ISAS - INTERNATIONAL SCHOOL
FOR ADVANCED STUDIES

Towards an improvement of phage libraries through the study of phage-bacteria interaction in two different systems.

Thesis submitted for the degree of

"Doctor Philosophiæ"

Candidate

Antonella Longo

Supervisors

Dott. Andrew Bradbury

Prof. Antonino Cattaneo

January 1996

SISSA - SCUOLA
INTERNAZIONALE
SUPERIORE
DI STUDI AVANZATI
TRIESTE
Strada Costiera 11

TRIESTE

2. Materials and methods. 67

1. Introduction.

Starting from the pioneering work by Smith (Smith, 1985), phage libraries have been widely used as one of the more promising available techniques to study protein-protein interaction. Fusion phage have been shown to be extremely useful for displaying peptide sequences (continuous epitopes) that may react with antibodies or a foreign protein. Fusion phage have also been constructed which display whole proteins such as antibodies.

There are however several limitations in using such fusion phage to identify altered peptides or proteins with new or enhanced binding properties, limits which will be discussed in section 1.3.1.1.

One of the limits of this technique is due to the procedure used to select the binding phage. The library is passed over the peptide or the protein bound to a solid support. This means that the peptide or the protein has to be purified or synthesized before it can be coupled to the solid support.

Starting from the observation that phage infect bacterial cells by binding to a bacterial protein which is used as an adhesion site, an alternative method has been developed and tested in our laboratory (Bradbury *et al.*, 1993). Bacteria expressing the antigen on their surfaces, within an outer membrane protein, LamB, are used to affinity purify

phage antibodies from a phage library. This method allows the selection to take place at the bacterial cell surface by simply cloning the peptide in a vector, without further purification. An obvious extension of this new approach is the possibility for the phage to infect the bacterial cell by simply binding to the peptide cloned in a bacterial outer membrane protein. If it were possible to clone an epitope on the phage-adsorption protein and the corresponding antibody on the phage-receptor protein, or vice versa, it may be possible to create a new port of entry for the phage. This would be more similar to the antibody binding to immune cells bearing the antigen and starting the chain of events of immunological response. In other words it would be possible to select the antibody by a simple infection event.

To reach this aim a good knowledge of both the phage-absorption protein and the phage-receptor protein is needed. It is worth noting that all type of molecules or structures accessible on the bacterial surface (lipopolysaccharide, protein, pili, flagella, etc.) can be exploited by a bacteriophage to establish the first contact with the bacterial cell. In other words any protein on the bacterial surface is a phage receptor for one or more than one phage. In order to realize the aim of linking the selection with the infection, it would thus be possible to choose among several couples of phage/bacterial protein.

In this thesis two of these possible couples have been analyzed with the final scope of engineering selection by infection: lambda phage/LamB and filamentous phage/pilus. While the phage-absorption protein for the Ff filamentous phage, gene III protein (gIIIp), is well known, the receptor protein on the F pilus is unknown. On the contrary much work has been done on the receptor protein of

phage lambda, LamB, but its absorption protein, J, has not been characterized in detail. Both gIIIp and LamB have been used to display epitopes and whole proteins, antibodies included. The available information about the two systems are both deficient in one part. My thesis deals with the first attempts to adapt the two unknowns (J and F/N adhesin) to a scheme of infection-selection.

I have modified the J protein of lambda by adding an epitope and an antibody to its C-terminus. I show that this site is not suitable for being used as a display vector. The choice of another site is made difficult by the fact that the knowledge of the structure of the J protein is very poor. Even the publication of the sequence of the host-range mutants (Werts *et al.*, 1994), which are localized in the *J* gene, have not clarified which are the residues involved in the interaction with the phage receptor, LamB (discussed in 1.1.3.2. and 3.1.5.).

Following this I have decided to concentrate on filamentous phage. A first step was the identification of the bacterial receptor for the phage. Although F pilus has been the object of a large amount of studies (Frost *et al.*, 1994), a putative adhesin on the tip of the pilus has not been identified. Furthermore there are theoretical and potential practical problems with the use of a homologous phage/bacterial pair (described in section 1.2.4.). Basing my work on experiments carried out in our laboratory (Marzari *et al* submitted) which show that fd phage can infect N pili when provided with the appropriate IKe receptor (discussed in section 1.2.4.) I decided to undertake studies at the N pilus with the final goal of identifying the structural protein(s) forming the pilus. I have obtained the almost complete sequence of the transfer region coding for the proteins having a role in pilus

assembly and structure. I have shown that the sequence of such proteins is homologuous to that of two pili belonging to different incompatibility groups (section 3.2.3.3.), and to that of other export systems (section 3.2.3.4.). Some hypothesis about which proteins can form the pilus structure are formed (section 3.2.3.5.).

Section 1.1. will deal with the basic biology of lambda phage. I will then focus on the mechanism of entry of the phage into its host cell, with particular emphasis on the proteins that have been shown to be involved in this process.

In section 1.2. the filamentous phage will be described, as well as their interaction with the particular bacterial structures they use for infection: pili

A short overview of the phage display technology is reported in section 1.3., with a particular emphasis to new developments in this field.

1.1. The lambda phage.

Lambda bacteriophage is an obligatory parasite: the phage particle itself is metabolically inert. Lambda was originally isolated from a culture of *Escherichia coli* K-12, which happened to be lysogenic for this phage (Lederberg, 1951). Since this discovery, lambda phage was usually propagated on nonlysogenic derivative of *E. coli* K-12 or more rarely in *E. coli* C (Bertani & Weigle, 1953), in strains of *Shigella flexneri* (Gemski *et al.*, 1972), and in hybrids between *E. coli* and *Shigella* (Gemski et al., 1972) or *Salmonella* strains (Baron *et al.*, 1972).

1.1.1. The lambda life cycle.

The genome of bacteriophage λ is a linear double-stranded DNA molecule, approximately 50 kb in length with a single-stranded 5′ extension of 12 bases at both ends: these extensions are complementary to each other and are called *cos* (*co*hesive end *s*ites). During infection, the right 5′ extension (*cosR*), followed by the entire genome, enters the host cell. Shortly after its penetration the DNA molecule is circularized: the *cos* ends are ligated each other by *E. coli* DNA ligase, forming a covalently closed circular molecule.

Lambda is a temperate phage: after infection, it may multiply in its host in either of two mutually exclusive modes, as an autonomous element, in the lytic mode, or as part of the bacterial chromosome, in the lysogenic mode (Figure 1.1).

In productive growth, the injected DNA molecule directs the synthesis of numerous gene products. These promote autonomous replication of the phage DNA, its packaging into phage particles, and

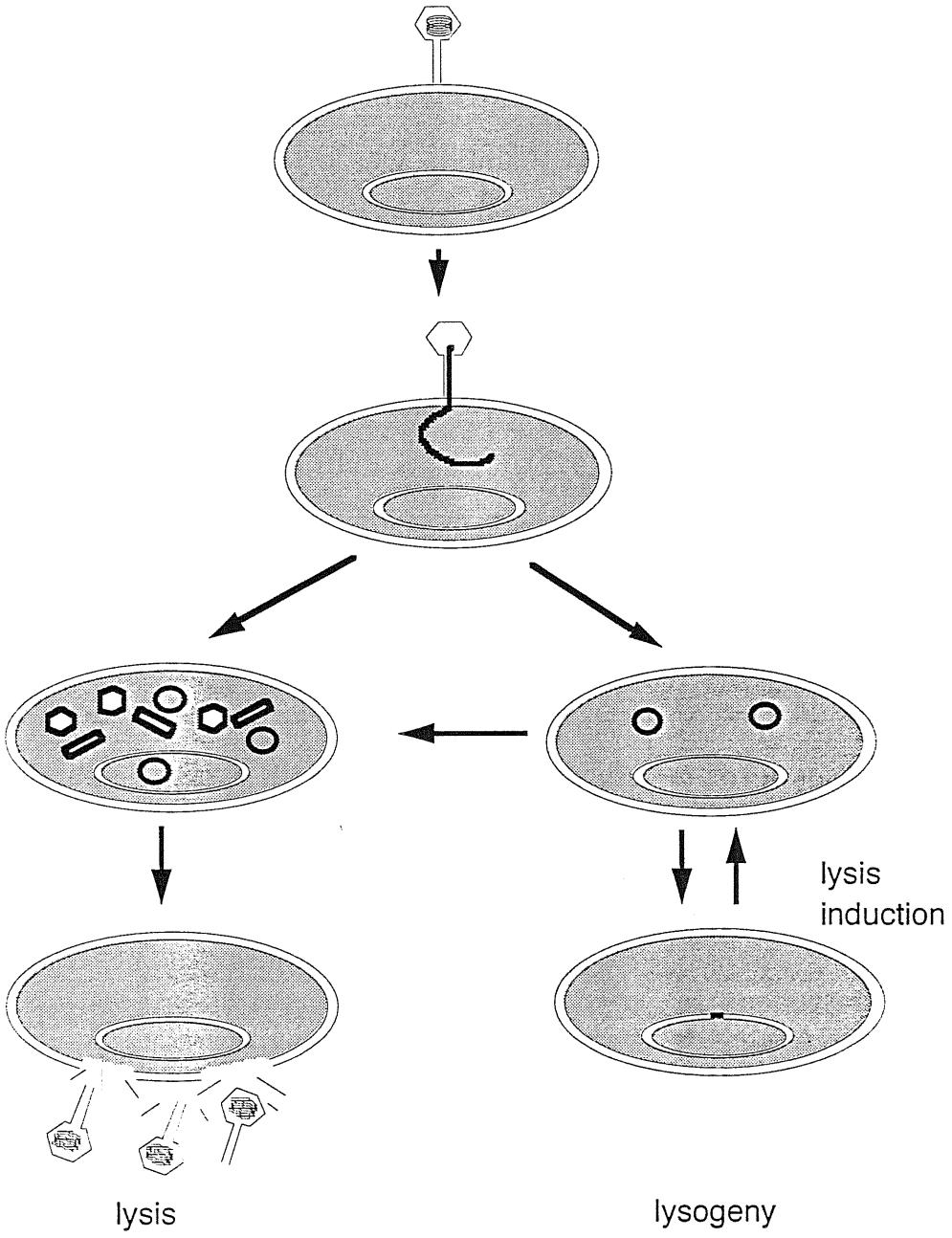


Figure 1.1. The life cycle of bacteriophage lambda. After entering the host cell, the phage can multiply by either a lytic or a lysogenic pathway. The bacteriophage growing in the lysogenic state can be induced to exit from the host chromosome and shift to lytic growth.

lysis of the cell to release about one hundred progeny particles, all within 50 minutes at 37°C. This sequence of events is called the productive or lytic cycle of phage growth.

Alternatively, lambda may enter a dormant stage and replicate as part of the bacterial chromosome. The injected DNA must first direct the synthesis of gene products that promote insertion of the phage chromosome into the DNA of the host, and then express other genes that act promptly to repress autonomous DNA replication and most phage functions. The inserted phage DNA is called prophage. Its replication contributes to viral growth indirectly, because every cell carrying a prophage is a potential centre for production of phage particles. A bacterial strain carrying a prophage is called a lysogen. The process beginning with infection and leading to the establishment of prophage is called lysogenization. A lysogen can be induced to enter the lytic cycle.

The λ genome expresses approximately 50 proteins. The complete genome showing the organization by functional units is drawn in Figure 1.2.. The various sets of λ genes can be designated according to their roles: head assembly, tail assembly, recombination, exclusion, regulation, replication, lysis.

1.1.2. Lambda phage morphogenesis.

Electron microscopy of negatively stained λ phage shows an icosahedral head from which projects a tubular tail of rather simple structure (Figure 1.3.a.).

Head and tail are synthesized and assembled in two different and

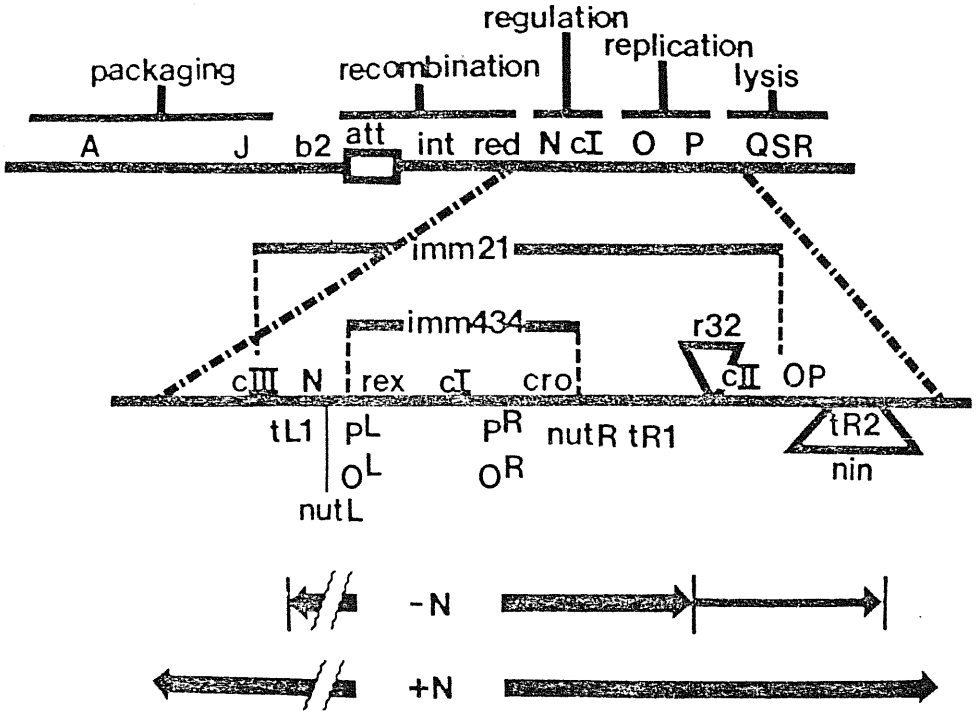
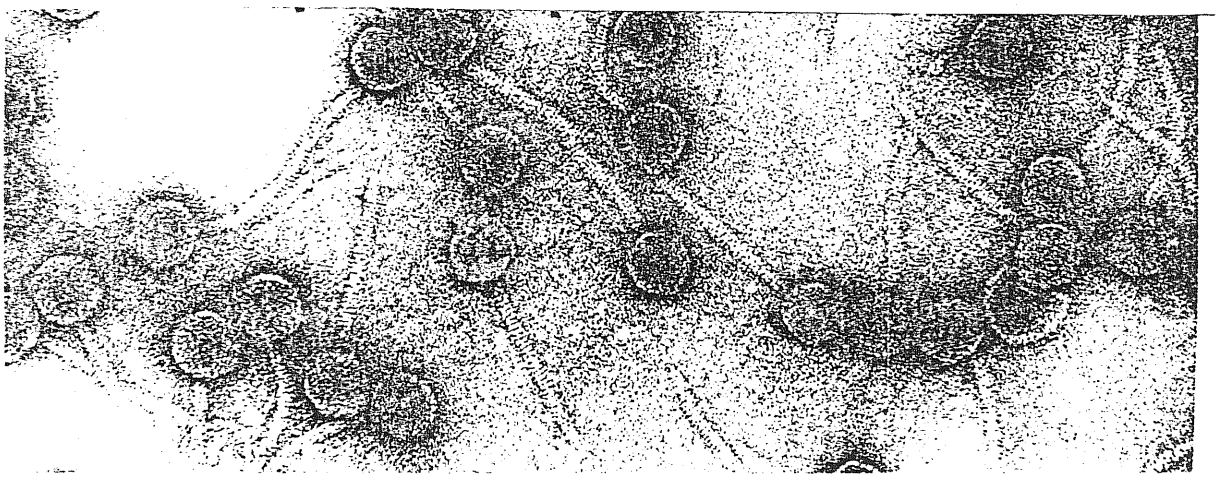


Figure 1.2. Map of lambda genome. A kilobase scale drawing of the lambda genome beginning and ending at the cohesive end sites (*cos*). Gene clusters are indicated as well as some of the major genes.



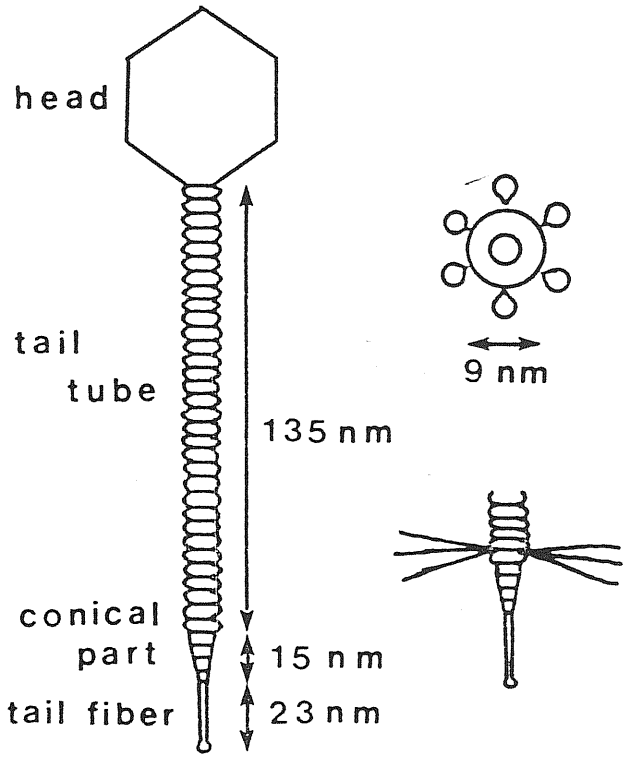


Figure 1.3. Lambda phage morphology. a) electron microscopy of negatively stained λ phage showing its organization in an icosahedral head from which projects a tubular tail of rather simple structure. b) scheme of the proteins forming the structure of the phage.

separate pathways: the products of genes *Nu1* through *FII* (Figure 1.2.) are required for head synthesis and assembly, while the products of genes *Z* to *J* (Figure 1.2.) are required for tail synthesis and assembly. Mutants of λ in head or tail genes produce free tails and free heads respectively, instead of infectious phage particles. These free heads and free tails are active: they can complement each other *in vitro* with high efficiency and yield infectious phage particles (Harrison *et al.*, 1973; Weigle, 1966).

The head is about 650 Å in diameter, hexagonal in outline and very probably icosahedral in shape (Figure 1.3.b.). The assembly of the λ head requires the products of at least ten phage genes and two host genes (*groEL* and *groES*) although the finished head contains amino acid sequences coded by only six of these genes (*W, B, C, D, E* and *FII*) and eight types of polypeptides (called gpE, gpD, gpW, gpFII, pB*, pX1 and pX2).

The tail of λ is classified as long non contractile. It consists of a uniform thin flexible tube (1350 Å in length), ending in a small conical part (150 Å in length) and a single tail fiber (230 Å in length) which is attached to the conical part at the distal end of the tail (Katsura, 1983) (Figure 1.3.b.). The hollow tube of the tail consists of 32 stacked disks, each of which is formed by six subunits of the major tail protein, gVp, arranged so that each disk has a central, 30 Å diameter hole. The conical part of the tail is apparently composed of three or four disks whose formation requires the products of the lambda tail genes *J, I, L, K, H, G* and *M*.

1.1.2.1. Tail proteins and assembly. The tail assembly pathway, i.e. the order of assembly of the tail gene products during the formation of the tail has been established by detecting the *in vitro* complementation activity and serum-blocking power of tail precursors that are present in the lysates of tail-gene mutants (Katsura, 1976b; Katsura & Kuhl, 1975). Assembly starts from gJp, the tail fiber located at the head-distal end in the finished phage particle (Figure 1.4.). Then the products of genes *I*, *L*, *K*, *H*, (*G*), and *M* act in this order on the tail fiber. gVp polymerizes on the initiator to form the tail tube. gUp apparently acts to terminate polymerization at the appropriate length, and gZp completes the tail in a form that can be attached to the head. The order of action of the gene products during assembly (*J*, *I*, *L*, *K*, *H*, *G*, *M*, *V*, *U*, *Z*) (Figure 1.4.) is similar to the arrangement of the genes on the λ genome (*J*, *I*, *K*, *L*, *M*, *H*, *G*, *V*, *U*, *Z*) (Figure 1.2.). The order of action is probably based on structural interactions between proteins involved in adjacent steps of the assembly pathway. Therefore, the similarity between order of functions and order of genes may have an evolutionary advantage, because it minimizes the number of new possibly unproductive interactions that might arise by recombination in the tail region between λ and related phage having partially homologous chromosomes (Casjens & Hendrix, 1974a; Casjens & Hendrix, 1974b).

gVp, the major tail protein which forms the hollow tube of the tail, has two folding domains, with the smaller C-terminal domain on the outside. Mutants of *V* gene with normally functioning tails have been found that lack up to one-third of pV; in these mutants, the domain of gVp that constitutes the outer knob is missing (Katsura, 1981). gV has been recently used to clone an epitope and a protein

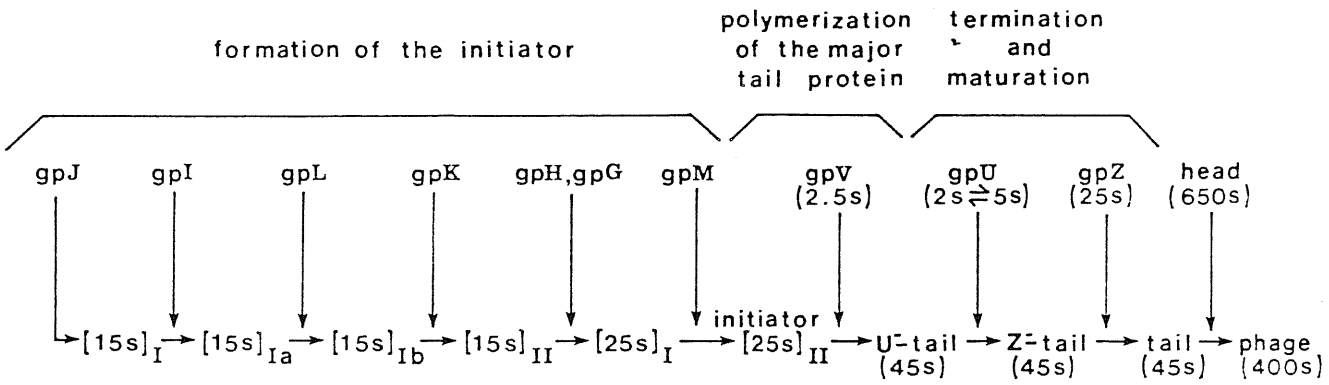


Figure 1.4. Synthesis and assembly of tail proteins.

for phage display (discussed in 1.3.1.1.) (Dunn, 1995; Maruyama *et al.*, 1994).

1.1.3. Host and phage proteins required for adsorption and penetration.

Adsorption of lambda on the bacterial cell surface is the first step in the infection of the host cell, followed by DNA injection. Adsorption is mediated by the tail fiber: gJp attaches the bacteriophage to the surface of the cell by binding to the outer membrane protein LamB, the so called lambda receptor (Randall-Hazelbauer & Schwartz, 1973; Schwartz & Le Minor, 1975b; Thirion & Hofnung, 1972). To penetrate the host cell, lambda DNA has to overcome a triple barrier, the outer membrane, the periplasmic space and the inner membrane. Penetration of phage DNA into the host cell is mediated by a specific interaction between components of the phage tail (gVp and gHp) and another *E. coli* gene, *ptsM*, which codes for an inner membrane protein (Elliott & Arber, 1978). The products of both *lamB* and *ptsM* have functions of sugar transport in addition to those in the infection of λ phage.

Host proteins that play an essential role in adsorption and penetration have been recognized mainly by genetic studies. The *E. coli* mutations in genes that do not allow the infection of λ phage are called λ^r (lambda-resistant) and *pel*$^-$ (*pe*netration of *l*ambda) for genes *lamB* and *ptsM*, respectively. Phage mutations that overcome these host mutations have been isolated and named λ*h* (host-range) (Appleyard *et al.*, 1956) and λ*hp* (host range of *pel*$^-$ bacteria) (Scandella & Arber, 1974). The λ*h* mutations cluster in the right end of *J* gene (Fuerst & Bingham, 1978; Mount *et al.*, 1968), and the λ*hp* mutations

are located in gene *H* and gene *V* (Scandella & Arber, 1976), which shows that the products of these tail genes function in adsorption and injection, respectively.

1.1.3.1. LamB, the phage receptor protein. The phage lambda receptor protein is an outer membrane protein encoded by the *E. coli lamB* gene (Randall-Hazelbauer & Schwartz, 1973; Schwartz, 1975; Thirion & Hofnung, 1972). LamB also serves as a specific receptor for several other bacteriophage (Charbit & Hofnung, 1985; Jacob & Wollman, 1954; Roa, 1979). The LamB protein forms pores which allow nonspecific permeation of low molecular mass molecules (less than 600 Daltons) but which display a distinct specificity for maltose and also facilitate the diffusion of maltodextrins containing as many as 7 or 8 glucose residues (Luckey & Nikaido, 1980; Neuhaus *et al.*, 1983; Wandersman *et al.*, 1979). For this reason it can also be called maltoporin (Neuhaus et al., 1983).

In the native state LamB is an oligomeric protein composed of three identical subunits each containing 421 amino acids. It has a relative molecular mass of 47.000 Daltons and exhibits some distinct similarities with unspecific porins in secondary, tertiary and quaternary structure but not primary structure. Its crystal structure has been recently solved (Schirmer *et al.*, 1995). The scaffold of the monomer is an 18-stranded antiparallel β-barrel forming a channel (Figure 1.5.a). Strands are connected to their nearest neighbours by long loops and turns. The loops are found at the end of the barrel that is exposed to the cell exterior. Site mutations that confer resistance to phage λ had been previously described (Charbit *et al.*, 1988; Gehring *et al.*, 1987). Only about half of the sites involved in recognition by phage

λ are exposed (residue numbers 155, 164, 259, 386, 387, 394, and 401), whereas the remaining sites appear to be buried at locations where the loops pack together to form the protrusions (Schirmer et al., 1995). All three monomers of a trimer are needed for phage adsorption (Marchal & Hofnung, 1983).

Before the x-ray structure was determined, the organization of the LamB channel had been studied by delection of putative surface loops, linker insertion mutants, and detection of foreign antigenic determinants (Charbit *et al.*, 1991; Hofnung, 1991; Hofnung *et al.*, 1981; Klebba *et al.*, 1994): a model for the folding of the protein was proposed on the basis of a combination of structural predictions from the amino acid sequence and of genetic and immunological data (Figure 1.5.b.). Comparison with the available x-ray model has confirmed that data obtained with these different methods have led to correct conclusions about tha topology of LamB (Hofnung, 1995).

Missense *lamB* mutants synthesize an altered phage receptor and are thereby resistant to wild type λ infection (Appleyard et al., 1956; Hofnung *et al.*, 1976; Randall-Hazelbauer & Schwartz, 1973; Schwartz, 1975). Among the *lamB* mutations yielding resistance to phage λh$^+$, the existence of three classes has been reported (Hofnung et al., 1976). One class (class III), composed of nonsense mutations and deletions, abolishes all of the known activities of the λ receptor. The other two classes do not abolish at least one of the *in vivo* activities of the λ receptor: class I allows growth of λh and λhh*, and class II allows growth of λhh*. A large number of class I and class II LamB missense mutants have been analyzed (Charbit *et al.*, 1984; Charbit et al., 1988; Clement *et al.*, 1983; Gehring et al., 1987). Genetic mapping of the

mutations showed that class I and class II mutations clustered mainly in a few regions of the *lamB* gene (Hofnung et al., 1981): class I mutants occur at positions 18, 148, 152, 163, 164, 245, 247, 249, 250, 259, and 382 of the mature LamB polypeptide, and all class II mutants corresponded to position 151 (Charbit et al., 1988). Recently the residue 151 has been proposed to belong to an accessibility gate controlling the access to the phage tight-binding site (Charbit *et al.*, 1994).

Binding of λ to the purified LamB from *Shigella sonnei* 3070 or certain strains of *E. coli* leads to ejection of DNA (Schwartz & Le Minor, 1975b). In contrast binding to LamB from *E. coli* K12 is reversible and DNA ejection does not take place unless chloroform or alcohol is present (Randall-Hazelbauer & Schwartz, 1973). The DNA sequence of the promoter-distal half of *lamB* from *Shigella sonnei* 3070 has been determined (Roessner & Ihler, 1987) and compared with the known sequence for the *E. coli* K12 gene (Clement & Hofnung, 1981). The only predicted amino acid changes in this region of LamB lie between positions 381 and 390, where seven of the ten amino acids are altered (Table 1.1.). This relatively short segment seems thus to be involved in triggering ejection of DNA from the bacteriophage.

	Residue									
	381	382	383	384	385	386	387	388	389	390
	THR	GLY	ASN	ALA	ASP	ASN	ASN	ALA	ASN	PHE
E. coli	ACC	GGT	AAC	GCT	GAT	AAC	AAC	GCG	AAC	TTC
S. sonnei	AAC	GGC	GAT	AGC	AAG	GTT	AAC	CCG	AAC	TAC
	ASN	GLY	ASP	SER	LYS	VAL	ASN	PRO	ASN	TYR

The nucleotide sequence of *S. sonnei* 3070 *lamB* is compared with the sequence of *E. coli* K12 *lamB* determined by Clement & Hofnung (1981) for the region corresponding to amino acid residues 381 to 390. There are 7 predicted amino acid changes (underlined) in the 8 affected codons. The only other nucleotide differences in the *Sal*I–*Sac*I fragment are T → C at position 766, G → T at position 909, T → C at position 930, and A → G at position 933; these changes do not alter the amino acid sequence.

Table 1.1. Comparison of the nucleotide sequences for residues 381 to 390 of *E. coli* K12 and *S. sonnei* 3070 *lamB*. From (Roessner & Ihler, 1987).

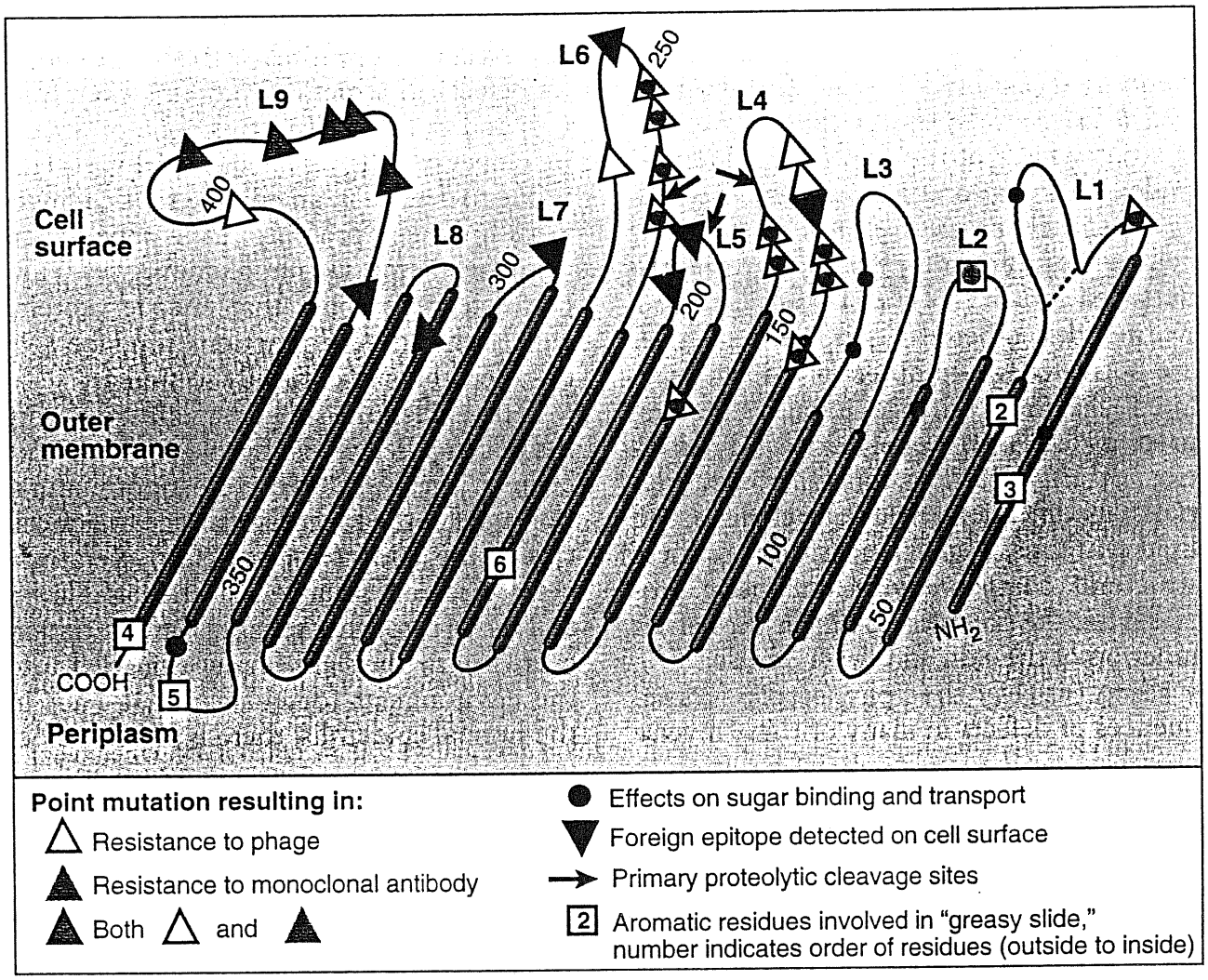


Figure 1.5.a. Schematic of the maltoporin monomer. The 421 amino acid residues are oriented from right (amino terminus) to left (carboxyl terminus). The barrel corresponding to a monomer has been opened and flattened on the page. Thick lines, residues in transmembrane β strands; thin lines, loops and turns. The triangles and arrows designate regions assumed to be near or at the cell surface. From (Hofnung, 1995).

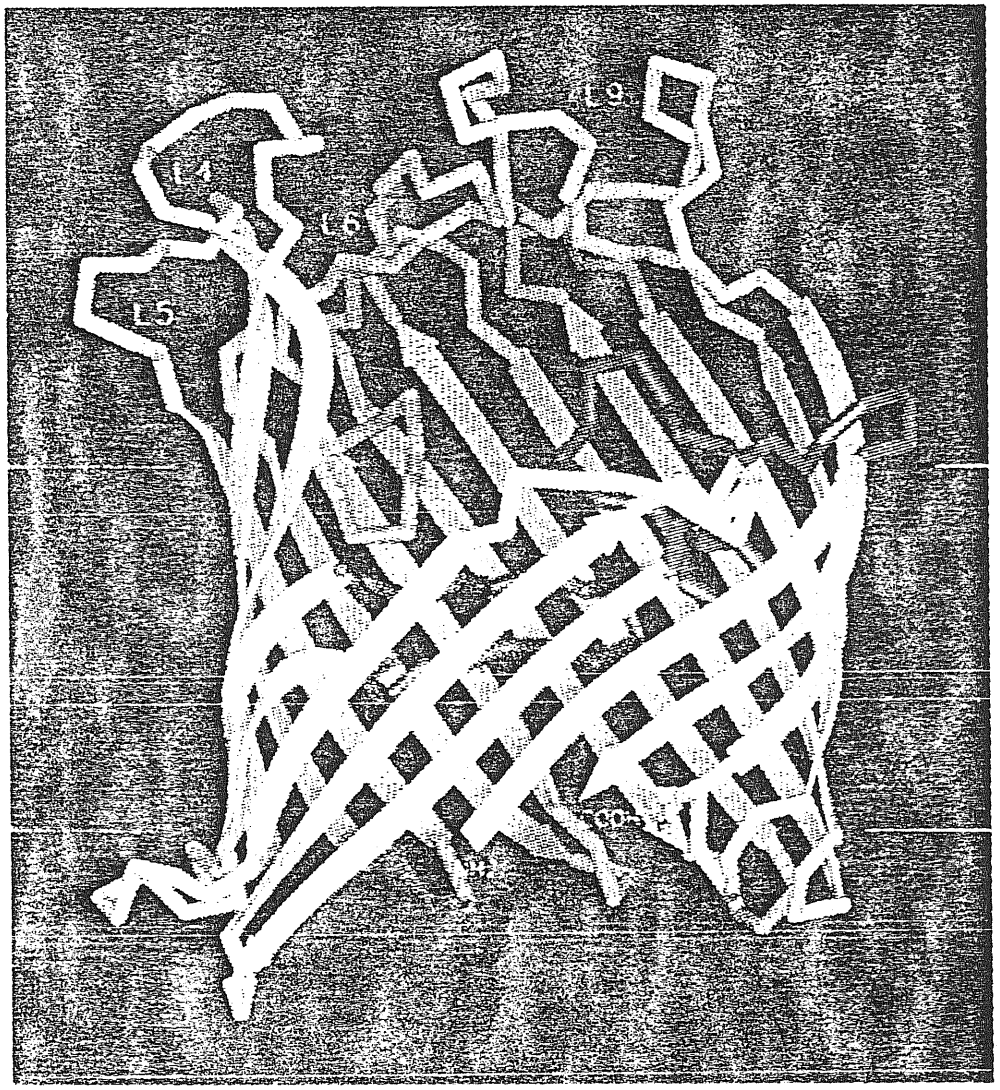


Figure 1.5. b. The x-ray structural model for the maltoporin monomer. The cell exterior is at the top and the periplasmic space at the bottom. The area of the subunit involved in trimer contacts is facing the viewer. The 18 antiparallel β strands of the barrel are represented by arrows. Strands are connected to their nearest neighbors by loops or regular turns. From (Schirmer et al., 1995).

1.1.3.2. Host range mutants and the J protein. The extended host-range mutants of λ (Appleyard et al., 1956) are selected using missense *lamB* mutants that are resistant to wild type λ infection (Hofnung et al., 1976; Randall-Hazelbauer & Schwartz, 1973; Schwartz, 1975). Various mutations in gene *J* were located in the right end of the gene (Fuerst & Bingham, 1978). The J product is a structural protein of 130.000 molecular weight present in two or three copies in the phage tail (Hendrix, 1971; Murialdo & Siminovich, 1971). It is believed to be the tail fiber seen in electron micrographs, although its identification with the tail fiber has never been made directly. It is the phage protein responsible for anti-λ serum blocking activity (Buchwald & Siminovich, 1969; Dove, 1966).

Since gJp is the target of neutralizing antibodies (Buchwald & Siminovich, 1969), and since host range mutations which alter the specificity of the attachment reaction map within gene *J* (Mount et al., 1968), it is clear that gJp is the protein which directly binds to LamB.

Recently the mutants of *J* gene responsible for one and two-step extended host range, have been sequenced (Werts et al., 1994). Seven residues at the distal part of the J protein have been identified which are involved in the extended host range properties of mutant λ phage (see Tables 1.2.a. and 1.2.b.).

Three different class I LamB mutants with mutations at residues 247 (Ser→Leu), 245 (Gly→Arg), and 148 (Glu→Lys), in the LamB protein, were used to select eleven one-step extended host range mutants of lambda (λh phage). For each of the 11 new λh phage, DNA sequence analysis revealed a single-amino-acid change in the distal

part of the *J* gene (Table 1.1.a.). All of the mutations were clustered in the distal 10% of the J protein as expected from a previous genetic work (Mount et al., 1968). The alterations in the 11 new λh phage affected only three residues within J. Five phage corresponded to a Val→Ala change at aa 1077, three phage corresponded to a Thr→Met change at aa 1040, and three phage corresponded to a Leu→Pro change at aa 1127. It is worth noting that several λh phage isolated from different class I LamB mutant strains corresponded to the same amino acid change in the J protein.

Two-step host range λhh* phage have been selected on the class II LamB mutant (a Gly→Asp change at site 151 of LamB) from three λh phage corresponding to three different missense mutations in λ DNA sequencing showed that for each selection, the four λhh* mutants corresponded to double and triple-point mutations, all located in the distal part of the *J* gene (Table 1.2.b.). In all cases, the λhh* derivatives corresponded to the presence of one or two additional mutations located upstream the parental (λh) mutation.

E. coli strain[a]	Alteration in LamB	Selected λh phage	DNA alteration in gene J[b]	Alteration in protein J[c]
pop7079	Ser-247 --> Leu	λh247-1	3230; GTA --> GCA	Val-1077 --> Ala
		λh247-2	3119; ACG --> ATG	Thr-1040 --> Met
		λh247-3	3230; GTA --> GCA	Val-1077 --> Ala
		λh247-4	3230; GTA --> GCA	Val-1077 --> Ala
		λh247-5	3230; GTA --> GCA	Val-1077 --> Ala
		λh247-6	3230; GTA --> GCA	Val-1077 --> Ala
		λh247-7	3380; CTG --> CCG	Leu-1127 --> Pro
pop7086	Gly-245 --> Arg	λh245-1	3119; ACG --> ATG	Thr-1040 --> Met
		λh245-2	3119; ACG --> ATG	Thr-1040 --> Met
		λh245-3	3380; CTG --> CCG	Leu-1127 --> Pro
pop7090	Glu-148 --> Lys	λh148-1	3380; CTG --> CCG	Leu-1127 --> Pro

[a] *E. coli* strains corresponding to class I *lamB* missense mutations.

[b] Each number corresponds to the position of the first base of the triplet in the nucleotide sequence of gene J.

[c] Each number indicates the position of the amino acid in the protein J sequence.

Table 1.1.a. DNA and amino acid sequence of one-step host range mutants. From (Werts et al., 1994).

Phage λh used for λhh* selection	Amino acid altered in protein J	λhh* phages selected	Alteration in gene J[a]	Amino acid altered in protein J[b]
λh245-2	Trh-1040 -->Met	λhh* 1040-1--λhh* 1040-4	3319; ACG --> ATC	Thr-1040 --> Met
			3031; AGC --> GGC	Ser 1011 --> Gly
λh148-1	Leu-1127 -->Pro	λhh* 1127-1--λhh* 1127-4	3380; CTG --> CCG	Leu 1127 --> Pro
			3233; CAG --> CGG	Gln 1078 --> Arg
λh247-6	Val1077 -->Ala	λhh* 1077-1--λhh* 1077-4	3230; GTA --> GCA	Val 1077 --> Ala
			3226; GCG --> TCG	Ala 1076 --> Ser
			3224; GAG --> GTG	Glu 1075 --> Val

[a]Each number is th position of the first base of the triplet in the gene J nucleotide sequence.

[b]Each number is the position of the amino acid in the protein J sequence.

Table 1.2.b. DNA and amino acid sequence of one-step host range mutants. from (Werts et al., 1994).

It is worth noting that each λh phage can grow on all class I LamB mutants as well as on the wild-type LamB strain, with an efficiency of plating identical to that of the parental wild-type phage. This has been interpreted as meaning that the mutated residues in the J protein and in the LamB mutants are not involved in allele-specific protein-protein interactions. Rather, the LamB mutations would block a step in phage absorption, and this block would be overcome by the mutations in the J protein.

The conclusion given by the authors is that "It is not possible to draw any obvious relationships between these sets of observations and our data which could lead to a general view of the mechanism of phage λ absorption".

1.1.3.3. PtsM, an inner protein necessary for DNA injection. The *E. coli* gene, *ptsM*, which codes for the mannose permease, an inner membrane protein, plays an essential role in the DNA injection of λ phage. It has been proposed that it mediates penetration of phage DNA across the cytoplasmic membrane (Elliott & Arber, 1978).

λ adsorbs normally to the *pel*⁻ host, but only 2-10% of the infected cells produce phage with a normal burst size or become lysogenic (Scandella & Arber, 1974). The remainder of the cells survive the infection. After adsorption of λ to *pel*⁻ cells, active phage do not elute spontaneously from the complex. Neither are adsorbed phage released in an inactivated form. Electron micrographs of λ infected cells show phage particles with empty heads attached to the surface of *pel*⁺ cells, while the particles attached to the *pel*⁻ cells appear to have full heads.

The *S. sonnei* LamB does not suppress the pel phenotype and infection is blocked prior to DNA injection. The conclusion from these works is that the *pel*⁻ mutant allows adsorption of λ but that it inhibits the subsequent injection of the phage DNA. A curious property of the *pel*⁻ block to injection is that it is much more severe for phage with genomes that are smaller than wild type, and the block can be largely overcome by DNA insertions or duplications, which increase the length of the genome (Emmons *et al.*, 1975).

The mutants λ*hp* and λ*hp*ts overcome the inhibition of DNA injection due to the bacterial *pel*⁻ mutation (Scandella & Arber, 1976). All the temperature-sensitive mutants are in gene *V*. The λ*hp*ts phage are heat stable and inject their DNA at 42°C the thermolabile step appears to be expressed late in infection, during phage tail assembly.

The non-temperature-sensitive mutants lie within gene *H*. These two proteins must therefore interact functionally with the bacterial *pel* product during injection.

The mannose permease consists of three different subunits, III^{Man}, $II\text{-}P^{Man}$ and $II\text{-}M^{Man}$. All three subunits are required for sugar transport and phosphorylation. Two ($II\text{-}P^{Man}$ and $II\text{-}M^{Man}$) are sufficient to mediate penetration of phage λ DNA (Erni *et al.*, 1987). III^{Man} (35 kDa) is a hydrophilic protein which is transiently phosphorylated and most likely contains the active site for sugar phosphorylation. $II\text{-}P^{Man}$ (28 kDa) contains long segments of hydrophobic amino acids of which some are completely free of charged residues. It might thus span the membrane several times and be mostly buried in the membrane bilayer. In $II\text{-}M^{Man}$ two regions of different hydrophobicity can be distinguished, the N-terminal third which is hydrophilic and the C-terminal two-thirds which contain four segments of hydrophobic residues. $II\text{-}M^{Man}$ can been expected to span the membrane once or twice only and have large domains exposed at the cytoplasmic and/or periplasmic membrane surface. The truncation of $II\text{-}M^{Man}$ markedly affects phage susceptibility.

1.1.4. Mechanism of phage absorption and penetration.

Adsorption and penetration are the first two steps in the life cycle of lambda bacteriophage.

Two steps have been recognized for adsorption, namely reversible and irreversible adsorption. Reversible adsorption of bacteriophage particles to the receptors is facilitated by magnesium ions and occurs efficiently and rapidly (within a few minutes) both at 37°C and at room

temperature, and even at lower temperatures. After this step the phage sediments together with the bacterium but it can still be inactivated by antiphage serum. Irreversible adsorption is dependent on temperature and the triggering of DNA injection and entry of the phage DNA into the cell does not occur until the temperature is raised to 37°C. The temperature dependent lag might reflect an enzymatic step; e.g., the tip of the phage tail might alter the cell wall of the bacterium to allow the DNA to be injected. Alternatively, a closed end of the phage tail might be opened.

After adsorption, the release of phage DNA is triggered and the DNA penetrates into the cytoplasm of the host leaving the protein coat attached to the surface of the cell. The triggering of DNA injection does not occur automatically at the time of irreversible binding between the phage tail and the bacterial receptor (McKay & Bode, 1976a). It occurs almost instantaneously at 37°C, takes a long time at 23°C, and in any case requires temperatures above 15°C. MacKay and Bode (McKay & Bode, 1976a) defined three steps after adsorption: lag, trigger and uptake. The lag function is followed by trigger or ejection of DNA, but if the condensed state of the DNA in the head is stabilized by putrescine the reaction stops between lag and trigger. In the normal pathway of DNA injection, the ejection of DNA is always accompanied by the up-take of DNA by the host cell. However if the phage-bacterium complex is treated with chloroform before the lag function, the uptake does not occur and the DNA is found in the culture medium.

Although even the free tail can adsorb to the bacterium, the lag reaction requires the head to be attached to the tail. It has been shown

(McKay & Bode, 1976a) that phage tails alone cannot complete lag functions. Biologically λ active tails were adsorbed to the host cells and incubated at 37° to allow these postulated temperature dependent reactions to occur (McKay & Bode, 1976a). After cooling to 4°C, purified λ heads were joined to the attached tails forming complete phage.

Analogous studies have been performed using isolated λ-receptor protein (McKay & Bode, 1976b; Roa & Scandella, 1976; Schwartz & Le Minor, 1975b) or membranes (Zgaga *et al.*, 1973) instead of the whole cell. The reactions between the phage and the isolated receptor show three steps: reversible interaction, irreversible binding (phage inactivation), and DNA ejection, which is temperature dependent. Curiously, chloroform or ethanol is required for the second step: wild type lambda is not inactivated when it is incubated with LamB isolated from most *E. coli* strains and DNA ejection does not take place unless chloroform (Randall-Hazelbauer & Schwartz, 1973) or 10% to 20% alcohol is present (Schwartz & Le Minor, 1975a). This requirement is suppressed by mutations in either gJp of phage λ or the receptor of the host cell. The binding to LamB from *Shigella sonnei* 3070 or certain strains of *E. coli* leads to ejection of DNA from bacteriophage lambda after its binding (Schwartz & Le Minor, 1975b). The same effect can be obtained with the host range mutants, λ*h*, which eject their DNA in the absence of chloroform after binding to the isolated *E. coli* receptor (Randall-Hazelbauer & Schwartz, 1973). Since both λ*h* and the *Shigella* receptor are able to ensure DNA ejection across a membrane in *in vitro* systems without the addition of $CHCl_3$, and since this ability is not required for injection of DNA into bacteria, some other triggering mechanism must operate *in vivo*. The requirement of organic solvent

probably corresponds to the finding in the studies using isolated membranes (Zgaga et al., 1973) that the outer membrane has to form a complex with the inner membrane to be absorbed by λ phage irreversibly.

The morphology of the complexes formed between bacteriophage lambda and structures containing the lambda receptor - either liposomes or rod-shaped particles - have been studied by the electron microscope (Roessner & Ihler, 1987; Roessner *et al.*, 1983). Two types of complexes can be seen. Complexes of type 1, which are reversible, involve binding only between the tips of the tail fibers and the receptors, so that the hollow tails are separated from the receptor-bearing structure by about 17 nm. In complexes of type 2, which are irreversible, the banded regions of the tails are found to be in direct contact with the surface of the liposome with no gap visible. The tail fiber is degradable by proteases in type 2 complexes, but not in type 1, suggesting that some change in the tail fiber may occur during the transition or after DNA ejection (Roessner & Ihler, 1984).

Complexes of type 1 form with the LamB protein isolated from *E. coli* K-12. When *Schigella* receptors are used, similar type 1 complexes can be visualized at 4°C, but upon warming they convert to complexes of type 2. Type 2 complexes also form between the *E. coli* K12 receptor and the extended host-range mutants of lambda, h and hh*.

It is unlikely that the tail fiber is drawn into the liposome by some type of ratchet mechanism in the transition from type 1 to type 2. If that were the case, gJp would be internal and protected by the liposome from external agents such as proteinases, just as DNA injected into

liposomes is protected from DNAses. As mentioned above, gJp is not protected, but rather becomes more sensitive to proteolytic degradation in type 2 complexes (Roessner & Ihler, 1984).

Hofnung speculates that λ distal tail fiber constituted by the J protein has to pass through an "accessibility gate" located at the surface of the LamB receptor to be able to interact tightly with the binding site of the receptor proteion and to allow further phage DNA injection (Werts et al., 1994). The effects of class I mutations in LamB would be to block the entry of the tail fiber by narrowing the accessibility gate. Extended host range mutations in gJp would affect the structure of the distal tail fiber by a global reduction of the diameter of the tip (Werts et al., 1994).

1.2. Filamentous phage.

The filamentous bacteriophage belonging to the genus *Inovirus* are a group of related viruses which only infect gram-negative bacteria and specifically adsorb to the tip of pili encoded by conjugative plasmids carried by the host cell. The virion consists of a circular single-stranded DNA molecule of about 6500 nucleotides encapsulated in a long cylindrical protein coat, with no organization into head or spikes.

On the basis of their X-ray diffraction pattern *Inovirus* can be separated into two symmetry classes: class I (Marvin *et al.*, 1974a) includes strains Ff (fd, f1, M13), If1, I2-2 and IKe, class II (Marvin *et al.*, 1974b) includes strains Pf1, and Xf.

This thesis deals with both Ff and IKe phage.

The best studied are the IncF-plasmid specific phage Ff, which infect *E. coli* via F pili (Model & Russel, 1988; Rasched & Oberer, 1986). These phage are very closely related, differing from one another by only a few nucleotides, with their 10 gene products being interchangeable under all conditions in which exchange has been attempted. They have been used as cloning vectors and, more recently, as a powerful screening tool in phage display (Cesareni, 1992; Scott & Smith, 1990).

Phage IKe infects *E. coli* cells carrying plasmids of the N- (Khatoon *et al.*, 1972), P- (Grant *et al.*, 1978) and I2-incompatibility group (Bradley *et al.*, 1983). IKe and Ff phage have an identical number and gene order and their genomes display 55% homology (Peeters *et al.*, 1985). The

extent of homology between the genes of IKe and Ff differs significantly from one gene to another. Genes that code for viral capsid proteins are less homologous than genes whose products are involved in the processes of DNA replication and phage morphogenesis (Bross *et al.*, 1988; Peeters et al., 1985).

1.2.1. Filamentous phage propagation.

Filamentous phage (thereafter any reference to filamentous phage is to be intended to Ff, unless specified) are continuously assembled and extruded from the bacterial cell without causing its lysis. Large numbers of progeny are produced (up to 1000 per cell in the first hour after infection) without seriously affecting the cell's ability to grow and divide.

The phage is a rod-shaped particle containing a circular, single stranded DNA genome oriented such that the packaging signal hairpin is located at the end of the particle that first emerges from the infected cell (Lopez & Webster, 1983; Webster *et al.*, 1981). The phage genome encodes ten proteins (for reviews see (Model & Russel, 1988; Rasched & Oberer, 1986). Three are involved with DNA synthesis (i.e. the product of genes *II, V , X*) and five are virion structural proteins (i.e. the product of genes *III, VI, VII, VIII,* and *IX*). The other two proteins are absolutely required for phage assembly (i.e. the product of genes *I,* and *IV*), even if their are not part of the mature virus particle: they probably interact to form an exit pore (Russel, 1994). In addition, morphogenesis requires host gene products (Lopez & Webster, 1985b; Russel & Model, 1983), of which thioredoxin is the best characterized (Lim *et al.*, 1985; Russel & Model, 1985).

The process of assembly and extrusion of phage particles (Figure 1.6) are coupled events which are initiated at the bacterial inner membrane and involve the virion DNA, the viral coat proteins, at least one host protein, and two phage-encoded morphogenetic proteins. These latter three proteins are not part of the phage particle but are absolutely required for its assembly (Russel, 1991). The precursor for the assembly process is newly synthesized circular single-stranded phage DNA in a complex with the phage gVp. A specific hairpin region of this DNA, present at one end of the DNA-gVp complex (Bauer & Smith, 1988), functions as a packaging signal (Dotto *et al.*, 1981) which appears to initiate assembly of the bacteriophage. The DNA is then extruded through the membrane, where the gVp is displayed and the five capsid proteins are assembly around the single-stranded DNA. This assembly process appears to occur at site of contact between the inner and outer membranes of the host bacterium (Lopez & Webster, 1985a) which resembles the adhesion zones described by Bayer (Bayer, 1968).

Despite binding to different receptors, i.e. N pili, the IKe phage undergo a similar life-cycle to Ff phage. Five of the 10 phage genes have been investigated with regard to their interchangeability between the two phage species (Table 1.3.). Only gVp, the single strand DNA binding protein, is interchangeable (Russel, 1991), while the morphogenetic proteins gIp and gIVp cannot be exchanged (Russel, 1992), unless they are transferred as a pair (Russel, 1993). Neither the replication protein gIIp (Peeters *et al.*, 1986) nor the minor coat protein gIXp (Russel, 1991) can be exchanged.

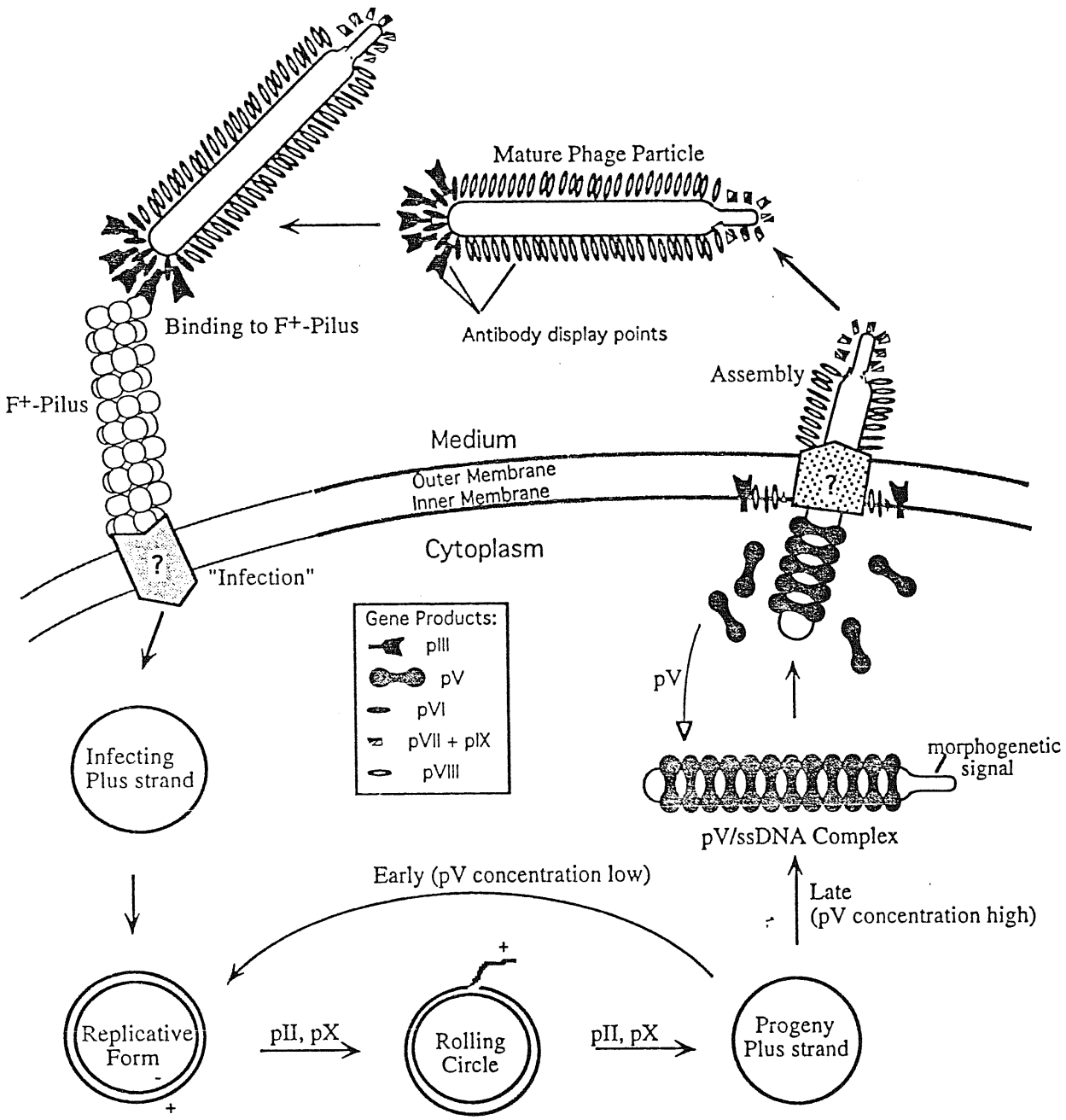


Figure 1.6. Speculative model for filamentous phage assembly.

gIIIp of phage IKe is not assembled into an Ff phage (Bross et al., 1988; Endemann *et al.*, 1993).

Protein	Function	No. aa (f1)	No. aa (IKe)	Identity	Similarity	Interchangeability
A. DNA replication						
II	Origin nickase	410	421	59·8	88·3	−
V	ssDNA binding	87	88	44·8	82·8	+
B. Virion proteins						
III	Adsorption (pVI end)	406*	415*	23·2	57·9	
VI	Stability (pIII end)	112	116	29·5	71·4	
VII	Initiation? (pIX end)	33	32	24·2	66·7	
IX	Initiation? (pVII end)	32	33	18·8	50·0	−
VIII	Major coat protein	50*	53*	40·0	76·0	
C. Assembly						
I	Initiation, elongation	348	365	53·4	84·5	−
IV	Exit. ?	405*	407*	47·9	87·9	−/+

Table 1.3. Summary of similarity and interchangeability of proteins from Ff and IKe phage. Protein length are derived from the DNA sequence of f1 (Hill & Petersen, 1982) and IKe (Peeters et al., 1985). An asterisk indicates proteins synthesized with a signal sequence. From (Russel, 1992).

1.2.2. Filamentous phage morphogenesis.

The general appearance of a filamentous phage is that of a flexible filament, about 900 nm long and 6 to 10 nm thick (Figure 1.6.). Its coat consists of a major protein, the product of gene *VIII* (gVIIIp) and four minor components, the products of genes *III*, *VI*, *VII*, and *IX*, (gIIIp, gVIp, gVIIp, gIXp) (Figure 1.6.). The virus contains no lipid or carbohydrate. The gVIIIp forms a cylindrical coat 1.5 to 2 nm thick around the DNA. The gVIIIp is arranged on a helical lattice with a 5-fold rotational axis and a 2-fold screw axis with a pitch of about 3.2 nm. The end of the virus assembled first (distal end) is capped by a "plug" made up of 5 copies each of gVIIp and gIXp. The end assembled last (proximal end) is capped by 5 copies of gVIp that attach 5 copies of the host binding protein gIIIp to the virion. The coat envelopes a DNA molecule that extends the total length of the virus and constitutes ca. 12% of its weight.

1.2.2.1. The major coat protein. The major coat protein, gVIIIp is only 50 a.a. long and is composed of three specific domains: a 19 amino acid long hydrophobic core is flanked by an acidic N-terminal part (residues 1-20) and a basic C-terminal part (residues 40-50). When associated in vitro with lipid bilayers or detergents, the coat protein can adopt two states: the α-oligomeric and β-polymeric states. The procoat protein inserts into the cell membrane (Chang *et al.*, 1978) where processing by bacterial leader peptidase removes the leader sequence from the N-terminus of the protein (Zimmermann *et al.*, 1982). Many copies of the membrane bound form accumulate within the membrane prior to their assembly into the outer coat of virus particles as they are

membrane bound form accumulate within the membrane prior to their assembly into the outer coat of virus particles as they are extruded through the cell membrane. The gVIII protein has been shown to exist in different conformations: in the intact phage it is at least 90% or more α-helical (Glucksman *et al.*, 1992; Marvin et al., 1974a); the membrane bound form the α-helix content decreases to 50-60%. The fd phage can be contracted by chloroform to short rods called I-form, which are compact structures about 1/3 as long as the original phage, at low temperature and to spheroidal particles called S-forms at high temperature (Roberts & Dunker, 1993). The structure of the coat protein in I-forms is the same as the structure of the coat protein in the native phage, while structure of the coat protein in S-forms is the same as the structure of the coat protein in the lipid bilayer.

1.2.2.2. The bacterial receptor protein, gene III protein. gIIIp is located together with g6p at one end of the phage particle. It mediates the recognition of the tip of the host sex pili (Caro & Schnos, 1966; Jacobson, 1972; Pratt *et al.*, 1969) and is involved in penetration of the single-stranded phage DNA into the host cell (Stengele *et al.*, 1990). Electron microscopy reveals a knob on stem structure pinned to the tip of the phage filament. The knob can be released by treatment with subtilisin, resulting in non infectious phage particles (Gray *et al.*, 1981).

The protein is synthesized as a precursor (424 a.a.) with an 18-residue signal sequence that is removed as the protein is inserted into the host membrane. Mature gIIIp is 406 amino acids long and functional sections within the primary sequence have been correlated with its different functions (Figure 1.7.). Residues 1 to 96 are sufficient

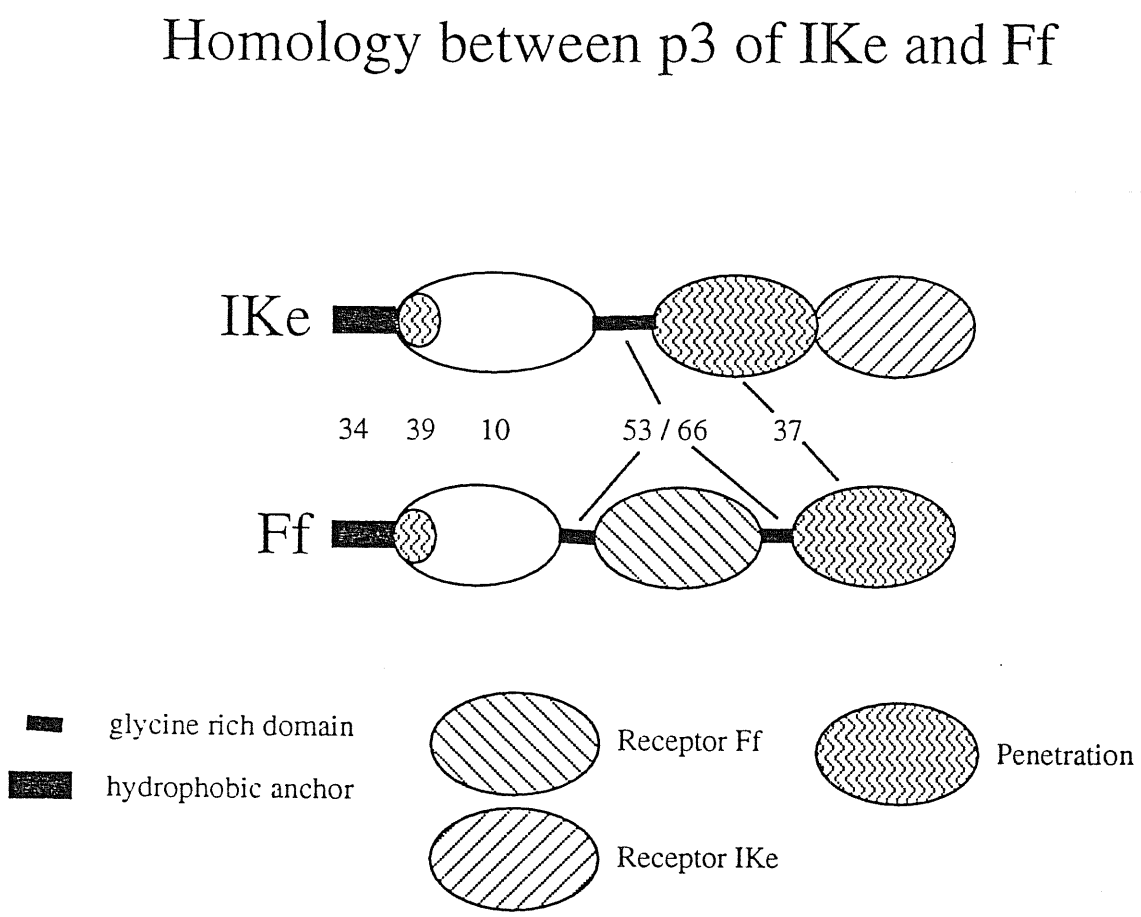


Figure 1.7. Schematic representation of the similarity in the structure of fd and IKe gene III proteins.

to cause the pleitropic effect (Boeke *et al.*, 1982). Deletion of amino acid residues 29 to 196 results in loss of infectivity (Nelson *et al.*, 1981) and transducing ability of F- cells (Russel *et al.*, 1988). Deletion of residues 29 to 339 results in loss of incorporation into phage (Crissman & Smith, 1984). Amino acid residues 379 through 401 form a hydrophobic region of 23 consecutive, uncharged aminoacids which serve as a membrane anchor (Davis *et al.*, 1985): removal of that region converts gIIIp from an integral membrane protein to a secreted one.

When present in the bacterial cell, either synthesized by infecting phage or by its gene cloned on a multicopy plasmid, gIIIp induces pleitropic effects in the host cell that closely mimic the tolerant phenotype conferred by the *tolA*, *tolQ*, or *tolR* mutation in *E. coli*. Thus these cells are resistant to super infection with other F-like specific DNA- and RNA-phage, show a higher sensitivity to detergents such as desoxycholate, have a higher tolerance to certain colicins, and show leakage of periplasmic proteins (Boeke et al., 1982; Rampf *et al.*, 1991; Smilowitz, 1974; Zinder, 1973). F-conjugation is also strongly inhibited.

Purified gIIIp has been shown to produce transmembrane channels in artificial lipid bilayed membranes (Glaser-Wuttke *et al.*, 1989). This is in accord with the finding that the protein contains several putative membrane spanning regions (Beck & Zink, 1981). The protein is able to oligomerize and to form a pore with an estimated diameter of 1.6 nm, which would allow the translocation of the circular single stranded DNA helix with a diameter of about 0.85 nm.

A hybrid protein consisting of the first 372 amino acids of gIIIp fused to the carboxy-terminal catalytic domain of colicin E3 was constructed (Jakes *et al.*, 1988). The hybrid protein killed cells that had the F pilus receptor for phage f1 but not F^- cells. This indicates that the receptor function of gIIIp are sufficient to cause retraction of F pilus when it binds. Therefore gIIIp alone if transferred to another molecule can mediate the entry of the bound molecule.

gIIIp has been used for display of peptides or proteins as discussed in section 1.3.1. The insertion of peptides between the two domains of gIIIp (Smith, 1985) or near the N-terminus (Parmley & Smith, 1988) does not destroy its functions in morphogenesis and infection.

gIIIp from IKe phage causes none of the pleitropic effects of Ff gIIIp. Its homology with Ff gIIp is weaker than between other proteins of the two phage. Homologous regions can be distinguished, but are found scrambled (Figure 1.7.b.). Two glycine rich domains are found in fd gIIIp, while IKe gIIIp contains only one, which is longer. The receptor domains, which bind to the pilus, are completely different in the two proteins and are found in different positions (Endemann *et al.*, 1992; Stengele et al., 1990). The penetration domain of IKe gIIIp is found just upstream of the glycine rich domain and appears to have similar functions (Endemann et al., 1992).

1.2.3. Pili.

Gram-negative bacteria have evolved a number of different genetic systems (*tra* systems) for bacterial conjugation, the process whereby DNA is transferred from one bacterial cell to another. *Tra* systems are encoded by plasmids which are classified in incompatibility groups (Table 1.4.), which include plasmids that can coexist stably in the same cell. Generally plasmids within an incompatibility group determine pili of uniform type.

The word "pilus" is derived from the Latin word meaning hair or hair-like structure. Morphologically, pili are too thin to be visible by ordinary light microscopy and must be visualized by electron microscopy. Conjugative pili have been classified into three morphological groups (Bradley, 1980): thin flexible (group 1), thick flexible (group 2), and rigid (group 3) (see Table 1.4.).

A pilus complex is almost completely analogous to a bacterial non-lytic virus. It is a rod-shaped or partially rod-shaped structure whose nucleic acid codes for, among other things, its own protein coat. It is infectious and requires a sensitive recipient cell. It adsorbs to a receptor on the recipient cell's surface and a copy of its DNA can exit from the host cell without lysing it. It is capable of autonomous replication inside the host cell.

Conjugative pilus morphology and pilus-specific phage sensitivity classifications have suggested that pilus-dependent transfer systems might belong to two evolutionary "families". One family

Plasmid incompatibility (Inc) group	Morphology	Receptor for phages	
		pilus tip	pilus side
B	Thin flexible		
C	Thick flexible		
D	Thick flexible	fd	C-1/D
FI	Thick flexible	fd	R17
FII	Thick flexible	fd	R17
H	Thick flexible		Hga1/PilHα
I_1	Thin flexible	Ifl/PR64FS	Iα
I_2	Thin flexible	Ifl/PR64FS	Iα
Iγ	Thin flexible	Ifl/PR64FS	Iα
J	Thick flexible		C-1
K	Thin flexible		
M	Rigid	X	M
N	Rigid	Ike,PP4	
P	Rigid	PP4,X	PRR1
T	Thick flexible		tf-1
U	Rigid	X	
V	Thick flexible		
W	Rigid	Pr4,X	
X	Thick flexible	X	

Table 1.4. Classification of plasmids into incompatibility groups.

includes the IncF complex and other plasmids with some F-like characteristics (e.g., IncD, IncC, and IncJ), whereas the second includes plasmids with phenotypic resemblance to RP4 (IncP).

1.2.3.1. Structure of F pilus. The most studied incompatibility system is that of IncF plasmid, recently reviewed in (Frost et al., 1994). All the sequences required for the conjugative trasmission of the *E. coli* K-12 fertility factor. F, are encoded within the 33.3-kb transfer region of the 100-kb IncFI plasmid. Figure 1.8. is a view of the organization of the F transfer region. At least 13 *tra* genes are expressed to form a membrane apparatus for pilus assembly and for conjugal DNA transfer. The *tra* gene products have been localized in the bacterial cell, as reported in Figure 1.9.a..

The F pilus is a hollow cylinder of 8 nm in outer diameter, and 2 nm in inner diameter, the length being very variable and ranging from 1-2 to 20 μm (Folkhard *et al.*, 1979; Marvin & Folkhard, 1986). The filament is formed by a single subunit, pilin, arranged in a helical array (Figure 1.9.b.). The helix contains five subunits per turn, with each subunit related by a fivefold rotation axis identical with the axis of the filament. With respect to its neighbors, each pentamer is rotated about the helix axis by about 29° and axially displaced by about 1.3 nm (Marvin & Folkhard, 1986).

The *traA* gene was originally shown by (Minkley *et al.*, 1976) to encode the pilus subunit, pilin. The gene sequence revealed an ORF of 121 codons (Frost *et al.*, 1984). The pilin subunit purified from F pili was determined to be a 7.2 kDa protein with an acetylated N terminus and a composition consistent with the last 70 aa of the sequence and is

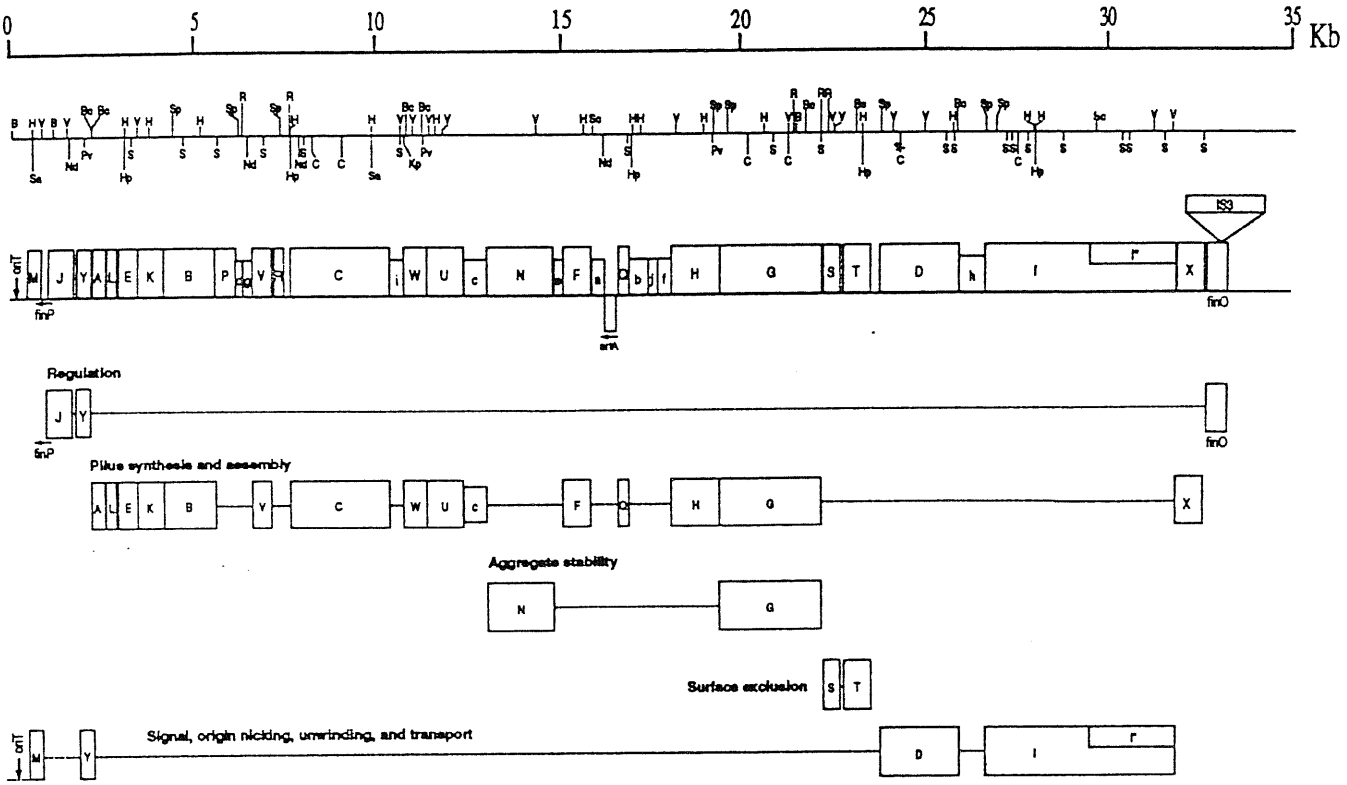


Figure 1.8. Organization of the F transfer region sequence. The top line gives the length in kilobases of the F transfer region. The second line indicates the position of some important restriction sites. The fhird line represents the genes and gene products; capital letters refer to *tra* genes, lowercase letter refer to *trb* genes, *artA* and *finP* are transcripts in the anti-orientation, *oriT* is the origin of transfer, IS3 refers to the IS3 element within the *finO* gene, and I* refers to a gene product encoded within the *traI* gene. The last five lines represent the functions of the *tra* genes identified to date.
From (Frost et al., 1994)

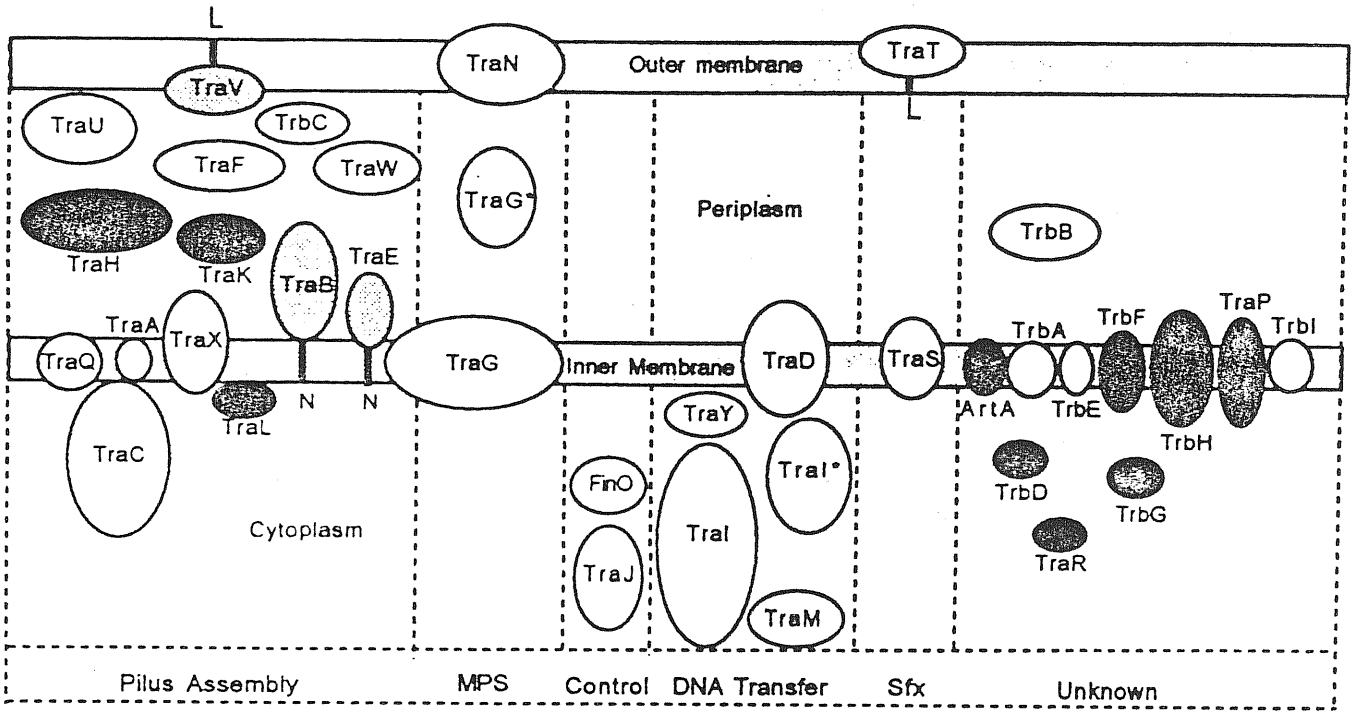


Figure 1.9. Diagram of the location of *tra* gene products. The *tra* gene products are separated according to function: pilus assembly, MPS (mating-pair stabilization), control (regulation of *tra* gene expression), DNA transfer, Sfx (surface exclusion), and unknown. Thin rectangles indicate single membrane-spanning regions, and L and N indicate lipoprotein and N-terminal anchor respectively.
From (Frost et al., 1994).

highly homologous in sequence to other F-like pilins (Frost *et al.*, 1985; Frost et al., 1984). The serological differences among the F-like pilins correlate with differences in sequence at the N-terminus and to a lesser extent the C-terminus of pilin (Finlay *et al.*, 1985; Frost et al., 1985). Using monoclonal antibodies specific for the acetylated N-terminus, it was demonstrated that the N-terminus was not exposed on the sides of the pilus tip (Frost *et al.*, 1986).

No missense mutation in the F pilin gene has been isolated which affects filamentous phage attachment to the pilus tip although F, R100-1 and ColB2 have different sensitivities to these phage (Frost et al., 1985; Willetts & Maule, 1986). The presence of an adhesin at the F pilus tip has been suggested (Anthony *et al.*, 1994), but the gene coding for the adhesin has not yet been identified.

1.2.3.2. *IncN group plasmids and the N pilus.* IncN plasmids have broad-host range, that is, they can conjugate to, and stably replicate in, a wide variety of gram-negative bacteria. They are isolated from *Shigella, Escherichia, Salmonella, Klebsiella, Proteus, Providence* (Hedges, 1972), *Enterobacter,* and *Aeromonas* (Aoki *et al.*, 1977; Arai & Ando, 1979). They were stable even in *Vibrio cholerae* (Arai & Ando, 1980; Arai *et al.*, 1974). Even among those bacteria in which they are not stably maintained without selection, transfer of the IncN plasmids from *E. coli* occurs at high efficiency.

Plasmids of this group are known to determine the synthesis and assembly of N pili, and as for several other groups of conjugative plasmids, there is a correlation between the presence of these pili and conjugation proficiency. *E. coli* strains that are N piliated are sensitive

1.2.3.2. IncN group plasmids and the N pilus. IncN plasmids have broad-host range, that is, they can conjugate to, and stably replicate in, a wide variety of gram-negative bacteria. They are isolated from *Shigella, Escherichia, Salmonella, Klebsiella, Proteus, Providence* (Hedges, 1972), *Enterobacter,* and *Aeromonas* (Aoki *et al.*, 1977; Arai & Ando, 1979). They were stable even in *Vibrio cholerae* (Arai & Ando, 1980; Arai *et al.*, 1974). Even among those bacteria in which they are not stably maintained without selection, transfer of the IncN plasmids from *E. coli* occurs at high efficiency.

Plasmids of this group are known to determine the synthesis and assembly of N pili, and as for several other groups of conjugative plasmids, there is a correlation between the presence of these pili and conjugation proficiency. *E. coli* strains that are N piliated are sensitive to two groups of N-pilus-specific bacteriophage, the filamentous phage IKe and I2-2 (Khatoon et al., 1972) and the icosahedral phage PRD1 (Olsen *et al.*, 1974) or closely related phage both of which attach to the tips of N pili. Bacteria carrying IncN plasmids are very poor donors in liquid culture, they conjugates far more efficiently in solid media (up to 10.000-fold higher) (Bradley, 1980; Dennison & Baumberg, 1975). This can be well explained by the fragile nature of N pili: these pili are readily detached from bacterial cells and found preferentially in the culture supernatant..

The morphology of N pili has been described (Bradley, 1979; Bradley, 1980): N pili are sharply pointed at their distal ends and are about 9.5 nm thick. Their protruding lengths vary between 0.1 and about 0.5 mm, but longer rods (> 1.0 mm) are present in cell-free

suspensions. Their most distinctive morphological characteristic is their inflexibility; they are straight rods and break rather than bend.

Genetic studies have been performed on several IncN plasmids; the transfer region of three of them has been analysed in detail. The *tra* genes of plasmid pKM101 are located in three discrete regions, two of which (TraI and TraII) are required in phage infectivity, for a total length between 16 and 20.5 kb (Winans & Walker, 1985). The transfer genes of plasmid pCU1 are within a single region 15.7 kb long, possibly interrupted by sequences not involved in the transfer functions (Thatte *et al.*, 1985). In plasmid N3, the transfer region is segregated in at least two domains (Glocker & Rasched, 1990).

These results indicate that the conjugal transfer region of IncN plasmids differs in many ways from the one of the more widely studied IncF plasmids (Frost, 1993). The genetic organisation and size of the Tra regions of the two types of plasmids are very different, being the *tra* genes of F encoded in a single contiguous region of DNA approximately 35 kb long.

Recently the complete sequence of the TraI and TraII regions of pKM101 conjugal transfer genes was determined (Pohlman *et al.*, 1994). Eighteen genes have been identified, at least eight and probably eleven of which are required for efficient conjugation. In the same paper the use of Tn*phoA* (Manoil & Beckwith, 1986) to construct translational fusions between pKM101 genes and the *E. coli phoA* gene, which encodes alkaline phosphatase, is described. This experiment provides evidence that at least 11 of the 18 genes are either fully or partially exported from the cytoplasm. Experiments to

visualise the products of these proposed genes have not been performed. Seven of these genes (*traA*, *traB*, *traC*, *traD*, *traF*, and *traG*) were previously shown to be required for conjugal transfer and for sensitivity to donor-specific bacteriophage (Winans & Walker, 1985).

The *tra* genes are homologous to and colinear with genes found in the virB operon of *Agrobacterium tumefaciens* Ti plasmid (Kuldau *et al.*, 1990; Shirasu & Kado, 1993; Shirasu *et al.*, 1990; Ward *et al.*, 1988; Ward *et al.*, 1990). Seven pKM101 *tra* genes are also homologous to *ptl* genes of *Bordetella pertussis*, which direct the export of pertussis toxin (Lessl *et al.*, 1992). No homology with F-pilus genes has been found.

1.2.4. Role of pilus in conjugation and in phage penetration.

Conjugation is a process by which transfer of DNA from a donor to a recipient cell occurs during cell-cell contact. For the initial mating contact to occur between conjugating bacteria, the donor cell must produce pili. It is controversial whether these structures are merely involved in a contractile mechanism to bring the cells into direct contact to permit membrane fusion and pore formation, or whether they act as actual canals for the transport of DNA. A recent study supports the canal hypothesis (Harrington & Rogerson, 1990): donor and recipient cells in a conjugation experiment were separated by a filter 6 microns thick with straight-through pores 0.01 to 0.1 micron in diameter. The filters are thick enough to prevent direct cell-to-cell contact and the pores allow the passage of extended F pili. Conjugative translocation was found, but the translocation rate was very low. This may suggest that DNA translocation through a pilus may be an exception. Other observation led to the opposite suggestion, namely

that the pilus is needed only for the initial contact between the mating partners and that DNA translocation occurs during close cell-to-cell contact. Disaggregation of pili by addition of SDS does prevent successful conjugation when SDS is added before the formation of close mating pairs but has no effect when it is added at the stage of close cell-to-cell contact (Achtman *et al.*, 1978).

A number of studies have strongly suggested that infection by filamentous phage is a multistep process (Jacobson, 1972; Marvin & Hohn, 1969). Initially, the particle binds to the tip of the pilus via an interaction with the phage-encoded gene III protein located at one end of the phage particle. It is postulated that the pilus subsequently depolymerizes into the membrane, thereby bringing the tip of the phage particle to the surface of the bacterium. There the particle must interact with additional proteins in order for the viral DNA to translocate into the cytoplasm (Sun & Webster, 1986; Sun & Webster, 1987). This step in the entry pathway is defined by mutations in the *tolQ*, *tolR*, and *tolA* genes of *E. coli* which render F^+ bacteria uninfectable by filamentous phage. These mutants are still fully competent to undergo conjugation and are able to bind filamentous phage to the tips of their pili (Sun & Webster, 1986).

Phage penetration involves insertion of the gVIIIp into the inner membrane of the host as the DNA enters the cell. The orientation of the phage relative to the host membrane is the same for penetration and subsequent assembly, which occur without cell lysis. Data suggest that phage penetration and phage assembly involve the same basic steps, but in reverse order. A molten globule intermediate for fd phage

penetration and assembly has been proposed (Dunker *et al.*, 1991a; Dunker *et al.*, 1991b).

It has been demonstrated that filamentous phage can infect F⁻ bacteria at extremely low frequencies, approximately 10^{-6} (Russel et al., 1988). These observations suggest that pili serve as the primary receptors for the filamentous phage, providing both an attachment site and a mechanism to bring the phage close to the bacterial membrane where the TolQ, TolR, and TolA proteins are located (see next section).

1.2.4.1. Tol proteins, inner proteins necessary for DNA injection. The *tolQ*, *tolR*, and *tolA* mutants of *E. coli* cannot be infected by phage fd and other phage which use the tips of conjugative pili as receptors (Bradley & Whelan, 1989; Russel et al., 1988). These mutants have a pleiotropic phenotype: they are hypersensitive to certain antibiotics and detergents and are tolerant to certain colicins. The precise functions of the Tol proteins are not known, but it seems that they stabilize the outer membrane and are perhaps involved in membrane assembly (Levengood-Freyermuth *et al.*, 1993). It is not known whether one or several of the Tol proteins can form pores for the translocation of colicins or phage DNA. It may well be that the Tol proteins only guide colicins or phage DNA from the receptor site to the actual site of translocation through the inner membrane. During this process, gIIIp of phage fd seems to interact with one or several Tol proteins. This was concluded by the observation that cells producing gIIIp from a plasmid can no longer be infected by fd and show an increased tolerance to colicins (Boeke et al., 1982). Apparently the plasmid-borne gIIIp can leave the cytoplasm and associate with the Tol

no longer free to interact with phage-derived gIIIp or colicin.

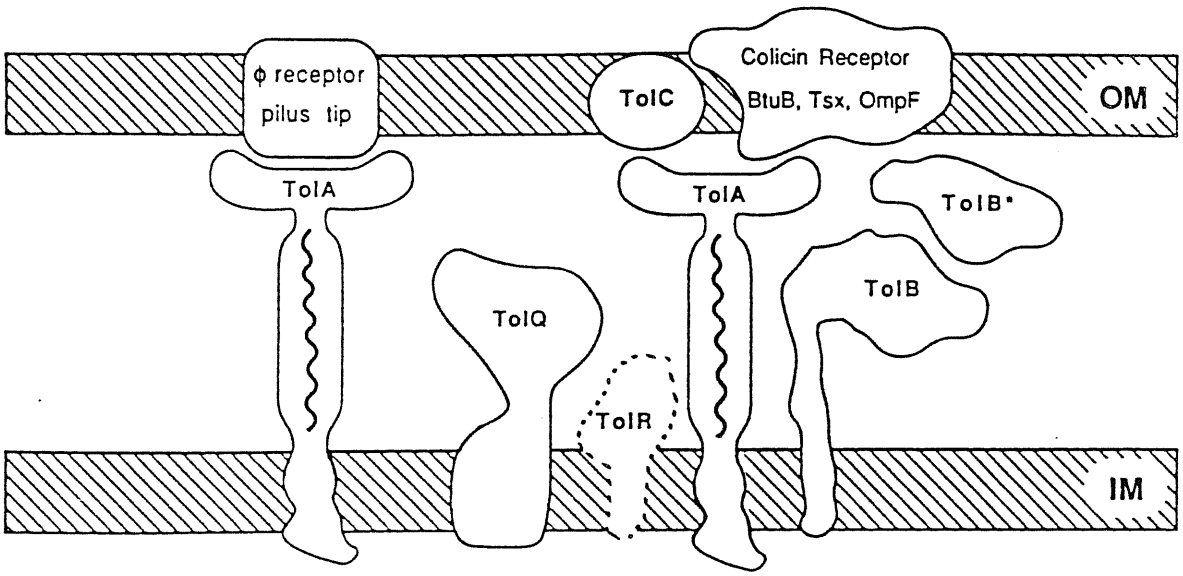


Figure 1.10. A schematic depiction of the possible location of the best characterized Tol proteins. A broken line is used to depict TolR since its position is based on its predicted amino acid sequence. OM, outer membrane; IM, inner membrane. From (Webster, 1991).

1.3. Surface display.

1.3.1. Filamentous phage libraries.

The last decade has seen the development of the phage display library. Filamentous bacteriophage such as fd have been used to display peptides and proteins on the viral surface. In particular, the expression of antibody fragments (McCafferty *et al.*, 1990) as fusions with gIIIp has allowed the purification of phage particles containing a specific fusion from an excess of phage particles lacking it, by affinity chromatography using immobilised antigen or receptor respectively. Phage particles isolated in this way contain the DNA encoding the fusion protein, and so the binding characteristics of molecules expressed on the surface can be used to isolate the genes encoding those molecules. Furthermore, there is the potential for manipulating this DNA by site-directed mutagenesis and then selecting improved molecules using an affinity matrix.

Foreign peptides can be engineered genetically into the N-terminal region of the major coat protein (gVIIIp). Up to six residues can be inserted close to the N-terminus. generating a recombinant virion in which all copies of gVIIIp display the peptide (Greenwood *et al.*, 1991; Minenkova *et al.*, 1993). In order to accomodate larger peptides it is necessary to construct a hybrid virion in which the modified coat proteins are interspersed with wild type protein subunits (Felici *et al.*, 1991; Greenwood et al., 1991).

Many proteins have now been functionally displayed (Table 1.5.).

Table 1.5.

1.3.1.1. New developments in phage libraries. Filamentous phage systems do not allow the construction of cDNA expression libraries because translational stop codons present at the 3′ end of non-translated regions of eukaryotic mRNA (Abelson1980) prevent construction of fusion proteins N-terminal to the gIII gene product. On the other hand fusions to the membrane bound C-terminus of gIII would hamper the incorporation of the fusion protein into the phage coat.

Two different solutions have been found to simplify the screening of cDNA libraries.

The first interesting display system is the so called pJuFo system (Crameri *et al.*, 1994; Crameri & Suter, 1993) based on the ability of the Leu zippers Jun and Fos to associate to each other. pJuFo is a phagemid vector constructed to fuse the modified Jun Leu zipper with the C-terminal domain of the pIII protein. The Jun leucine zipper is flanked by cysteine residues and expressed as an N-terminal fusion to the C-terminal domain of the gIII product. The Fos leucine zipper flanked by cysteine residues is used to produce Fos-cDNA fusions secreted via the *pelB* leader. The Jun and Fos Leu zippers, flanked by CysGlyGly and GlyGlyCys aa at the N- and C-terminus respectively, covalently link in the periplasmic space by proper disulfide bond formation. The pJuFo system allows isolation of cDNAs encoding proteins for which a ligand is available and circumvents immobilisation of the libraries on solid-phase supports which hamper selective enrichment of clones expressing the desired protein.

A second solution to the problem of functionally express cDNA libraries on the surface of filamentous phage has been recently described (Jespers *et al.*, 1995). Proteins have been covalently attached to to the C-terminus of the minor coat protein gVIp. Fusion to gVIp may be particularly useful for phage display of foreign proteins whose C-terminal ends participate in protein conformation or activity.

To be incorporated into active filamentous phage, polypeptides fused to coat proteins must be translocated into the host cell periplasm space before extrusion from the host as assembled phage particles (see section 1.2.1. on filamentous cell cycle). Many cytoplasmic proteins fused to the phage protein will interfere with the passage of the fusion product from the cytoplasm to the periplasm. Therefore, it is worthwhile to consider other phage systems for surface-display potential, e.g. lytic phage.

Recently bacteriophage lambda has been prospected as a possible phage system for surface-display by two different groups (Dunn, 1995; Maruyama et al., 1994). They both describe a gVp-fusion expression system for λ surface display. Brenner's group shows that the complete *E. coli* β-galactosidase and the plant *Bauhinia purpurea* agglutinin can be fused to the C-terminus of the truncated phage tail protein. Dunn used a modified gVp provided with a Ser-Gly linker. A basic C-terminal peptide sequence (RRASV, a target site for cAMP-dependent protein kinase) and β-galactosidase have been fused to the modified gVp. In both systems the fused proteins are active, and the peptide sequence is efficiently phosphorylated by the appropriate protein kinase. This system could be more appropriate for proteins that fold up

in the cytoplasm and could therefore complement the filamentous phage system.

1.3.2. Bacterial cell surface.

The expression of polypeptides on the surface of Gram-negative bacteria has been actively pursued for several years, in part because of interest in vaccine production. For reviews on bacterial cell surface display see (Georgiou *et al.*, 1993; Hofnung, 1991; Little *et al.*, 1994; Little *et al.*, 1993).

Gram-negative bacteria, such as *E. coli* and *Salmonella*, have two layers of membranes in the cellular envelope - the inner or cytoplasmic membrane and the outer membrane (Figure 1.11.).

Between these membranes is a periplasmic space containing a peptidoglycan layer that provides the cells with mechanical rigidity. In this periplasmic space there are also a variety of hydrolases and binding proteins. The composition of the outer membrane is somewhat unusual. This membrane bilayer is asymmetric, having an inner (periplasmic) leaflet composed of phospholipids and an outer (extracellular) leaflet formed by lipopolysaccharide (LPS). Unlike phospholipids having two acyl chains, LPS has six or seven saturated fatty acid chains.

There are two kinds of surface-exposed proteins in Gram-negative bacteria: (1) subunits of cellular appendanges such as pili or flagella which do not contain transmembrane domains; and (2) integral outer membrane proteins. In general the outer surface-exposed loops of

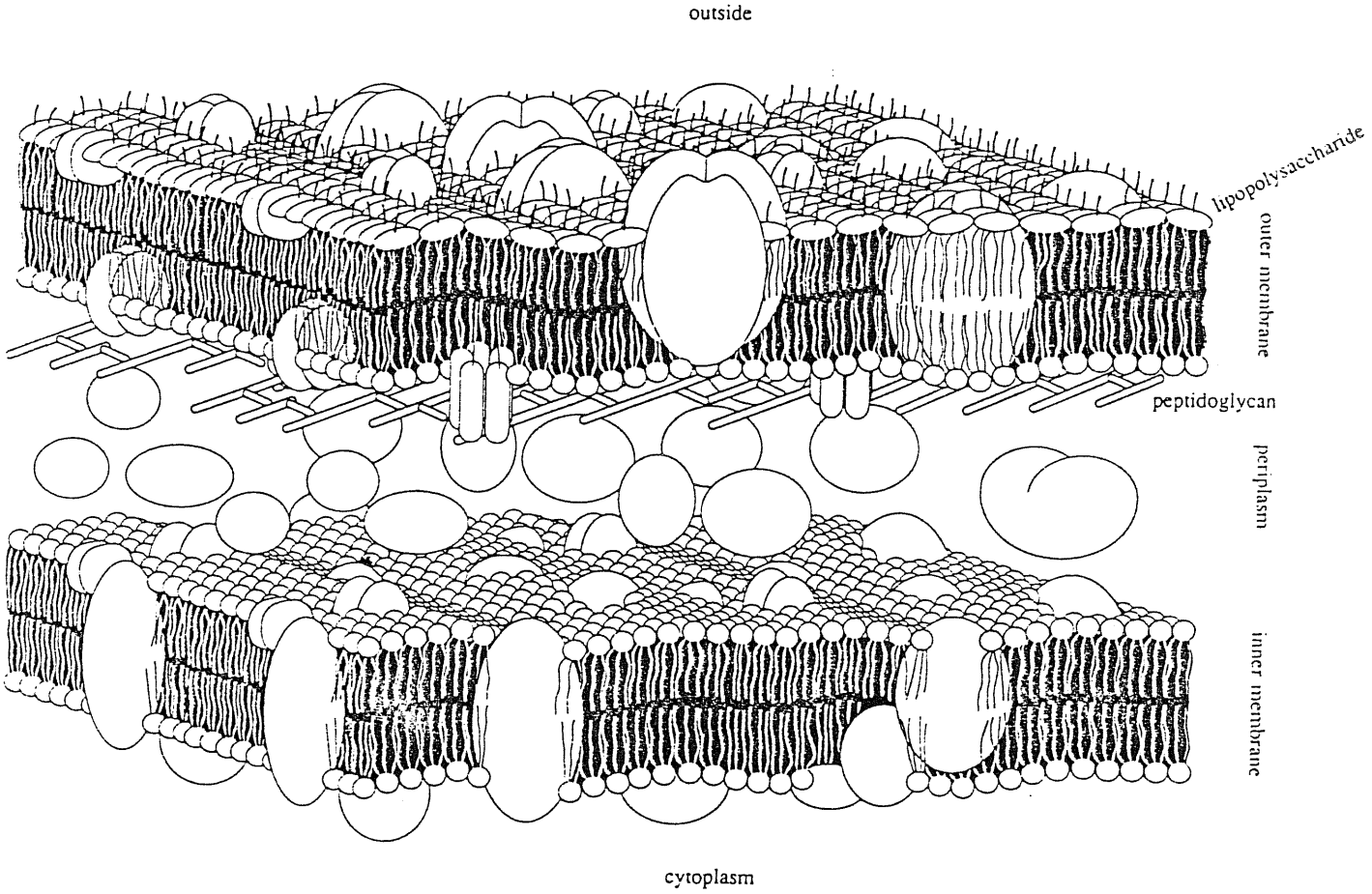


Figure 1.11. Schematic diagram of the organization of the cell envelope of Gram-negative bacteria. The lipopolysaccharide of *E. coli* K12 presents only short chains of O-polysaccharides, so that outer membrane proteins are relatively accessible from the outside. From (Hofnung, 1991)

outer membrane proteins can tolerate amino acid substitutions, insertions and deletions without perturbing the overall conformation.

Bacterial surfaces are more suitable for presenting higher numbers of proteins than in phage display and may also offer some advantages for the surface expression of peptides. To be displayed usefully on the surface of *E. coli*, a peptide or protein must cross the cytoplasmic membrane and the bacterial periplasm and then insert into, and be retained in the outer membrane in the correct conformation and orientation, such that it retains the desired biological properties and is accessible to the extracellular medium. A number of different approaches have been developed, according to the type of carrier used to target the peptide or protein to the cell surface (Figures 1. 12. and 1.13.).

One approach is to introduce a heterologous sequence into the surface-exposed loops of an outer-membrane protein. This system is well suited to the display of relatively short sequences: for instance LamB protein has been extensively used by Hofnung's group (Hofnung, 1991) and references therein]. The main limit to this approach is the need for the N- and C-termini of the insert to be close in three-dimensional space in order to preserve the tertiary structure of the carrier protein.

Another approach similar to the previous one is to use cell-surface structures, such as flagella or fimbriae. Recently bacterial flagella have been used to display random peptide libraries (Lu *et al.*, 1995). The entire coding sequence of *E. coli* thioredoxin (*trxA*) was inserted into a dispensable region of the gene for flagellin (*fliC*), the

Table 2. Proteins that have been functionally displayed on phage[a]

Protein	Molecular weight (kDa)	Fusion (pIII or pVIII)	Valency	Libraries constructed?	Refs
Cytokines/growth factors					
hGH	22	III	Mono	√	14, 17
IL-3	15	III	Poly		20
TNF	3 × 17	III	Mono		g
IL-8	2 × 8.4	III	Mono	√	h
Receptors/lectins					
CD4[d,e]	25	III	Poly		16
FcεRI α-chain[d,e]	19	III	Poly		21, 22
Ricin B-chain	32	III	Poly		23
PDGF receptor[d]	55	III	Poly		16
Protein A (B domain)	6.2	III	Mono		24
Enzymes					
Trypsin[c]	23	III/VIII	Poly		25
Staphylococcal nuclease	16.5	III	Poly		16
Alkaline phosphatase	2 × 60	III/VIII	Poly		26
Protease inhibitors					
BPTI	6.5	III	Poly	√	27
APPI	6.5	III	Mono	√	i
PAI	42	III	Mono	√[f]	28
Antibody fragments					
Fab	2 × 25[b]	III/VIII	Both	√	18
scFv	25	III	Both	√	29
DNA binding proteins					
Zif268	10	III	Mono	√	30, 31
Natriuretic peptides					
ANP	3.1	III	Mono	√	32

[a] Note that most successfully displayed proteins are normally extracellular and secreted, and compatible with secretion of pIII and pVIII through the oxidizing periplasmic environment during phage assembly. Although the intracellular zinc finger protein Zif268 can be displayed, intracellular proteins with 'free' (unliganded) cysteines may prove more recalcitrant. Protein size does not seem to be a major factor.

[b] Proteins that are non-covalent heterodimers (such as antibody Fab fragments) can be displayed by secreting one chain in free form and fusing the other to pIII or pVIII, so that association can occur in the periplasm.

[c] Trypsin phage co-purifies with ecotin, an endogenous *E. coli* protease inhibitor.

[d] Extracellular domains displayed.

[e] Individual domains also displayed.

[f] No selection performed.

[g] M. J. Zoller, pers. commun.

[h] H. B. Lowman, pers. commun.

[i] T. R. Hynes, A. A. Kossiakoff and J. A. Wells; M. Dennis and R. A. Lazarus (both unpublished).

Figure 1.12. Scheme of the proteins or peptides expressed on bacterial cells. The different possible approaches to protein display on the surface of *E. coli* are reported as discussed in the test. From (Hokney, 1994) and references therein.

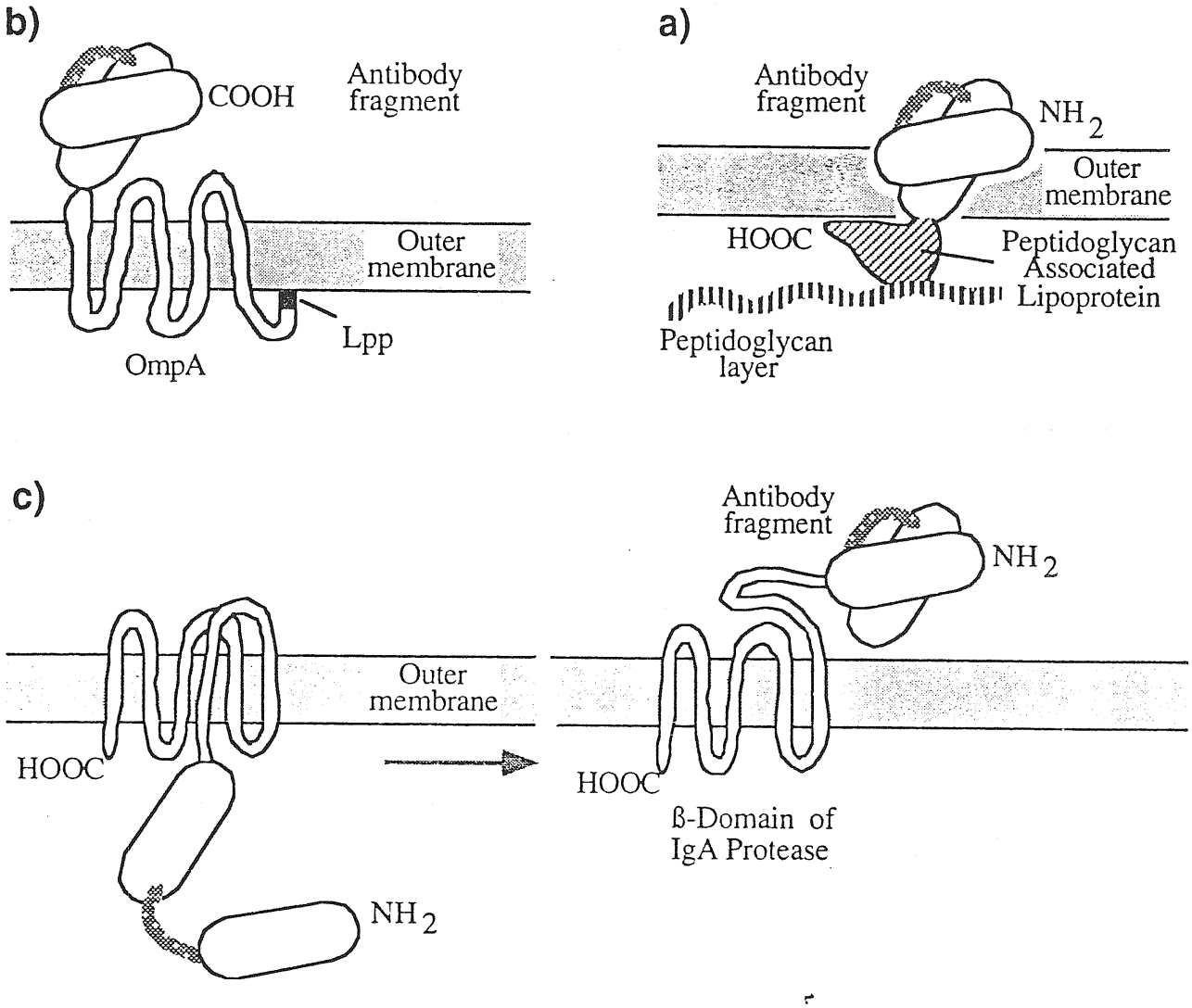


Figure 1.13. Bacterial surface display systems. Single chain antibody (scFv) fusions to: a, PAL (peptidoglycan associated lipoprotein); b, OmpA-Lpp; c, β-domain of IgA protease.
From (Little et al., 1994).

major structural component of *E. coli* flagella. The resulting fusion protein (FLITRX) was efficiently exported and assembled into partially functional flagella on the bacterial surface. A diverse library of random dodecapeptides were displayed in FLITRX on the exterior of *E. coli* as conformationally constrained insertions into the thioredoxin active-site loop, a location known to be a highly permissive site for the insertion of exogenous peptide sequences into native thioredoxin (LaVallie *et al.*, 1993). The inserted sequences are conformationally constrained.

The last approach is to fuse a large protein, up to 40-50 kDa, to the N- or C-terminus of a lipoprotein, lipoprotein outer-membrane protein fusion (Lpp-OmpA fusion), or IgA protease. The peptido-glycan-associated lipoprotein (PAL) has been used to target recombinant antibodies to the cell surface by fusing its N-terminus with single-chain variable domains (Fuchs *et al.*, 1991). The advantage of such a system is that the antigen-binding site projects outwards, away from the cell surface, and thus being very accessible.

1.4. About this thesis: selection by infection.

Even if phage display has been shown to be a powerful mean to select epitopes or proteins expressed on the phage surface, it has also shown some limitations. The panning with purified proteins or receptors may be very difficult or even impossible when the antigen has to be cloned or purified. For this reason the development of a simpler method which could join the selection to the infection would be of paramount importance.

Two systems have been considered in this thesis.

Initially the lambda phage which recognizes the LamB protein on the bacterial cell surface has been considered (section 3.1.). This system has been chosen because LamB have been used to clone epitopes mostly for vaccine production

The other system that has been considered is the filamentous phage/pilus system (section 3.2.). Experiments in our laboratory (Marzari *et al.*, ; Sblattero, 1995) have shown that it is possible to extent the host range of the filamentous phage fd by genetic manipulation of the gIII protein. Addition of the receptor domain from the related filamentous phage IKe to the N terminus of the fd gIIIp is sufficient to permit significant infection of bacteria bearing N pili (Figure 1.14.). In a selection by infection scheme an antibody fused to the gIIIp could mimic the receptor function if an epitope is fused to the pilus (Figure 1.15).

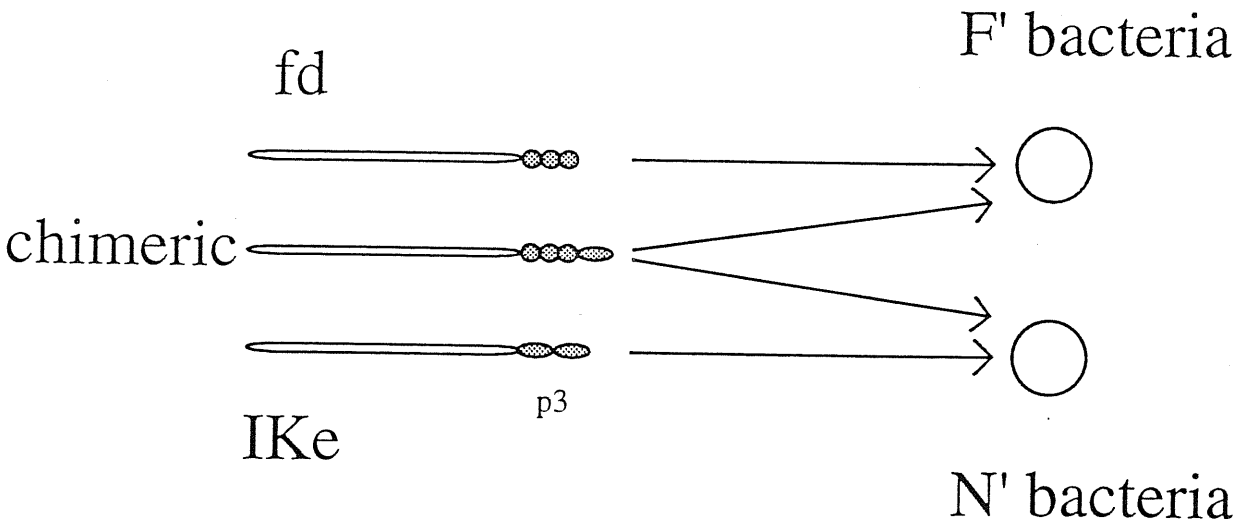


Figure 1.14. Scheme of the host ranges of the filamentous phage fd, IKe, and the fd containing chimeric gIIIp with parts of IKe gIIIp attached to the end of the fd gIIIp. fd infects F' bacteria, IKe infects N' bacteria, while chimeric phage are able to infect both.

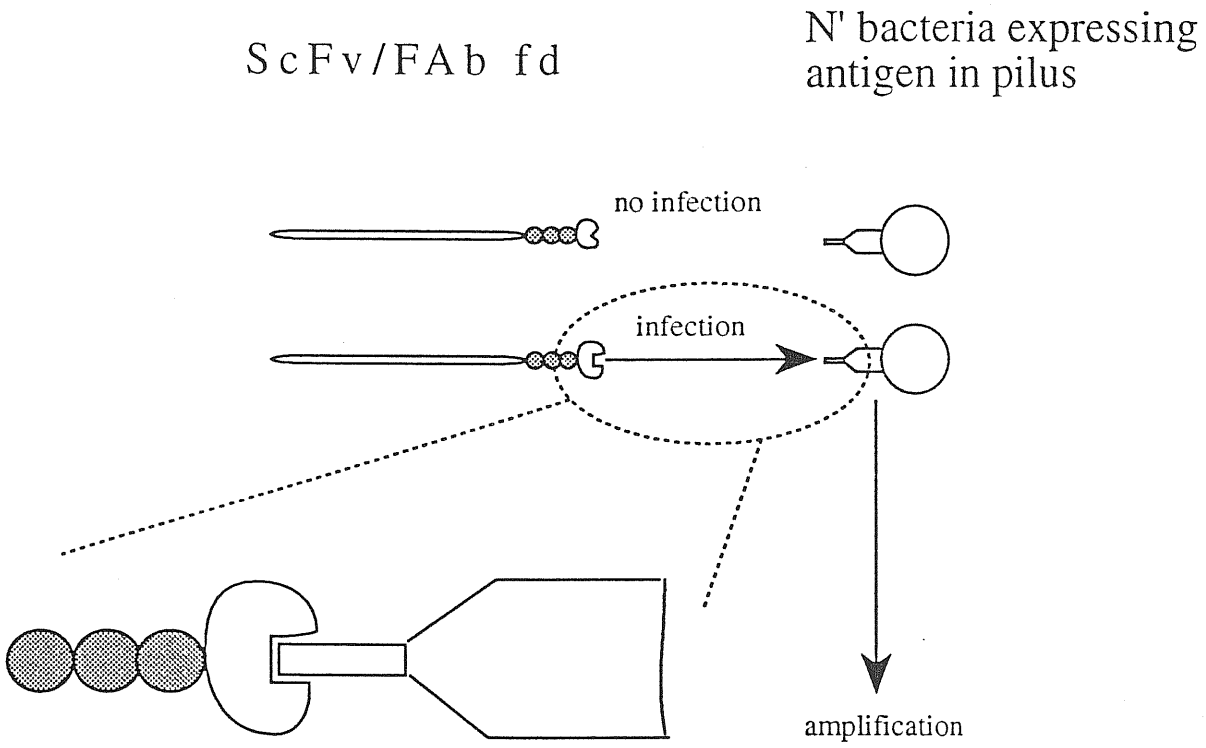


Figure 1.15. Scheme showing how selection by infection should work for the filamentous phage/pilus system. The same can be applied to different phage/bacteria systems.

During the course of this thesis a new system for selection and amplification of phages (SAP) for antibody libraries was developed (Duenas & Borrebaeck, 1994). It is based on the observation that the minor coat protein, gIIIp can be functionally separated into two parts, an N-terminal domain (N-pIII) binding to the F′ pili and allowing the infection of *E. coli*, and a C-terminal domain (pIII-C), anchored at the tip of the phage body. The truncated minor coat protein pIII-C by itself does not allow infection. However protein-protein interaction via specific target proteins can place the N-terminal part of the minor coat protein N-p3 on the surface of the phage in an appropriate position. thus restoring infectivity. Non infectious phage particles become infective by a specific interaction between displayed antibody fragment and a soluble protein consisting of the antigen fused to the N-terminal part of gIIIp.

The use of a soluble protein makes this procedure not so easily handful as it may seem. The presence of the soluble phage receptor can cause the pilus retraction (Jakes et al., 1988) such that further phage infections are impossible.

2. MATERIALS AND METHODS.

2.1. Abbreviations.

a.a., amino-acid
amp, amp^r, ampicillin, ampicillin resistance
APS, ammonium persolphate
BCIP, 5-bromo-4-chloro-3-indolyl-phosphate
bp, base pair
DDT, dithiothreitol
EDTA, disodium ethylenediaminetetraacetate
IPTG, Isopropyl-b-D-galactoside
kan, kan^r, kanamycin, kanamycin resistance
p.f.u., plaque forming unit
PEG, polyethylene glycol
SDS, sodium dodecyl sulfate
TEMED, N,N,N,N,Tetramethylethylendiamin
tet, tet^r, tetracycline, tetracycline resistance
X-gal, 5-bromo-4chloro-3-indolyl-b-D-galactoside

2.2. Media and solutions.

Different media and solutions have been used:
LB (Luria-Bertani): 10 g Bactotryptone, 5 g yeast extract, 10 g NaCl per liter;
2xYT: 10 g Bactotryptone, 5 g yeast extract, 5 g NaCl per liter;
SOB: 20 g Bactotryptone, 5 g yeast extract, 0.5 g NaCl per liter;
BCIP: BCIP was purchased from Sigma and incorporated into the agar at a final concentration of 100 mg/ml.
TFB (per litre): KCl (ultrapure) = 7.4g (100mM): MnCl.4H2O = 8.9g (45mM): CaCl2.2H2O = 1.5g (10mM): HACoCl3 (hexamine cobalt (III) trichloride) = 0.8g (3mM): K-MES 20ml of 0.5M stock (pH 6.3) (10mM). The solution is prepared and filtered after dissolving salts in 10mM K-MES and stored at 4°C
DnD (per 10ml): DTT = 3.3g (1M); DMSO = 9ml (90% final); Potassium acetate = 100µl of 1M stock (pH 7.5) (100mM). Stored at -20°C

PBS 0.1 % Tween20, 2% dry milk.

2.3. Bacterial strains, plasmids and bacteriophage.

Bacterial strains used in this thesis are derivatives of *E. coli* K-12:

DH5α [*supE*44 Δ*lacU*169 (80*lacZ* ΔM15) *hsdR*17 *recA*1 *gyrA*96 *thi*-1 *relA*1];

DH5F' [F' *supE*44 Δ*lacU*169 (80*lacZ* ΔM15) *hsdR*17 *recA*1 *gyrA*96 *thi*-1 *relA*1];

LE392 (*supE*44 *supF*58 *hsdR*514 *galK*2 *galT*22 *trpR*55 *metB*1 *lacY*1);

CC118 [*araD*139 Δ(*ara leu*)7697 Δ*lacX*74 *phoA*Δ20 *galE galK thi rpsE rpoB argE*(Am) *recA*1]

Plasmids and phage used in this thesis are listed:

M13 ( ), IKe ( ), IKK (a kanamycin-resistant IKe phage) (Endemann *et al.*, 1992), λTn*phoA*, λ5187.

2.4. Methods of molecular biology.

Standard methods of molecular biology were used as described in (Sambrook & Maniatis, 1989).

2.4.1. Extraction and purification of plasmid DNA.

Plasmids were purified from cultures grown in liquid medium containing the appropriate antibiotic. Cultures were inoculated with a single bacterial colony picked from an agar plate and grown overnight in a shaking bath at 37°C. The alkaline lysis method were used for almost all preparation (Sambrook & Maniatis, 1989).

To isolate large plasmids (more than 15 kb) the lysis by sodium dodecyl sulfate has been the method of choice (Sambrook & Maniatis, 1989).

When necessary the plasmid DNA was further purified by equilibrium centrifugation in CsCl-ethidium bromide gradients (Sambrook & Maniatis, 1989).

2.4.2. Gel electrophoresis and DNA extraction.

The Geneclean II kit from Bio101 has been used to remove and purify DNA from agarose gels, following the standard protocol.

2.4.3. Preparation and transformation of competent *E. coli.*

Different methods have been applied depending on the need for efficiency of transformation. For transformation of a ligation product Hanahan's protocol (Hanahan, 1983). Bacteria from a frozen stock are grown overnight on SOB plates. A single colony is picked and bacteria are grown in SOB in a flask with volume 10-30 times greater than volume of SOB, until OD550 0.2-0.4. Flask is placed on ice 10 min. From now on the bacteria must always be kept on ice. After a centrifugation of 5-10 min at 3000-5000 and 4°C, supernatant is poured off and remaining liquid removed with blue tip. The bacterial pellet is resuspend in a volume of TFB 1/12.5 of original SOB volume. Bacteria

are left on ice 10 minutes, then 35µl DnD for each 1ml TFB are added. After 10 minutes on ice another volume of DnD is added. The bacteria are now ready to use, and will remain good for at least 6 hours.

200µl of bacteria are used for transformation with DNA or ligation (volume less than 5µl). Bacteria are incubated with DNA in a Falcon 2063 tube, on ice 10-60 minutes. Heat shock for 1 minute at 42°C and leave on ice at least 2 minutes. If plasmid codes for ampicillin resistance plate immediately. If tet resistance then add 800µl SOB and leave at 37°C for 30 minutes.

Electroporation was used for transformation of lysogens which are sensitive to heat shock at 42°C (they enter the lytic cycle). Bacteria on are streaked on 2XTY or SOB (without Mg) plate and grown overnight in 2XTY or SOB (without Mg). The day after a single colony is grown to early to mid log about OD600 = 0.5-1 (best is about 0.75). Flack containing the bacteria is chilled flask for 15-30 mins on ice. Bacteria are placed into prechilled tubes and spin at 4000xg max for 15 mins in a prechilled rotor. Cells are resuspended in cold water in the same 2XTY volume. After spin, the pellet is again resuspended in cold water in the half 2XTY volume. The last step cells are resospended in 2ml of glycerol 10% for any 50 ml of original volume. For electroporating, 40µl cells are mixed with 1-2µl DNA and let sit on ice for 1 min. The machine is set to give 25µF,2.5kV with the pulse controller set to 200 ohms. The mixture of cells and DNA is transfered to a pre chilled 0.2 cm cuvette and shaked to the bottom. The cuvette is placed in the electroporetion chamber. Pulse once should yield a pulse with a time constat of 4.5 to 5 msec with a field strength of about 12.5kV Immediately add 1 ml of SOC and resuspend the cells, transfer to a 6032 tube and incubate as normal. Then plate.

2.4.4. Restriction and modification.

Standard protocols are used. Enzymes were from Pharmacia, Promega, Boehringer Mannheim and Biolabs.

2.4.5. Detection of proteins expressed from cloned genes.

Bacteria expressing the target protein were lysed directly in sample buffer. SDS-polyacrylamid gels were prepared as described by (Laemmli, 1970). Gels were fixed with methanol:glacial acetic acid and stained with Coomassie Brilliant Blue.

Western blot was performed by transferring proteins to a nitrocellulose mambrane, in a semi-dry system. The membrane was incubated in PBS for 1 hr with gentle agitation. After 5 washes in PBS the membrane was incubated in a solution of antibody for 1 hr with gentle agitation. After 5 washes a labelled secondary antibody (peroxidase conjugated) was added and incubated for one hour. Then the membrane, after 5 more washes, was transferred to the substrate solution (3-3'-diaminobenzidine tetrahydrochloride tablet dissolved in PBS containing 3% hydrogen peroxidase).

2.5. Experimental methods for use with lambda phage.

2.5.1. Growth of bacterial host strains.

The commonly used host for the growth of phage λ is LE392. For the titration and propagation of λ, bacteria are grown in a rich medium, such as SOB, with aeration at 37°C, in the presence of 0.4% maltose, but in the absence of glucose.

2.5.2. Titration of phage.

The phage lysate is diluted serially in growth medium. 0.2 ml of each dilution are added to 0.2 ml of indicator bacteria, (prepared as described in 2.5.1.). This preadsorption mixture is mixed by shaking and incubated for 15-20 min. at 37°C. About 4 ml of top agar (45°C) is then added to the tube and the total mixture is plated on LB plates. Plates are incubated overnight at 37°C.

2.5.3. Maintenance and testing of lysogens of bacteriophage lambda.

Master stocks of *E. coli* lysogens are stored at -80°C. The presence of a mutation which renders the bacteriophage *cI* gene product temperature-sensitive is checked each time by streaking from the master stock onto two LB agar plates. One plate is incubated at 30°C, the other at 45°C: the bacteria should grow only at 30°C. A single colony is picked and grown overnight at 30°C. The day after the cells are tested for lysogenicity (see 2.5.5.).

2.5.4. Induction of lysogenic lambda with ts repressor.

Lysogens harboring a prophage with temperature-sensitive repressor (cIts mutation should be incubated at 42-45°C for induction. Lysogenic bacterial strain is grown at 30°C. We usually dilute an overnight 1/100 and grown until OD=0.6. The temperature is than shifted to 42-45°C and incubation is continued for 15 minutes at this temperature. This heat shock inactivates the lambda repressor and allows the phage to enter a lytic mode. The culture is transferred to 37°C and further incubated until lysis occurs. Lysis starts after one hour. Once lysed a few drops of chloroform are added and shaking is continued at 37°C for a further 15 minutes. Debris is removed by centrifuging at 5000 rpm for 30' and the supernatant containing the phage can be stored at 4°C for years.

2.5.5. Purification of bacteriophage λ arms.

The arms of lambda were prepared by centrifugation through sucrose density gradient (Sambrook & Maniatis, 1989).

To avoid formation of stuffer fragments, the oligonucleotide complementary to one cohesive end is added when the arm from the other end has to be purified. L3 (5'-AGGTCGCCGCCC-3') is the oligonucleotide that is the 3' sticky end and is to use on 5' end, when desired product is 3' arm, while L5 (5'-GGGCGGCGACCT-3') is the oligonucleotide that is the 5' sticky end and is to use on 3' end, when desired product is 5' arm.

2.5.6. Electroporation of lysogens.

We observed that when it was necessary to introduce a plasmid in a lysogenic cell, the methods for transformation of bacteria in which there is the heat shock at 42°C was lethal for good lysogens. We thus decided to use electroporation to prepare competent lysogenic cells (see 2.4.3.).

2.5.7. In vitro packaging.

The Packagene lambda DNA packaging system from Pharmacia was used for preparation of lysogens.

2.6. Protocols for experiments with transposons.

2.6.1. The Tn*phoA* left end.

Tn*phoA* is a derivative of Tn5 that permits the generation of hybrid proteins composed of alkaline phosphatase lacking its signal peptide fused to amino-terminal sequences of other proteins Manoil, 1985 #341]. The DNA sequence of the left end of Tn*phoA* is shown in Figure 2.1. There are 50 bp of DNA between the left end of Tn*phoA* and the beginning of the 'phoA coding region. Of this sequence, 48 bp derive from $IS50_L$ and 2 bp derive from the Pst1 linker of pCH39. The amino acid residues encoded by this 50-bp sequence are present at the fusion joint of every hybrid protein generated by Tn*phoA* insertion. Tn*phoA* encodes all of alkaline phosphatase except the signal sequence and five amino acid residues of the mature protein.

2.6.2. Preparation of λ::Tn*phoA* phage stock.

Cells of LE392 are grown at saturation (OD 600 0.8) in LB and are infected with λ::Tn*phoA* (10^6 phage in 200 ml of cells). After mixing the cells are incubated at 37°C for 20 min. to allow phage absorption. To the cell and phage mixture 4 ml of melted top agar equilibrated at 45°C is added. The mixture is then spread per plate. An equivalent amount of LE392 cells with no phage added is plated as a control. The plates are incubated at 37°C for 6-8 h and inspected for lysis by comparing with the LE392 control. When confluent lysis is observed on the plates containing the λ::Tn*phoA* infected cells and the control plates show a visible bacterial lawn, the phage is harvested by adding 3 ml of LB onto plates and scraping the top agar into a sterile 50 ml Falcon tube. 100 ml of chloroform are added, the mixture is vortexed thoroughly, incubated on ice for 5 min., and centrifuged at 4 K for 5 min. After decanting the supernatant into a sterile screw-top tube, a few drops of chloroform are added and the lysate is stored at 4°C. The phage have been titered on LE392: we obtained a pfu= 3.10^{10} pfu/ml.

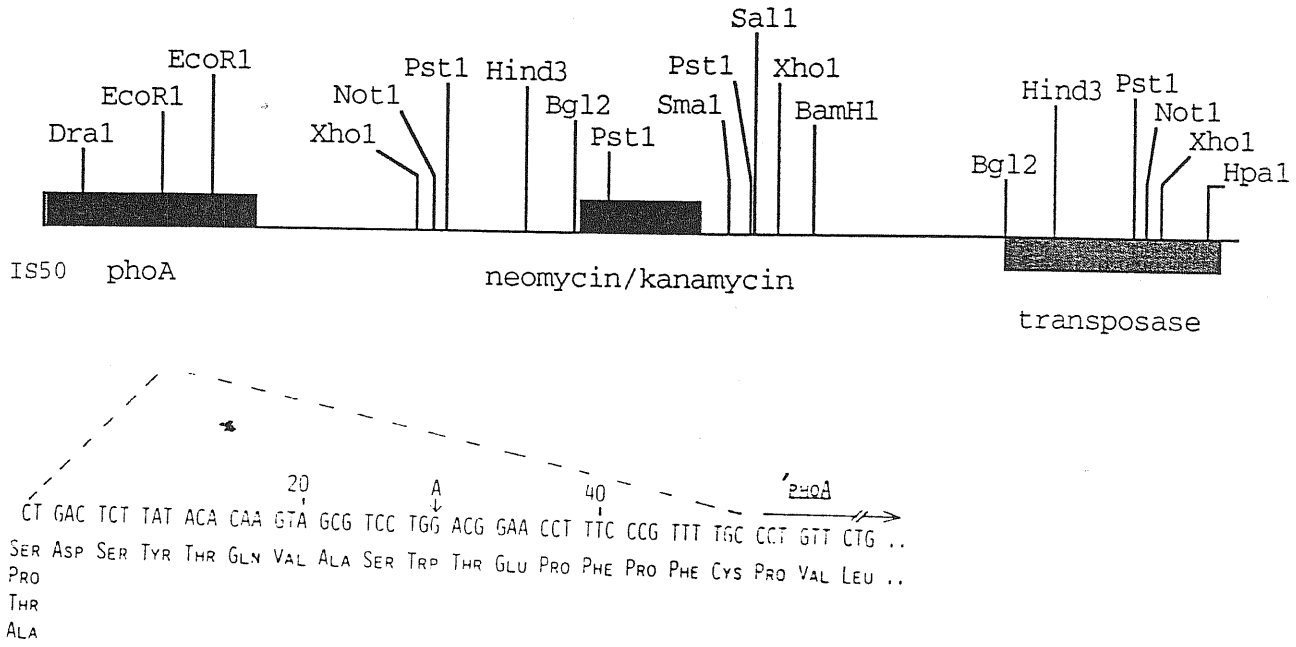


Figure 2.1. Structure of TnphoA, showing the nucleotide sequence of its left end and the deduced amino acid sequence.

2.6.3. Mutagenesis of plasmid DNA with transposon.

The plasmid to be mutagenized, pBGB, was trasformed into a Su°λ bacterial strain, CC118. Plasmid-bearing cells are grown to OD 0.7 in LB (supplemented with the antibiotic selective for the plasmid, amp, and containing 0.5% maltose) at 30°C. 1 ml of cells was incubated with 30 µl of λ::Tn*phoA*, incubated for 30 min. at 37°C and plated on LB agar plates containing amp and kanamycin (eleven plates). Colonies from each plates were scraped up, and plasmid DNA was isolated (as explained in 2.6.3.) and used to transform an *E. coli* Δ*phoA*- strain (CC118) to kanamicin and ampicillin resistance. This step unsures that all colonies contain the Tn*phoA* on the plasmid rather than in the chromosome. The medium used to plate transformants may contain the phosphatase chromogenic indicator BCIP for alkaline phosphatase activity detection (see 2.6.4). Plasmid DNA from single colonies was purified and digested with restriction endonucleases to roughly identify the locus of insertion of each transposon.

2.6.4. Rapid plasmid preparation using cells from an agar plate.

1 ml of LB is added to each plate and colonies are scraped up. After centrifugation at 4K, the bacterial pellet is placed into 200 µl resospension buffer (50 mM glucose, 10 mM EDTA, 25 mM Tris-HCl pH 8.0) in a 1.5 ml eppendorf tube. The tube is placed onto ice and incubated 5min. 400 µl of freshly prepared alkali lysing buffer (1% SDS, 0.2N NaOH) is added to the cells, the mixture is sucked back-and forth in the pipettman tip once or twice to mix the solutions and to help disperse any remaining lumps of cells. After 5 min. of incubation on ice, 300 µl of 3N NaAc pH 4.8 are added to each tube, that are vortexed vigorously and freezed on dry ice. Thaw and centrifuge 15 min. at 4°C in a microfuge. Supernatant is transfered to another eppendorf tube. The DNA is precipitated by adding 500 µl ethanol and placing it at -20°C. After centrifuging at 4°C for 5 min., the pellet is resuspended in 100 µl H_2O.

2.6.5. Active alkaline phosphatase fusions to plasmid gene products.

BCIP was used to detect the presence of active fusions of alkaline phosphatase and *tra* genes. Approximately 1% of the transformants expressed sufficient phosphatase activity to form blue colonies. Plasmid DNA from blue colonies was purified and digested with restriction endonucleases to roughly identify the locus of insertion of each transposon. The DNA was then sequenced.

2.7. Protocols for DNA sequencing.

2.7.1. Preparation of DNA.

Double stranded DNA is prepared with the alkaline lysis method. The MicroSpin Columns from Pharmacia is used to purify denatured DNA. The DNA from standard miniprep is dissolved in TE buffer. 5 µl of 2 M NaOH is added to 20 µl of DNA. The sample is incubated at room temperature for 10 minutes to denature the DNA. The column, prepared as indicated by Pharmacia, is placed in a new 1.5 µl tube and the sample is applied to the top center of the resin. The column is then spinned at 2000xg for 2 minutes. 20 µl of TE buffer are added to the column to eluate all DNA. The eluate is used directly for sequencing.

2.7.2. Sequence reaction.

Sequenase Version 2.0 T7 DNA polymerase form USB and T7 DNA polymerase from Pharmacia were used.

Sequencing was performed using 96 wells microtitre plates. The annealing reaction is started by putting 2µl of template DNA and 2µl of primer mix (6µl ddw, 1µl sequenase buff, 1µl oligo-0.5pmol/µl) in four wells for each reaction. After a brief spin, the plate is covered with Saranwrap, put at 55°C for 15 min., then let cool on bench for 10 min. Meanwhile a mixture (ddw 6.5, DTT 0.1M, 0.4, 35S dATP, labelling mix 0.4, Sequenase, 0.25) is prepared for the labeling reaction. The mixture is added, 2µl per well, the plate is spin and left 10 minutes at room temperature (less than 23°C). 2µl of appropriate termination mix is added to TCGA wells, spin and immediately float on 37°C water bath. Add 4µl of formamide dyes, heat at 80°C for 20 min and load on gel within 30 minutes.

3. Results.

3.1. Surface display on λ phage: gJp as a vector protein.

In this section the results obtained in order to engineer the J protein of λ phage to permit surface display on a lytic phage and, as a final step, to allow selection to happen through infection, are described.

Apart the ambitious project to realize the selection by a single infection event, another important reason for considering a new phage system for surface-display is to complement the filamentous phage system. In filamentous phage libraries, the epitope or the protein fused to the N-terminus of gIIIp has to pass through two membranes and the periplasmic space before being exported. Even if some cytoplasmic proteins have been fused on the coat protein of filamentous phage (Rebar & Pabo, 1994), the translocation into the host cell periplasm before extrusion could constitute a significant physical constrain for some fusion proteins. On the contrary a fusion library on a lytic phage would be more suitable for the display of cytoplasmic proteins, with the fusion phage available after cell lysis. Besides it is important to consider that lambda has been widely used as a cloning vector (see (Chauthaiwale *et al.*, 1992) for an excellent and complete review). Lambda offers a great advantage in cloning efficiency even when compared with the most efficient plasmid cloning system. A step in this direction is the ImmunoZAP™ 13 system, a λ vector system for the generation of combinatorial Fab libraries on the surface of filamentous phage (Hogrefe *et al.*, 1993).

During the course of this thesis, two papers were published in which gVp, the major protein of the tail tube, has been used as a vector for protein display (Dunn, 1995; Maruyama *et al.*, 1994), as I discussed in section 1.3.1.1.. However the main aim of this thesis remains the study of a phage/bacteria system in which it would be possible to realize the selection by infection procedure described in section 1.4.. For this reason another protein of lambda phage was chosen to engineer a phage-fusion system, i.e. gJp. This protein forms the tail fiber and is responsible for binding to the lambda phage receptor, LamB (see section 1.1.3.1.). In the presence of this well known phage receptor protein, it is obvious to work with the phage receptor for bacteria, gJp, the last being likely more appropriate to perform selection by infection.

Some constructs and preliminary experiments were performed in order to test if a functional J protein can be produced in *E. coli* when its gene is cloned in a plasmid vector (section 3.1.1.). The wild-type *J* gene was cloned in two *lac-tra* operon fusion plasmids, one with the gene for ampicillin resistance, the other with the gene for tetracycline resistance. A crude extract of bacteria with the cloned *J* gene was analyzed in a protein gel after IPTG induction: it was not possible to distinguish the gJp among the other proteins expressed by the bacteria. It was thus necessary to add a short peptide, which is recognized by a monoclonal antibody, at the C-terminus of J: the expression of the protein in the *E. coli* was detected by western-blotting.

The tolerance of bacteriophage λ for C-terminal fusions to the tail fiber protein, the *J* gene product, was investigated. A short peptide and

a ScFv antibody fragment were fused to the C-terminus of J (discussed in sections 3.1.4. and 3.1.3. respectively). The phage carrying the fusion of the epitope is infective, but the epitope is not exposed. On the contrary the fusion of the antibody fragment causes the loss of the phage infectivity. At the same time, lysogens with modified copies of J were complemented by wild-type *J* cloned in a plasmid under Lac promoter control. Due to problems of recombination, it was not possible to obtain clear information about the possibility to obtain chimaeric phage with wild-type and modified gJp.

At the same time other possible sites on J protein where surface display could be engineered were considered. The sequence of *J* gene is known, but no structural or biochemical information of its product are available. Besides *J* gene is 3399 bp long and the translated protein is 1133 a.a. long. No homology with other known protein has been found that could help to assign a role to this long sequence. The only interesting data is that the λ*h* mutations cluster in the right end of *J* gene (Fuerst & Bingham, 1978; Mount *et al.*, 1968) and alter the specificity of the attachment of pJ to LamB (Buchwald & Siminovich, 1969; Mount et al., 1968). The knowledge of those residues which are involved in the interaction with the receptor protein could be a good indication of possible cloning sites for peptide fusion. The sequencing of host range mutants λ*h* (Appleyard *et al.*, 1956) was thus planned, but such sequences were published during the course of this thesis (Werts *et al.*, 1994). On the basis of these last results, a tentative prediction of the residues involved in the interaction with LamB was attempted by

predicting the surface probability of both the wild-type J protein and the protein from extended host range mutants (section 3.1.5.).

3.1.1. Inducible synthesis of the tail fiber protein, gJp.

The gene for the tail fiber protein, i.e. gJp, was cloned in two *lac*-operon fusion plasmids, pBtKS which carries the gene for ampicillin resistance and pBR322 with tetracycline resistance. The two plasmids were called JLA and JLAtet respectively. The strategy of cloning is described in Figures 3.1. and 3.2..

In both plasmids the synthesis of gJ is induced by IPTG and the concentration of the inducer allows different levels of expression. The reason to have the *J* gene cloned in a plasmid under inducible control is to complement *J in vivo*, in a fashion similar to that carried out for phagemid display vectors. In these experiments wild-type *J* is provided in trans on a plasmid carrying a different selectable marker, i.e. tetracycline, with respect to the one present in the lysogens, that is ampicillin, as it will become clear after section 3.1.2. and 3.1.3..

After IPTG induction, the synthesis and production of J was tested by SDS-gel electrophoresis. Total lysates of small scale cultures of *E. coli* carrying the pJLA were prepared and charged on a SDS-polyacrylamide gel. The identification of the protein by Coomassie blue after IPTG induction was not possible due to the small amount of protein produced.

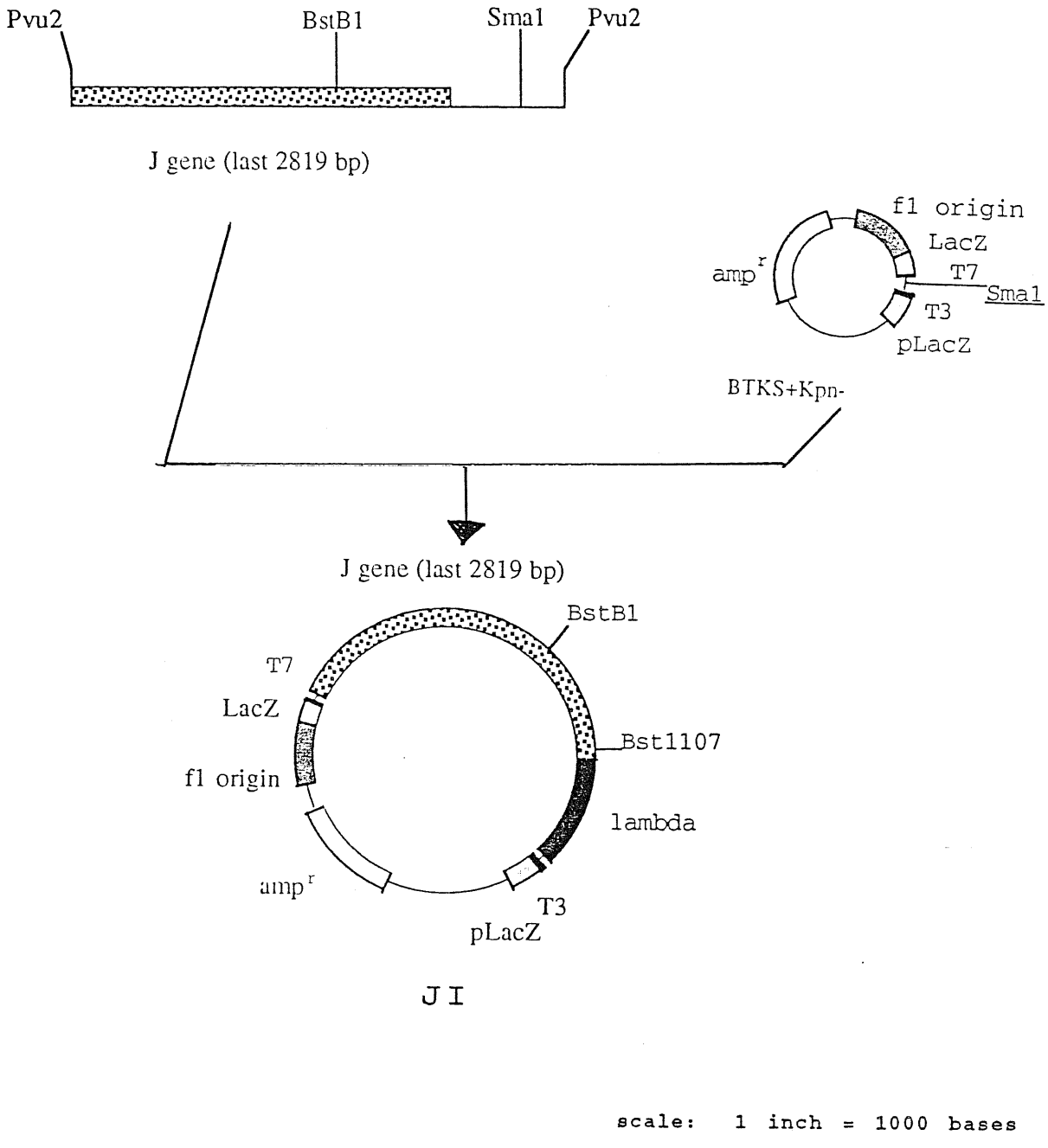


Figure 3.1.a. Construction of JI plasmid. Lambda DNA was cut with Pvu2 and the 3637 bp long fragment (16081-19718 in lambda sequence), which contains the last 2819 bp from *J* gene, was subcloned into the Sma1 site of BtKS^{+}(Kpn1^{-}). Before cloning the Sma1 fragment was dephosphorylated by calf alkaline phosphatase to avoid ri-ligating.

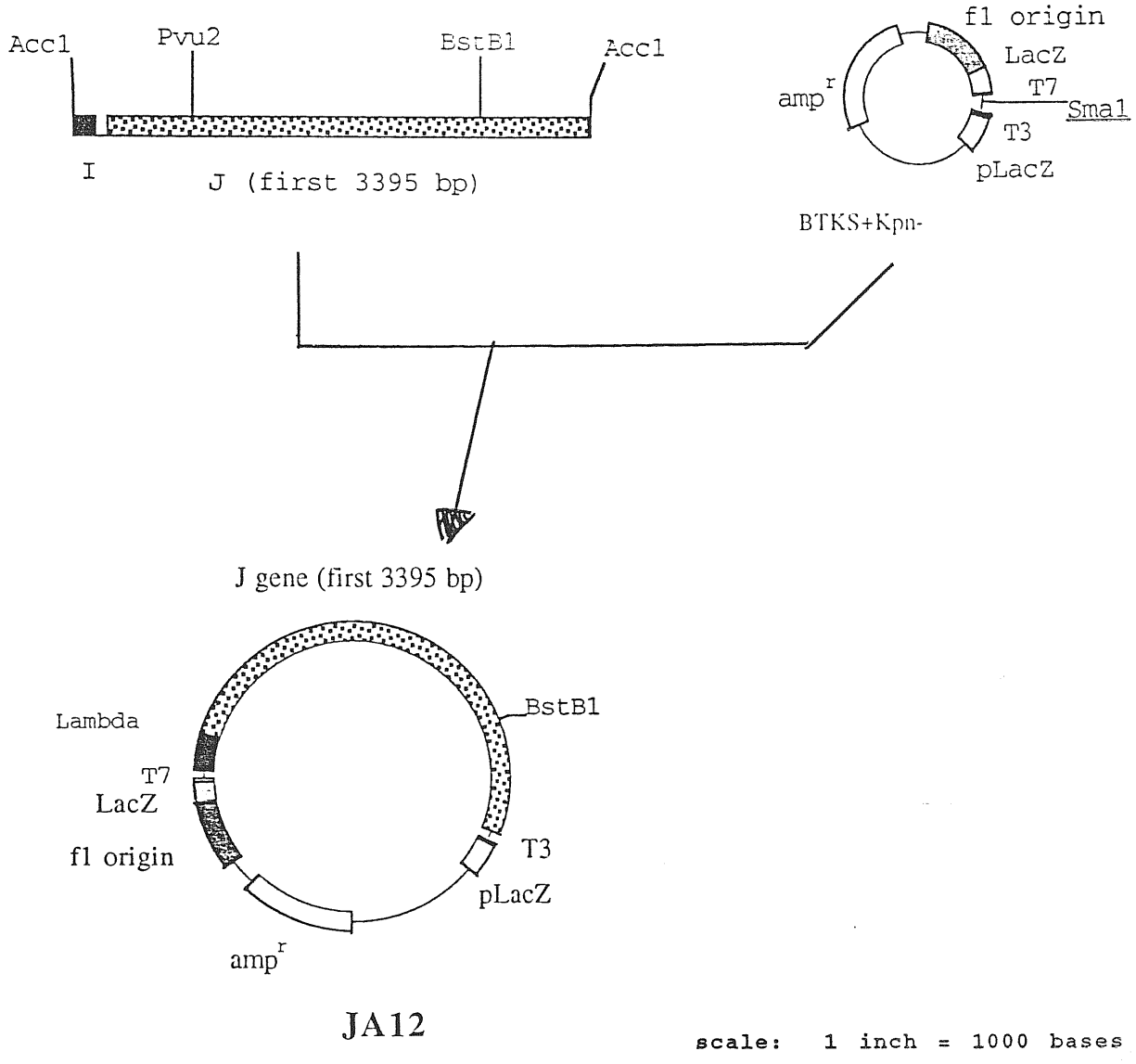


Figure 3.1.b. Construction of JA12 plasmid. Lambda DNA was cut with Acc1. The 3573 bp long fragment (15262-18835 in lambda sequence), containing the first 3395 bp from *J* gene, was rendered blunt by fill-in with Klenow DNA polymerase and was cloned in BtKS^{+}(Kpn^{-}) cut with Sma1. Before cloning the Sma1 fragment was dephosphorylated by calf alkaline phosphatase to avoid ri-ligating.

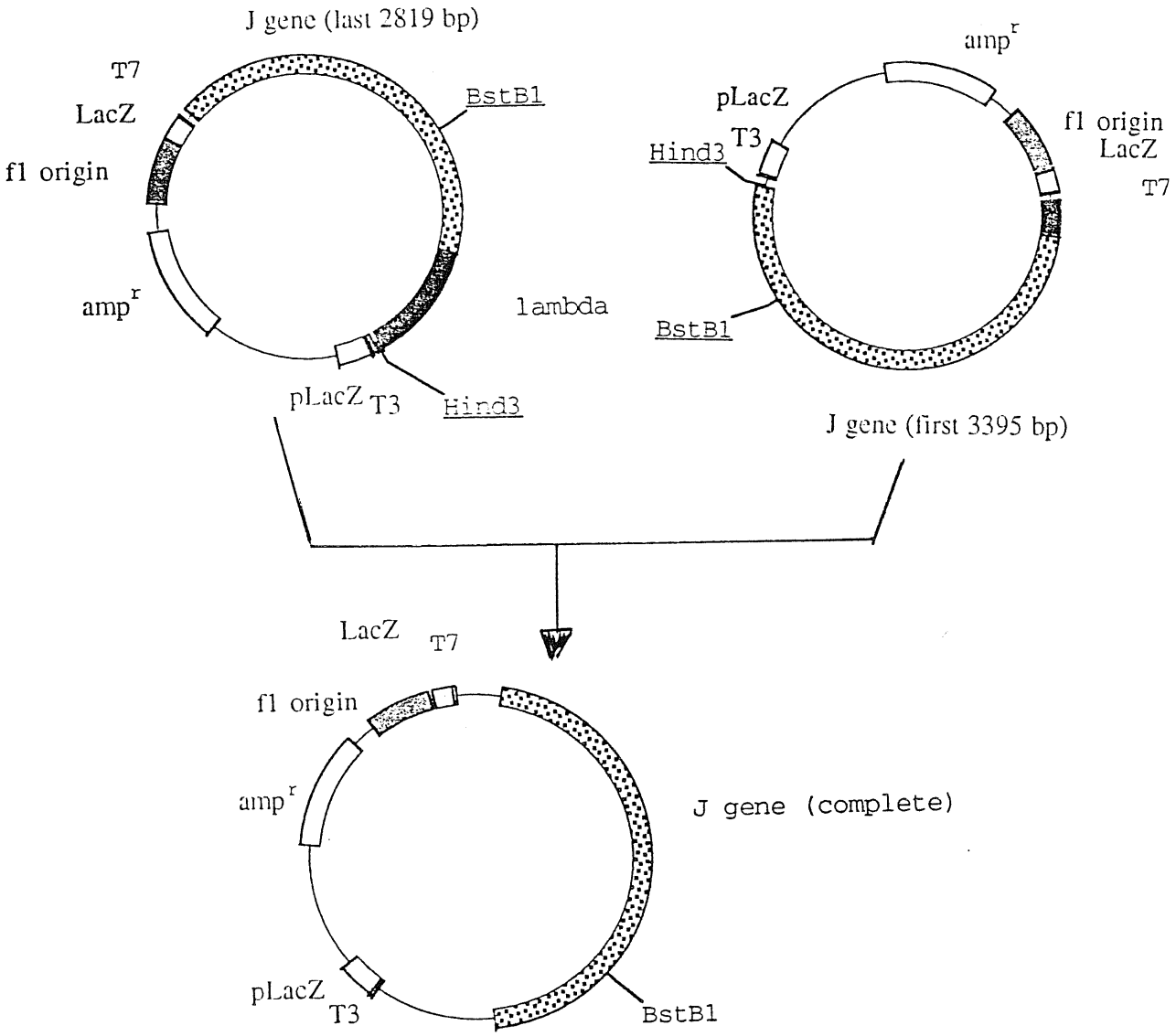


Figure 3.1.c. Construction of JL plasmid. The JL plasmid was constructed from JI and JA12. JI was restriction digested with BstB1 and Hind3: two fragments were obtained, with a length of 4.9 kbp and 1.7 kbp respectively. JA12 was restriction digested with BstB1 and Hind3, the 5.7 kbp fragment was purified and ligated with the smaller fragment of the previous digestion to give the final JL plasmid. JL is 6.6 kbp long and contains the whole gene for the J protein but in the inverse reading frame.

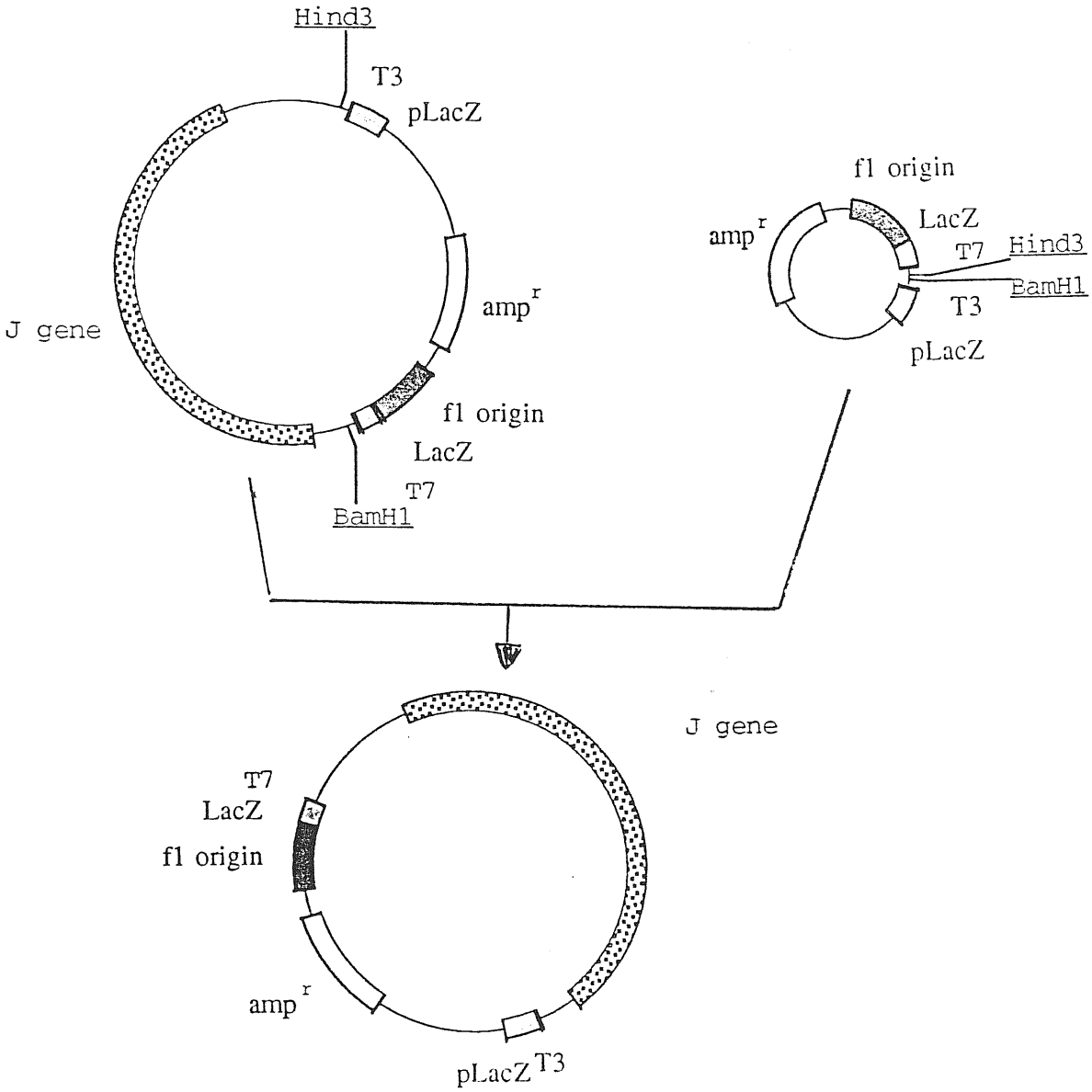
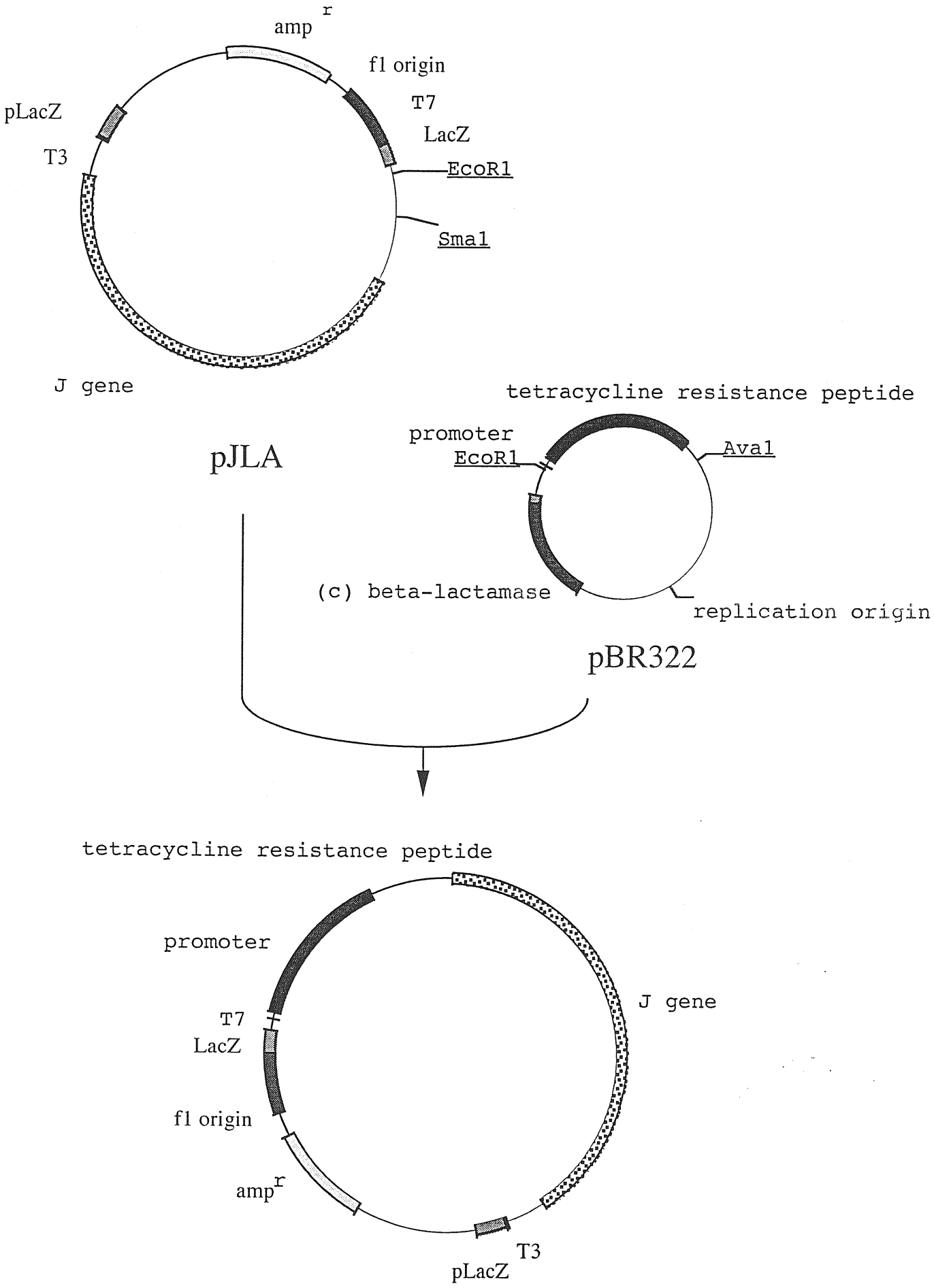


Figure 3.1.d. Construction of JLA plasmid. The JLA plasmid was constructed from JL and pBtSK+. JL was restriction digested with BamH1 and Hind3: two fragments have been obtained, with a length of 4.5 kbp and 2.9 kbp respectively. The bigger fragment was gel purified and ligated with the bigger fragment obtained by cutting pBtSK+ with BamH1 and Hind3. The final plasmid contains *J* gene with the right reading frame and can be selected for amp^r. The resulting plasmid was digested with different enzymes for restriction sites characteristic for *J* to test for the presence of the complete gene.

Figure 3.2. Construction of JLtet plasmid. The JLAtet plasmid was constructed from JLA and pBR322. JLA was restriction digested with Sma1 and EcoR1. pBR322 was restriction digested with Ava1 and EcoR1: two fragments were obtained, with a length of 2.9 kbp and 1.4 kbp. The smaller fragment was ligated with the JLA fragment which contains the *J* gene. The resulting plasmid contains the complete *J* gene with the right reading frame, and has a double resistance, amp^r and tet^r. During the ligation all restriction sites between Nae1 and EcoR5 included have been lost, maybe due to star activity of EcoR1.


amp r
f1 origin
T7
LacZ
EcoR1
Sma1
pLacZ
T3
J gene
pJLA
tetracycline resistance peptide
promoter
EcoR1
Ava1
(c) beta-lactamase
replication origin
pBR322
tetracycline resistance peptide
promoter
T7
LacZ
f1 origin
amp r
T3
pLacZ
J gene
pJLAtet

It was thus necessary to tag the gJ protein in a way that could allow the identification of the protein. The myc peptide which is recognized by the 9E10 antibody (Evan *et al.*, 1985) was cloned at the C-terminus of J protein by creating the JLAmyc plasmid (Figure 3.4.). Western blotting of the bacterial extracts with and without IPTG induction using the 9E10 antibody shows the presence of the protein (Figure 3.3.).

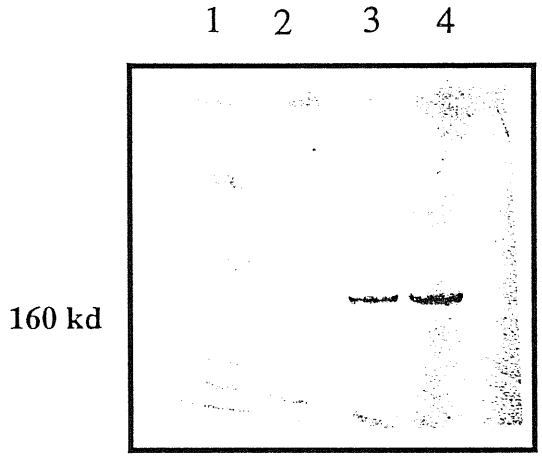


Figure 3.3. Western blot of bacterial extracts of cells containing pJLAmyc. The J protein was identified by the binding of the anti-p21^{ras} antibody (Y13-259) to myc-tag peptide fused with the C-terminus of J protein. Lines 1 and 2, without IPTG, lines 3 and 4, with IPTG.

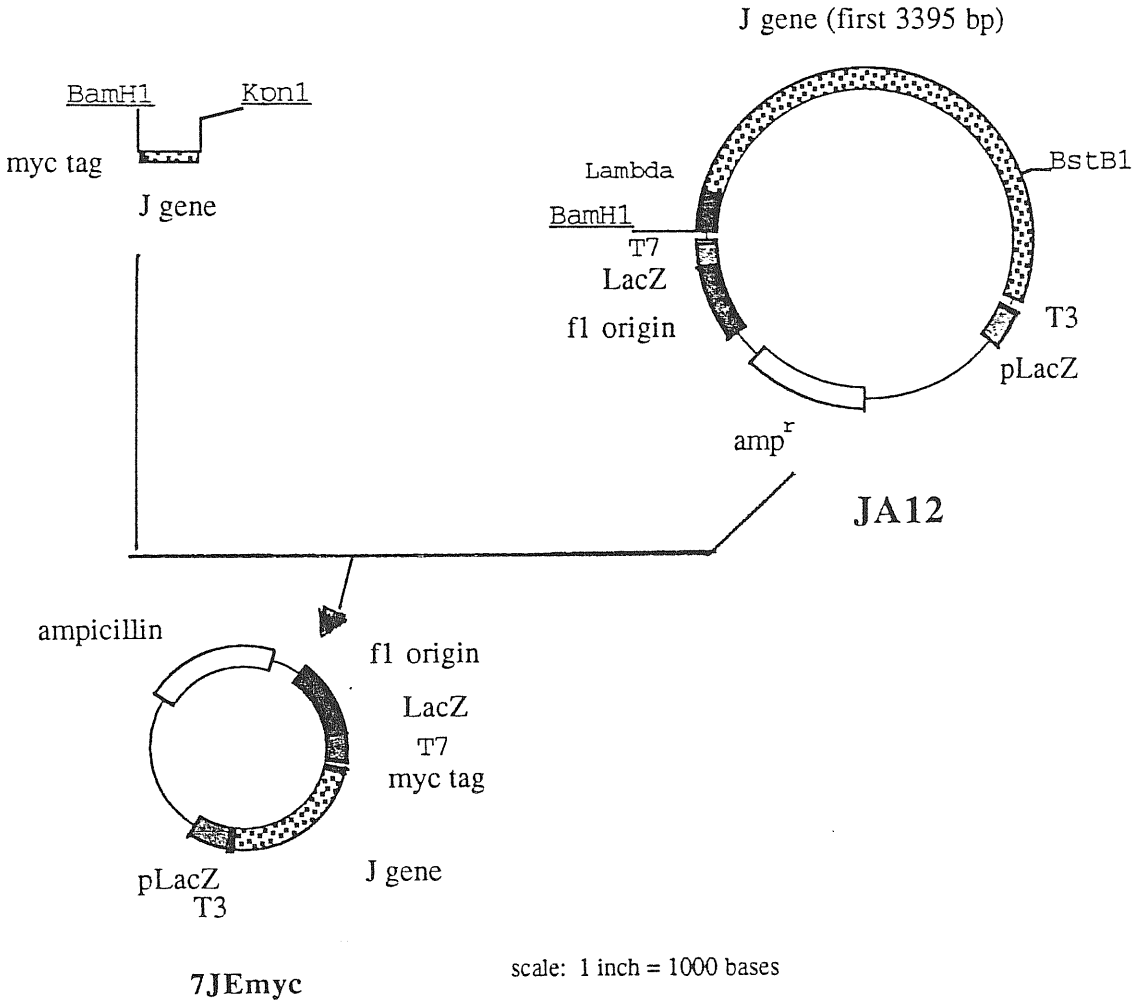


Figure 3.4.a. Construction of 7JEmyc. A PCR was run on lambda DNA with oligonucleotides J5Kpn1 (TGG CCG TCA GGT ACC CGT ACT) and Jmyc (TAG CGG ATC CGA GCT CAT TAA TTC AGA TCC TCT TCT GAG ATG AGT TTT TGT TCG ACT AGT GAG ATG CCC AGC GCC TGT TTC). J5Kpn1 is a 5' primer complementary to lambda sequence from 18547 to 18567 including Kpn1 site. Jmyc recognizes lambda sequence from 18873 to 18882 and encopasses the sequence encoding for myc-tag. A PCR fragment of 409 bp was obtained. The PCR fragment was digested with Kpn1 and BamH1 and cloned in JA12 (described in Figure 3.1.b.).

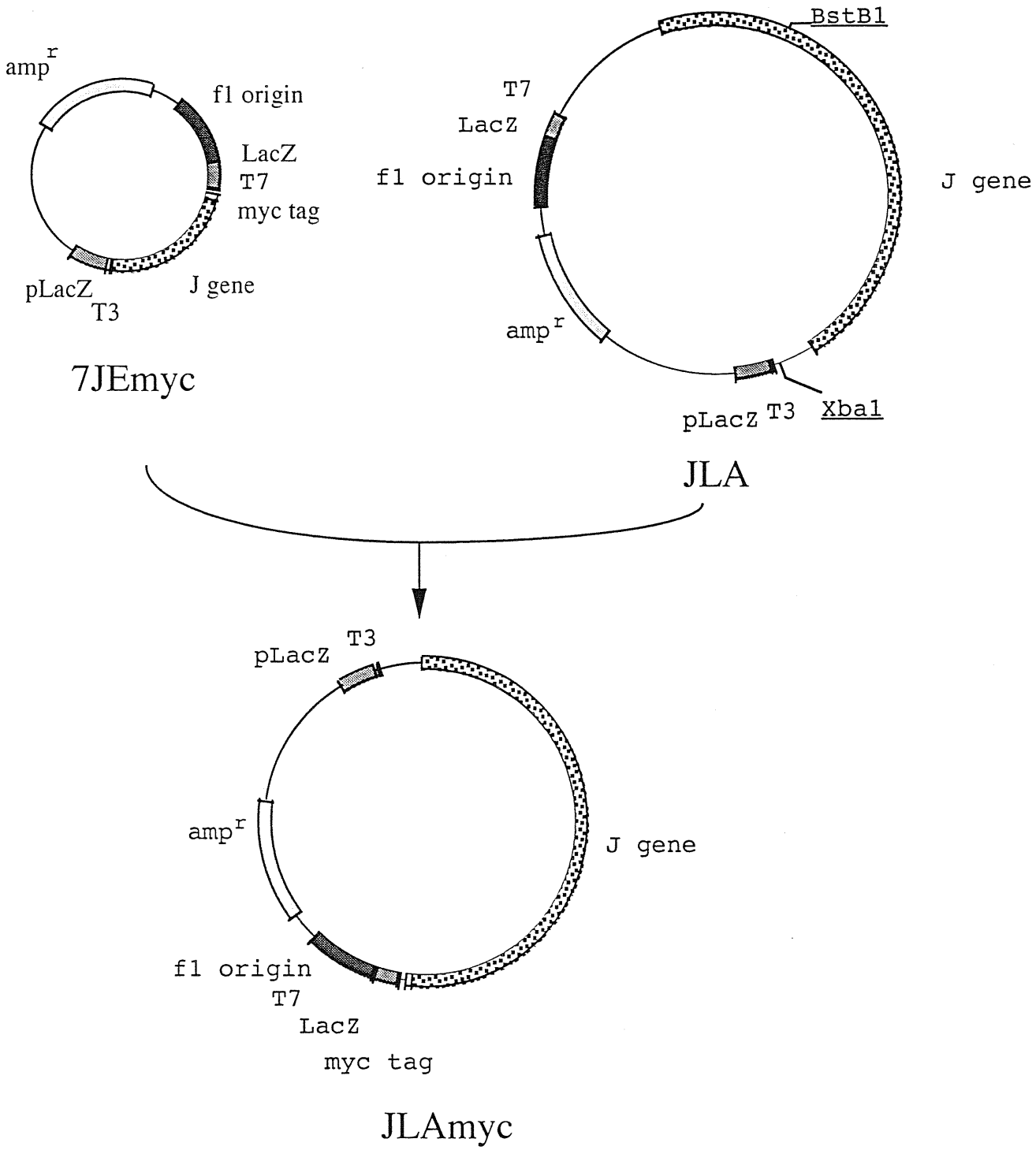


Figure 3.4.b. Construction of JLAmyc plasmid. The JLAmyc plasmid was obtained by digesting 7JEmyc and JLA with Xba1 and BstB1 and ligating the appropriate fragments.

3.1.2. An amp^r lambda phage.

The whole of pBlueScript (Stratagene) was cloned into λ downstream of the *J* gene, at the end of the region coding for the tail proteins in a region that has been shown to be not important for phage life. The strategy of cloning is shown in Figure 3.5: the lysogen, 6JK can be selected for ampicillin resistance. The resulting phage, λ6JK forms wild-type plaques when plated on LE392.

This is used as a normal control for experiments.

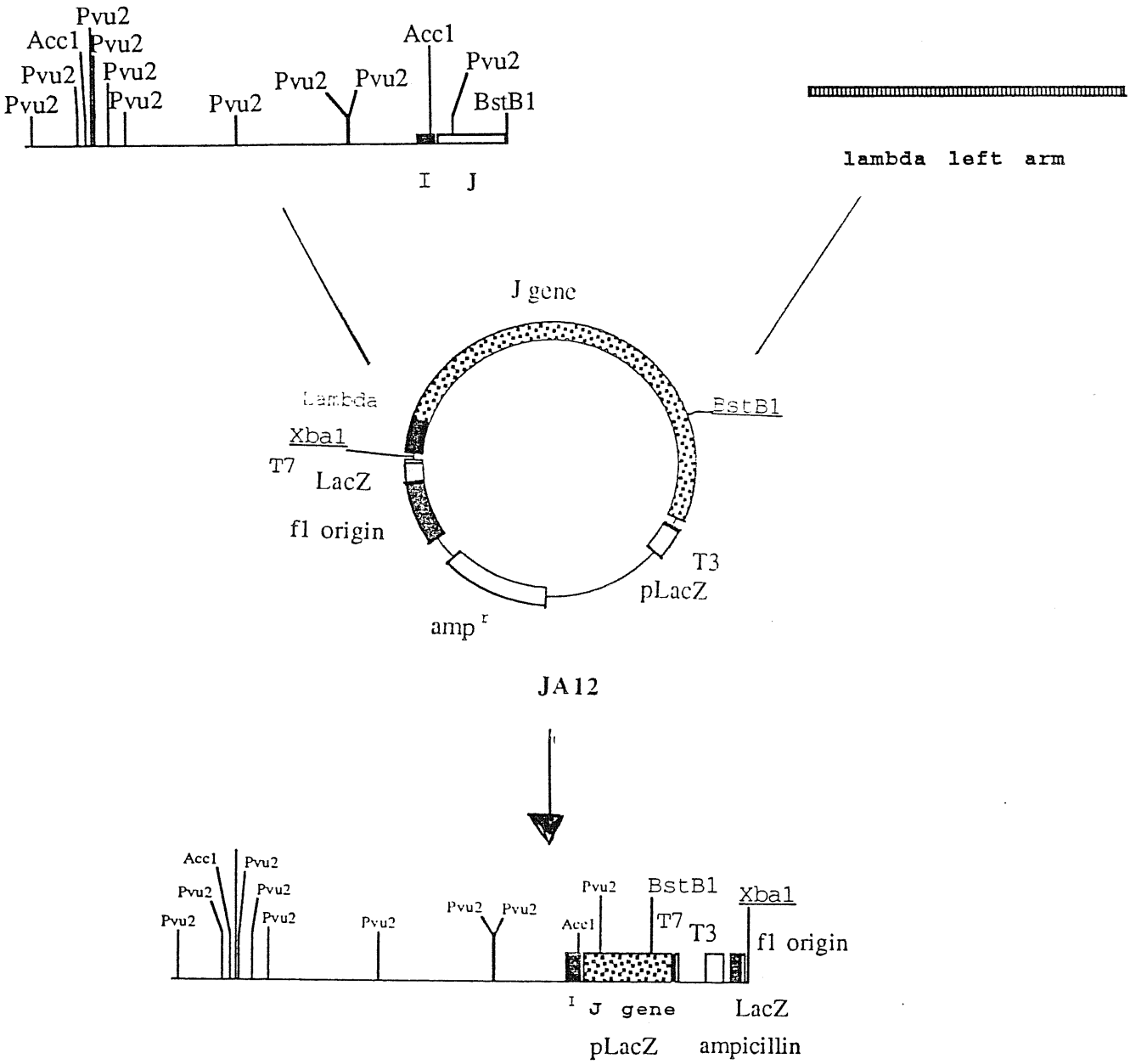


Figure 3.5. Construction of 6JK. JA12 (described in Figure 3.1.b.) was cut with BstB1 and Xba1 and inserted in the two lambda DNA arms, The left arm was obtained by cutting with BstB1 and Hind3 (which cuts several times in the right arm), the right arm was obtained by cutting with Xba1 and Mlu1 (which cuts several times in the left arm).

3.1.3. C-terminal antibody fusions with gpJ.

The C-terminus of J protein was chosen as the first site where to engineer a fusion lambda phage. The choice of cloning at this site was dictated by the observation (discussed in 1.1.3.2.) that the host range mutants are located in the last 10% of *J* gene (Mount et al., 1968). However no information is available about the sites in *J* gene that could affect tail assembly if modified. The only small contribution is that the insertion of a short oligonucleotide (for the purpose of eliminating the two Kpn1 restriction sites in *J* gene) apparently did not effect phage viability under normal conditions (Dunn & Blattner, 1987).

The heavy and light chain variable regions of Y13-259, a rat monoclonal antibody recognizing a mapped epitope (aa 60-76) within p21ras (Furth *et al.*, 1982; Sigal *et al.*, 1986) have been previously cloned (Werge *et al.*, 1992). The two regions were sequentially inserted into a pUC119 based ScFv expression vector, which provides a 5′ bacterial leader sequence, a linker, $(Gly_3Ser)_4$, which connects the two chains, and a 3′ tag recognized by the monoclonal antibody 9E10 (Evan et al., 1985). The Y13-259 was cloned at the C-terminus of the J protein, the strategy of cloning is shown in Figure 3.6.. The resulting lysogen, called JKY, is amp^r and when induced to enter the lytic cycle should produce a phage, λJKY, with the anti-$p21^{ras}$ ScFv antibody fragment and the myc-tag attached to the carboxy end of the J protein.

The lysogens are induced to enter the lytic cycle and the resulting phage are incubated with plating bacteria. After an overnight at 37°C, no plaques are present in the plates. Some different hypothesis can be proposed to explain the lack of plaques when λJKY is plated. One possibility is that the ScFv prevents a correct binding to the receptor protein, LamB, due for instance to steric hindrance. Or it may be that the presence of the antibody fragment fused to the the tail fiber protein interferes with the assembly process of the tail. In fact the assembly of functional tails is a process that is independent from the head assembly (as reviewed in section 1.1.2, tails and heads follow a different assembly pathway). At the same time the lytic cycle works even if no active phage are produced.

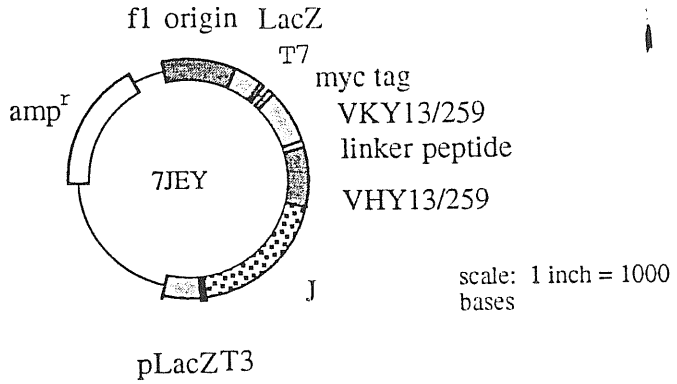


Figure 3.6.a. Construction of 7JEY. A PCR was run on lambda DNA with oligonucleotides J5Kpn1 (TCC AGA CTC CTG CAG TTG GAC CTG GAC CAC GCT GAT GCC CAG CGC CTG) and J3PstVH (TGG CCG TCA GGT ACC CGT ACT). J5Kpn1 is 5' primer complementary to lambda sequence from 18547 to 18567 including Kpn1 site. J3PstVH recognizes lambda sequence from 18877 to 18900 and allows the insertion of the ScFv.

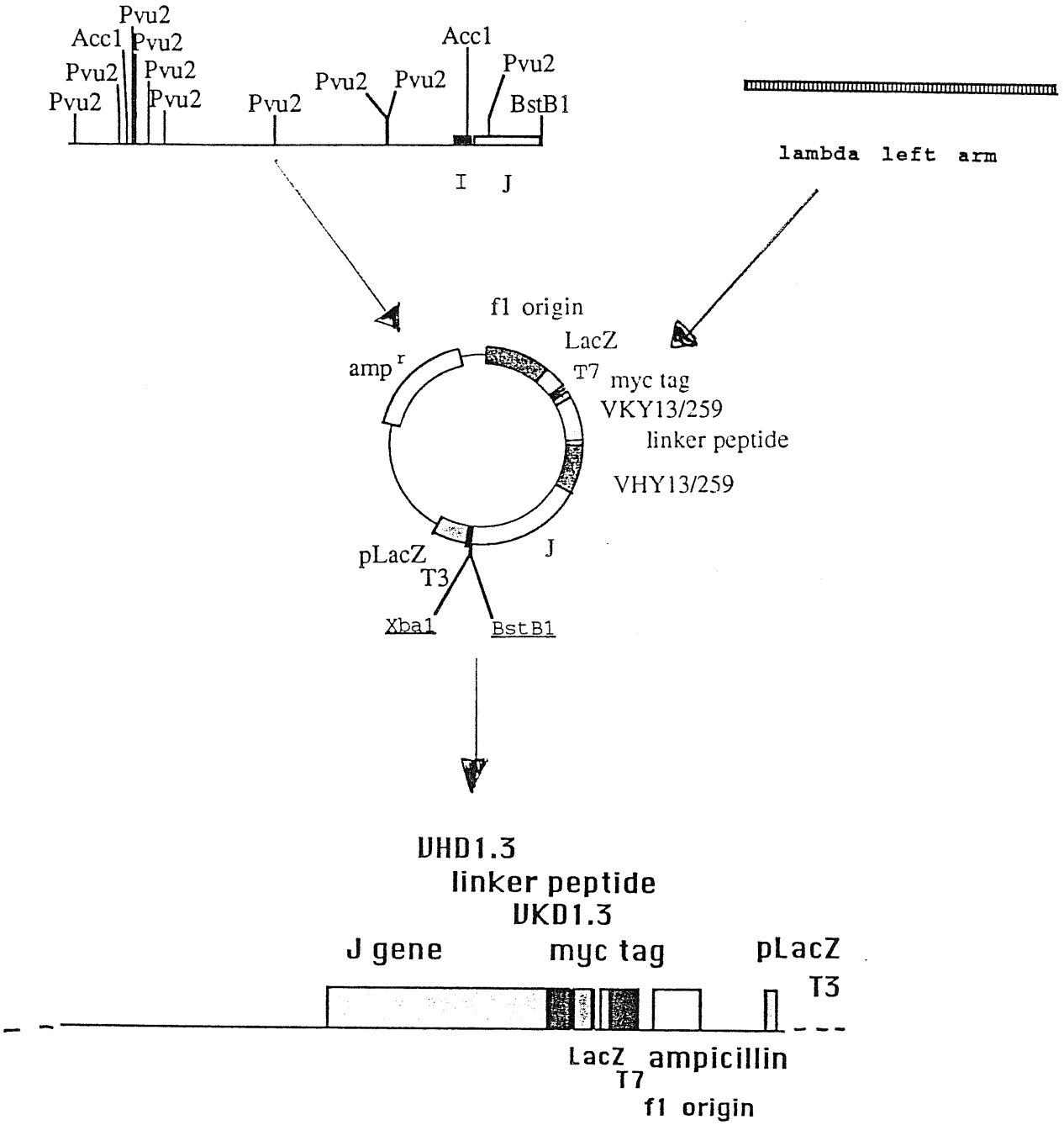


Figure 3.6.b. Construction of JKY lysogen. 7JEY was cut with BstB1 and Xba1 and cloned in the two lambda DNA arms, The left arm was obtained by cutting with BstB1 and Hind3 (which cuts only in the right arm), the right arm was obtained by cutting with Xba1 and Mlu1 (which cuts only in the left arm).

3.1.3.1. Infection with a modified phage receptor. One of the reason for the lack of infection when a ScFv fragment is fused to the J protein may be that the antibody fragment does not allow a correct binding to the phage receptor because it masks the amino acids involved in the phage receptor recognition. For this reason it was interesting to infect bacteria bearing a phage receptor in which the epitope recognized by the antibody fused to the gJp was cloned. Infection could thus happen to occur through a direct interaction of the antibody binding to the corresponding epitope.

λJKY phage was incubated with a bacterial strain containing vector pJD4 (Hofnung, 1991; Werge et al., 1992). This plasmid codes for a modified LamB protein, containing the p21ras epitope cloned in the BamH1 site. We thus have a phage carrying an antibody on its adhesin and the corresponding antigen cloned in the receptor for the same phage. But even in this case the λJKY phage was not able to infect the bacterial cell.

3.1.3.2. Complementation experiments. In order to see if it was possible to produce chimaeric phage containing a modified and wild type J protein, complementation experiments were carried out. It is not clear how many copies of J protein form the tail fiber: probably three copies for complementarity to the trimeric nature of the phage receptor. Phage infectivity could be preserved by constructing hybrid phage particles where most copies are wild-type while only one or a few are hybrid pJ carrying the N-terminal extension. This is achieved by phenotypic mixing of the hybrid pJ synthesized by a gene carried on the

lambda chromosome and the wild-type gene J product synthesized by a plasmid. I expect to obtain a population of infective phage with some of them containing at least one copy of the modified protein and the other copies of wild-type protein. These experiments are analogous to those carried out with filamentous phage (Bass *et al.*, 1990; Lowman *et al.*, 1991; Lowman & Wells, 1993): the coat protein fusion is expressed from a phagemid and a helper phage supplies a large excess of the wild-type coat protein. Therefore the phage are functional because the recombinant protein makes up only a small amount of the total coat protein.

The lysogen, JKY, was transfected together with pJLAtet (described in Figure 3.2.) DNA by electroporation and selected for double resistance, amp^r and tet^r. As a result, the double resistant cells should contain two copies of *J* gene: one copy, coding for the modified J protein, is on the λ chromosome, the other, coding for the wild-type protein, on the plasmid under Lac promoter control. Lysogens were induced to enter the lytic cycle and the resulting phage were purified.

If complementation works, these phage will have a mixed phenotype: some will have all copies of J wild-type, some will carry only the modified J, others will be a chimaera of both the modified and the wild-type protein. The latter phage will thus be able to infect as the wild-type phage.

All phage, even those with the wild-type phenotype, should maintain a genotype coding for the modified protein. Since the protein

produced by the phage is the modified one, these phage lose their ability to infect after the first infection. In other words, if the phage carrying the modified and the wild-type J are plated on LE392, their progeny will not infect, because the genotype will remain the modified one and the phage of the second cycle of infection will contain only modified J protein, which gives a phage that is not able to infect. On the contrary plaques will be visible only when they are plated with LE392 carrying the plasmid with the wild-type J (Figure 3.7.). The phage were thus incubated with both LE392 and LE392 with JLAtet (thereafter called LE392T). Unexpectedly wild-type plaques are present both on LE392 and LE392T. It is worth noting furthermore that the number of plaques is the same in the two plates.

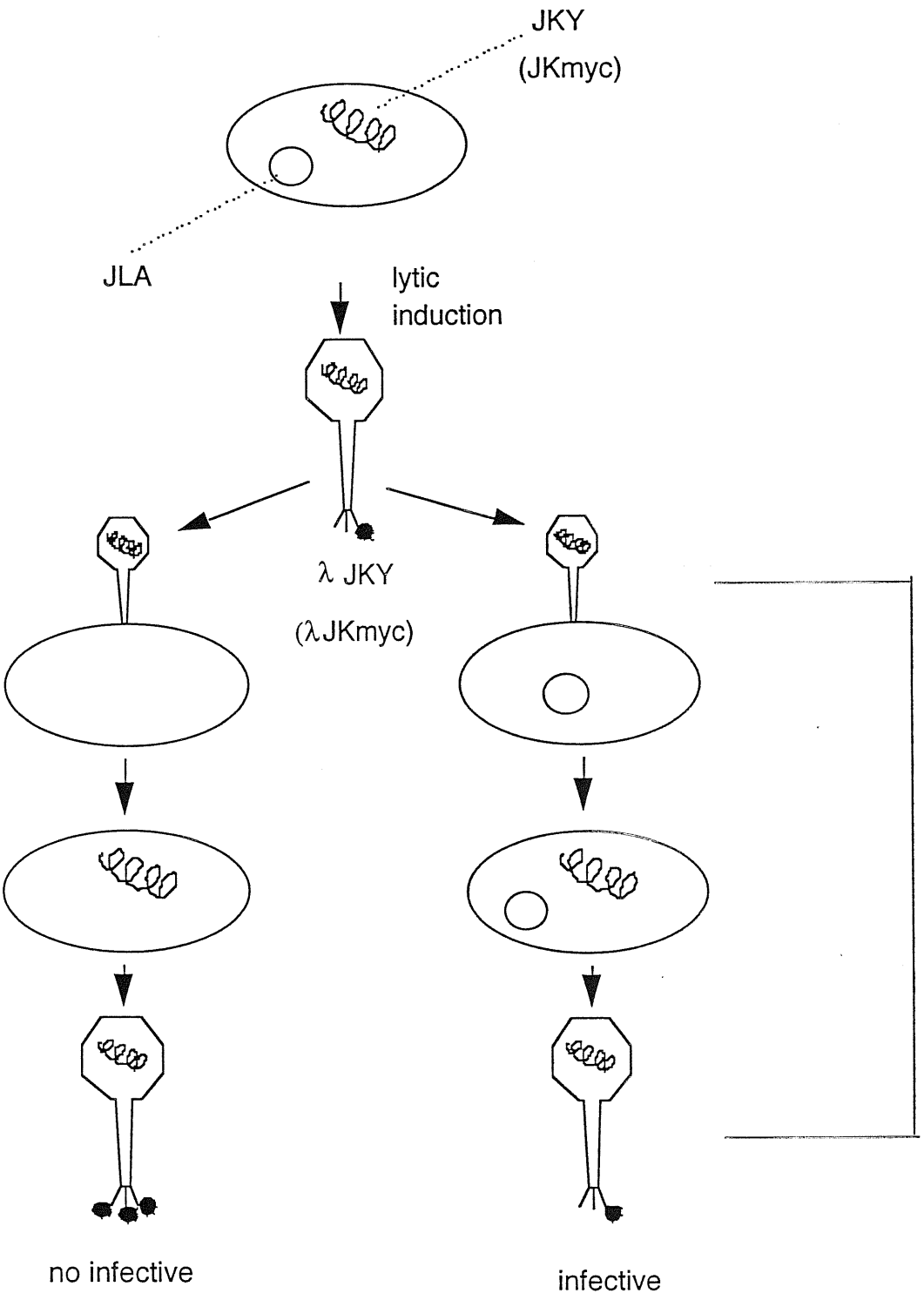


Figure 3.7. The expected phenotype for complementation experiments.

3.1.3.3. Identification of phage by PCR. In order to characterize the genotype of the phage obtained from the complementation experiments described in the previous section, PCR experiments where carried out.

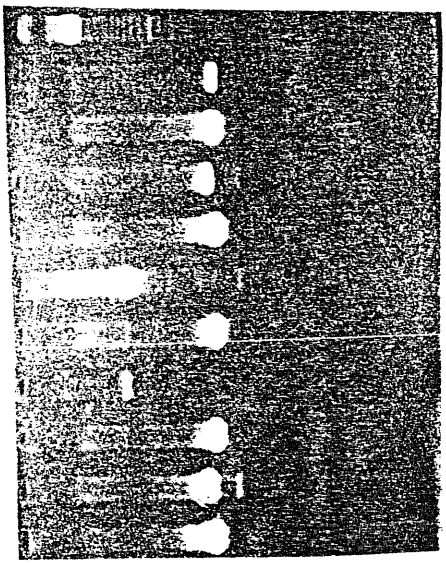
PCR was run using three different oligos: J5Kpn, a 3' oligonucleotide which includes the Kpn1 site of *J* gene at 18560; J31, a 5' oligonucleotide which recognizes the sequence 18943-18920 of the lambda genome, corresponding to the end of *J* gene (15505-18900); T7, a 5' oligonucleotide which contains the whole sequence of the T7 promoter. PCR was run by using oligonucleotides either mixed or in pairs (J5Knp1 with T7, J5Kpn1 with J31). Table 3.1. shows the expected length of the fragments (if any) that can be obtained.

At first PCR was run on the phage obtained after lytic induction, and then was carried out by picking the phage directly from the plaques they form after plating with LE392 and LE392T. The results of PCR experiments are reported in Figure 3.8.. Both phage containing the modified phenotype and wild-type phage are present. This mixed population can only be explained by a recombination event between the *J* gene on the plasmid and the *J* gene on the lambda chromosome.

	J5Kpn/T7	J5Kpn/J31
λ(6JK)	-	399
λ(JKY)	1166	-
λ(JKmyc)	436	-

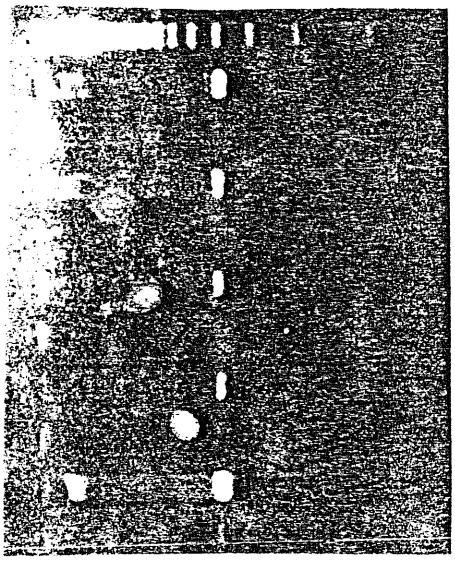
Table 3.1. Expected fragments from PCR experiments.

oligonucleotides used:
J5Kpn1/T7/J31
PCR of the phage



100 base pair ladder
6JK in LE392
6JK in LE392T
JKmyc in LE392
JKmyc in LE392T
JKY in LE392
JKY in LE392T
JKY in LE392
JKY in LE392T
JKmyc in LE392
JKmyc in LE392T

oligonucleotides used:
J5Kpn1/T7/J31
PCR of the phage



100 base pair ladder
JKmyc in LE392
JKmyc in LE392T
JKmyc in LE392
JKmyc in LE392T
JKY in LE392
JKY in LE392T
JKY in LE392
JKY in LE392T
6JK in LE392

Figure 3.8. Fragments obtained from PCR experiments.

From these results, a scheme of the genotype and phenotype for the phage at each stage can be hypothesized as shown in Figure 3.9.. Due to recombination, the phage that are released from the bacterial cell can be of two phenotypes (all modified J, e.g. not infective, one or more copies of wild-type J, e.g. infective) and two genotypes (modified and wild-type). It is worth noting that the phenotype of these phage is not related to the genotype because modified J protein can be assembled in a phage carrying wild type J DNA or vice versa. I expect only two of them to be infective, the ones bearing the wild-type J in one or more copies. When the miscellany of phage is incubated with LE392, only the ones having the wild-type J protein can infect. At this stage it is impossible to know if these active phage need all copies of J to be able to form wild-type plaques or not. But the ones having the λJKY genotype are lost after one cycle of infection. On the contrary if the miscellany is incubated with bacterial cells containing the wild-type genotype, the two population of phage will be maintained.

Recombination between plasmids and bacteriophage λ was studied previously (King & Richardson, 1986; Pogue-Geile *et al.*, 1980; Shen & Huang, 1986).

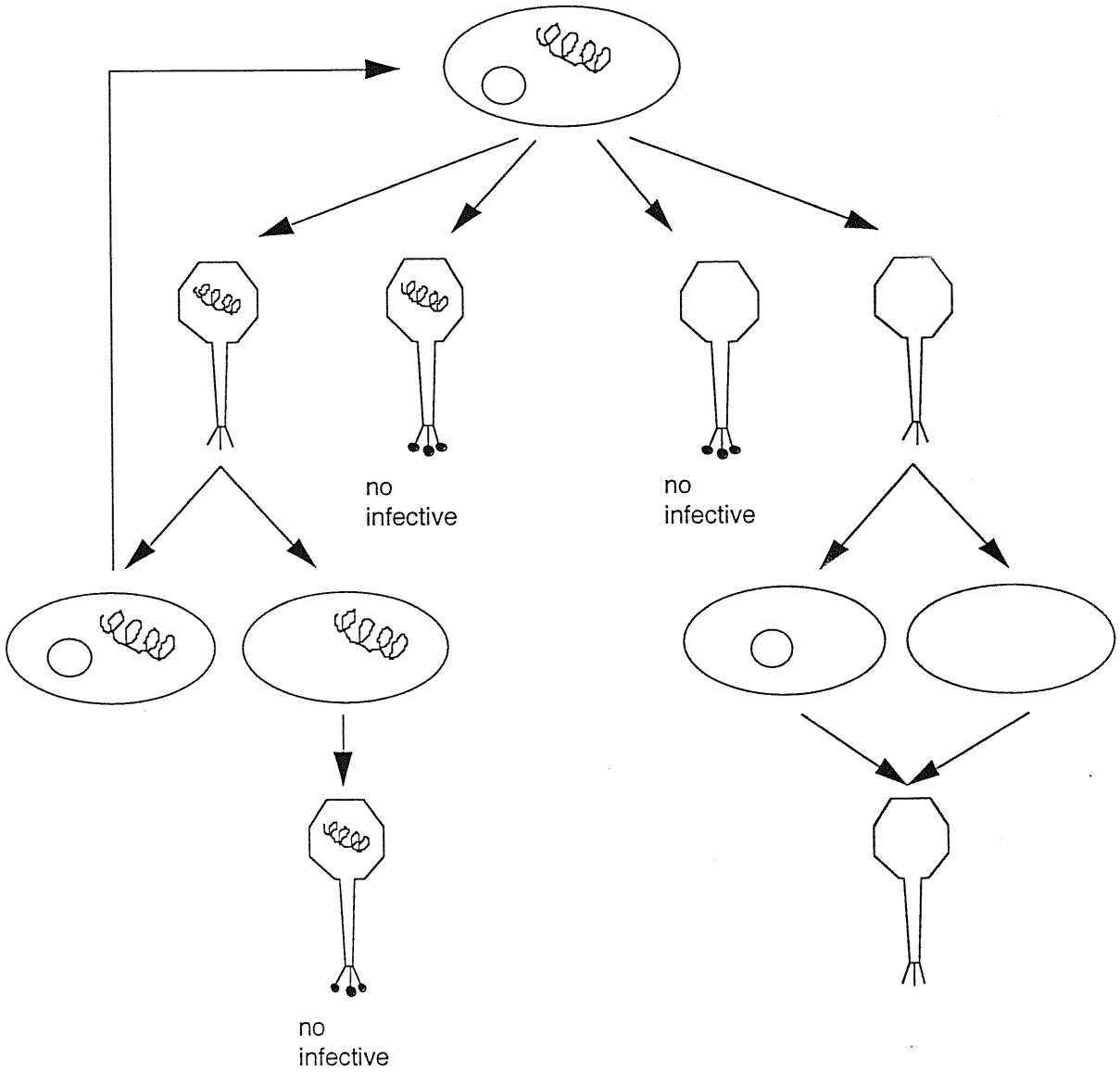


Figure 3.9. The expected phenotype after recombination. A detailed description is given in the text. The balls that have been drawn in the correspondence of the end of the tail represent the antibody or the myc-tag. The scheme can be applied to JKmyc as well, with the only difference that the phage bearing modified J are infective but give small plaques.

3.1.4. C-terminal peptide fusions with gJp.

At the same time as the C-terminal cloning of the ScFv fragment was performed, a short peptide was fused with the tail fiber protein. Myc-tag, whose amino acidic sequence is EQKLISEEDLN, is the peptide used in some ScFv cloning vectors: it is recognized by the monoclonal antibody 9E10 (Evan et al., 1985). The myc-tag peptide was attached to the carboxy end of the J protein with the strategy described in Figure 3.10.. The lysogen was called JKmyc and is amp^r.

The lysogens were induced to enter the lytic cycle, the phage isolated, λJKmyc, and then tested on LE392. Plaques smaller than those formed by wild-type phage can be observed. The titre is smaller than for wild-type phage and smaller burst size could be due to a reduced number of active phage particles produced.

The exposition of the epitope has been tested by panning in immuno tubes. A Nunc-immunotube with 9E10 antibody which recognizes the myc-tag are used. If the epitope is exposed it will bind to the 9E10 antibody and will be released only in the final washing step. The number of plaques obtained with λJKmyc is the same as with λ6JK (wild type J protein) after panning on 9E10 immunotube.

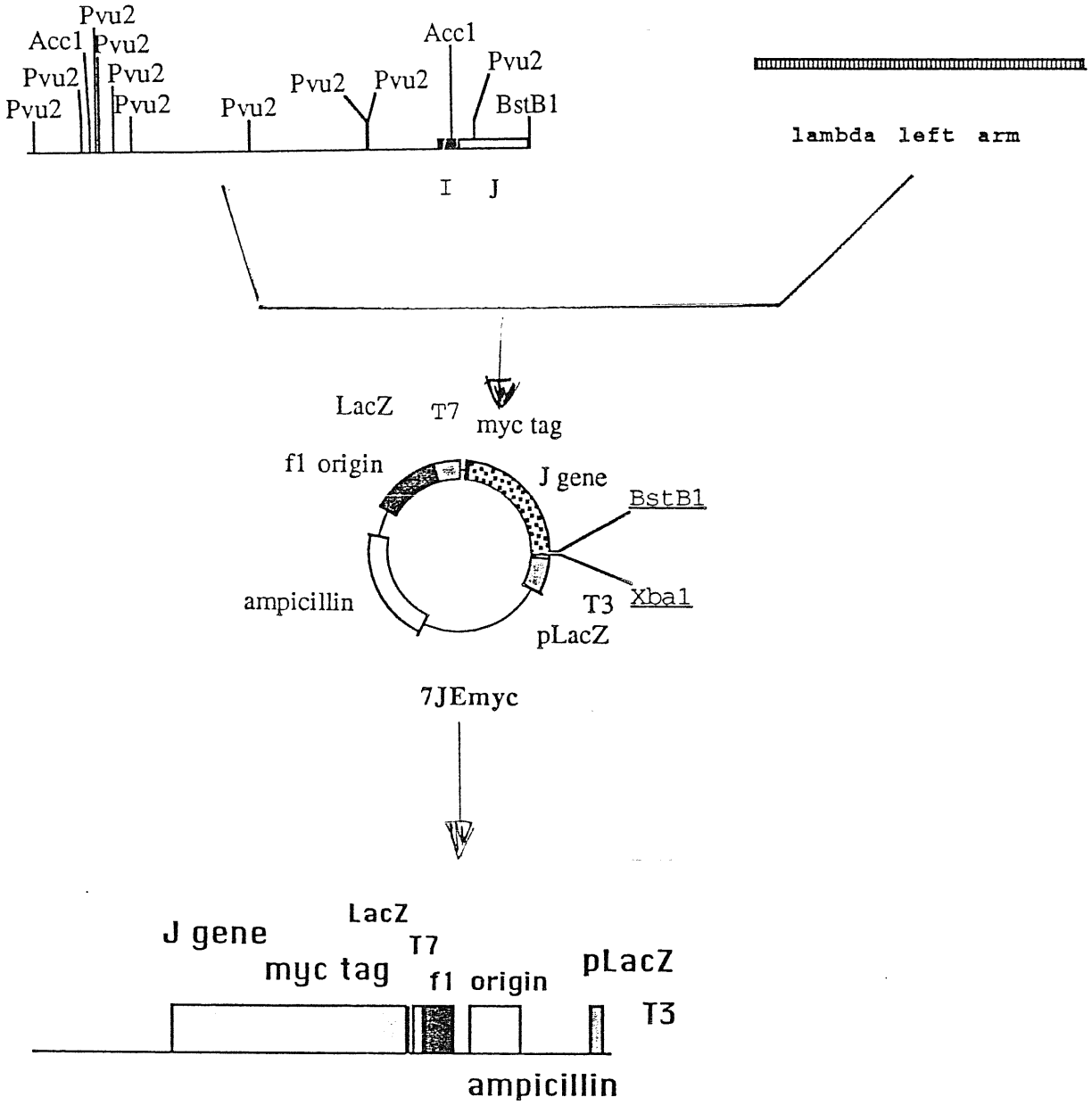


Figure 3.10. Construction of JKmyc. 7JEmyc was digested with BstB1 and Xba1. and cloned in the two lambda DNA arms, The left arm was obtained by cutting with BstB1 and Hind3 (which cuts only in the right arm), the right arm was obtained by cutting with Xba1 and Mlu1 (which cuts only in the left arm).

3.1.4.1. Complementation experiments. The JKmyc lysogen was transfected with pJLAtet DNA and then the lysis was induced. The plaques obtained by plating on incubating bacteria carrying the pJLAtet or without a plasmid-borne wild-type *J* gene are both wild-type and small sized. Complementation experiments with λJKmyc must take into account the different morphology of plaques produced after infection with wild type and myc-modified phage, discussed previously.

Wild-type and small plaques are present in both samples. This can be interpreted at the light of the results reported in the previous section, where recombination has been demonstrated to occur between λ and plasmid DNA. To confirm such a hypothesis, PCR was run with the same oligonucleotides as the ones used in the previously described experiments with λJKY (Table 3.1.). The results of PCR are in agreement with a recombination event (Figure 3.9.). The initial miscellany contains four types of phage, two of which are wild-type and two have myc-tag.

3.1.5. Prediction of gpJ structure.

The sequence of J protein (Figure 3.11.) was submitted for structure prediction using the protein structure prediction services available through World Wide Web, at the home page: http://www.embl-heidelberg.de/predictprotein/predictprotein.html.

These services allow to obtain either a secondary structure prediction or a surface accessibility profile. PHD, a profile fed neural network system, described by Rost (Rost & Sander, 1993), is used for the secondary structure prediction. It gives an expected accuracy of more than 72% for the three states, helix, strand and loop if, and only if, the alignment contains sufficient information. On the contrary, solvent accessibility is predicted by a neural network method rating at a correlation coefficient of 0.54 cross-validated on a set of 238 globular proteins (Rost & Sander, 1994) and it use no information from multiple alignments.

The submitted sequence was used to scan the sequence database SWISSPROT searching for similar sequences with the multiple sequence alignment (MaxHom) method described by (Sander & Schneider, 1991). Such data search has shown that there are no proteins homologous to gJ protein. As already explained, the accuracy of the prediction depends on how closely the protein of known structure resembles the set of known proteins used to derive the method. Due to the lack of available homologous proteins, the expected accuracy is about 6-10 percentage points less than the expected accuracy for

MGKGSSKGHTPREAKDNLKSTQLLSVIDAISEGPIEGPVDGLKSVLLNSTP
VLDTEGNTNISGVTVVFRAGEQEQTPPEGFESSGSETVLGTEVKYDTPITRTI
TSANIDRLRFTFGVQALVETTSKGDRNPSEVRLLVQIQRNGGWVTEKDITIK
GKTTSQYLASVVMGNLPPRPFNIRMRRMTPDSTTDQLQNKTLWSSYTEIID
VKQCYPNTALVGVQVDSEQFGSQQVSRNYHLRGRILQVPSNYNPQTRQYS
GIWDGTFKPAYSNNMAWCLWDMLTHPRYGMGKRLGAADVDKWALYVI
GQYCDQSVPDGFGGTEPRITCNAYLTTQRKAWDVLSDFCSAMRCMPVWN
GQTLTFVQDRPSDKTWTYNRSNVVMPDDGAPFRYSFSALKDRHNAVEVN
WIDPNNGWETATELVEDTQAIARYGRNVTKMDAFGCTSRGQAHRAGLWLI
KTELLETQTVDFSVGAEGLRHVPGDVIEICDDDYAGISTGGRVLAVNSQTRT
LTLDREITLPSSGTALISLVDGSGNPVSVEVQSVTDGVKVKVSRVPDGVAEY
SVWELKLPTLRQRLFRCVSIRENDDGTYAITAVQHVPEKEAIVDNGAHFDG
EQSGTVNGVTPPAVQHLTAEVTADSGEYQVLARWDTPKVVKGVSFLLRLT
VTADDGSERLVSTARTTETTYRFTQLALGNYRLTVRAVNAWGQQGDPASV
SFRIAAPAAPSRIELTPGYFQITATPHLAVYDPTVQFEFWFSEKQIADIRQVET
STRYLGTALYWIAASINIKPGHDYYFYIRSVNTVGKSAFVEAVGRASDDAE
GYLDFFKGKITESHLGKELLEKVELTEDNASRLEEFSKEWKDASDKWNAM
WAVKIEQTKDGKHYVAGIGLSMEDTEEGKLSQFLVAANRIAFIDPANGNET
PMFVAQGNQIFMNDVFLKRLTAPTITSGGNPPAFSLTPDGKLTAKNADISG
SVNANSGTLSNVTIAENCTINGTLRAEKIVGDIVKAASAAFPRQRESSVDWP
SGTRTVTVTDDHPFDRQIVVLPLTFRGSKRTVSGRTTYSMCYLKVLMNGA
VIYDGAANEAVQVFSRIVDMPAGRGNVILTFTLTSTRHSADIPPYTFASDVQ
VMIKKQALGISVV

Figure 3.11. The amino acidic sequence of gJp.

Figure 3,12. Surface accessibility profiles.

homologous proteins. The application of these algorithms for the prediction of the secondary structure of gJp is not suggested.

On the other side solvent accessibility prediction uses no information from multiple alignments. Figure 3.12. shows the surface accessibility profile for the gJ protein.

The first answer we can obtain from this profile is that the last 83 aa (from 1050 to the end) are very hydrophobic and the surface probability of this region is very low. We can speculate that an explanation of the failure to produce an active phage when a peptide is inserted at the C-terminus is due to the hydrophobic nature of this region.

From the accessibility profile it could be possible to obtain an indication of which region is more exposed and thus could be suitable for the expression of a peptide. Three regions seems to be more exposed: from 1006 to 1012, from 1028 to 1031, from 1043 to 1049. These regions are located in a very hydrophobic region and are the only ones predicted to be the exposed on the surface.

The same method has been applied to the mutants described by Werts et al (Werts et al., 1994).

The Val→Ala mutation at 1077 has been analyzed. The profile of this region, from 1060 to 1100 seems to be very hydrophobic. There are only a few amino acids. which are hydrophilic, from 1071 to 1073 (GAA), but the mutation seems to extend this region (GAANEAA).

3.1.6. Conclusions.

In this chapter I have discussed the possibility of using the J protein from lambda bacteriophage as a cloning site for phage display of epitopes or proteins. To our knowledge this is the first attempt to engineer the tail tube protein as a fusion vector for creating a phage library.

The tolerance of bacteriophage λ for C-terminal additions to the tail fiber protein, the *J* gene product, was investigated. The cloning of a short peptide of 17 amino acids (myc-tag) is compatible with phage assembly and production but some kind of interference takes place during phage assembly or phage absorption to its host, since small plaques form after plating on *E. coli* (discussed in 3.1.4.). These experiments have shown that a ScFv antibody fragment expressed on gpJ yealds a non-infective phage (discussed in 3.1.3.). The C-terminus, to which peptide was fused, is a very hydrophobic region and this may be one reason for the failure of the epitope to be exposed.

It is important to underline that PCR experiments and morphology of plaques are not sufficient to clarify if the modified protein is present in the infective phage. In order to conclude this study of complementation, it would be necessary to do a protein gel analysis (SDS/polyacrylamide gel) of purified bacteriophage lambda (CsCl banding) with or without incorporation of J fused to myc-tag and Y13-259. It would be necessary to perform a western blotting with 9E10 of the same preparation. Even in this case it could be necessary to do

electron microscopy of λJKmyc and λJKY grown in LE492 or in LE392 with the wild-type J.

It is worth noting that the insertion of foreign proteins on the lambda surface has been shown to be possible on gpV groups (Dunn, 1995; Maruyama et al., 1994), but in both papers the modified protein was not present in many copies. It has been observed that more than two molecules on a virion particle may have an inhibitory effect on the coat assembly and explain the poor growth of fusion phages in hosts with strong suppressor activity (Maruyama et al., 1994).

The knowledge of the residues responsible for recognition of the phage-receptor protein, LamB, could also help in our "guessing". A couple of papers from Hofnung's group (Charbit *et al.*, 1994; Werts et al., 1994) and the resolution of the x-ray structure of the LamB protein (Schirmer *et al.*, 1995) became available during the course of this thesis. These new results show the gJp/LamB interaction to be a very complex one, and the residues directly involved in the interaction between the two proteins have not yet been identified. It seems that mutations on gJ proteins may modify the shape of the tail fiber in a way that the phage can not interact with a tight binder which is localized in the inner part of the LamB protein (Werts et al., 1994). This binder could be the 151 residue of LamB (Charbit et al., 1994).

At the light of these last results it seems to me very difficult to find a site in LamB where the insertion of an epitope could allow the binding antibody fused to the lambda gJp to start the reversible and

irreversible steps which are the first two ones of the infection events. The very last answer could anyway come from further biochemical and biophysical studies. Using the words that Hofnung spent when he was commenting on the resolved structure of LamB, the protein that he is investigating since the seventies: "as usual a nice achievement in science provides more questions as answers." (Hofnung, 1995).

3.2. The N pilus.

This section deals with the results obtained in the study of the transfer region of the IncN plasmid N3.

As discussed in section 1.2.3.2., the IncN plasmids are broad host range plasmids coding for N-pilus. In the IncN plasmid N3, the genes coding for the structural protein of the pilus are localized in a region of about 10 kb called Tra1 (Glocker & Rasched, 1990). The filamentous phage IKe (Khatoon *et al.*, 1972), which can specifically bind to the tip of the pilus and thus infect the host cell, can be used to identify the presence of the pilus. Resistance to the infection of this phage is diagnostic of the absence of the pilus, due to a mutation or deletion in some of its genes. The binding of the filamentous phage to the tip of the pilus is mediated by a minor-coat protein, the gene III product (gIIIp). The gIIIp of IKe phage has been compared to the same protein from Ff phage: the two proteins share common characters as seen in section 1.2.2.2. (Figure 1.14). Fragments of the gIIIp from the filamentous phage IKe were fused to the same protein from Ff phage (Marzari *et al.*, ; Sblattero, 1995). The Ff phage, which usually infects *E. coli* by binding to the F pili, acquires the ability to grow on cells bearing only N-pili (Marzari et al., ; Sblattero, 1995). This shows that when the appropriate protein domain is fused to the N-terminus of fd gIIIp, the host specificity can be changed.

At the light of these results, the possibility that the fusion of an epitope on the tip of the pilus could allow a direct interaction with the

corresponding antibody fused to the infective phage and thus realise the infection is credible.

The main difficulty in the realization of such a project is that the phage receptor protein on the pilus has not been identified. Although the F pilus has been studied for more than twenty years, as reviewed in (Frost *et al.*, 1994), the existence of an adhesin protein at the tip of the pilus is still an hypothesis. At the same time a new pilus system such as the N pilus would allow to switch from the F pilus to the N pilus by simply changing the bacterial receptor protein (Figure 1.14.).

In order to obtain information about the structure of the N pilus, the sequencing of the pilus gene cluster started. The sequence could reveal some interesting homology with other transfer systems and help in finding the pilin and eventually the adhesin protein.

Among the isolated IncN plasmids, whose transfer region has been investigated (section 1.2.3.2.), I have worked on the IncN N3, which has been first isolated in 1964 (Watanabe *et al.*, 1964). The boundaries of the IncN plasmid N3 transfer region have been defined by λ::Tn5 insertion mutagenesis and successively screened for sensitivity to phage IKe infection (Glocker & Rasched, 1990).

At first restriction sites that could allow to isolate the transfer region were tested. Unluckily no restriction sites were available for our scope (paragraph 3.2.1.). Tn*phoA*, a derivative of Tn5 (Manoil & Beckwith, 1985), has been used to construct translational fusions between IncN N3 genes and the transposon in order to obtain new

restriction sites for sub-cloning the transfer region (paragraph 3.2.2.) and investigate further the subcellular localization of these gene products. After having obtained insertions of Tn*phoA* in the transfer region, sequencing started from one end of the inserted transposons (paragraph 3.2.4).

While this work was underway the complete sequence of a cluster of pKM101 conjugal transfer genes was published (Pohlman *et al.*, 1994). The 12.4 kb region contains all the genes required for phage sensitivity. The sequence of IncN N3 transfer region obtained up to that moment was compared with the published sequence to see the level of homology between the N3 transfer region and that of pKM101. Interesting results were obtained that show the two plasmids to have a very high homology. Differences in nucleotide sequence are point, missence mutations. Some of these mutations causes the lose of a restriction site. At the light of these results the complete sequencing of the complete transfer region for the N pilus from a different plasmid from the already published sequence was considered to be worthwhile.

3.2.1. Genetic map of the transfer region of the IncN plasmid N3.

The transfer region from IncN N3 harbouring the genes for pilus synthesis and assembly was isolated by transposon mutagenesis (Glocker & Rasched, 1990). A large region of about 20 kbp, containing the defined transfer region, was cloned in pJE279 (Helsberg *et al.*, 1985) and pLK15 (Bross *et al.*, 1988), plasmids pEG21 and pBG21 were obtained respectively (Glocker & Rasched, 1990). Cells carrying these

new plasmids are sensitive to phage IKe and produce pili, but are conjugation-deficient.

pEG21 was obtained from Prof. Rasched and used in this study of the N pilus.

At first a detailed restriction map for the pEG21 plasmid was obtained and reported in Figure 3.13.. It is clear that few restriction sites are available that could be used to isolate the transfer region. As the cloning vector is unknown and details of its construction have proved difficult to obtain, in order to facilitate the mapping a new construct was obtained.

The Pst1-BamH1 region, about 18 kbp long, containing all the Tra1 region, was cloned in pBtKS$^+$ cut with the same two enzymes (Figure 3.14.). The new construct, pBGB, was selected for amp^r. *E. coli* cells harbouring pBGB were infected with IKe: wild type plaques were obtained showing that all the functions for pilus production are preserved. A restriction map of the new plasmid is shown in Figure 3.15.. The region responsible for conferring sensitivity to IKe phage is highlighted.

A comparison with the restriction sites present in the transfer region of pKM101, both from genetic studies (Winans & Walker, 1985) and from the recently available nucleotide sequence (Pohlman et al., 1994), shows that most sites are common to the two plasmids (Figure 3.16.). The restriction maps of plasmids N3 and pKM101 display a striking similarity to the enzyme Hpa1. However two Bgl1 sites are

missing in the transfer region of plasmid pKM101. Two Kpn1 are present in pKM101, while only one is present in N3. This suggests some similar organisation, but perhaps different detailed sequence.

Another construct was obtained by cloning the Kpn1-BamH1 region in pBtKS. Bacteria bearing this plasmid are not sensible to IKe. This means that a region necessary for pilus synthesis is missing. Such an observation is in agreement with the data from pKM101 which show the Kpn1 site to be inside the transfer region.

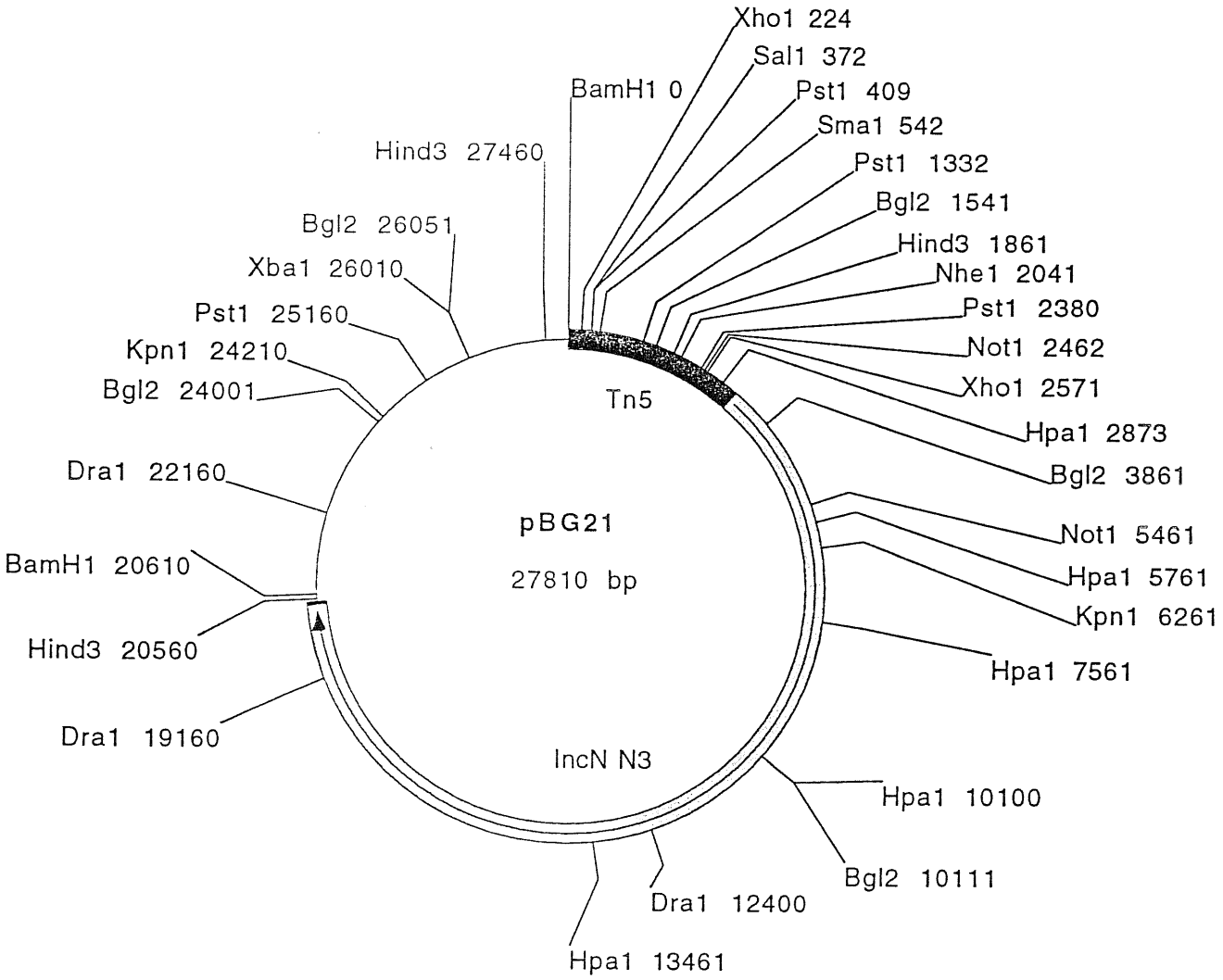


Figure 3.13. Restriction map for pEG21.

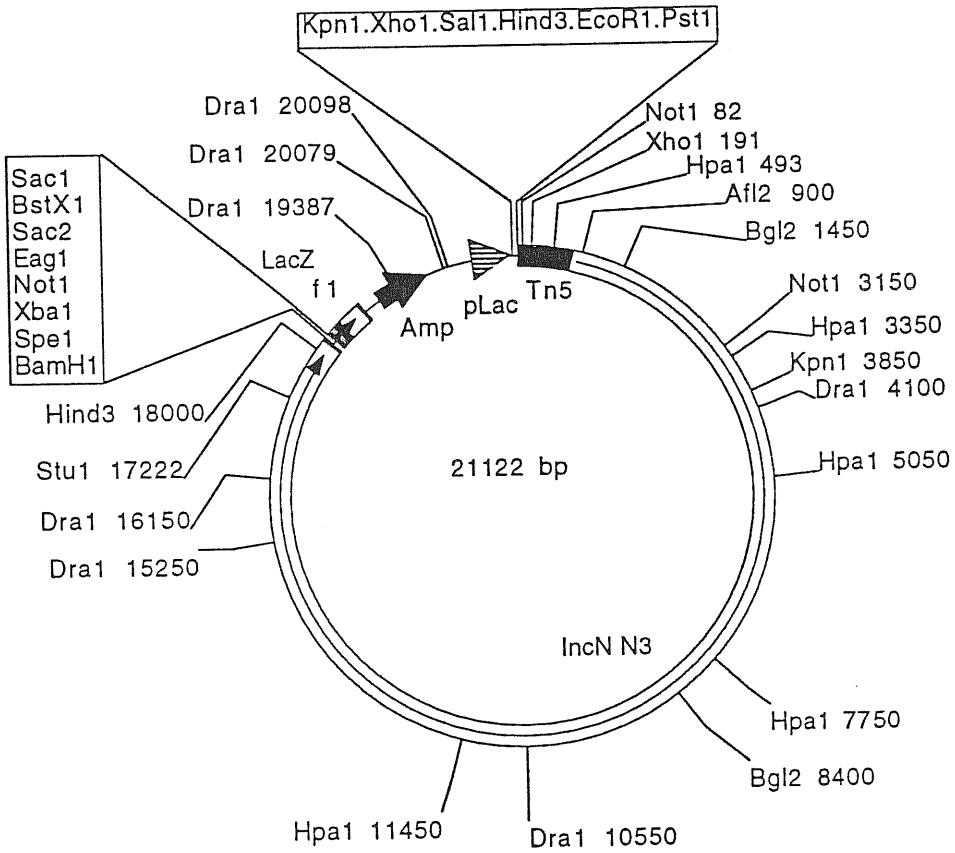


Figure 3.14. Restriction map for pBGB.

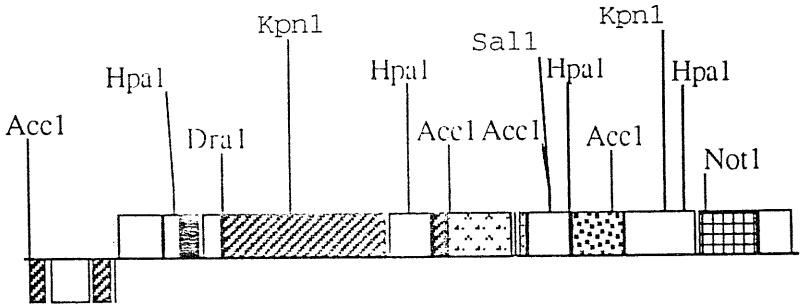


Figure 3.15. Restriction map of the IncN transfer region from pKM101.

From the above data, it is evident that it is not possible to define and isolate the transfer region by using the restriction sites present in the plasmid. Restriction sites that could allow to clone a shorter transfer region are not available. To overcome this difficulty, restriction sites were introduced artificially into this region of the plasmid by transposon insertion as it will be described in the next section.

3.2.2. Isolation and mapping of Tn5 insertions.

The minimal region of pBGB DNA sufficient for pilus synthesis was determined by λ::Tn*phoA* insertion mutagenesis and screening for IKe sensitivity.

200 insertion derivatives were isolated.

They first have been analyzed for their ability to confer sensitivity to IKe phage infection. When the transposon is inserted in a region flanking the transfer region, the pilus is produced and the infectivity of IKe is not effected, e.g. infection plaques are present. These clones are said to be IKe^s. When infection plaques are not present, the clone is resistant to the phage and thus the transposon is inserted into the transfer region: these clones are IKe^r. We note that about half clones are resistant to IKe infection. This is in agreement with the expected ability of the transposon to randomly insert into a plasmidic DNA. In fact pBGB is about 21 kbp long, while the total transfer region is expected to have a length of about 10 kbp (Figure 3.14.).

After testing the infectivity of IKe, the DNA was extracted from these cells and digested with Hind3 or/and BamH1. Figure 3.16. shows the different restriction patterns generated by the latter two enzymes for pBGB and pBGB with inserted Tn*phoA*. Hind3 cuts twice in the transposon and twice in pBGB: thus two constant fragments will be always present, one 3 kbp long from the vector, the other 3.9 kbp from the transposon. Double digestion with Hind3 and BamH1 allow to identify the orientation of the inserted transposon. The map shown in Figure 3.17. was obtained. IKe^r and IKe^s insertions are indicated and the region coding for the structural proteins of the N pilus is thus defined.

Figure 3.16. Patterns generated by Hind3 or/and BamH1 for pBGB and pBGB with inserted Tn*phoA*.

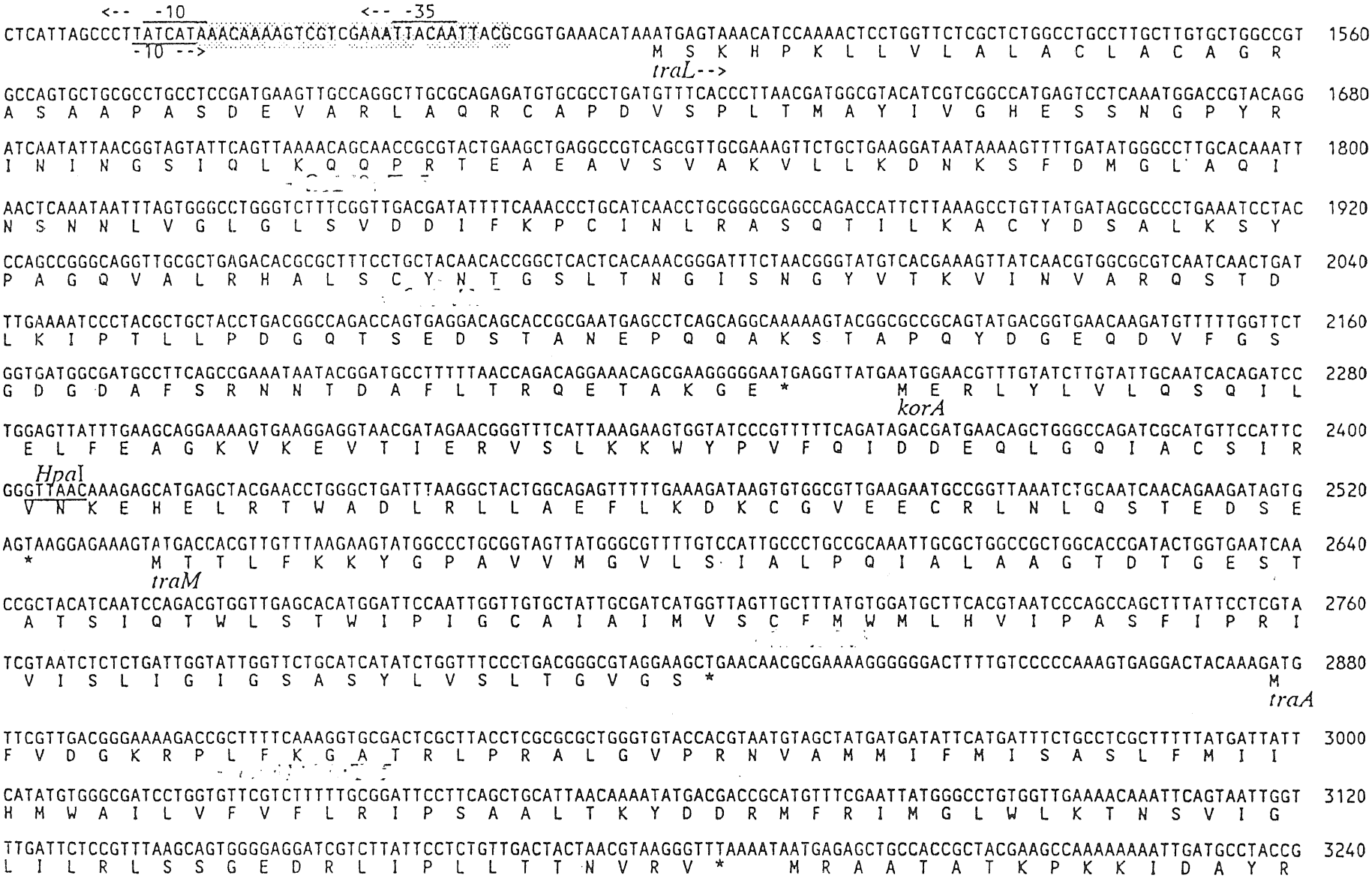
Figure 3.17. Map of the transposon insertions. IKe^r and IKe^s insertions are indicated and the region coding for the structural proteins of the N pilus is thus defined.

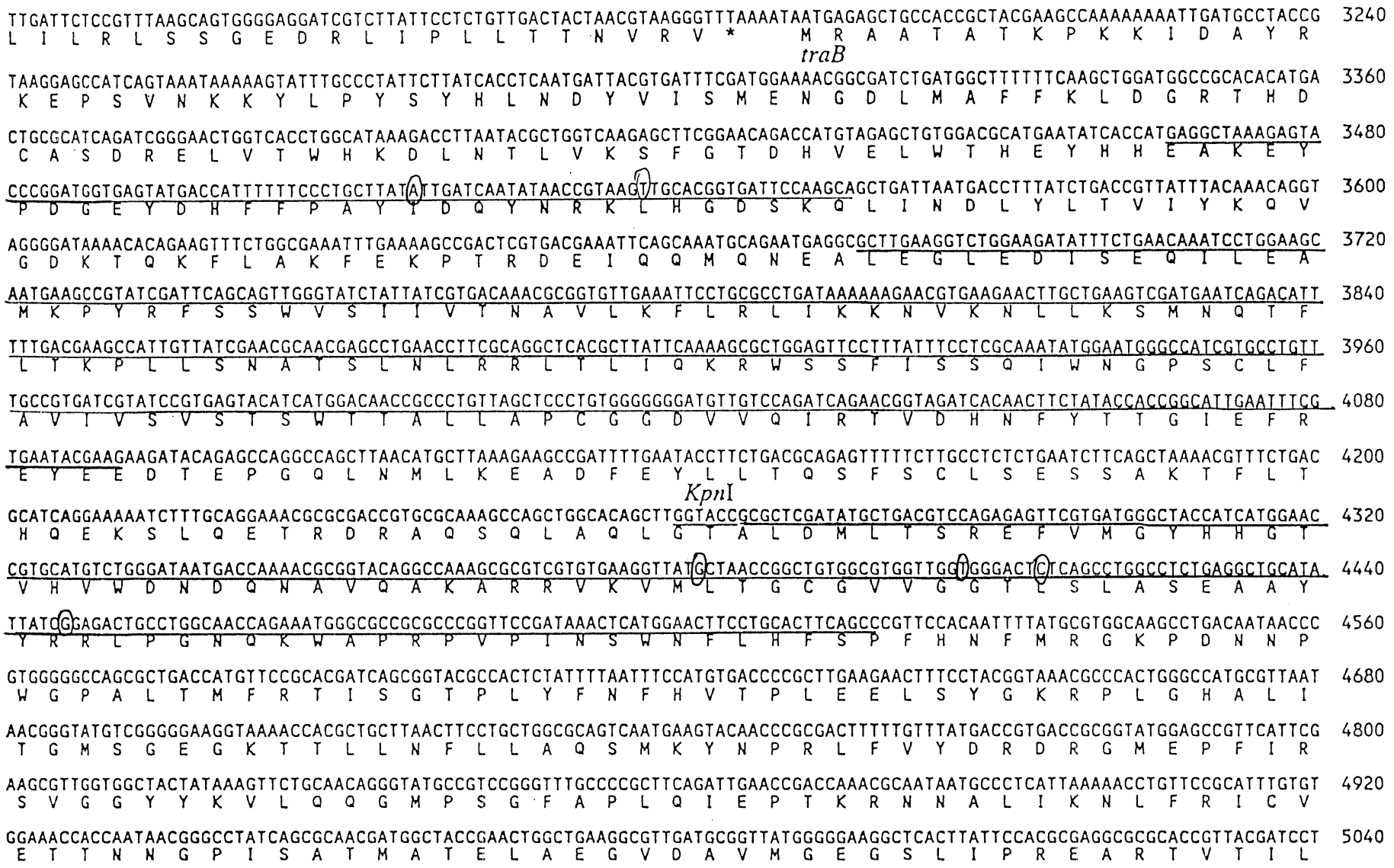
3.2.3. DNA sequence.

A series of Tn*phoA* insertions have been isolated in the *tra* region (see previous paragraph). Some of these have been chosen as starting point for sequencing. A synthetic oligonucleotide that is complementary to one end of Tn*phoA* was used to determine the sequence of DNA immediately adjacent to each insertion. Tnpho-back (GCC GGG TGC AGT AAT ATC) was used to amplify the 5′ region with respect to the inserted transposon. It was not possible to obtain an oligonucleotide amplifying the 3' region, because of the presence of 680 bp from Tn5 in pBGB. The Pst1 site used to clone the transfer region is situated inside the transposon which was used in the previous work to isolate the Tra1 region (Glocker & Rasched, 1990).

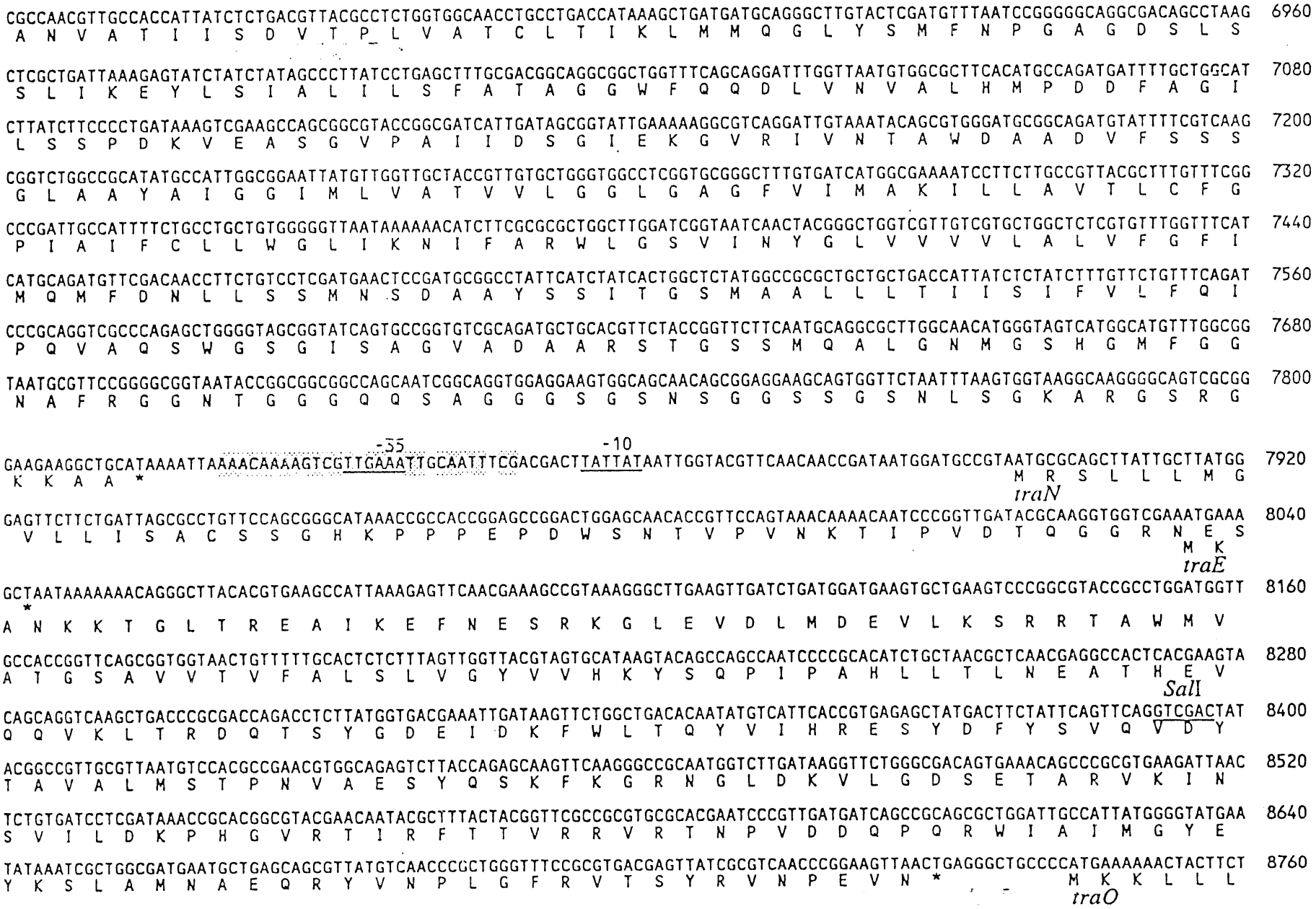
3.2.3.1. Sequence homology to another IncN plasmid. Figure 3.18. reports complete sequence of the Tra1 region from pKM101 and the sequence obtained in this study, that is from IncN N3, is reported as. The nucleotides which differ from one sequence to the other are highlighted. In Figure 3.24. the amino acid residues which result to be changed for the presence of missence mutations is reported. Comparison of the sequence obtained for the transfer region of IncN N3 with the published sequence of the transfer region of pKM101 has revealed that the two plasmids are almost completely homologous but some point mutations are present.

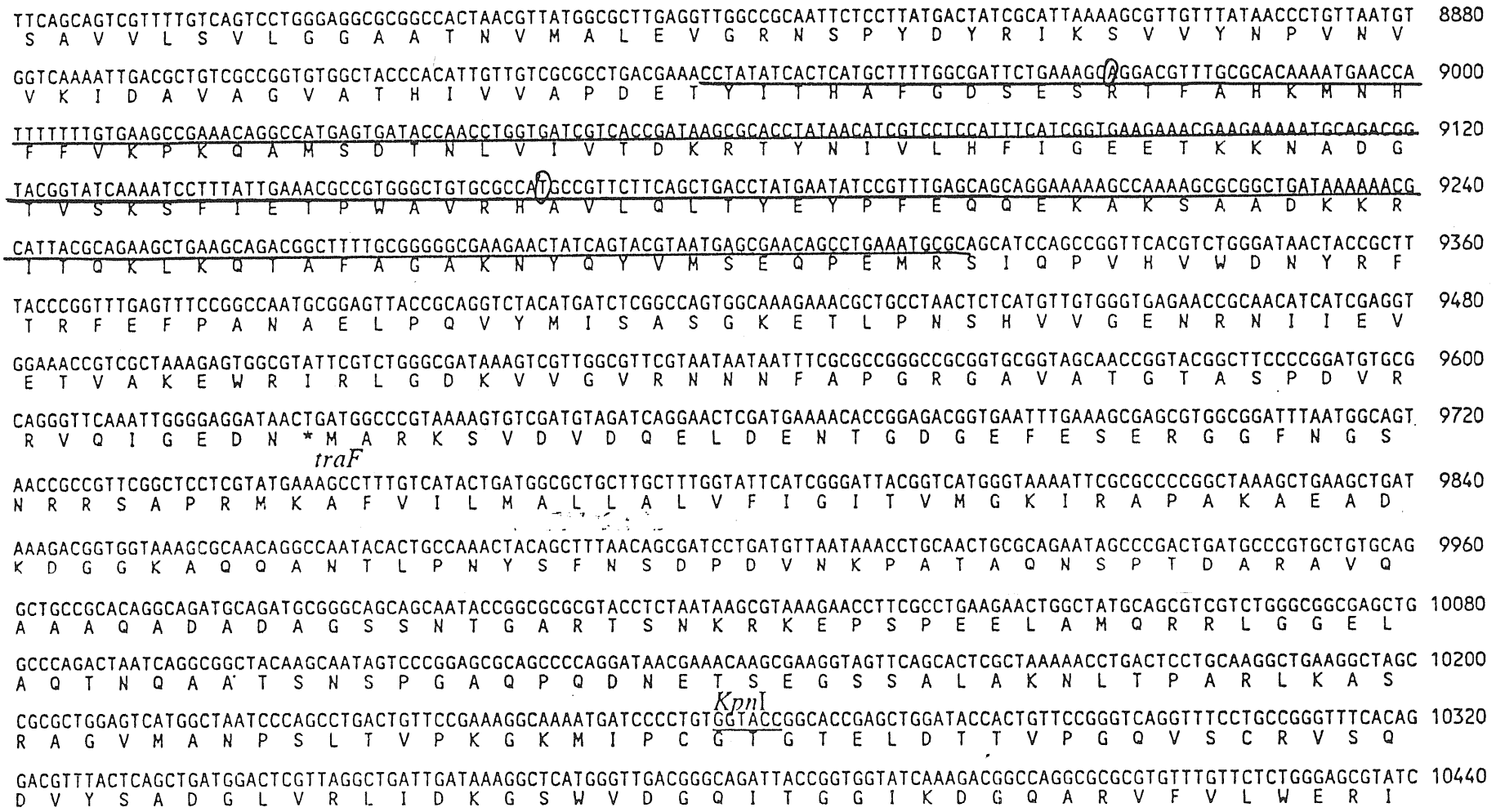
Figure 3.23. In the following pages the complete sequence of the Tra1 region from pKM101 (Pohlman et al., 1994) is reported and compared to the available IncN N3 sequence (underlined nucleotides). The nucleotides which differ from one sequence to the other are highlighted. The open reading frame of each protein is reported.

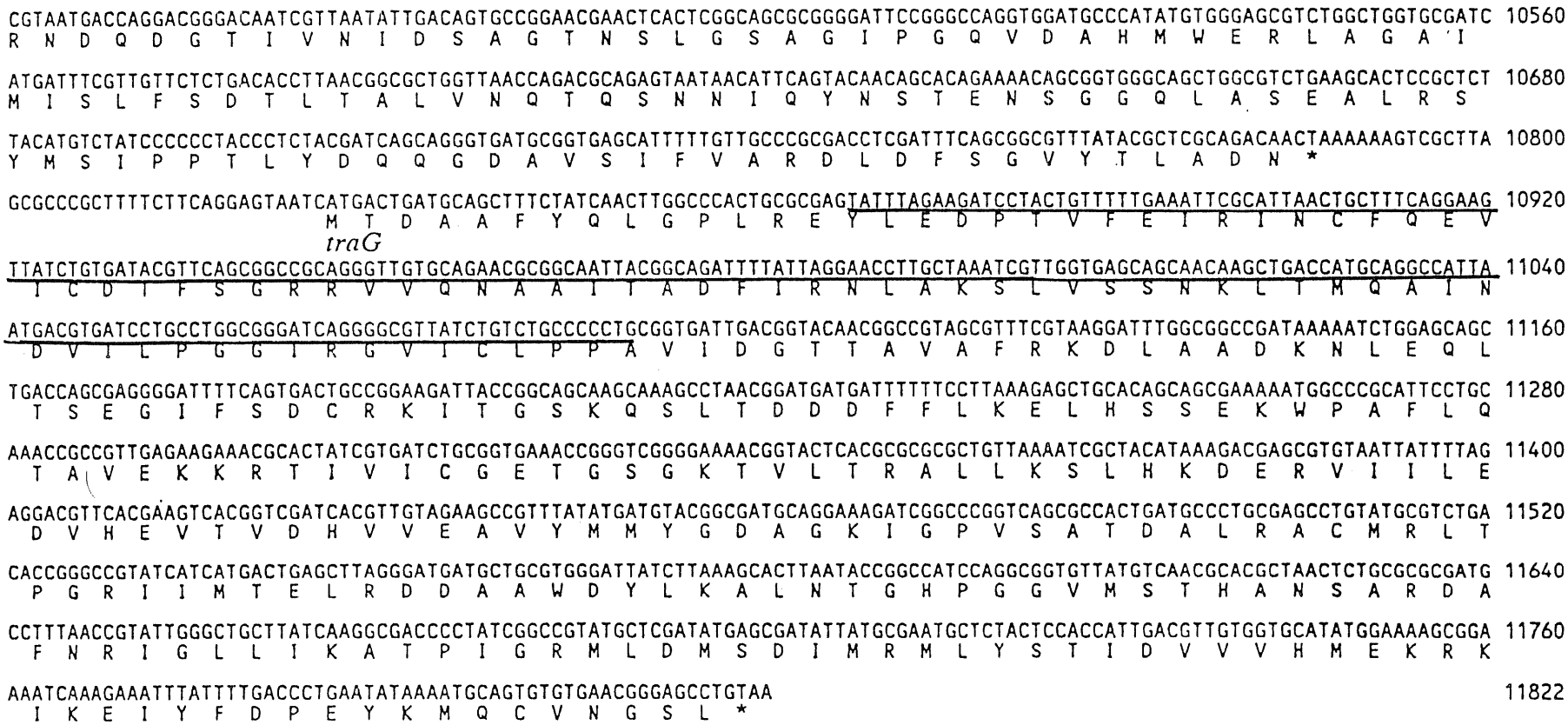

<-- -10
-10 -->
<-- -35
traL-->
korA
HpaI
traM
traA


traB
KpnI


GGACGGGTACGTGAATGAAGTTGTGGAAAATGGCGTATCACTGAAAGGGCTGCTGCGCGAATGGACGCGCGAAGGCCAGTATGGCTGGCTGTTTGACAATGATAAAGACAGCCTGGATCT 5160
D G Y V N E V V E N G V S L K G L L R E W T R E G Q Y G W L F D N D K D S L D L
CAGCGCGAATGATATTTTTGGCTTCGATTTATCCGAGTTTATCGCAGCCAAAGAGGAAGTATCCAGCCCGGCCCGTACTCCGCTCATAGTGTACCTTCTGTACCGGGTACGTGACTCCAT 5280
S A N D I F G F D L S E F I A A K E E V S S P A R T P L I V Y L L Y R V R D S I
CGACGGCAAACGCCGCGTCATTCAGTGCTTTGACGAGTTCCACGCCTACCTTGACGATCCGGTTATCGAGCGTGAAGTTAAGCGTGGTATCAAAACTGACCGTAAGAAAGACGCTATCTA 5400
D G K R R V I Q C F D E F H A Y L D D P V I E R E V K R G I K T D R K K D A I Y
TGTGTTTGCCACGCAGGAGCCGAACGATGCGCTGTCCAGCCGTATTGGCCGCACGATCATGTCGCAGACCGTCACAAAAATCTGCCTGCGCGATCCGGAAGCTATCCGAGAGGATTATGC 5520
V F A T Q E P N D A L S S R I G R T I M S Q T V T K I C L R D P E A I R E D Y A
CTTCCTTACTGATGCTGAATACGACGCGCTGATGTCGATTACCGAACACTCCAGACAGTTCCTGGTTAAACAAGGGCAACAGTCTGCGATTGCTTCTTTCAATCTCTACCCTCGCAACAG 5640
F L T D A E Y D A L M S I T E H S R Q F L V K Q G Q Q S A I A S F N L Y P R N S
CGACGATATTGATGCAGATATTAAGACAATGGACAACGTACTTAGCGTGTTGTCCGGTGAACCACAAAACGCCGAAATTGCGCATGAGCTGGTTGAACGGCTCGGTAATGACCCTGAAGT 5760
D D I D A D I K T M D N V L S V L S G E P Q N A E I A H E L V E R L G N D P E V
ATGGCTCAAAGAATACTTGGCGCCTGACGGCTTAACAACGAGGCAAAACTCCATGAAAAAATCACTTACGGCAGTGCTGCTGACCACCGGCTTGATTCTTGGGGGCGCGCAAAGGCCCTCC 5880
W L K E Y W R L T A * M K K S L T A V L L T T G L I L G G A Q R P S
traC
GCTGGCATCATCGTATCCAACCCCACCGAGCTGATTAAACAGGGTGAACAGCTTGAACAAATGGCTCAACAGCTGGAGCAGCTTAAAAGCCAGCTGGAGACGCAGAAAAATATGTATGAG 6000
A G I I V S N P T E L I K Q G E Q L E Q M A Q Q L E Q L K S Q L E T Q K N M Y E
TCGATGGCTAAAACAACCAACCTGGGCGATCTGCTGGGGACGTCAACCAATACGCTGGCGAATAACCTGCCAGACAACTGGAAAGAGGTGTACAGCGACGCGATGAACTCCAGTTCGTCG 6120
S M A K T T N L G D L L G T S T N T L A N N L P D N W K E V Y S D A M N S S S S
GTCACGCCTTCTGTTAACAGCATGATGGGCCAGTTTAATGCTGAGGTTGACGACATGAGCCCCAGGCAAGCAATCGCCTATATGAAACAAAAGCTGGACGGAAAAAGGTGCTTACGACCG 6240
V T P S V N S M M G Q F N A E V D D M S P R Q A I A Y M K Q K L D G K R C L R P
TGTGATGGCCGAAAAGCCTACAACAACCAGATGCAAGAGCTGAGCGATATGCAGGCGCTGACGGAGCAGATTAAATCGACGCCAGACCTGAAATCCATTGCCGACCTGCAAGCACGTATC 6360
C D G R K A Y N N Q M Q E L S D M Q A L T E Q I K S T P D L K S I A D L Q A R I
CAGACGTCACAAGGCGCTATACAGGGTGAGCAGGCGACATGGAATTTGATGAATATGTTACAGCAATCGCAGGACAAACTTTTGCGCGCACAGAAGGATCGTGCGACCCGTAATTTTGTT 6480
Q T S Q G A I Q G E Q A T W N L M N M L Q Q S Q D K L L R A Q K D R A T R N F V
TTTGGAACCGGTGGGGACGTTACCGCGTCACCTTCAATTAACTGAGGAAATTATGAAAAAACTACTGCTGCTTATCCCTTTCTTCCTGGTGGCCTGCGACGCTTCGCATGACGTTGATTG 6600
F G T G G D V T A S P S I N * M K K L L L L I P F F L V A C D A S H D V D W
eex
GTACAAAAAGCATGACAAAGAGCGCAAGGCAACGATCGAGGAGTGCAAGAAGAATGCTGATGAACTCCAGAAGCCAGACTGCAAAAACGCGCGTGAGGCTGATCGCCAGCTGTTTGTTCT 6720
Y K K H D K E R K A T I E E C K K N A D E L Q K P D C K N A R E A D R Q L F V L
CGGTAAAGAGGACGGTGAAATTAACTCACCGAAAATTTAAGAGTAAGGAGGCAATATGGCATTCACCCTGGTACAAGACATTTTCGCAAAAGTAGACGGGGCGATTACGTCAATGGTGAG 6840
G K E D G E I N S P K I * M A F T L V Q D I F A K V D G A I T S M V S
traD


-35
-10
traN
traE
SalI
traO


traF
KpnI


CGTAATGACCAGGACGGGACAATCGTTAATATTGACAGTGCCGGAACGAACTCACTCGGCAGCGCGGGGATTCCGGGCCAGGTGGATGCCCATATGTGGGAGCGTCTGGCTGGTGCGATC 10560
R N D Q D G T I V N I D S A G T N S L G S A G I P G Q V D A H M W E R L A G A I
ATGATTTCGTTGTTCTCTGACACCTTAACGGCGCTGGTTAACCAGACGCAGAGTAATAACATTCAGTACAACAGCACAGAAAACAGCGGTGGGCAGCTGGCGTCTGAAGCACTCCGCTCT 10680
M I S L F S D T L T A L V N Q T Q S N N I Q Y N S T E N S G G Q L A S E A L R S
TACATGTCTATCCCCCCTACCCTCTACGATCAGCAGGGTGATGCGGTGAGCATTTTTGTTGCCCGCGACCTCGATTTCAGCGGCGTTTATACGCTCGCAGACAACTAAAAAAGTCGCTTA 10800
Y M S I P P T L Y D Q Q G D A V S I F V A R D L D F S G V Y T L A D N *
GCGCCCGCTTTTCTTCAGGAGTAATCATGACTGATGCAGCTTTCTATCAACTTGGCCCACTGCGCGAGTATTTAGAAGATCCTACTGTTTTTGAAATTCGCATTAACTGCTTTCAGGAAG 10920
M T D A A F Y Q L G P L R E Y L E D P T V F E I R I N C F Q E V
traG
TTATCTGTGATACGTTCAGCGGCCGCAGGGTTGTGCAGAACGCGGCAATTACGGCAGATTTTATTAGGAACCTTGCTAAATCGTTGGTGAGCAGCAACAAGCTGACCATGCAGGCCATTA 11040
I C D T F S G R R V V Q N A A I T A D F I R N L A K S L V S S N K L T M Q A I N
ATGACGTGATCCTGCCTGGCGGGATCAGGGGCGTTATCTGTCTGCCCCCTGCGGTGATTGACGGTACAACGGCCGTAGCGTTTCGTAAGGATTTGGCGGCCGATAAAAATCTGGAGCAGC 11160
D V I L P G G I R G V I C L P P A V I D G T T A V A F R K D L A A D K N L E Q L
TGACCAGCGAGGGGATTTTCAGTGACTGCCGGAAGATTACCGGCAGCAAGCAAAGCCTAACGGATGATGATTTTTTCCTTAAAGAGCTGCACAGCAGCGAAAAATGGCCCGCATTCCTGC 11280
T S E G I F S D C R K I T G S K Q S L T D D D F F L K E L H S S E K W P A F L Q
AAACCGCCGTTGAGAAGAAACGCACTATCGTGATCTGCGGTGAAACCGGGTCGGGGAAAACGGTACTCACGCGCGCGCTGTTAAAATCGCTACATAAAGACGAGCGTGTAATTATTTTAG 11400
T A V E K K R T I V I C G E T G S G K T V L T R A L L K S L H K D E R V I I L E
AGGACGTTCACGAAGTCACGGTCGATCACGTTGTAGAAGCCGTTTATATGATGTACGGCGATGCAGGAAAGATCGGCCCGGTCAGCGCCACTGATGCCCTGCGAGCCTGTATGCGTCTGA 11520
D V H E V T V D H V V E A V Y M M Y G D A G K I G P V S A T D A L R A C M R L T
CACCGGGCCGTATCATCATGACTGAGCTTAGGGATGATGCTGCGTGGGATTATCTTAAAGCACTTAATACCGGCCATCCAGGCGGTGTTATGTCAACGCACGCTAACTCTGCGCGCGATG 11640
P G R I I M T E L R D D A A W D Y L K A L N T G H P G G V M S T H A N S A R D A
CCTTTAACCGTATTGGGCTGCTTATCAAGGCGACCCCTATCGGCCGTATGCTCGATATGAGCGATATTATGCGAATGCTCTACTCCACCATTGACGTTGTGGTGCATATGGAAAAGCGGA 11760
F N R I G L L I K A T P I G R M L D M S D I M R M L Y S T I D V V V H M E K R K
AAATCAAAGAAATTTATTTTGACCCTGAATATAAAATGCAGTGTGTGAACGGGAGCCTGTAA 11822
I K E I Y F D P E Y K M Q C V N G S L *

	sequence from pKM101		sequence from N_3	
gene name	codon	amino acid	codon	aminoacid
traB				
	GAG	glu	CAA	gln
	GCT	alanine	GCA	alanine
	AAA	lysine	AAG	lysine
	TCA	serine	TCT	serine
	AAT	asparagine	AGC	serine
	GCG	alanine	GCA	alanine
	AAC	asparagine	AAT	asparagine
	CTG	leucine	TTG	leucine
	CCA	proline	CCG	proline
	AAA	lysine	AAG	lysine
	GTG	valine	ATC	ilsoleucine
	GCG	alanine	GCC	alanine
	TCG	serine	TCT	serine
	TCG	serine	TGG	tryptophane
	TCT	serine	TCA	serine

Figure 3.24. In the following tables the difference in the *tra1* gene sequence from N3 and pKM101 is reported. Only those mutations which results in a change in amino acid sequence are reported.

	sequence from pKM101		sequence from N_3	
gene name	codon	amino acid	codon	aminoacid
traB				
	ATT	isoleucine	GTT	valine
	TTG	leucine	CTG	leucine
	ATG	methionine		
	GGT	glycine		
	CTC	leucine		
	GTA	valine	GTT	valine
	TGG	tryptophan	GGG	cysteine
traO				
	AGG	arginine	TGG	tryptophan

3.2.3.2. Amino acid sequence homology to other Inc groups. The gene products of the IncN conjugal transfer region show significant similarities to proteins encoded by plasmids from other incompatibility groups, i.e. IncW and IncP. These plasmids are called promiscuous conjugative plasmids, because of their broad host range. They conjugatively transfer to many different genera of Gram-negative bacteria.

Comparative analysis between the *tra* genes of IncN and those of the Tra2 operon of RP4 (Lessl *et al.*, 1992) and of the PilW operon of R388 (Bolland *et al.*, 1990) revealed striking similarity. In Figure 3.25. the structure of the cluster of genes for the IncN and IncP transfer region is reported while in Figure 3.26. the sequence alignment of the IncN transfer regions from pKM101, and the IncP from RP4 is shown.

Transfer system of the IncP plasmids have an extremely wide host range, and appear to be capable of mediating DNA transfer into virtually any Gram-negative bacterium. The IncP can mediate horizontal gene transfer from Gram-negative bacteria to phylogenetically remote organisms including even yeast (Heinemann & Spague, 1989). Transfer genes of the IncP plasmid RP4 are grouped in two separate regions, designed Tra1 and Tra2. *Tra2* gene products are proposed to be mainly responsible for the formation of mating pairs in conjugating cells. The nucleotide sequence of the entire RP4 Tra2 region has shown twelve open reading frames (Lessl et al., 1992). Based on hydropathy plot analysis, most of the Tra2 open reading frames encode proteins that may interact with membranes. TrbB, TrbC, TrbE,

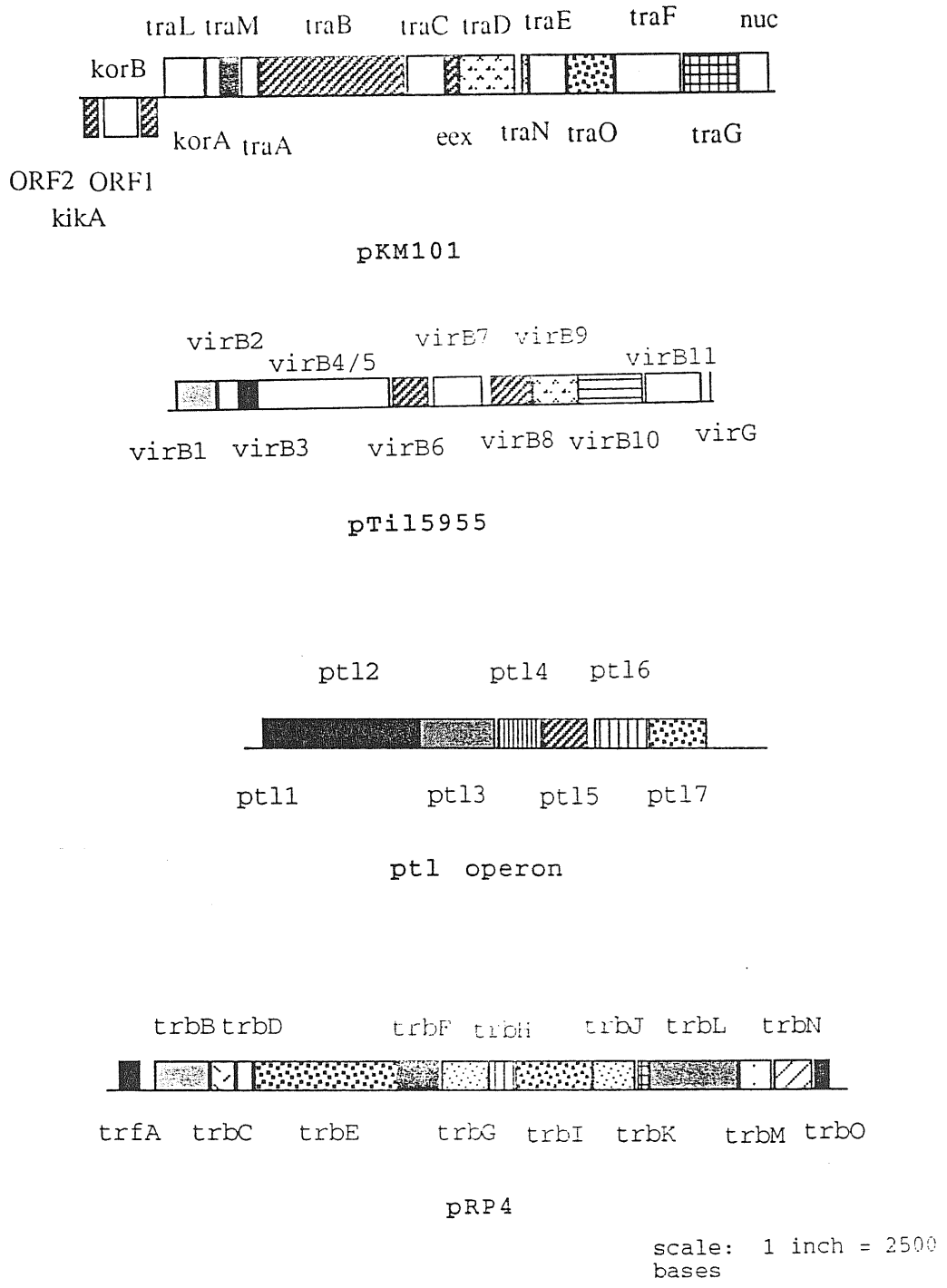


Figure 3.25. Comparison of the structure of the cluster genes of the pKM101 *tra* genes, *Agrobacterium tumefaciens* virB genes, *Bordetella pertussis ptl* genes, and IncP *tra* genes.

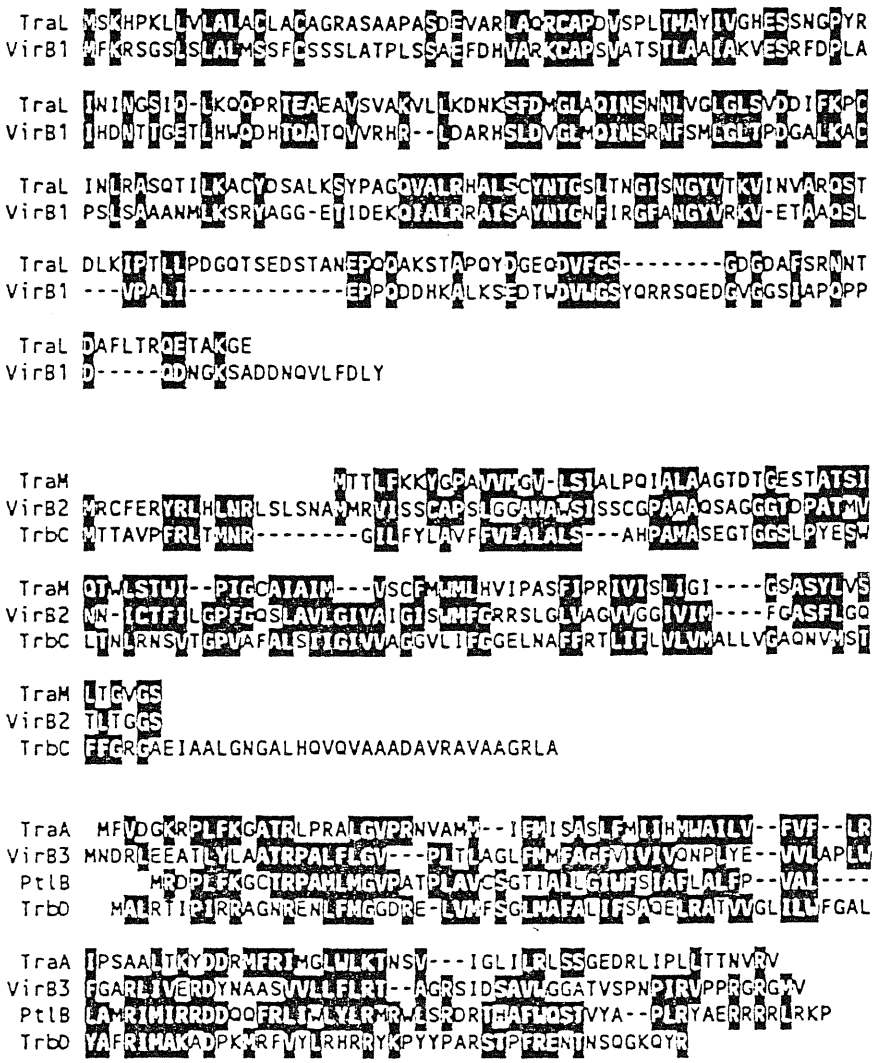
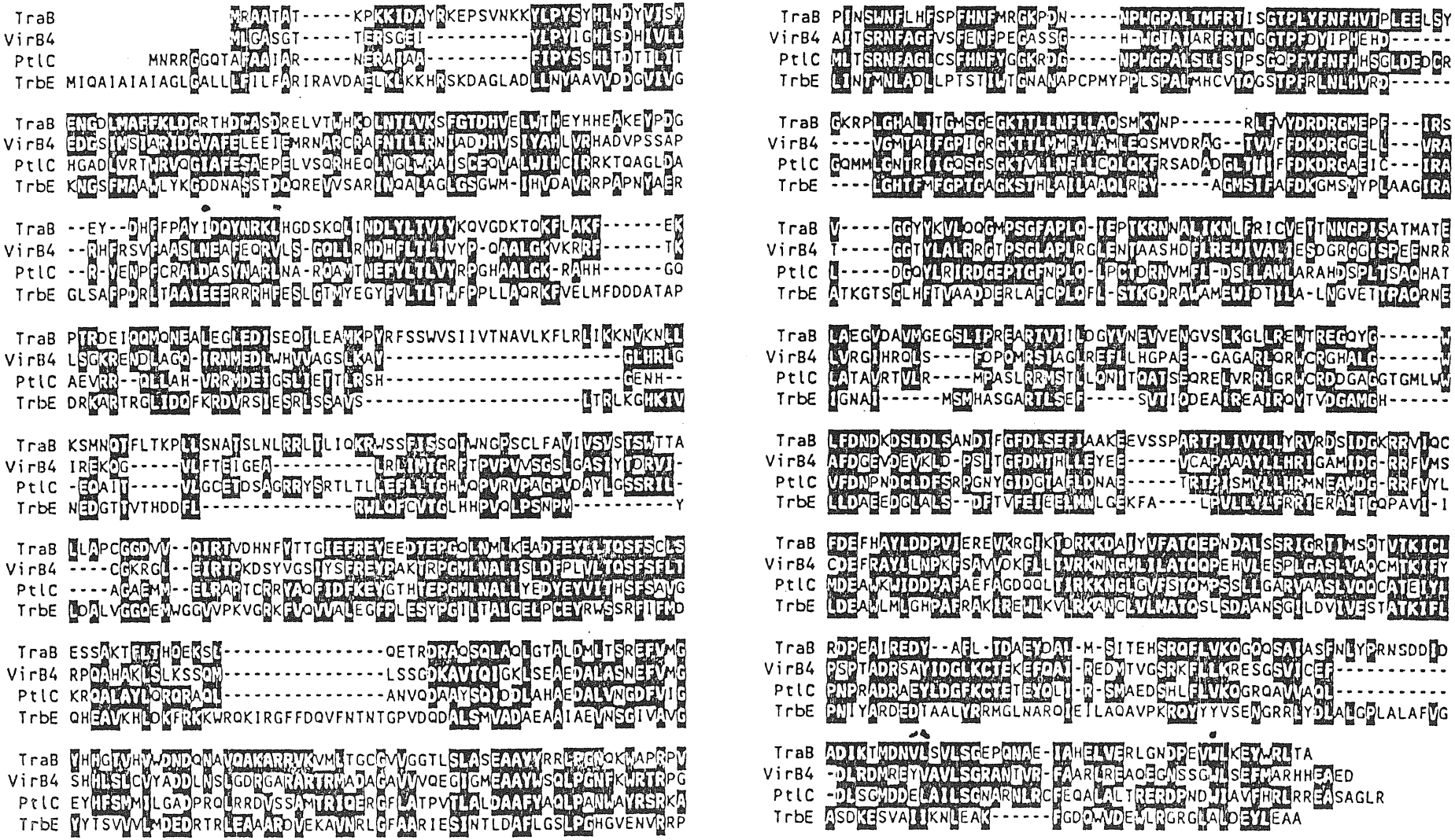
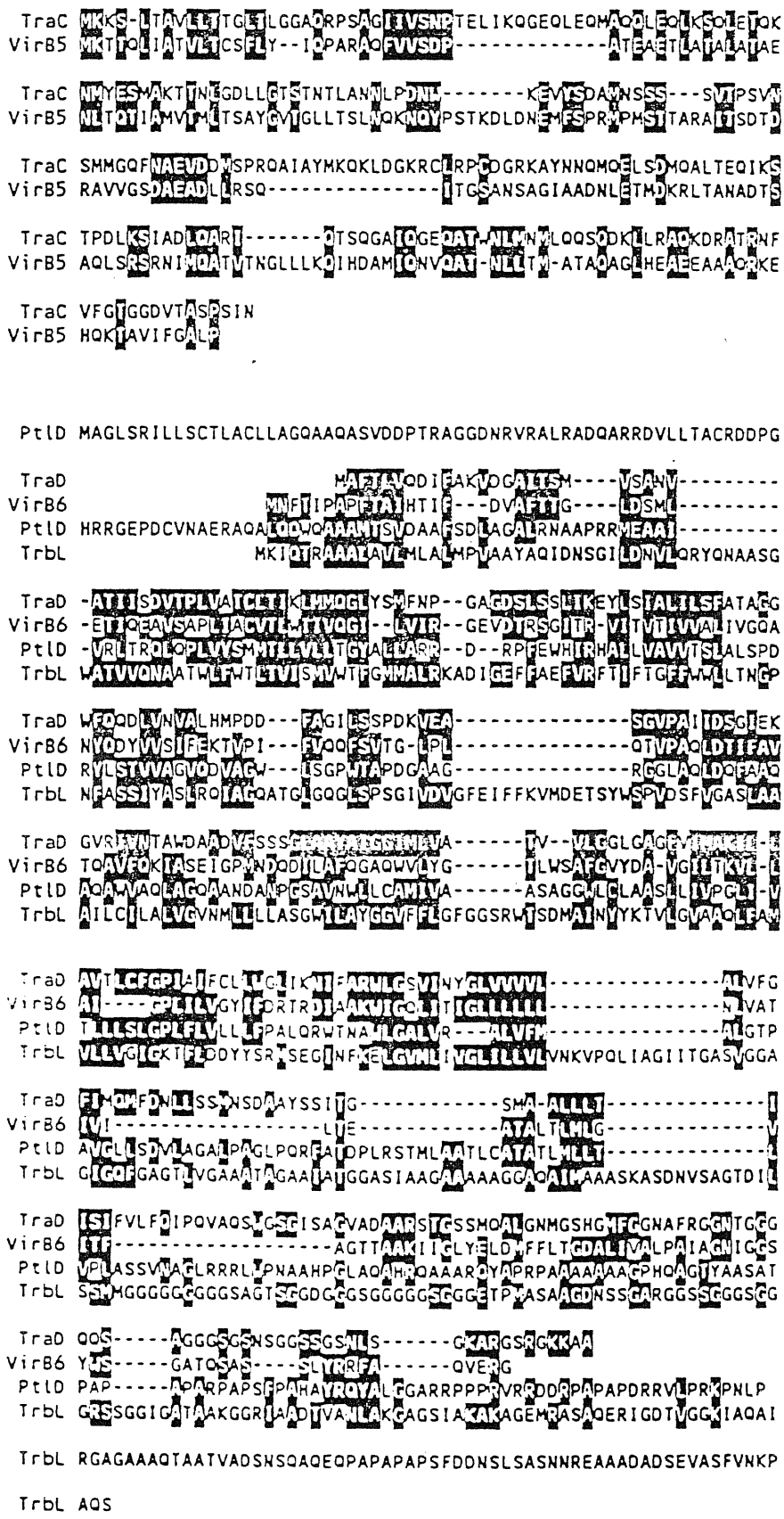
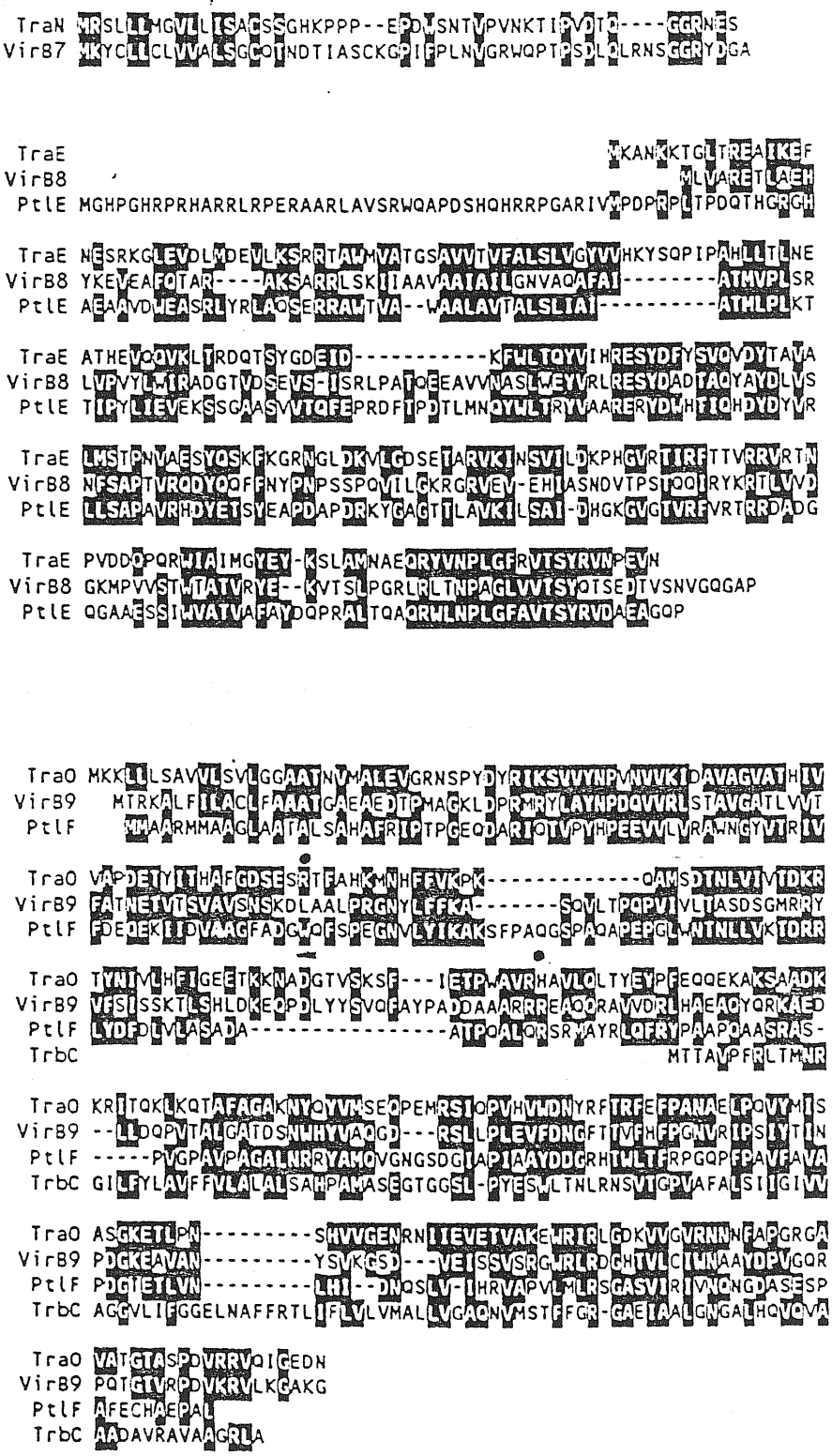
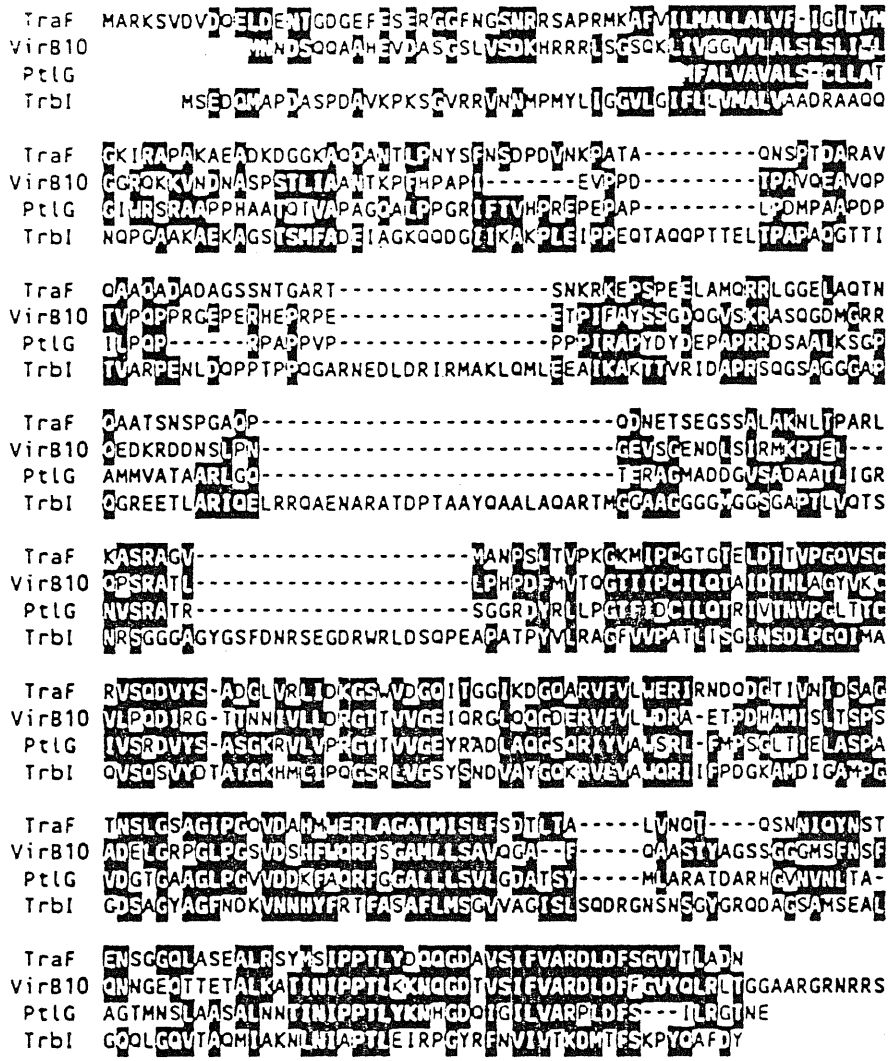


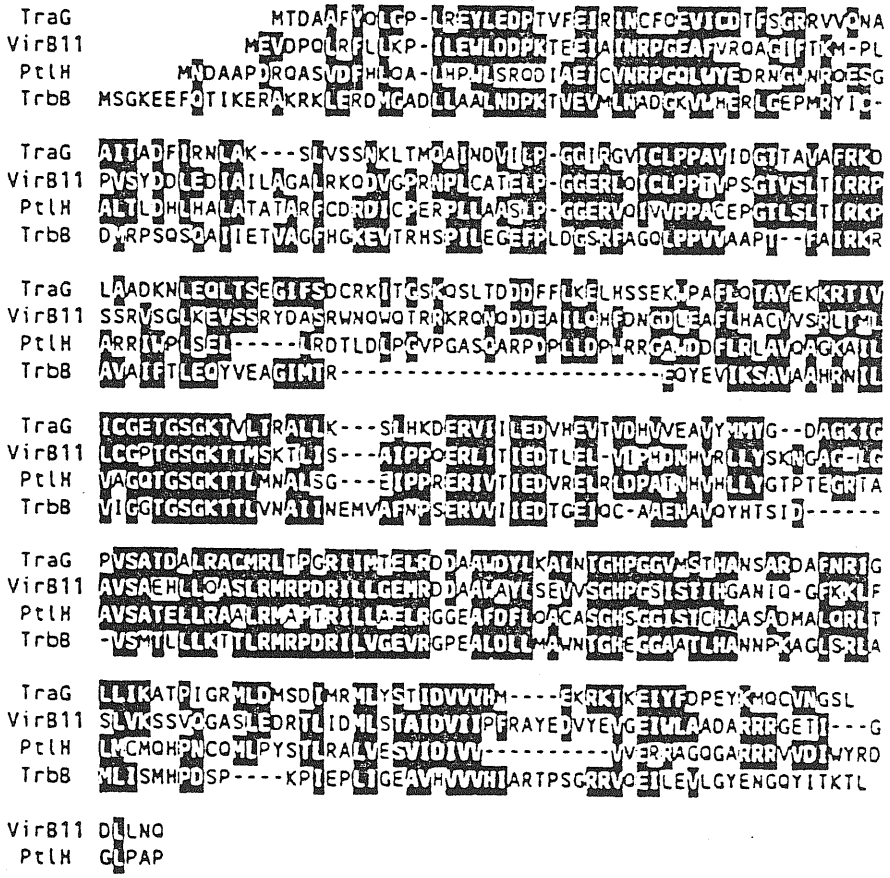
Figure 3.26. Comparison of the sequence of proteins produced by *tra* genes of IncN (pKM101), *virB* genes of *Agrobacterium tumefaciens*, *ptl* genes of*Bordetella pertussis*, and *tra* genes of IncP (RP4). From (Pohlman et al., 1994).


TraB
VirB4
PtlC
TrbE


TraC
VirB5
PtlD
TraD
VirB6
TrbL


TraN
VirB7
TraE
VirB8
PtlE
TraO
VirB9
PtlF
TrbC


TraF
VirB10
PtlG
TrbI


TraG
VirB11
PtlH
TrbB

TrbG, and TrbL have been proposed to be membrane components and have been shown to be essential for conjugation and absorbance by donor-specific phage (Lessl et al., 1992)

It has been shown that W-pili are morphologically similar (Bradley, 1980), both are sensitive to the pilus-specific phage PRD1 (Olsen *et al.*, 1974), but do not cross react immunologically. The pilW operon of IncW plasmid R388 is required for conjugative transfer of this plasmid between bacteria and, more specifically, for synthesis and assembly of the conjugative pilus (Bolland et al., 1990). Relatedness is also shown by the ability of the TRA plasmids to complement mutations in PIL_W (Bolland et al., 1990). Sequence analysis of the PilW operon of the IncW plasmid R388 is reported but the sequence has not been published, nor has it been submitted to data bank. It seems that 50.8% similarity between 10 of the predicted products of the PilW operon and virB operon was shown.

3.2.3.4. Amino acid sequence homology to other export proteins.

Extensive protein sequence similarities between IncN transfer region from pKM101 and both the products of the *virB* operon of *Agrobacterium tumefaciens* , and the *ptl* operon of *Bordetella pertussis* were previously reported (Pohlman et al., 1994) (Figure 3.26.).

Agrobacterium tumefaciens is a gram-negative soil bacterium. It causes crown gall disease by transferring a segment of DNA, the T-DNA (transferred DNA), from its Ti (tumour inducing) plasmid to susceptible plant cells where it is integrated and maintained. For this

interkingdom DNA transfer process to occur, a set of Ti plasmid-encoded virulence (*vir*) genes is required. These genes are organized in six essential operons, of which the largest is the 9.5-kb *virB* operon. The *virB* operon have been sequenced (Kuldau *et al.*, 1990; Shirasu *et al.*, 1990; Thompson *et al.*, 1988; Ward *et al.*, 1988; Ward *et al.*, 1990a). It is composed of 11 genes (*virB1* to *virB11*). The 11 gene products of the *virB* operon, together with the VirD4 protein, are proposed to form a membrane complex which mediates the transfer of T-DNA to plant cells. The VirB1, VirB2, VirB5, VirB7, and VirB9 proteins each contain a signal peptide-like sequence which suggests they are membrane components. Nine of the VirB proteins and VirD4 have been localized to the inner and outer membrane (Shirasu & Kado, 1993; Shirasu *et al.*, 1994; Thorstenson *et al.*, 1993; Ward *et al.*, 1990b).

Current data suggest that VirB proteins form a structure in the bacterial membrane to allow passage of the T-DNA complex to the plant cell, and it is likely that there is close association of the inner and outer membranes at the point where the VirB proteins accumulate. The process of T-DNA transfer from to plant cells can be considered as a conjugation-like mechanism. The transfer of T-DNA presumably requires energy, possibly produced through the action of a transport ATPase. Recent studies have shown that the *vir* region can function to promote the conjugative transfer of certain broad-host range plasmids between *Agrobacterium* cells and the *virB* operon is essential for this process (Beijersbergen *et al.*, 1992). Anyway the appearance of pili following *vir* gene induction has not been reported in *A. tumefaciens*.

The deduced amino acidic sequence of 6 of the *trb* genes (IncP), 10 of the *trw* genes (IncW), and 11 of the *trn* genes (IncN) share significant sequence identity with the VirB proteins. Not only are the *virB* genes similar in sequence to the conjugation genes but they are also functionally similar.

The second related transport system, encoded by the *Bordetella pertussis ptl* operon, directs the export not of DNA-protein complexes but of a complex of six proteins comprising pertussis toxin (PT). PT is composed of five types of subunits, S1-S5, found in a 1:1:1:2:1 ratio. Each PT subunit is synthesized with a secretion signal sequence which should target it to the periplasm where the subunits may then associate with membrane proteins. A region of the *B. pertussis* chromosome, termed the *ptl* locus, is critical for efficient secretion of PT (Weiss *et al.*, 1993). The *ptl* locus of *Bordetella pertussis* contains eight open reading frames which are predicted to encode proteins (PtlA to PtlH) that are essential for secretion of pertussis toxin from the bacterium (Weiss et al., 1993). Three of the Ptl proteins, PtlE, PtlF, and PtlG, have been identified and their cellular localization have been determined (Johnson & Burns, 1994). This study indicates that all three Ptl proteins are associated with the membranes of *B. pertussis*, suggesting that the Ptl proteins form a gate or channel which facilitates transport of pertussis toxin. The same experiments carried out on related *B. pertussis* strains *Bordetella parapertussis* and *Bordetella bronchiseptics*, which are known to be defective in the secretion of PT. show that they do not contain detectable levels of PtlE or PtlF. This lack of detectable

Ptl protein may provide an explanation for previous observations which indicated that introduction of the genes encoding pertussis toxin subunits from *B. pertussis* into other *Bordetella spp.* results in production of the toxin but not secretion of the toxin.

The *ptl* pathway seems to be necessary to complete the secretion process. PT presumably only needs assistance through the outer membrane of *B. pertussis*, perhaps explaining why only 7 of the 11 *virB* or *tra* genes have a *ptl* homology. An hypothesis could be that the genes which are missing are those which code for proteins that form a structure like a pilus. In fact the toxin only needs to exit the cell, and it can recognize its target cell by its own.

3.2.3.4. A role for the proteins. coded by the transfer region .

In the very last months some papers investigating the role of these proteins have appeared. The creation of a DNA/protein transfer channel has been proposed. This channel could allow the passage of not only the prepilin, but also the DNA.

Two of the proteins in each of the studied trasfer systems contain predicted type A nucleotide-binding sites (Walker boxes): TraG/TrbB/VirB11 and TraB/TrbE/VirB4. This consensus domain places these proteins among a superfamily of mononucleotide binding proteins involved in bacterial transport processes (Higgins *et al.*, 1986; Walker *et al.*, 1982). Proteins of this superfamily are thought to hydrolize ATP to provide energy for the transport processes (Higgins, 1990). Preliminary experiments indicate that TrbB, VirB11 and VirB4

possess ATPase and protein kinase activity (Fullner *et al.*, 1994; Lederberg, 1951; Lessl *et al.*, 1993). Interestingly, TraG-like proteins exist in numerous export systems of various bacteria (Hobbs & Mattick, 1993). These either contribute to formation of type IV fimbriae (PilB, PilT, PilF), secretion of exoproteins (XcpR, PulE, OutE, ExeE, XpsE, EspE, PilH), or uptake of DNA (ComG1). The phylogenetic tree of TraG-related proteins suggests that four proteins (TraG, TrbB, PtlH, and VirB11) form a subgroup in the TraG family. They obviously share a higher degree of similarity with each other than with the rest of the proteins listed. This is surprising since *B. pertussis* Ptl operon is proposed to be responsible for toxin export (Weiss et al., 1993) and hence does not seem to fit into a class of DNA transfer systems encoded by N3 Tra2, RP4 Tra2, and Ti VirB.

The *virB2* gene product is 52% identical in amino acid sequence to pro-pilin from F-plasmid (pro-TraA). The *virB2* gene product is processed to a 7.2 kDa subunit similar to TraA (Shirasu & Kado, 1993). Thus, by analogy to the RP4 transfer system, a pilus-like transport structure may be built in *Agrobacterium tumefaciens*.

3.2.4. Tn*phoA* mutagenesis.

In order to determine whether the products of the transfer region genes are fully or partially exported, the transposon Tn*phoA* was used to make translational fusions between the *tra* genes which are the object of our study and *phoA*, which encodes alkaline phosphatase (Manoil & Beckwith, 1986). Tn*phoA* is a derivative of transposon Tn5

that contains a truncated form of the alkaline phosphatase gene *phoA* in the left inverted repeat of the transposon. The phosphatase gene lacks both a promoter and the signal sequences necessary for transport of the enzyme in the periplasmic space where it is functional. This construct permits the generation of hybrid proteins composed of alkaline phosphatase fused to amino-terminal sequences of other proteins.

Alkaline phosphatase which is normally located in the periplasm of *E. coli,* is efficiently secreted to the periplasm when fused either to a signal sequence from another periplasmic protein or to a signal sequence from the outer membrane proteins. These heterologous signal sequences are processed during secretion. In the absence of a complete signal sequence phosphatase becomes localized in the cytoplasm and is inactive. Phosphatase fusion proteins lacking up to 13 amino-terminal amino acids beyond the signal sequence show the same specific activity as that of the wild type enzyme. PhoA can be detected, even at low levels, in bacterial colonies by use of a chromogenic substrate, 5-bromo-4-chloro-3-indolyl-phosphate (XP). Colonies formed by *E. coli* cells with exported alkaline phosphatase are blue.

Some independent insertion derivatives that form blue colonies on BCIP indicator media were identified and localized. DNA of dark blue colonies was isolated and digested with BamH1 or/and Hind3 for rough estimation of the position of TnphoA. Subsequently, double-stranded DNAs of the fusion constructs were sequenced for exact

determination of TnphoA on the plasmid construct using the TnphoA-back primer (-98) 5'- GCCGGGTGCAGTAATATC - 3' (-81). A comparison with the 57 insertions obtained by (Pohlman et al., 1994) shows a complete agreement with our data, with one exception (Figure 3.27.). We obtained an enzymatically active Tn*phoA* insertion in *traB* gene. No insertions in *traB* were found in the work on pKM101 (Pohlman et al., 1994): this was interpreted as an evidence that the TraB protein is cytoplasmically localized. Such hypothesis was said to be supported by the hydropathy profile of the protein. It seems to me that this point is not clear, because the protein is very hydrophilic, and for this reason it is expected to be exported in the periplasm. On the other hand I observe that a class of membrane proteins, i.e. porins, is characterized by having a hydrophilic profile even if they are membrane proteins. TraB is homologous to VirB4, which even if lacking properties typical of a membrane protein has a cytoplasmic membrane localization (Berger & Christie, 1993; Shirasu et al., 1994).

Figure 3.27.

3.2.5. Conclusions.

The partial sequence of the IncN N3 transfer region obtained in this thesis has been compared with the published pKM101 sequence (Pohlman et al., 1994). While the complete sequencing of the transfer region is in progress, the data available to this moment are sufficient to be able to say that the two plasmids are different. The comparison has revealed interesting differences between the two plasmids. A reason for such differences can derive from a different evolutionary story: plasmids maintained in various bacteria belonging to different genera are subject to various site-specific restriction endonucleases. In fact only those plasmids which could preserve the portion which controls all the important Inc group-specific functions can survive in such a variety of bacteria.

It is interesting that proteins from conjugative pilus assembly share similarities with T-DNA transfer and protein secretion systems. This similarity may reflect the involvement of these systems in the passage of macromolecules across the bacterial membranes. Both the processes of T-DNA transfer and conjugation clearly involve passage of a DNA-protein complex through the bacterial membrane, and there is considerable evidence to shown an evolutionary relationship between these systems.

The production of a monoclonal antibody against the pilus is essential in order to identify the protein(s) forming the pilus structure. The presence of an adhesin (if it does exist) at the pilus tip could be

anyway difficult to be demonstrated since this protein may be present in a low copy number. After this it would be possible to hypothesize tolerant sides where to clone an epitope.

References

References

Achtman, M., Morelli, G. & Schwuchow, S. (1978). Cell-cell interactions in conjugating *Escherichia coli*: role of F pili and fate of mating aggregates. *J Bacteriol* **135**(3), 1053-61.

Anthony, K. G., Sherburne, C., Sherburne, R. & Frost, L. S. (1994). The role of the pilus in recipient cell recognition during bacterial conjugation mediated by F-like plasmids. *Mol Microbiol* **13**(6), 939-953.

Aoki, T., Kitao, T. & Arai, T. (1977). R plasmids in fish pathogens. In *Plasmids, medical and theoretical aspects* (Mitsuhashi, S., Rosival, L. & Kremery, V., eds.), pp. 39-45. Springer-Verlag, Berlin.

Appleyard, R. K., McGregor, J. F. & Baird, K. M. (1956). Mutations to extended host range and the occurrence of phenotypic mixing in the temperature coliphage lambda. *Virology* **2**, 565.

Arai, T. & Ando, T. (1979). Plasmids which make their host bacteria mutable as well as resistant to ultraviolet irradiation and their mechanisms. In *Microbial drug resistance* (Mitsuhashi, S., ed.), Vol. 2, pp. 197-203. Japan. Sci. Soc. Press, Tokyo.

Arai, T. & Ando, T. (1980). Plasmids which make their host bacteria mutable as well as resistant to ultraviolet irradiation. *Keio J. Med.* **29**, 46-52.

Arai, T., Sugawara, K. & Aoki, T. (1974). On the stability of various R factors in *Vibrio cholerae*. *Japan. J. Bacteriol.* **29**, 615.

Baron, L. S., Ryman, I. R., Johnson, E. M. & Gemski, P. (1972). Lytic replication of coliphage lambda in *Salmonella typhosa* hybrids. *J.Bacteriol.* **110**, 1022-1031.

Bass, S., Greene, R. & Wells, J. A. (1990). Hormone phage: an enrichment method for variant proteins with altered binding properties. *Proteins* **8**(4), 309-14.

Bauer, M. & Smith, G. P. (1988). Filamentous phage morphogenetic signal sequence and orientation of DNA in the virion and gene-V protein complex. *Virology* **167**(1), 166-75.

Bayer, M. E. (1968). Adsorption of bacteriophages to adhesions between wall and membrane of *Escherichia coli*. *J.Virol.* **2**, 346-356.

Beck, E. & Zink, B. (1981). Nucleotide sequence and genome organisation of filamentous bacteriophages f1 and fd. *Gene* **16**, 35-58.

Beijersbergen, A., Schilperoort, R. A. & Hooykaas, P. J. (1992). Conjugative transfer by the virulence system of *Agrobacterium tumefaciens*. *Nature* **256**, 1324-1327.

Berger, B. R. & Christie, P. J. (1993). The Agrobacterium tumefaciens virB4 gene product is an essential virulence protein requiring an intact nucleoside triphosphate-binding domain. *J Bacteriol* **175**(6), 1723-34.

Bertani, G. & Weigle, J. J. (1953). Host controlled variation in bacterial viruses. *J.Bacteriol.* **65**, 113-121.

Boeke, J. D., Model, P. & Zinder, N. D. (1982). Effects of bacteriophage f1 gene III protein on the host cell membrane. *Mol Gen Genet* **186**(2), 185-92.

Bolland, S., Llosa, M., Avila, P. & de la Cruz, F. (1990). General organization of the conjugal transfer genes of the IncW plasmid R388 and interactions between R388 and IncN and IncP plasmids. *J Bacteriol* **172**(10), 5795-802.

Bradbury, A., Persic, L., Werge, T. & Cattaneo, A. (1993). Use of living columns to select specific phage antibodies. *Bio/Tech.* **11**, 1565-1569.

Bradley, D. E. (1979). Morphology of pili determined by the N incompatibility group plasmid N3 and interaction with bacteriophages PR4 and IKe. *Plasmid* **2**(4), 632-6.

Bradley, D. E. (1980). Morphological and serological relationships of conjugative pili. *Plasmid* **4**(2), 155-69.

Bradley, D. E., Coetzee, J. N. & Hedges, R. W. (1983). IncI2 plasmids specify sensitivity to filamentous bacteriophage IKe. *J. Bacteriol.* **154**, 505-507.

Bradley, D. E. & Whelan, J. (1989). *Escherichia coli* tolQ mutants are resistant to filamentous bacteriophages that adsorb to the tips, not the shafts, of conjugative pili. *J Gen Microbiol* **135**(Pt 7), 1857-63.

Bross, P., Bussmann, K., Keppner, W. & Rasched, I. (1988). Functional analysis of the adsorption protein of two filamentous phages with different host specificities. *J Gen Microbiol* **134**(Pt 2), 461-71.

Buchwald, M. & Siminovich, L. (1969). Production of serum-blocking material by mutants of the left arm of the lambda chromosome. *Virology* **38**, 1.

Caro, L. G. & Schnos, M. (1966). The attachment of the male-specific bacteriophage F1 to sensitive strains of Escherichia coli. *Proc Natl Acad Sci U S A* **56**(1), 126-32.

Casjens, S. R. & Hendrix, R. W. (1974a). Locations and amounts of the major structural protein bacteriophage lambda. *J. Mol. Biol.* **88**, 535.

Casjens, S. R. & Hendrix, R. W. (1974b). Comments on the arrangement of the morphogenetic genes of bacteriophage lambda. *J. Mol. Biol.* **90**, 20.

Cesareni, G. (1992). Peptide display on filamentous phage capsids. *FEBS* **307**(1), 66-70.

Chang, C., Blobel, G. & Model, P. (1978). Detection of prokaryotic signal peptidase in an *E. coli* membrane fraction: endoproteolytic cleavage of nascent f1 precoat protein. *Proc. Natl. Acad. Sci. USA* **75**, 361-365.

Charbit, A., Clement, J. M. & Hofnung, M. (1984). Further sequence analysis of the phage lambda receptor site. Possible implications for the organization of the LamB protein of *Escherichia coli* K12. *J. Mol. Biol.* **175**, 395-401.

Charbit, A., Gehring, K., Niakido, H., Ferenci, T. & Hofnung, M. (1988). Maltose transport and starch binding in phage resistant point mutants of maltoporin. *J. Mol. Biol.* **201**, 487-496.

Charbit, A. & Hofnung, M. (1985). Isolation of different bacteriophages using the LamB protein for adsorption on *Escherichia coli* K-12. *J. Virol.* **53**, 262-275.

Charbit, A., Ronco, J., Michel, V., Werts, C. & Hofnung, M. (1991). Permissive sites and topology of an outer membrane protein with a reporter epitope

[published erratum appears in J Bacteriol 1991 Dec;173(24):8014]. *J.Bacteriol.* **173**, 262-275.

Charbit, A., Werts, C., Michel, V., Klebba, P. E., Quillardet, P. & Hofnung, M. (1994). A role for residue 151 of LamB in bacteriophage lambda adsorption: possible steric effect of amino acid substitutions. *J Bacteriol* **176**(11), 3204-9.

Chauthaiwale, V. M., Therwath, A. & Deshpande, V. V. (1992). Bacteriophage lambda as a cloning vector. *Microbiol. Rev.* **56**(4), 577-591.

Clement, J. M. & Hofnung, M. (1981). Gene sequence of the lambda receptor, an outer membrane protein of *E. coli* K12. *Cell* **27**(3 Pt 2), 507-14.

Clement, J. M., Lepouce, E., Marchal, C. & Hofnung, M. (1983). Genetic study of a membrane protein: DNA sequence alterations due to 17 *lamB* point mutations affecting adsorption of phage lambda. *EMBO J.* **2**, 77-80.

Crameri, R., Jaussi, R., Menz, G. & Blaser, K. (1994). Display of expression products of cDNA libraries on phage surfaces. A versatile screening system for selective isolation of genes by specific gene-product/ligand interaction. *Eur J Biochem* **2264**(12-24), 53-58.

Crameri, R. & Suter, M. (1993). Display of biologically active proteins on the surface of filamentous phages: a cDNA cloning system for selection of functional gene products linked to the genetic information responsible for their production. *Gene* **137**(1), 69-75.

Crissman, J. W. & Smith, G. P. (1984). Gene-III protein of filamentous phages: evidence for a carboxyl- terminal domain with a role in morphogenesis. *Virology* **132**(2), 445-55.

Davis, N. G., Boeke, J. D. & Model, P. (1985). Fine structure of a membrane anchor domain. *J. Mol. Biol.* **181**, 111-121.

Dennison, S. & Baumberg, S. (1975). Conjugational behaviour of N plasmids in *Escherichia coli* K12. *Molec. Gen. Genet.* **138**, 323-331.

Dotto, G., Enea, V. & Zinder, N. (1981). Functional analysis of bacteriophage f1 intergenic region. *Virology* **114**, 463-473.

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

Duenas, M. & Borrebaeck, C. A. K. (1994). Clonal selection and amplification of phage displayed antibodies by linking antigen recognition and phage replication. *Bio/Tecnol.* **12**, 999-1002.

Dunker, A. K., Ensign, L. D., Arnold, G. E. & Roberts, L. M. (1991a). A model for fd phage penetration and assembly. *FEBS Lett* **292**(1-2), 271-4.

Dunker, A. K., Ensign, L. D., Arnold, G. E. & Roberts, L. M. (1991b). Proposed molten globule intermediates in fd phage penetration and assembly. *FEBS Lett* **292**(1-2), 275-8.

Dunn, I. S. (1995). Assembly of functional bacteriophage lambda virions incorporating c-terminal peptide or protein fusions with the major tail protein. *J. Mol. Biol.* **248**, 497-506.

Dunn, I. S. & Blattner, F. R. (1987). Charons 36 to 40: multi enzyme, high capacity, recombination deficiemnt replacement vectors with polylinkers and polystuffers. *Nucl. Acid. Res.* **15**, 2677-2698.

Elliott, J. & Arber, W. (1978). *Mol.Gen.Genet.* **161**, 1-8.

Emmons, S. W., MacCosham, V. & Baldwin, R. L. (1975). Tandem genetic duplication in phage lambda. III. The frequency of duplication mutants in two derivatives of phage lambda is

indipendent of known recombination systems. *J. Mol. Biol.* **91**, 133.

Endemann, H., Bross, P. & Rasched, I. (1992). The adsorption protein of phage IKe. Localization by deletion mutagenesis of domains involved in infectivity. *Mol Microbiol* **6**(4), 471-8.

Endemann, H., Gailus, V. & Rasched, I. (1993). Interchangeability of the adsorption proteins of bacteriophages Ff and IKe. *J Virol* **67**(6), 3332-7.

Erni, B., Zanolari, B. & Kocher, H. P. (1987). The mannose permease of *Escherichia coli* consists of three different proteins. *J.Biol.Chem.* **262**, 5238-5247.

Evan, G. I., Lewis, G. K., Ramsay, G. & Bishop, J. M. (1985). Isolation of monoclonal antibodies specific for human c-myc proto-oncogene product. *Mol. Cell Biol.* **5**, 3610-3616.

Felici, F., Castagnoli, L., Musacchio, A., Jappelli, R. & Cesareni, G. (1991). Selection of antibody ligands from a large library of oligopeptides expressed on a multivalent exposition vector. *J. Mol. Biol.* **222**, 301-310.

Finlay, B. B., Frost, L. S., Paranchych, W., Parker, J. M. & Hoodes, R. S. (1985). Major antigenic determinants of F and ColB2 pili. *J. Bacteriol.* **163**, 331-335.

Folkhard, W., Leonard, K. R., Malsey, S., Marvin, D. A., Dubochet, J., Engel, A., Achtman, M. & Helmuth, R. (1979). X-ray diffraction and electron microscope studies on the structure of bacterial F pili. *J Mol Biol* **130**(2), 145-60.

Frost, L. S. (1993). Conjugated pili and pilus-specific bacteriophages. In *Bacterial conjugation* (Clewell, D. B., ed.), pp. 189-221. Plenum Press, New York.

Frost, L. S., Finlay, B. B., Opgenorth, A., Paranchych, W. & Lee, J. S. (1985). Characterization and sequence analysis of pilin from F-like plasmids. *J. Bacteriol.* **164**, 1238-1247.

Frost, L. S., Ippen-Ihler, K. & Skurray, R. A. (1994). Analysis of the sequence and gene products of the transfer region of the F sex factor. *Microbiol Rev* **58**(2), 162-210.

Frost, L. S., Lee, J. S., Scraba, D. G. & Paranchych, W. (1986). Two monoclonal antibodies specific for different epitopes within the amino-terminal region of F pilin. *J Bacteriol* **168**(1), 192-8.

Frost, L. S., Paranchych, W. & Willets, N. S. (1984). DNA sequence of the F traALE region that includes the gene for F pilin. *J. Bacteriol.* **160**, 395-401.

Fuchs, P., Breitling, F., Dubel, S., Seehaus, T. & Little, M. (1991). Targeting recombinant antibodies to the surface of *Escherichia coli*: fusion to a peptidoglycan associated lipoprotein. *Biotechnology (N Y)* **9**(12), 1369-72.

Fuerst, C. R. & Bingham, H. (1978). Genetic and physiological characterization of the J gene in bacteriophage lambda. *Virology* **87**, 437-458.

Fullner, K. J., Stephens, K. M. & Nester, E. W. (1994). An essential virulence protein of *Agrobacterium tumefaciens*, VirB4, requires an intact mononucleotide binding domain to function in transfer of T-DNA. *Mol Gen Genet* **245**(6), 704-715.

Furth, M. E., Davis, L. J., Fleurdelys, B. & Scolnic, E. M. (1982). Monoclonal antibodies to the p21 products of the transforming gene of Harvey murine sarcoma virus and the cellular *ras* gene family. *J. Virol.* **43**, 294-304.

Gehring, K., Charbit, A., Brissaud, E. & Hofnung, M. (1987). Bacteriophage lambda receptor site on *Escherichia coli* K-12 LamB protein. *J. Bacteriol.* **169**, 2103-2106.

Gemski, P., Alexeichik, J. A. & Baron, L. S. (1972). Behavior of coliphage lambda in *Shigella flexneri*. *J.Virol.* **10**, 668-674.

Georgiou, G., Potschke, H. L., Stathopoulos, C. & Francisco, J. A. (1993). Practical applications of engineering Gram-negative bacterial cell surfaces. *Tibtech* **11**, 6-10.

Glaser-Wuttke, G., Keppner, J. & Rasched, I. (1989). Pore-forming properties of the adsorption protein of filamentous phage fd. *Biochim Biophys Acta* **985**(3), 239-47.

Glocker, B. & Rasched, I. (1990). Genetic analysis of the transfer region of the IncN plasmid N3. *Res Microbiol* **141**(6), 621-31.

Glucksman, M. J., Bhattacharjee, S. & Makowsli, L. (1992). Three dimensional structure of a cloning vector. X-ray diffraction studies of filamentous bacteriophage M13 at 7 Å resolution. *J. Mol. Biol.* **226**, 455-470.

Grant, R. B., Whiteley, M. H. & Shapley, A. J. (1978). Plasmids of incompatibility group P code for the capacity to propagate bacteriophage IKe. *J Bacteriol* **136**(2), 808-11.

Gray, C. W., Brown, R. S. & Marvin, D. A. (1981). Adsorption complex of filamentous fd virus. *J. Mol. Biol.* **146**, 621-627.

Greenwood, J., Willis, A. E. & Perham, R. N. (1991). Multiple display of foreign peptide on a filamentous bacteriophage. *J. Mol. Biol.* **220**, 821-827.

Hanahan, D. (1983). Studies of transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* **166**, 557.

Harrington, L. C. & Rogerson, A. C. (1990). The F pilus of *Escherichia coli* appears to support stable DNA transfer in the absence of wall-to-wall contact between cells. *J Bacteriol* **172**(12), 7263-4.

Harrison, D. P., Brown, D. T. & Bode, V. C. (1973). The lambda head-tail joining reaction: Purification, properties and structure of biologically active heads and tails. *J. Mol. Biol.* **79**, 437.

Hedges, R. W. (1972). Phenotypic characterization of fi- R factors determining the restriction and modification HspII specificity. *Mol. Gen. Genet.* **115**, 225-233.

Heinemann, J. A. & Spague, G. F. (1989). Bacterial conjugative plasmids mobilize DNA transfer between bacteria and yeast. *Nature* **340**, 205-209.

Helsberg, M., Ebbers, J. & Eichenlaub, R. (1985). Mutations affecting replication and copy number control in plasmid mini-F both reside in the gene for the 29-kD protein. *Plasmid* **14**, 53-63.

Hendrix, R. W. (1971). Identification of proteins coded in phage lambda. In *The bacteriophage lambda.* (Hershey, A. D., ed.), pp. 355. Cold Spring Harbor Laboratory, New York.

Higgins, C. F. (1990). The role of ATP in binding-protein -dependent transport systems. *Res. Microbiol.* **141**, 353-360.

Higgins, C. F., Hiles, I. D., Salmond, G. P. C., Gill, D. R. & Downie, J. A. (1986). A family of related ATP-binding subunits coupled to many distant biological processes in bacteria. *Nature323*(4448-4450).

Hill, D. F. & Petersen, G. B. (1982). Nucleotide sequence of bacteriophage f1 DNA. *J. Virol.* **44**, 32-46.

Hobbs, M. & Mattick, J. S. (1993). Common components in the assembly of type 4 fimbriae, DNA transfer systems, filamentous phage and protein-secretion apparatus: a general system for the formation of surface-associated protein complexes. *Mol Microbiol* **10**(2), 233-243.

Hofnung, M. (1991). Expression of foreign polypeptides at the *Escherichia coli* cell surface. *Methods Cell Biol.* **34**, 77-105.

Hofnung, M. (1995). An intelligent channel (and more). *Science* **2676**(51974), 473-474.

Hofnung, M., Jezierska & Braun-Breton. (1976). *lamB* mutations in *E.coli* K12: growth of lambda host range mutants and effect of nonsense suppressors. *Mol.Gen.Genet.* **145**, 207-213.

Hofnung, M., Lepouce, E. & Braun-Breton, C. (1981). General method for the fine mapping of the *Escherichia coli* K12 *lamB* gene: location of missense mutations affecting bacteriophage lambda adsorption. *J. Bacteriol.* **148**, 853-860.

Hogrefe, H. H., Mullinax, R. L., Lovejoy, A. E., Hay, B. N. & Sorge, J. A. (1993). A bacteriophage lambda vector for the cloning and expresion of immunoglobulin Fab fragments on the surface of filamentous phage. *Gene* **128**, 119-126.

Hokney, R. C. (1994). Recent developments in heterologous protein production in *Escherichia coli. TIBTECH* **12**, 456-463.

Jacob, F. & Wollman, E. L. (1954). Etude genetique d'un bacteriophage tempere' d'*Escherichia coli. Ann. Inst. Pasteur* **87**, 653-673.

Jacobson, A. (1972). Role of F pili in the penetration of bacteriophage fl. *J Virol* **10**(4), 835-43.

Jakes, K. S., Davis, N. G. & Zinder, N. D. (1988). A hybrid toxin from bacteriophage f1 attachment protein and colicin E3 has altered cell receptor specificity. *J. of. Bacteriol.* **170**(9), 4231-4238.

Jespers, L. S., Messens, J. H., De Keyser, A., Eeckhout, D., van den Brande, I., Gansemans, Y. G., Lauwereys, M. J., Vlasuk, G. P. & Stanssens, P. E. (1995). Surface expression and ligand-based selection of cDNAs fused to filamentous phage gene VI. *Bio/Tech* **13**, 378-382.

Johnson, F. D. & Burns, D. L. (1994). Detection and subcellular localization of three Ptl proteins involved in the secretion of pertussis toxin from *Bordetella pertussis. J Bacteriol* **176**(17), 5350-5356.

Katsura, I. (1976b). Morphogenesis of bacteriophage lambda tail, polymorphism in the assembly of the major tail protein. *J.Mol.Biol.* **107**, 307-326.

Katsura, I. (1981). Structure and function of the major tail protein of bacteriophage lambda. *J. Mol. Biol.* **146**, 493-512.

Katsura, I. (1983). Tail assembly and injection. In *Lambda II* (Cold Spring Harbor Laboratory, C. S. H., ed.), pp. 331-346, New York.

Katsura, I. & Kuhl, P. W. (1975). Morphogenesis of the tail of bacteriophage lambda III.Morphogenetic pathway. *J.Mol.Biol.* **91**, 257-273.

Khatoon, H., Iyer, R. V. & Iyer, V. N. (1972). A new filamentous bacteriophage with sex-factor specificity. *Virology* **48**(1), 145-55.

King, S. R. & Richardson, J. P. (1986). Role of homology and pathway specificity for recombination between plasmids and bacteriophage lambda. *Mol.Gen.Genet.* **204**, 141-147.

Klebba, P. E., Hofnung, M. & Charbit, A. (1994). A model of maltodextrin transport through the sugar-specific porin, LamB, based on deletion analysis. *Embo J* **13**(19), 4670-4675.

Kuldau, G. A., De Vos, G., Owen, J., McCaffrey, G. & Zambryski, P. (1990). The virB operon of *Agrobacterium tumefaciens* pTiC58 encodes 11 open reading frames. *Mol Gen Genet* **221**(2), 256-66.

Laemmli, U. K. (1970). *Nature* **227**, 680-685.

LaVallie, E. R., DiBlasio, E. A., Kovacic, S., Grant, K. L., Schendel, P. F. & McCoy, J. M. (1993). A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm. *Biotechnology (N Y)* **11**(2), 187-93.

Lederberg, E. M. (1951). Lysogenicity in *E.coli* K12. *Genetics* **36**, 560.

Lessl, M., Balzer, D., Pansegrau, W. & Lanka, E. (1992). Sequence similarities between the RP4 Tra2 and the Ti VirB region strongly support the conjugation model for T-DNA transfer. *J Biol Chem* **267**(28), 20471-80.

Lessl, M., Balzer, D., Weyrauch, K. & Lanka, E. (1993). The mating pair formation system of plasmid RP4 defined by RSF1010 mobilization and donor-specific phage propagation. *J Bacteriol* **175**(20), 6415-25.

Levengood-Freyermuth, S. K., Click, E. M. & Webster, R. E. (1993). Role of the carboxyl-terminal domain of TolA in protein import and integrity of the outer membrane. *J Bacteriol* **175**(1), 222-8.

Lim, C., Haller, B. & Fuchs, J. (1985). Thioredoxin is the bacterial protein encoded by fip that is required for filamentous bacteriophage f1 assembly. *J. Bacteriol.* **161**, 799-802.

Little, M., Breitling, F., Micheel, B. & Dubel, S. (1994). Surface display of antibodies. *Biotech. Adv.* **12**, 539-555.

Little, M., Fuchs, P., Breitling, F. & Dubel, S. (1993). Bacterial surface presentation of proteins and peptides: an alternative to phage technology? *Trends Biotechnol* **11**(1), 3-5.

Lopez, J. & Webster, R. E. (1983). Morphogenesis of filamentous bacteriophage f1: orientation of extrusion and production of polyphage. *Virology* **127**(1), 177-93.

Lopez, J. & Webster, R. E. (1985a). Assembly site of bacteriophage f1 corresponds to adhesion zones between the inner and outer membranes of the host cell. *J. Bacteriol.* **163**, 1270-1274.

Lopez, J. & Webster, R. E. (1985b). *fipB* and *fipC*: two bacterial loci required for morphogenesis of the filamentous bacteriophage f1. *J. Bacteriol.* **163**, 900-905.

Lowman, H. B., Bass, S. H., Simpson, N. & Wells, J. A. (1991). Selecting high-affinity binding proteins by monovalent phage display. *Biochemistry* **30**(45), 10832-8.

Lowman, H. B. & Wells, J. A. (1993). Affinity maturation of human growth hormone by monovalent phage display. *J Mol Biol* **234**(3), 564-78.

Lu, Z., Murray, K. S., van Cleave, V., LaVallie, E. R., Stahl, M. L. & McCoy, J. M. (1995). Expression of thioredoxin random peptide libraries on the *Escherichia coli* cell surface as functional fusions to flagellin: a system designed for exploring protein-protein interaction. *Bio/Tech.* **13**, 366-372.

Luckey, M. & Nikaido, H. (1980). Specificity of diffusion channels produced by lambda phage receptor protein of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **77**, 167-171.

Manoil, C. & Beckwith, J. (1985). Tn*phoA*: a transposon probe for protein export signals. *Proc. Natl. Acad. Sci. U.S.A.* **82**(23), 8129-8133.

Manoil, C. & Beckwith, J. (1986). A genetic approach to analyzing membrane protein topology. *Science* **233**, 1403-1408.

Marchal, C. & Hofnung, M. (1983). *EMBO J.* **2**, 81.

Maruyama, I. N., Maruyama, H. I. & Brenner, S. (1994). lambdafoo: a lambda phage vector for the expression of foreign proteins. *Proc. Natl. Acad. Sci. USA* **91**, 8273-8277.

Marvin, D., Pigram, W., Wiseman, R., Wachtel, E. & Marvin, F. (1974a). Filamentous bacterial viruses. XII. Molecular architecture of the class I (fd, If1, IKe) virion. *J. Mol. Biol.* **88**, 581-598.

Marvin, D., Wiseman, R. & Wachtel, E. (1974b). Filamentous bacterial viruses. XI. Molecular architecture of the class II (Pf1, Xf) virion. *J. Mol. Biol.* **82**, 121-138.

Marvin, D. A. & Folkhard, W. (1986). Structure of F-pili: reassessment of the symmetry. *J Mol Biol* **191**(2), 299-300.

Marvin, D. A. & Hohn, B. (1969). Filamentous bacterial viruses. *Bacteriol. Rev.* **33**, 172-209.

Marzari, R., Sblattero, D., Righi, M. & Bradbury, A. Extending filamentous phage host range by the grafting of a heterologous receptor binding domain. *submitted.*

McCafferty, J., Griffiths, A. D., Winter, G. & Chiswell, D. J. (1990). Phage antibodies: filamentous phage displaying antibody variable domains. *Nature* **348**(6301), 552-4.

McKay, D. J. & Bode, V. C. (1976a). Events in lambda injection between phage adsorption and DNA entry. *Virology* **72**, 154-166.

McKay, D. J. & Bode, V. C. (1976b). Binding to isolated phage receptors and λ DNA release in vitro. *Virology* **72**, 167-181.

Minenkova, O. O., Ilyichev, A. A., Kishchenko, G. P. & Petrenko, V. A. (1993). Design of specific immunogens using filamentous phage as the carrier. *Gene* **128**(1), 85-8.

Minkley, E. G., Polen, S., Brinton, C. C. & Ippen-Ihiler, K. (1976). Identification of the structural gene for F-pili. *J. Mol. Biol.* **108**, 111-121.

Model, P. & Russel, M. (1988). Filamentous bacteriophage. In *The bacteriophages* (Calendar, R., ed.), Vol. 2, pp. 375-456. Plenum Press, New York.

Mount, D. W. A., Harris, A. W., Fuerst, C. R. & Siminovich, L. (1968). Mutation in bacteriophage lambda affecting particle morphogenic. *Virology* **35**, 134.

Murialdo, H. & Siminovich, L. (1971). The morphogenesis of bacteriophage lambda. III. Identification of genes specifying morphogenetic proteins. In *The bacteriophage lambda.* (Hershey, A. D., ed.), pp. 711. Cold Spring Harbor Laboratory, New York.

Nelson, F. K., Friedman, S. M. & Smith, G. P. (1981). Filamentous phage DNA cloning vectors: a noninfective mutant with a nonpolar deletion in gene III. *Virology* **108**(2), 338-50.

Neuhaus, J.-M., Schindler, H. & Rosenbush, J. P. (1983). The periplasmatic maltose-binding protein modifies the channel forming characteristics of maltoporin. *EMBO J.* **2**, 1987-1991.

Olsen, R. H., Siak, J. & Gray, R. (1974). Characteristic of PRD1, a plasmid-dependent broad-host-range DNA bacteriophage. *J. Virol.* **14**, 689-699.

Parmley, S. F. & Smith, G. P. (1988). Antibody-selectable filamentous fd phage vectors: affinity purification of target genes. *Gene* **73**(2), 305-318.

Peeters, B. P., Peters, R. M., Schoenmakers, J. G. & Konings, R. N. (1985). Nucleotide sequence and genetic organization of the genome of the N- specific filamentous bacteriophage IKe. Comparison with the genome of the F-specific filamentous phages M13, fd and f1. *J Mol Biol* **181**(1), 27-39.

Peeters, B. P., Schoenmakers, J. G. & Konings, R. N. (1986). The gene II proteins of the filamentous phages IKe and Ff (M13, fd and f1) are not functionally interchangeable during viral strand replication. *Nucleic Acids Res* **14**(12), 5067-80.

Pogue-Geile, K. L., Dassarma, S., King, S. R. & Jaskunas, S. R. (1980). Recombination between bacteriophage lambda and plasmid pBR322 in *Escherichia coli*. *J. Bacteriol.* **142**, 992-1003.

Pohlman, R. F., Genetti, H. D. & Winans, S. C. (1994). Common ancestry between IncN conjugal transfer genes and macromolecular export systems of plant and animal pathogens. *Mol Microbiol* **14**(4), 655-668.

Pratt, D., Tzagoloff, H. & Beaudoin, J. (1969). Conditional lethal mutants of the small filamentous coliphage M13. II. Two genes for coat proteins. *Virology* **39**, 42-53.

Rampf, B., Bross, P., Vocke, T. & Rasched, I. (1991). Release of periplasmic proteins induced in *E. coli* by expression of an N-terminal proximal segment of the phage fd gene 3 protein. *FEBS Lett* **280**(1), 27-31.

Randall-Hazelbauer, L. & Schwartz, M. (1973). Isolation of the bacteriophage lambda receptor from *Escherichia coli*. *J.Bacteriol.* **116**, 1436-1446.

Rasched, I. & Oberer, E. (1986). Ff coliphages: structural and functional relationships. *Microbiol Rev* **50**(4), 401-27.

Rebar, E. J. & Pabo, C. O. (1994). Zinc finger phage: affinity selection of fingers with new DNA-binding specificities. *Science* **263**(5147), 671-3.

Roa, M. (1979). Interaction of bacteriophage K10 with its receptor, the lamB protein of *Escherichia coli*. *J.Bacteriol.* **140**, 680-686.

Roa, M. & Scandella, D. (1976). Multiple steps during the interaction between coliphage lambda and its receptor protein in vitro. *Virology* **72**, 182-194.

Roberts, L. M. & Dunker, A. K. (1993). Structural changes accompanying chloroform-induced contraction of the filamentous phage fd. *Biochemistry* **32**(39), 10479-88.

Roessner, C. A. & Ihler, G. M. (1984). Proteinase sensitivity of bacteriophage lambda tail protein gpJ and pH* in complexes with the lambda receptor. *J.Bacteriol.* **157**, 165-170.

Roessner, C. A. & Ihler, G. M. (1987). Sequence of amino acids in lamB responsible for spontaneous ejection of bacteriophage lambda DNA. *J.Mol.Biol.* **195**, 963-966.

Roessner, C. A., Struck, D. K. & Ihler, G. M. (1983). Morphology of complexes formed between bacteriophage lambda and structures containing the lambda receptor. *J.Bacteriol.* **153**, 1528-1534.

Rost, B. & Sander, C. (1993). Improved prediction of protein secondary structure by use of sequence profiles and neural networks. *Proc. Natl. Acad. Sci. USA* **90**, 7558-7562.

Rost, B. & Sander, C. (1994). Conservation and prediction of solvent accessibility in protein families. *Proteins* **20**, 216-226.

Russel, M. (1991). Filamentous phage assembly. *Mol. Microbiol.* **5**(7), 1607-1613.

Russel, M. (1992). Interchangeability of related proteins and autonomy of function: the morphogenetic proteins of filamentous phage f1 and IKe cannot replace one another. *J. Mol. Biol.* **227**, 453-462.

Russel, M. (1993). Protein-protein interactions during filamentous phage assembly. *J. Mol. Biol.* **231**, 689-697.

Russel, M. (1994). Phage assembly: a paradigm for bacterial virulence factor export? *Science* **2652**(5172), 612-614.

Russel, M. & Model, P. (1983). A bacterial gene, *fip*, required for filamentous bacteriophage f1 assembly. *J. Bacteriol.* **154**, 1064-1076.

Russel, M. & Model, P. (1985). Thioredoxin is required for filamentous phage assembly. *Proc. Natl. Acad. Sci. USA* **82**, 29-33.

Russel, M., Whirlow, H., Sun, T. P. & Webster, R. E. (1988). Low-frequency infection of F^- bacteria by transducing particles of filamentous bacteriophages. *J Bacteriol* **170**(11), 5312-6.

Sambrook, J. & Maniatis, T. (1989). *Molecular cloning. A laboratory manual.* Cold Spring Harbor (Press., C. S. H. L., Ed.), New York.

Sander, C. & Schneider, R. (1991). Database of homology-derived protein structures and the structural meaning of sequence alignment. *Proteins* **9**(1), 56-68.

Sblattero, D. (1995). Elaborazione del gene III del fago fd per la selezione di librerie di anticorpi. Bachelor Degree Thesis, University of Trieste.

Scandella, D. & Arber, W. (1974). An *Escherichia coli* mutant which inhibits the injection of phage lambda DNA. *Virology* **58**, 504-513.

Scandella, D. & Arber, W. (1976). Phage lambda injection into *Escherichia coli* pel— mutants is restored by mutations in phage genes V or H. *Virology* **69**, 206-215.

Schirmer, T., Keller, T. A., Wang, Y. F. & Rosenbusch, J. P. (1995). Structural basis for sugar translocation through maltoporin channels at 3.1 A resolution. *Science* **267**, 512-514.

Schwartz & Le Minor, L. (1975a). *J. Virol.* **15**, 679-685.

Schwartz, M. (1975). Reversible interaction between coliphage lambda and its receptor protein. *J.Mol.Biol.* **99**, 185-.

Schwartz, M. & Le Minor, L. (1975b). Occurrence of the bacteriophage lambda receptor in some enterobacteriaceae. *J.Virol.* **15**, 679-685.

Scott, J. K. & Smith, G. P. (1990). Searching for peptide ligands with an epitope library. *Science* **249**, 386-390.

Shen, P. & Huang, H. V. (1986). Homologous recombination in *Escherichia coli*: dependence on substrate length and homology. *Genetics* **112**, 441-457.

Shirasu, K. & Kado, C. I. (1993a). Membrane location of the Ti plasmid VirB proteins involved in the biosynthesis of a pilin-like conjugative structure on *Agrobacterium tumefaciens*. *FEMS Microbiol Lett* **111**(2-3), 287-94.

Shirasu, K. & Kado, C. I. (1993b). The virB operon of the *Agrobacterium tumefaciens* virulence regulon has sequence similarities to B, C and D open reading frames downstream of the pertussis toxin-operon and to the DNA transfer-operons of broad- host-range conjugative plasmids. *Nucleic Acids Res* **21**(2), 353-4.

Shirasu, K., Koukolikova-Nicola, Z., Hohn, B. & Kado, C. I. (1994). An inner-membrane-associated virulence protein essential for T-DNA transfer from *Agrobacterium tumefaciens* to plants exhibits ATPase activity and similarities to conjugative transfer genes. *Mol Microbiol* **11**(3), 581-588.

Shirasu, K., Morel, P. & Kado, C. I. (1990). Characterization of the virB operon of an *Agrobacterium tumefaciens* Ti plasmid: nucleotide sequence and protein analysis. *Mol Microbiol* **4**(7), 1153-63.

Sigal, I. S., Gibbs, J. S. & Scolnik, E. M. (1986). Identification of effector residues and a neutralizing epitope of H-ras-encoded p21. *Proc. Natl. Acad. Sci. USA* **83**(4725-4729).

Smilowitz, H. (1974). Bacteriophage f1 infection and colicin tolerance. *J. Virol.* **13**, 100-106.

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

Stengele, I., Bross, P., Garces, X., Giray, J. & Rasched, I. (1990). Dissection of functional domains in phage fd adsorption protein. Discrimination between attachment and penetration sites. *J Mol Biol* **212**(1), 143-9.

Sun, T. P. & Webster, R. E. (1986). *fii*, a bacterial locus required for filamentous phage infection and its relation to colicin-tolerant tolA and tolB. *J. Bacteriol.* **165**, 107-115.

Sun, T. P. & Webster, R. E. (1987). Nucleotide sequence of a gene cluster involved in entry of E-colicin and single-stranded DNA of infecting filamentous bacteriophage into *Escherichia coli*. *J. Bacteriol.* **169**, 2667-2674.

Thatte, V., Bradley, D. E. & Iyer, V. N. (1985). N conjugative transfer system of plasmid pCU1. *J Bacteriol* **163**(3), 1229-36.

Thirion, J. P. & Hofnung, M. (1972). On some aspects of phage lambda resistance in *E. coli* K12. *Genetics* **71**, 207.

Thompson, D. V., Melchers, L. S., Idler, K. B., Schilperoort, R. A. & Hooykaas, P. J. (1988). Analysis of the complete nucleotide sequence of the *Agrobacterium tumefaciens* virB operon. *Nucleic Acids Res* **16**(10), 4621-36.

Thorstenson, Y. R., Kuldau, G. A. & Zambryski, P. C. (1993). Subcellular localization of seven VirB proteins of *Agrobacterium tumefaciens*: implications for the formation of a T-DNA transport structure. *J Bacteriol* **175**(16), 5233-41.

Walker, J. E., Sarste, M., Runswick, M. J. & Gay, N. J. (1982). Distantly related sequences in the a- and b-subunits of ATP-synthase, myosin, kinases, and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* **1**, 945-951.

Wandersman, C., Schwartz, M. & Ferenci, T. (1979). *E. coli* mutants impaired in maltodextrin transport. *J. Bacteriol.* **140**, 1.

Ward, J. E., Akiyoshi, D. E., Regier, D., Datta, A., Gordon, M. P. & Nester, E. W. (1988). Characterization of the virB operon from an *Agrobacterium tumefaciens* Ti plasmid. *J Biol Chem* **263**(12), 5804-14.

Ward, J. E., Akiyoshi, D. E., Regier, D., Datta, A., Gordon, M. P. & Nester, E. W. (1990a). Correction: characterization of the virB operon from *Agrobacterium tumefaciens* Ti plasmid. *J Biol Chem* **265**(8), 4768.

Ward, J. E., Jr., Dale, E. M., Nester, E. W. & Binns, A. N. (1990b). Identification of a virB10 protein aggregate in the inner membrane of Agrobacterium tumefaciens. *J Bacteriol* **172**(9), 5200-10.

Watanabe, T., Nishida, H., Ogata, C., Arai, T. & Sato, S. (1964). Episome-mediated transfer of drug resistance in *Enterobacteriaceae*. VII. Two types of naturally occurring R factors. *J. Bacteriol.* **88**, 716-726.

Webster, R., Grant, R. & Hamilton, L. (1981). Orientation of the DNA in the filamentous bacteriophage f1. *J. Mol. Biol.* **152**, 357-374.

Webster, R. E. (1991). The *tol* gene products and the import of macromolecules into *Escherichia coli*. *Mol Microbiol* **5**(5), 1005-11.

Weigle, J. (1966). Assembly of phage lambda in vitro. *Proc. Natl. Acad. Sci.* **55**, 1462.

Weiss, A. A., Johnson, F. D. & Burns, D. L. (1993). Molecular characterization of an operon required for pertussis toxin secretion. *Proc Natl Acad Sci U S A* **90**(7), 2970-4.

Werge, T., Bradbury, A., Di Luzio, A. & Cattaneo, A. (1992). A recombinant cell line, expressing a form of the anti p1ras antibody, Y13-259, which binds protein A and may be produced as ascites. *Oncogene* **7**, 1033-1035.

Werts, C., Michel, V., Hofnung, M. & Charbit, A. (1994). Adsorption of bacteriophage lambda on the lamB protein of *Escherichia coli* K-12: point mutations in gene *J* of lambda responsible for extended host range. *J. of Bacteriol.* **174**(4), 941-947.

Willetts, N. & Maule, J. (1986). Specificities of IncF plasmid conjugation genes. *Genet Res* **47**(1), 1-11.

Winans, S. C. & Walker, G. C. (1985). Conjugal transfer system of the IncN plasmid pKM101. *J Bacteriol* **161**(1), 402-10.

Zgaga, V., Medic, M. & Salaj-Smic, E. (1973). Infection of *Escherichia coli* envelope membrane complex with lambda phage: adsorption and penetration. *J.Mol.Biol.* **79**, 697-.

Zimmermann, R., Watts, C. & Wickner, W. (1982). The biosynthesis of membrane-bound M13 coat protein. *J. Biol. Chem.* **257**, 6529-6536.

Zinder, N. (1973). Resistance to colicins E3 and K induced by infection with bacteriophage f1. *Proc. Natl. Acad. Sci. USA* **70**, 3160-3164.