INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction..

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

University of Alberta

CONJUGAL AND GENETIC ANALYSIS OF THE INCHI1 PLASMID, R27

by

Craig K. Sherburne 

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Doctor of Philosophy

In

Bacteriology

Department of Medical Microbiology and Immunology

EDMONTON, ALBERTA

Fall, 2000

University of Alberta

Library Release Form

Name of Author: Craig K Sherburne

Title of Thesis: Conjugal and genetic analysis of the IncHI1 plasmid, R27.

Degree: Doctor of Philosophy

Year this Degree Granted: 2000

Craig Sherburne
3680-33 Street
Edmonton,Alberta
Canada

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled *Conjugal and Genetic Analysis of the IncHI1 Plasmid R27* in partial fulfillment of the requirements for the degree of *Doctor of Philosophy* in Bacteriology.

Dr. Diane E. Taylor (Supervisor)

Dr. John Elliott

Dr. Laura Frost

Dr. Brenda Leskiw

Dr. Kenneth Sanderson
(External Examiner)

5 September 2000
Date

ABSTRACT

The complete DNA sequence of R27 has been compiled and analyzed. The 180,361bp plasmid contains 210 open reading frames, 70 of which have been previously identified or contain significant homology to other plasmid or prokaryotic open reading frames. Two transfer regions separated by over 64kb are observed, with the gene order and individual genes in transfer region 2 (Tra2) displaying significant homology to the F plasmid transfer region

The frequency of transfer and ability to form mating aggregates, using isogenic *S. typhimurium* LPS mutant recipients and donors, were assessed for the IncHI1 plasmids R27 and pDT2454, and the IncHI2 plasmid R478. It was observed that only a specific truncation of the outer core, by mutation of the *rfaF* LPS biosynthesis gene, increased conjugal transfer. Stabilization and initiation of conjugal transfer is, in most cases, negatively affected by LPS truncation.

The major component of the R27 conjugal pilus was identified as a 7.6kDa peptide, likely generated by the *trhA* gene. Although visualized in a cell-free radiolabeled expression system, the *trhA* gene product is resistant to denaturation and protein staining in SDS-PAGE analysis. No mutations of the *trhA* gene in R27 through antibiotic cassette insertion were obtained. It is possible that insertion events in *trhA* destabilize R27 by an unknown mechanism.

The gene responsible for entry exclusion in R27 has been identified. R27 transfer has been shown to significantly decrease into recipient cells containing ORF018, encoding a 14.1 kDa protein. The protein contains no significant homology to entry

exclusion proteins of other systems, although it does contain a strong N-terminal transmembrane sequence theoretically sequestering it to the bacterial membrane.

R27 contains a large number of elements in common with the F plasmid family, although it encodes a pilin with strong similarity to the IncP plasmids, a separation of the transfer regions similar to the IncP plasmids, and an entry exclusion gene which shows no homology to either system. These results demonstrate the apparent hybrid nature of R27, with its current state appearing to be the result of genetic exchanges between ancestors of many different plasmid systems forming the unique IncHI1 system observed today.

ACKNOWLEDGMENTS

I would like to thank Dr Fred Blattner, Erik Grotbeck, Debra Rose and Valerie Burland for generation of the complete R27 sequence (University of Wisconsin) in Section V. Trevor Lawley for transposon analysis, and Matt Gilmour for their contributions to the sections entitled Transposable elements and Temperature sensitive conjugation respectively, included in Section V. Michelle Rooker for aid in DNA preparation and purification of R27 used in Section V and sequence assembly of the Tra2 region initiated by Dr. Peter Newnham, presented in Section IV. MALDI TOF spectroscopy was performed by the Department of Chemistry, University of Alberta.

The R27 sequence was generated as part of the Bacterial Pathogen Genome Initiative, funded by NIH grant # AI44387. Other work was funded by the Medical Research Council of Canada. I acknowledge support from the Faculty of Medicine 75th Anniversary Award, 1994-95 and the Province of Alberta Graduate Student Fellowship 1995-96.

My thanks to Dr Diane Taylor for her support, guidance and understanding throughout this project and Dr John Elliott and Dr Laura Frost for their part in the supervision of my doctoral thesis. As well, Dr Ken Whelan, for teaching me of the art of molecular biology as much as he did the science and Michelle Rooker for her aid throughout the years.

The completion of this work is a testament to the support I received from friends and family. My father, Richard Sherburne, for his encouragement and constant support throughout the difficult times, without his strength and example none of what I have

accomplished would be possible. I am indebted to my mother Eileen, brother Christopher, and sister Jeanette who were always a source of laughter and joy. Roy Thomas, who has stood by me, as much literally as figuratively, through my life and put up with as much from me as any other. Finally to Joan Munro, my wife to be, who has weathered the worst of the storm that is a doctorate, and still stands strong.

TABLE OF CONTENTS

LIST OF TABLES

Table VI-3 Transfer of pDT2454 into recipients containing cloned genes from the *Xba*D region. 134

LIST OF FIGURES

ABBREVIATIONS

ATP	adenosine triphosphate
BHI	brain heart infusion
BSA	bovine serum albumin
dATP	deoxyadenine triphosphate
dCTP	deoxycytosine triphosphate
dGTP	deoxyguanosine triphosphate
dNTP	deoxynucleoside triphosphate
dTTP	deoxythymidine triphosphate
DNA	deoxyribonucleic acid
DTT	dithiothreitol
EDTA	ethylene-diaminetetra-acetic acid
IPTG	isopropylthiogalactoside
LB	Luria broth
PB	phosphate buffer
PCR	polymerase chain reaction
PEA	o-phosphorylethanolamine
PEG	polyethelyne glycol
RNA	ribonucleic acid
TAE	Tris-acetate EDTA
TBS	Tris buffered saline
UV	ultra violet

Chapter I Introduction

1. Plasmids

Bacterial plasmids were initially identified as extra-chromosomal genetic elements in the early 1960s (94). Building on previous observations of the "fertility factor", small, circular, double-stranded DNA molecules were shown to transfer from a host bacterium to a recipient cell, (47). Capitalizing on the mobilization of portions of the chromosome by F plasmid integration, mapping of the *E. coli* chromosome was initiated and thus advanced our knowledge to such an extent that *E. coli* is the organism of choice for most molecular biological applications.

Plasmids have been found throughout the bacterial world and nearly all bacterial genera have been shown to harbour them (32). The first instance of plasmid-mediated antibiotic resistance was observed in Japan with *Shigella* in 1957 (4). Initially plasmids were found to encode a number of metabolic functions benefiting the host cell. Since then the study of plasmids has increased dramatically as more plasmids are discovered to be involved in antibiotic resistance transfer between pathogenic bacteria (168).

2. Plasmid incompatibility

It was found that in the absence of selective pressure, related plasmids were unable to be stably maintained in the same bacterial host (reviewed in 28, 57). This effect, termed incompatibility, was exploited to classify plasmids, with those unable to stably coexist classified as members of the same incompatibility group. The diversity of plasmids is remarkable, with over 25 different incompatibility groups identified to date in

Enterobacteriaceae alone (28, reviewed in 147). A further 13 incompatibility groups have been observed in the *Pseudomonas* family of bacteria (147). Observation and analysis of incompatibility led to a greater understanding of plasmid maintenance, and the plasmid-encoded machinery involved. Two causes of incompatibility were identified within plasmids. The first arises from the plasmid ensuring that daughter cells contain the plasmid following bacterial replication (partitioning), and the second arises from the plasmid ensuring that it is efficiently replicated with as little metabolic burden on the host cell as possible (replication) (104). Plasmid incompatibility arises as a side effect of related partitioning or replication systems interfering with each other's normal action (reviewed in 104).

Bacterial plasmid replication is generally controlled by plasmid-encoded components initiating the replication event at a site-specific origin (*ori*) (73). The plasmid relies heavily on the host cell for DNA replication functions, in contrast replication initiation and therefore plasmid copy number are tightly regulated by plasmid genes. Based upon the method of regulation of replication initiation (103), plasmid replication control can be classified into one of two systems.

The first method, a negative feedback control system, relies on regulation of initiation events through production of a negative regulator. Examples include the ColE1 plasmids (RSF1030, pBR322, pMB1) and the IncFII plasmids (R1, NR1, R6-5) (73). In the ColE1 system replication is initiated by a 700bp RNA primer (RNAII), which hybridizes immediately upstream of the origin of replication. Hybridization of RNAII to the DNA near the origin is dependant on proper secondary structure being formed within the RNA molecule and subsequent specific cleavage by RNaseH. Control of plasmid

replication occurs through the production of the RNAI molecule, a 108bp antisense RNA to RNAII, which binds to the 5' region of RNAII, disrupting its secondary structure and preventing replication initiation (159). The plasmid-encoded protein Rom (RNA I modulator) is observed to increase the stability of the RNAI:RNAII complexes 100 fold (158). Plasmid replication, and therefore plasmid copy number, is controlled by a combination of the ratio of RNAI to RNAII, the rate of synthesis of RNAI, RNAI stability in the bacterial host, and the stability of the RNAI:RNAII complex (73).

A second method of replication control is based upon binding of a replication initiation protein to a series of oligonucleotide repeats, called iterons (73, 103). It is believed that the iterons serve to titrate the initiation protein, effectively sequestering it and preventing initiation of plasmid replication (73). The binding of the replication protein to iterons flanking the replication gene can also act as an auto-regulator, preventing transcription of the gene (25, 73, 101). Deletion of iterons is found to increase plasmid copy number, whereas increasing the number of iterons in a plasmid decreases copy number (160). Though consistent with the titration of the replication initiator protein by iterons, increasing concentrations of the replication protein supplied *in trans* does not increase copy number of the plasmid (160).

A second observed property of initiator proteins, ie. that they can couple iterons by forming a bridge between them, formed the basis for a second model of iteron based control, in which steric hinderance of coupled iterons reduces initiation events (103). Recent observations of the P1 plasmid system and its replication protein, RepA, have demonstrated that plasmid replication clears the *repA* promoter of bound RepA, allowing

repA gene transcription (97). It is proposed that replication-induced transcription is necessary to ensure RepA availability after the replication event (97).

3. The conjugation process

Many plasmids are able to actively transfer their DNA from one bacterial cell to another, a process termed conjugation. Acting as a relatively efficient means of genetic exchange in a population of cells, bacterial conjugation is of major evolutionary, ecological and medical importance. Although spontaneous mutations within the bacterial cell can result in antibiotic resistance, the rate of multiple antibiotic resistance in a bacterial cell has been estimated to be 10^{-16} events per cell (168). In contrast the transfer rate of conjugative plasmids can reach frequencies of 10^{-1} transconjugants per organism (168). Therefore not only are plasmids able to confer antibiotic resistance in bacteria, more importantly conjugative plasmids are able to mediate horizontal transfer of various factors, including multiple-antibiotic resistance in a population of cells (168).

Conjugation in bacteria carrying the F factor has been separated into five distinct stages; (I) pilus binding to a receptor, (II) formation of wall-to-wall contact, (III) aggregate stabilization, (IV) DNA transfer, and (V) aggregate disassociation (3). There is no evidence of the existence of pilin in Gram-positive bacteria with cell to cell contact mediated in some genera by plasmid-encoded aggregation factors (39). Almost no homology is observed between the transfer genes of the Gram-positive and Gram-negative plasmid systems. Therefore there is no indication that the stages for conjugation are similar between the systems.

Following attachment of the pilus to a recipient cell and cell contact being made, the mating aggregate is stabilized. In the F plasmid system the process is mediated by plasmid-encoded, membrane bound genes, TraG and TraN (8). Recent evidence indicates that recognition of the recipient cell is mediated by these two genes (8). Therefore, in the F-system, the pilus acts as a non-specific adhesion factor, mediating the close association necessary for TraG and TraN to recognize an acceptable recipient cell and initiate aggregate stabilization (8).

Once contact is made and the aggregate stabilized, the plasmid DNA undergoes processing and replication resulting in a single strand of DNA being transferred from the donor to the recipient cell. A single sequence-specific nick is made in the plasmid mediated by plasmid-encoded proteins. The strand is then displaced into the recipient cell in a 5' to 3' direction using the replicative machinery of the host cells, termed DCDS (Donor Conjugative DNA Synthesis) (171). A number of proteins are encoded within the stretch of DNA which initially enters the recipient cell (the leading strand). These genes, which show little homology among plasmid systems, encode proteins aiding the establishment of the plasmid in the recipient strand such as the F plasmid *ssb*, a single-stranded DNA-binding protein (75). An excellent example of a leading strand gene is the *ardA* anti-restriction locus. Found in the IncB, FV, I1, K, and N plasmids, *ardA* is found to alleviate typeI DNA restriction preventing plasmid restriction upon entering the recipient cell (36).

The organization of the genes responsible for the mobilization of the plasmid during conjugation can lead to a broad categorization of Gram-negative plasmid systems. In the majority of plasmid conjugation systems the genes responsible for both pilus

assembly/mating stabilization and conjugal mobilization are found in a contiguous stretch of genes named the transfer region. Other plasmids, such as the IncP, and Ti plasmids, have two separate and distinct transfer regions within the plasmid, termed Tra1 and Tra2. Tra1 represents the region in which conjugal transfer of the plasmid is initiated (*oriT*), containing genes required for conjugal DNA processing and the leading strand genes. Tra2 contains the genes necessary for pilin biogenesis and mating aggregate stabilization (111).

4. Conjugative pili

Cell to cell contact is required for conjugal transfer (33), yet the pilus plays an initial role in bringing the cells together and is necessary for efficient DNA transfer to the recipient cell (9). The F pilus, the best studied of the plasmid-encoded pili, is formed by a helical array of subunits, or pilin. The helix contains five subunits per turn generating a mature pilus of 8nm diameter (2nm inner core) ranging from 1 to 20μm in length (112). The major roles of the pilus are in the recognition of recipient cells and the ability to bind to various receptors in the outer membrane. After attachment to the recipient cell, the pilus is thought to depolymerize, bringing the two cells into cell to cell contact leading to the establishment of the mating bridge (29). Although there is some evidence for transfer of DNA through the pilus in the absence of close contact (61, 109), the majority of the evidence supports the necessity for cell to cell contact in conjugal transfer (173). Once contact is established the fully polymerized pilus is not necessary, as addition of SDS at concentrations that dissociate the pilus does not inhibit DNA transfer once the cells are joined as a mating aggregate (2).

In the F system the pilin protein is produced as 120 amino acids, inserted in the membrane by a pilin specific chaperone, TraQ, and processed by the host leader peptidase to 70 amino acids (93). Following acetylation the pilin subunits are stored in the inner membrane as a large pool, until assembled through the action of the pilin assembly genes (8, 95). A total of 15 of the 23 genes found in the F transfer region have been identified as being involved in the processing and polymerization of the pilus (49), an observation consistent with other plasmid systems (111, 124, 156). It is a measure of the necessity of the pilus in the efficient transfer of the plasmid that it has survived the evolutionary pressure towards a more efficient, pilus-free transfer system.

The pilin products of the IncP plasmids are closely related to the TrbC gene of the *Agrobacterium tumefaciens* Ti plasmid and the VirB2 genes in the Ti plasmid virulence region. Recently it has been observed that the pilin product of these plasmids undergo a number of processing events, leading to the eventual cyclization of the peptide rendering it N-terminally blocked for amino acid analysis by Edman degradation (40, 59). A host-encoded C-terminal truncation of residues is observed in the IncP and Ti plasmid pilin proteins (59). These related pilin proteins undergo further processing, with the removal of four C-terminal amino acids by the plasmid encoded TraF protease, followed by cyclization of the peptide (40).

5. Conjugation system relationships

Recent studies have identified homology between the conjugal transfer system of the Ti plasmid (*trb*) and the *icm/dot* regions of *Legionella pneumophila* responsible for human macrophage killing and intracellular growth of the bacteria. Both systems are

able to mobilize the plasmid RSF1010 (164), an IncQ plasmid encoding the genes necessary for conjugative DNA mobilization but not for conjugal-pair formation. Mobilization of RSF1010 is an effective measure of the presence of conjugal-formation genes in the donor cell (135). The *icm/dot* genes are not normally involved in DNA transfer from one cell to another, instead they play a role in transfer of inhibitor macromolecules from *L. pneumophila* into macrophages to subvert the endocytic pathway (164). Therefore it appears that the conjugal DNA processing genes present in RSF1010 are able to utilize the *icm/dot* protein transfer system to transfer plasmid DNA, implicating the systems as arising from a common ancestor. Further evidence of common ancestry among these virulence regions and conjugation systems has been reported in the IncI plasmid R64 and Colb-P9 (Figure I-1A). The complete sequence of the transfer region has been compiled, and 10 genes necessary for conjugal transfer of the plasmid have been shown to have significant homology to *L. pneumophila icm/dot* genes (76).

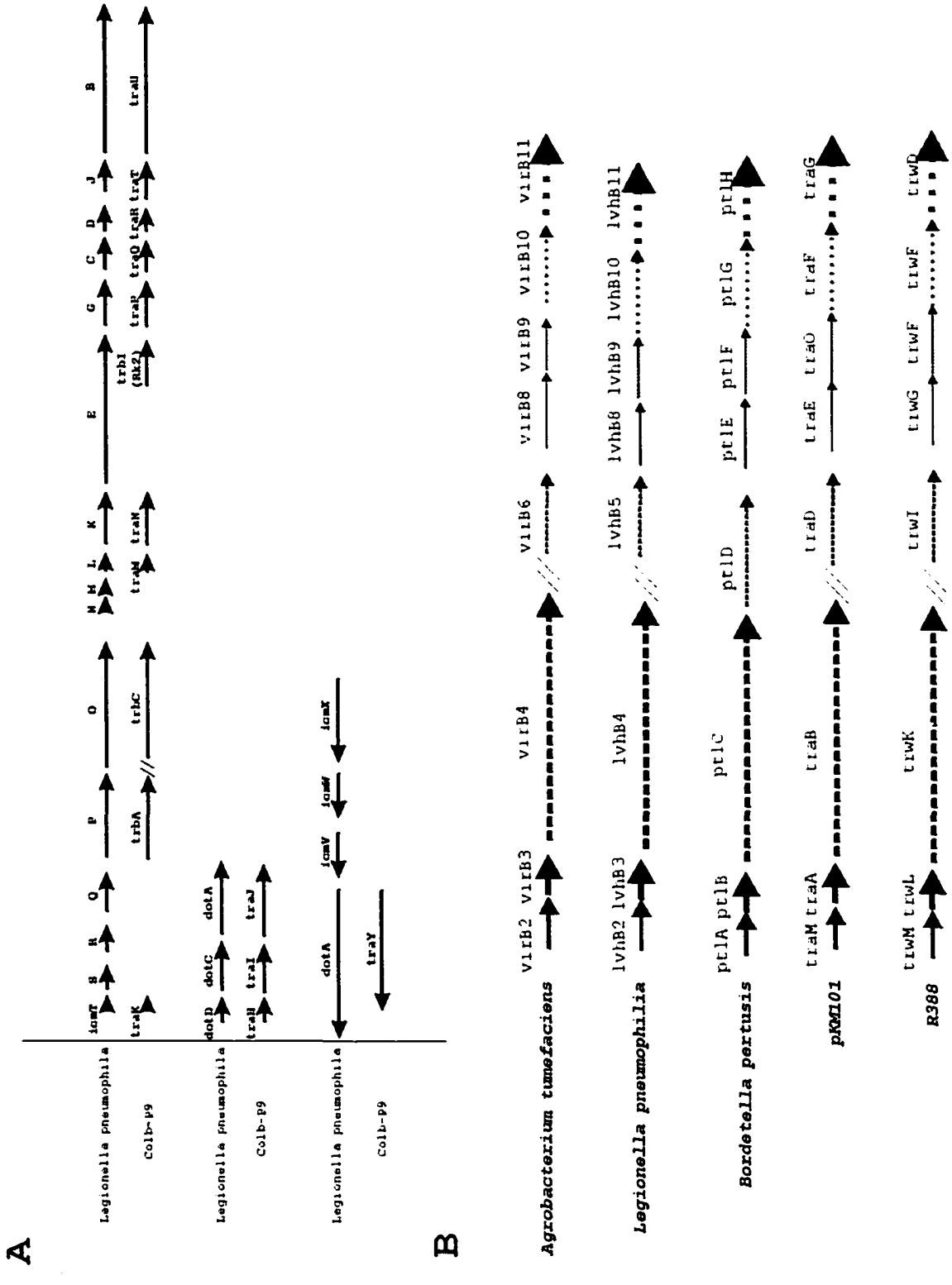
There also appears to be a relationship among the Ti plasmid *trb* and *vir* genes, the *icm/dot* region of *L. pneumophila*, a second secretion system in *L. pneumophila* required for intracellular growth (*lvh*), *Bordetella pertussis* pertussis toxin (*ptl*) and the *Helicobacter pylori cag* pathogenicity island (Figure I-1B)(136). These systems have the ability to transport proteins out of bacteria, and are considered to belong to the type IV secretion system (41, 136). Recent comparisons between the *virB* system of Ti plasmids and the *trb* system of RP4 has identified a single gene, the *trbI* of RP4 and *virB10* of the Ti plasmids, common among all type IV secretion systems (84, 136).

Figure I-1 Relationships between Type IV secretion systems

Schematic comparison of type IV secretion systems and plasmid conjugal DNA transport systems.

(A) Comparison of the *icm/dot* (intracellular multiplication/defective for organelle trafficking) region of *Legionella pneumophila*, needed for macrophage killing and intracellular multiplication, and the conjugation related proteins of the IncI plasmid colb-P9 (137).

(B) Comparison of homologous genes observed throughout the *Agrobacterium tumefaciens* Ti plasmids vir region, the lvh secretion system of *L. pneumophila*, the pertusis toxin export system (*ptl*) of *Bordetella pertussis*, the *tra* conjugative system of the IncN plasmid pKM101 and the *tra* conjugative system of the IncW plasmid R388 (136).



Both the TrbI and VirB10 proteins are necessary for conjugal transfer of DNA to the recipient bacterium in the RP4 *trb* system and to the host plant in the Ti *virB* system (14, 60). It is believed that *virB10* plays a role in assembling the mating pore complex, a necessary stage in the transfer of DNA across the membrane. The role of *trbI* homologues in other type IV secretion systems is unknown. The presence of a *virB10/trbI* homologue in all type IV secretion systems does not necessarily indicate similar function across all systems. In the Ti system VirB10 anchors other VirB proteins, including the VirB7 and VirB9 proteins, two proteins without homologues in the RP4 *trb*, or other type IV secretion systems (84). Therefore exact duplication of function for *trbI* is unlikely.

In addition to appearing to be related to type IV secretion systems, a strong relationship is observed between the Ti and IncP conjugal transfer systems (5, 82). Most of the predicted proteins encoded within the Tra1 and Tra2 regions of the Ti plasmid showed 30-50% amino acid identity to the IncPα prototype plasmid RP4 (5). Furthermore, the conjugal transfer genes of the Ti plasmids showed significant homology to the *vir* (virulence) genes located elsewhere in the Ti plasmid. The *vir* genes, responsible for the transfer of plasmid DNA into host plants, are required for the inter-kingdom transfer of genetic information (174). The similarity in the systems, although possibly indicating a common ancestor, may be indicative of the limited options available in transport of macromolecules. In effect, the transport of DNA from a Gram-negative donor to recipient requires the transfer of extremely large molecules (DNA) across 4 surface barriers which is a complex process.

6. Entry exclusion.

The presence of a plasmid in a recipient cell limits the establishment of a related incoming plasmid (80). This phenomenon has been found to be determined by two separate processes: reduced stability of the plasmid in the recipient (incompatibility) and the inhibition of the incoming plasmid from entering the cell (entry exclusion) (105). Plasmids that exhibit entry exclusion usually belong to the same incompatibility group but not all plasmids within an incompatibility group display the same degree of entry exclusion (58).

The entry exclusion system of the F and F-like plasmids is the best studied (42). It consists of a two protein system encoded by the *traS* and *traT* genes (8). The TraT protein, present in the outer membrane of the bacterial host, inhibits mating pair stabilization, but does not affect pilus binding. The TraS protein, present in the inner membrane, blocks a plasmid-specific signal necessary for initiation of DNA synthesis and transport (8). There is evidence that a two protein entry exclusion system exists in the IncI plasmid R64, though both proteins appear to be transcribed as a single gene, with the two proteins resulting from an inframe reinitiation of translation (51).

There are also systems in which entry exclusion is mediated by a single gene product, usually localized to the inner membrane. These include the *eex* gene of the IncN plasmid pKM101 (118), the *trbK* gene of the IncP plasmid RP4 (58) and the *trbK* gene of the Ti plasmids (84). The genes all encode small peptides of approximately 6-8kDa, containing a predicted N-terminal transmembrane helix, although only the TrbK protein of RP4 has been localized to the cytoplasmic membrane (58, 84, 118). There is little to no homology among the entry exclusion genes of these plasmids, with considerable

sequence divergence observed between plasmid families, even though function is conserved (83). With most entry exclusion genes, the presence of the gene is not necessary for plasmid transfer (8, 58, 118), but likely plays a role in stability of the plasmid in the environment. There is evidence that the entry exclusion gene of the F-mobilizable plasmid ColE1, *mbeD*, is necessary for transfer, suggesting it is a multi-functional protein (175).

7. Bacterial lipopolysaccharide.

Gram-negative bacteria contain a double membrane separated by a thin layer of peptidoglycan. The outer membrane is asymmetric with the inner monolayer composed primarily of glycerophospholipids and the outer monolayer composed primarily of lipopolysaccharide (LPS) (as reviewed in121). The LPS has been implicated as an important barrier to large (>650Da) hydrophobic molecules, with the outer membrane permeability to these molecules shown to be 1-2% that of the simple phospholipid bilayers (117). LPS, in combination with the outer phospholipid monolayer, forms a gel phase bilayer, separated by a relatively large fluid space, with LPS molecules forming tightly packed carbohydrate chains observed to be more dense than even phospholipid bilayers (143).

Divalent cations are found to be important in stability and outer membrane permeability and are localized to the inner core, proximal to the membrane monolayer, the 3-deoxy-D-manno-octulosonic acid (KDO) subunits present in both *E. coli* and *Salmonella enterica* sv. Typhimurium (*Salmonella typhimurium*) LPS, and the phosphate substituations on the initial heptose sugars (Figure I-2) (143). Recent evidence suggests

that cations play a role in negative-charge neutralization, allowing tighter packing of the LPS molecules, rather than cross-linking of the individual LPS molecules (143).

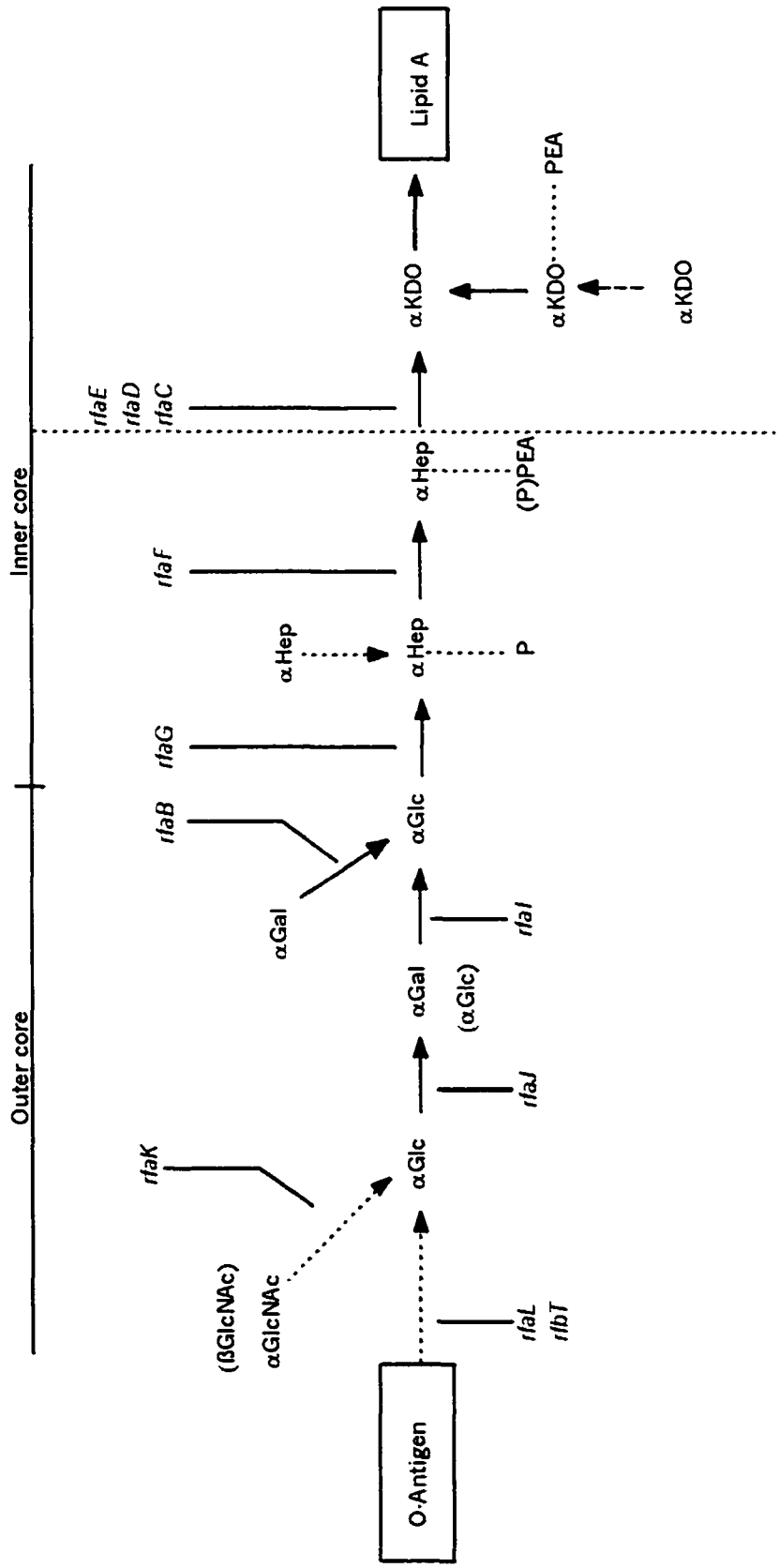
The structural features of bacterial lipopolysaccharide are well established (120) with an excellent understanding of the genetic and molecular mechanisms for its biosynthesis in *E.coli* and *S. typhimurium* (62). Bacterial LPS consists of a lipid A "anchor" to the outer membrane, a short core oligosaccharide which is in turn divided into an inner and outer core, and the O-polysaccharide. LPS is variable, with variation mostly observed in the O-polysaccharide which is highly polymorphic (108).

Lipid A, the hydrophobic portion of LPS, forms the outer leaflet of the outer membrane and is responsible for the observed endotoxic properties of the LPS (120). A minimal LPS structure has been observed in the laboratory, consisting of the lipid A molecule glycosylated with two KDO residues. The lipid A, formed independently of the core oligosaccharide, is necessary for formation of a viable membrane, allowing mutation and study of the individual LPS genes responsible for the biosynthesis of the LPS structure distal to the minimal core (62).

The inner core is essential in establishing the barrier functions of the outer core membrane (62). There is a limited variation observed in the core structures between *E. coli* and *S. typhimurium*, likely representing its importance in membrane stability. Phosphate substitution of the inner core heptose residues is found to be essential for membrane stability and proper bacterial growth (62, 120). Loss of either the phospholipid derivatives or the heptose moiety is responsible for the "deep rough" phenotype, with strains showing increased phopholipid-protein ratio, a 50% reduction in OmpA, increased membrane phosphatidylethanolamine content, and increased colonic

Figure I-2 LPS core of *S. typhimurium*

The LPS core of *S. typhimurium* illustrating core mutations and location of LPS truncation (121, 167). The abbreviations GlcNAc, Glc, Hep, P, and KDO stand for *N*-acetyl-D-glucosamine, D-glucose, D-galactose, L-D-glycero-D-mannoheptose, phosphate, and 2-deoxy-D-manno-octulosonic acid respectively.



acid production (113, 133). The "deep-rough" mutants are observed to be supersensitive to hydrophobic compounds, with a "leaky" outer membrane resulting in significant leakage of the periplasmic enzymes into the surrounding medium (113).

The O-polysaccharide side chain acts as a protective layer in Gram-negative bacteria and is the first element of the bacterial cell to interact with the environment, in turn encouraging antigenic variation in the O-side chain (62). This is generated by the diversity of the individual sugar components, order, length and even sugar linkages (122). Loss of this repeating polysaccharide chain due to mutation in the biosynthetic genes, termed "rough" mutants, results in a characteristic colony morphology. Increasing membrane permeability with LPS truncation has been found to actually increase adhesion of *E. coli* O157:H7 adhesion to epithelial cells, indicating that the full length LPS may actually mask adhesion factors on the bacterial cell surface (15). It is likely that the role of LPS as a protective barrier, with an intact full length O-polysaccharide, is of greater benefit to the bacteria superceding any interference in surface protein interactions arising from full length O-polysaccharide.

8. LPS in conjugation.

Efforts to identify components on the recipient cell which interact with the pilus have implicated lipopolysaccharide (LPS) and membrane protein components of the recipient bacteria (2). An OmpA outer membrane protein mutant has been shown to be a defective recipient for the IncF plasmid transfer in liquid medium but not for R100 or F donors in plate matings (9). Rough mutant recipient strains, arising from mutations in genes responsible for LPS outer core synthesis, were more proficient as recipients than

smooth recipients with O polysaccharide. Plasmids from incompatibility groups FI, FII, M, J, and Iβ demonstrated higher frequencies of conjugation with rough than with smooth strains, whereas no difference was noted with plasmids of the incompatibility groups T, L, P, N and W. Reduced fertility of smooth strains is believed to occur as a result of the O polysaccharide of LPS shielding the recipient from contact with the pilus ultimately reducing the frequency of mating aggregate stabilization (130). The LPS αKDO is essential for bacterial growth. Mutations affecting the inner core of *S. typhimurium* LPS are usually poor recipients for IncF and other plasmids (130).

Recent findings by Anthony *et.al.* (9) have shown that the initial heptose (HepI) is essential for transfer of certain IncF plasmids, and the presence of extracellular *O*-phosphorylethanolamine (PEA) interferes with mating. Within the LPS inner core, the initial heptose substituted with pyrophosphatidylethanolamine (PPEA) arises from the transfer of PEA from phosphatidylethanolamine to phosphate-substituted heptose (HepI-P), producing HepI-PPEA. Following LPS translocation to the outer membrane, TolC, an outer membrane protein, is thought to remove PEA from HepI-PPEA, giving rise to HepI-P (133). The removal of PEA by TolC is inefficient, producing heterogeneity in the cellular LPS. To better illustrate the heterogeneity of the molecule, the initial phosphate moiety in substituted heptose residues will be designated (P)PEA. Studies of various F-like plasmids (9) have shown that the (P)PEA substituent plays an integral role in efficient transfer of some plasmids. Yet the role of (P)PEA in mating is poorly understood, because of other physiological changes accompanying (P)PEA loss.

9. IncH plasmids

Originally identified as responsible for chloramphenicol resistance in *Salmonella typhi* isolated from an outbreak of typhoid fever in Mexico in 1972 (7), the incompatibility group H (IncH) plasmids have been frequently associated with antibiotic resistance in *S. typhi* (152). The plasmids of the IncH family were unusual in their transfer properties, with optimum transfer occurring between 26°C to 30°C and undetectable transfer at physiological temperatures of 37°C (154). IncH plasmids are large, with members of the group ranging in size from 160kb to 260kb in size (147).

The plasmids belonging to the IncH family were identified by DNA hybridization as belonging to two subgroups, $IncH_1$ and $IncH_2$ (141), later classified as subgroups of IncHI plasmids: IncHI1, IncHI2 and IncHI3 (170). Plasmids of the IncHI1 group are distinguishable from IncHI2 and IncHI3 due to their incompatibility with F plasmids (142), mediated by a RepFIA replicon homologue (132). Plasmids isolated from *Klebsiella aerogenes*, specifying H pili but compatible with previously identified H plasmids, were placed into a separate group, designated IncHII (21). Incompatible with one another, the two IncH groups are observed to have very little DNA homology, although a high degree of homology exists among the plasmids within the subgroups (170).

Using restriction endonuclease mapping, identification of regions within R27 responsible for conjugal transfer (151), citrate utilization (149), tetracycline resistance (150), and incompatibility (53) was possible (Figure I-3).

The copy number of R27 within the host cell has been found to be low, approximately one to two copies per chromosomal equivalent (147). Although of

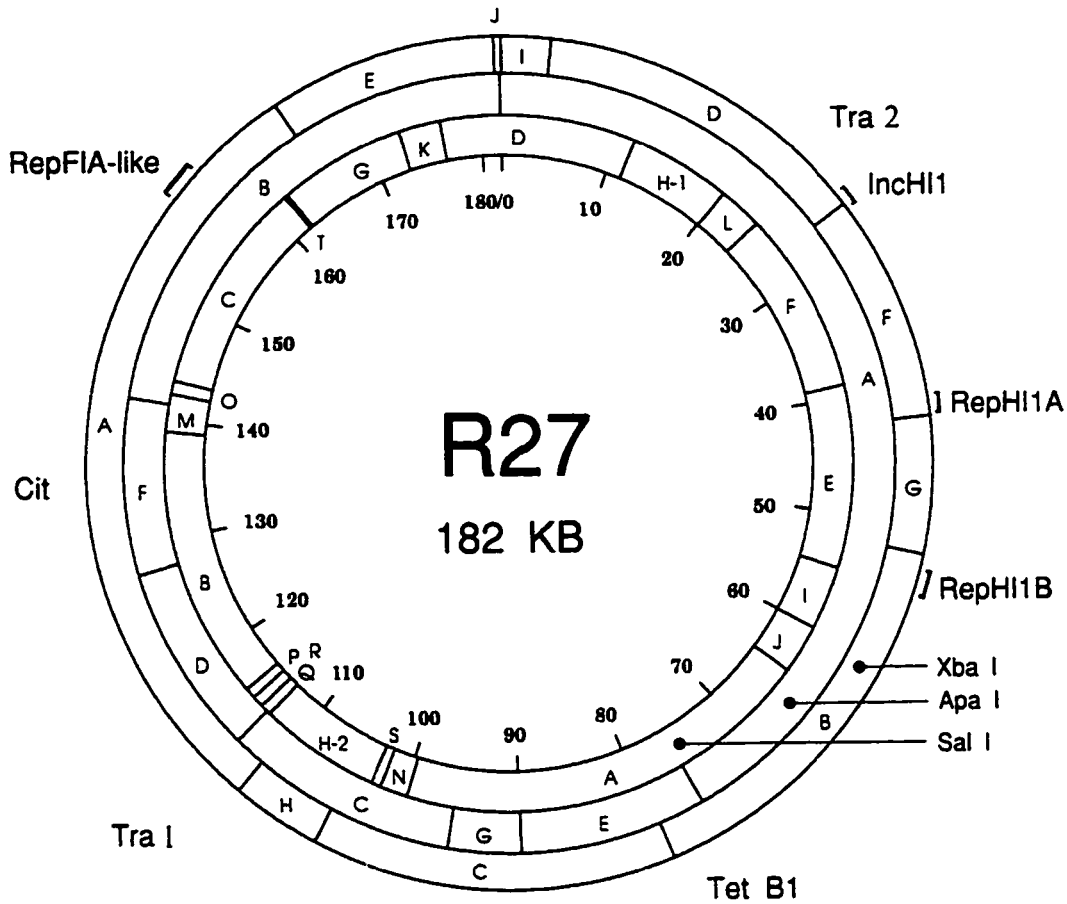


Figure I-3 Circular representation of the IncHI1 plasmid R27

A circular represention of the R27 plasmid with the transfer regions (Tra1 and Tra2), replication (RepHI1A and RepHI1B), incompatibility (IncHI1), tetracycline resistance (TetB1), citrate utilization (Cit) and IncF-like replicon (RepFIA-like) regions identified. The plasmid has been mapped with the restriction endonucleases *Sal*I, *Apa*I and *Xba*I (53).

relatively low copy number, the plasmids are stably maintained throughout successive generations, indicative of a rigorous partitioning copy-number control mechanism (104). R27 contains two distinct replicons, RepHIA and RepHIB, responsible for the efficient replication of the plasmid (25, 101). Both utilize an iteron method of replication control, in which the replication proteins (RepHIA-RepA and RepHIB-RepA) bind to specific sequences upstream and downstream of the *repA* genes, acting as negative self-regulators (25, 101). Binding of the RepHIB-RepA protein not only initiates transcription of its gene, it has also been shown to be involved in formation of the initiation complex 223bp upstream of the *repA* gene (25).

The two R27 replicons contain significant sequence and organizational similarity to one another, yet the RepHIA and RepHIB replicons operate independent of one another (25). Compatibility of the replicons arises from differences in the iteron binding sequences flanking the replication genes, with RepA from the RepHIA unable to bind, and therefore unable to affect the RepHIB replicon, and vice versa (25). Neither RepHIA nor RepHIB exhibit incompatibility with the parent plasmid when cloned separately. Interference in the functioning of one replicon does not disrupt the normal replication of the plasmid, as replication of the plasmid is then performed by the second replicon (52). Only when the function of both replicons, RepHIA and RepHIB, is suppressed simultaneously, due to the presence of these replicons exogenous to the core plasmid, is incompatibility observed (52).

The IncH plasmids, with the exception of the IncHI3 plasmid MIP233, encode thick flexible conjugative pili (20). The degree of piliation of IncH plasmid hosts is relatively low, ranging from one to two pili per cell for the IncHII plasmid pHH1508a to

less than one pili per 1000 cells with the IncHI1 plasmid R27 (90, 147). The R27 derivative, pDT2454, in which a Tn*7* insertion into the *htdA* (H-transfer determinant) gene resulted in 1000-fold increase in conjugation frequency, also results in constitutive expression of H-pilus (169). A similar phenotype was observed with the Tn*lacZ* derivative of R27, pDT1942 (91). Producing one to two pili per cell, the R27 derivatives pDT2454 and pDT1942, enabled analysis of the R27 pilus.

Utilizing the H-pilus specific bacteriophage Hgal (91), both the kinetics and assembly of the H-pilus were examined (90). Biosynthesis of the H-pilus is extremely slow in comparison to other conjugal pili, requiring 15 minutes for maturation compared to 30 seconds required for the F plasmid conjugative pilus (90, 106). It was also observed that the pilin subunits are assembled at the base of the pilus, rather than by extrusion of the subunits from the pilus tip. Furthermore the mature pilus was stable at temperatures non-permissive for conjugation (37°C) although mating aggregate formation was inhibited at 37°C (90).

Purification of the R27 pilus resulted in the identification of a single 3.2kDa peptide by SDS PAGE separation (99). The pilin subunits were resistant to both N-terminal amino acid degradation and proteolytic digestion. Although small in comparison to other conjugal pilin proteins in other plasmid families, no other protein was identified within a purified pilin preparation using high-resolution SDS PAGE analysis and both Coomassie Blue, and silver staining protocols (99). Random mini-Tn*10* transposon mutagenesis of pDT1942, generating transfer-deficient plasmids, allowed identification of genes necessary for pilus biosynthesis, but no mutant was isolated in which pilin subunit synthesis was disrupted (99). Although unsuccessful in

identifying the gene responsible for synthesis of the pilin subunits, these mutants have been very useful in analysis of the transfer regions.

Chapter II Materials and Methods

1. Antibiotics and media

Media used for growth of bacteria was the following: Brain Heart Infusion (BHI) (Difco laboratories, Detroit, MI), Luria Broth (Difco), and Penassay Broth (Difco). For solid media, agar was added prior to autoclave sterilization to a final concentration of 1.6% (w/v). Antibiotics were used at the concentrations listed in Table II-1.

Table II-1 Antibiotics used in this study.

Antibiotic	Concentration (μg/ml)
Carbenicillin (Cb)	100
Chloramphenicol (Cm)	20
Kanamycin (Km)	50[a]
Naladixic Acid (Nal)	100
Rifampicin (Rif)	100
Streptomycin (Str)	100
Tetracycline (Tc)	20

[a] This is the final concentration of kanamycin base in the medium. Purified kanamycin contains only 80% of the active component, kanamycin base.

2. Bacterial strains and plasmids.

Bacterial strains and plasmids used in this study are listed in Table II-2 and Table II-3 respectively. All strains were stored at -80°C in 20% glycerol in BHI

Table II-2 Bacterial strains used in this study

Strain	Genotype or description	Reference or Source
Salmonella tyhphimurium		
SL1027	*metA*22 *trpC*2 H1-bH2-e,n,x (cured of Fels 2) *fla*66 *rpsL*120 *xyl*404 *metE*551	(172)
SA2205	*rfaL446* Str^R, SL1027 parent	K. Sanderson[a]
SL1905	*rfaJ4041* Str^R, SL1027 parent	(146)
SA2206	*rfaI1417* Str^R, SL1027 parent	K. Sanderson[a]
SL1060	*rfaH481* Str^R, SL1027 parent	(145)
SL1183	*rfaF546* Str^R, SL1027 parent	(166)
SL733	*rfaK933* Str^R, LT2 parent	K. Sanderson[a]
SL3600	*rfaD657* Str^R, SL1027 parent	(81)
SL1102	*rfaE543* Str^R, SL1027 parent	(123)
Escherichia coli		
J53-1	*pro met* Nal^R	(10)
RG192-2	*ara leu* lac^- Rif^R	(148)
JE2571-1	*leu thr fla pil* str Rif^R	(19)
DT1801	lacZM15 F'$lacI^q$ *proA*	F. Rentier-Delrue[b]
SMR121	*recB*21 *recC*22 *sbcB*15 *sbcC*	S. Rosenberg[c]
SMR130	C600-RecD1009	S. Rosenberg[c]
DH10B[d]	F^- *mcrA* Δ(*mrr-hsd* RMS-*mcrBC*) Φ80DlacZ Δm15 Δ*lac*X74 *deoR recA*1 *endA*1 *araD*139 Δ(*ara, leu*)7697 *galU galK* λ- *rpsL hupG*	GibcoBRL, Burlington, Ont.

[a] Salmonella Genetic Stock Center, University of Calgary.

[b] Laboratoire central de Génie génétique, Université de liége, Belgium

[c] University of Alberta, Department of Biochemistry

[d] The lab strain also specifies a low-level resistance to Nal (20μg/ml)

Table II-3 Plasmids used in this study

Plasmid	Resistance Marker	Use	Reference or source
R27	Tc	IncHI1 Prototype	(150)
pDT2454	Tc	Mutant of R27, derepressed for transfer	(169)
R478	Cm Km Tc	IncHI2 Prototype	(154)
pDT1233	Tc Cb	Cosmid subclone of R27 *SalI*-FE	(99)
pDT1942	Tc Km	Mutant of R27, derepressed for transfer	(91)
pGEM-T	Cb	PCR cloning vector	Promega
pZERO-2	Km	*ccdB* based cloning vector	Invitrogen
pUC119	Cb	Cloning vector	(162)
pOK12	Km	Cloning vector	(161)

broth. Strains retrieved from cold storage were revived at 37°C overnight by growth on solid medium containing antibiotics selective for the strain, if appropriate.

3. Conjugation and assay of transfer ability.

Donor and recipient cells were grown to late log phase in Penassay broth for assay of LPS effect on transfer and in BHI broth for all other conjugation experiments. An inoculum of 50µl of late log phase donor and recipient cells were mixed and cells allowed to mate overnight at 28°C in 20ml of BHI broth. Following mating, the numbers of total recipient cells and transconjugants were determined by serial dilution of the cell suspension in 0.05M PB and plating the dilutions on appropriate selective medium. Solid surface matings were performed using equal numbers of late log phase cells (50µl) mixed and spread evenly on BHI agar (1.6%), with cells scraped off following overnight incubation at 28°C using 1ml of 0.05M PB.

Mating efficiency in the presence of 0.1M PEA (Sigma Chemical Company, St. Louis, MO) was determined by preincubating donor cells for 2 hours with 0.1M PEA and allowing the donor and recipient cells to mate on BHI agar containing 0.1 M PEA. Following overnight incubation at 28°C, cells were suspended in 0.05M PB, serially diluted and plated on appropriate selective media to determine transfer efficiency.

4. Determination of bacterial mating aggregation.

Donor cultures were grown for two hours at 27°C in BHI broth, 10µl were then removed using a wide-bore pipette tip and placed on the inside surface of a sterile plastic petri dish. The plate was inverted and 10µl of recipient culture grown for two hours at

27°C was added to the hanging drop using a wide-bore pipette tip. The hanging drop suspension was placed over a small volume of water, to limit droplet evaporation, and incubated for 30 minutes at 27°C. As a control 20µl of the two hour donor culture was incubated by the same hanging drop method.

A 200-mesh electron microscope copper grid (Fisher Scientific, 12-624-103), freshly coated with Formvar, was placed on the underside of the hanging drop and left at room temperature for one minute. The grid was carefully removed, edge blotted with Whatman #1 filter paper, and negatively stained with 1% phosphotungstic acid (w/v) (pH7.0) (Sigma) for 1 second.

Aggregate formation was assayed as a percentage of all cells observed using the Philips Model 410 transmission electron microscope, at an accelerating potential of 80kV. Cell wall contact was considered aggregation and adjacent cells were carefully scrutinized at varying magnifications to differentiate random bacterial arrangements from aggregates. Aggregation due to mating was assessed by comparison of the percentage of aggregates observed between plasmid-free *E. coli* J53-1 and the appropriate LPS mutant recipient, to the aggregation observed between *E. coli* J53-1 (pDT2454) and appropriate LPS mutant recipient. LPS specific alteration of mating aggregation was determined by comparison of the percentage of aggregation observed with LPS mutant recipients and wild-type donors to LPS mutant donors and wild-type recipients.

5. Agarose gel electrophoresis

DNA fragments were separated using standard agarose (Promega, Madison, WI) at concentrations ranging from 0.6%-1.5% w/v. Electrophoresis was performed at 5-10 V/cm for various time periods with TAE running buffer (0.04M Tris, 0.04M acetic acid,

1.0mM EDTA, pH 8.3) and loading buffer (0.05% bromophenol blue, 0.05% xylene cyanol FF, 8% w/v sucrose). DNA present in the gels was visualized by immersion of the gel in 5μg/ml ethidium bromide in water for 20-30 minutes, followed by illumination with a UV transilluminator (260nm).

6. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

Polyacrylamide gels containing 1% of the denaturant sodium dodecyl sulfate (SDS) were prepared as described by Laemmli (78). Polyacrylamide concentrations varied from 10-15% using a 40% acrylamide solution (19:1) with 1% ammonium persulfate as catalyst. A stacking gel of 4% acrylamide was layered above the actual separating gel for increased band resolution. A voltage of approximately 8 V/cm was applied to the gel until the dye front had passed the interface between the stacking and resolving gels, at which time the voltage was increased to approximately 15 V/cm for the duration of the run.

After electrophoresis, gels were immersed in Coomassie Blue stain (40% methanol, 10% acetic acid, 0.2% Coomassie Brilliant Blue R250) for at least 4 hours. Gels were destained by washing in 30% methanol with 10% acetic acid until the required contrast was achieved. Alternatively silver staining of the proteins was used (128). After overnight fixation in 40% ethanol and 5% acetic acid, the gels were washed 3 times in Milli-Q water for 10 minutes. The gels were then incubated in 0.1% $AgNO_3$ (w/v) for 30 minutes at room temperature. They were then washed a further 3 times in Milli-Q water and developed by immersion in 2.5% Na_2CO_3 with 0.02% formaldehyde until bands were visible. Further development was prevented by addition of 1% acetic acid.

7. High resolution SDS-PAGE.

To resolve the small pilin polypeptide, high resolution polyacrylamide gel electrophoresis were used according to the method of Shägger and vonJagow (138) in which three separate polyacrylamide layers were used. A 5% acrylamide stacking gel was separated from a 16.5% separating gel by a 10% acrylamide spacer. In addition separate cathode (0.1M Tris, 0.1M Tricine, 0.1% SDS, pH 8.25) and anode (0.2M Tris, pH8.9) buffers were used during electrophoresis, however the voltage used was identical to that used in normal SDS-PAGE analysis.

8. Pilus purification.

Purified conjugative pili were prepared from *E. coli* JE2571-1 (pDT1942) using previously identified methods (99) and analyzed by high-resolution SDS-PAGE. Coomassie blue stain was used for gross assay of purity of the preparations and silver staining for visualization of the purified pilin.

9. Plasmid purification

Large scale DNA purification was performed using a modification of the Birnboim and Doly protocol (16). An overnight culture (500 ml) of overnight broth growth of the *E. coli*, containing the plasmid of interest, was centrifuged at 5,000RPM (Beckman JA-14 rotor, 4,000xg) for 5 minutes. Excess broth was removed, and the bacteria were carefully suspended in 20ml of solution 1 (40mM Tris, 2.5mM EDTA, 10% glucose, pH 8.0), with special attention given to prevent premature lysis of the cells and

shearing of the DNA. Following, 40ml of solution 2 (0.2N NaOH, 1% SDS) was added and then mixed by careful inversion until clearing occurred. The sample was mixed with 30ml of solution 3 (3M sodium acetate, pH 5.0) and allowed to stand in ice water for a minimum of 45 minutes. The resulting precipitate was removed by centrifugation at 10,000RPM (JA-14 rotor, 15,000xg) for 20 minutes, and the supernatant mixed with an equal volume of isopropanol and incubated at -20°C for a minimum of 45 minutes. The solution was then centrifuged at 10,000RPM (JA-14 rotor, 15,000xg) for 10 minutes, the resulting pellet was then dried, and dissolved in 4ml of TE buffer (1mM EDTA, 10 mM Tris, pH 7.5) containing 20μg/ml DNAase free RNAase, and 5.2g of $CsCl_2$ dissolved into it. After transfer to a Beckman Quick-Seal centrifuge tube (Beckman Instruments Inc, Palo Alto, CA), 40μl of 10mg/ml ethidium bromide solution was added. The tubes were then centrifuged in a Beckman TL-100 ultracentrifuge for 18-24 hours at 80,000 RPM (TLA 100.2 rotor, 227,000xg)

Medium scale plasmid DNA preparations were performed by a modified version of the large scale DNA purifications, using 20ml overnight broth culture of the bacteria containing the plasmid. The volumes of the solution were proportional to the volume of the cells: 400μl of solution 1, 800μl of solution 2, and 600μl of solution 3. Following addition of solution 3 and incubation in ice-water for 45 minutes, 0.7ml of supernatant was removed, 0.8ml of isopropanol added and incubated for a further 45 minutes at -20°C. The solution was centrifuged at 14,000 RPM (Eppendorf centrifuge 5414C, 10,000xg) for 20 minutes, the resulting pellet dried and dissolved in 400μl of solution 4 (50mM Tris, pH8.0, 0.1M sodium acetate). Two volumes of ethanol were added, the solution incubated at -20°C for 30 minutes and DNA precipitated by centrifugation at

14,000RPM (Eppendorf centrifuge 5415C, 10,000xg) for 10 minutes. The resulting pellet was redissolved in 200µl of TE buffer and 1µl of DNAase free RNAase added.

Small scale DNA preparations were used for PCR screening of colonies to identify the presence of correct insertions in recombinant vectors. Procedures were similar to those used in medium scale preparations, except the bacteria were suspended in 150µl of solution1, cells lysed by addition of 300µl of solution 2 followed by 225µl of solution 3. Following centrifugation the pellet was dissolved in 450µl of solution 4, and the DNA ethanol precipitated.

10. Transformation of bacterial cells using $CaCl_2$.

Plasmid DNA and ligation mixtures were transformed into cells by the method of Cohen *et al* (27). A 100µl aliquot of *E. coli* DH10B culture grown in BHI broth overnight, was inoculated into 20ml of BHI broth in a 40ml flask, which was incubated at 37°C with constant shaking for 3 hours. The culture was aseptically transferred to a sterile 50ml centrifuge tube, and centrifuged at 5,000RPM (JA-20 rotor, 3,000xg) for 10 minutes. The cells were then resuspended in 10ml of ice-cold 50mM $CaCl_2$, stored at 4°C overnight and pelleted by centrifugation at 5,000RPM (JA-20 rotor, 3,000xg) for 5 minutes. The supernatant was decanted and the cells resuspended in 2ml of ice-cold 50mM$CaCl_2$. Using a sterile pipette 100µl of the cellular suspension was added to approximately 50ng of DNA, mixed carefully and incubated at 0°C for 30 minutes. The suspension was then subjected to a heat shock at 42°C for 90 seconds, chilled in an ice bath for 2 minutes and 800µl of BHI broth added. Cells were incubated at 37°C for a

minimum of 1 hour before plating on media containing appropriate antibiotic(s) to select for the presence of the plasmid within the cells.

11. Transformation of bacterial cells by high-voltage electroporation

Transformation of *E. coli* cells was performed using the method of Dower and Miller, with high-voltage applied to the prepared cells allowing entry of exogenous DNA (37). Electro-competent cells were prepared by inoculating 20ml of LB broth with 100µl of overnight culture and then incubated at 37°C with constant shaking. The cells were transferred to a sterile centrifuge tube and pelleted by centrifugation at 5,000RPM (JA-20 rotor, 3,000xg) for 5 minutes. Cells were washed 3 times in ice-cold sterile Milli-Q water, followed by one wash in ice-cold sterile 10% glycerol in Milli-Q water. Following the final pelleting the cells were resuspended in 2ml ice-cold 10% glycerol in Milli-Q water.

DNA solutions used for electroporation require very low concentrations of salt and therefore must be de-salted, though conditions varied depending on the DNA concentration within the sample. Addition of two volumes of room temperature ethanol to the DNA sample, brief incubation at -20°C and centrifugation at 14,000RPM (Eppendorf centrifuge 5415C, 10,000xg) for 10 minutes was usually sufficient. The DNA could be further de-salted if necessary following centrifugation by careful washing of the intact pellet with 70% ethanol at room temperature. The DNA pellet was then dissolved in 10µl of sterile Milli-Q water.

A volume of 40µl of the bacterial cells, prepared as described above, was mixed with 10µl of DNA, incubated at 0°C for 1 minute and placed in a chilled Biorad

GenePulser cuvette (Cat# 165-2086). The cuvette was inserted into the Biorad GenePulser and subjected to electroporation using 2.4 Volts potential, 200 Ohms resistance and 25μFd capacitance. After electroporation 1ml of LB broth was added, the mixture was incubated for 1 hour at 37°C and plated on appropriate antibiotic medium for selection of the plasmid.

12. Restriction endonuclease digestion

Restriction endonucleases were obtained from two sources (Life Technologies, Rockville, MD; New England Biolabs, Canada). Digestions were carried out according to specifications supplied by the manufacturers using the restriction buffers supplied with the enzymes. Restricted DNA was analyzed by agarose gel electrophoresis.

When possible, multiple restriction digestions were performed using compatible buffers. If compatible buffers were not available following restriction digestion with one enzyme, the enzyme was heat inactivated by incubation for 10 minutes at 65°C, then precipitated in 2 volumes of ice-cold ethanol. The DNA was dissolved in water, a compatible buffer added and then restricted using the second enzyme.

Individual bands were isolated and purified after gel electrophoresis, allowing only a brief exposure to the UV light source to minimize DNA damage. The bands of interest were cut from the agarose gel, finely diced, placed in a Spin-X centrifuge filter (Corning Inc., Corning, NY) and centrifuged at 14,000RPM (Eppendorf Centrifuge 5415C, 10,000xg) for 20 minutes to separate the liquid and DNA from the agarose. To the resulting liquid 2 volumes of ethanol were added. The ethanol mixture was stored at

-20°C for 30 minutes and the DNA pelleted by centrifugation at 14,000RPM (Eppendorf centrifuge 5415C, 10,000xg) for 10 minutes.

13. DNA ligation

DNA restriction fragments with compatible ends were ligated by reaction with T4 DNA ligase (Life Technologies, Rockville, MD) overnight at room temperature using the manufacturer's recommendations for buffer conditions (50mM Tris, 10mM $MgCl_2$, 1mM DTT, 1mM ATP, 5% PEG-8000, pH 7.5). Total volume was kept to 20µl and 1-2 units of T4 DNA ligase used for each reaction.

For those reactions requiring blunt ends for ligation using DNA containing 5' or 3' overhangs, Klenow fragment was used to generate blunt ends. The DNA fragment was suspended in water with 10µl of 10X Klenow buffer (10mM Tris, 5mM $MgCl_2$, 7.5mM DTT, pH 7.5), 1µl 0.1M DTT, 0.5µl BSA, 2µl dNTP mix (1.25mM dATP, dGTP, dCTP, dTTP), 1µl 10mM ATP and the total volume brought to 100µl with sterile Milli-Q water. To this mixture 1µl of DNA polymerase I (Klenow fragment) (Life Technologies, Rockville, MD) and 1µl of T4 Kinase (Life Technologies, Rockville, MD) was added and the tube containing the reaction incubated at 37°C for 1 hour. After incubation 1µl of 0.5M EDTA was added to stop the reaction and the enzymes digested by addition of 10µl of 1mg/ml Proteinase K and 10µl 1% SDS, then incubated at 42°C for 1 hour. Protein was removed by phenol-chloroform extraction and DNA precipitated by addition of 2 volumes of ethanol. The tube was placed at -20°C for 30 minutes and the DNA pelleted by centrifugation at 14,000RPM (Eppendorf centrifuge 5415C, 10,000xg) for 10 minutes. Ligation conditions were similar to those used in 5' and 3' overhang ligations except 3µl

of T4 ligase was used and incubated at 4°C for 6 hours, followed by 16°C for 16 hours and then 22°C and 2 hours.

14. Polymerase chain reaction

The polymerase chain reaction was used for amplification of fragments for both analysis of recombinant vectors and for generation of specific linear fragments from larger constructs for cloning. Template DNA was either purified (for cloning) or prepared using the small scale plasmid preparation (for recombinant vector analysis). Primers used in this study are listed in Table II-4 .

Approximately 1pM of each primer was added to approximately 0.1μg of template DNA containing 2μl of 10mM dNTP mixture, 5μl of 10X PCR buffer (50mMKCl, 100mM Tris-HCl, 1.5mM $MgCl_2$, pH 8.3), 1.5μl of $MgCl_2$, 1.5μl of Taq DNA polymerase (Life Technologies, Rockville, MD). The reaction mixture was brought to a total volume of 50μl with sterile distilled water. Samples were transferred to the Perkin-Elmer 9600 thermocycler or the BioRad GeneCycler and amplification performed for 35 cycles with the following parameters: 94°C, 40 sec; 50°C, 30 sec; 72°C, 45 sec.

The pGEM-T vector system (Promega Corporation, Madison, WI) was used for cloning of those PCR fragments generated by Taq polymerase PCR, not incorporating unique restriction sites within the primers. In such cases the PCR product was separated and purified by agarose gel electrophoresis and the product ligated with the pGEM-T vector using usual 5' and 3' overhang ligation conditions. The fidelity of the resulting

Table II-4 Oligonucleotide primers used in this study

Primer name	Sequence[a]
CRAIG1	AGGCTCTGGGAGGCAG
CRAIG2	CATGCCGTCTGTGATGG
CRAIG3	TTCCGATGATGGACGCTTC
CRAIG4	ATGATTTTACCCGGCGCACC
CRAIG5	CGCTTTTGCATTTTTGAAGTCC
CRAIG10b	GAATTCAAATCAACACGGAG
CRAIG11b	CCCCCTCATTGTTATTGTTC
CRAIG15	CGCGGATCCATGGAACTGACATTGAATACTAACGTT
CRAIG16	TTCCAAGTCATTGGCTGCAGTTCACAGAGGAATACCAGCATCCAGGAA
CRAIG19	CCGGAATTCAATGGAACTGACATTGAATAC
CRAIG20	CGGGATCCGTCAATGGTGATGGTGATGGTGAGCCAGAGGAATACCAGCATCCAG
CRAIG34	CCACACCCCCCTTGGTTGGTTCTTG
CRAIG35	GTTTCAAGCCTGCCGATGTG
CRAIG36	TGCACGCTTTACCGTTCG
CRAIG37	GACTGATGTGACTGTGG
CRAIG38	TACCGGCACAGTTCGTTTTCCTG
CRAIG39	TCGGTCATGCTGCAAC
CRAIG31	CAGGATCCCAACCCTAGATTCATGTACACCTG
CRAIG33	GCTGAAGCACAGAAAACCG
CRAIG40	CCCAAGCTTGGGATCAGAAATCAGTAGCCCAG
CRAIG41a	CGGGATCCGCGAGTAACCAGACATCAG
CRAIG52	AAAAATTCACCCTGCCCC
CRAIG60	CCCAAGCTTCACTCACAAAAATTACCCC
CRAIG61	CGCGGATCCCCTAAAAATGCCAGCAAC
CRAIG62	CCCAAGCTTAAAAATTCACCCTGCCCC
CRAIG63	CGCGGATCCTTGCTGACATCAACTTCAACC
CRAIG64	CCCAAGCTTCGAGATGGACTCACATTGCC
CRAIG65	CGCGGATCCGTCTACGCATTACCAAGTCC
CRAIG72	CCGGAATTCATGAGAGGCTCGCATCACCAT

[a] Restriction sites incorporated in the oligonucleotide primers are underlined

clone was ensured by DNA sequencing of the insert using primers specific to upstream (M13For) and downstream (M13Rev) of the insertion.

Those PCR amplified products with unique restriction sites available for cloning were amplified using the *Pfu* DNA polymerase (Stratagene) for increased fidelity of the final product. Buffer and reaction conditions were used as recommended by the manufacturer.

15. DNA sequencing

DNA sequencing was performed using the ThermoSequenase, radiolabeled terminator, cycle sequencing kit (Amersham), with ^{33}P labeled nucleotides. Sequenced products were visualized using a 40% urea PAGE gel running at 135W. Glycerol tolerant buffer at a final concentration of 0.8X to 1X (Per litre: 21.6g Tris base, 7.2g Taurine, 4g EDTA) was used to compensate for the high glycerol content of the pre-diluted enzyme.

16. Cell free protein expression

Expression of polypeptides from cloned DNA fragments in a cell free background was performed using the Promega *E. coli* T7 S30 extract system for circular DNA, a coupled transcription/translation system, using ^{35}S methionine for labeling. 5µl of $CsCl_2$ purified recombinant vector DNA (approximately 4µg) was added to 5µl of the amino acid mixture (without methionine), containing 20µl of S30 premix, 15µl of T7 S30 extract and 3µl of ^{35}S methionine. The mixture was gently mixed and incubated at 37°C

for 1.5 hours. The reaction was stopped by immersion in an ice bath and the reaction contents separated by SDS-PAGE analysis. The subsequent gel was dried, under vacuum, at 65°C for 4 hours and exposed to Kodak X-ray film (X-OmatAR) overnight at room temperature to visualize the polypeptides.

17. Immunoblotting procedures.

Proteins separated by SDS-PAGE were electrophoretically transferred to nitrocellulose filter using a tank electroblotting device (BioRad laboratories, Hercules, CA) for 6 hours at 27V (25mM phosphate buffer, pH7.4). The membrane was then transferred to a small dish and washed for 1 hour by slight shaking in PBS-Tween (25mM PBS, 0.0005% Tween-20, pH7.4) with 10% skim milk powder (milk) and then rinsed twice in PBS-Tween (10% milk). A 1:2000 dilution of the primary antibody in 40ml PBS-Tween (5% milk) was added and incubated with slight shaking, at room temperature, for one hour. After draining, the membrane was rinsed twice in PBS-Tween (5% milk) and washed twice for 10 minutes in PBS-Tween (5% milk). To the membrane 40ml of a 1:5000 dilution of horseradish peroxidase-labelled secondary antibody, in PBS-Tween (5% milk), was added to the membrane and incubated at room temperature for 15 minutes. The antibody solution was then drained and washed twice for 10 minutes in PBS-Tween (5% milk). The membrane was transferred to a new dish, immersed in ECL mixture (Enhanced Chemiluminescence, Amersham Life Science: 1.5ml ECL Solution1, 1.5ml ECL Solution2, 3ml water) and left to sit for one minute. Excess fluid was drained, the membrane wrapped in Saran Wrap, exposed to Kodak X-Omat Blue film for 20 seconds to two minutes, and the film developed.

18. DNA transfer to nitrocellulose membranes

Following agarose gel electrophoresis of DNA fragments, the gel was transferred to a glass baking dish and extraneous areas of the gel cut away. A small nick was placed at the lower left corner for later orientation of the gel. The DNA was denatured by soaking of the gel in 1.5M NaCl and 0.5M NaOH for 45 minutes, washed briefly in deionized water and the pH neutralized by incubation in 1M Tris (pH7.4) with 1.5M NaCl twice for 30 minutes. The gel was drained and then centered on a piece of Whatman 3MM paper wrapped about a glass plate support. The wrapped support was then placed over a glass dish and the dish filled with transfer buffer (1.5M NaCl, 0.15M Sodium Citrate) such that the ends of the 3MM paper were immersed. A piece of nitrocellulose filter was cut to the exact size of the gel, moistened briefly in deionized water and immersed in transfer buffer for five minutes. The filter was then placed over the agarose gel, all air bubbles carefully removed and the gel completely surrounded, but not covered, with Saran Wrap. Two pieces of 3MM paper were wet in 2X SSC buffer (0.3M NaCl, 0.03M sodium citrate, pH7.0) and placed over the nitrocellulose membrane. A 10cm stack of paper towels was then placed over the gel/membrane/3MM paper, weighted down by approximately 500g and left overnight.

The following day the stack was disassembled, the nitrocellulose filter removed and carefully washed for five minutes in 6X SSC buffer (0.9M NaCl, 0.09M sodium citrate, pH7.0). The membrane was then air-dried and the DNA fixed to the membrane by baking under vacuum for one hour at 80°C.

19. DNA probe labeling.

Probes were constructed by PCR with appropriate primers and template DNA, generating linear DNA specific to the sequence to be probed. The PCR product was purified by agarose gel electrophoresis, precipitated and dissolved in deionized water. An aliquot of approximately 25ng was removed, denatured in boiling water for five minutes and cooled on ice. To the cooled DNA, 2μl of nucleotide solution (0.5mM dATP, dTTP, dGTP) and 5μl of 10X Klenow buffer was added, and the total volume brought to 50μl with deionized water. As a radioactive label, 1-2μl of ^{32}P dCTP was added, immediately followed by addition of 1μl of DNA polymerase I. The mixture was incubated for one hour at 37°C and the DNA precipitated by addition of 3μl of sodium acetate (3M, pH5.0) and 100μl of ice-cold ethanol, then incubated for 30 minutes at ¯20C. The DNA was pelleted by centrifugation for 10 minutes at 14,000 RPM (Eppendorf Centrifuge 5415C, 10,000xg), dried, and re-dissolved in 20μl of TE buffer. A 2μl sample was removed and radioactivity assessed in the Wallac 1414 Liquid Scintillation Counter using 10ml of CytoScint (ICN Pharmaceuticals, Costa Mesa, CA) liquid scintillation cocktail.

20. Sequential scanning of amino acid reading frames encoded within the Tra2 region.

Using the Borland C^{++} V4.5 programming language (Borland International, Scotts Valley, CA), a computer program was constructed in which formatted amino acid sequence was input, the amino acid composition within a 25 amino acid frame calculated, compared to previously determined values and the frame shifted downstream by one

amino acid. This procedure was repeated until all input amino acids were sequentially scanned.

21. Bacteriophage isolation

Bacteriophage preparations of Hgal were obtained by inoculating 20 ml of BHI broth with 50µl of 6 hour culture of *E. coli* J53-1 (pDT1942) and addition of 50µl of Hgal stock ($6.4x10^8$ pfu/ml). The suspension was incubated overnight at 28°C, vortexed briefly and centrifuged at 5,000RPM (JA-20 rotor, 3,000xg) for 10 minutes. The supernatant was decanted, 0.5ml of chloroform added and the phage suspension stored at 4°C.

22. Bacteriophage assay.

A bacteriophage assay was performed as described previously (128). A volume of 50µl of *E. coli* J53-1 (pDT1942), grown at 28°C for 6 hours in BHI broth, was added to 8ml of molten BHI broth containing 0.8% agar equilibrated to 45°C. The bacterial suspension was carefully mixed, spread onto a BHI agar plate and allowed to solidify at room temperature. After solidification 10µl of phage suspension was spotted onto the plate and the plate incubated overnight at 28°C. Plates were examined for zones of clearing at the point of phage spotting.

Assay of the plaque forming units present within a sample was determined by addition of 100µl of serial dilution of the bacteriophage sample to *E. coli* J53-1 (pDT1942), grown for 6 hours at 28°C, in 8ml of molten BHI broth containing 0.8% agar equilibrated to 45°C. The phage mixture was carefully spread onto a BHI agar plate and

incubated overnight at 28°C. Individual plaques were counted, and the total number of plaque forming units per ml (pfu/ml) determined.

23. Radiolabeling of phage

As a method of identifying successful pilin expression within transfer-deficient mutants of R27, the H-pilus specific bacteriophage Hgal was radiolabeled, purified and then used for screening of previously generated mutants (99). Labeled bacteriophage were generated by incorporation of a radioactive label by growth on defined media containing ^{35}S methionine.

Labeling media was prepared as follows: 100ml of M9 minimal media (128) containing 200µl $MgSO_4$ (1M), 20µl $CaCl_2$ (1M), 1ml proline (5mg/ml), 1ml thiamine (1.5mg/ml), 1ml threonine (5mg/ml), 1ml arginine (1.5mg/ml) and either 1.6g (base) or 0.8g (soft overlay) agar. The medium was sterilized by autoclaving and 1ml glucose (20% w/v) and 50µl of 1,000 Ci/mM ^{35}S methionine were added (NEN Life Science Products, Boston, MA). A volume of 25ml of 1.6% agar was poured into a sterile petri dish and allowed to solidify. Following solidification, 100µl of a 6 hour culture *E. coli* J53-1 (pDT1942), grown at 28°C, was added to 4ml of molten supplemented M9 medium containing 0.8% agar equilibrated to 45°C. Immediately, 50µl phage stock ($6.4x10^7$ pfu/ml) and 5µl ^{35}S methionine was added, carefully mixed and poured on to the solidified agar base. The plate was then sealed with parafilm and incubated at 28°C.

After an 18 hour incubation, 1ml of LB medium was placed on the overlay at room temperature for 5 minutes and then the cells were carefully scraped off the plate. The resulting suspension was lysed at room temperature by treatment for 20 minutes in a

Bransonic 220 sonicating water bath (Tetrochem laboratories Ltd., Edmonton, Canada) at room temperature and LB medium added to bring the total volume to 5 ml. The bacteriophage suspension was purified as described in the next section.

24. Bacteriophage Hgal concentration.

Purification of the bacteriophage Hgal was based upon the methods of Sambrook *et al.* (128). DNase I and RNase were added to the 5ml phage suspension to a final concentration of 1μg/ml. The enzymes were allowed to digest their substrates for 30 minutes at room temperature. Then 2.9g of NaCl was carefully dissolved in the suspension which was incubated at 0°C for 1 hour. The contents were transferred to microfuge tubes and centrifuged at 14,000RPM (Eppendorf centrifuge 5415C, 10,000xg) for 10 minutes at 4°C. To the supernatant 0.5g of PEG was dissolved in the supernatant and then cooled to 0°C for 1 hour. Bacteriophage were pelleted by centrifugation at 14,000RPM (Eppendorf centrifuge 5415C, 1,000xg) for 10 minutes, the supernatant removed, and the remaining fluid carefully drained. Final titre of the purified bacteriophage was determined to be $6x10^9$ pfu/ml with a measured radioactivity of $7.3x10^6$ cpm/ml.

25. Screening of pilin-deficient mutants.

Cells containing transfer-deficient R27 mutants were grown overnight at 28°C and 100μl of culture was either used directly or first lysed using the dismembranator/probe sonicator (Fisher Scientific) set at 60% power, 3 times for one minute. The Bio-Dot microfiltration apparatus (Bio-Rad laboratories, Richmond, CA)

was used for transfer of the cells to a nitrocellulose membrane (0.45μm, Micron Separations Inc, Westborough, MA). The membrane was clamped between the gaskets and the 96 well sample template, 100μl of culture or lysate were added to individual wells and allowed to filter through the membrane by gravity only. Wells not containing samples were covered with parafilm to create a seal for later steps requiring vacuum. The wells were washed with 100μl TBS (20mM Tris, 500mM NaCl, pH7.5) containing 0.05% Tween 20 using gravity only, followed by a second wash of 100μl using vacuum to speed the process.

Once dry, the nitrocellulose filter was removed and washed for 5 minutes on a rotary shaker with 0.05% Tween 20 in TBS, followed by blocking of the membrane by incubation overnight at room temperature in 1% BSA in TBS with slight shaking. The membrane was then washed with 0.05% BSA-TBS, and incubated overnight in 20mL 0.05% BSA-TBS containing 150μl of radiolabeled phage. The membrane was then washed three times for 15 minutes in 0.05% BSA-TBS, followed by two washes in TBS. The membrane was wrapped in Saran wrap and exposed to Kodak X-ray film (X-Omat AR) at room temperature until an appropriate exposure was obtained (approximately 36 hours).

26. Construction of histidine tagged TrhA protein.

A series of 6 histidine residues (6xHis) was used to tag the TrhA protein at the N- and C- terminal ends of the protein. C-terminal 6xHis additions were generated using primers constructed such that 3 sets of the DNA sequence "CATCAC" encoding two histidine residues,were generated on the "+" reading strand, followed by a "TGA" stop

codon (CRAIG19 and CRAIG20), using the cosmid pDT1233 as template. The unique restriction sites *Eco*RI and *Bam*HI were incorporated into the upstream and downstream primers respectively for use in directional cloning of the amplified DNA into the expression vector pMS119.

Addition of a N-terminal histidine label was performed by either of two methods. A N-terminal 6xHis label was incorporated using constructed PCR primers similar to that done with the C-terminal tags (CRAIG72 and CRAIG73) with an "ATG" methionine start codon encoded into the upstream primer, followed by 3 sets of the "CATCAC" nucleotides. The resulting PCR product was cloned into the expression vector pMS119.

A second method of N-terminal histidine-labeling utilized the expression plasmid pBT1T-Tϕ , a modified version of the prokaryotic expression vector pT7-7. pBT1T-Tϕ contains, amongst other features, a T7 gene10 ϕ10 promoter, 6xHis tag and a T7 transcription terminator. Primers specific to the *trhA* gene were constructed, incorporating a unique upstream *Bam*HI and a downstream *Pst*I restriction site into the PCR product for directional cloning into the pBT1T-Tϕ, resulting in a proper fusion protein.

27. Procedure for generation of *trhA* mutants

The complete *trhA* gene and flanking sequence was amplified by PCR and cloned into pGEM-T generating the plasmid pCS15. The primers used for the PCR reaction were specific to approximately 400bp upstream and downstream (nt 37197-38463)of the *trhA* gene (MIC49 and MIC50).

Using a unique *Eco*RV restriction site located 156 nucleotides (nt37882) downstream of the "ATG" start of the *trhA* gene, two separate insertion-mutants were generated by blunt-end ligation. The Tn*5* kanamycin-resistance (Km^R) cassette was excised from the cloning vector pUC4-KIXX (12) by restriction with the *Sma*I enzyme. The fragment was agarose gel purified and blunt-end ligated into the unique *Eco*RV site within the cloned *trhA* gene in pCS15. A second mutant was generated by insertion of a *Campylobacter coli* chloramphenicol cassette (Cm^R) excised from pUOA20 using *Hin*cII and blunt-end ligated into the unique *Eco*RV restriction site within pCS15. The resulting plasmids were transformed into *E. coli* DH10B and plated on BHI agar selecting for the pCS15 plasmid (Cb^R) and the presence of the inserted cassette (Km^R or Cm^R). Correct insertion of the cassettes was determined by PCR using primers flanking the *Eco*RV site (CRAIG3 and CRAIG4) and by sequencing using the flanking primers.

As a positive control for the following recombination experiments, 3.9kb *Hin*dIII fragment carrying the mini-Tn*10* insertion was excised and agarose gel purified from the pDT1942 transfer-deficient mutant, pDT2955 (99). The position of this mutation has previously been identified as being within the *trhC* gene and the mutant has a transfer-deficient phenotype. The purified fragment was then cloned into the unique *Hin*dIII site within the multiple cloning site of pUC119, generating the plasmid pCS21.

28. Procedure for generation of *trhA* mutants within R27 using linear fragment recombination.

As labeling and therefore effective purification of the TrhA protein was unsuccessful, mutation of the *trhA* gene within R27 by recombination, to prove the role

of *trhA* in pilus biogenesis, was attempted. Linear fragments of the Km or Cm cassette disrupted *trhA* gene were generated using primers specific to approximately 400 bp upstream and downstream of the *trhA* gene (MIC49 and MIC50)(nt 37197-38463), the *trhA* gene proper (CRAIG10b and CRAIG11b)(37658-38044) or primers bracketing the *Eco*RV cassette insertion site by approximately 100 nucleotides (CRAIG4 and CRAIG5)(37833-37994).

The linear PCR fragments were purified by agarose gel electrophoresis and introduced by electroporation into *E. coli* TG1 (R27), *E. coli* SMR130 (R27) and *E. coli* RG192-2 (R27). After incubation at 37°C for 1 hour in 1ml BHI broth, the cells were plated on BHI agar containing antibiotics to select for the presence of R27 (Tc) and the disrupting cassette (Cm or Km). As a positive control the 3.9kb *Hin*dIII fragment was excised from pCS21, agarose gel purified and introduced by electroporation into the three strains.

29. Procedure for generation of transfer mutants using vector based recombination.

Two linear fragments were generated by PCR of the disrupted *trhA* gene using two sets of primers. One set was specific for the boundaries of the *trhA* gene (CRAIG10b and CRAIG11b) and one set specific to 400bp upstream and downstream of the gene (MIC49 and MIC50). The fragments were then ligated into pGEM-T. As a positive control the 3.9kb *Hin*dIII fragment of pCS21 was excised by endonuclease digestion and blunt-end ligated into the insertion site of pGEM.

These recombinant plasmids were introduced by electroporation into *E. coli* SMR130 (R27), incubated for 1 hour at 37°C in 1ml BHI broth and the cells were mated

with 50µl of mid-log recipient, *E. coli* J53-1 in 20ml of BHI broth, overnight at 28°C. The presence of the disrupted *trh* gene in the recipient was selected by growth on BHI agar plates containing antibiotics to select for the presence of R27 and the recombined cassette within the recipient cells.

30. Procedure for selection of *trhA* recombinants by lethal gene induction.

As linear and vector based recombination mutation of the *trhA* gene did not generate site-specific insertion of the disrupting cassette, selection of *trhA* recombinants by utilizing the presence of the lethal *ccdB* gene within the pZERO cloning vector was attempted. Copies of the PCR generated *trhA* genes containing the Cm cassette disruption were cloned into the unique *Stu*I site found downstream of the *ccdB* gene. The 3.9kb pCS21 *Hin*dIII fragment containing mini-Tn*10* was excised using *Hin*dIII, purified, blunt-ends were generated and the fragment cloned into the *Stu*I site within pZERO-T.

These constructs were introduced by electroporation into *E. coli* RR1. Cells were incubated for 1 hour at 37°C in 1ml BHI broth and mated with 50µl of mid-log *E. coli* J53-1 which were incubated overnight at 27°C. The mating mixture was plated on BHI agar plates containing CmTcNal to select for the presence of the disrupted *trhA* gene in the recipient and 50mM IPTG to induce the *ccdB* gene if present within the recipient cell.

31. Generation of the complete sequence of R27.

Plasmid DNA was isolated from *Escherichia coli* (J53-1) by alkaline lysis and purified by cesium chloride density gradient centrifugation. DNA was prepared for library construction by nebulization, end-repair and size fractionation (92). Recovered

DNA fragments were ligated into M13Janus (23). Library subclones were picked as plaques, from which template DNAs were prepared and then sequenced by Prism Dye-terminator Cycle Sequencing chemistry and analyzed on ABI377 automated sequencers in the laboratory of Dr. F. Blattner, University of Wisconsin (17). The annotated sequence is deposited in GenBank, Accession number AF250878.

32. Annotation

Sequence assembly and identification of open reading frames (ORFs) was performed as described previously (24). Further identification of the ORF start sites was based on ribosomal binding site (RBS) matches upstream of potential starts. In the absence of strong RBS prior to putative start sites, codon usage data was utilized to identify likely starts.

Searches of the protein databases for amino acid similarities were performed using DeCypher II system (TimeLogic Inc.) to assign known functions or suggest functions for new ORFs (17). Further analysis of the ORFs was performed using BLAST sequence analysis tools (6) with subsequent comparison of ORFs showing significant homology (greater than 10^{-3} significance) performed using a Lipman-Pearson algorithm (114). Significant similarity was defined as at least 30% identity observed over 60% of the ORF, though those ORFs showing lower than 30% identity over greater than 60% of the protein were also included. ORFs are labelled ORF001 to ORF210.

G+C content was calculated using the Windows program from the GCG Software Suite Version 9.1 (Genetics Computer Group, Inc.) with the following parameters: window of 1000 nucleotides with a 100 nucleotide shift.

Prediction of transmembrane helices was completed using the TMPred internet-based program (www.ch.embnet.org/software/TMPRED_form.html).

Analysis of ORF158 and comparison to TlpA was done using the COILS2 program (http://www.ch.embnet.org/software/COILS_form.html) (88) (28 residue window, MDITK sequence profile, weighting on) and PARCOIL (http://nightingale.lcs.mit.edu/cgi-bin/score) (13).

33. Cloning of ORFs015-018

Primers incorporating restriction endonuclease sites specific to upstream (*Hin*dIII) and downstream (*Bam*HI) of ORF015 (CRAIG60+61), ORF016 (CRAIG40+41a), ORF017 (CRAIG62+63) and ORF018 (CRAIG64+65) were constructed. Using the A22 cosmid, containing the cloned *Xba*D fragment of R27, the individual genes were PCR amplified using *Pfu* polymerase. The linear fragments were agarose gel purified, digested with *Bam*HI and *Hin*dIII restriction endonucleases and ligated into pZERO-2.

The four ORFs (ORF015-018) were cloned as a single unit using primers incorporating restriction endonuclease sites upstream (*Bam*HI) (CRAIG31) and downstream (*Hin*dIII) (CRAIG65) of the putative transcripition start of ORF018. The restriction sites were reversed so as to allow directional cloning into the pZERO vector opposite to the vectors internal P_{lac} promoter. The genes were PCR amplified using the A22 cosmid as template with *Pfu* polymerase. The resulting fragment was agarose gel purified, endonuclease digested with *Bam*HI and *Hin*dIII, and ligated into pZERO, generating a clone of the four ORFs, without its native transcription promoter site, under the control of P_{lac}. Using the primers CRAIG60 and CRAIG65 with the A22 cosmid as

template, the four ORFS (ORF015-018) were PCR amplified, endonuclease digested (*Bam*HI, *Hin*dIII) and ligated into pZERO, resulting in the four ORFs being under the control of the P_{lac} promoter of the vector.

34. Mutation of ORF018

ORF018 contains a single *Nsi*I restriction site, used for creation of an insertion-mutant of the gene. Unfortunately the pZERO vector, in which the ORF was originally cloned, contains two separate *Nsi*I sites upstream of the *Bam*HI and *Hin*dIII sites used in the original cloning. Therefore it was necessary to transfer the ORF to a vector which does not contain a *Nsi*I restriction site (pUC119).

ORF018 was PCR amplified using the CRAIG64+65 primers with the A22 cosmid as template, digested with *Bam*HI and *Hin*dIII endonuclease, agarose gel purified and ligated into the unique *Bam*HI/*Hin*dIII site within pUC119. The resulting construct was transformed by electroporation into *E. coli* DH10B, and colonies containing recombinant plasmids were selected. A single colony was chosen, and PCR screened for the presence of ORF018 using the CRAIG64 and CRAIG65 primers. This plasmid was then purified, and digested with the *Nsi*I endonuclease.

A linear Cm^R cassette was generated using primers (CRAIG36+37) internal to the mini-Tn*10* transposon, amplifying a 780nt fragment (nt238-1038) containing the Cm^R gene. This was then blunt-end ligated into the unique *Nsi*I site present within the pUC119 cloned ORF018. The plasmid was transformed into *E. coli* DH10B and transposon insertions selected by growth on BHI agar containing Cb and Cm. A 1.4kb

fragment was then released by digestion with *Bam*HI and *Hin*dIII endonucleases and cloned into the unique *Bam*HI and *Hin*dIII sites present in pZERO.

Sequence-specific recombination between the disrupted ORF018 and R27 was performed as detailed previously.

Chapter III LPS Effect on Recipient Ability*

As LPS truncation affected the frequency of conjugation in some plasmid systems, the effects of LPS truncation on the conjugation of the IncHI1 plasmids was assessed. Isogenic *S. typhimurium* containing mutations within the LPS biosynthetic genes provided a range of well-defined LPS truncations which were available for study. As it was possible to use the LPS truncated mutants as both donor and recipient, the role of LPS in the process of conjugation was further elucidated. Observing the change in aggregation of mating mixtures using selected LPS truncated mutants as either donor or recipient was possible and further enhanced our understanding of the role of LPS in IncHI1 plasmid conjugation.

1. Effect of LPS inner core mutations on recipient ability of *S. typhimurium*.

Conjugation experiments in liquid medium were carried out with the IncHI1 plasmids R27, its transfer-derepressed derivative pDT2454, and the IncHI2 plasmid R478; using *E. coli* J53-1 as the donor. Isogenic LPS mutants of *S. typhimurium* SL1027 were used as recipients and transfer frequency determined. The transfer frequencies of these plasmids were reduced to the limit of detection (less than 10^{-8} transconjugants per recipient cell) in the *S. typhimurium* recipients with *rfaD* and *rfaE* mutations (Table III-1). In contrast, mutation of the inner core by *rfaF* increased transfer frequency, relative to wild-type, for the R27, pDT2454, and R478 plasmids. Complementation of the *rfaF* mutation by introducing, by electroporation, a plasmid carrying the wild-type *rfaF* gene, pKZ1100 (140), reduced transfer frequency to the observed wild-type levels.

Table III-1. Conjugative transfer of IncH plasmids to *S. typhimurium* SL1027 derivatives in liquid medium

Strain	Partial Phenotype	Transfer frequency at 27°C[a]		
		IncHI		IncHI2 (R478)
		R27	PDT2454	
SL1027	Wild-type	$<1 \times 10^{-8}$	3.4×10^{-6}	6.3×10^{-6}
SA2205	*rfaL*	$<1 \times 10^{-8}$	3.6×10^{-6}	1.6×10^{-6}
SL733	*rfaK*	$<1 \times 10^{-8}$	3.7×10^{-6}	2.2×10^{-6}
SA1905	*rfaJ*	$<1 \times 10^{-8}$	7.5×10^{-6}	9.8×10^{-6}
SA2206	*rfaI*	$<1 \times 10^{-8}$	4.3×10^{-6}	2.0×10^{-6}
SL1060	*rfaH*	$<1 \times 10^{-8}$	3.5×10^{-7}	7.7×10^{-6}
SL1183	*rfaF*	2.6×10^{-6}	5.5×10^{-4}	4.0×10^{-3}
SL1102	*rfaE*	$<1 \times 10^{-8}$	$<1 \times 10^{-8}$	$<1 \times 10^{-8}$
SL3600	*rfaD*	$<1 \times 10^{-8}$	$<1 \times 10^{-8}$	8.2×10^{-7}

[a] Transconjugants per recipient cell and average of three experiments. Determined by an 18 hour mating in Penassay broth (see Materials and Methods)

The *rfaI*, *rfaJ*, *rfaL*, and *rfaK* mutations, resulting in truncation of the outer core, LPS had little effect on recipient ability. The *rfaH* mutation, believed to be a positive regulator required for expression of several enzymes including those responsible for addition of galactose (Gal) from UDP-Gal to Gal-deficient LPS cores (133), had no obvious effect on H plasmid transfer.

The transfer frequency of the plasmid to the recipient was assessed for matings on solid medium (Table III-2). Differences in transfer frequencies, compared to wild-type levels, correspond to those observed with matings in liquid medium. Complementation of the *rfaF* mutation with pKZ110, reduced transfer to the observed wild-type levels on solid medium, as was observed with liquid matings.

2. Effect of PEA on transfer efficiency.

The addition of exogenous PEA to matings is known to significantly decrease transfer frequency for the F plasmids (9), however its method of action is poorly understood. The effect of PEA addition on the recipient ability of *rfaF* mutant recipients was assessed. Mating between *E.coli* J53-1 (pDT2454) and wild-type *S.typhimurium* in the presence of 0.1M PEA resulted in a decrease from 1.1×10^{-6} to 6.8×10^{-8} transconjugants per recipient cell, whereas *rfaF* mutant recipient transfer frequency decreased from 1.4×10^{-3} to 1.8×10^{-7}.

Table III-2. Conjugative transfer of IncH plasmids to *S. typhimurium* SL1027 derivatives on solid medium.

Strain	Partial genotype	Transfer frequency at 27°C[a]		
		IncHI1 (R27)	IncHI1 (pDT2454)	IncH2 (R478)
SL1027	Wild-type	1.5×10^{-7}	1.1×10^{-6}	8.4×10^{-3}
SA2205	*rfaL*	$<1 \times 10^{-8}$	$<1 \times 10^{-8}$	5.3×10^{-5}
SL733	*rfaK*	$<1 \times 10^{-8}$	2.0×10^{-7}	2.2×10^{-3}
SL1183	*rfaF*	3.6×10^{-5}	1.4×10^{-3}	4.8×10^{-3}
SL1102	*rfaE*	$<1 \times 10^{-8}$	$<1 \times 10^{-8}$	$<1 \times 10^{-8}$

[a]Transconjugants per recipient cell and average of three experiments. Determined by an 18 hour mating on BHI agar plates (see Materials and Methods). All transfer frequencies within the triplicate assays were within the same log.

3. Transfer from LPS mutant S. *typhimurium* donors to wild-type *S.typhimurium* recipients.

Previous experiments failed to identify if changes in transfer frequency due to LPS mutation was specific to the recipient phenotype. Therefore the frequency of transfer from *S. typhimurium* SL1183 (pDT2454) to *S. typhimurium* SL1027 was determined. Transfer from *rfaF* mutant donors to wild-type recipients was observed to be 1.4×10^{-3}, compared with 3.0×10^{-3} transconjugants per recipient as observed in transfer from wild-type donors to *rfaF* mutant recipients. In the context of 2-log increases or decreases in transfer frequency observed between *rfaF* mutant recipients and wild-type recipients, this change is not considered significant. Transfer using *rfaK*, *rfaL*, *rfaH* and *rfaE* mutant strains as donors to *S. typhimurium* SL1027 resulted in transfer frequencies nearly identical to that observed in plasmid transfer to LPS mutant recipients

4. Determination of aggregate formation by *S. typhimurium* LPS mutants.

Formation of wall-to-wall contact, visualized as bacterial aggregates, is an important step in the conjugal transfer of plasmids. Changes in the outer membrane structure, such as truncation of the LPS, may lead to a change in the affinity of the cells for each other. Therefore assessment of the bacterial aggregation, both in a mating mixture and in a pure culture may lead to a better understanding of the role of LPS in conjugation.

Transmission electron microscopy was used to determine aggregate formation between a donor culture of *S. typhimurium* SL1027 (pDT2454) and four *S. typhimurium* recipients: SL1027 (wild-type), SL1183 (*rfaF*), SL1022 (*rfaE*) and SL1905 (*rfaJ*). The

aggregation of the three LPS mutants carrying pDT2454, with *S. typhimurium* SL1027 was also determined. In turn the percentage of cells observed to be within aggregates was compared to the cells in pure culture and to that observed in matings between the LPS mutant bacteria containing pDT2454 and *S. typhimurium* SL1027 recipient. The results are presented in Table III-3. Changes in aggregation between LPS truncated donors to wild-type recipients, and wild-type donors to LPS truncated recipients were only considered as relevant to this study.

5. Conclusions

Alteration of the surface structures caused by mutations leading to LPS truncation results in changes in transfer frequency of R27, pDT2454 and R478 to a *S. typhimurium* SL1027 recipient. Changes in transfer frequency of IncHI plasmids into LPS mutant recipients do not correspond to observations made in the IncF and IncP systems (9, 130), which may be due to donor and recipient interactions specific for the incompatibility group. Therefore the proteins involved in the initial steps in conjugation may not be conserved throughout the different plasmid systems.

Mutations in the LPS biosynthetic pathway which leave intact a minimal inner core structure, such as the *rfaF* mutation in *S. typhimurium*, appears to allow transfer of the IncHI plasmids at higher frequencies. Approximately 100 fold greater transfer frequency was observed in an *rfaF* mutant recipient compared to wild-type. Other mutations in the LPS biosynthetic pathway decreased (*rfaD*, *rfaE*) or had little effect (*rfaL*, *rfaK*, *rfaJ*, *rfaI*, *rfaH*) on transfer frequency. This finding indicates that the LPS

Table III-3. Mating specific aggregation between LPS mutant recipients and *S. typhimurium* SL1027

Partial genotype	Observed Aggregation[a] (number of cells counted)			Change in aggregation specific to recipient LPS
	Recipient alone[b]	To mutant[c]	From mutant[d]	
wild-type[e]	43 +/-3	49 +/-3	n/a[f]	n/a[f]
rfaE	37 +/-2%	52 +/-3	36 +/-2	+16
rfaF	62 +/-2	68 +/-2	56 +/-3	+12
rfaJ	60 +/-2	65 +/-3	55 +/-3	+10

[a] Aggregation, expressed as a percentage of all cells in an aggregate within the counted population of cells. Margin of error based upon sample size and within at a 95% confidence level.

[b] Aggregation observed in negative control containing recipient only.

[c] Between wild-type *S. typhimurium* SL1027 (pDT2454) donor and LPS mutant recipient

[d] Between LPS mutant strains carrying pDT2454, acting as donors, and wild-type *S. typhimurium* SL1027 recipients.

[e] Mating between *S. typhimurium* SL1027 (pDT2454) and *S. typhimurium* SL1027 (Rif^R)

[f] As both recipient and donor were of the same phenotype (SL1027) there is no orientation of the transfer with respect to recipient.

oligosaccharide, outer core or inner core may interfere with one of the steps prior to, or during, H-plasmid transfer. Complementation of the *rfaF* phenotype with a wild-type gene indicates that the increase in transfer frequency is strictly due to the LPS phenotype.

The *rfaD* and *rfaE* mutant recipients are completely heptose deficient, and this clearly interfered with recipient ability for the IncH plasmids tested. The *rfaD* and *rfaE* mutations are responsible for heptose synthesis and attachment to the nascent inner core, and these mutations result in LPS truncation following the 3-deoxy-D-manno-octulosonic acid residue. Of these inner core mutations resulting in heptose-deficient cores, only *rfaD* mutant recipients displayed any detectable transfer. *S. typhimurium* strains with the *rfaD* mutation have been observed previously to be somewhat leaky (133, 140). Mutations in the *rfaD* gene cause a defect in the epimerization of ADP-L-glycero-D-mannoheptose to ADP-L-glycero-L-mannoheptose, which is normally transferred to the inner core by the *rfaC* gene product. The *rfaD* mutation causes a buildup of ADP-l-glycero-D-mannoheptose, which has been observed occasionally to be incorrectly incorporated into the inner core. The incorporated ADP-l-glycero-D-mannoheptose allows continued core synthesis and leads to a heterogeneous population of LPS molecules, which contains some full-length LPS molecules (129). Therefore, the observed increase in the transfer frequency of the IncHI2 plasmid R478 into *rfaD* mutant recipients, compared to *rfaE* mutant recipients, may be a result of the leaky nature of the *rfaD* mutation.

Increased aggregation was observed in mating mixtures of all mutants tested, compared to that observed in wild-type donor alone (43%), indicating a plasmid specific interaction between the donor and recipient cells, increasing aggregation within the

mating mixture. Aggregation increased in suspensions of LPS-mutant recipients alone, compared to that observed in wild-type recipient suspensions, indicates an increased affinity for LPS mutants to aggregate. Aggregation consistently increased with the introduction of donor cells harboring pDT2454.

In a conjugation mixture recipients with truncated LPS have a greater affinity to form aggregates compared to donor cells with identical LPS truncation. These results indicate that there is a recipient factor that is uncovered due to LPS truncation. As mentioned previously, conjugation can be separated into five stages. It is likely that the truncation of the LPS leads to increased plasmid-encoded pilus binding, leading to a greater percentage of wall-to-wall contact in the cells. Increased pilus binding could likely be the result of the unmasking of a cryptic receptor on the recipient in LPS truncated mutants. As increased aggregation does not necessarily lead to increased transfer frequency, it is clear that even though LPS mutation is of benefit in the first two steps of the conjugation process, with the exception of *rfaF* mutant cells, it interferes with later steps.

The effect of PEA addition on cellular physiology is poorly understood, but its effect on transfer of F plasmids well characterized (9). Addition of PEA to the medium may lead to increased presence of HepI-(P)PEA in the inner core. This may be caused by inhibition of inefficient cleavage of the HepI-(P)PEA by membrane bound TolC, via a shift in the kinetic equilibrium by the relatively large PEA concentration in the medium.

Reciprocal transfer of pDT2454 from LPS mutant *S. typhimurium* donors to wild-type *S. typhimurium* recipients showed no significant variation from that observed in wild-type hosts donating to LPS mutant recipients. Therefore changes in mating

efficiency due to LPS mutation are not dependant on truncation of the recipient LPS, but on truncation occurring on either the donor or recipient. The LPS does not affect the recipient ability of the cells; although it does result in increased plasmid initiated aggregation. This observation indicates that increase in transfer frequencies is not the result of increased receptor availability.

It is notable that only the *rfaF* mutant recipients displayed increased transfer frequency compared to wild-type cells. In turn the *rfaF* mutant recipient was decreased 100 fold more by addition of PEA than the observed decrease in transfer observed in conjugations with wild-type donors and recipients. Yet mating-specific aggregation of *rfaF* mutant recipients is consistent with that observed in other strains with truncated inner and outer LPS cores. It therefore appears that the (P)PEA moiety, uncovered by the *rfaF* mutation is highly related to transfer efficiency. The exact relationship of aggregation to transfer frequency is unknown, with even lower transfer rates due to LPS mutation possible in the absence of increased aggregation.

Micellar membrane fractions containing homogeneous HepI-(P)PEA LPS have been shown to have a decreased binding to divalent cations compared to similar preparations of wild-type membranes (133). Therefore, the (P)PEA substituent appears to mask the charge of the phosphate group. The masking of the overall outer membrane charge by increased PEA presence in the LPS may be responsible for the decreased transfer frequencies observed in the presence of 0.1M PEA. It is possible that plasmid transfer benefits from a charge interaction between HepI-P and components present in either the donor or recipient cell. In turn, the masking of the phosphate's charge through

increased HepI-PPEA content in the membrane may be detrimental to efficient plasmid transfer.

Molecular modeling of the *S. typhimurium* LPS by Katowsky et. al. (72) is consistent with the results of our transfer experiments. Two of their conclusions are of particular relevance to our study. (I) The LPS core is a rigid structure, perpendicular to the outer membrane, with the O polysaccharide at a 90° angle to the core, and is believed to form a felt-like network over the cell surface. (II) The phosphate group on the HepI is always exposed to the media.

The felt-like network formed by LPS containing O polysaccharide does not appear to physically interfere with the efficient transfer of the IncH plasmid, as truncation of the LPS prior to the O polysaccharide has little effect on H plasmid transfer. This study has shown the importance of the HepI(P)PEA residue and effects of altered phosphorylation of the HepI residue due to increased PEA in the media, while molecular modeling demonstrates that this group will always be exposed and may possibly interact with the partner cell. Complete removal of the HepI, and therefore removal of the HepI-(P)PEA via the *rfaD* and *rfaE* mutations, may decrease transfer due to absence of this (P)PEA moiety. Regardless of the exact affect of PEA on substitution of HepI-(P)PEA, PEA hinders IncH plasmid transfer, indicating that the charge of the HepI substitution is involved in transfer.

Chapter IV R27 Conjugal Pilin

Using sequence analysis and data obtained from matrix assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry, the pilin subunit was found to be encoded by ORF034 of R27, which was denoted as *trhA*. The presence of a homologue to a peptidase found in the IncP plasmids in R27, allowed a greater understanding of the potential processing of the R27 pilin. Using sequence specific recombination, mutation of the *trhA* gene in R27 was attempted. Using a chloramphenicol disrupted 3.9kb fragment of the *trhC* gene as a positive control, mutation of the *trhC* gene in R27 was possible through sequence specific recombination. R27 plasmids with mutations in the *trhA* gene were not isolated using identical methods.

1. Physical assay of purified pilin.

H-specific pilin was purified from the pil$^-$, fla$^-$ host *E. coli* JE2571-1 (pDT1942) by previously described methods. The resulting product was assayed by high resolution SDS PAGE and visualized using silver or Coomassie blue staining. Silver staining produced a single band, 3.2kDa in size whereas bands were not visible following Coomassie blue staining (Figure IV-1). These results are identical to those obtained in previous H-pili purifications (99).

The purified pilus sample was then assayed by MALDI-TOF (165), performed in the Department of Chemistry at the University of Alberta. Two peaks were observed, a major peak at 7,234Da with a minor 3,630Da peak (Figure IV-2).

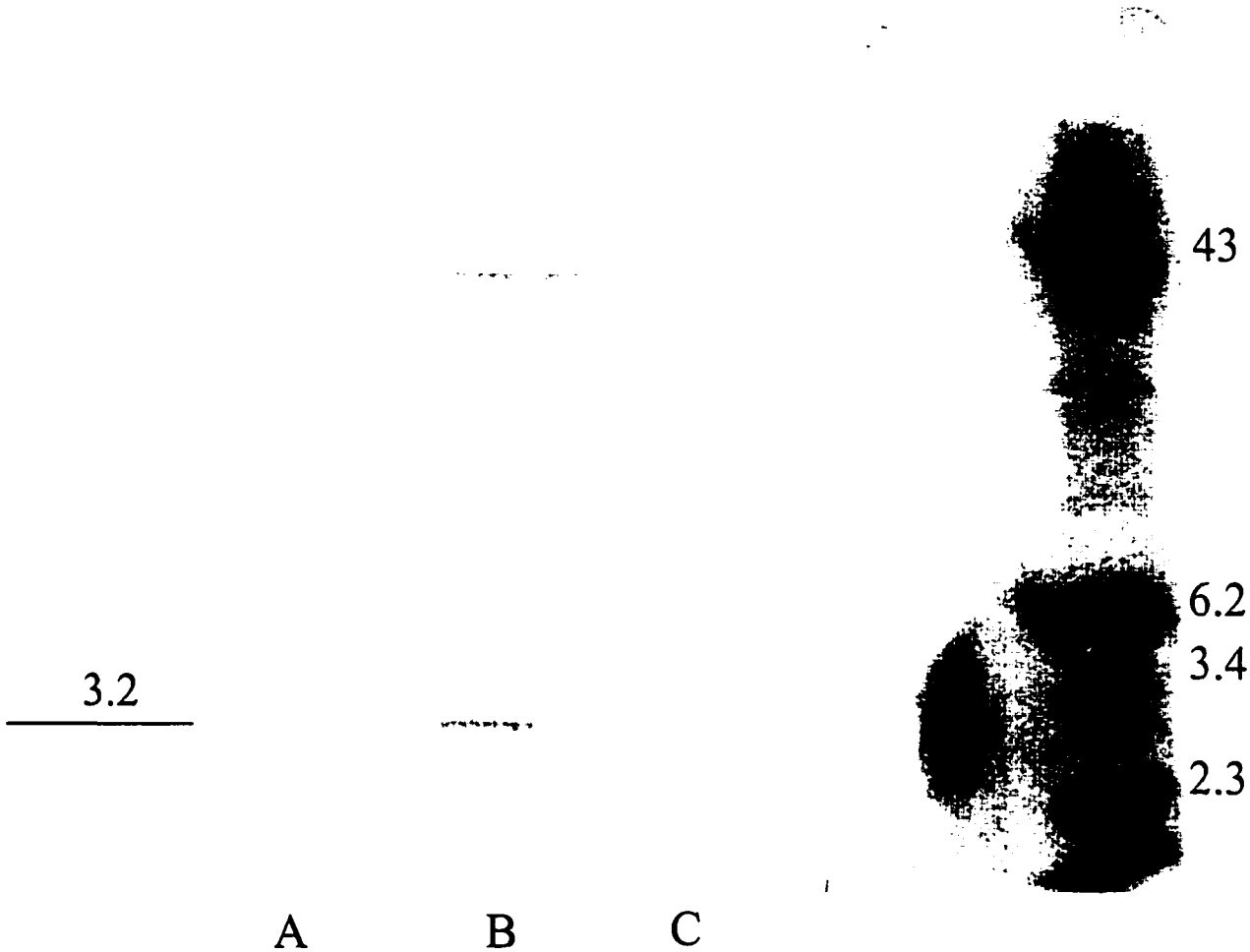


Figure IV-1 Silver staining of purifiedR27 pilin

Pili purified from *E. coli* JE2571-1 (pDT2454) were purified as previously described, SDS PAGE separated and silver stained as described in Materials and Methods. Marker sizes are listed in kDa

Lane A Final purified pilin

Lane B Concentrated pilin, prior to purification by repeated Ammonium sulphate precipitation.

Lane C Loading buffer, negative control.

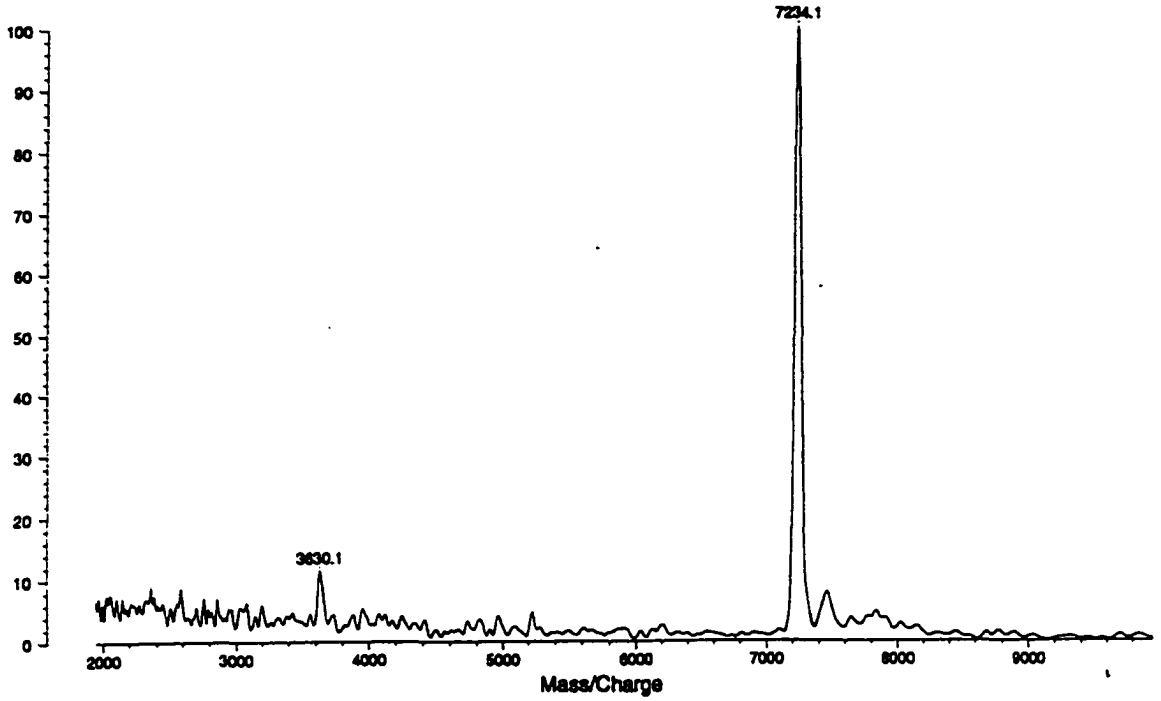


Figure IV-2 MALDI-TOF mass spectrometry of pilin.

Pili were purified from *E. coli* JE2571-1 (pDT2454) and purified as previously described. The purified preparation, displaying a single 3.2kDa band following SDS PAGE silver-staining, was assayed by MALDI-TOF spectrometry. Mass/charge (m/z) is listed in Daltons.

Though the presence of a double-charged protein is common in mass spectrometry (appearing as a peak at half the original protein's weight), the peak is 13Da lighter than would be expected if it was doubly charged. Only a 3.2kDa band is visualized in SDS PAGE, likely the 3630Da peak observed in MALDI spectroscopy of the purified pilin sample. Therefore the major constituent of the purified pilin sample is not separated or visualized in SDS PAGE analysis, even using 6M Urea as a denaturant.

2. Scanning for the pilin gene using amino acid content.

Over 30,000bp of DNA sequence was compiled (126) encompassing the complete Tra2 region of the plasmid R27, sequence flanking the Tra2 region, and the *Sal*I-G fragment (previously identified as being involved in pilin expression through miniTn10 mutagenesis). This DNA sequence was then translated into the corresponding 6 frames of amino acid sequence (3 forward, 3 backward) and compiled as a string of over 180,000 single letters, representing each amino acid and a stop codon. This was then used as input for the program constructed for sequential scanning and assay of amino acid content using a frame size of 25 and the previously determined amino acid composition of the purified pilin (2xAsx, 2xThr, 4-5xSer, 2xGlx, 1xGly, 2xAla, 2xVal, 1xIle, 2xLeu, 2xLys) (99).

Various stringencies were used, allowing for some error in the DNA sequence or possible errors in the amino acid content of the purified pilin. No frame of encoded amino acids was identified in the sequence resembling the amino acid content of the purified pilin. Purification of H-specific pilin was only successful in identifying the size of the pilus subunits. N-terminal sequencing of the purified pilin was unsuccessful (99) and the

extremely low production of R27 pili and its derepressed derivative pDT1942, prevented the use of C-terminal or mass spectrometric amino acid sequencing. Even the use of the previously determined amino-acid content of the purified pilin proved ineffective in identification of the pilin gene through sequence analysis, therefore attention was turned to mutation of the pilin gene by random transposon mutagenesis.

3. Assay of pilin production through phage adhesion

To identify genes involved in pilus synthesis and to derive a new screening assay for detecting pilin–deficient mutants, 45 transfer deficient mutants, previously generated by mini-Tn*10* insertion (99), were tested by phage adhesion assay. *E. coli* J53-1 was used as a host for each of the 45 transfer-deficient mutants tested for the ability of the H-pilus specific bacteriophage Hgal, with a ^{35}S labeled methionine incorporated, to adhere. Of the 45 transfer-deficient mutants tested, 37 were positive in the phage adhesion test and 8 were observed to interfere with absorption of Hgal bacteriophage using a crude phage preparation. These 8 *E.coli* J53-1 strains containing plasmid mutants were then retested using purified radiolabeled Hgal bacteriophage. Of these only two mutant-containing *E.coli* strains were identified as unable to absorb the bacteriophage. The plasmids, designated pDT2972 and pDT2973 (Figure IV-3), had been previously identified as having insertions in the *trhB* and *trhC* genes respectively.

Previous typing of the 45 transfer-deficient mutants of R27 had identified only 11 which allowed formation of plaques on addition of the Hgal bacteriophage, even though 43 of the 45 were competent for radiolabeled-phage adhesion (Table IV-1). Yet the bacteriophage are clearly H-pilus specific. Although the bacteriophage are able to adhere

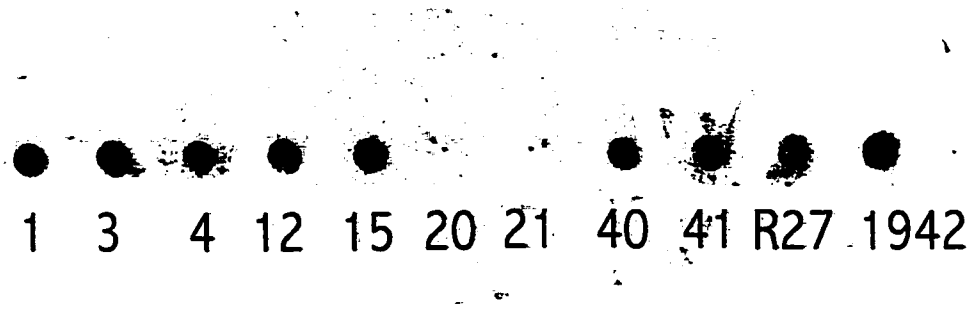


Figure IV-3 Adhesion of radiolabeled phage to cells containing transfer-deficient pDT1942.

E. coli RG192-2 was adhered to a nitrocellulose filter and incubated with concentrated, radiolabeled Hgal bacteriophage. After washing the filter was exposed to Xray film for 36 hours. The numbers correspond to transfer-deficient mutants, 1 (pDT2953) to 45 (pDT2997) of the derepressed for transfer derivative of R27, pDT1942 generated previously (100). Only transfer-deficient mutant 20 (pDT2972) and 21 (pDT2973) prevented binding of radiolabeled phage. Results were consistent throughout four binding assays.

Table IV-1 Plaque formation, pilin expression and phage adhesion of transfer deficient R27

Gene Insertion	Plasmid[a,b]	Plaque Formation[a]	Pilin Expression[c]	Phage Adhesion[d]
trhB	pDT2970	-	+	+
	pDT 2972	-	+/-	-
	pDT 2985	-	+	+
trhC	pDT 2955		+/-	+
	pDT 2957	-	+	+
	pDT 2958	-	+	+
	pDT 2959	-	+	+
	pDT 2962	-	+	+
	pDT 2969	+/-	+	+
	pDT 2973	-	+/-	-
	pDT 2978	-	+	+
	pDT 2990	-	+	+

[a] Newnham, 1995(99)

[b] Rooker *et al*, 1999 (126)

[c] Pilin expression was established by observation of antibody labeling of a 3.2kDa band following SDS PAGE separation of whole cell protein. Antibodies were raised in Rabbits inoculated with purified pilin.

[d] This study

to cells carrying mutant R27, when tested using radiolabeled phage adhesion, they are unable to sustain a lytic infection, thereby implicating the transfer apparatus of R27 as an integral part of the infectious process of Hgal. Of note was the observation that pDT2973 (#20) and pDT2972 (#21) specify weak pilin expression as determined by SDS-PAGE immunoassay (Table IV-1). Although located in genes involved in pilus assembly, these specific insertion events have apparently disrupted pilin expression in an unknown manner. Other mini-Tn*10* insertions within the *trhB* (pDT2970, 2985) and *trhC* (pDT2955, 2957, 2962, 2979) genes prevented lytic bacteriophage infection, but did not affect pilin expression or phage adhesion.

Previous analysis of transfer-deficient R27 mutants utilized antibodies generated in rabbits inoculated with purified pilin (99). These antibodies appeared to recognize a 3.2kDa band in SDS PAGE separation of whole cell protein preparations. Yet there is a correspondence between the transfer-deficient plasmids pDT2972 and pDT2973, which produced low amounts of this 3.2kDa protein and also did not allow Hgal bacteriophage adhesion. This effect is not directly due to the insertion of the mini-Tn*10* transposon into the *trhB* and *trhC* genes, as other mutants of these genes express the 3.2kDa band. Although the 3.2kDa band appears to be a minor component of the pilin sample it is likely associated with pilin expression in R27 as phage adhesion indicates.

4. Identification of the pilin gene through sequence analysis

Sequence analysis in conjunction with transposon mutagenesis (99) identified two regions responsible for conjugative transfer of the plasmid, transfer region 1 (Tra1) and transfer region 2 (Tra2). Sequence similarity to the IncF plasmid group clearly identified

Tra2 as being responsible for pilin synthesis and mating pair formation (see below). Therefore the Tra2 region was the primary focus in the search for the pilin gene.

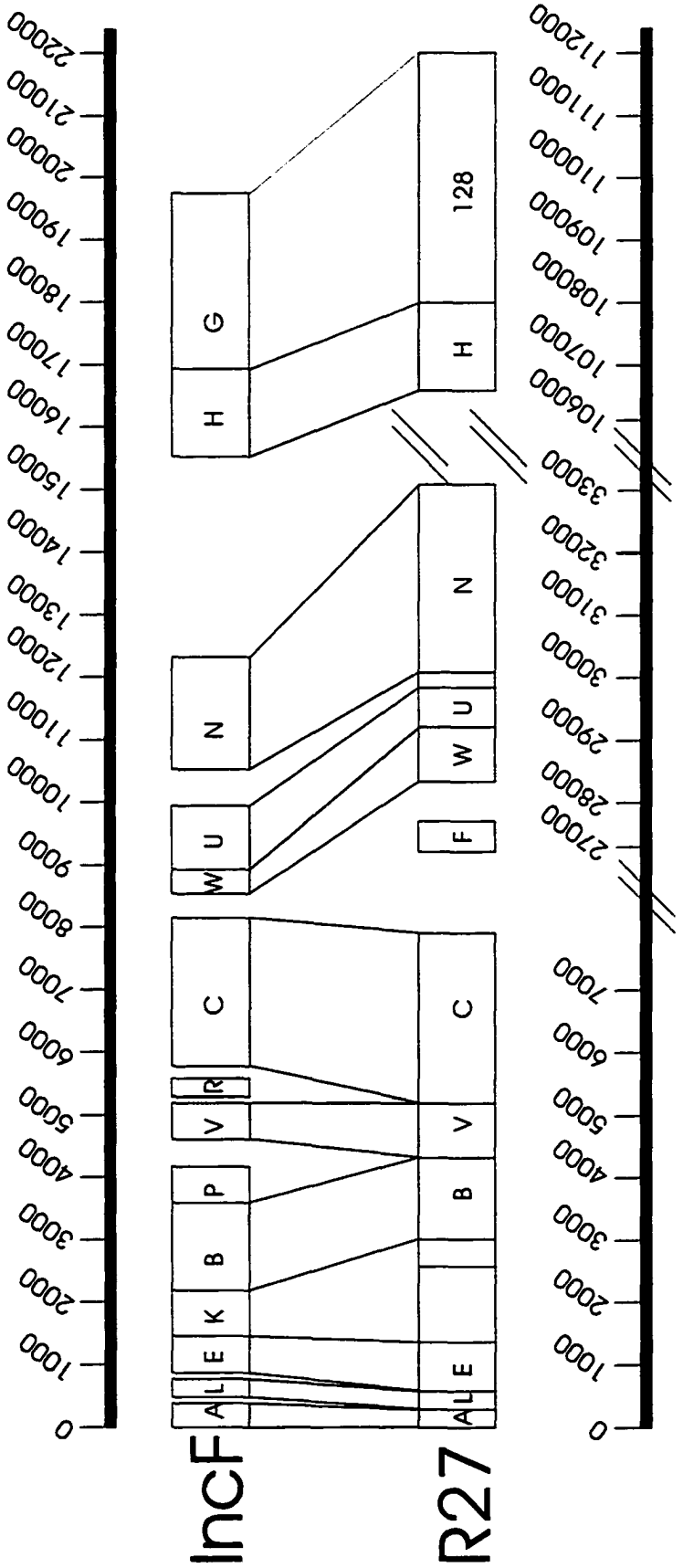
There is a strong similarity in the gene order between the transfer regions of the IncF family of plasmids (with notable exceptions outlined later) and the Tra2 region of R27 (Figure IV-4). As the Tra2 region has a similar order to that seen in the F system, initial attention was turned to ORF034, the first gene in the Tra2 region and designated as *trhA* (see later), as the most likely candidate for the pilin gene in R27 (126).

ORF034, encoding a protein of 188a.a., is similar in size to the pilin genes of related systems: 121 a.a. encoded by the IncF pilin gene *traA* and 145 a.a. encoded by the IncP pilin gene *trbC*. Even though the Tra2 region of R27 shows stronger similarity, on an individual gene basis and overall organization to the IncF plasmid group, ORF034 shows stronger similarity to the IncP pilin TrbC, 19.4% amino acid identity over 93 amino acids, compared to 20.5% over only 38 amino acids for the IncF pilin TraA (Figure IV-5). In addition ORF034 contains a predicted N-terminal leader peptidase cleavage site at amino acid 43 (102) and a nearly identical hydropathicity profile to the IncP plasmid family pilin: TrbC of the IncPα plasmid RP4 and the IncPβ plasmid R751 (Figure IV-6).

Taking into consideration the N-terminal signal peptidase cleavage sites in ORF034 and TrbC of RP4 and the C-terminal processing TrbC is known to undergo (40), the resulting peptides are extremely similar. Of particular note is the presence of the four C-terminal amino acids (amino acids 114-117) of ORF034. Within the IncP plasmid RP4, the plasmid-encoded protease TraF (40), cleaves the four C-terminal amino acids from TrbC (amino acids 119-122) after host-encoded proteolysis. ORF034 appears to be

Figure IV-4 Comparison of the transfer region of the F plasmid and the R27 Tra2 region.

The mating pair formation and pilus synthesis genes of the F plasmid aligned with the Tra2 region of R27. The Tra2 of R27 is disrupted by a 20kb insertion between *trhC* and the *traF* homologue, as indicated by the slashes. There appears to be further homology to the F transfer genes in the Tra1 region of R27 as indicated by the inclusion of *trhH*. ORF128 contains motifs consistent with the F plasmid *traG* gene.



truncated compared to TrbC, with no obvious C-terminal motifs for recognition by host-encoded proteases. Nevertheless there does exist a TraF homologue (ORF008, see later) in R27. Removal of the four C-terminal amino acids in ORF034, as observed in RP4, would result in a protein of 7.22kDa, which coincides with the observed mass of the purified pilin (Figure IV-7), as determined by MALDI time-of-flight spectroscopy. Based on these results the gene corresponding to ORF034 was named *trhA* and is the likely candidate for the pilin gene within R27.

The putative N-terminal cleavage product of the *trhA* gene (amino acids 1-42) would result in a peptide of 4.7kDa. This is markedly larger in size than the 3.6kDa pilin-associated peptide observed in MALDI spectroscopy and the 3.2kDa band observed in SDS PAGE analysis of purified pilin.

5. Labelling of TrhA

Primers specific to upstream and downstream regions of the *trhA* gene were constructed (MIC32 and MIC33), the *trhA* gene was amplified and cloned into the pGEM-T cloning vector. The plasmid was then used as the template for the Promega cell-free coupled transcription/translation system, incorporating ^{35}S methionine label. The *trhA* gene product was expressed, separated by SDS-PAGE and visualized by exposure of the gel to X-Ray film (Figure IV-8). Although clearly expressed in the cell free system, over expression within *E. coli* did not render a band corresponding to either the full sized product of 12.4kDa or the N-terminally processed protein of 7.2kDa follwing Coomassie staining.

Previous attempts to stain purified pilin, clearly visible by electron microscopy, resulted in the staining of a single 3.2 kDa band, which MALDI-TOF spectrometry had

Figure IV-5 Alignment of ORF034 to the IncP pilin precursor, TrbC and IncF TraA.

An alignment of the putative pilin precursor of R27 (ORF034) to the pilin precursors (TrbC) of the IncPα plasmid RP4 (RP4_TrbC) and/or IncPβ plasmid R751 (R751_TrbC).

PanelA: An alignment using the algorhythm of Lipman and Pearson (FASTA) (85).

PanelB: Clustal alignment of the RP4 TrbC (Ser37-Gly114) following N-terminal signal peptidase, host encoded C-terminal peptidase and TraF modification (Figure IV-7)(40) and ORF034 (Lys29-Lys117) following predicted N-signal peptidase cleavage

A. Lipman and Pearson alignment of ORF34 and pilin genes.

```
          v30        v40        v50        v60        v70        v80        v90
ORF34     LFIL-LVLANCSFAYAGSDDGAFG-DIW-AYMSEALTGAPG---KIIACGMLFSVAYFGVVKPNLGLA
          :F:L L.L:. :   :.:.:G::  : W : :.:::TG: :   .II:  :  :V  FG    .:  :
RP4_TrbC  FFVLALALSAHPAMASEGTGGSLPYESWLTNLRNSVTGPVAFALSIIGIVVAGGVLIFGGELNAFFRT
                 ^30        ^40        ^50        ^60        ^70        ^80

                      v100       v110
ORF34     LV-SALMMLVMANGEKIISSFLDAG
          L:  .L:M ::..::::::S:F:: G
RP4_TrbC  LIFLVLVMALLVGAQNVMSTFFGRG
             ^90        ^100       ^110
```

```
          v30        v40        v50        v60        v70        v80        v90
ORF34     LFILLVLANCSFAYAGSDDGAFGDIW-AYMSEALTGAPG---KIIACGMLFSVAYFGVVKPNLGLA
          L ::L: :: :FA .G::::   : W : :.:::TG: :   .II:  :  ::  FG    .:  :
R751_TrbC LAFFLLAPQHAFASEGTGGSLPYESWLTNLRNSVTGPVAFALSIIGIVVAGGILIFGGELNGFLRT
               ^30        ^40        ^50        ^60        ^70        ^80
                      v100       v110
ORF34     LV-SALMMLVMANGEKIISSFLDAG
          L:  .L:M ::..::::::S:F:: G
R751_TrbC LIFIVLVMGLLVGAQNMMSTFFGRG
          ^90        ^100       ^110
```

```
                v60        v70        v80        v90
ORF34     WAYMSEALTGAPGKIIACGMLFSVAYFGVVKPNLGLAL
          W. ::E.L.GA   ::: .: F :A F:::.  :::::
IncF_TraA WVVLAEVLVGAVMYMMTKNVKF-LAGFAIISVFIAVGM
          ^80        ^90        ^100       ^110
```

B. Clustal Alignment of RP4 TrbC and ORF034 of R27, post processing.

```
ORF34         LFILLVLANCSFAYAGSDDGAFGDIWAYMSEALTGAPG------KIIACGMLFSVAYFG- 53
              :*:* :  ..  * *... *.     :::::  ....*      .**.  :  .*  **
RP4_TrbC      FFVLALALSAHPAMASEGTGGSLPYESWLTNLRNSVTGPVAFALSIIGIVVAGGVLIFGG 60

ORF34          VVKPNLGLALVSALMMLVMANGEKIISSFLDA 85
               ::. :   :. .*:* ::..::::*:*:. *
RP4_TrbC       ELNAFFRTLIFLVLVMALLVGAQNVMSTFFGR 92
```

Figure IV-6 Hydropathy profile of ORF34 as compared to the pilin precursor genes of the IncP plasmids.

The calculated hydropathy of ORF034 of R27, TrbC of the IncPα plasmid RP4 and TrbC of the IncPβ plasmid R751, based on a Kyte and Doolittle scale (77) (Window=9). Computation was performed using the ProtScale Web-based software (http://www.expasy.ch/cgi-bin/protscale.pl). Comparison of the hydropathy profile of ORF034 to random peptides, or other pilin proteins (IncF, IncW, Ti) displayed no similarity in profiles.

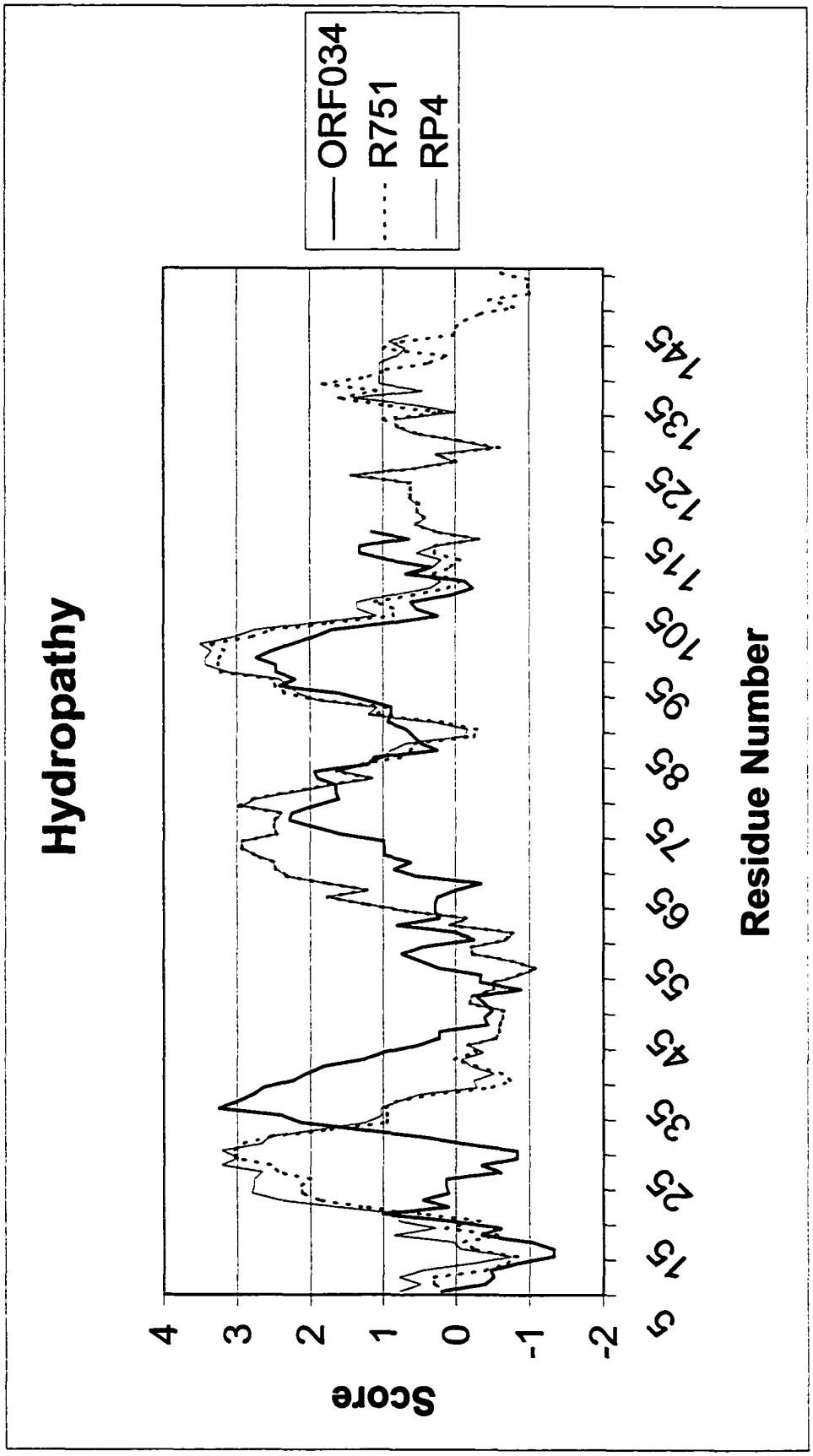
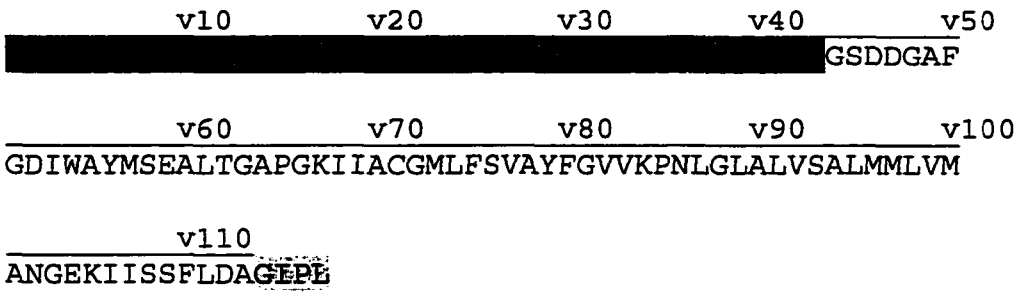


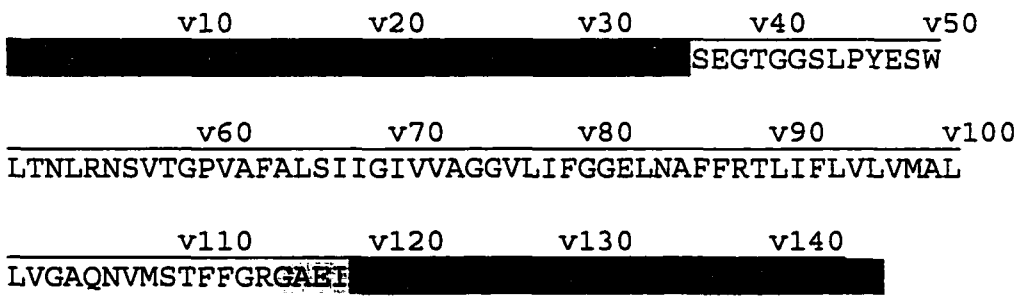
Figure IV-7 Truncation of the R27 ORF034 and the pilin precursor, TrbC, of the IncPα plasmid RP4

Depiction of the hypothesized truncation of ORF034 (TrhA) of R27 and the observed cleavage of the IncPα plasmid RP4 (40). Peptide fragments cleaved by host encoded N- or C-terminal peptidases are identified by dark shading, with the C-terminal cleavage event due to the action of the RP4 plasmid encoded protein TraF (59) shown in light grey. A hypothetical C-terminal cleavage by the R27 encoded TrhF protein is also shaded in light grey. No C-terminal cleavage by host peptidases is believed to be required in R27. ORF34 has a post-processing predicted mass of 7217Da compared to the predicted mass of 8135Da for the processed RP4 TrbC.

R27 ORF034



RP4 TrbC



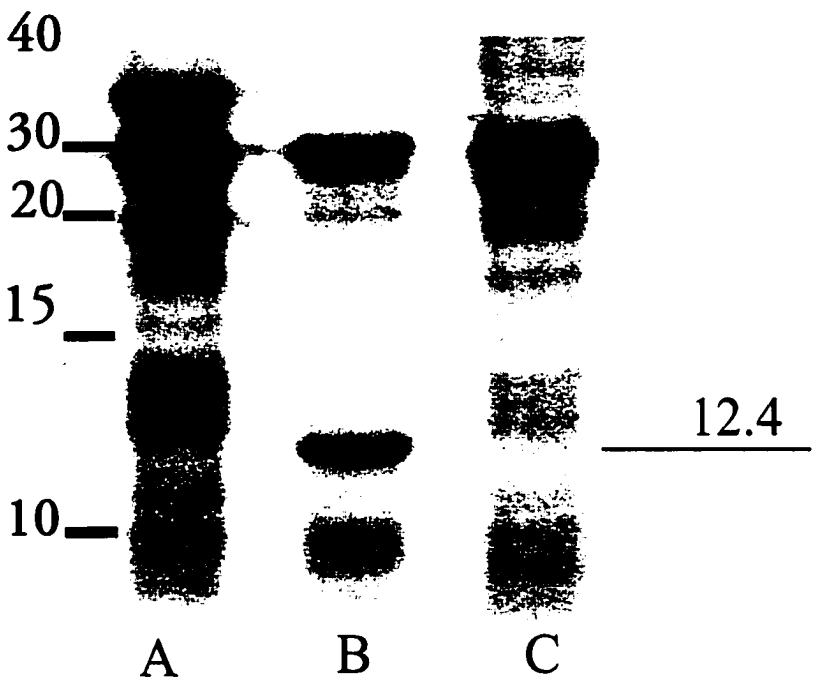


Figure IV-8 Cell free expression of cloned *trhA*.

The cloned *trhA* gene in the pGEM-T vector, expressed in a cell-free coupled transcription/translation system, incorporating ^{35}S label. The 12.4kDa band in LaneB represents the unprocessed *trhA* gene product. Molecular weights listed as kDa.

Lane A: pGEM-T vector without insert.

Lane B: *trhA* cloned into the MCS of pGEM-t

Lane C: No vector.

identified as a minor component of in the sample of purified H pilus. Therefore attempts at labeling the TrhA protein were made, as an alternate method of pilin visualization. Both a N- and C-terminal histidine hexamer (6xHis) labeled TrhA proteins were constructed. Neither of the labeled proteins were observed following expression in whole cells by Coomassie blue or silver staining, or by Western blot analysis using anti-6xHis antibodies. The constructs encoded proteins of the appropriate size (13kDa) which were visible in a cell free expression system containing ^{35}S methionine label. Therefore the protein constructs were expressible and intact in the whole cell system but not visualized by SDS PAGE separation.

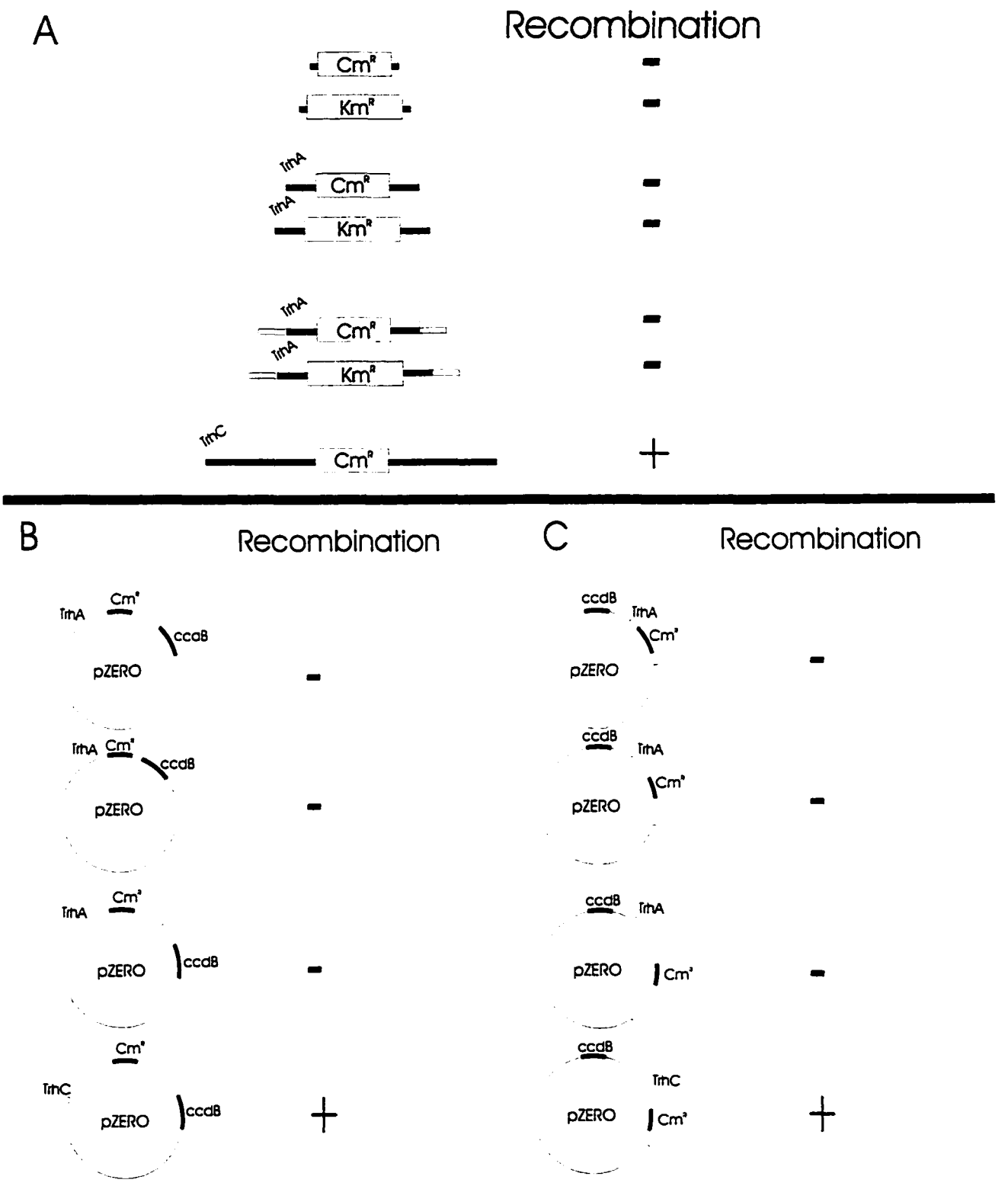
6. Attempts to generate mutations in *trhA*.

The *trhA* gene cloned into the MCS of the pGEM-T vector, pCS19, was disrupted by insertion of either a Km resistance cassette, or a Cm resistance cassette by ligation into the unique *Eco*RV restriction site within the *trhA* gene. These two plasmids, containing disrupted *trhA* genes, were then used in attempts to generate *trhA* mutants within R27 utilizing sequence specific recombination.

Linear fragments of the disrupted *trhA* gene, generated by PCR and the 3.9kb *Hin*dIII fragment from pCS21 containing a Cm^R disrupted *trhC* gene, were introduced into *E.coli* SMR121 (R27) by electroporation (Figure IV-9). *E. coli* SMR121 is hyper-recombinant due to mutations in *recBC*, increasing the recombination rate within the host compared to other host cells. The strain also carries mutations in the *sbcB* and *sbcC* genes increasing the stability of linear DNA within the host. The use of linear DNA for mutation of chromosomal mutants has been documented previously in *Acinetobacter*,

Figure IV-9 Sequence specific recombination attempted in R27.

A diagramatic representation of sequence specific recombination between linear fragments (Panel A) (Section II-28), fragments cloned into the MCS of pZERO-2 disrupting the lethal *ccdB* gene (Panel B) (Section II-29) and fragments cloned downstream of the *ccdB* gene (Panel C) (Section II-30). Fragments cloned downstream of the *ccdB* gene allow induction of the lethal gene, eliminating single recombination events or erroneous presence of the pZERO construct in assumed recombination, which occur when using a vector based template for recombination. Detection of successful sequence specific recombination is listed next to each construct. The positive control, mini-Tn*10* disrupted 3.9kb *trhC* fragment (Cm^R) is included as the last entry in each panel and represents the positive control



therefore we believed that *E. coli* SMR121 could be used to obtain site-specific insertion mutants of R27 (74).

Though no Cm^R or Km^R colonies were observed following introduction of the linear fragments of the disrupted *trhA* genes, Cm^R colonies were formed by the introduction of the 3.9kb Cm^R *Hin*dIII fragment from pCS21. The colonies were selected and the ability of R27 to initiate conjugative transfer was assayed. Of 16 colonies selected nine were transfer-deficient, with seven able to successfully transfer to a recipient host. Of those able to initiate conjugative transfer, none carried the Cm^R to the recipient following conjugation.

A single colony of the transfer-deficient Cm^R hosts was selected and the R27 plasmid purified by the medium-scale plasmid preparation in 20ml of LB broth containing 20μg/ml Cm. The resulting DNA preparation was digested overnight with *Xba*I restriction endonuclease, the endonuclease inactivated, and the resulting DNA fragments ligated into the unique *Xba*I site within pOK12. The recombinant plasmid was then introduced into *E. coli* DH5α by electroporation and grown on BHI agar plates containing Cb and Cm to select for the presence of the vector and the Cm resistant insert respectively.

A single colony was selected and the plasmid isolated by $CsCl_2$ plasmid purification. The plasmid was digested with various restriction endonucleases, transferred to a nitrocellulose membrane, and probed with a mini-Tn*10* specific ^{32}P-labeled DNA probe generated using mini-Tn*10* specific primers (CRAIG1+CRAIG2) (Figure IV-10). The binding of the mini-Tn*10* specific probe corresponds to an insertion

Figure IV-10 Determination of sequence specific recombination using mini-Tn*10* disrupted *trhC*

Hybridization of a ^{32}P labeled mini-Tn*10* specific probe to the *Xba*I fragment cloned from the transfer deficient mutant of R27 derived from recombination with a vector based mini-Tn*10* disrupted *trhC* fragment. The arrow corresponds to the site of recombination of the disrupted *trhC*, as indicated by the hybridization. The observed hybridization pattern indicates that the Cm^R disrupted *trhC* fragment specifically recombined within the R27 based *trhC* gene. DNA markers are listed in kb.

Lane A: *Xba*I digestion

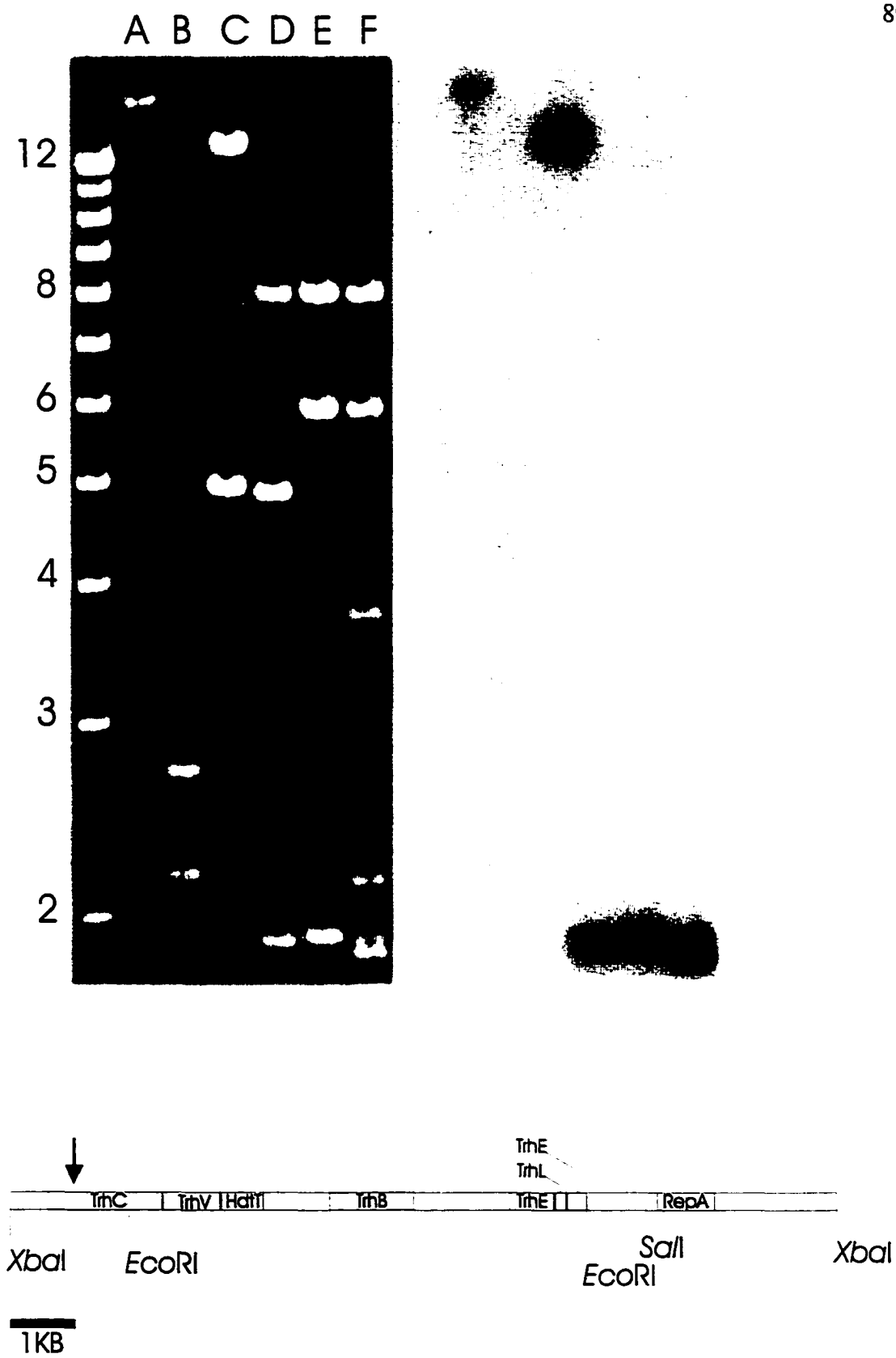
Lane B: *Sal*I *Xba*I digestion

Lane C: *Sal*I digestion

Lane D: *Eco*RI *Sal*I digestion

Lane E: *Eco*RI

Lane F: *Eco*RI *Xba*I



within the 1.7kb *Xba*I *Eco*RI fragment, with an identical restriction profile to that observed with the Tra$^-$ #3 from which pCS21 was derived (Figure IV-10).

Therefore sequence specific recombination between the linear fragments of DNA did occur. Using primers upstream and downstream of the mini-Tn*10* insertion (CRAIG34+35, CRAIG36+37, CRAIG38+39) linear fragments with 900, 400, and 200nt of *trhC* specific sequence flanking the mini-Tn10 insertion, were generated from pCS21. These fragments were then introduced by electroporation into *E.coli* SMR121 (R27) and Cm^R colonies, in which the R27 plasmids were transfer-deficient, could be produced with as little as 200bp of sequence flanking the mini-Tn*10* insertion in *trhC*.

7. Vector-based recombination studies.

The disrupted *trhA* fragments resulting in 600, 150, and 25 bp of *trhA* specific sequence flanking a Cm-resistance cassette were ligated into the insertion site of the pGEM-T PCR cloning vector (Figure IV-9). The plasmids were then introduced into *E.coli* SMR130 (R27) and conjugated overnight with *E.coli* J53-1. The following day the cell suspension was plated on BHI agar containing CmTcRif selecting for the presence of the R27 plasmid in the recipient with the Cm resistance cassette. Although Cm^{R}Tc^{R}Nal^R colonies were generated all the colonies tested (>60) were also Cb^R. This indicates that transfer of the Cm^R cassette due to recombination into the R27 plasmid prior to, or during, conjugation had not occurred. Alternatively the disrupted *trhA* gene, along with significant amounts of the pGEM-T plasmid (including the Cb^R) had recombined into R27 prior to, or during, conjugation.

As a positive control the 3.9kb *Hin*dIII fragment from pCS21 was blunt-end ligated into the insertion site of pGEM-T. As a method of quantitating recombination events as a function of conjugation of R27, the constructs were introduced by electroporation into the hyper-recombinant strain *E. coli* SMR130 (R27), the Rec^+ strain *E. coli* TG1 (R27) and the Rec^- strain *E. coli* RG192-2 (R27). These were then conjugated overnight with *E. coli* J53-1. Rates of conjugative transfer of R27 to the recipient and the percentage of those plasmids carrying the Cm^R were calculated and are presented in Table IV-2. Of the $Cm^RTc^RNal^R$ colonies, 60 were testcd for Cb resistance and all found to be sensitive indicating no abnormal transfer of the vector during conjugation. A sample of 16 colonies were tested for transfer ability and all were observed to be transfer deficient and all carried the mini-Tn*10* sequence as shown by PCR screening using the miniTn*10* specific primers CRAIG1 and CRAIG2 (Figure IV-11).

As recombination events with vector-based disrupted *trhA* genes resulted in carry-over of the pGEM-T vector, or significant quantities of vector DNA, an alternative method of vector-based recombination was used. The pZERO system utilizes the lethal *ccdB* gene, derived from the ccd (control of cell death) locus in the F plasmid, under the control of a Lac promoter. The pZERO-2 vector contains the *ccdB* gene, incorporating a MCS site near the 5' end of the gene. Insertions within the MCS result in disruption of the Lac promoter controlled *ccdB* gene therefore preventing expression of the *ccdB* gene and leading to a selection method for recombinant plasmids. As the *ccdB* gene is under the control of the P_{lac} promoter maintenance of the native vector is possible in a $LacI^q$ background.

Table IV-2 Frequency of transfer, and rate of sequence specific recombination in three *E. coli* strains.

	Colony Forming Units		
	SM130	TG1	RG192
Nal	$8.7x10^{7}$	$2.1x10^{7}$	$1.3x10^{2}$
Tc Nal	$1.5x10^{4}$	$3.5x10^{4}$	$3.7x10^{5}$
Tc Nal Cm	$1.3x10^{2}$	21	$0x10^{0}$
Transfer Frequency	$1.7x10^{-2}$	$1.7x10^{-1}$	$2.8x10^{-3}$
Frequency of transfer with Cm^{R}	$1.5x10^{-7}$	$1.0x10^{-8}$	0
Efficiency of Recombination	$8.8x10^{-6}$	$5.8x10^{-8}$	n/a

[a] Determined following overnight conjugation at 28°C

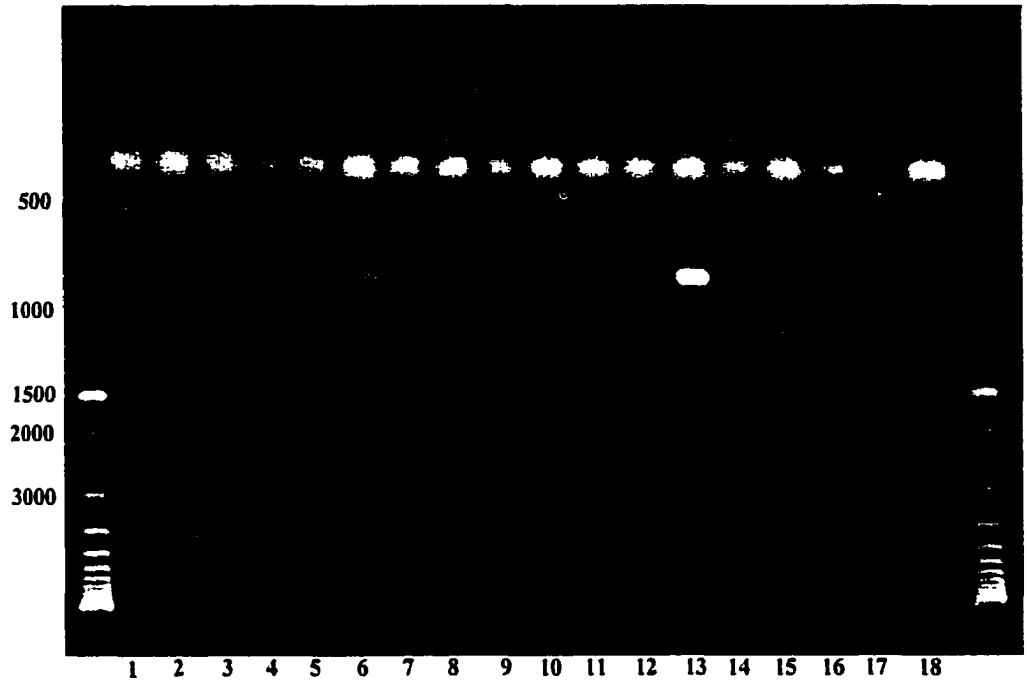


Figure IV-11 Screening of transfer-deficient colonies obtained from recombination of a vector-based disrupted *trhC* gene.

Using the primers CRAIG1 and CRAIG2, specific for a 300kb fragment of mini-Tn*10*, 16 transfer deficient colonies were screened (Lanes 1-16) with a plasmid free *E. coli* J53-1 as a negative control (Lane 17) and $CsCl_2$ purified mini-Tn*10* plasmid (Lane 18). Molecular weight markers are listed in bp. All colonies tested generated a 300bp amplified fragment, though colonies 6, 13, and 14 also amplified anomalously and were not used in further analysis.

The three linear fragments of the Cm resistance cassette disrupted *trhA* genes generated for linear fragment recombination, and the 3.9kb *Hin*dIII fragment of pCS21, were blunt-end ligated into the unique *Stu*I site of pZERO-2 (Figure IV-9). Insertion of the fragments into the *Stu*I site, located downstream of the terminus of the *ccdB* gene, resulted in a recombinant plasmid with an intact and P_{lac} inducible *ccdB* gene. These constructs, were then introduced by electroporation into *E.coli* DT1801 (R27), a Rec^+ host which also contains a chromosomally encoded $LacI^q$ gene. These cells were then immediately mated with *E. coli* RG192-2. Transconjugants carrying Cm^R due to sequence-specific recombination with the disrupted *trhA* gene, or the 3.9kb pCS21 *Hin*dIII fragment were selected for growth on BHI agar containing CmRifTc and 0.1mM IPTG. The presence of IPTG in the selection of the transconjugants should ensure that any recipient cells containing vector DNA are killed by the induction of the *ccdB* gene.

No colonies resulted from mating with disrupted *trhA* fragments while the 3.9kb pCS21 *Hin*dIII positive control generated many colonies. Of the Cm^R colonies tested, all were transfer-deficient. All the colonies tested also contained the mini-Tn*10* transposon as assayed by PCR screening. This clearly indicates that any sequence-specific recombination event within the *trhA* gene present in R27 resulted in the plasmid being unable to transfer or being destabilized once transferred regardless of selective pressure due to antibiotic presence. The disrupted *trhA* constructs were stable in the multi-copy vector pGEM-T as in pCS16 and pCS17, and therefore it is unlikely that the antibiotic cassette disrupted *trhA* genes are lethal to the cells.

To rule out the possibility of promoters internal to the antibiotic resistance cassettes causing aberrant transcription of genes downstream of *trhA*, the orientation of

the Cm and Km resistance cassette disruption was determined. Using the CRAIG5 primer, specific to 90nt upstream the *Eco*RV restriction site within *trhA*, the exact orientation of the cassettes was determined. It was found that the Cm resistance cassette was oriented in the direction of transcription of the Tra2 genes, whereas the Km resistance cassette was oriented opposite to the Tra2 transcript. Therefore the lack of insertions within the *trhA* gene, in R27, is unlikely to be caused by alteration of the transcription of downstream genes.

8. Conclusions.

It therefore appears that the *trhA* gene product encodes a protein which following post-transcriptional modification, corresponds to the 7.2kDa peptide observed in MALDI spectrophotometry of purified pilus. Furthermore both the purified pilus, and the expressed *trhA* gene product were not visible using SDS PAGE separation. The *trhA* gene product is SDS PAGE separable in cell free extracts, but whole cell expression, even in the presence of denaturants, is not visualized in SDS PAGE gels. The creation of N- or C-terminal 6xHis fusion proteins did not aid in the visualization of the *trhA* gene product with SDS PAGE analysis.

These observations with the *trhA* gene product correspond with the analysis of the purified pilus from JE2571-1 (pDT1942). MALDI spectroscopy identified two components to the purified pilus, a 3.6kDa fragment which is visualized in SDS PAGE separations of purified pilin. The second component, a 7.2kDa peptide, is almost exactly the size of the peptide that would result from N-terminal signal peptidase and a C-terminal cleavage hypothetically mediated by TrhF. SDS PAGE separation of purified

pilus resulted in only a single 3.2kDa band, and enzymatic degradation of the purified pili was unsuccessful (99). Furthermore the amino acid analysis of the purified pilin provided ratios which do not coincide with any known gene product in the Tra2 region, nor does it match with the observed amino acids in ORF034/TrhA. It is possible that the previously derived amino acid analysis of the purified pilin was in err, or represents a protein product not encoded within the Tra2 region.

Attempts to mutate the *trhA* gene were unsuccessful, though the cause is unknown as insertional mutation of the isolated *trhA* gene is not harmful to the host cell. Generation of mutations by recombination between Cm^R disrupted *trhC* (pCS21) and R27 was relatively efficient and specific for the gene. It is clear that disruption of the *trhA* within R27 is either lethal to the cell or destabilizing to the plasmid as mutation of other transfer genes within R27 was successful. In addition, a large number of transfer-deficient mutants have been isolated using random transposon mutagenesis, of which none have been observed to occur within the *trhA* gene further indicating that transposon disruption of the gene is disruptive to the plasmid or host.

These results explain previous studies and analysis of R27 pilus. Both the pilin subunits and purified pilus appear to form strong aggregates, which are not visualized in SDS PAGE analysis. Therefore the previous identification of a 3.2kDa band as the pilin subunit appears to be incorrect. Instead a 7.2kDa peptide, possibly originating from a N- and C-terminally cleaved *trhA* gene product, is the pilin subunit, though it does not appear to be circularized as is observed in the RP4 pilin. Circularization was identified in the IncP and Ti plasmids by a decrease of molecular mass by 18Da, a single water molecule, as determined by MALDI compared to calculated molecular mass of the

processed protein (40). Yet in R27 there is an observed increase in the molecular mass of purified pilin as determined by MALDI (7234Da) compared to the calculated pilin mass following processing (7217Da). It is possible that the pilin undergoes further modification, complicating interpetation of the MALDI data.

This 7.2kDa peptide is the major component of purified pilus. The previously identified SDS PAGE visible 3.2kDa peptide does appear to be pilus related as previous Western blot analysis of transfer-deficient R27 using antibodies raised against purified pili correspond to bacteriophage Hgal adhesion studies. As the 3.2kDa fragment does not correspond to the size of the N-terminal portion of TrhA following N-terminal processing, it is possible that this 3.2kDa peptide, a minor component of the mature pilus, is an immunogenic tip protein, perhaps involved in recognition of, and binding to, other cells. Alternatively it is possible that circularization of the processed pilin subunit decreasing its cross-section when traveling through an acrylamide gel, in combination with the high acidity of the peptide, leads to increased running speeds of the pilin through a SDS PAGE gel. Therefore it is possible that the 3.2kDa SDS PAGE band, and the 7.2kDa peptide observed in MALDI spectrometry are the same peptide.

Chapter V Complete DNA Sequence and Analysis of R27*

The entire IncHI1 plasmid R27 was sequenced, and assembled. Analysis of the sequence obtained, allowed identification of the general structure of the plasmid, and allowed comparison of the two transfer regions of R27 to other plasmid systems. It was also possible to extrapolate from sequence data to propose a number of possible mechanisms involved in the temperature-sensitive conjugation of R27.

1. General properties of the R27 plasmid sequence

The IncHI1 plasmid R27 contains 180,461 bp (Figure V-1) which correspond to 210 ORFs initiated by 186 putative ATG, 4 TTG, and 20 GTG start codons. Table V-1 lists the exact position and size of the previously identified ORFs and those ORFs from this study that show similarity to protein sequences available in the public databases. For the latter, the percentage of amino acid identity between the R27 ORF and the prospective homologue is also included. Excluded from this table are the 140 ORFs (66.7%) that do not demonstrate similarity to any known ORFs. A linear map of R27 indicating functional organization is shown in Figure V-2 with all ORFs labeled. The putative ORFs account for 84.4% of the total plasmid, leaving 28,098 bp of intergenic regions. No ORFs directly related to pathogenesis were identified.

The calculated G+C content of R27 is 45.8%, compared with 50-53% for *S. typhi* (64), the usual host for this plasmid, and 50.8% for *E. coli* (17). Within R27, regions of notably different G+C content can be identified (Figure V-3). A peak located at 26,300 bp indicates the high G+C content corresponding to repeats involved in plasmid

* A version of this work has been published as: Sherburne CK, Lawley TD, Gilmour MW, Blattner FR, Burland V, Grotbeck E, Rose DJ, Taylor DE. 2000 The complete DNA sequence and analysis of R27, a large IncHI plasmid from *Salmonella typhi* that is temperature sensitive for transfer. Nucleic Acids Res. 2000 May 15;28(10):2177-2186

Figure V-1 Circular representation of the IncHI1 plasmid R27

A circular represention of the R27 plasmid with the replication and mating stabilization, conjugal transfer, transposon and insertion elements, drug resistance, and citrate utilization regions identified.

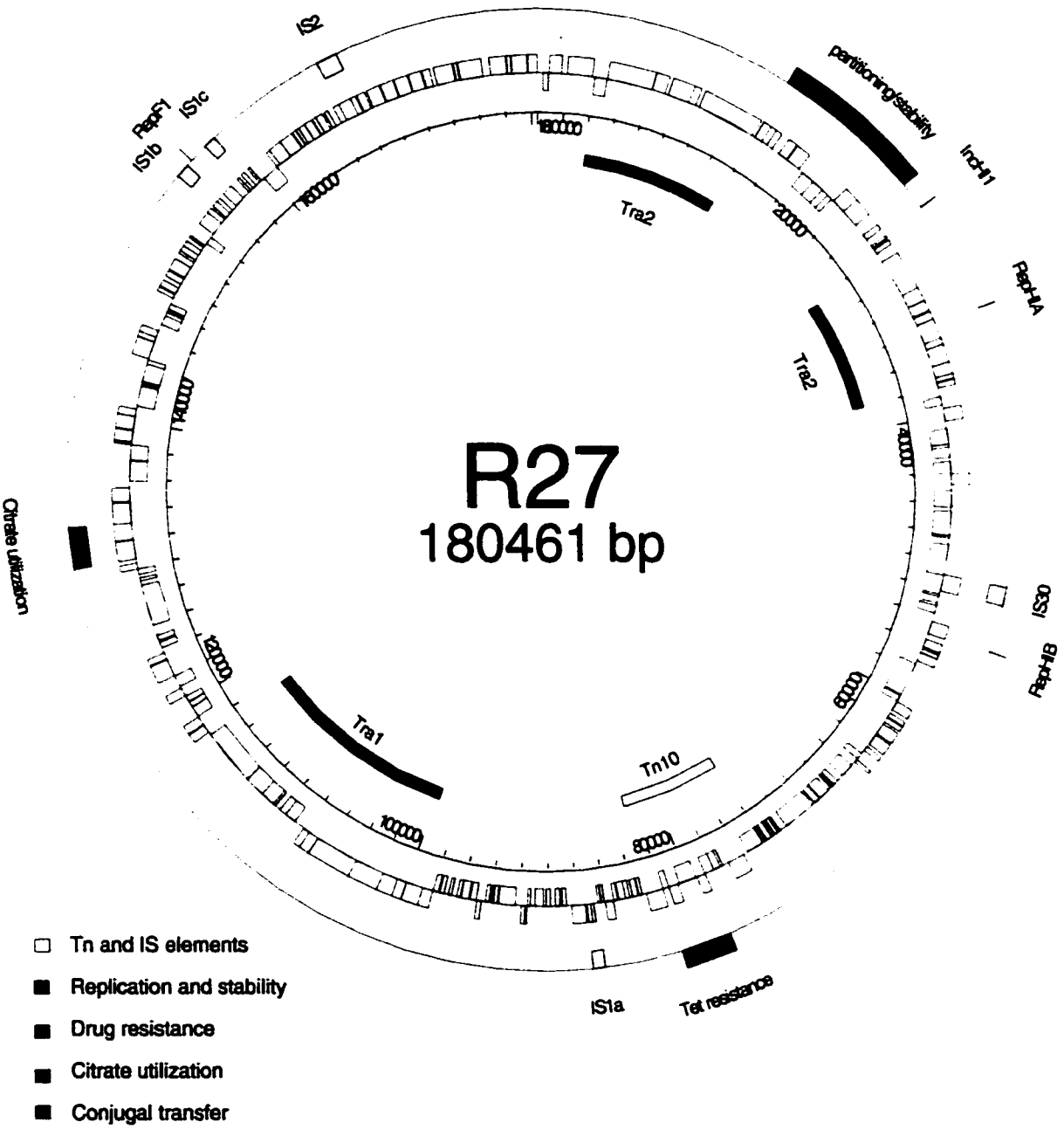


Table V-1 Previously characterized ORFs (termed Tra, Trh, and Htd) and R27 ORFs that have significant homology to other prokaryotic ORFs.

ORF	Coding region	Start codon	Length[a]	Homology or Designation[b]	Identity %/ Range[c]	Reference	GI accession #[d]
003	2116-3900	ATG	594	**TrhI**		(126)	
005	4982-8158	ATG	1058	**TrhN**		(126)	
006	8181-9128	ATG	315	**TrhU**		(126)	
007	9378-10886	ATG	502	**TrhW**		(126)	
008	10873-11385	ATG	170	**TrhF**		(126)/this paper	
010	16077-16640	ATG	187	HtdK. IncHI2 plasmid R478.	100%/79	K. F. Whelan *et al*, unpublished	1695863
011	16637-17071	ATG	144	HtdF. IncHI2 plasmid R478.	100%/132	K. F. Whelan *et al*, unpublished	1695862
012	17061-17513	ATG	150	**HtdA**		(169)	
013	18079-18756	ATG	225	Orfb. IncHI2 plasmid R478.	58.2%/110	K. F. Whelan *et al*, unpublished	1695866
014	18800-19834	ATG	344	StbA. IncHI2 plasmid R478.	79.9%/344	K. F. Whelan *et al*, unpublished	1695865
019	23081-24088	ATG	335	ParB. Prophage/plasmid P1.	27.9%/315	(1)	129631
020	24085-25338	ATG	417	ParA. Prophage/plasmid P1.	25.2%/260	(1)	129629
024	28465-31146	ATG	893	**TrhC**		(126),	

[a] predicted number of amino acids of R27 ORF
[b] previously identified R27 ORFs are in bold, while those R27 ORFs being analyzed for the first time have homologues listed as well as the source organism or genetic element of the homologue
[c] the number of amino acids in a contiguous stretch from which the identity is calculated
[d] if ORF is previously uncharacterized then accession number for the homologous ORF is listed

						(11)	
025	31155-32105	ATG	316	**TrhV**		(126)	
027	32115-32867	ATG	250	HtdT. IncHI2 plasmid R478.	75.0%/92	(110)	4104003
029	33494-34852	ATG	452	**TrhB**		(126)	
032	36519-37304	ATG	261	**TrhE**		(126)	
033	37316-37633	ATG	105	**TrhL**		(126)	
034	37685-38038	ATG	117	**TrhA**		(126)	
036	39480-40355	ATG	291	**RepHIA**		(53)	
041	44007-44780	ATG	257	DNA adenine methylase. *Escherichia coli.*	31.7%/257	(22)	118682
043	46612-48288	ATG	558	Restriction methylase. *Serratia marcescens.*	70.4%/558	D. T. Page *et al*, unpublished.	1809268
046	51035-52186	ATG	383	Transposase for insertion sequence element is*30.*	100%/383	(30)	2851554
050	54356-55237	ATG	293	**RepHIB**		(52)	
073	73125-73301	ATG	58	Unknown. *Myxococcus xanthus*	56.7%/30	(127)	2105134
076	74086-75294	ATG	402	Transposase (insertion sequence IS*10*). *Escherichia coli.*	99.0%/402	(18)	2126090
077	75660-76865	ATG	401	Sodium/glutamate symport carrier protein. *Haemophilus influenzae.*	55.3%/395	(44)	1169972
078	77309-77629	ATG	106	Hypothetical 13.0 kDa protein in	24.8%/103	Y. Kobayashi *et al*,	1731086

				glnQ-ansR intergenic region. *Bacillus subtilis.*		unpublished.	
081	78680-79306	ATG	208	Tetracycline repressor protein TetR. *Escherichia coli.*	100%/208	R.A. Griep *et al*, unpublished.	3435310
082	79385-80590	ATG	401	Transposon Tn*10* TetA protein. *Escherichia coli.*	99.8%/401	(65)	43701
083	80703-81371	ATG	222	Transposon Tn*10* TetC protein. *Escherichia coli.*	99.5%/197	(134)	135616
084	81384-81800	ATG	138	Transposon Tn*10* TetD protein. *Escherichia coli.*	99.8%/138	(134)	135617
085	81810-83018	ATG	402	Transposase (insertion sequence IS*10*) - *Escherichia coli.*	100%/402	(18)	2126090
088	84215-84523	ATG	102	L0009. *Escherichia coli.*	46.9%/49	(115)	3414877
094	86785-87078	ATG	97	Transposase . *Escherichia coli.*	99%/97	(38)	2104506
095	86997-87272	GTG	91	InsA protein . *Escherichia coli.*	97.8%/91	(107)	76271
096	87426-88376	ATG	316	Mg-transport system I.	98.3%/300	(31)	421131

				Escherichia coli.			
097	88746-89027	ATG	93	Hypothetical 27.9 kDa protein in uvrd-cora intergenic region. *Escherichia coli.*	100%/93	(31)	267496
101	91263-91478	ATG	71	*stm* gene product. *Escherichia coli.*	72.7%/44	(56)	42375
111	96195-96431	ATG	78	Hypothetical protein. *Escherichia coli* O157:H7.	35.2%/69	(24)	3822199
119	101815-103899	ATG	694	TraG		This paper	
120	103899-106934	ATG	1011	TraI		This paper	
126	110834-111880	ATG	348	Pilus assembly and synthesis protein. *Sphingomonas aromaticivorans.*	24.1%/212	(124)	3378366
127	111870-113285	ATG	471	TrhH protein precursor. *Escherichia coli.*	22.9%/445	(43)	464925
130	118193-118708	GTG	211	BfpH. *Escherichia coli.*	34.9%/126	M. S. Donnenberg, unpublished.	1122409
131	118856-119605	ATG	249	Thiol:disulfide interchange protein	35.3%/218	(139)	729372

				DsbC precursor. *Erwinia chrysanthemi.*			
135	121245-122054	ATG	269	Outer membrane protein. *Coxiella burnetii.*	22.2%/262	(63)	477583
136	122106-123314	ATG	402	MucB protein. *Salmonella typhimurium*	46.9%/398	R.M. Hall *et. al.*, unpublished.	127507
137	123349-123783	ATG	144	MucA protein. *Salmonella typhimurium.*	47.6%/141	(116)	127504
141	128983-129204	ATG	73	Edolysin. Bacteriophage A511.	29.8%/57	(87)	2493918
144	130558-131853	ATG	431	Citrate-proton symport protein CitA. *Escherichia coli.*	99.3%/431	(131)	116479
145	131908-133047	ATG	379	Citrate utilization protein CitB. *Escherichia coli.*	99.2%/379	(131)	141324
146	133034-134437	ATG	467	Hypothetical protein. *Rhodobacter capsulatus.*	35.3%/464	(163)	3128345
147	134535-135461	ATG	308	Nitrogen assimilation control protein NAC.	32.6%/305	(98)	2495386

				Escherichia coli.			
148	135926-137332	ATG	468	Putative transposase y4bf. *Rhizobium* sp. NGR234.	47.1%/425	(46)	2497379
153	140402-140917	ATG	171	ParB protein. *Escherichia coli.*	39.2%/150	(55)	129520
154	141140-142567	ATG	475	Restriction methylase EcoRII. *Escherichia coli.*	66.0%/447	(144)	127448
156	142891-144210	ATG	439	Unknown. *Sphingomonas aromaticivorans.*	22.4%/331	(124)	3378329
157	144128-144427	ATG	99	Conserved hypothetical protein. *Borrelia burgdorferi.*	34.8%/65	(45)	2688062
158	144482-145702	ATG	406	TlpA. *Salmonella typhimurium.*	21.9%/370	(68)	321935
164	148225-148629	ATG	134	DNA-binding protein H-NS. *Escherichia coli.*	63.2%/134	(119)	123441
177	155773-156276	ATG	167	Hypothetical protein A-167 (insertion sequence IS*1*). *Escherichia coli.*	93.4%/167	(107)	76274
178	156195-	GTG	91	InsA	98.9%/91	(107)	76271

	156470			protein. *Escherichia coli.*			
179	156848-157603	ATG	251	RepE IncFIA	94.8%/249	(53)	
180	158185-158688	ATG	167	Putative protein. Insertion sequence IS*1*	98.6%/167	(71)	149061
181	158607-158891	GTG	94	InsA. *Escherichia coli.*	100%/51	(89)	294461
182	159190-159375	ATG	61	Hha protein. *Escherichia coli.*	46.7%/60	D. Roberts *et al,* unpublished.	1773143
195	166661-167566	GTG	301	Insertion element IS*2* hypothetical 34.4 kDa protein. *Escherichia coli.*	99.7%/301	(125)	140808
197	167524-167889	ATG	121	Insertion element IS*2* hypothetical 13.4 kDa protein. *Escherichia coli.*	100%/121	(125)	140806

Figure V-2 ORFs of R27.

Linear representation of R27, showing all ORFs, their orientation and with color coding by functional groups. The upper rectangles are ORFs transcribed from right to left and the lower rectangles are transcribed left to right. Color coding of functional groups is as follows: yellow, Tn and IS elements; black, replication and stability; red, drug resistance; blue, citrate utilization; magenta, conjugal transfer. The GenBank entry of ORFs were assigned unique identifiers in the form Rxxxx. This figure was created using GeneScene (DNASTAR).

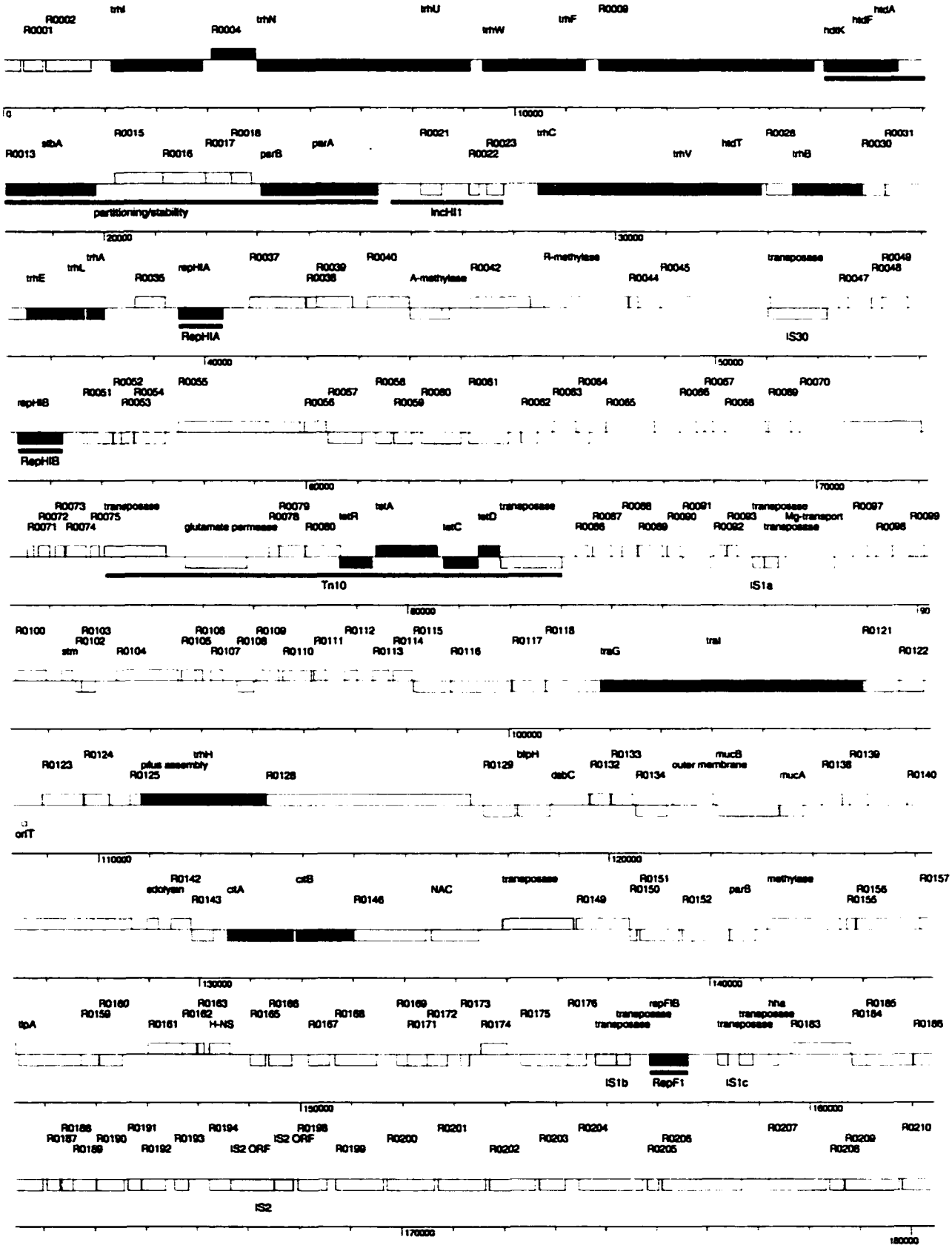
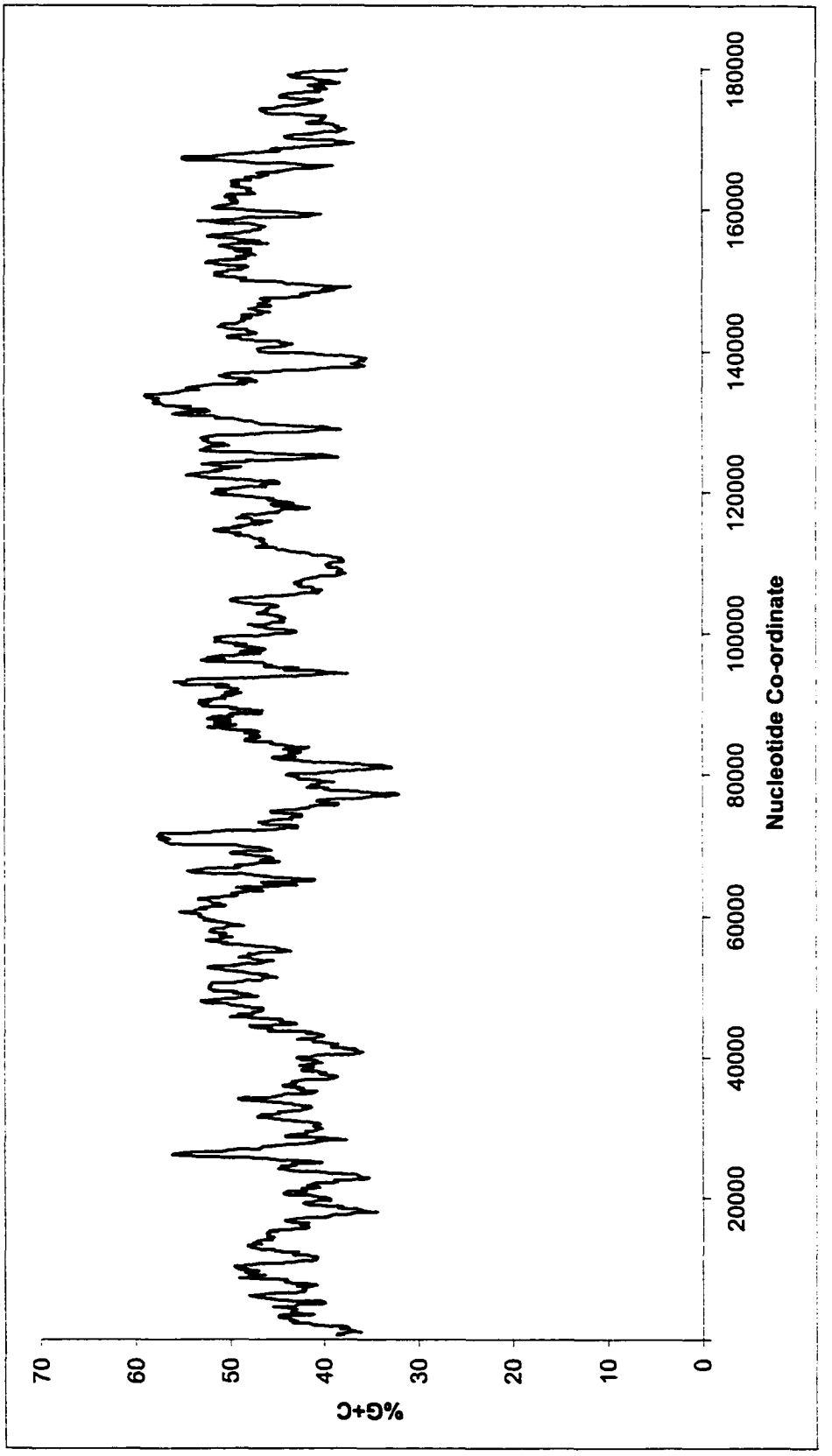


Figure V-3 Graphical representation of the G+C% of the R27 plasmid

G+C profile of the complete DNA sequence of R27. Of note are the peaks at 26,300bp and 133,000bp representing the incompatibility regions (IncHI1) and citrate utilization genes respectively. The troughs observed at 76,300bp and 82,900bp represent the insertion points of the transposon Tn*10*. The figure was plotted using data generated from the GCG program Window



incompatibly (see below). A pair of troughs in the range of 76,300-82,900 bp correspond to the transposon Tn*10*. The intergenic region between ORF077 and 078, and the 5' end of ORF083 (*tetC*) are responsible for these two troughs in Figure V-3 (29.9, and 31.4 %G+C respectively), possibly representing the composition of the original source of the transposon. The peak at 132,400-133,900 bp corresponds to the genes encoding citrate utilization proteins (*citA*, *citB*) and a gene whose product shows significant similarity to a hypothetical *Rhodobacter capsulatum* protein.

2. Transposable elements

R27 contains five insertion sequence elements and one transposon, Tn*10*. In addition, ORF148 shows 47.1% identity to a putative transposase from the large *Rhizobium* sp. plasmid pNGR234 (46). Three copies of IS*1* (designated a-c) are present. IS*1*a is located approximately 13 kbp from the TraI region on the putative lagging strand. One of the coding regions within IS*1*a (*insB*) has accumulated several nonsense mutations. As this would result in premature termination of the protein, this copy of IS*1* is probably inactive. IS*1*b and IS*1*c are found on either side of a DNA segment containing both ORF179 and the repeat sequences that flank ORF179, which displays 94.8% similarity to the F plasmid replication protein E (132).

Two other intact elements, IS*30* and IS*2*, share 100% and 99% identity respectively with previously reported counterparts. Both appear to encode full-length transposases and are predicted to be active. All the IS elements are oriented in the same direction in R27 with respect to their transposases although the significance of this is unknown.

The tetracycline resistance determinant of R27 is encoded within the full Tn*10* transposon by the *tetACD* genes and is controlled by the Tet repressor (ORFs 082-084, and ORF081 respectively) (79). Both the tetracycline determinants in Tn*10* and the transposases (ORFs 076 and 085) are well conserved with respect to their counterparts in *E. coli*, although transposase activity has not been demonstrated within R27.

3. Replication regions

R27 contains two IncHI1-specific replicons and a partial IncFIA replicon (52, 99), with copy number control dependent on an iteron-binding mechanism, in which the replication initiation protein binds to repeated oligonucleotide sequences in the origin. The two IncHI1-specific replication determinants are both able to support active replication of the plasmid and both operate independently of one another (101). Each replicon independently specifies incompatibility (101), where incompatibility is defined as the inability of two related plasmids to replicate in the same host cell. The one-way incompatibility with IncF plasmids is due to ORF179, the R27 replication initiator protein RepE. R27 contains multiple replicons which could give it a competitive advantage over plasmids with only a single replicon and thus increase its host range within the *Enterobacteriaceae*.

The specific IncHI replicons, RepHI1A and RepHI1B, are similar in their overall organization, but the amino acid sequences of their replication initiation proteins are only 34.9% similar (101). The two RepHI proteins are comparable in size (292 aa and 294 aa respectively) and organization of flanking replication protein binding sites: 4 direct repeat sequences upstream and 3 downstream for RepHIA, and 4 upstream and 4 downstream

for RepHIB. The repeats surrounding the gene for RepHIA have been shown to act as iterons for the RepHIA protein (99). The same mechanism is presumed to occur between the RepHIB protein and eight surrounding repeats. The independence of the two replicons in incompatibility is believed to be due to differences in the repeats. The repeat sequences of RepHIA and RepHIB are: "5'-AAAAGCTTTGnATGAATG-3'" and "5'-ATCCACTATACCGGGTA-3'", respectively. Both replicons contain recognition sequences for DNA adenine methylases (GATC) (54) upstream of the Rep genes, and in both cases the recognition sequence is embedded in a 7 nucleotide repeat (AGGATCAA).

4. Transfer regions 1 and 2

The Tra1 region of R27 encodes proteins required for DNA translocation across the membrane during conjugation as well as initial replication events after the plasmid has entered the recipient cell. This is an energy-dependent process in which R27 utilizes host DNA replication machinery and Tra1 encoded proteins. The limits of the Tra1 region were identified by transposon mutagenesis (99). A major element of the Tra1 region is the origin of transfer (*oriT*) from which transfer is initiated by nicking, unwinding, and replication of the leading strand (171). A potential *oriT* is present between ORF122-123 (a 758 bp untranslated region). Within this region are a series of four 33bp inverted repeats, with two potential transcription initiator sequences oriented in opposite directions near the repeats, a common pattern for *oriT* sequences (171).

Although there are no significant database homologies in the region surrounding *oriT*, the R27 Tra1 region has clear organizational similarities to both the IncP and IncW plasmid transfer regions. On the lagging strand (the last genes to enter in conjugation) of

the IncP and IncW plasmids are genes encoding for a relaxase/nickase (*traI* and *trwC*, respectively) and a coupling protein (*traG* and *trwB*) (86). The relaxase/nickase protein introduces a site specific single-stranded nick in *oriT*, with the coupling protein stabilizing the complex and possibly aiding in the recruitment of the cellular DNA replication machinery. Plasmid DNA transfer is thought to utilize the host DNA replication machinery to drive the single-stranded replication product across the membranes into the recipient cell, where it is circularized, and the complementary strand synthesized by host encoded proteins (86).

A pattern of motifs has been reported in the class of related anchoring/coupling proteins of which TrwB of R388, TraD of R100, VirD4 of Ti, and TraG of RP4 are members (11, 86). All of these proteins contain N-terminal transmembrane segments followed by one or more ATP binding domains (86). BLAST analysis of ORF119 identified low similarity (BLAST score 0.78) to the TrwB protein of the IncW plasmid R388. The deduced amino acid sequence of ORF119 contains two predicted transmembrane segments spanning residues 22-40 and 46-64, with an ATP/GTP binding motif "A" spanning residues 205-212. The location of this ORF near *oriT*, its requirement for transfer, as indicated by transposon mutagenesis (99), and the presence of the transmembrane segments followed by the NTP binding domain lead us to propose that ORF119 is an anchoring/coupling protein, and to name it TraG following the IncP nomenclature.

Immediately following TraG in the IncP system is TraI, the protein responsible for the nicking and cleavage of the double-stranded DNA at *oriT* (11). A number of studies have identified a catalytic tyrosine residue in the nicking enzymes found in

different families of plasmids (11, 86). Llosa *et. al.* have identified three regions present in all the proteins mediating nicking and initiation of DNA transfer: I) an active tyrosine residue, II) the residue sequence of "H(nX)S", followed by III) "H(10X)HXHXXXXN" (86). In ORF120 we have found regions in the N-terminus corresponding to each of the three essential motifs for nicking proteins (Figure V-4). We have named the gene for ORF120 *traI* based on the three motifs identified in nicking enzymes and the necessity of this ORF for transfer indicated by transposon mutagenesis (99). TraI is more closely related to the nicking enzymes of the IncP system since it does not contain an identifiable C-terminal helicase-specific motif (11) seen in IncW TrwC and IncF TraI.

ORF126 shows similarity (24.1% over 212 amino acids) to a *Sphingomonas aromaticivorans* pilus synthesis and assembly protein and is necessary for transfer (99). ORF127 has 22.9% identity over the entire length compared to TraH of IncF. *traG* and *traH* of the F plasmid are thought to encode proteins involved in pilus assembly and aggregate stabilization in the F system (11). The similarity between ORF127 and IncF TraH, and the necessity of ORF127 for transfer, based on transposon mutagenesis (99), led us to designate it TrhH. This ORF was not called "Tra" because this designation is reserved for only the DNA processing proteins of TraI, whereas all proteins responsible for mating pair formation and pilus assembly will be named "Trh". Adjacent to *trhH* is ORF128, which was also identified as necessary for transfer by transposon mutagenesis (99). This ORF encodes a protein of 1329aa predicted to have 8 strong transmembrane helices. These characteristics are similar to that of TraG of the IncF plasmids, encoded immediately downstream of *traH* (49). We did not expect to encounter pilus assembly proteins (ORF126, TrhH and ORF128) within the DNA processing region of R27.

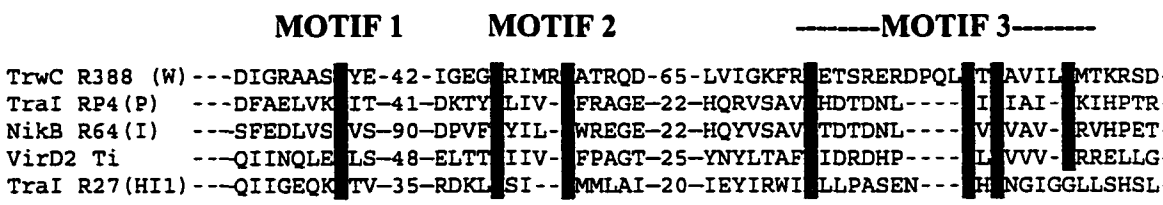


Figure V-4 Alignment of the TraI homologue of R27 (ORF120) and those of other plasmids

Alignment of the TraI (ORF120) with nickase/relaxase proteins from other conjugative plasmids (IncW R388, IncP RP4, IncI R64, Ti, and IncHI1 R27). Amino acid residues that follow the guidelines of the nickase/relaxase distinguishing motifs, as set out in the discussions, are highlighted.

The transfer region 2 (Tra2), previously characterized by Rooker *et al* (126), contains the genes responsible for mating pair formation and pilus biosynthesis. Nomenclature of the Tra2 region was established previously (126), with gene names following the F plasmid transfer system to which the R27 Tra2 region shows the greatest similarity in regards to gene organization. Several ORFs within the Tra2 region which show extremely high similarity to the H transfer determinants (Htd) identified in the *Serratia marcescens* IncHI2 plasmid R478 (110). The organization of four R27 genes (*trhB*, *htdT*, *trhV*, and *trhC*) is similar to that of the R478 *htd* homologues. Though Page *et al* were unable to identify two separate transfer regions in R478 via transposon mutagenesis, we find that the Tra2-equivalent regions of the two IncHI plasmids are related (110).

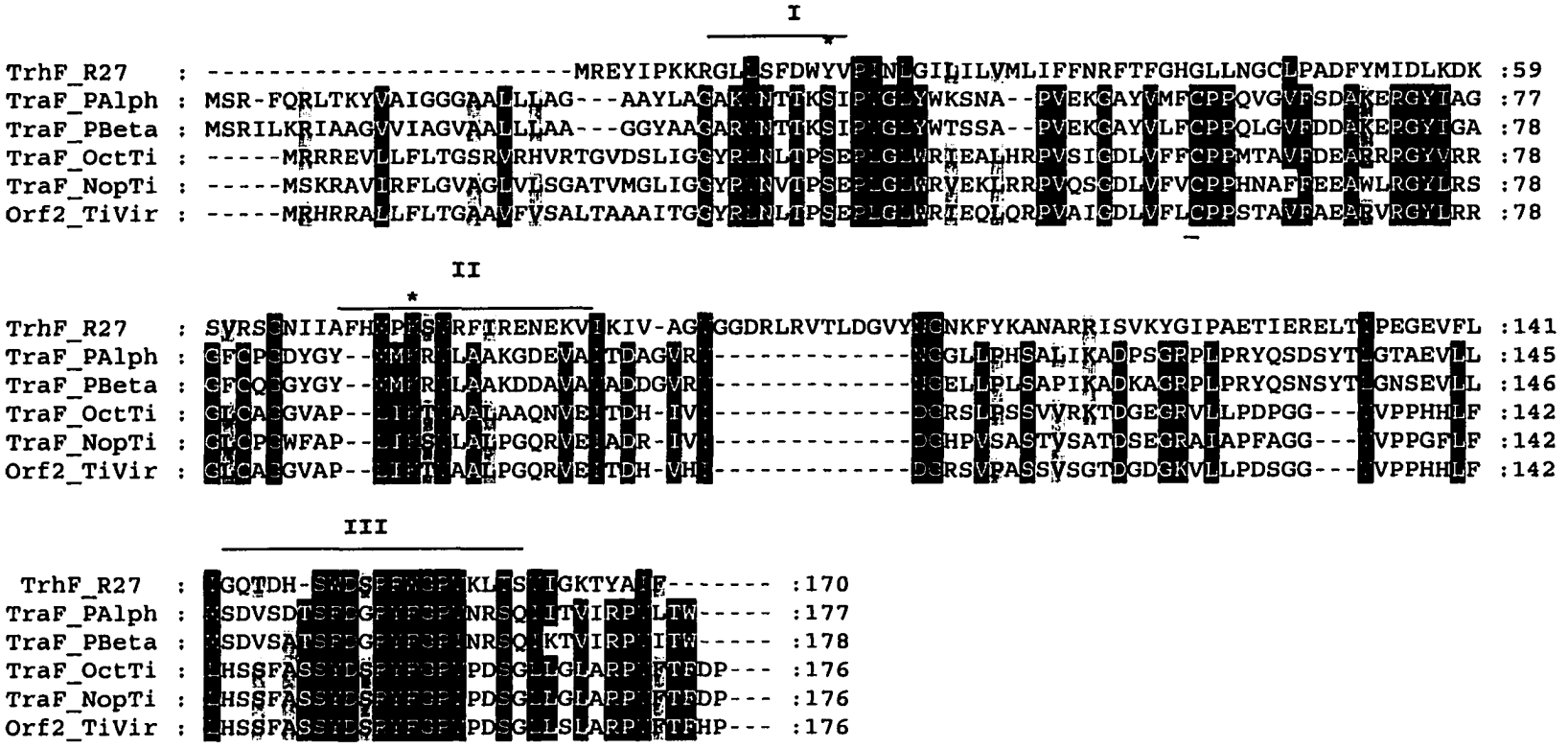
The Tra2 region is interrupted by partitioning and stability genes, leaving the *trhWUN* transfer genes separately located from the rest of the Tra2 genes that encode transfer proteins. In other prototypical transfer systems the *trhWUN* homologues are contiguous with the other transfer related genes. The 3127bp segment interrupting Tra2 is found between the *trhC* and *parA* genes. This region has also been shown to be responsible for IncH incompatibility (53). A DNA segment containing five 31 bp direct repeats (27541-27786) is sufficient for incompatibility (53). Another twenty-one 50 nucleotide repeats (25597-26746), each containing an identical 31 bp direct repeat is located downstream. When subcloned, the five 31 bp repeats are sufficient for some level of incompatibility (53). These repeats, along with the three adjacent partitioning system homologues (ORF014, 019, and 020), constitute a partitioning module that

ensures each daughter cells receives R27, which has a very low copy number, during division of the host cell. ORF014 is very similar to the StbA protein of plasmid R478 (Table V-1) and also StbA encoded on the enteropathogenic *E. coli* EAF plasmid (38% identity over 324 amino acids) (157). ORF019 and 020 are similar to the ParA and ParB partitioning proteins of the *E. coli* prophage/plasmid P1. During P1 partitioning ParB interacts with direct repeats (*parS*) forming a nucleoprotein structure (1, 34), which is functionally equivalent to a plasmid centromere (34). The role of ParA during partitioning of P1 is less clear, but it is known to regulate its own expression (48). Interruption of the Tra2 region by the partitioning region indicates that R27 could have undergone an intramolecular rearrangement during its evolution.

The IncP plasmid RP4 has been found to encode a sequence specific peptidase, TraF, that cleaves the pre-pilin subunit TrbC at the carboxy terminus (59). This activity is required for pili processing. The peptidase activity is attributed to the catalytic domains I and II, whereas domain III is believed to play a role in the conformation of the mature protein. Comparison of ORF008 and the *traF* gene product from RP4 shows conservation of catalytic domains II and III, along with an N-terminal transmembrane segment, potentially anchoring the protein to the membrane (Figure V-5). In particular, the catalytic lysine (lys-89 in TraF and lys-73 in ORF008) of domain II is conserved. These similarities have prompted us to term ORF008 as TrhF. Nevertheless in domain I the catalytic serine residue (ser-37) found in peptidases of the IncP and Ti systems is replaced by a tyrosine (tyr-17) in TrhF. The predicted transmembrane helical domain of TrhF spans residues 17 to 35 (data not shown), effectively burying the domain I-equivalent region of TrhF in the cytoplasmic face of the inner membrane. This would

Figure V-5. Alignment of TraF homologues and ORF008 (TrhF)

PILEUP comparison of the TraF homologues identifying a conserved motif between the proposed TrhF and other genes related to the TraF protein of the IncPα plasmid RP4 (11, 86). The illustrated sequences are from R27 (TrhF_R27), the IncPα plasmid RP4 (TraF_PAlph), the IncPβ plasmid R751 (TraF_PBeta), the *Agrobacterium tumefaciens* octopine-type Ti plasmid (TraF_OctTi), the *A. tumefaciens* nopaline-type Ti plasmid (TraF_NopTi), and ORF2 from the Vir2 region of the Ti plasmid (Orf2_TiVir). Catalytic residues described in the text are denoted above the alignment with an asterisk. Catalytic domains I, II, and III as described by Haase and Lanka (59), are denoted above the alignment with a line. Shading is according to agreement between aligned proteins, with variation of dark and light corresponding to level of correspondance.



itself prevent catalytic function, if any indeed remained after the serine substitution to tyrosine. If domains II and III of TrhF alone are sufficient for activity is not yet clear. Haase and Lanka found that when the catalytic ser-37 of RP4 domain I is replaced by an alanine, this cloned *traF* construct is able to partially complement a *traF* mutant (59). Like the homologues in the *Agrobacterium tumefaciens* Ti conjugative transfer region, the TraF protein in the IncP system is required for C-terminal processing of the pilin subunits leading to synthesis of the mature pili, thus the R27 homologue may also have an important role in pilin maturation.

The pilin subunit of R27 is most likely to be ORF034 (TrhA). This is encoded by the first gene in the Tra2 operon and has a deduced post-cleavage size similar to TraA, the F pilin subunit (7.6kDA for TrhA compared to 7.2kDa for TraA) (49). TrhA shows 19.4% identity over 93 of its 121 amino acids to TrbC, the RP4 pilin (50). Peptidase cleavage sites are predicted to be present in TrhA similar to those found in TrbC.

5. Temperature sensitive conjugation

A process that resembles bacterial conjugation and is permissive at low temperature is the T-DNA transfer system of the *A. tumefaciens* (26, 50), where crown gall tumor formation in the recipient plant is induced at 19°C, but not above 29°C (176). Studies on T-DNA transfer show that thermosensitivity is a property of the T-DNA transfer proteins, in which elevated temperatures result in repression of T-DNA transfer gene expression and instability of at least one essential transfer protein (59, 70). In addition, conjugal transfer of the Ti plasmid is also thermosensitive (155).

Transfer of IncHI1 plasmids, including R27, is also temperature sensitive and is inhibited at 37°C and above (154). Transfer of R27 is dependent on the temperature in which the donor cells have been grown prior to mating. (154). There are several possible underlying mechanism(s) of thermosensitivity, including transfer gene mRNA instability, transfer protein or pilus instability, the loss of requisite protein:protein or protein:DNA interactions, or differential binding of repressor protein(s) through a range of temperatures leading to transfer gene repression.

Several observations support these hypotheses. Firstly, Maher *et.al.* found that donor cultures grown exclusively at 37°C are non-piliated, but when 27°C cultures are switched to 37°C for one hour H-pili are retained (90). This implies that there is either repression of transfer gene expression or transfer gene mRNA instability at 37°C, rather than instability of H-pili structure. This is further supported by the observation that when cultures grown at 37°C are switched to 27°C, pili are not observed until three hours have elapsed (90). A three hour regeneration period might be necessary to transcribe transfer genes, translate transfer mRNA into protein, and assemble a functional transfer protein complex and pilus. Support for a second hypothesis is from the following data: when mating cultures are shifted to 37°C for 30 min, aggregation of donor and recipient cells no longer occurs (90). Also, when a mating culture is briefly exposed to high temperature (i.e. a one minute pulse at 37°C) there is a reduction in transfer frequency (99). This rapid response to temperature suggests inactivation of preformed transfer elements, possibly pili or other transfer proteins involved in cell-cell contact.

Sequence analysis of R27 has identified ORFs whose homologues are involved in thermorepressed expression in other systems. Homologues of the temperature-regulated

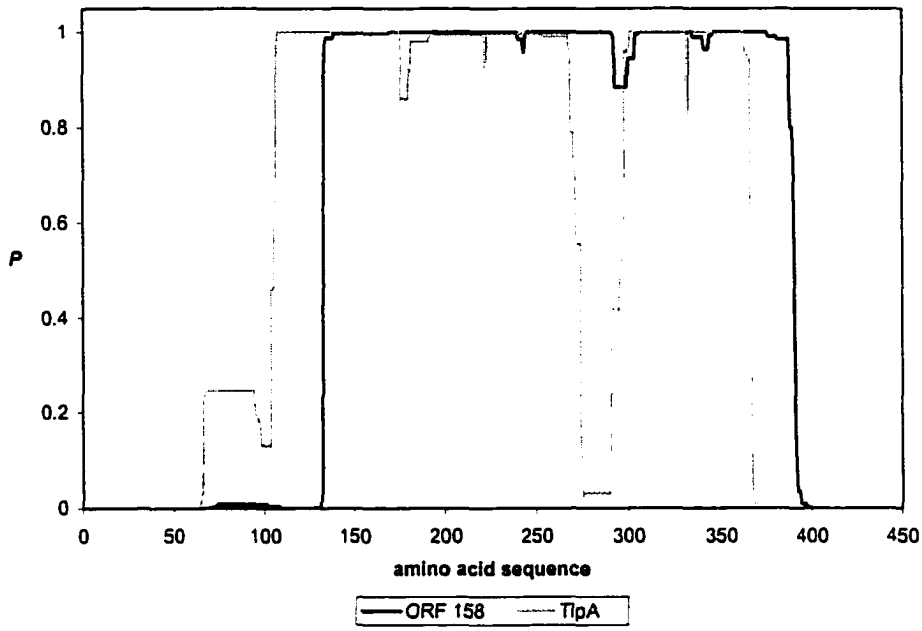
DNA-binding proteins TlpA and H-NS have been discovered in the R27 sequence. The TlpA protein encoded on the *S. typhimurium* virulence plasmid is a thermosensing repressor (69). The carboxy terminus of TlpA forms coiled coils, while deletion of the amino terminal residues 31 to 43 prevent DNA binding (68). ORF158 show 21.9% amino acid identity over 370 amino acids, and distinguishable similarity to the TlpA DNA-binding domain (Figure V-6A). Both ORF158 and TlpA show very similar profiles from coiled-coil prediction algorithms (Figure V-6B) (13, 96). The TlpA and H-NS proteins both bind DNA at low temperature and permit expression at higher temperatures, the opposite of the thermosensitivity seen in R27. Additional studies are required to determine the roles played by TlpA and H-NS homologues in R27, including their potential influence on transfer. As in other recently sequenced plasmids (1, 22), we observed a mosaic-like structure in the R27 conjugative plasmid.

6. Conclusions.

The current state of the R27 plasmid appears to have arisen from a series of insertion events, both small and large in nature. Though containing two separate transfer regions, similar to the IncP plasmids, the individual genes in both the Tra1 and Tra2 regions display considerably more homology to the genes of the single transfer region in the F plasmid family. The separation of the transfer regions in R27 (over 70kb) is greater than that observed in the IncP plasmids (111), though not at all uncommon in the Ti plasmids of *A. tumefaciens*. (35, 67). Furthermore there are a number of genes observed in the Tra1 region (ORF0126, *trhH*, ORF128, ORF130) which appear to be involved in pilus synthesis and assembly, functions normaily associated with the Tra2 region.

A

```
        TlpA : 30 LFNQVGSNPTSLRQI 45
      ORF158 : 74 LRLKM-NSFALAQEM 88
```



B

Figure V-6 Alignment of the TlpA DNA Binding Domain

(**A**) Amino acid alignment of the TlpA DNA-binding domain (69), underlined, with surrounding sequences, and the proposed equivalent in ORF158.

(**B**) The COILS2 output for ORF158 and TlpA indicating the probability (*P*) of distinct regions of each protein having a coiled-coil conformation. ORF158 and TlpA also have similar outputs using the PARCOIL program (13) (data not shown).

A simple insertion event in the far history of R27, dividing the single transfer region into the two transfer regions, does not account for the present plasmid organization. The orientation of the pilus synthesis and assembly genes observed in the Tra1 region are opposite in orientation to the Tra2 genes. Therefore if the R27 plasmid is an ancient ancestor of the F plasmid family, as individual protein homology seems to indicate, there have been not only a series of insertions in the plasmid, but internal inversions of the transfer region(s) as well.

The number of insertion elements found in R27 identifies a possible mechanism for the evolution of the R27 plasmid, particularly the presence of the Tn*10* transposon, conferring the observed Tc^R to the plasmid. The partially assembled *S. typhi* CT18 genome sequence, recently released through the world wide web (http://www.sanger.ac.uk/Projects/S_typhi/) has identified a 218kb multidrug resistance plasmid pCHM1. Comparison between R27 and the single "contig" representing pCHM1, identified only 3836bp of DNA within R27 (0.8%) that does not show near DNA homology to pCHM1. The sequence disparity is confined to six regions within R27, of which four appear to be related to transposon or insertion element activity. This clearly outlines the importance of insertion elements and transposons in the evolution of large plasmids, even for plasmids so clearly related as pCHM1 and R27 appear to be and serves as a reminder of the potent combination transposons and conjugative plasmids represent in pathogenic bacteria.

Chapter VI Entry Exclusion of R27

Entry exclusion proteins mediate the exclusion of similar plasmids from entry into cells containing a pre-existing plasmid. Screening a previously constructed cosmid library of R27 for the entry exclusion phenotype, it was possible to identify the genes as being contained on a single 22kb *Xba*I fragment. Sequence analysis and cloning of the genes contained within the fragment allowed identification of a single ORF, ORF018, as responsible for entry exclusion in R27.

1. Isolation of entry exclusion phenotype.

In the search for the gene(s) responsible for entry exclusion in R27, the transfer frequency of R27 from host *E. coli* J53-1 (R27) to an *E. coli* RG192-2 recipient containing various R27 fragments was assessed. A set of previously constructed *Sal*I (100) and *Xba*I cosmids(Table VI-1), chosen so as to represent the majority of R27, were introduced into *E. coli* RG192-2 by electroporation. Frequency of transfer of R27 into each recipient was then determined after overnight matings. Selection of transconjugants was performed using antibiotics to select for the recipient host (Rif^R) and incoming plasmid (Tc^R), so as to limit the effect of plasmid incompatibility or destabilizing of the incoming R27 plasmid by factors contained within the cosmid clone.

Those recipients containing cloned XbaI-D fragments consistently displayed decreased transfer efficiency of R27. The effect of these cosmid clones on the transfer of the IncHI2 plasmid, TP116, was tested (Table VI-2). Previous studies (153) had observed a 10-fold reduction in transfer frequency of IncHI2 plasmids into IncHI1 containing recipients. A similar reduction was observed with the three cosmid clones

Table VI-1 Transfer frequency of R27 into *E. coli* RG192-2 recipients carrying R27 cosmids.

Cosmid[a]	Fragments contained	Transfer Frequency[b]
RG192-2	None	1.9×10^{-4}
pRG1241	IncHI1[c]	2.1×10^{-6}
A15	*Xba*I-G	3.2×10^{-4}
A22	*Xba*I-D	6.8×10^{-6}
A24	*Xba*I-D+F	0.9×10^{-6}
A35	*Xba*I-E+D	1.2×10^{-5}
DT1602	*Sal*I-C+G+T	3.9×10^{-4}
DT1693	*Sal*I-B+M+O+R+Q+P	2.5×10^{-4}

[a] Cosmids are contained in *E. coli* RG192-2 host.

[b] Transfer frequency calculated as number of transconjugants per recipient after overnight conjugation at 28°C in BHI broth. Transconjugants were selected using antibiotics specific to the incoming plasmid (Tc^R) and recipient cell (Rif^R)

[c] Recipient cells contain the entire IncHI1 plasmid

Table VI-2 Transfer frequency of the IncHI2 plasmid TP116 into recipients carrying the *Xba*D fragment of R27.

Recipient[a]	Transfer Frequency[b]
RG192-2	1.8×10^{-4}
RG192-2 (R27)	3.2×10^{-5}
RG192-2 (A22)	5.9×10^{-5}
RG192-2 (A24)	3.9×10^{-5}
RG192-2 (A35)	7.2×10^{-5}

[a] Recipient strain *E. coli* RG192-2. Fragments present in the cosmid clones are listed in Table V-1.

[b] Transfer frequency calculated as number of transconjugants per recipient following overnight conjugation at 28°C in BHI broth. Transconjugants were selected for using antibiotics specific to the incoming plasmid (Cm^R) and recipient cell (Rif^R)

containing the R27 *Xba*I-D fragment. Therefore the reduction of transfer frequency of R27 into recipients was caused by a gene product or element present in the *Xba*I-D fragment.

2. Analysis of the *Xba*I-D fragment

As the whole of the *Xba*I-D fragment had been sequenced prior to the acquisition of the complete DNA sequence of R27 (126), analysis of the genes located within it was possible (Figure VI-1). Most ORFs encoded within the region contained significant homology to genes not found to be related to those involved in entry exclusion in other plasmid systems(Figure VI-1). The homologues to the IncHI2 H-transfer determinants (*htdA*, *htdF*, *htdK*) have been observed to be necessary for transfer using transposon mutagenesis (K.F. Whelan, unpublished), a feature not generally associated with entry exclusion genes.

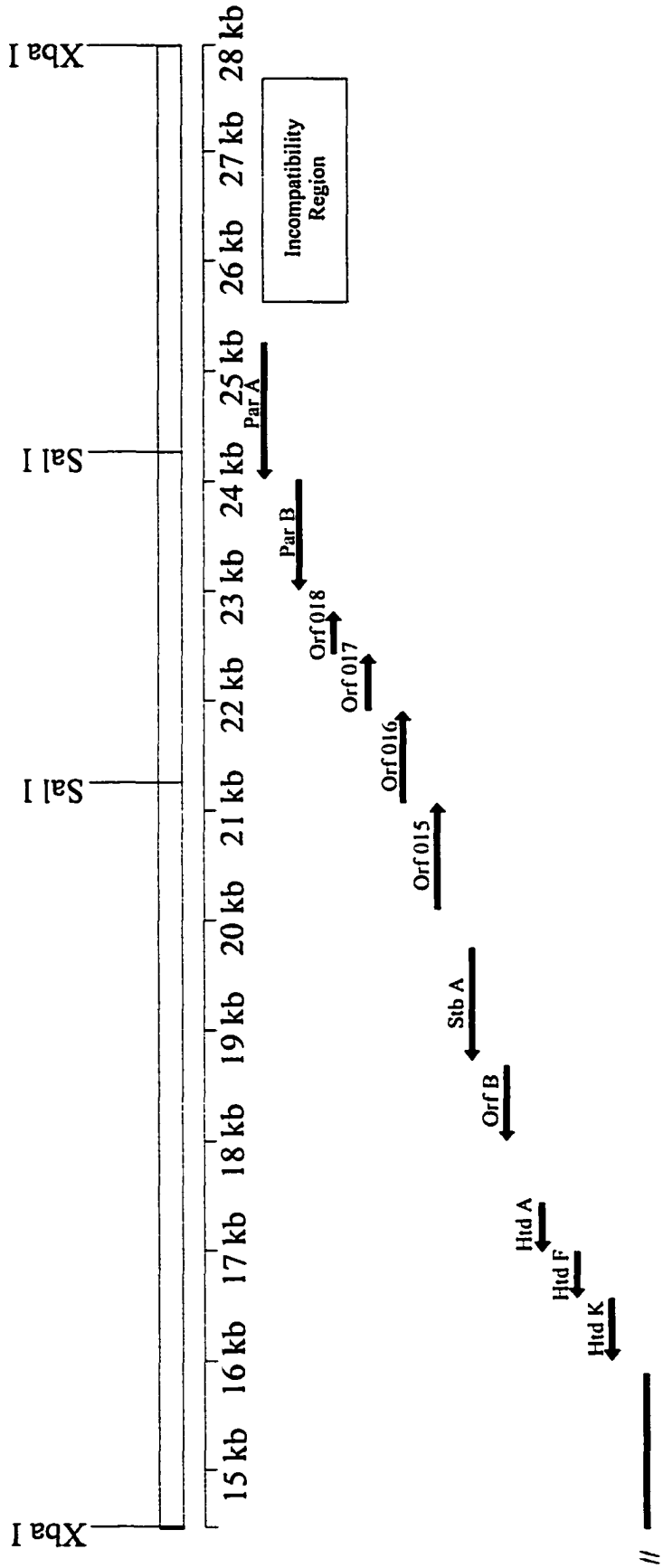
Of the five ORFs displaying no homology to known proteins, four were clustered between the *stb*A and *par*A homologues (ORF015-018) with the fifth truncated by the *Xba*I-D restriction site. The putative start codons for these four ORFs overlapped with upstream genes (0, 9, and 3 nucleotides respectively) with a strong –10 and –35 transcription initiation upstream of ORF015 (nt 19,995-20,029). Therefore the search for the entry exclusion gene was initiated with these four genes.

3. Cloning of ORFs 015-018

The four ORFs were cloned both individually and as a group both under the control of the putative promoter contained upstream of ORF015 and under the control of

Figure VI-1 Structure of the *Xba*I-D fragment of R27.

The genes contained within the *Xba*I-D fragment of R27, responsible for the entry exclusion phenotype. The direct repeats responsible for incompatibility are identified, along with the homologues to each ORF. ORF009 is truncated by the *Xba*I site. All ORFs except for ORF015-018 display significant homology to proteins unrelated to entry exclusion and the corresponding gene names are listed. ORF015-018 display no homology to any known proteins.



the pZERO-2 P_{lac} promoter. The transfer frequency of R27 from *E. coli* RG192-2 (pDT2454) into *E. coli* DH10B containing these cloned genes was determined and is presented in Table VI-3.

ORF015-017 had no detectable effect on entry of R27 into the cells, even under IPTG induced expression from the P_{lac} promoter. Transfer of R27 into recipients containing either cloned ORF018, or the four genes (ORF015-018) cloned sequentially under the control of the pZERO-2 P_{lac} promoter decreased transfer by a factor of 10^3. Even with repression of the P_{lac} promoter by addition of 50mM glucose in the nutrient rich BHI broth, transfer frequency was not reduced, though transfer frequency of R27 did return to normal with the two constructs (ORF018, 015-018 P_{lac}) in an *E. coli* DT1801 host, which constitutively produces the *lac* repressor protein $LacI^q$. The DNA fragment containing ORF015-018 cloned into pZERO-2 opposite of the vectors P_{lac} promoter, including the putative RNA transcription site, reduced transfer of R27 into the host to $6.7x10^{-5}$, similar to the transfer frequency of R27 into recipients containing the cosmid A24.

In order to clarify the role of ORF018 in entry exclusion, a *Nsi*I site within ORF018, unique to the ORF and in the vector pUC119, was exploited to generate an insertion mutant of the gene. Transfer of pDT2454 to *E. coli* DH10B recipients carrying the disrupted ORF018 gene was observed to occur at rates similar to that observed in the transfer of the plasmid into a recipient containing no IncHI1 plasmid.

Exploiting previous results with sequence specific recombination disrupting genes within R27, the Cm^R cassette disrupted ORF018 present in pZERO was introduced into

Table VI-3 Transfer of pDT2454 into recipients containing cloned genes from the *Xba*D region.

Recipient	Transfer Frequency[a]		
	50mM Glucose[b]	LB	0.05mM IPTG[c]
ORF015	3.4×10^{-3}	2.3×10^{-3}	2.1×10^{-3}
ORF016	4.7×10^{-3}	3.4×10^{-3}	4.1×10^{-3}
ORF017	1.16×10^{-3}	5.8×10^{-3}	3.1×10^{-3}
ORF018	2.0×10^{-6}	$<1 \times 10^{-8}$	$<1 \times 10^{-8}$
ORF015-018 P_{lac}	4.7×10^{-6}	$<1 \times 10^{-8}$	$<1 \times 10^{-8}$
ORF015-018 Native	5.3×10^{-5}	4.3×10^{-5}	
A24 cosmid	7.9×10^{-5}	6.7×10^{-5}	

[a] Transfer frequency calculated as number of transconjugants per recipient following an eight hour conjugation at 28°C in BHI broth. Transconjugants were selected for using antibiotics specific to the incoming plasmid (Tc^R) and recipient cell (Rif^R)

[b] Recipient cells were pre-grown and conjugations performed in BHI broth containing 50mM glucose.

[c] Recipient cells were pre-grown and conjugations performed in LB broth containing 0.05mM IPTG.

E. coli TG1 (R27) by electroporation and the resulting strain immediately mated with *E. coli* J53-1. $Tc^{R}Cm^{R}$ colonies were selected, all of which were Km^{S}, indicating that the Cm^{R} had been transferred as part of R27. Sequence specific recombination was ascertained by PCR screening of the plasmid using primers upstream (CRAIG52) and downstream (CRAIG33) of the disrupted ORF018. Transfer frequency of pDT2454 into *E. coli* J53-1 host containing R27 with ORF018 disrupted was found to be $1.4x10^{-2}$ compared to a transfer frequency of $5.7x10^{-2}$ for transfer of pDT2454 into *E. coli* TG1 with no plasmid.

4. Conclusions

The transfer frequency of R27 is markedly decreased by the presence of ORF018 and is dependent on the expression of the gene. Therefore it is clear that the reduction in transfer of R27 to recipients is due to the peptide produced by ORF018. It is possible that the reduction in transfer arises from destabilization of the nascent R27 strand arriving in the recipient due to the ORF018 RNA transcript. This is unlikely as even minor transcription of the gene, as observed with cloned ORF018 in the recipient cell, with the vectors P_{lac} promoter repressed by addition of 50mM glucose, still reduces the transfer frequency of R27. Due to the transitory nature of RNA in the bacterial cell it is unlikely that the few copies of RNA in the bacterial cell resulting from ORF018 expression would be sufficient to reduce the transfer frequency to that observed.

Therefore the entry exclusion effect in R27 appears to be due to the action of the ORF108 gene product, acting as a single element. The protein product, at a size of 14.1kDa is significantly larger than the other previously observed gene products from

single element entry exclusion systems. The ORF does not show any homology to entry exclusion genes of other systems, although like the entry exclusion proteins from the Ti, IncN, and IncP plasmids, R27 contains a predicted N-terminal trans-membrane sequence in the initial N-terminal portion of the protein (amino acids 3-19) (66). There is no established nomenclature for entry exclusion genes in other systems, with gene naming following according to gene order or order of discovery rather than function. As ORF018 falls into what appears to be an insert between partitioning genes in the Tra2 region, gene order and order of discovery is not a viable scheme for naming of the gene.

Chapter VII Concluding Remarks

Comparison of mating specific aggregation to transfer frequency in LPS mutant recipients and donors allows a better understanding of the transfer process. Truncation of bacterial LPS appears to play a role in the stabilization of the mating aggregates, as opposed to increased conjugal transfer arising from increased pilus binding. Furthermore, increased stabilization of the mating aggregate is not dependent upon exposing of recipient proteins, rather it may result from the removal of one layer of O-polysaccharide, either on the recipient or the donor.

Molecular modeling of the LPS provides an explanation as to why truncation of the LPS can lead to increased conjugal DNA transfer. The O-polysaccharide presents a very dense barrier decreasing bacterial permeability (72, 143), and can actually interfere with *E. coli* adhesion to epithelial cells (15). It is perhaps not at all surprising that the O-polysaccharide can also interfere with the formation of mating aggregates necessary for efficient transfer of R27 or interaction of membrane bound transfer proteins with recipient cells. Yet truncation of the inner core, or alteration of the charge within the LPS core through addition of PEA can decrease transfer frequency of R27, although the mechanism of action is unknown. It is likely that LPS truncation affecting changes in the outer membrane or outer membrane proteins plays a role in the conjugation process, complicating attempts to identify LPS contribution to mating efficiency.

The use of LPS mutant recipients with donors containing mutations in the R27 transfer genes may further elucidate the role of transfer proteins in conjugation. In future studies it will be interesting to determine if the TrhN and ORF128 proteins in R27,

homologues of the TraN and TraG proteins of the F system, play as important a role in the specificity and stabilization of the mating aggregates in the IncH plasmids as they do in the F plasmid system and if they are affected by LPS truncation (8).

The entry exclusion phenotype was localized to a 14kb cosmid-cloned fragment of R27 and further delineated to the activity of a single gene, ORF018. ORF018 is found in a stretch of four contiguous genes appearing to be under the control of a single transcriptional initiator upstream of ORF015, all of which display no significant homology to previously identified genes. The surface exclusion gene of R27 has no homology to previously identified entry exclusion genes from other plasmid systems, though it does contain a strong N-terminal transmembrane segment similar to single-protein entry exclusion systems found in other plasmids.

Previous amino acid analysis of purified pili appears to have been complicated by the inability to visualize the true pilin subunit in SDS PAGE separations. The possibility exists that the 3.2kDa band visible in SDS PAGE separation represents the 7.2kDa peptide observed in MALDI analysis of purified pilin, with the erroneous transfer through the gel arising from a combination of charge and altered protein cross section due to circularization of the peptide. Alternatively the SDS PAGE visualized 3.2kDa band is a minor component of purified pili, perhaps representing a tip adhesin. Yet it is clear that the 3.2kDa protein visualized in SDS PAGE separation is associated with pilin production, as phage adhesion correlates with the disappearance of this 3.2kDa band as assessed by Western blots (99). The cloned *trhA* gene does produce an appropriately sized protein as visualized using a cell-free expression system, though no band is observed with SDS PAGE separations of whole cells carrying the cloned gene. It

therefore appears that the TrhA protein forms tight protein aggregates, resistant to denaturation and staining.

Mutagenesis of the *trhA* gene met with limited success. It was possible to generate insertion-mutants of vector-cloned *trhA*, yet site-specific recombination of the disrupted gene into R27 was unsuccessful. Therefore it appears that disruption of *trhA* in R27, the lead gene in the Tra2 region, interferes with plasmid stability. It is unlikely that this is due to interruption of transcription of downstream genes as antibiotic resistance cassette insertions into *trhA*, oriented in opposite directions, were unable to undergo site-specific recombination into R27. Yet the methodology used for recombinational mutation was successful for other transfer genes and the gene responsible for entry exclusion, proving to be an effective tool in mutational analysis of R27.

It is interesting to note that lack of bacteriophage Hgal replication does not indicate lack of adhesion to the bacteria. All evidence is that Hgal is pilus specific for adhesion with previous studies utilizing this feature to good effect (90, 91). It appears that Hgal also requires an extended pilus for infection. The bacteriophage is able to adhere to cells containing R27 plasmids which have mutations in genes displaying significant homology to the pilin processing and assembly genes of the IncF system (126), however it is unable to initiate a lytic infection (99). It is likely that the extended pilus is required for RNA transfer from the bacteriophage to the host cell.

Genetic analysis of R27, allowing gene naming and proposal of gene function based upon homologies to other systems, has given some insight into the relationship of the IncH plasmids to other plasmid families. Nearly half (10 of 21) of the transfer genes identified to date show significant homology to transfer genes of the IncF plasmid. Of

the genes showing homology to the F plasmid transfer genes, most are contained in the Tra2 region. In contrast most of the genes found within the Tra1 region contain no significant homology to previously identified genes in any plasmid systems. Amino acid motifs within the genes found in the Tra1 region, in combination with their necessity for transfer, have allowed some assignment of function to the genes.

Even though there are two transfer regions within the IncH plasmid, the plasmid does not appear to be related to the two transfer system IncP or Ti plasmids. The lack of a *trbI* or a *virB10* homologue, found amongst all type IV secretion systems, suggests that the transfer system of R27 is not closely related to type IV secretion systems. The strong homologies observed among the transfer genes of the IncH and F plasmid systems, along with the remnant of a F replicon does indicate that either the IncH plasmid diverged from the F plasmid family, or exchanged significant amounts of DNA with an F like plasmid in the distant past. The separation of the two transfer regions in R27 is likely due to a series of insertion events in the plasmid, with the possibility of complete loss of the original conjugal DNA processing genes, replaced with the Tra1 region as it is found today.

References

1. **Abeles, A. L., S. A. Friedman, and S. J. Austin.** 1985. Partition of unit-copy miniplasmids to daughter cells. The DNA sequence and functional organization of the P1 partition region. Journal of Molecular Biology. **185:**261-272.

2. **Achtman, M., G. Morelli, and S. Schwuchow.** 1978. Cell-cell interactions in conjugating *Escherichia coli*: role of F pili and fate of mating aggregates. Journal of Bacteriology. **135:**505-515.

3. **Achtman, M., and R. Skurray.** 1977. A redefinition of the mating phenomenon in bacteria, p. 235-279. *In* J. Reissig (ed.), Microbial interactions. Chapman and Hall, London, England.

4. **Akiba, T., K. Koyame, Y. Ishiki, and S. Kumira.** 1960. On the mechanism of the development of multiple drug-resistance clones of *Shigella*. Japanese Journal of Microbiology. **4:**219-227.

5. **Alt-Morbe, J., J. L. Stryker, C. Fuqua, P. Li, S. K. Farrand, and S. C. Winans.** 1996. The conjugal transfer system of *Agrobacterium tumefaciens* octopine-type Ti plasmids is closely related to the tranfer system of an IncP plasmid and distanly related to Ti plasmid *vir* genes. Journal of Bacteriology. **178:**4248-4257.

6. **Altschul, S. D., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. 1990 Basic local alignment search tool. Journal of Molecular Biology. **215:**403-410.

7. **Anderson, E. S., and H. R. Smith.** 1972. Chloramphenicol resistance in the typhoid bacillus. British Medical Journal. **3:**329-332.

8. **Anthony, K. G., W. A. Klinke, J. Manchak, and L. S. Frost.** 1999. Comparison of proteins involved in pilus synthesis and mating pair stabilization from the related plasmids F and R100-1: insights into the mechanism of conjugation. Journal of Bacteriology. **17:**5149-5159.

9. **Anthony, K. G., C. Sherburne, R. Sherburne, and L. S. Frost.** 1994. The role of the pilus in recipient cell recognition during bacterial conjugation mediated by F-like plasmids. Molecular Microbiology. **13:**939-953.

10. **Bachman, B. J.** 1972. Pedigrees of some mutant strains of *Escherichia coli* K-12. Bacteriological Reviews. **36:**525-557.

11. **Balzer, D., P. Werner, and E. Lanka.** 1994. Essential motifs of relaxase (TraI) and TraG proteins involved in conjugative transfer of plasmid RP4. Journal of Bacteriology. **176:**4285-4295.

12. **Barany, F.** 1985. Single-stranded hexameric linkers: a system for in-phase insertion mutagenesis and protein engineering. Gene. **37:**111-123.

13. **Berger, B., D. B. Wilson, E. Wolf, T. Tonchev, M. Milla, and P. S. Kim.** 1995. Predicting coiled coils by use of a pairwise residue correlations. Proceedings of the National Academy of Sciences. **92:**8259-8263.

14. **Berger, B. R., and P. J. Christie.** 1994. Genetic complementation analysis of the *Agrobacterium tumefaciens virB* through *virB11* are essential virulence genes. Journal of Bacteriology. **176:**3646-3660.

15. **Bilge, S. S., J. C. Vary, S. F. Dowell, and P. I. Tarr.** 1996. Role of *Escherichia coli* O157:H7 O side chain in adherance and analysis of an *rfb* locus. Infection and Immunity. **64:**4795-4801.

16. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screeninng recombinant plasmid DNA. Nucleic Acids Research. **7**:1513-1523.

17. **Blattner, F. R., G. I. Plunkett, C. A. Bloch, N. T. Perna, V. Burland, R. M., J. Collado-Vides, J. D. Glasner, C. K. Rode, G. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, and Y. Shao.** 1997. The complete genome sequence of *Escherichia coli* K-12. Science. **277**:1453-74.

18. **Bogosian, G., K. Bilyeu, and J. P. O'Neil.** 1993. Genome rearrangements by residual IS*10* elements in strains of *Escherichia coli* K-12 which had undergone Tn*10* mutagenesis and fusaric acid selection. Gene. **133**:17-22.

19. **Bradley, D. E.** 1986. Determination of pili by conjugative bacterial drug resistance plasmids of incompatibility groups B, C, H, J, K, M, V, and X. Journal of Bacteriology. **141**:63-71.

20. **Bradley, D. E.** 1980. Morphological and serological relationships of conjugative pili. Plasmid. **4**:155-169.

21. **Bradley, D. E., V. M. Hughes, H. Richards, and N. Datta.** 1982. R plasmids, of a new incompatibility group determine constitutive production of H pili. Plasmid(7):230-238.

22. **Brooks, J. E., R. M. Blumenthal, and T. R. Gingeras.** 1983. The isolation and characterization of the *Escherichia coli* DNA adenine methylase (dam) gene. Nucleic Acids Research. **11**:837-851.

23. **Burland, V., D. Daniels, G. r. Plunkett, and F. Blattner.** 1993. Genome sequencing on both strands:the Janus strategy. Nucleic Acids Research. **21:**3385-90.

24. **Burland, V., Y. Shao, N. T. Perna, G. Plunkett, H. J. Sofia, and F. R. Blattner.** 1998. The complete DNA sequence and analysis of the large virulence plasmid of *Escherichia coli* O157:H7. Nucleic Acids Research. **26:**4196-4204.

25. **Chahdi, A. Z., P. Gabant, and M. Couturier.** 1997. Molecular analysis of RepHIB, a replicon specific to IncHI1 plasmids. Molecular and General Genetics. **255:**477-486.

26. **Christie, P. J.** 1997. *Agrobacterium tumefaciens* T-complex transport apparatus: a paradigm for a new family of multifunctional transporters in eubacteria. Journal of Bacteriology. **179:**3085-3094.

27. **Cohen, S. N., A. C. Y. Chang, and L. Hsu.** 1972. Nonchromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R-factor DNA. Proceedings of the National Academy of Sciences. **69:**2110-2114.

28. **Couturier, M., F. Bex, P. L. Bergquist, and W. K. Maas.** 1988. Identification and classification of bacterial plasmids. Microbiological Reviews. **52:**375-395.

29. **Curtiss, R.** 1969. Bacterial conjugation. Annual Reviews in Microbiology. **23:**69-136.

30. **Dalrymple, B., P. Caspers, and W. Arber.** 1984. Nucleotide sequence of the prokaryotic mobile genetic element IS30. EMBO. **3:**2145-2149.

31. **Daniels, D. L., G. I. Plunkett, V. Burland, and F. R. Blattner.** 1992. Analysis of the *Escherichia coli* genome: DNA sequence of the region from 84.5 to 86.5 minutes. Science. **257:**771-778.

32. **Datta, N.** 1975. Epidemiology and classification of plasmids, p. 9-15. *In* D. Schlessinger (ed.), Microbiology-1974. American Society for Microbiology, Washington, D.C.

33. **Davis, B. D.** 1950. Non-filterability of the agents of genetic recombination in *E. coli*. Journal of Bacteriology. **60:**507.

34. **Davis, M. A., and S. J. Austin.** 1988. Recognition of the P1 plasmid centromere analog involves binding of the ParB protein and is modified by a specific host factor. EMBO Journal. **6:**1881-1888.

35. **de Greve, H., H. Decraemer, J. Seurink, M. von Montagu, and J. Schell.** 1981. The functional organization of the octopine *Agrobacterium tumefaciens* plasmid pTiB6S3. Plasmid. **6:**235-248.

36. **Delver, E. P., V. U. Kotova, G. B. Zavilgelsky, and A. A. Beloguruv.** 1991. Nucleotide sequence of the gene (*ard*) encoding the antirestriction protein of plamid ColIb-P9. Journal of Bacteriology. **173:**5887-5892.

37. **Dower, W. J., J. F. Miller, and C. W. Ragsdale.** 1988. High efficiency transformation of *E. coli* by high-voltage electroporation. Nucleic Acids Research. **16:**6127-6145.

38. **Drummelsmith, J., P. A. Amor, and C. Whitfield.** 1997. Polymorphism, duplication, and IS1-mediated rearrangement in the chromosomal *his-rfb-gnd*

region of *Escherichia coli* strains with group IA and capsular K antigens. Journal of Bacteriology. **179**:3232-3238.

39. **Ehrenfeld, E. E., R. E. Kessler, and D. B. Clewell.** 1968. Identification of pheromone-induced surface proteins in *Streptococcus faecalis* and evidence of a role for lipoteichoic acid in formation of mating aggregates. Journal of Bacteriology. **168**:6-12.

40. **Eisenbrandt, R., M. Kalkum, E. Lai, R. Lurz, C. Kado, and E. Lanka.** 1999. Conjugative pili of IncP plasmids, and the Ti plasmid T pilus are composed of cyclic subunits. Journal of Biological Chemistry. **274**:22548-22555.

41. **Finlay, B. B., and S. Falkow.** 1997. Common themes in microbial pathogenicity revisited. Microbiology and Molecular Biology Reviews. **61**:136-169.

42. **Finlay, B. B., and W. Paranchych.** 1986. Nucleotide sequence of the surface exclusion genes *traS* and *traT* from the IncF*lac* plasmid pED208. Journal of Bacteriology. **166**:713-721.

43. **Firth, N., and R. Skurray.** 1992. Characterization of the F plasmid bifunctional conjugation genes. Molecular and General Genetics. **232**(1):145-153.

44. **Fleischmann, R. D., M. D. Adams, O. White, R. A. Clayton, E. F. Kirkness, A. R. Kerlavage, C. J. Bult, J.-F. Tomb, B. A. Dougherty, J. M. Merrick, K. Mckenney, G. Sutton, W. Fitzhugh, C. A. Fields, J. D. Gocayne, J. D. Scott, R. Shirley, L.-I. Liu, A. Glodek, J. M. Kelley, J. F. Weidman, C. A. Phillips, T. Spriggs, E. Hedblom, M. D. Cotton, T. R. Utterback, M. C. Hanna, D. T. Nguyen, D. M. Saudek, R. C. Brandon, L. D. Fine, J. L. Fritchman, J. L. Fuhrmann, N. S. M. Geoghagen, C. L. Gnehm, L. A. Mcdonald, K. V. Small,**

C. M. Fraser, H. O. Smith, and J. C. Venter. 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science. **269**:496-512.

45. **Fraser, C. M., S. Casjens, W. M. Huang, G. G. Sutton, R. A. Clayton, R. Lathigra, O. White, K. A. Ketchum, R. Dodson, E. K. Hickey, M. Gwinn, B. Dougherty, J.-F. Tomb, R. D. Fleischmann, D. Richardson, J. Peterson, A. R. Kerlavage, J. Quackenbush, S. Salzberg, M. Hanson, R. van-Vugt, N. Palmer, M. D. Adams, J. D. Gocayne, J. Weidman, T. Utterback, L. Watthey, L. McDonald, P. Artiach, C. Bowman, S. Garland, C. Fujii, M. D. Cotton, K. Horst, K. Roberts, B. Hatch, H. O. Smith, and J. C. Venter.** 1997. Genomic sequence of a Lyme disease spirochete, *Borrelia burgdorferi*. Nature. **390**:580-586.

46. **Freiberg, C., R. Fellay, A. Bairoch, W. J. Broughton, A. Rosenthal, and X. Perret.** 1997. Molecular basis of symbiosis between *Rhizobium* and legumes. Nature. **387**:394-401.

47. **Freifelder, D.** 1968. Studies with *E. coli* sex factors. Cold Spring Harbor Symposium on Quantitative Biology. **33**:425-434.

48. **Friedman, S. A., and S. J. Austin.** 1988. The P1 plasmid-partition system synthesizes two essential proteins from an autoregulated operon. Plasmid. **2**:103-112.

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

50. **Fullner, K. J., and E. W. Nester.** 1996. Temperature affects the T-DNA transfer machinery of *Agrobacterium tumefaciens*. Journal of Bacteriology. **178:**1498-1504.

51. **Furuya, N., and T. Komano.** 1994. Surface exclusion gene of $IncI_1$ plasmid R64: nucleotide sequence and analysis of deletion mutants. Plasmid. **32:**80-84.

52. **Gabant, P., A. O. Chahdi, and M. Couturier.** 1994. Nucleotide sequence and replication characteristics of RepHI1B: a replicon specific to the IncHI1 plasmids. Plasmid. **31:**111-120.

53. **Gabant, P., P. Newnham, D. Taylor, and M. Couturier.** 1993. Isolation and location on the R27 map of two replicons and incompatibility determinant specific for IncHI1 plasmids. Journal of Bacteriology. **175:**7697-7701.

54. **Geiger, G. E., and P. Modrich.** 1979. Recognition sequence of the *dam* methylase of *Eschericia coli* K12 and mode of cleavage of *DpnI* endonuclease. Journal of Biological Chemistry. **254:**1408-1413.

55. **Gerlitz, M., O. Hrabak, and H. Schwab.** 1990. Partitioning of broad-host-range plasmid RP4 is a complex system involving site-specific recombination. Journal of Bacteriology. **172:**6194-6203.

56. **Golub, E. I., and H. A. Panzer.** 1988. The F factor of *Escherichia coli* carries a locus of stable plasmid inheritance *stm*, similar to the *parB* locus of plasmid RI. Molecular and General Genetics. **214:**353-357.

57. **Grindley, N. D. F., J. N. Grindley, and E. S. Anderson.** 1972. Molecular studies of the R factor compatibility groups. Molecular and General Genetics. **119:**287-297.

58. **Haase, J., M. Kalkum, and E. Lanka.** 1996. TrbK, a small cytoplasmic membrane lipoprotein, functions in entry exclusion of the IncPα plasmid RP4. Journal of Bacteriology. **178:**6720-6729.

59. **Haase, J., and E. Lanka.** 1997. A specific protease encoded by the conjugative DNA transfer systems of IncP and Ti plasmids is essential for pilus biosynthesis. Journal of Bacteriology. **179:**5728-5735.

60. **Haase, J., R. Lurz, A. M. Grahn, D. H. Bamford, and E. Lanka.** 1995. Bacterial conjugation mediated by plasmid RP4: RSF1010 mobilization, donor-specific phage preparation, and pilus production require the same Tra2 core compnonents of a proposed DNA transport complex. Journal of Bacteriology. **177:**4779-4791.

61. **Harrington, L. C., and A. C. Rogerson.** 1990. The F pilus of *Escherichia coli* appears to support stable DNA transfer in the absence of wall-to-wall contact between cells. Journal of Bacteriology. **172:**7263-7264.

62. **Heinrichs, D. E., J. A. Yethon, and C. Whitfield.** 1998. Molecular basis for structural diversity in the core regions of the lipopolysaccharides of *Escherichia coli* and *Salmonella enterica*. Molecular Microbiology. **30:**221-232.

63. **Hendrix, L. R., L. P. Mallavia, and J. E. Samuel.** 1993. Cloning and sequencing of *Coxiella burnetii* outer membrane protein gene *com1*. Infection and Immunity. **61:**470-477.

64. **Hill, L. R.** 1966. An index to deoxyribonucleic acid base composition of bacterial species. Journal of General Microbiology. **44:**419-437.

65. **Hillen, W., and K. Schollmeier.** 1983. Nucleotide sequence of the Tn*10* encoded tetracycline resistance gene. Nucleic Acids Research. **11**:525-539.

66. **Hofmann, K., and W. Stoffel.** 1993. TMbase- A database of membrane spanning proteins segments. Biological Chemistry Hoppe-Seyler. **374**:166.

67. **Holsters, M., B. Silva, F. van Vilet, C. Genetello, M. de Block, P. Dhaese, A. Depicker, D. Inze, G. Engler, R. Villarroel, M. van MOntagu, and J. Schell.** 1980. The functional orginzation of the nopaline *A. tumefaciens* plasmid pTiC58. Plasmid. **3**:212-230.

68. **Hurme, R., K. D. Berndt, E. Normork, and M. Rhen.** 1996. DNA binding exerted by a bacterial gene regulator with an extensive coiled-coil domain. Journal of Biological Chemistry. **271**:12626-12631.

69. **Hurme, R., K. D. Berndt, E. Normork, and M. Rhen.** 1997. A proteinaceous gene regulatory thermometer in Salmonella. Cell. **90**:55-64.

70. **Jin, S., Y. N. Song, W. Y. Deng, M. P. Gordon, and E. W. Nester.** 1993. The regulatory VirA protein of *Agrobacterium tumefaciens* does not function at elevated temperatures. Journal of Bacteriology. **175**:6830-6835.

71. **Johnsrud, L.** 1979. DNA sequence of the transposable element IS*1*. Molecular and General Genetics. **169**:213-218.

72. **Katowsky, M., T. Gutberlet, and H. Bradaczek.** 1992. Molecular modeling of three-dimensional structure and conformational flexibility of bacterial lipopolysaccharide. Journal of Bacteriology. **174**:4798-4806.

73. **Kittell, B. L., and D. R. Helinski.** 1993. Plasmid Incompatibility and replication control., p. 223-242. *In* D. B. Clewell (ed.), Bacterial conjugation. Plenum Press, New York, London.

74. **Kok, R. G., D. A. D'Argenio, and L. N. Ornston.** 1997. Combining localized PCR mutagenesis and natural transformation in direct genetic analysis of a transcriptional regulator gene, *pobR*. Journal of Bacteriology. **179:**4270-4276.

75. **Kolodkin, A. L., M. A. Capage, E. I. Golub, and K. B. Low.** 1983. F sex factor of *E. coli* K-12 codes for a single-stranded DNA binding protein. Proceedings of the National Academy of Sciences, U.S.A. **80**(4422-4426).

76. **Komano, T., T. Yoshida, K. Narahara, and N. Furuya.** 2000. The transfer region of IncI1 plasmid R64: similarities between R64 tra and legionella *icm/dot* genes. Molecular Microbiology. **35:**1348-1359.

77. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. Journal of Molecular Biology. **157**(105-32).

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

79. **Lawley, T. D., V. Burland, and D. E. Taylor.** 2000. Analysis of the complete nucleotide sequence of the tetracycline-resistance transposon Tn*10*. Plasmid. **43:**235-239.

80. **Lederberg, J., L. Cavalli, and E. M. Lederberg.** 1952. Sex compatibility in *Escherichia coli*. Genetics. **37:**720-730.

81. **Lehmann, V., G. Hämmerling, M. Nurminen, I. Minner, E. Ruschmann, O. Lüderitz, T. Kuo, and B. Stocker.** 1973. A new class of heptose-deficient

mutants of *Salmonella typhimurium*. European Journal of Biochemistry. **32:**268-275.

82. **Lessl, M., D. Balzer, W. Pansegrau, and E. Lanka.** 1992. Sequence similarities between the RP4 Tra2 and the Ti VirB region strongly support the conjugation model for T-DNA transfer. Journal of Biological Chemistry. **267:**29471-20480.

83. **Li, P. L., D. M. Everhart, and S. K. Farrand.** 1998. Genetic and sequence analysis of the *trb* locus on pTiC58, a mating-pair formation system related to members of the type IV secretion family. Journal of Bacteriology. **180:**6164-6172.

84. **Li, P. L., I. Hwang, H. Miyagi, H. Ture, and S. K. Farrand.** 1999. Essential components of the Ti plasmid *trb* system, a type IV macromolecular transporter. Journal of Bacteriology. **181:**5033-5041.

85. **Lipman, D. J., and W. R. Pearson.** 1985. Rapid and sensitive protein similarity searches. Science. **227:**1435-41.

86. **Llosa, M., S. Bolland, and F. de la Cruz.** 1994. Genetic organization of the conjugal DNA processing region of the IncW plasmid R388. Journal of Molecular Biology. **235:**448-464.

87. **Loessner, M. J., G. Wendlinger, and S. Scherer.** 1995. Heterogeneous endolysins in *Listeria monocytogenes* bacteriophages: a new class of enzymes and evidence for conserved holin genes within the siphoviral lysis cassettes. Molecular Microbiology. **16:**1231-1241.

88. **Lupas, A., M. VanDyke, and J. Stock.** 1991. Predicated coiled coils from protein sequences. Science. **252:**1162-1164.

89. **Machida, Y., C. Machida, and E. Ohtsubo.** 1984. Insertion element IS*1* encodes two structural genes required for its transposition. Journal of Molecular Biology. **177:**229-245.

90. **Maher, D., R. Sherburne, and D. E. Taylor.** 1993. H-pilus assembly kinetics determined by electron microscopy. Journal of Bacteriology. **175:**2175-2183.

91. **Maher, R., R. Sherburne, and D. E. Talylor.** 1991. Bacteriophages for incompatibility group H plasmids: morphological and growth characteristics. Plasmid:141-146.

92. **Mahillon, J., H. Kirkpatrick, H. Kijenski, C. Bloch, C. Rode, G. Mayhew, D. Rose, G. r. Plunkett, V. Burland, and F. Blattner.** 1998. Subdivision of the *Escherichia coli* K-12 genome for sequencing: manipulation and DNA sequence of transposable elements introducing unique restriction sites. Gene. **223:**47-54.

93. **Majdalani, N., and K. Ippen-Ihler.** 1996. Membrane insertion of the F-pilin subunit is Sec independent but requires leader peptidase B and the proton motive force. Journal of Bacteriology. **178:**3742-3747.

94. **Marmur, J., R. Rownd, S. Falkow, L. S. Baron, C. Schildkraut, and P. Doty.** 1961. The nature of intergeneic episomal infection. Proceedings of the National Academy of Sciences, U.S.A. **47:**972-979.

95. **Moore, D., D. C. Hamilton, K. Maneewannakul, Y. Mintz, L. S. Frost, and K. Ippen-Ihler.** 1993. The *Escherichia coli* K-12 F plasmid gene *traX* is required for acetylation of F-pilin. Journal of Bacteriology. **175:**1375-1383.

96. **Mourino, M., C. Madrid, C. Balsalobre, A. Prenafeta, F. Munoa, J. Blanco, M. Blanco, J. E. Blanco, and A. Juarez.** 1996. The Hha protein is a modulator

of expression of virulence factors in *Escherichia coli.* Infection and Immunity. **64:**2881-2884.

97. **Mukhopadhyay, S., and D. K. Chattoraj.** 2000. Replication-induced transcription of an autorepressed gene: The replication initiator gene of plasmid P1. Proceedings of the National Academy of Sciences, U.S.A. **97:**7142-7147.

98. **Muse, W. B., and R. A. Bender.** 1998. The nac(nitorgen assimilation control) gene from *Escherichia coli.* Journal of Bacteriology. **180:**1166-1173.

99. **Newnham, P. J.** 1995. PhD. Thesis. University of Alberta, Edmonton, Alberta.

100. **Newnham, P. J., and D. E. Taylor.** 1990. Genetic analysis of transfer and incompatibility functions within the IncHI plasmid R27. Plasmid. **23**(2)**:**107-118.

101. **Newnham, P. J., and D. E. Taylor.** 1994. Molecular analysis of RepHI1A, a minimal replicon of the IncHI1 plasmid. Molecular Microbiology. **11:**757-768.

102. **Nielsen, H., J. Engelbrecht, B. Brunak, and G. von Heijne.** 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Engineering. **10:**1-6.

103. **Nordstrom, K., S. Molin, and J. Light.** 1984. Control of replication of bacterial plasmids: genetics, molecular biology and physiology of the plasmid R1 system. Plasmid. **12:**71-90.

104. **Novick, R.** 1987. Plasmid Incompatibility. Microbiological Reviews. **51:**381-395.

105. **Novick, R. P.** 1969. Extrachromosomal inheritance in bacteria. Journal of Bacteriology. **33:**210-235.

106. **Novotny, C., J. Carnahan, and C. C. J. Brinton.** 1969. Mechanical removal of F pili, type I pili, and flagella from Hfr and TF donor cells and the kinetics of their reappearance. Journal of Bacteriology. **98:**1294-1306.

107. **Ohtsubo, H., and E. Ohtsubo.** 1978. Nucleotide sequence of an insertion element IS*1*. Proceedings of the National Academy of Sciences, U.S.A. **75:**615-619.

108. **Orskov, F., and I. Orskov.** 1992. *Escherichia coli* serotyping and disease in man and animals. Canadian Journal of Microbiology. **38:**699-704.

109. **Ou, J. T., and T. F. Anderson.** 1970. Role of pili in bacterial conjugation. Journal of Bacteriology. **102:**648-654.

110. **Page, D. T., K. F. Whelan, and E. Colleran.** 1999. Mapping studies and genetic analysis of transfer genes of the multi-resistant IncHI2 plasmid R478. FEMS Micriobiology Letters. **179:**21-29.

111. **Pansegrau, W., E. Lanka, P. T. Barth, D. H. Figurski, D. G. Guiney, D. Haas, D. R. Helinski, H. Schwab, V. A. Stanisich, and C. M. Thomas.** 1994. Complete nucleotide sequence of Birmingham IncPα plasmids. Journal of Molecular Biology. **239:**623-663.

112. **Paranchych, W., and L. S. Frost.** 1988. The physiology and advanced biochemistry of pili. Advances in Microbial Physiology. **29:**53-114.

113. **Parker, C. T., A. W. Kloser, C. A. Schnaitman, M. A. Stein, S. Gottesman, and B. W. Gibson.** 1992. Role of the *rfaG* and *rfaP* genes in determining the lipopolysaccharide core structure and cell surface properties of *Escherichia coli*. Journal of Bacteriology. **174:**2525-2538.

114. **Pearson, W. R., and D. J. Lipman.** 1988. Improved tools for biological sequence comparison. Proceedings of the National Academy of Sciences, U.S.A. **85:**2444-2448.

115. **Perna, N. T., G. F. Mayhew, G. Posfai, S. Elliott, M. S. Donnenberg, J. B. Kaper, and F. R. Blattner.** 1998. Molecular evolution of a pathogenicity island from enterohemorrhagic *Escherichia coli*. Infection and Immunity. **66:**3810-3817.

116. **Perry, K. L., S. J. Elledge, B. B. Mitchell, L. Marsh, and G. C. Walker.** 1985. *umuCD* and *mucAB* operons whose products are required for UV light- and chemical-induced mutagenesis: UmuD, MucA, and LexA proteins share homology. Proceedings of the National Academy of Sciences, U.S.A. **82:**4331-4335.

117. **Plesiat, P., and H. Nikaido.** 1992. Outer membranes of gram-negative bacteria are permeable to steroid probes. Molecular Microbiology. **6:**1323-1333.

118. **Pohlman, R. F., H. D. Geneti, and S. C. Winans.** 1994. Entry exclusion of the IncN plasmid-pKM101 is mediated by a single hydrophillic protein containing a lipid attachment motif. Plasmid. **31:**158-165.

119. **Pon, C. L., R. A. Calogero, and C. O. Gualerzi.** 1988. Identification, cloning, nucleotide sequence and chromosomal map location of *hns*, the structural gene for *Escherichia coli* DNA-binding protein H-NS. Molecular and General Genetics. **212**(2):199-202.

120. **Raetz, C. R. H.** 1996. Bacterial lipopolysaccharides: A remarkable family of bioactive macroamphiphiles., p. 1035-1063. *In* F. C. Neidhardt, R. Curtiss, J. L. Ingraham, E. C. C. Lin, and K. B. Low (ed.), *Escherichia coli*, and *Salmonella*.

Cellular and Molecular BIology. American Society for Microbiology Press, Washington, D.C.

121. **Ratez, C. R. H.** 1990. Biochemistry of endotoxins. Annual Reviews in Biochemistry. **59:**129-170.

122. **Reeves, P. R., M. Hobbs, M. A. Valvano, M. Skurnik, C. Whitfield, D. Coplin, N. Kido, J. Klena, D. Maskell, C. R. H. Raetz, and P. D. Rick.** 1996. Bacterial polysaccharide synthesis and gene nomenclature. Trends in Microbiology. **4:**495-503.

123. **Roantree, R., T. Kuo, and D. MacPhee.** 1977. The effect of defined lipopolysaccharide core defects upon antibiotic resistances of *Salmonella typhimurium*. Journal of General Microbiology. **103:**223-234.

124. **Romine, M. F., L. C. Stillwell, K.-K. Wong, S. J. Thurston, E. C. Sisk, C. W. Sensen, T. Gaasterland, J. D. Saffer, and J. K. Fredrickson.** 1999. Complete sequence of a 184-Kilobase catabolic plasmid from *Sphingomonas aromaticivorans* F199. Journal of Bacteriology. **181:**1585-1602.

125. **Ronecker, H. J., and B. Rak.** 1987. Genetic organization of insertion element IS*2* based on a revised nucleotide sequence. Gene. **59:**291-296.

126. **Rooker, M., C. Sherburne, D. Lawley, and D. E. Taylor.** 1999. Characterization of the Tra2 region of the IncHI1 plasmid R27. Plasmid. **41:**226-39.

127. **Salmi, D., V. Magrini, P. L. Hartzell, and P. Youderian.** 1998. Genetic determinants of immunity and integration of temperate *Myxococcus xanthus* phage Mx8. Journal of Bacteriology. **180:**614-621.

128. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning, a laboratory manual. 2nd ed. Cold Spring Harbor Laboratory, New York.

129. **Sanderson, K. E.** 1993. University of Calgary, Personal Communication.

130. **Sanderson, K. E., J. Janzer, and J. Head.** 1981. Influence of lipopolysaccharide and protein in the cell envelope on recipient capacity in conjugation of *Salmonella typhimurium*. Journal of Bacteriology. **148:**283-893.

131. **Sasatsu, M., T. K. Misra, L. Chu, R. Laddaga, and S. Silver.** 1985. Cloning and DNA sequence of a plasmid-determined citrate utilization system in *Escherichia coli*. Journal of Bacteriology. **164:**983-993.

132. **Saul, D., D. Lane, and P. L. Bergquist.** 1988. A replication region of the IncHI plasmid is highly homologous with the RepFIA of F. Molecular Microbiology. **2:**219-225.

133. **Schnaitman, C. A., and J. D. Klena.** 1993. Genetics of lipopolysaccharide in enteric bacteria. Microbiological Reviews. **57:**655-682.

134. **Schollmeier, K., and W. Hillen.** 1984. Transposon Tn*10* contains two structural genes with opposite polarity between *tetA* and IS*10*R. Journal of Bacteriology. **160:**499-503.

135. **Scholz, P., V. Haring, B. Whittmann-Liebold, K. Ashman, M. Bagdasarian, and E. Scherzinger.** 1989. Complete nucleotide sequence and gene organization of the broad-host-range plasmid RSF1010. Gene. **75:**271-288.

136. **Segal, G., J. J. Russo, and H. A. Shuman.** 1999. Relationships between a new type IV secretion system and the *icm/dot* virulence system of *Legionella pneumophila*. Molecular Microbiology. **34:**799-809.

137. **Segal, G., and H. A. Shuman.** 1999. Possible origin of the *Legionella pneumophila* viulence genes and their relation to *Coxiella burnetti*. Molecular Microbiology. **33:**669-670.

138. **Shägger, H., and G. vonJagow.** 1987. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Analytical Biochemistry. **166:**368-379.

139. **Shevchik, V. E., G. Condemine, and J. Robert-Baudouy.** 1994. Characterization of DsbC, a periplasmic protein of *Erwinia chrysanthemi* and *Escherichia coli* with disulfide isomerase activity. EMBO Journal. **13:**2007-2012.

140. **Siresena, D. M., P. R. MacLachlan, S. Liu, A. Hessel, and K. E. Sanderson.** 1994. Molecular analysis of the *rfaD* gene, for heptose synthesis, and the *rfaF* gene, for heptose transfer, in lipopolysaccharide synthesis in *Salmonella typhimurium*. Journal of Bacteriology. **176:**2379-2385.

141. **Smith, H. R., N. D. F. Grindley, G. O. Humphreys, and E. S. Anderson.** 1973. Interactions of group H resistance factors with the F factor. Journal of Bacteriology. **115:**623-628.

142. **Smith, H. W.** 1974. Thermosensitive transfer factors in chloramphenicol-resistant strains of *Salmonella typhi*. Lancet. **11:**281-282.

143. **Snyder, S., D. Kim, and T. J. McIntosh.** 1999. Lipopolysaccharides bilayer structure effect of chemotype, core mutations, divalent cations, and temperature. Biochemistry. **38:**10758-10767.

144. **Som, S., A. S. Bhagwat, and S. Friedman.** 1987. Nucleotide sequence and expression of the gene encoding the EcoRII modification enzyme. Nucleic Acids Research. **15:**313-332.

145. **Stocker, B., B. Males, and W. Takano.** 1980. *Salmonella typhimurium* mutants of RfaH- phenotype: genetics and antibiotic sensitivities. Journal of General Microbiology. **116:**17-24.

146. **Stocker, B., M. Nurminen, and P. Mäkelä.** 1979. Mutants defective in the 33K outer membrane protein of *Salmonella typhimurium.* Journal of Bacteriology. **139:**376-383.

147. **Taylor, D. E.** 1989. General properties of resistance plasmids., p. 325-357. *In* L. E. Bryan (ed.), Handbook of experimental pharmacology, vol. 91. Springer-Verlag, Berlin, Heidelberg.

148. **Taylor, D. E.** 1983. Transfer defective and tetracycline sensitive mutants of the incompatibility group HI plasmid R27 generated by insertion of transposon 7. Plasmid. **9:**227-239.

149. **Taylor, D. E., and E. C. Brose.** 1986. Location of plasmid-mediated citrate utilization determinant in R27 and incidence in other H incompatibility group plasmids. Applied and Environmental Microbiology. **52:**1394-1397.

150. **Taylor, D. E., and E. C. Brose.** 1985. Restriction endonuclease mapping of R27(TP116), an incompatibility group HI subgroup 1 plasmid from *Salmonella typhimurium*. Plasmid. **13:**75-77.

151. **Taylor, D. E., E. C. Brose, S. Kwan, and W. Yan.** 1985. Mapping of transfer regions within incompatibility group HI plasmid R27. Journal of Bacteriology. **162:**1221-1226.

152. **Taylor, D. E., J. C. Chumpitaz, and F. Goldstein.** 1985. Variability of IncHI1 plasmids from *Salmonella typhi* with special reference to Peruvian plasmids encoding resistance to trimethoprim and other antibiotics. Antimicrobial Agents and Chemotherapy. **28:**452-455.

153. **Taylor, D. E., and R. B. Grant.** 1977. Incompatibility and surface exclusion propoerties of H1 and H2 plasmids. Journal of Bacteriology. **131:**174-178.

154. **Taylor, D. E., and J. G. Levine.** 1980. Studies of temperature-sensitive transfer and maintenance of H incompatibility group plasmids. Journal of General Microbiology. **116:**475-484.

155. **Tempe, J., A. Petit, M. Holsters, M. vanMontagu, and J. Schell.** 1977. Thermosensitive step associated with transfer of the Ti-plasmid conjugation: possible relation to transformation in crown gall. Proceedings of the National Academy of Sciences, U.S.A. **74:**2848-2859.

156. **Thorsted, P. B., D. P. Macartney, P. Akhtar, A. S. Haines, N. Ali, P. Davidson, T. Stafford, M. J. Pocklington, W. Pansegrau, B. M. Wilkins, E. Lanka, and C. M. Thomas.** 1998. Complete sequence of the IncPβ plasmid R751:Implications for the evolution and orginisation of the IncP backbone. Journal of Molecular Biology. **282:**969-90.

157. **Tobe, T., T. Hayashi, C. G. Han, G. K. Schoolnik, E. Ohtsubo, and C. Sasakawa.** 1999. Complete DNA sequence and structural analysis of the

enteropathogenic Escherichia coli adherence factor plasmid. Infection and Immunity. **67**:5455-5462.

158. **Tomizawa, J.** 1990. Control of ColE1 plasmid replication. Interaction of Rom protein with an unstable complex formed by RNA I and RNA II. Journal of Molecular Biology. **212**:695-708.

159. **Tomizawa, J., T. Itoh, G. Selzer, and T. Som.** 1981. Inhibition of ColE1 RNA primer formation by a plasmid-specified small RNA. Proceedings of the National Academy of Sciences, U.S.A. **78**:1421-1425.

160. **Uga, H., F. Matsunaga, and C. Wada.** 1999. Regulation of DNA replication by iteron: an interaction between the *ori2* and *incC* regions mediated by RepE-bound iterons inhibits DNA replication of mini-F plasmid *Escherichia coli*. EMBO Journal. **18**:3856-3867.

161. **Vieira, J., and J. Messing.** 1991. New pUC-derived cloning vectors with different selectable markers and DNA replication origins. Gene. **100**:189-194.

162. **Viera, J., and J. Messing.** 1987. Production of single-stransded plasmid DNA. Methods of Enzymology. **153**:3-11.

163. **Vlcek, C., V. Paces, N. Maltsev, J. Paces, R. Haselkorn, and M. Fonstein.** 1997. Sequence of a 189-kb segment of the chromosome of *Rhodobacter capsulatus* SB1003. Proceedings of the National Academy of Sciences, U.S.A. **94**:9384-9388.

164. **Vogel, J. P., H. L. Andrews, S. K. Wong, and R. R. Isberg.** 1998. Conjugative transfer by the virulence system of *Legionella pneumophila*. Science. **279**:873-875.

165. **Vorm, O., P. Roepstorff, and M. Mann.** 1994. Improved resolution and very high sensitivity in MALDI TOF of matrix surfaces made by fast evaporation. Analytical Chemistry. **66:**3281-3287.

166. **Wang, L., and P. R. Reeves.** 1994. Involvement of the galactosyl-1-phosphate transferase encoded by the *Salmonella enterica rfbP* gene in O-antigen sub-unit processing. Journal of Bacteriology. **176:**4348-4356.

167. **Wang, Y., and D. E. Taylor.** 1990. Chloramphenicol resistance in *Campylobacter coli*: nucleotide sequence, expression, and cloning vector construction. Gene. **94:**21-28.

168. **Waters, V. L.** 1999. Conjugative transfer in the dissemination of Beta-Lactam and aminoglycoside resistance. Frontiers in Bioscience. **4:**416-439.

169. **Whelan, K. F., D. Maher, E. Colleran, and D. E. Taylor.** 1994. Genetic and nucleotide sequence analysis of the gene *htdA*, which regulates conjugal transfer of IncHI plasmids. Journal of Bacteriology. **176:**2242-2251.

170. **Whiteley, M., and D. E. Taylor.** 1983. Identification of DNA homologies among H incompatibility group plasmids by restriction enzyme digestion and Southern transfer hybridization. Antimicrobial Agents and Chemotherapy. **24:**194-200.

171. **Wilkins, B., and E. Lanka.** 1993. DNA processing and replication during plasmid transfer betwenn gram-negative bacteria, p. 105-129. *In* D. B. Clewell (ed.), Bacterial Conjugation. Plenum Press, New York.

172. **Wilkinson, R., P. Gemski, and B. Stocker.** 1972. Non-smooth mutants of *Salmonella typhimurium*: differentiation by phage sensitivity and genetic mapping. Journal of General Microbiology. **70:**527-554.

173. **Willets, N.** 1993. Bacterial Conjugation: A historical perspective, p. 1-22. *In* D. B. Clewell (ed.), Bacterial Conjugation. Plenum Press, New York, New York.

174. **Winans, S. C.** 1992. Two-way chemical signaling in *Agrobacterium*-plant interactions. Microbiological Reviews. **56:**12-31.

175. **Yamada, Y., M. Yamada, and A. Nakazawa.** 1995. A ColE1-encoded gene directs entry exclusion of the plasmid. Journal of Bacteriology. **177:**6064-6068.

176. **Zupan, J. R., and P. Zambryski.** 1995. Transfer of T-DNA from *Agrobacterium* to the plant cell. Plant Physiology. **107:**1041-1047.

Appendix:Code for sequential scanning of input amino acids for specific amino acid content

```
1     #include <stdio.h>
2     #include <iostream.h>
3     #include <fstream.h>
4     #include <string.h>
5     #include <conio.h>
6     #include <math.h>
7     #include <io.h>
8     #include <alloc.h>
9     #include <fcntl.h>
10    #include <process.h>
11    #include <sys\stat.h>
12
13    const size=20000;
14
15    /* variables */
16    char firstread[size]; /* initial array input read into memory */
17    char buffer(size];
18    char test_seq ;/* invoking the variable so as to be able to point to
it. */
19    char* test_seqpoi ; /* pointer to actual sequence in the memory,
20                 will be the basis for all other maniuplations */
21    const int frame = 25 ; /* size of the reading frame to be assayed
22    char file-name(20] ; /* name of the sequence file */
23    int pos ; /* position marker for incrementing through the test-
seg string
24    int handle,bytes,counter;
25    char temp_string[frame];
26
27    int acid-want[20];     /*   acids input wanted
28    int acid-count[20];     /* acids counted in a given frame
29    int nearness ; /* integer representing the closeness of a given temp
string's
30                  acid count to the acid want */
31    int close_pos [600]; /* an array of pointers to the start of a frame
of a
32              pre-determined closeness */
33    int file_open=0; /* 0 if not opened, 1 if succesful open */
34
35
36    void read frame(char*,int,int);
37    int acid_assay(char[frame], int[20]);
38    float acid  evaluate (int[20],int[20]);
39    void get acids();
40    void initialization();
41    void read_file();
42    void buffer_enter();
43
44
45
46
47
```

```
48
49
50    void main()
51       {printf("Please enter file name of output");
52        scanf("%s",file_name);
53        FILE *stream;
54        handle=open(file name,O_CREAT | O_TEXT,S_IREAD | S_IWRITE);
55        stream=fdopen(handle, "w");
56        read_file()
57        buffer_enter();
58        fprintf(stream,"test_seqpoi=%s \n",test_seqpoi);
59
60
61
62
63       /* now the sequence has been read, and is residing at the address
test_seqpoi
64          waiting to be accessed */
66           get_acids() ;      /* input wanted acids */
67           clrscr();
68
69    /* loop for frame reading */
70              /* for (pos=1; pos<60; pos++) */
71       for (pos=0; pos<((strlen(test_seqpoi)-frame)); pos++)
72                  {
73                  initialization();
74                  read_frame(test_seqpoi,pos,frame);
75                  /*printf ("frame %s \n", temp_string); */
76                  acid_assay(temp_string,acid_count);
77                  nearness=acid_evaluate(acid_count,acid_want);
78                  if(pos%1000 ==0)
79                     {printf("position %d reached. \n",pos);
80                     }
81              int nn;
82              if (nearness<14)
83                     {  close_pos(nn] = test_seq+pos;
84                        fprintf(stream,"frame is %s ", temp_string);
85                        fprintf(stream,"position=%d ",pos);
86                        fprintf(stream,"nearness=%d \n",nearness);
87                        printf("frame is %s ", temp_string);
88                        printf("position=-%d ",pos);
89                        printf("nearness=%d \n",nearness);
90                        nn++;
91                     };
92                  }
93           ;
94    }
95
96
97   void read file{}
98     { printf("please enter the file name in the current working
directory:");
99         scanf("%20s", file_name);
100
101 if ((handle =open(file_name, O_RDONLY | O_BINARY, S_IWRITE | S_IREAD)) ==-1)
102       {
103          printf("Error Opening File\n");
```

```
104          file_open=0;
105          exit(1);
106       }
107
108        if ((bytes = read(handle, firstread, size)) ==-1)
109          {
110          printf("Read Failed.\n");
111          file_open=0;
112          exit(1);
113          }
114       else  {
115          printf("Read: %d bytes read.\n", bytes);
116          file_open=1;
117             }
118     }
119
120  void buffer_enter()
121    {
122       counter =0;
123       int n ;
124      for(n=0;n<strlen(firstread);n++)
125      {if (firstread[n] !=' ' && firstread[n] !='\r'&& firstread[n] !='\n')
126            {
127             buffer[counter]=firstread[n];
128             counter++;
129             }
130           };
131       test_seqpoi=new char [(strlen(buffer)) +1] ;
132
133        test_seqpoi=buffer;
134   }
135
136
137
138   int acid_assay(char* temp_string, int acid_count[20])
139      {    char acid;
140          int count;
141          for (count=0; count <= frame; count++)
142             { acid=*(temp_string+count);
143             switch (acid)
144                 { case 'A':
145                    acid_count[0] = acid_count[0]+1;
146                    break;
147                   case 'R':
148                    acid_count[1] = acid_count[1]+1;
149                    break;
150                   case 'N':
151                  acid_count[2] = acid_count[2]+1;
152                    break;
153                   case 'D':
154                    acid_count[2] = acid_count[2]+1;
155                    break;
156                   case 'C':
157                    acid_count[4] = acid_count[4]+1;
158                    break;
159                   case 'E':
160                    acid_count(5] = acid_count[5]+1;
```

```
161                     break;
162                    case 'Q':
163                 acid_count[6] = acid_count[5]+1;
164                     break;
165                    case 'G':
166                 acid _count[7] = acid_count[7]+1;
167                     break;
168                    case 'H':
169                 acid_count[8] = acid_count[8]+1;
170                     break;
171                    case 'I':
172                 acid_count[9] = acid_count[9]+1;
173                     break;
174                    case 'L':
175              acid_count[10] = acid_count[10]+1;
176                     break;
177                    case 'K':
178                 acid_count[11] = acid_count[11]+1;
179                     break;
180                    case 'M':
181              acid_count[12] = acid_count[12]+1;
182                     break;
183                    case 'F':
184                 acid_count[13] = acid_count[13]+1;
185                     break;
186                    case 'P':
187                     acid_count[14] = acid_count[14]+1;
188                     break;
189                    case 'S':
190                 acid_count[15] = acid_count[15]+1;
191                     break;
192                    case 'T':
193                 acid_count[16] = acid_count[16]+1;
194                     break;
195                    case 'W':
196                  acid_count[17] = acid_count[17]+1;
197                     break;
198                    case 'Y':
199                  acid_count[18] = acid_count[18]+1;
200                     break;
201                     case 'V':
202                  acid_count[19] = acid-count[19]+1;
203                     break;
204               };
205          };
206           return *acid_count;
207     }
208
209   void initialization()
210             int i;
211       for (i=0; i<=20; i++)
212       { acid_count[i] = 0;
213       };
214        }
215
216   void read_frame (char* test_seqpoi, int entering_pos, int frame)
217       {  int n = 0 ;
```

```
218         int posn=entering_pos;
219          for (posn; posn<(entering_pos+frame); posn++)
220              {   temp_string[n] = *(test_seqpoi + posn);
221                 n++ ;};
222       }
223
224   void get_acids()
225        {
226  cout <<  "Please ensure you enter a value in each catagory, even if 0. \n";
227                    cout  << "  \nA? "  ; scanf("%d",&acid_want[0]);
228                    cout  << "  \nR? "  ; scanf("%d",&acid_want[1]);
229                    cout  << "  \nN+D? "; scanf("%d",&acid_want[2]);
230                    cout  << "  \nC? "  ; scanf("%d",&acid_want[4]);
231                    cout  << "  \nE+Q? "; scanf("%d",&acid_want[5]);
232                    cout  << "  \nG? "  ; scanf("%d",&acid_want[7]);
233                    cout  << "  \nH? "  ; scanf("%d",&acid_want[8]);
234                    cout  << "  \nI? "  ; scanf("%d",&acid_want[9]);
235                    cout  << "  \nL? "  ; scanf("%d",&acid_want[10]);
236                    cout  << "  \nK? "  ; scanf("%d",&acid_want[11;
237                    cout  << "  \nM? "  ; scanf("%d",&acid_want[12]);
238                    cout  << "  \nF? "  ; scanf("%d",&acid_want[13]);
239                    cout  << "  \nP? "  ; scanf("%d",&acid_want[14]);
240                    cout  << "  \nS? "  ; scanf("%d",&acid_want[15]);
241                    cout  << "  \nT? "  ; scanf("%d",&acid_want[16]);
242                    cout  << "  \nW? "  ; scanf("%d",&acid_want[17]);
243                    cout  << "  \nY? "  ; scanf("%d",&acid_want[18]);
244                    cout  << "  \nV?"   ; scanf("%d",&acid_want[191);
245         }
246
247
248    float acid_evaluate(int* acid_count,int* acid_want)
249       {
250          int temp,counter,internal_nearness,tester;
251          internal_nearness=0;
252          temp = 0 ;
253          for(counter=0; counter <20; counter++)
254             {  /* printf ("%d-",acid_count[counter])
255                printf(I"%d=",acid_want[counter]);
256                tester = abs(acid_count[counter]-
acid_want[counter]);
257                printf ("%d", tester);*/
258                temp = abs(acid_count[counter]-acid_want[counter]);
259
260
261              internal_nearness = temp+internal_nearness;
262             /* printf(" internal_nearness=%d",internal_nearness) ;
*/
263            };
264         return internal_nearness;
265     }
```