Conjugal DNA transfer by the *Escherichia coli* K-12 sex factor F is effected through the expression of 20 or more *tra* genes, most of which are combined in a single operon approximately 30 kilobases long termed the *traYZ* operon (16). Two *tra* genes, *traM* and *traJ*, and the origin of transfer gene *oriT* are located upstream and adjacent to this large operon (2, 49). Sex factor F belongs to the incompatibility group IncFI. Members of the related incompatibility group IncFII, which includes plasmids such as R5, R6, R1, R100, R136, and R6-5, also engage in conjugal DNA transfer. Some of these have been compared to sex factor F by a very thorough series of electron microscope heteroduplex studies (39). In these studies the F factor was first compared to R6-5 and R1, and then R6-5 and R1 were compared to R100-1 (a mutant of R100). The hybridizations showed that the *tra* genes of sex factor F were approximately 85% homologous with the *tra* genes of R6-5. The *tra* genes of R6-5 were 100% homologous with those of R100-1, and the *tra* genes of R1 were approximately 90% homologous with the F factor and 85% homologous with R6-5. Although the limit of detectability of nonhomology with the electron microscope method is 30 to 50 base pairs (bp), these studies formed the basis for an implied consensus among those who work with R100 and R6-5 that the *tra* genes in the two plasmids are identical. A current interpretation of these data is that these two plasmids differ only by the gain or loss of a few insertion sequences or transposons and none of these occurs in the transfer genes themselves. The homologous genes in these two IncII plasmids could be identical, but the genes would have to be sequenced to prove identity.

Complementation tests of the *tra* genes of sex factor F and plasmid R100 have confirmed the functional identity anticipated from the electron microscope studies for most of the genes (51). A few *tra* genes are plasmid specific, and in general, their positions on the plasmids coincide with the regions of nonhomology.

The expression of the *tra* operons of sex factor F and plasmid R100 is under the control of the product of three trans-acting genes, namely *traJ*, *finP*, and *finO*. The *tra* and *finP* genes from both R100 and the F factor are plasmid specific, but the *finO* genes of other IncFII plasmids such as R136 and R6-5 appear identical to the *finP* gene of R100 (9, 46). Sex factor F does not have a functional *finO* gene, but the product of the *finO* gene of R100 is able to function fully with its own genes and with those of sex factor F to control transfer of both plasmids.

The control of expression of the *traYZ* operon in both plasmid R100 and sex factor F requires the *finO* gene product to interact with the plasmid-specific *finP* gene or gene product to negatively control the expression of the *traJ* gene. The *traJ* gene product is a positive control element needed for the expression of the *traYZ* operon. (This model of control is called the FinOP model [10, 48].) Sex factor F is naturally promiscuous because it lacks its own *finO* gene. Strains with point mutations in the *finP* genes of sex factor F and plasmid R6-5 are also promiscuous, even in the presence of a functional *finO* gene.

The *finP* gene of sex factor F has been subcloned and sequenced (19, 45). Mutations in the *finP* gene all map inside the *traJ* gene of factor F thereby establishing that the genes either overlap or that the *finP* gene is transcribed from the opposite strand from that used for transcription of the *traJ* gene (11). An analysis of the sequence of the *finP* gene of factor F has led to the suggestion that the gene encodes a small protein (45) and also to the preliminary suggestion that *finP* may encode an antisense RNA (32).

The nature of the interaction between the *finP* gene or gene product and the *finO* gene product remains unexplained. Experiments have been reported in which a small *BglII*-*SalI* fragment from factor F, which appeared to contain all of the *finP* activity of factor F, was linked to the
galactokinase gene of *E. coli* and the effects of the *finO* gene products on the *finP* promoter were measured (32). None were seen. Examination of the sequence, however, suggests that alternate promoters for the *finP* gene may exist further upstream and these were removed by the digestion with *Bgl*II.

We report here the cloning and sequencing of plasmid R100 DNA that contains the entire R100 *finP* gene, most of the *traM* gene, and the promoter-proximal region of the *traJ* gene. The *traM* gene is another plasmid-specific gene that may encode a protein that binds specifically to the plasmid DNA near the site of single-stranded nicking that has been shown to precede plasmid transfer (8). The sequences of the *traM* genes from sex factor *F* (45) and the IncFI1 plasmid R1 (23) have already been determined and compared.

**MATERIALS AND METHODS**

**Bacterial strains and plasmids.** *E. coli* HB101 (F-*λ*- *hsdS20* recA13 lacY1 proA2 leu arai4 galK2 rpsL20 xylS 1 *metI* supE44) was used as the host for all of the plasmids tested. Strain HB101 was obtained from D. Miller. Strain JC3272 (*his trp lys str gal λ*- [Δdef]) and its *Nal* derivative ED3818 were used as the recipients in quantitative crosses. The plasmids used as cloning vehicles were pBR322 and pBR325 obtained from Bethesda Research Laboratories, Inc., and pH34 obtained from P. Prentki and H. M. Krisch. pH34 is identical to pBR322 except that it has a 10-bp insert in the *EcoRI* site that contains a *SmaI* site. This allows blunt-end cloning into the *SmaI* site and removal of the cloned fragment with *EcoRI* (35). Plasmids isolated in this work are shown in Table 1. R136 *finP* - was obtained from N. Grindley (13). It was originally called 240 *drp2* (14).

**Media and culture conditions.** All bacterial cultures, except those used in the procedure for large-scale preparation of plasmid DNA, were grown in L broth or on ML plates (25). For large-scale plasmid preparation, M9 medium supplemented with 0.5% Casamino Acids (Difco Laboratories) and 0.4% glucose was used (44). M9 was also used for minimal medium agar plates. Supplemental amino acids required were added to a concentration of 0.002%. Antibiotics were used at the following concentrations: tetracycline (10 μg/ml), ampicillin (100 μg/ml), streptomycin (200 μg/ml), and nalidixic acid (40 μg/ml). The temperature for all incubations was 37°C.

**Testing for FinP.** Testing for FinP was performed by a standard donor ability test from stable heterozygotes (3). The test measured the amount of transfer of a *finP* mutant of the FII plasmid R136 from a cell containing both the cloned test fragment and the R136 mutant by measuring the number of tetracycline-resistant recipients. Counterselection of donors was done with nalidixic acid, and the nalidixic-acid-resistant strain ED3818 was used as the recipient. Preliminary experiments established that the optimal mating time at 37°C for these crosses was 90 min rather than the more normal 30 to 40 min used in the original work.

**Testing for the oriT gene.** HB101 isolates were tested for plasmids carrying the R100 *oriT* gene by first forming a stable R100-1 heterozygote (selecting for spectinomycin resistance) and then screening for *amp* transfer to strain JC3272 from the heterozygotes, selecting against the donors by their amino acid requirements. The mating time was 30 min at 37°C.

**Cloning.** The cloning vehicles for this work were *EcoRV*-digested pBR322, *EcoRI*-digested pBR322, *EcoRI*-digested pBR325, *EcoRV*- and *EcoRV*-digested pBR322, and *SmaI*- digested pH34. Before use, all of the plasmids were dephosphorylated with calf intestinal phosphatase, as described by Maniatis et al. (26). Clones containing blunt-end fragments in the pH34 *SmaI* site were detected by screening Tc' colonies for plasmids that contained *EcoRI* releas-

| TABLE 1. Description of plasmids constructed |
|---------------------------------------------|
| R100 genes     | Plasmid | Cloning vehicle | Restriction site used for insertion or deletion | Fragment inserted or deleted (Δ) | Size of cloned fragment (bp) |
|----------------|---------|----------------|-----------------------------------------------|---------------------------------|-----------------------------|
| tra            | pWD1    | pBR322         | *EcoRI*                                       | *EcoRI-B of R100*               | 12,660                      |
|                | pWD8    | pBR322         | *EcoRI*                                       | *EcoRI-C of R100*               | 11,600                      |
|                | pWD10   | pBR322         | *EcoRI*                                       | *EcoRI-F of R100*               | 6,370                       |
|                | pWD22   | pBR325         | *EcoRI*                                       | *EcoRI-D of R100*               | 11,300                      |
|                | pWD23   | pBR325         | *EcoRI*                                       | *EcoRI-E of R100*               | 7,590                       |
| finP and oriT  | pBF1    | pBR322         | *EcoRV*                                       | *EcoRV-E of *EcoRI*-D of R100* | 1,240                       |
|                | pBF3*   | pBR322         | *EcoRV*                                       | *EcoRV-E of *EcoRI*-D of R100* | 1,240                       |
|                | pBF3    | pBR322         | *EcoRV*                                       | *EcoRV-E of *EcoRI*-D of R100* | 1,240                       |
|                | pBF3    | pBR322         | *EcoRV*                                       | *EcoRV-E of *EcoRI*-D of R100* | 1,240                       |
|                | pBF4    | pBF3           | *EcoRV*                                       | *EcoRV-E of *EcoRI*-D of R100* | 1,240                       |
|                | pBF4    | pBF3           | *EcoRV*                                       | *EcoRV-E of *EcoRI*-D of R100* | 1,240                       |
|                | pWD28   | pBR322         | *EcoRV*                                       | *EcoRV-E of *EcoRI*-D of R100* | 1,240                       |
|                | pWD34   | pH34           | *SmaI*                                        | *HaeIII-A of pBF1*             | 751                         |
|                | pWD35*  | pH34           | *SmaI*                                        | *HaeIII-A of pBF1*             | 751                         |
|                | pWD36   | pH34           | *SmaI*                                        | *HaeIII-A of pBF1*             | 751                         |
|                | pWD37*  | pH34           | *SmaI*                                        | *HaeIII-A of pBF1*             | 751                         |

* Identical to the plasmid immediately above, but the cloned fragment is in opposite orientation.

**TABLE 2. *EcoRV* fragment sizes of R100 *tra* plasmids**

| *EcoRV* fragment | pWD22 | pWD8 | pW23 | pWD10 | pWD1 |
|------------------|-------|------|------|-------|------|
| A                | 6.3   | 6.0  | 5.9  | 6.3   | 8.1  |
| B                | 3.15  | 3.6  | 4.6  | 1.95  | 3.7  |
| C                | 2.9   | 2.4  | 2.5  | 1.3   | 1.8  |
| D                | 1.75  | 1.5  | 0.58 | 0.85  | 1.6  |
| E                | 1.24  | 0.45 | 0.33 | 1.3   | 1.3  |
| F                | 0.76  | 0.40 |     | 0.52  |      |
| G                | 0.52  | 0.36 |     |       |      |
| H                | 0.33  |      |     |       |      |
| I                | 0.30  |      |     |       |      |
| J                | 0.23  |      |     |       |      |

* The net values (kilobase pairs) for fragment sizes, determined by subtracting the size of pBR322 or pBR325, as appropriate (Table 1), from the sum of the fragment sizes were as follows: pWD22, 10.63; pWD8, 11.21; pWD23, 7.59; pWD10, 6.37; pWD1, 12.66.
able fragments of the predicted sizes. The identity of the fragments was established by digestion of the plasmids with other restriction enzymes.

Transformation. The CaCl₂ transformation procedure described by Maniatis et al. (26) was used. The transformation mixtures were allowed to grow for 1 h in L broth with 0.2% glucose before being plated on ML agar plates containing either ampicillin (100 μg/ml) or tetracycline (10 μg/ml).

DNA preparation. Small preparations of plasmids were made from 1 to 2 ml of shaken overnight L broth cultures by the Holmes and Quigley procedure (18). For restriction analysis of these preparations, a sample equivalent to 0.1 ml of culture provided enough DNA for one track on an agarose gel. Large amounts of pWD35 DNA were initially prepared by the method of Mukhopadhyay and Mandal (31). Later, this and all other plasmids were prepared as described by Thompson et al. (44) by chloramphenicol amplification and reversed-phase column chromatography on NACS37.

DNA sequencing. For DNA sequencing, the rapid method described by Bencini et al. (4) was used. This method is a modified version of the original Maxam and Gilbert method (27). Labeling of fragments at the 3' end was done with the Klenow fragment of DNA polymerase by the method of Drouin (7); the 5'-end-labeling method was adapted from the method of Maniatis et al. (26). For sequence analysis, the Beckman Microgenie Sequence Analysis Program was used. This computer program is described by Queen and Korn (36).

Enzymes. Calf intestinal phosphatase was obtained from P-L Biochemicals, Inc., and the Klenow fragment of polymerase I, polymerase I, and DNase I were all obtained from Bethesda Research Laboratories, Inc. All restriction endonuclease were from New England BioLabs, Inc. Digestions were all performed in the buffers recommended by the supplier.

Restriction mapping. DNA fragments for mapping and sequencing were obtained from restriction endonuclease digestion mixtures by electroelution from agarose gels onto DEAE paper, as described by Dretzen et al. (6). Elution of the DNA from the paper was done in three consecutive 30-min incubations in 150 μl of the salt solution described, for all sizes of paper up to 10 cm². Before and after each elution step the solvent was removed by a 10-min centrifugation of the paper in a 500-μl Eppendorf tube with pin holes in the bottom and top. The eluates were collected in 1.5-ml Eppendorf tubes. Yields were 80% or better. Restriction mapping was performed on isolated DNA fragments in single and multiple digests as necessary and by probing Southern gel blots with labeled pBR322 or a labeled EcoRV fragment. All agarose gels were horizontal in Tris borate with ethidium bromide.

RESULTS

The transfer gene region of plasmid R100 is carried on five adjacent EcoRI fragments. In earlier studies this region was largely undigested by the then commonly available restriction enzymes (30, 43). This led to size estimates of the transfer region that were based on a summation of the sizes of the EcoRI fragments, all of which fell outside the range of accurately measurable fragments. To get a more accurate size estimate for these EcoRI fragments and to aid in

TABLE 3. Donor ability of R136 fin⁺ from strain HB101 containing cloned R100 fragments

| Coresident plasmid | Te⁺ transconjugants/100 donors |
|--------------------|-------------------------------|
| pBF1               | 7.6 × 10⁻⁴ (5)               |
| pBF2               | 2.3 × 10⁻⁴ (2)               |
| pBF3               | 3.0 (3)                      |
| pBF4               | 0.17 (3)                     |
| pBR322             | 3.6 (9)                      |
| pWD28              | 2.7 (2)                      |
| pWD34              | 5.2 × 10⁻⁴ (2)               |
| pWD35              | 1 × 10⁻⁴ (2)                 |
| pWD36              | 2.9 (2)                      |
| pWD37              | 3.0 (2)                      |

* Number in parentheses is number of independent assays.

TABLE 4. Mobilization by R100-1 of pBR322 carrying R100 DNA

| Resident plasmid | Ap⁺ transconjugants/100 donors |
|------------------|-------------------------------|
| pWD22            | 4 × 10⁻³ (2)                  |
| pBF1             | 6 × 10⁻³ (5)                  |
| pWD34            | <1 × 10⁻⁴ (5)                 |
| pWD35            | <1 × 10⁻⁴ (2)                 |
| pWD36            | 4 × 10⁻³ (5)                  |
| pWD37            | 1 × 10⁻² (6)                  |
| None             | <1 × 10⁻⁴ (4)                 |

* Number in parentheses is number of separate measurements.
subcloning the transfer genes, we cloned the five fragments (D, C, E, F, and B) and screened a series of enzymes for those that would digest the cloned R100 DNA into a modest number of pieces. Of the enzymes tested, EcoRV, BanI, and AhaIII gave the best size estimates for each of the five fragments. EcoRV was then chosen because it cut pBR322 only once. The fragment sizes generated from the five plasmids and the net size of each of the EcoRI fragments cloned are shown in Table 2. By using these sizes and a value of 87.2 for the R100 coordinate of the EcoRI B-H fragment junction, as determined by sequencing (37), and the assigned value of 89.3 for the IS/1b resistance determinant reference point (29), the EcoRI D-C fragment junction was located at 49.4 kilobases.

The EcoRI D fragment of plasmid R100 (R100-1) contains the oriT, traM, finP, and traJ genes and the first part of the traYZ operon (46) in addition to IS10R and a portion of the tet gene of Tn10. The DNA sequences for IS10R (15) and the tet gene (17, 33) are known and were used to identify restriction sites in the right half of this fragment. A partial restriction map of Tn10 (20) was similarly used. A restriction map of the rest of the EcoRI D fragment was prepared to allow cloning and mapping of the finP, oriT, and traJ genes and the beginning of the traYZ operon. The map is shown in Fig. 1. All of the EcoRV fragments were then isolated.

EcoRV fragments A, E, and F were cloned into the EcoRV site of plasmid pBR322 and fragments B and C, each bearing an EcoRI end and an EcoRV end, were cloned into pBR322, which had been digested with both enzymes. Only the plasmids in which a perfect EcoRV restriction site was reformed were saved. EcoRV subclones of the R100 EcoRI D fragments were then tested for the oriT and finP genes. The FinP test was based upon the assumption that the cloned finP gene could act in trans to reduce the transfer of a coresident finO+ finP- IncFII plasmid. The data (Table 3) showed that the EcoRV E fragment contained the finP gene and that orientation did not matter. The oriT test was based upon the assumption that a coresident R100-1 could mobilize any pBR322 vehicle that carried an intact R100 oriT region. Similar tests were used previously for R100 oriT clones (50), R1 oriT (34), and F plasmid oriT (19). The data (Table 4) showed that the EcoRV E fragment contained all the oriT activity of the much larger R100 EcoRI D fragment from which it was derived.

The plasmids containing the EcoRV E fragment in both orientations in the EcoRV site of pBR322 were called pBF1 and pBF2. Derivatives of both of these plasmids were made by digesting them with BamHI and religating. The products, pBF3 and pBF4 (Fig. 2), contained the leftmost 152 bases of the pBF1 insert and the rightmost 1,090 bases, respectively. pBF3 lost all of the original finP activity, indicating that the finP gene included a site sensitive to the BamHI endonuclease.

HaeIII digested the cloned fragment into two nearly equal
portions (Fig. 2). To further subclone the finP and oriT genes, a HaeIII digest of pBF1 was made and the top two HaeIII bands were cloned in both directions into Smal-digested pHPI34 to form the plasmids pWD34, pWD35, pWD36, and pWD37 (Fig. 2).

Data for the FinP and oriT tests on these smaller plasmids are shown in Tables 3 and 4. The data showed that pWD34 and pWD35 carried the R100 finP gene, whereas pWD36 and pWD37 contained all of the oriT activity associated with the larger fragments from which they were derived. The marked decrease in absolute transfer levels (Table 4) compared with what was seen in cloned R100 oriT carried in other strains (50) appears to derive entirely from host and recipient strain differences. The donor strains used were modificationless, whereas the recipients carried wild-type restriction. Other donor strain differences are under investigation.

The R100 DNA in the finP plasmid pWD34 (and pWD35) mapped between coordinates 46.51 and 47.24. In the remainder of this work this fragment is called the P fragment. The R100 DNA in the oriT plasmid pWD36 mapped between coordinates 45.94 and 46.51. The reference point for these determinations was the EcoRI restriction site in R100 between the C and D fragments mentioned above.

The scheme for sequencing the P fragment of pWD35 by the Maxam-Gilbert technique is shown in Fig. 3. The sequence itself is shown in Fig. 4. The orientation is such that the oriT gene is to the left of the P fragment and the traYZ operon is to the right. The 22-base sequence before the first HaeIII site (GGCC) at base 23 was taken from the sequence of the R100 DNA in the oriT clone pWD36 (S. A. McIntire, B. E. Fee, and W. B. Dempsey, manuscript in preparation).

The potential open reading frames (orfs) in this sequence are listed in Table 5. All those listed were preceded by at least a 3-base homology with the Shine-Dalgarno sequence AAGAGG within 10 bases of the translation start site. This analysis showed that the top strand had one large orf (orf2) that opened at base 21. orf2 contained three additional start codons, which would allow three subsets of orf2 to be translated in frame. The top strand had another reading frame (orf1) which opened at base 598 but did not close in this sequence. A similar search of the bottom strand showed two orfs, orf3 and orf4. Each of the orfs shown in Table 5 used a high percentage of rare codons. These values are shown in Table 5, as well as the expected sizes of the protein products and their start and stop points. The rare codons searched for were those identified by Konigsberg and God-
A similar high occurrence of rare codons is found in this same region of sex factor F (11, 45).

When we compared the sequence in Fig. 4 with the F factor sequence published by Thompson and Taylor (45) and the R1 sequence published by Koronakis et al. (23), we concluded that orf2 encodes the R100 traM protein and orf1 encodes the beginning of the R100 traJ protein. The amino acid sequence of orf2 and the differences between it and the sequence of the traM product from factor F are shown in Fig. 5. Only 14 of the 127 amino acid residues of orf2 were different in the two proteins. Seven of these were conservative differences. In contrast, a comparison between the sequence of orf2 and the recently published traM sequence of plasmid R1 (23) showed 28 mismatches, only half of which were conservative (data not shown).

Although the homology between orf2, the F factor traM
gene, and the R1 traM gene is remarkable, we cannot unconditionally assign the identity of orf2 to R100 traM. The assignment of the F factor traM gene to the F sequence is based in part upon complementation of F factor traM mutants by cloned F fragments carrying this DNA and in part upon detection of a unique 13-kilodalton protein in extracts of cells containing the cloned F traM gene (21). No protein of this size was detected when the R100 traM analog, R6-5, was so examined (1). We have not yet been successful in subcloning an R100 traM fragment that makes an unequivocally identifiable product nor have traM mutants of R100 been isolated.

A computer analysis of the traM proteins from factor F, plasmid R1, and orf2 showed they had nearly identical hydrophobicity patterns (Hopp and Woods method). Similar α-helical contents were also seen for all three proteins (method of Garner et al.). Each traM protein showed a run of at least 35 α-helical residues immediately before residue 73 and a run of at least 22 α-helical residues after residue 99. Each traM sequence also showed a 10-residue-long α-helical region in the 25 residues between residues 74 and 98, and this α-helical region was set apart from other patterns. These identities strongly suggest that orf2 is the R100 traM gene.

Because the data in Table 3 show that digestion with BamHI and relaxation interrupt finP function, neither orf2 nor any of the potential orfs within orf2 can be the finP gene because all of these orfs begin and end left of the BamHI site. No such conclusion can be made about the two small orfs on the lower strand. A transcript carrying these could easily begin on the distal side of the BamHI site.

The DNA sequence was screened for initiation signals. One set was found on the top strand. A Pribnow box with the sequence TATAT was found, beginning at base 485. The −35 sequence following to this was ATGAC, beginning at base 462. This is presumed to be the set of signals to initiate the traJ message at either base 493 or base 495.

Upstream of the BamHI restriction site, three sets of sequences were found that could serve as initiation signals on the bottom strand. Set 1 has a Pribnow box, TATGAT, beginning at base 703, and a −35 sequence TTGTAG; set 2 has a Pribnow box, TATATT, beginning at base 644 and a −35 sequence TCGACA; and set 3 has a Pribnow box, TAGGAT, beginning at base 616 and a −35 sequence TTGACG. Although all of these sets satisfy the four rules of McClure (28) for procaryotic promoters, no runoff transcripts were detected when the P fragment was incubated with RNA polymerase and the four ribonucleotide triphosphates (5). Other possible promoter combinations exist, but they were not included because they contained too many down mutations of the type described by von Hippel et al. (47).

Computer searches for inverted repeats that might serve as transcription termination signals showed several in the region between bases 400 and 650. Two of these had stems whose Gibbs' free energy values (G) were below −15 kcal (ca. −63 kJ/mol). These were at bases 468 to 505 (G = −17 kcal [ca. −71 kJ/mol]) and at bases 535 to 564 (G = −25 kcal [ca. −105 kJ/mol]). The first of these may serve as a rho-dependent terminator for the traM transcript in the top strand. A similar inverted repeat in this region is not seen in sex factor F. The second of these inverted repeats could also serve as a termination signal for traM in the top strand, and in the bottom strand it has the stem-loop strength and run of thymidine residues downstream that are characteristic of rho-independent terminators. Transcripts starting at the three sets of putative start sites on the bottom strand and ending at this terminator would have lengths of 163, 104, and 74 bases. In contrast to the first inverted repeat, this one is conserved in sex factor F (45). The positions of these two inverted repeats are shown in Fig. 4. Two other stem-loops were found that overlap the one at bases 535 to 564.

We reported previously (5) that incubation of the P fragment with RNA polymerase and ribonucleotide triphosphates did not result in the synthesis of a detectable product, whereas control experiments with another piece of R100 DNA gave readily detectable RNA transcripts. The P fragment also was ligated to the strong TAC promoter in both orientations, but we were unable to detect any proteins larger than 3,000 daltons being made during the induction of this promoter (data not shown). Smaller peptides were not seen either but have not been unequivocally eliminated as possibilities.

**DISCUSSION**

This work reports the sequence of the plasmid R100 DNA that encodes the finP gene and parts of the adjacent genes traM and traJ. The cloned DNA is known to carry all of the R100 finP gene, because the cloned DNA was in itself a sufficient source of all the trans-acting finP protein needed to inhibit transfer in a finP finO′ test system. The presence of parts of the traM and traJ genes was established only by the very strong homology their sequences show to the same transfer genes in the related plasmids F and R1. The finP gene crosses the BamHI restriction endonuclease site in the fragment. This site is far to one side of the DNA, just inside orf1, the orf in the top strand that corresponds to the traJ reading frame in plasmid F. This location establishes limits for the finP gene. If it is in the top strand, it must overlap traJ, the gene whose expression it controls. If it is in the bottom strand, then finP is antisense to traJ. It is well established that the finP gene controls transcription of the traJ gene and not translation (12, 48). Transcriptional regulation by a trans-acting overlapping gene would almost certainly be mediated through a protein. Transcriptional regulation by antisense transcripts, on the other hand, is now well established and does not necessarily require a protein.

In the top strand of the P fragment, the BamHI restriction site is downstream of orf2 and the shorter orf it contains and inside the orf1 or traJ reading frame. This rules out all but the traJ orf as a potential orf for a finP protein if the finP gene is encoded on the top strand. If the traJ and finP genes share the same orf, then one would be a truncated version of the other. This seems very unlikely to us in view of the data that show that finP controls transcription of traJ (12, 48). We feel instead that the finP gene must be encoded on the bottom strand. On the bottom strand, the BamHI site is

| Site(s) | Start (base) | End (base) | Protein size (daltons) | % of: 8 rarest codons | 23 infrequent codons |
|---------|-------------|-----------|-----------------------|----------------------|---------------------|
| orf1    | 598         | 616       | 14,510                | 22.7                 | 34.1                |
| orf2    | 21          | 404       | 1,296                 | 8.7                  | 37.6                |
| orf3    | 562         | 524       | 7.7                   | 23.7                 |                     |
| orf4    | 573         | 493       | 2,744                 | 11.1                 | 37                  |
| Non-reading frames* |        |           |                       | 11.6                 | 33.6                |
| Regulatory proteins* |     |           |                       | 9.3                  | 24                  |
| Nonregulatory proteins* |      |           |                       | 4.2                  | 12                  |

* Values are from Konigsberg and Godson (22).
upstream of the two orfs and the potentially very strong rho-independent terminator and downstream of all three potential transcription initiation sites. That leaves either or both of the bottom strand orfs as potential sources of any \textit{finP} protein and the transcript as an antisense RNA. There are preliminary reports that the R100 \textit{finP} gene is in the bottom strand and of evidence for the existence in vivo of antisense transcripts from it (B. E. Fee, S. A. McIntire, and W. B. Dempsey, Abstr. 13th Int. Congr. Biochem., abstr. no. MO-062, 1985; W. B. Dempsey, Fed. Proc., 45:1879, abstr. no. 2324, 1986).

A similar location for the \textit{finP} gene of plasmid F has been put forth by Fowler et al. (11). They found that a number of \textit{finP} point mutations in the beginning of the \textit{traJ} orf of that plasmid, and they suggested that a low-molecular-weight protein was part of the \textit{finP} gene products. In a preliminary publication, Mullineaux and Willets have confirmed that location for the plasmid F \textit{finP} gene and have argued that the product of the gene is an antisense RNA (32).

We did not establish in the present work that the product of the \textit{finP} gene is an antisense RNA, and accordingly we will not discuss the details of possible antisense mechanisms. Several control systems that use antisense RNA have been described (24, 38, 40, 41), and they all have a common property that does not seem to be shared with the FinOP control system: they show some regulation whenever the antisense component is present alone.

There is a very important difference between the FinOP system and these antisense systems. In the FinOP system transfer is not inhibited in all the presence of either the \textit{finO} or \textit{finP} gene alone, whereas the regulatory effects can be additive in other systems. One explanation may be that the measured entity in this comparison, namely transfer of the whole plasmid, is misleading. If expression of the transfer operon requires only a few molecules of the \textit{traJ} gene product, then a 10-fold difference in \textit{traJ} gene transcription may be all that is necessary to go from no expression to full expression on the \textit{traYZ} operon. The measurement of transfer would then be a kind of amplification of the \textit{traJ} gene expression. Another explanation may be that the \textit{finO} and the \textit{finP} products do interact with each other first to form an active inhibitor. Some resolution of this may be possible by quantitating the amounts of translatable \textit{traJ} message or measuring \textit{traJ} promoter strengths under various conditions. Using this last technique for sex factor, Mullineaux and Willets found that the \textit{traJ} promoter does vary only about 10-fold between fully inhibited and uninhibited states (32).

We think that the region from base 462, the presumed -35 RNA polymerase binding site for synthesis of the \textit{traJ} transcript, to base 760 at the right end of the sequence contains all of the R100 \textit{finP} gene and its most likely site of action and possibly contains the site of action of the \textit{finO} gene product. We searched this region for unusual sequences. There was no homology with the sequence of the R100 \textit{finO} gene (S. A. McIntire and W. B. Dempsey, unpublished data). We did find overlapping stem-loops and orfs in the bottom strand that looked like a translationally controlled transcription termination system. This attenuator-like system is described here. A stem-loop that has all of the required characteristics of a rho-independent transcription termination signal can be formed on the bottom strand by bases 564 to 535 (stem of 12 bp, eight G·C pairs, G = −25 kcal [−105 kJ/mol], loop of six bases, and a run of five Ts downstream of the stem). A second stem-loop can be formed from bases 597 to 547 (stem of 15 bp, seven G·C pairs, G = −9 kcal [ca. −38 kJ/mol], loop of 21 bases) and a third stem-loop can be formed from bases 600 to 571 (stem of 11 bp, four G·C pairs, G = −6 kcal [ca. −25 kJ/mol], loop of eight bases). Clearly, computer analysis of sequences generally shows a number of potential stem-loops. What we think is unusual about this set is first that their stems overlap exactly as has been described for attenuator sequences (42) and second that two orfs, orf3 and orf4, totally overlap the strong terminator (Table 5). If either orf is translated, the transcript would not stop at the rho-independent terminator but would instead go to the −17-kcal stem-loop at bases 505 to 468 further downstream. This would mean that two transcripts being made off of the bottom strand should be found. These in fact have been found by using SP6-directed probes of the area (W. B. Dempsey, manuscript in preparation).

We were unable to detect proteins larger than 3,000 daltons, even when the F fragment was ligated to the strong TAC promoter (smaller proteins were not detected either,

**FIG. 5.** Derived amino acid sequences of the putative R100 \textit{traM} protein. The complete bottom sequence represents the derived R100 polypeptide. The upper sequence represents the F plasmid amino acids where plasmids R100 and F differ; a blank space represents homology.
but these have not been unequivocally excluded. Our current explanation for not seeing some protein is that all of the orfs have very rare codons. Another reason is that in one orientation, the P fragment begins one base inside orf2, so that orf2 is not translatable. In this case, the transcript from TAC probably ends either at the traM terminator (bases 468 to 505) or at the strong stem-loop that occurs in the leader of the traM message (bases 535 to 564), so the J transcript is not made. We presume based on the presence of a strong terminator sequence in the leader region of the traJ mRNA that the control of traJ expression may involve premature termination of this RNA (see also reference 32). In the opposite orientation, the P fragment, under the control of the TAC promoter, has the same problem. orf3 and orf4 both contain very rare codons, and what probably happens is that this results in ribosome pausing at a number of codons. This would eventually yield small peptides of various sizes. The pausing would also uncouple transcription and allow termination of the transcript at the strong rho-independent terminator (bases 486 to 520). Experiments are in progress to resolve these issues.

ACKNOWLEDGMENTS

This work was supported by the Research Service of the Veterans Administration, and by Public Health Service grant AI 17402 from the National Institutes of Health.

The technical help of Kathryn Sims is gratefully acknowledged. S. A. McIntire provided continual interest and useful discussions.

LITERATURE CITED

1. Achtman, M., K. Kusecek, and K. N. Timmis. 1978. tra cistrons and proteins encoded by the Escherichia coli antibiotic resistance plasmid R6-5. Mol. Gen. Genet. 163:169–179.

2. Achtman, M., R. A. Skurray, R. Thompson, R. Helmhut, S. Hall, L. Beutin, and A. J. Clarke. 1978. Assignment of tra cistrons to EcoRI fragments of F sex factor DNA. J. Bacteriol. 133:1383–1392.

3. Achtman, M., N. Willett, and A. J. Clarke. 1971. Beginning a genetic analysis of conjugal transfer determined by the F factor in Escherichia coli by isolation and characterization of transfer-deficient mutants. J. Bacteriol. 106:529–538.

4. Beninc, D. A., G. A. O’Donovan, and J. R. Wild. 1984. Rapid chemical degradation sequencing. Biotechniques 2:4–5.

5. Dempsey, W. B. 1985. Molecular details of the genes finO and finP which control the infectious transfer of the antibiotic resistance plasmid R100, p. 939. In D. R. Helenski, S. N. Cohen, D. B. Clewell, D. A. Jackson, and A. Hollaender (ed.), Plasmids in bacteria. Plenum Publishing Corp., New York.

6. Dretten, G., M. Bellard, P. Sassone-Corsi, and P. Chambron. 1981. A reliable method for the recovery of DNA fragments from agarose and acrylamide gels. Anal. Biochem. 112:295–298.

7. Drouin, J. 1980. Cloning of human mitochondrial DNA in Escherichia coli. J. Mol. Biol. 140:15–34.

8. Everett, R., and N. S. Willett. 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.

9. Finnegan, D., and N. Willett. 1972. The nature of the transfer inhibitor of several F-like plasmids. Mol. Gen. Genet. 119:57–66.

10. Finnegan, D., and N. Willett. 1973. The site of action of the F transfer inhibitor. Mol. Gen. Genet. 127:307–316.

11. Fowler, T., L. Taylor, and R. Thompson. 1983. The control region of the F plasmid transfer operon: DNA sequence of the traJ and traI genes and characterization of the traI-F-Z promoter. Gene 26:79–89.

12. Gaffney, D., R. Skurray, and N. Willett. 1983. Regulation of the F conjuction genes studied by hybridization and tra-lacZ fusion. J. Mol. Biol. 168:103–122.

13. Grindley, N. D. F., E. S. Anderson, H. R. Smith, and J. N. Grindley. 1971. The effects of Salmonella typhimurium on derepressed mutants of &-like factors. Genet. Res. 17:89–93.

14. Grindley, N. D. F., J. N. Grindley, H. R. Smith, and E. S. Anderson. 1973. Characterization of derepressed mutants of an &-like factor. Mol. Gen. Genet. 120:27–34.

15. Halling, S. M., R. W. Simons, J. C. Way, R. B. Walsh, and N. Kleckner. 1982. DNA sequence organization of IS10-right of Tn10 and comparison with IS10-left. Proc. Natl. Acad. Sci. USA 79:2606–2612.

16. Helmuth, R., and M. Achtmann. 1975. Operon structure of DNA transfer cistrons of the F sex factor. Nature (London) 257:652–656.

17. Hillen, W., and K. Scholmeier. 1983. Nucleotide sequence of the Tn10 encoded tetracycline resistance gene. Nucleic Acids Res. 11:25–539.

18. Holmes, D. S., and M. Quigley. 1981. A rapid boiling method for the preparation of bacterial plasmids. Anal. Biochem. 114:193–197.

19. Johnson, D., R. Everett, and N. Willett. 1981. Cloning of F DNA fragments carrying the origin of transfer oriT and the transfer functions. Proc. Natl. Acad. Sci. USA 78:580–585.

20. Jongensen, R. A., and W. S. Reznikoff. 1979. Organization of structural and regulatory genes that mediate tetracycline resistance in transposon Tn10. J. Bacteriol. 138:705–714.

21. Kennedy, N., L. Beutin, M. Achtmann, R. Skurray, U. Rahmsdorf, and P. Herrlich. 1977. Conjugation proteins encoded by the F sex factor. Nature (London) 270:580–583.

22. Konigsberg, W., and G. N. Godson. 1983. Evidence for the use of rare codons in the dnaG gene and other regulatory genes of Escherichia coli. Proc. Natl. Acad. Sci. USA 80:687–691.

23. Koronakis, V., E. E. Bauer, and G. Hogenauer. 1983. The traM gene of the resistance plasmid R1: comparison with the corresponding sequence of the Escherichia coli F factor. Gene 36:79–86.

24. Lacetena, R. M., and G. Cesareni. 1981. Base pairing of RNA I with its complementary sequence in the primer precursor inhibits CoEI replication. Nature (London) 294:623–626.

25. Lennox, E. S. 1955. Transduction of linked genetic characters of the host by bacteriophage P1. Virology 1:190–206.

26. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

27. Maxam, A. M., and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. 65:599–650.

28. McClure, W. R. 1985. Mechanism and control of transcription initiation in prokaryotes. Annu. Rev. Biochem. 54:171–204.

29. Mickel, S., E. Ohtsubo, and W. Bauer. 1977. Heteroduplex mapping of small plasmids derived from R-factor R12: in vivo recombination occurs at IS1 insertion sequences. Gene 2:193–210.

30. Miki, T., A. M. Easton, and R. H. Rownd. 1978. Mapping of the resistance genes of the R plasmid NR1. Mol. Gen. Genet. 158:217–224.

31. Muckpadya, M., and N. C. Mandal. 1983. A simple procedure for large-scale preparation of pure plasmid DNA free from chromosomal DNA from bacteria. Anal. Biochem. 133:265–270.

32. Mullineaux, P., and N. S. Willett. 1983. Molecular analysis of the transfer region of plasmid P, p. 605–614. In D. R. Helenski, S. N. Cohen, D. B. Clewell, D. A. Jackson, and A. Hollaender (ed.), Plasmids in bacteria. Plenum Publishing Corp., New York.

33. Nguyen, T. T., K. Postle, and K. P. Bertrand. 1983. Sequence homology between the tetracycline-resistance determinants of Tn10 and pBR322. Gene 25:553–562.

34. Ostermann, E., F. Krcek, and G. Hogenauer. 1984. Cloning the origin of transfer region of resistance plasmid R1. EMBO J. 3:1731–1735.

35. Prentki, P., and H. M. Kriseh. 1982. A modified pBR322 vector with improved properties for the cloning of DNA, and sequencing of blunt-ended DNA fragments. Gene 17:189–196.

36. Queen, C., and L. J. Korn. 1984. A comprehensive sequence analysis program for the IBM personal computer. Nucleic Acids
Res. 12:581–599.

37. Rosen, J., T. Ryder, H. Inokuchi, H. Ohtsubo, and E. Ohtsubo. 1980. Genes and sites involved in replication and incompatibility of an R100 plasmid derivative based on nucleotide sequence analysis. Mol. Gen. Genet. 179:527–537.

38. Rownd, R. H., D. D. Womble, X. Dong, V. A. Luckow, and R. P. Wu. 1985. Incompatibility and IncFII plasmid replication control, p. 335–354. In D. R. Helinski, S. N. Cohen, D. B. Clewell, D. A. Jackson, and A. Hollaender (ed.), Plasmids in bacteria. Plenum Publishing Corp., New York.

39. Sharp, P. A., S. N. Cohen, and N. Davidson. 1973. Electron microscope heteroduplex studies of sequence relations among plasmids of Escherichia coli. II. Structure of drug resistance (R) factors and F factors. J. Mol. Biol. 75:235–255.

40. Simons, R. W., and N. Kleckner. 1983. Translational control of 1S10 transposition. Cell 34:683–691.

41. Stougaard, P., S. Molin, and K. Nordstrom. 1981. RNAs involved in copy-number control and incompatibility of plasmid R1. Proc. Natl. Acad. Sci. USA 78:6008–6012.

42. Stroynowski, I., and C. Yanofsky. 1982. Transcript secondary structures regulate transcription termination at the attenuation of S. marcescens tryptophan operon. Nature (London) 298:34–38.

43. Tanaka, N., J. H. Cramer, and R. H. Rownd. 1976. EcoRI restriction endonuclease map of the composite R plasmid NR1. J. Bacteriol. 127:619–636.

44. Thompson, J. A., R. W. Blakesley, K. Doran, C. J. Hough, and R. D. Wells. 1983. Purification of nucleic acids by RPC-5 ANALOG chromatography: peristaltic and gravity-flow applications. Methods Enzymol. 100:368–399.

45. Thompson, R., and L. Taylor. 1982. Promoter mapping and DNA sequencing of the F plasmid transfer genes traM and traJ. Mol. Gen. Genet. 188:513–518.

46. Timmis, K. N., I. Andres, and M. Achtman. 1978. Fertility repression of F-like conjugal plasmids: physical mapping of the R6-5 finO and finP cistrons and identification of the finO protein. Proc. Natl. Acad. Sci. USA 75:5836–5840.

47. von Hippel, P. H., D. G. Bear, W. D. Morgan, and J. A. McSwiggen. 1984. Protein-nucleic acid interactions in transcription: a molecular analysis. Annu. Rev. Biochem. 53:389–446.

48. Willetts, N. 1977. The transcriptional control of fertility in F-like plasmids. J. Mol. Biol. 112:141–148.

49. Willetts, N., J. Maule, and S. McIntire. 1976. The genetic locations of traO, finP, and tra-4 on the E. coli K12 sex factor F. Genet. Res. 26:255–263.

50. Willetts, N., and B. Wilkins. 1984. Processing of plasmid DNA during bacterial conjugation. Microbiol. Rev. 48:24–41.

51. Willetts, N. S. 1971. Plasmid specificity of two proteins required for conjugation in E. coli K12. Nature (London) New Biol. 230:183–185.