegrate pathway would result in the linearization of the chromosome and subsequent cell death. This can be observed by examination of Fig. 15. If the vector DNA in the donor molecule is not continuous, and replication of the strand transferred product occurs, then the resulting product is linear. However, since Mu is capable of forming lysogens, this problem must be circumvented by the phage.

Mu uses replicative transposition to amplify its genome during lytic infection (Chaconas et al., 1980, 1981a), generating replicon fusions or cointegrates. Therefore, the early favorite integration model involved replicative transposition to insert the Mu DNA into the chromosome, followed by recombination to regenerate a circular host chromosome (discussed by Toussaint and Résibois, 1983). However, experimental evidence which demonstrated that infecting Mu DNA could integrate into a plasmid, pBR322, without replication, provided early evidence that Mu integration could proceed via conservative transposition (Liébart et al, 1982). An experiment performed by Akroyd and Symonds (1983) showed that information from both infecting Mu DNA strands was passed on to the progeny phage. Experiments reported by Chaconas et al. (1983), demonstrated that integration of Mu resulted in the formation of simple insertions. These experiments combined, illustrated that integration of infecting Mu DNA did in fact, proceed conservatively. Subsequently, Harshey (1984), utilized a fully *dam* methylated phage to infect a *dam⁻ E. coli* strain. Harshey probed the integrated DNA with
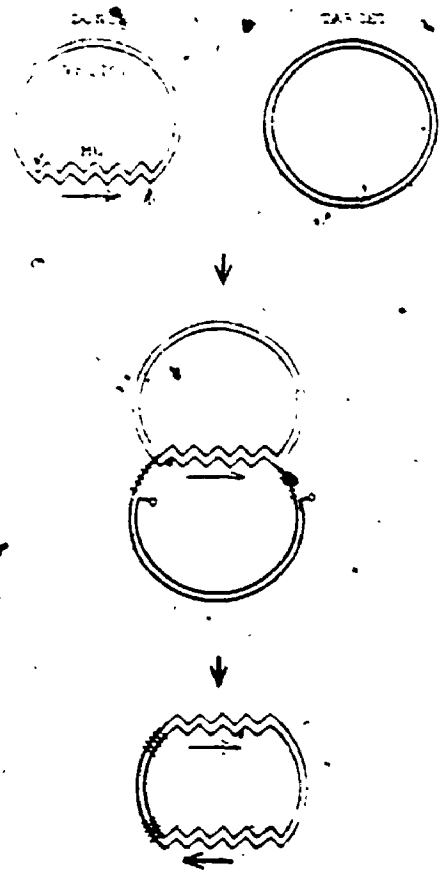
FIGURE 15 TRANSPOSITION FROM A LINEAR DONOR PRODUCES A LINEAR PRODUCT

This figure is a modification of Fig. 1.3, and shows how a transposon produces a linear product when the vector sequences are not continuous. As before, simultaneous nicks at opposite ends of the transposon (wavy lines) on opposite strands are produced. The target molecule (thick lines) is opened by a double stranded cut that produces overhanging ends. The nicks at the ends of the transposon are ligated to the overhanging target sequences, generating the strand transferred structure shown in the middle that resembles an open θ. DNA replication proceeds through the transposon stopping once the transposon has been replicated. The end result of this transposition is a structure containing 2 transposons joined by the target molecule, with the vector ends not being connected, thus producing a linear end product. In the diagram, the nick sites are shown as small arrows, the hydroxyl group in the strand transferred product, from which DNA replication is initiated, is shown as the small open circle. The lines crossing the target DNA represent the base pairs duplicated in the replicated end product. The orientation of the transposon is shown by the arrow parallel to it.

restriction endonucleases that were differentially affected by the methylation state of the DNA. This experiment clearly illustrated that the first transposition event following infection did not involve substantial DNA replication of the infecting Mu DNA, and was therefore conservative.

The integration of Mu into the host chromosome, therefore, represents a special challenge to the phage. Mu must be able to switch from the normal replicative transposition pathway and utilize conservative transposition during infection.

Protein requirements for the replicative transposition of Mu *in vivo* and *in vitro* include host functions (Toussaint and Faelen, 1974; McBeth and Taylor, 1982, 1983; Mizuuchi, 1983; Craigie and Mizuuchi, 1985b). The *in vivo* studies of Faelen et al. (1978) and Wijffelman and Lotterman (1977), established that the Mu *A* gene product was required for replicative transposition. This dependence on the A protein has also been found *in vitro* (Mizuuchi, 1983). *In vivo* replication of Mu cannot proceed in the absence of the Mu *B* gene product (Wijffelman and Lotterman, 1977). Genetic assays have demonstrated that, when gp*B* is absent, replicative transposition is reduced to 1% of the level normally found (Coehlo et al., 1982; Chaconas et al., 1984). Replicative transposition of Mu has not been observed *in vitro* in the absence of the B protein (Mizuuchi, 1983; Chaconas et al., 1985a). A function that stimulates Mu replication *in vivo*, *arm*, (Goosen et al., 1982) is not required *in vitro* (Mizuuchi, 1983).

The conservative integration reaction of Mu is less well characterized than its replicative counterpart since its discovery is much more recent. The Mu *A* gene product is absolutely required for lysogeny although the Mu *B* gene product is not (O'day et al., 1978). Mu *B* gene mutants, however, show approximately a 10 fold reduction in lysogeny. The isolation of Mu *A* gene mutants that allow lysogenization but not integration of the bulk infecting DNA, (Toussaint et al., 1987) indicate that the requirements for lysogeny may not accurately reflect those for integration. In addition, only 1% - 10% of infecting Mu DNA forms lysogens (Howe and Bade, 1975).

Chaconas et al. (1985b) analyzed the fate of bulk infecting DNA using various Mu *A* and *B* amber mutants. They found that *Aam* mutants did not integrate to an appreciable level, confirming the previous studies on the necessity of the *A* gene product. *Bam* mutants lacking 66 or more amino-acids, integrated less than 5% of the infecting DNA under non-permissive conditions, however, *Bam* mutants producing a *B* protein truncated by 18 amino-acids, integrated their DNA at a normal level. Neither class of *Bam* mutants were able to replicate their DNA. The *B* protein is known to amplify Mu replication approximately 100 fold (Wijffelman and Lotterman, 1977) and this amplification was previously thought to be its major function. The study of Chaconas et al. (1985b) demonstrated that the *B* protein was also necessary for normal levels of Mu integration. The fact that the carboxyl terminal 18 amino-acids were dispensable for integration, but not replication, suggested an essential role for

these residues in replicative transposition. Chaconas et al.
(1985b) speculated that these residues may be involved in the
choice between the two transposition pathways utilized by Mu. This
choice could be made by a specific interaction between the carboxyl
terminal amino-acids of the B protein and a protein present only
during infection.

A further problem with bacteriophage Mu integration is the fact
that infecting DNA must be protected from host exonucleases as the
virion DNA is linear and non-permuted. Ljungquist and Bukhari
(1979) utilized sucrose gradients to analyze the fate of ^{32}P
labeled Mu DNA infecting unlabeled host cells. They observed that,
in both lysogens and sensitive cells, the infecting DNA was
converted to a form that sedimented twice as rapidly as linear Mu
DNA. In the sensitive host this fast sedimenting form was transient
and at a lower level than in the lysogenic infections, presumably
because, in the lysogens superinfecting Mu DNA could not integrate.

Chase and Benzinger (1982) observed that a small fraction of Mu
DNA, isolated from freeze-thawed virions, had a 65 KDa protein
bound noncovalently to the ends. They further demonstrated that the
protein-DNA complex was capable of transfecting E. coli
spheroplasts 2-3 orders of magnitude more efficiently than
deproteinized Mu DNA. Finally, they showed that one function of the
protein bound to the ends of the Mu DNA was to protect the
infecting DNA from exonucleases,

Harshey and Bukhari (1983) demonstrated the presence of a 64 kDa protein attached to the ends of infecting Mu DNA. The protein-DNA complex was observed to be transient in Mu sensitive *E. coli* cells, but, could accumulate and persist in cells lysogenic for Mu. The DNA in this protein-DNA complex was resistant to *in vitro* exonuclease degradation by both a 3' and a 5' exonuclease, showing that both strands of the infecting Mu DNA were protected by the 64 kDa protein. In addition, Harshey and Bukhari repeated the experiments of Chase and Benzinger (1982), and demonstrated that the 64 kDa protein Mu DNA complex was able to transfect *E. coli* several orders of magnitude more efficiently than naked Mu DNA.

Puspurs et al. (1983) infected *E. coli* minicells with Mu phage and also isolated a 64 kDa protein-Mu DNA complex. This study contained some important findings about this infection complex. First, this study observed that all the infecting DNA was in the form of the protein-DNA complex. Second, they demonstrated, by virtue of using ^{35}S labeled phage, that the protein was injected along with the infecting DNA. Third, they showed that the protein-DNA complex did not contain the Mu gp*H* protein. This was relevant as gp*H* was the only Mu gene product of 64 kDa known at the time. Fourth, this study showed that there was only one 64 kDa protein in the virion. Finally, they demonstrated that no protein synthesis was required for the production of the protein-DNA complex.

The experiments described above form the starting point for the work reported in this thesis. In Chapter 2, both published (Gloor and Ghaconas, 1986) and unpublished data that assign the gene for the Mu 64 kDa protein as the Mu *N* gene product, are presented. Chapter 3 describes the determination of the sequence of both the Mu *N* and *P* genes. This chapter also describes the overproduction of each gene product to a level of greater than 5% of total cell protein. Chapter 4 presents analysis of the transposition of the gp*N*-Mu DNA complex *in vitro*, and shows that the strand transferred product produced by this complex resembles a typical Shapiro structure.

CHAPTER 2

THE INJECTED, 64 kDa Mu VIRION PROTEIN

IS THE Mu *N* GENE PRODUCT

INTRODUCTION

Ljungquist and Bukhari (1979) demonstrated the presence of a rapidly sedimenting fraction of Mu DNA isolated by sucrose gradient sedimentation of *Escherichia coli* cells lysed soon after infection. This fraction was much more abundant when Mu lysogens, rather than sensitive cells, were infected. The DNA in the fraction was observed to be linear when purified. They proposed that the rapidly sedimenting fraction was an intermediate in the integration of Mu into the chromosome.

Chase and Benzinger (1982) found a 65 kDa protein bound noncovalently to the ends of a small fraction of Mu DNA isolated by freeze-thaw disruption of virions. This protein-DNA complex sedimented slightly faster in sucrose gradients than deproteinized Mu DNA. They observed that transfection of *E. coli* spheroplasts was greatly stimulated by the presence of the protein, and postulated that one of the functions of this protein was the protection of the Mu ends prior to integration.

Harshey and Bukhari (1983) more fully characterized the infecting protein-DNA complex. They purified the rapidly sedimenting DNA from Mu infected Mu lysogens and showed that there was a 64 kDa protein bound non-covalently to the ends of the infecting DNA. They further demonstrated that the purified protein-DNA complex was resistant to both 5' and 3' exonucleases. This group repeated the transfection experiments of Chase and

Benzinger (1982) and found that their protein-DNA complex was capable of transfecting *E. coli* spheroplasts several orders of magnitude more efficiently than deprote..nized Mu DNA.

Puspurs et al. (1983) purified a 64 kDa protein-DNA complex from Mu infected minicells. This protein-DNA complex was less stable than Harshey and Bukhari's, probably because the minicells were lysed by 0.1% SDS. The study of Puspurs et al. demonstrated that there was only one 64 kDa protein in the virion and that this protein did not correspond to any known Mu gene product. This study also demonstrated that no protein or RNA synthesis was required for the production of the complex. The rapid sedimentation was a consequence of supercoiling of the infecting DNA in this complex, since inhibitors of DNA gyrase abolished its formation.

The experiments described above all concluded that there was a protein-DNA complex involved in Mu integration. The protein in this complex was injected along with the infecting DNA into the cell, protecting it from degredation. Mu DNA is always found covalently associated with the DNA of the last host that it infected (Bukhari et al., 1976). Therefore, the Mu DNA in this protein-DNA complex was expected to be in a similar conformation to that of lysogenic Mu DNA; that is, supercoiled and surrounded by host DNA. Based on the studies outlined above, we reasoned that the conservative, simple insertion of Mu DNA into the chromosome during integration (Liebart et al., 1982; Akroyd and Symonds, 1983; Chaconas et al., 1983; Harshey, 1984), might be a consequence of having the 64 kDa

protein bound to the ends of the infecting Mu DNA. As a first step to understanding the role of this protein in Mu integration it was necessary to identify the gene coding for the protein. Results presented in the following chapter shows that the 64 kDa protein is the product of the Mu *N* gene. Most of the work described has been published (Gloor and Chaconas, 1986) previously.

MATERIALS AND METHODS

MEDIA AND CHEMICALS

LB broth was prepared as described by Miller (1972) except that N-Z amine (Scheffield Products) was substituted for Bactotryptone. LB agar plates were supplemented with glucose to a final concentration of 0.1% and with $MgSO_4$ and $CaCl_2$ to final concentrations of 2.5 mM and 1 mM respectively. Strains containing plasmids were always grown in the presence of 100 μg/ml of ampicillin in liquid and 100 μg/ml of carbenicillin on plates.

Chemicals for polyacrylamide gels and western blotting were obtained from Bio-Rad except for SDS* (BDH). Tris was from Sigma or Terrochem, glycine and the antibiotics were purchased from Sigma. All other chemicals were obtained from Baker or BDH or as noted.

BACTERIA AND BACTERIOPHAGE STRAINS

The bacterial strain C600 (*thr, leu, fhuA, glnV, gal, lacY, rpsL*) carrying a λc+ prophage and pCQV2 was used as the source of the cloning vector (Queen, 1983). MM294 (*endA, thi, hsdR, glnV*), CA275 (Hfr, *lacZam, trp, rel, spoT, tyrT*) and an isogenic Su⁻ strain CA274 were obtained from the *E. coli* genetic stock center. Bu5029 (Δpro⁻ *lac, trp, recA, rpsL*) and the isogenic *rec*+ strain 40 were from the late Dr. A. Bukhari's (Cold Spring Harbor Laboratory) strain collection. P678-54 (*thr, leu, minA, minB, thi*

ara, lacY, gal, malA, xyl, mtl, rpsL, fhuA, azi Adler et al,
1966) was acquired from Dr. D. Denhardt (University of Western
Ontario). The bacteriophage strain Muc*ts*62 and the Muc*ts*62 amber
mutant phage; *Yam*1001, *Yam*1027, *Yam*1034, *Nam*7365, *Nam*1995, *Nam*7106,
Pam 1008, and *Pam*1012 were obtained from Dr. M. Howe (O'Day et al.,
1979). Muc*ts*62amp is the mutant isolated by Leach and Symonds and
is maintained as a lysogen in GC103 (Δ*pro-lac, leu, met,* Mu^R,
D108^R).

PREPARATION AND ANALYSIS OF NUCLEIC ACIDS

Restriction enzymes were purchased from Bethesda Research
Laboratories (BRL), New England Biolabs (NEB), Pharmacia-PL (PPL)
and Beohringer-Mannheim (BM). All restriction digests were
performed using the suppliers recommended reaction buffer, except
where noted. All other enzymes were also used in the manufacturers
suggested buffer, except where noted. Plasmids were routinely
prepared by the method of Birnboim and Doly (1979) and were
extracted with an equal volume of ∅–CHCl$_3$ (phenol-chloroform-
isoamyl alcohol 25:24:1; equilibrated with 0.5 M Tris-HCl pH 7.8).
Bacteriophage DNA was prepared as previously described (Chaconas et
al 1981b, 1983). One unit of T1 ribonuclease (BRL) was routinely
added to restriction digests of plasmids to degrade RNA. T4 DNA
ligase was from PPL and ligations were carried out in the
recommended buffer.

E. coli DNA polymerase 1 was purchased from BM or BRL. When 5'
overhanging ends derived from restriction of the DNA were to be
filled in, 1 unit of the Klenow fragment and 50 pMole each of the 4
dNTP's were added directly to the restriction reaction, and
incubated at room temperature for 15 min.

When further manipulations were to be done with the restricted
plasmids the reactions were terminated by the addition of an equal
volume of ϕ-CHCl$_3$. Ethanol precipitation, done by adding 1/5 volume
10 M ammonium acetate pH 6.0, or 3 M sodium or potassium acetate pH
5.0 and 2 volumes ethanol, was done to remove the ϕ-chloroform and
to concentrate the DNA. When restriction digests were to be
analyzed by gel electrophoresis, the reactions were terminated by
the addition of 1/9 volume of 10X EDTA loading dye (60% sucrose,
200 mM EDTA, 0.01% bromophenol blue), A 10X SDS loading dye (1%
SDS, 60% sucrose and 0.01% bromophenol blue) was used where noted.
Gel electrophoresis of nucleic acids was carried out according to
standard techniques in TAE buffer (Maniatis et al., 1982).

All transformations were by the method of Hanahan (1983), using
cultures in mid to late log phase, in the recommended RbCl
containing buffer. Most strains gave transformation frequencies of
>1 X 10^8 transformants/ug of pUC9, except AA102 which was about 20
fold less efficient with this technique.

CONSTRUCTION OF Mu LIBRARY IN pCQV2

The cloning vector pCQV2 was linearized with BamH1, extracted with φ-CHCl$_3$ and ethanol precipitated. The vector DNA was subsequently dephosphorylated with bacterial alkaline phosphatase, phenol extracted two times and ethanol precipitated twice. Mu DNA was partially digested with Sau3A until the average size of the fragments was about 5 kbp as determined by agarose gel electrophoresis. Ligation was carried out in a final volume of 15 ul at 15° overnight. The reaction contained 2 Weiss units of T4 DNA ligase, 1 µg of vector DNA and 2 µg of Mu DNA. MM294 was made competent by the procedure of Hannahan (1983) and transformed with the ligated DNA. The library, which contained 50% inserts as judged by colony hybridization with ^{32}P labeled Mu DNA (Grunstein and Hogness, 1975), was screened with antiserum against the virion 64 kDa protein by a procedure similar to that of Helfman et al. (1983). Briefly, cells were plated to a density of about 3 000/plate and grown at 32° overnight. The cells were lifted off the plate onto nitrocellulose discs (Nucleopore) and grown at 43° for a further 12 hours to induce protein production. The cells were exposed to chloroform vapor for twenty minutes and to lysis buffer containing 2 µg/ml of DNase I (Sigma) and 80 µg/ml lysozyme (Sigma) for 1 hour. The nitrocellulose discs were then incubated with a 1:100 dilution of preabsorbed antiserum overnight and the filters were washed and processed as described under western blotting.

PREPARATION OF BACTERIOPHAGE FOR ANTIGEN PRODUCTION

Mu phage was grown as previously described (Chaconas et al., 1983). Phage were purified from lysates by a modification of a procedure described in Silhavy et al. (1984) for purifying phage lambda. Briefly, 40 liters of GC103 was grown in a fermentor and induced to make phage by thermal shift. The lysate was collected, cooled on ice, and stored overnight on ice. Bacterial debris was removed by centrifugation for 10 minutes, in a Beckman JA-10 rotor, utilizing a Beckman J2-21 centrifuge. The phage were concentrated using a Pellicon concentrator (Durapore filter, 100 000 molecular weight cutoff) to a final volume of <1 liter. The phage were pelleted from this concentrated lysate by centrifugation in a Beckman type 19 rotor, at 18 000 rpm, for 2 hours. The pellets were overlayed with 5 ml of Mu buffer (150 mM NaCl, 20 mM Tris-HCl pH 7.5, 20 mM $MgSO_4$, 2 mM $CaCl_2$) and left overnight at 4°. The pellets were resuspended in Mu buffer and loaded onto a step gradient (top to bottom: 15 ml of 15% glycerol, 25 ml 40% glycerol, 5 ml of CsCl (BDH), density =1.5g/ml, all made up in Mu buffer) prepared in Beckman 45Ti tubes. The step gradients were centrifuged at 100 000 X g for 1 hour at 20°. The phage, found at the 40% glycerol-CsCl interface, were removed and centrifuged to equilibrium in a solution of CsCl (ρ=1.5 g/ml) in Mu buffer at 85 000 X g for 24 hours in a Beckman SW 41 rotor, at 20°. The phage were recovered by side puncture, pooled, and recentrifuged to equilibrium to further concentrate them. The concentrated phage were dialyzed against two changes of Mu buffer and stored at 4°. The concentration of the

phage was estimated by absorbance at 260 nM and corrected for light scattering at 320 nM (Freifelder, 1982). The concentration of phage was usually near 0.5 nMol/ml.

PREPARATION OF ANTIBODY TO THE 64 kDa PROTEIN

The protein was isolated from whole phage particles by three serial preparative SDS-polyacrylamide gel electrophoresis runs. Briefly, about 500 picomoles of phage particles were denatured for 5 minutes at 100° in loading buffer and loaded in a single gel sized well on a 15 by 15 by 0.2 cm, 10 % polyacrylamide gel with a four percent stacking gel (Laemmli, 1970). After electrophoresis, the protein bands were stained with coomassie blue R-250 and then destained in 10% acetic acid for 2-4 hours. The 64 kDa band was cut out of the gel and equilibrated with several changes of loading buffer until bromophenol blue, used as an indicator dye, no longer turned yellow. This indicated that most of the acetic acid had been washed out. The gel slices from 8 gels were cut into small cubes and loaded onto one 8% gel with a 4% stacking gel. The gel was run at 50 V overnight to ensure that proper stacking occurred. The protein band was excised and run on a third gel to eliminate trace amounts of contaminating proteins. Usually about 100 to 200 µg of 64 kDa protein were obtained from eight first dimension gels.

The gel slices containing the purified 64 kDa protein were equilibrated with water and then forced through a series of needles

until the suspension would easily pass through one which was 21 ga. The suspension was then mixed with an equal volume of Freund's complete adjuvant (Gibco) for the first injection and an incomplete adjuvant for the succeeding injections. A 1.5 Kg, female, New Zealand White rabbit was injected with a total of 100 μg of 64 kDa protein, using 6 injections in the thigh and calf muscles. Thirty-four days later the rabbit was injected at several locations intradermally, with a total of 75 μg of protein. After a further thirty-one days, the rabbit was boosted with a total of 50 μg of 64 kDa protein intradermally, and 14 days later the rabbit was sacrificed by heart puncture, from which the serum was prepared. The titre of anti-64 kDa antibody did not change significantly over the last 14 days, as measured by probing western blots of total phage proteins. The serum seemed to contain many antibodies which bound to E. coli proteins when western blots were performed. These antibodies were found to be present in serum from the pre-immune rabbit. The high background problems were minimized by incubating 20 mls of 1:100 diluted serum with 3×10^{10} E. coli cells that had been boiled for 15 minutes in 0.1 ml of water. The entire mixture of heat killed cells and antiserum was used to probe the blots.

PREPARATION OF ^{35}S-METHIONINE LABELED Mu PHAGE AND MINICELL INFECTION

Strain GC103 was grown in LB broth to a density of 9×10^8 cells per ml at 32° and then shifted to 43° for 18 minutes. The culture (25 ml) was filtered and washed on a nitrocellulose

membrane with a small volume of 0.85% NaCl. The cells were
resuspended in 25 ml of labeling medium containing 1X M9 medium
(Davis et al., 1980), the 18 non-sulfur containing amino acids at
the concentrations recommended by Davis et al. (1980), and
dephosphorylated yeast extract (Chaconas et al., 1985) at 2% of
the concentration found in LB broth. Then, 1.7 mCi of [^{35}S]-met
(NEN >800 Ci/mM) was added to the culture and incubation at 43° was
continued for another 10 minutes. The culture was shifted to 38°
until lysis was complete. The phage were pelleted and resuspended
in Mu buffer as previously described (Chaconas et al., 1983),
except that the Mu buffer lacked calcium. The phage were
chromatographed on a Sepharose CL-4B (Pharmacia) column (1 X 30 cm)
equilibrated with calcium-deficient Mu buffer to remove any
contaminating proteins. The recovered phage (95-100%) were titered
and used for infecting minicells.

Minicell infections were done exactly as described by Puspurs
et al. (1983), except that the cells were lysed by two cycles of
freezing and thawing followed by the addition of Triton X-100 to a
final concentration of 0.5% and two more cycles of freeze-thaw. The
minicell lysates were cleared by centrifugation in an Eppendorf
centrifuge for 10 minutes prior to sucrose gradient sedimentation.
The DNA-containing peaks were followed by the inclusion of a small
amount of ^{32}P-labeled Mu phage (Chaconas et al., 1985b) in the
infection experiments. We found that the complex isolated in this
manner was as salt-stable as the complex isolated by Harshey and
Bukhari (1983) and much more stable than the complex of Puspurs et

al. (1983; data not shown). This is probably because Puspurs et al.
(1983) lysed their minicells with SDS rather than a non-ionic
detergent.

SDS-POLYACRYLAMIDE GEL ELECTROPHORESIS AND WESTERN BLOTTING

Polyacrylamide gels for blotting were always 10% gels with a 4%
stacking gel (Laemmli, 1970) and were run at constant voltage
overnight. Protein bands were transferred to Schleicher and Schuell
nitrocellulose using a Bio-Rad Trans Blot cell according to the
manufacturers instructions (Towbin et al., 1979). The blot was
washed with TBS (150 mM NaCl, 20 mM Tris-HCl pH 7.5) + Tween 20
(Bio-Rad; Batteiger et al., 1982) to remove acrylamide. Twenty ml
of 1:100 diluted antiserum in the above buffer was preabsorbed with
3×10^{10} heat killed E. coli cells for 2-4 hours on ice. Blots were
incubated overnight at room temperature and washed 3 times with TBS
+ Tween 20. Blots were subsequently incubated, for 2 hours, with
goat antirabbit peroxidase conjugated antiserum (BM; 1:2 000
dilution in TBS + Tween 20). The second antibody was washed off by
three washes in the same buffer. The blots were washed once in TBS
and developed using 4-chloro-1-napthol (Sigma; 75 mg in 25 ml
methanol) and H_2O_2 (Baker, 50 µl in 100 ml H_2O). Alternatively,
goat antirabbit phosphatase conjugated second antibody (Sigma;
1:500 dilution) was used. This antibody was developed using 1 mg/ml
α-napthol phosphate (Sigma) and 1 mg/ml fast blue RR (Sigma) in 50

mM Tris-HCl pH 8.0, 100 mM NaCl, and 5 mM $MgCl_2$ (Dr. S. Dunn, personal communication).

RESULTS

PREPARATION OF ANTISERUM

The 64 kDa protein was purified and antiserum to it was prepared as described in Materials and Methods. A comparison of the rabbit's preimmune and immune serum is shown in Fig. 21. The preimmune serum did not bind to the 64 kDa protein whereas the immune serum reacted to give a strong signal. Both the preimmune serum and the immune serum bound to a minor extent with the major tail protein. In addition, both sera contained antibody capable of binding many *E. coli* proteins on western blots (not shown). Antibodies to the *E. coli* proteins were reduced by preabsorption of the sera prior to use by an incubation of the diluted sera with heat killed *E. coli* cells. After preabsorption of the 64 kDa antiserum, it was suitably specific for both screening of the recombinant clones and for western blots of phage infected cells.

PREPARATION OF THE Mu LIBRARY IN pCQV2

A library of *Sau*3A restriction endonuclease cut Mu DNA was cloned into the expression vector pCQV2. This vector expresses a temperature sensitive λcI857 repressor which regulates the expression of the λ rightward promoter, P_R, which is also carried on the plasmid. Downstream of the promoter sequence there is a unique *Bam*HI site for insertion of foreign DNA (Queen, 1983). The *Bam*HI sticky ends are complementary to the *Sau*3A cut Mu DNA,

```
            A   B          C

92.5-

66.2-        —  ←        →

45.0-

31.0-

21.5-

    Immunodetection      Coomassie
                         Blue Stain
```
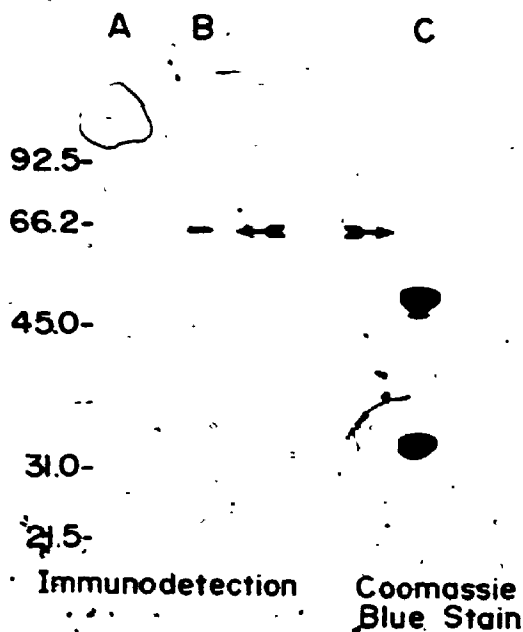
FIGURE 21: SPECIFICITY OF ANTIBODY TO THE 64 kDa PROTEIN

Hu virion proteins were run on a 10% SDS-polyacrylamide gel and blotted onto nitrocellulose. After probing the blots were developed with a peroxidase conjugated second antibody as described in Materials and Methods. The blots were probed with a 1/100 dilution of preimmune serum (A) or immune serum (B). Molecular mass marker positions are shown on the left and a coomassie blue stain of the virion proteins is also shown (C). The 64 kDa protein is indicated by the arrows. We estimate that this polypeptide is present in about 3 copies per phage particle based on staining intensity compared with the major tail protein (gpL) which is present at 192 copies per virion (Admiraal and Mellema, 1976).

allowing for efficient ligation. At 32° the promoter is repressed
by the active repressor. Expression of genes cloned into the *Bam*HI
site of the vector can be induced by shifting the temperature of
the bacterial culture to 43°.

Screening of the recombinant DNA library was performed as
described in Materials and Methods. Thirty-six independent clones
were isolated and rescreened by the above procedure; several of
these clones gave a strong signal when they were screened. These
clones were grown overnight and the sizes of the recombinant
plasmids were analyzed by cell lysis and direct agarose gel
electrophoresis of the mini-lysates (Chaconas et al., 1981a). Four
clones were isolated (pGG409, pGG413, pGG417, pGG418) that
contained plasmids with insert sizes between 2 and 5 kilobases.
Protein synthesis directed by the promoter, P$_R$, was induced in
these strains by shifting the temperature of the cultures to 43°
for two hours. The proteins were run on a 10% SDS-polyacrylamide
gel and transferred onto nitrocellulose. The results of this
experiment are shown in Fig. 2.2. All four of the positive clones
produced a protein that reacted with the antiserum to the 64 kDa
protein. Furthermore, in all cases the protein produced by the
plasmids had the same apparent molecular mass as the 64 kDa protein
found in the Mu virion. This protein was not observed upon
induction of the vector, pCGV2. In comparison with the amount of
this protein produced during lytic development of Mu (see Fig.
2.6), an increase in expression of at least 10-fold was observed in

1 2 3 4 5 6

92.5-

66.2-

45.0-

31.0-

21.5-
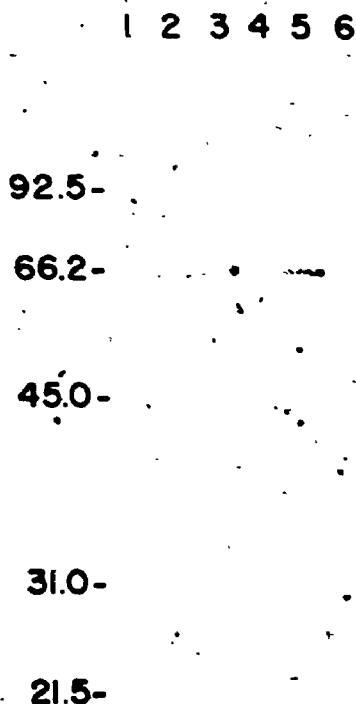
FIGURE 22: IMMUNOBLOT ANALYSIS OF POLYPEPTIDES PRODUCED BY POSITIVE CLONES

Bacterial strains containing pGG409, pGG413, pGG417, pGG418 or the vector pCQV2 were grown in LB broth containing 100 ug/ml ampicillin at 32° to a density of about 5×10^8 cells/ml. The cultures were then shifted to 43° and a sample was removed from each culture 2 1/2 hours after temperature shift. The cells were pelleted, resuspended and denatured in loading buffer, and run on a 10% SDS-polyacrylamide gel. The proteins were transferred onto nitrocellulose, probed with antibody to the 64 kDa protein and developed with a second antibody as described in Materials and Methods. Lane 1 contains total virion proteins, lanes 2-5 contain samples from the cultures harboring plasmids pGG409, pGG413, pGG417, pGG418 and the vector, pCQV2, respectively. Molecular mass marker positions are shown on the left. The intensely staining band with an apparent molecular mass of 64 kDa is the protein of interest.

the recombinant clones. We could not, however, detect the
overproduced 64 KDa protein by coomassie blue staining.

IN VITRO ANALYSIS OF PROTEIN PRODUCTION

The purified plasmids were also tested for their ability to
stimulate production of peptides in an *E. coli* coupled, *in-vitro*
transcription-translation system (Zubay, 1973; performed by Dr.
George Mackie of this department). The results of this experiment
are shown in Fig. 2.3. The only detectable polypeptides common to
all the recombinant plasmids were the 64 KDa polypeptide and
β-lactamase. The largest polypeptide ran as a dimer of 64 and 62
KDa. The dimer bands could not be made to coelectrophorese when
the amount of sulfhydryl reducing agent was varied, indicating that
the migration of the bands was not a property of their sulfhydryl
oxidation state. In the absence of suitable labeled molecular mass
markers, a portion of the protein produced *in vitro* from pGG418 was
coelectrophoresed with an induced-cell lysate from a strain
containing pGG418. The protein was blotted onto nitrocellulose and
probed with the anti-64 KDa protein antiserum. Subsequently the
blot was autoradiographed. The upper band on the autoradiograph
coelectrophoresed with the 64 KDa protein produced *in vivo*. Under
no conditions have I observed the 62 KDa band on western blots of
induced cell cultures.

The production of other gene products varied with the
recombinant plasmid used as the template. Plasmid pGG413 and pGG409
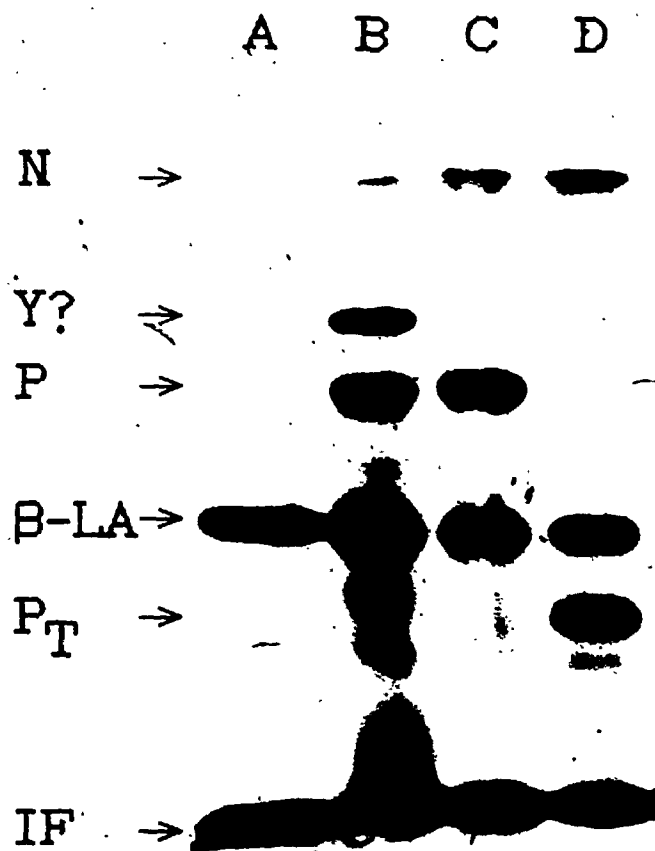
FIGURE 2.3: PROTEINS PRODUCED FROM RECOMBINANT PLASMIDS *IN VITRO*

The recombinant plasmids were used as templates for an *in vitro* coupled transcription-translation system (Zubay, 1973), and was performed by Dr. G. Mackie of this department. [^{35}S]-methionine was the labeled amino-acid. Proteins from the reaction were precipitated by the addition of 2 volumes of 10 mM β-mercaptoethanol and 15 volumes of acetone, followed by incubation at -70° overnight. The precipitate was collected by centrifugation and the pellet was resuspended in 100 ul of SDS-acrylamide gel loading buffer. Proteins were denatured by boiling for 5 minutes and loaded onto a 10% SDS-acrylamide gel with 4% stacking gel (Laemmli, 1970), the gel (20 cm) was run for 16 hours at 50 v. The gel was dried autoradiographed. Lane A shows the protein profile derived from the vector, pCQV2, lanes B-D shows the proteins produced by the recombinant plasmids, pGG409, pGG418, pGG417, respectively. The migration of the gene products; gp*N* (N), gp*P* (P) and β-lactamase (β-LA), are indicated on the left. The ion front is also indicated (IF). Y? indicates the presumed partial *Y* gene product. P$_T$ indicates the truncated *P* gene product.

were observed to produce identical protein profiles by this
analysis. Plasmid pGG417 did not produce a 43 KDa band
corresponding to the Mu F gene product (Giphart-Gassler et al.,
1981). Rather, this plasmid generated a smaller band of ≈28 KDa. In
addition, pGG409 and pGG413 generated a polypeptide product of
≈49 kDa. These plasmids have a substantial amount of DNA upstream
of the other two plasmids and rescue some, but not all, Yam mutants
(see below), indicating that the ≈49 kDa polypeptide may correspond
to a truncated Y gene product.

PHYSICAL MAPPING OF RECOMBINANT PLASMIDS

In order to determine the limits of the cloned DNA from each
plasmid on the physical map of Mu, the plasmids were labeled with
^{32}P by nick translation and hybridized to Southern blots of
restriction endonuclease cleaved Mu DNA (Chaconas et al., 1980).
The results of this experiment are summarized in Fig. 2.4 and show
that the plasmids all contained Mu DNA from the Y, N and P gene
regions (Schumman et al., 1980). Plasmids pGG409 and pGG413
hybridized identically to all the restriction fragments and both
are referred to below as pGG409. Plasmids pGG409, pGG417 and pGG418
hybridized with the same efficiency to the restriction fragments
represented by the thick lines in Fig. 2.4, and did not hybridize
to the fragments represented by the thin lines. The restriction
fragments represented as open boxes hybridized with differing
efficiencies to the recombinant plasmids. Plasmids pGG417 and
pGG418 hybridized less efficiently to the large $HindIII$-$SalI$
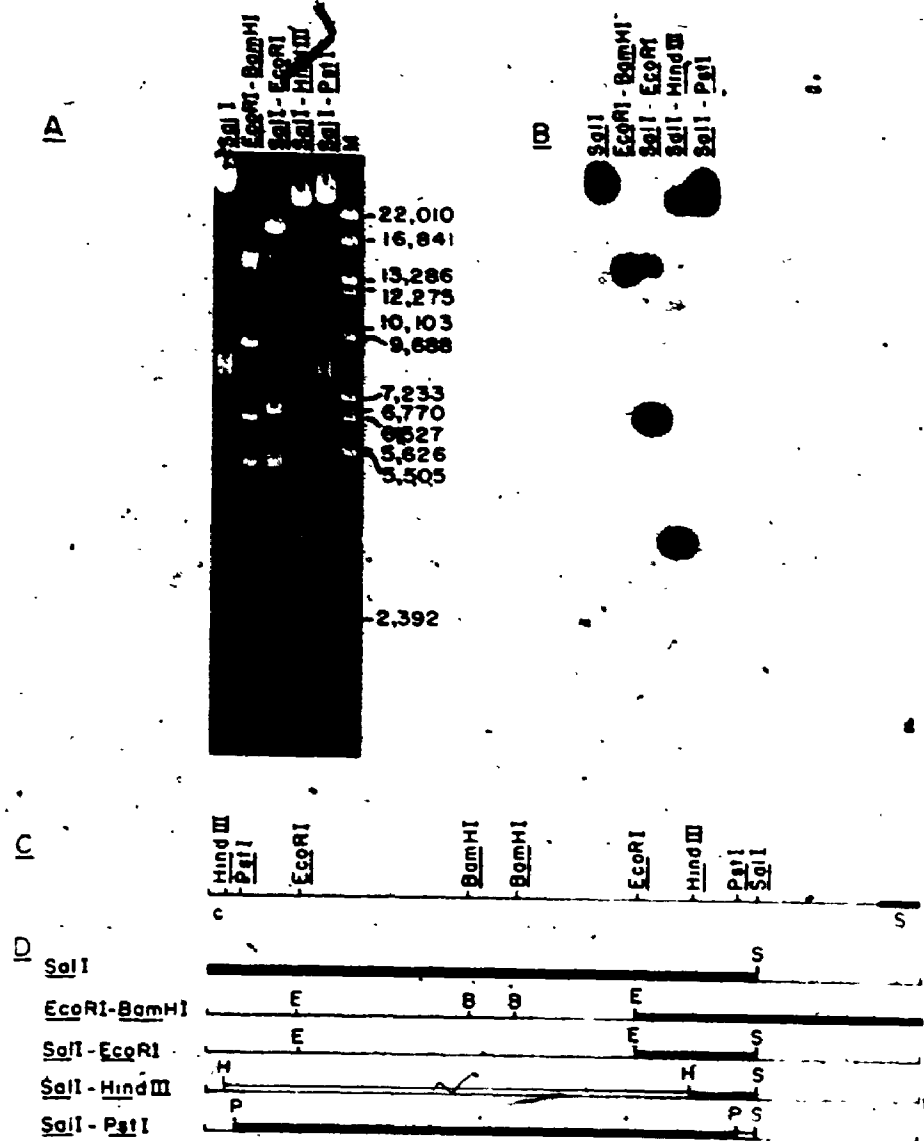
FIGURE 2.4: PHYSICAL MAPPING OF CLONES PRODUCING THE 64 kDa PROTEIN

Mu DNA was cleaved with the restriction enzyme(s) indicated and run on a 0.5% agarose gel. An ethidium bromide stained gel is shown in panel A. The marker lane (M) contains a mixture of BamHI and BgIII cut λ DNA, and the size of each fragment is shown in bp. The Mu DNA Southern blotted and hybridized to ^{32}P labeled pCQV2, pGG409, pGG413, pGG417 or pGG418 as previously described (Chaconas et al, 1980). An autoradiogram of restriction enzyme cut Mu DNA hybridized with pGG417 is shown in panel B. A physical map of the Mu genome, with the pertinent restriction sites, is shown in C (Marrs and Howe, 1983). A summary of the hybridization data is shown in D. The restriction fragments that did not hybridize to the plasmids are shown as thin lines, the thick lines represent restriction fragments which hybridized to all of the plasmids and the open boxes represent restriction fragments which hybridized with varying efficiency to the 4 plasmids.

restriction fragment than did pGG409. The lower hybridization
efficiency indicated that these plasmids contained less homologous
DNA in this region of the Mu genome as compared to pGG409. Plasmid
pGG417 did not hybridize to the 1.6 kbp *PstI-SalI* restriction
fragment, however, pGG409 and pGG418 did hybridize to the fragment.
This indicated that pGG417 did not contain any detectable DNA
sequences from this restriction fragment. The vector, pCQV2, did
not hybridize to any Mu DNA restriction fragments. These results
indicate that the four recombinant plasmids contain the ≈2 600 base
pairs of DNA between the rightmost *HindIII* and *PstI* restriction
sites of the Mu genome as common regions. This area must therefore
contain most, if not all, of the Mu DNA sequences required to code
for the 64 kDa protein.

A more precise physical analysis of the plasmids was carried
out by restriction endonuclease mapping. The data are summarized in
Fig. 2.5. These results are in agreement with, and supplement the
hybridization data shown in Fig. 2.4. The similar plasmids pGG409
and pGG413 could not be resolved by restriction mapping and have
been assumed to be identical or nearly so.

The plasmid pGG417 was observed to behave, when restriction
mapped, as if ≈150 bp of DNA had been deleted adjacent to the *PstI*
site near the right end of the insert. Sequencing of the Mu *N* and
P genes (Chapter 3) revealed two *Sau3A* sites in the cloned DNA. One
of these sites was about 540 bp upstream of the *PstI* site. DNA
sequence upstream of the *PstI* site was compared to known sequences

- λ Sequences. ▬▬ pBR322 Sequences. ——— Mu Sequences. ▼ - BamHI - Sau3A junction
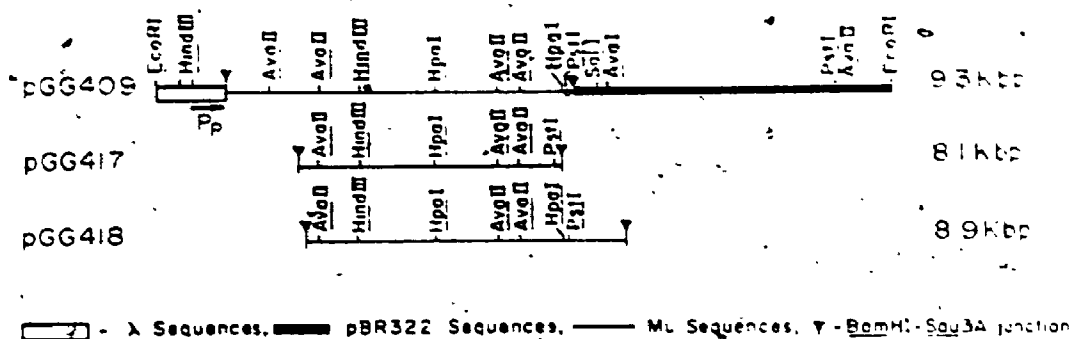
FIGURE 2.5: RESTRICTION MAP OF DERIVED PLASMIDS

Plasmids pGG409, pGG413, pGG417 and pGG418 contain Sau3A fragments of Mucts62amp cloned into the BamHI site of pCGV2. Plasmid pGG413 was found to be identical to pGG409 by restriction mapping and is not shown. The thin lines represent Mu derived DNA, the thick line shows pBR322 derived DNA and the open box shows derived DNA sequences. The arrow beneath the sequence shows the location and orientation of the inducable promoter, P_R. Plasmids pGG417 and pGG418 are shown without the attached pBR322 sequences and are aligned at the HindIII site found in all the recombinant plasmids. The derived restriction maps are in agreement with the previously published map of this region (Marrs and Howe, 1983), except for the deletion in pGG417.

in the NIH sequence databank, and it was found that at least part of the sequence was complementary to the sequence in the Mu G region. Since the Mu library was produced by random cleavage with Sau3A, it is probable that one or more pieces of DNA, not from the Mu N and P gene regions, were ligated to this site prior to joining the vector.

GENETIC MAPPING

A genetic analysis of the plasmids was carried out using complementation and marker rescue experiments (Table 2.1). Lawns of isogenic recA and rec⁺ E. coli, harboring the plasmids pGG409, pGG417, pGG418 or pCGV2, were spotted with lysates from Mu phage containing amber mutations in the essential Mu genes: H, N, P, and Y. These genes were known to map in the general vicinity of the cloned fragments (Schumman et al., 1980). As expected, clearing was not detectable in the absence of a complementing plasmid and marker rescue was observed only in a rec⁺ background. The data indicated that the only complementing gene common to all the plasmids was the N gene, suggesting that it might encode the virion 64 kDa protein.

Plasmid pGG417, which contains the smallest insert, was found to complement only the N gene. This plasmid could also rescue some, but not all, mutations in genes Y and P on spot tests. Whole plate assays were done for the Yam and Pam mutants that could not be rescued by spot tests. Using whole plate assays, the Pam mutation in deletion group 65 was rescued at a level of about 500 times

TABLE 2.1: GENETIC MAPPING OF RECOMBINANT CLONES

STRAIN	RELEVANT GENOTYPE	PLASMID	AMBER MUTANT/DELETION GROUP					
			Y1027 53	Y1001 54	Y1034 57	N7365 58	F1008 62	F1012 65
CA274	Su-		-	-	-	-	-	-
CA275	Su+		+	+	+	+	+	+
BU5029	recA	pCQV2	-	-	-	-	-	-
40	rec+	pCQV2	-	-	-	-	-	-
BU5029	recA	pGG409	-	-	-	+	+	+
40	rec+	pGG409	-	x	x	+	+	+
BU5029	recA	pGG417	-	-	-	+	-	-
40	rec+	pGG417	-	-	x	+	x	o
BU5029	recA	pGG418	-	-	-	+	+	+
40	rec+	pGG418	-	-	x	+	+	+

- indicates no marker rescue or complementation
+ indicates complementation
x indicates marker rescue observed on spot tests
o indicates marker rescue observed on whole plate assays only

Complementation and marker rescue experiments were performed by
spot tests; amber mutants not rescuable by spot tests were
re-examined by whole plate assays. Mu phage lysates were prepared
with titers on the Su+ indicator strains between 3×10^7 and 7×10^9
pfu/ml. For spot tests, phage dilutions spanning 4 orders of
magnitude were spotted onto lawns of the desired isogenic strain
pairs. Duplicate plates were incubated at 32, 37 and 43°. At 32°,
complementation was observed with MuPam mutants but not MuNam or
MuYam mutants. The results for 37 and 43° were identical and are
shown in the table. Whole plate assays were done by plating 3 serial
dilutions of the phage lysates, starting with a 10^{-2} dilution. The
plates were incubated at 37° overnight and scored. This allowed us to
observe recombinants occurring with a frequency of recombination
greater than 2.3×10^{-7} of input pfu's. Marker rescues that were
observed on whole plate assays but not spot tests were found to
rescue at a level ≈500 fold reduced from marker rescues observed on
spot tests. Marker rescues were verified by picking individual
plaques and testing for efficient growth on both Su- and Su+ host
strains. No complementation or marker rescue was found for mutants
before deletion group 54.

lower than the level of the control marker rescues and about 5 times greater than the background. The whole plate assays of the Nam mutants which were tested, showed no detectable marker rescues. The markers that were rescuable by spot tests were in deletion groups 57 (Yam1034) and 62 (Pam1008), which directly flank deletion groups 58-61 within gene N (O'Day et al, 1979). We infer from the data that this plasmid carries a functional N gene and only part of the flanking Y and P genes.

As mentioned above, this plasmid carries a deletion of the P gene. This deletion probably begins about 540 bp from the PstI site and obliterates the complementation of Pam phage. It also results in the inability to rescue Pam1012, in deletion group 65, on spot tests but not on the whole plate assays. We, therefore, conclude that part-or all of this deletion resides in gene P and that the deletion resides very close to the P1012 amber mutation site. Titers of Nam phage on this plasmid or any of the other plasmids, except the vector, were indistinguishable from those found using a Su+ strain.

The genetic data in Table 21 establishes the left hand limits of all the recombinant plasmids and the right hand limit of pGG4{7, the smallest construct. The right hand limits of the larger plasmids were not determined since this had no bearing on the coding region of the 64 kDa protein. The two larger plasmids should carry at least some, if not all of gene Q, and perhaps gene W

sequences, since these genes follow gene F and map to the right of the PstI site (Schumman et al., 1980).

To unambiguously demonstrate that gene N encodes the 64 kDa virion protein, infection experiments were performed with MuNam phage mutants. Control experiments with Mucts62amp demonstrated that the 64 kDa protein could not be detected with the antiserum before 35 minutes after infection. The level of the 64 kDa protein peaked at 45 minutes post infection and remained steady until lysis (data not shown). This demonstrated that detectable quantities of the 64 kDa protein were not produced until late in the lytic cycle. Both Nam mutations (Nam1995, Nam7106) were from deletion group 61 (O'Day et al., 1979), which is the deletion group closest to the carboxyl-terminus of the protein. These mutations have been sequenced, as described in the Results section of Chapter 3. Upon infection of a Su$^-$ strain these mutant phage should express truncated forms of the 64 kDa protein. This was tested by infecting isogenic Su$^-$ (CA274) and Su$^+$ (CA275) strains and running the total cellular proteins at 50 minutes post infection on an SDS-polyacrylamide gel as previously described (Chaconas et al., 1985b). The proteins were transferred to nitrocellulose and probed with 64 kDa antiserum. The results of this experiment are shown in Fig. 2.6. Panels A and B are duplicate blots incubated with a peroxidase conjugated second antibody and an alkaline phosphatase conjugated second antibody, respectively. Lane M shows the migration of the full size 64 kDa protein as found in the virion. In the Su$^+$ strain (lanes 1 and 2), a 64 kDa protein can be
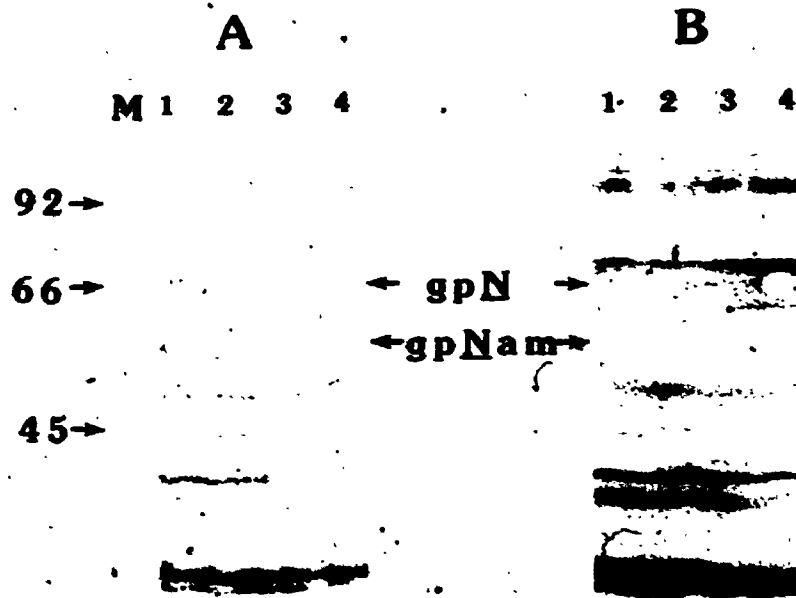
FIGURE 2.6: IDENTIFICATION OF gpN AS THE 64 kDa PROTEIN

Isogenic Su⁻, Su⁺ strains were infected at an m.o.i. of 3 with
Nam1995 and Nam7106 phage as described previously (Chaconas et al.,
1985b). Total cellular proteins were run on a 10% polyacrylamide gel,
transferred onto nitrocellulose and probed with antibody to the 64
kDa protein as described in Materials and Methods. Duplicate blots
were incubated with peroxidase conjugated second antibody or alkaline
phosphatase conjugated second antibody and are shown in panels A and
B respectively. Lane M shows the mobility of the wild type protein as
found in intact viral particles. Lanes 1 and 2 show the polypeptides
produced with the mutant phage in the Su⁺ strain. Lanes 3 and 4 show
the polypeptides produced in the Su⁻ strain. The migration position
and size of the molecular mass markers is shown on the left. The
positions of the intact and truncated forms of gpN are indicated by
the arrows. The low intensity of the bands was expected since the
intracellular concentration of the 64 kDa protein is predicted to be
low (100 phage per cell X 3 copies per phage = 300 copies per cell).

detected. In the Su⁻ strain this band disappears for both Nam
mutants and is replaced by truncated forms with molecular masses of
55 kDa and 54 kDa respectively (lanes 3 and 4). This establishes
the identity of gp$N$ as being the 64 kDa protein to which we raised
the antisera that was used for immunoscreening the Mu expression
library.

Several other points which are peripheral to the subject of
this paper are worthy of mention here. The first is that in
contrast to the results for gene N, complementation of gene P
occurred at 32°, suggesting the presence of a Mu-promoter between
the end of gene N and the start of gene P. Physical evidence for a
promoter in this region has been obtained by S.F. Stoddard and M.M.
Howe (personal communication; Howe, 1987). The second point is
that our genetic data, and the in $vitro$ translation data, are not
compatible with the previous assignment of gp$Y$ as a 12 kDa protein
(Giphart-Gassler et al., 1981). Plasmids pGG417 and pGG418 rescue a
Y mutation in deletion group 57 but not in deletion groups 53 or
54. The plasmid pGG409 has an additional 900-1 000 base pairs of
DNA beyond the left extremity of the inserts in pGG417 and pGG418.
This larger insert efficiently rescues the Yam mutation in deletion
group 54 but does not rescue the Yam mutation in deletion group 53.
This suggests that the product of the Y gene may be a polypeptide
of greater than 35 kDa. It is also noteworthy that we have observed
a 49 kDa polypeptide which may correspond to a portion of gp$Y$ in
the cell free transcription-translation system (Fig. 2.3).

DEMONSTRATION OF THE IDENTITY OF gpN WITH THE INJECTED VIRION PROTEIN

To prove the identity of the 64 kDa protein in the intracellular Mu protein-DNA complex as being the Mu N gene product, the 64 kDa protein from the intracellular complex and from gpN manufactured in-vitro were isolated and subjected to peptide mapping (Cleveland et al., 1977). The results of this experiment are shown in Fig. 2.7. Lanes 1, 2 and 3 respectively show the partial proteolysis profiles of the 64 kDa protein isolated from the intracellular protein-DNA infection complex, the virion 64 kDa protein and the N protein produced in-vitro. In each lane it is evident that the pattern is essentially the same. There is a slight difference between the profiles, in the low molecular weight region of the gel, in lane 3 compared to lanes 1 and 2; this may indicate some proteolytic decay of the protein produced in-vitro. Lane 4 is a control showing the partial proteolysis pattern produced using the Mu major tail protein as the substrate. It is evident that the pattern is entirely different than the one observed in lanes 1-3.

This experiment illustrates that there is only one polypeptide of 64 kDa in the Mu virion. If there were more than one 64 kDa protein in the virion, the partial proteolysis profile of this protein (lane 2) would not match that produced by the N protein produced in the in-vitro transcription-translation system (lane 3). From the experimental evidence presented above, we conclude that the Mu N gene encodes the 64 kDa virion protein, which is injected
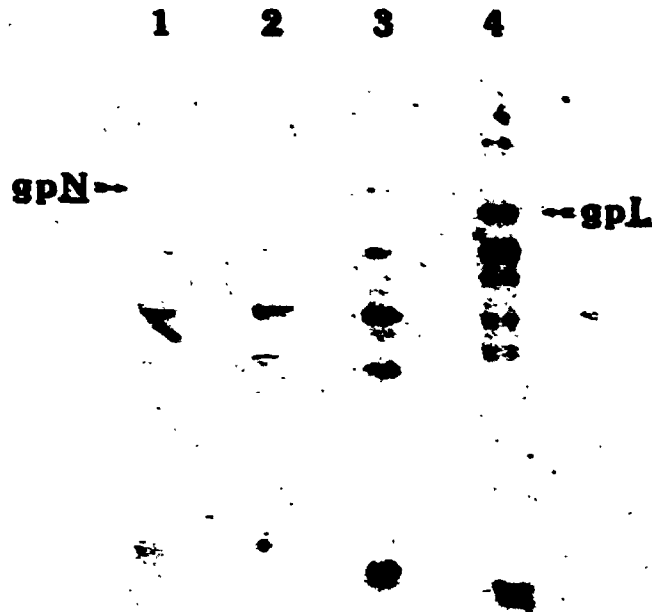
FIGURE 2.7: COMPARISON OF gpN AND THE INTRACELLULAR INTERMEDIATE BY PARTIAL PROTEOLYSIS

The intracellular protein-DNA complex was isolated in essentially the same manner as by Puspurs et al. (1983); exceptions are noted in Materials and Methods. Bacteriophage from GC103 were (^{35}S)-methionine labeled as described in Materials and Methods. [^{35}S]-Methionine labeled N protein was prepared in a coupled *in vitro* transcription-translation system (Zubay, 1973) using pGG418 as the template. The proteins from the virion and the intracellular Mu protein-DNA complex migrating at the same molecular weight on SDS-polyacrylamide gels as the N protein produced *in vitro* were cut out of the gel and equilibrated with loading buffer. Limited proteolysis was done in the stacking gel using V-8 protease as described by Cleveland et al. (1977), except that 1 ug of unlabeled BSA was added to each well to normalize the total protein concentration. Lane 1 contains the 64 kDa protein from the intracellular complex, lane 2 contains the 64 kDa protein from the Mu virion, lane 3 contains gpN produced *in vitro* and lane 4 contains Mu major tail protein from the virion. The arrows indicate the migration of the uncleaved proteins on the SDS-gels.

into the host cell and forms a stable protein-DNA complex with the
infecting Mu DNA.

DISCUSSION

Antiserum directed against the Mu virion 64 kDa protein has been prepared. This antiserum was used to probe a Mu expression library and recombinant plasmids were recovered which produced the 64 kDa protein. Southern blotting and restriction analysis of the DNA cloned into the expression vector demonstrated that the recombinant plasmids contained Mu DNA near the Mu Y, N, and P gene regions. Genetic analysis of the recombinant plasmids showed that the only functional Mu gene contained on all the constructs was the Mu N gene. Infection of Su⁻ E. coli cells with MuNam phage has illustrated that the antiserum recognizes the truncated forms of gpN. Partial proteolysis of the protein isolated from the intracellular protein-DNA complex and the cloned gene product has demonstrated that they are identical. These experiments indicate that the injected 64 kDa protein found in the Mu virion is the product of the Mu N gene.

The Mu N gene product has previously been assigned a molecular mass of 60 kDa (Giphart-Gassler et al., 1981), based on analysis of proteins synthesized in infected minicells. The difference between the two molecular mass assignments is larger than the error usually observed for SDS-polyacrylamide gel electrophoresis. There is only a small amount of N protein produced upon whole cell infection, and perhaps this small amount could not be adequately observed in minicell infections. Giphart-Gassler et al. observed the presence of another 64 kDa protein coded for by a Mu late gene, gpH. This

gene product is made in greater abundance than gp*N*. The *Ham* mutant utilized by this group to identify gp*H* as a 64 kDa protein produced a truncated polypeptide of approximately 63 kDa. The relative intensity of this truncated product may have obscured the gp*N* 64 kDa band, thus allowing it to remain undetected.

Two other points should be mentioned in this discussion. First, it is probable that the Mu *Y* gene product is considerably larger than the 12 kDa protein identified by Giphart-Gassler et al. (1981). Plasmids pGG409 and pGG413 produced a ≈49 kDa polypeptide when the plasmids were transcribed and translated *in vitro*. However, the entirety of the Mu *Y* gene was not present on these plasmids as defined by either complementation or marker rescue, indicating that the polypeptide observed *in vitro* was not a full length gene product. In addition, contrary to our published conclusions (Gloor and Chaconas, 1986) on the size of the deletion in pGG417, it is likely that this plasmid was deleted during the production of the expression library, by partial digestion with *Sau*3A. The deletion probably originates at the *Sau*3A site, deleting ≈15 kDa of potential *P* gene coding region. It is also probable, since the plasmid is not truncated at this site, that DNA not from this region has been added downstream of the partial *P* gene. This conclusion is more consistent with the known sequence of the region (Chapter 3) and the observed size of the truncated gene product observed *in vitro*.

The N protein appears to be a multipurpose one. In addition to binding to the ends of infecting Mu DNA (allowing circularization, supercoiling and exonuclease protection) the protein seems to be an integral part of phage tail structure. Grundy and Howe (1985) have found that Mu*Nam* phage are not able to produce tail structures. A very tight interaction between the N protein and gp*L*, the major tail protein (Gloor and Chaconas, unpublished), has been observed. However, infection of sensitive *E. coli* cells is not inhibited following incubation of Mu virions with the polyclonal gp*N* antiserum (data not shown). This suggests that the N protein is largely buried in the tail structure.

The determination of gp*N* as the injected 64 KDa protein that circularizes the infecting Mu DNA, explains the reason that mutants defective in integration mapping to this cistron have not been identified. These mutants would be genetically invisible because mutations that truncate the polypeptide (eg. *Nam* mutants) would not allow phage to be produced (Grundy and Howe, 1985), nor would mutants that do not allow the proper folding of the polypeptide chain (eg. *Nts* mutants.

CHAPTER 3

OVERPRODUCTION OF THE Mu N AND P PROTEINS

AND

SEQUENCING OF THE N AND P GENES

INTRODUCTION

In Chapter 2, the Mu virion 64 kDa protein was identified as the product of the Mu *N* gene. Further characterization of the *N* gene and its product is described in this chapter. A segment of DNA carrying the *N* and *P* genes was subcloned into a more potent expression vector, resulting in amplification of the corresponding gene products. The *N* and *P* genes were also cloned into a double stranded sequencing vector and the complete sequence of both genes has been determined.

OVERPRODUCTION OF THE *N* AND *P* GENES

The original constructs (pGG409, pGG417 and pGG418 cloned into the vector pCQV2) produced approximately 10 fold more N protein than was found in infected cells (Gloor and Chaconas, 1986). This level of overproduction was not sufficient to allow detection of the N protein on coomassie blue stained SDS gels, although the P protein was observed. Overproduction of the Mu *N* and *P* genes was achieved by deleting DNA upstream of the Mu *N* gene with *Bal*31 exonuclease and subsequently, cloning the genes into pKK223-3, a more potent expression vector. This vector contains a *tac* hybrid promoter under the control of the *lac* repressor. The -35 region is derived from the *E. coli trp* operon and the -10 region is from the *lac* operon (de Boer et al., 1983). Following the promoter is a polylinker derived from pUC-8 (Vieira and Messing, 1982) and the *E. coli rrnB* operon termination sequence (Brosius et al., 1981). The

presence of this termination sequence ensures that no transcripts originating from the *tac* promoter will interfere with the origin of plasmid replication (Brosius, 1984). Utilizing this vector, in some constructs, the N and P proteins accumulated after induction with IPTG, each forming approximately 5% of total cell protein.

SEQUENCING OF THE *N* AND *P* GENES

The vector chosen for sequencing, pAA3-7X (see Fig. 3.4), was constructed by Ahmed (1985, 1987) to generate random deletions from a fixed end point into a piece of cloned DNA. The vector contains TN9 with the IS*1*R deleted at the *Pst*I site, the remaining active IS*1* module provides the endpoint for formation of adjacent deletions (an intrinsic property of IS*1*; Iida at al., 1983). In addition, the plasmid contains part of the *E. coli gal* operon which codes for the *galK* and *galT* genes. A portion of the *galE* gene is also present. The two functional galactose operon genes code for galactokinase and galactose-1-puridyl transferase enzymes, which form UDP-galactose from UTP and galactose. Accumulation of UDP-galactose is lethal to the cell. Therefore in the absence of UDP galactose-4-epimerase, the product of the *galE* gene which converts UDP-galactose to UDP-glucose, the cell dies. In a strain deficient in galactose epimerase, growth on ampicillin and galactose allows selection of deletions extending into the cloned DNA, originating at the left end of IS*1*.

The ability to utilize a supercoiled double stranded DNA template to generate DNA sequence (Chen and Seeburg. 1985) allows the direct determination of the sequence of the deleted plasmid DNA.

MATERIALS AND METHODS: OVERPRODUCTION

CHEMICALS, ENZYMES AND DNA TECHNIQUES

See Chapter 2; Materials and Methods

SUBCLONING THE Mu N GENE INTO pKK223-3

Plasmid pGG418, (Gloor and Chaconas, 1966; Chapter 2) was used as the source of the Mu *N* gene. The plasmid (25 μg) was linearized by digestion with the restriction enzyme *Eco*RI. The restriction enzyme was then inactivated by heating at 65° for 10 minutes, the DNA was ethanol precipitated and resuspended in 40 μl of TE. An equal volume of 2X *Bal* 31 nuclease buffer was added with 2.4 units of *Bal*31 (NEB) exonuclease. The reaction was incubated at 37° and at time points of 6, 7, 8, 9, 10, and 11 minutes, aliquots of the reaction were removed and added to an equal volume of stop mix (100 mM EDTA, 100mM EGTA). A small sample of each time point was removed and analyzed by agarose gel electrophoresis. Following analysis, the remainder of the appropriate time points were pooled, extracted twice with φ-CHCl$_3$ and the DNA was ethanol precipitated. The DNA was resuspended in TE and the 5' or 3' extensions remaining from the *Bal*31 nuclease were removed by Mung Bean Nuclease (PPL). This reaction was done at room temperature for 30 minutes in a volume of 200 μl, and contained 25 units of the enzyme. The enzyme was removed by extraction with φ-CHCL$_3$, the DNA was ethanol precipitated twice and resuspended in 10 μl of TE.

EcoRI linkers (dG-G-A-A-T-T-C-C) were obtained as a gift from Dr. George Mackie. The linkers (100 pMole) were phosphorylated in a reaction buffer (10 µl) containing 50 mM Tris pH 7.8, 10 mM $MgCl_2$, 0.1 mM Spermidine (Sigma), 0.1 mM EDTA, 5 mM DTT, 1mM ATP, and 10 units of polynucleotide kinase (prepared by Dr. George Chaconas). The reaction was incubated at 37° for 1 hour. The linkers were purified by one extraction with ø-$CHCl_3$ followed by two extractions with $CHCl_3$. Prior to use in ligation reactions, the linkers were annealed by three sequential 5 minute incubations; at 37°, room temperature and on ice.

Five µl of pGG418, prepared as above, was ligated to the phosphorylated EcoRI linker in a 1:10 molar ratio at 20°, in a volume of 13 µl, for 16 hours. The DNA was extracted once with ø-$CHCl_3$, ethanol precipitated and resuspended in 43 µl of H_2O. The plasmid was then digested with 10 units of the restriction enzymes EcoRI and PstI. The reaction was incubated at 37° for 2 hours in the recommended buffer for EcoRI. The DNA was ø-$CHCl_3$ extracted, ethanol precipitated and resuspended in 20 µl of TE.

The expression vector utilized for overproduction of the N gene product was pKK223-3 (PPL). The plasmid (2.1 µg) was digested by both EcoRI and PstI in EcoRI buffer. The reaction was terminated by extracting with ø-$CHCl_3$,the DNA was recovered by ethanol precipitation and was resuspended in 20 µl of TE.

Plasmids pGG418 and pKK223-3, prepared as described above, were ligated together in 20 ul of ligation buffer containing either A; 0.1 ug of pKK223-3 and 2 ug of pGG418, or B; 0.2 ug of pKK223-3 and 5 ug of pGG418, in a volume of 20 ul. The reactions contained 1 Weiss unit of T4 DNA ligase and were incubated for 3 hours at 22^{C}. One ul of the ligation reaction was used to transform JM105 (Δpro-lac, supE, thi, rpsL, endA, sbcB, hsdR, F' traD, proAB, lacI^{q}lacZΔM15; received from PPL). Five percent of the transformation was plated directly onto LB plates containing carbenicillin, the remainder was diluted 1/50 in LB plus ampicillin and grown overnight to produce a liquid library. Seventy-five colonies grew on the transformation test plate for ligation A, and 38 colonies for ligation B.

SCREENING THE RECOMBINANT LIBRARY

Fifty colonies from ligation A and 38 colonies from ligation B were patched onto LB plates containing carbenicillin and grown overnight. The colonies were lysed and their DNA fixed to a nitrocellulose filter as described by Grunstein and Hogness (1975). The filters were hybridized with ^{32}P-labeled Mu DNA prepared by nick-translation as described in Chapter 2, and the filters were processed for autoradiograpy. The hybridizations indicated that approximately 40% of ligation A contained a Mu DNA insert and 25% of ligation B contained a Mu DNA insert.

The libraries were screened for production of the Hu N protein in a similar manner to that described in Chapter 2. Two hundred individual colonies were picked from streaks of the two libraries, patched onto LB plates containing carbenicillin and grown for 6 hours at 37°. The patches were lifted from the plates by a nitrocellulose disk and placed 'colony side up' on an similar plates containing 1 mM IPTG (Sigma). The colonies were grown overnight on the filters and processed for immunoscreening as described in Chapter 2. Two colonies which bound the anti-N protein antiserum were detected in library A and 12 were detected in library B. The positive colonies (14) were inoculated into growth tubes containing 2 ml of LB broth and ampicillin and grown at 37° until they were estimated to be in mid log phase of growth. At that time IPTG was added to 1 mM and the incubation was carried out for an additional 2 hours. One and one half ml of cells were pelleted in a microfuge, resuspended in 200 μl of 1X SDS gel loading dye and boiled for 10 minutes. Aliquots of the lysed cells (25 μl) were run on a 10% SDS acrylamide gel with a 4% stacking gel (Laemmli 1970) to assess the production of the Hu N and P proteins, whose genes were on the cloned piece of DNA.

PHYSICAL MAPPING

The 14 positives clones from the immunoscreen were inoculated into 5 ml of LB broth containing ampicillin and grown overnight at 32°. The cells were analyzed by cracking gel electrophoresis

(Chaconas et al., 1981a) on a 0.5% agarose gel along with size
standards.

The SDS-acrylamide gel and the cracking gel of the recombinant
clones demonstrated that 10 of the 14 possible clones contained
single Mu DNA insertions. DNA was prepared from these clones
(Birnboim and Doly, 1979) and the plasmids were restriction mapped
by BamHI, PstI and EcoRI-HindIII double digests. In addition, the
sequence of the 10 plasmids was determined for the junction between
the promoter sequence on the vector and the inserted Mu DNA (see
sequencing section).

GENETIC MAPPING

The 10 recombinant plasmids were transformed into strain X7102
(trp, pro, rpsL, lacI^q; Zehnbauer and Markovitz, 1980), which is a
suppressor minus lac I^q strain. Lawns of the transformed E. coli
strain containing one of the recombinant plasmids, or the vector
alone, were spotted with dilutions of Muctsamp, MuAam3011,
MuYam1034, MuNam7365 and MuPam1008 phage (phage strains described
in Chapter 2) in the presence and absence of 1 mM IPTG. The plates
were grown for 16 hours at 37° and scored for the ability of the
recombinant plasmids to complement the various phage mutants.

RESULTS: OVERPRODUCTION

SUBCLONING THE N AND P GENES INTO pKK223-3

The Mu N and P genes were subcloned into the potent expression vector pKK223-3. A map showing the vector and two representative plasmids derived from it are shown in Fig. 3.1. In order to clone the Mu N and P genes, the vector was linearized by double digestion with the restriction endonucleases EcoRI and PstI, to generate non-complementary cohesive ends. These were ligated to the Mu N and P genes, excised from pGG416 prepared as described in Materials and Methods. Since both fragments of DNA contained the same non-complementary cohesive ends, the N and P genes were expected to be cloned in the proper orientation and in high yield. In fact this was observed. Two ligation reactions were performed with the vector-insert ratio varied; 40% Mu DNA inserts were obtained in ligation A and 25% inserts were obtained in ligation B, as judged by hybridization of randomly selected clones to a Mu DNA probe.

N AND P PROTEIN PRODUCTION

The clones were screened for their ability to overproduce the Mu N gene product by direct colony immunoscreening, as described in Chapter 2. There were significantly fewer clones able to produce the N protein than contained Mu DNA insertions: 1% in ligation A and 6% in ligation B. This probably reflects the inability to
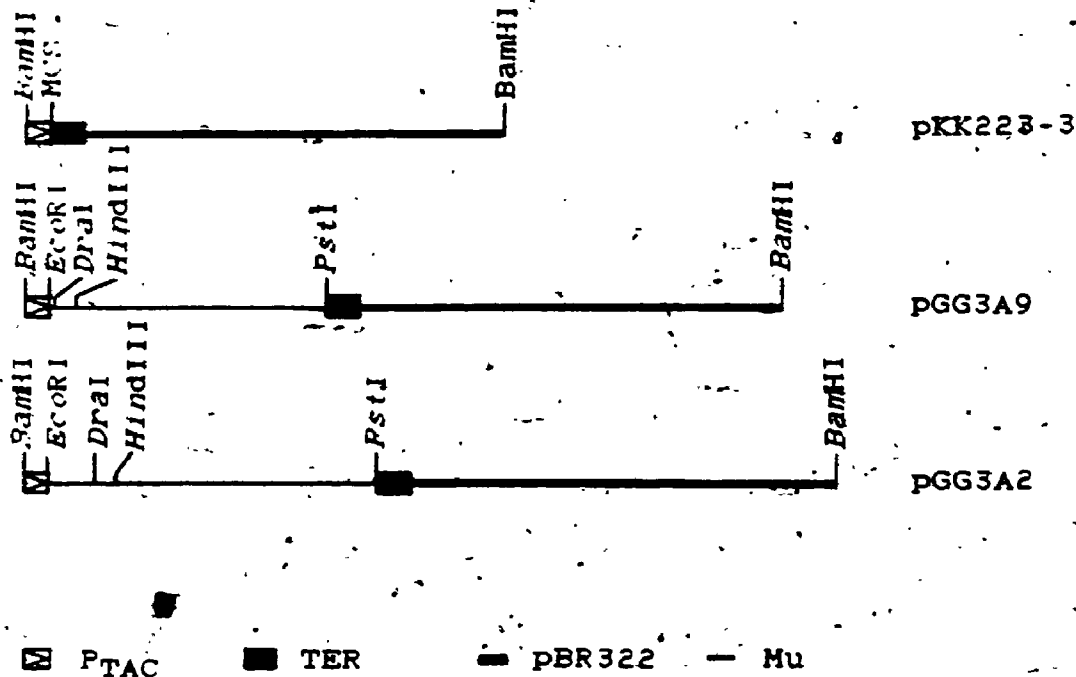
FIGURE 34: MAP OF pKK223-3 AND DERIVED PLASMIDS

The vector plasmid, pKK223-3, contains a strong promoter designated as P_{tac}, a multiple cloning site (MCS) containing the restriction sites EcoRI, SmaI, BamHI, SalI, PstI and HindIII respectively. A strong termination sequence and the β-lactamase gene allow stable maintenace of the plasmid. The derivatives of this plasmid containing the Mu N and P genes cloned into the EcoRI-PstI cut vector were constructed as described in materials and methods. Plasmid pGG3A2 was the largest clone recovered while pGG3A9 was the smallest. Plasmid pGG3B2 (not shown) contains a duplicated multiple cloning site. The restriction sites in this plasmid prior to the cloned Mu DNA are: EcoRI, SmaI, BamHI, SalI, PstI, SalI, BamHI, SmaI and EcoRI. This inverted duplication probably arose during the cloning.

produce a significant amount of N protein if a large part of the gene had been deleted by the Bal31 treatment of pGG418.

.'The positive clones were also examined for the amount of N protein overproduction by SDS-polyacrylamide gel electrophoresis of IPTG induced cells. The results of this analysis are shown in Fig. 3.2. The amount of N protein produced by the clones varies widely, from less than 1% of total cell protein, e.g. pGG2A2, to over 5% total cell protein, e.g. pGG3A7. The amount of total cell protein was determined by densitometry of the wet gel on an LKB densitometer as described by the manufacturer. Peaks were cut out and weighed in order to obtain the percentage of total cell protein. Two clones were observed not to make either the N or P proteins when observed on this gel, pGG2A1 and pGG3A3, and are not described further. In addition, clones with multiple insertions of Mu DNA were not further characterized. It is also apparent that the Mu P gene product is overproduced by these plasmids. The amount of overproduction is nearly independent of the construct and is also nearly independent of induction with IPTG. At best, a 2-fold induction of P protein when IPTG is added to the growth medium was observed.

RESTRICTION MAPPING OF OVERPRODUCING CLONES

Plasmids from the clones producing N protein were mapped by restriction endonuclease digestion. The EcoRI and PstI restriction endonuclease sites were present in all of the mapped plasmids,

UNINDUCED | PGG2A1 | PGG2A2 | PGG3A1 | PGG3A2 | PGG3A3 | PGG3A4 | PGG3A5 | PGG3A6 | PGG3A7 | PGG3A8 | PGG3A9 | PGG3B1 | PGG3B2 | PGG3B3

92.5 →
66.2 → ← gp*N*
45 → ← gp*P*
31 →
21.5 →
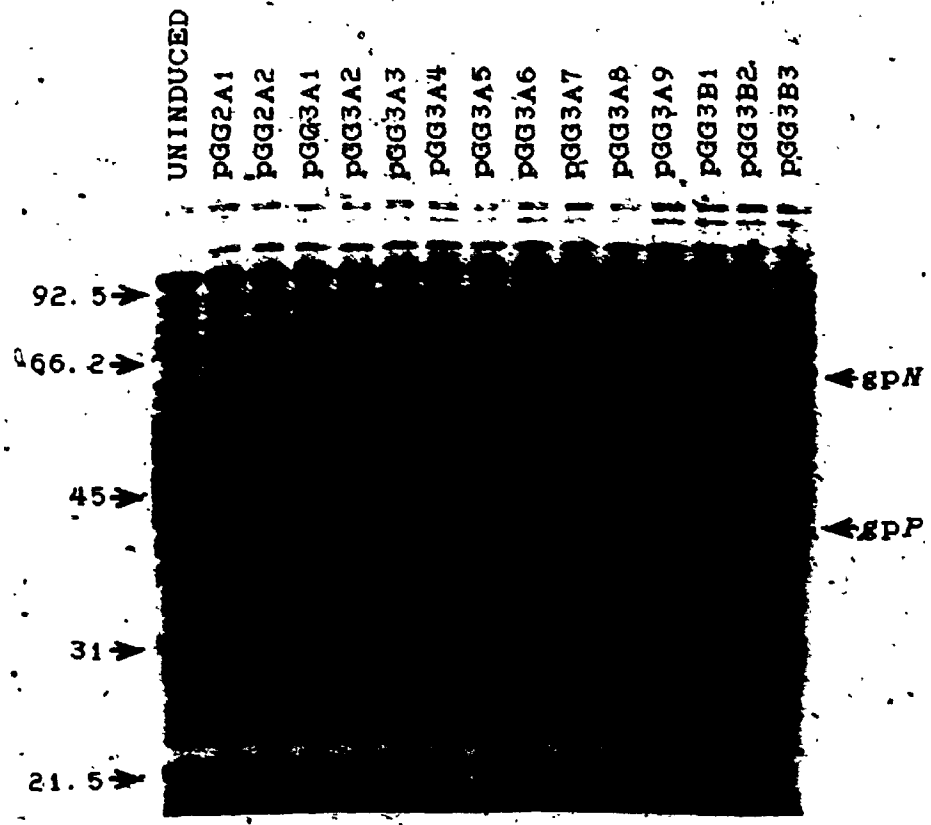
FIGURE 32: SDS-POLYACRYLAMIDE GEL OF INDUCED CLONES

The bacterial strain JM105 containing the plasmids indicated were
grown in LB broth containing ampicillin at 37° to approximately 5×10^8 cells/ml, at this point IPTG was added to 1 mM and cell growth
was permitted to proceed for a further 2 hours. One ml of the cells
were pelleted in a microfuge, resuspended in 150 ul of 1X SDS-gel
loading dye and boiled for 5 minutes. Twenty-five ul of the resulting
cell lysate was loaded in each lane. The first lane shows an
uninduced culture of pGG2A2, the remaining lanes show induced
cultures. The migration positions of molecular mass marker proteins
(kDa) is shown on the left. The migration positions of the N and P
proteins is shown on the right.

indicating that there was no undesired exonucleolytic degredation of the plasmids during the construction. Since the recovered clones are identical, except for the region left of the *Hin*dIII site, restriction maps of the largest (pGG3A2) and the smallest (pGG3A9) of the plasmids are shown in Fig. 3.1. In addition, the top strand near the *Eco*RI junction of all the plasmids was sequenced by double stranded DNA sequencing, utilizing a primer complementary to the pKK223-3 promoter sequence. The sequencing data complement and extend the restriction maps of these plasmids. The location of the deletion end points for these plasmids is noted in Table 3.1.

The amount of overproduction of the N protein coincides approximately with the deletion size generated by *Bal*31 during the subcloning of these plasmids. A graph correlating the amount of protein produced by the various plasmids and where the deletion end points map is shown in Fig. 3.3. This figure illustrates that there is a steady increase in the amount of protein produced by the plasmids up to and including the deletion end point of pGG3B3. After this point, the amount of protein produced by the constructs decreases. There are several explanations for the variation in the protein production of the plasmids. For example, the vector is known to contain a strong ribosome binding site and the location of this relative to the start codon of the *N* gene may influence the relative amounts of protein produced. There may also be inhibitory secondary structures in the messenger RNA, or possibly other inhibitory sequences further upstream of the start codon. In any

TABLE 3.1: DELETION END POINTS IN OVERPRODUCING CLONES

PLASMID	START NT	BP FROM AUG	SEQUENCE
pGG3A2	1	-471	TTGTG
pGG3A4	185	-287	GGCAG
pGG3A1	278	-194	CAGAA
pGG3B2	289	-183	GGCCA
pGG3B1	349	-123	GAACT
pGG2A2	361	-111	GGGTT
pGG3A7	397	-75	TGGTA
pGG3A8	402	-70	GGGGC
pGG3B3	418	-54	GCGGG
pGG3A9	448	-24	GACCA

The location of the deletion end points relative to both the initial nucleotide (NT) sequenced and the AUG codon for the *N* gene is shown above, for the overproducing plasmids based on the vector pKK223-3. The first 5 nucleotides of Mu DNA in the plasmid insertion is also given.
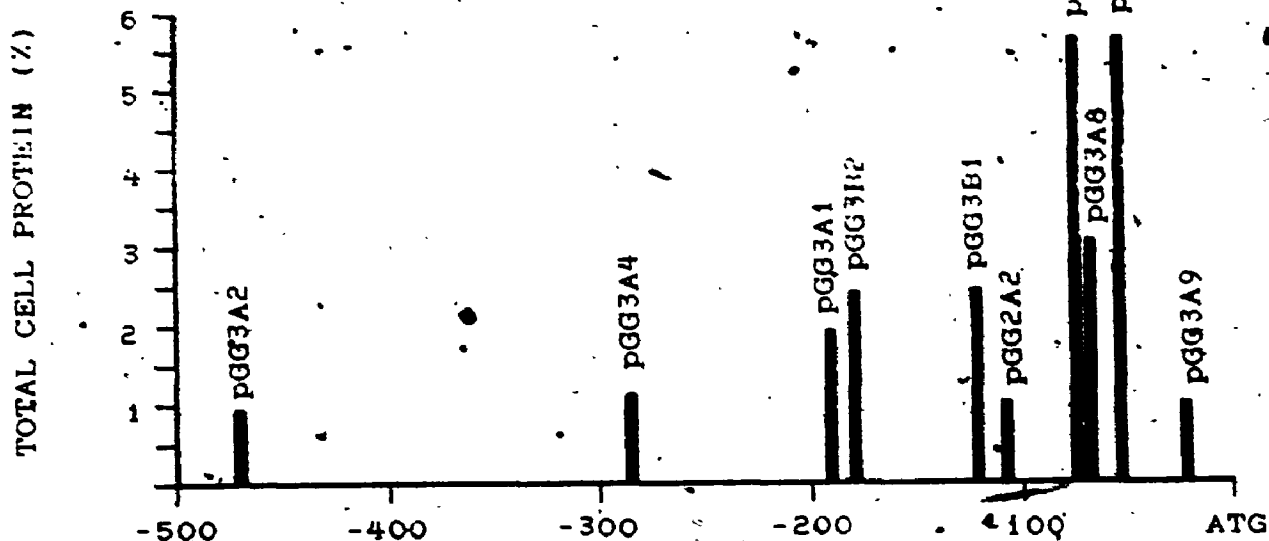
FIGURE 3.3: PROTEIN PRODUCTION BY OVERPRODUCING PLASMIDS

Strains overproducing the Mu N and P proteins were induced to make protein and run on an SDS-polyacrylamide gel. The gel was stained with coomassie blue and thoroughly destained. Densitometry of the wet gel on an LKB densitometer was done to determine the approximate amount of N and P protein produced by the various constructs. The bar graph above shows the results for the N protein. P protein production was nearly constant between the constructs at about 5% of total cell protein. Overproducing plasmids are arranged in order of distance from the N protein initiation codon and the junction between the Mu and vector sequence. The N protein as a % of total cell protein is charted. The two best overproducing clones were pGG3A7 and pGG3B3.

event, the mechanism of maximal overproduction was not of primary interest but rather the achievement of it.

GENETIC MAPPING

The 10 recombinant plasmids producing the Mu N protein were examined for their ability to produce a functional Mu N gene product. The plasmids were introduced into a suppressor minus lacI^q strain and various dilutions of Mu Nam, Pam and Aam phage were spotted onto the lawns. The production of a functional protein would allow the mutant phage to grow and produce a clear area where the phage were spotted. The results of this experiment are shown in Table 3.2. All 10 of the recombinant plasmids produced a functional N and P gene product. None of the plasmids could support the growth of the Mu Aam phage, but all of the plasmids could support the growth of Muctsamp. The amount of N protein could not be observed on an SDS-polyacrylamide gel when the cells were grown in the absence of IPTG (Fig. 3.2, lane 1); however, it is noteworthy that all the plasmids expressed enough N protein to complement the MuNam phage tested. This indicates that the tac promoter is not fully repressed even in a $lacI^q$ strain.

TABLE 3.2: GENETIC MAPPING OF N AND P OVERPRODUCING CLONES

PLASMID	PHAGE SPOTTED			
	Mu*A*am3011	Mu*N*am7365	Mu*P*am4008	Mu*c*tsamp
pGG3A2	−	+	+	+
pGG3A4	−	+	+	+
pGG3A1	−	+	+	+
pGG3B2	−	+	+	+
pGG3B1	−	+	+	+
pGG2A2	−	+	+	+
pGG3A7	−	+	+	+
pGG3A8	−	+	+	+
pGG3B3	−	+	+	+
pGG3A9	−	+	+	+

Mu phage containing the mutations indicated were spotted
onto a lawn of Su⁻ *E.coli* cells containing the plasmids indicated
at dilutions of 10^{-1}, 10^{-3} and 10^{-5}. For more specific details
see Materials and Methods. In the table above, − indicates that no
phage growth was observed; + indicates phage growth as indicated
by clearing of the spot.

MATERIALS AND METHODS: SEQUENCING

CHEMICAL, ENZYMES AND DNA TECHNIQUES

See Chapter 2: Materials and Methods.

CLONING THE MU N AND P GENES INTO pAA3-7X

PLASMID pGGSB

Plasmid pGGSB was derived from the ligation of pAA3-7X and pGG3A2 as described below. A restriction map of this plasmid and the other plasmids constructed is shown in Fig. 3.4. Plasmid pGG3A2, the largest of the plasmids used to overproduce the N gene product, was restricted with PstI and EcoRI, using the recommended buffer for EcoRI. The restriction endonucleases were removed by extraction with ϕ-CHCl$_3$ and the DNA was recovered by ethanol precipitation. The cohesive ends resulting from the restriction endonuclease digestions were removed by subsequent digestion with Mung Bean Nuclease (approximately 0.5 units/μg DNA) at 22^0 for 30 minutes. The Mung Bean Nuclease was removed by ϕ-CHCl$_3$ extraction and the DNA was ethanol precipitated. BamHI linkers (dC-G-C-G-G-A-T-C-C-G-C-G) were obtained as a gift from Dr. George Mackie and were phosphorylated as described previously, using polynucleotide kinase prepared by Dr. George Chaconas. The kinase reaction contained 1 unit of kinase per 25 pMole linker and was performed at 37^0 for 90 minutes (conditions as described in

previous section). The kinase was removed by extraction of the reaction, twice with ϕ-CHCl$_3$ and twice with CHCl$_3$. The linkers were annealed as described, and ligated to pGG3A2 in the ratio of 50:1 linkers:plasmid ends. The ligation was performed at 17° with 5 Weiss units of T4 DNA ligase for 18 hours in a final volume of 16 ul. After ligation, the reaction was diluted to 100 ul with 1X BamHI buffer, 10 units of BamHI were added and the restriction was incubated at 37° for 1 hour. The reaction was terminated by ϕ-CHCl$_3$ extraction and the DNA was recovered by ethanol precipitation. The DNA was resuspended in 10 ul of TE.

The plasmid pAA3-7X was cut with BamHI, and dephosphorylated with 2 units of bacterial alkaline phosphatase (PRL) per ug of DNA at 37° for 1 hour. The two plasmids were ligated together using final DNA concentrations of about 50 ug/ml of pAA3-7X and 100 ug/ml of pGG3A2. Five Weiss units of T4 DNA ligase were added to the reaction in a final volume of 10 ul. The ligation was incubated at room temperature for 2½ hours.

The E. coli strain AA102 (recA, pro, thi, supE, endA, hsdR, Δ(gal-chlD-pgl-att λ; Ahmed, 1985) was transformed with 5 ul of the ligation mix. The cells were plated on LB agar plates containing carbenicillin, and incubated at 37° overnight. Single colonies were picked off the transformation plates, patched onto LB plus carbenicillin plates, and the size of the recombinant plasmids analyzed by cracking gel electrophoresis of the plasmid DNA (Chacona± et al, 1981a). Several plasmids containing inserted Mu

DNA were mapped by restriction analysis and were found to contain the Mu *N* and *P* genes. However, only one orientation was recovered.

PLASMID pAATER

Recombinant plasmids in one orientation only (pGGSB) were recovered after analysis of >100 colonies from three separate cloning attempts. In order to clone the opposite orientation, pAA3-7X, was modified by inserting a strong transcriptional termination sequence in the tetracycline gene.

Plasmid pAA3-7X was restricted with *Sal*I, extracted with ∅-CHCl₃, ethanol precipitated, and the 5' overhanging end was filled in with *E. coli* DNA polymerase I (Klenow fragment) in the presence of the four dNTP's. The plasmid was extracted with ∅-CHCl₃ and ethanol precipitated. The DNA pellet was resuspended, digested with *Bam*HI, ∅-CHCl₃ extracted and ethanol precipitated. The resulting plasmid contained a "blunt" end (derived from the *Sal*I site) and a "sticky" end (derived from the *Bam*HI site) thus ensuring that the two ends would not ligate to each other. The transcriptional termination sequence was donated by the plasmid pKK223-3, which was previously used to overproduce the *N* gene product. Plasmid pKK223-3 was restricted with *Ssp*I and *Bam*HI in *Ssp*I buffer. The plasmid was ∅-CHCl₃ extracted, ethanol precipitated and resuspended in 10 µl of TE.

Plasmid pAA3-7X and pKK223-3, both prepared as above, were ligated in a 1:10 molar ratio using 3 units of ligase and 100 ng of pAA3-7X in a final volume of 10 µl. The ligation mixture was used to transform DH20 (*endA, hsdR, supE, recA, gyrA, relA?, F' lacI^q, lacZ, proAB*; Hannahan, 1983) and the transformed cells were grown in LB broth plus chloramphenicol (25 µg/ml) in order to select for plasmids derived from pAA3-7X. The liquid library prepared in this way contained >50% of the desired product, as judged by restriction analysis. Plasmids from 10 single colonies were restriction mapped, and all were found to be the desired product, which was named pAATER.

PLASMID pGGST

Plasmid pGGST was derived from the ligation of pAATER and pGG418 as follows. Plasmid pAATER was digested with *Bam*HI, ø-CHCl$_3$ extracted and ethanol precipitated. After the 5' single-stranded DNA overhanging ends were filled in with *E. coli* DNA polymerase 1 (Klenow fragment) and the four dNTP's, the plasmid was ø-CHCl$_3$ extracted, ethanol precipitated, digested with *Sal*I, again ø-CHCl$_3$ extracted and ethanol precipitated. Plasmid pGG418 was restricted to completion with *Dra*I and then NaCl was added to a final concentration of 150 mM. The restriction endonuclease, *Sal*I, was added and incubation was continued for another 2 hours. The DNA was extracted with ø-CHCl$_3$ and ethanol precipitated. The plasmids were ligated together in a 10 µl reaction mixture containing approximately 50 ng of each plasmid and 4 units of T4 DNA ligase.

The ligated plasmid mixture was used to transform DH20 and the transformed cells were grown in LB plus chloramphenicol. The recombinant plasmids were extracted from the liquid library, run on a 1% agarose gel overnight and the region of interest was cut out. The DNA was purified from the gel by the use of Gene Clean Bio 101) and used to transform DH20. The transformed cells were plated onto LB agar plates containing carbenicillin and grown overnight. Twenty colonies were picked and grown in LB broth containing ampicillin. Plasmid DNA was extracted from the cultures and restriction mapped. Four plasmids were found to have insertions in the desired orientation, one of which was saved and named pGGST.

PLASMID pGGDH

The 5' end of the top strand was poorly represented in deletion libraries formed from pGGST; therefore, it was necessary to clone this end in the sequencing vector without the 3' end. The plasmid pGG3A7 was digested to completion with *Dra*I and *Hpa*I, in the recommended buffer for *Hpa*I. The restricted plasmid was run on a 1% agarose gel overnight and the *Dra*I-*Hpa*I fragment derived from the Mu DNA insert was recovered with Gene Clean. Plasmid pAA3-7X was hydrolysed with *Bam*HI and the 5' overhanging ends were filled in with the Klenow fragment of DNA polymerase I and the four dNTP's. The DNA was CHCl$_3$ extracted and ethanol precipitated. The Mu DNA fragment and the vector were ligated together in a 5:1 molar ratio with pAA3-7X at a concentration of about 30 µg/ml, using 5 Weiss units of T4 DNA ligase in a reaction volume of 10 µl. The

reaction was incubated at room temperature for 8 hours. DH20 was transformed with an aliquot of the ligation and grown as a liquid library overnight in LB broth containing ampicillin. The DNA was extracted from the library and run on a 1% agarose gel. The library contained less than 1% insertions as judged by agarose gel electrophoresis. To avoid a large scale screening of the library the region of interest was excised from the gel, purified by Gene Clean and used to retransform DH20. Individual colonies of transformed cells were sized by cracking gel electrophoresis. The retransformed library contained 10% insertions. Restriction mapping showed that all the recombinant plasmids were in the same orientation as pGGST and were probably siblings, since the original cloning was not very efficient. This plasmid was named pGGDH.

OLIGONUCLEOTIDE PURIFICATION

Oligonucleotides were synthesized Dr. G. Mackie of this department. They were deblocked and supplied as a dried powder and were resuspended in a small volume of TE before use. Their concentration was determined by absorbance at 260 nM (1 $A_{260} \approx 20$ µg/ml Maniatis et al., 1982). The primer purification procedure evolved from that described by Maniatis and Efstratiadis (1980). A 20% acrylamide slab gel was prepared as described for DNA sequencing. The oligonucleotides were denatured by incubation at 65° in 50% formamide loading dye for 3 minutes, and ≈ 5 A_{260} units of oligonucleotide was loaded per lane. The gel was run at $\approx 50^{\circ}$ until the bromophenol blue dye was 2/3 down the gel. The

oligonucleotides were then visualized by shadowing with long wave
ultraviolet light (320 nM) using parafilm as a fluorescent
background. The central portion of the slowest migrating band was
cut out of the gel and diced. The gel pieces were overlayed with 2
volumes of extraction buffer (150 mM NaCl, 10 mM EDTA, 10 mM
Tris-HCl pH 7.8) and rotated at 4° overnight. The mixture was
centrifuged in a microfuge for 5 minutes, the supernatant removed
and lyophilized until the volume was less than 100 µl. The purified
oligonucleotides were desalted on a 0.5 ml spin column (Silhavey et
al., 1984) containing Sephadex G-25 in TE. The approximate
concentration of the oligonucleotides was determined by absorbance
at 260 nM. The purified product was stored at -70° until use.

SELECTION OF DELETIONS FOR SEQUENCING

Transformants of AA102 containing the cloned DNA of interest
were streaked on LB agar plates containing carbenicillin. Single
colonies were picked off the plates and patched onto MacConkey's
indicator agar plates containing 0.1% galactose and carbenicillin.
Generally, about 80% of the patches contained one or more viable
cells following incubation overnight at 37°. The viable cells were
picked and streaked for single colonies on the MacConkey's
galactose-carbenicillin plates. Usually, about 75% of the cells
that were restreaked were capable of forming single colonies; the
remainder grew as mixed populations and did not contain deletions
(Ahmed, 1987). The single colonies were patched onto LB plus
carbenicillin plates and grown overnight. The size of the plasmid

in the patches was assessed by cracking gel electrophoresis
(Chaconas et al., 1981a). Those plasmids containing deletions of
the appropriate size were grown in LB broth containing ampicillin
and were processed for sequencing as described below.

DOUBLE STRANDED SEQUENCING PRIMERS

The sequence of the primers used for double stranded sequencing is
shown below. All sequences are in the 5' to 3' orientation. The
origin and strand that sequence from the primer is obtained from is
also indicated.

PRIMER SEQUENCE	COMPLEMENTARY TO:	ORIGIN	STRAND
ATGCGTCCGGCGTAGA	pBR322 POS'N 396	NEB #1219	B
GTGTGGAATTGTGAGCGG	pKK223-2 POS'N 44	GM-9	T
AAAACACCATCATACACTAA	IS1 POS'N 45	GC-3 (NEB #1225)	T
CCACACTGAACTACGTCATC	N&P POS'N 1339	GC-6	T
CCCCCATCAGCGAGGAGGCC	N&P POS'N 1185	GC-7	B
TTCAGACAGTGGTACAGCC	N&P POS'N 1587	GC-8	T
GTATCGTATCCCGGAAAGG	N&P POS'N 2619	GC-9	T

DOUBLE STRANDED DNA SEQUENCING

Supercoiled DNA was prepared by a modification of the method of
Birnboim and Doly (1979). Although readable sequence could be
obtained from plasmids prepared for other purposes, the following
procedure gave the most reproducible results. A 25 ml overnight

culture grown in LB broth containing ampicillin was pelleted at 7 500 rpm in a Beckman JA-20 rotor in a J2-21 centrifuge for 5 minutes. The pellet was resuspended in 0.9 ml of resuspension buffer (25 mM Tris-Cl pH 8.2, 10 mM EDTA, 50 mM glucose) and 0.1 ml of a 10 mg/ml solution of lysozyme (Sigma) was added, the mixture was incubated on ice for approximately 5 minutes. The cells were lysed by the addition of 2 ml of lysing solution (0.2 M NaOH, 1% SDS), gently mixed by slowly rolling the tube and incubated at room temperature for about 5 minutes. The lysed cell suspension was neutralized by the addition of 1.5 ml of a 3 M potassium acetate (pH 5.0) solution, mixed gently and placed on ice for 15-20 minutes. Protein and high molecular weight DNA were pelleted by centrifugation at 15 000 rpm in the same rotor and centrifuge, for 5 minutes, 1.5 ml of the supernatant was divided between 3 large eppendorf tubes and ethanol precipitated. The precipitate was collected by centrifugation for 10 minutes in a microfuge; each pellet was resuspended in 67 µl of TE and pooled. Fifty µl of 10 M ammonium acetate, pH 6.0, was added to the pooled plasmid DNA, and the DNA was incubated on ice for 15 minutes. The mixture was centrifuged for 5 minutes in an eppendorf centrifuge to precipitate most of the high molecular weight RNA. The supernatant was extracted with φ-CHCl₃, the aqueous layer was removed to another tube and precipitated with ethanol. The nucleic acid pellet was resuspended in 90 µl of TE and 10 µl of a 1 mg/ml solution of boiled RNAse A (Sigma) was added. The mixture was incubated at room temperature for 15 minutes, extracted with φ-CHCl₃ and the aqueous layer was ethanol precipitated. The purified DNA pellet was

resuspended in 25 µl of TE and stored at 4° until used in sequencing reactions. The concentration of plasmid DNA was probably between 500 and 800 µg/ml. The absolute concentration was not important since the sequencing reactions were reproducible.

The sequencing reactions were performed by a modification of established protocols (Sanger et al., 1977; Biggin et al., 1983; Chen and Seeburg, 1985; Heinrich, 1986). Five µl of DNA was added to 3 µl of H_2O and 2 µl of a 2 M NaOH, 2 mM EDTA solution. The mixture was incubated at room temperature for 5-10 minutes to denature the plasmid DNA. The DNA was precipitated by the addition of 70 µl of 95% ethanol and 3 µl of a 3 M sodium acetate pH 5.0 solution. The precipitate was collected by centrifugation for 15 minutes in a microfuge, the supernatant was removed and the pellet was overlayed with 100 µl of 95% ethanol. The DNA was centrifuged for a further 5 minutes, drained, dried for 1-2 minutes at 42° and resuspended in 10 µl of primer solution (1.5 µl primer @ 63 µg/ml, 1.2 µl 10X RT buffer (200 mM Tris-Cl pH 8.2, 50 mM $MgCl_2$, 50 mM DTT, 500 mM NaCl), 7.3 µl H_2O). The primer and plasmid DNA were annealed by incubation at 42° for 10 minutes. After annealing, 3.5 µl of [α-thio ^{35}S]-dATP (NEN, >500 Ci/mMole) and 8 units of AMV-reverse transcriptase (BM) were added. Three µl aliquots of the primed plasmid mixture was added to 3 µl of the A, C, G, and T mixes (A mix: dCTP, dGTP, dTTP each at 100 µM, ddATP at 4 µM; C mix: dGTP, dTTP each at 100 µM, dCTP at 10 uM, ddCTP at 5 µM; G mix: dCTP, dTTP each at 100 µM, dGTP at 10 µM, ddGTP at 4 µM; T mix: dCTP, dGTP each at 100 µM, dTTP at 5 µM; ddTTP at 4 µM, all

1.0

1.1

1.25

1.4

1.6

1.8

2.0

2.2

2.5

2.8

3.2

3.6

4.0
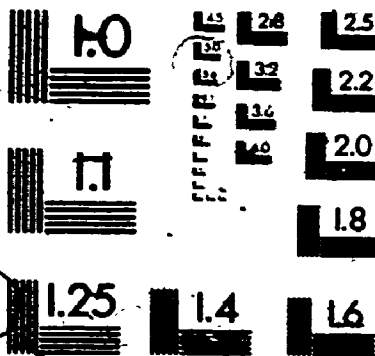
made up in 1X RT buffer) and incubated at 42° for 30 minutes. When 7-deazaguanosine triphosphate (BM) was included, it was used under identical conditions to dGTP in the G mix. After incubation, 0.8 ul of chase (dATP at 200 uM, dCTP, dGTP and dTTP each at 100 uM in TE containing 0.03 units/ul of DNA polymerase I (Klenow fragment, BRL), was added to each reaction and the incubation continued for 20-30 minutes at the same temperature.

When the reactions were performed with the Klenow fragment of DNA polymerase the mixes contained: A mix: dCTP, dGTP, dTTP each at 100 uM, ddATP at 100 uM ;C mix: dGTP, dTTP each at 100 uM, dCTP at 10 uM, ddCTP at 100 uM; G mix: dCTP, dTTP each at 100 uM, dGTP at 12.5 uM, ddGTP at 120 uM; T mix: dCTP, dGTP at 100 uM, dTTP at 5 uM, ddTTP at 250 uM. Sequencing reactions were performed with the Klenow fragment in 1X RT buffer as for reverse transcriptase, except incubations were performed at 37° for 20 minutes and Klenow was not included in the chase.

Sequenase (U.S.Biochemicals) was used as per the manufacturers instructions, except that template was limiting and the denatured plasmid (2 ul prepared as above) was used. The primer and template were annealed as described above, using 0.5 ul of primer to account for the smaller amount of input DNA.

Sequencing reactions were terminated by the addition of 4 ul of formamide loading dye (97% formamide, 2.5 mM EDTA, 0.01%

bromophenyl blue and 0.01% xylene cyanol, heated for 3 minutes at
100° and immediately placed on ice.

Sequencing gels were run according to Miller et al, (1985) with
some modifications suggested by Heinrich (1986); and were 6%
acrylamide, 0.3% bis acrylamide gels containing 8 M urea and run in
1X TBE (89 mM Tris, 89 mM boric acid, 2 mM EDTA, Maniatis et al
1982). Gels were either 0.6 mm slabs, 0.4-1.2 mm wedge gels or 0.5X
TBE to 4X TBE buffer gradient gels (Biggin et al, 1983). The gels
were prewarmed and run at a constant temperature of 60° (45-46 mA
for the slab gels, 50-52 mA for the wedge gels and 38-40 mA for the
buffer gradient gels). Running the gels at constant current
resulted in the most stable temperature. The bromophenol blue dye
ran about 20 nucleotides ahead of the junction between IS1 and the
cloned Mu DNA. Following electrophoresis, the gels were placed in
5% acetic acid for 10-15 minutes and then in two changes of H_2O for
5 minutes each. Subsequently the gels were dried and exposed to
film for 16-48 hours

SEQUENCING OF MUTANT PHAGE STRAINS

Bacteriophage Mu was prepared by two sequential CsCl gradients
and the phage DNA was extracted as previously described (Chaconas
et al, 1983). Ten to twenty µg of phage DNA was denatured by NaOH
exactly as described for plasmid sequencing. The denatured phage
DNA was then processed for sequencing identically to the plasmid
template. The denatured plasmid pellet recovered from the ethanol

wash was resuspended in 10 µl of Sequenase annealing buffer containing 1 µl of the desired primer (primer at 65 µg/ml). The phage DNA and the primer were annealed at 37° for 10 minutes and 1 µl of [α-thio ^{35}S]dATP was added along with 2 µl of 1:10-1:20 diluted labelling mix, 1 µl 0.1 M DTT and 2 µl of 1:8 diluted Sequenase (For more specific details see the Sequenase brochure supplied by U.S. Biochemicals). The mixture was incubated for 5 minutes at room temperature and then aliquoted into 2.5 µl of the A, C, G and T termination mixes which had been prewarmed to 37°. The termination reaction was carried out for 5 minutes at 37° and stopped by the addition of 4 µl of formamide loading dye. Sequencing reactions were subsequently run on a 6% sequencing gel as described above.

RESULTS: SEQUENCING

SEQUENCING OF THE N AND P GENES

Mu DNA clones in the double stranded sequencing vector, pAA3-7X, and its derivative, pAATER, were the source of most of the sequence data presented. A primer complementary to the pBR322 BamHI site was used to generate sequence near the PstI site of plasmid pGG409. Plasmid pGG3A7 was used to generate sequence when primers specific to the Mu N gene were utilized. In the discussion, below, the top strand (T) refers to the strand that has the same polarity as the mRNA and the bottom strand (B) refers to the mRNA template strand. Three constructs containing Mu DNA were produced: pGGSB contained Mu DNA in the orientation necessary to sequence the mRNA template strand, pGGST contained Mu DNA in the orientation necessary to sequence the complement of the template strand while, pGGDH, contained a small fragment of the original Mu DNA and was in the same orientation as pGGST.

There were several constructs made for the sequencing of the Mu N and P genes. Restriction maps of these plasmids and of the parent vector are shown in Fig. 3.4.

During the cloning of the N and P genes into pAA3-7X for sequencing, it was observed that one orientation of the DNA was preferentially selected. The orientation that was not recovered was that where the P promoter would transcribe through the plasmid
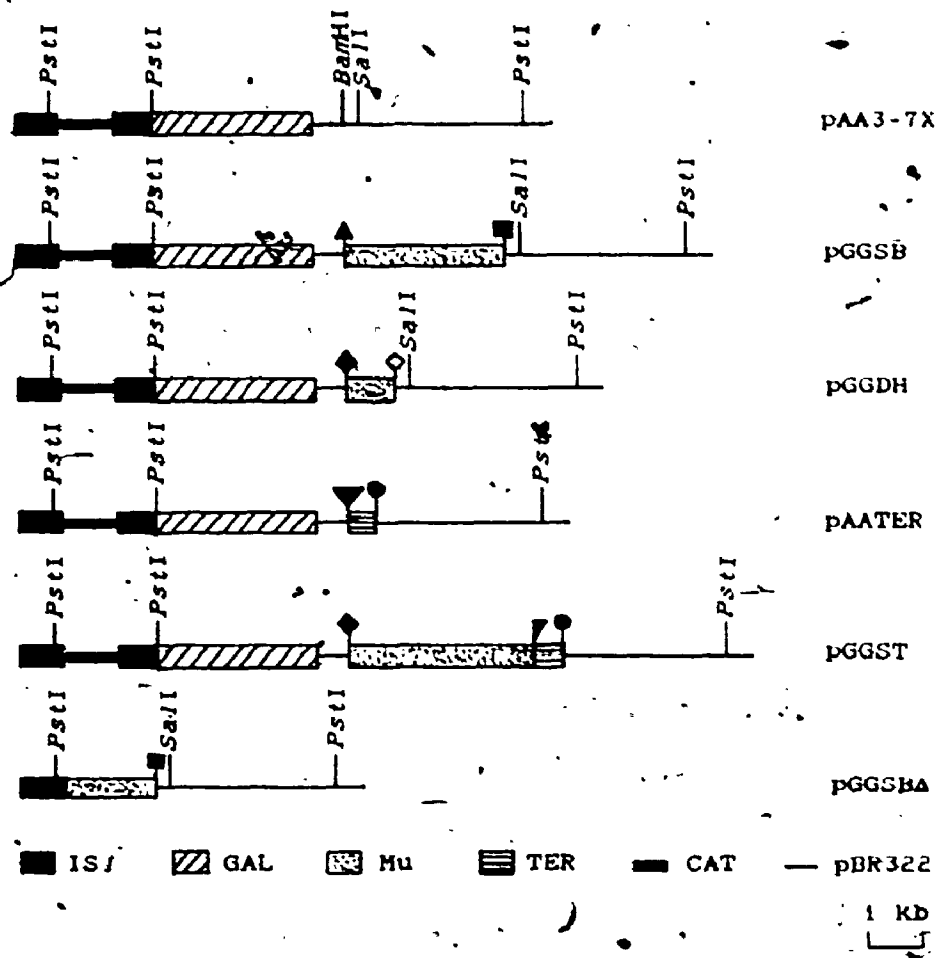
FIGURE 3.4: PLASMIDS UTILIZED TO SEQUENCE THE Mu *N* AND *P* GENES

A restriction map of the plasmids utilized to sequence the Mu *N* and *P* genes is shown. The parent vector for all the constructs was pAA3-7X and is the top plasmid. Plasmids containing Mu DNA that were directly derived from this vector (pGGSB, pGGDH) are shown below. Plasmid pAATER, a derivative of pAA3-7X containing a strong terminator is shown next. The multiple cloning site in pAATER contains (l-r) the recognition sites for the following restriction enzymes: BamHI, SaII, PstI and HindIII. Plasmid pGGST, the sequencing derivative of this vector is shown below pAATER. The final construct shown has the typical structure of a deleted plasmid from which sequence in the cloned DNA could be obtained. Intervening DNA between the active IS*I* module and the cloned DNA has been deleted allowing a primer complementary to IS*I* to read into the cloned DNA. The origin shown for the restriction maps is the left end of the active IS*I* module on the plasmid, from which the deletions originate. Symbols shown are:

▲, BamHI-PstI junction sequence; ■, EcoRI-BamHI junction sequence;
◆, BamHI-DraI junction sequence; ◇, HpaI-BamHI junction sequence;
▼, multiple cloning site; ●, SspI-SaII junction site;
▽, SaII-multiple cloning site junction (SaII site present)

origin of replication located near the cloning site. It had been inferred previously (see Chapter 2 and above) that there was a relatively strong promoter located in this region of DNA which was responsible for constitutive production of gp$P$ at high levels. When an attempt was made to clone the genes into pAA3-7X by directional cloning, microcolonies were produced upon extended incubation, but cell growth on ampicillin was far too slow to allow efficient recovery of the plasmids. This slow growth is probably indicative of interference with plasmid maintenance by transcripts from the P promoter.

To alleviate this problem, pAATER was constructed as described in the Materials and Methods. This plasmid allowed efficient recovery of the desired orientation, presumably because there was now a strong termination sequence located between the promoter and the plasmid origin.

The derivative of pAATER used for sequencing, pGGST, was preferentially deleted near the termination sequence when incubated on ampicillin and galactose and, therefore, the inserted DNA less than 2 kbp from the termination sequence was rarely represented in the deletion libraries. Therefore, to sequence the first kb, the DraI-HpaI fragment of the Mu N gene was cloned into pAA3-7X in the orientation necessary to sequence the top strand, and two primers were made to fill in the remaining gaps. One gap was in the top strand and included the putative P promoter. The other gap was in the bottom strand and was only 130 bp in size.

The complete sequence of both strands (≈3 000 bp) was determined. A diagram showing the end point and orientations of the sequence obtained from the various deletions recovered is shown in Fig. 3.5. Approximately 200 base pairs was the usual practical limit for reading sequence off a double stranded template. When extreme care was taken in preparing the template DNA an extra 100 base pairs was occasionally possible. The use of Sequenase tended to allow slightly more bases to be read. The sequence obtained from the two primers made to fill in the gaps are shown as the bold arrows. The complete nucleotide sequence of this region and some restriction sites are shown in Fig. 3.6.

There are only two long open reading frames in this sequence. The first frame is presumed to code for the Mu *N* gene, the second for the Mu *P* gene. These two open reading frames overlap by 8 base pairs and are different. All open reading frames are translated in Fig. 3.6. The putative *N* and *P* amino-acid sequences are shown in bold print. The DNA sequence where the two reading frames overlap is also in bold print.

The open reading frame in the Mu *N* gene was confirmed by sequencing of two amber mutants; Mu*N*am1995 and Mu*N*am 7106. These mutants produce truncated polypeptides of 55 and 54 kDa respectively, when Su⁻ *E. coli* cells are infected (Gloor and Chacoñas, 1986; Chapter 2). The mutation sites reside at positions 1415 and 1438 nucleotides in the sequence (Fig. 3.6), and are in the frame necessary to truncate the long open reading frame present
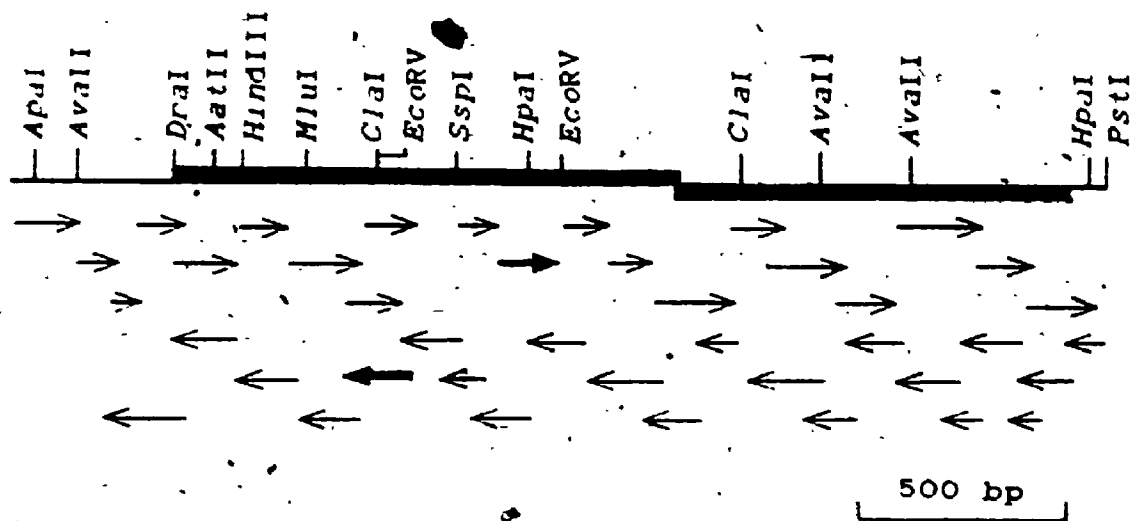
FIGURE 35: STRATEGY FOR SEQUENCING THE Mu *N* AND *P* GENES

A restriction map of the complete region sequenced is shown above. Plasmids with deletions extending from the right end of IS*I* into the cloned Mu DNA were selected on galactose and ampicillin. Deletions derived from pGGST and pGGDH were utilized to generate overlapping sequence of the top strand and are the arrows pointing right. Plasmids selected in the same manner derived from pGGSB were used to generate overlapping sequence of the bottom strand, and are the arrows pointing left. The bold arrows represent sequence obtained from primers specifically synthesized to fill in the remaining gaps. The *N* and *P* gene open reading frames are shown as the dark boxes above and below the central line, respectively.

Figure 3. 6: SEQUENCE AND TRANSLATION OF THE Mu *N* AND *P* GENES

The sequence of the Mu genome containing the *N* and *P* genes and
the translation of the two long reading frames, is shown following.
Some restriction endonuclease recognition sites are underlined in
the DNA sequence. The DNA sequence where the two genes overlap is
shown in bold print. The amino-acid sequence of the proteins has
been underlined at the proposed DNA binding sites. The sequence
has been determined in both strands starting at nucleotide 285 and
continuing to the end.

```
TTGTGACAAA AGGTGCCGGG GCGCTGGGCT GGATGGGGAA AGGTGCCGGT CGTGCGTTGT 60

CACGTCTGGG GGGCCCGGTA ATGGGGGCAC TCCAGCTTGC CCGCGTCCTG ATGGATGAGC 120

AGGCGTCAAC CCATGAAAAA GCAGGCGCAA TTGGCAGTAC AGCCGGTGCA TGGCTCGTGC 180
                   AvaII
CGTTGGCAGC CTTGCCGGAC CGCTGGGTAC GGTTGCTGGT GCCACGCTGG GCAGTGGTCG 240

CCGGGGAGTA TCTGGGCGGT TTTGTAACCG ACCTGTATCA GAAATGGACG GCCACGGATA 300

AGGAACCGCA GGAACACAAA AAGTCAATGC GGAAGCCTCG CTGCGCGTCG AACTCGGCGA 360

GGGGTTACGT CTGACCAGTT CCCGCGTCAC CGAGGATGGT ATGGGGCTGA ATATTTACGC 420
                                                        DraI
GGGCGATAAC TACATTACGG GCTGGTAAGA CCATGTTTGA AGATGCTTTA AATGCCGTTA 480
                                                     MetProLeu

ATGCTGTCCG GGATAAAACC GGTGGAGGCA GGAAAAGAAC CGGCAAAGGC ACGTTCGGTA 540
MetLeuSer GlyIleLys ProValGluAla GlyLysGln ProAlaLys AlaArgSerVal
                                                            BstNI
ACGTGCCGTT TCTCGTCATC ATCGAGGAGG CAAAAACAGG CTGGCGGGCC ACGTCGCCTG 600
ThrCysArg PheSerSer SerSerArgArg GlnLysGln AlaGlyGly ProArgArgLeu
                                        BstNI      HindIII
GTTAAACGCG AGTACCCGTT ACGTGATACC GGCGGCGTCA ATGACCTGGG GAAAAAGCTT 660
ValLysArg GluTyrPro LeuArgAspThr GlyGlyVal AsnAspLeu GlyLysLysLeu

CGCTCCCGTA CATTCAGCGC CTGCATTCTG AACAGCAACG CAGAAACAGC CAGAGATGAA 720
ArgSerArg ThrPheSer AlaCysIleLeu AsnSerAsn AlaGluThr AlaArgAspGlu

GCGGGTCCGC TGATGGATGC TCTTGATGCT CCGGGTAGCG GTGAGCTGGT ACATCCTGAT 780
AlaGlyAla LeuMetAsp AlaLeuAspAla ProGlySer GlyGluLeu ValHisProAsp

TTCGGCACTG TGGACGTCAT GGTGGATTCA TGGGAATCCC GCACTAAAGC GGATGAACTG 840
PheGlyThr ValAspVal MetValAspSer TrpGluCys ArgThrLys AlaAspGluLeu
         MluI
AATTATTACG CGTTCACCGT TACCGTTTAT CCGTCGTTGC AGGATACTGC CCCGGACGCA 900
AsnTyrTyr AlaPheThr ValThrValTyr ProSerLeu GlnAspThr AlaProAspAla

GAGACAGACA CCAGTGCAGC CGTACCGGCA CAGGCCGTTG CTGTAACCGG TTCTCTGGGA 960
GluThrAsp ThrSerAla AlaValProAla GlnAlaVal AlaValThr GlySerLeuGly

GATACGCTGT CCTCTGTCTG GCAAACCGTA AAAGATGGCA CTGCGGCGGC AACCGCCGTG 1020
AspThrLeu SerSerVal TrpGlnThrVal LysAspGly ThrAlaAla AlaThrAlaVal
            ClaI    EcorV
ATGGAAGCTG TAACCGGTGT CATCGATGAT ATCAGTGATG CGGTGGACAA TCTGGGGGTT 1080
MetGluAla ValThrGly ValIleAspAsp IleSerAsp AlaValAsp AsnLeuGlyVal
```

Sau3A

ACGCAGACTG TCAGCGGTCT GATGGGATCG CTTTCTGCGA TGAAAGGCTC TGTGACCAGC 1140
ThrGlnThr ValSerGly LeuMetGlySer LeuSerAla MetLysGly SerValThrSer

CTGATTAACC AGCCTCCCAT GCTGGCCTCC TCGCTGATGG GGGCGCTGTC CGGCGTTTCA 1200
LeuIleAsn GlnProAla MetLeuAlaSer SerLeuMet GlyAlaLeu SerGlyValSer

TCGTTATGCG ATACCCGGAC AGCATTTTCC ACATGGAACC GTCTGGCGCA GCGATTCGAA 1260
SerLeuCys AspThrArg ThrAlaPheSer ThrTrpAsn ArgLeuAla GlnArgPheGlu

CGTCGCCATG CCGCCACCGC AGGCAGACAG GGGACAATCA CAACCTCGTA CAACAGTCCG 1320
ArgArgHis AlaAlaThr AlaGlyArgGln GlyThrIle ThrThrSer TyrAsnSerPro

GTTGCAGAAA AAAATATTGC CACACTGAAC TACGTCATGC TGGCAGCGGC GCAGACATAC 1380
ValAlaGlu LysAsnIle AlaThrLeuAsn TyrValMet LeuAlaAla AlaGlnThrTyr
        BstNI

CGGGCAGAAG CTGCCAGCCA GGCACTGACT GCGGCACTGG ATTTCAGTCG CCGGATGGAT 1440
ArgAlaGlu AlaAlaSer GlnAlaLeuThr AlaAlaLeu AspPheSer ArgArgMetAsp

AATGCCGCCC GTGCACCTGT ACTGGATGCC CCGTCCACCA CAACCGGCAC AGCCAGCGGG 1500
AsnAlaAla ArgAlaPro ValLeuAspAla ProSerThr ThrThrGly ThrAlaSerGly
                                                    HpaI

GCCAGCAGCA CATCTGCTAC CGTCACACAG GGACAGTTAC AGTTAACTGC CATAACCCCG 1560
AlaSerSer ThrSerAla ThrValThrGln GlyGlnLeu GlnLeuThr AlaIleThrPro

GACGGCGGCT TTTCACAGGT ATCCTTTTCA GACAGTGGTA CAGCCACGCC CCCGGTATTT 1620
AspGlyGly PheSerGln ValSerPheSer AspSerGly ThrAlaThr ProProValPhe
        EcoRV

GAAAGTGTGT CCGATATCGA AAAAACCACT GCCATGCTGG GGGCGGCGCT GGATAGCGTC 1680
GluSerVal SerAspIle GluLysThrThr AlaMetLeu GlyAlaAla LeuAspSerVal

ATTCTGACGG CATCTGAGCA GGGTTTTTCG ACAGACAGTG TTCAGCTTAC GCAACTGCGT 1740
IleLeuThr AlaSerGlu GlnGlyPheSer ThrAspSer ValGlnLeu ThrGlnLeuArg

CTGCTGGTTG TTGCCGACCT GGAAAAACGC GGGCTGCAAC TGGCGGGTAG TGAATCACAC 1800
LeuLeuVal ValAlaAsp LeuGluLysArg GlyLeuGln LeuAlaGly SerGluSerHis

CACCTGCCAG AAACGCTCCC GGCAATGGTT GCACTGTACC GGTTCACCGG AAACAGCCGG 1860
HisLeuPro GluThrLeu ProAlaMetVal AlaLeuTyr ArgPheThr GlyAsnSerArg

AACTGGCAAC GGCTGGCCCG CAGGAACGGT ATCAGCAACC CGTTATTTGT TCCCGGTGGT 1920
AsnTrpGln ArgLeuAla ArgArgAsnGly IleSerAsn ProLeuPhe ValProGlyGly

GTCAGTATTG AGGTGATTAA TGAGTAATAC CGTCACACTG CGAGCGGATG GCAGGCTGTT 1980
ValSerIle GluValIle AsnGlu *
                MetSerAsnThr ValThrLeu ArgAlaAsp GlyArgLeuPhe

TACCGGCTGG ACGTCAGTCT CTGTCACCCG CTCGATTGAA TCCGTAGCCG GATATTTTGA 2040
ThrGlyTrp ThrSerVal SerValThrArg SerIleGlu SerValAla GlyTyrPheGlu

GCTGGGGGTG AACGTGCCAC CGGGCACGGA TTTATCCGGG CTGGCTCCGG GGAAGAAGTT 2100
LeuGlyVal AsnValPro ProGlyThrAsp LeuSerGly LeuAlaPro GlyLysLysPhe
                                                    ClaI

CACGCTGGAA ATCGGGGGGC AGATTGTCTG CACCGGTTAT ATCGATTCAC GGCGACGCCA 2160
ThrLeuGlu IleGlyGly GlnIleValCys ThrGlyTyr IleAspSer ArgArgArgGln

```
GATGACCGCT GACAGTATGA AAATCACTGT CGCCGGACGT GACAAAACGG CTGACCTGAT 2220
MetThrAla AspSerMet LysIleThrVal AlaGlyArg AspLysThr AlaAspLeuIle

TGACTGTGCT GCCGTTTACA GTGGCGGACA GTGGAAAAAC CGCACACTGG AGCAGATTGC 2280
AspCysAla AlaValTyr SerGlyGlyGln TrpLysAsn ArgThrLeu GluGlnIleAla

GCGTGACCTG TGCGCTCCTT ATGGCGTTAC CGTTCGCTGG GAGCTTTCCG ATAAGGAAAG 2340
ArgAspLeu CysAlaPro TyrGlyValThr ValArgTrp GluLeuSer AspLysGluSer
```

                                Ava II
```
TTCGGCAGCT TTTCCCGGCT TCACGCTGGA CCATTGAGAA ACCGTTTATG AGGCGCTGGT 2400
SerAlaAla PheProGly PheThrLeuAsp HisSerGlu ThrValTyr GluAlaLeuVal

GCGTGCCTCC CGCGCACGCG GTGTACTGAT GACCAGCAAT GCCGCCGGAG AGCTGGTATT 2460
ArgAlaSer ArgAlaArg GlyValLeuMet ThrSerAsn AlaAlaGly GluLeuValPhe

CAGCCGGGCT GCCAGCACAG CCACTGATGA GCTGGTTCTC GGAGAAAATC TGCTGACACT 2520
SerArgAla AlaSerThr AlaThrAspGlu LeuValLeu GlyGluAsp LeuLeuThrLeu

GGATTTTGAG GAAGACTTCC GCGACCGGTT CAGCGAATAC ACCGTCAAGG GGTATGCCCG 2580
AspPheGlu GluAspPhe ArgAspArgPhe SerGluTyr ThrValLys GlyTyrAlaArg
```
                                                        Ava II
```
CGCAAATGGT GCTGAGGGTG ATGATATTGA TGCGAAAAGT ATCGTATCCC GGAAAGGGAC 2640
AlaAsnGly AlaGluGly AspAspIleAsp AlaLysSer IleValSer ArgLysGlyThr
```
                                        Bcl II Sau3A
```
CGCCACTGAC AGTGATGTGA CCCGTTACAG ACCGATGATC ATCATTGCTG ACAGCAAGAT 2700
AlaThrAsp SerAspVal ThrArgTyrArg ProMetIle IleIleAla AspSerLysIle

TACGGCGAAG GATGCACAGG CCCGCGCCCT GCGTGAGCAA CGCCGCAGAC TGGCAAAATC 2760
ThrAlaLys AspAlaGln AlaArgAlaLeu ArgGluGln ArgArgArg LeuAlaLysSer

CATCACCTTT GAGGCAGAAA TTGACGGATG GACTCGCAAG GACGGGCAAC TCTGGATGCC 2820
IleThrPhe GluAlaGlu IleAspGlyTrp ThrArgLys AspGlyGln LeuTrpMetPro

GAACCTGCTG GTCACTATTG ATGCCTCGAA ATATGCCATC AAAACCACGG AATTACTGGT 2880
AsnLeuLeu ValThrIle AspAlaSerLys TyrAlaIle LysThrThr GluLeuLeuVal
```
                                BstNI
```
CAGCAAAGTC ACCCTGATAC TGAATGACCA GGACGGGCTG AAAACCCGCG TCAGCCTTGC 2940
SerLysVal ThrLeuIle LeuAsnAspGln AspGlyLeu LysThrArg ValSerLeuAla

ACCACGCGAA GGCTTTCTGG TGCCGGTTGA AAGCGACCGT AAAAACAGGA AAGGCGGCGA 3000
ProArgGlu GlyPheLeu ValProValGlu SerAspArg LysAsnArg LysGlyGlyAsp

CAGTAACGGC GGTATTGATG CGCTGGTTGA AGATTATTAT CGCAGACACC CGGAGAAAAC 3060
SerAsnGly GlyIleAsp AlaLeuValGlu AspTyrTyr ArgArgHis ProGluLysThr

GCCGCCGTGG AAAGAGTAAA TGATTCCGCC TTGAACCGCC TGCTGACGCC GCTGATGCGT 3120
ProProTrp LysGlu *
```
                                HpaI
```
CGTGTGCGCC TGATGCTTGC GCGCGCTGTT GTTAACGTGA TTAACGACGG GCGAAAGGTT 3180
```
    PstI
```
CAGAACCTGC AGGT
```

in this region. Both mutations were found to be C-T transitions. The mutations are shown in fig 3.7. The reading frame for the Mu P gene was determined by sequencing of p33417 through the second Sau3A site in the sequence shown in Fig 3.6. This plasmid produces a truncated P protein in vitro (Gloor and Chaconas, 1986; Chapter 2) of 28 kDa. The termination codon (opal) resides 30 nucleotides from the Sau3A site and is in the proper frame to terminate the P gene open reading frame. There is a third open reading frame extending from the start of the sequence determined and ending at nucleotide 371. The Mu Y gene is known to map in this region (Gloor and Chaconas, 1986; Chapter 2) and the open reading frame may correspond to the carboxyl terminus of this gene. This open reading frame does not overlap with the N gene.

The putative amino-acid content and molecular mass predictions for the two proteins are shown in Table 3.3. The mass predictions for the Mu N protein are substantially lower than expected. The N protein migrates on SDS-polyacrylamide gels as a 64 kDa protein, while the sequence predicts a mass of 51 548 da. This is unusual but not unprecedented (de Jong et al., 1978; Smith and Nicolas, 1983; Ferguson et al., 1984), as several other proteins also migrate anomalously in this gel system. Indeed, it has been observed that changing only one amino acid can alter the migration of a 20 kDa protein by up to 2 kDa (de Jong et al., 1978).

The mass predictions for the P protein are very close to the expectations, the Mu P protein has an apparent mass of 43 kDa on

FIGURE 3.7: SEQUENCE OF TWO MuNam MUTANTS

The reading frame of the Mu *N* gene was confirmed by sequencing of two amber mutations in this gene. The mutations *Nam*1995 and *Nam*7106 were sequenced by double stranded sequencing using the linear phage DNA and sequenase. Both mutations were found to be C to T transitions. The location of the mutant base is shown by the arrow marked with a dot, the corresponding base in the other mutant phage that agrees with the known sequence of this region is shown by the unmarked arrow.

TABLE 3.3: PREDICTED AMINO-ACID CONTENT OF gpN AND gpP

	N PROTEIN		P PROTEIN		AVERAGE
AA	No.	%	No.	%	%
Ala	64	13.03	36	9.50	8.6
Arg	29	5.91	32	8.44	4.9
Asn	16	3.26	10	2.64	4.3
Asp	28	5.70	30	7.92	5.5
Cys	4	0.81	3	0.79	2.9
Gln	21	4.28	8	2.11	3.9
Glu	20	4.07	24	6.33	6.0
Gly	37	7.54	30	7.92	8.4
His	4	0.81	2	0.53	2.0
Ile	13	2.65	19	5.01	4.5
Leu	45	9.16	32	8.44	7.4
Lys	13	2.65	21	5.54	6.6
Met	13	2.65	6	1.58	1.7
Phe	13	2.65	11	2.90	3.6
Pro	21	4.28	12	3.17	5.2
Ser	52	10.59	28	7.39	7.0
Thr	48	9.78	32	8.44	6.1
Trp	4	0.81	6	1.58	1.3
Tyr	8	1.63	11	2.90	3.4
Val	38	7.74	26	6.86	6.6
END	1	0.20	1	0.26	

N PROTEIN

molecular weight = 51 518
number of amino acids = 491
Arg + Lys = 42
Asp + Glu = 48

P PROTEIN

molecular weight = 41 738
number of amino acids = 379
Arg + Lys = 53
Asp + Glu = 54

The amino-acid content of the N and P proteins predicted from translation of the DNA sequence is shown above along with the values for an average of 314 *E. coli* proteins of known sequence (Dayhoff et al., 1978). There are no striking differences between the average values and the values predicted for the N and P proteins, both proteins conform fairly closely to the average. In the table above, No refers to the number of amino-acids predicted to be in the protein and % refers to the mole percent of the amino-acid predicted to be in the protein. The table below contains the sum of the masses of the amino-acids for the N and P proteins, as well as the number of basic and acidic amino-acids predicted to be in each protein. The N protein is expected to be acidic and the P protein is expected to be nearly neutral (Pusrurs et al., 1983; Gloor and Chaconas, unpublished). The N protein (18%) is considerably less charged than either the average protein (25%) or the P protein (28%). However, the number of aliphatic amino-acids in N (43%) does not significantly vary from either the average protein (39%) or the P protein (39%).

SDS-polyacrylamide gels and the predicted molecular mass is 41 732 da. This is well within experimental error for determining molecular mass on this gel system

The predicted amino-acid sequences of the N and P proteins were scanned for the presence of any potential DNA binding domains conforming to either the helix-turn-helix motif (Pabo and Sauer, 1984) or to the zinc finger motif (Berg, 1986).

Sequences near to the carboxyl terminus of the N and the middle of the P protein were found to contain the consensus sequence which Pabo and Sauer (1984) found in many DNA binding proteins. These amino-acid sequences are underlined in Fig. 3.6. A comparison of the consensus sequence and the sequence of the Mu A, B, c, N and P proteins is shown in Fig 3.8. Positions of the protein conforming to Pabo and Sauer's consensus sequence are underlined in the figure. The Ala and Gly residues at positions 5 and 9 respectively are highly conserved, the residue at position 15 may be either Ile or Val. In addition there are strong tendencies toward certain amino-acid types at other positions. Positions 1-3, 6-7,11-14, and 16-17 tend to be polar amino-acids. Positions 4-5, 8, 10, 15, 18 and 19 tend to be aliphatic amino-acids. In the DNA binding proteins of known structure the hydrophilic residues tend to be solvent exposed, while the aliphatic amino-acids tend to be buried in the interior of the protein (Pabo and Sauer, 1984).

CONSENSUS SEQUENCE

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
P	P	P	A	Ala	P	P	A	Gly	A	P	P	F	P	Ile	P	P	A	A	
				Gly										Val					

gpN
Val-Ile-Leu-Thr-<u>Ala-Ser-Glu</u>-Gln-<u>Gly-Phe</u>-Ser-Thr-Asp-Ser-Val-Gln-Leu-Thr-Gln-Leu

gpF
Val-<u>Lys-Gly</u>-Tyr-<u>Ala-Arg</u>-Ala-Asn-<u>Gly-Ala</u>-Glu-Gly-Asp-Asp-Ile-Asp-Ala-Lys-Ser-Ile

gpC
Gln-<u>Lys-Lys</u>-Glu-<u>Gly</u>-Val-<u>Lys</u>-Gly-<u>Gly-Lys</u>-Ala-Val-<u>Glu-Tyr-Asp-Val</u>-Met-<u>Ser-Met-Pro</u>-Thr

gpA (35-55)
<u>Arg</u>-Trp-<u>Arg-Ala-Gly</u>-Val-<u>Lys-Gly-Gly</u>-Lys-Ala-Ile-<u>Glu-Tyr</u>-Asn-Ala-<u>Asn-Ser-Leu-Pro</u>-Val

gpB
Phe-<u>Lys-Gln-Ile-Ala</u>-Leu-<u>Glu</u>-Ser-<u>Gly-Leu</u>-Ser-Thr-Gly-Thr-Ile-Ser-Ser-Phe-Ile-Asn

FIGURE 3.6: POTENTIAL NUCLEIC ACID BINDING DOMAINS OF gpN AND gpP

The sequence of the N and P gene products both contain regions of homology to the helix-turn-helix motif in some DNA binding proteins. The consensus sequence is shown with highly conserved amino-acids indicated, less conserved regions are indicated as: P indicating polar amino-acids predominate; A, indicating that aliphatic amino-acids predominate. The amino-acids in the proteins that conform to the consensus sequence are underlined. The assignment of amino-acids as polar or nonpolar is based on Lehninger (1975). The figure is redrawn from Pabo and Sauer (1984).

DISCUSSION

OVERPRODUCTION

The Mu N and P genes have been subcloned into a potent expression vector. Clones have been recovered which produced each of the N and P proteins to about 5% of total cell protein. The expression of the N gene depended on the position of the presumed N start codon; in general, as the tac promoter and start codon were brought closer more protein was produced. N protein production went from undetectable levels on SDS-polyacrylamide gels, to the major protein observed in some constructs. The amount of P protein produced by the various constructs was nearly independent of the plasmid recovered. P protein production was induced only slightly, in the presence of IPTG. This was probably due to the presence of a promoter reading out of the N gene into P, that was previously inferred (Gloor and Chaconas, 1986; Howe, 1987). The tac promoter could not be completely repressed in a lacI^q strain, as genetically detectable amounts of the N gene product were observed in the absence of inducer.

All the clones produced an N gene product that was functional in complementing MuNam phage mutants in vivo. Therefore, attempts were made to purify the N protein from the overproducing isolates.

The N protein was found to be insoluble when cells induced to make protein were lysed; this was observed for all overproducing

clones including those in the vector pCGV2. The insoluble N protein
could be solubilized in >6 M urea; when the urea was removed the
protein was found to be soluble. However, the solubilized protein
was extremely susceptible to proteolysis, producing a proteolysis
product of ≈50-55 kDa. Further purification of the protein was
attempted by column chromatograpy of the solubilized protein.
However, multiple forms were observed when chromatography of the
protein was performed. These multiple chromatographing forms could
be unified by completely reducing the protein with DTT or BME. This
probably indicates that the protein folded incorrectly in vivo
allowing random links to be formed between the proteins cysteine
residues. Complete reduction of these residues therefore allowed
one conformation of the protein to predominate in solution.

The partially-purified N protein was tested for DNA binding
activity as well as the ability to modify the in vitro strand
transfer reaction (Craigie et al., 1985; Surette et al., 1987). The
protein exhibited no activity in vitro. This indicates that the
predominant conformation assumed by the reduced protein is not the
functional one. Since the protein is capable of assuming multiple
conformations in vivo this may indicate that the correct folding of
this protein requires a template or cofactor for correct folding to
occur.

SEQUENCING

The complete DNA sequence of the region of the Mu genome coding for the N and P genes has been determined in both strands. Two long open reading frames were found, completely contained in the derived sequence, when it was translated. These frames were assumed to code for the Mu N and P gene products. The reading frame for the N gene was confirmed by sequencing of two Nam mutants that truncate the protein. The P gene reading frame was inferred by sequencing of a deleted P gene that produces a truncated gene product, when translated in vitro.

The gene order for the area of the Mu genome sequenced has previously been determined to be: Y, N, P, Q (O'Day et al., 1979). The region of the Mu genome between the second HindIII and second PstI sites had previously been shown to contain the Mu N and P genes (Schumann et al., 1980). Mutants in the first deletion group of the N gene and the last deletion group of the P gene were used in assigning the genes to this region. Thus confirming that the majority of the N and P genes lie between these two sites.

The derived amino-acid sequence of the N gene predicts that the N protein will have a molecular mass of 51 518 da. However, there is compelling evidence that the full length sequence for the N gene is as presented. First, the plasmid pGG3A9, which contains Mu DNA within 25 bp of the proposed start site, complements Nam phage. Second, this plasmid produces a full length protein as determined

on SDS-polyacrylamide gels. Third the sequence of both strands at
the carboxyl terminus of the *N* gene is unambiguous and identical.
Fourth, the nucleotide sequence of Mu*N*am7106 and Mu*N*am1995 phage
was determined. These amber mutants truncate gp*N*, and produce
protein products on SDS-polyacrylamide gels with molecular masses
of 54 and 55 kDa respectively (Gloor and Chaconas, 1986; Chapter
2). The size of the truncated products is consistent with the
amount of N predicted to be deleted from the carboxyl terminus.
Taken together this evidence rules out the possibility that the *N*
gene is larger than indicated, and argues that the protein's
migration on SDS-polyacrylamide gels is an intrinsic property of
the protein.

The *N* gene start codon can be assumed to be one of the first 2
in frame ATG codons in the *N* open reading frame, since the third in
frame ATG codon is much further into the gene. The first ATG codon
in the *N* open reading frame has a slightly more favorable
Shine-Dalgarno sequence than the second although the former
seqeunce is not particularly strong (Kozak, 1983). The context of
both the first ATG codons is also unfavorable (Hui et al., 1984).
The *N* gene product is expressed rather poorly and there is no
compelling evidence that one of the first 2 ATG codons is the
correct translation initiation site. The precise start site will
have to be determined when a purified active protein is obtained.

The initiation codon for the P protein can be inferred. The
translation product of the *P* gene open reading frame has a

predicted mass of 41,738 da. which is very close to the mass of 43 kDa observed on SDS-polyacrylamide gels. The next ATG codon in this frame is 75 amino-acids away; the translation product starting at this codon would be approximately (41,738 da - (75 X ≈110 da) = 33 488 da) which is much smaller than the mass observed. Since the vast majority of proteins migrate near their true mass on SDS-polyacrylamide gels, this would be a less likely mass for the protein than that predicted by initiation at the first ATG codon in the reading frame.

The *N* and *P* gene products both contain a region of homology to many site specific DNA binding proteins which bind in the major grove of the DNA. The N protein is known to bind to the ends of the infecting Mu DNA very tightly (Harshey and Bukhari, 1983; Gloor and Chaconas, unpublished). This tight binding implies multiple contacts between the N protein and the DNA, probably much tighter than can be achieved with one DNA binding domain. The protein does not contain any amino-acid sequence indicative of a zinc finger motif DNA binding protein. Perhaps the tight association of the N protein with Mu DNA is the result of wrapping the DNA around the protein (or vice versa). The role of the potential site specific DNA binding domain remains to be established. It should be noted that the position of the N protein bound at the ends of the protein-DNA complex is not known. Available experimental evidence does not allow us to discriminate between binding to the ends of the DNA or binding at a site near to the ends. Grundy and Howe (1985) found that Mu *Nam* phage produced normal full heads but no

phage tails. This rules out the possibility that the potential DNA binding domain has a role in packaging of the phage DNA. This result also implies that the N protein becomes bound to the phage DNA following packaging, perhaps as late as during the passage of the DNA through the tail.

The potential DNA binding site in the Mu *P* gene product is more enigmatic. The P protein is known to be a structural part of the phage tail, no other functions have been assigned to this protein (Grundy and Howe, 1985).

The genes coding for the Mu N and P proteins were found to overlap by 8 nucleotides. The P promoter is known to be located between the second *Eco*RV and first *Hpa*I site in the sequence determined (Margolin, Rao, and Howe personal communication). This places the *P* gene promoter in the carboxyl terminal 1/3 of the *N* gene. The Mu genome is not a genome that has been selected for compactness, since there is about 2 kbp of non-utilized DNA that is packaged in the phage head, located at the right end (Bukhari et al., 1976). This implies that the small overlap in these genes serves some function, perhaps in regulation of the level of production of the *N* or *P* gene products.

During the course of this sequencing project 3 different enzymes were utilized to sequence from double stranded DNA templates. The enzymes were: Klenow fragment of DNA polymerase I, AMV reverse transcriptase and Sequenase. Sequenase was found to

give the most satifactory results. The Klenow fragment was observed
to read GC rich regions poorly generating multiple bands in all
four lanes in these regions. AMV reverse transcriptase tended to
terminate in AT rich regions, especially following a run of T's.
This could be circumvented by adding a small amount of Klenow to
the chase as described in Materials and Methods. In addition,
reverse transcriptase was found to be slower than the other
enzymes, often requiring reaction times of >30 minutes for 200
bases to be read off the gel. Sequenase was not found to have the
drawbacks of the other two enzymes and future double stranded
sequencing should be done with this enzyme.

CHAPTER 4

IN VITRO TRANSPOSITION OF THE N PROTEIN-Mu DNA COMPLEX

INTRODUCTION

The transposition pathway followed by a given Mu DNA molecule
is controlled in cis by the source of the DNA. Integration of
infecting Mu DNA proceeds via a conservative transposition
mechanism (Liébart et al., 1982; Akroyd and Symonds, 1983; Harshey,
1984) and generates simple insertions (Chaconas et al., 1983).
These end products are radically different than the replicative
transposition end products formed during the Mu lytic cycle
(Chaconas et al 1980, 1981b). One of the major differences between
infecting and lysogenic Mu DNA is the presence of a 64 kDa protein
bound non-covalently to the heterogeneous ends of the former
(Harshey and Bukhari, 1983; Puspurs et al., 1983). This protein has
been shown to be the product of the Mu *N* gene (Gloor and Chaconas,
1986; Chapter 2). The N protein-Mu DNA complex is extremely
stable; Harshey and Bukhari 1983) found that it was resistant to 6
M NaCl. The complex was demonstrated to be non-covalent since it
could be disrupted either by SDS, or by high concentrations of urea
or guanidine-HCl. The N protein is a candidate for the *cis*
controlling factor which is active during Mu integration.

Figure 4.1 shows the DNA reactions involved in two potential
models for conservative transposition. As described in Chapter 1,
the Shapiro model of replicative transposition (Shapiro, 1979)
predicts that the ends of the transposon are nicked and the nicks
are ligated to a staggered, double stranded break found at the
target site. Subsequently, specific replication of the transposon
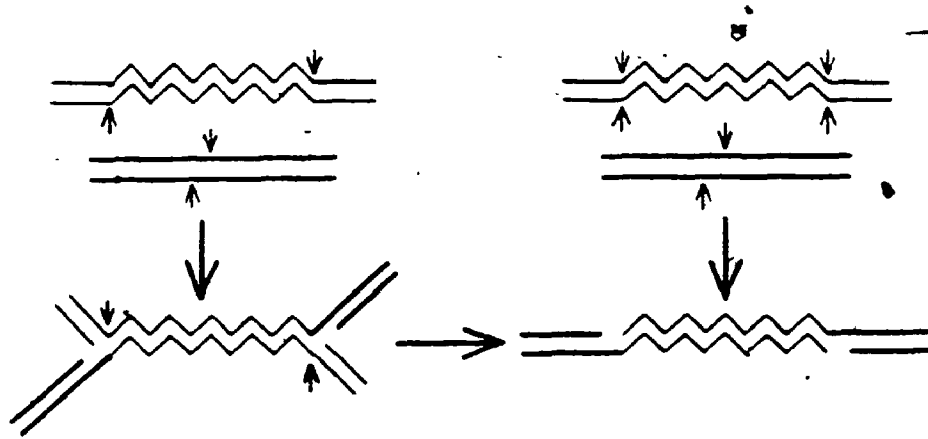
FIGURE 4.1 TRANSPOSITION MODELS OF N PROTEIN-Mu DNA COMPLEX

A conservative simple insertion into a target site could arise
from two dissimilar transposition mechanisms. In the two step model,
shown on the left, one strand at each end of the transposon is nicked
and ligated to the cut target DNA. This generates a typical Shapiro
strand transferred product. Subsequently, the opposite strand in the
strand transferred product is nicked to complete the reaction. In the
one step model, shown on the right, both strands at the end of the
transposon are cut and ligated to the cut target DNA. The end
product, a conservative, simple insertion, that has not been gap
repaired, is shown on the bottom right. The thin lines represent the
flanking DNA at the original transposon site, the thick lines
represent the target DNA and the wavy lines represent the transposon
(Mu) DNA sequences. The small arrows show the positions of the single
stranded cuts necessary for the reactions to occur. This figure is
based on the transposition model of Shapiro (1979).

generates two copies of the transposon, one of which is inserted
into a new site. Conservative transposition would proceed if the
Shapiro transposition strand transfer product could be processed
nonreplicatively, for example, by nicking near the ends of the
transposon in the strand opposite to the original nick site
(Ohtsubo et al., 1981). This second nick could be introduced
either before any replication or shortly after replication has
begun, to produce either a true conservative transposition or a
primarily conservative transposition, respectively. Another method
that may be used to generate conservative transpositions would be a
concerted double stranded cut (Berg, 1977) at the ends of the
transposon followed by joining of the transposon to the target
site. Experiments done to date *in vivo* do not allow us to
discriminate between the concerted and alternate processing
conservative transposition models. Indeed, rapid processing of a
Shapiro strand transferred product *in vivo* would be difficult to
distinguish from concerted double stranded cuts.

The two proposed conservative transposition models make
contrasting predictions. The model where the Shapiro transposition
intermediate is processed predicts that, in the donor complex and
the strand transferred end product, the ends of the Mu genome will
be found attached to either the host sequences at the ends of Mu or
to the target DNA. The double stranded cut model predicts that,
following integration of Mu into the target molecule, one of the
phage's DNA strands will be unattached to the target, since the
latter is opened by a staggered cut (see Fig. 4.1).

An in vitro transposition system for Mu has been described (Mizuuchi, 1983). This system uses the Mu transposition proteins, gpA and gpB, and a crude DNA replication extract (Fuller et al., 1981) to supply host factors required to catalyze the transposition of a donor DNA molecule. The donor plasmid contains functions necessary for plasmid maintenance as well as the reactive Mu DNA transposition sequences found at the Mu ends, but is deleted for many of the internal phage sequences. These donor plasmids are termed mini-Mu's. The target molecule in this system is double stranded øX174 RF DNA although other double stranded plasmids may be used. This crude system permits the joining of the Mu sequences to the target site, followed by further processing of the strand-transferred products (i.e. replication, ligation, etc). Furthermore, when DNA replication is blocked, a product accumulates whose structure is identical to that predicted by Shapiro (1979) for the product of the strand-transfer reaction in transposition (Craigie and Mizuuchi, 1985; Miller and Chaconas, 1986). The donor molecule used by Mizuuchi gave both replicative and conservative transposition end products. In our hands with a different donor molecule, only replicative end products have been detected (Gloor and Chaconas, unpublished).

The in vitro system was subsequently refined (Craigie et al., 1985) so that the initial strand transfer reaction could proceed in the absence of host proteins required for DNA replication. This refined system requires the Mu A and B proteins along with the E. coli HU protein. Cofactors for this reaction are Mg^{++} and ATP

(or ATPγS; Maxwell et al., 1987). The donor DNA for the reaction must be supercoiled and carry the Mu left (L) and right (R) ends in the proper orientation. The target DNA must be double stranded but may be linear, open circular or supercoiled. IHF is able to stimulate the reaction under conditions of reduced supercoiling (Surette and Chaconas, unpublished).

The supercoiling requirement of the reaction is not for the strand transfer itself but rather for the formation of an active intermediate in the reaction, a type 1 transpososome. The type 1 transpososome consists of the Mu A and probably E. coli HU proteins tightly bound to the Mu ends thus bringing them together. The donor molecule in this structure has an open circular non-Mu domain and a supercoiled Mu domain (Surette et al., 1987). Magnesium ions are required for the reaction to proceed. Type 1 transpososomes may be preformed and subsequently chased to strand transferred products by addition of the Mu B protein, a target DNA molecule and ATP (or ATPγS). Strand transferred products that contain the bound transposition proteins are referred to as type 2 transpososomes. The strand transferred products may be replicated by addition of a crude DNA replication extract (Craigie and Mizuuchi, 1985b).

The N protein-Mu DNA complex is assumed to be the integrative precursor in the insertion of Mu into the host chromosome for the following reasons. First, this protein-DNA complex accumulates in host cells lysogenic for Mu (Ljungquist and Bukhari, 1979; Harshey and Bukhari, 1983) and in cells infected with MuAam mutants (Gloor

and Chaconas, unpublished, this thesis) which are defective in integration (O'Day et al., 1978). However, only small, transient levels of this protein-DNA complex are found when integration is allowed to proceed following infection (Ljungquist and Bukhari, 1979; Harshey and Bukhari, 1983). Second, in vitro studies have shown that transposition of Mu requires a supercoiled DNA substrate (Craigie et al., 1985). The N protein-Mu DNA complex can be isolated in supercoiled form from infected cells (Harshey and Bukhari, 1983; Puspurs et al., 1983; Gloor and Chaconas, unpublished, this thesis). In addition, the infecting protein-DNA complex is capable of protecting the infecting Mu DNA from host exonucleases. Chase and Benzinger (1982) showed that a non-covalent 65 kDa protein-DNA complex isolated from freeze-thawed Mu particles was resistant to E. coli exonuclease V. Harshey and Bukhari (1983) demonstrated, in vitro, that the isolated protein-DNA complex was resistant to both T7 exonuclease and E. coli exonuclease III which are 5' and 3' specific exonucleases, respectively. The isolated protein-DNA complex has been shown to transfect E. coli spheroplasts approximately 3 orders of magnitude more efficiently than deproteinized Mu DNA (Chase and Benzinger, 1982; Harshey and Bukhari, 1983).

The N protein-Mu DNA integrative precursor might be expected to behave differently in vitro than a covalently closed circular mini-Mu molecule. Specifically, the integrative precursor might be expected to insert into the target molecule by the concerted double stranded cut model of conservative transposition (Berg, 1977). The

ability to form double stranded cuts at the ends of Mu would reflect an intrinsic ability of the protein-DNA complex to transpose conservatively.

The N protein-Mu DNA complex was tested for its ability to form a strand transferred end product *in vitro*. Cells infected with Mu*Aam* phage were chosen as the source of the protein-DNA complex since they do not integrate in Mu sensitive cells (O'Day et al., 1978). Infected cells were gently lysed by several freeze-thaw cycles and the protein-DNA complex was isolated by sucrose gradient sedimentation. Supercoiled fractions were utilized as the donor molecules in the defined *in vitro* Mu strand transfer system (Craigie et al., 1985). The strand transferred products of the reaction were analyzed for similarities and differences to those of previously characterized mini-Mu derivatives (Craigie and Mizuuchi, 1985b; Surette et al., 1987).

MATERIALS AND METHODS

PREPARATION OF BACTERIOPHAGE Mu FOR INFECTION

Bacteriophage Mu was labeled with ^{32}P similarly to the preparation method described by Chaconas et al. (1983). Briefly, *E. coli* cells lysogenic for Mu (GC419 *lac, thr, leu, fhuA, suII, Mu^R, D108^R, Mucts71Aam3011*; Chaconas et al., 1985) were grown overnight at 30° in LB broth; 1 ml of cells were then pelleted and washed twice with 0.85% NaCl. The cells were resuspended in 0.1 ml of 0.85% NaCl and added to 25 ml of PFB, prepared as described (Chaconas et al., 1985b), except that CaCl$_2$ was omitted. The cells were grown to a density of about 2.5 x 10^8/ml and 0.5 mCi of [^{32}P]-orthophosphate (NEN; approximately 1 in 2 Mu genomes contain 1 ^{32}P molecule under these labeling conditions) was added. The cell culture was incubated at the same temperature until it had doubled, at which time it was shifted to 42.5° for 30 minutes and then placed at 38° until lysis was complete. The cells were always grown with vigorous aeration to ensure maximal cell growth and phage production.

The lysed cells were placed on ice and cell debris was pelleted by centrifugation for 10-15 minutes at 4°, at 10 000 rpm in a Beckman J2-21 centrifuge using a JA-20 rotor. Phage Mu was pelleted out of the supernatant fraction by centrifugation at 16 500 rpm for 2 hours at 4° in the same rotor and centrifuge. The phage pellet was overlaid with a small volume (<1 ml/25 ml culture)

of Mu buffer (25 mM Tris-HCl pH 7.5, 25 mM $MgSO_4$, 150 mM NaCl) and incubated at 4° for at least 2 hours. The phage pellet was gently resuspended and titered on the Su^- $E.$ $coli$ strain, CA274, at 10^{-6} dilution and the isogenic Su^+ strain, CA74 ($Hfr.$ $lac,$ $trp,$ $relA,$ $spoT,$ $suI;$ Brenner and Beckwith, 1965; obtained from the $E.$ $coli$ genetic stock center), at 10^{-7}, 10^{-8} and 10^{-9} dilutions. The titre of phage was usually greater than 10^{11} pfu/ml.

PURIFICATION OF THE N PROTEIN - Mu DNA COMPLEX :

The $E.$ $coli$ strain 40 ($\Delta pro/lac,$ $trp,$ $rpsL;$ Bukhari and Metlay, 1973) was grown in LB broth + 25 mM $MgSO_4$ and 1 mM $CaCl_2$ to a density of approximately 5×10^8 cells/ml at 37°. Bacteriophage Mu, prepared as described above, was added in drops over 1 minute to a multiplicity of infection between 20 and 30, while the culture was being swirled in the water bath to thoroughly mix the cells and the phage. The infected cells were removed from the water bath and incubated at 22°, without mixing, for 15 minutes, followed by incubation at 37° with aeration for a further 20 to 30 minutes. The infected cells were pelleted in a JA-20 rotor in a Beckman J2-21 centrifuge at 10 000 rpm for 5 minutes at 4°. The cells were vigorously resuspended in 10 ml of ice cold buffer 1 (10 mM Tris-HCl pH 8.2, 10 mM EDTA, 10% sucrose) and pelleted as above. This step was repeated once. The second pellet was resuspended in 10 ml of ice cold buffer 2 (10 mM Tris-HCl pH 8.2, 10 mM EDTA, 50 mM NaCl) and pelleted as above. The pellet was

resuspended in a final volume of 0.9 ml of buffer 2 per 12.5 ml of
original cell culture.

The cell suspension was lysed by adding to each 0.9 ml aliquot:
1 μl of 100 mM PMSF (Sigma), and 0.1 ml of 10 mg/ml egg white
lysozyme (Sigma), and was immediately placed at -20°. The cells
were frozen and thawed twice. After the second freeze-thaw cycle,
100 μl of 10% Triton X-100 (Bio-Rad) was added. Again, the cells
were frozen and thawed at -20° and centrifuged for 10 minutes in a
microfuge at 4°. It should be noted that at least one of the
freeze-thaw cycles was left for an extended period of time (>8 _____
hours). This seemed to improve the yield substantially, although
further extended incubations at -20° had no effect.

The supernatant fraction from the lysed cells was loaded onto a
5%-20% sucrose gradient prepared in buffer 3 (25 mM Tris-HCl pH
8.2, 1 mM EDTA, 0.1 mM PMSF/containing 1 M NaCl. The sucrose
gradient was either run in a Beckman SW-50.1 (0.15 ml of
supernatant, 45 000 rpm, 2 hours, 4°) or an SW-27 (1.0 ml of
supernatant, 25 000 rpm, 5 hours, 4°) rotor. Approximately thirty
fractions were collected by bottom puncture and Cerenkov counts
were determined. When the peak fractions were to be concentrated,
the complex was dialyzed overnight against buffer 3 containing 0.5
M NaCl, to remove the sucrose.

Concentration of the purified complex was performed using a
Conflit 250/HF concentrator, as described by the manufacturer.

Briefly, the concentrating apparatus was assembled and washed 3 times with deionized water. Five ml of the pooled and dialyzed peak fractions were loaded into the concentrator and the apparatus was spun for 10-12 minutes at about 1 000 x g at 4° in a tabletop centrifuge. Further loadings of the concentrator were done if the original volume was greater than 5 ml. The concentrated fractions were recovered in a volume less than 0.5 ml with recovery being greater than 90%.

TRANSPOSITION OF THE PURIFIED PROTEIN-DNA COMPLEX *IN VITRO*

Transposition reactions were performed essentially as described by Surette et al. (1987). The Mu A protein was prepared by either Mike Surette or Subba Reddi Palli of this lab, the *E. coli* HU protein was purified by Mike Surette (Surette et al., 1987) and the Mu B protein was purified by Dr. George Chaconas (Chaconas et al., 1985a). *E. coli* IHF protein was prepared by Mike Surette from an overproducing strain supplied by Dr. H. Nash (K5746; Nash et al., 1987).

The reactions contained variable amounts of the N protein-Mu DNA complex, depending on the preparation, but these amounts were never saturating. Reactions were run in 25 mM Tris-HCl pH 7.5, 10 mM MgCl$_2$, 140 mM NaCl. The final concentrations of the other components were: gpA 7 μg/ml, gpB 5 μg/ml, HU 3 μg/ml, target DNA 20 μg/ml, ATP 2 mM. IHF was used in the reactions noted at about about 350 ng/ml. When ATPγS replaced ATP, the Mu B protein was

preincubated with the target DNA and 25 uM ATPγS for 15 minutes, on ice, in order to allow the B protein to bind to the target DNA (Maxwell et al., 1987). The concentrations of the components of this binding reaction were 10X that of the final transposition reaction. Transposition reactions were incubated for 20 minutes at 30° and terminated by the addition of 1/9 volume of either 10X EDTA loading dye (60% sucrose, 0.2 M EDTA, 0.1% bromophenol blue dye) or 10X SDS loading dye (60% sucrose, 1% SDS, 0.1% bromophenol blue dye). Any preincubations with A and/or HU and/or IHF were done for 5 minutes at 22°. Relaxation of supercoils was done by adding 1 unit of calf thymus topoisomerase 1 (BRL) to the reaction. Relaxation was complete in less than 5 minutes.

Alkaline gel analysis of the reaction products was done by the addition of 1/4 volume of alkaline loading dye (10% ficoll, 0.01% bromocresol green, 100 mM EDTA, 0.25 M NaOH). The DNA was electrophoresed on agarose gels in alkaline gel buffer (30 mM NaOH, 2 mM EDTA) at 40 v (≅150 mA) for 16 hours (Davis et al., 1983). The buffer was recirculated by a pump at a rate sufficient to exchange 1/3 of the anodic and cathodic buffer reservoirs each hour. The gels were soaked in H_2O for 30 minutes, prior to drying, to remove most of the NaOH.

RESULTS

PURIFICATION OF THE N PROTEIN-Mu DNA COMPLEX

Purification of the N protein-Mu DNA complex was attempted by several different methods. The published procedure of Harshey and Bukhari (1983) relied on low salt sucrose gradients performed at pH 10.5, followed by concentration of the DNA on 2 sequential CsCl gradients and subsequent sedimentation through a sucrose gradient in 1 M NaCl. This procedure was tedious and tended to give varying amounts of supercoiled complex in the end. The purification procedure of Puspurs et al. (1983), where the protein-DNA complex was purified out of infected minicells, proved to give lower yields compared to whole cell infections. Individual minicell preparations varied widely in their ability to produce supercoiled N protein-Mu DNA complex.

The method of choice for N protein-Mu DNA complex purification was by sucrose gradients in 1 M NaCl as described in the Materials and Methods. This method proved to be simpler and more reproducible than the published procedures. Peak fractions across a typical sucrose gradient, run from right to left, are shown in Fig 4.2, along with the ^{32}P counts across the gradient. There is good separation of supercoiled, open circular and linear Mu DNA on this gradient. Separations performed on the SW 50.1 rotor, as described in Materials and Methods, were slightly better, but the amount of complex purified was limited by the small size of the rotor.
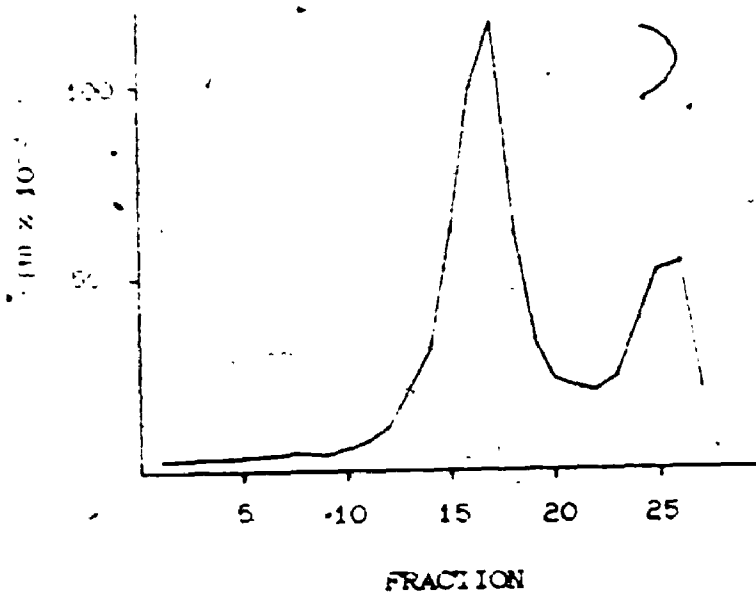
FRACTION

13 14 15 16 17 18 19 20 21 22

OC→

L →

SC→

FIGURE 42 PURIFICATION OF gpN-Mu DNA COMPLEX

The cells were lysed and run on a sucrose gradient in an SW-27 rotor as described in Materials and Methods. The graph shows the ^{32}P counts across the gradient. The peak fractions indicated, were electrophoresed on a 0.35% agarose gel, dried and autoradiographed. The migration positions of linear (L), open circular (OC) and supercoiled (SC) N protein-Mu DNA complex are indicated.

A fraction containing the N protein-Mu DNA complex was examined for the presence of nucleases. Incubation of this fraction with linear, supercoiled or single stranded DNA, in the in vitro reaction buffer at 37° for 1 hour, had no effect on the electrophoretic mobility of the input DNA (data not shown). From this, I concluded that, under these conditions, there was no detectable nuclease activity in these protein-DNA complex preparations.

BEHAVIOR OF CONTROL PLASMIDS *IN VITRO*

The refined *in vitro* Mu strand transfer system (Craigie et al., 1985) was used in all of the following experiments. This system utilizes the Mu transposition proteins to cut and join the ends of Mu to a target molecule. As mentioned above, the *in vitro* strand transfer system generates a structure consistent with the Shapiro model of transposition (Craigie et al., 1985). It was necessary to determine whether the N protein-Mu DNA donor complex was able to be utilized in this *in vitro* transposition system. Furthermore, we needed to determine if the infecting protein-DNA complex behaved similarly to a mini-Mu substrate in this reaction.

The plasmids pGG215 (a 4.15 Kb *XhoI/PvuII* fragment of pCL222 (Chaconas et al., 1981b) cloned into a 2.95 Kb *SalI/PvuII* fragment of pBR322, Gloor and Chaconas, unpublished) or pMC321 (Chaconas et al., 1981b) were used as donors. The behavior of pGG215 *in vitro* has been well documented by Surette et al., (1987), and is the

prototypical mini-Mu used in this lab to generate Shapiro type transposition strand transferred products.

The in vitro behavior of pMC321 which contains the entire Mu genome, inserted into pSC101, was compared to pGG215 (data not shown; for a diagram summarizing the reactions referred to below see Fig. 1.4). Plasmid pGG215 is known to form a protein-DNA complex - a type 1 transpososome - following incubation with the Mu A and *E. coli* HU proteins. This was also observed for pMC321. If the proteins in the type 1 transpososome are stripped off the donor plasmid with SDS, prior to electrophoresis, the resulting plasmid is found to be nicked. This occurred with both pGG215 and pMC321. Plasmid pGG215, in the form of a type 1 transpososome, may be chased to a type 2 transpososome by the addition of the Mu B protein, a target molecule and ATP (or ATPγS). A pMC321 type 1 transpososome could also be efficiently chased to a type 2 transpososome by addition of the same components. The type 1 transpososome of pGG215 contains a supercoiled Mu domain and relaxed vector domain. Relaxation of the supercoiled domain in the type 1 transpososome does not affect the ability of this structure to be chased to a type 2 transpososome for either pGG215 or pMC321. Both ends of Mu are always found attached to either the vector or target sequences except in the nicked species observed when type 1 transpososomes are deproteinized.

REQUIREMENTS FOR THE IN VITRO TRANSPOSITION REACTION UTILIZING THE
N PROTEIN-Mu DNA COMPLEX AS DONOR

Fig. 4.3 shows the relative migration positions of the starting N protein-DNA complex and the strand transferred end products formed in vitro. The electrophoretic mobility of the starting material is shown in the first lane of panel A, where the N protein-Mu DNA complex is mainly supercoiled; addition of SDS (lane 2) removes the N protein and the DNA is linearized. The supercoiled donor fraction containing both A and HU shows that the mobility of the supercoiled species is slowed slightly. Much of the DNA has been nicked by a persistent contaminant in the gpA preparation (M. Surette, personal communication). There is a minor, SDS stable product formed when only A and HU are added, and its formation varies depending on the batch of N protein-Mu DNA complex that is used. This stable product could not be chased into the strand transferred product (data not shown). The complete reaction containing the A, B and HU proteins, along with the target DNA (∅X174 RF), produces three bands. One band runs intermediate between the supercoiled and linear bands. Another runs similarly to the relaxed band, and most probably is unreacted N protein-Mu DNA complex. The slowest band runs slightly behind the relaxed band. Addition of SDS to the complete reaction results in only two bands. The band running with linear Mu represents unreacted donor complex, whereas the slowest migrating band is the strand transferred product where the N protein-DNA complex has been joined to the target DNA.
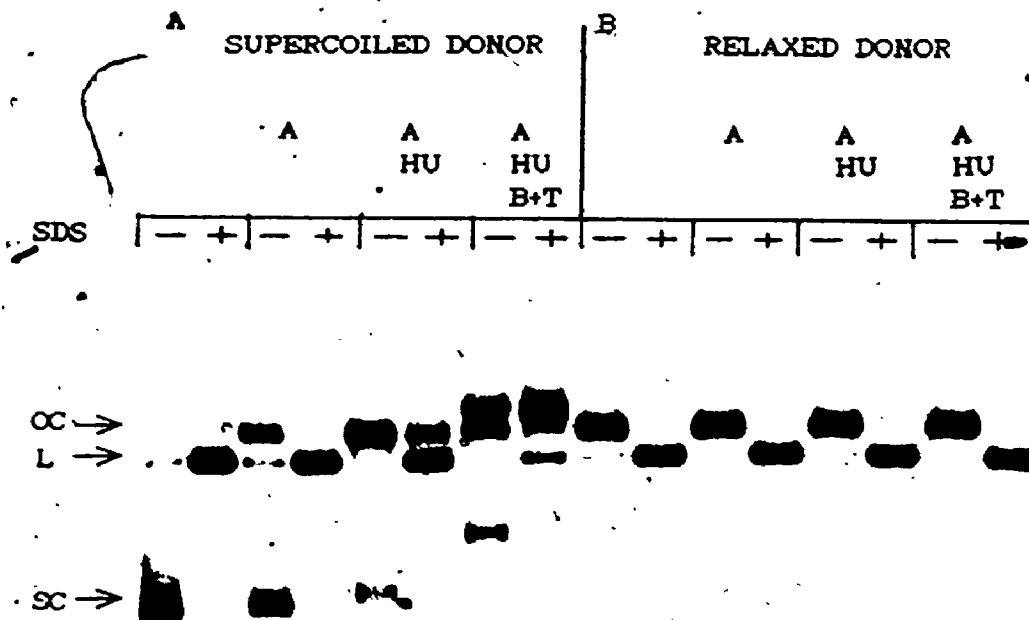
FIGURE 4.3 REQUIREMENTS FOR *IN VITRO* TRANSP...

Purified gpN-Mu-DNA complex was incubated with components of the *in vitro* reaction noted. The reactions were run as described in Materials and Methods, run on a 0.35% agarose gel, dried and autoradiographed. B+T indicates the addition of the Mu B protein and the target DNA. The first set of reactions (A) were performed with a supercoiled fraction, the second set (B) were done with a relaxed fraction off the same sucrose gradient. The lanes loaded in SDS loading dye are indicated, while the other lanes were loaded in EDTA loading dye. SDS loading dye disrupts protein-DNA interactions and allows the migration of the DNA alone to be observed. The migration positions of supercoiled (SC), open circular (OC) and linear (L) N protein-Mu-DNA complex are noted.

The experiment shown in Fig. 4.3 also details the protein and donor supercoiling requirements of the strand transfer reaction. Fractions off a sucrose gradient containing either supercoiled donor complex or relaxed donor complex are compared for their ability to drive the *in vitro* strand transfer reaction This figure illustrates that donor supercoils are required for the formation of the strand transferred end product since the supercoiled substrate is efficiently chased to the strand transferred product while the relaxed substrate is not. Furthermore, one should note that no reaction occurs in the absence of the Mu A protein

Two controls that were done are not shown in Fig. 4.3. The first control demonstrated that if the supercoiled fraction is converted to a relaxed donor complex by topoisomerase 1, prior to the addition of the transposition proteins, this fraction is no longer able to form the strand transferred product. In fact, the topoisomerase 1 relaxed supercoiled fraction behaves identically to the relaxed fraction. The second control was to run an *in vitro* reaction using pGG215 as the donor molecule, including the supercoiled or relaxed fraction at the concentration specified for the reactions shown in the figure. The mini-Mu was able to form both type 1 and type 2 transpososomes when either fraction was included. These controls demonstrated that the relaxed fraction did not contain an inhibitor of the *in vitro* reaction.

Craigie et al. (1985) demonstrated that the strand transfer reaction in vitro required a supercoiled donor molecule but that the target molecule could be supercoiled, open circular or linear. The target requirements for transposition of the N protein-Mu DNA complex were examined by using supercoiled, relaxed or linear øX174 DNA as the target. The N protein-Mu DNA complex was able to utilize all of these target molecules with approximately equal efficiency (data not shown).

TYPE 1 TRANSPOSOSOME FORMATION

This experiment illustrates that the N protein-Mu DNA complex can form a type 1 transpososome, where the A and possibly the HU proteins form a protein-DNA complex at the ends of the Mu DNA. This is the active intermediate in the mini-Mu strand transfer reaction (Surette et al, 1987). In Fig. 4.4, the donor complex has been preincubated with Mu A protein, HU and $MgCl_2$. A portion of the preincubated material was relaxed with topoisomerase I. Following relaxation, the Mu B protein and target DNA were added. The final two reactions are the ones to note in this figure. The second to last reaction was relaxed with topoisomerase I prior to addition of the B protein and target DNA while, the final reaction was not relaxed before addition of these components. The donor complex, that was relaxed after addition of the A and HU proteins, is capable of forming the strand transferred product. This reaction is indicative of type 1 transpososomes. Relaxation of the donor
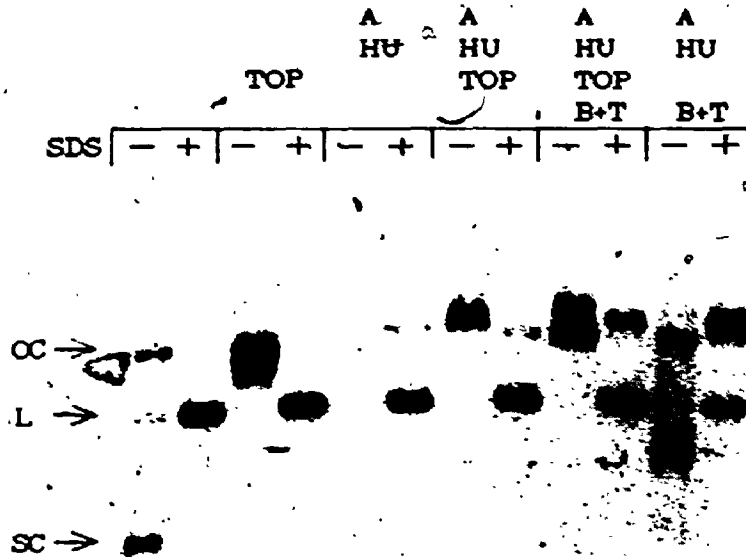
FIGURE 4.4 EVIDENCE FOR TYPE 1 TRANSPOSOMES

Purified gpN-Hu DNA complex was incubated with the components of the *in vitro* reaction noted. The reactions were run as described in the Materials and Methods, run on a 0.35% agarose gel, dried and autoradiographed. All reactions containing the A and Hu proteins were preincubated for 5 minutes. Topoisomerase 1 (TOP) was added where noted, following the preincubation with A and HU, and incubated a further 5 minutes. Following all preincubations, gpB, ATP and target DNA was added and incubation was continued for a further 20 minutes. Reactions electrophoresed in the lanes indicated were stopped by the addition of SDS loading dye, the other lanes contain parallel samples loaded in EDTA loading dye. The migration positions of supercoiled (SC), open circular (OC) and linear (L) N protein-Hu DNA complex are noted.

complex prior to addition of the A and HU proteins results in no
strand-transfer reaction occurring (shown in Fig. 4.3).

EFFECT OF IHF

The small *E. coli* protein IHF was tested for its ability to
substitute for HU in the *in vitro* strand transfer reaction since it
was found that IHF could stimulate transposition *in vitro* under
reduced levels of donor plasmid supercoiling (Surette and Chaconas,
unpublished). Fig. 4.5 shows that IHF is unable to substitute for
HU under the conditions tested, since complete reaction mixtures
containing IHF, but not HU, do not form a strand transferred
product. There was a slight decrease in the amount of the SDS
stable product formed when A and HU were incubated with the complex
when IHF was also present in the reaction.

IHF allows the strand transfer reaction to proceed *in vitro*
when the donor molecule is less supercoiled than normally required
(Surette and Chaconas, unpublished). The results shown here
illustrate that, *in vitro*, IHF has no effect on the products of the
strand transfer reaction using the protein-DNA complex as donor. An
interesting possibility is that IHF could affect Mu integration by
allowing the infecting DNA-protein complex to be integrated before
it is fully supercoiled. This would enable the infecting DNA to
integrate more rapidly and escape host mechanisms for degrading
exogenous DNA.

FIGURE 4.5 EFFECT OF IHF ON TRANSPOSITION

Purified gpN-Mu DNA complex was incubated with components of the
in vitro reaction noted. The reactions were run as described in
Materials and Methods, stopped, run on a 0.35% agarose gel, dried and
autoradiographed. IHF was added to the reactions indicated. The set
of reactions on the left were loaded in SDS loading buffer, the other
set of reactions were loaded in EDTA loading buffer. The migration
positions of linear (L), open circular (OC) and supercoiled (SC) N
protein-Mu DNA complex are shown on the left.

FATE OF THE *E. coli* ENDS *IN VITRO*

Figure 4.1 describes two alternate modes for transposition, both of which could produce a conservative, simple insertion of the transposon into the target DNA. In the first model, a Shapiro strand transferred intermediate is formed and chased into a conservative transposition by a second nick introduced in the unnicked transposon strand. The strand transferred end product which would be observed in this model would contain both Mu ends attached to either the heterogeneous host DNA or to the target molecule. Subsequent steps would process this intermediate toward a conservative transposition. The second model illustrated in Fig. 4.1 produces conservative transpositions by the concerted cutting of both strands of the transposon prior to insertion into the target molecule. The strand transferred end product in this model would contain one strand not attached to either the target molecule or to the donor molecule sequences.

The integration of Mu *in vitro* may be examined by the use of alkaline gels to separate the two DNA strands, in order to observe whether the strands are attached to the host or target DNA, or if one strand is free in the strand transferred product. Fig. 4.6 shows the results for the Mu R end. The protein-DNA complex was allowed to form either: type 1 transpososomes, where one strand is nicked by the Mu A protein; or type 2 transpososomes, where the complete strand transfer reaction has proceeded. In the absence of the transposition proteins, both strands are attached to the
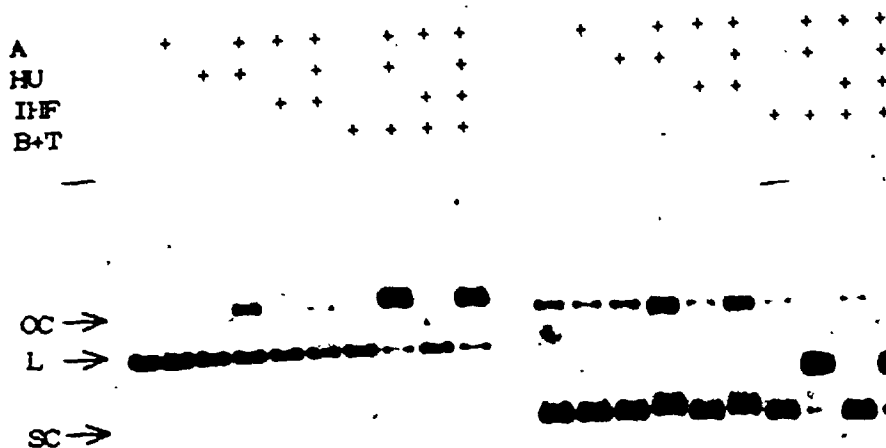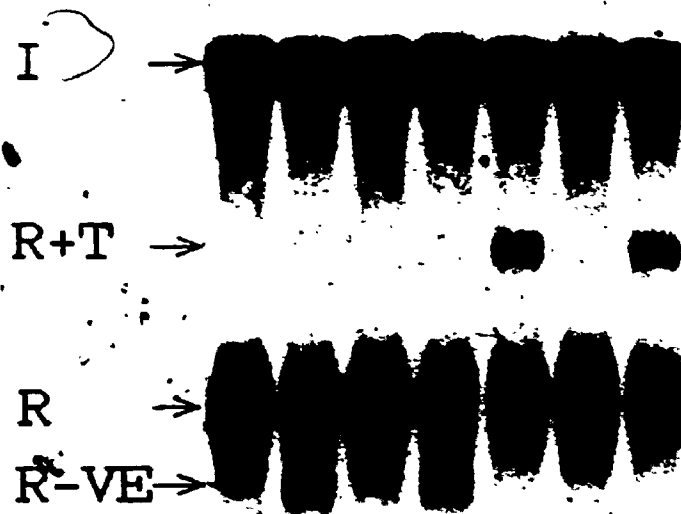
FIGURE 4.6 BEHAVIOUR OF THE RIGHT END OF Mu *IN VITRO*

Purified gp*N*-Mu DNA complex was incubated with components of
the *in vitro* reaction noted. The reactions were run as described
in Materials and Methods. Completed reactions were extracted once
with ø-CHCl₃, ethanol precipitated and resuspended in 18 ul of H₂O.
Two ul of 10X *Hind*III buffer and 5 units of *Hind*III were added; the
mixture was incubated at 37° for a further 15 minutes. Following
restriction, 6 ul of 5X alkaline loading dye was added and the
samples were electrophoresed at 40 V for 18 hours in a 0.5%
alkaline agarose gel. The gel was dried and autoradiographed. The
migration positions of the internal and right end *Hind*III fragments
are shown as I and R respectively. When the right end is nicked the
variable ends of Mu are separated from the right end Mu fragment
under denaturing conditions, this nicked right end fragment is
indicated as R-VE. The nicked strand that has been joined to the
target DNA is shown as R+T.

heterogeneous E. coli ends, as evidenced by the smeared R end band.
The reaction containing the A and HU proteins where the end of the
Mu DNA is nicked, shows that one strand is freed of the E. coli DNA
by the nick introduced at the end of the Mu genome and now migrates
as a discrete smaller band. The complete integration reaction
illustrates that, following the addition of target DNA and the Mu B
protein, both strands are again attached. One strand is still
attached to the heterogeneous E. coli sequences; the other strand
is attached to the target DNA and migrates more slowly on the gel.
In summary, this demonstrates that, in vitro, the integration of
the R end of the Mu N protein-DNA complex proceeds via formation of
a Shapiro strand transferred intermediate.

The integration of the Mu L end is examined in Fig. 4.7. The
amount of heterogeneous host DNA at the left end of Mu is much less
than that found at the right end and so resolution of the host
attached and nicked ends is much more difficult. The small amount
of host DNA necessitates cutting quite close to the left end of Mu
and high percentage gels are essential to adequately resolve the
small fragments. The Mu DNA was restricted with HindIII which cuts
about 1000 nucleotides from the Mu left end (Preiss et al., 1982)
and this figure shows the relevant portion of a 2% agarose gel run
in alkaline gel buffer to denature the DNA. The first lane shows
the migration of unreacted Mu left end sequences which migrate as a
slight smear. The second lane shows the the Mu left end sequences
found in a type 1 transpososome; there is a shift of a portion of
the DNA toward a smaller species however, the original band is also

A   +   +

HU   +   +

B+T   +

LE ⟶
L-VE ⟶

FIGURE 4.7 NICKING OF THE LEFT END OF Mu *IN VITRO*

The A protein-Mu DNA complex was incubated with the A and HU and/or IHF proteins to form type 1 transposomes where the end of Mu is nicked, in order to observe if the left end behaves in a similar manner to the right. Following the incubation with the components noted, the reactions were extracted with CHCl₃, ethanol precipitated, and resuspended in 18 ul of TE. Two ul of 10X *Hind*III buffer were added along with 5 units of *Hind*III. The DNA was restricted for 15 minutes at 37° and 6 ul of alkaline loading dye was added. The reactions were electrophoresed on a 2% alkaline agarose gel, dried and autoradiographed. The first lane shows the migration of the left end without any additions, the second lane shows the migration of the left end DNA strands following incubation with the A and Hu proteins, the third lane shows the left end DNA strands as they are found in the strand transferred end product. LE denotes the Mu left end while, L-VE denotes the left end which has been nicked thus losing the variable host sequences.

present. The third lane shows the migration of the Mu left end sequences found in a type 2 transpososome, the smaller band has disappeared, presumably linked to the target DNA.

Linkage of the Mu left end sequences to the target DNA in the type 2 transpososome is confirmed in the next figure. Fig. 4.8 demonstrates that the left end is attached to the target DNA when the complete strand transfer reaction is performed. In this figure, the Mu DNA was cleaved with *Bg*III following the strand transfer reaction, this enzyme produces two fragments, the smaller of which is the left end. In the complete reaction mixture, the left end is covalently joined to the target molecule (lanes A and C), and therefore runs more slowly. This figure also illustrates that in the absence of all the requirements for the complete reaction (the second lane contains no HU), no linkage of the left end and target molecule occur.

These figures together show that the left end behaves similarly to the right end *in vitro*, that is the both the left and right ends are nicked in the type 1 transpososome and the nicked species are then joined to the cut target DNA.
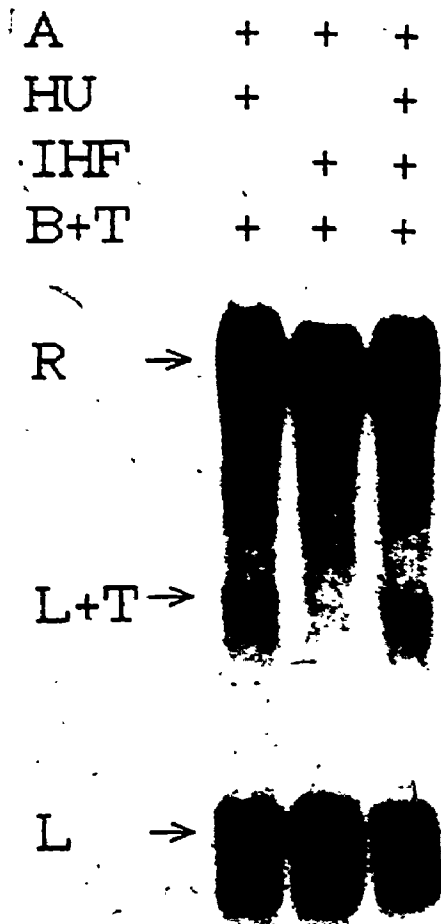
FIGURE 4.8 JOINING OF THE LEFT END AND TARGET DNA *IN VITRO*

Purified gpN-Mu DNA complex was incubated with components of the *in vitro* reaction noted. The reactions were run as described in Materials and Methods. Reactions were extracted once with ϕ-CHCl$_3$, ethanol precipitated and resuspended in 18 ul of H$_2$O. Two ul of 10X BgIII buffer and 25 units of BgIII was added; the mixture was incubated at 37° for 15 minutes. Following the restriction reaction, 6 ul of 5X alkaline loading dye was added and the samples were electrophoresed at 40 V for 18 hours in a 0.5% alkaline agarose gel. The gel was dried and autoradiographed. The migration positions of the left and right end BgIII fragment are shown as L and R respectively. The nicked left end strand that has been joined to the target DNA is shown as L+T.

DISCUSSION

A working hypothesis described in the introduction to this
chapter was the possibility that the N protein-Mu DNA integrative
precursor could possess the intrinsic potential to transpose
conservatively by the introduction of double stranded cuts at the
ends of the Mu DNA. However, the experimental evidence provided in
this chapter shows that the N protein-Mu DNA complex was joined to
the target molecule in a manner consistent with the Shapiro (1979)
model of transposition. That is, by the introduction of nicks at
the ends of the Mu DNA followed by the joining of the nicked ends
to a double stranded cut target molecule.

The strand transferred product observed for mini-Mu's *in vitro*
possesses replication potential, there is a free 3' hydroxyl group
at the end of the target molecule that is not attached to the
transposon (Mizuuchi, 1984) and the mini-Mu strand transferred
product may be replicated by the addition of crude *oric* type
replication extracts (Craigie and Mizuuchi, 1985b). It therefore
follows that the strand transferred product observed with the
protein-DNA complex as the donor molecule should also have
replication potential. The N protein-DNA complex was observed to be
unstable in these crude extracts due to the large amount of
nucleases present. *In vivo* integrating Mu DNA is observed to be
replicated to a minor extent, 4-8% of infecting Mu DNA is
replicated; however, complete replication of the integrated DNA is
not observed (Harshey, 1984). Chaconas et al. (1985b) found that a

mutant of the Mu B protein lacking the carboxyl terminal 18 amino-acids could integrate at normal levels but could not subsequently replicate the inserted phage DNA. This implies that these residues on the B protein are intimately involved in the replication of Mu DNA, but not the integration of Mu. Mutations at other sites of gpB drastically reduce integration. These carboxyl terminal residues are not specifically removed during infection, thus indicating that they may be blocked by some other protein. The amount of DNA replication observed *in vitro* is consistent with both the concerted model followed by repair or the processing model with replication being stopped at an early point before the genome is replicated.

It should be noted that *in vivo* Mu DNA replication is believed to be asymmetric, beginning at the left, and proceeding to the right end (Wijffelman and van de Putte, 1977; Goosen, 1978; Higgins et al., 1983). This asymmetric replication could be a consequence of the A protein binding sites at the left and right ends of Mu being considerably different; there are two tight A protein binding sites at the left end and one at the right end (Craigie et al., 1984). This implies that there are differences in protein content and organization at the two Mu ends during transposition. Recently, Craigie and Mizuuchi (1987) found that a linear molecule, containing one copy of the Mu right end at each terminus, could be efficiently utilized in the *in vitro* strand transfer reaction. Similar constructs with either the two left ends or one left and one right end at the termini of the linear molecule were not

active. These results suggest that the requirements for transposition of the Mu left end are more stringent than those for the right end. The in vitro results of Craigie and Mizuuchi (1987) when taken with the replication results described above, suggest that the Mu left end is the controlling end in transposition and that the function required of the right end is the ability to be recognized by the transposase.

If one assumes that the left end is the end necessary for efficient replication as well as much of the specificity of the reaction, a further assumption would be that the factor responsible for the switch to conservative transposition would act at or near the Mu left end. The Mu N protein is less than 150 bp away from the Mu left end in the infecting protein-DNA complex (George and Bukhari, 1981). One should recall that there is a protein-DNA complex, including the Mu A and B proteins, located at the ends of the strand transferred product (Surette et al., 1987). Given the tight interaction between the N protein and the ends of the infecting Mu DNA (Harshey and Bukhari, 1983; Gloor and Chaconas, unpublished), one could imagine lengthy contacts between the DNA and the N protein extending mear to the Mu left end. Alternatively, the N protein could act as a multimer allowing it to touch Mu proteins located at the left end. Therefore it is not unreasonable to assume that the N protein could be within range of contacting the Mu B protein located at the left end during infection. This contact could block the carboxyl terminus of the B protein and therefore mimic the situation found with the mutant B

protein lacking the C-terminal eighteen amino acids by preventing DNA replication to proceed and forcing integration to proceed conservatively.

A further arguement for the N protein being involved in the switch to conservative transposition is a temporal one. Ideally, the controlling factor should be present only during integration so that later replicative events are not interfered with. The N protein is injected along with the infecting DNA and is not synthesized until late in the phage life cycle, thus fitting this predicted pattern.

A further consideration that should be addressed is the DNA requirements for conservative transposition. Dubow and Lalumiere (Plasmid, in press) examined the fate of mini-Mu molecules containing progressively less heterologous DNA between the Mu ends. They observed that when the ends were separated by less than 406 bp of DNA that > 90% of the transposition events observed were simple insertions. When separated by 750 bp, 30-40% simple insertions were observed and with 1.6 kbp of DNA separating the ends, 10-20% simple insertions were observed. The protein-DNA complex used in the experiments described in this chapter contain between 1-2 kbp of DNA between the Mu ends due to the attached heterogeous host DNA (Bukhari et al., 1976). This would imply that there is a limited intrinsic propensity for Mu integration to proceed conservatively, generating simple insertions, since with this amount of DNA between the ends <10% of the transpositions should be simple insertions.

However, 100% of the infecting Mu DNA integrates conservatively
(Harshey, 1984) generating simple insertions. (Chaconas et al.
1987) therefore, there must be a further mechanism in operation to
generate the observed frequency.

The last issue that must be addressed is the control of the
specificity of the second nick required to produce a conservative
simple insertion. The nonreplicated strand transferred product is
not likely to be randomly attacked, at the junction of Mu and
target sequences, by host nucleases for two reasons: First, random
attack of the strand transferred product by an endonuclease would
result in the excision of the integrated Mu DNA in a large
percentage of the integration events, this can be verified by
examination of Fig. 4.1. Second, the type 2 transpososome is
insensitive to digestion by S1 nuclease (Surette and Chaconas,
unpublished), whereas the deproteinized DNA is. This indicates that
the proteins forming the type 2 transpososome shield the sensitive
single stranded DNA at the Mu-target DNA junction. Inspection of
Fig. 4.1 shows that the strand that would be have to be nicked in
Mu to produce a conservative transposition would be the lagging
strand, if replication is assumed to be initiated from the target
DNA's free hydroxyl group. Perhaps, here is a role for one of the
semi-essential early genes of Mu. If replication of the strand
transferred product is blocked, an accessory protein with the
necessary specificity to produce the nicks required for
conservative processing of the Shapiro intermediate could be
activated. In the presence of replication, this opportunity would

not occur and replicative transposition end products would be formed. This hypothesis predicts that a Mu gene product that gives a low level of integration but a normal burst size when a Mu lysogen is induced, could be found. Alternatively, an unreplicated lagging single strand could be degraded by host enzymes. This would be followed by repair of the leading strand to produce a simple insertion.

In summary, the experiments described in this chapter show that the *in vitro* strand transfer reaction, utilizing the N protein-Mu DNA complex, produces Shapiro type strand transferred products. These strand transferred products are similar to those formed by both a mini-Mu and a whole Mu control plasmid. The strand transferred end product formed with the protein-DNA complex as donor molecule, has both strands of Mu DNA attached to either the heterogeneous host sequences or to the target molecule. Therefore, it is likely that Mu integration is forced toward conservative transposition by a replication block rather than by the production of concerted double stranded cuts at the ends of Mu. It is possible that the Mu N protein could act by inhibiting replication, thus allowing nucleases to cut the strand transferred products and produce conservative simple insertions of the infecting Mu DNA.

SUMMARY

The purpose of this thesis was to characterize the 64-kDa protein-Mu DNA complex isolated from Mu-infected cells. This was achieved by examining the protein bound to the ends of the infecting Mu DNA as well as by examining the behaviour of the protein-DNA complex in an in vitro transposition system

Work reported in this thesis described the identification of the injected 64-kDa Mu virion protein as the Mu *N* gene product. This identification was achieved by a combination of physical and genetic means.

The Mu *N* gene was further characterized by the determination of its sequence; the adjacent *P* gene was also sequenced. The amino-acid sequence of both proteins was observed to contain a region of homology to the helix-turn-helix domain found in many DNA binding proteins (Pabo and Sauer, 1984). The semi-purified N protein was not observed to bind DNA *in vitro*. Both proteins were overproduced to greater than 5% total cell protein. An activity for the N protein could not be detected *in vitro*.

The infecting protein-DNA complex was examined for the product of the strand-transfer reaction formed *in vitro*, to determine if the observed conservative integration of Mu DNA into the host chromosome, occurs by double-strand breaks at the end of the transposon or via processing of a Shapiro strand-transferred product. *In vitro* the infecting protein-DNA complex was observed to form a normal strand-transferred product analogous to that formed

by mini-Mu derivatives of Mu. This result implies that the *in vivo* reaction also occurs via processing of a Shapiro strand-transferred product. The mechanism of this processing could not be examined in vitro due to the instability of the protein-DNA complex in the crude extracts necessary for the analysis.

The work described in this thesis should provide a good basis for further studies on Mu integration. Site-specific mutagenesis of the Mu *N* gene could be utilized to dissect the regions of the protein required for phage morphology from those required for integration and DNA binding, thus alleviating the genetic invisibility problem observed with this gene. In addition, the strand-transferred product recovered using the N protein-Mu DNA complex could be used in a purified replication system to observe the effect (if any) the N protein has on the ability of this complex to replicate. Furthermore, if it is possible to obtain a fraction of the N protein that exerts a measureable activity, interactions between the N and B proteins could be examined.

REFERENCES

146

1. Adler, H. I., Fisher, W. d., Cohen, A. and Hardigree, A. A. (1965) Miniature *Escherichia coli* cells deficient in DNA. *Proc. Natl. Acad. Sci.* 57: 321-326.

2. Admiraal, G. and Mellema, J. E. (1976) The structure of the contractile sheath of bacteriophage Mu. *J. Ultrastruct. Res.* 56: 48-64.

3. Ahmed, A. (1985). A rapid procedure for DNA sequencing using transposon-promoted deletions in *Escherichia coli*. *Gene* 39: 305-310.

4. Ahmed, A. (1987) *Meth. Enzym.* 154: In Press

5. Akroyd, J. and Symonds, N. (1983) Evidence for a conservative pathway of transposition of phage Mu. *Nature* 303: 84-86.

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

7. Allet, B. (1979) Mu insertion duplicates a 5 base pair sequence at the host inserted site. *Cell* 16: 123-129.

8. Arthur, A. and Sherratt, D. J. (1979) Dissection of the transposition process: A transposon-encoded site-specific recombination system. *Mol. Gen. Genet.* 175: 267-274.

9. Batteiger, B., Newhall, W. J. and Jones, R. B. (1982) The use of Tween-20 as a blocking agent in the immunological detection of proteins transferred to nitrocellulose membranes. *J. Immunol. Methods* 55: 297-307.

10. Berg, D. E. (1977) Insertion and excision of the transposable kanamycin resistance determinant in Tn5. In *DNA insertion elements, plasmids and episomes* (ed. A. I. Bukhari, J. A. Shapiro and S. L. Adhya) pg. 205-212. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

11. Berg, J. M. (1986) Potential metal-binding domains in nucleic acid binding proteins. *Science* 232: 485-487.

12. Biggin, M. D., Gibson, T. J. and Hong, G. F. (1983) Buffer gradient gels and [^{35}S]-label as an aid to rapid DNA sequence determination. *Proc. Natl. Acad. Sci.* 80: 3963-3965.

13. Birnboim, H. C. and Doly, J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* 7: 1513-1523.

14. Brenner, S. and Beckwith, J. R. (1965) *Ochre* mutants, a new class of suppressible nonsense mutants. *J. Mol. Biol.* 13: 629-637.

15. Brosius, J., Dull, T. J., Sleeter, D. D. and Noller, H. F. (1981) Gene organization and primary structure of a ribosomal RNA operon from *E. coli*. *J. Mol. Biol.* *148:* 107-127

16. Brosius, J. (1984) Toxicity of an overproduced foreign gene product in *Escherichia coli* and its use in plasmid vectors for the selection of transcription terminators. *Gene 27:* 161-172

17. 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. 72:* 4399-4403.

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

19. Bukhari, A. I. and Metlay, M. (1973) Genetic mapping of prophage Mu. *Virology 54:* 109-116.

20. Chaconas, G., Gloor, G. and Miller, J. L. (1985a) Amplification and purification of the bacteriophage Mu encoded *B* transposition protein. *J. Biol. Chem 260:* 2662-2669.

21. Chaconas, G. Harshey, R. M. and Bukhari, A. I. (1980) Association of Mu-containing plasmids with the *Escherichia coli* chromosome upon prophage induction. *Proc. Natl. Acad Sci. 77:* 1778-1782.

22. Chaconas, G., Kennedy, D. L. and Evans, D. (1983) Predominant integration end-products of infecting Mu DNA are simple insertions with no preference for integration of either Mu DNA strand. *Virology 128:* 48-59.

23. Chaconas, G., Giddens, E. B., Miller, J. L. and Gloor, G. (1985b) A truncated form of the bacteriophage Mu B protein promotes conservative integration, but not replicative transposition, of Mu DNA. *Cell 41:* 857-865.

24. Chaconas, G., Harshey, R. M., Sarvetnick, N. and Bukhari, A. I. (1981a) Predominant end-products of prophage Mu DNA transposition during the lytic cycle are replicon fusions. *J. Mol. Biol. 150:* 341-359.

25. Chaconas, G., Gloor, G., Miller, J. L., Kennedy, D. L., Giddens, E. B. and Nagainis, C. R. (1984) Transposition of bacteriophage Mu DNA: Expression of the A and B proteins from λP_L and analysis of infecting Mu DNA. *Cold Spring Harbor Symp. Quant. Biol. 49:* 279-284.

26. Chaconas, G., de Bruijn, F. J., Casadaban, M. J., Lupski, J. R., Kwoh, T. J., Harshey, R. M., DuBow, M. S. and Bukhari, A. I. (1981b) *In vitro* and *in vivo* manipulations of bacteriophage Mu DNA: Cloning of Mu ends and construction of mini-Mu's carrying selectable markers. *Gene 13:* 37-46.

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

28 Chen, E. J. and Seeburg, P. H. (1985) Supercoil sequencing: A fast
    simple method for sequencing plasmid DNA. *DNA 4:* 165-170.

29 Cleveland, D. W., Fischer, S. G., Kirschner, M. W. and Laemmli, U. K.
    (1977) Peptide mapping by limited proteolysis in sodium
    dodecyl sulphate and analysis by gel electrophoresis.
    J. Biol. Chem 252: 1102-1106.

30 Coehlo, A., Maynard-Smith, S. and Symonds, N. (1982) Abnormal
    cointegrate structures mediated by gene *B* mutants of phage
    Mu: Their implications with regard to gene function. *Mol.
    Gen. Genet. 185:* 356-362.

31. Craigie, R. and Mizuuchi, K. (1985a) Cloning of the *A* gene of
    bacteriophage Mu and purification of its product, the Mu
    transposase. *J. Biol. Chem 260:* 1832-1835.

32. Craigie, R. and Mizuuchi, K. (1985b) Mechanism of transposition
    of bacteriophage Mu: Structure of a transposition
    intermediate. *Cell 41:* 867-876.

33. Craigie, R. and Mizuuchi, K. (1986) Role of DNA topology in Mu
    transposition: Mechanism of sensing the relative
    orientations of the two DNA segments. *Cell 45:* 793-800.

34. Craigie, R. and Mizuuchi, K. (1987) Transposition of Mu DNA:
    Joining of Mu to target DNA can be uncoupled from cleavage
    at the ends of Mu. *Cell 51:* 493-501.

35. Craigie, R., Arndt-Jovin, D. J. and Mizuuchi, K. (1985) A defined
    system for the DNA strand-transfer reaction at the
    initiation of bacteriophage Mu transposition: Protein and
    DNA substrate requirements. *Proc. Natl. Acad. Sci. 82:*
    7570-7574.

36. 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.

37. Dayhoff, M. O., Hunt, L. T. and Hurst-Calderone, S. (1978) in
    Atlas of Protein Sequence and Structure (ed. M. O. Dayhoff)
    Vol. 5, Suppl. 3: 363-369

38. Daniell, E., Abelson, J., Kim, J. S. and Davidson, N. (1973)
    Heteroduplex structures of bacteriophage Mu DNA. *Virology
    51:* 237-239.

39. 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.

40. Davis, R. W., D. Botstein, and J. R. Roth, eds. (1980) *Advanced bacterial genetics.* Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

41. deBoer, H. A., Comstock, L. J. and Vasser, M. (1983) The *tac* promoter: A functional hybrid derived from *trp* and *lac* promoters. *Proc. Natl. Acad. Sci.* 80: 21-25

42. de Jong, W. W., Zweers, A. and Cohen, L. H. (1978) Influence of single amino-acid substitutions on electrophoretic mobility of SDS-protein complexes. *Biochem Biophys. Res. Comm* 82: 532-539.

43. Dubow, M. S. and Lalumiere (1987) The construction of a mini-Mu plasmid whose transposition outcomes are not cointegrates. *Plasmid* (in press).

44. Derbyshire, K. M. and Grindley, N. D. F. (1986) Replicative and conservative transposition in bacteria. *Cell* 47: 325-327.

45. Faelen, M. and Toussaint, A. (1973) Isolation of conditional defective mutants of temperate phage Mu-1 and deletion mapping of the Mu-1 prophage. *Virology 54:* 117-124.

46. Faelen, M., Huisman, O. and Toussaint, A. (1978) Involvement of phage Mu-1 early functions in Mu-mediated chromosomal rearrangements. *Nature 271:* 580-582.

47. Ferguson, B., Jones, N., Richter, J. and Rosenberg, M. (1984) Adenovirus *E1a* gene product expressed at high levels in *Escherichia coli* is functional. *Science 224:* 1343-1346.

48. Freifelder, D. (1982) *Physical Biochemistry*, pg. 504-506, Freeman Publications, San Francisco.

49. Fuller, R. S., Kaguni, J. M. and Kornberg, A. (1981) Enzymatic replication of the origin of the *Escherichia coli* chromosome. *Proc. Natl. Acad. Sci.* 78: 7370-7374.

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

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

52. 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.

53. 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.

54. Goosen, T. (1978) Replication of bacteriophage Mu: Direction and possible location of the origin. in *DNA synthesis: Present and future* p. 121-126. (ed. I. Molineux and M. Kohiyama). Plenum Press, New York.

55. Goosen, T., Giphart-Gassler, M. and van de Putte, P. (1982) Bacteriophage Mu DNA replication is stimulated by non-essential early functions. *Mol. Gen. Genet.* 186: 135-139.

56. Grindley, N. D. F. and Reed, R. R. (1985) Transpositional recombination in prokaryotes. *Ann. Rev. Biochem.* 54: 863-895.

57. Groenen, M. A. M., Timmers, E. and van de Putte, P. (1985) DNA sequences at the ends of the genome of bacteriophage Mu essential for transposition. *Proc. Natl. Acad. Sci.* 82: 2087-2091.

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

59. Grunstein, M. and Hogness, D. S. (1975) Colony hybridization: A method for the isolation of cloned DNA's that contain a specific gene. *Proc. Natl. Acad. Sci.* 72: 3961-3965.

60. Hannahan, D. (1983) Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* 166: 557-580.

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

62. Harshey, R. M. and Bukhari, A. I. (1983) Infecting bacteriophage Mu DNA forms a circular protein-DNA complex. *J. Mol. Biol.* 167: 427-441.

63. Heinrich, P. (1986) *Guidlines for quick and simple plasmid sequencing.* Boehringer-Mannheim technical manual, Boehringer Mannheim GmbH Biochemica, Mannheim, West Germany.

64. Helfman, D. M., Feramisco, J. R., Fiddes, J. C., Thomas, G. P. and Hughes, S. K. (1983) Identification of clones that encode chicken tropomyosin by direct immunological screening of a cDNA library. *Proc. Natl. Acad. Sci.* 80: 31-35.

65. Higgins, N. P., Moncecchi, D., Manlapaz-Ramoz, P. and Olivera, B. M. (1983) Bacteriophage Mu DNA replication *in vitro.* *J. Biol. Chem.* 258: 4293-4297.

56. Howe, M. M. (1973) Prophage deletion mapping of bacteriophage Mu-1. *Virology* 54: 93-101.

67. Howe, M. and Bade, E. G. (1975) Molecular biology of bacteriophage Mu. *Science* 190: 624-632.

68. Howe, M. M. (1987) Phage Mu: An overview. in *Phage Mu* (eds. N. Symonds, A. Toussaint, P. van de Putte and M. M. Howe) pg. 25-39. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

69. Hsu, M. T. and Davidson, N. (1974) Electron microscope heteroduplex study of the heterogeneity of Mu phage and prophage DNA. *Virology* 58: 229-239.

70. Hui, A., Dinkelspiel, K. and DeBoer, H. A. (1984) Mutagenesis of the three bases preceeding the start codon of the β-galactosidase mRNA and its effect on translation in *Escherichia coli*. *EMBO J.* 3: 623-629.

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

72. Iida, S., Meyer, J. and Arber, W. (1983) Prokaryotic IS elements. In *Mobile genetic elements* (ed. J. A. Shapiro), pg. 159-221. Academic Press, New York.

73. Kahmann, R., Rudt, F. and Mertens, G. (1984) Substrate and enzyme requirements for *in vitro* site-specific recombination in bacteriophage Mu. *Cold Spring Harbor Symp. Quant. Biol.* 49: 285-294.

74. Kamp, D., Kahmann, R., Zipser, D., Broker, T. R. and Chow, L. T. (1978) Inversion of the G DNA segment of phage Mu controls phage infectivity. *Nature* 271: 577-580.

75. Kozak, M. (1983) Comparison of initiation of protein synthesis in prokaryotes, eukaryotes and organelles. *Microbiol. Rev.* 47: 1-45.

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

77. Leach, D. and Symonds, N. (1979) The isolation and characterization of a plaque-forming derivative of bacteriophage Mu carrying a fragment of Tn3 conferring ampicillin resistance. *Mol. Gen. Genet.* 172: 179-184.

78. Liebart, J. C., Ghelardini, P. and Paolozzi, L. (1982) Conservative integration of bacteriophage Mu DNA into pBR322 plasmid. *Proc. Natl. Acad. Sci.* 79: 4362-4366.

79. Ljungquist, E. and Bukhari, A. I. (1979) Behavior of bacteriophage Mu DNA upon infection of *E. coli* cells. *J. Mol. Biol. 133:* 339-357.

80. Maniatis, T. and Efstradiatis, A. (1980) Fractionation of low molecular weight DNA in polyacrylamide gels containig 98% formamide or 7 M urea. *Meth. Enzym. 65:* 299-305.

81. Maniatis, T., E.F. Fritsch and J. Sambrook, eds. (1982) *Molecular cloning: a laboratory manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

82. Marrs, C.R. and Höwe, M.M. (1983) *Ava*II and *Bgl*II restriciton maps of bacteriophage Mu. *Virology 126:* 563-575.

83. Maxwell, A., Craigie, R. and Mizuuchi, K. (1987) B protein of bacteriophage Mu is an ATPase that preferentially stimulates intermolecular DNA strand transfer. *Proc. Natl. Acad. Sci. 84:* 699-703.

84. McBeth, D.L. and Taylor, A.L. (1982) Growth of bacteriophage Mu in *Escherichia coli dnaA* mutants. *J. Virol. 44:* 555-564.

85. McBeth, D.L. and Taylor, A.L. (1983) Involvement of *Escherichia coli* K12 DNA plymerase I in the growth of bacteriophage Mu. *J. Virol. 48:* 149-156.

86. McClintock, B. (1950) The origin and behavior of mutable loci in Maize. *Proc. Natl. Acad. Sci. 36:* 344-355.

87. Miller, J.H. ed. (1972) *Experiments in molecular genetics.* Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

88. Miller, J.L. and Chaconas, G. (1986) Electron microscopic analysis of *in vitro* transposition intermediates of bacteriophage Mu DNA. *Gene 48:* 101-108.

89. Miller, J.L., Anderson, S.K., Fujita, D.J., Chaconas, G., Baldwin, D. and Harshey, R.M. (1984) The nucleotide sequence of the B gene of bacteriophage Mu. *Nucleic Acids Res. 12:* 8627-8638.

90. Mise, K. (1971) Isolation and characterization of a new generalized transducing bacteriophage different from P1 in *Escherichia coli. J. Virol. 7:* 168-175.

91. Mizuuchi, K. (1983) *In vitro* transposition of bacteriophage Mu: A biochemical approach to a novel replication reaction. *Cell 35:* 785-794.

92. Mizuuchi, K. (1984) Mechanism of transposition of bacteriophage Mu: Polarity of the strand transfer reaction at the initiation of transposition. *Cell 39:* 395-404.

93. Nash, H. A., Robertson, C. A., Flamm, E., Weisberg, R. A. and Miller, H. I. (1987) Overproduction of *Escherichia coli* integration host factor, a protein with nonidentical subunits. *J. Bact. 169:* 4124-4127.

94. O'Day, K. J., Schultz, D. W. and Howe, M. M. (1978) Search for integration-deficient mutants of bacteriophage Mu. In *Microbiology: 1978* (ed. D. Schlessinger), p. 48-51. American Society for Microbiology, Washington, D. C.

95. O, Day, K. J., Schultz, D. W., Ericson, W., Rawluk, L. and Howe, M. M. (1979) Correction and refinement of the genetic map of bacteriophage Mu. *Virology 93:* 320-328.

96. Ohtsubo, E., Zenilman, M., Ohtsubo, H., McCormick, M., Machida, C. and Machida, Y. (1981) Mechanism of insertion and cointegration by IS*1* and Tn*3*. *Cold Spring Harbor Symp. Quant. Biol. 45:* 283-295.

97. Pabo, C. O. and Sauer, R. T. (1984) Protein-DNA recognition. *Ann. Rev. Biochem. 53:* 293-321.

98. Preiss, H., Kamp, D., Kahmann, R., Brauer, B. and Delius, H. (1982) Nucleotide sequence of the immunity region of bacteriophage Mu. *Mol. Gen. Genet. 186:* 315-321.

99. Puspurs, A. H., Trun, N. J. and Reeve, J. N. (1983) Bacteriophage Mu circularizes following infection of *Escherichia coli*. *EMBO J. 2:* 345-352.

100. Queen, C. (1983) A vector that uses phage signals for efficient synthesis of proteins in *Escherichia coli*. *J. Mol. Appl. Genet. 2:* 1-10.

101. Sanger, F., Micklen, S. and Coulson, A. R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. 74:* 5463-5467.

102. Shapiro, J. A. (1979) Molecular model for the transposition and replication of bacteriophage Mu and other transposable elements. *Proc. Natl. Acad. Sci. 76:* 1933-1937.

103. Schumann, W., Bade, E. G., Forgie, R. A. and Howe, M. M. (1980) Cloning of DNA fragments of the right end of phage Mu and location of the *Hind*III, *Sal*I, *Pst*I and *Bam*HI restriction sites on the genetic map of Mu. *Virology 104:* 418-425.

104. Silhavey, T. J., Berman, M. L. and Enquist, L. W. (1984) *Experiments with Gene Fusions* pg. 95-96, Cold Spring Harbor Laboratory, cold Spring Harbor, New York.

105. Smith, B. J. and Nicolas, R. H. (1983) Analytical electrophoretic techniques in protein chemistry. In *Techniques in molecular biology* (ed. J. M. Walker and W. Gaastra) pg. 25-48. Croom Helm Ltd. Beckenham, Great Britain.

106. Surette,, M. G., Buch, S. J. and Chaconas, G. (1987) Transposasomes: Stable protein-DNA complexes involved in the *in vitro* transposition of bacteriophage Mu DNA. *Cell 49:* 253-262.

107. Taylor, A. L. (1963) Bacteriophage-induced mutation in *E. coli*. *Proc. Natl. Acad. Sci. 50:* 1043-1051.

108. Toussaint, A., Desmet, L., Faelen, M., Alazard, R., Chandler, M. and Pato, M. (1987) *In vivo* mutagenesis of bacteriophage Mu transposase. *J. Bact. 169:* 5700-5707.

109. Toussaint, A. and Faelen, M. (1974) The dependence of temperate phage Mu-1 upon the replication functions of *E. coli* K12. *Mol. Gen. Genet. 131:* 209-214.

110. Toussaint, A. and Résibois, A. (1983) Phage Mu: Transposition as a life style. In *Mobile Genetic Elements* (ed J. A. Shapiro), pg 105-158. Academic Press, New York.

111. Towbin, H., Staehelin, T. and Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. *Proc. Natl. Acad. Sci. 76:* 4350-4354.

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

113. Wijffelman, C. A., Westmaas, G. C. and van de Putte, P. (1973) Similarity of vegetative map and prophage map of bacteriophage Mu. *Virology 54:* 125-134.

114. Wijffelman, C. A., Gassler, M., Stevens, W. F. and van de Putte, P. (1974) On the control of transcription of bacteriophage Mu. *Mol. Gen. Genet. 131:* 85-96.

115. Wijffelman, C. and Lotterman, B. (1977) Kinetics of Mu DNA synthesis. *Mol. Gen. Genet. 151:* 169-174.

116. Wijffelman, C. A. and van de Putte, P. (1977) Asymmetric hybridization of Mu strands with short fragments synthesized during Mu DNA replication. In *DNA insertion elements, plasmids and episomes* (ed A. I. Bukhari, J. A. Shapiro and S. L. Adhya) pg. 329-333. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

117. Williams, J. G. K. and Radding, C. M. (1981) Partial purification and properties of an exonuclease inhibitor induced by bacteriophage Mu-1. *J. Virol. 39:* 548-558.

118. Zehnbauer, B. A. and Markovitz, A. (1980) Cloning of gene *kim* (*capR*) and identification of polypeptides specified by the cloned deoxyribonucleic acid fragment. *J. Bact.* 143: 852-863.

119. Zubay, G. (1973) In vitro synthesis of proteins in microbial systems. *Ann. Rev. Genet.* 7: 267-287.
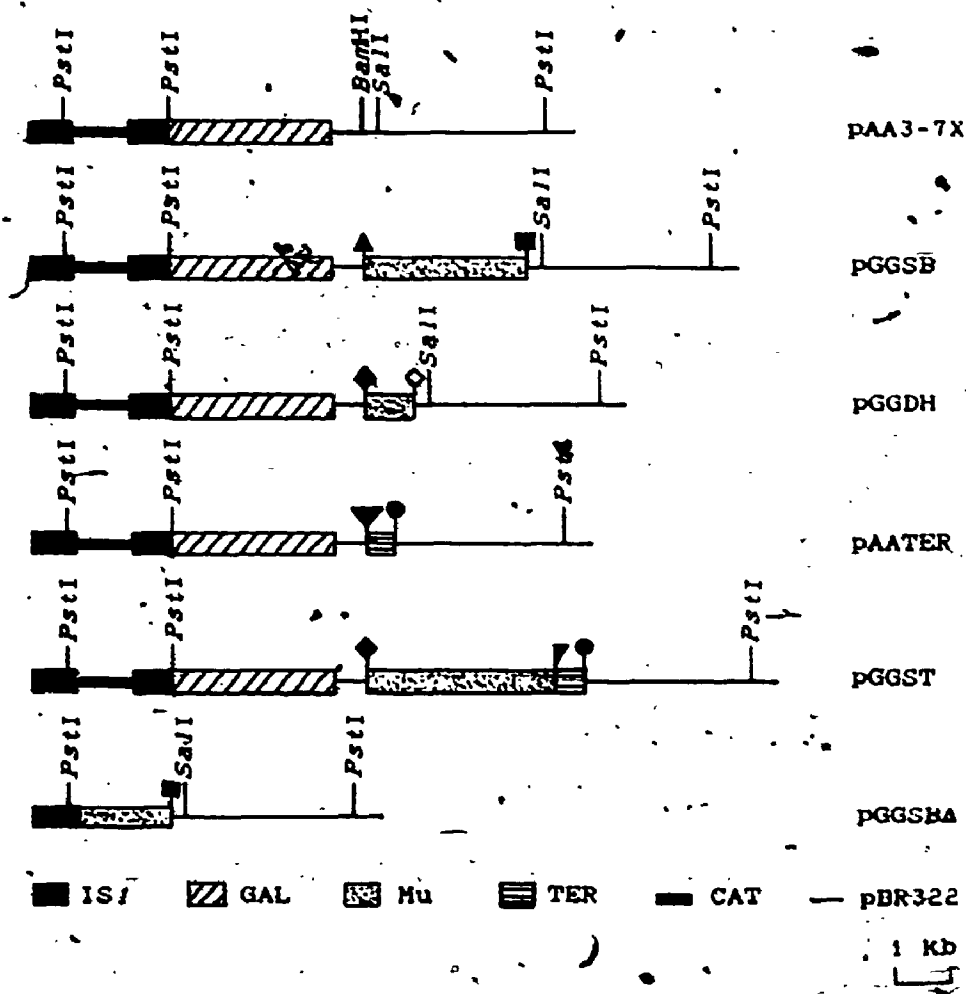
FIGURE 3.4: PLASMIDS UTILIZED TO SEQUENCE THE Mu *N* AND *P* GENES

A restriction map of the plasmids utilized to sequence the Mu *N* and *P* genes is shown. The parent vector for all the constructs was pAA3-7X and is the top plasmid. Plasmids containing Mu DNA that were directly derived from this vector (pGGSB, pGGDH) are shown below. Plasmid pAATER, a derivative of pAA3-7X containing a strong terminator is shown next. The multiple cloning site in pAATER contains (l-r) the recognition sites for the following restriction enzymes: BamHI, SalI, PstI and HindIII. Plasmid pGGST, the sequencing derivative of this vector is shown below pAATER. The final construct shown has the typical structure of a deleted plasmid from which sequence in the cloned DNA could be obtained. Intervening DNA between the active IS*I* module and the cloned DNA has been deleted allowing a primer complementary to IS*I* to read into the cloned DNA. The origin shown for the restriction maps is the left end of the active IS*I* module on the plasmid, from which the deletions originate. Symbols shown are:

▲, BamHI-PstI junction sequence; ■, EcoRI-BamHI junction sequence;
◆, BamHI-DraI junction sequence; ◇, HpaI-BamHI junction sequence;
▼, multiple cloning site; ●, SspI-SalI junction site;
▽, SalI-multiple cloning site junction (SalI site present)