Sequence Similarities between the RP4 Tra2 and the Ti VirB Region Strongly Support the Conjugation Model for T-DNA Transfer*

Monika Lessl, Dietmar Balzer, Werner Pansegrau, and Erich LankaS

From the Max-Planck-Institut für Molekulare Genetik, Abteilung Schuster, Ihnestrasse 73, D-1000 Berlin 33, Federal Republic of Germany

The conjugative transfer system of plasmid RP4 consists of two distinct regions, Tra1 and Tra2, separated by the Par/Mrs (Multimer resolution system) region, the flaA (fertility inhibition of IncW plasmids) locus, IS8 and the kanamycin resistance gene aphA. The Tra1 region contains the origin of transfer and mainly encodes functions of the DNA transfer and replication system. A detailed study of Tra1 has been carried out by nucleotide sequence analysis and subsequent purification and functional analysis of Tra proteins (Ziegelin et al., 1991; Miele et al., 1991; Guiney and Lanka, 1989; Wilkins and Lanka, 1992). However, functions encoded by Tra2 have not yet been characterized extensively. Recently, the precise dimensions of Tra2 were determined by deletion analysis (Lessl et al., 1992). Accordingly, the Tra2 core region, essential for intraspecific E. coli matings, was located between the RP4 coordinates 18.03-29.26 kb,1 flanked by the traA operon and the flaA region (Thomas and Helinski, 1989, Fong and Stanisich, 1989). It was proposed that functions encoded by Tra2 are mainly responsible for the formation of mating pairs between conjugating cells. However, it was shown recently that at least one function located in Tra1 is supposed to be involved in the formation of the mating pair system (Waters et al., 1992).

Mating requires the coordinated interaction of a number of cell-surface components (sex-pili, outer and inner membrane proteins) to form a so-called mating bridge between donor and recipient cells (Willets and Wilkins, 1984). Since this process is only poorly understood, a detailed study of Tra2 was carried out. The molecular characterization of genes encoded by Tra2 facilitates their functional analysis. To reveal the genetic organization of this region the nucleotide sequence was determined. Potential structural genes were identified by codon preference analysis and verified by expression using a phage T7-promoter system. Our results show that the Tra2 core region encompasses twelve orfs, most of which appear to encode membrane-associated proteins. A data base search resulted in the detection of a remarkable relationship of Tra2 proteins to proteins encoded by the VirB operon of the Ti plasmid. Since these proteins are proposed to form a membrane-spanning pore or channel to mediate the exit of T-DNA (Engström et al., 1987), a corresponding role for the RP4 Tra2 gene products might be assumed. Furthermore, evidence for an overall functional relationship between T-DNA transfer and conjugal DNA transfer are discussed.

EXPERIMENTAL PROCEDURES

Bacterial Strains, Bacteriophages, and Plasmids—E. coli strain XL1-Blue (Bullock et al., 1987) was used as host for plasmids and for the propagation of Mi3 mpl8/mpl9 derivatives. The abbreviations used are: kb, kilobase(s); MOPS, 3-(N-morpholino)propanesulfonic acid; orf, open reading frame; nt, nucleotide position(s); SD, Shine-Dalgarno.

*This work was financially supported by Sonderforschungsbereich Grant 344/B2 of the Deutsche Forschungsgemeinschaft (to E. L.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

This work is dedicated to Professor Heinz Schuster on the occasion of his 65th birthday.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) M93669.

To whom correspondence should be addressed. Tel.: 30-8307-242; Fax: 30-8307-383.

The conjugal transfer of plasmid RP4 consists of two distinct regions, Tra1 and Tra2, separated by the Par/Mrs (Multimer resolution system) region, the flaA (fertility inhibition of IncW plasmids) locus, IS8 and the kanamycin resistance gene aphA. The Tra1 region contains the origin of transfer and mainly encodes functions of the DNA transfer and replication system. A detailed study of Tra1 has been carried out by nucleotide sequence analysis and subsequent purification and functional analysis of Tra proteins (Ziegelin et al., 1991; Miele et al., 1991; Guiney and Lanka, 1989; Wilkins and Lanka, 1992). However, functions encoded by Tra2 have not yet been characterized extensively. Recently, the precise dimensions of Tra2 were determined by deletion analysis (Lessl et al., 1992). Accordingly, the Tra2 core region, essential for intraspecific E. coli matings, was located between the RP4 coordinates 18.03-29.26 kb,1 flanked by the traA operon and the flaA region (Thomas and Helinski, 1989, Fong and Stanisich, 1989). It was proposed that functions encoded by Tra2 are mainly responsible for the formation of mating pairs between conjugating cells. However, it was shown recently that at least one function located in Tra1 is supposed to be involved in the formation of the mating pair system (Waters et al., 1992).

Mating requires the coordinated interaction of a number of cell-surface components (sex-pili, outer and inner membrane proteins) to form a so-called mating bridge between donor and recipient cells (Willets and Wilkins, 1984). Since this process is only poorly understood, a detailed study of Tra2 was carried out. The molecular characterization of genes encoded by Tra2 facilitates their functional analysis. To reveal the genetic organization of this region the nucleotide sequence was determined. Potential structural genes were identified by codon preference analysis and verified by expression using a phage T7-promoter system. Our results show that the Tra2 core region encompasses twelve orfs, most of which appear to encode membrane-associated proteins. A data base search resulted in the detection of a remarkable relationship of Tra2 proteins to proteins encoded by the VirB operon of the Ti plasmid. Since these proteins are proposed to form a membrane-spanning pore or channel to mediate the exit of T-DNA (Engström et al., 1987), a corresponding role for the RP4 Tra2 gene products might be assumed. Furthermore, evidence for an overall functional relationship between T-DNA transfer and conjugal DNA transfer are discussed.

EXPERIMENTAL PROCEDURES

Bacterial Strains, Bacteriophages, and Plasmids—E. coli strain XL1-Blue (Bullock et al., 1987) was used as host for plasmids and for the propagation of Mi3 mpl8/mpl9 derivatives. The abbreviations used are: kb, kilobase(s); MOPS, 3-(N-morpholino)propanesulfonic acid; orf, open reading frame; nt, nucleotide position(s); SD, Shine-Dalgarno.
The RP4 Tra2 Region

Fig. 1. Sequencing strategy of the RP4 Tra2 region. The horizontal arrows indicate the extent and direction of readable sequence information on the autoradiograms of sequencing gels. The numbers below the black line refer to the kilobase scale of the RP4 standard map (Lanka et al., 1983). Restriction endonuclease sites have been verified by appropriate enzyme digestion.

DNA Sequence Analysis of the RP4 Tra2 Region—The nucleotide sequence of the RP4 Tra2 region was determined on both strands ranging from restriction sites SplI (19.70 kb) to XhoI (31.23 kb). The sequencing strategy is shown in Fig. 1. In order to provide the complete nucleotide sequence of the RP4 Tra2 region, we present sequence information between restriction sites EarI (18.03 kb) and XhoI (31.23 kb; Fig. 2).

The overall GC content of the sequenced region is about 63%, which is slightly less than that of the RP4 Tra1 region (65%, Ziegelin et al., 1991). However, there are local differences: the coding region of trbA exhibits a GC content of only 55%, whereas trbH, L, M, N, and O show a GC content of 65-66%.

Putative promoter sequences, identified as σ70 consensus sequences (according to the algorithm of Harley and Reynolds, 1987) are shown in Fig. 2. The consensus sequences P2, P3, and P4 were confirmed by RNA-polymerase binding studies in the electron microscope (Lessa et al., 1992). No additional promoter-like sequence could be identified up to the Sall site (34.4 kb), neither by computer search, nor by RNA-polymerase binding assays. This is consistent with our assumption that the entire Tra2 region is transcribed in clockwise direction as a single transcriptional unit (Lessa et al., 1992). Examination of the nucleotide sequence for the occurrence of rho-independent transcriptional terminators (according to the algorithm of Brendel and Trifonov, 1984) revealed a typical GC-rich segment of dyad symmetry (Fig. 2, nt 13131-13155) followed by a short stretch of thymidine residues. The free energy of this hairpin structure is ΔG = −18.1 kcal/mol. Since no promoter consensus sequence could be identified near this terminator-like structure, a role in the transcriptional regulation of the downstream genes might be assumed. Binding of an unknown factor might allow the RNA-polymerase to read through the termination signal, a mechanism known as antitermination (Rosenberg and Court, 1979).

Binding sites for RP4 key regulatory proteins KorA and KorB (Smith et al., 1984; Young et al., 1985; Balzer et al., 1992) are indicated in Fig. 2 (Oa, Ob). Two of the KorB binding sites (Fig. 2, nt 176–188 and 627–639) are located near promoter consensus sequences (P3, P4), whereas the other two are located either within an orf (trbJ, Fig. 2, nt 8728–8740) or in between two orfs (trbN and trbO, Fig. 2, nt 12652–12664). Specific binding of KorB protein to these Oa sequences has been shown in all cases (Balzer et al., 1992), assuming that KorB might play an important role in the regulation of the Tra2 gene expression. Recent sequences similarities between TraB, KorA, and other regulatory proteins were found, suggesting that TraB acts as a transcriptional regulator (Jagura-Burdzy et al., 1992). Confirmation of this assumption was achieved by experiments demonstrating that TraB represses transcription of both the trfA and the Tra2 operons. This indicates that vegetative plasmid replication and conjugative DNA transfer might be coregulated. Furthermore, a 23-base pair inverted repeat structure at nt 118–141 (Fig. 2), which might function as a protein recognition site, could be detected.

Parts of the sequence presented here were also determined by other groups (Fig. 2). A comparison revealed differences to our data at several positions: next to nt 5195 the G, to nt 5206 the C, and to nt 5445 the C is missing, resulting in three frame-shifts. This alters the reading frame of trbF (Fig. 2) between nt 5195 and 5445. However, the coding preference...
The nucleotide sequence shown starts at -35 and ends at -10 upstream of the ORF. The ORF begins with the start codon ATG. The sequence contains the sequence of interest, which is underlined.

[Image of the nucleotide sequence with underlines indicating the start and stop of the ORF.]
The RP4 Tra2 Region

Fig. 2—continued
plot (Gribsov et al., 1984) of the altered trbF orf is quite unusual for RP4 coding regions (data not shown, Ziegelin et al., 1991). Since the codon preference plot of the trbF orf presented here coincides with that of RP4 coding regions and nucleotide sequence of both strands did not reveal any ambiguity, we suggest that the nucleotide sequence shown here is correct. Additional differences identified were as follows: nucleotides CATCGT appear next to C at nt 6055, the triplet GCG next to A at nt 9830 is missing, and at nt 13071 a C is correct. Additional differences identified were as follows:

| 9473-9498 |
| 9689-9714 |
| 6790-6815 |
| 1820-1845 |
| 212-237 |

*G-U base pairing was considered.*
Signal peptidase cleavage sites predicted for the gene products of the RP4 Tra2 region. Potential protein export signal sequences of predicted orfs are shown (Fig. 2). The amino acids of the signal sequences are indicated by capital letters (von Heijne et al., 1986). The signal sequence signals are potential candidates for the labeling procedure. For labeling host cells, carrying the RP4 plasmid pDB7756 or appropriate deletion derivatives were applied (Fig. 3). These plasmids, except pDB775, still carry the original RP4 Tra2 control region (Fig. 3). Since stable maintenance of Tra2 requires kor functions, the korA-korB region was supplied in order to verify the predicted orfs of Tra2. Proteins were specifically labeled using the phage T7 expression vector system (Tabor et al., 1987). Because Tra2 protein sequences contain either methionine or cysteine in internal positions, all of them are potential candidates for the labeling procedure. For labeling host cells, carrying the RP4 plasmid pDB7756 or appropriate deletion derivatives were applied (Fig. 3). These plasmids, except pDB775, still carry the original RP4 Tra2 control region (Fig. 3). Since stable maintenance of Tra2 requires kor functions, the korA-korB region was supplied in trans (pML50) in the host cell, carrying pT7-7 derivatives (Lessl et al., 1992).

Assignment of Tra2 proteins to orfs was carried out considering two parameters: (i) apparent molecular masses of protein bands visible on the autoradiogram were compared to calculated molecular masses of predicted Tra2 proteins and (ii) bands absent in extracts of cells harboring deletion derivatives of pDB7756 were related to gene products encoded by the deleted orfs. Proteins labeled by [55S]methionine/cysteine were visualized by autoradiography of the SDS-polyacrylamide gel (Fig. 5). In extracts of cells harboring pT7-7 and pML50, no signal could be detected on the autoradiogram (Fig. 5). Thus, the labeling procedure is specific for Tra2-encoded proteins. Applying the rules, mentioned above, 11 out of 15 orfs shown in Fig. 2 could be identified tentatively. Bands migrating at molecular masses of 90, 61.5, and 58 kDa were related to orfs trbE, I, and L, encoding the three largest Tra2 gene products (predicted sizes: TrbE, 94 kDa; TrbI, 49 kDa; TrbL, 52 kDa). Additionally, in extracts of cells harboring pDB7756ΔA a truncated trbE gene product was detectable (TrbE*, 70 kDa, Fig. 5), because in this derivative 747 nucleotides of the trbE orf were deleted.

Two bands of 35.5 and 31 kDa (Fig. 5) might correspond to TrbG and TrbF (predicted sizes: TrbG, 33 kDa; TrbF, 27 kDa). Beneath the band, tentatively assigned to TrbG an additional band was visible, possibly representing TrbG lacking its signal sequence (TrbG*; Fig. 5). The trbF (exxA) gene product was overproduced previously, using a tac expression vector (Lessl et al., 1991). The size of the overproduced protein coincides well with a band appearing at 27.8 kDa. The 26-

"The molecular mass (m) and the isoelectric point (pl) are theoretical values that have been calculated by the UWGCG computer programs PEPTIDESORT and ISOELECTRIC."
products tentatively assigned to the bands appearing on the autoradiogram shows radioactively labeled extracts of the protein. Molecular mass markers: myosin (200.0 kDa), phosphorylase b (PHOSb, 97.4 kDa), bovine serum albumin (BSA, 69.0 kDa) ovalbumin (OVA, 46.0 kDa), carbonic anhydrase (CAR, 30.0 kDa), trypsin inhibitor (TRY, 21.5 kDa), lysozyme (LYS, 14.3 kDa). Tra2 gene products tentatively assigned to the bands appearing on the autoradiogram are listed next to the indicated molecular masses. The kDa polypeptide might represent the processed version of TrbC, a weaker one was visible, tentatively assigned to a band migrating at 21.8 kDa on the polyacrylamide gel (predicted size: 22 kDa). In extracts of cells harboring pDB7756Δ3 (Fig. 5) two bands are disappearing, showing sizes of 20.2 and 17.5 kDa. An exact assignment of these bands is not possible, since they could correspond either to TrbH (predicted size: 17.0 kDa) or TrbK, even if TrbK exhibits a predicted molecular mass of only 7 kDa.

Bands showing sizes of 12 kDa and 16.5 kDa (Fig. 5) might be related to gene products of trbA and C (predicted sizes: TrbA, 14 kDa; TrbC, 15 kDa). Beneath the band matching TrbC, a weaker one was visible, tentatively assigned to a processed TrbC gene product (TrbC*, Fig. 5). According to this study, for 4 of the 15 orfs shown in Fig. 2, namely trbB, D, N, and O, no corresponding gene product could be assigned. Concerning trbD and O, this might be due to the inefficient SD sequences preceding the start codons (Table I).

In the case of TrbB, the protein was overproduced and purified after replacement of the original SD sequence by that of T7 gene 10. Afterward the orf was identified by immunological cross-reaction and sequencing of the NH2-terminus (Motallebi-Veshareh et al., 1992).2 TrbB serum was also used to detect TrbB in the labeled cell extracts (data not shown). Thus, TrbB was detectable in all extracts except in that of pDB7756, a deletion derivative lacking trbB (Fig. 3). Because TrbB is only detectable by antibody reaction and not by labeling, we conclude that during the short labeling time virtually no synthesis of TrbB took place. Hence, turnover of this proteins seems to be very low, requiring regulation on the translational level. This assumption is supported by the occurrence of a large stem-loop structure preceding the TrbB start codon (Motallebi-Veshareh et al., 1992). A translational repressor, potentially the TrbB protein itself, might bind to this hairpin structure and prevent synthesis of TrbB. Such a translational feedback regulation has also been described for the L11 ribosomal protein operon of E. coli (Baughman and Nomura, 1984). In summary, all orfs located in the Tra2 core region, except that of trbD could be confirmed either by the T7 system or by expression using a P trb vector system.

Induction of host cells harboring plasmids pDB7756 and pDB7756Δ1 resulted in a pattern identical to that shown in Fig. 5. This is surprising, because pDB7756Δ1 lacks the Sall-NorI fragment (Fig. 3), encoding four orfs with predicted sizes of 27, 15, 26, and 26 kDa, respectively. Even subcloning of the PstI-Sphi fragment (30.9–33.2 kb) fragment in pTT7-7 followed by the labeling procedure did not yield bands detectable by autoradiography (data not shown). However, analysis of the nucleotide sequence revealed a rho-independent terminator-like structure between the PstI and the Xhol restriction site (nt 13131–13155, Fig. 2). Only deletion of this region resulted in expression of downstream genes. Thus, a clone with the Xhol-Sphi fragment (31.2–33.2 kb, Fig. 3) inserted in pTT7-7 and subsequent labeling gave rise to visible bands that correspond in size to those of predicted polypeptides (nucleotide sequence data of pTiA6, because gene products of octopine Ti plasmids seem to exhibit a slightly higher positional identity compared to that of Tra2, than those of the nopaline type. Percentages of positional similarities and identities of the corresponding gene products are listed in Table III.

The similarity between TrbB (also called KilBI) and VirB11 was described previously by Motallebi-Veshareh et al. (1992). These authors presented a set of six proteins showing extended sequence similarities to the deduced TrbB amino acid sequence. A characteristic shared by all these proteins is that they are largely hydrophilic and apparently do not possess any potential membrane-spanning domains or signal sequences for protein export. Therefore, we assume that TrbB is located within the cytoplasm or associated to the cytoplasmic membrane. VirB11 of pTiA6 was identified within the cytoplasmic membrane fraction by immunoblot analysis (Christie et al., 1989). Furthermore, it is remarkable that all these proteins carry a highly conserved type I nucleotide binding motif (Walker et al., 1982), usually connected to a NTPase activity. For gene products of virB11 and trbB an ATPase and protein kinase function was shown in vitro (Christie et al., 1989).2 Nucleotide binding motifs were also

---

2 D. Balzer and E. Lanka, unpublished data.

3 D. Balzer, E. Scherzinger, and E. Lanka, unpublished data.
detected in the highly hydrophobic gene products of \textit{trbE} (GXXGXXGTK; Fig. 2) and \textit{virB4} (GXXGXXGKS; Ward \textit{et al.}, 1988 and 1990a), indicating that VirB1 and TrbB, as well as VirB4 and TrbE, might be involved in an activating process during DNA transfer.

VirB2 and TrbC are both strikingly hydrophobic throughout the entire sequence. At the NH\textsubscript{2} terminus of TrbC a potential signal sequence for protein export was identified (Fig. 4), whereas it is lacking in VirB2. The amino termini of TrbK, TrbH, and VirB7 have features characteristic for lipoprotein signal sequences (VirB7: LSG/C; Fig. 4). Nevertheless, no amino acid similarity was found within these proteins except at their NH\textsubscript{2} termini. Shiraiu \textit{et al.} (1990) detected similarities of the VirB7 gene product to various lysis proteins, supporting the former idea that Tra2 gene products are essential for mobilization of the small non-conjugative IncQ plasmid RSF1010 (Lesl\textit{e} et al., 1992). Analyzing the deduced amino acid sequences of Tra2 products, we recognized that nearly all of them are highly hydrophobic and show isoelectric points in the range of 8–11 (Table II). In addition, amino-terminal regions of TrbC, G, H, J, K, L, and M have features characteristic for bacterial export signal sequences (Fig. 4). Considering these findings we suggest a membrane or membrane-associated localization of these proteins, supporting the former idea that Tra2 gene products are mainly involved in facilitating the cell to cell contact during mating-pair formation. In addition, mutations in a section of Tra1 led to resistance to donor specific bacteriophages (Dps+; Barth \textit{et al.}, 1978; Watanet \textit{et al.}, 1980). Recently it was shown that mutations in \textit{traF} (Ziegelin \textit{et al.}, 1991; Wates \textit{et al.}, 1992) resulted in a Dps− phenotype, indicating that at least one Tra1 function seems to be part of the mating pair apparatus.

\begin{table}[h]
\centering
\begin{tabular}{llll}
\hline
\textbf{Table III} & \\
Comparison of the predicted amino acid sequences of the RP4 Tra2 & and the pTiA6 VirB gene products & \\
\hline
& \\
\textbf{Trb} & \textbf{VirB} & \textbf{\% similarity*} & \textbf{\% identity*} & \\
B & 11 & 52.5 & 30.0 & \\
C & 2 & 49.6 & 23.4 & \\
D & 3 & 39.8 & 26.5 & \\
E & 4 & 47.6 & 23.2 & \\
F & 5 & 39.8 & 15.4 & \\
I & 10 & 52.9 & 27.2 & \\
\hline
\end{tabular}
\caption{Comparison of the predicted amino acid sequences of the RP4 Tra2 and the pTiA6 VirB gene products.}
\end{table}

* Determined according to the UWGCG program GAP.

Regarding the relationship between Tra2 and VirB, it is remarkable that similarity is not only restricted to the amino acid level but also extends to the level of gene arrangement (Fig. 6). The consecutive order of orfs encoding for related proteins is identical in both regions, except for \textit{trbB} and \textit{virB11}, which are arranged at the opposite ends of the operons. Remarkably, orfs of both regions are closely packed, sometimes with overlapping stop and start codons. Such overlaps were found between \textit{virB 2/3/4} and 9/10 of \textit{pTiA6} (Ward \textit{et al.}, 1988 and 1990a), as well as between \textit{trbB}/E/F of RP4. Promoter regions are located at the left end of both regions, indicating that both regions are transcribed as a single transcriptional unit.

\section*{DISCUSSION}

Transfer systems of the IncP plasmids have an extremely wide host range, and appear to be capable of mediating DNA transfer into virtually any Gram-negative bacterium. To provide information on the whole set of RP4-encoded transfer genes, we determined the nucleotide sequence of the entire RP4 Tra2 region. Accordingly, the total nucleotide sequence of \textit{Tra} regions of two conjugative plasmids, namely \textit{P} and \textit{RP4} (Ziegelin \textit{et al.}, 1991; Miele \textit{et al.}, 1991; Motallebi-Veshareh \textit{et al.}, 1992; Waters \textit{et al.}, 1992) is known now. Even though \textit{tra} genes of both systems fulfill a similar function, significant similarities were not detectable in either the nucleotide or amino acid level, except a weak similarity between gene products of RP4 \textit{traG} (Ziegelin \textit{et al.}, 1991) and \textit{P traD} (Jalajakumari and Manning, 1989).

Considering sequence data along with RNA-polymerase binding studies (Lesl\textit{e} \textit{et al.}, 1992), we suggest that Tra2 is transcribed as a single transcriptional unit. By deletion studies we defined a 11.23-kb Tra2-core region, which is essential for matings between \textit{E. coli} cells. A codon preference analysis revealed 12 orfs within this region (\textit{traBA-L}), whereby 11 out of this twelve were confirmed by specific expression in \textit{E. coli} using a T7 or a P\textsubscript{promoter system. Genes located downstream of the Tra2 core region up to the \textit{fia}A region (Fong and Stanisch, 1989) might be involved in conjugative DNA transfer in hosts other than \textit{E. coli}.

It was proposed that proteins encoded by Tra2 are mainly responsible for formation of mating pairs in conjugating cells (Guiney and Lanks, 1989). This suggestion was based on the observations that mutations in \textit{tra2} led to resistance of host cells against donor specific bacteriophages (Barth \textit{et al.}, 1978; Palombo \textit{et al.}, 1980; Lesl\textit{e} \textit{et al.}, 1992) and that Tra2 gene products are essential for mobilization of the small non-conjugative IncQ plasmid RSF1010 (Lesl\textit{e} \textit{et al.}, 1992). Considering these findings we suggest a membrane or membrane-associated localization of these proteins, supporting the former idea that Tra2 gene products are mainly involved in facilitating the cell to cell contact during mating-pair formation. In addition, mutations in a section of \textit{Tra} led to resistance to donor specific bacteriophages (Dps+) (Barth \textit{et al.}, 1978; Watanet \textit{et al.}, 1980). Recently it was shown that mutagenesis of \textit{traF} (Ziegelin \textit{et al.}, 1991; Wates \textit{et al.}, 1992) resulted in a Dps− phenotype, indicating that at least one \textit{Tra1} function seems to be part of the mating pair apparatus.

\textsuperscript{K. Ippen-Ihler, L. Frost, and R. Skurray, personal communication.}
RP4-like sequences have been detected in a variety of plasmids that were previously regarded as unrelated (Fig. 7). Sequences similar to the nick region of RP4 oriT exist in the conjugal plasmid R64 (IncW), the mobilizable plasmid pTF-FC2 of *Thiobacillus ferrooxidans* and also in the border sequences of Ti plasmids (Waters et al., 1991; Pansegrau and Lanka, 1991). Related sequences were also found in small mobilizable plasmids of *Staphylococcus aureus*, encoding relaxases, which seem to be similar to the RP4 tral gene product. The RP4 primase (TralC) contains a common sequence motif, which could be identified in the primase of an I-type plasmid (Collib-P9) as well as in the α-protein of phage P4 (Strack et al., 1992).

Comparisons of amino acid sequences deduced from the Tra2 nucleotide sequence led to the finding of resemblances between Tra2 and VirB proteins specified by the IncP and Ti plasmid transfer systems, respectively (Fig. 6, Table III). VirB gene products are proposed to form a pore in the *Agrobacterium* envelope through which the T-DNA would be passed into the plant cell. This suggestion is supported by the highly hydrophobic character of VirB proteins and the finding that three proteins encoded by the 5′-portion of the VirB operon fractionate with the cell envelope (Engström et al., 1987). Additionally, resulting from fusion studies, VirB10 was located in the inner membrane (Ward et al., 1990b). Considering the sequence relationships between Tra2 and VirB gene products, as well as their similar properties (hydrophobicity, basicity, signal sequences), we assume that for both processes a similar transfer machinery is used.

Since gene products TrbB and VirB11, as well as TrbE and VirB4, possess potential NTP binding sites, they might fuel the transfer process. Supporting this idea, we found stretches of similarity in the COOH-terminal part of TrbB and VirB4 to the Fil protein of *Salmonella typhimurium* (Vogler et al., 1991). This protein is proposed to be a part of an ATP-driven protein translocase responsible for export of flagellar proteins via the flagellum-specific export pathway. Furthermore, for TrbB and VirB11 an ATPase and protein kinase activity has been shown in vitro (Christie et al., 1989). Until now, six proteins from different organism were identified to be similar to VirB11 and TrbB (Motallebi-Veshareh et al., 1992). All of them are involved in translocation of either DNA or protein through the bacterial envelope, indicating the existence of a more universal transport mechanism.

Surprisingly, the RP4 Tra2 region is not only similar to VirB, but also to the Pilw region of the IncW plasmid R388. Gene products encoded by this region are also involved in conjugal transfer of R388, most likely in the formation of mating pairs (Bolland et al., 1990). Therefore, RP4 Tra2, Ti VirB and R388 Pilw seem to belong into one class of operons involved in DNA transfer that is defined by the following four criteria: (i) exclusive sequence similarities, (ii) partially identical gene order, (iii) similar properties of encoded proteins, and (iv) proposed equivalent roles of gene products in the transfer process.

Data presented here provide an additional indication for an extended relationship between IncP-mediated bacterial conjugation and T-DNA transfer. As reported previously, comparison of the IncP nick regions and the Ti/Ri border sequences led to the extraction of a 12-nucleotide consensus sequence (Waters et al., 1991; Pansegrau and Lanka, 1991). IncP oriT, as well as the right border-repeat of the Ti-plasmid seem to promote the transfer of DNA in an unidirectional manner (Miranda et al., 1992; Grunter et al., 1988; Al-Doori et al., 1982). Nicking enzymes of both plasmids (RP4-Trai; Ti-VirD2) share similar sequences (Pansegrau and Lanka, 1991) and are functionally related. As an intermediate during DNA transfer, TraI and VirD2 both become covalently attached to the 5′-end of the corresponding transferred DNA strand (Pansegrau et al., 1990; Ward and Barnes, 1988; Young and Nester, 1988; Herrera-Estrella et al., 1988; Dürenberger et al., 1989; Howard et al., 1988) A tyrosine near the NH2 terminus, Tyr-22 (TraI) and Tyr-29 (VirD2), seems to be the most likely candidate for the formation of the covalent protein-DNA complex (Vogel and Das, 1992). VirD2, however, is not the only protein related to an RP4 TraI gene product; VirD4 and RP4 TraG also show similarities (Ziegelin et al., 1991). Interestingly, the traG gene product also exhibits stretches of similarities to the F TraD protein (Jalajakumari and Manning, 1989). Note that all these proteins are hydrophobic and seem to be membrane-located, as was shown recently for the VirD4 protein of the Ti plasmid (Okamoto et al., 1991). Moreover, the relationship between TraI and VirD products also extends to the level of gene arrangement, since the order of traG–traD in the TraI region resembles the VirD operon of the Ti plasmid (Ziegelin et al., 1991). In addition, it was shown recently that VirB and VirD4 proteins, which are normally involved in transfer of DNA from bacteria to plants, can direct conjugal transfer of an IncQ plasmid between bacteria (Beijersbergen et al., 1992).

All data listed strongly support the conjugation model for T-DNA transfer proposing a mechanically analogous process for T-DNA transfer to plant cells and bacterial conjugation (Stachel and Zambryski, 1986). Therefore, a detailed study of conjugal transfer on one hand, and T-DNA transfer on the other hand, might help to shed light on mechanistic details of DNA transfer for both systems.

Some transfer systems apparently contain modules with different evolutionary histories. For example, the IncW plasmid R388 contains elements related to the RP4 tral gene, the F traI gene and the Ti VirB operon. It became evident that one replicon can be associated with evolutionarily distinct tra genes. There is no obligatory coupling between the DNA transfer system and the replicon type. Therefore, conjugation systems might have developed by recombinational rearrangement of components from different systems during evolution.

---

5. F. de la Cruz, M. Lessal, and E. Lanka, unpublished data.


7. F. de la Cruz, personal communication.