#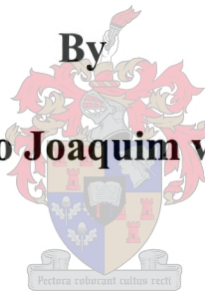 Analysis of the Mobilization Region of the Broad Host-Range IncQ-like Plasmid, pTC-F14, and its Ability to Interact with a Related Plasmid, pTF-FC2.

By

Leonardo Joaquim van Zyl

Thesis presented in partial fulfilment for the degree of

Master of Sciences at the University of Stellenbosch

Supervisor: Prof. D.E. Rawlings

December 2003

Declaration

I, the undersigned, hereby declare that the work contained in this thesis is my own original work and that I have not previously in its entirety or in part submitted it at any university far a degree.

Abstract

The 14.2 kb plasmid pTC-F14 was isolated from the moderately thermophilic (45°-50°C), highly acidophilic (pH 1.5 to 2.5), chemolithotrophic bacterium *Acidithiobacillus caldus* and has a replicon that is closely related to the promiscuous, broad host-range, IncQ-family of plasmids. The region containing the mobilization genes was sequenced and encoded five Mob proteins and an origin of transfer, which are related to the DNA processing (Tra1) region of IncP1 plasmids, rather than to the three Mob protein systems of the IncQ-1-group plasmids (e.g. plasmids RSF1010 or R1162). Plasmid pTC-F14 is the third example of an IncQ family plasmid that has five *mob* genes, with the others being pTF-FC2 and pRAS3.1. The minimal region that was essential for mobilization included the *mobA*, *mobB* and the *mobC* genes as well as the *oriT*. The *mobD* and *mobE* genes were non-essential, but together enhanced the mobilization frequency by approximately 300-fold. The *repB* gene increased the mobilization frequency but was not essential for mobilization.

Mobilization of pTC-F14 between *Escherichia coli* strains by a chromosomally integrated RP4 plasmid was more than 3500-fold less efficient than the mobilization of pTF-FC2. When both plasmids were co-resident in the same *E. coli* host, pTC-F14 was mobilized at almost the same frequency as pTF-FC2. This enhanced pTC-F14 mobilization frequency was due to the presence of a combination of the pTF-FC2 *mobD* and *mobE* gene products, the functions of which are still unknown. pTF-FC2 could mobilize the *oriT* of pTC-F14 whereas pTC-F14 could only mobilize the pTF-FC2 *oriT* if provided with some of the mobilization genes from the pTC-FC2 mobilization region. Unexpectedly either the *mobEDC* genes or the *mobAB* genes would allow the mobilization of the pTF-FC2 *oriT* by pTC-F14 even though there was no common gene between the two subsets. No evidence for any negative effect on the transfer of one plasmid by the related, potentially competitive plasmid was obtained.

Opsomming

Die 14.2 kb plasmied, pTC-F14, is uit die matig termofiliese (45ºC tot 50ºC), hoogs asidofiliese (pH 1.5 tot 2.5), chemolitooutotrofiese bakterium *Acidithiobacillus caldus* geisoleer en beskik oor `n replikon wat verwant is aan die vanaf die IncQ-familie van plasmiede. Hierdie plasmiede is alom bekend vir hulle promiskuïteit tydens konjugasie asook hul vermoë om in `n groot aantal verskillende gasheer organismes te kan repliseer. DNA volgorde analise van die mobiliserings area het `n oordrags oorsprong asook vyf oop leesrame onthul wat nader verwant is aan die DNA prosseserings gene van die Tra1 area op die IncP1 plasmiede, as die van die mobiliserings stelsel van die IncQ-1-groep plasmiede. Plasmied pTC-F14 is die derde voorbeeld, saam met pTF-FC2 en pRAS3.1, van `n IncQ-tipe plasmied met `n vyf-geen mobiliserings sisteem. Die kleinste area op die plasmied nodig vir mobilisering van pTC-F14 is bepaal, en het die *mobA*, *mobB* en *mobC* gene sowel as die oordrags oorsprong ingesluit. Saam, was die *mobD* en *mobE* gene verantwoordelik vir `n 300-voud toename in die mobilisasie frekwensie van pTC-F14 alhowel die gene nie absoluut nodig was vir mobilisering van die plasmied nie. Die *repB* geen het ook bygedra tot die frekwensie waarteen die volledige plasmied gemobiliseer was, maar hierdie geen was ook nie nodig vir mobilisering van die pTC-F14 plasmied nie.

Die frekwensie waarteen pTC-F14 tussen *Escherichia coli* rasse beweeg het tydens konjugasie, terwyl gebruik gemaak is van `n chromosomaal geintegreerde RP4 plasmied, was ongeveer 3500-voud laer as die van pTF-FC2. Indien beide pTC-F14 en pTF-FC2 in dieselfde *E. coli* gasheer aangetref word, word beide plasmiede teen ongeveer dieselfde frekwensie gemobiliseer. Die verhoogde frekensie vir pTC-F14 was as gevolg van die teenwoordigheid van beide die *mobD* en *mobE* gene van die pTF-FC2 plasmied, waarvan die funksies nog onbekend is. Plasmied pTF-FC2 kon die oordrags oorsprong van pTC-F14 mobiliseer waarteenoor plasmied pTC-F14 die oordrags oorsprong vanaf pTF-FC2 slegs kon mobiliseer indien een van twee dele van die pTF-FC2 mobiliserings gene voorsien word (al was daar geen oorvleuling tussen die twee nie). Alhoewel die plasmiede moontlik kon kompeteer op die vlak van plasmied oordrag is geen negatiewe kompetesie waargeneem tussen dié twee verwante plasmiede nie.

Table of Contents

Acknowledgements

I would like to thank my supervisor, Professor Douglas Rawlings for the opportunity and privilege to work in his laboratory. His boundless enthusiasm for the subject not only inspired me to deliver my best, but also served as encouragement whenever we experienced difficulties during the project.

My appreciation goes to both Professor Rawlings and Dr Shelly Deane who were always willing to advise and discuss this work no matter how busy they were.

Thanks to everybody in the Biomining Research Group for making my time working on this project a pleasant experience and a special thanks to Dr Nicolette Coram for her friendship and support.

My gratitude goes to my parents for their moral and financial support.

I acknowledge financial support from the National Research Foundation.

The research was funded by BHP Billiton.

Abbreviations

A	adenosine
A	alanine
aa	amino acids
A + T-rich	adenosine and thymidine rich sequence
ADP	adenosine 5'-diphosphate
Amp	ampicillin
ATP	adenosine 5'-triphosphate
bp	base pair (s)
°C	degrees Celsius
C	cytosine
C-terminal	carboxyl-terminus
Cm	chloramphenicol
Cys	cysteine
D	aspartic acid
Da	Daltons
DNA	deoxyribonucleic acid
DsDNA	double stranded deoxyribonucleic acid
dso	double strand origin
Dtr	DNA transfer and replication functions
G	guanine
G + C-rich	guanosine and cytosine rich sequence
Glu	glutamic acid
GSP	general secretory pathway
GTP	guanine 5'-triphosphate
h	hour (s)
H	histidine
His	histidine
IHF	integration host factor
Inc	incompatibility group (s)
IPTG	isopropyl-β-D-thiogalactopyranoside
IVSS	type IV secretion system (s)
kb	kilobase pair (s) or 1000-bp
kDa	kilodaltons or 1000 daltons
Km	kanamycin
LA	luria agar
M	methionine
mg	milligram

vii

min	minute (s)
ml	milliliter
Mpf	mating pair formation function
MW	molecular weight
μg	microgram
μl	microlitre
NBD	nucleotide binding domain
NCBI	National Center for Biotechnology Information
N-terminal	amino-terminus
ng	nanogram
nm	nanometer
ORF	open reading frame
oriT	origin of transfer
oriV	origin of vegetative replication
p	plasmid
PAI	pathogenicity island
PCR	polymerase chain reaction
PT	pertussis toxin
RBS	ribosome binding site
RCR	rolling circle replication
rpm	revolutions per minute
rRNA	ribosomal ribonucleic acid
S	serine
ssDNA	single stranded deoxyribonucleic acid
sso	single strand origin
T	thymine
T-DNA	transfer deoxyribonucleic acid
TMH	trans membrane helix
Tyr	tyrosine
UV	ultraviolet
w/v	weight per volume

Chapter One

Literature review

Contents

1.1 Introduction.

Bacterial plasmids are extrachromosomal species of DNA (Linear-dsDNA, circular-dsDNA or ssDNA) that replicate independently of the host chromosome and although they often carry genes that are beneficial to the host they are non-essential for host survival (Novick *et al*, 1980). They range in size from 2kb to several hundred kilobase-pairs and are found in most bacterial populations: some bacteria carrying more than one type of plasmid or more than one copy of a certain plasmid. Because of this it is thought that they can constitute between 1% and 10% of the genome of a bacterial species.

Bacterial plasmids are highly mobile genetic elements moving not only between bacteria of the same species but across species, genera and even kingdoms (Buchanan-Wollaston *et al*, 1987; Gormley *et al*, 1991; Margensin *et al*, 1997; Rawlings *et al*, 2001). This allows for the spread of traits that help bacteria adapt rapidly to changes in their environment. Some of these traits include antibiotic resistance, resistance to heavy metals, virulence determinants, the ability to degrade aromatic compounds, the ability to fix nitrogen, UV resistance (plasmid pAD1), resistance to radiation and increased mutation frequency (Clewell *et al*, 1986).

Conjugation is one of the processes through which bacteria take up plasmid DNA and occurs both in gram-positive and gram-negative bacteria. This involves the transfer of DNA from a donor to a recipient cell through an as yet unspecified conduit. For plasmids isolated from gram-negative bacteria, three plasmid encoded elements are required for mobilization to take place: an origin of transfer (*oriT*, required in *cis*), a mating pair formation system (Mpf) required for donor-recipient contact, as well as a DNA transfer and replication system (Dtr) for processing of the DNA at the *oriT* prior to transfer.

Two goups of mobile plasmids have emerged. Plasmids that carry both Mpf and Dtr determinants are said to be conjugative or self-transmissible, which affords them great autonomy when moving in the gene pool, whilst plasmids lacking the Mpf genes are

referred to as mobilizable. The latter can be transferred in the presence of a conjugative plasmid or "helper" plasmid that supplies the Mpf gene products.

Although nearly twenty incompatibility groups (Inc) exist for *Escherichia coli* plasmids, few have been studied in detail with regards to their mobilization systems. Of these, five variants have been identified and it would appear that these gram-negative plasmid transfer mechanisms have many features in common with the biggest difference discerning them being a lack of overall homology between these systems (Zechner et al, 2000).

1.2 Bacterial conjugation: three model systems for gram-negative plasmid transfer

1.2.1 IncPα plasmid RP4

The IncP group consists of conjugative, broad-host-range plasmids that produce thick inflexible pili, which conjugate preferentially on semi-solid media. These plasmids are divided into two groups: IncPα and IncPβ represented by their respective archetypal plasmids RP4 and R751 (Pansegrau *et al*, 1994a; Thorsted *et al*, 1998). They are of particular interest because of their ability to promote gene spread between diverse bacterial species as well as their own promiscuity that serves as a model for movement of plasmid-borne antibiotic resistance genes in clinical environments (Lowbury *et al*, 1969; Fulbrook *et al*, 1970; Sykes *et al*, 1970; Datta *et al*, 1971; Holloway *et al*, 1973; Ingram *et al*, 1973; Chandler *et al*, 1974; Guiney *et al*, 1984; Breton *et al*, 1985; Greener *et al*, 1992; Giebelhaus *et al*, 1996; Bates *et al*, 1998; Hoffman *et al*, 1998; Thorsted *et al*, 1998; Samuels *et al*, 2000; Christie *et al*, 2001). Other than the F plasmid mobilization system, the RP4 (IncPα) mechanism of transfer has been intensively studied and has thus become one of the model systems for describing gram-negative plasmid transfer.

1.2.1.1 DNA processing at the origin of transfer

The first step in the current model for gram-negative plasmid transfer holds that to produce the single stranded DNA molecule for transfer, a site and strand specific nick has to be introduced. This reaction occurs within a protein complex, designated the relaxosome, found to form at the *oriT* and is composed of the Dtr system gene products.
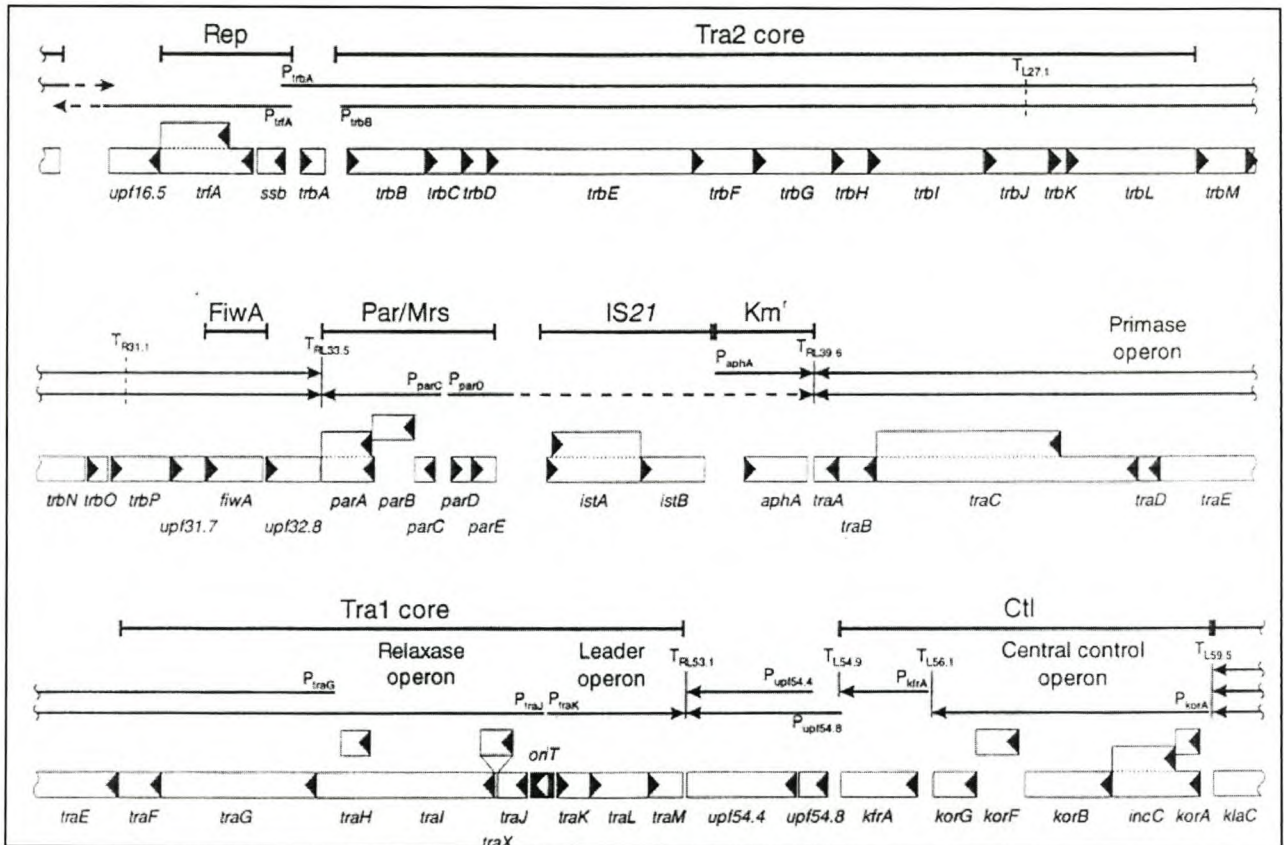
Figure 1.2.1.1.1: Genetic organization of the IncPα transfer system. Distinct regions on the IncPα map are marked by bold lines. The extension of transcripts is indicated by light arrows. Where transcription endpoints are not defined, lines are broken. Boxes with triangles mark the position of the genes. Filled triangles indicate the direction of translation. Black boxes represent intergenic regions of defined function, i.e. origin of transfer DNA replication. Open triangles within black boxes indicate the direction of transfer replication (Pansegrau *et al*, 1994b).

The mobilization region of plasmid RP4 consists of two core regions: Tra1 and Tra2 separated by a kanamycin resistance gene, IS21, a partitioning system and the *fiwA* locus (Figure 1.2.1.1.1) (Lessl *et al*, 1992a; Lessl *et al*, 1993). Tra2 encompasses the Mpf genes whilst Tra1 encodes the DNA processing machinery. Tra1 consists of 17 ORF's and includes the *oriT*. Not all of the ORF's in Tra1 are involved with DNA processing or are needed for transfer to take place. Deletion analysis of this section has shown that the regions from *traA* to *traE* and *traL* to *traO* are not absolutely required for transfer, at least between *Escherichia coli* cells. If *traM* is supplied, the transfer efficiency is increased suggesting that it could assist conjugation. It is also hypothesized that the *tra* genes that are apparently not needed might be used when plasmid transfer takes place between bacterial hosts other than *E. coli* (Lessl *et al*, 1993).   TraE has recently been shown to be a homologue of the *E. coli* DNA topoisomerase III and could complement insertional inactivation of the *topB* gene in *E. coli* K38 (Li *et al*, 1997). A possible reason as to why deletion or inactivation of the plasmid-borne copy would not result in a transfer deficient phenotype, could be that the hosts (*E. coli.* in most studies) Topo III masks this deficiency. Two possible roles for *traE* were put forward: firstly as a mechanism to ensure replicating plasmids are separated prior to partitioning and secondly to assist in unwinding of the DNA helix during conjugative transfer of the plasmid. No helicase to perform this function has been identified in RP4. Being a broad-host-range system this would appear to be a safety feature to guard against entering a cell without such a mechanism.

The *traC* gene encodes two protein products: TraC1 (118kDa) and TraC2 (80kDa). Both have been shown to have primase activity and are involved in the vegetative replication of the plasmid. One of the questions in plasmid transfer is how the plasmid is reconstituted in the recipient in terms of circularization and replacement of the double strand. Complementary strand synthesis is thought to occur as the T-strand enters the cell, and because the DNA enters in the 5' to 3' direction the synthesis of the complementary strand would have to be discontinuous requiring multiple primers. One hypothesis is that the plasmid-encoded primase is transferred along with the single-stranded DNA to the recipient where it generates the primers from which the complementary strand can be synthesized. The primases of RP4 and the Collb-P9 (IncI1) plasmid are indeed transferred to the recipient, although for RP4 the primase is not required for plasmid transfer to take place (Rees *et al*, 1990). It could be that a

host-encoded (*E. coli*) primase might be able to mimic the RP4 primase function as TraC1 can suppress *E. coli dnaG* mutants (Strack *et al*, 1992). A proposed role for the RP4 TraC1 in conjugation is that it could facilitate plasmid replication in a recipient, which does not have a homologous equivalent much like the *traE*, acting as a "fail safe" (Lanka *et al*, 1981; Nash *et al*, 1988). An observation that lends credence to this idea is that some narrow-host-range systems like the F group of plasmids do not encode or transfer a primase and rely on host encoded primases for replication, which may limit the variety of hosts they may be transferred to. The Collb-P9 plasmid primase, Sog, is needed for DNA transfer, but is not reliant on co-transfer with DNA to be exported to the recipient cell (Wilkins *et* al, 2000). This would suggest that the primase could coat the T-strand and facilitate its transport to the recipient (section 1.3). The two genes found in Tra1 that are not involved with Dtr are *traF* and *traG*. Both these genes are essential transfer components and will be discussed later.

The relaxases are the best characterized of the relaxosome proteins and are responsible for introducing the site and strand specific nick prior to transfer. They are also involved in the termination reaction catalyzing the circularization of the transferred strand. Inspection of various relaxase proteins shows three conserved domains/motifs within the N-terminal part that contains the catalytic domain(s) (Figure 1.2.1.1.2). The C-terminal shows great variability and is thought to interact with other components of the transfer machinery (TraG or other relaxosome
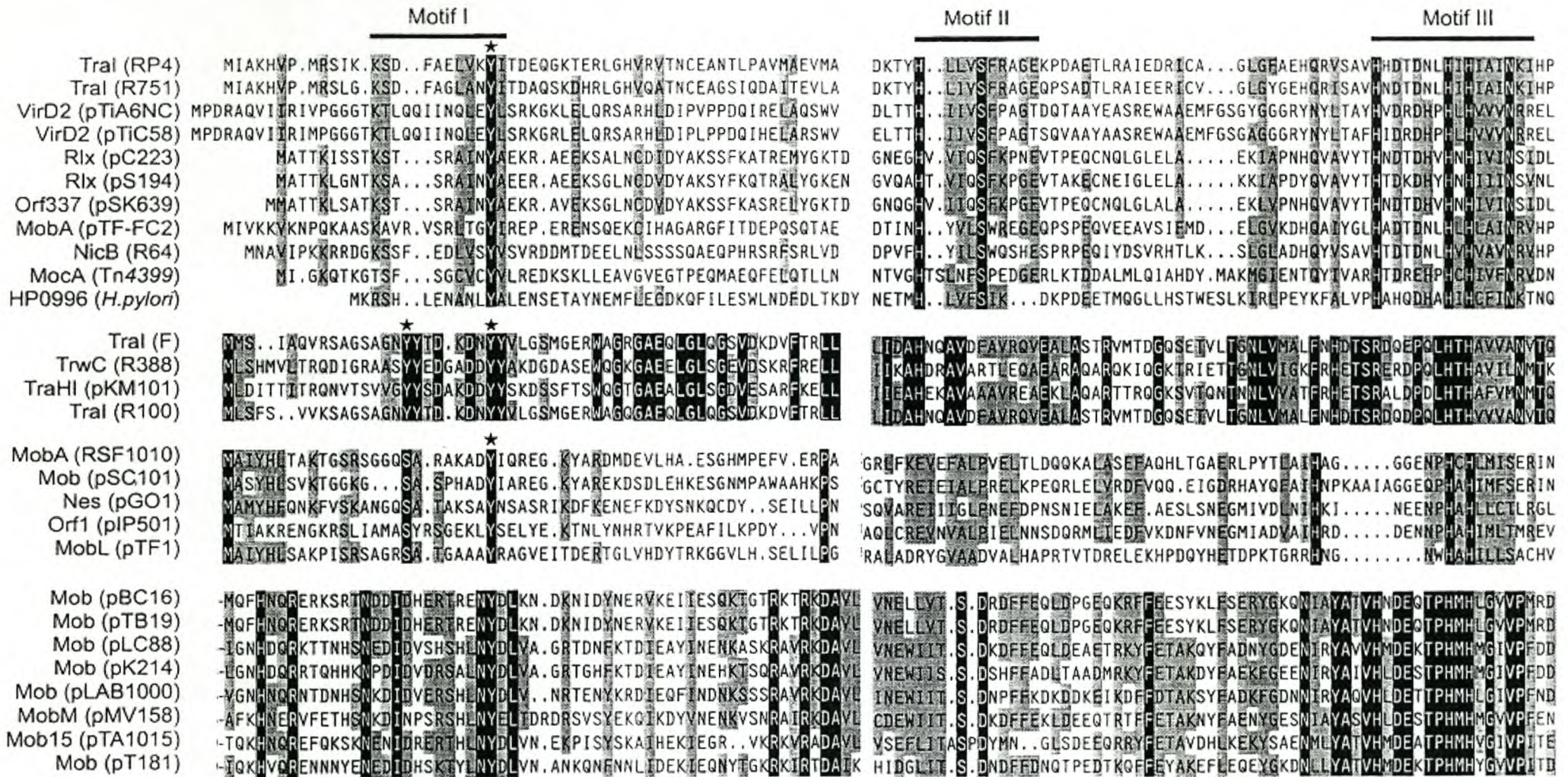
Motif I   Motif II   Motif III

```
TraI (RP4)      MIAKHVP.MRSIK.KSD..FAELVKYITDEQGKTERLGHVRVTNCEANTLPAVMAEVMA    DKTYH..LLVSFRAGEKPDAETLRAIEDRICA...GLGFAEHQRVSAVHDTDNLHIHIAINKIHP
TraI (R751)     MIAKHVP.MRSLG.KSD..FAGLANYITDAQSKDHRLGHVQATNCEAGSIQDAITEVLA    DKTYH..LIVSFRAGEQPSADTLRAIEERICV...GLGYGEHQRISAVHNDTDNLHIHIAINKIHP
VirD2 (pTiA6NC) MPDRAQVIIRIVPGGGTKTLQQIINQLEYLSRKGKLELQRSARHLDIPVPPDQIRELAQSWV    DLTTH..IIVSFPAGTDQTAAYEASREWAAEMFGSGYGGGRYNYLTAYHVDRDHPLHVVVNRREL
VirD2 (pTiC58)  MPDRAQVIIRIMPGGGTKTLQQIINQLEYLSRKGRLELQRSARHLDIPLPPDQIHELARSWV    ELTTH..IIVSFPAGTSQVAAYAASREWAAEMFGSGAGGGRYNYLTAFHIDRDHPLHVVVNRREL
Rlx (pC223)     MATTKISSTKST...SRAINYAEKR.AEEKSALNCDIDYAKSSFKATREMYGKTD    GNEGHV.VTQSFKPNEVTPEQCNQLGLELA.....EKIAPNHQVAVYTHNDTDHVHNHIVINSIDL
Rlx (pS194)     MATTKLGNTKSA...SRAINYAEER.AEEKSGLNCDVDYAKSYFKQTRALYGKEN    GVQAHT.VIQSFKPGEVTAKECNEIGLELA.....KKIAPDYQVAVYTHTDKDHYHNHILINSVNL
Orf337 (pSK639) MMATTKLSATKST...SRAINYAEKR.AVEKSGLNCDVDYAKSSFKASRELYGKTD    GNQGHV.IIQSFKPGEVTPEQCNQLGLALA.....EKLVPNHQVAVYTHNDTDHVHNHIVINSIDL
MobA (pTF-FC2)  MIVKKYKNPQKAASKAVR.VSRLTGYIREP.ERENSQEKCIHAGARGFITDEPQSQTAE    DTINH..YVLSWREGEQPSPEQVEEAVSIEMD...ELGVKDHQAIYGLHADTDNLHLAINRVHP
NicB (R64)      MNAVIPKKRRDGKSSF..EDLVSYVSVRDDMTDEELNLSSSSQAEQPHRSRFSRLVD    DPVFH..YILSWQSHESPRPEQIYDSVRHTLK...SLGLADHQYVSAVHTDTDNLHVHVAVNRVHP
MocA (Tn4399)   MI.GKQTKGTSF...SGCVCYVLREDKSKLLEAVGVEGTPEQMAEQFELQTLLN    NTVGHTSLNFSPEDGERLKTDDALMLQIAHDY.MAKMGIENTQYIVARHTDREHPHCHIVFNRVDN
HP0996 (H.pylori) MKRSH..LENANLYALENSETAYNEMFLEGDKQFILESWLNDFDLTKDY    NETMH..LVFSIK...DKPDEETMQGLLHSTWESLKIRLPEYKFALVPHAHQDHAHIHCFINKTNQ
```

```
TraI (F)        MMS..IAQVRSAGSAGNYYID.KDNYYVLGSMGERWAGRGAEQLGLQGSVDKDVFTRLL    ITDAHNQAVDFAVRQVEALASTRVMTDGQSETVLTGNLVMALFNHDTSRDQEPQLHTHAVVANVTQ
TrwC (R388)     MLSHMVLTRQDIGRAASYYEDGADDYYAKDGDASEWQGKGAEELGLSGEVDSKRFRELL    ITKAHDRAVARTLEQAEARAQARQKIQGKTRIETTGNLVIGKFRHETSRERDPQLHTHAVILNMTK
TraHI (pKM101)  MLDITTITRQNVTSVVGYYSDAKDDYYSKDSSFTSWQGTGAEALGLSGDVESARFKELL    ITEAHEKAVAAAVREAEKLAQARTTRQGKSVTQNTNNLVVATFRHETSRALDPDLHTHAFVFMNMTQ
TraI (R100)     MLSFS..VVKSAGSAGNYYID.KDNYYVLGSMGERWAGQGAEQLGLQGSVDKDVFTRLL    IDAHNQAVDFAVRQVEALASTRVMTDGQSETVLTGNLVMALFNHDITSRDQDPQLHTHVVYANVTQ
```

```
MobA (RSF1010)  NAIYHITAKTGSRSGGQSA.RAKADYIQREG.KYARDMDEVLHA.ESGHMPEFV.ERPA    GRLFKEVEFALPVELTLDQQKALASEFAQHLTGAERLPYTLAIHAG.....GGENPICHLMISERIN
Mob (pSC101)    MASYHLSVKTGGKG...SA.SPHADYIAREG.KYAREKDSDLEHKESGNMPAWAAHKPS    GCTYREIEIALPRELKPEQRLELYRDFVQQ.EIGDRHAYQEAIHNPKAAIAGGEQPHAIMFSERIN
Nes (pGO1)      MAMYHFQNKFVSKANGQSA.TAKSAYNSASRIKDFKENEFKDYSNKQCDY..SEILLPN    SQVAREIIGLPNEFDPNSNIELAKEF.AESLSNEGMIVDLNIHKI.....NEENPHAHLLECTLRGL
Orf1 (pIP501)   NTIAKRENGKRSLIAMASYRSGEKLYSELYE.KTNLYNHRTVKPEAFILKPDY...VPN    AQLCREVNVALBIELNNSDQRMLIEDFVKDNFVNEGMIADVAIHRD.....DENNPHAHIMLTMREV
MobL (pTF1)     NAIYHLSAKPISRSAGRSA.TGAAAVRAGVEITDERTGLVHDYTRKGGVLH.SELILPG    RALADRYGVAADVALHAPRTVTDRELEKHPDQYHETDPKTGRRHNG........NWHAHILLSACHV
```

```
Mob (pBC16)     -MQFHNQRERKSRTNDDIDHERTRENYDLKN.DKNIDYNERVKEIIESQKTGTRKTRKDAVL    VNELLVT.S.DRDFFEQLDPGEQKRFFEESYKLFSERYGKQNIAYATVHNDEQTPHMHLGVVPMRD
Mob (pTB19)     -MQFHNQRERKSRTNDDIDHERTRENYDLKN.DKNIDYNERVKEIIESQKTGTRKTRKDAVL    VNELLVT.S.DRDFFEQLDPGEQKRFFEESYKLFSERYGKQNIAYATVHNDEQTPHMHLGVVPMRD
Mob (pLC88)     -LGNHDQRKTTNHSNEDIDVSHSHLNYDLVA.GRTDNFKTDIEAYINENKASKRAMRKDAVL    VNEWIIT.S.DKDFFEQLDEAETRKYFETAKQYFADNYGDENIRYAMVHMDEKTPHMHMGIVPFDD
Mob (pK214)     -LGNHDQRRTQHHKNPDIDVDRSALNYDLVA.GRTGHFKTDIEAYINEHKTSQRAMRKDAVL    VNEWIIS.S.DSHFFADLTAADMRKYFETAKDYFAEKFGEENIRYAIVHLDESTPHMHMGIVPFDD
Mob (pLAB1000)  -VGNHNQRNTDNHSNKDIDVERSHLNYDLV..NRTENYKRDIEQFINDNKSSSRAVRKDAVL    INEWIIT.S.DNPFFKDKDDKEIKDFFDTAKSYFADKFGDNNIRYAQVHLDETTPHMHLGIVPFND
MobM (pMV158)   -AFKHNERVFETHSNKDINPSRSHLNYELIDRDRSVSYEKGIKDYVNENKVSNRAIRKDAVL    CDEWIIT.S.DKDFFEKLDEEQTRTFFETAKNYFAENYGESNIAYASVHLDESTPHMHMGVVPFEN
Mob15 (pTA1015) -TQKHNGREFQKSKNENIDRERTHLNYDLVN.EKPISYSKAIHEKIEGR..VKRKVRADAVL    VSEFLITASPDYMN..GLSDEEQRRYFETAVDHLKEKYSAENMLYATVHMDEATPHMHVGIVPITE
Mob (pT181)     -IQKHVQRENNNYENEDIDHSKTYLNYDLVN.ANKQNFNNLIDEKIEQNYTGKRKIRTDAIK    HIDGLIT.S.DNDFFDNQTPEDTKQFFEYAKEFLEQEYGKDNLLYATVHMDEKTPHMHYGVVPITD
```

Figure 1.2.1.1.2: Alignment of conjugative DNA relaxases. Amino acid positions that are conserved throughout are drawn with a black background. A shaded background marks positions where conservative replacements may occur. The conserved motif's identified in TraI of RP4 are indicated by black lines above sequence blocks. An asterisk marks the active tyrosine (Zechner *et al*, 2000

7

components). IncP type relaxases would appear to be the most widely distributed occurring in gram-positive and gram-negative bacteria, conjugative transposons, mobilizable elements (pTF-FC2) and the Agrobacterial T-DNA transfer system (section 1.3) (Zechner *et al*, 2000).
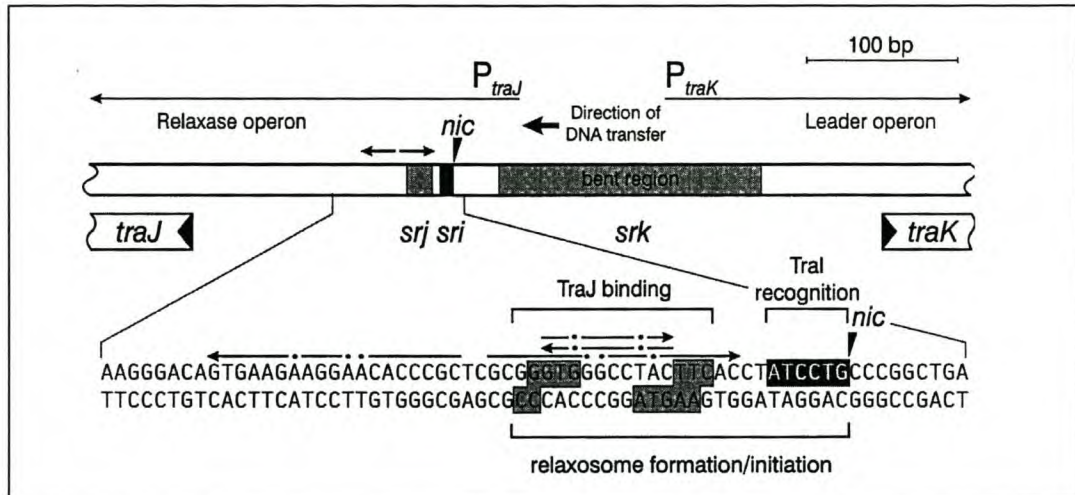

Figure 1.2.1.1.3: Structural organization of the RP4 transfer origin. Arrows P_{traJ} and P_{traK} indicate the divergent transcription start sites and direction for the leading region and the relaxase operon. Bold arrows indicate the inverted repeat adjacent to the *nic* site (wedge). The shaded bars (*sri*, *srj* and *srk*) show binding sites for the respective transfer gene products. Dots within the arrows above the enlarged *oriT* sequence indicate deviations from the symmetry. Shaded boxes in the TraJ binding area mark nucleotides, which are protected by TraJ during hydroxyl radical footprinting (Zieglin *et al*, 1989). The wedges indicate the position of the *nic* (cleavage) site. Nucleotides that are recognized by TraI, which are next to the *nic* site, are indicated as white text on a black background (Pansegrau *et al*, 1996).

TraI has been identified as the RP4 relaxase and recognizes a highly conserved nucleotide hexamer adjacent to the cleavage site (*sri*) (Figure 1.2.1.1.3) (Pansegrau *et al*, 1993). Sequencing of a tetrapeptide, found after treatment of TraI bound to an *oriT*-containing oligonucleotide with proteinase K, shows tyrosine 22 located in motif I to bind covalently to the DNA (Pansegrau *et al*, 1993). Cleavage of the *oriT* sequence is abolished by replacement of Tyr-22 with leucine or arginine and also results in transfer deficient phenotypes, lending further evidence to Tyr-22 being the catalytic active-site residue in TraI (Balzer *et al*, 1994). Mutation of serine 74 to alanine (motif II) leads to an increase in the topoisomerase activity and results in all of

the super coiled input *oriT* DNA being converted to the covalently closed relaxed form, I^O. The result demonstrates that the cleave-joining ability of the mutant is not affected and is explained by impaired binding of the TraI S74A to the *oriT* leading to the occasional release of the 3' end at *nic* when the plasmid is in the cleaved state. Following spontaneous relaxation of the plasmid DNA, the relaxase could reseal the cleaved strand. This confirmed the role of motif II in recognition and tight binding of the 3' terminus (*sri*) at the *oriT*, but it is probably not involved in the cleave-joining activity of the relaxase (Pansegrau *et al*, 1994b; Pansegrau *et al*, 1996). Motif III, together with motif I, is thought to form the catalytic center of TraI. Exchange of His-116 with a serine residue, results in the loss of detectable cleaving activity, whilst retaining the ability to form stable relaxosomes. In motif III, amino acids 111-113 constitute a "DTD" motif, which could coordinate Mg^{2+} ions essential for catalytic activity (cleaving-joining reactions) of TraI but are not required for relaxosome formation (Argos *et al*, 1988; Pansegrau *et al*, 1990b). Mutation of either of the asparagine residues severely reduces its cleave-joining ability but does not affect relaxosome formation (Pansegrau *et al*, 1996). This is said to confirm the role of motif III as being part of the catalytic center (Figure 1.2.1.1.4).
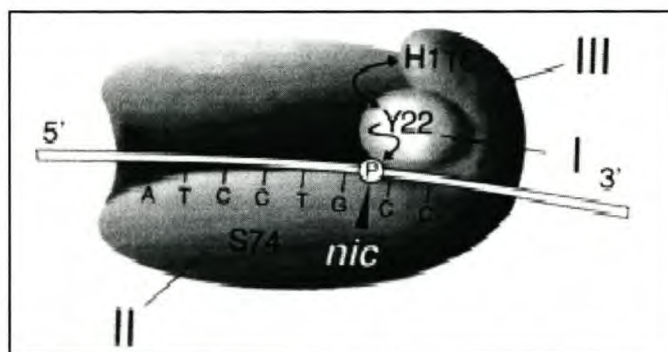
Figure 1.2.1.1.4: A model of the TraI catalytic center. Motif's I-III identified in TraI are indicated. Arrows indicate interaction between amino acids or between an amino acid residue and the DNA. The *nic* site is indicated with a wedge (Pansegrau *et al*, 1994b).

His-116 is thought to activate the hydroxyl group on Tyr-22 by proton abstraction giving an efficient nucleophile (the tyrosyl oxygen) to attack the phosphodiester backbone at *nic* (Byrd *et al*, 1997). TraI thus becomes covalently bound to the 5' end of the cleaved DNA through Tyr-22 (Guiney *et al*, 1975; Pansegrau *et al*, 1990a). It is thought that the IncP-type relaxases function as a dimer, which enables them to

catalyze a second cleavage reaction when a greater than unit length substrate is produced by strand transfer replication during conjugal transfer (Figure 1.2.1.1.5). There appear to be many similarities between the mechanisms involved in rolling circle replication and plasmid transfer. These will be discussed in detail in section 1.5
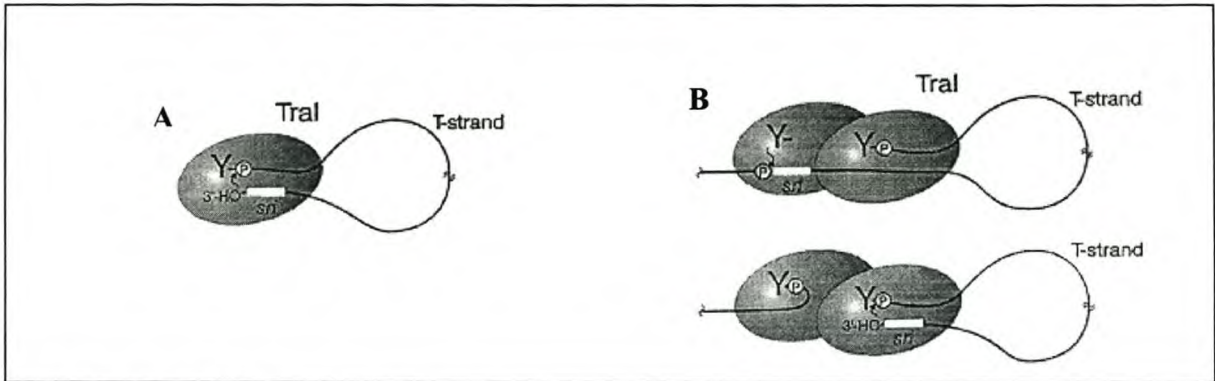
and section 1.5.1.

Figure 1.2.1.1.5: Model for termination of transfer DNA replication. A, closing of the T-strand through joining of a free 3'-OH⁻ end at *nic* directly to the 5' bound moiety. B, proposed mechanism of termination for the TraI dimer via second cleavage and circularization (Pansegrau *et al*, 1996).

Three gene products from the Tra1 region have been shown to be absolutely required for stable relaxosome formation: TraJ, TraI and TraH (Pansegrau *et al*, 1990b). The first step in relaxosome formation is recognized as the binding of TraJ to a 10-bp palindromic sequence (*srj*) to the right half of a 19-bp inverted repeat sequence in *oriT* as shown by DNAseI footprinting analysis (Figure 1.2.1.1.3) (Furste *et al*, 1989; Zieglelin *et al*, 1989). Next the TraI recognizes and binds to the TraJ•*srj* probably through protein-protein interactions between TraJ and TraI as well as protein-DNA interaction in recognizing the *sri* (Pansegrau *et al*, 1990a; Pansegrau *et al*, 1990b). Although the acidic oligomeric protein, TraH, is not essential for plamsid transfer it's binding, together with TraJ and TraI, to supercoiled DNA (containing the *oriT* sequence) decreases the electrophoretic mobility in non-denaturing gels, which is indicative of stable complex formation. It also increases the amount of specifically relaxed DNA observed when reconstructing the relaxosome *in vitro*. TraH is thought to stabilize the relaxosome by binding to both the TraJ and TraI, as it has no apparent *oriT* binding properties and complexes with these proteins in the absence of any *oriT*

DNA (Pansegrau *et al*, 1990b). Another essential transfer component, TraK, binds to and wraps ~200bp of DNA (*srk*) around a core of multimeric TraK subunits. This is proposed to influence the superhelicity at the *oriT* region and as such allows for better positioning of relaxosome components increasing the fraction of DNA that can be captured in the nicked state (Ziegelin *et al*, 1992).

Promoters *traJp* and *traKp* are regulated in two ways: (i) The relaxosome complex that forms at the *oriT* denies access to the RNA polymerase or, (ii) The TraI and TraJ proteins prevent the RNA polymerase from entering the elongation phase after initiation at either promoter. These promoters are also autoregulated by their various transcriptional products. TrbA, coded for by the Tra2 region, represses *traJp*, *traKp* and *traGp* to keep expression of the Tra1 region coordinated with Tra2 as well as replication and inheritance functions (Figure 1.2.1.2.4) (Jagura-Burdzy *et al*, 1992; Zatyka *et al*, 1994; Zechner *et al*, 2000).

1.2.1.2 Mating pair formation

To be able to transfer plasmid DNA from one cell to another, some form of contact has to be established between donor and recipient. The DNA processing takes place through the Dtr machinery, but which system is involved in mediating contact between cells? The Mpf system of plasmid RP4 has been well studied and although evidence points to a system suited to protein transport, the way in which nucleoprotein complexes cross the cell membrane remains unclear. These functions have been narrowed to a region of plasmid RP4 called Tra2 (Figure 1.2.1.1.1) (Lessl *et al*, 1992). It consists of 16 ORF's (*trbA – trbP*), thought to be transcribed as a single unit. The criteria used to define the Mpf system of RP4 include: mobilization of the non-selftransmissible plasmid RSF1010 (IncQ), pilus production and donor specific phage propagation. Because RSF1010 encodes only relaxosomal components and needs a Mpf system from a helper plasmid, this makes it ideally suited to studying the Mpf functions while the pilus is required for contact with the recipient and should be an integral part of the Mpf system. A number of bacterial viruses exploit cells containing conjugative plasmids as hosts for their reproduction, and are referred to as donor specific phages. The pilus of RP4 harboring cells has been shown to be the attachment site for phages PRR1 and Pf3 and is the proposed binding site for PRD1,

suggesting a role for the pilus as a means for the viral genome entering the cell (Grahn *et al*, 1997). This knowledge has helped in elucidating which gene products are involved in pilus biosynthesis and formation of the Mpf complex (Kotilainen *et al*, 1993; Daugelavičius *et al*, 1997). Five (*trbA* and *trbM–trbP*) of the Tra2 gene products have been shown not to be required for any of the above-mentioned functions. The remaining twelve ORF's make up the Tra2 core region (Haase *et al*, 1995). The same set of Mpf genes, *trbB-trbL*, is required for pilus assembly, phage reproduction and the formation of conjugative junctions together with TraF and TraG from the Tra1 region (Zechner *et al*, 2000). However, TraG is not required for phage reproduction or formation of conjugative junctions (section 1.4) (Haase *et al*, 1995).

The RP4 pili have been described as being inflexible compared with those of the IncF and IncI (thin pilus) plasmids and are best suited to promote mating on semi-solid media (Zechner *et al*, 2000). These pili are easily removed from the cell surface and it has been proposed that their role in RP4 mediated conjugation is not as important, in initiating conjugation and bringing cells into close contact, as in the F plasmid. Although junctions that look morphologically like those found in wild type RP4 matings have been seen using donors lacking functional pili, DNA transfer does not occur under these conditions (Figure 1.2.1.2.1).
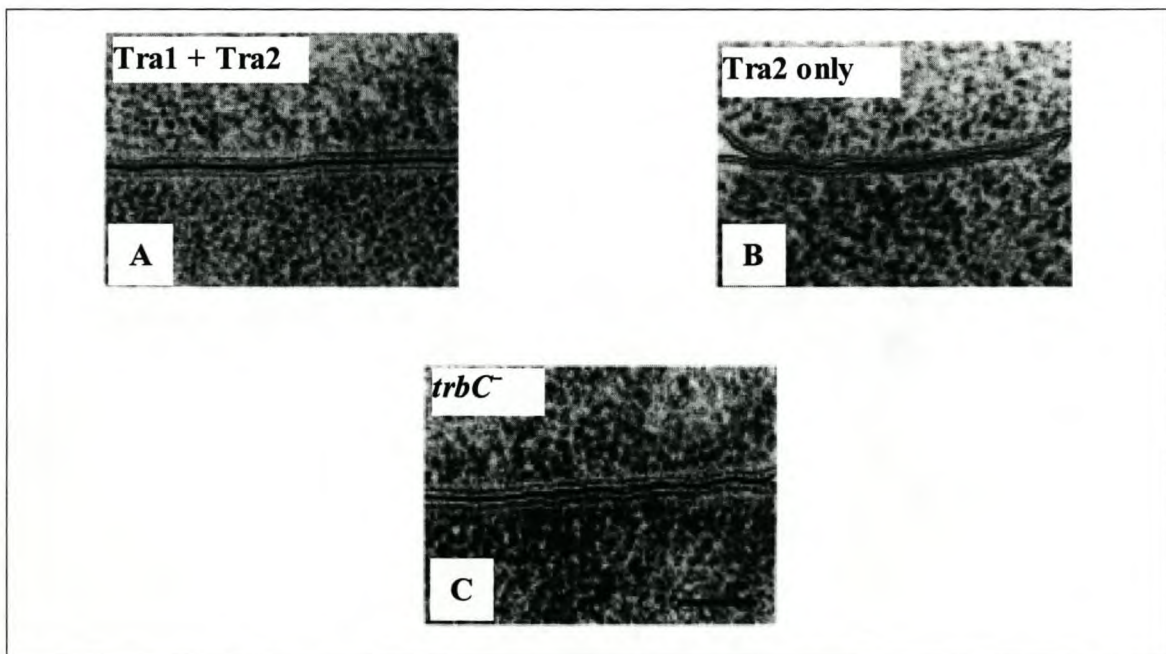
Figure 1.2.1.2.1: Cell-cell association between intraspecific *E. coli* with various donor strains carrying derivatives of the RP4 mating system (Tra1 and Tra2). Panel A represent a reconstituted RP4-mediated mating. The donor carries both Tra1 and Tra2, which together supply all the necessary functions to allow normal pilus formation (*traF* and Tra2). An electron dense area with very few gaps, indicative of a normal mating junction, can be seen between the cells. Panel B shows junctions where the donor carries only the Tra2 region (lacking functional pili). Once again a junction type similar to the reconstituted mating can be seen. Panel C shows the junctions between donor and recipient in the absence of the *trbC* gene. Junctions resembling those of the reconstituted mating can be seen (Samuels *et al*, 2000).

The pili are proposed to stabilize the junctions formed allowing DNA transfer (Haase *et al*, 1995; Samuels *et al*, 2000). TrbC was tentatively identified as the pillin subunit based on sequence similarities with VirB2 from the Ti-plasmids in *Agrobacterium tumefaciens* and TraA from plasmid R1-19 known to be pillin biosynthetic proteins (Lessl *et al*, 1992; Shirasu *et al*, 1993). Antiserum raised against purified RP4 pili recognize the processed forms of TrbC in solid phase immunoassays as well as on pili in immuno gold labeling experiments, indicating that the pillin subunit is encoded by the *trbC* gene (Eisenbrandt *et al*, 1999). Mutational analysis of *traF* showed its product to be an essential transfer component needed for pilus synthesis, mobilization of RSF1010 and phage sensitivity (Waters *et al*, 1992; Lessl *et al*, 1993; Haase *et al*, 1995). Haase and coworkers demonstrated that TraF plays a vital role in the multistep pillin maturation process. Firstly the 15kDa TrbC prepillin is N-terminally modified by removal of the signal peptide through proteolytic cleavage mediated by a host factor, presumable the *E. coli* signal peptidase, Lep. This is followed by removal of 27aa at the C-terminal by an unidentified host-encoded factor (Figure 1.2.1.2.2). The TraF then catalyzes the final step by removing four amino acids from the C-terminal and circularizing the remaining 78aa peptide, joining serine 37 and glycine 114, to form the mature pillin (Figure 1.2.1.2.3) (Haase *et al*, 1997).
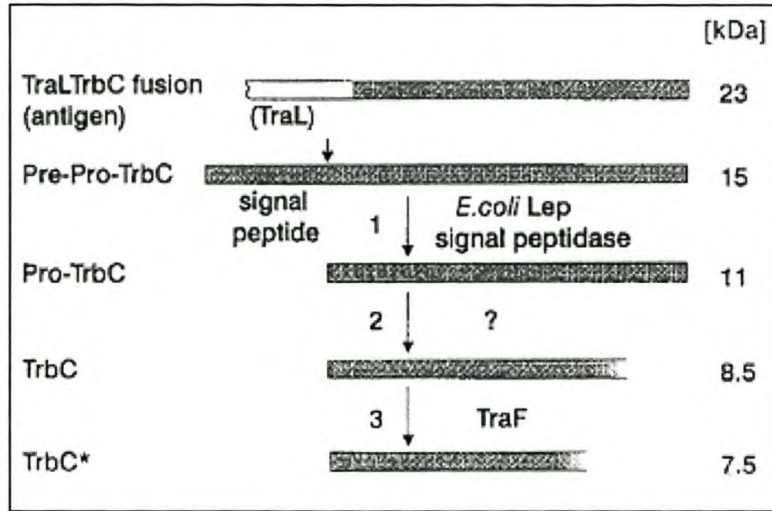
13

Figure 1.2.1.2.2: Model of the TrbC maturation process (Haase *et al*, 1997)

This type of processing and cyclization has been shown for the TrbC homologue VirB2 of Agrobacterial plasmid pTiC58 (IncRH1). The TraF analog identified on pTiC58 is not essential for T-DNA transfer, and both processing and cyclization was shown to occur in *A. tumefaciens* without the plasmid being present. It is possible that the C-terminal of VirD4 could catalyze this reaction as it has some conserved amino acid residues in common with signal peptidases (Haase *et al*, 1997; Eisenbrandt *et al*, 1999). It is thought that an Agrobacterial chromosomally encoded enzyme is involved in forming the peptide linkage as VirB2 does undergo processing but not cyclization in *E. coli*. TraF is proposed to belong to the family of serine proteases similar to Lex-A whose catalytic activity depends on a serine-lysine dyad-like mechanism (Black *et al*, 1993; Eisenbrandt *et al*, 2000). Although the *traF* gene sequence shows similarities with prokaryotic and eukaryotic signal peptidases, it displays high substrate specificity and cannot complement a temperature sensitive *lepB E. coli* mutant (Inanda *et al*, 1989). TraF also contains an excellent signal sequence and is located in the inner membrane.
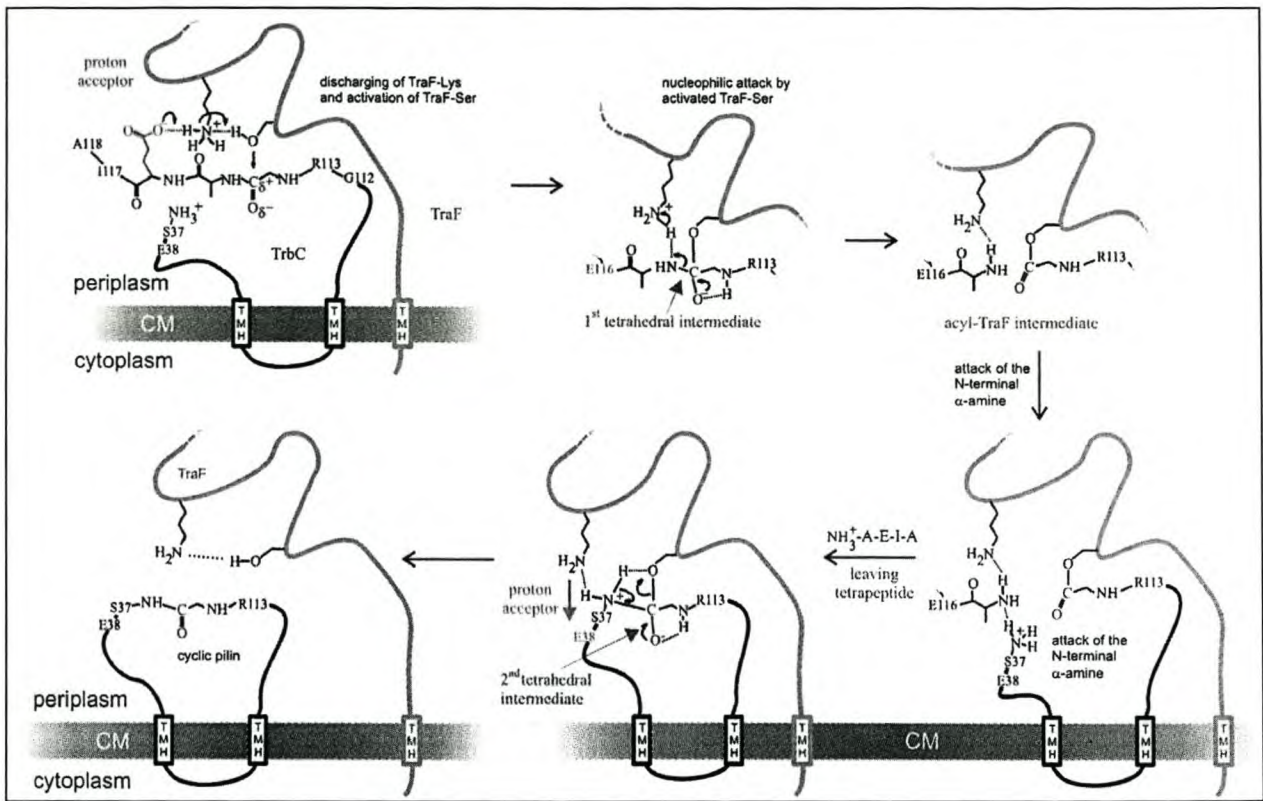
Figure 1.2.1.2.3: Proposed model for the TraF-catalyzed formation of an internal peptide bond in TrbC (Eisenbrandt *et al*, 2000).

The RP4 TrbB protein and its homologues, R388 TrwD and the HP0525 protein from the *cag* pathogenicity island (PAI) in *Helicobacter pylori* form hexameric ring structures that, apart from the TrbB, which is found in the cytoplasm, are associated with the inner membrane (Krause *et al*, 2000a; Krause *et al*, 2000b). These proteins belong to the PulE/VirB11 family of the *sec*-dependant general secretion pathway (GSP) and display weak NTPase activity in vitro, similar to chaperones like DnaK and ClpA (Zylicz *et al*, 1983; Hwang *et al*, 1988; Motallebi-Veshareh *et al*, 1992). Binding, but not hydrolysis, of NTP's to TrbB stabilizes the ring structure which is proposed to act either as a chaperone for unfolded Mpf components helping in the assembly of the complex, or to facilitate transfer of the nucleoprotein complex. The TrbE is a proposed NTPase that is related to the *traC* (IncF) / *virB4* (IncRH1) family of proteins which are tightly associated with the inner membrane. The conserved NTP binding motif is essential for it's function and TrbE is believed to be involved

15

with the transport or positioning of other Mpf components or with energizing the DNA transfer process itself (Lessl *et al*, 1992; Berger *et al*, 1993; Shirasu *et al*, 1994). TrbH is an outer membrane lipoprotein that may play a role similar to that of VirB7 (Ti plasmids) and/or TraV (F-plasmid) in helping to anchor the transmembrane complex in the outer membrane (Grahn *et al*, 2000; Harris *et al*, 2001). TrbK, a small inner membrane lipoprotein, has been identified as the only plasmid encoded protein involved in entry/surface exclusion. This prevents non-productive donor-donor matings and is implemented in two ways: (i) By preventing the formation of stable mating pairs possibly by blocking the pilus binding site (TraT from the F plasmid and Asc10 from pCF10) and, (ii) By preventing DNA from entering the cell (TraS from the F plasmid and TrbK). TrbK expressed by a donor is thought to interact with one or more of the Mpf components, in another donor, perhaps including TraF and TraG but not TrbK itself thereby preventing the formation of a DNA entry pore in the recipient (Haase *et al*, 1995; Haase *et al*, 1996). TrbN appears to belong to a family of transglycosylases that includes TraL from pKM101 (IncN) and VirB1 from pTiA6 (IncRH1). These muramidase-like proteins are thought to be involved with local lysis of the peptidoglycan cell wall allowing other components of the Mpf system to stretch across the cell membrane (Dijkstra *et al*, 1996). Another role may be to facilitate the passage of the nucleoprotein complex through this layer (Bayer *et al*, 1995; Baron *et al*, 1997).

Figure 1.2.1.2.4 shows the complex regulatory circuits for both the Tra1 and Tra2 regions. The global regulatory protein TrbA as stated before represses promoters' *traJp*, *traKp* and *traGp* whilst also repressing *trbBp*, which is responsible for the expression of the entire Tra2 operon. The strong *trbBp* is subject to strict control being shut down by KorB, TrbA, KorA (lifting repression of *trbAp*) and competition from the *trfAp*. The *trbAp* is silenced by the strongly expressed *trfAp* with which it overlaps (Jagura-Burdzy *et al*, 1994). TrbA autoregulates *trbAp* while KorB also represses this promoter (Jagura-Burdzy *et al*, 1997). This complex regulatory circuit is thought to allow for expression of the Tra2 region during the initial stages of plasmid establishment and allows the plasmid to spread through a population after which a steady state is reached as TrbA and KorB accumulate and shut down expression.
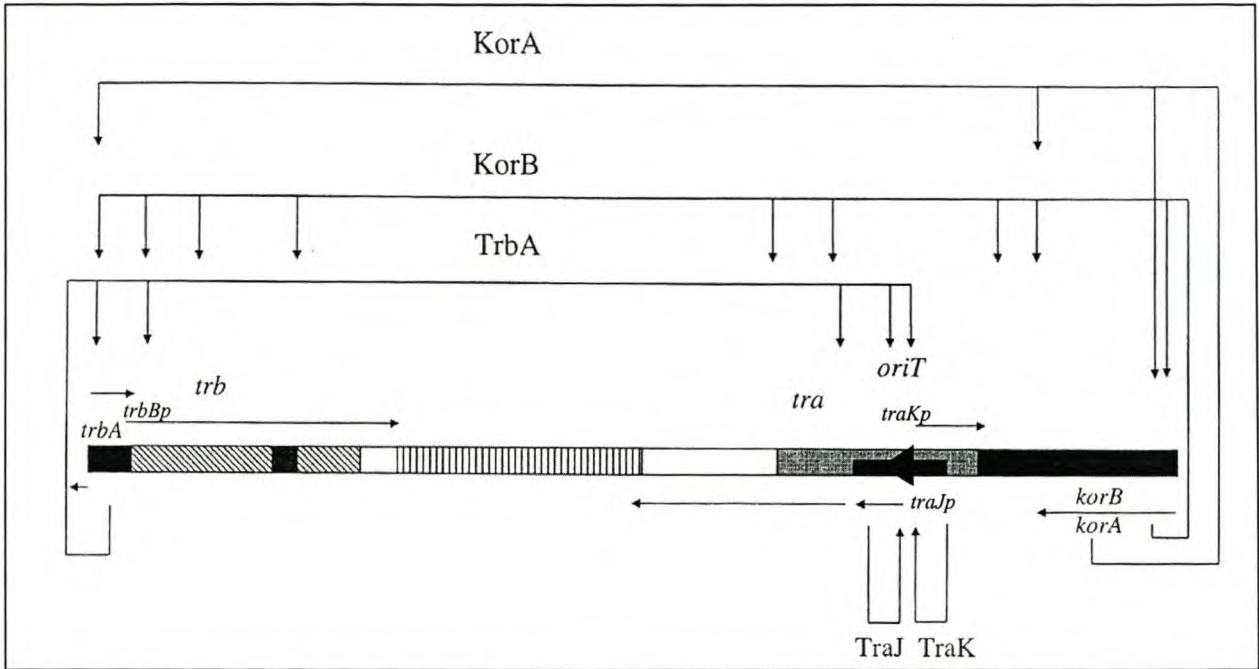
Figure 1.2.1.2.4: Schematic diagram showing the control of IncP transfer genes. Three global regulators supply the major control. KorB and TrbA repress expression of transfer genes directly. KorA is needed to derepress *trbA* , providing a way of shutting down transcription of the *tra* and *trb* genes once the plasmid is established. The assembly of the relaxosome is autogenously controlled by TraJ and TraK which bind to *oriT* and repress the promoters in this region. Regulatory genes, black; DNA processing genes, light grey; mating pair formation, diagonal hatching; white, non-essential or unknown function; vertical hatching, unrelated genes between the two regions (Zechner *et al*, 2000).

1.2.2 IncQ plasmids

The IncQ group/family of plasmids are characterized as being relatively small, and highly mobilizable by several conjugative plasmids (Frey and Bagdasarian *et al*, 1989). These features together with its broad host range make these plasmids highly promiscuous. They are particularly well mobilized by plasmids from the IncP1 (RP4 and R751) family, which affords them a very wide host range, capable of inter-kingdom transfer to plant and animal cells as well as a range of gram-positive and gram-negative baceteria (Meyer *et al*, 1982; Rawlings *et al*, 1984; Buchanan-Wollaston *et al*, 1987; Yoshida *et al*, 1997). This is due in part to the Mpf system coded for by the helper plasmid, which can establish contact with such diverse recipients, but also because these plasmids encode their own replication machinery.

This feature allows the plasmid be maintained in a host independent of some host encoded replication factors (Scherzinger *et al*, 1991; Yoshida *et al*, 1997).

The best-studied representatives of IncQ plasmids are RSF1010, R1162 and R300B (Barth *et al*, 1974; Guerry *et al*, 1974; Meyer *et al*, 1982). Although RSF1010 is considered to be the representative IncQ plasmid, R1162 has been described as identical or nearly identical to RSF1010 and the mobilization region/functions of R1162 have been extensively studied and will be discussed here.

1.2.2.1 DNA processing at the origin of transfer

Four groups have been identified among the IncQ plasmids based on plasmid incompatibility testing, the presence or absence of a plasmid addiction system, amino acid sequence comparisons of the *repA*, *repB* and *repC* genes and whether a three-gene or five-gene mobilization system is present (Rohrer *et al*, 1992; Tietze *et al*, 1998; Rawlings and Tietze *et al*, 2001).

The mobilization system of R1162 consists of three genes namely *mobA*, *mobB* and *mobC*. The genes are divergently transcribed from the origin of transfer, which consists of a 38bp minimal region and features an inverted repeat adjacent to the *nic* site (Brasch *et al*, 1987). The inverted repeat is thought to play an important role in termination of transfer, where it forms a hairpin structure once the plasmid DNA is in a single-stranded form (T-strand), but not in the determination of the cleavage site *in vivo* (Battacharjee *et al*, 1992; Scherzinger *et al*, 1993; Becker *et al*, 2000). The *oriT* appears to be processed in much the same way as for those of the IncP, IncF and related plasmid *oriT*'s: A site and strand specific nick is introduced and the DNA is transferred in a polar fashion from 5′ to 3′ then circularized once in the recipient (Kim *et al*, 1989; Meyer *et al*, 1989).

All three proteins make up the relaxosome at the *oriT*. Although neither the MobB nor MobC is absolutely required for plasmid mobilization they are needed for the efficient transfer of the plasmid. The MobA can cleave both ss- and dsDNA *oriT* substrates by itself, and appears to have a single large domain for contacting the *oriT* consisting of at least 184aa including the N-terminal (Becker *et al*, 2002). This

domain recognizes both the inner arm of the inverted repeat as well as the conserved bases between the inverted repeat and the *nic* site. This is unlike the TraI of the RP4 plasmid where TraJ is needed bind to the inverted repeat first before it can bind to the TraJ•*srj*. The C-terminal of MobA seems to be involved with recognition/binding to the MobB (Perwez *et al*, 1999). Tyrosine-25 has been shown to bind covalently to the 5′ end of the *nic* site and mobility shift assays suggest that only one molecule of MobA binds per relaxosome. In studies done to determine whether a rolling circle replication type mechanism is responsible for strand replacement synthesis on the donor copy of a plasmid involved in conjugal transfer, molecules were constructed that contained two directly repeated copies of R1162 *oriT*. One *oriT* would be mutated so that transfer strand replication could not be terminated (ter), while the other *oriT* would not allow initiation at the nick site (*nic*). After mating plasmid DNA was extracted from the transconjugants and the majority had three copies of *oriT* (Figure 1.2.2.1.1). Two of these were the original mutants while the third was a recombinant, fully functional *oriT*. If the 3′ end of *nic* had not been extended by an RC-type mechanism this construct would not be observed. This provided evidence that a RCR mechanism is involved in transfer strand replication but also that the MobA can terminate this process by a second cleavage mechanism (Erickson *et al*, 1993). If only one molecule of MobA is involved in relaxosome formation and thus cleavage of the *oriT*, there must be a second nucleophile to carry out the second cleavage reaction. Thus far no second nucleophile has been found as evidenced by the inability of MobA, with the active tyrosine replaced with phenylalanine, to demonstrate any cleavage ability (Zechner *et al*, 2000). It is suggested that as with the TraI of RP4 a second molecule of MobA is recruited for this function or that a second, cryptic nucleophile awaits discovery. The MobA can also ligate single-stranded *oriT* DNA cleaved at the *nic* site and this could be another mechanism for circularizing the plasmid in the recipient (Battacharjee *et al*, 1991; Battacharjee *et al*, 1992).
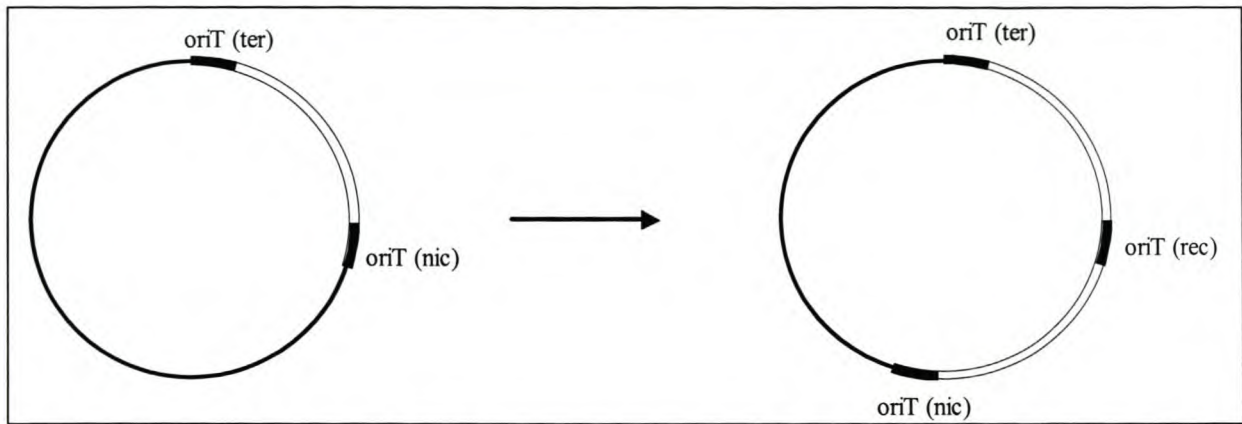
Figure 1.2.2.1.1: Strand transfer replication starts at the *oriT* (ter) and continues past the *oriT* (nic). Termination at this point would result in a molecule that cannot be replicated and would be broken down. When *oriT* (nic) is reached the second time, termination results in a molecule like the one displayed on the right side of the arrow. The *oriT* (rec) is generated through a conjugation dependent mechanism and is thought to be made up of the 5' end of *oriT* (ter) and the 3' end of *oriT* (nic) (Erickson et al, 1993).

MobC, a 94aa protein, is proposed to act as a molecular wedge assisting MobA in local strand separation of the AT-rich region next to the *nic* site and extending this separation to the nucleotide bases at the *nic* site helping to increase the mobilization frequency ± 50 fold (Figure 1.2.2.1.2) (Zhang *et al*, 1997). In the absence of MobC, nicking and strand separation become more sensitive to the level of gyrase present in the cell with the amount of nicked DNA decreasing with lower levels of gyrase present. MobA-strand separation could depend more on the helical distortion created by supercoiling when MobC is unavailable (Zhang *et al*, 1997). The proposed mechanism by which the MobB stabilizes the relaxosome is by first binding to MobA and then coupling with other *oriT*-bound relaxosomes. These stable relaxosomes do not allow for transcription from the promoters located within the *oriT* region and helps promote plasmid cleavage and conjugal transfer. The MobB is also thought to interact either directly with the transfer machinery coded for by the helper plasmid or indirectly by modifying the conformation of MobA to better recognize these components thus helping to target the relaxosome to the transfer apparatus (Perwez *et al*, 1996; Perwez *et al*, 1999).
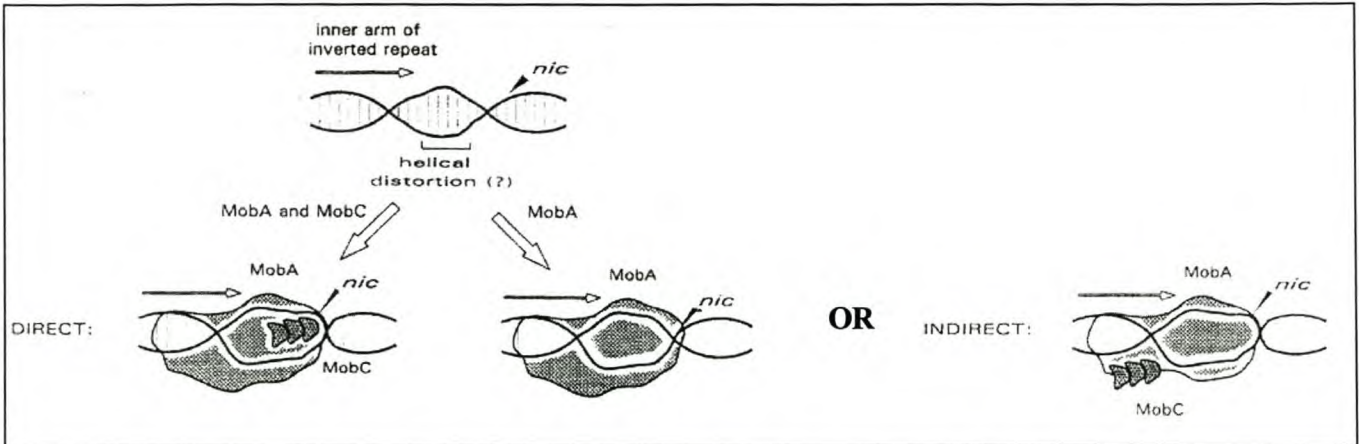
20

Figure 1.2.2.1.2: Strand separation of *oriT* DNA in the presence or absence of MobC. Many copies of MobC may congregate at the relaxosome, but only three are shown here for clarity (Zhang *et al*, 1997).

Because a DNA single strand is transferred, a priming system for its complement is required. An interesting feature found on all plasmids with IncQ and IncQ-like replicons examined so far, is the fusion of the relaxase to the plasmids primase. These can either be made as separate proteins or a fusion protein. It is logical to assume that this fusion happened only after acquisition of the mobilization genes and was selected because it offered/offers an improved mobilization frequency. This becomes important once the transfer frequency becomes low due to factors such as being mobilized by a conjugative plasmid (pOX38 a derivative of the F plasmid) not well suited to mobilization of, in this case, IncQ plasmids. For R1162 the long form (fusion protein MobA/RepB) of the primase has been shown to be the only replication protein active in conjugal transfer (Henderson *et al*, 1996; Henderson *et al*, 1999). This is however not always the case. For plasmid pTF-FC2, which has a five-gene mobilization system, with mob genes more closely related to those from the IncP1family of Dtr genes, a fusion protein is also present but the primase can be removed without affecting the plasmids ability to be mobilized efficiently (Rohrer and Rawlings, 1992; Rawlings and Tietze, 2001). Like the C-terminal helicase of plasmids F and R388 (see below), the fusion may help to increase the local concentration of the primase. Whether priming at the *oriV* takes place before (donor), during or after transfer (recipient) is not known but it is suggested that the fusion

protein is immobilized at the conjugal pore where it scans the T-strand for the priming site.

The mechanism of transfer for R1162 seems to involve the following steps: (i) The MobA recognizes and perhaps makes specific contact with the TAA nucleotides next to the *nic* site and causes local melting of the DNA helix at this A + T-rich region, (ii) MobC then helps extend the region of strand separation to the *nic* site giving a ssDNA molecule which is the substrate for the MobA, (iii) After proteolytic cleavage of the nic site the MobA becomes covalently bound to the 5′ end, (iv) The MobA/RepB possibly scans the transferred strand for the priming site and makes the primer used in replacement strand synthesis, as well as stuttering when the hairpin formed by the inverted repeat is reached allowing it to catalyze the second cleavage reaction and ligate the 5′ and 3′ ends, (v) This is followed by complementary strand synthesis on the T-strand in the recipient from a primer made at the *oriV* and rolling circle replication to replace the T-strand on the donor copy of the plasmid.

1.2.3 IncF-like plasmids

Lederberg and Tatum were the first to report conjugation (1946). The movement of chromosomal markers, observed by them, between *E. coli* strains was later linked to integration of the F plasmid into the host's chromosome. Bacteria containing F were sensitive to a number of phages such as M13. Grouping of other plasmids into the F-like plasmid group was originally done on the basis that these plasmids shared the same sensitivity to these so-called F specific phages. The F-like plasmid group has now been further subdivided into seven incompatibility groups based on plasmid incompatibility resulting from similarities within the replicons of these plasmids (Bergquist *et al*, 1987; Firth *et al*, 1996).

The F-like plasmids are found throughout the *Enterobacteriaceae* family and carry clinically significant determinants such as antibiotic resistance as well as for the production of hemolysin and toxins. These are relatively large, conjugative plasmids that have a narrow host range. They produce long, thick and flexible pili allowing them to conjugate in liquid culture. Plasmids F, R100 and R1 have received most attention with regard to studying the F-like conjugative transfer functions.

22

1.2.3.1 DNA processing at the origin of transfer

Unlike the transfer system of RP4 described earlier the F plasmid *tra* region is not broken up into separate areas (Tra1 and Tra2), but the genes for Dtr and Mpf are found together. It consists of 36 ORF's either known or predicted to encode products. Apart from *artA*, all translated genes are encoded on the same DNA strand, with a large operon predicted for the most of the *tra* region from *traY* to *traX* (Frost *et al*, 1994; Firth *et al*, 1996). The *oriT* is located next to the *traM* gene outside the *tra* region (Figure 1.2.3.1.1).

Twenty-seven of these ORF's have been identified as being essential for conjugation. Those that are not required include *artA*, *traR* and *trbA, B, E, H* and *trb* J. The products of *trbD, F*, and *trbG* have yet to be identified or assigned a function and may be involved in conjugation (Firth *et al*, 1996).

The proteins thought to be involved with F plasmid DNA metabolism are TraI, TraY, TraM and integration-host-factor (IHF). IHF histone-like protein is a heterodimer involved in processes such as replication, transcription and recombination and is coded for by the chromosomal *himA* and *hip* genes (Drlica and Rouvière-Yaniv *et al*, 1987). DNaseI foot-printing analysis has identified three sites within the F *oriT* that are protected by purified TraM (Figure 1.2.3.1.2). Although binding of F TraM to the *oriT* does not stimulate cleavage of the *nic* site *in vitro*, recent evidence shows that *in vivo* the TraM from R1 promotes nicking at *oriT* (Everett *et al*, 1980; Kupelweiser *et al*, 1998; Verdino *et al*, 1999). F TraM is required for DNA transfer but not for pilus formation (Achtman *et al*, 1972; Everett *et al*, 1980). This contrasts with the role of TraM in plasmid R1 where it forms part of a complex regulatory circuit, which controls expression from the *traJp* that in turn switches on transcrption from *traYp* (the first gene in the large *tra* operon). Thus R1 TraM controls expression of both genes involved with DNA metabolism and Mpf (Pölzleitner *et al*, 1997).

23
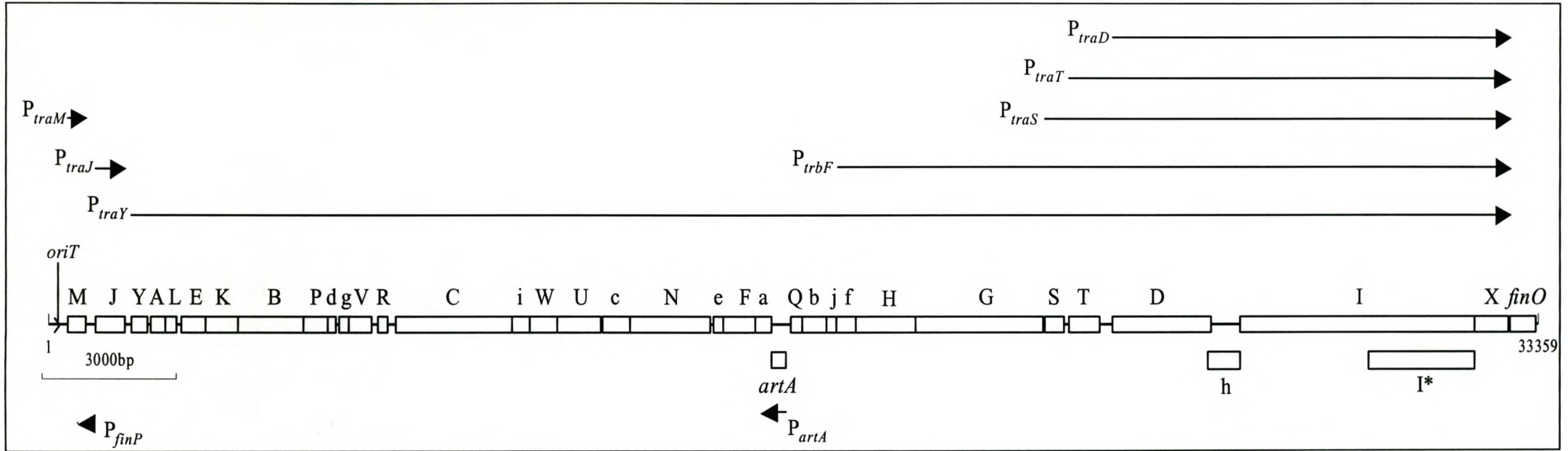
Figure 1.2.3.1.1: Physical and genetic map of the F-plasmid transfer region. Arrowed lines denote transcripts initiating at the indicated promoters. The extent of indicated genes are represented by boxes. Capital and lowercase letters are used to label *tra* and *trb* genes, respectively. An arrowhead indicates the position of the origin of transfer (*oriT*) and the direction of conjugal transfer. The map is adapted from Firth *et al* 1996.

TraM is highly plasmid specific binding only to the sites within its cognate *oriT* as is demonstrated by studies involving the mobilization of chimeric *oriT* DNA by heterologous transfer systems. R100-1 TraM could not complement a *traM* mutation in the F plasmid (pOX38-*traM*K3), while purified F TraM displayed a low affinity for R100-1 TraM binding sites using mobility shift assays (Fekete *et al*, 2000). Although TraM does not form part of the F plasmid relaxosome one of the roles proposed for this protein is that its binding to the *oriT* forms a nucleosome-like complex, which could affect the local superhelical density promoting cleavage at the *nic* site (Howard *et al*, 1995). F TraM has been shown to interact with TraD (coupling protein) *in vitro* and this suggests that TraM, via TraD, may couple the relaxosome to the bacterial membrane or to the Mpf complex (Disque-Kochem *et al*, 1997). Evidence for this is the isolation of both TraM and TraI from the inner membrane fraction in the presence of TraD *in vivo* (DiLaurenzio *et al*, 1992; Panicker *et al*, 1992). In light of the proposed nucleosome complex formed by TraM, it would appear that it plays a similar role to that of TraK in RP4 although here it also acts as the interface for the F relaxosome with the transfer machinery where TraI fulfills this role in RP4.

The *traI* gene codes for the F plasmid relaxase which, together with TrwC from R388 and TraHI from pKM101, as a group, was shown to be unique among the conjugative DNA relaxases in that they have two active-site tyrosines. At least for TraI and TrwC it has been shown that both tyrosines take part in the cleavage reaction. The reaction(s) possibly catalyzed by these relaxases are described in more detail in section 1.5.1. The F TraI recognizes and binds to a G + T-rich region on the *traM* side of the *nic* site (Figure 1.2.3.1.1). Here it catalyzes the reversible transesterification reaction that cleaves the *nic* site and becomes covalently bound to the 5' end of the cleaved DNA (Matson *et al*, 1993; Sherman *et al*, 1994). F TraI can efficiently cleave a single stranded *oriT* DNA substrate on its own, as well as a superhelical substrate although with reduced efficiency (negative supercoiling can drive transient melting of duplex DNA). Binding to the *oriT* by both TraY and IHF are required for TraI to bind and cleave open circular or linear dsDNA *in vitro* (Inamoto *et al*, 1994; Howard *et al*, 1995). The need to add TraY and IHF first to *oriT* sequence to observe maximal relaxosome formation, suggests that these proteins bind to *oriT* before TraI does and that their role is to prepare the DNA for TraI binding.

Like the F TraM, TraI and TraY are plasmid specific and bind only to their cognate *oriT* recognition sequences (Fekete *et al*, 2000). Binding of TraY causes the DNA to bend ~50º, while IHF binding to a site (*ihfA* that is intrinsically bent) between *nic* and *sbyC* induces further bending of ~140º (Tsai *et al*, 1990; Luo *et al*, 1994; Rice *et al*, 1996; Byrd *et al*, 1997). As with other relaxosomes, protein-protein interactions are thought to play an important role in formation of the F plasmid relaxosome. The bending introduced by TraY and IHF is suggested to: (i) Bring TraY and TraI into closer contact to enable them to interact, and (ii) To present the TraI binding site (*ibs*) region as ssDNA that is more easily bound and cut by TraI (Howard *et al*, 1995; Nelson *et al*, 1995; Zechner *et al*, 2000).
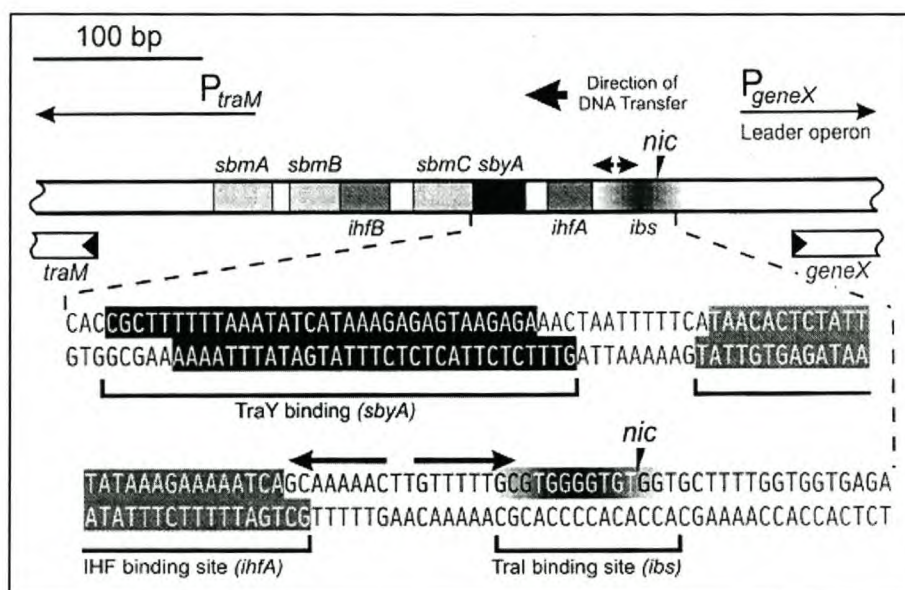

Figure 1.2.3.1.2: Structural organization of the F-plasmid transfer origin. Divergent arrows adjacent to the *nic* site mark inverted repeat sequences. Shaded or black regions have been identified as binding areas for the respective proteins. On the expanded sequence shown at the bottom shaded or black areas represent those nucleotides that are protected during Dnase I treatment. For other symbols see Figure 1.2.1.1.3 (Zechner *et al*, 2000).

Although all conjugative DNA relaxases belong to one family of Mob proteins, as opposed to those involved in rolling circle replication, the relaxases from IncF and IncW plasmids belong to a separate subfamily compared to those of the IncP plasmids (Ilyina *et al*, 1992; Byrd *et al*, 1997). In the case of R388, TrwC and F TraI the C-terminal acts as an ATP-dependant 5' - 3' DNA helicase. The helicase domain of TrwC and F TraI has been shown to be essential for conjugative transfer with no

chromosomally encoded proteins being able to complement the helicase mutations (Llosa *et al*, 1996; Byrd *et al*, 1997; Matson *et al*, 2001). Although independent expression of the N-terminal relaxase and C-terminal helicase domains allow conjugal DNA transfer, the frequency is several orders of magnitude lower than that of the WT. Two reasons for the development of a fused protein have been put forward: (i) Linkage of the N-terminal relaxase domain and the C-terminal helicase may increase the local concentration of the helicase needed for strand separation (ii) In addition to the former the helicase could also provide the motive force for DNA transport across the cell membrane (Abdel-Monem *et al*, 1983; Grandoso *et al*, 1994; Llosa *et al*, 1996).

1.2.3.2 Mating pair formation

Some of the F-pilus assembly machine components share homology with the *vir* system from the *A. tumefaciens* Ti plasmids and therefore are part of the type IV secretion system (IVSS) family, section 1.3 (Christie *et al*, 2001).

Three proteins have been shown to interact to form the mature F-pilin subunit. TraA is the 121 amino acid prepilin protein that consists of two regions: a 51 amino acid leader peptide at its N-terminal and a 70 amino acid C-terminal, which constitutes the mature pilin (Minkley *et al*, 1976; Majdalani *et al*, 1996). Moore and coworkers first described the accumulation of F-pilin in the inner membrane, proposed to act as a reservoir to be "drained" during pilus assembly and "refilled" during pilus retraction (Moore *et al*, 1981; Sowa *et al*, 1983). The mature F-pilin can be divided into four structural domains. Domain IV (the C-terminal 20 amino acids) can insert itself into the inner membrane independent of other Tra proteins (Harris *et al*, 1999). TraQ is proposed to be a chaperone for TraA entry into the inner membrane. This was confirmed by using a yeast two-hybrid system to show a transient interaction between TraA and TraQ (Wu and Ippen-Ihler, 1989; Harris *et al*, 1999). The cytoplasmic C-terminal of TraQ is thought to interact with the TraA through its domain IV. This interaction is thought to slow-down entry of TraA into the membrane allowing correct folding of the protein (Figure 1.2.3.2.1). The TraA leader peptide is cleaved by host encoded leader peptidase B and possibly anchored in the cytoplasmic membrane via interaction with membrane phospholipids (Majdalani *et al*, 1996). From the yeast

two-hybrid analysis it would appear if no other proteins are involved with the efficient targeting of TraA to the inner membrane, apart from TraQ. However this system is not as effective as F in restoring membrane F-pilin levels to that of the wild type (WT) (Paiva *et al*, 1996).
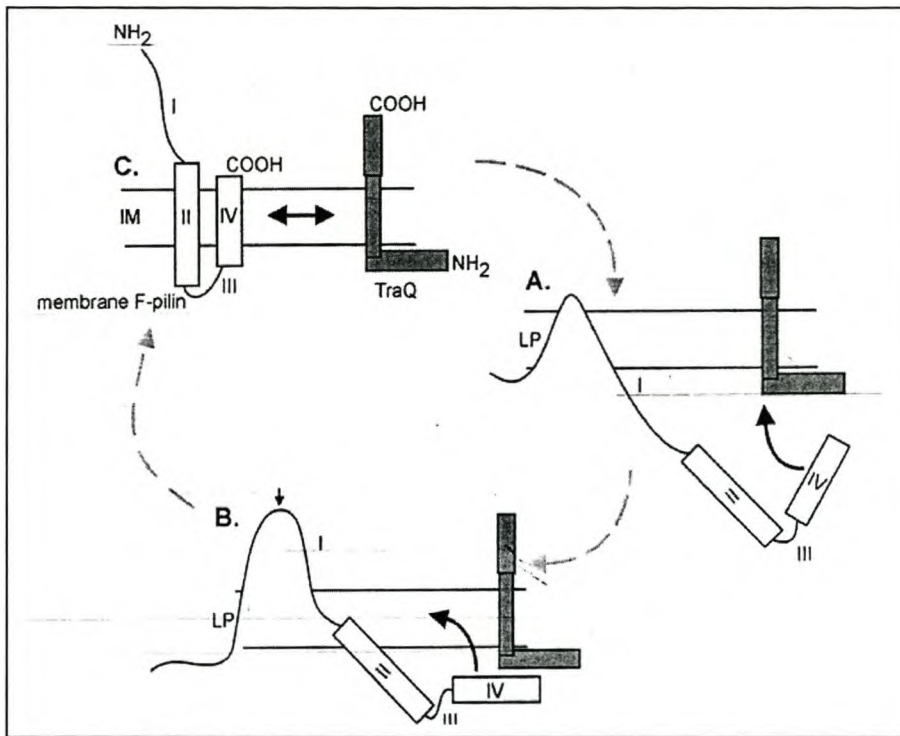
Figure 1.2.3.2.1: The TraQ cycle for accumulation of inner membrane F-pilin. TraQ (shaded) is depicted as a bitopic inner membrane protein with its N-terminal (horizontal segment) cytoplasmic. Interaction with TraA domain IV (panel A) is required for proper membrane insertion of preceding TraA segments. This interaction is transient, with domain IV itself partitioning into the membrane (panel B), thereby releasing TraQ for another cycle. Cleavage between TraA leader peptide and domain I (panel B, downward arrowhead) completes the formation of membrane F-pilin (panel C) (Paiva *et al*, 1992).

TraX is required for N^{α}-acetylation of F-pilin polypeptide, however lack of acetylation in TraX mutants does not result in a transfer deficient- or phage resistant phenotype (Moore *et al*, 1993). In the absence of TraQ, TraA is quickly degraded (Maneewannakul *et al*, 1993). Co-expression of *traX* and *traA* seemed to stabilize TraA and prevent rapid breakdown. As TraX was shown not to associate with TraA via protein-protein interactions, stabilizing the F-pilin in such a manner therefore seems unlikely. Thus the stabilizing effect could most probably be mediated by N-

terminal acetylation of TraA (Maneewannakul *et al*, 1993; Harris *et al*, 1999). A small percentage of pilin subunits appear to be modified and as such migrates more slowly on SDS-page gels. The purpose of having irregular pilin subunits inserted in the pilus is proposed to increase the flexibility (Maneewannakul *et al*, 1993). The nature of the modification is still unknown.
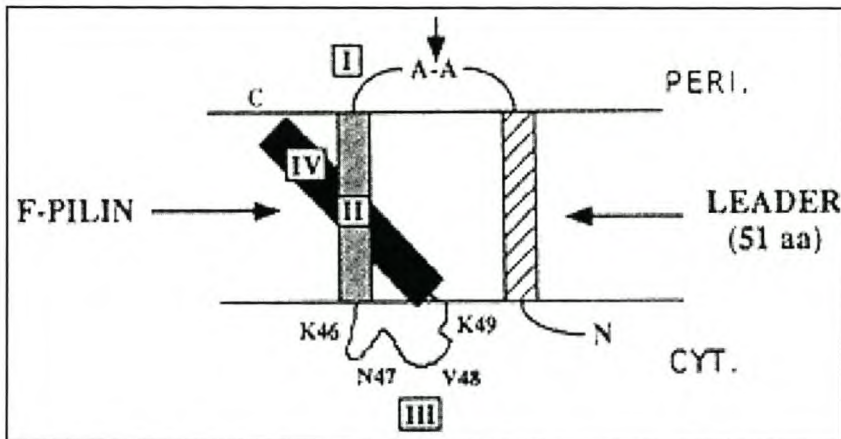
Figure 1.2.3.2.2: Relationship between TraA processing and the structure proposed for F membrane pilin. The TraA leader extends to the vicinal alanines cleaved proteolytically during the formation of F-pilin (arrowhead). The four structural domains are indicated (I-IV) (Paiva *et al*, 1992).

The F-pilin pool seems to localize to specific areas in the periplasm. Figure 1.2.3.2.3 shows a proposed structure for the assembled F-pilus, while Figure 1.2.3.2.2 shows its topology in the cytoplasmic membrane (Paiva *et al*, 1992). While domain I is exposed to the periplasm in membrane pilin it does not appear to be exposed in the F-pilin filament. Monoclonal antibodies to domain I do not attach to the length of the F-pilus but do react with the membrane form as well as the pilus tip. Therefore it seems as though the N-terminal of the F-pilin is tucked into the F-pilus structure (Grossman *et al*, 1990; Paiva *et al*, 1992). Residues at or near the C-terminus as well as amino acids 12-22 of mature F-pilin are important for RNA phage binding and are therefore suggested to be exposed on the outside of the pilus (Frost *et al*, 1988). Four hydrophilic residues (46-49) form the so-called KNVK turn comprising domain III, which is exposed to the cytoplasm (membrane pilin) or to the lumen of the assembled pilus (Paiva *et al*, 1992).
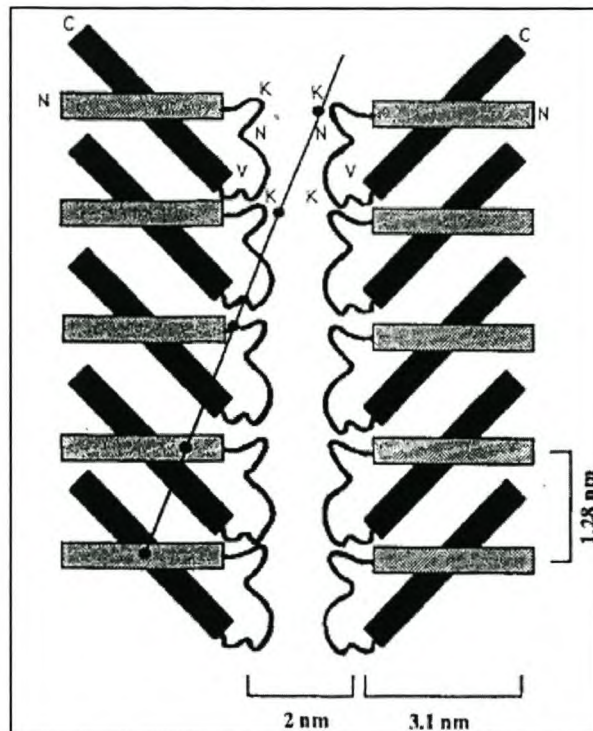
Figure 1.2.3.2.3: Proposed structure of filament F-pilin. The rotation of successive F-pilin layers about the helix is not shown (28.8° with respect to the disc below). The diagonal line indicates the surface lattice expected from that rotation. Black segments mark domain IV while gray segments indicate domain II. The KNVK turn (domain III) is shown as well (Paiva *et al*, 1992).

Several mutants from the *tra* region that allow accumulation of pilin in the inner membrane but that fail to produce functional pili have been identified. Although the sub-cellular location of most of these gene products are known, very little is known about the specific roles each plays. For TraB, TraE, TraG, TraL and TrbI hydropathy analysis suggests an inner membrane location, while TraW, TraU, TraF, TraK, TraH and TraC are thought to be located in the periplasm (Frost *et al*, 1994). TraV was shown to be an outer membrane lipoprotein, which despite a lack of overall homology was thought to act like VirB7, from the *vir* system in *A. tumefaciens*, in forming intramolecular disulphide bonds with other Mpf components (refer to section 1.3) (Doran *et al*, 1994). The internal cysteine residues important in VirB7 for the formation of a functional Mpf complex were found not to be important for pilus formation in TraV but rather that these amino acids are important for the targeting of TraV to the outer membrane (Harris *et al*, 2002). TraV, TraK and TraB were found,

by yeast two-hybrid analysis, to form a protein interaction group. Although TraK and TraB in the absence of a functional TraV fractionate with the periplasm and outer membrane respectively, the formation of the TraV/B/K complex results in all these proteins being located in the outer membrane (Firth *et al*, 1996; Harris *et al*, 2002). Thus TraV appears to anchor the complex in the outer membrane.

TraG is not only required for pilus synthesis but is also involved in formation of stable mating pairs. Only the N-terminal portion of TraG is required for pilus biogenesis and is thought to include a large periplasmic domain that is anchored in the cytoplasmic membrane (Firth *et al*, 1996). The actual assembly process is proposed to be energy dependent, which is evidenced by the retraction of pili when cells are exposed to respiratory poisons such as cyanide and arsenate (Novotny *et al*, 1974). The energy for this process could be from ATP hydrolysis by the VirB4 homologue TraC in which Walker A nucleotide binding motifs have been identified (Frost *et al*, 1994). Another candidate is TraH that also features these ATP/GTP binding site motifs. The pilin subunits are assembled into discs of five subunits each and are stacked such that each disc is rotated by 28.8° with respect to the disc below. The prepilin can be seen as a substrate for a protein secretion system that is actively pumped across the bacterial cell envelope with the ability to be taken up by the bacterium again (Firth *et al*, 1996).

The first step in formation of stable mating pairs (shear resistant) was thought to be the recognition of outer membrane protein A (OmpA) and constituents of the lipopolysaccharide layer (LPS) by the pilus tip (Anthony *et al*, 1994). This is followed by pilus retraction and eventually formation of stable wall-wall contacts. Recent evidence has indicated that it is the *traN* product that is responsible for interaction with OmpA and LPS and not the pilus tip (Klimke *et al*, 1998). F TraN mutants showed severely reduced transfer frequencies when mated with *ompA* mutant recipients with a functional copy of TraN being expressed *in trans*. However when mated with WT recipients the transfer frequencies were restored. R100-1 TraN mutants did not display the same phenotype with high transfer frequencies to both *ompA* mutant- and WT recipients. The same result is observed when using *rfa* (LPS formation) mutant recipients. Therefore although TraN of F seems to recognize and

bind OmpA and/or LPS the same is not true for R100-1. The N- and C-termini of F and R100-1 TraN shows a high degree of similarity while the central region does not and may represent an area of allelic-specific receptor interactions (Klimke *et al*, 1998).

TraN also interacts with TraG. As described earlier only the N-terminal portion of TraG is required for pilus formation. The C-terminal region is required for mating pair stabilization (MPS) and consists of a large periplasmic domain anchored in the inner membrane that has been suggested to function in stabilization independent of the N-terminus (Firth *et al*, 1992). The exact receptor for the pilus tip is still unknown. The phenotypes of TraN and TraG mutants as well as membrane location suggests an interaction in the periplasm that results in the formation of stable mating pairs (Manning et al, 1981)

1.3 Type IV secretion systems

Mating pair formation systems and other macromolecular transfer systems that share a common ancestry are referred to as the type IV secretion system (IVSS) thus distinguishing the conjugation-related systems from other bacterial secretion pathways (Salmond *et al*, 1994). The Mpf systems are unique in that they have adapted to not only transfer proteins but nucleoprotein complexes: a prerequisite for conjugal DNA transfer. Representatives of systems dedicated to conjugal DNA transfer include the Tra2 region of plasmid RP4, the Tra system of F, the IncN plasmid pKM101 as well as the *pil* genes from R388, R64 and the *magB* genes from plasmid pVT745 (Llosa *et al*, 1994; Pohlman *et al*, 1994; Yoshida *et al*, 1999; Galli *et al*, 2001; Sakai *et al*, 2002). Many of the macromolecular secretion systems that are not involved in conjugation play a role in virulence through secretion of effector proteins or toxins and have been identified as critical to survival of several intracellular pathogens in their respective hosts. Examples of these include the type IVSS responsible for intracellular survival and multiplication of *Brucella melitensis* biovar *suis* (O'Callaghan *et al*, 1999 MM 33: 1210-1220), secretion of pertussis toxin from *Bordetella pertussis* (Covacci *et al*, 1993), *H. pylori*'s *cag* PAI (Covacci *et al*, 1999), *Brucella abortus* (Sieira *et al*, 2000) and the *dot/icm* system of *Legionella pneumophila* (Vogel *et al*, 1998).
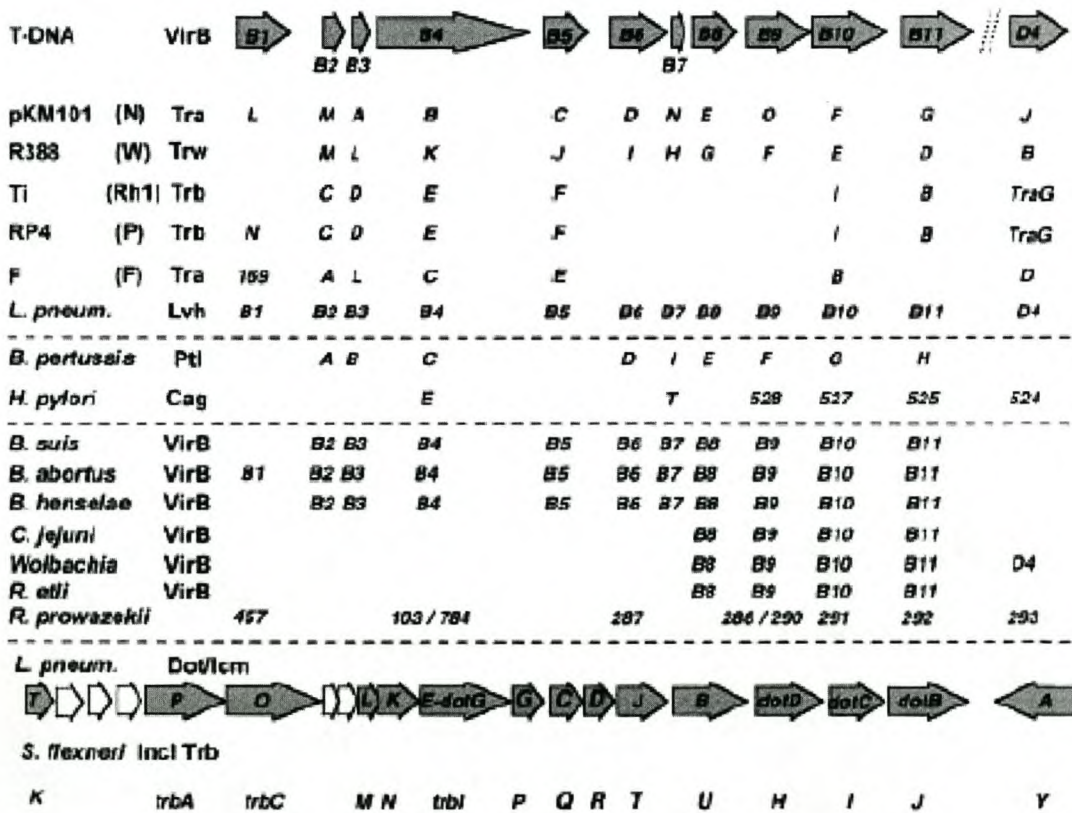
The *vir* system, which encodes a type IVSS, found on the IncRH1 plasmids in *A. tumefaciens* mediates transfer of a discrete portion of the plasmid (T-DNA) to susceptible plants. Here the T-DNA is integrated into the host chromosome and expression of genes located on this segment ultimately leads to tumorous growths caused by uncontrolled cell proliferation. The main Dtr functions are coded for by the *virC* and *virD* operons located in the *vir* regulon. On the plasmid the T-DNA is flanked on either side by 24bp imperfect direct repeats or so-called border sequences. These are closely related, at a nucleotide sequence level, to the *nic* site within the *oriT* of plasmid RP4 (Pansegrau *et al*, 1991). The T-DNA is removed from the plasmid by nicking of the border regions by VirD2, a TraI homologue. The VirD2 molecule then becomes covalently attached to the 5' end of the single stranded T-DNA and this complex is subsequently transferred by the Mpf system, coded for by the *virB* genes, to the recipient (Dürrenberger *et al*, 1989). Not only do the *vir* genes display a high degree of amino acid similarity with the *tra* region proteins from pKM101 (IncN), but the mechanism by which the T-DNA is transferred to the plant cell, is strikingly similar to conjugal plasmid transfer (Zambrynski *et al*, 1992). A second system found on IncRH1 plasmids is responsible for conjugative transfer of the plasmid itself. The *tra* system consists of a section, *trb*, which codes for the Mpf transfer machinery and a second region also termed *tra* that codes for the Dtr functions. The *trb* genes are homologous to the Tra2 region genes from plasmid RP4 while the *oriT* region and the relaxase, TraA, are more closely related to those found in RSF1010 (Li *et al*, 1998; Lessl *et al*, 1992b). A TraG and TraF, which are homologues of those found in RP4, are also part of the *tra* region (Farrand *et al*, 1996).

The *virB* encoded Mpf system is well studied and has become the reference point for type IVSS. No analogues of *virB6 – virB9* are found in the Tra2 region of RP4 and their role in the type IVSS of *A. tumefaciens* will therefore be discussed (Table 1.3.1). It appears that VirB8 is involved with the initial steps in Mpf complex formation as it localizes to a few unique sites on the bacterial membrane. VirB9 and VirB10 were recently shown to form a protein complex with VirB8, and their cellular location thus depended on where VirB8 was located (Kumar *et al*, 2000; Kumar *et al*, 2001). The VirB7 lipoprotein forms a homodimer as well as a heterodimer with VirB9 that is

proposed to form a nucleation center or scaffold around which the transfer complex is assembled. The VirB7-VirB9 heterodimer subunits are linked with disulfide bridges between Cys-24 of VirB7 and Cys-262 of VirB9 with the help of disulfide bond formation proteins (Bardwell *et al*, 1994; Anderson *et al*, 1996; Spudich *et al*, 1996). Although no physical contact could be detected between the VirB9 and VirB10 proteins, the VirB7-VirB9 complex is needed for the accumulation of VirB10 to WT levels. These interactions are thought to stabilize or position VirB10 that allows it to interact appropriately with either itself or other proteins to possibly form a complex that spans both bacterial membranes. The VirB9 stabilizing effect on VirB10 could be indirect via another protein, which stabilizes VirB10, or directly through transient interaction with VirB10 to stabilize/position it long enough to form VirB10 multimers observed after chemical cross-linking (Beupre *et al*, 1997). The function of the VirB5 protein is not known, however its co-purification with T-pili from *A. tumefaciens* strains carrying octapine and nopaline plasmids suggest several possible roles. VirB5 may form part of a complex at the pilus base or VirB5 could be a minor component of the pilus localized at the tip and, which mediates adhesion to recipient cells. VirB5 will only accumulate to WT levels in the presence of VirB6. This stabilizing effect is likely through binding of VirB6 to a periplasmic protein, which in turn stabilizes VirB5 or binds to a protein responsible for VirB5 breakdown. The absence of VirB6 only affects the levels of Vir proteins involved with T-pilus assembly (VirB3, VirB5 and the VirB7 homodimer). This protein is now thought to be a key regulator of T-pilus assembly through its stabilizing interaction with VirB5 or to serve as an inner membrane channel for the transfer of T-pilus components to the periplasm (Hapfelmeier *et al*, 2000). The VirB7 homodimer, which was proposed to be an intermediate in VirB7-VirB9 heterodimer formation, localizes extra cellularly (Berger *et al*, 1994; Fernandez *et al*, 1996; Sagulenko *et al*, 2001). In VirB6 deletion mutants the level of VirB7 monomer and VirB7-VirB9 heterodimer were unaffected whereas the VirB7 homodimer was not detected. The VirB7 homodimer also co-fractionates with VirB5 and VirB2 and associates tightly with the pilus. These results together imply that the homodimer does not serve as an intermediate in VirB7-VirB9 heterodimer formation but that it forms a physiologically relevant complex with the T-pilus. The proposed function of this homodimer is that it might assemble at the outer membrane as an oligomeric ring to facilitate substrate translocation or play a role in pilus assembly (Sagulenko *et al*, 2001). The heterodimer is suggested to help

anchor the pilus to the bacterial membrane but could also form a structure similar to the one proposed for the homodimer (Zhu *et al*, 2000).

Table 1.3.1: The type IVSS known or postulated to translocate macromolecular substrates intercellularly. The *A. tumefaciens virB* gene products shown across the top assemble as the T-DNA transfer system. The underlying examples of type IVSS are composed of homologues of some of the VirB proteins. The top group corresponds to systems shown to transfer DNA between bacteria; the *B. pertussis* and *H. pylori* systems deliver known substrates (PT and CagA respectively) to mammalian cells, and the third group corresponds to systems whose substrates are presently unknown but are postulated to be effector proteins. The *L. pneumophila dot/icm* gene products shown across the bottom are homologues of the *Shigella flexneri* CollIb-P9 (IncI) transfer proteins. This system can conjugally transfer DNA, but its proposed role in virulence is to export effector proteins (Christie *et al*, 2001).

T-DNA	VirB	B1	B2 B3	B4	B5	B6 B7	B8	B9	B10	B11	D4
pKM101 (N)	Tra	L	M A	B	C	D N E	O	F		G	J
R388 (W)	Trw		M L	K	J	I H G	F	E		D	B
Ti (Rh1)	Trb		C D	E	F			I		B	TraG
RP4 (P)	Trb	N	C D	E	F			I		B	TraG
F (F)	Tra	169	A L	C	E				B		D
L. pneum.	Lvh	B1	B2 B3	B4	B5	D6 D7 D8	D9	D10		D11	D4
B. pertussis	Ptl		A B	C		D I E	F	G		H	
H. pylori	Cag			E		T	528	527		525	524
B. suis	VirB		B2 B3	B4	B5	B6 B7 B8	B9	B10	B11		
B. abortus	VirB	B1	B2 B3	B4	B5	B6 B7 B8	B9	B10	B11		
B. henselae	VirB		B2 B3	B4	B5	B6 B7 B8	B9	B10	B11		
C. jejuni	VirB						B8	B9	B10	B11		
Wolbachia	VirB						B8	B9	B10	B11	D4
R. etli	VirB						B8	B9	B10	B11	
R. prowazekii		457	103 / 784		287		286 / 290	291	292		293

L. pneum. Dot/Icm

S. flexneri IncI Trb

| K | trbA | trbC | M N | trbI | P | Q R T | U | H | I | J | | V |

The pertussis toxin (PT), from the gram-negative organism *B. pertussis*, belongs to the A/B family of toxins and is composed of 5 subunit types S_1-S_5 found in a 1:1:1:2:1 ratio (Tamura *et al*, 1982). The toxin interrupts signal transduction in mammalian cells by ADP-ribosylating the α-subunits of GTP-binding regulatory proteins (Katada *et al*, 1982; Gilman *et al*, 1987). Each toxin subunit is synthesized with a signal sequence and is transferred across the cytoplasmic membrane by the GSP (Type II, *sec*-dependent) to the periplasm. The holotoxin is then assembled and exported across the outer membrane by a type IVSS (Farizo *et al*, 2000). Although *ptl*-encoded pilus formation has not yet been visualized, the pilin homologue in *B. pertussis* PtlA is needed for efficient pertussis toxin secretion (Craig-Mylius *et al*, 1999). When compared with the *virB* system, the PT export system lacks *virB1* and *virB5* homologues and only transports substrates across the outer membrane. It also lacks a TraG-like coupling protein (section 1.4).

One of the questions raised here is: Why, if the PT secretion system has all the components needed for a transmembrane export structure, is the transport across the inner membrane mediated by the GSP (Table 1.3.1)? In answer to this question, the GSP seems to replace the role that a TraG-like coupling protein together with the inner membrane section of the transmembrane complex would have played. Without the coupling protein the substrate can perhaps not gain access or be processed by the inner membrane segment. Reasons as to why these seemingly vestigial components are present might be that the PT export system represents a protein transfer system being tuned for DNA transfer or *vice versa*, or it could be that they are needed for the structural stability of the transmembrane complex. In light of studies done to determine the functional role of these components it would seem that they are not exclusively required for the latter. Secondly, which component(s) would allow a protein secretion system to transport a nucleoprotein complex? DNA transfer systems are capable of translocating protein independent of DNA or conjugation (VirE2 and VirF in T-DNA transfer and TraC1 in RP4). For this reason it has been postulated that DNA transfer by these "dedicated protein export systems" is accidental (Christie *et al*, 1997; Vergunst *et al*, 2000; Wilkins *et al*, 2000). Evidence for this is the indiscriminate transfer of DNA by the T-DNA transfer system as long as the recognition site for VirD2 is present (Miranda *et al*, 1992). Although the T-strand is

only bound at the 5′-end by a relaxase protein the naked trailing DNA strand would have to pass through the export pathway, which could suggest some form of interaction between the transport components and the DNA (VirE2 single stranded binding protein only coats the T-strand once inside the recipient). The PT export system also crucially lacks a TraG-like coupling protein and cannot transfer DNA. Although not conclusively proven, there is some evidence for horizontal gene transfer in *H. pylori* via a DNase-resistant, unidirectional mechanism, which requires cell-to-cell contact (Kuipers *et al*, 1998; Ando *et al*, 2000). The *H. pylori cag* system also lacks *virB1*, *virB2*, *virB3*, *virB5* and *virB6* homologues, but does code for a coupling protein. These coupling proteins may have co-evolved with the relaxosome components in an effort to become independent of the host encoded secretion pathways but also as an essential component for feeding the nucleoprotein complex into the Mpf system as the relaxosome components may not have evolved from secreted proteins (section 1.4).

It is now evident from examples such as the amino acid sequence similarity and gene arrangement of plasmid pVT745, as well as the phylogenetic analysis of type IVSS, that a protein export system was the progenitor of these systems (Christie *et al*, 2001; Galli *et al*, 2001). These results also suggest that they evolved from a single precursor system without exchange of genes between divergent systems. Although when compared to the *virB* region many type IVSS seem to have lost essential Mpf components (VirB6-VirB9). However other proteins whose function have not been elucidated but are essential for pilus formation and conjugal transfer are still present in these systems (TrbG, TrbJ and TrbL of RP4 and TraK and TraV from the F plasmid). The question of whether DNA transfer across the bacterial cell membrane is a one or two-step process has not been answered. A one step process is supported by the idea that the VirD2 and VirE2, which can be transported across the membrane by the type IVSS independent of DNA transfer, do not have characteristic N-terminal signal sequences and therefore need a dedicated transport machine to recognize their transport signals. The cumulative evidence for the structure of the *virB* encoded protein complex suggests a transmembrane channel that could allow the one-step transfer of the nucleoprotein complex across both cell envelopes. A two-step mechanism has also been proposed where the T-strand is first taken across the inner

membrane to the periplasm after which it is exported. This is analogous to the PT export pathway and warrants consideration as no coupling between TraG-like proteins and the Mpf complex has been shown and DNA transfer through the transmembrane complex has not been proven (Llosa *et al*, 2002).

The conservation of a VirB10 homologue in many systems hints at a central role in the formation of the transmembrane structure. The ATPases, VirB4 and VirB11, seem to be as conserved underscoring their importance in providing the energy for either pilus assembly, assembly of the transmembrane complex or energizing the DNA transfer reactions. The *virB8*, *virB9* and *virB10* gene cluster also seems to be conserved in many Mpf systems as well as in protein export systems, which may suggest a structure that functions in a unified manner, possibly comprising the transmembrane conduit.

1.4 TraG-like coupling proteins

Two sets of functions have been defined for plasmid conjugal transfer thus far: Dtr and Mpf. A third function, the TraG-like (named from RP4) coupling proteins, is responsible for establishing a connection between the relaxosome and the Mpf complex. Complementation and mutational analysis indicate that TraG is not involved with relaxosome formation nor is it needed for establishing cell-to-cell contact or for phage growth (Fürste *et al*, 1989; Pansegrau *et al*, 1990; Waters *et al*, 1992). It is, however, absolutely required for RP4 self-transfer and mobilization of RSF1010 by RP4 (Lessl *et al*, 1993). A coupling role was therefore suggested. Table 1.3.1 shows some of the TraG-like proteins that have been studied. Others include TrsK from the gram-positive plasmid pGO1 and the MobB from plasmid CloDF13. The latter is of interest as it is the only known mobilizable plasmid to specify it's own TraG-like protein (Cabezon *et al*, 1997). TraG-like proteins are bound to the inner membrane. They have two conserved motifs, which specify a nucleotide-binding domain (NBD) and mutational analysis indicates that both are essential for transfer. These motifs are similar to the Walker A and B nucleotide binding motifs of NTPases and ABC transporters (Walker *et al*, 1982; Schneider *et al*, 1998; Panicker *et al*, 1992).

The crystal structure for TrwB lacking its N-terminal transmembrane segment, called TrwBΔN70, has recently been solved (Gomis-Rüth *et al*, 2002). The wild type TrwB behaves like both a monomer and a hexamer during purification, whereas TrwBΔN70 behaves only as a monomer throughout this process. However the TrwBΔN70 crystal structure unveiled a molecule with six equivalent protein units. This seemed to be caused by the precipitation conditions during purification rather than occurring naturally in solution. The dissociation constants for the monomer and hexamer forms of WT TrwB for NTP's do not differ enough to suggest an active/inactive half binding-site on the monomer. The lack of the N-terminal (TrwBΔN70) shows that this transmembrane domain appears to play an important role in the architecture of TrwB and suggests a dynamic equilibrium between the monomeric and hexameric forms. Although TrwBΔN70 (see below) binds to DNA non-specifically and no sequence specific interaction with the *oriT* has been recognized, its crystal structure revealed a TraM-like (F plasmid) NBD. The TrwB is thought to have incorporated a TraM-like NBD into its structure (Gomis-Rüth *et al*, 2002). It is postulated that these proteins have NTPase activity and although it does bind NTP's no NTP hydrolyzing activity has been found for TraG, HP0524, TraD or TrwBΔN70 (Balzer *et al*, 1994; Moncalián *et al*, 1999; Schröder *et al*, 2002). An activation step involving accessory factors, possibly from the DNA transport complex, needed for switching on of the NTPase activity has now been suggested (Moncalián *et al*, 1999; Hormaeche *et al*, 2002; Schröder *et al*, 2002).
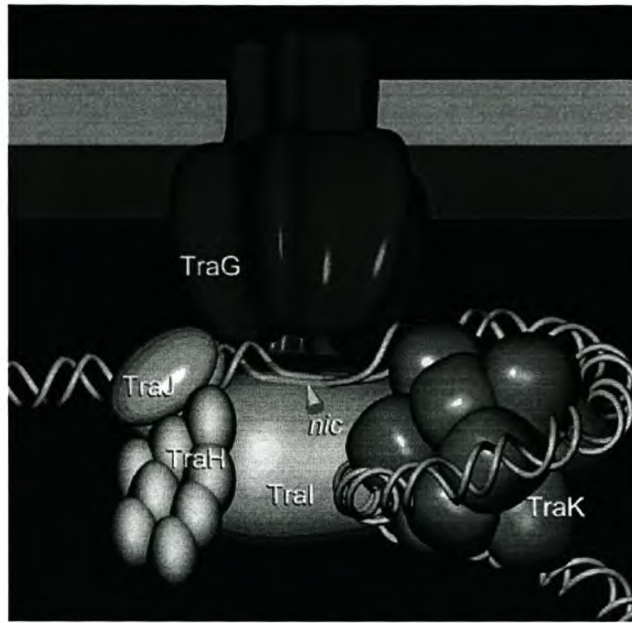
Figure 1.4.1: Proposed model for the RP4 relaxosome. TraG is a membrane anchored, multimeric protein probably forming a pore-like structure that could serve as a channel for translocation of the transferred ssDNA (T-DNA). The relaxase TraI and the plasmid DNA both bind to this TraG pore. TraI cleaves the *oriT* sequence of RP4 at the *nic* site and is covalently attached to the 5' end of the DNA single strand. TraJ binds to the sequence upstream of the *nic* site (*srj*) and is required for relaxase activity. TraH is a homomultimer that stabilizes the TraI-TraJ-DNA complex, probably by bridging TraJ and TraI. TraK binds to a sequence down-stream of the *nic* site and functions as a DNA chaperone, facilitating the formation of the TraI-DNA adduct. A helicase for displacement of the transferred strand has not been identified in the RP4 system (Schröder *et al*, 2002).

Comparison of the affinity of purified TrwbΔN70 for ATP with that of purified wild type TrwB monomer reveals that the absence of the N-terminal segment reduces the nucleotide binding affinity by about 10-fold, which hints at the stabilizing effect the anchoring of this domain in the inner membrane has on the soluble cytoplasmic domain (Hormaeche *et al*, 2002). The location and hexameric structure proposed are suggestive of a translocation portal or passageway for exported substrates.

The TraG of RP4 has been shown to bind specifically to the relaxase TraI from RP4 as well as to the relaxase, Mob, from the gram-positive mobilizable plasmid pBHR1 (Figure 1.4.1) (Szpirer *et al*, 2000; Schröder *et al*, 2002). The F plasmids TraD interfaces with TraM, which forms part of the relaxosome (Disque-Kochem *et al*,

1997). Recently the carboxyl terminus of TraD was shown to add specificity and efficiency to F-plasmid conjugative transfer. Deletion of 140 amino acids from the C-terminal increased transfer frequencies of R388 *trwB* mutants when provided in trans as well as for RSF1010 in this mutant background. The deletion also caused a decrease in transfer frequency of the F plasmid itself while increasing the frequency for RSF1010 in an F *traD* mutant background. This specialization is thought to have occurred due to the stable environment the plasmid finds itself in being narrow host range (Sastre *et al*, 1998). Direct evidence of a link between TraG-like proteins and components of the Mpf transfer machinery is still lacking.

An interesting question that comes from the observation that one conjugative plasmid can mobilize a non-selftransmissble plasmid at a higher frequency than another, or *vice versa* is: Which specific interactions are involved between the relaxosomes of these mobilizable plasmids and the Mpf system used to transfer them? As mentioned previously, RSF1010 can be mobilized by a number of different conjugative plasmids and other type IVSS. For mobilization of RSF1010 by RP4 the Tra2 genes plus the TraF and TraG from the Tra1 region are sufficient (Eisenbrandt *et al*, 1999). Similarly the *virB* region plus VirD4 are all that is required from the Ti plasmids to mobilize RSF1010 (Fullner *et al*, 1998). The PILw (analogous to *virB* and Tra2 regions) and TrwB are all that is required for RSF1010 mobilization by R388 (Cabezon *et al*, 1994). The Ti plasmid *oriT* and relaxase TraA are related to the analogous components of the RSF1010 *mob* system, however the Ti plasmid *tra* (*trb* and *tra* regions together) system does not transfer RSF1010 (Cook *et al*, 1992). This is surprising as the TraG and Mpf (*trb*) system of the Ti plasmid is closely related to their counterparts in RP4, which mobilizes RSF1010 at high frequency.

Table 1.4.1: Interactions between Mpf and Dtr systems mediated by IncRh1 and IncP coupling proteins

Coupling protein	Mating bridge[a]			Relaxosome
	IncRh1 *trb*	IncP	IncRh1 *vir*	
TraG$_{pTiC58}$	+	−		IncRh1 *trb*
				IncP
	−			IncQ
			−	IncRh1 *vir*
TraG$_{RP4}$	−	−		IncRh1 *trb*
	+	+		IncP
	+	+		IncQ
			−	IncRh1 *vir*
VirD4	−			IncRh1 *trb*
		−		IncP
	+	−	+	IncQ
			+	IncRh1 *vir*

[a] +, effective interface; −, ineffective interface.

Table 1.4.1 summarizes the results of swapping the TraG-like proteins between these systems. TraG of RP4 can also substitute for TrwB of R388 in RSF1010 transfer but cannot for R388 self transfer (Cabezon *et al*, 1994). This demonstrates that TraG-RP4 can interface with the R388 Mpf system but does not recognize the R388 relaxosome. Likewise the results from the limited set of data in Table 1.4.1 indicate that in most cases specificity is conferred through interactions between the coupling protein and the relaxosome. However, Ti plasmid *trb* mutants fail to transfer from donors carrying IncP1 plasmids indicating that the TraG-pTiC58 cannot couple the Ti relaxosome to the IncP1 Mpf complex. Secondly VirD4 can couple the RSF1010 relaxosome with both the *vir* and *trb* systems of IncRH1 plasmids but cannot substitute for TraG during RP4 mediated transfer of RSF1010. These observations point to specificities with respect to the coupling protein and the mating bridge. Thus, the coupling protein determines the recognition of a relaxosome by any given mating bridge (Mpf system). The specificity of this protein for any of the relaxosomal or mating bridge components will determine whether or not the nucleoprotein complex will be transported to the recipient (Hamilton *et al*, 2000).

1.5 Rolling circle replication and conjugative plasmid transfer

Ohki and coworkers first suggested a model for conjugative DNA transfer based on the rolling circle (RC) mode of DNA replication, mainly because a DNA single strand is transferred unidirectionally from donor to recipient (Ohki *et al*, 1969).

Rolling circle replication (RCR) was thought to be limited to ssDNA coliphages and small multicopy plasmids from gram-positive bacteria. Plasmids have now been isolated from gram-negative bacteria, cyanobacteria and archaea that replicate via a RC-mechanism. The *Staphylococcal* plasmids pT181, pC221, pUB110 and pC194 as well as the *Streptococcal* plasmid pMV158 have been the focus of studies into the molecular mechanisms involved in RC-replication. The first step in the model proposed for RC-replication is initiation by the plasmid encoded Rep protein. Rep introduces a strand and site-specific break at a region called the double stranded origin (*dso*). The *dso* has two regions: the *bind* region, which is recognized and bound non-covalently by the Rep protein, and a *nic* region where the strand break is introduced. The *nic* sites are highly conserved among members of a family whereas differences are found between *bind* sites within the same family. These differences make the *dso* highly specific for their cognate initiator proteins. The Rep *bind* sequence may also include inverted repeats and although a few RCR-plasmids contain two or three directly repeated sequences (pMV158), iterons are generally not present in the *dso* of most RCR-plasmids. These inverted repeats form hairpins with the single stranded *nic* site exposed at the end.

After cleavage the initiator proteins become covalently bound to the 5'-end of the *nic* site via a phosphotyrosine bond. The 3'-OH end produced by the cleavage reaction is then used as a primer for leading strand synthesis, and elongation from this end continues until the replisome reaches the reconstituted *dso*. The Rep, which may act as a monomer or dimer, then catalyzes a second cleavage reaction at the *nic* site and rejoins the newly formed 3'-end and the bound 5'-end to give two products (see below). One of the termination products is a dsDNA molecule made up of one of the parent strands and the newly synthesized strand. The second product is a ssDNA molecule made from the displaced parental strand. For pT181 termination of leading

strand synthesis only requires 18bp (*nic* site) of the *dso* while initiation requires a much larger region including the Rep *bind* site (Zhao *et al*, 1996). This mechanism does not seem to be conserved in pMV158 (see below).

Generating a complementary strand for the ssDNA product released after termination of leading strand synthesis initiates and terminates at the single stranded origin (*sso*). The *sso* is sequence and orientation specific and is normally located a short distance upstream of the *dso* (Gruss *et al*, 1987). Several *sso* variants have been identified based on their secondary structure sso_A, sso_U, sso_T, and sso_W. The host RNA polymerase synthesizes an RNA primer from both sso_A and sso_U type origins while some like sso_W support replication in both an RNA polymerase-dependent and independent manner (Seegers *et al*, 1995). The host DNA polymerase I initiates DNA synthesis from the RNA primer and DNA polymerase III holoenzyme continues lagging strand synthesis (Kramer *et al*, 1997).

1.5.1 RCR initiator proteins and conjugative DNA relaxases

The RCR-plasmids have been divided in to several families based on sequence similarity between their *dso* regions, and homology between motifs found in their Rep initiator proteins (Table 1.5.1.1) (Khan *et al*, 1997). These conserved motifs are shared with the Rep proteins from ssDNA coliphages and geminivirusses as well as the DNA relaxases, involved in conjugative DNA transfer, from both gram-positive and gram-negative plasmids (Waters *et al*, 1993; Pansegrau *et al*, 1991; Ilyina *et al*, 1992). At least two domains are involved in nicking of the RCR-plasmid *dso*. The first of these, motif III, contains the active tyrosine residue, which becomes covalently bound to the 5'-end of *nic* after cleavage. An exception is the RepB of plasmid pMV158. In this case the Tyr-99 does not form a stable covalent bond with the DNA. A transient bond between the initiator RepB and the *nic* site, which may be long enough to initiate replication, has been proposed (Moscoso *et al*, 1995).

Table 1.5.1.1: Conserved motifs in initiator proteins of RC replicons [a]

Family [b]	Sequence of [c]:	
	Motif II	Motif III
ΦX174	GRLHFHAVHFM	VGFYVAKYVNKKSDM
cons	gxuHUHuxuux	UgxYuakYuxkxxxx
pC194	YNPHFHVLIAV	ELYEMAKYSGKDSDY
cons	yxxHUHvLUxV	xxxExxKYxxKxxDU
pMV158	KKAHYHVLYIA	NVENMYLYLTHESKD
cons	KkxHYHUUUxx	XxxgxUxYUtHxxxD
pT181		SNRFIRIYNKKQERK
cons		SdRFIRIYNKKqERK

[a] The two motifs found in Rep proteins of plasmids representing the main plasmid families, and their comparison with those found in the initiator proteins of ssDNA coliphages (Koonin *et al*, 1993)

[b] Below each Rep protein, a derived consensus is shown.

[c] Motif II, termed the HUH motif, is thought to be the metal-binding domain of the Rep proteins. This motif is not found in the Rep proteins of pT181 family or in the initiator gpII protein of filamentous coliphages. Motif III (the catalytic motif) contains the conserved tyrosine residue involved in the nucleophilic attack to the plasmid DNA at initiation of replication. Capital letters denote totally conserved residues; lowercase letters denote residues conserved in many but not all protein within the family; U and u denote hydrophobic residues; x denotes any residue (del Solar *et al*, 1998).

Most conjugative DNA relaxases, including those form gram-positive plasmids, carry only one nucleophilic tyrosine (Figure 1.2.1.1.2 as well as sections 1.2.1.1.; 1.2.2.1 and 1.2.3.1) However, TraI from the F and R100 plasmids as well as TrwC from R388 were shown to have two clusters of two potentially active tyrosines (Tyr-16 or Tyr-17 and Tyr-23 or Tyr-24 in TraI from plasmid F). Recent evidence indicates that one tyrosine from each cluster actively participates in DNA cleavage (Zechner *et al*, 2000; Byrd *et al*, 1997). The reactions catalyzed by these two residues may resemble those involved in bacteriophage ΦX174 replication. The ΦX174 CisA replication initiator protein also has two active tyrosine residues spaced three amino acids apart. During phage replication CisA introduces a site and strand specific nick in the plus strand and binds covalently to the 5'-end of the cleaved DNA through a

phosphodiester bond with one of these tyrosine residues. A protein complex consisting of CisA and a host encoded helicase unwind the DNA while the 3'-OH⁻ end is extended by DNA polymerase III HE-catalyzed RCR (Yarranton *et al*, 1979). Once the CisA/helicase complex reaches the reconstituted origin it cleaves this *nic* site in a transesterification reaction using the free tyrosine. The tyrosine involved in the original cleavage reaction now catalyses the ligation of the bound 5' moiety to the newly formed 3'-OH⁻ (Fig 1.5.1.1). One of the products released is a circular ssDNA molecule comprised of the plus strand. This tandem arrangement of active sites enables the CisA to reinitiate RCR after a full round of replication leading to the production of a large number of single stranded phage genome copies (Hanai *et* al, 1993).
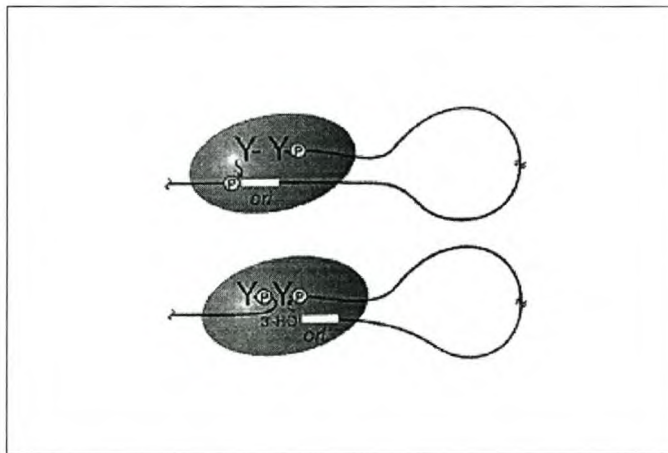
Figure 1.5.1.1: Model for the termination of bacteriophage ΦX174 replication: the reaction catalyzed by the tandem arrangement of active-site tyrosines. Ellipsoids represent protein subunits (CisA). The encircled P represents the phosphodiester moiety, of the tyrosine(s) Y, at the *nic* site. Bent arrows indicate nucleophilic attacks (Pansegrau *et al*, 1996).

In plasmids pT181 and pC221 the phosphodiester bond 5'-ApT-3' is cleaved by Tyr-191 and Tyr-188 respectively, while in the RepA of pC194 Tyr-214 needs two other residues Glu-142 and Glu-210 for catalytic activity (Thomas *et al*, 1990; Noirot-Gros *et al*, 1994). Tyr-214 is involved with nicking of the DNA during initiation while a Glu-210 activated water molecule cleaves the reconstituted *nic* site during termination. Because this second cleavage is not a transesterification reaction, the

RepA is released and does not reinitiate replication. This may be a mechanism that has evolved to prevent runaway replication, which could be deleterious to the host.

How do the RCR initiator proteins and DNA relaxases, with a single active tyrosine, terminate replication or the strand transfer reactions? A mechanism proposed by Rasooly and coworkers, based on the RepC from pT181 is depicted in Figure 1.5.1.2. Their solution is that the RepC acts as a homodimer. Recent evidence has shown that RepC may be a monomer in solution but dimerizes on the DNA (Khan *et al*, 1997). After catalyzing the second cleavage reaction the RepC becomes bound to the 5'-end of a 10bp oligonucleotide produced by elongation of the leading strand past the *nic* site during replication of the leading strand. The result is a RepC::RepC-DNA covalent complex which is inactive and cannot reinitiate replication (Rasooly *et al*, 1993; Jin *et al*, 1997). This could be yet another mechanism to curb runaway replication.

As discussed earlier the MobA of the IncQ plasmid R1162 can terminate transfer DNA replication via a second cleavage mechanism. Similar results have also been obtained for TraI from the F plasmid (two nucleophilic tyrosines) (Gao *et al*, 1994). Studies using an immobilized RP4 TraI monomer have shown that it is capable of cleavage of at the *nic* site during initiation and can join the free 3'-end of *nic* with the bound 5'-end. The TraI monomer could however not catalyze a second cleavage reaction on a newly introduced intact *nic* site while already bound to the 5'-end of *nic*. If the 3'-end at *nic* were to be extended by a RCR mechanism a TraI monomer would not be able to cleave the reconstituted *nic* site. It was therefore postulated that the relaxase acts as a dimer similar to the Rep initiator proteins of RC-plasmids pT181 and pC221 (Pansegrau *et al*, 1996).
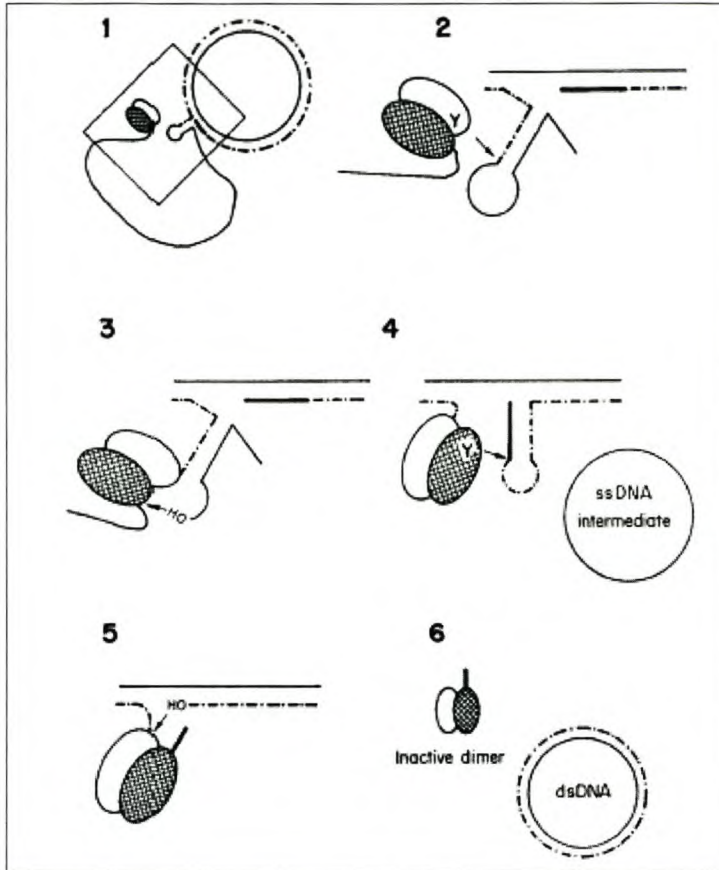
Figure 1.5.1.2: Model for termination of RC replication. Arrows indicate nucleophilic attacks exerted by the OH groups of the Tyr residue of Rep (Y) or by 3' -OH groups of the DNA (OH). Solid lines, parental DNA; broken lines, newly synthesized DNA. The thick solid line indicates the nucleotides that are newly synthesized past the reconstituted *dso* and that will remain covalently bound to the Rep protein to generate a Rep/Rep* inactive dimer. The two subunits of the RepC dimer are differently depicted (del Solar *et al*, 1998).

The second domain involved with nicking of the origin (*dso*) in RCR is a DNA binding domain. This recognizes and interacts via specific non-covalent binding of the less conserved *bind* region of the *dso*. In RepD from pC221 the *bind* domain and active tyrosine residue are separated by 80 amino acids. This indicates that the final structure of the protein has to be such that these two domains are close together to interact with the *dso* where the *bind* and *nic* sites are contiguous. This is similar to the arrangement of motif I and II form TraI of RP4 discussed earlier. The DNA binding domain is located towards the C-terminal end of the initiator proteins from the pT181 family. Six amino acids located in this motif determine DNA binding specificity, which allows only the initiator and *bind* site from the same plasmid to recognize each

other. Replacement of these amino acids in RepC from pT181 with the corresponding residues from RepD (pC221) and RepE (pS194) allowed the mutant RepC to interact with the *bind* region specific to each of the donor proteins. The amino acids on either side of these six residues also seem to be important in the binding at the *dso* or for the correct folding of the DNA binding domain. A 26 amino acid peptide, which included this six-amino-acid region, was found to be inactive in DNA binding.

There are however differences between DNA relaxases and RCR initiator proteins. One of these is that the relaxases have a high affinity for the 3'-terminal region of the substrate. This helps keep the superhelicity of the plasmid DNA intact allowing it to exist stably as a relaxosome without impairing other plasmid functions. The RCR initiators, on the other hand, release the 3'-terminus directly after cleavage. The active site tyrosines of RCR initiators are also always located toward the C-terminal while the nucleophilic tyrosines in all conjugative relaxases described thus far are located within the first 50 amino acids from the N-terminal.

1.6 Aims of this project

The first aim was to determine and analyze fully the nucleotide sequence of the mobilization region of pTC-F14 (Chapter2). This would not only lay the groundwork for future research on this region but also allow for comparison with other systems.

The second aim was to identify which genes are required and which are accessory to mobilization. This would be the first step in elucidating the function of mobilization genes as well as indicating, which gene products should be the focus of future research.

A third aim was to determine if, because pTC-F14 (Isolated from *Acidithiobacillus caldus* in a bacterial consortium undergoing pilot scale testing for the commercial extraction of nickel from ores) and pTF-FC2 (Isolated from *Acidithiobacillus ferrooxidans* found in the leach liquor of a South African gold mine) were isolated from a shared environment, these plasmids compete with or cross complement each

other at the level of mobilization. This would help to better understand the evolutionary relationship between these two IncQ-2-group plasmids.

Chapter Two

Delineation and sequence analysis of the functional mobilization region

Content

2.1 Introduction

Movement/transfer of conjugative or mobilizable plasmids within the horizontal gene pool as well as their ability to be maintained in a variety of hosts results in the spread of various genes including antibiotic resistance markers and virulence determinants. These complicate traditional antimicrobial treatment in human, animal and plant diseases. The study of mobilization systems therefore serves to broaden our understanding of the mechanisms involved in plasmid transfer, which in turn allows us to either exploit these e.g. in developing genetic systems (transferable vectors) for study and manipulation of organisms, or to find ways of curbing plasmid transfer.

The 14.2 kb plasmid pTC-F14 was isolated from the moderately thermophilic (45°-50°C), highly acidophilic (pH 1.5 to 2.5), chemolithotrophic *Acidithiobacillus caldus* strain F (Hallberg *et al*, 1994). This organism was one of two dominant organisms in a bacterial consortium undergoing pilot scale testing for the commercial extraction of nickel from ores (Rawlings *et al*, 1999). pTC-F14 was shown to have an IncQ-like replicon that was closely related, and compatible with the broad host range plasmid pTF-FC2 (Gardner *et al*, 2003). Plasmid pTF-FC2 had been previously isolated from a different, but related iron and sulfur-oxidizing bacterium *Acidithiobacillus ferrooxidans* found in the leach liquor of a South African gold mine (Rawlings *et al*, 1984; Dorrington *et al*, 1990; Rohrer *et al*, 1992). pTC-F14 was also shown to replicate autonomously not only in *E. coli* but also in *Pseudomonas putida* and *Agrobacterium tumefaciens* LBA4404 suggesting a broad host range for this plasmid (Gardner *et al*, 2001).

As mentioned in section 1.2.2.1, four groups have been identified among the IncQ plasmids. These are made up of two major groups of IncQ and IncQ-like plasmids with the main distinguishing characteristic between the major groups being whether they have a three-gene IncQ-like mobilization system or a five-gene IncP-like mobilization system (Rawlings and Tietze *et al*, 2001). RSF1010 together with pIE1107, pIE1115, pIE1130 and pDN1 form one of the major groups and have three-mob protein IncQ-type systems (Tietze *et al*, 1998; Whittle *et al*, 2000). The other group now designated IncQ-2-group, thus far, is comprised of only two plasmids:

pTF-FC2 and pRAS3 (Rohrer *et al* 1992; L'Abée-Lund *et al*, 2002). The latter being isolated from the fish pathogen *Aeromonas salmonicida* in Norway. The mobilization genes of group 2 are more closely related to the Dtr genes of the Tra1 from the region of IncP1 group of plasmids such as RP4 and R751, rather than the other IncQ-like plasmids. Another plasmid isolated from *At. ferrooxidans*, pTF1, has a two-gene mobilization system made up of *mobL* and *mobS*. These genes are most closely related to *mobA* and *mobC* from RSF1010, respectively. MobL (378 amino acids) is unique in that unlike its IncQ counterparts, it does not have a C-terminal primase domain although a heptapeptide sequence, A-Q-R-Q-Q-E-K found in the primase domain of MobA from RSF1010 is also found toward the C-terminal of MobL (Drolet *et al*, 1990).

Sequencing and analysis of the pTC-F14 mobilization region would show the arrangement and relation of the mobilization genes to other systems or could reveal a novel mobilization system. Although the environments within most bacterial cells are similar this could also shed light on specific adaptations of a mobilization system (Dtr genes) to different ecological niches or hosts (temperature and/or available energy). Sequencing is also a prerequisite for accurately identifying sites required, in *cis*, for plasmid transfer e.g. the *oriT* or IHF binding sites.

2.2 Methods

2.2.1 Bacterial strains and plasmids

The bacterial strains used in this study are listed in Appendix 1. Plasmids used as cloning vectors and for making various constructs are also listed in Appendix 1. The *E. coli* strain DH5α was routinely used for maintenance of plasmids and for preparation of plasmid DNA.

2.2.2 Media and growth conditions

Escherichia coli strains were grown in either Luria-Bertani broth or on LA plates at 37ºC, supplemented with antibiotics as required at the following concentrations: ampicillin (100 μg/ml), chloramphenicol (34 μg/ml), streptomycin (50 μg/ml), kanamycin (50 μg/ml), nalidixic acid (50 μg/ml) (Sambrook *et al*, 1989).

2.2.3 DNA manipulations, sequencing and bioinformatics

For cloning of fragments generated by PCR, the pGEM®-T vector system was used (Promega Corporation, Madison, Wisconsin, USA, catalogue number A3600). DNA fragments excised from agarose gels used for cloning were extracted using a GFX™ kit (Amersham® Biosciences Corporation, catalogue number 27960201). Clones to be sequenced were isolated and purified from a 5 ml overnight culture using a High Pure Plasmid Isolation Kit™ (Roche®, Basel, Switzerland, catalogue number 1754785). The concentration of purified plasmid DNA was determined by a single wavelength reading at 260 nm in a UV spectrophotometer, and diluted to give a final concentration of 120 ng/µl used in the sequencing reaction. The DNA sequence was determined by sequencing from the ends of a number of subclones to obtain overlapping sequence from both strands (Figure 2.3.2.1). Sequencing was by the dideoxy chain termination method, using an ABI PRISM™ 377 automated DNA sequencer and the sequence was analysed using a variety of software programmes but mainly the PC-based DNAMAN (version 4.1) package from Lynnon Biosoft. The nucleotide sequence of the 5.54 kb region sequenced was submitted to the EMBL database (accession number NC_004734/AF325537). Searches for sequences related to Mob proteins were performed using the gapped-BLAST program of the National Center of Biotechnology (NCBI) at www.ncbi.nih.nlm.gov (Altschul *et al*, 1997). Sequence alignments (based on CLUSTAL W) were carried out using the multiple alignment programme. Amino acid sequence homology trees were constructed using the tree output programme within the DNAMAN package. Other techniques were performed according to standard procedures or to the manufacturers' recommendations (Sambrook *et al*, 1989; Ausubel *et al*, 1993).

2.2.4 Mating procedures

A two-way mating system was employed using either *E. coli* S17-1, which has an RP4 derivative integrated into the chromosome to provide the Mpf genes, or HB101 as donors and CSH56 as recipient. Donor and recipient cells were cultured separately overnight with appropriate antibiotic selection. Cells were harvested by centrifugation for 2 min at 8000 rpm, washed three times in 0.85% (w/v) NaCl solution, and mixed in a donor-to-recipient ratio of 1:10. 100 µl of this mixture was

spotted onto a LA plate and incubated at 37ºC for 1 h. The agar plug was excised and suspended in 5 ml of saline solution and vigorously shaken to dislodge mating cells and disrupt conjugation. Cells were pelleted by a 2-min spin in a microcentrifuge and resuspended in 1 ml 0.85% NaCl solution. If ampicilin was used for selection, cells were washed again three times in saline solution to remove any extracellular β-lactamase and to prevent the growth of satellite colonies at low dilutions. Serial dilutions were then plated onto media that selected for donor and transconjugant cells. The transfer frequency was calculated as the number of transconjugants per donor during the 1 h mating period.

2.2.5 PCR amplification of the minimum mobilization region clones

PCR reactions were performed using the Expand™ High fidelity *Taq* DNA polymerase (catalogue number 1732641) from Roche® with a Hybaid® PCR Sprint cycler. The template for these reactions was plasmid pMmob, which consists of the 5.54 kb *Bam*HI-*Xba*I fragment from pTC-F14 cloned into pUC19. The primers used to give pMmob1-pMmob9 (Figure 2.3.3.1) are listed in Appendix 3. After an initial denaturation of 60 s at $94^{o}C$, 25 cycles of 30 s at $55\text{-}60^{o}C$ (depending on primer set) and an elongation step of up to 4 min at $72^{o}C$ were performed. A final extension step of 120 seconds at $72^{o}C$ before cooling to $4^{o}C$ completed the reaction. The various PCR products were cloned into the non-mobilizable vector pUC19 to give the deletion series pMmob1-pMmob9 (Figure 2.3.3.1).

2.2.6 In-vitro transcription translation analysis

A prokaryotic, DNA-directed, *E. coli* S30 extract-based *in vitro* transcription-translation kit for circular DNA (Promega Corporation, catalogue number L1020, lot number 158097) was used to identify translation products from the minimum mobilization region clones (pMmob1 through pMmob8) as well as from pMmob and the pTC-F14Cm plasmid. The kit was used as per manufacturers recommendations except for using half reaction volumes to double the amount of labeling reactions that could be performed. The [^{35}S]-methionine-labeled translation products were separated on a 15% sodium dodecyl sulfate-polyacrylamide electrophoresis gel and detected by autoradiography.

2.3 Results and Discussion

2.3.1 Mobilization of pTC-F14

A selectable chloramphenicol resistance gene was cloned into plasmid pTC-F14 using the unique *Bam*HI site to produce plasmid pTC-F14Cm (Appendix 2). When mobilized from *E. coli* S17-1 the transfer frequency of pTC-F14Cm was approximately 2.8×10^{-3} transconjugants per donor. To determine whether the type of conjugative plasmid affected the mobilization frequency, we compared mobilization frequencies using two self-transmissible plasmids different from the RP4 (IncPα) that was integrated into the chromosome of *E. coli* S17-1. Plasmid pTC-F14Cm was mobilized by R751 (IncPβ) from *E. coli* HB101 at a frequency of about 1.3×10^{-5}, about 100-fold lower than by the RP4 derivative in *E. coli* S17-1, while mobilization by R388 (IncW) was not detectable. A 5.54 kbp *Bam*HI-*Xba*I fragment from pTC-F14 was subcloned into the non-mobilizable vector, pUC19 (pMmob), and was mobilized by *E. coli* S17-1 at frequencies that approached saturation, demonstrating that all the necessary features needed for mobilization were present. Saturation indicates that after 1 h of mating at a donor to recipient ratio of 1 to 10, the number of transconjugants was approximately equal to the number of recipients (Table 2.3.3.1). This 5.54 kb fragment therefore contained all the components needed for mobilization and was sequenced (Figure 2.3.2.1).

2.3.2 DNA sequence analysis of the *Bam*HI-*Xba*I subclone

The 5.54 kb insert of pMmob was first mapped for various restriction enzyme sites that could be used for subcloning. The insert DNA was digested to give fragments of ±200 bp – ±1.3 kb, which were cloned into the pUCBM21 vector, and sequenced using the M13 forward and reverse universal primer set (Figure 2.3.2.1).
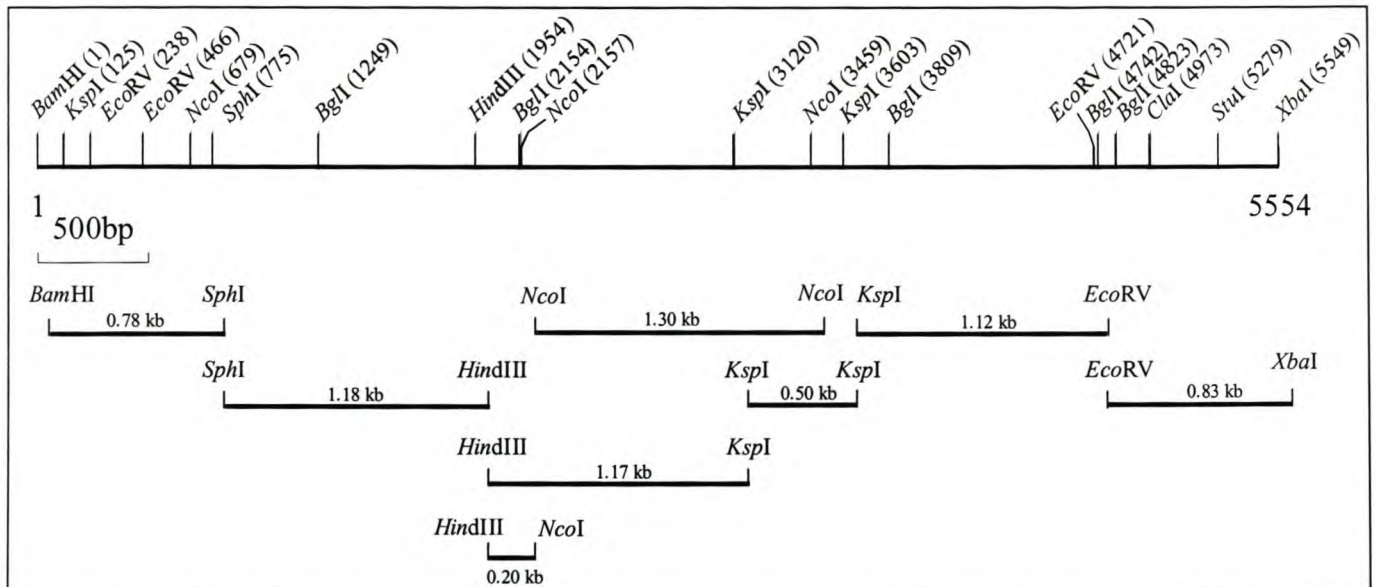
Figure 2.3.2.1: Linear map of the 5.54 kbp *Bam*HI-*Xba*I clone. Restriction enzyme positions that could be determined by digestion are indicated. Subclones used for sequencing as well as their respective sizes are shown.

The PCR generated minimum mobilization region clones, pMmob1-pMmob8, (section 2.3.3) were also sequenced from both ends to ensure that all necessary open reading frames (ORF's) were included. This sequence was compared with the sequence data from sub-cloning to confirm the sequence accuracy over larger subclones such as the 1.18kb *Sph*I-*Hin*dIII fragment (Figure 2.3.3.1).

From the compiled sequence, five putative ORF's were identified based on sequence similarity with known genes. The highest amino acid sequence similarity found, was to the mobilization genes of plasmid pTF-FC2: *mobA* through *mobE*. This nomenclature was therefore used to describe the pTC-F14 mobilization genes. The arrangement of the ORF's was similar to those of pTF-FC2 and pRAS3.1 being divergently transcribed from the *oriT* (Figure 2.3.3.1). All potential ORF's had

identifiable putative Shine-Dalgarno ribosome binding sites (Table 2.3.2.1)(Shine and Dalgarno *et al*, 1974).

Table 2.3.2.1: Potential translation initiation sites as indicated by the Shine-Dalgerno sequence.

ORF	Translation initiation region[a]		
mobA	CGCCAUC**GAGGGG**CUGAGUC	AUG	AUC
mobB	CCGCCGC**GAAAGGAGA**CACG	AUG	CCA
mobC	GGCAC**AGGAGGAA**CGACAGG	AUG	GCA
mobD	AAACCGGAUUC**GGAGGA**CUU	AUG	AGC
mobE	GACCGCAUU**GAAGG**CGGCUU	AUG	AGU

[a] Putative Shine-Dalgarno sequences are in bold type.

Three of the Mob proteins (MobA, MobB and MobC) of pTC-F14, pTF-FC2 and pRAS3.1 had a greater than 20% amino acid sequence identity to the N-terminal ±400 amino acid portion of TraI, and the complete TraJ and TraK proteins of the IncPα plasmid RP4 and the IncPβ plasmid R751 respectively. When compared with the published amino acid sequence as well as the sequence available on the NCBI database for MobA from pTF-FC2 and pRAS3.1, the pTC-F14-MobA N-terminal was found to extend further toward the MobB by 56 amino acids, with its start codon overlapping the MobB C-terminal. Upon closer inspection of the DNA sequence of pTF-FC2 and pRAS3.1 (Genbank accession AY043299.1/NC_003124.1) the MobA proteins for each of these plasmids was also found to be truncated at the N-terminal by 57 amino acids with a GTG start codon for the pTF-FC2 MobA. The "extended" form of these proteins was used for comparison (Figure 2.3.2.2). The pTC-F14-MobA displays all three motifs conserved within other DNA relaxases with a potential active-site tyrosine, Tyr-24, aligning with the active-site tyrosine identified in RP4 TraI (Pansegrau *et al*, 1994). No ORF's, internal to MobA, were identified unlike the TraI of RP4 from within which the *traH* gene is transcribed.

```
                              Motif I
                                        *
pTF-FC2   MIVKKVKNPQKAASKAVRVSRLTGYIREPERENSQEKCIHAGARGFITDEPQSQTAEMIALSQEAVRSK.     6!
pTC-F14   MIVKKVKS.NRTKGKAASIRDLTNYIREPQNRNPNEKVLYANGRGFISDTHAAQREEMVALAAEAVRSR.     6!
pRAS3.1   MIVKKVPNPKKSASKAQRIGQLTSYVRSPESESPQEKCLYAGARGFMMDDPKSQTAEMIALSQEAVRSK.     6!
pRA2      MLAKVPPKRADGKTSFKSLAKYACERDHIDPETG.AVERRECSTETNCLDKDTAWREMKAVSDMNGRVK.     6!
R751      MIAKHVPMRSLGKSDFA...GLANYITDAQSKD..HRLGHVQATNCEAGSIQDAITEVLATQHTNTRAKG     6!
RP4       MIAKHVPMRSIKKSDFA...ELVKYITDEQGKT..ERLGHVRVTNCEANTLPAVMAEVMATQHGNTRSEA     6!
Consensus    m  k                                                    e   a        r
```

```
           Motif II                                        Motif III
pTF-FC2   DTINHYVLSWREGEQPSPEQVEEAVSIFMDELGVKDHQAIYGLHADTDNLHLHLAINRVHPETLKVVKIN    13!
pTC-F14   NPVNHYILSWREGEQPSPEQVEEAVSIFLDELGLQEHQVIYALHKDTDNLHLHIAVNRVHPETLKCVEIN    138
pRAS3.1   DTINHYVVSWREGEQPSPEQVEEAVSIFMDELGWKDHQAIYGLHSDTDNIHLHIVINRVHPETLKIVEKN    13!
pRA2      DPVYHFTVSWPAHEKPTDAQVFEAGREGMKSLGMEGHQYLAAVHRDTDNVHGHFMVNRVNPETYKAVYPD    138
R751      DKTYHLIVSFRAGEQPSADTLRAIEERICVGLGYGEHQRISAVHNDTDNLHIHIAINKIHPTRHTMHEP.    134
RP4       DKTYHLLVSFRAGEKPDAETLRAIEDRICAGLGFAEHQRVSAVHHDTDNLHIHIAINKIHPTRNTIHEP.    134
Consensus       h   s     e p            lg   hq      h dtdn h h    n   p
```

Figure 2.3.2.2: Alignment of IncP-like conjugative DNA relaxases. Conserved amino acids are indicated with white text on a black background while a shaded background marks positions where conservative replacements may occur. The asterisk above the sequence indicates the active tyrosine identified in the RP4 and R751 TraI proteins. The conserved motif's identified in TraI of RP4 are indicated by black lines above the sequence blocks (Figure 1.2.1.1.2).

MobD and MobE had weaker but detectable sequence identity (17-18%) to TraL and TraM of RP4 and R751 respectively. The Walker A ATP/GTP binding site (P-Loop) found in TraL is also conserved in MobD, located between amino acids 14 and 21 (G-K-G-G-V-G-K-S) (Appendix 4) (Zieglin *et al*, 1991). Using the DAS (http://www.sbc.su.se/~miklos/DAS/maindas.html) and SOSUI (http://sosui.proteome.bio.taat.ac.jp/sosuimenuo.html) web based transmembrane prediction servers, one potential transmembrane helix (TMH) was identified in MobE that stretches from amino acid 107 to 129. This feature seems to be conserved in MobE from pTF-FC2 and pRAS3.1 (Figure 2.3.2.3). However, in TraM from RP4 and R751, amino acids 122-144 in the C-terminal appear to function as a membrane anchor. If the MobE is a membrane bound protein, it appears to have a domain on either side of the membrane, unlike the TraM. This may represent an adaptation to a different environment/host or perhaps to a different conjugative system (Chapter 4). As yet no function has been assigned to the TraL or TraM of the IncP1 plasmids nor are they required for transfer of these plasmids between *E. coli* strains. The same applies for the MobD and MobE of pTC-F14 (section 2.3.3).
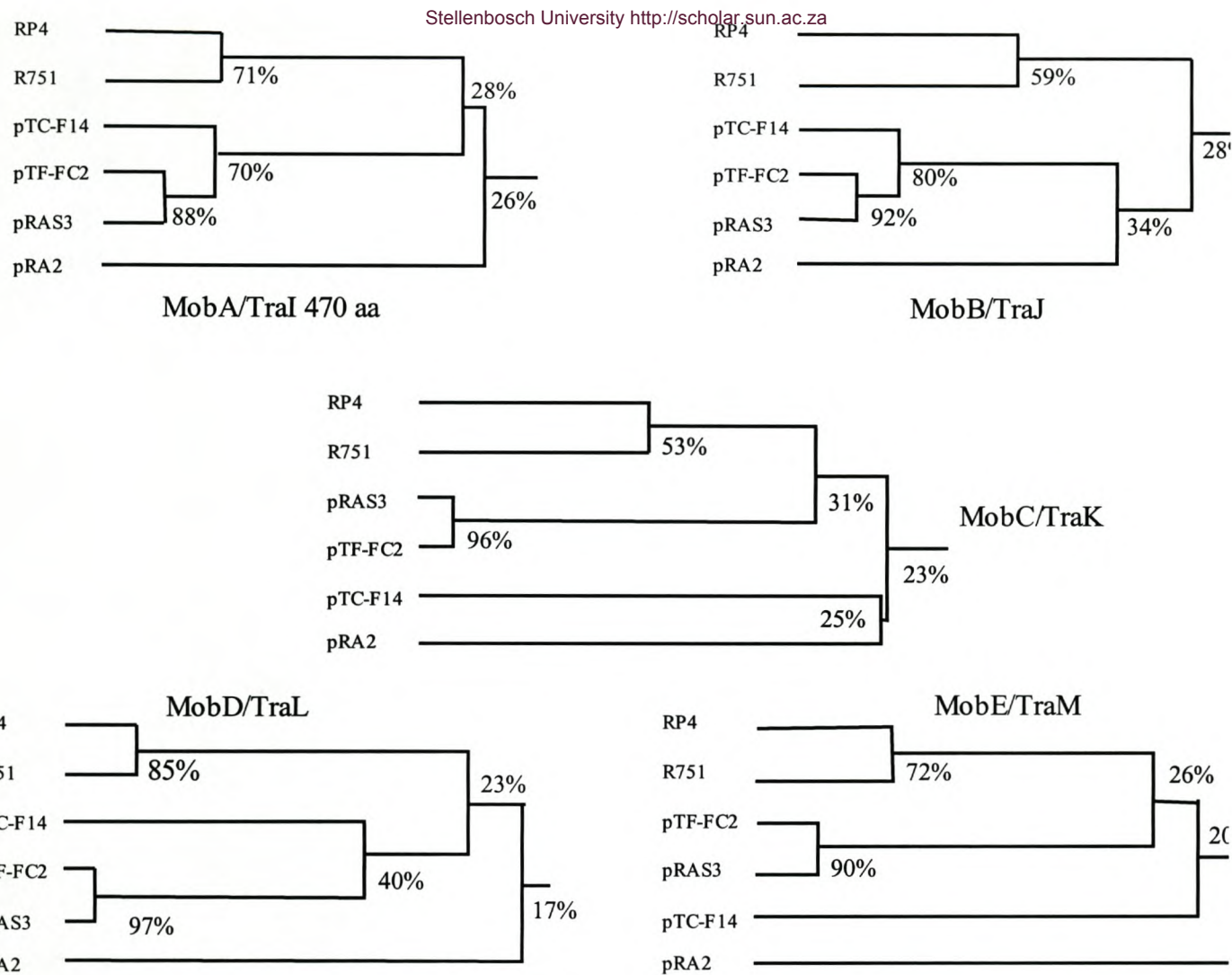
Figure 2.3.2.3: Phylogenetic relationship between the MobA/TraI, MobB/TraJ, MobC/TraK, MobD/TraL and MobE/TraM proteins of the IncP and IncQ-2-type plasmids as well as plasmid pRA2. Since in the IncQ-related plasmids the MobA protein exist as a MobA-RepB fusion, only the N-terminal 400 aa was considered for comparison. Percentages are percent amino acid sequence identities. GenBank accession numbers are RP4,(X54459);R751,(X54458); pTF-FC2,(M57717); pTC-F14,(NC004734); pRAS3.1,(NC003123); pRA2,(U88088).

These Mob proteins clearly belong to the IncP-like family of conjugation-associated DNA processing proteins (Dtr) and a dendrogam showing the relationship between proteins of this family is presented in Figure 2.3.2.3 (Pansegrau *et al*, 1994; Thorsted *et al*, 1998). One may speculate that, at least MobA, MobB and MobC, perform similar functions to their counterparts in the IncP1 plasmids as discussed in section 1.2.1.1.

```
                                                                  TMHa
pTC-F14    KIASNAEEAAKGRIS.....AASS.ELLSKAGQQLVDAFRSDLGKTLWSRSVWAGLVLVVVGIGSYAAGE    12
pTF-FC2    NFKVTADATVKASAE....AAKADLAQAVAAAAQEVAHNTSAKQMWQWAAGCIAVAFLCVALFGWYMHSS    13
pRAS3.1    NFKVTADATVKASAE....AAKADLAQAVAAAAQEVAHNTSAKQMWQWAAGCIAVAFLCVGLFGWYMHSS    13
RP4        GIAHRWGEDAKAKAERMLNAALAASKDAMAKVMKDSA.............................      9
R751       SIAHRWGDDAKGKAERTLNAALAASKDAMTRGMQEGAK.............................      9


pTC-F14    VDIWWHSKEVRQVEQKLQRLKAEEAAIQTRIGGIVQECPAFDGIQSGPCVPIDVAANVANNFGYEKHDGK   19
pTF-FC2    GKDSGYQGRIPAGYTEAKDEKAAAAWANTPEGRT.AYRFAQSGE......LQRLARCSG........KGW   19
pRAS3.1    GKDSGYQAGYGAGYTEAKDEKAAAAWANTPEGRL.AYRFAQTGS......LASLAKCDR........PGW   19
RP4        ....................AQAAEAIRREI..DDGLGRQLAA......KVADARRV.........MMN   12
R751       ....................A.AAEAVRREV...EAVTAQLVA......PIRBARRVA........MMN   12


pTC-F14    TVYYGRLDIGAIEQAETKQHQ...                                               22
pTF-FC2    KVEKCACYPYPVANEGTYGWALP.                                               21
pRAS3.1    YVEKGVCYVKP.ASDGTYGWRL.P                                               21
RP4        MIAGCM.VLFAAALVVWASL....                                               14
R751       MVAACMAVV.AAGLALWASL....                                               14
                     TMHb
```

Figure 2.3.2.4: Alignment of MobE-like proteins showing the transmembrane helices (lines above and below sequence blocks). [a] Trans membrane helix conserved in MobE from pTC-F14, pTF-FC2 and pRAS3.1. [b] Trans membrane helix conserved in TraM from RP4 and R751

Some of the predicted amino acid sequences and characteristics of the mobilization proteins of pTC-F14 differed substantially from their counterparts in pTF-FC2. Compared with the pTF-FC2 sequence the MobA/RepB fusion and MobB proteins were the best conserved at 73.3 and 77.8% amino acid sequence identity, while the MobC, MobD and MobE proteins were poorly conserved at 26.5, 39.8 and 21.2% amino acid sequence identity respectively (Table 2.3.2.2). Although all pairs of Mob proteins were of comparable sizes, the predicted pI values of the MobE proteins differed by almost 3 pH units.

Surprisingly, plasmid pRAS3.1 has Mob proteins that are considerably more closely related to pTF-FC2, than pTC-F14 is to pTF-FC2. The sequences of the MobA, MobB, MobC, MobD and MobE proteins of pRAS3.1 are 93.8, 88.8, 94.1, 97.4 and 88.8% identical to pTF-FC2 respectively, whereas they are only 72.7, 74.5, 25.8, 40.7 and 20.8% identical to pTC-F14.

Table 2.3.2.2: A comparison between the predicted mobilization proteins of pTC-F14 and pTF-FC2.

Predicted gene product	pTC-F14			pTF-FC2			Percentage amino acid identity
	Length (aa)	Predicted mass (kDa)	pI	Length (aa)	Predicted mass (kDa)	pI	
MobA/RepB	889	102,182	9.5	888	101,212	9.6	73.3
MobB	103	11,198	9.7	106	11,605	9.8	77.8
MobC	131	13,969	10.0	118	12,941	10.0	26.5
MobD	226	24,696	6.6	227	25,274	5.3	39.8
MobE	220	23,811	5.5	213	23,093	8.2	21.2

The *oriT* region of pTC-F14 was identified by inspection of the sequence and found to be located on a 203 bp *NcoI-Hin*dIII fragment. This fragment was cloned into the non-mobilizable pUC19 vector (pOriT-F14) and transformed into *E. coli* S17-1 that contained a chloramphenicol resistant pTC-F14Cm to provide the Dtr functions. pOriT-F14 was mobilized at a frequency that was about 500-fold greater than that of pTC-F14Cm, indicating the presence of a functional *oriT* (Table 2.3.3.1). The *oriT* regions of IncPα and IncPβ plasmids, as well as the four selected plasmids which have mobilization regions related to the IncP plasmids are compared in Figure 2.3.2.5. The four mobilizable plasmids each contained an inverted repeat sequence that has been shown to be the site where the TraJ of plasmid RP4/RK2 binds prior to nicking at the *oriT* (Ziegelin *et al*, 1989; Whittle *et al*, 2000). The highly conserved nucleotide hexamer recognized by TraI in RP4, which immediately precedes the *nic* site, is also shown (Appendix 4). No potential integration host factor (IHF) binding sites were identified unlike in the *oriT* sequence of pTF-FC2 (Rohrer and Rawlings *et al*, 1993), although IHF binding in pTF-FC2 was shown only to have a minor effect on the mobilization frequency. In contrast to the mobilization proteins where plasmids pTF-FC2 and pRAS3.1 were the most closely related, the *oriT* regions of pTC-F14 and pRAS3.1 were considerably more closely related (42/50 bp matches)

than pRAS3.1 and pTF-FC2 (32/50 bp matches) or pTC-F14 and pTF-FC2 (30/50 bp
matches) (Figure 2.3.2.5).

```
A                 ───── ──   ─────>  <───── ── ─ ─────          ▼
RP4       GTGAAGAAGGAACACCCGCTCGCGGGTGGGCCTACTTCACCTATCCTGCC

          ───── ──   ─────>  <───── ── ─ ─────
R751      GTGAAGATAGATAACCGGCTCGCCGGTTAGCTAACTTCACACATCCTGCC


B                 ─────   ─────── ─────>  <───── ── ── ──────
pTF-FC2   CCGTTGGAACCACGTGGCGAAGCCATGTGTGTTACAACGGTCATCCTGTA

          ─ ───── ─   ─────>  <───── ─── ── ─
pTC-F14   GGCTTGGCGGGAACATGCGTAG CATGTGTATGACAAGTCCCATCCTGTT

          ───── ──   ─────>  <───── ── ──────
pRAS3.1   GGATTGGCAGGAACATGCGTAG CATGTGTATGACAATCCGCATCCTGTA

          ── ─ ─── ─────>  <───── ── ─ ──
pRA2      CCGGTGGCACACTGCCGCTTTAGCGGCGGTGGGACATTCCCTACCCTGTC
Mob consensus     TGG A         GC      C    GT    ACA       A CCTGT
Joint consensus     G                             AC        A CCTG
```

Figure 2.3.2.5: Comparison of the *oriT* regions of A., the IncP conjugative plasmids and B.,
the IncQ-2-group and pRA2 mobilizable plasmids. Imperfect inverted repeat sequences are
shown using arrows, while the highly conserved hexamer preceding the nick site is boxed.
The small vertical arrowhead shows the nick site as determined for RP4/RK2 (Appendix 4).

As was proposed for pTF-FC2, it appears that plasmid pTC-F14 is a natural hybrid of
an RSF1010-like replicon and an RP4-like mobilization region, which clearly belongs
to the IncQ-2-group plasmid family.

2.3.3 Delineation of the region required for mobilization

A series of PCR based deletions of the pTC-F14 mobilization region was constructed to: (i) Determine the smallest region that is mobilizable at the frequency of the intact *mobE-repB* region, (ii) To test the necessity of certain genes in mobilization (whether they are required) and, (iii) To be used in identifying translation products from predicted ORF's (Figure 2.3.3.1). The latter will be dealt with in section 2.3.4.
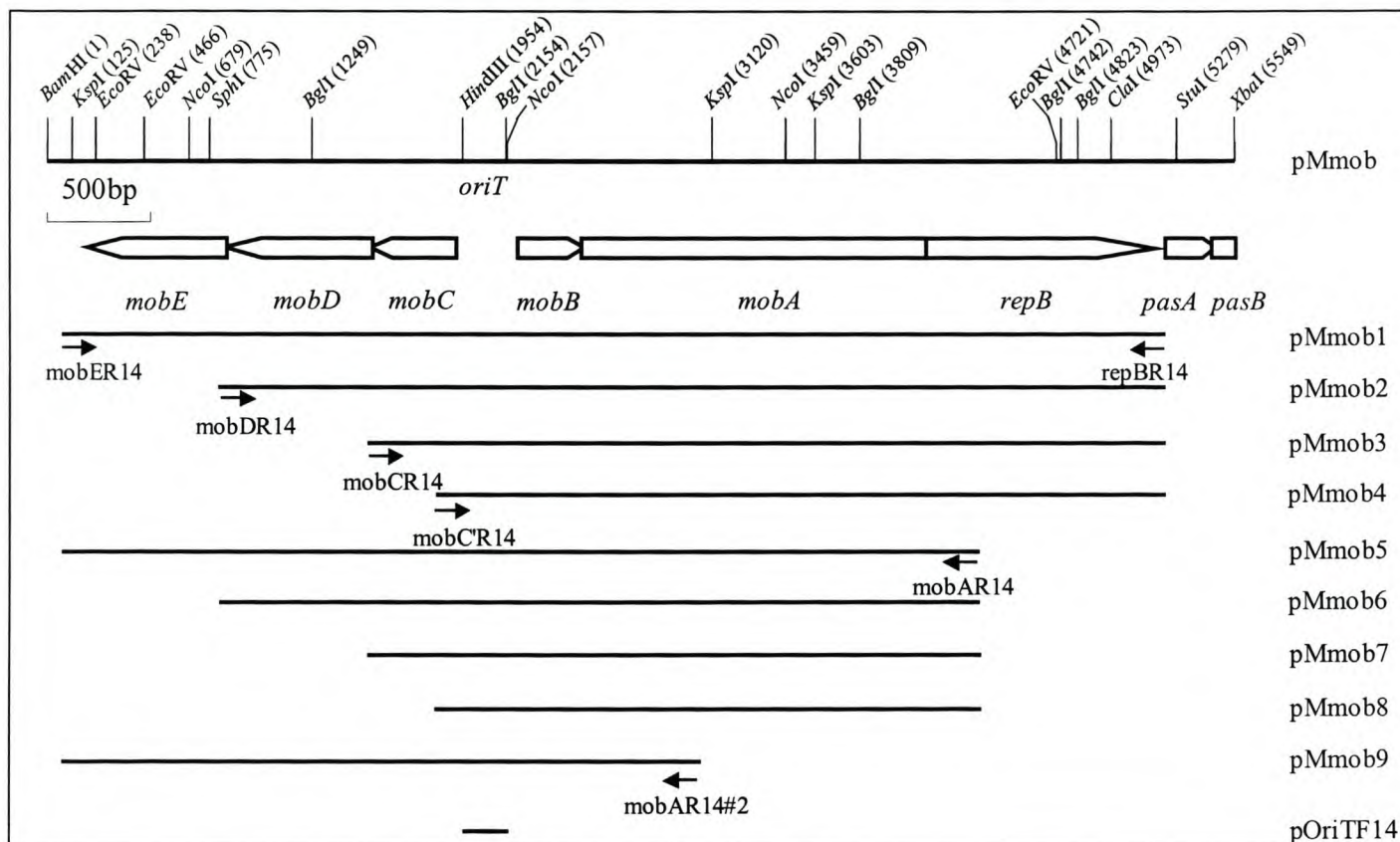
Figure 2.3.3.1: The 5.54 kb *Bam*HI-*Xba*I region (pMmob) of pTC-F14 showing the locations of *mob*, *repB* and *pas* genes as well as the *oriT*. Short horizontal arrows show the positions of the primers used to amplify and construct certain subclones.

When the entire *mobE-repB* region of pTC-F14 was cloned into the ColE1-replicon based pUC19 vector (pMmob1), the mobilization frequency was at the level of saturation. This was an increase of more than 3000-fold relative to the frequency obtained with the mobilization genes linked to its natural replicon. To test whether this increase in mobilization frequency was due to placement of the *mobE-repB* region in the pUC19 vector (up to 500 copies), the *mobE-repB* region of pMmob1 was

64

cloned into vectors pACYC184 (p15A replicon; 20-30 copies) and pBR322 (ColE1 replicon; 25-50 copies). Both of these constructs (pMmob1184 and pMmob1322) had mobilization frequencies ~50-fold less than the pUC19 construct (pMmob1) but still ~100-fold higher than that of the parent plasmid (Table 2.3.3.1). Reasons for the increased mobilization frequency could be: (i) A higher dosage of Mob proteins, (ii) More copies per cell could mean that the relaxosomes may be more readily available for transfer, (iii) A combination of these factors. The higher transfer frequency of pOriTF14 when compared with pTC-F14Cm (copy number 12-16) shows the contribution of having more copies of *oriT* in the cell, while having the same amount of Mob protein available as the natural plasmid (pTC-F14Cm) per cell. Increased transfer frequencies have been reported for the entire F plasmid transfer region when cloned into a multicopy plasmid as well as for the Tra1 core region of RP4 when fused with a multicopy ColD replicon (Johnson *et al*, 1980; Balzer *et al*, 1994). Another example is the shift of equilibrium of an F-TraI mediated *oriT* cleavage reaction toward the relaxed species when TraM is provided in trans from a multicopy plasmid (Fekete *et al*, 2000). This suggested that the increase in mobilization frequency was associated with the placement of the mobilization region within the high copy number vector pUC19.

Deletion of most of *repB* gene reduced the mobilization frequency by about 30-fold, indicating that in the MobA/RepB polyprotein the primase domain assisted, but was not essential for mobilization. Using the *mobE-mobA* (pMmob5) construct as a starting point, sequential deletion of the *mobE* (pMmob6), *mobED* (pMmob7) and *mobEDC* (pMmob8) genes was carried out. Deletion of *mobE* had no discernable effect on the mobilization frequency, while deletion of both *mobE* and *mobD* (pMmob7) reduced the mobilization frequency by approximately 600-fold, while there was no detectable mobilization of the *mobE-mobC* deletion (pMmob8). This is in contrast to pTF-FC2 where deletion of *mobE* reduced mobilization 150-fold with no mobilization detected on deletion of both *mobE* and *mobD* (Rohrer *et al*, 1992). Deleting most of *mobA* from pMmob5 (pMmob9) resulted in a construct with a mobilization frequency below the detection limit. This demonstrates that, although the MobA is related to TraI of RP4, the IncP1 DNA relaxase cannot complement the MobA deficiency to a level where mobilization can be detected (S17-1 donor).

The increase in mobilization frequency of the minimum mobilization region clones made the assay more sensitive allowing detection of mobilization and/or differences in mobilization that would normally fall below the detection limit. It was however not sensitive enough to detect possible mobilization of pMmob8 or pMmob9. This, together with not having deletion or mutant constructs targeted specifically at *mobA*, *mobB* or *mobC* (leaving other *mob* genes intact) does not allow us to conclude whether or not these are absolutely required for mobilization. However, as the MobA is almost certainly the DNA relaxase in this system, the protein would most likely be required for mobilization. We can also conclude that although the *mobD*, *mobE* and *repB* gene products enhance the mobilization frequency, they are not absolutely required for mobilization between *E. coli* strains, and when coupled to the natural replicon, the *mobC-mobA* construct would probably be mobilized albeit at a very low frequency.

Table 2.3.3.1: Mobilization frequencies of plasmids and constructs used in chapter 2

Test plasmid	Plasmid in *trans*	Mobilization[a] frequency of test plasmid
pTC-F14Cm		$2.8 \pm 1.8 \times 10^{-3}$
pTC-F14Cm[b]	R751	1.3×10^{-5}
pTC-F14Cm[b]	R388	$< 10^{-6}$
pMmob (F14, *mobEDCBArepBpasA*)		≥ 10
pMmob1 (F14, *mobEDCBArepB*)		≥ 10
pMmob5 (F14, *mobEDCBA*)		$3.3 \pm 3.1 \times 10^{-1}$
pMmob6 (F14, *mobDCBA*)		$1.2 \pm 1.0 \times 10^{-1}$
pMmob7 (F14, *mobCBA*)		$5.3 \pm 3.3 \times 10^{-3}$
pMmob8 (F14, *mobBA*)		$< 10^{-6}$
pMmob9 (F14, *mobEDCB*)		$< 10^{-6}$
pMmob8 (F14, *mobBA*)	pTC-F14Cm	2.3 ± 1.3
pMmob1184		$2.0 \pm 0.8 \times 10^{-1}$
pMmob1322		$2.3 \pm 1.6 \times 10^{-1}$
pOriTF14		$< 10^{-6}$
pOriTF14	pTC-F14Cm	1.4 ± 0.9

[a] mobilization frequency is the number of transconjugants per donor during a 60 min mating with a donor:recipient ratio of 1:10 using *E. coli* S17.1as donor and *E. coli* CSH56 as recipient. A mating frequency of ≥ 10 is indicated when the number of transconjugants equalled the number of recipients. Mating frequencies were the average of at least three independent experiments and standard deviations are indicated.

[b] *E. coli* HB101 was used as the donor strain.

2.3.4 Protein analysis

To determine whether or not polypeptides were produced from the predicted mobilization gene ORF's, as well as to ensure that they were being expressed in *E. coli* (pTC-F14 isolated from *At. caldus*), protein products from the deletions used in section 2.3.3 (pMmob1-pMmob8) were examined and compared with protein products from pMmob and pTC-F14Cm (Figure 2.3.4.1). This would also ensure that the differences in the mobilization frequencies observed in the previous section were not because of certain translation products being absent from the deletion constructs.
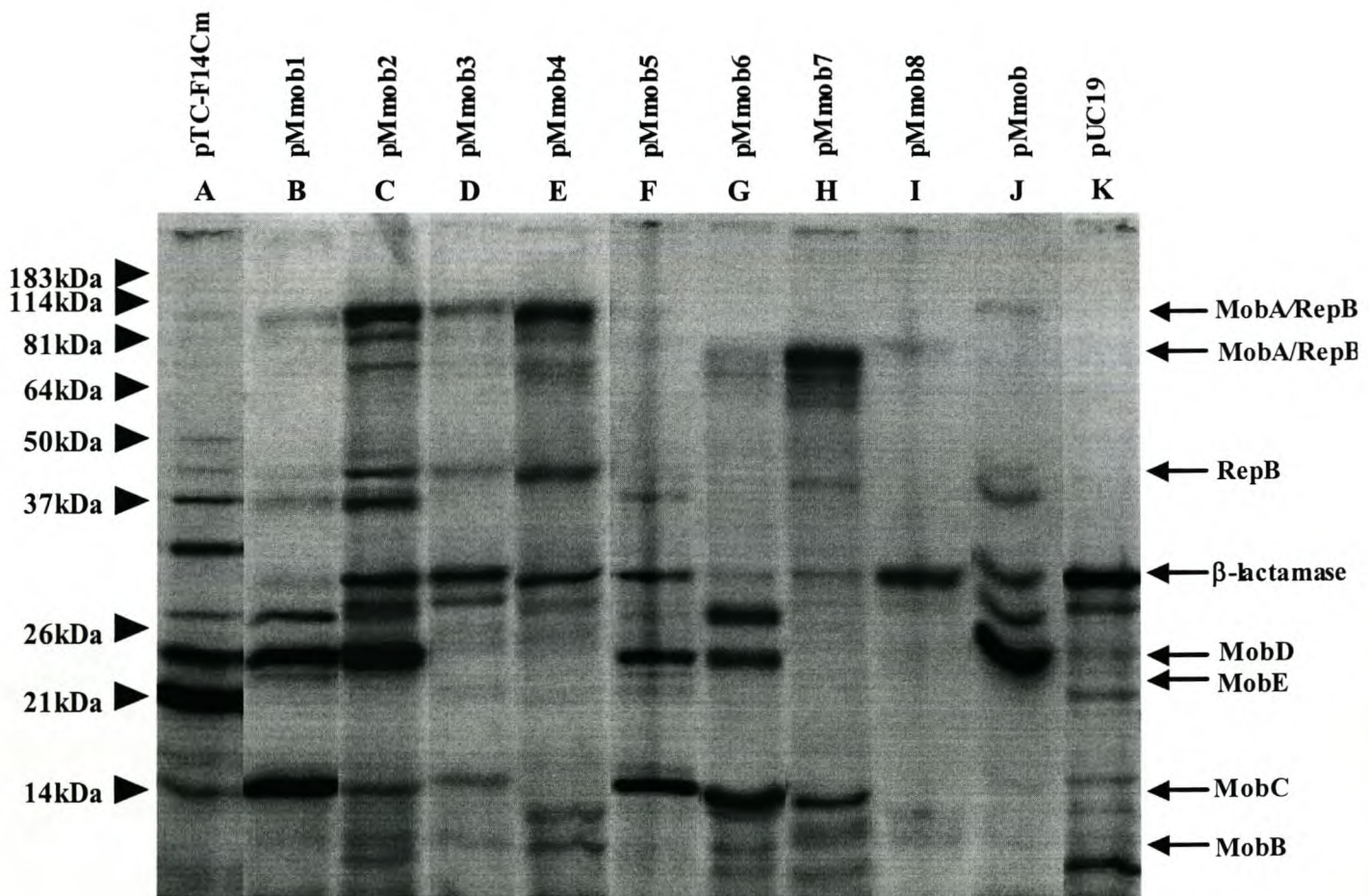
Figure 2.3.4.1: SDS-PAGE analysis of the proteins expressed from pTC-F14 (lane A), pMmob (lane J), the minimum mobilization region deletion series (lanes B-I), as well as the pUC19 vector control (lane K) by using an *E. coli*-derived, in vitro transcription-translation system.

A characteristic of all IncQ and IncQ-like plasmids is that the *mobA* and *repB* genes are fused is such a way that the MobA (nickase) and the RepB (primase) may be synthesized as separate proteins or as MobA/RepB fusion protein. All three polypeptides have been detected for plasmids RSF1010 (Scholtz *et al*, 1989) and pTF-FC2 (Dorrington *et al*, 1991; Rawlings personal communication). The difference between deletion set pMmob1-pMmob4 and pMmob5-pMmob8 is the removal of 262 amino acids from the C-terminal of the RepB primase in the latter. As the MobA and RepB may be synthesized as a fusion protein this would result in a reduction of the size of the predicted fusion protein and might also lead to a reduction in size or disappearance of the remainder of the RepB (N-terminal 90 amino acids) if still expressed as a separate protein. A protein corresponding to the predicted MW of the MobA/RepB fusion protein (±102.2 kDa) is observed in lanes A, J and B-E together with a ±40 kDa protein, which corresponds to the MW of the RepB primase (Table 2.3.2.2). Neither of these proteins is observed in deletions pMmob5-pMmob8 (lanes F-I), and the ±73 kDa protein observed in lanes G and H corresponds to the size of the truncated fusion protein (MobA-RepB'). The poorly expressed ±11.2 kDa protein seen in all samples corresponds to the predicted size of the MobB protein. Because of the uneven migration of the bands and a potential vector band in this region, the presence of the MobB cannot be certain. A ±14 kDa protein observed in lanes B, C and D as well as lanes F, G and H correlates with MobC as it is no longer present in pMmob4 and pMmob8, after deletion of the ORF thought to code for the protein, while also being present in lanes A and J. A protein corresponding to the predicted MW of MobD (±24.7 kDa) is seen in lanes B, C, F and G. The disappearance of the protein again coincides with the deletion of the *mobD* while also being present in lanes A and J. The MobE protein (±23.8 kDa) can be seen just below MobD in lanes B and F as well as a diffuse band in the same position in lanes J and A. The ±30 kDa signal (lanes B-J) corresponds to the β-lactamase complementing fragment while the band at ±28 kDa (lanes B-J) appears to originate from the vector (pUC19).

Carrying out Western blots using antibodies raised to the purified Mob proteins would be the method of choice for identifying proteins from a particular ORF, and should be considered in future studies to verify these results.

Chapter Three

Interaction of the pTC-F14 mobilization region with that of pTF-FC2

Contents

3.1 Introduction

Although great advances have been made in establishing evolutionary relationships between organisms, using tools such as comparing 16S-rRNA sequence, classifying plasmids has been and remains difficult. The primary reason for this is that there is no one shared element like the 16S-rRNA sequence, which exist among plasmids that could be used for this purpose. Another important factor that has hamstrung classification is a lack of information with relatively few plasmids having been fully sequenced.

Systems for assessing the relationships between plasmids have been proposed and include the following criteria: (i) Phenotypic characterization, (ii) Physical characterization, (iii) Incompatibility, (iv) Replicon typing, (v) Whole plasmid sequencing. The structure of plasmids can also be described as mosaic as they are often made up of different "modules" that can either have an essential function such as replication, maintenance or transfer, or they may be accessory functions as those that contribute to the phenotype of the host (Osborn *et al*, 2000). Not all plasmids carry transfer, maintenance or accessory functions and the absolute minimum requirement for a plasmid is the **ability** to replicate, and is the reason why plasmid replicons have been extensively used for classification. This minimalist approach essentially only describes evolutionary relationships between such replicons and does not allow for inclusion of other plasmid systems and the events, which result in the acquisition, or movement of the accessory genes. When comparing two closely related plasmids, a useful tool in establishing the evolutionary relationship between them could be to determine whether related plasmid systems other than the replicons can still interact. This may provide a rough guide as to how long the respective plasmids have been separated and may also indicate which adaptations result in a fitter plasmid. These changes are thought to occur through mechanisms such as point mutations, insertions and deletions, rearrangements such as inversion and translocation, insertion and excision of mobile genetic elements and co-integrate formation. Evolutionary events that result in the origin of a new incompatibility (Inc) group are termed macroevolution, while microevolution refers to events that do not change the Inc group.

The IncQ group of plasmids are thought to be highly evolved because of the compactness of the essential genetic information, characterized by overlapping genes and by gene products that exert multiple functions in mobilization, replication and regulation. As these plasmids are highly mobilizable and have a broad host range, they would be expected to have been exposed and adapted to a variety of cell environments as well as other plasmids with which they would have to compete for not only replication space but also perhaps for access to a conjugative system. An example of the latter is the reduced virulence observed in *A. tumefaciens* strains harboring the pTiC58 plasmid with an RSF1010 plasmid co-resident. Although this phenomenon was first attributed to the conjugal intermediate of RSF1010 competing with and thus interfering with the VirB dependent export of the VirE2 protein from *A. tumefaciens*, recent studies suggest that this inhibition is not due to the reduction of VirE2 export but perhaps by influencing a later step in the transport process or causing a more general inhibition not restricted to any specific step (Binns *et al*, 1995; Stahl *et al*, 1998; Bravo-Angel *et al*, 1999; Chen *et al*, 2000). Another level at which competition exists for the Mpf system is the affinity that the coupling protein exhibits for a particular relaxosome component(s) (Chapter 1 section 1.4).

The proteins involved with DNA metabolism at the *oriT* are frequently plasmid-specific. For example, despite the high degree of similarity between the DNA-processing transfer proteins and the *oriT* regions of the IncP plasmids RP4/RK2 and R751, the *oriT* of RP4/RK2 cannot be transferred by R751 (Fürste *et al*, 1989). Transfer of the RP4/RK2 *oriT* took place only when the specific *traJ* and *traK* genes of RP4/RK2 were present with *traI* also being required although this was not plasmid specific. Plasmid RP4 TraJ (Ziegelin *et al*, 1989) and TraK (Ziegelin *et al*, 1992) proteins bind specifically to different features of the *oriT* region of RP4 but not R751. There are, however examples of interaction of the Dtr functions of closely related plasmids with each other's *oriT*'s. The *oriT* of pTiC58 is mobilized by genetically distinct *Agrobacterium* plasmids pAtK84b, pTiT37 and pTiI5955 and these interactions have been used to support the claim that these plasmids share a common ancestor (Cook *et al*, 1992). The inability of plasmids pTiC58 and RSF1010 to cross complement showed that although the *oriT* homologies suggest that these transfer

71

systems share a common ancestor they have diverged from each other to the point that they are no longer cross functional.

As pTC-F14 and pTF-FC2 are IncQ-like promiscuous plasmids that were isolated from acidiphilic, iron- and/or sulfur-oxidizing, chemolithotrophic bacteria that share a similar habitat, it is not unlikely that the plasmids may come into contact with each other. Plasmids pTC-F14 and pTF-FC2 have diverged sufficiently for their replicons to be compatible, which should allow them to coexist in the same host cell (Gardner *et al*, 2001; Gardner *et al*, 2003). This raised questions such as, have the *mob* genes diverged sufficiently to be plasmid-specific or will they complement the mobilization activity of each other? Was there competition between plasmids at the level of mobilization? That is, had one of the plasmids evolved a mobilization system that would allow it to dominate the horizontal transfer process, thereby giving it a selective advantage in preference to the other or actively hinder transfer of the other? Here we report on the ability of the mobilization systems of pTC-F14 and pTF-FC2 to interact with each other.

3.2 Methods

The media and growth conditions, DNA manipulations and mating procedures used in this Chapter are the same as those described in Chapter 2. The bacterial strains used are listed in Appendix 1 and the plasmid constructs used in this Chapter are listed in Appendix 2.

3.2.1 PCR of pTF-FC2 *mobC*, *mobD* and *mobE* deletions

To generate pmobE, pmobDE and pmobCDE, primers mobER2, mobEF2, mobDEF2 and mobCDEF2, listed in Appendix 3, were used with pDER412 as template. After an initial denaturation of 60 s at 94°C, 25 cycles of 30 s at 55-60°C (depending on primer set) and an elongation step of up to 4 min at 72 °C were performed. A final extension step of 120 seconds at 72 °C before cooling to 4 °C completed the reaction. These PCR products were cloned into the pKK223-3 expression vector with an IPTG inducible *tac* promoter to ensure expression (Figure 3.3.1.1). The PCR products were sequenced from both ends to ensure that all necessary ORF's were included and that the gene fusions were accurate.

3.3 Results

3.3.1 Comparison of the mobilization efficiencies and interaction between the mobilization systems of pTC-F14 and pTF-FC2

The mobilization frequencies of plasmids containing the *mob* genes of pTC-F14 and pTF-FC2, when mobilized by an IncPα plasmid and associated with their natural replicons, were compared. A selectable kanamycin resistance marker was cloned into plasmid pTC-F14 to produce plasmid pTC-F14Km. This plasmid was mobilized from *E. coli* S17-1 donor cells to CSH56 recipient cells at a frequency of 2.83 X 10^{-3} transconjugants per donor, which was similar to that of pTC-F14Cm. However, this frequency was more than 3000-fold less than plasmid, pDER412 that contained the pTF-FC2 mobilization genes (see Table 3.2.2.1). To test whether mobilization of one plasmid was affected by co-residence of the other, both pTC-F14Km and pTF-FC2 (pDER412) were placed into *E. coli* S17-1 cells and the frequency of transfer was measured. The frequency of mobilization of pTC-F14Km was enhanced almost to saturation in the presence of pTF-FC2 while the presence of pTC-F14 had no discernable effect on the mobilization of pTC-FC2. To determine what property of pTF-FC2 was required for this mobilization frequency enhancement, plasmid constructs containing combinations of pTF-FC2 *mob* genes, subcloned into the vector pACYC184 were introduced into *E. coli* S17-1 (pTC-F14Km) cells. Co-resident plasmids, pAC221 (containing pTF-FC2 *mobA* and *mobB*), and pAC209 (containing *mobA*, *mobB*, *mobC*, *mobD* and a truncated *mobE*) did not increase the frequency of mobilization. In contrast, pAC105, which contained a *mobC*, *mobD* and a complete *mobE*, enhanced the mobilization frequency of pTC-F14 by about 100-fold, though this was about 10-fold less than when the whole of pDER412 was present. To determine whether this result was due to the *mobE* of pTF-FC2, the gene was amplified by PCR and cloned behind the IPTG inducible *tac* promoter in plasmid pKK223-3 (construct pmobE). This construct did not improve the mobilization frequency of pTC-F14Cm. When a combination of pTF-FC2 *mobD* and *mobE* genes (pmobDE) were placed *in trans* and induced with IPTG, mobilization of pTC-F14Cm reached saturation. IPTG induction of a combination of the *mobCDE* genes (pmobCDE) enhanced pTC-F14Cm mobilization by approximately 150-fold to about

the same level as with pAC105. This indicated that it was the combination of *mobD* and *mobE* from pTF-FC2 that enhanced pTC-F14 mobilization.
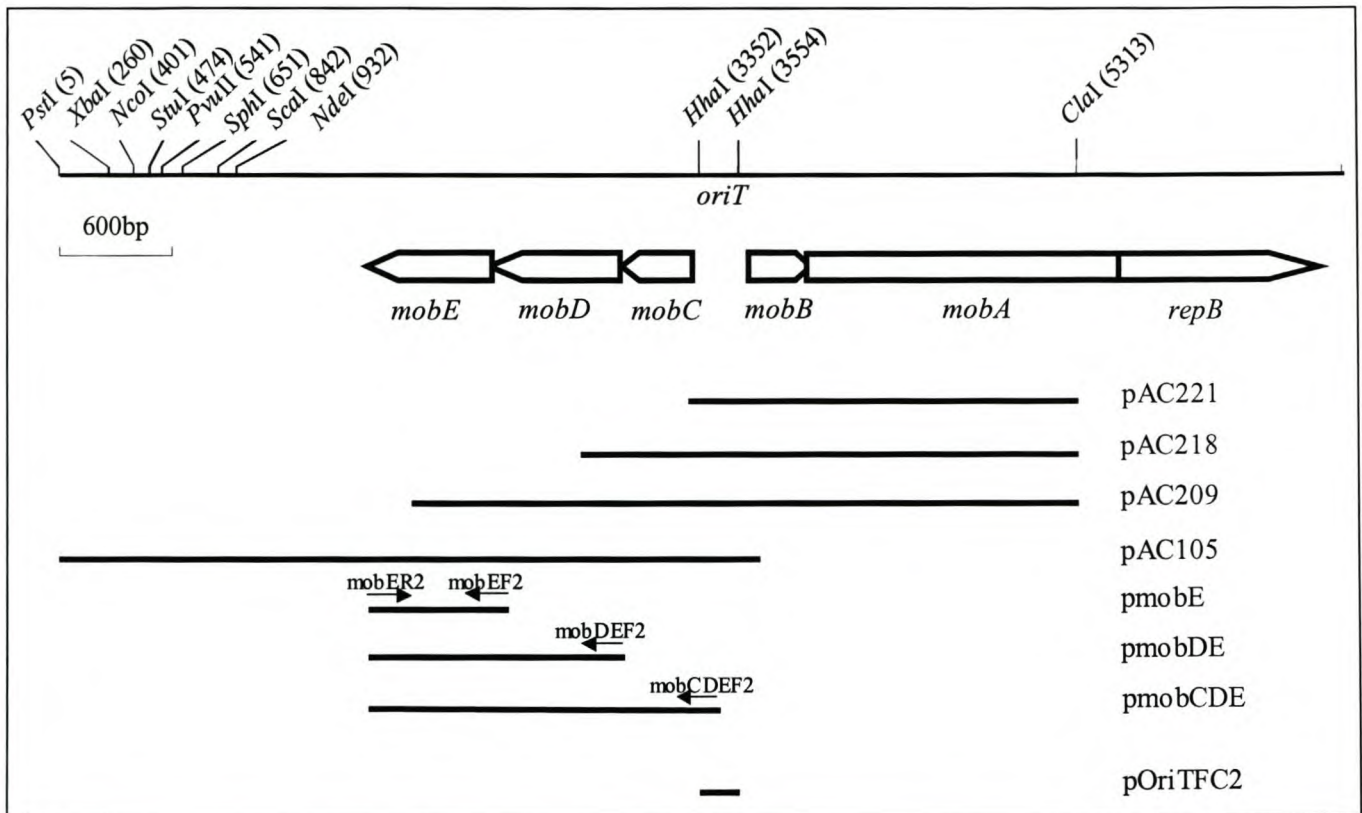
Figure 3.3.1.1: The genetic and restriction endonuclease cleavage map as well as the previously reported *mob* region subclones (pAC series) used in this study of the mobilization region of plasmid pTF-FC2 (Rohrer *et al*, 1993). Short horizontal arrows show the positions of the primers used to amplify and construct certain subclones.

3.3.2 Interaction at the *oriT* regions

To test whether the mobilization proteins of the two plasmids could act specifically on the *oriT* regions of each other, plasmids containing one of the cloned *oriT* fragments pOriT-F14 (203 bp *Hind*III-*Nco*I fragment Appendix 2) and pOriT-FC2 (208 bp *Hha*I-*Hha*I fragment Appendix 2) were transformed into *E. coli* S17-1 containing either pTC-F14Cm or pDER412. Both cloned *oriT* regions were functional as they were mobilized by their respective parent plasmids at a frequency comparable to, or greater than, the parent plasmid (Table 3.3.2.1). The construct containing *oriT* region of pTC-F14 (pOriT-F14) was mobilized at a frequency of 1.48 transconjugants per donor when pTC-F14 was placed *in trans*, but only at 3.48×10^{-2} with pDER412 *in*

trans. With both pTC-F14 and pDER412 *in trans*, the mobilization frequency of pOriT-F14 reached saturation.

In the pTF-FC2 *oriT* complementation experiments, pDER412 was able to mobilize a construct containing its own *oriT* (pOriT-FC2) at a saturation frequency while mobilization by pTC-F14 was below the limit of detection. Complementation of the *oriT* regions was therefore thought to be unidirectional with pTF-FC2 able to mobilize the *oriT* of pTC-F14, but not *vice versa*. We tested whether there was any detectable interaction between the *mob* genes of pTC-F14 and the *oriT* of pTF-FC2, by providing two subsets of the *mob* genes of pTF-FC2 *in trans*. Plasmid pTC-F14Km was able to mobilize pOriT-FC2 at a frequency of 3.54×10^{-1} when in the presence of the pTF-FC2 *mobCDE* genes (pAC105) and at a frequency of 1.67×10^{-2} when the *mobAB* genes (pAC221) were present. This result was surprising and suggested that at least one of the *mobCDE* as well as one of the *mobAB* gene products of pTF-FC2 gene products is able to independently assist in the recognition of the heterologous pTF-FC2 *oriT* by pTC-F14.

Table 3.3.2.1: Mobilization frequencies of plasmids and constructs used in chapter 3

Test plasmid	Plasmid(s) in *trans*	Mobilization[a] frequency of test plasmid
pTC-F14Km		$2.7 \pm 1.5 \times 10^{-3}$
pDER412		≥ 10
pDER412	pTC-F14Km	≥ 10
pTC-F14Km	pDER412	8.5 ± 0.5
pTC-F14Cm	pAC105 (FC2, *mobEDC*)	$7.9 \pm 2.1 \times 10^{-1}$
pTC-F14Cm	pAC209 (FC2, *mobDCBA*)	$2.3 \pm 3.5 \times 10^{-3}$
pTC-F14Cm	pAC221 (FC2, *mobBA*)	$8.9 \pm 2.4 \times 10^{-4}$
pTC-F14Cm	pmobE (FC2)	$4.4 \pm 2.1 \times 10^{-3}$
pTC-F14Cm	pmobDE (FC2)	≥ 10
pTC-F14Cm	pmobCDE (FC2)	$4.7 \pm 2.9 \times 10^{-1}$
pTC-F14Cm	pAC218 (FC2 *mobC*) + pmobE (FC2)	$1.1 \pm 2.6 \times 10^{-3}$
pOriTF14		$< 10^{-6}$
pOriTF14	pTC-F14Cm	1.5 ± 1.0
pOriTF14	pDER412	$3.5 \pm 0.04 \times 10^{-2}$
pOriTF14	pTC-F14Km + pDER412	≥ 10
pOriTFC2		$< 10^{-6}$
pOriTFC2	pTC-F14Cm	$< 10^{-6}$
pOriTFC2	pDER412	≥ 10
pOriTFC2	pTC-F14Km + pAC105 (FC2, *mobEDC*)	$3.5 \pm 4.7 \times 10^{-1}$
pOriTFC2	pTC-F14Km + pAC221 (FC2, *mobBA*)	$1.7 \pm 0.8 \times 10^{-2}$

[a] mobilization frequency is the number of transconjugants per donor during a 60 min mating with a donor:recipient ratio of 1:10 using *E. coli* S17-1 as donor and *E. coli* CSH56 as recipient. A mating frequency of ≥10 is indicated when the number of transconjugants equaled the number of recipients. Mating frequencies were the average of at least three independent experiments and standard deviations are indicated.

3.4 Discussion

The discovery that when pTF-FC2 was co-resident with pTC-F14, the mobilization of the latter plasmid was increased by about 3000-fold was unexpected. We further discovered that the presence of the combination of the pTF-FC2 *mobD* and *mobE* genes, but not the individual *mobD* and *mobE* genes, was responsible for this increase. This suggests that the apparently dispensable pTC-F14 *mobE* gene does play a role in mobilization but that the pTC-F14 *mobE* gene is not optimally functional in the IncPα plasmid based mating system used (see below). The functions of MobD and MobE proteins are unknown and the same applies to the related TraL and TraM proteins of the IncPα and IncPβ plasmids. TraL has been found to have an ATP- or GTP-binding Walker A box (Thorsted *et al*, 1998) and this box is present and highly conserved in the MobD proteins of the IncQ-like plasmids (Chapter 2 section 2.3.2). The role of MobD and MobE in facilitating the mobilization of one plasmid by another found in this current study emphasizes the need to discover the function of the proteins.

In the present study plasmid pTF-FC2 was clearly much more readily mobilized between *E. coli* strains than plasmid pTC-F14. Furthermore, a co-resident pTF-FC2 could mobilize a plasmid containing the *oriT* of pTC-F14 (though not as efficiently as pTC-F14), while a co-resident pTC-F14 could not mobilize a plasmid containing the *oriT* of pTF-FC2. Based on these results plasmid pTF-FC2 might be expected to be a more promiscuous plasmid than pTC-F14. However, it must be taken into account that mobilization studies were carried out between *E. coli* strains using the chromosomally located IncP plasmid RP4 as a conjugative helper plasmid. It is possible that pTF-FC2 is more suited to mobilization by this plasmid than pTC-F14, while there may be an as yet unknown helper plasmid that mobilizes pTC-F14 better than pTF-FC2. The reason for the unexpected observation that the *mobD* and *mobE* genes of pTF-FC2 were better able to assist pTC-F14 mobilization than its own genes,

could be because the MobD and MobE proteins of pTF-FC2 are better suited to work with RP4, while the equivalent proteins of pTC-F14 may be better suited to function with a different conjugative plasmid. The pRAS3.1 mobilization region, which seems more similar to that of pTF-FC2 than pTC-F14 suggests that although all three plasmids shared a common ancestor the pTC-F14 mobilization region may have experienced selection pressure to change (Chapter 2).

The interpretation of experiments on the ability of plasmids containing the cloned *oriT* regions to be mobilized by the *mob* genes of the other plasmid is not fully clear. The *oriT* of pTC-F14 could be mobilized by its own *mob* proteins and this mobilization frequency was enhanced in the presence of pTF-FC2. This result was consistent with the ability of pTF-FC2 to enhance the mobilization frequency of pTC-F14. Plasmid pTC-F14 was not able to mobilize a plasmid containing the *oriT* of pTF-FC2 unless some of the pTF-FC2 genes were present. What was surprising is that when we attempted to determine which of the pTF-FC2 *mobAB* genes or the *mobCDE* genes were required, we found that either set of genes partly enhanced mobilization. A possible explanation is that more than one of the products of the pTF-FC2 *mob* genes is likely to enhance binding of the pTC-F14 mobilization complex to the *oriT* of pTF-FC2. In the case of pTF-FC2, it is likely that MobB and MobC of pTF-FC2 could possibly bind to its own *oriT* and thereby assist the otherwise *oriT* specific proteins of pTC-F14 to recognize the *oriT* of pTF-FC2. These results also demonstrate that the MobA protein from pTC-F14 can functionally interact with the pTF-FC2 *oriT*. Assuming MobB performs the same function as TraJ in RP4, the MobB of pTC-F14 could recognize the inverted repeat at the *oriT* of pTF-FC2. Rohrer and coworkers had shown that the presence of only MobA and MobB of pTF-FC2 were not sufficient for mobilization thus the pTC-F14 MobC, MobD and/or MobE may interact successfully, via protein-DNA or protein-protein interactions, with the MobA and/or MobB of pTF-FC2 to form a functional relaxosome (Rohrer *et al*, 1992).

Part of the motivation for this study was to gain an understanding of the evolution of mobilization systems. The sequence similarity between the proteins associated with plasmid replication and mobilization, clearly indicate that plasmids pTC-F14 and pTF-FC2 share a common ancestor. It has been reported that plasmids pTC-F14 and

pTF-FC2 are compatible in *E. coli* and this indicates that replicons of the plasmids have diverged sufficiently for them to function as independent units. Pressure for the replicons to diverge may have arisen because the two broad host range, promiscuous plasmids occur in bacteria that share a similar ecological niche. This means they may have frequently encountered each other and divergence to the point of compatibility would mean that the plasmids will not exclude each other from the same host cell and thereby each would have an increased 'replication space'. It was of interest to discover whether mobilization systems of these related plasmids would compete with each other. If one plasmid has a more dominant mobilization system it would presumably be transferred horizontally to more host cells than the competing plasmid and thereby dominate an ecosystem. No reduction in mobilization frequency of one plasmid when co-resident with the other plasmid was detected. In contrast, a co-resident pTF-FC2 appeared to assist the mobilization of pTC-F14. Similarly, the cloned *oriT* region of pTF-FC2 could be mobilized by a co-resident pTC-F14, when some but not all of the pTF-FC2 *mob* genes were present.

Chapter Four

General Conclusions

Plasmid pTC-F14 is the third member of the five-*mob* gene IncQ-like plasmid family to have its mobilization system investigated. Two other members of this IncQ-like plasmid family, now designated IncQ-2-group, are pTF-FC2 and pRAS3.1, although no report on the biology of pRAS3.1 mobilization has been published. The observation that the amino acid sequence relationship between the mobilization proteins of plasmid pTF-FC2 and pRAS3.1 is much closer than between pTF-FC2 and pTC-F14 is remarkable (Figure 2.3.2.3). The implication is that all three plasmids shared the same common ancestor but that pTF-FC2 and pRAS3.1 either diverged more recently than pTF-FC2 and pTC-F14 or that pTC-F14 has experienced selection pressure to change. Since divergence, pTF-FC2 and pRAS3.1 are now found in bacteria as different as the obligately acidophilic chemolithotrophic *At. ferrooxidans* strain FC isolated in South Africa and the neutrophilic, heterotrophic *A. salmonocida* strain isolated in Norway. This serves to illustrate the highly promiscuous nature of the IncQ plasmid family. The 32.7 kb plasmid, pRA2, is another example of a plasmid containing a set of five *mob* genes that are related to the Tra1 system of the IncP plasmids (Kwong *et al*, 2000). However, plasmid pRA2 has a unique replicon, with no similarity to those of the IncQ-like plasmids, and this suggests that the five *mob* gene system is a mobilization module that may also be acquired by different, otherwise unrelated plasmids.

Although a novel mobilization system was not found, analysis of the pTC-F14 mobilization region does add to our growing knowledge of plasmid mobilization systems. An interesting discovery was that unlike the MobA and MobB proteins, the amino acid sequences of the MobC, MobD and MobE proteins of pTF-FC2 and pTC-F14 were rather unrelated to each other. Furthermore, although deletion of the *mobD* and *mobE* genes of pTC-F14 did not reduce the mobilization frequency of pTC-F14 by RP4, if the *mobD* and *mobE* genes of pTF-FC2 were provided in *trans*, the mobilization frequency of pTC-F14 increased by about 3000-fold. This indicated that the MobD and MobE of pTF-FC2, but not the MobD and MobE of pTC-F14 played

an important role in optimising mobilization of pTC-F14 by the IncPα conjugative system (Chapter 3). Since the heterologous MobD and MobE proteins were more effective at increasing the frequency of mobilization than the homologous proteins, this suggests that the MobD and MobE proteins of pTF-FC2 are more suited to functioning with the conjugative plasmid RP4 than the pTC-F14 proteins. The functions of the MobD and MobE proteins are unknown as are the equivalent proteins in RP4 (TraL and TraM). This current study has emphasized the need to determine what the functions of these proteins are.

The addition of the sequence to plasmid databases would also allow us to, in future, get the best approximation of the evolutionary history of these plasmids (Chapter 3). Analysis of this mobilization region is also not only of interest from a fundamental point of view, but conjugation may prove more effective in transferring modified DNA molecules across the *Acidithiobacillus caldus* membrane, than conventional transformation techniques. This has already been demonstrated for *Acidithiobacillus ferrooxidans*, *Acidithiobacillus thiooxidans* and *Acidithiobacillus neapolitanus* (Liu *et al*, 2000; Jin *et al*, 1992; Kulpa *et al*, 1983). A better understanding of the workings of the mobilization region could help in development of more efficient cloning vectors.

Although we currently do not know whether pTF-FC2 and pTC-F14 ever occur in the same host cell in their natural environment, if they were to do so, we know that both their replication and mobilization systems have developed so as not to compete. This raises the question of which factors have driven the evolution of these two plasmids? If they were to occur in the same cell there would be pressure for their replicons to diverge, so as not to be incompatible with each other, but this would not necessarily apply to their mobilization systems. As both pTC-F14 and pTF-FC2 are mobilizable they need a conjugative system for plasmid transfer to take place. In the case of pTF-FC2 a ±30 kb plasmid was identified in the same strain (FC) that this plasmid was isolated from, while a ±45 kb plasmid was identified together with pTC-F14 in the F strain from which it was isolated. Although speculative, it is possible that these larger plasmids may encode conjugative systems to which the respective mobilizable plasmids have adapted and which has resulted in sequence divergence especially in

the *mobD* and *mobE* genes (also *mobC*). This scenario would be plausible if the mobilizable plasmids were found in separate hosts most of the time. While if the plasmids frequently encountered each other, they would presumably also frequently encounter the same conjugative plasmid and would therefore be under pressure to both adapt to similar conjugative plasmids. The speculation as to whether the plasmids frequently encounter each other and evolve accordingly is inconclusive. One may interpret the observation that the plasmid replicons are compatible as an indication that they frequently encounter each other in the same host cell, but the observation that they are not equally tuned to mobilization by the same conjugative plasmid as evidence that they do not encounter the same group of conjugative plasmids. What would help to strengthen the speculation that each mobilizable plasmid is optimally tuned to a specific conjugative plasmid is the discovery of a conjugative plasmid that mobilizes pTC-F14 more efficiently than pTF-FC2. The 30 kb and 45 kb plasmids found within *At. ferrooxidans* and *At. caldus* are of a comparable size to the smallest known conjugative plasmids. For example, plasmids belonging to the *E. coli* IncW group are within the 30-40 kbp range. Possibly the 45 kb plasmid resident in the *At. caldus* 'f' strain is an example of such a plasmid. Attempts are being made to isolate this plasmid.

Future studies would need to determine the role of each of the pTC-F14 mobilization gene products in relaxosome formation but perhaps also in interfacing with the coupling protein or stabilizing the relaxosome at the cell membrane through interaction with the Mpf components. This would also confirm that they indeed perform an equivalent function to their RP4 counterparts. These protein-protein interactions could be examined using the two-hybrid system used so effectively for elucidating the interactions of some of the Mpf components from the *Agrobacterium tumefaciens* Ti plasmid system (Chapter 1). As the interaction with the coupling protein is a required step in the conjugation process, the *mobD* and *mobE* gene products, which were not required for pTC-F14 mobilization, could probably be ruled out as the coupling protein interface with the relaxosome (Chapter 1 section 1.4). They may serve to stabilize the relaxosome by possibly binding to some of the other relaxosome components similar to TraH from plasmid RP4 or possibly interact with Mpf components helping to anchor the relaxosome at the cell membrane thus ensuring more efficient transfer. Also important is the regulation of the respective mobilization

genes not only as a separate mobilization module but also in concert with the replication genes as the existence of a MobA-RepB fusion suggests that regulation of the replication and mobilization is linked (Gardner *et al*, 2003).

Appendix 1

Bacterial strains and plasmid cloning vectors used in this study

Strain or plasmid	Genotype or description	Reference/Source
Strain		
DH5α	F′/endA1 hsdR17 ($r_K^-m_K^+$) supE44 thi-1 reacA1 gyrA (Nal^r) relA1 Δ(lacZYA-argF)U169 (φ80dlacΔ(lacZ)M15)	Promega Corp., Madison, WI
S17-1	recA pro hsdR (RP4-2 Tc::Mu Km::Tn7)	Simon et al, 1983
CSH56	F$^-$ ara Δ(lacpro) supD nalA thi	Cold Spring Harbor, NY
HB101	F$^-$ Δ(mcrC-mrr) hsdS20 recA13 ara-14 proA2 lacY1 λ$^-$ galK2 rpsL20(Sm^r) Xyl-5 mtl-1 supE44	Boyer et al, 1969
Plasmids		
pUC19	Amp^r lacZ′, ColE1 replicon, cloning vector	Yanisch-Perron et al, 1985
pACYC184	Tc^r Cm^r, p15A replicon, cloning vector	Chang et al, 1978
pBR322	Amp^r Tc^r, ColE1 replicon, cloning vector	Bolivar et al, 1977
pKK223-3	Amp^r , ColE1 replicon, expression vector	Brosius et al, 1984

Appendix 2

Plasmid constructs containing portions of pTC-F14 or pTF-FC2, generated or used during this study

Construct	Genotype or description	Reference/Source
pTC-F14Cm	Cm^r, natural pTC-F14 plasmid with a chloramphenicol resistance gene inserted at the single *Bam*HI site	Gardner *et al*, 2001
pTC-F14Km	Km^r, pTC-F14Cm with the chloramphenicol resistance gene replaced by the kanamycin resistance gene from Tn*5*	This study
pDER412	Cm^r, natural pTF-FC2 plasmid with chloramphenicol resistance gene cloned into the Tn*5467* transposon	Rawlings *et al*, 1984
pMmob	Amp^r, 5554bp *Bam*HI-*Xba*I fragment of pTC-F14 containing all mobilization genes and the *repB* primase, cloned into pUC19	This study
pMmob1-pMmob9	Amp^r, PCR based deletions of the mobilization region of pTC-F14 cloned into pUC19. Refer to Fig. 1.	This study
pMmob1184	Cm^r, minimum mobilization region one of pTC-F14 cloned into the tetracycline resistance marker of pACYC184	This study
pMmob1322	Amp^r, minimum mobilization region one of pTC-F14 cloned into pBR322	This study
pmobE	Amp^r, PCR product of *mobE* gene of pTF-FC2 cloned into pKK223-3	This study
pmobDE	Amp^r, PCR product of *mobDE* genes of pTF-FC2 cloned into pKK223-3	This study
pmobCDE	Amp^r, PCR product of *mobCDE* genes of pTF-FC2 cloned into pKK223-3	This study
pAC105	Cm^r, exonuclease III shortening of the pTF-FC2 mobilization region containing *mobC*, *D* and *E* cloned into pACYC184	Rohrer *et al*, 1993
pAC209	Cm^r, exonuclease III shortening of the pTF-FC2 mobilization region containing *mobA*, *B*, *C* and *D* and a truncated *mobE* cloned into pACYC184	Rohrer *et al*, 1993
pAC218	Cm^r, exonuclease III shortening of the pTF-FC2 mobilization region containing *mobA*, *B* and *C* with *mobD* and *E* removed also in pACYC184	Rohrer *et al*, 1993
pAC221	Cm^r, exonuclease III shortening of the pDER412 mobilization region containing *mobA*, *B* and the *oriT* cloned into pACYC184	Rohrer *et al*, 1993
pOriTF14	Amp^r, a 203 bp *Hind*III-*Nco*I fragment of pTC-F14 containing the *oriT* cloned into pUC19	This study
pOriTFC2	Amp^r, a 208 bp *Hha*I-*Hha*I fragment of pTF-FC2 containing the oriT cloned into pUC19	This study

Appendix 3

Nucleotide sequence of PCR primers used in this study

Name	Sequence	Reference/Source
mobEF2	(*Eco*RI)5´-TACAGAATTCAGCAAGCGCATGAGC-3´	This study
mobDEF2	(*Eco*RI)5´-TACAGAATTCCCAAAACCCGACAGC-3´	This study
mobCDEF2	(*Eco*RI)5´-TATAGAATTCCCACGTGGCGAAGCC-3´	This study
mobER2	(*Xba*I) 5´-TACATCTAGAATGTTGAGCGCGTCG-3´	This study
mobAR14	(*Eco*RI) 5´-TACAGAATTCCGGTCCATGTCGTCG-3´	This study
repBR14	(*Eco*RI) 5´-TACAGAATTCGGGTAATCGGATGGC-3´	This study
mobC´R14	(*Pst*I) 5´-TATACTGCAGCTTTCCCGCCTTTGC-3´	This study
mobCR14	(*Pst*I) 5´-TATACTGCAGTTGCCACCACCGACG-3´	This study
mobDR14	(*Pst*I) 5´-TATACTGCAGTCGGGTGTCGGTTCC-3´	This study
mobER14	(*Pst*I) 5´-TACACTGCAGCTGTCCGAAAGTAGG-3´	This study
mobAR14#2	5´-TGGCGTCGCTTGTTTGGTTC-3´	Gardner *et al*, 2003

Appendix 4

Plasmid-map and complete sequence of pTC-F14 (Accession number AF325537)
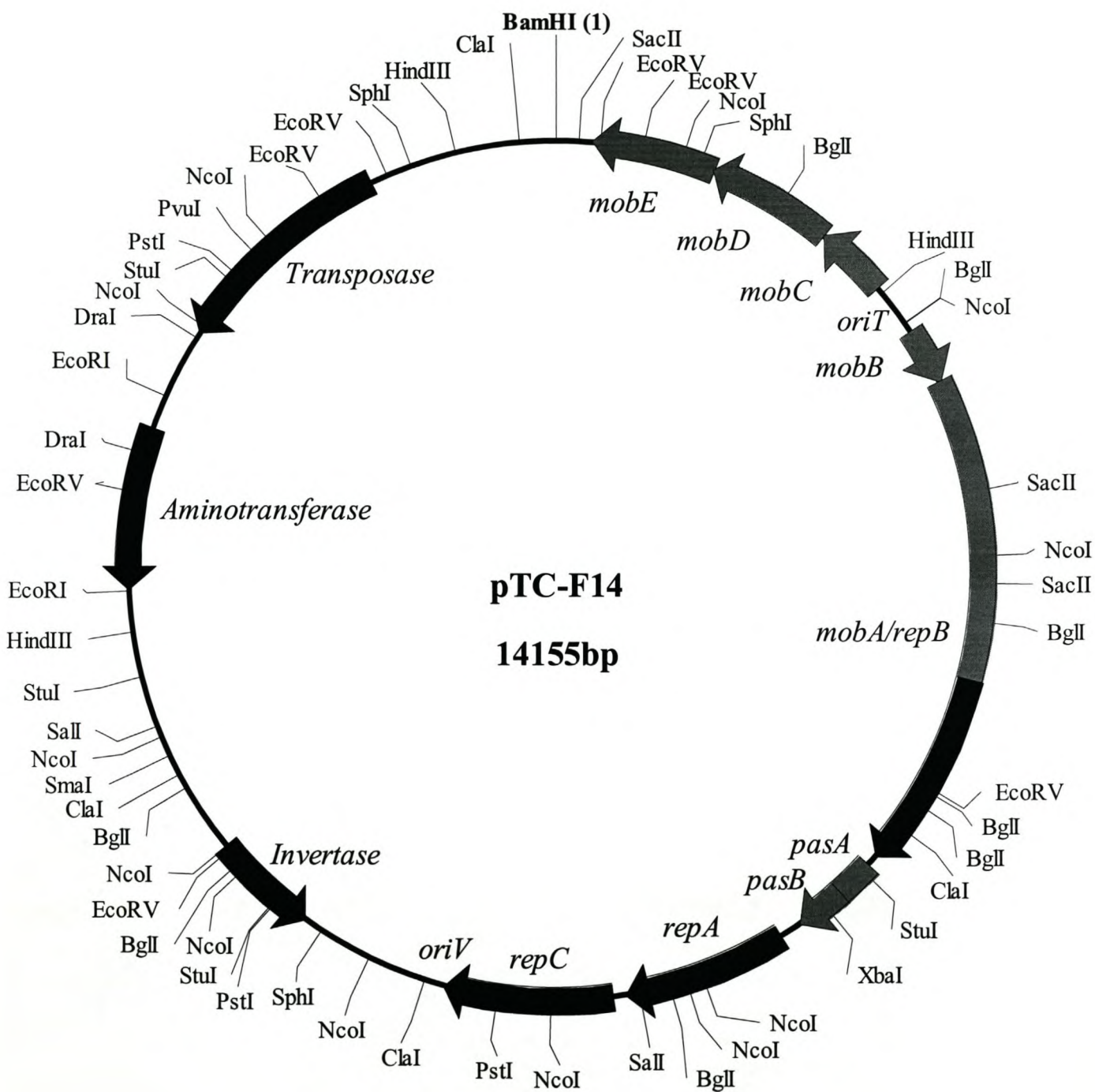
Figure A: Circular map of pTC-F14 compiled by L. van Zyl, M. Gardner and G. Goldschmidt. The unique BamHI site is used as bp=1.

Appendix 4: (cont.)

Annotation of the complete pTC-F14 sequence, using the unique BamHI as bp=1

```
Universal code
           BamHI
1          GGATCCTTGTGGAGAAATTGTCAGACAGGGCAGTTGCAAGATCGAGGCGCACCCACTGTC


61         CGAAAGTAGGCCAAGCGCCTTCTGGACAGGAGCAGAGATCGCACCTCTGGGGCGGGGCGG


121        GCCGCGGGGTGTGTCTGCAAGGGCGAAGCCCGCGCAGCGTAGCGAAGTTTTTGCGGGACT


181        TGTCCGCCCAATTTGTTCACTGATGCTGCTTTGTCTCTGCCTGCTCGATGGCTCCGATAT
                     *  Q  H  Q  K  T  E  A  Q  E  I  A  G  I  D

241        CCAGGCGGCCATAGTATACGGTTTTCCCATCGTGTTTCTCGTATCCAAAGTTATTTGCTA
            L  R  G  Y  Y  V  T  K  G  D  H  K  E  Y  G  F  N  N  A  V

301        CGTTGGCTGCCACATCAATGGGTACGCAAGGGCCAGACTGGATGCCGTCGAACGCAGGGC
            N  A  A  V  D  I  P  V  C  P  G  S  Q  I  G  D  F  A  P  C

361        ACTCTTGGACAATACCACCGATGCGCGTTTGGATGGCCGCTTCCTCGGCCTTGAGACGTT
            E  Q  V  I  G  G  I  R  T  Q  I  A  A  E  E  A  K  L  R  Q

421        GCAGCTTTTGTTCGACCTGCCGGACTTCTTTTGAGTGCCACCAGATATCGACCTCTCCGG
            L  K  Q  E  V  Q  R  V  E  K  S  H  W  W  I  D  V  E  G  A

481        CAGCGTATGACCCAATCCCCACGACTACAAGCACCAGCCCGGCCCAGACGGATCTCGACC
            A  Y  S  G  I  G  V  V  V  V  L  V  L  G  A  W  V  S  R  S

541        AAAGCGTTTTTCCCAAGTCACTGCGGAAGGCATCGACGAGCTGCTGTCCAGCTTTGGACA
            W  L  T  K  G  L  D  S  R  F  A  D  V  L  Q  Q  G  A  K  S

601        GCAGTTCTGATGACGCGGCGGAGATTCTTCCCTTGGCTGCTTCTTCGGCGTTACTGGCGA
            L  L  E  S  S  A  A  S  I  R  G  K  A  A  E  E  A  N  S  A

661        TTTTATTTGCCGTGTTCTCCATGGCCTTGGGCAGTTTGCTGAACGTCCCATAGTAGGCAT
            I  K  N  A  T  N  E  M  A  K  P  L  K  S  F  T  G  Y  Y  A

721        CCAAGGTGACCAGGATGGGAAACATGGCATCGTTCTGCGGAATGTCCAGGGCATGCGCGA
            D  L  T  V  L  I  P  F  M  A  D  N  Q  P  I  L  A  H  A  I

781        TGGCGGAGACCCTGTTGATGTCTTCGGGTGTCGGTTCCCGCCCGGTGATGAGTCCGATGG
            A  S  V  R  N  I  D  E  P  T  P  E  R  G  T  I  L  G  I  A
                     mobE
841        CTTTCTTCAGATCACTCATAAGCCGCCTTCAATGCGGTCTGTGCGGCCTCTCGATAGCGG
            K  K  L  D  S  M
                     *  E  Y  A  A  K  L  A  T  Q  A  A  E  R  Y  R

901        CGCAGTGCGGAGCGTTCCGAGATGGAGAGCTTGGTATCGGCGCTGCCCAGAGAAAGCCGG
            R  L  A  S  R  E  S  I  S  L  K  T  D  A  S  G  L  S  L  R

961        TTGTCCACGAGTTTGTCGGCCACCAGATCGTTTAGCTCTGGAAATATCACGGTTCCGGAA
            N  D  V  L  K  D  A  V  L  D  N  L  E  P  F  I  V  T  G  S
```

```
1021   ACACGGTTTTTCAGAGTGCTGTTGGCAAACCGGGCAAATTTTTCAGGTGCCCCGAAATAC
        V   R   N   K   L   T   S   N   A   F   R   A   F   K   E   P   A   G   F   Y

1081   GTGTTCAACACGGCATAGGTGGCATGGTAGGGCTTGGCCGTTTCCAGAAAATCTTTCAAC
        T   N   L   V   A   Y   T   A   H   Y   P   K   A   T   E   L   F   D   K   L

1141   AACTCCAGCGAGTCCCGTTGCCGGTTGATGGGCCAGAGCATGATGAGTTCCCGGTTCTGC
        L   E   L   S   D   R   Q   R   N   I   P   W   L   M   I   L   E   R   N   Q

1201   TCCTGTGCCACATCGTTCAGGATGCCGCCATACTGGATGAGTGCCGAAGTGGCCCGCGCT
        E   Q   A   V   D   N   L   I   G   G   Y   Q   I   L   A   S   T   A   R   A

1261   GCGGTATTGACCACGATGCTGTTTTTGGGGTGGTTCTCGATGATGTTACCGAGATGGATA
        A   T   N   V   V   I   S   N   K   P   H   N   E   I   I   N   G   L   H   I

1321   TATCCCTCTTCGGTATCCAGGTTGCAGATATGGCAAGGCAGCAGGTCTTTGAGGGCCTTG
        Y   G   E   E   T   D   L   N   C   I   H   C   P   L   L   D   K   L   A   K

1381   TAGGTATCGGGATTGCTGTCGTCACTCTCCACCAGAACACAGGAATGACTCCCGGAGAGC
        Y   T   D   P   N   S   D   D   S   E   V   L   V   C   S   H   S   G   S   L

1441   AGTGCATCGACGAGAGCCATTGTCACGGTGGATTTTCCGACGCCCCCCTTGCCACCACCG
        L   A   D   V   L   A   M   T   V   S   K   G   V   G   G   K   G   G   G
```

mobD   RBS
```
1501   ACGACCAGTATCTTGTTGAGAGACTTGCTCATAAG[TCCT]CCGAATCCGGTTTGACCGTAA
        V   V   L   I   K   N   L   S   K   S   M
                                            *   L   D   E   S   D   P   K   V   T   F

1561   AGGTTCCAGCCTGCGGGATAGGTGCTGTGTTCGCTTGGCCGGACGCTGTTGGCGCCGGTG
            T   G   A   Q   P   I   P   A   T   N   A   Q   G   S   A   T   P   A   P   T

1621   TGGTGGATTTTTTCTTTCTGGTGACTGTTGATTTGTTCTGCTTCTTTCCTGCAACGATCT
            T   S   K   K   K   R   T   V   T   S   K   N   Q   K   K   G   A   V   I   E

1681   CAGTGATGGTTTTCGGTAGTATACCGAACACGTCATCATGCTTCAGGGCATCCGCAATCT
            T   I   T   K   P   L   I   G   F   V   D   D   H   K   L   A   D   A   I   Q

1741   GTTTGCTGGTGTAGCCTTTCTTGATCAAGGCTTCCAGGTCACTCTTGATCGCTCTGATGA
            K   S   T   Y   G   K   K   I   L   A   E   L   D   S   K   I   A   R   I   V

1801   CATCGGCTTTCCCGCCTTTGCCGGTGACTTCTCCGGGCTGCTGGGTTTTCTCCAGATTTT
            D   A   K   G   G   K   G   T   V   E   G   P   Q   Q   T   K   E   L   N   E

1861   CCAGTGCAGTCCTTGCTTTTTTCGCGTATGCGTCTATATCGTAGAGTTGTTTCGAGCTTG
            L   A   T   R   A   K   K   A   Y   A   D   I   D   Y   L   Q   K   S   S   A
```

mobC   RBS   HindIII
```
1921   CCATCCTGTCGTTC[CTCCT]GTGCCGGTTTTCTTAAGCTTTGCATAGATTCTACTTCATTG
        M
```

Proposed *oriT* nick site ____
```
1981   ACTTTCGTAAAGAAAGCACTACAATACGTATAAGCTATATTTTGA[ACAGGATG]GGACTTG
```

```
2041   TCATACACATGCTACGCATGTTCCCGCCAAGCCACGCCCAACCATTCGGCGCTACGCTTC
```

NcoI
```
2101   TCACGGTCCGGCTCATCCTGCAAAGGGCAAGCCCCTTGACCCCGGCCAGCGCGGCGCCAT
```

RBS   mobB
```
2161   GGCTTCTTGCTGACCCAGACACACCCCGCCGCCGCGAA[AGGAG]ACACGATGCCATTTACC
                                                        M   P   F   T
```

```
2221    GTGCAAGGACTGGAACCTCTGGATGCGGTCGTCAATGTCCGGCTAACCGCATCGGAGAAA
         V  Q  G  L  E  P  L  D  A  V  V  N  V  R  L  T  A  S  E  K

2281    GCGCGTCTACGGGAGGATGCGGACCTCGCCGGATTGAGTGTTTCCGAGCTGGTGCGCCGC
         A  R  L  R  E  D  A  D  L  A  G  L  S  V  S  E  L  V  R  R

2341    CGCTACTTTGGCAGGCCGATTGTGGCTCATGCGGATGCGGTCTTGCTCAAAGAGCTACGC
         R  Y  F  G  R  P  I  V  A  H  A  D  A  V  L  L  K  E  L  R

2401    CGCATCGGGGGATTGCTCAAGCATGTACACAATGAAAGCGGTGGAGCGTACAGCCAGCAA
         R  I  G  G  L  L  K  H  V  H  N  E  S  G  G  A  Y  S  Q  Q
                                                          mobA
2461    ACTGCCGCCGTTCTGGTCACTTTGAAAGCCGCCATCGAGGGGCTGAGTCATGATCGTTAA
         T  A  A  V  L  V  T  L  K  A  A  I  E  G  L  S  H  D  R  *
                                                          M  I  V  K
2521    AAAGGTCAAGAGCAATCGAACAAAGGGCAAAGCGGCGAGTATACGCGATCTGACTAACTA
         K  V  K  S  N  R  T  K  G  K  A  A  S  I  R  D  L  T  N  Y

2581    CATCCGGGAGCCGCAGAACCGGAATCCAAATGAGAAGGTACTTTACGCGAACGGACGGGG
         I  R  E  P  Q  N  R  N  P  N  E  K  V  L  Y  A  N  G  R  G

2641    TTTCATCAGCGACACTCATGCCGCCCAGCGGGAGGAAATGGTAGCCCTGGCAGCGGAAGC
         F  I  S  D  T  H  A  A  Q  R  E  E  M  V  A  L  A  A  E  A

2701    GGTACGCAGCCGCAATCCGGTCAATCACTACATCCTGAGCTGGCGGGAAGGGGAGCAGCC
         V  R  S  R  N  P  V  N  H  Y  I  L  S  W  R  E  G  E  Q  P

2761    TAGCCCGGAGCAGGTGGAAGAAGCGGTAAGCATCTTTCTGGATGAACTTGGCTTGCAGGA
         S  P  E  Q  V  E  E  A  V  S  I  F  L  D  E  L  G  L  Q  E

2821    ACACCAGGTTATCTACGCCCTGCACAAGGACACGGATAACCTGCATCTGCATATCGCCGT
         H  Q  V  I  Y  A  L  H  K  D  T  D  N  L  H  L  H  I  A  V

2881    CAATCGCGTACACCCAGAAACGCTCAAGTGTGTGGAGATCAACAAGGGCTTTGACCTTGA
         N  R  V  H  P  E  T  L  K  C  V  E  I  N  K  G  F  D  L  E

2941    ATCCGCGCATCGAGCTATTGCCCGGATTGAACATGCACAGGGATGGCAACGGGAGCAGAA
         S  A  H  R  A  I  A  R  I  E  H  A  Q  G  W  Q  R  E  Q  N

3001    CGGGCGCTATGAAGTGTTAGAGAATGGCGAGCTAGGACGAGAACACCTGGAACCAAACAA
         G  R  Y  E  V  L  E  N  G  E  L  G  R  E  H  L  E  P  N  K

3061    GCGACGCCAACCGGAGCAGCGCAAGCGGGACAAGGAGAACCGCACCGGGGAGAAATCCGC
         R  R  Q  P  E  Q  R  K  R  D  K  E  N  R  T  G  E  K  S  A

3121    GGAACGTATCGCCATTGAGATCGGTGCGCCGATCATCAAGCAAGCGCAAAGCTGGGAACA
         E  R  I  A  I  E  I  G  A  P  I  I  K  Q  A  Q  S  W  E  Q

3181    GTTGCACCGGGAACTGGCAGCACAAGGGATGCGCTACGAGCAGAAGGGCAGTGGCGCATT
         L  H  R  E  L  A  A  Q  G  M  R  Y  E  Q  K  G  S  G  A  L

3241    GCTTTGGGTCGGTGAGGTAGCGGTCAAAGCCAGCAGTGCCGACCGGGAAGCGAGCCTCGG
         L  W  V  G  E  V  A  V  K  A  S  S  A  D  R  E  A  S  L  G

3301    CAAACTGCAAAAGCGGCTGGGCGCCTATGAACCCGCGCAAGCACCTTCGCCTGTGGCGCA
         K  L  Q  K  R  L  G  A  Y  E  P  A  Q  A  P  S  P  V  A  Q

3361    ACGGAAGCCGGAACCGCTTCAACCCGACAGGCCGGAGTGGGAGGACTTCATGGCTGGACG
         R  K  P  E  P  L  Q  P  D  R  P  E  W  E  D  F  M  A  G  R

3421    CAAAATGCACTACGCAGAAAAGAACGCGGCCAAGCTCTCCATGGACCAGCGGCAGGAACA
```

89

```
              K   M   H   Y   A   E   K   N   A   A   K   L   S   M   D   Q   R   Q   E   Q

3481   GGAACGCAAGGCATTGCAAGCGCGGCAGCAGGAACAACGCAAGGTACTCCTGGGCGGACG
              E   R   K   A   L   Q   A   R   Q   Q   E   Q   R   K   V   L   L   G   G   R

3541   ATGGAATGGCAAGGGAGAAGCGCTGAACGCGCTGCGCGGCGTGCTTGCCGCAGAACAGGC
              W   N   G   K   G   E   A   L   N   A   L   R   G   V   L   A   A   E   Q   A

3601   CGCGGAGAAAGCCGCCCTGAAAGAACGTCACCAGCAGGAACGGCAACAGTGGCGGCAACA
              A   E   K   A   A   L   K   E   R   H   Q   Q   E   R   Q   Q   W   R   Q   Q

3661   GTACCGCCCGTATCCCGACTTCGAGCAGTGGTTACGACGCGAGCACGGGGCAGAACAGGC
              Y   R   P   Y   P   D   F   E   Q   W   L   R   R   E   H   G   A   E   Q   A

3721   GACTAGGTGGCGCTACCGGGAGGCGGAACCGCAACGCATCGAGGGCGACTCGACCGAAGC
              T   R   W   R   Y   R   E   A   E   P   Q   R   I   E   G   D   S   T   E   A

3781   GCCGAAACCTCGGGACATTCGCGCCTACCGGGCGGAGATCGTGGGACAGGAGGTCCGTTA
              P   K   P   R   D   I   R   A   Y   R   A   E   I   V   G   Q   E   V   R   Y

3841   CACCCCCAAGAGTGGCGCGGGCGGCGGGCCGGGGGGTGTGTCTTTTGTGGACAAGGGCAG
              T   P   K   S   G   A   G   G   G   P   G   G   V   S   F   V   D   K   G   R

3901   GATCATCGAAATCCACGATTGGCGGAACCAGGACACCACTCTTGCGGCGCTCCAGCTCTC
              I   I   E   I   H   D   W   R   N   Q   D   T   T   L   A   A   L   Q   L   S

3961   GGCGCAGAAATGGGGCAAGTTCACCGTGACCGGGAACGACGAATACAAAGCGCTGTGCGT
              A   Q   K   W   G   K   F   T   V   T   G   N   D   E   Y   K   A   L   C   V

4021   GAAGCTGGCGGTGGAGCACGGTTTCCAGATCACGAACCCGGAGCTTCAGGAGGTCATTCG
              K   L   A   V   E   H   G   F   Q   I   T   N   P   E   L   Q   E   V   I   R
                                       RBS                     repB
4081   GCAGGAACGGCAACGGATGCGGCAGGAGAGGGCACAGGCGATGAAATCGGAGCAGATCAA
              Q   E   R   Q   R   M   R   Q   E   R   A   Q   A   M   K   S   E   Q   I   K

4141   GCCGTTCGAGCGATACGCCGAAGCGGTTGGCGCCGAGCGCTACCGGGTGACAAGCATCAA
              P   F   E   R   Y   A   E   A   V   G   A   E   R   Y   R   V   T   S   I   K

4201   AATGCGGCCAGATGGCAGCAAGCAAACCTTCATTCTCGATAAGCGGGACGGTATCACGCG
              M   R   P   D   G   S   K   Q   T   F   I   L   D   K   R   D   G   I   T   R

4261   CGGCTTCACGCCGGAAGAAATCGCCCAAAGGACGCCCGAAATGCAGCGTTTACAGCGCCG
              G   F   T   P   E   E   I   A   Q   R   T   P   E   M   Q   R   L   Q   R   R

4321   GGGCGAAAACCTGTATTACACGCCGCTCTCCGAAGGGAAACACCATATCCTGATCGACGA
              G   E   N   L   Y   Y   T   P   L   S   E   G   K   H   H   I   L   I   D   D

4381   CATGGACCGTGAGAAACTGGACCGGCTGATTCGTGATGGCTATCAGCCCGCCGTCGTGCT
              M   D   R   E   K   L   D   R   L   I   R   D   G   Y   Q   P   A   V   V   L

4441   GGAATCCAGCCCCGGCAACTATCAGGCGGTCATCACCATTCCGAAGCTGGGGACCCCCTT
              E   S   S   P   G   N   Y   Q   A   V   I   T   I   P   K   L   G   T   P   F

4501   CGACAAGGACGTGGGGAACCGCCTGAGCGATGCGCTCAACCGCGAATACGGCGACCCCAA
              D   K   D   V   G   N   R   L   S   D   A   L   N   R   E   Y   G   D   P   K

4561   GCTGTCTGGTGCCATCCATCCGCACCGCGCTCCCGGCTACGAGAACCGCAAGCCCAAGCA
              L   S   G   A   I   H   P   H   R   A   P   G   Y   E   N   R   K   P   K   H

4621   CCGGCGGGAAGATGGCAGCTATCCCGAAGTGCGTTTGCTCAAGGCCGAGCGGCGCGAGTG
              R   R   E   D   G   S   Y   P   E   V   R   L   L   K   A   E   R   R   E   C
```

90

```
                                                         EcoRV
4681   CGGCAAGACGCTGGCGCTTTCCCGCGAGATCGACGCCGGGATATCAGCGGCAGGCCGCCGA
        G  K  T  L  A  L  S  R  E  I  D  A  G  Y  Q  R  Q  A  A  E

4741   GAAGGCCCTGAAAACGCCCGTAGAGCAATTTTCGGAGCCAAACAGTACCCAGACCACCCC
        K  A  L  K  T  P  V  E  Q  F  S  E  P  N  S  T  Q  T  T  P

4801   TGTCTCTGAAAAAACCGCCACAGAGGCTTATTGGCGGCATTACCGCGATGTTCGCAAGCG
        V  S  E  K  T  A  T  E  A  Y  W  R  H  Y  R  D  V  R  K  R

4861   CCAACGCGGAATGCTGGACCTGTCCCGCGTGGACGCCATGATTGCCGTGCGGATGCGCGT
        Q  R  G  M  L  D  L  S  R  V  D  A  M  I  A  V  R  M  R  V
                                                         ClaI
4921   TACCGGCTTTGACCAGTCGGCTATTGAGGGCGCCATCTACCAGTGTGCCCCATCGATCCG
        T  G  F  D  Q  S  A  I  E  G  A  I  Y  Q  C  A  P  S  I  R

4981   GGAGCAGCAGGAAAGCCGGGACTGGACCGACTACGCCCGCCGCACGGCCCGCTATGCCTA
        E  Q  Q  E  S  R  D  W  T  D  Y  A  R  R  T  A  R  Y  A  Y

5041   CAGCGCCGCAGGTGACCGGCAGGCCGCCGACCTGGGTAAGTACCGCCAGCAGTGGGAGAA
        S  A  A  G  D  R  Q  A  A  D  L  G  K  Y  R  Q  Q  W  E  K

5101   GCTGGAGGGGCGCGAGCGGCAGCAGGAGCAGGCCAAGGCGCGGGAGATTGAGCGCAACGG
        L  E  G  R  E  R  Q  Q  E  Q  A  K  A  R  E  I  E  R  N  G

5161   GCCGAGTATGAGTCGTTAGTATTCTTCACAGCGGCATGATATACTTGTATATCATTTTGA
        P  S  M  S  R  *
        RBS          pasA                                    StuI
5221   GCAGGAGCTAAACATGCTTGCCATCCGATTACCCACCGAAGTGGAAAATCGCCTTGAGGC
                      M  L  A  I  R  L  P  T  E  V  E  N  R  L  E  A

5281   CTTGGCGCAGGCCACAGGACGCACGAAAACCTTTTACGCCCGCGAAGCGATCCTTGAGCA
        L  A  Q  A  T  G  R  T  K  T  F  Y  A  R  E  A  I  L  E  H

5341   CCTGGACGACCTCGAAGATTTGTACCTTGCAGAGCAACGCCTGATCGACATTCGCGCAGG
        L  D  D  L  E  D  L  Y  L  A  E  Q  R  L  I  D  I  R  A  G
                                                 pasB
5401   CCGAAGCCGTACCTACACGCTAGAGGAAGTGGAGCGCGATCTTGGCTTGGCGGATTGAGT
        R  S  R  T  Y  T  L  E  E  V  E  R  D  L  G  L  A  D  *
                                                 L  A  W  R  I  E

5461   TTGATGACAAGGCCAAGAAAGACTTGGCTGCGCTGGATAAGAGCGTTGCCAAACGCATCA
        F  D  D  K  A  K  K  D  L  A  A  L  D  K  S  V  A  K  R  I
                                   XbaI
5521   CGGCTTTTCTGCGAGAGCGCGTCGCCCATCTAGACGACCCGCGCAGCATTGGCGAAGCCC
        T  A  F  L  R  E  R  V  A  H  L  D  D  P  R  S  I  G  E  A

5581   TCAAAGGCTCCAAGCTGGGAGACTTCTGGAAGTACCGCGTAGGGGATTGGCGAATCATTG
        L  K  G  S  K  L  G  D  F  W  K  Y  R  V  G  D  W  R  I  I

5641   CCAGCATTGAGGATGAAGCCTTGCGTATTCTTGTCGTGCGTATTGGTAACCGCCGAGAGG
        A  S  I  E  D  E  A  L  R  I  L  V  V  R  I  G  N  R  R  E

5701   TATACCGAAAGTAAACAACCCCGGCATGAAGGGCCGGGGTTTCCCGCGCAACCGGATGAA
        V  Y  R  K  *

5761   AGGACGGCGCGAGCCGGAGTGTCAGTCCCCGGCAGCGCCTAACCACGCCAGTTCTGGAAA
        RBS          repA
5821   GGAGAACAGCATGGCTTTGGACATTATGGCGGCATTCACCAACGAACCGCCAGAACTGGA
                      M  A  L  D  I  M  A  A  F  T  N  E  P  P  E  L  D
```

91

```
5881   TTTCATCTGGCCGGGATTCCTGGCCGGGACCGTGGGCGCACTGGTTGCACCTGGCGCGAC
        F  I  W  P  G  F  L  A  G  T  V  G  A  L  V  A  P  G  A  T

5941   CGGCAAGAGCTTCTGGGCCTTGGAAGCGGCGATGGCCGTCGCGTGCCGGGAGCCTGGCGG
        G  K  S  F  W  A  L  E  A  A  M  A  V  A  C  R  E  P  G  G

6001   AGACCTGCTGGGGCTTGCGCCGAAGTACACCGGGCGCGTTGTCTATCTGGCCGGAGAAGA
        D  L  L  G  L  A  P  K  Y  T  G  R  V  V  Y  L  A  G  E  D

6061   TCCAGAACCAGCCTTAATCCGGCGCGTTCATGCCATAGGCCAGCATCTCAACCAGTCGGC
        P  E  P  A  L  I  R  R  V  H  A  I  G  Q  H  L  N  Q  S  A

6121   CCGCGAAGCCATTGCTGAGAATCTGACCCTTGAGCCGATCATGGGCAAGCGGCTGAACAT
        R  E  A  I  A  E  N  L  T  L  E  P  I  M  G  K  R  L  N  I

6181   CATGGACGAGGCCCACCTGCATCGCATCATCGAGGCCAGCTCCGGGGCACGGCTGATCGT
        M  D  E  A  H  L  H  R  I  I  E  A  S  S  G  A  R  L  I  V
                                                      NcoI
6241   GCTGGACACCCTGAGCCGTATTCACAATCTGGACGAGAACAGCAACGGTGCCATGGCGCA
        L  D  T  L  S  R  I  H  N  L  D  E  N  S  N  G  A  M  A  H

6301   TCTGGTGGCAACCCTGGAGCATATTGCCGCCAGTACCGGCGCATCGGTGCTGTACCTGCA
        L  V  A  T  L  E  H  I  A  A  S  T  G  A  S  V  L  Y  L  H
              NcoI
6361   TCATGTCAGCAAAGGCAGCGCCATGGCCGGGCAGACGGATCAGCAGCAAGCGGCGCGGGG
        H  V  S  K  G  S  A  M  A  G  Q  T  D  Q  Q  Q  A  A  R  G

6421   GGCATCAGCCCTGATCGACAATGCCCGTTGGTGCGGCTACGTTGCCAAAATGGCAGAGCA
        A  S  A  L  I  D  N  A  R  W  C  G  Y  V  A  K  M  A  E  H

6481   TGAGGCCGCGCGACTGAGCGACCGTAGCTTTGACCGTGCGCCCATTGGCGACCGGCGCGG
        E  A  A  R  L  S  D  R  S  F  D  R  A  P  I  G  D  R  R  G

6541   CTTCTTCGTGCGCTTCGGTGTGAGCAAACAGAACTACGACGCAACATCGCTCGAAAGCTG
        F  F  V  R  F  G  V  S  K  Q  N  Y  D  A  T  S  L  E  S  W
                                                  SalI
6601   GTATCAGCGAGTGGATGGCGGGGTGTTGATGCCCGTCGACCTGGTGGAAGCAACGAGCAA
        Y  Q  R  V  D  G  G  V  L  M  P  V  D  L  V  E  A  T  S  N

6661   TGGGAAGGTAAAAAACGATGGCAAAGGAAAGCGCGATGAGATCTGACCCTGTGCTATCGC
        G  K  V  K  N  D  G  K  G  K  R  D  E  I  *
                                                      repC
6721   ACCCCCTGCCGATCCCCTTTTTGTGTGGCTATACCCAGGGAGGCAAGCCATGAAGAAGGG
                                                           M  K  K  G

6781   AAGCGCGATGACCTTCGACCTCACCCACGCAAGACACGATCCGGCTCACTGTCTGGCTCC
        S  A  M  T  F  D  L  T  H  A  R  H  D  P  A  H  C  L  A  P

6841   TGGTCTTTTTCGCAGCCTCAAGCGTGGCGAGCGAAAAAAACTCAAGTTGGACGTGACCTA
        G  L  F  R  S  L  K  R  G  E  R  K  K  L  K  L  D  V  T  Y

6901   CACTTACGGTAAGGATCGTGTTGAGTTTAGCGGTCCAGAGCCGCTTGGTGTTGATGATAT
        T  Y  G  K  D  R  V  E  F  S  G  P  E  P  L  G  V  D  D  M

6961   GCGTGTCCTGCAAGGTCTCGTTGCCGTCGCCGCCACCTGTGGGTCGGAAAGTCGGGGAAT
        R  V  L  Q  G  L  V  A  V  A  A  T  C  G  S  E  S  R  G  I

7021   CATGCTGCGCCCTGAGCCAAAAAGTGAAGCGGGACATCAGTTGAGGCTGTTGCTTGATTT
        M  L  R  P  E  P  K  S  E  A  G  H  Q  L  R  L  L  L  D  L
              NcoI
```

```
7081    AGAATGGGACGCCAAGGAAAAAGATGCCATGGTCGCTAAAGGCAGCTTCCGGCAACTTGC
         E   W   D   A   K   E   K   D   A   M   V   A   K   G   S   F   R   Q   L   A

7141    AAGCGAGCTTGGATATGCTACAGACAGTGGCGGCGCGTTTCGTGTCATTCGAGAAAGTAT
         S   E   L   G   Y   A   T   D   S   G   G   A   F   R   V   I   R   E   S   I

7201    CGAACGACTATGGAAGGTCTCCGTGATCGTTCAGCGCGAGGGGAAACGACAGGGATTCCG
         E   R   L   W   K   V   S   V   I   V   Q   R   E   G   K   R   Q   G   F   R

7261    CATCCTCTCCGAATATGCCAGCGACGAAGACACCGGGAAGCTATTTGTTGCCTTGAACCC
         I   L   S   E   Y   A   S   D   E   D   T   G   K   L   F   V   A   L   N   P

7321    CAGACTGGCTGAGGCAATCATCGGAGCACGTCCACACACCCGTATCGAGATGGCAGAGGT
         R   L   A   E   A   I   I   G   A   R   P   H   T   R   I   E   M   A   E   V

                PstI
7381    GCGAGCACTGCAGACTGACCCTGCCCGTCTCATTCATCAGCGGCTATGCGGTTGGATCGA
         R   A   L   Q   T   D   P   A   R   L   I   H   Q   R   L   C   G   W   I   D

7441    CCCCGGCAAGTCCGGGCGCGTCGAGCTGGACACGCTTTGCGGTTATATCTGGCCCGATCA
         P   G   K   S   G   R   V   E   L   D   T   L   C   G   Y   I   W   P   D   Q

7501    GGCCAGCGGGTCAACCATGCGTATGCGCCGACAGACTGCTCGTAATGCGCTGCCCGAGCT
         A   S   G   S   T   M   R   M   R   R   Q   T   A   R   N   A   L   P   E   L

7561    TACCACTGTAGGCTGGACGGTGAACGAATACGCCAAGGGAAAATGGGAAATCAGCAGACC
         T   T   V   G   W   T   V   N   E   Y   A   K   G   K   W   E   I   S   R   P

                22/22 bp                23/22 bp                22/22 bp
7621    GAAGGCCCTTGCCCTAACGTTGACGAGCCCCTGCCCTAACGTTGACGAGCCCCTGCCCTA
         K   A   L   A   L   T   L   T   S   P   C   P   N   V   D   E   P   L   P   *

                22/22 bp                21/22 bp
7681    ACGTTGACGAGCCCCTGCCCTAACGTTGACGAGCCCCTGCCCTAACGTTGACACCACAAA

                                                                ClaI    _____
7741    ACACGCTAAGCCAGCATTGACGGGGCTTTGCGGAAGATTTGAAAAATCGATCCAAGATAT
         8 bp inverted repeats, weak stem loop

A+T rich region 29/40   14/15bp region highly conserved in IncQ-like plasmids
7801    TCTCCAAGAATATCCACTATGCGCGGCCTTCAGCCGCCGCTTGATGCGTCGTCAGAAGCC
                                                2 X 12 bp tandem repeats

7861    TTGCACTTCGGCCTTCGGCCTGCGCGGGCTACGCCCTTGCAGACACACCCCGCCGCCCGC
                                        12 bp inverted repeat ΔG-15.2

7921    CCGCCCGCCGCCCGCTTGTACGTTCTCCAAGGTGGATAACGTGACTTACGTTATCCACCA

7981    ACCATGCCGCCGATAGTGAGCACCTACTGGGTAGACGAGCGGCTGATGGTGTAAATAACC
                                        weak 8 bp inverted repeat

                                                                        NcoI
8041    TTGACAAATAGATGTAAATTTACTATATTACATTTACAATGTAATTACAGGAGAGTGCCA

        Start 124aa ORF, strong RBS, no meaningful homology
8101    TGGAAGCAGGACCGATCACACAACAAGACGTTGCCGCCGCCGCTGACCAGCTTGGCGGCC
         M   E   A   G   P   I   T   Q   Q   D   V   A   A   A   A   D   Q   L   G   G

8161    CCGCCAACACCAATGCCGCGAAAATCCGGGAGATTCTGGGCCGGGGGGAGTCTCGCCACCA
         P   A   N   T   N   A   A   K   I   R   E   I   L   G   R   G   S   L   A   T
```

```
8221    TTCAACGCCATTTGCAGGCCCTTAGAGACGCTCAGAGGGCTCCAGAGCTACCAGAGGCCG
        I  Q  R  H  L  Q  A  L  R  D  A  Q  R  A  P  E  L  P  E  A

8281    TCCAGACGATTCCAACGCCTCCTGAGGCCGTTTCTGAGGCTTCCAGAGGCATTTGGGCGG
        V  Q  T  I  P  T  P  P  E  A  V  S  E  A  S  R  G  I  W  A

                                                          SphI
8341    CGGCCTGGGCAATGGCGGAGCAACGCCACGCGGAAAGTCTGGCGCATGCCCCCTGCTGGG
        A  A  W  A  M  A  E  Q  R  H  A  E  S  L  A  H  A  P  C  W

8401    TTATAACAGGTATTACTTTTGCGTACTCTGGCAACGGATGCCAGTCTCTATCTCCCAATA
        V  I  T  G  I  T  F  A  Y  S  G  N  G  C  Q  S  L  S  P  N

                 Stop of 124 aa ORF      Stop of invertase ORF 189aa
8461    CAACCCTGCTTTAGGGCGAAAATCCATGCTATGAACCATTCCCGCCACGCTTACGCTGCA
        T  T  L  L  *                  *  S  G  N  G  G  R  K  R  Q  L

8521    ACTCGCGATAGAGGACGGTTTTACTGATCTTGAGGCGGGTGGCGGCCTCGCGGACGCTGA
         E  R  Y  L  V  T  K  S  I  K  L  R  T  A  A  E  R  V  S  L

8581    GGCCCTGGGCGATCAACTGCTGGGCGCGTTGCTGCTTCTCTGGGGTCATGACGGGTTTGC
         G  Q  A  I  L  Q  Q  A  R  Q  Q  K  E  P  T  M  V  P  K  R

8641    GTCCCCCTTTGCGGCCACGGGCGACTGCAGCGGCGAGGCCTGCCAAGGTGCGCTCCCGGA
         G  G  K  R  G  R  A  V  A  A  A  L  G  A  L  T  R  E  R  I

8701    TAATGTCACGCTCGAACTGTCCGAGCGCCCCAAAGAGATGGAAGGTCAGTCTCCCGGCTG
         I  D  R  E  F  Q  G  L  A  G  F  L  H  F  T  L  R  G  A  P

8761    GGGTGGTGGTGTCAATGGATTCGGTCAGGGACCGAAAACCGCACCCGCGTTGTTCCAGGT
         T  T  T  D  I  S  E  T  L  S  R  F  G  C  G  R  Q  L  E  D

8821    CGTTCAGGACCTCGATGAGGTGGGGGGAGAGAACGGGCCAGGCGATCCAGTCGCCAGACGG
         N  L  V  E  I  L  H  P  L  S  R  A  L  R  D  L  R  W  V  T

8881    TGAGCACGTCGCCAGGGCGGAGGAAGGCCATGGCCTGCACCAATCCGGTGCGTTCCGCCT
         L  V  D  G  P  R  L  F  A  M  A  Q  V  L  G  T  R  E  A  K

8941    TGGCCCCGCTGGCGACATCCTCAAAGGTGTGTTCGCACCCGGCATGGTTCAGGGCATCCC
         A  G  S  A  V  D  E  F  T  H  E  C  G  A  H  N  L  A  D  R

                                 Start of invertase-like ORF 189aa
9001    GCTGTAGGTCGAGGTTCTGGTCTCCGGTGGATATCCGCGCATAACCGATCTGTGCCATGG
         Q  L  D  L  N  Q  D  G  T  S  I  R  A  Y  G  I  Q  A  M

9061    GTGTCGCCTGTGTGCGTCAAGTCGTCCCACAATCATAGCGTCCGTCTATCCGTTGAAACA

9121    GGGGGTTGTGGGACGCTTTTGGGGGTGTCCGTCAGAGGATCGGGATCACCCTTCGTCTGC

9181    CAGAGCCGGAGTTTCAGCCATCCGTCCCCATAGCGCCAGACCCGCCGATAGGTTTCGCGC

9241    TCCCAGTCATCCATCCGGTTGTGGGCTTGACGGCGCAGCGCCCAGACCGTGAGATCCAGG

                              Stop of A.tumer hyp ORF
9301    CGAAAGGGCGGTCGCGGGCTGAGGCGGAAGTAGTATCCCATCTATCGGTCCAGATCCGGC
                                                    *  R  D  L  D  P

9361    CGATAACTGCCGGCGGGGGTACGGAAGGCAATGGCTAGGCGGTTCCAGGCGTTGATGGCC
         R  Y  S  G  A  P  T  R  F  A  I  A  L  R  N  W  A  N  I  A

9421    ACGACGGCCAAGGTCAGATCCACCAGTTCCTTCTCCTCGAAATAACGGCGGACGTTTTCA
         V  V  A  L  T  L  D  V  L  E  K  E  E  F  Y  R  R  V  N  E
```

94

```
9481    TAGAGCGCATCCGAGACGCCTTCATCGATGCGGGTGATAGCATCGGTCCACGCCAGCGCC
        Y  L  A  D  S  V  G  E  D  I  R  T  I  A  D  T  W  A  L  A

9541    GCGCTTTCCCGCTCGCTGAAAAATGGCGCTTCCCGCCAAGCGGCGAGGGCATAAAGACGC
        A  S  E  R  E  S  F  F  P  A  E  R  W  A  A  L  A  Y  L  R

9601    TGCTCCATCTCGCCCCGGGCGCGGGCTTCCTTGCTGTGCATGTCGATGCAAAAGGCGCAG
        Q  E  M  E  G  R  A  R  A  E  K  S  H  M  D  I  C  F  A  C

9661    CCGTTGAGGATGGATGCGCGCAGCTTCACCAGTTCCAGCAGCGCTGGTTCCAAGTTGCTT
        G  N  L  I  S  A  R  L  K  V  L  E  L  L  A  P  E  L  N  S

9721    CCATGGAGATAAGTCTCCAGATCCCGCAGGCTGCGGATGCCGTTGGGGGATACCAAGCCA
        G  H  L  Y  T  E  L  D  R  L  S  R  I  G  N  P  S  V  L  G
                    Start of A. tumer hyp ORF 153aa
9781    TAGTCGACGCGTTTTTCCATCATGGA**CTCCT**GTGGGAACCTTCTTGGAGCGGAAGGGGAT
        Y  D  V  R  K  E  M  M

9841    CTGGGGCCGGATACCAGATCCCGCCATCGGGACGCCGCCGCGCCCAAGGGCTCAAACGAT

9901    GCCTCTGGTCATGCCCTGGCGTCAGCAGGTGATGCCACGGTAAAACCCTGGGCGGTAAGG

9961    ACGTCAGCAATGAGAAAGCGATGACAGCGCCACGGGAGGGTGCTCGGCGCACAGGAGGAC

10021   CGTCCTGCGCCTACGCGACCAAGGTGCTCCAAGGCCTCCAATCCTTTCTGGAAAGGCGCC

10081   GTCTGCGCAAGCGACCATATCGGCAGATACTGCCAGGATAGAGCCAGAAGACTGATCTAC

10141   CGGCACAGCGCGACCATTCTAATCACGGATGTTCAGAGGTTCCAAAGGAATATCATGAGC

10201   GCCACATATCACGATCCACGATACGCTAGGCAGGATGGCTAGACAGGCCGCCTCCGCAGA

10261   CGGTATCTGATCCTGAGCGAGCGTTGCCCATAGTAAGCTTGCGGGGTATCCTCTAACTGA

10321   AGAACGACGGCTTGAGACTGGAAGCAGACGTTCGGTTGGTGCGCCAAGAAAGGCGCATCG

10381   AATCTCCTGCCAGTCCAAGGTTAGTCAGCATTGGGCATGATCTGTAAGCTCGCTTGCTAT

10441   CACACCTCCTAATAATCAACTATGATTGACGCTGGCGGTACAATCCGATGAGGAATGGCA

10501   ACGTGAACAAGCAAGAATTC**CTA**TCCTGCTTCGTTAAGGGCCTCTAAACATAGCTTCTCA
                            *  G  A  E  N  L  A  E  L  C  L  K  E

10561   TATCCTAGTGGTGTCAATTCAGGTTCTTCTATAAACTGTGTAATTTCAACCTTATGTTTT
        Y  G  L  P  T  L  E  P  E  E  I  F  Q  T  I  E  V  K  H  K

10621   TTTATTTCATCTGCCAATACATTTTCAAGCATATCGTACAGCATATTAAGACTTAGTTGT
        K  I  E  D  A  L  V  N  E  L  M  D  Y  L  M  N  L  S  L  Q

10681   TTAATCCCAGAATAGAACGATACCGAAATGAATTTTGAAAATATCTCATTTTCAGCAAGA
        K  I  G  S  Y  F  S  V  S  I  F  K  S  F  I  E  N  E  A  L

10741   AGCACTTTTATCTTATCCACACAATCATCTTTTATACTATTTATGTTATCTTTGCTGTCT
        L  V  K  I  K  D  V  C  D  D  K  I  S  N  I  N  D  K  S  D

10801   GGAAATGCTCCCTTAGCATATTCCTTAAAATCCAATGAGCTTGTATTTAGGAACAACGAT
        P  F  A  G  K  A  Y  E  K  F  D  L  S  S  T  N  L  F  L  S

10861   TTTATTTCTTTTATTGTGTTTTCTATATCCTCTCCGAATTTATTGTATATCACATTTTCA
        K  I  E  K  I  T  N  E  I  D  E  G  F  K  N  Y  I  V  N  E
```

```
10921   AAAAAATATTCCTTCTCTTTTAACCATTCCTGACCGTCAAGAACATCATATTTGACAACC
        F  F  Y  E  K  E  K  L  W  E  Q  G  D  L  V  D  Y  K  V  V

10981   CTTACCCCCTCGAAAAATCTTTCTGACAGCTCGCTGTAATGTGTCTTGATGATATCTGAA
        R  G  V  E  F  F  R  E  S  L  E  S  Y  H  T  K  I  I  D  S

11041   ATTTTTTGTTTGTACTCGTTATTGTTTATGGTGTATTCTTTAGTCAGTTTTGCAACAGAA
        I  K  Q  K  Y  E  N  N  N  I  T  Y  E  K  T  L  K  A  V  S

11101   AACGCCATAGATGTGATTATACTCTTTATTTGTTCTTGCTCAAAATAAGCATTTGAATCA
        F  A  M  S  T  I  I  S  K  I  Q  E  Q  E  F  Y  A  N  S  D

11161   TTATCTTCATCCTCTATATCAACATCGCCATAGTCAGTTATGTCATCCTCGTCATCACCG
        N  D  E  D  E  I  D  V  D  G  Y  D  T  I  D  D  E  D  D  G

11221   TCTTCAAAAAAGAACTTTAAATCGCCATCCTTTAGTTGTTGAAGGGCTATTATTATATCG
        D  E  F  F  F  K  L  D  G  D  K  L  Q  Q  L  A  I  I  I  D

11281   TCAGATGAGCATGTCGTGCTTTCCTTGGCGTTATCATACTGCGTCAAGTCTTTCTCTAGT
        D  S  S  C  T  T  S  E  K  A  N  D  Y  Q  T  L  D  K  E  L
                            Start of aminotransferase-like ORF 286aa
11341   GTTGGAGTAATACGTGTTGCCTTTTTTTTTCCAGAAAAA**CAT**ATAAAAGG**CTCCT**AGATTT
        T  P  T  I  R  T  A  K  K  K  W  F  F  M

11401   TTTATTGAGAGCGATACCAGACAAAATGCATCCTTTGGTCCGGTTCCGGCCGGGGACCGG

11461   AGCAGAACCCATTGATTCCATGAGAGCCGAGAGCCGCAGAAGGCCGGCAAACATCTGTCC

11521   TTACGGTGAACATGGTCATTTCCACACCGAATCCCCGAATTCACCGTATAGAAGACCTGA

11581   TGGCACCCTAGGATATGAGAGCTATGTTAGTGCCTTAACGAAGCAGGATGGATTCTGAAC

11641   TAACAAAATCGGACCGGCGACCAAGGTGCTGATATAATTGCAAAAAAGAATGATAATTCT

11701   TTGGTGGTGCAGTGCAAATATTACTCCGGTGGGCGTTGAACCGAAAGCAGTCCAGGAAGT

11761   TATCGCTGCGCGTACATATTATTGTGCTGATTATGCCACCGTTGTATCAAACACATACTT

11821   CACACCTGGCGCGAAAGATTTAGCAAAAAGAGGCGATGTTTTACTATTGGACGAAAACAA

                                         **direct repeat of IS element**
11881   GTTGAAGAATATTGACTTTTTTTTTAAAAAACAGTTA<u>GGCTCTTCGACGTTTCATGGGGT</u>
                                            *  T  P

11941   <u>AGCCTCAGATCCAGCTTGCCCAGGATGAGG</u>TAAGCCATGGCGATGAAGTTCTTGTGGGTG
        L  R  L  D  L  K  G  L  I  L  Y  A  M  A  I  F  N  K  H  T

12001   CGGTAGCCACGGGCTTTCGCCTTGGCGGACTGGAGGAGGCTGTTGAAGCCCTCCAGGATG
4001    R  Y  G  R  A  K  A  K  A  S  Q  L  L  S  N  F  G  E  L  L

12061   CCATTGCTGACTTGGCTCTCGAACCAGCGGAGGACCCCGTCCCAATGGTTCATGAGGGTG
4021    G  N  S  V  Q  S  E  F  W  R  L  V  G  D  W  H  N  M  L  D

12121   TAGGCGACCCTGACGATAGGCGGCAGATCGCTGGTTCTGGCGTTTTCCAACCAGGCTTTC
4041    Y  A  V  R  V  I  P  P  L  D  S  T  R  A  N  E  L  W  A  K

12181   AAGAGGGTAGCCCCCTGGTGGCGATTCTTGATCGTGAAGATGTCCTGAAAGGTCAGGCGG
        L  L  D  A  G  Q  H  R  N  K  I  T  F  I  D  Q  F  T  L  R

12241   AACTGGTAGGCCTGCGCCGTCTTGAGGTTCTGGTCTTTGAGCAATTCCTGCAGCTTTTCT
```

96

```
              F   Q   Y   A   Q   A   T   K   L   N   Q   D   K   L   L   E   Q   L   K   E

12301   TTCTGCTTCATTTTGAGGTTGCAATCGTTCTTGAGCCAAAGCCAGCGGGTCTTTTTGAGA
              K   Q   K   M   K   L   N   C   D   N   K   L   W   L   W   R   T   K   K   L

12361   TTTGGCTGGGTGAGGACTTCCCCCCTTGCGCACGTCGTCTACGGCCTCGTTGACGAGCTTC
              N   P   Q   T   L   V   E   G   K   R   V   D   D   V   A   E   N   V   L   K

12421   ATGAGGTGAAAACGATCGAAAGTGATCTCCGCATTGGGCAGGTGCTCGGCAGCCCCTTTC
              M   L   H   F   R   D   F   T   I   E   A   N   P   L   H   E   A   A   G   K

12481   TGGAAGGCCGGCGAGAGGTCCATGCTCACATCGGTGATCGCTTCCGCGCTACCACCATGG
              Q   F   A   P   S   L   D   N   S   V   D   T   I   A   E   A   S   G   G   H

12541   GCCTGAAGATCTCCGGAGAATTTCTCAAAGGTCTTGGCATCCTTGCCGGGAGTAGCGAAC
              A   Q   L   D   G   S   F   K   E   F   T   K   A   D   K   G   P   T   A   F

12601   AAGAGTCGCCGGGCATTCAGATCCACGAAGAGCGTGATGTAGTCATGTCCATGCCGACTG
              L   L   R   R   A   N   L   D   V   F   L   T   I   Y   D   H   G   H   R   S

12661   CTGGTTTCATCGACGCCGACGGCATGGACATTGGCCATATCCACCGCAGCACGAGCTTCG
              S   T   E   D   V   G   V   A   H   V   N   A   M   D   V   A   A   R   A   E

12721   GGCACATAATGGTCAATCACTCGCCACAGGAGCTTGTCGGTCTCACCGACCATGCGAGCT
              P   V   Y   H   D   I   V   R   W   L   L   K   D   T   E   G   V   M   R   A

12781   GCCGTCAATACCGGCATCTCCCGCACCAGGGTCATGATCAGCGCTTCAAAGAGCAGGGTG
              A   T   L   V   P   M   E   R   V   L   T   M   I   L   A   E   F   L   L   T

12841   AAACGCGAGCCTTCCCGCGCCCAGGGAACAGATATCTGATGCACTCCATGCTCCTGGCAC
              F   R   S   G   E   R   A   W   P   V   S   I   Q   H   V   G   H   E   Q   C

12901   TTCACACGAGGTACACGGGCATGGAGATAGGCTTCATGCTGGAAGAAATCCATGTGCCGC
              K   V   R   P   V   R   A   H   L   Y   A   E   H   Q   F   F   D   M   H   R

12961   CAGGTATGGTCACGGGTATCATGTACCGGACACTCCTCACCACAGACGGAGCAAGCAAAG
              W   T   H   D   R   T   D   H   V   P   C   E   E   G   C   V   S   C   A   F

13021   CGACTACCTTTGGGAAAGTTGATGTGCAGATCCAGGCGTTTCTCCTCCACCGTGAAAGTC
              R   S   G   K   P   F   N   I   H   L   D   L   R   K   E   E   V   T   F   T

13081   ACATGATCCACCAACCACGGCGGAACCAATCCTAACGCGAGAGAAAACAGCTCTTCAGGG
              V   H   D   V   L   W   P   P   V   L   G   L   A   L   S   F   L   E   E   P
        **Transposase**                              **direct repeat of IS element**
13141   ACCATCAGCTAA**CTCCT**ATTCAGGGCACGACCGCTCCACAGCATTCTACCCCC̲C̲T̲A̲C̲C̲C̲A̲C̲
              V   M

13201   T̲C̲A̲A̲T̲C̲T̲G̲A̲C̲G̲A̲A̲G̲A̲G̲C̲C̲ATAAGGAAGGATCTCGATATCATAAGTTTCCAGTGATGTGAA

13261   TCGTTCTTTGGTGCCATCCAACTCCCATAGAATCTGGCTGTCTCTAAGCGCACGTAATCC

13321   ATACAGCCAGAAATCACCGGGCATGTCGCTCGGGTCTAACCTAAACTGCATGCTCCCGGC

13381   ATGGATGCTGAAGACAATCTGATGCCTTTCCCCATCTAGTGGATATGGTTTACTATTACT

13441   TCTGTGCTCATCGAAACCATCTTTCTTTTTGGACCAGTATAAGGTTGTGTTTGCGCTCAA

13501   TGTCCTATTATTTGTAATCACTGCGGGAGCAGTGATTAGGCTGGCCTTTTCCTGCTGTAC

                                              stop of 86 aa ORF
13561   CGCCTCCAATATAAACTGGTAGACGGACCCATCT**TCA**GAACAGAAGATGTTCTGGACTAA
```

```
                                    *   F   L   L   H   E   P   S   F
13621  AGCTTCTGGACAAGGTAAAAAAAGCGAAATGTGCGTCTGATCCAAGCAAGTCCTTTAGGG
         S   R   S   L   T   F   F   A   F   H   A   D   S   G   L   L   D   K   L   P

13681  GTCGGGTAGCGGAAAGTTCTCCATCCAAATAGTTCAAGGATGACGCCGATATGCCCAACA
         R   T   A   S   L   E   G   D   L   Y   N   L   S   S   A   S   I   G   L   M

13741  TTAGGGATGAAATCAAGAATCTACCTTCAGCTATTAGTAGGTCTTGCAGGAAGCTGCAAC
         L   S   S   I   L   F   R   G   E   A   I   L   L   D   Q   L   F   S   C   C
                                       Start of 86 aa ORF, no homology
13801  AATCTTCCCGGTTCCCGAAGGTGCTCCAGAATGTCGGCGCAGACGATATAGT**CAT**AACGC
         D   E   R   N   G   F   T   S   W   F   T   P   A   S   S   I   T   M

13861  AGCTTCTGAAGATTGCGCGGAAGATGCTCCCGCTCCAGATCCCACTCCCATAAGTGACGA

13921  TAGAAGGGTCGTGCCGCTTCTAGTTCAATGTGATTGGCATCGATACCATCCACCACGCAG

13981  TGCAGACTTTCACTCAGGTACTTCCCCAACGTGCCTACACCACAACCAACGTCCAGGACA

14041  AGGCTGCCGGGCTGAATGCGTCGGGCAATCTTTGCAAGACTATCTTGGCCCTGAGGATCT

14101  ATCGGGGGACGATTATAAACATGCGGGGTATGTGGGACTGATTCATACGGGATCC
```

Notes

When joined at *Bam*HI site no meaningful ORFs detected
Also no meaningful ORFs if IS element deleted

98

References

Abdel-Monem, M., G. Taucher-Scholz, and M. Q. Klinkert. 1983. Identification of *Escherichia coli* DNA helicase I as the *traI* gene product of the F sex factor. Proc. Natl. Acad. Sci. **80:**4659-4663

Achtman, M., N. Willetts, and A. J. Clark. 1972. Conjugational complementation analysis of transfer-deficient mutants in F*lac* in *Escherichia coli*. J. Bacteriol. **110:**831-842

Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller. and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res. **25:**3308-3402

Anderson, L. B., A. V. Hertzel, and A. Das. 1996. *Agrobacterium tumefaciens* VirB7 and VirB9 form a disulfide-linked protein complex. Proc. Natl. Acad. Sci. **93:**8889-8894

Ando, T., Q. Xu, M. Torres, K. Kusugami, D. A. Israel, and M. J. Blaser. 2000. Restriction-modification system differences in *Helicobacter pylori* are a barrier to interstrain plasmid transfer. Mol. Microbiol. **37:**1052-1065

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. Mol. Microbiol. **13:**939-953

Argos, P. 1988. A sequence motif in many polymerases. Nucleic Acids. Res. **16:**9909-9916

Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. A. Smith, J. G. Seidman, and K. Struhl. 1989. Current Protocols in Molecular Biology. John Wiley, New York

Balzer, D., W. Pansegrau, and E. Lanka. 1994. Essential motifs of relaxase (TraI) and TraG proteins involved in conjugative transfer of plasmid RP4. J. Bacteriol. **176:**4285-4295

Bardwell, J. C. A. 1994. Building Bridges - Disulphide Bond Formation in the Cell. Mol. Microbiol. **14:**199-205

Baron, C., M. Llosa, S. Fhou, and P. C. Zambrynski. 1997. VirB1, a component of the T-complex transfer machinery of *Agrobacterium tumefaciens*, is processed to a C-terminal secreted product, VirB1*. J. Bacteriol. **179:**1203-1210

Barth,P. T., and N. J. Grinter. 1974. Comparison of the deoxyribonucleic acid molecular weights and homologies of plasmids conferring linked resistance to streptomycin and sulfonamides. J. Bacteriol. **120:**618-630

Bates, S., A. M. Cashmore, and B. M. Wilkins. 1998. IncP plasmids are unusually effective in mediating conjugation of *Escherichia coli* and yeast: involvement of the Tra2 mating system. J. Bacteriol. **180:**6538-6543

Battacharjee, M. K., and R. J. Meyer. 1991. A segment of a plasmid gene required for conjugal transfer encodes a site-specific, single stranded DNA endonuclease and ligase. Nucleic Acids Res. **19:**1129-1137

Battacharjee, M., X. M. Rao, and R. J. Meyer. 1992. Role of the origin of transfer in termination of strand transfer during conjugation. J. Bacteriol. **174:**6659-6665

Bayer, M., R. Eferl, G. Zellnig, K. Teferle, A. Dijkstra, G. Koraimann, and G. Högenauer. 1995. Gene *19* of plasmid R1 is required for both efficient conjugative DNA transfer and bacteriophage R17 infection. J. Bacteriol. **177:**4279-4288

Becker, E., and R. Meyer. 2000. Recognition of *oriT* for DNA processing at termination of a round of conjugal transfer. J. Mol. Biol. **300:**1067-1077

Becker, E., and R. Meyer. 2002. MobA, the DNA strand transferase of plasmid R1162: The minimal domain required for DNA processing at the origin of transfer. J. Biol. Chem. **277:**14575-14580

Berger, B. R., and P. J. Christie. 1993. The *Agrobacterium tumefaciens* virB4 gene product is an essential virulence protein requiring an intact nucleoside triphosphate-binding domain. J. Bacteriol. **175:**1723-1734

Berger, B. R., and P. J. Christie. 1994. Genetic complementation analysis of the *Agrobacterium tumefaciens virB* operon: *virB2* through *virB11* are essential virulence genes. J. Bacteriol. **176:**3646-3660

Bergquist, P. L. 1987. Incompatibility, p. 37-78. *In* K. G. Hardy (ed.), Plasmids: a practical approach. IRL Press, Oxford.

Beupre, C. E., J. Bohne, E. M. Dale, and A. N. Binns. 1997. Interactions between VirB9 and VirB10 membrane proteins involved in movement of DNA from *Agrobacterium tumefaciens* into plant cells. J. Bacteriol. **179:**78-89

Binns, A. N., Beaupré, and E. M. Dale. 1995. Inhibition of VriB-mediated transfer of diverse substrates from *Agrobacterium tumefaciens* by the IncQ plasmid RSF1010. J. Bacteriol. **177:**4890-4899

Black, M. T. 1993. Evidence that the catalytic activity of prokaryote leader peptidase depends upon the operation of a serine-lysine catalytic dyad. J. Bacteriol. **175:**4957-4961

Bolivar, F., R. L. Rodriguez, P. J. Greene, M. C. Betlach, H. L. Heyneker, and H. W. Boyer. 1977. Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. Gene **2:**95-113

Boyer, H. W., and D. Roulland-Dussoix. 1969. A complementation analysis of the restriction and modification DNA in *Escherichia coli*. J. Mol. Biol. **41**:459-472

Brasch, M. A., and R. J. Meyer. 1987. A 38 base-pair segment of DNA is required in *cis* for conjugative mobilization of broad host-range plasmid R1162. J. Mol. Biol. **198**:361-369

Bravo-Angel, A. M., V. Gloeckler. B. Hohn, and B. Tinland. 1999. Bacterial conjugation protein MobA mediates integration of complex DNA structures into plant cells. J. Bacterial. **181**:5758-5765

Brosius, J., and A. Holy. 1984. Regulation of ribosomal RNA promoters with a synthetic *lac* operator. Proc. Natl. Acad. Sci. **81**:6929-6933

Buchanan-Wollaston. V., J. E. Passiatore, and F. Cannon. 1987. The *mob* and *oriT* mobilization functions of a bacterial plasmid promote its transfer to plants. Nature **328**:172-175

Byrd. D. R., and S. W. Matson. 1997. Nicking by transesterification: the reaction catalyzed by a relaxase. Mol. Microbiol. **25**:1011-1022

Cabezón, E., E. Lanka, and F. de la Cruz. 1994. Requirements for mobilization of plasmid RSF1010 and ColE1 by the IncW plasmid R388: *trwB* and RP4 *traG* are interchangeable. J. Bacteriol. **176**:4455-4458

Cabezón, E., J. I. Sastre, and F. de la Cruz. 1997. Genetic evidence of a coupling role for the TraG protein family in bacterial conjugation. Mol. Gen. Genet. **254**:400-406

Chandler, P. M., and V. Krishnapillai. 1974. Phenotypic properties of R factors of *Pseudomonas aeruginosa*: R factors readily transferable between *Pseudomonas* and the Enterobacteriaceae. Genet. Res. Camb. **23**:239-250

Chang, A. C. Y., and S. N. Cohen. 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the p15A cryptic miniplasmid. J. Bacteriol. **134**:1141-1156

Chen, L., C. M. Li, and E. W. Nester. 2000. Transferred DNA (T-DNA)-associated proteins of *Agrobacterium tumefaciens* are exported independently. Proc. Natl. Acad. Sci. **97**:7545-7550

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

Christie, P. J. 2001. Type IV secretion: intracellular transfer of macromolecules by systems ancestrally related to conjugation machines. Mol. Microbiol. **40**:294-305

Clewell, D. B., P.K. Tomich, M. C. Gawron-Burke, A. E. Franke, M. Madion, J. H. Shaw, R. Wirth, F. An, M. Mori, C. Kitada, M. Fujino, and A. Suzuki. 1986. Sex-pheromone systems in *Streptococcus faecalis*. Banbury Report 24. Antibiotic resistance genes: ecology, transfer, and expression. Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.

Cook, D. M., and S. K. Farrand. 1992. The *oriT* region of the *Agrobacterium tumefaciens* Ti plasmid pTiC58 shares DNA sequence identity with the transfer origins of RSF1010 and RK2/RP4 and with the T-region borders. J. Bacteriol. **174**:6238-6246

Covacci, A., and R. Rappouli. 1993. Pertussis toxin export requires accessory genes located downstream from the pertussis toxin operon. Mol. Microbiol. **8**:429-434

Covacci, A., J. L. Telford, G. del Giudice, J. Parsonnet, and R. Rappuoli. 1999. *Helicobacter pylori* virulence and genetic geography. Science **289**:1328-1333

Craig-Mylius, K. A., and A. A. Weiss. 1999. Mutants in the *ptlA-H* genes of *Bordetella pertussis* are deficient for pertussis toxin secretion. FEMS Microbiol. Lett. **179**:479-484

Datta, N., R. W. Hedges, E. J. Shaw, R. B. Sykes, and M. N. Richmond. 1971. Properties of an R factor from *Pseudomonas aeruginosa*. J. Bacteriol. **108**:1244-1249

Daugelvičius, R., J. K. H. Bamford, A. M. Grahn, E. Lanka, andD. H. Bamford. 1997. The IncP plasmid-encoded cell envelope-associated DNA transfer complex increases cell permeability. J. Bacteriol. **179**:5195-5202

del Solar, G., R. Giraldo, M. J. Ruiz-Echevarría, M. Espinosa, and R. Díaz-Orejas. 1998. Replication and control of circular bacterial plasmids. Microbiol. Mol. Biol. Rev. **62**:434-464

Dijkstra, A. J., and W. Keck. 1996. Peptidoglycan as a barrier to transenvelope transport. J. Bacteriol. **178**:5555-5562

Di Laurenzio, L., L. S. Frost, and W. Paranchych. 1992. The TraM protein of the conjugative plasmid F binds to the origin of transfer of the F and ColE1 plasmids. Mol. Microbiol. **6**:2951-2959

Disque-Kochem. C., and B. Dreiseikelmann. 1997. The cytoplasmic DNA-binding protein TraM binds to the inner membrane protein TraD *in vitro*. J. Bacteriol. **179**:6133-6137

Doran, T., S. Loh, N. Firth, and R. Skurray. 1994. Molecular analysis of the F plasmid *traVR* region: *traV* encodes a lipoprotein. J. Bacteriol. **176**:4182-4186

Dorrington, R.A., and D.E. Rawlings. 1990. Characterization of the minimum replicon of the broad-host-range plasmid pTF-FC2 and similarity between pTF-FC2 and the IncQ plasmids. J. Bacteriol. **172**:5697-5705

Dorrington, R. A., S. Bardien and D. E. Rawlings. 1991. The broad-host-range plasmid pTF-FC2 requires a primase-like protein for autonomous replication in *Escherichia coli.* Gene **108:**7-14

Drolet, M., P. Zanga, and P. C. K. Lau. 1990. The mobilization and origin of transfer regions of a *Thiobacillus ferrooxidans* plasmid: relatedness to plasmids RSF1010 and pSC101. Mol. Microbiol. **4:**1381-1391

Dürrenberger, F., A. Crameri, B. Hohn, and Z. Koukolíková-Nicola. 1989. Covalently bound VirD2 protein of *Agrobacterium tumefaciens* protects the T-DNA from exonucleolytic degradation. Proc. Natl. Acad. Sci. **86:**9154-9158

Eisenbrandt, R., M. Kalkum, E. Lai, R. Lurz, C. I. Kado, and E. Lanka. 1999. Conjugative pili of IncP plasmids, and Ti plasmid T pilus are composed of cyclic subunits. J. Biol. Chem. **274:**22548-22555

Eisenbrandt, R., M. Kalkum, R. Lurz, and E. Lanka. 2000. Maturation of IncP pilin precursors resembles the catalytic dyad-like mechanism of leader peptidases. J. Bacteriol. **182:**6751-6761

Erickson, M. J., and R. J. Meyer. 1993. The origin of greater-than-unit-length plasmids generated during bacterial conjugation. Mol. Microbiol. **7:**289-298

Everett, R. and N. Willetts. 1980. Characterization of an *in vivo* system for nicking at the origin of conjugal DNA transfer of the sex factor F. J. Mol. Biol. **136:**129-150

Farizo, K. M., T. Huang, and D. L. Burns. 2000. Importance of holotoxin assembly in Ptl-mediated secretion of pertussis toxin from *Bordetella pertussis.* Infect. Immun. **68:**4049-4054

Farrand, S. K., I. Hwang, and D. M. Cook. 1996. The *tra* region of the nopaline type Ti plasmid is a chimera with elements related to the transfer system of RSF1010, RP4 and F. J. Bacteriol. **178:**4233-4247

Fekete, R. A., and L. S. Frost. 2000. Mobilization of chimeric *oriT* plasmids by F and R100-1: role of relaxosome formation in defining plasmid specificity. J. Bacteriol. **182:**4022-4027

Fernandez, D., T. T. Dang, G. M. Spudich, X. Zhou, B. R. Berger, and P. J. Christie. 1996. The *Agrobacterium tumefaciens* virB7 gene product, a proposed component of the T-complex transport apparatus, is a membrane-associated lipoprotein exposed at the periplasmic surface. J. Bacteriol. **178:**3156-3167

Firth, N., and R. Skurray. 1992. Characterization of the F plasmid bifunctional gene, *traG*. Mol. Gen. Genet. **232:**145-153

Firth, N., K. Ippen-Ihler, and R. A. Skurray. 1996. Structure and function of the F factor and mechanism of conjugation, p. 2377-2401 *In* F. C. Neidhardt, R. Curtiss III, J. L. Ingraham E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schhaecter, and H.E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology, 2nd ed. ASM Press, Washington, D.C.

Frey, J., and M. Bagdasarian. 1989. The biology of IncQ plasmids, p. 79-93 *In* C. M. Thomas (ed.), Promiscuous plasmids of gram-negative bacteria. Academic Press, Ltd., London, United Kingdom.

Frost, L S., and W. Paranchych. 1988. DNA sequence analysis of point mutations in *traA*, the F pilin gene, reveal two domains involved in F-specific bacteriophage attachment. Mol. Gen. Genet. **213**:134-139

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

Fulbrook, P. D., S. W. Elson, and B. Slocombe. 1970. R-factor mediated β–lactamase in *Pseudomonas aeruginosa*. Nature **226**:1054-1056

Fullner, K. J. 1998. Role of *Agrobacterium virB* genes in transfer of T complexes and RSF1010. J. Bacteriol. **180**:430-434

Fürste, J. P., W. Pansegrau, G. Ziegelin, M. Kröger, and E. Lanka. 1989. Conjugative transfer of promiscuous IncP plasmids: interaction of plasmid-encoded products with the transfer origin. Proc. Natl. Acad. Sci. **86**:1771-1775

Galli, D. M., J. Chen, K. F. Novak, and D. J. Leblanc. 2001. Nucleotide sequence and analysis of conjugative plasmid pVT745. J. Bacteriol. **183**:1585-1594

Gao, Q., Y. Luo, and R. C. Deonier. 1994. Initiation and termination of DNA transfer at F plasmid *oriT*. Mol. Microbiol. **11**:449-458

Gardner, M. N., S. M. Deane, and D. E. Rawlings. 2001. Isolation of a new broad-host-range IncQ-like plasmid, pTC-F14, from the acidophilic bacterium *Acidithiobacillus caldus* and analysis of the plasmid replicon. J. Bacteriol. **183**:3303-3309

Gardner, M. N. 2003. An investigation into the replicon of a broad host range mobilizable plasmid isolated from the moderately thermophillic bacterium *Acidithiobacillus caldus*. Ph.D. thesis, University of Stellenbosch, Stellenbosch, South Africa.

Giebelhaus, L. A., L. Frost, E. Lanka, E. P. Gormley, J. E. Davies, and B. Leskiw. 1996. The Tra2 core of the IncPα plasmid RP4 is required for intergeneric mating between *Escherichia coli* and *Streptomyces lividans*. J. Bacteriol. 178:6378-6381

Gilman, A. G. 1987. G proteins: transducers of receptor-generated signals. Annu. Rev. Biochem. **56**:615-649

Gomis-Rüth, F. X., G. Moncalián. F. de la Cruz, and M. Coll. 2002. Conjugative plasmid protein TrwB, an integral membrane type IV secretion system coupling protein: detailed structural features and mapping of the active site cleft. J. Biol. Chem. **277**:7556-7566

Gormley, E., and J. Davies. 1991. Transfer of plasmid RSF1010 by conjugation from *Escherichia coli* to *Streptomyces lividans* and *Mycobacterium smegmatis*. J. Bacteriol. **173**:6705-6708

Grahn, A. M., J. Haase, E. Lanka, and D. H. Bamford. 1997. Assembly of a functional phage PRD1 receptor depends on 11 genes of the IncP plasmid mating pair formation complex. J. Bacteriol. **179**:4733-4740

Grahn, A. M., J. Haase, D. H. Bamford, and E. Lanka. 2000. Components off the RP4 conjugative transfer apparatus form an envelope structure bridging inner membrane and outer membranes of donor cells: implications for related macromolecule transport systems. J. Bacteriol. **182**:1564-1574

Grandoso, G., M. Llosa, J. C. Zabala, and F. de la Cruz. 1994. Purification and biochemical characterization of TrwC, the helicase involved in plasmid R388 conjugal DNA transfer. Eur. J. Biochem. **226**:403-412

Greener, A., S. Lehman, and D. R. Helinski. 1992. Promoters of the broad host range plasmid RK2: analysis of transcription (initiation) in five species of gram-negative bacteria. Genetics **130**:27-36

Grossman, T. H., L. S. Frost, and P. M. Silverman. 1990. Structure and function of conjugative pili: monoclonal antibodies as probes for structural variants of F pili. J. Bacteriol. **172**:1174-1179

Gruss, A., H. F. Ross, and R. P. Novick. 1987. Functional analysis of a palindromic sequence required for normal replication of several staphylococcal plasmids. Proc. Natl. Acad. Sci. **84**:2165-2169

Guerry, P., J. van Embden, and S. Kalkow. 1974. Molecular nature of two nonconjugative plasmids carrying drug resistance genes. J. Bacteriol. **117**:619-630

Guiney, D. G., and D. R. Helinski. 1975. Relaxation complexes of plasmid DNA and protein. III. Association of protein with the 5' terminus of the broken DNA strand in the relaxed complex of plasmid ColE1. J. Biol. Chem. **250**:8796-8703

Guiney, D. G., Jr. 1984. Promiscuous transfer of drug resistance in gram-negative bacteria. J. Infect. Dis. *149*:320-329

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 propagation, and pilus production require the same Tra2 core components of a proposed DNA transport complex. J. Bacteriol. **177**:4779-4791

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

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 synthesis. J. Bacteriol. 179:5728-5735

Hallberg, K. B., and E. B. Lindström. 1994. Characterization of *Thiobacillus caldus* sp. nov., a moderately thermophilic acidophile. Microbiology 140:3451-3456

Hamilton. C. M., H. Lee, P. Li, D. M. Cook, K. R. Piper, S. B. von Bodman, E. Lanka, W. Ream, and S. K. Farrand. 2000. TraG from RP4 and TraG and VirD4 from Ti plasmids confer relaxosome specificity to the conjugal transfer system of pTiC58. J. Bacteriol. 182:1541-1548

Hanai, R., and J. Wang. 1993. The mechanism of sequence specific cleavage and strand transfer by φx174 gene A* protein. J. Biol. Chem. 268:23830-23836

Hapfelmeier, S., N. Domke, P. C. Zambryski, and C. Baron. 2000. VirB6 is required for stabilization of VirB5 and VirB3 and formation of VirB7 homodimers in *Agrobacterium tumefaciens*. J. Bacteriol. 182:4505-4511

Harris, R. L., A. Scholl, M. Conrad, M. Dresser, and P. Silverman. 1999. Interaction between the F plasmid TraA (F-pilin) and TraQ proteins. Mol. Microbiol. 34:780-791

Harris, R. L., V. Hombs, and P. M. Silverman. 2001. Evidence that TraV, TraK and TraB assemble into an envelope-spanning structure in *Escherichia coli*. Mol. Microbiol. 42:757-766

Harris, R. L., and P. M. Silverman. 2002. Roles of internal cysteines in the function, localization, and reactivity of the TraV outer membrane lipoprotein encoded by the F plasmid. J. Bacteriol 184:3126-3129

Henderson, D., and R. Meyer. 1996. The primase of broad-host-range plasmid R1162 is active in conjugal transfer. J. Bacteriol. 178:6888-6894

Henderson, D., and R. Meyer. 1999. The MobA-linked primase is the only replication protein of R1162 required for conjugal mobilization. J. Bacteriol. 181:2973-2978

Holloway, B. W., and M. H. Richmond. 1973. R-factors used for genetic studies in strains of *Pseudomonas aeruginosa* and their origin. Genet. Res. Camb. 21:103-105

Hormaeche, I., I. Alkorta, F. Moro, J. M. Valpuesta, F. M. Goñi, and F. de la Cruz. 2002. Purification and properties of TrwB, a hexameric, ATP-binding integral membrane protein essential for R388 plasmid conjugation. J. Biol. Chem. 277:46456-46462

Howard, M. T., W. C. Nelson, and S. W. Matson. 1995. Stepwise assembly of a relaxosome at the F plasmid origin of transfer. J. Biol. Chem. **270**:28381-28386

Ilyina, T. V., and E. V. Koonin. 1992. Conserved sequence motifs in the initiator proteins for rolling circle DNA replication encoded by diverse replicons from eubacteria, eucaryotes and archaebacteria. Nucl. Acids. Res. **20**:3279-3285

Inamoto, S., H. Fukuda, T. Abo, and E. Ohtsubo. 1994. Site- and strand specific nicking at *oriT* of plasmid R100 in a purified system: enhancement of the nicking activity of TraI (helicase I) with TraY and IHF. J. Biochem. **116**:838-844

Inanda, T., D. L. Court, K. Ito, and Y. Nakamura. 1989. Conditionally lethal amber mutations in the leader peptidase gene of *Escherichia coli*. J. Bacteriol. **171**:585-587

Ingram, L. C., M. H. Richmond, and R. B. Sykes. 1973. Molecular characterization of the R factors implicated in the carbenicillin resistance of a sequence of Pseudomonas *aeruginosa* strains isolated from burns. Antimicrob. Agents. Chemother. **3**:279-288

Jagura-Burdzy, G., F. Khanim, C. A. Smith, and C. M. Thomas. 1992. Crosstalk between plasmid vegetative replication and conjugative transfer: repression of the *trfA* operon by *trbA* of broad host-range plasmid RK2. Nucleic Acids Res. **20**:3939-3944

Jagura-Burdzy, G., and C. M. Thomas. 1994. KorA protein of promiscuous plasmid RK2 controls a transcriptional switch between divergent operons for plasmid replication and conjugative transfer. Proc. Natl. Acad. Sci. **91**:10575-10597

Jagura-Burdzy, G., and C. M. Thomas. 1997. Dissection of the switch between genes for replication and transfer of promiscuous plasmid RK2: basis of the dominance of *trfAp* over *trbAp* and specificity for KorA in controlling the switch. J. Mol. Biol. **265**:507-518

Jin, S. M., W. M. Yan, and Z. N. Zang. 1992. Transfer of IncP plasmids to extremely acidophilic *Thiobacillus thiooxidans*. Appl. Environ. Microbiol. **58**:429-430

Jin, R., A. Rasooly, and R. P. Novick. 1996. *In vitro* inhibitory activity of RepC/C*, the inactivated form of the pT181 plasmid initiation protein, RepC. J. Bacteriol. **179**:141-147

Johnson, D. A., and N. S. Willetts. 1980. Construction and characterization of multicopy plasmids containing the entire F transfer region plasmid. Plasmid **4**:292-304

Katada, T., and M. Ui. 1982. Direct modification of the membrane adenylate cyclase system by islet-activating protein due to ADP-ribosylation of a membrane protein. Proc. Natl. Acad. Sci. **79**:3129-3133

Khan, S. A. 1997. Rolling-circle replication of bacterial plasmids. Microbiol. Mol. Biol. Rev. **61**:442-455

Kim, K., and R. J. Meyer. 1989. Unidirectional transfer of broad host-range plasmid R1162 during conjugative mobilization. Evidence for genetically distinct events at *oriT*. J. Mol. Biol. **208**:501-505

Klimke, W. A., and L. S. Frost. 1998. Genetic analysis of the role of the transfer gene TraN of the F plasmid in mating pair stabilization during conjugation. J. Bacteriol. **180**:4036-4043

Koonin,E. V., and T. V. Ilyina. 1993. Computer-assisted dissection of rolling circle DNA replication. BioSystems **30**:241-268

Kotilainen, M. M., A. M. Grahn, J. K. H. Bamford, and D. H. Bamford. 1993. Binding of an *Escherichia coli* double-stranded DNA virus PRD1 to a receptor coded by an IncP-type plasmid. J. Bacteriol. **175**:3089-3095

Kramer, M. G., S. A. Khan, and M. Espinosa. 1997. Plasmid rolling circle replication: identification of the RNA polymerase-directed primer RNA and requirement of DNA polymerase I for lagging strand initiation. EMBO J. **16**:5784-5795

Krause, S., M. Bárcena, W. Pansegrau, R. Lurz, J. M. Carazo, and E. Lanka. 2000a. Sequence-related protein export NTP-ases encoded by the conjugative transfer region of RP4 and by the *cag* pathogenicity island of *Helicobacter pylori* share similar hexameric ring structures. Proc. Natl. Acad. Sci. **97**:3067-3072

Krause, S., W. Pansegrau, R. Lurz, F. de la Cruz, and E. Lanka. 2000b. Enzymology of type IV macromolecules secretion systems: the conjugative transfer regions of plasmids RP4 and R388 and the *cag* pathogenicity island of *Helicobacter pylori* encode structurally and functionally related nucleoside triphosphate hydrolases. J. Bacteriol. **182**:2761-2770

Kuipers, E. J., D. A. Israel, J. G. Kusters, and M. J. Blaser. 1998. Evidence for a conjugation-like mechanism of DNA transfer in *Helicobacter pylori*. J. Bacteriol. **180**:2901-2905

Kulpa, C. F., M. T. Roskey, and M. T. Travis. 1983. Transfer of plasmid RP1 into chemolithotrophic *Thiobacillus neapolitanus*. J. Bacteriol. **156**:434-436

Kumar, R. B., Y. H. Xie, and A. Das. 2000. Subcellular localization of the *Agrobacterium tumefaciens* T-DNA transport pore proteins: VirB8 is essential for the assembly of the transport pore. Mol. Microbiol. **36**:608-617

Kumar, R. B., and A. Das. 2001. Functional analysis of the *Agrobacterium tumefaciens* T-DNA transport pore protein VirB8. J. Bacteriol. **183**:3636-3641

Kupelweiser, G., M. Schwab, G. Högenauer, G. Koraimann, and E. Zechner. 1998. Transfer protein TraM stimulates TraI-catalyzed cleavage of the transfer origin of plasmid R1 *in vivo*. J. Mol. Biol. **275**:81-94

L'Abée-Lund, T. M., and H. Sørum. 2002. A global non-conjugative Tet C plasmid pRAS3, from *Aeromonas salmonicida*. Plasmid **47**:172-181

Lanka, E., and P. T. Barth. 1981. Plasmid RP4 specifies a deoxyribonucleic acid primase involved in its conjugal transfer and maintenance. J. Bacteriol. **148**:769-781

Lessl, M., D. Balzer, R. Lurz, V. L. Waters, D. G. Guiney, and E. Lanka. 1992a. Dissection of IncP conjugative plasmid transfer: defenition of the transfer region Tra2 by mobilization of the Tra1 region in *trans*. J. Bacteriol. **174**:2493-2500

Lessl, M., D. Balzer, W. Pansegrau, and E. Lanka. 1992b. Sequence similarities between the RP4 Tra2 and the Ti VirB region strongly support the conjugation model for T-DNA transfer. J. Biol. Chem. **267**:20471-20480

Lessl, M., D. Balzer, K. Weyrauch, and E. Lanka. 1993. The mating pair formation system of plasmid RP4 defined by RSF1010 mobilization and donor-specific phage propagation. J. Bacteriol. **175**:6415-6425

Li, Z., H. Hiasa, U. Kumar, and R. J. DiGate. 1997. The *traE* gene of plasmid RP4 encodes a homologue of *Escherichia coli* DNA topoisomerase III. J. Biol. Chem. **272**:19582-29587

Liu, Z., N. Guiliani, C. Appia-Ayme, F. Borne, J. Ratouchniak, and V. Bonnefoy. 2000. Construction and characterization of a *recA* mutant of *Thiobacillus ferrooxidans* by marker exchange mutagenesis. J. Bacteriol. **182**:2269-2276

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

Llosa, M., G. Grandoso, M. A. Hernando, and F. de la Cruz. 1996. Functional domains of protein TrwC of plasmid R388: dissected DNA strand transferase and DNA helicase activities reconstitute protein function. J. Mol. Biol. **264**:56-67

Llosa, M., F. X. Gomis-Rüth, M. Coll, and F. de la Cruz. 2002. Bacterial conjugation: a two-step mechanism for DNA transport. Mol. Microbiol. **45**:1-8

Lowbury, E. J. L., A. Kidson, H. A. Lilly, G. A. Ayliffe, and R. J. Jones. 1969. Sensitivity of *Pseudomonas aeruginosa* to antibiotics: emergence of strains highly resistant to carbenicillin. Lancet **2**:448-452

Luo, Y., Q. Gao, and R. C. Deonier. 1994. Mutational and physical analysis of F plasmid TraY protein binding to *oriT*. Mol. Microbiol. **11**:459-469

Nash, J., and V. Krishnapillai. 1988. Role of IncP-1 plasmid primase in conjugation between *Pseudomonas* species. FEMS Microbiol. Lett. **49**:257-260

Nelson, W. C., M. T. Howard, J. A. Sherman, and S. W. Matson. 1995. The *traY* gene product and integration host factor stimulate *Escherichia coli* DNA helicase I-catalyzed nicking at the F plasmid *oriT*. J. Biol. Chem. 270:28374-28380

Novick, R. R. 1980. Plasmids. *Scientific American* Dec. 77-90.

Majdalani, N., D. Moore, S. Maneewannakul, and K. Ippen-Ihler. 1996. Role of the propilin leader peptide in the maturation of F-pilin. J. Bacteriol. **178:**3748-3754

Maneewannakul, K., S. Maneewannakul, and K. Ippen-Ihler. 1993. Synthesis of F-pilin. J. Bacteriol. **175:**1384-1391

Manning, P. A., G. Morelli, and M. Achtman. 1981. *traG* protein of the F sex factor of *Escherichia coli* K-12 and its role in conjugation. Proc. Natl. Acad. Sci. **78:**7487-7491

Margensin, R., and F. Schinner. 1997. Heavy metal resistant *Arthrobacter* sp. — a tool for studying conjugal plasmid transfer between gram-negative and gram-positive bacteria. J. Basic Microbiol. **37:**217-227

Matson, S. W., W. C. Nelson, and B. S. Morton. 1993. Characterization of the reaction product of the *oriT* nicking reaction catalyzed by *Escherichia coli* DNA helicase I. J. Bacteriol.. **175:**2599-2606

Matson, S. W., J. K. Sampson, and D. R. N. Byrd. 2001. F plasmid conjugative transfer: the TraI helicase is essential for DNA strand transfer. J. Biol. Chem. **276:**2372-2379

Meyer, R., M. Hinds, and M. Brasch. 1982. Properties of R1162, a broad host range, high copy number plasmid. J. Bacteriol. **150:**552-562

Meyer, R. 1989. Site-specific recombination at *oriT* of plasmid R1162 in the absence of conjugative transfer. J. Bacteriol. **171:**799-806

Minkley, E. G., S. Polen, C. C. Brinton, and K. Ippen-Ihler. 1976. Identification of the structural gene for F-pilin. J. Mol. Biol. **108:**111-121

Miranda, A., G. Janssen, L. Hodges, E. G. Peralta, and W. Ream. 1992. *Agrobacterium tumefaciens* transfers extremely long T-DNAs by unidirectional mechanism. J. Bacteriol. **174:**2288-2297

Moncalián, G., E. Cabezón, I. Alkorta, M. Valle, F. Moro, J. M. Valpuesta, F. M. Goñi, and F. de la Cruz. 1999. Characterization of ATP and DNA binding activities of TrwB, the coupling protein essential in plasmid R388 conjugation. J. Biol. Chem. **274:**36117-36124

Moore, D., B. A. Sowa, and K. Ippen-Ihler. 1981. Location of an F-pilin pool in the inner membrane. J. Bacteriol. **146:**251-259

Moore, D., C. M. 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. J. Bacteriol. **175:**1375-1383

Moscoso, M., G. del Solar, amd M. Espinosa. 1995. Specific nicking-closing activity of the initiator of replication protein RepB of plasmid pMV158 on supercoiled or single-stranded DNA. J. Biol. Chem. **270:**3772-3779

Motallebi-Veshareh, M., D. Balzer, E. Lanka, G. Jagura-Burdzy, and C. M. Thomas. 1992. Conjugative transfer functions of broad host-range plasmid RK2 are coregulated with vegetative replication. Mol. Microbiol. **6:**907-920

Noirot-Gros,M. F., V. Bidnenko, and S. D. Ehrlich. 1994. Active site of the replication protein of the rolling circle plasmid pC194. EMBO J. **13:**4412-4420

Novotny, C. P., and P. Fives-Taylor. 1974. Retraction of F pili. J. Bacteriol. **117:**1306-1311

O'Callaghan, D., C. Cazevieille, A. Allardet-Servent, M. L. Boschiroli, G. Bourg, V. Foulongne, P. Frutos, Y. Kulakov, and M. Ramuz. 1999. A homologue of the *Agrobacterium tumefaciens* VirB and *Bordetella petrussis* Ptl type IV secretion systems is essential for intracellular survival of *Brucella suis*. Mol. Microbiol. **33:**1210-1220

Osborn, M., S. Bron, N. Firth, S. Holsappel, A. Huddleston, R. Kiewiet, W. Meijer, J. Seegers, R. Skurray, P. Terpstra, C. M. Thomas, P. Thorsted, E. Tietze, and S. L. Turner. 2000. The evolution of bacterial plasmids, p. 301-361. *In* C. M. Thomas (ed.), The horizontal gene pool, bacterial plasmids and gene spread. Harwood Academic Publishers, Amsterdam, The Netherlands.

Paiva, W., T. Grossman, and P. Silverman. 1992. Characterization of F-pilin as an inner membrane component of *Escherichia coli* K12. J. Biol. Chem. **267:**26191-26197

Paiva,W., and P. Silverman. 1996. Effects of F encoded components and F-pilin domains on the synthesis and membrane insertion of TraA$^{\prime}$-PhoA fusion proteins. Mol. Microbiol. **19:**1277-1286

Panicker, M. M., and E. G. Minkley, Jr. 1992. Purification and properties of the F sex factor TraD protein, an inner membrane conjugal transfer protein. J. Biol. Chem. **267:**12761-12766

Pansegrau. W., G. Ziegelin, and E. Lanka. 1990a. Covalent association of the *traI* gene product of plasmid RP4 with the 5'-terminal nucleotide at the relaxation nick site. J. Biol. Chem. **265:**10637-10644

Pansegrau. W., D. Balzer, V. Kruft, R. Lurz, and E. Lanka. 1990b. *In vitro* assembly of relaxosomes at the transfer origin of plasmid RP4. Proc. Natl. Acad. Sci. **87:**6555-6559

Pansegrau, W.. and E. Lanka. 1991. Common sequence motifs in DNA relaxases and nick regions from a variety of DNA transfer systems. Nucleic Acids. Res. **19:**3455

Pansegrau, W., F. Schoumacher, B. Hohn, and E. Lanka. 1993. Site-specific cleavage and joining of single-stranded DNA by VirD2 protein of *Agrobacterium tumefaciens* Ti plasmids: analogy to bacterial conjugation. Proc. Natl. Acad. Sci. **90:**11538-11542

Pansegrau, W., W. Schröder, E. Lanka. 1994a. Concerted action of three distinct domains in the DNA cleaving-joining reaction catalyzed by relaxase (TraI) of conjugative plasmid RP4. J. Biol. Chem. **269:**2782-2789

Pansegrau, W., E. Lanka, P. T. Barth, D. H. Figurski, D. G. Guiney, D. Haas, D. R. Helinski, H. Schwab, V. A. Stanisch, and C. M. Thomas. 1994b. Complete sequence of Birmingham IncPα plasmids, compilation and comparative analysis. J. Mol. Biol. **239:**623-663

Pansegrau, W., and E. Lanka. 1996. Mechanisms of initiation and termination reactions in conjugative DNA processing. J. Biol. Chem. **271:**13068-13076

Perwez, T., and R. Meyer. 1996. MobB protein stimulates nicking at the R1162 origin of transfer by increasing the proportion of complexed plasmid DNA. J. Bacteriol. **178:**5762-5767

Perwez, T., and R. Meyer. 1999. Stabilization of the relaxosome and stimulation of conjugal transfer are genetically distinct functions of the R1162 protein MobB. J. Bacteriol. **181:**2124-2131

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

Pölzleitner, E., E. Zechner, W. Renner, R. Fratte, B. Jauk, G. Högenauer, and G. Koraimann. 1997. TraM of plasmid R1 controls transfer gene expression as an integrated control element in a complex regulatory network. Mol. Microbiol. **25:**495-507

Rasooly, A., and R. P. Novick. 1993. Replication-specific inactivation of the pT181 plasmid initiator protein. Science **262:**1048-1050

Rawlings, D. E., I. Pretorius, and D. R. Woods. 1984. Expression of a *Thiobacillus ferrooxidans* origin of replication in *Escherichia coli*. J. Bacteriol. **159:**737-738

Rawlings, D. E., N. J. Coram, M. N. Gardner, and S. M. Deane. 1999. *Thiobacillus caldus* and *Leptospirillum ferrooxidans* are widely distributed in continuous-flow biooxidation tanks used to treat a variety of metal-containing ores and concentrates, p. 777-786. *In* R. Amils and A. Ballester (ed.), Biohydrometallurgy and the environment toward the mining of the 21st century. Elsevier Press, Amsterdam.

Rawlings. D. E., and E. Tietze. 2001. Comparitive biology of the IncQ and IncQ-like plasmids. Microbiol. Mol. Biol. Rev. **65:**481-496

Rees, C. E. D., and B. M. Wilkins. 1990. Protein transfer into the recipient cell during bacterial conjugation: studies with F and RP4. Mol. Microbiol. **4:**1199-1205

Rice, P. A., S. Yang, K. Mizuuchi, and H. A. Nash. 1996. Crystal structure of an IHF-DNA complex: a protein induced DNA U-turn. Cell **87:**1295-1306

Rohrer, J. 1992. Characterization of the mobilization region of the broad host-range plasmid pTF-FC2. Ph. D. thesis, University of Cape Town, Cape Town, South Africa.

Rohrer, J., and D. E. Rawlings. 1992. Sequence analysis and characterization of the mobilization region of a broad-host-range plasmid, pTF-FC2, isolated from *Thiobacillus ferrooxidans*. J. Bacteriol. **174:**6230-6237

Rohrer, J., and D. E. Rawlings. 1993. Regulation of mobilization of the broad-host-range plasmid pTF-FC2. Mol. Microbiol. **9:**1051-1059

Sagulenko, V., E. Sagulenko, S. Jakubowski, E. Spudich, and P. J. Christie. 2001. VirB7 lipoprotein is exocellular and associates with the *Agrobacterium tumefaciens* T pilus. J. Bacteriol. **183:**3642-3651

Sakai, D., and T. Komano. 2002. Genes required for plasmid R64 thin-pilus biogenesis: identification and localization of products of the *pilK, pilM, pilO, pilP, pilR,* and *pilT* genes. J. Bacteriol. **184:**444-451

Salmond, G. P. C. 1994. Secretion of extracellular virulence factors by plant pathogenic bacteria. Annu. Rev. Phytopathol. **32:**181-200

Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular Cloning: A laboratory manual, 2nd edition. Cold Spring Harbor Laboratory Press: New York.

Samuels, A. L., E. Lanka, and J. E. Davies. 2000. Conjugative junctions in RP4-mediated mating of *Escherichia coli*. J. Bacteriol. **182:**2709-2715

Sastre, J. I., E. Cabezón, and F. de la Cruz. 1998. The carboxyl terminus of protein TraD adds specificity and efficiency to F-plasmid conjugative transfer. J. Bacteriol. **180:**6039-6042

Scherzinger, E., V. Haring, R. Lurz, and S. Otto. 1991. Plasmid RSF1010 DNA replication *in vitro* promoted by purified RSF1010 RepA, RepB and RepC proteins. Nucleic Acids Res. **19:**1203-1211

Scherzinger, E., V. Kruft, and S. Otto. 1993. Purification of the large mobilization protein of plasmid RSF1010 and characterization of its site-specific DNA-cleaving/DNA-joining activity. Eur. J. Biochem. **217:**929-938

Scholtz, P., V. Haring, B. Wittmann-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

Schröder, G., S. Krause, E. L. Zechner, B. Traxler, H. Yeo, R. Lurz, G. Waksman, and E. Lanka. 2002. TraG-like proteins of DNA transfer systems and of the *Helicobacter pylori* type IV secretion system: inner membrane gate for exported substrates? J. Bacteriol. **184**:2767-2779

Seegers, J. F. M. L., A. C. Zhao, S. A. Khan, L. E. Pearce, W. J. J. Meijer, G. Venema, and S. Bron. 1995. Structural and functional analysis of the single-stranded origin of replication from the lactococcal plasmid pWVO1. Mol. Gen. Genet. **249**:43-50

Sherman, J. A., and S. W. Matson. 1994. *Escherichia coli* DNA helicase I catalyzes a sequence specific cleavage ligation reaction at the F plasmid origin of transfer. J. Biol. Chem. **269**:26220-26226

Shine, J., and L. Dalgarno. 1974. The 3'-terminal sequence of *E. coli* 16S rRNA: complementary to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. **71**:1342-1346

Shirasu, K., and C. I. Kado. 1993. Membrane location of the Ti plasmid VirB proteins involved in the biosynthesis of a pilin-like conjugative structure on *Agrobacterium tumefaciens*. FEMS Microbiol. Lett. **111**:287-294

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

Sieira, R., D. J. Comerci, D. O. Sanchez, and R. A. Ugalde. 2000. A homologue of an operon required for DNA transfer in *Agrobacterium* is required in *Brucella abortus* for virulence and intracellular multiplication. J. Bacteriol. **182**:4849-4855

Simon, R., U. Priefer, and A. Puhler. 1983. A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in gram-negative bacteria. Biotechnology **1**:784-790

Sowa, B. A., D. Moore, and K. Ippen-Ihler. 1983. Physiology of F-pilin synthesis and utilization. J. Bacteriol. **153**:962-968

Spudich, G. M., D. Fernandez, X. Xhou, and P. J. Christie. 1996. Intermolecular disulfide bonds stabilize VirB7 homodimers and VirB7/VirB9 heterodimers during biogenesis of the *Agrobacterium tumefaciens* T-complex transport apparatus. Proc. Natl. Acad. Sci. **93**:7512-7517

Stahl, L. E., A. Jacobs, and A. N. Binns. 1998. The conjugal intermediate of plasmid RSF1010 inhibits *Agrobacterium tumefaciens* virulence and VirB-dependent export of VirE2. J. Bacteriol. **180**:3933-3939

Strack, B., M. Lessl, R. Calendar, and E. Lanka. 1992. A common sequence motif, -E-G-Y-A-T-A-, identified within the primase domains of plasmid-encoded I- and P-type DNA primases and the α protein of the *Escherichia coli* satellite phage P4. J. Biol. Chem. **267**:13062-13072

Sykes, R.B., and M. H. Richmond. 1970. Intergenic transfer of a β-lactamase between *Ps. aeruginosa* and *E. coli*. Nature **22**:952-954

Szpirer, C. Y., M. Faelen, and M. Couturier. 2000. Interaction between the RP4 coupling protein TraG and the pBHR1 mobilization protein Mob. Mol. Microbiol. **37**:1283-1292

Tamura, M., K. Mogimori, S. Murai, M. Yajima, K. Ito, T. Katada, M. Ui, and S. Ishii. 1982. Subunit structure of Islet-Activating protein, pertussis toxin, in conformity with the A-B model. Biochemistry **21**:5516-5522

Tietze, E. 1998. Nucleotide sequence and genetic characterization of the novel IncQ-like plasmid pIE1107. Plasmid **39**:165-181

Thomas, C. D., D. F. Balson, and W. V. Shaw. 1990. *In vitro* studies of the initiation of staphylococcal plasmid replication. J. Biol. Chem. **265**:5519-5530

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 evolution and organization of the IncP backbone. J. Mol. Biol. **282**:969-990

Tsai, M. M., Y. H. Fu, and R. C. Deonier. 1990. Intrinsic bends and integration host factor binding at the F plasmid *oriT*. J. Bacteriol. **172**:4603-4609

Verdino, P., W. Keller, H. Strohmaier, K. Bischof, H. Lindner, and G. Koraimann. 1999. The essential transfer protein TraM binds to DNA as a tetramer. J. Biol. Chem. **274**:37421-37428

Vergunst, A. C., B. Schrammeijer, A. den Dulk-Ras, C. M. de Vlaam, T. J. Regensburg-Tunik, and P. J. Hooykaas. 2000. VirB/D4 dependent protein translocation from *Agrobacterium* to plants. Science **290**:979-982

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**:873-876

Walker, J. E., M. Saraste, M. J. Runswick, and N. J. Gay. 1082. Distantly related sequences in the α- and β-subunits of ATP synthase, myosin, kinases, and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J. **1**:945-951

Waters, V. L., B. Strack, W. Pansegrau, E. Lanka, and D. G. Guiney. 1992. Mutational analysis of essential IncPα plasmid transfer genes *traF* and *traG* and involvement of *traF* in phage sensitivity. J. Bacteriol. **174**:6666-6673

Waters, V. L., and D. G. Guiney. 1993. Processes at the nick region link conjugation, T-DNA transfer and rolling circle replication. Mol. Microbiol. **9:**1123-1130

Whittle, G., M. E. Katz, E. H. Clayton, and B. F. Cheetham. 2000. Identification and characterization of a native *Dichelobacter nodosus* plasmid, pDN1. Plasmid **43:**230-234

Wilkins, B. M., and A. T. Thomas. 2000. DNA-independent transport of plasmid primase protein between bacteria by the I1 conjugation system. Mol. Microbiol. **38:**650-657

Wu, J. H., and K. Ippen-Ihler. 1989. Nucleotide seuqnce of *traQ* and adjacent loci in the *Escherichia coli* K-12 F plasmid transfer operon. J. Bacteriol. **171:**213-221
Yanisch-Perron, C., J. Vieira, and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequence of the M13mp18 and pUC19 vectors. Gene **33:**103-119

Yarranton, G. T., and M. L. Gefter. 1979. Enzyme-catalyzed DNA unwinding: studies on *Escherichia coli* rep protein. Proc. Natl. Acad. Sci. **76:**1658-1662

Yoshida, K., T. Takegami, A. Katoh, M. Nishikawa, and M. Albritton. 1997. Construction of a novel conjugative plasmid harboring a GFP reporter gene and its introduction into animal cells by transfection and trans-kingdom conjugation. Nucleic Acids Symp. Ser. **37:**157-158

Yoshida, T., S. Kim, and T. Komano. 1999. Twelve *pil* genes are required for biogenesis of the R64 thin pilus. J. Bacteriol. **181:**2038-2043

Zambryski, P. C. 1992. Chronicles from the *Agrobacterium*-plant cell DNA transfer story. Plant. Mol. Biol. **43:**465-490

Zatyka, M., G. Jagura-Burdzy, and C. M. Thomas. 1994. Regulation of transfer genes of promiscuous IncPα plasmid RK2: repression of Tra1 region transcription both by relaxosome proteins and the Tra2 regulator TrbA. J. Microbiol. **140:**2981-2990

Zechner, E. L., F. de la Cruz, R. Eisenbrandt, A. M. Grahn, G. Koraimann, E. Lanka, G. Muth, W. Pansegrau, C. M. Thomas, B. M. Wilkins, and M. Zatyka. 2000. Conjugative-DNA transfer processes, p. 87-174. *In* C. M. Thomas (ed.), The horizontal gene pool, bacterial plasmids and gene spread. Harwood Acedemic Publishers, Amsterdam, The Netherlands.

Zhang, S., and R. Meyer. 1997. The relaxosome protein MobC promotes conjugal plasmid mobilization by extending DNA strand separation to the nick site at the origin of transfer. Mol. Microbiol. **25:**509-516

Zhao, A. C., and S. A. Khan. 1996. An 18-bp sequence is sufficient for the termination of rolling circle replication of plasmid pT181. J. Bacteriol. **178:**5222-5228

Zhu, J., P. M. Oger, B. Schrammaijer, P. J. Hooykaas, S. K. Farrand, and S. C. Winans. 2000. The bases of crown gall tumorigenesis. J. Bacteriol. **182:**3885-3895

Ziegelin, G., J. P. Fürste, and E. Lanka. 1989. TraJ protein of plasmid RP4 binds to a 19-base pair invert sequence repetition within the transfer origin. J. Biol. Chem. **264:**11989-11994

Ziegelin, G., W. Pansegrau, R. Lurz, and E. Lanka. 1992. TraK protein of conjugative plasmid RP4 forms a specialized nucleoprotein complex with the transfer origin. J. Biol. Chem. **267:**17279-17286