MICROBIOLOGICAL REVIEWS, June 1994, p. 162–210
0146-0749/94/$04.00+0

Vol. 58, No. 2

Analysis of the Sequence and Gene Products of the Transfer Region of the F Sex Factor

LAURA S. FROST,[1]* KARIN IPPEN-IHLER,[2] AND RONALD A. SKURRAY[3]

Department of Microbiology, University of Alberta, Edmonton, Alberta, Canada T6G 2E9[1]; Department of Medical Microbiology and Immunology, Texas A&M University, College Station, Texas, 77843-1114[2]; and School of Biological Sciences, University of Sydney, NSW 2006, Australia[3]

INTRODUCTION

Bacterial conjugation is the process by which DNA is transferred from one bacterium to another, usually through a mating bridge requiring close contact between the donor and recipient cells. DNA transfer is initiated at a specific sequence called the origin of transfer (*oriT* or *nic*) in response to an unidentified signal generated by the genes involved in mating-pair formation (*tra*) and transferred to the genes involved in DNA transfer (*tra*) or mobilization (*mob*). As such, the events resulting in DNA transfer represent a prototype for cellular recognition and communication between cells.

The mechanism of F-mediated conjugation has also been the subject of several reviews (118, 120, 266, 267). Briefly, the F pilus, an extracellular filament expressed by the transfer region of the plasmid-bearing donor cell, creates a specific contact with one or more recipient cells, leading to the formation of a mating pair or aggregate. Once the mating aggregate is stabilized, a single strand of DNA is transferred in a 5′-to-3′ direction, beginning at the nick site of the origin of transfer (*oriT*). The entire process of F plasmid transfer (100 kb), which takes approximately 5 min, results in recircularization of the transferred strand and synthesis of complementary DNA in both the donor and recipient cells; both are then capable of plasmid transfer. With the realization that intergeneric (101) and interkingdom (110) transfer of the F plasmid is possible (reviewed in reference 174), the nature of the mating bridges involved in this unusual DNA transport process, as well as the evolutionary significance of conjugation, have become questions of primary importance.

All the sequences required for the conjugative transmission of the *Escherichia coli* K-12 fertility factor, F, are encoded within the 33.3-kb transfer region of this 100-kb IncFI plasmid.

* Corresponding author. Mailing address: Department of Microbiology, M330 Biological Sciences Building, University of Alberta, Edmonton, Alberta, Canada T6G 2E9. Phone: (403) 492-0458. Fax: (403) 492-2216.

F is considered the paradigm for plasmid-specified transfer systems and has been the subject of study for almost 50 years. Bacterial conjugation was first identified as a function of the F plasmid, and a history of the genetic analysis of F transfer has recently been written by Willetts (262). The F plasmid was used to develop many of the genetic techniques commonly used to dissect procaryotic systems, and F product analysis has been central in elucidating the basic mechanisms of plasmid replication and transmission. Although progress has also been made toward understanding other plasmid transfer systems, such as those specified by plasmids of incompatibility groups IncI, IncP, IncN, and IncW, the focus of this review will be the F plasmid and the related plasmids in the IncF complex. Although F (IncFI) is the only plasmid in the IncF complex for which the sequence of the entire transfer region has been determined, sequences of portions of the transfer regions of R386 (IncFI), R1, R100, R136, ColB2-K77, ColB4-K98, R6-5, R538-1 (IncFII), pSU316 (IncFIII), R124 (IncFIV), pED208 (IncFV), and pSU233 (IncFVII) are available for comparison as discussed in the appropriate sections below. In most cases studied, the transfer regions of these plasmids exhibit a high degree of homology with each other and with F. Some of their transfer genes are allelic with the F transfer genes in that they clearly are derived from a common ancestor but are unable to complement mutations in the corresponding gene of the F plasmid. This variation in sequence is the basis of the specificity of the transfer systems of these plasmids.

This article is meant to serve as a reference for the complete sequence of the F transfer region and a compilation of the properties of the gene products that it is known to encode. The information currently available for the size, cellular distribution, and function of F transfer region gene products is summarized, and, when possible, comparative data for the allelic sequences and analogous products of other F-like transfer systems are presented. This information should be of use in the purification and characterization of the various components of the F transfer system and the elucidation of the mechanism of mating-pair formation and DNA transfer.

Conjugative pilus morphology and pilus-specific phage sensitivity classifications have suggested that pilus-dependent transfer systems might belong to two evolutionary "families" (see references 75, 118, and 206). One family includes the IncF complex and other plasmids with some F-like characteristics (e.g., IncD, IncC, and IncJ), whereas the second includes plasmids with phenotypic resemblance to RP4 (IncP). Sequence analysis has shown that the Ti plasmid of *Agrobacterium tumefaciens* also belongs to the RP4 family, and homology, occasionally extensive, has been detected among the transfer regions of broad-host-range plasmids from the IncP, IncN, IncW, and IncI groups, the *vir* region of the Ti plasmid of *A. tumefaciens*, and the mobilization regions of plasmids pTF-FC2, RSF1010, pTF1, and pSC101 (33, 52, 86, 102, 143, 147, 203, 221, 226). Recently, fundamental functional relationships have been revealed between these two families by using multiple-alignment algorithms. Domains in two F *tra* proteins involved in DNA transfer (TraD and TraI) were shown to exhibit homology to domains in analogous products encoded by the *tra* regions in RP4 (IncP) and R388 (IncW) and the *vir* region of the Ti plasmid (42, 126, 144, 256). A further demonstration of homology between F TraD and the staphylococcal conjugative plasmid pSK41 TraK protein (69) now suggests that the mechanism of DNA transfer is conserved throughout the eubacteria. There also appears to be a distant relationship between the products of F *traA* (F pilin) and F *traL* and pairs of products encoded by the Ti plasmid *virB* operon (VirB2 and VirB3) and the region of RP4 involved in pilus biosynthesis (TrbC and TrbD) (126, 143, 230). Similarity between F TraC and proteins required for Ti plasmid virulence, toxin biosynthesis in *Bordetella pertussis*, and transfer in RP4 (IncP) and R388 (IncW) has been reported (126, 231). By using similar techniques, F TraY has also been identified as a member of the family of proteins related to the P22 phage repressors Arc and Mnt (23). One can expect that relationships between F transfer proteins and products expressed by the other systems will continue to emerge as additional sequence and functional data accumulate.

ORGANIZATION OF THE F TRANSFER REGION SEQUENCES

Figure 1 summarizes the coding sequences and functions of gene products in the F transfer region compiled from the DNA sequence published or made available by various laboratories over the last 10 years. As drawn in Fig. 1, transcription of *tra* (tall boxes, capital letters), *trb* (short boxes, lowercase letters), and *finO* gene sequences proceeds from left to right: the opposite strand contains the gene (*finP*) for a small regulatory antisense RNA, FinP, and an open reading frame (ORF) for a gene of unknown function, *artA*. Kilobase coordinates in Fig. 1 correspond to those of the sequence in Fig. 2, where nucleotide (nt) 1 is the first nucleotide in the *Bgl*II site, located at position 66.6F on the 100-kb F plasmid map (266). The orientation of the origin of transfer, *oriT*, located immediately downstream from this site, determines the direction of conjugal DNA transport. The first DNA sequences to be transported to the recipient cell include the *Bgl*II site and ORF169 in the leading region (see Leading Region of Transfer, below); sequences encoding the transfer region enter the recipient last.

The distribution of loci for the different functions within the F *tra* region is also indicated in Fig. 1. The transfer systems of most F-like plasmids express both of the regulatory genes *finO* and *finP* and are typically repressed; an IS*3* insertion in the F *finO* gene is responsible for the constitutive transfer system characteristic of the F plasmid. FinOP repression normally reduces expression of the TraJ regulatory protein, which is required for efficient initiation of transcription at the major transfer region promoter, P_Y, located between *traJ* and *traY*. Studies on the structure and DNA-binding activity of TraY, encoded by the first gene in the *tra* operon, suggest that it also acts as a transcriptional regulator of *tra* protein expression. The IncFV plasmid F_0 *lac* appears to have a different system from that of F for controlling *traJ* expression since there is no apparent homolog for the *finP* gene. Additional functional groups include the genes involved in synthesis and assembly of F pilus filaments, surface exclusion, mating-pair or aggregate stabilization, and transfer of the single strand of DNA to the recipient; the functions of the other loci remain unclear. Maximal expression of these genes (with the exception of *traM* and *traJ*) occurs when the P_Y promoter is active and stimulated by TraJ.

The F *tra* region DNA sequence presented in Fig. 2 is a compilation of segments of sequence starting from the *Bgl*II site following coordinate 0 in Fig. 1 and includes *oriT*, *traM*, *traJ*, *traY* (72, 243, 244), *traA*, *traL*, *traE* (83), *traK*, *traB*, *traP*, *trbD* (74), *trbG*, *traV*, *traR* (51), *traC*, *trbI*, *traW* (164, 224), *traU*, *trbC*, *traN*, *trbE* (161, 163, 188), *traF*, *trbA*, *artA*, *traQ*, *trbB*, *trbJ* (269–271), *trbF*, *traH*, *traG* (70, 105), *traS*, *traT* (104, 123), *traD*, *trbH*, *traI* (25, 124, 274), *traX*, and *finO* (35). The entire sequence can be accessed through the GenBank accession number U01159.

Promoter regions, binding sites, and other features noted in Fig. 2 are discussed in corresponding sections of the text. The

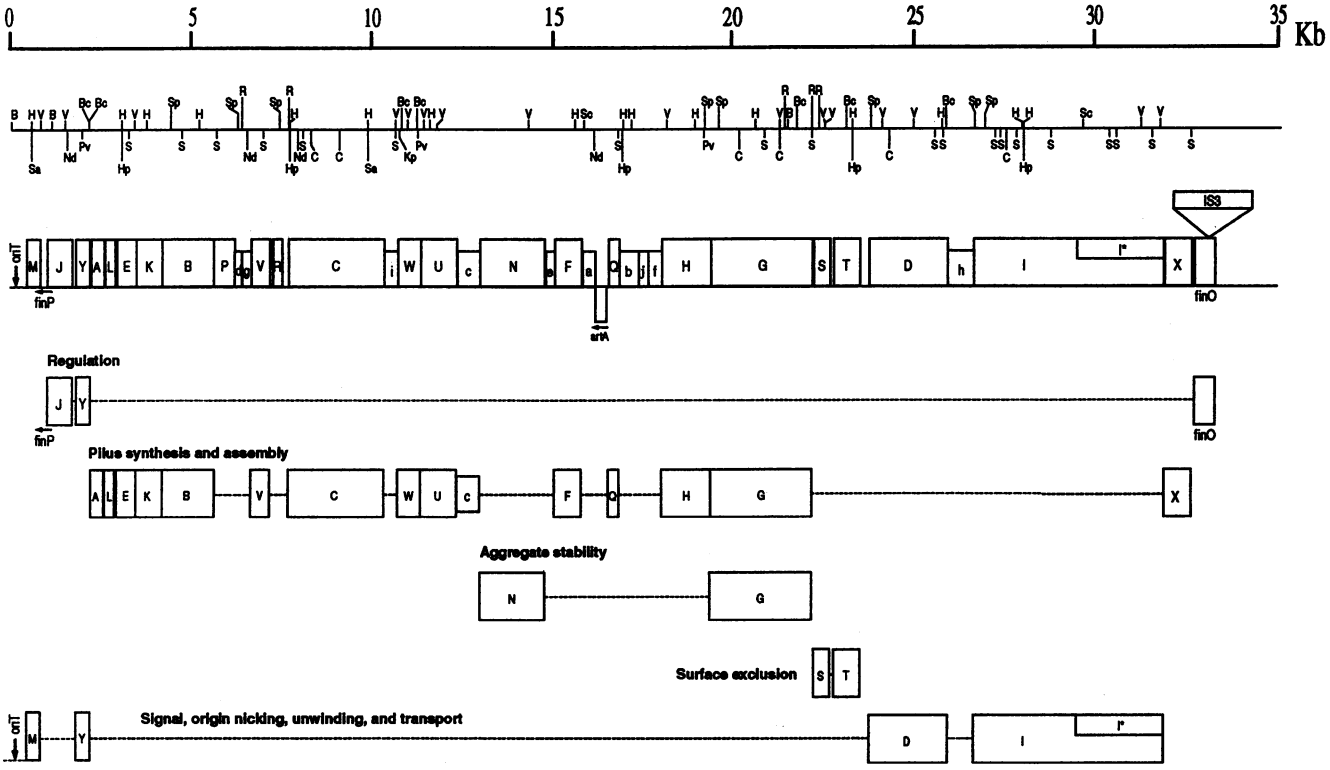


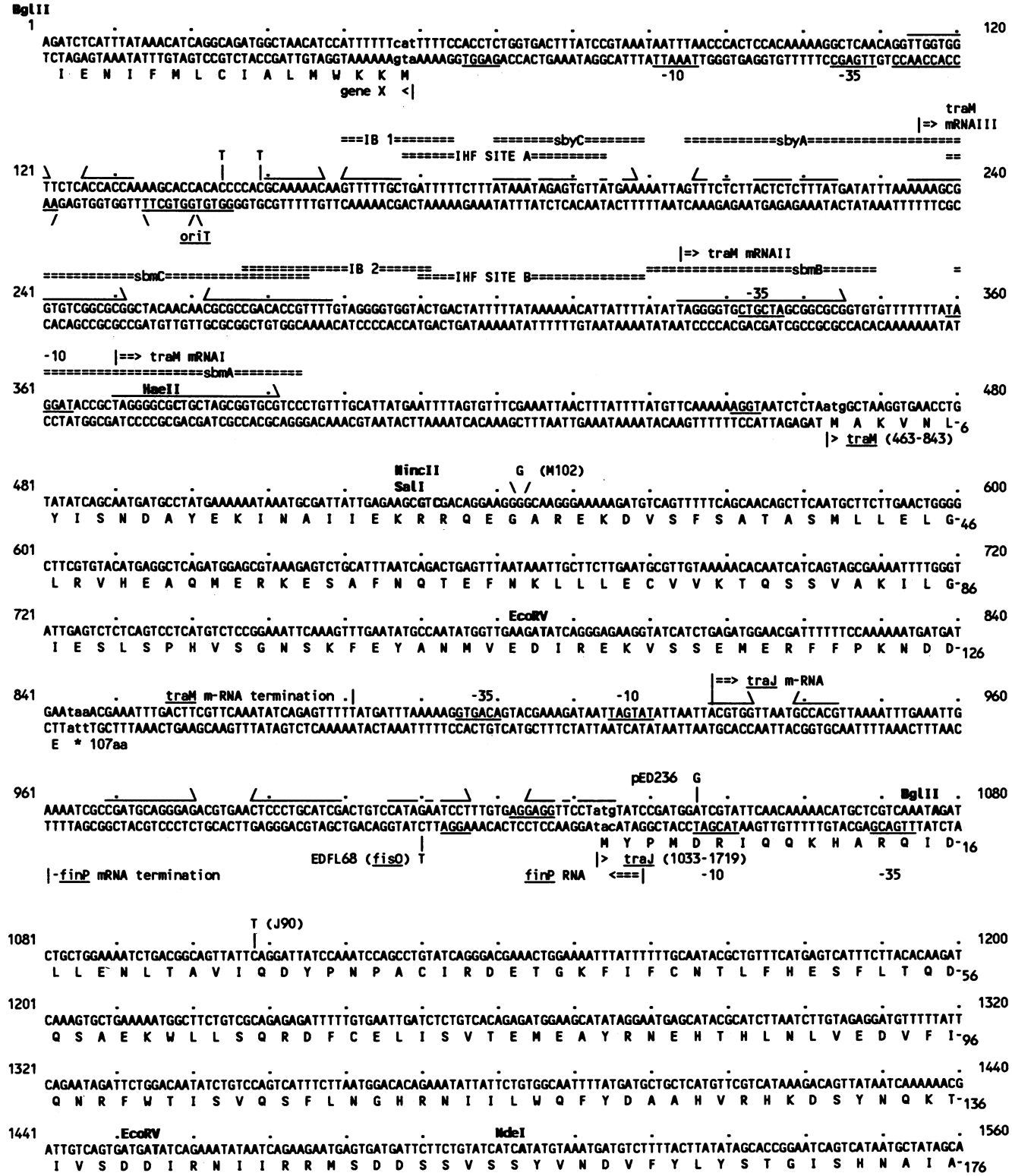
FIG. 1. View of the transfer region. The details of the figure are explained in the text. The top line gives the length in kilobases of the F transfer region. The second line indicates the position of restriction sites referred to in the text and other figures. Abbreviations: B, *Bgl*II; Sa, *Sal*I; V, *Eco*RV; Nd, *Nde*I; Bc, *Bcl*I; H, *Hin*cII; Hp, *Hpa*I; S, *Sma*I; Sp, *Sph*I; C, *Cla*I; R, *Eco*RI; Kp, *Kpn*I. The third line represents the genes and gene products; capital letters refer to *tra* genes, lowercase letters refer to *trb* genes, *artA* and *finP* are transcripts in the anti-orientation, *oriT* is the origin of transfer, IS*3* refers to the IS*3* element within the *finO* gene, and I* refers to a gene product encoded within the *traI* gene. The last five lines represent the functions of the *tra* genes identified to date.

sequence alterations associated with various nonsense and missense mutations are marked above the sequence. Restriction enzyme cleavage sites mapped in Fig. 1 are indicated in boldface type in Fig. 2; positions of a few additional sites mentioned in the text are noted in parentheses. Prominent inverted and direct repeats are indicated, as well as known and potential binding sites for DNA-binding proteins. An enormous number and variety of subclones carrying portions of this sequence have been generated, and readers desiring a particular construct are advised to contact the investigator involved directly. An early enumeration of subclones in plasmids and lambda lysogens was compiled by Willetts and Skurray (265); however, the list has grown dramatically in recent years.

TRANSCRIPTION OF THE F TRANSFER REGION

The putative −35 and −10 binding sites and transcriptional initiation sites that have been identified within the *tra* region are also indicated in Fig. 2. These sites include those originally described for P_M, the *traM* promoter (M mRNAI, nt 332 to

FIG. 2. Sequence of the F transfer region starting from the first base of the sequence defining the *Bgl*II site. The details of the genes and gene products are given in the text. The sequence is available through GenBank as accession number U01159. The nucleotide sequence is marked at every 10th nucleotide, when possible, and the number of nucleotides is given at the end of each line. The single-letter code for the sequence of the most probable ORFs is indicated below the nucleotide sequence, and a running total of the number of amino acid residues is given as a subscript at the end of each line except for the last line of sequence for any given protein. * indicates a stop codon. The sequences of both strands of the DNA are given for intergenic regions as well as the anti-*tra* oriented gene *artA*. ⊢ indicates the start of the translation; the name of the gene product and the first and last nucleotide in the coding sequence are given in brackets. Mutations in the proteins are indicated by a vertical line, with the base change, name of the mutation, and resultant amino acid substitution given above the line. ==> indicates the start of transcription; the −10 and −35 sequences for promoters are underlined, and the name of the transcript and the position of the end of the transcript are also indicated. Direct and inverted repeats in the sequence are indicated by arrows, and binding sites for DNA-binding proteins are given as double-dashed lines. The core stability region of the transfer mRNA (nt 2633 to 2798) is indicated by a dashed line, and the smaller inverted repeats within this region are marked as a double-dashed line. Signal sequence cleavage sites are indicated by vertical arrows below the line. Features of interest, including homology to signature sequences mentioned in the legends to Fig. 3 and 4, are underlined and described below the line. The nick site (/\, nt 140 to 141 according to reference 173), the site of the IS*3* insertion, and the positions of frameshift mutations (\/) are indicated below and above the line, respectively. The restriction sites for common restriction endonucleases are given in boldface type above the line, and the nucleotide on the 5′ side of the cut site is also in boldface type. Nucleotides in parentheses above the line in the *traX-finO* region (nt 32466 to end) represent the sequence of the F plasmid as reported by Yoshioka et al. (274).


BglII
gene X
traM
mRNAIII
sbyC
sbyA
IHF SITE A
oriT
traM mRNAII
sbmC
IB 2
IHF SITE B
sbmB
traM mRNAI
sbmA
MaeII
traM (463-843)
HincII
SalI
G (M102)
traM m-RNA termination
EcoRV
traJ m-RNA
pED236
BglII
EDFL68 (fisO)
traJ (1033-1719)
finP mRNA termination
finP RNA
T (J90)
NdeI

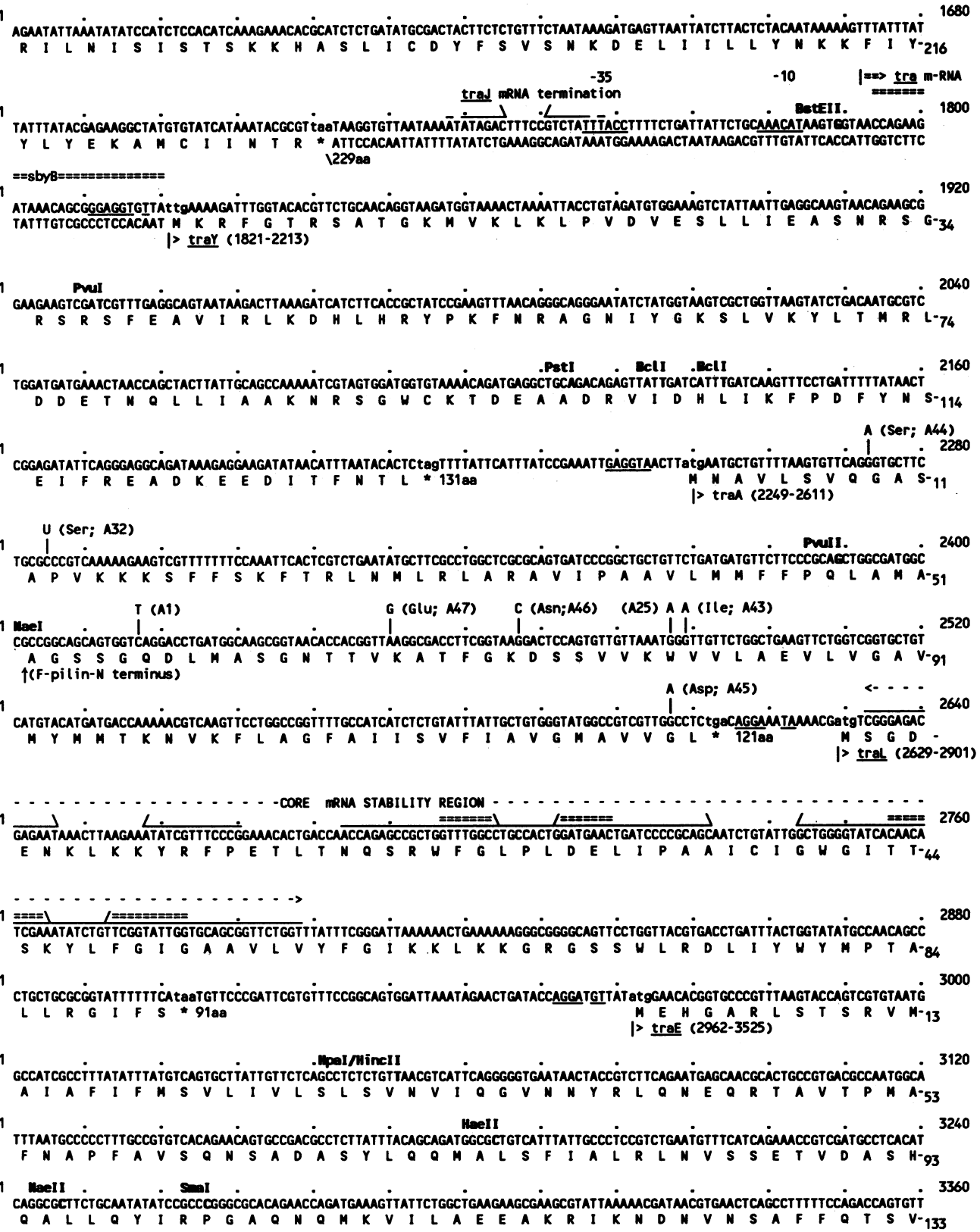


FIG. 2—*Continued.*

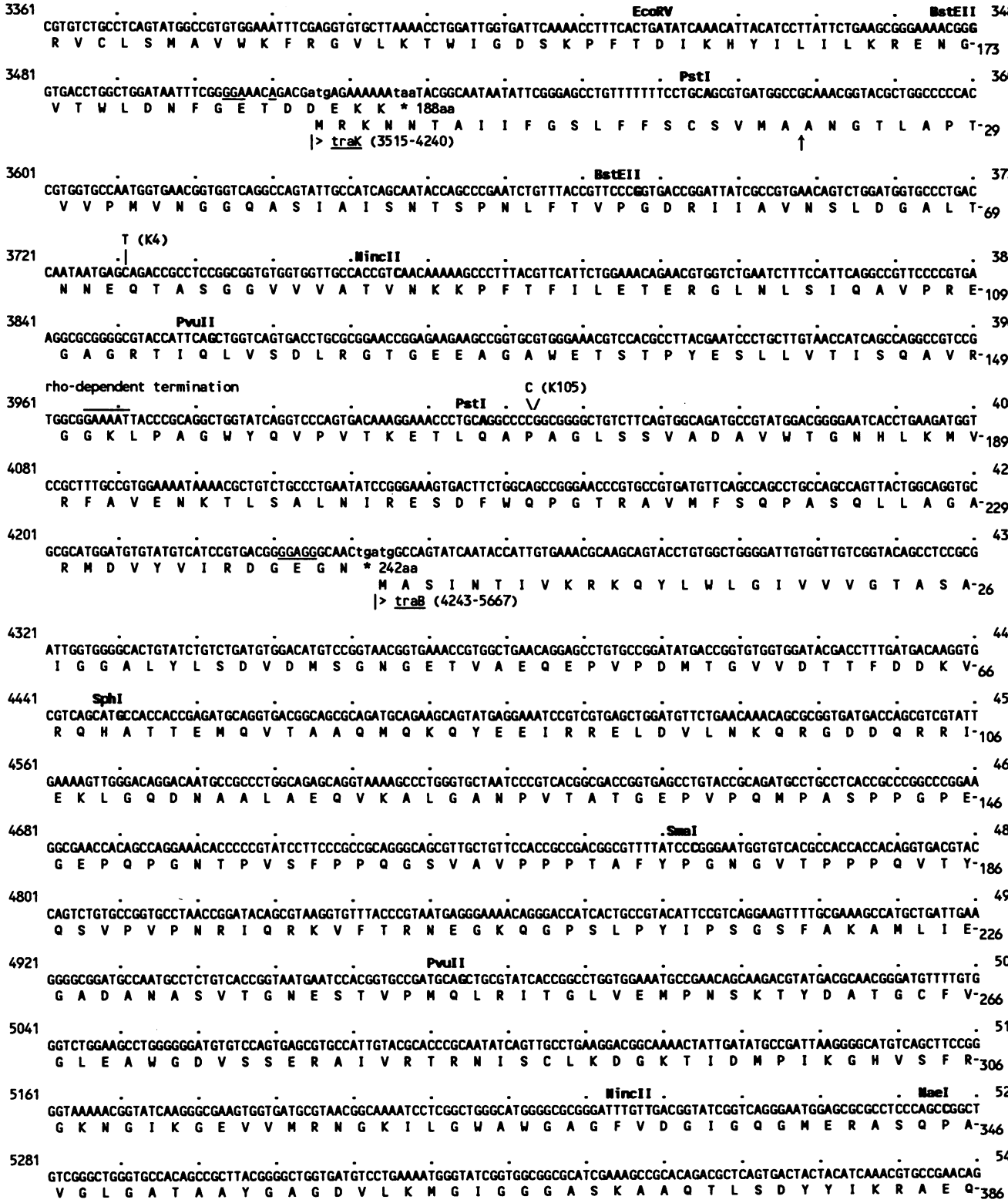


FIG. 2—*Continued.*

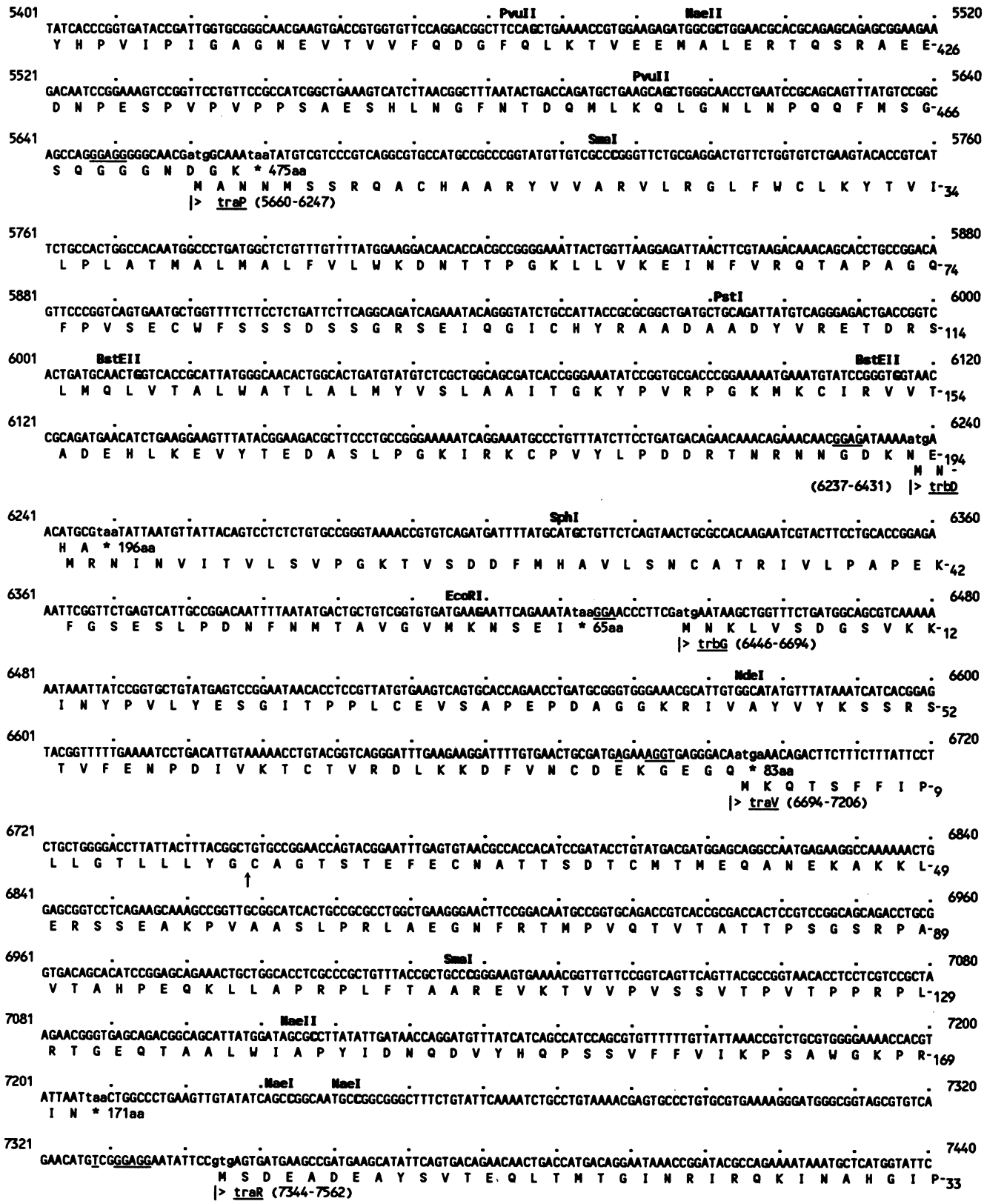


FIG. 2—*Continued.*

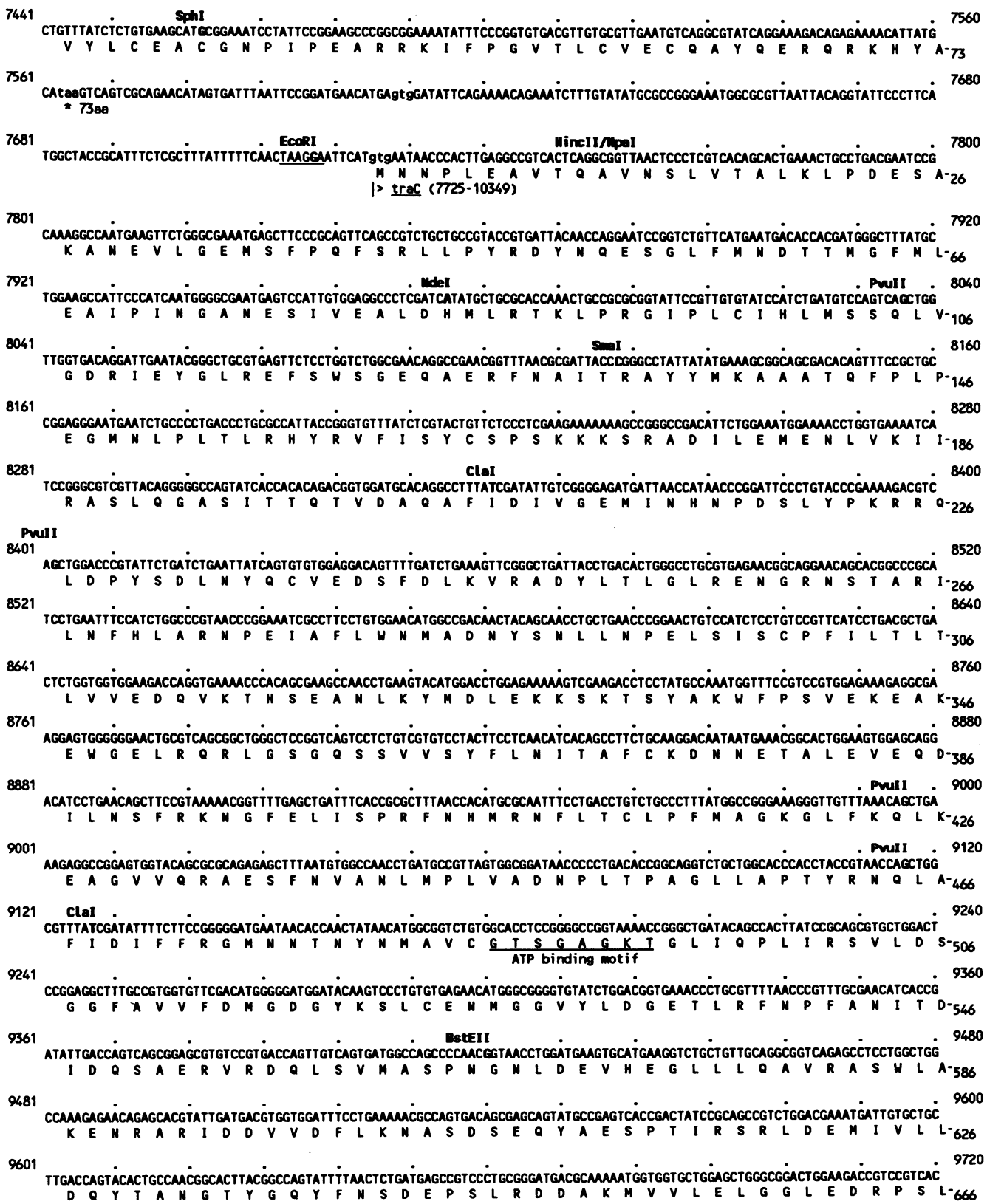


FIG. 2—*Continued.*

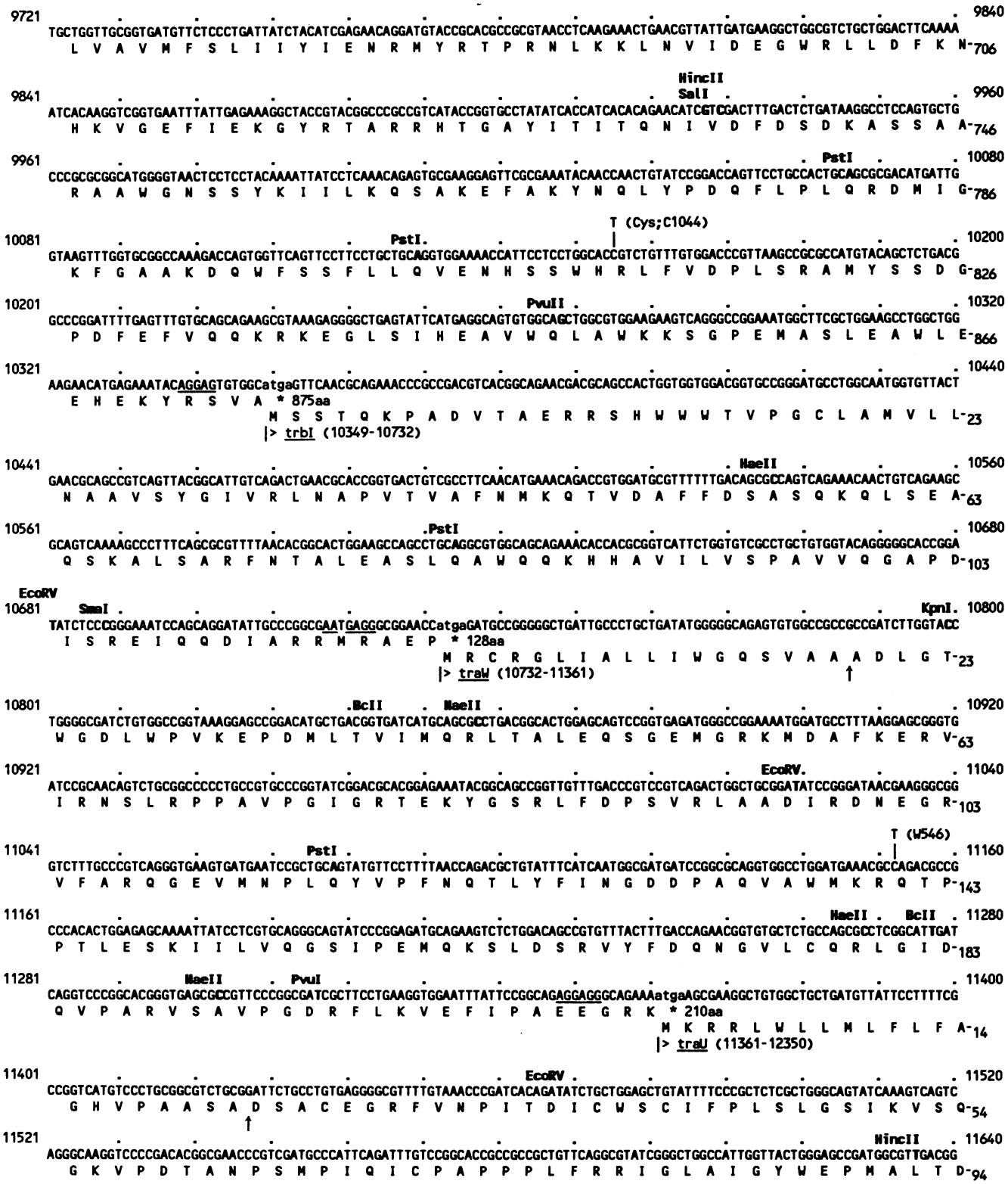


FIG. 2—*Continued.*

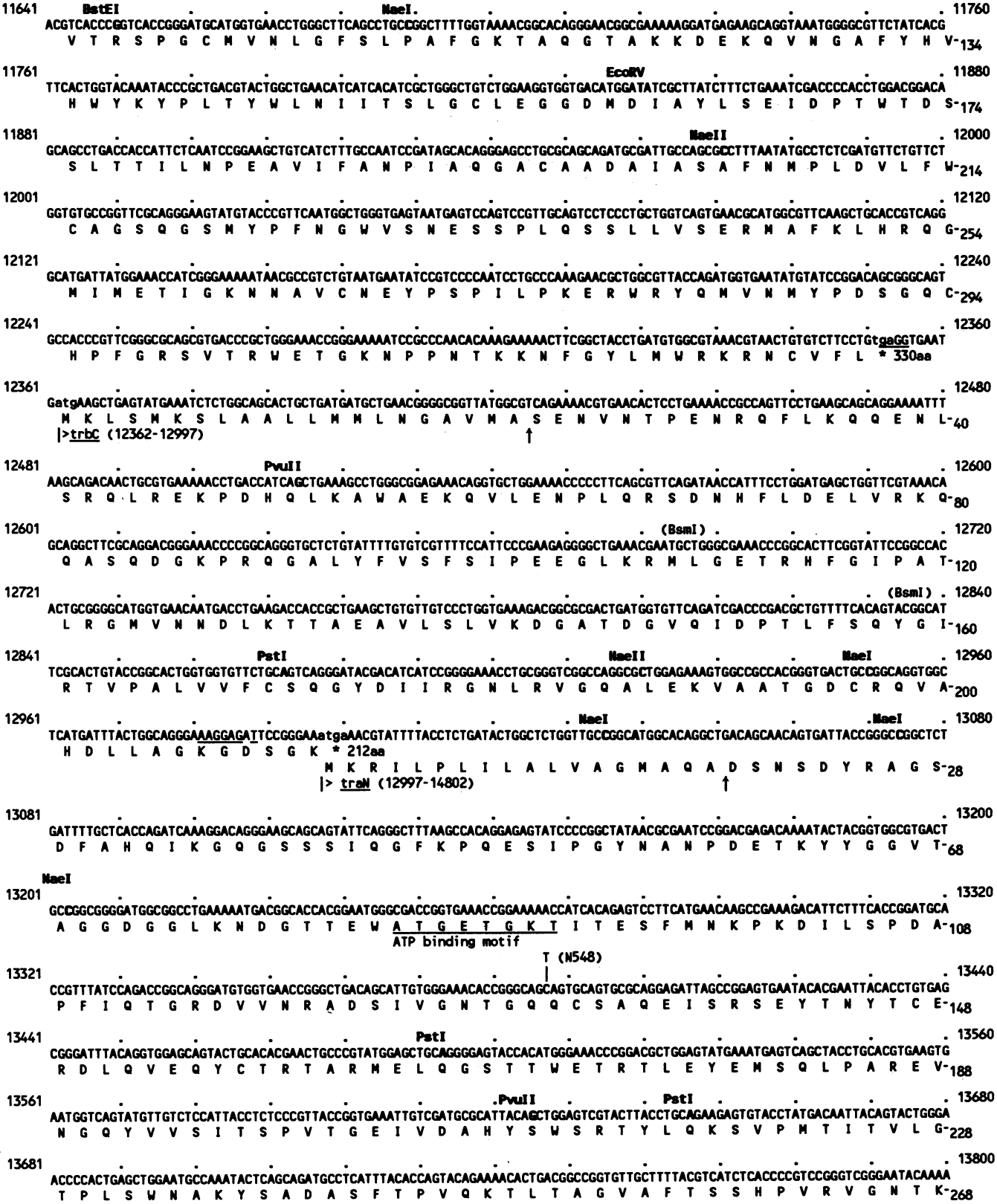


FIG. 2—*Continued.*

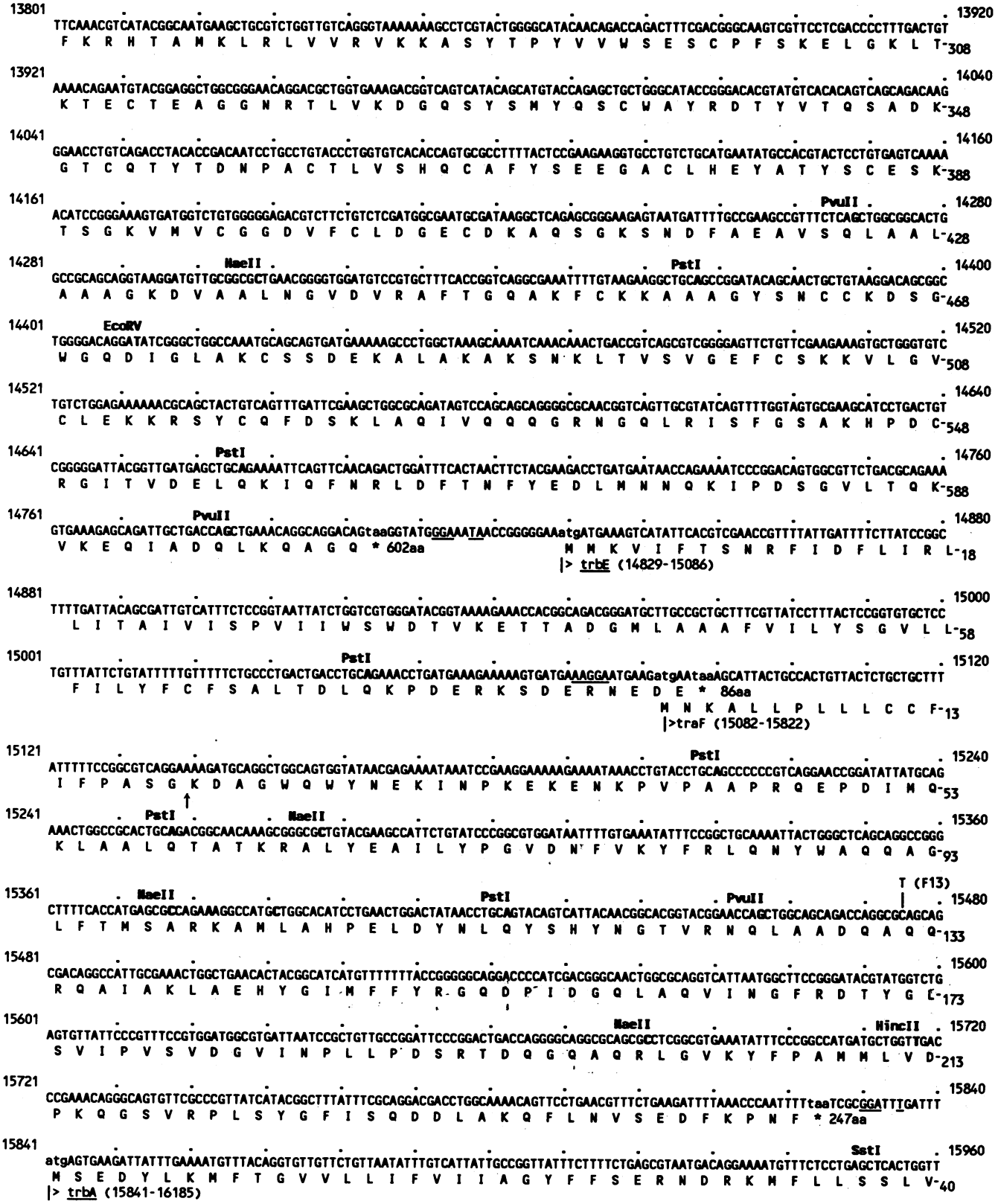


FIG. 2—*Continued.*

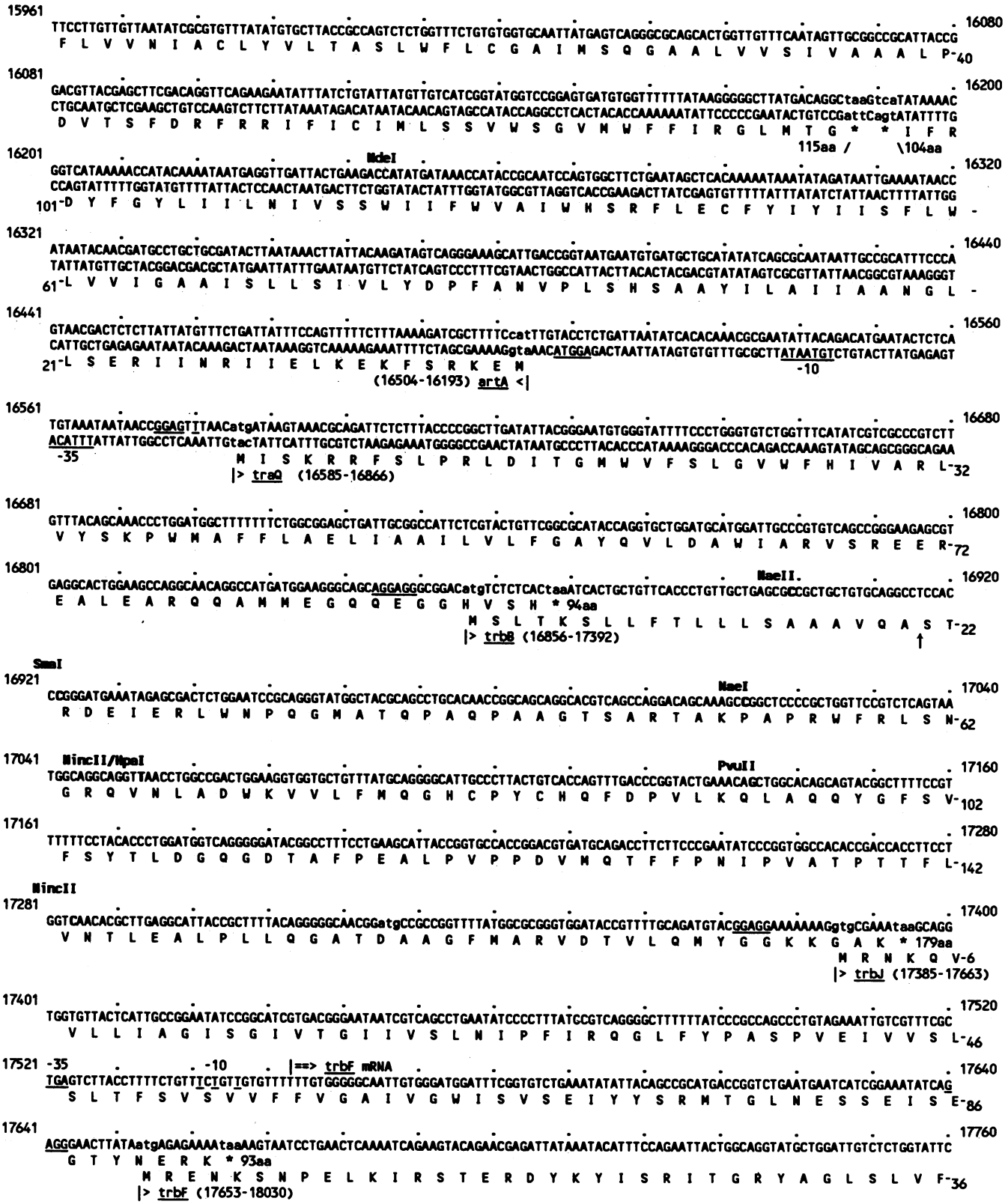


FIG. 2—*Continued.*

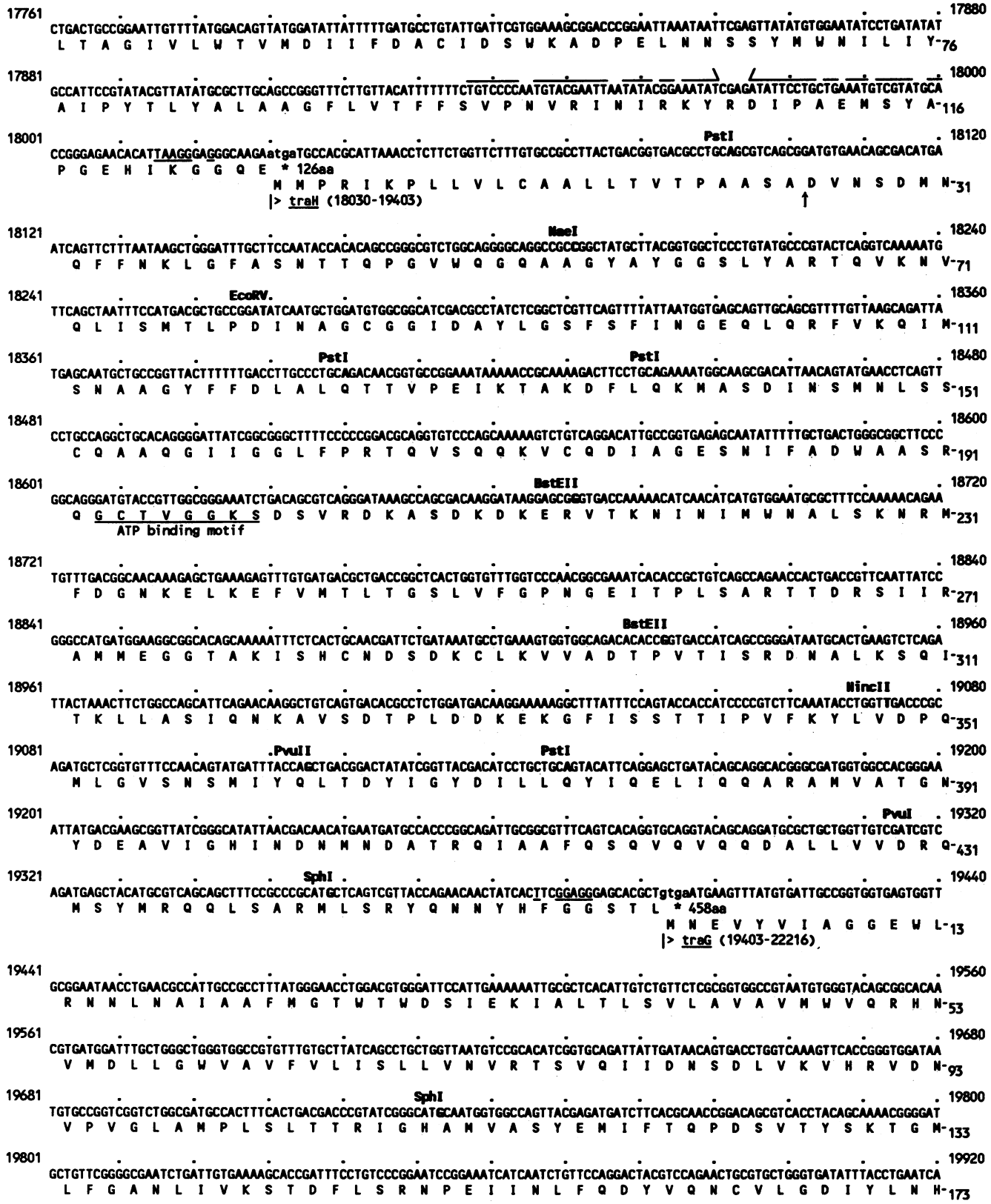


FIG. 2—*Continued.*

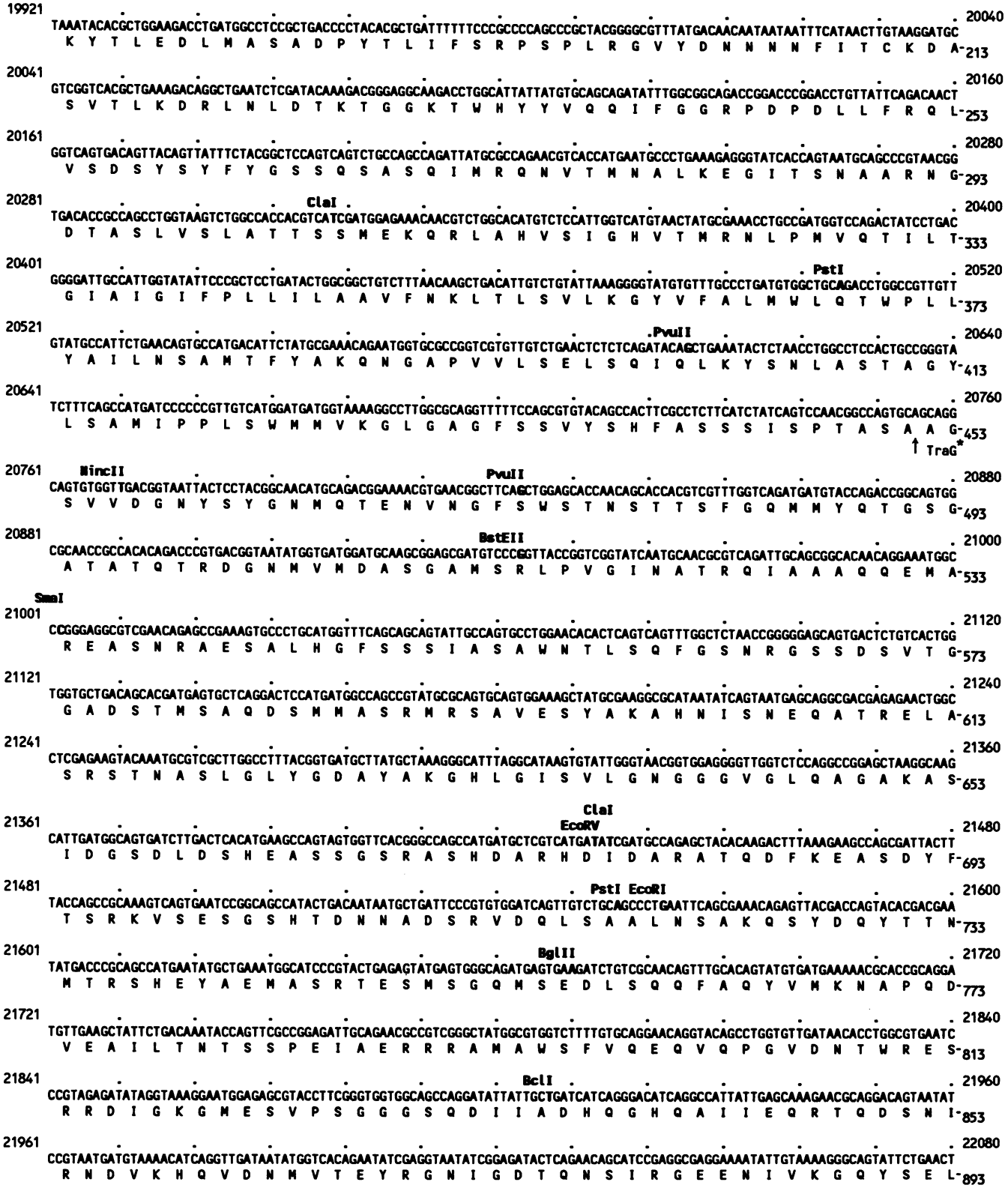


FIG. 2—*Continued.*

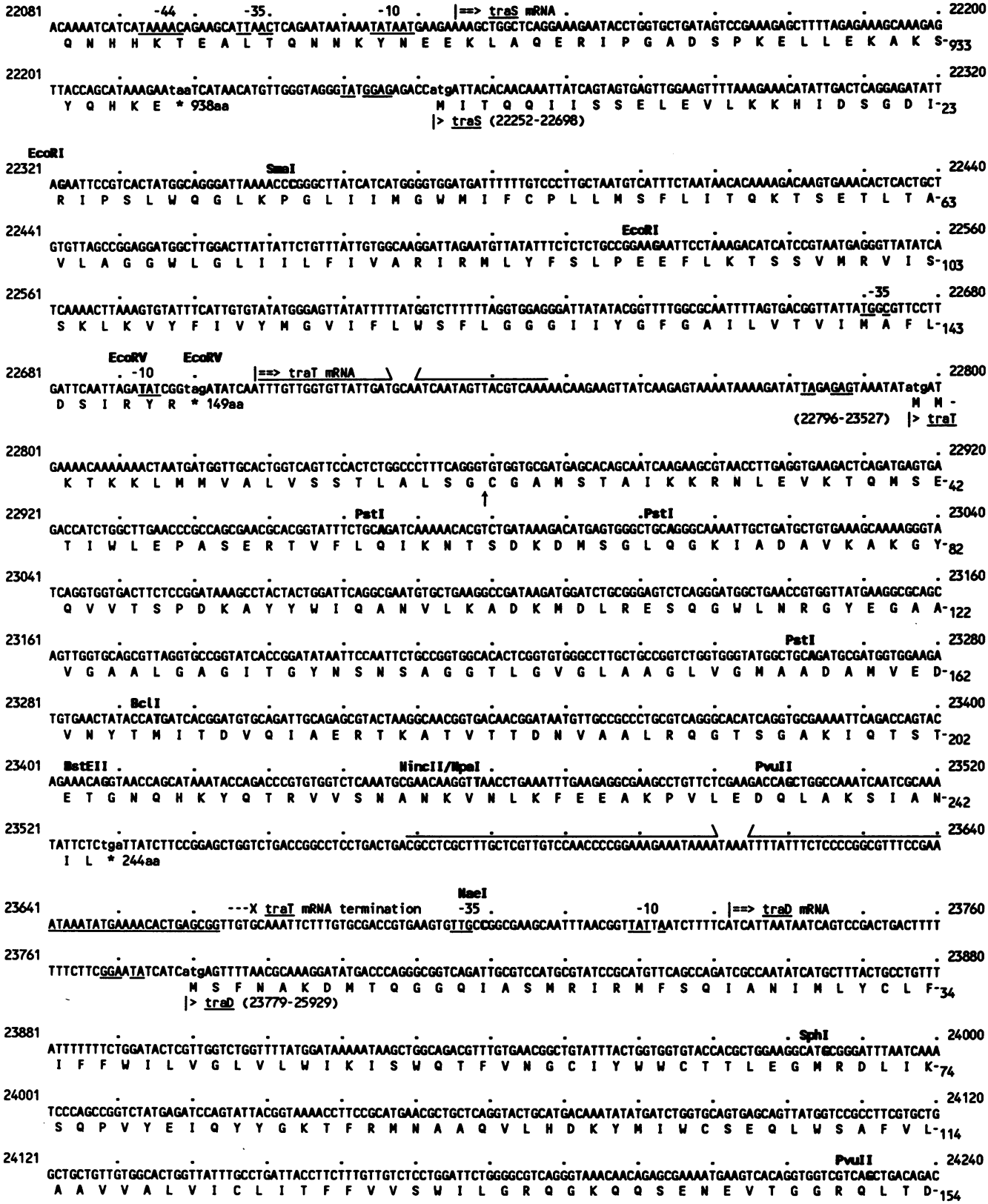


FIG. 2—*Continued.*

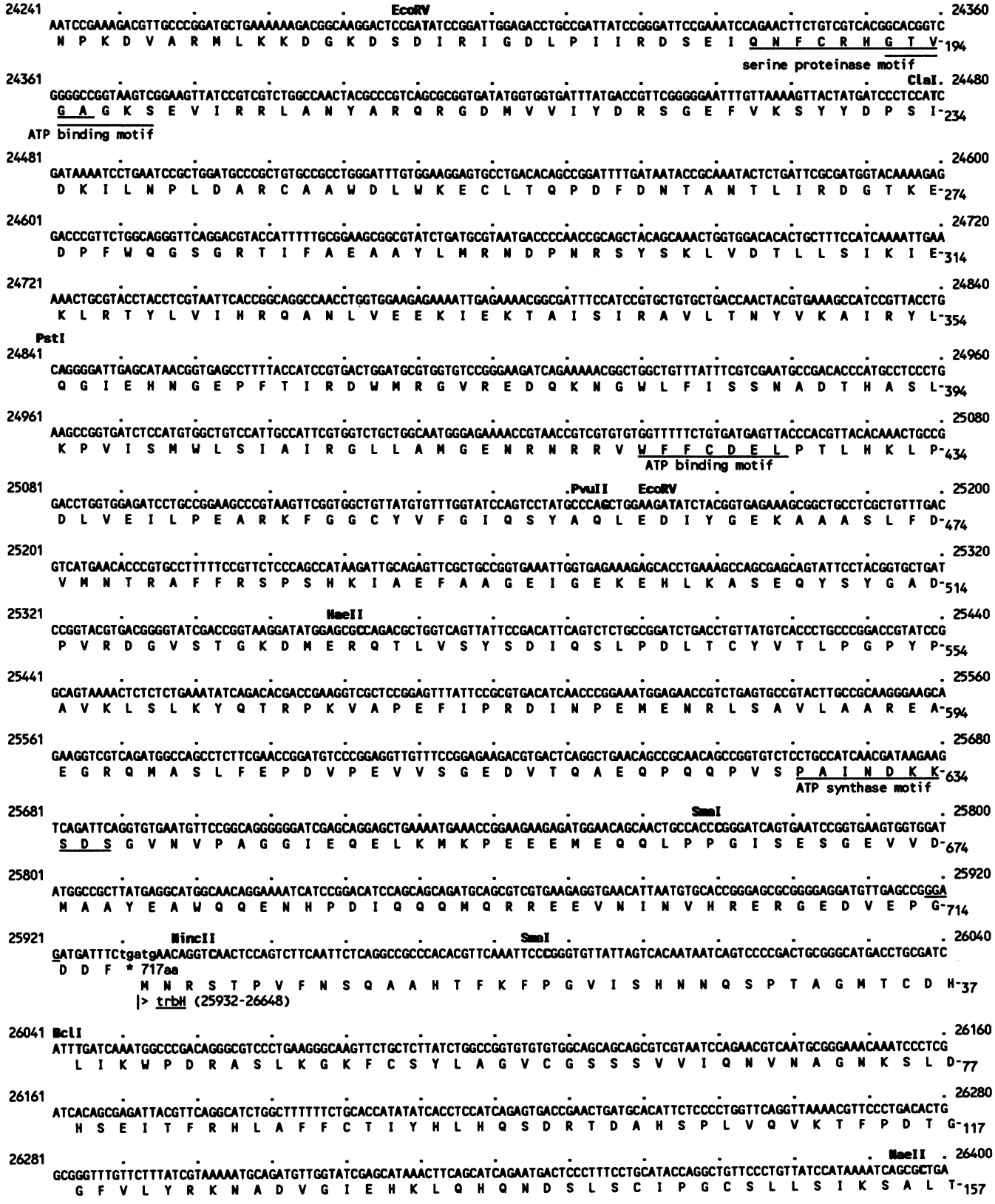


FIG. 2—*Continued.*

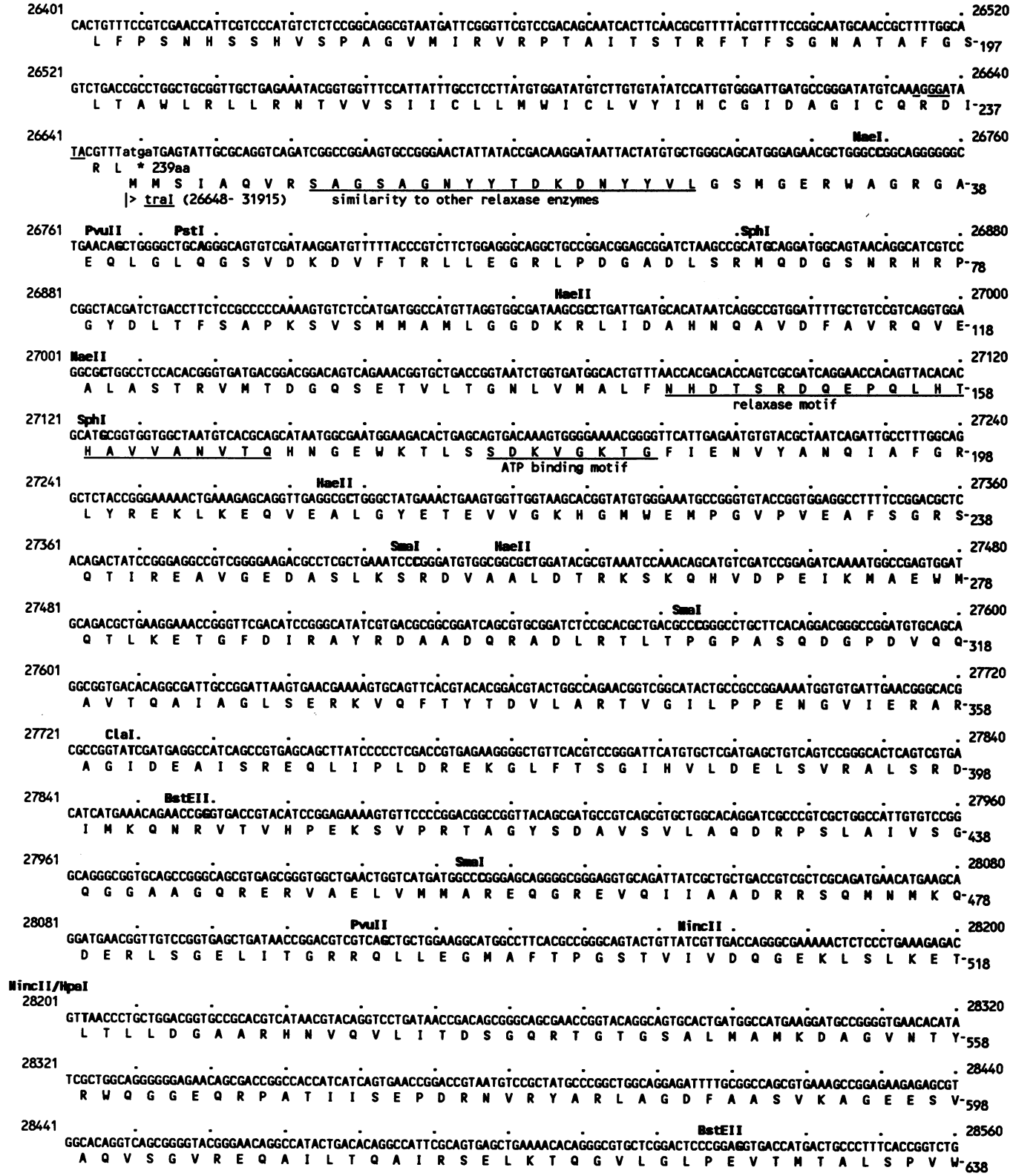


FIG. 2—*Continued.*

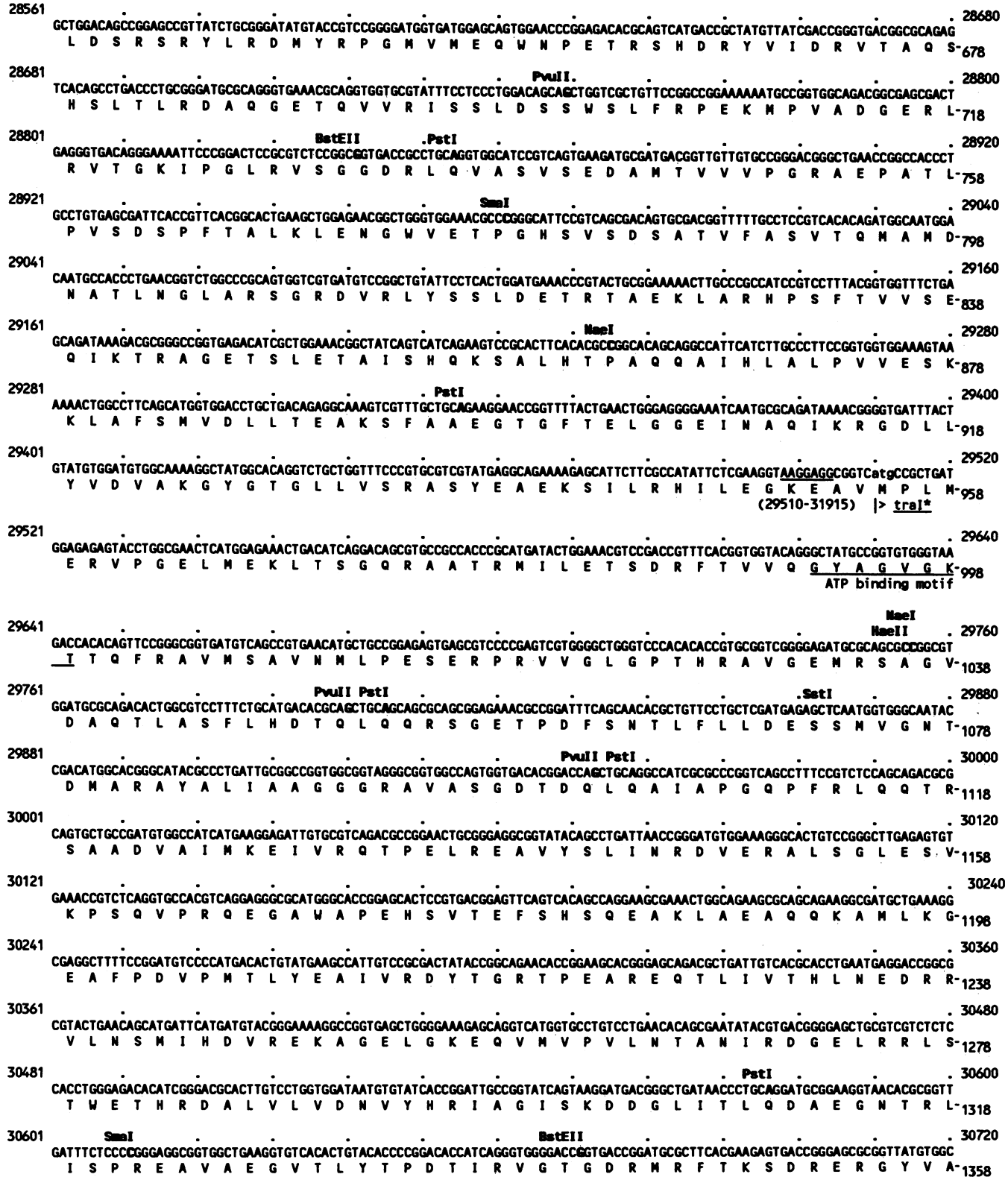


FIG. 2—*Continued.*

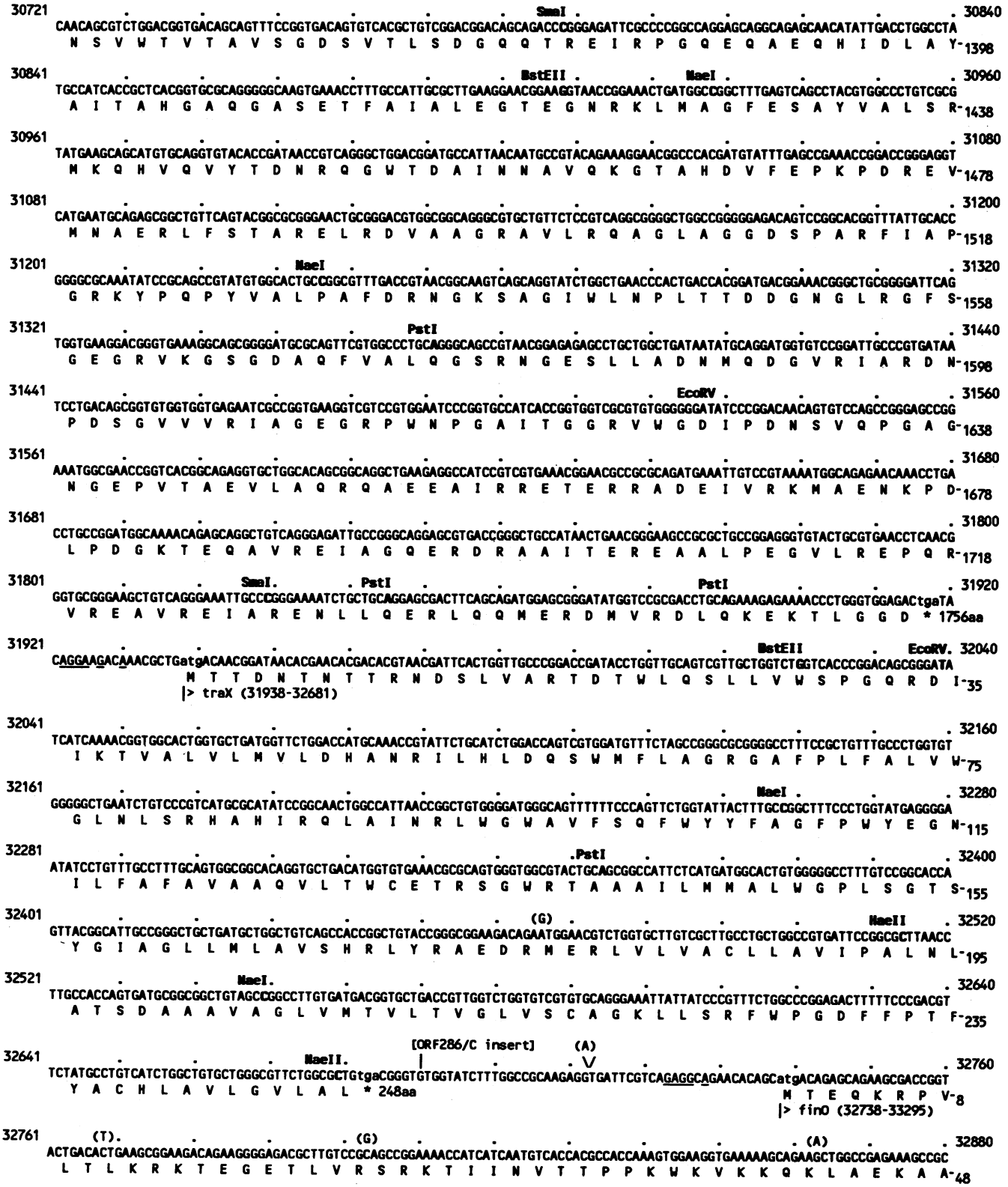


FIG. 2—*Continued.*

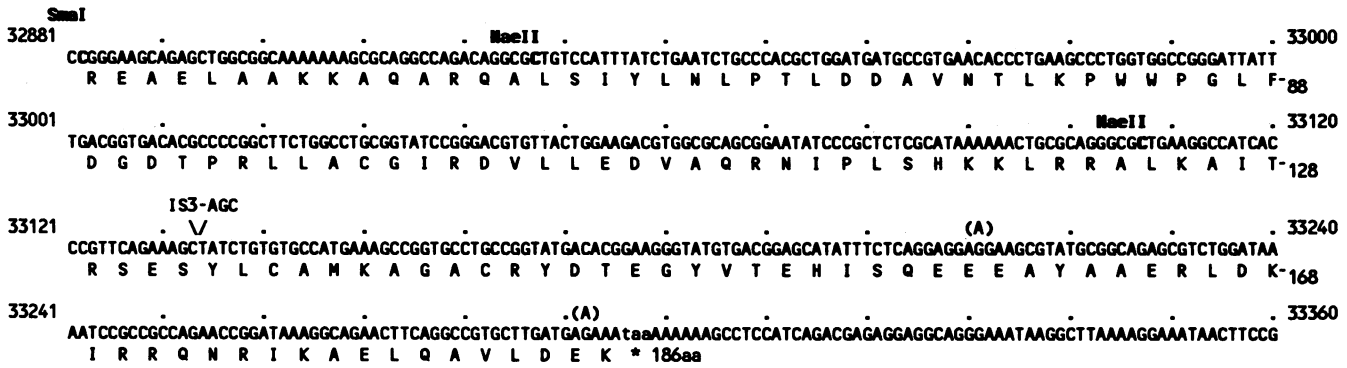


FIG. 2—*Continued.*

370) (243), as well as the initiation sites recently described which are closer to *oriT* (M mRNAII, nt 291 to 324; M mRNAIII, ≅nt 200 to 235) (207); for P_J, the *traJ* promoter (nt 895 to 928) (243); for P_{finP}, the promoter for the antisense RNA gene, *finP* (nt 1074 to 1039) (72, 251); for P_Y, the major transfer operon promoter (nt 1756 to 1792) (90, 234, 235), for P_{artA} the anti-*tra* oriented gene, *artA* (nt 16566 to 16532) (269); and for P_{trbF}, P_{traS}, P_{traT}, and P_{traD}, promoters preceding *trbF* (nt 17521 to 17553) (105), *traS* (nt 22093 to 22135) (104), *traT* (nt 22670 to 22708) (104), and *traD* (nt 23695 to 23732) (104), respectively. Interestingly, the promoters for *trbF* and *traS* are located within the upstream coding sequences for *trbJ* and *traG*, respectively. In the R1 plasmid, the *traY* promoter has been mapped within the C-terminal coding region of *traJ*, suggesting an in *cis* element of control in the regulation of the P_Y promoter (134).

Early studies indicated that genes *traM* and *traJ* are expressed from individual transcripts and that expression of the majority of *tra* genes is regulated at P_Y (the *tra* operon promoter) (111, 257). The polarity of phage insertion mutations, together with the phenotype of *traJ* mutants, suggested that a single TraJ-dependent transcript for the entire *tra* operon extends from *traY* to *traI* (120, 257). The recently characterized genes *traX* (35) and *finO* (250) might also be included in such a transcript. Thus, the maximum length of the P_Y transcript is not completely defined. *traM* expression is negatively autoregulated by TraM (207), and the finding that there is a transcript for the *traM* gene originating from the region bound by TraY in *oriT* (Fig. 2; sbyA, nt 195 to 238) suggests that TraJ may also have an indirect effect on transcription of *traM* (discussed in the section on the TraM protein, below) (207).

Recent studies have identified promoters from which distal *tra* genes can be expressed, and there is convincing evidence for P_Y-independent, constitutive expression of the surface exclusion genes *traS* and *traT* (31, 104, 124, 212) and of genes *traD* (104) and *traI* (25, 194). A promoter preceding *trbF* may also contribute to distal-gene expression (105). Whether transcription from P_Y extends through to enhance the expression of these genes is now in question. The identification of a strong transcriptional termination loop just distal to *traT* (nt 23569 to 23663) has led to the suggestion that the transcript from the P_Y promoter may not extend past this site and that transcription of *traD*, *traI*, and *traX* may be entirely dependent on distal promoters (35, 104). However, recent evidence indicates that deletion of the P_Y promoter region results in a significant decrease in the amounts of TraD and TraI protein that can be detected, even in the presence of *traJ*. This suggests that the normal level of their expression depends either on transcription from the P_Y promoter or on a product from the P_Y transcript (157).

TraJ-independent transcription from distal promoters could be important in the expression of gene products required for establishment of the plasmid immediately after entry into the recipient cell. For instance, expression of the genes for surface exclusion immediately after transfer could ensure against possibly damaging "supertransfer" from neighboring donor cells. Expression of TraD and TraI may facilitate establishment of a new relaxosome complex in the recipient cell and may play a role in termination of transfer and establishment of a membrane-binding site for the newly transferred DNA strand. Similarly, newly created donor cells, carrying F-like plasmids which are normally repressed (F is continually derepressed), become temporarily derepressed (high frequency of transfer) for several generations because of the TraJ protein synthesized before FinOP repression of *traJ* is complete (179, 261). This would allow for maximal expression of genes for transfer and surface exclusion and would promote conjugation at high levels to suitable recipient cells but would protect the donor cell from the lethal zygosis-like activity of other donor cells. The role of transient derepression in matings on solid surfaces has recently been discussed by Simonsen (237).

The P_Y transcript is expected to be long and to undergo posttranscriptional processing. In this case, gene products such as F pilin, which are needed in large quantities, should be expressed from the most stable segments of this transcript. The transcript for the R1 *traA* (pilin) gene has been shown to undergo both 5′ and 3′ processing to a stable core RNA (the equivalent F region is nt 2632 to 2799), protected at the 3′ end by a large hairpin structure in the *traL* sequence, which is conserved in the F sequence (nt 2684 to 2799) (131). The transcripts for the *traS* and *traT* genes also appear to be stable, with the *traT* gene bracketed by large inverted repeats between nt 22708 to 22747 and nt 23569 to 23663 (104).

A highly polar mutation, *traK4* (nt 3731) (11, 180, 263), has been shown to occur immediately prior to a potential intragenic rho-dependent terminator or *rut* sequence (nt 3966 to 3971) (208, 219). This conforms to the observation that polar mutations involving premature termination occur at sequences where C is more abundant than G, which define transcription termination elements (13). That such transcription termination elements occur regularly throughout the F *tra* sequence (208) fits with the suggestion that intragenic terminators may be invoked during times of stress to halt unnecessary transcription prematurely (219). Antitermination has also been implicated in the transcription of the *tra* operon. Mutations in the *rfaH*

gene (formerly *sfrB*) affect F transfer and can be suppressed in *rho* mutants, lending support to this idea (22). Thus, *rfaH* may encode a gene product which ensures the complete transcription of long operons such as those for transfer and lipopolysaccharide biosynthesis, possibly by altering the host RNA polymerase (22, 87, 210).

TRANSLATIONAL PRODUCTS OF THE F TRANSFER REGION

Figure 2 also shows the translational products of the most reasonable open reading frames in the F *tra* region (start and stop codons are given in lowercase type), including one for the *finO* sequence translated after deletion of the IS*3* element that interrupts it. The properties predicted are summarized in Table 1. For consistency, predictions for all products were made by using one set of computer programs. Table 1 provides the coordinates for each open reading frame, together with the number of amino acid residues, relative molecular weight (M_r), and isoelectric point of the polypeptide product predicted from that sequence by using the Genetics Computer Group Sequence Analysis Software Package program PEPTIDESORT (47). Except when noted in the table, a protein approximately the same size as the predicted product has also been identified experimentally. In cases when protein processing is known or thought to occur, numbers in brackets provide the corresponding data predicted for the mature product. The subcellular location is given, both as predicted by using computer algorithms based on DNA sequence information and as experimentally determined when such information is available. The experimental location is based on immunological (A) or radiological (L) detection in F^+ cells or analyses of maxicells or other hosts in which *tra* genes were expressed from multicopy vectors (X). Only the most recent data are presented; older, conflicting results are not cited, for the sake of clarity. Table 1 also contains comments for each protein, including a summary of relevant experimental data, references, and predictions about the protein structure generated by using the data base PROSITE (19) and PCGENE programs PSIGNAL and SURFACEPLOT (84, 249). The signal sequences of polypeptides predicted to undergo processing and modification are shown in Fig. 2 and compiled in Table 2. When possible, results confirming these predictions are noted. A summary of the predicted and experimentally determined locations of the entire array of *tra* products is shown in Fig. 3. With the exception of pilin, the *traA* gene product, which is discussed below, the predicted locations of *tra* products have been consistent with the experimentally determined locations.

Figure 4 aligns the start codons and compares putative ribosome-binding sites and translational start sequences for all of the *tra* region ORFs thought to be translated. Stop codons within these sequences are marked with asterisks, and coordinates and distances between stop and start codons are indicated. An examination of the start codons (Fig. 4) reveals that the AUG codon is preponderant; however, alternative codons are also used. The F and R100 *traY* coding sequences initiate at a UUG codon (116, 136), whereas other F-like plasmids use GUG to initiate *traY* (98, 134). Other F *tra* genes with a GUG start codon are *traC* (224), *trbJ* (156), *traG* (70), and, most probably, *traR* (51).

Although there are intervening sequences between a number of *tra* genes, nesting of sequences appears to be the rule. Sets of genes that exhibit a stop/start codon (or greater) overlap are *traE-traK-traB-traP-trbD*; *trbG-traV*; *traC-trbI-traW-traU*; *trbC-traN*; *trbE-traF*; *traQ-trbB-trbJ-trbF-traH-traG*; and *traD-trbH-traI*. Besides the sizeable intergenic regions between *traM* and *traJ*, *traJ* and *traY*, *artA* and *traQ*, *traS* and *traT*, and *traT* and *traD*, which accommodate promoter sequences, intergenic gaps are also present upstream from *traA*, *traR*, and *traC* for unknown reasons. A second translational initiation is thought to occur in the *traI* gene and result in the production of TraI* (25, 247), whereas TraG* is probably not produced by a translational restart but by proteolytic cleavage in the periplasm (see below) (70).

LEADING REGION OF TRANSFER

Although the F transfer region can be defined as the sequence between the origin of transfer (*oriT*) and the end of the fertility inhibition gene, *finO* (Fig. 1 and 2), genes bordering these sequences are highly conserved among all IncF plasmids studied and probably also have a role in plasmid transfer. However, this role is subtle, and mutations in most genes examined to date do not affect transfer between *E. coli* strains.

The portion of F defined as the leading region of transfer (coordinates 53.3F to 66.7F on the F map) is a highly conserved 13-kb segment of DNA immediately adjacent to *oriT* and extending into the primary replication-incompatibility region, RepFIA (36). The leading-region sequences are believed to be the first to enter the recipient cell during conjugation (for a model of the mechanism, see references 255 and 266). Transcriptional and translational studies have shown that this region is expressed and encodes at least eight polypeptides (36, 122, 214).

The functions of four gene products of the leading region have now been identified: *ssb* encodes a single-stranded DNA-binding protein with sequence homology to, and ability to complement mutations in, the *E. coli* chromosomal *ssb* product (29, 130); *psiB* encodes a product which is thought to inhibit the RecA protein coprotease activity and thereby bring about a demonstrated interference with the SOS response in the host cell (18, 54); *flm* is a complex region composed of three genes, *flmA*, *flmB*, and *flmC*, involved in the maintenance of unstable plasmids in a population (150); and *orf169* (formerly *geneX*) is the first gene transferred into the recipient cell and has been shown to be important for transfer (20) (see below).

Both *ssb* and *psiB* sequences are conserved among a range of transmissible plasmids (94, 95, 112) and are considered likely to play a role in establishing newly transferred plasmid DNA in the recipient cell. Significantly, both *ssb* and *psiB* on the IncI1 plasmid ColIb-P9 and *psiB* on F have been shown to be zygotically induced in the recipient cell after transfer (16, 125). The *flm* sequence, also previously referred to as *parL* (150, 215) or *stm* (96), which is largely identical to the *hok-sok-mok* (*parB*) region of plasmid R1 (92), is also conserved among a number of IncF conjugative plasmids (151). This plasmid maintenance system may also play a postconjugational role in ensuring the predominance of plasmid-carrying cells in the recipient population.

Sequences potentially involved in the synthesis of the complement to the transferred single strand of DNA in the recipient have been identified in the leading region. Analysis of available leading-region nucleotide sequences has revealed potential integration host factor (IHF) recognition sites and primosome assembly sites (148). Three of the latter are located within functional single-stranded initiation sequences, *ssiD* and *ssiE* (196); primers for recipient conjugal DNA synthesis could be initiated at these sites.

Among other highly conserved genes and ORFs in the leading region are *psiA*, which appears to be coexpressed with *psiB* but may not be involved in suppressing the SOS response

TABLE 1. Properties of F transfer region products

Gene	Function	Coordinates of reading frame	Amino acid length encoded [Mature][a]	Polypeptide M_r [Mature M_r][b]	pI[b]	Subcellular location			Comment[c]
						Predicted[c]	Experimental[c]	Basis[d]	
artA		16504–16193	104	12,131	8.42	IM			Encoded by anti-*tra* strand. Product not yet detected, but *artA*::*lacZ* fusion protein was expressed (269). Predicted to be an integral membrane protein with membrane-spanning segments at nt 46–71, 79–100, and possibly 21–40.
finO	Regulation	32738–33295	186	21,221	10.48	C			Not expressed from F owing to IS*3* insertion between nt 33133 and 33134 (31, 35, 274, 275); original protein of size indicated is predicted to have been hydrophilic. No membrane-associated regions predicted.
traA	F pilin subunit	2249–2611	121 [70]	12,768 [7,186]	[10.13]	OM	IM	A, L, X	Signal sequence cleaved between 51 and 52 (83). Maturation is *traQ* dependent; acetylation of the pilin N terminus (Ala-52) depends on *traX* (137, 158, 186, 191). Unassembled F pilin subunits are located in the IM of F^+ strains (186, 190, 200), and membrane F pilin is oriented such that both its N-terminal and C-terminal regions face the periplasm (200). Transmembrane segments are predicted to be at 77–95 and 101–121.
traB	F pilus assembly	4243–5667	475	50,460	5.18	IM			Predicted to be a predominantly hydrophilic protein with a probable N-terminal membrane anchor at residues 13–33 (74, 192).
traC	F pilus assembly	7725–10349	875	99,191	5.90	C (M)	IM	A	A cytoplasmic protein but membrane associated in F *lac* cells (224, 225). There is an ATP-binding motif at 486–493; aa 664–680 and 293–309 are hydrophobic and may be involved in peripheral membrane or internal associations.
traD	DNA transport	23779–25929	717	81,683	5.27	IM	IM	X	Purified TraD binds to blue dextran, indicating nonspecific DNA-binding activity (202). Its sequence (124, 274) is predominantly hydrophilic. There is a strong prediction for N-terminal membrane segments at 27–47 and 113–133, and an additional hydrophobic segment is at 392–412. There are ATP-binding motifs at 192–199 and 421–426 and an ATP synthase-like sequence at 628–637. A subtilisn-like serine proteinase motif (predicted to have a histine-active site at residue 191) is also present.
traE	F pilus assembly	2962–3525	188	21,161	10.06	IM	IM	X	Predicted to be predominantly hydrophilic with an N-terminal membrane anchor at residues 12–33 (6, 83, 119, 137).
traF	F pilus assembly	15082–15822	247 [228]	27,973 [25,939]	[9.49]	P/OM	P	X	Precursor and product observed; mature protein released from maxicells by osmotic shock and predicted to be hydrophilic (270).
traG	F pilus assembly and aggregate stability	19403–22216	938 [487]	102,471 [52,719]	6.30 [5.30]	IM P	IM P	A A	Only the N-terminal sequences of TraG are required for pilus assembly, but TraG C-terminal sequences are required for aggregate stabili-

Continued on following page

TABLE 1—*Continued*

Gene	Function	Coordinates of reading frame	Amino acid length encoded [Mature][a]	Polypeptide M_r [Mature M_r][b]	pI[b]	Subcellular location			Comment[e]
						Predicted[c]	Experimental[c]	Basis[d]	
									zation (70, 167). Product localization with antibodies indicates that TraG may be cleaved (suggested site is between residues 451 and 452); both a 100-kDa IM product and a C-terminal 50-kDa periplasmic polypeptide (TraG*) have been detected (70). Potential transmembrane segments occur at ca. 54–73, 327–348, 364–385, 407–426, and 428–447; protease accessibility studies with maxicells suggest that large segments of TraG are located in the periplasm (70).
traH	F pilus assembly	18030–19403	458 [434–435]	50,244 [47,781–47,851]	[7.22]	P/OM			Mature protein predicted to be predominantly hydrophilic. ATP-binding motif at 193–200 (70, 105, 167).
traI	*oriT* nicking unwinding	26648–31915	1756	192,014	5.94	C	C	A	Helicase I (1). Nicks *oriT*, becomes attached to the 5′ end, and unwinds in the 5′-to-3′ direction (172, 173, 218, 247). No predicted membrane-associated regions; ATP-binding motifs are at residues 177–184 and 992–998 (25, 40).
*traI**		29510–31915	802	87,882	5.41	C	C	A	Same polypeptide sequence as TraI C terminus (25, 246). ATP-binding motif at residues 48–55.
traJ	Regulation	1033–1719	229	27,061	6.78	C	C	A	Insoluble aggregates expressed from multicopy vectors, but cytoplasmic in F^+ cells (38, 39, 243). No membrane-associated regions are predicted.
traK	F pilus assembly	3515–4240	242 [221]	25,627 [23,307]	[5.38]	P/OM			Mature protein predicted to be predominantly hydrophilic (74).
traL	F pilus assembly	2629–2901	91	10,379	10.64	C (M)			Considerable hydrophobic character and weak prediction for membrane association of residues 31–61. Possibly a peripheral membrane protein (6, 83).
traM	Mating signal	463–843	127	14,507	5.17	C	C		Protein has been purified from cytoplasmic fractions. Binds to multiple sites in the DNA sequence between *oriT* and *traM* (2, 3, 49, 50, 227, 243). No predicted DNA-binding domain or membrane-associated regions.
traN	Aggregate stability	12997–14802	602 [584]	65,714 [63,823]	[7.57]	P/OM	OM	X	After processing, mature protein expressed in maxicells is in OM fractions and partly exposed to protease K digestion of intact cells (161). Only short hydrophobic regions are predicted; ATP-binding motif at 84–91.
traP		5660–6247	196	21,961	9.37	IM			Transmembrane segments predicted at 25–49 and 118–138 (74, 192).
traQ	F pilin synthesis	16585–16866	94	10,866	9.38	IM	IM	X	Required for efficient *traA* product utilization (154, 158, 269, 271). Two probable transmembrane segments at 17–34 and 37–57 might be stabilized by an interhelix ion pair between Arg-31 and Glu-45.
traR		7344–7562	73	8,313	7.20	C			Polypeptide encoded would be hydrophilic; no membrane-associated regions predicted (51, 156, 192).
traS	Surface exclusion	22252–22698	149	16,860	10.37	IM	IM	X	IM location when expressed in minicells (6). Transmembrane segments

Continued on following page

TABLE 1—*Continued*

Gene	Function	Coordinates of reading frame	Amino acid length encoded [Mature][a]	Polypeptide M_r [Mature M_r][b]	pI[b]	Subcellular location			Comment[e]
						Predicted[c]	Experimental[c]	Basis[d]	
									predicted at 34–53, 61–79 and 108–143. The latter long segment could double back into a hairpin at the Gly-Gly-Gly sequence at 124–126 (123).
traT	Surface exclusion	22796–23527	244 [233]	26,018 [23,795]	9.04	OM	OM	A, L	A major OM protein in F^+ cells (5, 183). Lipoprotein processing and N-terminal sequence confirmed (181, 185, 209). Hydrophobic regions occur at 23–40 and 121–131 (123).
traU	F pilus assembly	11361–12350	330 [308]	36,786 [34,275]	[7.68]	P/OM	P	X	*traU* mutants do express some F pili. Both precursor and mature products have been observed; the latter was released from maxicells by osmotic shock (188). In addition to the signal sequence, regions 37–50, 99–116, 139–148, and 184–209 are hydrophobic, so there may be peripheral membrane associations (188).
traV	F pilus assembly	6694–7206	171 [153]	18,588 [16,564]	[9.38]	OM			Two products, approximately 17 and 21 kDa, have been observed (51, 192). Sequence suggests lipoprotein processing and OM location.
traW	F pilus assembly	10732–11361	210 [192–193]	23,630 [21,690–21,761]	[8.39]	P/OM	P	X	Precursor and product observed; mature protein found in periplasmic fractions (162, 164). Signal sequence cleavage predicted between 17/19 or 18/19.
traX	F pilin acetylation	31938–32681	248	27,535	8.93	IM	IM	X	Predicted to be an integral membrane protein with multiple transmembrane segments and a highly hydrophobic overall character (35, 155, 186).
traY	*oriT* nicking	1821–2213	131	15,183	10.19		C		Required in a lambda-F *oriT* nicking assay (57) but not in an in vitro reaction (172, 218). Protein has been purified from cytoplasmic fraction, and DNA-binding sites have been analyzed (115, 116, 136, 195). No membrane-associated regions predicted (72, 84).
trbA		15841–16185	115	12,944	8.36	IM	IM	X	Protein found in inner membrane fractions (127, 269, 271). Predicted to be highly hydrophobic in character; segments with membrane-spanning potential at 5–25, 39–59, 62–80, and 90–109. Leucine zipper motif at 35–56 of unknown significance.
trbB		16856–17392	179 [159]	19,505 [17,445]	[8.10]	P/OM	P	X	Precursor and product observed; mature protein is released by osmotic shock from maxicells (127, 269, 271) and predicted to be quite hydrophilic.
trbC	F pilus assembly	12362–12997	212 [191]	23,433 [21,226]	[8.12]	P/OM	P	X	Precursor and product observed; mature protein is released from maxicells by osmotic shock (163), but short hydrophobic regions could indicate peripheral membrane associations.
trbD		6237–6431	65	7,053	5.73	C			Product not yet detected; sequence suggests a hydrophilic protein; no membrane-associated regions predicted (74).

Continued on following page

TABLE 1—*Continued*

Gene	Function	Coordinates of reading frame	Amino acid length encoded [Mature][a]	Polypeptide M_r [Mature M_r][b]	pI[b]	Subcellular location			Comment[e]
						Predicted[c]	Experimental[c]	Basis[d]	
trbE		14829–15086	86	9,910	4.39	IM	IM	X	Product analysis indicates an intrinsic inner membrane protein (161). Transmembrane segments predicted at 11–33 and 44–68.
trbF		17653–18030	126	14,480	8.40	IM			Product not yet detected. Predicted to be an integral membrane protein with two transmembrane segments at 29–51 and 69–97 (105, 146).
trbG		6446–6694	83	9,126	6.48	C			Product not yet detected; no membrane-associated regions predicted (51).
trbH		25932–26648	239	26,259	8.95	IM			Product not yet detected; C-terminal membrane anchor predicted at 209–228; several other short or weakly hydrophobic segments are present (25, 156).
trbI		10349–10732	128	14,132	10.12	IM	IM	X	Product analysis in maxicells indicates an intrinsic IM protein (164); sequence predicts N-terminal anchor at 18–40.
trbJ		17385–17663	93	10,160	6.80	IM			Predicted to be an inner membrane protein with membrane-spanning regions at 6–29 and 42–67 (156, 269).

[a] Except as indicated, values given in brackets for amino acid length and other properties of processed proteins are based on predicted signal sequence cleavage sites.
[b] M_r and pI values were obtained with the Genetics Computer Group program PEPTIDESORT (47). Values in brackets are derived from the predicted sequence of the unmodified, processed polypeptide.
[c] C, cytoplasm; IM, inner membrane; P, periplasm; OM, outer membrane. (M) indicates peripheral membrane association.
[d] A, location of product in F^+ cells detected with antibody; L, location of labeled protein in F^+ cells; X, location of labeled product expressed from multicopy plasmids.
[e] Except as noted, protein products consistent in size with the predicted products have been identified. Predictions concerning protein character were based on results obtained with PROSITE and the PCGENE program SURFACEPLOT (249). Numbers refer to amino acid residues in the product encoded.

in the recipient (17, 18, 149), and ORFs, ORF169 and ORF273 (59, 97, 148, 151). *orf169*, also designated *geneX*, is immediately adjacent to *oriT* and crosses the *Bgl*II site at position 1 in Fig. 2. It begins at nt 49 and is transcribed in the opposite direction to the transfer region. The beginning of the *orf169* transcript near nt 74 and the first *traM* transcript (M mRNAIII, nt 235) could result from expression from divergent promoters flanking *oriT* (207). Expression of ORF169 is affected by IHF in the R100 plasmid system (46), and recent evidence demonstrates that expression of the homolog in the R1 transfer region, *gene19*, is affected by RNase III cleavage (133). *gene19* can be considered a transfer gene since mutations in its coding sequence result in drastically reduced levels of transfer and interfere in the release of the male-specific phage R17 (20). ORF273 is transcribed in the opposite orientation to ORF169 and in the same direction as the majority of genes in the transfer region (148). The ubiquity of these ORFs and other leading-region sequences among transmissible plasmids suggests an importance for the whole region in the evolution of conjugative plasmids. This topic remains the subject for another review once the entire nucleotide sequence and results of further functional analysis are available for the leading region.

NICKING AND INITIATION OF TRANSFER AT THE ORIGIN OF TRANSFER

The biochemistry of the "nicking" reaction has indicated that the TraI protein, which nicks the DNA, is a transferase that catalyzes the covalent attachment of the 5′ end of the DNA to the protein. Thus this reaction is not a true nick which results in free 5′ and 3′ ends; however, for the sake of simplicity and for historical reasons, we have continued to use this term. The initiation of DNA transfer is thought to involve recognition of a signal, presumed to be generated when mating-pair formation has been successfully completed, by the DNA transfer machinery, which is known as a relaxosome or "transferosome." Subsequent steps include the completion of the nicking reaction, unwinding of the DNA, and transport of the nicked strand into the recipient cell. The timing of the nicking reaction has not been established; an equilibrium between nicking and religation may occur in vivo, or nicking may occur in response to a signal generated during mating-pair formation. The equilibrium model is currently favored since inefficient nicking has been shown to occur in vitro in the absence of the *tra* gene products required for mating-pair formation (173, 218). The *tra* proteins involved in DNA metabolism during transfer are TraM, TraY, TraI, and TraD (57). It has been suggested that TraM may process the signal that a mating-pair or aggregate has formed (266) and that TraY may direct the nicking enzyme, TraI, to *oriT* (247). TraI, which is DNA helicase I (1), apparently has two functions, nicking and unwinding (218), whereas TraD is involved in pumping the DNA into the recipient cell (201). A discussion of the processing of DNA at *oriT* during plasmid transfer is provided by Wilkins and Lanka (256) and includes a comparison of the F system with those specified by other conjugative plasmids in gram-negative bacteria.

TABLE 2. Processing sites of F *tra* region products, including the processing sites for leader peptidases I and II

Gene	Signal length[a]	Signal sequence	Mature N terminus[b]	Comments
F pilin precursor				
traA	51	MNAVLSVQGASAPVKKKSFFSKFTRLN- MLRLARAVIPAAVLMMFFPQLAMA	AGSSGQDLMA-	Mature protein is in inner membrane if assembly proteins are absent (186, 190). Acetylated N terminus of mature F pilin confirmed (83).
Periplasmic proteins				
traF	19	MNKALLPLLLCCFIFPASG	KDAGWQWYNE-	Mature protein released by osmotic shock from maxicells carrying TraF$^+$ plasmids (270).
traG	(451)[c]	. . . MMVKGLGAGFSSVYSHFASSSISPTASA	AGSVVDGNYS-	Protein TraG* has been detected in periplasmic fractions; suggested TraG* N terminus corresponds to amino acid 452 in TraG (70).
traH	23	MMPRIKPLLVLCAALLTVTPAAS	ADVNSDMNQF-	Subcellular location predicted (105).
	24	or MMPRIKPLLVLCAALLTVTPAASA	DVNSDMNQFF-	
traK	21	MRKNNTAIIFGSLFFSCSVMA	ANGTLAPTVV-	Subcellular location predicted (74).
traU	22	MKRRLWLLMLFLFAGHVPAASA	DSACEGRFVN-	Mature protein released by osmotic shock from maxicells carrying TraU$^+$ plasmids (188).
traW	17	MRCRGLIALLIWGQSVA	AADLGTWGDL-	Mature protein released by osmotic shock from maxicells carrying TraW$^+$ plasmids (164).
	18	or MRCRGLIALLIWGQSVAA	ADLGTWGDLW-	
trbB	20	MSLTKSLLFTLLLSAAAVQA	STRDEIERLW-	Mature protein released by osmotic shock from maxicells carrying TrbB$^+$ plasmids (271).
trbC	21	MKLSMKSLAALLMMLNGAVMA	SENVNTPENR-	Mature protein released by osmotic shock from maxicells carrying TrbC$^+$ plasmids (163).
Outer membrane proteins				
traN	18	MKRILPLILALVAGMAQA	DSNSDYRAGS-	Mature protein is found in outer membrane fractions of maxicells carrying TraN$^+$ plasmids and partially degraded by protease treatment of whole cells (161).
traT	21 (LP)[d]	MMKTKKLMMVALVSSTLALSG	CGAMSTAIKK-	Location, lipoprotein peptidase processing, and N terminus have been confirmed (181, 185).
traV	18 (LP)[e]	MKQTSFFIPLLGTLLLYG	CAGTSTEFEC-	Tentative location, suggested by possible lipoprotein peptidase site (51).

[a] As suggested by the PCGENE program PSIGNAL (249). LP indicates a lipoprotein peptidase site not identified by this program.
[b] Except as noted, the N-terminal sequence is based on a PSIGNAL prediction.
[c] The amino-terminal segment of TraG includes inner membrane domains that are essential for expression of F pili; only the residues immediately preceding the putative processing site are shown.
[d] This signal sequence was used in deriving the PROSITE consensus sequence for lipoprotein peptidase sites.
[e] Processing was not predicted by PSIGNAL; PROSITE did not identify this lipoprotein-processing site unless a one-residue mismatch was allowed.

oriT Region

The *oriT* region is arbitrarily defined as the region between the *Bgl*II site at nt 1 (Fig. 2) and the beginning of the *traM* gene (nt 463). This region was first sequenced by Thompson et al. (244) and contains the site where nicking occurs and transfer of single-stranded DNA is initiated, in a 5′-to-3′ manner, into the recipient cell (113). The lower strand of the sequence presented in Fig. 2 is nicked, and the first sequence transferred is that for ORF169, which is part of the leading region as described above.

Deletion analysis of cloned F *oriT* sequences has revealed that this region can be divided into two sections. One is essential for nicking, and the second is essential for transfer, provided that nicking has occurred (85). Clones that contained the sequence between nt 118 and 215 were nicked in an in vivo nicking assay; however, normal mobilization of these *oriT* clones required that the sequence be extended to nt 253. This suggests that the IHF-binding site (site A) and part of the TraY-binding site (sbyA) are required for nicking, while sbyA and half of the inverted repeat which defines the third, weakest TraM-binding site, sbmC, are required for transfer (Fig. 2, nt 118 to 253). Deletion of sequences from the *Bgl*II site toward *oriT* resulted in the formation of plasmid multimers in the recipient cell, suggesting a defect in the resolution of the plasmid during transfer termination (85). Recently, evidence has been presented that initiation and termination can be separated into distinct functions on the basis of a deletion analysis of the region immediately downstream from *oriT* (89). Whereas initiation of transfer requires IHF site A and sbyA, efficient termination requires the presence of IHF site A only. Thus termination has fewer requirements than initiation and involves the inverted repeats immediately upstream of *oriT* and the sequences downstream of *oriT* up to and including IHF site A. These results may be related to the finding that recombination between F *lac* and a λ*plac5* is enhanced and depends on

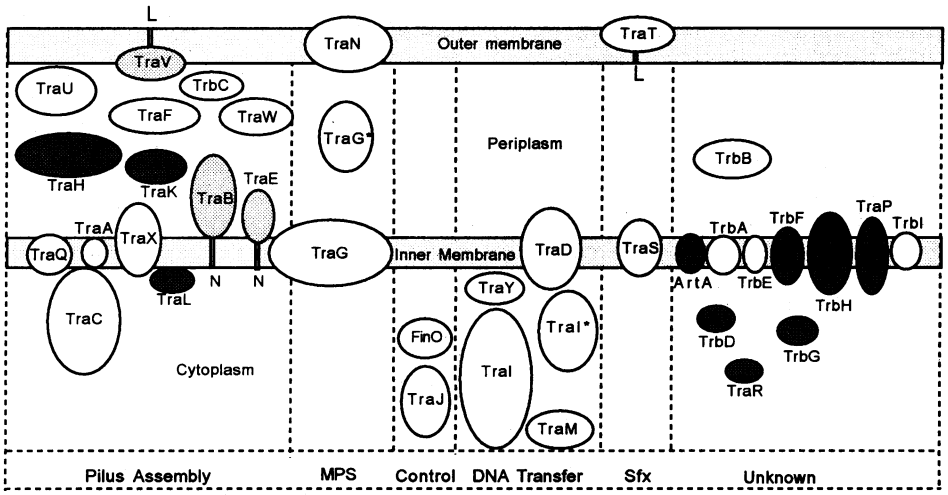


FIG. 3. Diagram summarizing the information on the *tra* gene products with respect to their location in the bacterial cell, taken from Table 1. The *tra* gene products are separated according to function: pilus assembly, MPS (mating-pair stabilization), control (regulation of *tra* gene expression), DNA transfer, Sfx (surface exclusion), and unknown. TraD may also have a role in pilus assembly (not shown; see the text). TraG, which has two functions, is shown overlapping the boxes for MPS and pilus assembly. The deeply shaded *tra* products are positioned according to computer predictions. The more lightly shaded products representing TraB, TraE, and TraV are oriented arbitrarily in the membrane with the bulk of the protein exposed in the periplasm. Thin rectangles indicate single membrane-spanning regions, and L and N indicate lipoprotein and N-terminal anchors, respectively.

the presence of *oriT* in the plasmid as well as functional TraY and the N-terminal portion of TraI (28).

The site of nicking was determined by Thompson et al. (245), using a strand interruption assay, to be after position 140 (Fig. 2 and 5) on the transferred (bottom) strand and has been confirmed in vitro by Matson et al. (172, 173). Reygers et al. (218) have reported the nick site to occur after nt 141, but their data also showed prominent nicking between nt 140 and 141. The nick site in R100 has also been shown to correspond to position 140 (117). Examination of the sequences of the first four alleles (I to IV) of *oriT* (259) (Fig. 5) reveals almost complete conservation of sequence between nt 1 and 173, suggesting that the allelic specificity of *oriT* is conferred by the interaction of transfer proteins with DNA sequences bordering this region. Two mutations in *oriT* have been characterized and sequenced (nt 144 and 149) and were found to be immediately downstream of the nick site (244). This is the same region that contains a 2-base difference in sequence between F and R100 (nt 146 and 148), which may define the allelic specificity of TraI (Fig. 5). Recently, sequences of the *oriT-traM-finP* regions of three other F-like plasmids were reported: P307 (98), pSU223 (222), and pSU316 (152), which is very closely related to F (see Fig. 5 for a comparison of *oriT*). These new examples confirm that the arrangement of sequences in *oriT* is important and is strongly conserved.

The *oriT* region is subject to bending both as a consequence of its sequence and through the action of proteins such as the chromosomally encoded IHF and, perhaps, various *tra* proteins such as TraY (see below). An intrinsic bend, which is usually caused by runs of homopolymeric dA · dT bases repeated in phase along the DNA sequence (37), is centered at position 245 and is a composite of two bends near positions 171 and 267 (248). IHF causes sharp bends (140°) in the DNA (273) and binds strongly to a site (site A) in *oriT* extending from bp 160 to 195 and with less affinity to a second site (site B) extending from nt 290 to 320 (248). IHF also affects the efficiency of F plasmid transfer by affecting the transcription of F transfer functions (88). It binds to equivalent sequences in R100 (114, 175) and affects the initiation and length of the *orf169*, *traM*, and *traJ* transcripts (44, 46). pED208 *oriT* has four IHF-binding sites, two of which correspond to those in F *oriT* (see Fig. 5) (50). The remaining two sites in pED208 were identified in the presence of potassium glutamate, which enhanced DNA binding (50). The presence of binding sites for IHF at corresponding positions in F has not been ruled out.

A comparison of the eight available *oriT* sequences from plasmids of the IncF complex reveals precisely defined regions which can be assigned different functions (Fig. 5). The first 109 bp contains the promoter and translational start for *orf169* (also called *geneX* or *gene19*), which is transcribed in the opposite orientation to most of the genes in the transfer region. This sequence is highly conserved among the different *oriT* sequences of the F-like plasmids (Fig. 5). The region defined by nt 110 to 177 contains the nick site in F (between nt 141 and 142) within a G+T-rich sequence on the bottom strand (Fig. 2 and 5). Again, this region is almost identical among the first seven alleles of *oriT* and is highly conserved in the eighth allele for pED208. The sequence between nt 114 and 145 contains the sequence G/TTG(G)TGGTG/T three times, once on the top strand between nt 114 and 122 and twice on the bottom strand (nt 133 to 125 and nt 144 to 137), where the second sequence contains the nick site (244). The two overlapping inverted repeats containing this sequence (nt 114 to 132 and nt 112 to 145 [Fig. 2]) may be important for the initiation and termination of transfer; interestingly, the inverted repeat at nt 114 to 132 is the first sequence transferred into the recipient cell and has been implicated in transfer termination (85). A third, small inverted repeat between nt 150 and 167 is conserved within the first seven alleles of *oriT* but is not present in pED208 (Fig. 5).

The sequences corresponding to the F segment nt 161 to 237 (which include an intrinsic bend at position 168) are all A+T rich. Since the group II and III alleles are almost identical in this region and express the same *traY* gene product, the *oriT* binding site for TraY was suggested to be in this A+T-rich segment (61). This prediction has been confirmed for the F

TraY protein, which binds between nt 205 and 241 (sbyA), containing half of an imperfect inverted repeat (nt 180 to 224 [Fig. 2 and 5]) (136, 195). Similarly, a binding site for the R100 TraY protein has been found within this region; the sequence recognized is TAA(A/T)T, which resembles the ATAAA sequence found in F sbyA (nt 220 to 224, bottom strand) (115, 195). Recently, TraY has been shown to bend the DNA at an angle of 50 to 55°, and this bend is centered at nt 218 (153) (see the section on the TraY protein, below).

Immediately after position 239, the sequences for the different plasmids diverge (Fig. 5), but they characteristically include a complex series of direct and inverted repeats specific for each allele. F contains an inverted repeat (nt 234 to 278) and two imperfect direct repeats (nt 324 to 345 and nt 370 to 391), each containing the sequence CGGC/TGCG (Fig. 2) (61, 244). These repeats define the binding sites for the F TraM protein and have been named sbmA, sbmB, and sbmC (Fig. 2 and 6) (49).

Besides the promoter identified by runoff transcription in vitro (243), two other promoters for *traM* have been identified by using primer extension (207). Two of these promoters (M mRNAI and M mRNAII [Fig. 2]) initiate at identical positions within the direct repeats that define the sbmA and sbmB binding sites (compare the sequence surrounding nt 326 to 371 [Fig. 2]). Primer extension analysis and immunoblots of TraM protein suggest that TraM regulates its own expression from both these promoters. A third promoter for *traM* initiates within the sbyA binding site for TraY near *oriT* (M mRNAIII; nt 235) and has increased activity in *traY* and *traJ* mutants (202) (see the section on the TraM protein, below). Recently, autoregulation has been reported for the *traM* gene of the R1 plasmid by using a single-copy *lacZ* reporter gene linked to the *traM* promoter region and incorporated into the chromosome by lysogenization (228). By using similar techniques, two promoters have been mapped for the *traM* gene of R100, one of which is constitutive while the other is autoregulated by R100 TraM and is partially repressed by IHF (Fig. 5) (3). The positions of these promoters differ somewhat from those reported previously for R100, one of which is within the *traM* coding region (45).

The TraM protein of pED208 binds to four sites in the *oriT* region, which contain 15 of the 16 short 5-bp repeats (GANTC) in phase along the DNA (Fig. 5) (48). Similarly, the group III *oriT* segment of the R1 plasmid, which includes two large binding sites for the R1 TraM protein, also contains *Hin*fI-like sequences, although they are not as frequent or equally spaced as in pED208 (Fig. 5) (227). The *oriT* region of R100 contains four binding sites for the R100 TraM protein (2) with a consensus binding sequence reminiscent of F (compare TAGGGtCGctaCTa/gG [R100] with TAGGGGCGCTGCT AG [F; nt 370 to 384]). Interestingly, the *oriT* region of ColE1 contains an F TraM-binding site (49), although TraM is not required for ColE1 mobilization (258). This suggests that F TraM may have a subtle role in ColE1 mobilization not detected in the mobilization assay or that one of the ColE1 *mob* genes is an analog of F TraM, despite the absence of apparent homology (24, 74).

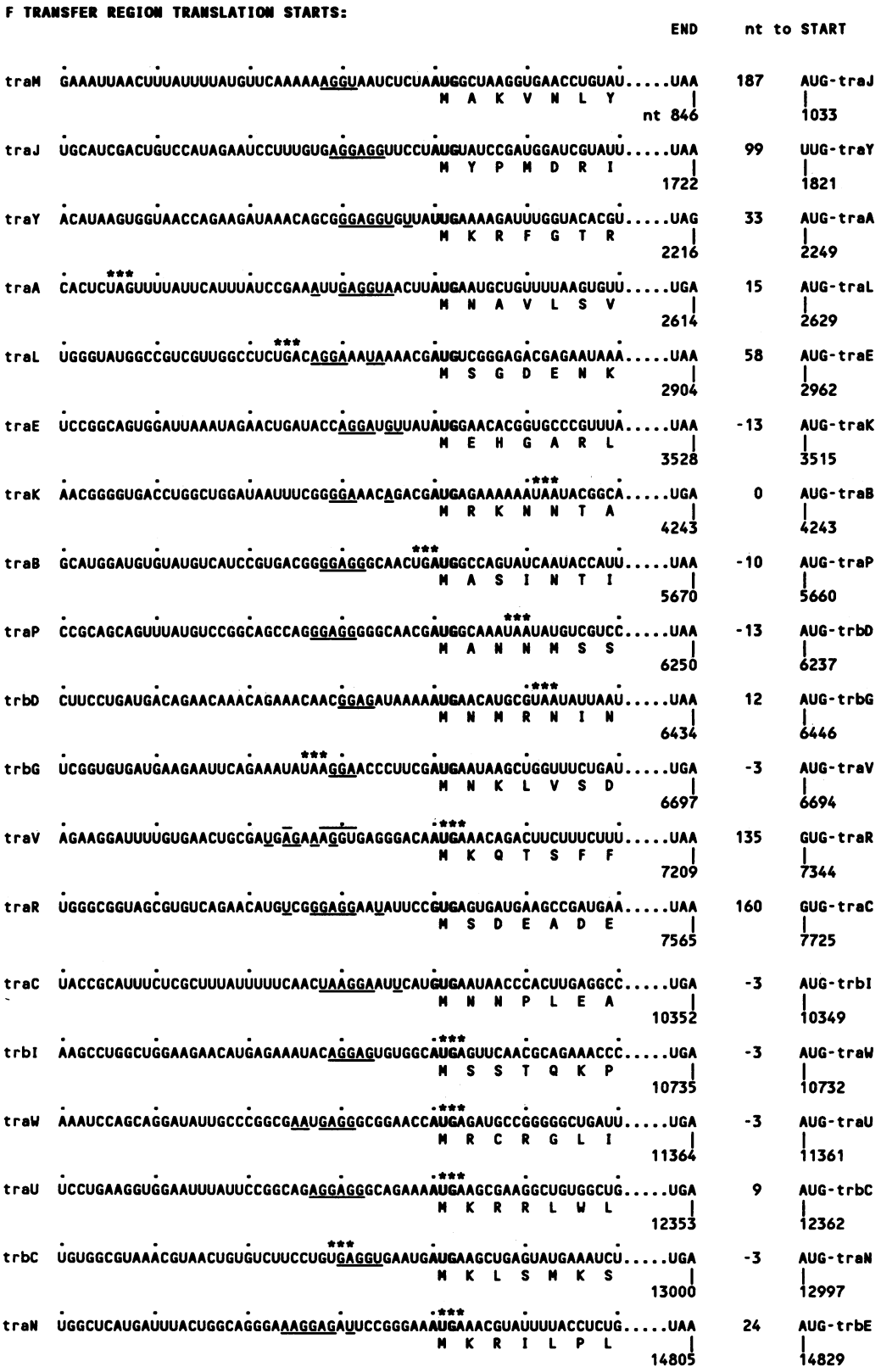
A primosome (n′)-binding site has been proposed for the sequence at positions 236 to 241 (AAGCGG), which lies within the inverted repeat that forms sbmC and coincides with the initiation of transcription from the third *traM* promoter (nt 235, M mRNAIII [Fig. 2 and 5]) (267). Unlike the IncI plasmid ColIb-P9 (30, 216) or the IncP plasmid RP4 (138, 217), the F plasmid does not encode a plasmid-specific primase and relies on the host primase to initiate DNA replication in the recipient cell. Potential *ssi* (single-strand initiation) sites in the leading region have been mapped, as have a number of *ssi* sites throughout the plasmid (171, 196). The presence of one site in the F-f5 fragment (fifth largest *Eco*RI fragment of F) has been shown to allow conversion of single-stranded DNA to a replicative form in vitro by using a mechanism similar to that for φX174. Thus the *ssi* sites may be used in DNA synthesis in the recipient cell, whereas the primosome assembly site and possibly the *traM* promoter under TraY control could be used in the initiation of transfer. Recent evidence suggests that a noncovalent association between the TraI protein and the 3′-OH generated by nicking at *oriT* may prevent rolling-circle replication from this end; in an in vitro system, treatment with sodium dodecyl sulfate (SDS) and proteinase K was required to obtain extension with DNA polymerase I (173).

TraY Protein

The F *traY* gene (nt 1821 to 2213) (72) was initially thought to encode a 13.8-kDa protein. However, the gene was later shown to initiate with the rare codon TTG and encode a protein of 15.2 kDa, which more closely resembled its apparent molecular mass on SDS-polyacrylamide gels (116, 136). The F *traY* gene appears to be an example of gene duplication in which the N-terminal and C-terminal halves of the protein show homology to each other and to the smaller products of other *traY* alleles (116) (Fig. 6).

The F TraY protein was shown to bind to the upper strand of the *oriT* region between nt 205 and 241 by band retardation assays and DNA footprinting techniques (Fig. 2 and 5) (195); the R100 TraY protein binds to the corresponding region of the group IV *oriT* allele (115). In both cases, binding of TraY has also been demonstrated near the transfer operon promoter preceding *traY* (P_Y) and has been named sbyB in F (nt 1794 to 1820) (115, 195). Recently, a third TraY-binding site, sbyC, overlapping the IHF-binding site, site A (nt 180 to 199 [Fig. 2]), has been proposed (153). In the absence of IHF, binding of two additional monomers of TraY to sbyC can be demonstrated in vitro at high concentrations of TraY (153). Comparison of the three TraY-binding sites, sbyA, sbyB, and sbyC, has revealed that the most highly conserved sequence is CtCT NTTTAT, which may be important in TraY recognition (153). In vitro analyses further indicate that each arm of the imperfect inverted repeat (nt 180 to 224 [Fig. 2]) includes the most critical sequence required for TraY recognition of the two adjacent binding sites, sbyC and sbyA (153). The precise boundaries of the sbyC binding site and the in vivo conditions under which TraY rather than IHF might bind to sbyC remain to be characterized.

TraY has been predicted to be related to the P22 phage repressors Arc and Mnt (23). These repressors act as tetramers with their N-terminal domains in an extended conformation, establishing contact with their operator sequences. Since the primary differences among the four TraY alleles occur within the first 24 residues, this domain could be involved in TraY binding to the DNA. Recently, the structure of the *met* repressor, which is homologous to the Mnt and Arc repressors, has been solved (239). This protein uses a "ribbon-helix-helix" motif to form contacts with the DNA where the two dimers of the *met* repressor bind short sequences 8 bp apart. The authors suggested that F TraY probably shares this protein conformation and that one TraY protein could serve as a dimer (239). This proposal has recently been confirmed since sbyA binds two TraY monomers (153). Interestingly, there is a short repeated sequence spaced 8 bp apart in sbyA (TCTCTTA CTCTCTT; nt 208 to 221) which is similar to the spacing in the recognition sequence for the *met* repressor (239). This se-


F TRANSFER REGION TRANSLATION STARTS:
END
nt to START
traM GAAAUUAACUUUAUUUUAUGUUCAAAAAAGGUAAUCUCUAAUGGCUAAGGUGAACCUGUAU.....UAA 187 AUG-traJ
M A K V N L Y
nt 846
1033
traJ UGCAUCGACUGUCCAUAGAAUCCUUUGUGAGGAGGUUCCUAUGUAUCCGAUGGAUCGUAUU.....UAA 99 UUG-traY
M Y P M D R I
1722
1821
traY ACAUAAGUGGUAACCAGAAGAUAAACAGCGGGAGGUGUUAUUGAAAAGAUUUGGUACACGU.....UAG 33 AUG-traA
M K R F G T R
2216
2249
traA CACUCUAGUUUUAUUCAUUUAUCCGAAAUUGAGGUAACUUAUGAAUGCUGUUUUAAGUGUU.....UGA 15 AUG-traL
M N A V L S V
2614
2629
traL UGGGUAUGGCCGUCGUUGGCCUCUGACAGGAAAUAAAACGAUGUCGGGAGACGAGAAUAAA.....UAA 58 AUG-traE
M S G D E N K
2904
2962
traE UCCGGCAGUGGAUUAAAUAGAACUGAUACCAGGAUGUUAUAUGGAACACGGUGCCCGUUUA.....UAA -13 AUG-traK
M E H G A R L
3528
3515
traK AACGGGGUGACCUGGCUGGAUAAUUUCGGGGAAACAGACGAUGAGAAAAAAUAAUACGGCA.....UGA 0 AUG-traB
M R K N N T A
4243
4243
traB GCAUGGAUGUGUAUGUCAUCCGUGACGGGGAGGGCAACUGAUGGCCAGUAUCAAUACCAUU.....UAA -10 AUG-traP
M A S I N T I
5670
5660
traP CCGCAGCAGUUUAUGUCCGGCAGCCAGGGAGGGGGCAACGAUGGCAAAUAAUAUGUCGUCC.....UAA -13 AUG-trbD
M A N N M S S
6250
6237
trbD CUUCCUGAUGACAGAACAAACAGAAACAACGGAGAUAAAAAUGAACAUGCGUAAUAUUAAU.....UAA 12 AUG-trbG
M N M R N I N
6434
6446
trbG UCGGUGUGAUGAAGAAUUCAGAAAUAUAAGGAACCCUUCGAUGAAUAAGCUGGUUUCUGAU.....UGA -3 AUG-traV
M N K L V S D
6697
6694
traV AGAAGGAUUUUGUGAACUGCGAUGAGAAAGGUGAGGGACAAUGAAACAGACUUCUUUCUUU.....UAA 135 GUG-traR
M K Q T S F F
7209
7344
traR UGGGCGGUAGCGUGUCAGAACAUGUCGGGAGGAAUAUUCCGUGAGUGAUGAAGCCGAUGAA.....UAA 160 GUG-traC
M S D E A D E
7565
7725
traC UACCGCAUUUCUCGCUUUAUUUUUCAACUAAGGAAUUCAUGUGAAUAACCCACUUGAGGCC.....UGA -3 AUG-trbI
M N N P L E A
10352
10349
trbI AAGCCUGGCUGGAAGAACAUGAGAAAUACAGGAGUGUGGCAUGAGUUCAACGCAGAAACCC.....UGA -3 AUG-traW
M S S T Q K P
10735
10732
traW AAAUCCAGCAGGAUAUUGCCCGGCGAAUGAGGGCGGAACCAUGAGAUGCCGGGGGCUGAUU.....UGA -3 AUG-traU
M R C R G L I
11364
11361
traU UCCUGAAGGUGGAAUUUAUUCCGGCAGAGGAGGGCAGAAAAUGAAGCGAAGGCUGUGGCUG.....UGA 9 AUG-trbC
M K R R L W L
12353
12362
trbC UGUGGCGUAAACGUAACUGUGUCUUCCUGUGAGGUGAAUGAUGAAGCUGAGUAUGAAAUCU.....UGA -3 AUG-traN
M K L S M K S
13000
12997
traN UGGCUCAUGAUUUACUGGCAGGGAAAGGAGAUUCCGGGAAAUGAAACGUAUUUUACCUCUG.....UAA 24 AUG-trbE
M K R I L P L
14805
14829

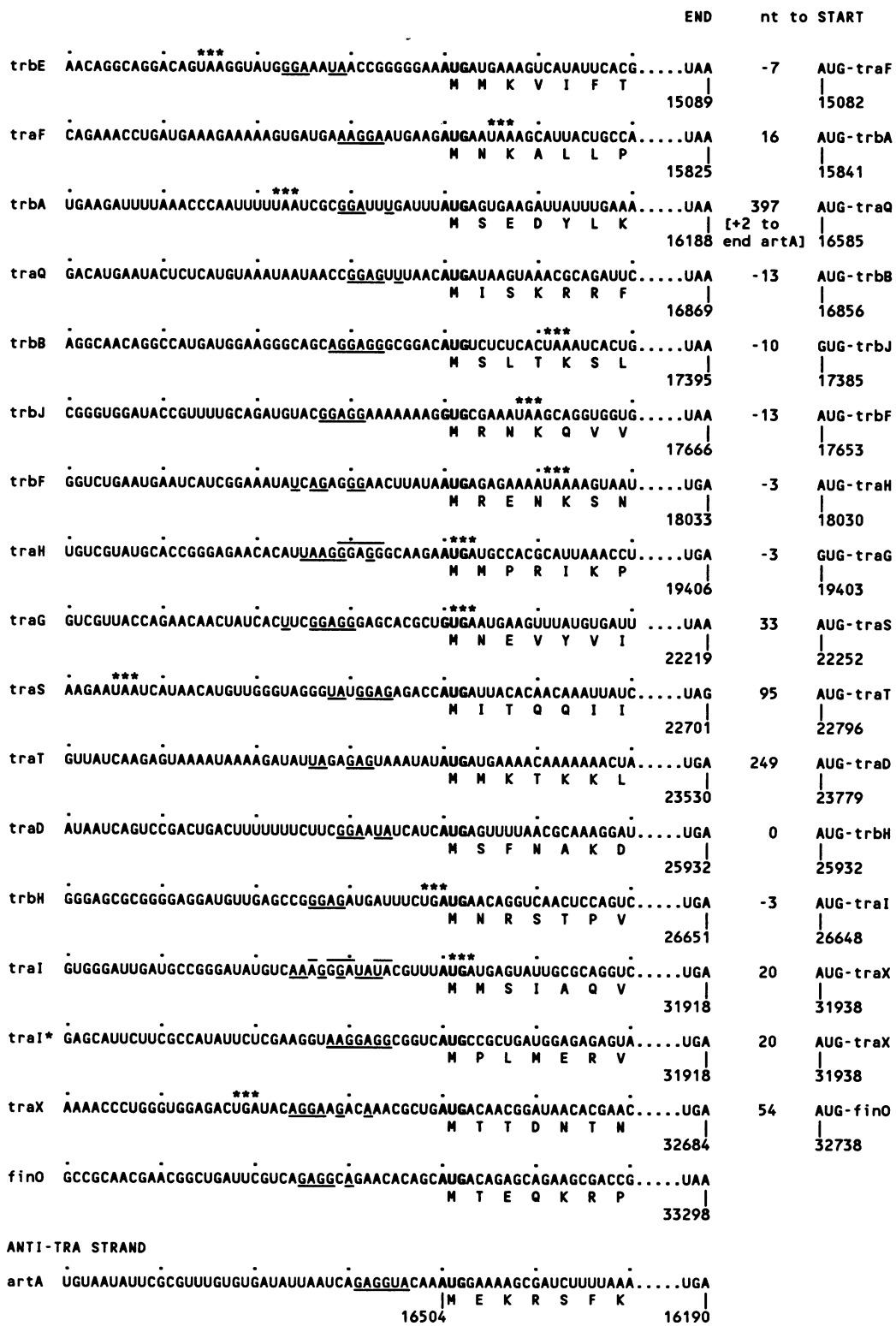


FIG. 4. Summary of the F transfer region translational start sites. The sequence of the regions at the initiation of translation for each gene is aligned with respect to the initiating codon for methionine. The possible ribosome-binding sites are underlined or overlined, and the stop codons for the previous gene (if applicable) are shown as ***. The sequence of the stop codon for each gene, as well as the sequence of the succeeding start codon, is given. The position for the stop codon for each gene is indicated below the line, as is the position of the start codon in the following gene. The number of intervening base pairs is shown; a negative number indicates that the genes overlap.

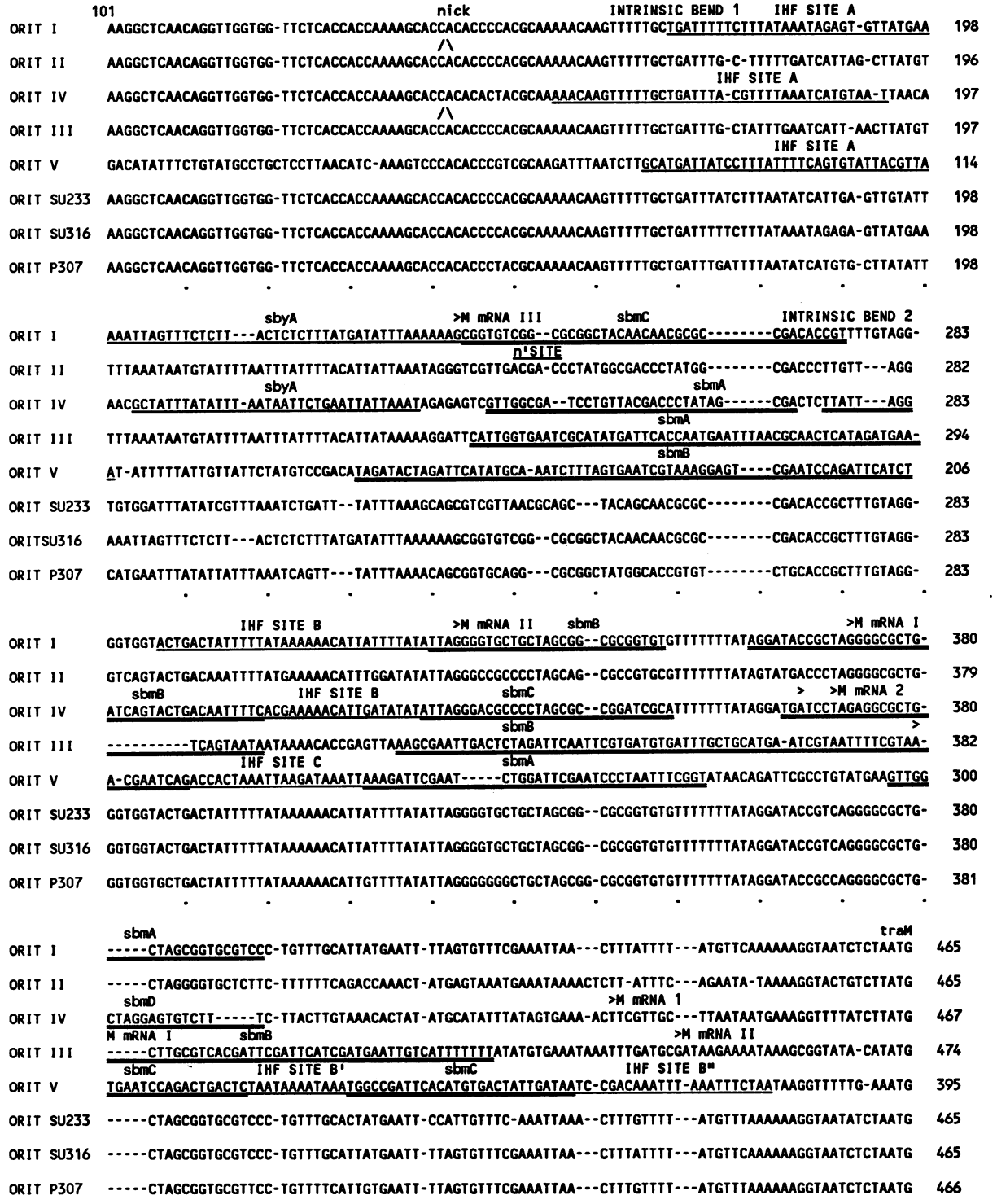


FIG. 5. Alignment of the *oriT* sequences of F-like plasmids. The alignment was performed by using the CLUSTAL program of PCGENE. For simplicity, the prototype plasmid for allele I is F, II is ColB4-K98, III is R1 or R1-19, IV is R100 or R100-1, and V is pED208 or F_0 *lac* unless otherwise stated, as suggested by Willetts and Maule (259). The sequence alignments gave rise to gaps marked by dashes in the sequences, and these gaps were taken into account when enumerating the nucleotides in the different sequences. Every 10th position is marked by a dot below the sequence, and the number of residues in a sequence is noted as a running total in the right margin. The sequence in F(I) from nt 101 (Fig. 2) to the start of the *traM* gene (nt 465) (243, 244) is compared with equivalent sequences in ColB4-K98 (II) (61), R1-19 (III) (62), R100 (IV) (58, 61), and pED208 (V) (48). In addition, the sequences for the *oriT* regions of pSU233 (222), pSU316 (152), and P307 (98) are presented. The proposed nick site for the bottom strand is marked as nick (/\) (173). The general location of the two intrinsic bends in the F *oriT* near positions 171 and 267 (248) are indicated, and details are given in the text. The proposed IHF-binding sites are underlined for F (248), R100 (46), and pED208 (50), where IHF binds more strongly to SITE A than SITE B. The binding site for TraY has been determined for F (136, 153, 195) and

quence is consistent with the proposed consensus sequence for TraY recognition discussed above.

The finding that TraY binds near its own promoter suggests that TraY may have a role in the control of *tra* gene expression. Although no *traY* amber mutant has been available, the method of Kathir and Ippen-Ihler (127) for introducing kanamycin resistance gene insertions into F *lac* or the F deletion derivative pOX38 (103) has recently been applied to *traY*. Interestingly, an F *traY* mutant that contains a kanamycin resistance gene cassette insertion between the *Bcl*I sites in the second half of *traY* (nt 2124 and 2126) is very deficient in pilus expression (128), whereas in constructs in which a T7 promoter has been substituted for P_Y, *traY* insertion mutations do not affect piliation (155). In both cases, the *traY* insertion mutations decrease the level of transfer to low but measurable frequencies (155). Thus, TraY appears to be essential for transfer and pilus expression, and a truncated *traY* gene product may have some residual activity within its first uninterrupted domain. This has also been postulated to be the case for pED208, in which an IS*2* element interrupts *traY* and a possible fusion peptide can substitute for intact TraY protein (63).

TraM Protein

Since TraM was not required for nicking but was required for replacement strand synthesis in the donor cell, it was assigned the function of signaling that DNA transfer should begin (267). Sequence analysis of the F *traM* gene (243) predicted a protein of 14.5 kDa (nt 463 to 843). The F TraM protein has been purified and shown to correspond to this ORF, although its apparent molecular mass as determined by SDS-polyacrylamide gel electrophoresis (PAGE), corresponds to a protein of 10 kDa (49). The *traM102* frameshift mutation (12) has been sequenced (nt 541) (74), and a derivative of pOX38 with an interrupted *traM* gene has been generated. These *traM* mutations reduce transfer to a background level but do not affect piliation or surface exclusion (12, 207). The level of TraM expression is not greatly affected by TraJ, since immunoblots of cells carrying F plasmids with and without the *traJ90* amber mutation produce equivalent amounts of TraM protein (236); however, *traM* is negatively regulated by TraM and TraY. Mutations in *traM* increase expression from the *traM* promoters mRNAI and mRNAII, whereas mutations in *traY* increase transcription from the mRNAIII promoter (Fig. 2) but result in a decrease in the amount of TraM protein (207).

Interestingly, the *traJ90* mutation, which prevents normal *traY* expression, causes an increase in the production of *traM* mRNA from promoter mRNAIII, as expected; however, the level of TraM protein expression is unaffected (207). These results suggest that a lack of TraY protein causes increased transcription from the *traM* promoter, which overlaps sbyA, a potential primosome-binding site and the beginning of sbmC. This increased transcription is coupled to decreased TraM protein production, although the other two promoters (mRNAI and mRNAII) for *traM* expression are unaffected (207). These results suggest that TraY may affect *traM* expression at two levels; it acts as a transcriptional regulator (207), and it affects TraM posttranslationally, perhaps by promoting the incorporation of TraM into the transferosome, hence protecting it from rapid degradation (49).

The TraM proteins of R1 (III) (227), R100 (IV) (2), and pED208 (V) (48) have been purified, and their *oriT* binding sites have been determined (Fig. 2 and 5) (see above). The eight known TraM sequences define five alleles of *traM* which, with the possible exception of the fifth allele for pED208, are highly homologous if conservative substitutions are allowed (Fig. 7). The sequences of the *traM* genes from pSU233, pSU316, and P307 are almost identical to that for F and contain only minor alterations (98, 152, 222). However, it is not known whether these genes can complement a *traM* mutation in F. The small number of differences at the N terminus of the first four alleles suggests that this domain is involved in binding to *oriT*. Mutational analysis of the R1 TraM protein has confirmed this prediction and has shown that mutations that affect TraM action reside in an α-helical segment of the protein, which may contact the major groove of the DNA (228).

The F and pED208 TraM proteins are highly structured tetrameric proteins with an α-helical content of 79 and 66%, respectively. The remainder of the protein is involved in β-sheet and β-turns, as determined by circular dichroism. TraM is a cytoplasmic protein, and a small amount is found in the inner membrane fraction of F and pED208, suggesting that it may be part of a membrane-associated relaxosome complex (49). The TraM protein has no obvious homology to DNA-binding proteins with a helix-bend-helix motif (242), suggesting that its DNA-binding domain is unusual and may involve a single α-helical segment (228). F TraM, unlike the other F-like TraM proteins, binds the top strand of the DNA more than to the other strand, as assayed by DNase I protection experiments (49). This property is shared by the TraK protein of RP4, which has also been demonstrated to generate nucleosome-like structures in the *oriT* region of that plasmid (276). It is not known whether F TraM also bends and wraps the DNA into a higher-order structure; results with pED208 TraM suggest that this allele of TraM binds both DNA strands equally and does not bend or compact the DNA (50).

TraD Protein

The TraD protein is a large, 81.7-kDa protein (nt 23779 to 25929) found in the inner membrane of F^+ cells (202). The *traD* genes of F and R100 were sequenced by Jalajakumari and Manning (124) and by Yoshioka et al. (274), respectively. The F TraD protein is predicted to be hydrophilic, with an acidic pI (<6.0); however, it is soluble only in the presence of detergents (202) and has three hydrophobic domains that may anchor it in the membrane (Table 1). Analysis of TraD revealed a putative ATP/GTP-binding motif (GTVGAGKS; nt 24352 to 24376) (274) and a region resembling the signature sequence for the ATP synthase alpha and beta subunits (PAINDKKSDS; nt 25645 to 25689) (249). An interesting difference between the F

R100 (115), and the nomenclature sbyA for the TraY-binding site has been adopted from Inamoto and Ohtsubo (115). The primosome assembly site (n′ site), proposed by Willetts and Wilkins (267), is shown and corresponds to the sequence AAGCGG. The TraM-binding sites are double-underlined and correspond to information provided for F TraM (49), R1 TraM (227), R100 TraM (2), and pED208 TraM (48, 50). The nomenclature sbmA, sbmB, etc., denotes the order of the affinity of TraM for these sites in F and pED208. In pED208 (V), the IHF-binding site SITE B has been divided into two binding sites SITE B′ and SITE B″ which bracket sbmD. Three sites of initiation of transcription of *traM* have been reported for F (207, 243); two for R1 (130) and two for R100 (3, 45) are indicated as >M mRNA I, II, or III. TraM indicates the start of the *traM* coding sequence.

```
                    .         .         .         .         .         .         .
TRAY IN             MKRFGTRSATGKMVKLKLPVDVESLLIEASNRSGRSRSFEAVIRLKDHLHRYPKF-NRAGNIYGKS--------  65
TRAY IC             ---------LVKYLTMRLDDETNQLLIAAKNRSGWCKTDEAADRVIDHLIKFPDFYNSEIFREADKEEDITFNTL  66
TRAY II             MRRRNARGGISRTVSVYLDEDTNNRLIKAKDRSGRSKTIEVQIRLRDHLKRFPDFYNEEIFREVTEESESTFKEL  75
TRAY III            MRRRNARGGISRTVSVYLDEDTNNRLIRAKDRSGRSKTIEVQIRLRDHLKRFPDFYNEEIFREVIEENESTFKEL  75
TRAY IV             MSRNIIRPAPGNKVLLVLDDATNHKLLGARERSGRTKTNEVLVRLRDHLNRFPDFYNLDAIKEGAEETDSIIKDL  75
TRAY V    MKEPKSNIRKLIDIGGLIGKKVNISCSLDEAIDELLMESALKSGWSKKREAELRLEDHLRRFSLVPAEEQYIEKKVD--------  77
TRAY P307           MSRGRTRAAPGNKVLLILDETTNQKLLAARDRSGRTKTNEVFIRLKDHLNRFPDFYNSSLVKEEAEGIDDI----  71
                    . . . . .   .  *:.... .*. :  :**:.:. *:  *: *** :::.:.. .   :   .   .   .
```

FIG. 6. Comparison of the available sequences of TraY. The N-terminal (IN) and C-terminal (IC) halves of the F TraY protein (72, 116) have been compared with each other and with the sequences of ColB4-K98 (61), R1-19 (62), R100 (61), and pED208 (48, 63). Note that the sequences of TraY from ColB4-K98 and R1 are virtually identical. In addition, the sequences for the TraY protein of P307 (98) is also presented. Identical residues are marked with an asterisk, positions with two possible residues are marked with a colon, and positions with three different possible amino residues are marked with a period. Four or more different sequences are left as blanks. The sequence alignments gave rise to gaps (marked by dashes) in the sequences, and these gaps were taken into account when enumerating the residues in the different sequences. Every 10th position is marked by a dot below the sequence, and the number of residues in a sequence is noted as a running total in the right margin.

and R100 TraD proteins is the repeated sequence QQP, which occurs eight times consecutively in R100 and only once in F (nt 25653; amino acids [aa] 622 to 624) (274). Homology between F TraD and TrwB of R388 (IncW) (42), TraG of RP4 (IncP), and VirD4 of the Ti plasmid (144) has been found. The nucleoside triphosphate (NTP)-binding site mentioned above was initially found to be conserved in F and R100 but not TraG and VirD4. Instead, a type B NTP-binding site is conserved throughout these sequences and is represented in F by the sequence WFFDEL at nt 25039 to 25059 (aa 420 to 426) (144). Recently, homology between TraD of F and TraK of the staphylococcal conjugative plasmid pKS41 has been observed, and realignment of the all the TraD homologs raises the possibility that the type A and type B NTP-binding sites in F are conserved among all of these proteins (69).

TraD has been assigned a role in actively transporting the transferred strand into the recipient cell. This is based on the observation that *traD* mutants are very transfer deficient but active in *oriT* nicking and DNA displacement assays (57, 129) and are able to express pili and to form functional stable mating aggregates (161, 167, 193, 201). Recently, the TraD protein has been purified and shown to bind to a DNA-cellulose column in a pH-sensitive manner, indicating nonspecific binding to nucleic acids (202). TraD is also involved in the transport of other nucleic acids; it is required for the penetration of RNA from the F-specific RNA phages R17, MS2, and f2 but not Qβ or the filamentous phages (204). It is also required for the mobilization of the ColE1 plasmid but not the closely related plasmid CloDF13, although copy number mutants of CloDF13 inhibit TraD function through the FinC fertility inhibition system (258, 267). Cells carrying F *traD* mutations are slightly multipiliated, suggesting that TraD may also have a function in pilus assembly and retraction (14, 26).

TraI Protein

The TraI protein is a large soluble protein of 192 kDa (nt 26648 to 31915), also known as helicase I, whose functional properties have been extensively characterized by Hoffmann-Berling and colleagues. In 1983, helicase I was found to be the product of the *traI* gene in the *E. coli* Hfr strains used in its purification (1). F *traI* was sequenced by Bradshaw et al. (25); the TraI protein is an ATP-dependent helicase which requires a single-stranded region of DNA of about 200 nt as a substrate. That it processively unwinds DNA in a 5′-to-3′ direction is consistent with its role in the transfer of the single-stranded DNA (256). Recent studies have further characterized its enzymatic activity as a helicase, including the number of TraI molecules required for optimal activity and the kinetics of ATP hydrolysis (21, 40, 172).

TraI is essential to transfer. An early finding, consistent with a requirement for helicase I for DNA strand displacement, was that replacement strand synthesis did not occur in mating aggregates formed by F *traI* mutants (129). More recently, the TraI protein has been demonstrated to be required for nicking at *oriT* (247); it functions as a phosphodiester transferase catalyzing a reaction in which the 5′ end of the nicked DNA becomes covalently bound to TraI (195). The 3′-OH terminus of nt 140 is free to act as a substrate for DNA synthesis by DNA polymerase I in vitro, provided that the DNA has been treated with SDS and proteinase K. Whether this is the result of a noncovalent association of the 3′-OH with TraI or represents physical blocking of the nick site remains to be determined (173). The requirement for RNA primer synthesis to initiate replacement strand synthesis in the donor in vivo argues against simple rolling-circle replication from the free 3′-OH and suggests a more complex mechanism for initiating replacement strand synthesis (129, 267).

In vitro, the nicking reaction requires a supercoiled DNA substrate and Mg^{2+} but not the addition of ATP or the presence of TraY (172, 218). The effect of TraM or TraD has not been tested. Sequence comparisons of other relaxase enzymes that also attach covalently to the 5′ terminus suggest that the two pairs of tyrosines at the N terminus of TraI (aa 16, 17 and 23, 24; see nt 26693 and 26714 [Fig. 2]) may be important for nicking and for correct termination after one round of transfer (42). This is reminiscent of the A protein in ϕX174, in which two tyrosines spaced 4 residues apart are essential for initiating and terminating the packaging of a DNA strand from the replicative form of the viral DNA (253). A model for nicking and religation involving a single tyrosine residue in the TraI protein of the RP4 transfer system has been presented by Wilkins and Lanka (256). The F TraI and R388 TrwC (IncW) proteins share 33% sequence identity, and the TrwC sequence has been used to show weak homology between F TraI and RP4 TraI (42). Unlike RP4 TraI, F TraI and TrwC also have helicase activity within their C-terminal domains. In addition, a short sequence within TraI (aa 145 to 167; nt 27080 to 27148 [Fig. 2]) and TrwC can be aligned with a "relaxase motif" found in a number of other enzymes that catalyze site-specific cleavage reactions (42, 203).

Deletion and insertion analysis of the *traI* gene confirmed that the nicking activity is in the N-terminal domain of the protein (247). The product expressed by the *traIdel29* mutant, which lacks the first (N-terminal) 14% of *traI*, had lost nicking activity but retained strand separation activity (218). The TraI*

```
TraM I      MAKVNLYISNDAYEKINAIIEKRRQEGAREKDVSFSATASMLLELGLRVHEAQMERKESAFNQTEFNKLLLECVVKTQSSVAKILGIESLSPHVSGNSKF 100
TraM P307   MAKVNLYISNDAYEKINAIIEKRRQEGAREKDVSFSATASMLLELGLRVYEAQMERKESAFNQTEFNKLLLECVVKTQSTVAKILGIESLSPHVSGNPKF 100
TraM SU233  MAKVNLYISNDAYEKINAIIEKRRQEGAREKDVSFSATASMLLELGLRVYEAQMERKESAFNQAEFNKLLLECVVKTQSSVAKILGIESLSPHVSGNPKF 100
TraM SU316  MAKVNLYISNDAYEKINAIIEKRRQEGAREKDVSFSATASMLLELGLRVHEAQMERKEFAFNQAEFNKLLLECVVKTQSSVAKILGIESLSPHVSGNSKF 100
TraM IV     MARVILYISNDVYDKVNAIVEQRRQEGARDKDISVSGTASMLLELGLRVYEAQMERKESAFNQTEFNKLLLECVVKTQSSVAKILGIESLSPHVSGNPKF 100
TraM III    MARVNLYISNEVHEKINMIVEKRRQEGARDKDISLSGTASMLLELGLRVYDAQMERKESAFNQTEFNKLLLECVVKTQSTVAKILGIESLSPHVSGNPKF 100
TraM II     MAKVQAYVSDEIVYKINKIVERRRAEGAKSTDVSFSSISTMLLELGLRVYEAQMERKESAFNQAEFNKVLLECAVKTQSTVAKILGIESLSPHVSGNPKF 100
TraM V      MPKIQTYVNNNVYEQILDLVTIRKQEGE---EASLSNVSSMLLELGLRVYMIQQEKREGGFNQMEYNKLMLENVSRVRAMCTEILKMSVLNQESIASGNF  97
            *:::. *:::....::  ::: *:.**:. .:.*.* .::*********:.:*:*::*.:***.*:**::**::::::::.:::**::::*:::::::..:*

TraM I      EYANMVEDIREKVSSEMERFFPKNDD--E  127
TraM P307   EYSNMVEDIREKVSVEMERFFPKNDD--E  127
TraM SU233  EYANMVEDIREKVSVEMERFFPKNDD--E  127
TraM SU316  EYANMVEDIREKVSSEMERFFPKNDD--E  127
TraM IV     EYANMVEDIREKVSSEMERFFPKNDE--E  127
TraM III    EYASMVDDIREKVSVEMDRFFPKNDD--E  127
TraM II     EYANMVEDIRDKVSSEMERFFPENDE--E  127
TraM V      DYAVIFPAIDKFAREQVSIFFPDDEDDQE  126
            :*:.::.:*:.:::.::.:***.::::::*
```

FIG. 7. Comparison of the available sequences of TraM. The sequences of TraM are from F (I) (243), pSU233 (222), pSU316 (152), P307 (98), R100 (IV) (58), R1-19 (III) (62), ColB4-K98 (II) (61), and pED208 (V) (48). The sequences are arranged in order of their similarity to the F TraM sequence, and the notation in the figure is the same as in Fig. 6.

protein, encoded by the same ORF as TraI, is predicted to begin at nt 29510 (Fig. 2) (246). TraI* corresponds to the polypeptide once called TraZ and originally identified as protein 2b (166); although TraI* was suggested in early models to be required for nicking, it is now clear that it has no role in the nicking reaction. Its function is currently unknown (246, 247). Both TraI and TraI* contain a potential ATP/GTP-binding motif (GYAGVGKT; nt 29621 to 29642; aa 992 to 999) (249) identified by sequence analysis; the correlation between this site and ATP-dependent helicase activity has not been tested, although ATPase activity for the C-terminal domain of TraI has been demonstrated (247). Bradshaw et al. (25) pointed out a second possible NTP-binding motif (nt 27176 to 27199; aa 177 to 184), and Dash et al. (40) have detected two rates of ATP hydrolysis depending on the ATP concentration used in the assay. These authors have speculated that ATP hydrolysis may be required for both breakage of hydrogen bonds between base pairs and translocation of the helicase molecules along the DNA strand. In addition, the level of aggregation of TraI may affect its ability to hydrolyze ATP, with maximal hydrolysis occurring as TraI aggregates on its DNA substrate (40).

Areas of Investigation

Currently the model for the initiation of F transfer involves a series of unlinked processes. The *oriT* region appears to be organized into a nucleosome through the action of TraM and IHF, which are thought to alter the superhelical density (48, 49). The possibility that adjustments in DNA topology modulate initiation events should be investigated. TraY seems to have a critical function both in controlling a redundant *traM* promoter near *oriT* (M mRNAIII), which overlaps a possible primosome assembly site, and in activating the major transfer operon at P_Y. Interruption of TraY activity could lead to decreased expression of pilus synthesis genes and increased transcription from the M mRNAIII promoter, perhaps in preparation for transfer. TraI is clearly involved in nicking and unwinding of DNA during transfer; however, the role of TraI* is unknown. Prior to conjugation, TraI is a nicking (-religating) enzyme; its conversion (activation) to a helicase is of primary importance to transfer initiation. Certainly, the A+T-rich character of the IHF site A and sbyA sequences coupled with the promoter for *traM* in this region suggests a step which opens the DNA in preparation for entry of the bulky TraI helicase. If the TraI protein remains associated in some way with both the 5′ and 3′ ends of the transferred strand, a mechanism for feeding the DNA presumably along or past a complex containing TraD into the recipient cell must be uncovered. In addition, the establishment of a new transferosome in the recipient cell and the mechanism for terminating transfer require further study. To understand the "signals" that initiate transfer, more information is needed about the interactions among the proteins forming the transferosome in the envelope of the donor cell.

REGULATORY REGION OF THE F TRANSFER OPERON

The regulatory region includes the *traJ* and *finP* genes and the P_Y promoter. TraJ is a regulatory protein required for the initiation of high levels of transcription from the P_Y promoter. *finP* encodes an antisense RNA, FinP, that negatively regulates TraJ expression in the presence of the *finO* gene product, which is encoded at the distal end of the transfer region (68, 81). When FinO is present, FinP is stabilized and blocks *traJ* expression by interacting with the leader region of the *traJ* transcript, possibly by preventing translation from the ribosome-binding site for *traJ*. In F, the *finO* gene is interrupted by an IS*3* element (31, 275), leading to constitutive expression of *traJ*, derepressed transfer operon expression, and pilus synthesis. However, in the presence of active FinO protein in the host, the FinOP repression system causes a three- to fivefold decrease in the amount of *traJ* transcript, and the F transfer operon becomes repressed (141).

Dempsey has proposed a "latch relay" system for control of transfer operon expression in R100 (45). If the amount of FinP falls below a critical level, *traJ* is transcribed and the transfer operon is expressed transiently. In addition, Dempsey has proposed two R100 *finP* transcripts, both of which terminate at the same position but one originates further within the *traJ* gene than the F transcript indicated in Fig. 2 (43). Two TraM-associated transcripts which affect the levels of FinP have been identified for R100 (45) and R1 (132, 135) (Fig. 2). Analysis of the transcripts from the control region of F has shown that there is only one *finP* transcript and that there are no observable transcripts from *traM* that extend into the

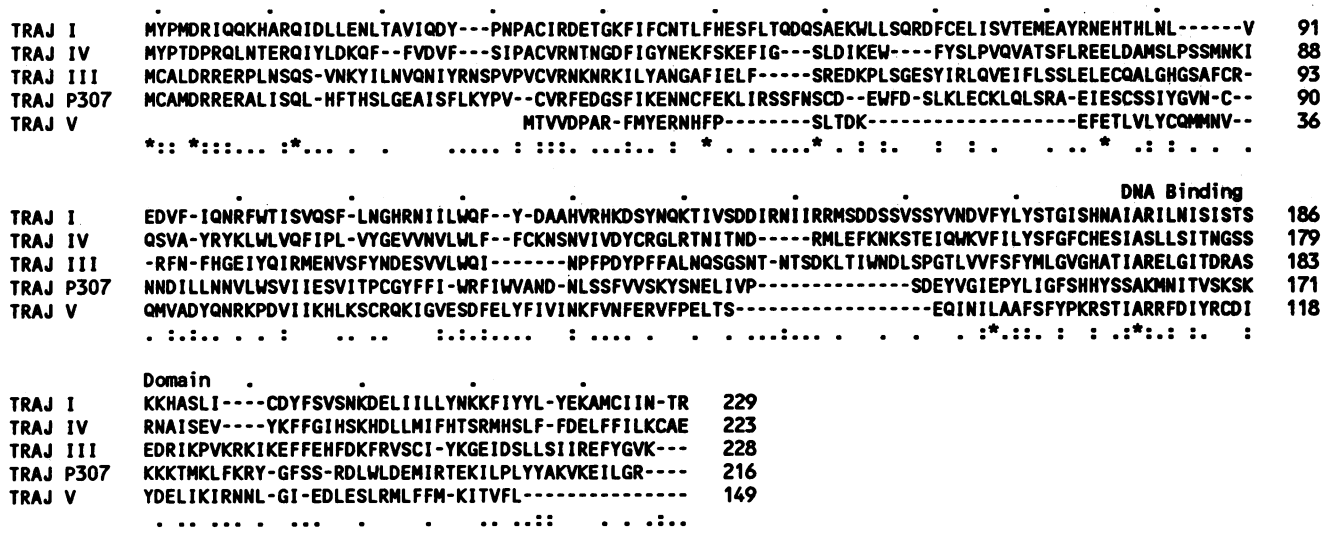


FIG. 8. Comparison of the available sequences of TraJ. The sequences of TraJ are from F (I) (72), P307 (98), R1-19 (III) (62), R100 (IV) (116), and pED208 (V) (48). The region containing the proposed DNA-binding domain is indicated above the F sequence. The sequences are arranged in order of their similarity to the F TraJ sequence, and the notation in the figure is the same as in Fig. 6.

finP-traJ genes that could contribute to the control of *traJ* expression (43, 141, 207).

TraJ Protein

Expression of the F transfer region from the P_Y promoter is controlled by the *traJ* gene product, which is a positive regulator of transcription (257). The sequence of the F *traJ* gene revealed a protein of 27 kDa (nt 1033 to 1719) (72, 243). Initially assigned to the outer membrane, TraJ was later shown to be a cytoplasmic protein in F^+ cells. This discrepancy was due to the insolubility of the TraJ protein expressed by multicopy vectors, which led to its appearance in the outer membrane fraction of detergent-solubilized cell extracts (38, 39). An amber mutant, *traJ90*, has been characterized, and the mutation has been localized (Fig. 2, nt 1108) (243).

Four alleles of *traJ* have been characterized (Fig. 8); the first three were described by Willetts and Maule (259), and the fourth, from pED208, which is 35% shorter than F *traJ*, was described subsequently (48). The four variants of TraJ are remarkably dissimilar from one another, unlike other *tra* genes which have multiple alleles (*traM* or *traY*, for instance). The area that has the greatest homology is in a region that could be a DNA-binding domain, since it resembles the consensus sequence for DNA-binding proteins having a helix-bend-helix motif (242) (Fig. 8). The TraJ proteins are characterized by a fairly high aromatic amino acid content (12 to 15%). A fifth variant of TraJ, expressed by plasmid P307, has also been described and is also presented in Fig. 8 (98).

The mode of action of TraJ has been puzzling. F TraJ has been shown to be a positive regulator of transcription originating from the P_Y promoter with either *lacZ* (87) or *galK* (194) as the reporter gene. However, TraJ has not yet been shown to bind to an operator sequence in the P_Y promoter region, although deletion analysis of this region has demonstrated that this is its site of action (see the section on the P_Y promoter, below).

traJ Promoter and Antisense RNA, FinP

The *traJ* promoter has been mapped for the F (243), R1 (134), and R100 (43) plasmids. The untranslated portion of the F *traJ* transcript is 105 nt long, and the complementary antisense RNA, FinP, is approximately 75 bases long (Fig. 2, nt 1038 to ca. 965) (43, 81, 132, 194). A small inverted repeat within the leading region of the *traJ* transcript (ACGTGGT TAATGCCACGT from nt 928 to 945 [Fig. 2]) resembles the consensus sequence for the cyclic AMP receptor protein-cyclic AMP complex (55). The transfer operon is known to be sensitive to catabolite repression, as evidenced by the poor expression of the transfer operon in strains of *E. coli* carrying the *cya* mutation, which have increased pilus production on the addition of exogenous cyclic AMP (108, 176). Perhaps the discrepancy between the length of FinP and the untranslated portion of the *traJ* transcript would accommodate control of *traJ* by cyclic AMP receptor protein-cyclic AMP complex (205).

The F *finP* transcript has been predicted to fold into two stem-loop structures in which the first stem is complementary to the ribosome-binding site for *traJ* (Fig. 2, nt 968 to 1035) (65, 132). Recently, the structure of the FinP RNA and the first 211 bp of the *traJ* mRNA has been demonstrated to correspond to the predicted structure (252). Among the six alleles of FinP originally identified by Willetts and Maule (259), only five different sequences of the *finP* gene were obtained (Fig. 9) (65). These differences reside principally in the loops of the predicted stem-loop structures and define the specificity of each FinP for its cognate *traJ* mRNA (65, 132). The *finP* region of the R100 plasmid (group IV; Fig. 9) has an extra predicted stem-loop which is not present in the other alleles of *finP*. Dempsey (43, 45) has proposed a more complex model for R100 *traJ* repression based on three possible stem-loop structures. Two other variants of FinP have recently been isolated from plasmids pSU233 and P307; they also show the characteristic variation in sequence in the predicted loops (98, 222).

The specificity of FinP action may be loop specific, in a manner similar to the interaction of RNAI with RNAII that regulates replication of ColE1-like plasmids (56). Koraimann et al. (132) demonstrated that mutations affecting the loops of the R1 FinP RNA reduced the ability of cloned *finP* to repress the transfer of R1 in *trans*. Duplex formation between FinP and *traJ* RNA has been demonstrated in vitro, and the association constant is similar to that reported for other

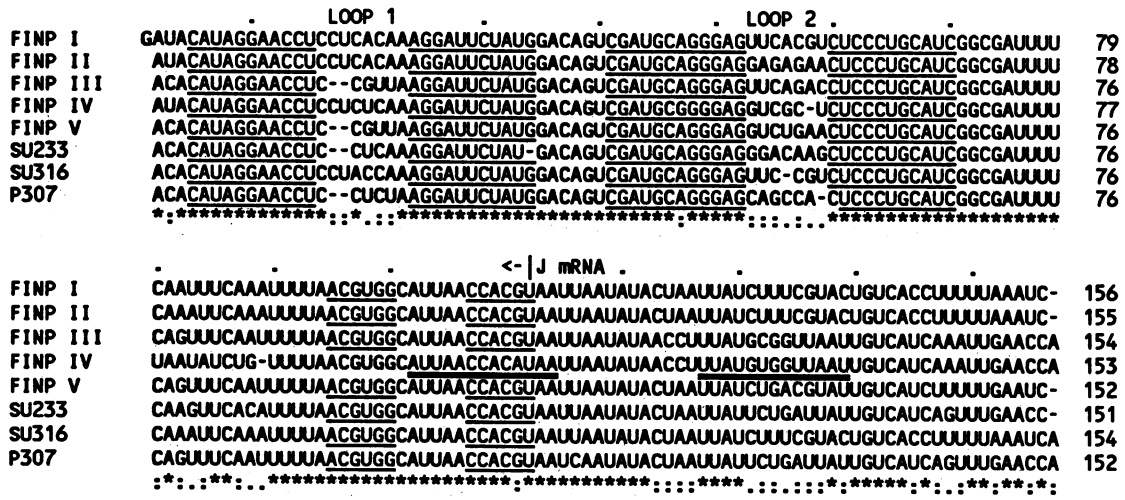


FIG. 9. Comparison of the sequences of the *finP* RNAs and the region containing the *traJ* promoter taken from Finlay et al. (65), where allele I corresponds to F; II corresponds to ColB2Fdr, ColVB*trp*, and R124; III corresponds to R1-19; IV corresponds to R100-1 and R386; and V corresponds to ColB4-K98. In addition, the sequences of the corresponding *finP* RNAs and *traJ* promoter region for pSU233 (222), pSU316 (152), and P307 (98) are given. The 5′ terminus of FinP has been mapped for F (252), R1-19 (132), and R100 (43). The 5′ termini of the other FinP RNAs are arbitrarily chosen. The inverted repeats forming stem-loops 1 and 2 are underlined in the top half of the figure, and Loop 1 and Loop 2 indicate the positions of the loops. The double-underlined inverted repeat in the bottom half of the figure is unique to R100 (FinP IV) (43). The small, single-underlined repeat at the beginning of the *traJ* transcript is the putative cycle AMP-cyclic AMP receptor protein-binding site mentioned by Paranchych et al. (205). The position of the start of the *traJ* transcript in F is given as <−|J mRNA. The sequence presented represents the bottom strand in a 5′-to-3′ direction such that the orientation of the *traJ* mRNA is the reverse of that presented in Fig. 2. While the sequence to the right of the *traJ* start of transcription represents the intergenic region between *traM* and *traJ* as well as the end of the *traM* transcript (in F), the sequence is presented as RNA for the sake of simplicity.

antisense RNA systems (252). An F *fisO* (site of action of FinOP) mutant, which is constitutive for transfer even in the presence of FinO, was found to carry a point mutation in *finP* that causes FinP RNA to be inactive and insensitive to FinO (C-to-T at nt 1010) (67, 68, 81). Recent experiments have shown that FinO stabilizes FinP RNA even in the absence of the *traJ* gene, suggesting that FinO acts on FinP itself (141). In vitro experiments have demonstrated that FinO binds perfectly matched duplex RNA and consequently binds stem-loop II of both TraJ and FinP RNAs; it also accelerates duplex formation between FinP and *traJ* mRNA fivefold in vitro (251). FinO extends the half-life of FinP RNA from 2 to >40 min (141, 251); the FinP RNA carrying the *fisO* mutation is unaffected and degrades quickly (141). When the *fisO* mutation is combined with additional mutations that increase the free energy for forming the first stem in FinP, the phenotype remains derepressed, indicating that the *fisO* site either affects the interaction of FinO with FinP RNA or increases the degradation of FinP by an as yet uncharacterized intracellular RNase (81).

A mutant of F, SLF20 (141), carries the same mutation in the *finP* promoter first described for the FinP$^-$ mutant of ColB2Fdr named pED236 (nt 1046) (65). This mutant produces no detectable FinP and maintains a *traJ* mRNA level threefold higher than that in wild-type F; however, it does not produce significantly greater numbers of pili (five per cell). The *traM-traJ* intergenic region in pED208 is unrelated to the corresponding sequence in F, and no inverted repeats which would be the equivalent of those in FinP have been identified (48).

FinO Protein

The FinO protein (21.2 kDa) is encoded by the last known gene of the transfer regions of F-like plasmids (Fig. 1; nt 32738 to 33295) (175, 274, 275) and is essential for FinP activity (67). The FinO protein is predicted to be a soluble, cytoplasmic protein with a basic nature (pI = 10.5). Two alleles have been reported for *finO* (259) on the basis of their levels of repression of F and other F-like derepressed plasmids, although there is little difference in sequence between group I (R100, R6-5) (35, 175, 274) and group II (ColB2) (250) alleles (Fig. 10). The allelic difference in *finO* is dependent on the presence of a gene upstream from *finO*, found only on plasmids from group I alleles and called *orfC* (R100) (274) or *orf286* (R6-5) (35). The F plasmid has no gene equivalent to *orf286/orfC*; however, the position of insertion of this gene in other F-like plasmids is shown in Fig. 2 (nt 33133). FinO is expressed very poorly from the T7 ϕ10 promoter in an overexpression system because of the short half-life of the mRNA. However, if the gene *orf286* is supplied in *cis* but not in *trans*, an accumulation of FinO protein in the cells is observed and the half-life of *finO* mRNA is increased. This increase in mRNA stability has been localized to a region of complementarity between the *orf286* and *finO* sequences which may base pair with each other at the mRNA level and protect the *finO* transcript from degradation (250). Therefore, the level of repression of the transfer region is highly dependent on the concentration of FinO as well as that of FinP.

Downstream from *finO* are two genes, *orfB* and *orfA*, which have been completely sequenced for R100 and partially sequenced for F; the *orfB* sequences are related, whereas the *orfA* sequences diverge (274). The *orfA* gene product is related to a family of nucleases including the *Staphylococcus aureus* nuclease (34) and nucleases of the pSa plasmid (32) and *parB* locus of RP4 (93). It is interesting that this region of F contains *finO* and *orfC/orf286* (in F-like plasmids), which stabilize RNA, *orfA*, which encodes a potential nuclease, and *srnB*, a gene involved in F plasmid maintenance, which may allow RNase I

```
FINO I F    MTEQKRPVLTLKRKTEGETLVRSRKTIINVTTPPKWKVKKQKLAEKAAREAELAAKKAQARQALSIYLNLPTLDDAVNTLKPWWPGLFDGDTPRLLACGI  100
FINO I B2   MTEQKRPVLTLKRKTEGETPVRSRKTIINVTTPPKWKVKKQKLAEKAAREAELAAKKAQARQALSIYLNLPSLDEAVNTLKPWWPGLFDGDTPRLLACGI  100
FINO IIR65  MTEQKRPVLTLKRKTEGETPTRSRKTIINVTTPPKWKVKKQKLAEKAAREAELTAKKAQARQALSIYLNLPTLDEAVNTLKPWWPGLFDGDTPRLLACGI  100
FINO IIR100 MTEQKRPVLTLKRKTEGETPTRSRKTIINVTTPPKWKVKKQKLAEKAAREAELTAKKAQARQALSIYLNLPSLDEAVNTLKPWWPGLFDGDTPRLLACGI  100
            ********************::**********************************:*****************:**:************************

FINO I F    RDVLLEDVAQRNIPLSHKKLRRALKAITRSESYLCAMKAGACRYDTEGYVTEHISQEEEAYAAERLDKIRRQNRIKAELQAVLDEK  186
FINO I B2   RDVLLEDVAHGNIPLSHKKLRRALKAITRSESYLCAMKAGACRYDTEGYVTEHISQEEEAYAAERLDKIRRQNRIKAELQAVLDEK  186
FINO IIR65  RDVLLEDVAQRNIPLSHKKLRRALKAITRSESYLCAMKAGACRYDTEGYVTEHISQEEEVYAAERLDKIRRQNRIKAELQAVLDEQ  186
FINO IIR100 RDVLLEDVAQRNIPLSHKKLRRALKAITRSESYLCAMKAGACRYDTEGYVTEHISQEEEVYAAERLDKIRRQNRIKAELQAVLDEQ  186
            *********::**************************************************:**********************:
```

FIG. 10. Comparison of two sequences of *finO*, allele I, with two sequences of *finO*, allele II, illustrating the near identity between the sequences. The F FinO sequence is taken from Yoshioka et al. (275), ColB2 is taken from van Biesen and Frost (250), R6-5 is taken from Cram et al. (35), and R100 is taken from Yoshioka et al. (274, 275) and McIntire and Dempsey (175). The sequences are arranged in order of their similarity to the F FinO sequence, and the notation in the figure is the same as in Fig. 6.

access to the cytoplasm as a result of membrane damage (91, 198).

P_Y Promoter

The P_Y promoter, formerly called the P_{YZ} promoter, was originally thought to be the site of initiation of a 32-kb mRNA transcript which encoded all the genes in the F transfer region (*traY* to *traZ*) except *traM* and *traJ* (111, 257). However, because the *traZ* gene no longer exists (the *traZ* gene product is TraI*) (246) and because there is uncertainty about the length of this transcript, the name for the promoter has been shortened to P_Y.

An inverted repeat immediately following the *traJ* gene and preceding the P_Y promoter could act as a rho-independent terminator for the *traJ* transcript (nt 1738 to 1758). The P_Y promoter has been mapped by Fowler et al. (72), Mullineaux and Willetts (194), and Silverman et al. (234), and the initiation of transcription has been shown to be located within the *Bst*EII site at nt 1789 to 1795. Data obtained with deletion mutants also led Fowler and Thompson (73) to propose that a pair of overlapping shadow promoters are active in the absence of the primary promoter. More recently, Silverman et al. (234) positioned the 5′ terminus of the predominant mRNA species expressed in F *lac* and reporter plasmid hosts at nt 1792, but they also observed weak signals for transcripts as much as 20 nt longer or 14 nt shorter and suggested that several promoters may be active in this region. Alternatively, such variation may reflect "transcriptional slippage," since Xiong and Reznikoff (272) have shown that transcripts initiated at a run three or more T · A base pairs may carry variable oligo(U) sequences at their 5′ ends. These authors also identified a pattern of highly conserved bases that seems to compensate for the absence of a typical −35 sequence in some *E. coli* sigma 70 promoters. Interestingly, the same pattern is present in the F P_Y promoter region, including the TXTGX (nt 1873 to 1877) immediately prior to the suggested −10 region (aAacAT). Although TraJ binding to P_Y has not been demonstrated, the TraY protein has been shown to bind to the P_Y region in F between the *Bst*EII site and the beginning of the *traY* gene (sbyB; see the section on the TraY protein, above) (195); the R100 TraY protein binds to a similar site in the R100 P_Y promoter region (115). The P_Y promoter of the R1 plasmid and transcription initiation site appear to lie within the 3′ end of the reading frame for the R1 *traJ* gene, suggesting that, on this plasmid, *traJ* expression could affect transcriptional activation of the P_Y promoter in *cis* (134).

As well as requiring TraJ for initiation, transcription from the P_Y promoter depends on the product of the chromosomally encoded gene *sfrA* (also known as *arcA*, *cpxC*, *dye*, and *fexA* [27, 121, 142, 233]), a member of the family of two-component regulators (240). CpxA is probably the sensor that acts in conjunction with SfrA (254); no function has been assigned to a third gene affecting F expression, *cpxB*. Interestingly, mutations in SfrA/ArcA that affect the control of aerobic respiration do not affect F transfer. Consequently, the name SfrA continues to be used in the context of F transfer operon expression. The decrease in P_Y activity in an $SfrA^-$ host was not attributable to a change in TraJ expression; in addition, transcription from P_Y was decreased in strains carrying mutations in IHF, suggesting a role for IHF in transcription from this promoter (88, 234). *traJ*-independent or *sfrA*-independent mutants have been isolated, and the mutations were found to be clustered in a region upstream from the site of transcription initiation (235). Recently, Gaudin and Silverman (90) have suggested that the P_Y promoter is sensitive to levels of supercoiling and that a nucleoprotein complex of unspecified composition alters the superhelical density at this promoter. TraJ and perhaps other activators could disrupt this interaction, leading to transcription of the transfer operon (90). Thus, the P_Y promoter may prove to be very complex, with levels of control imposed by the host through the SfrA and IHF proteins as well as the *tra* gene products, TraJ and TraY, which may alter the superhelical density of the promoter region.

GENES FOR PILUS SYNTHESIS AND ASSEMBLY

The F pilus, an extracellular filament that extends from the surface of the bacterial cell, recognizes various receptors. Its roles include recognition of and attachment to suitable recipient cells (the first step toward conjugation), recognition of surface exclusion proteins (leading to disengagement from mating contacts with cells carrying a closely related plasmid), and provision of attachment sites for pilus-specific bacteriophages. Current evidence indicates that once extended, F pili can also depolymerize into the donor or recipient cell. Thus, recipient cells or phages to which these filaments have attached are carried to the donor cell surface as the F pilus depolymerizes.

F pilus filaments consist of a single subunit (pilin) arranged in a helical array to give a fiber 8 nm in diameter with a 2-nm hollow core. The helix contains five subunits per turn, with a rise of 1.28 nm and a repeat distance of 32 nm (170, 206). The length of the pilus can be very variable, ranging from 1–2 to 20 μm under unusual growth conditions. The pili of the F-like plasmids can be distinguished from each other serologically

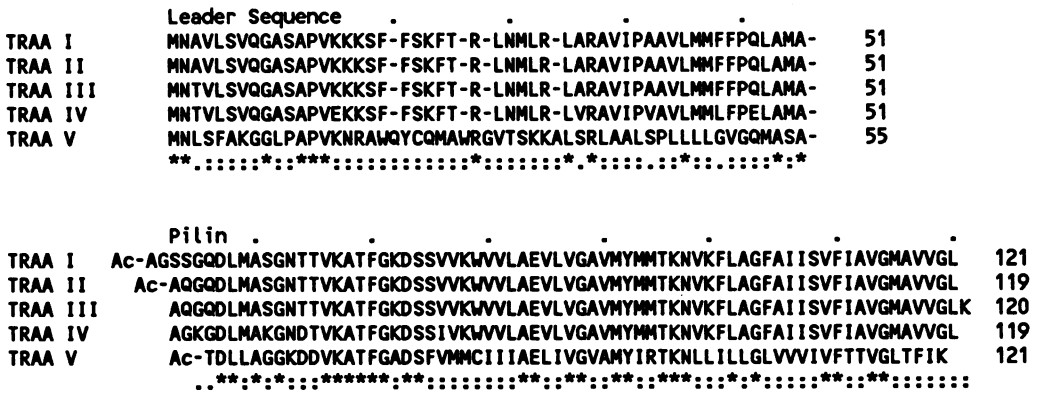


FIG. 11. Comparison of the five types of pilin encoded by *traA*. The sequences are taken from Frost et al. (79, 83) and Finlay et al. (63). The leader sequences and the mature pilin sequences are compared separately. The N termini of F (I), ColB2 (II), and pED208 (V) have been shown to be acetylated (Ac-) (60, 77, 83), while the N termini of R1-19 (III) and R100-1 (IV) have been determined to be blocked, but the presence of an acetyl group has not been demonstrated (79). The sequences are arranged in order of their similarity to the F pilin sequence, and the notation in the figure is the same as in Fig. 6.

(139) and by their pattern of F-specific phage attachment (79). Synthesis of the mature F-pilin subunit depends on the three *tra* products expressed by the widely separated genes *traA*, *traQ*, and *traX* (Fig. 1).

Product of the Pilin Gene, *traA*

The *traA* gene was originally shown by Minkley et al. (184) to encode the pilus subunit, pilin. The gene sequence revealed an ORF of 121 codons (nt 2249 to 2611) (83). The pilin subunit purified from F pili was determined to be a 7.2-kDa protein with an acetylated N terminus and a composition consistent with the last 70 aa in the sequence (83). These results corresponded well to other results demonstrating that the 13-kDa *traA* product (propilin) could be processed to a 7-kDa pilin polypeptide in cells expressing *traQ* (119, 191). The 51-aa propilin signal sequence (Table 2) is unusually long, although the sequence context for its cleavage appears typical of *E. coli* signal peptidase target sites. The sequence of the signal peptide is conserved among the five alleles of F pilin studied (Fig. 11) (79), and mutations that affect F pilus synthesis and function have been found to occur in the signal sequence as well as in the coding region (Fig. 2) (82). F pilin subunits appear to be assembled from an inner membrane pool, and the mature subunits are found in inner membrane fractions when pilus assembly is defective (137, 189, 190).

Serological differences and variations in phage sensitivity patterns exhibited by F-like pili were postulated to reflect differences in the sequence of the pilin subunit (139, 140, 178, 259, 264). These were determined to occur primarily at the N and C termini of pilin (79). The major antigenic determinant was discovered to be the acetylated N terminus of the pilin protein (64, 80). The N-terminal antigen is not exposed along the length of normal pilus filaments, and no evidence that it is exposed at the pilus tip has been provided (80, 268). Like propilin, the 7-kDa pilin polypeptide produced from clones expressing *traA* and *traQ* reacts poorly with polyclonal antiserum raised against purified pili (137). Since this 7-kDa product [Ap7(Q)] also migrated slightly faster than mature F pilin, it became clear that an additional F product was needed for N-terminal acetylation of F pilin (137, 271). Recent results demonstrate that the *traX* product is required for this reaction (186).

Mutations that affected RNA phage sensitivity fell into two classes. Those that affected R17 attachment were located near the beginning of the mature pilin protein (aa 11 to 22 in mature pilin, aa 62 to 73 in propilin) (Fig. 11), whereas those that strongly affected Qβ attachment were at the C terminus and involved a change in charge. The sequences with which filamentous phages interact are less clear-cut but may be near the N terminus (82).

Circular dichroism studies suggest that the mature pilin protein has a highly α-helical character (14, 41), and secondary-structure predictions suggest that the protein has two hydrophobic segments in addition to the first 26 residues at the N terminus. Fiber diffraction confirmed the α-helicity of the pilin protein but suggested that the helices were not necessarily parallel to the axis of the filament (71). Therefore the protein may form a hairpin with the residues near the C terminus exposed on the sides of the pilus and the internal basic residues (aa 97 and 100 [Fig. 11]), which interrupt the hydrophobic character of the protein, exposed in the lumen of the pilus. Immunological and Tn*phoA* fusion analysis of the orientation of pilin sequences in the cytoplasmic membrane have suggested the same structure for membrane pilin subunits with both N-terminal and C-terminal residues located on the periplasmic side (200). The N-terminal domain may be exposed on the distal surface of assembled pilins and, consequently, may be exposed at the pilus tip. However, the presence of another protein or substituent at the tip that acts as an adhesin has not been disproven.

Efforts to demonstrate that the F pilin subunit is posttranslationally modified with phosphate or glucose, as originally reported (26), have shown that these moieties, if present at all, are not covalently linked (15). Recently, electron spectroscopic imaging has shown that F and F-like pili are devoid of substantial amounts of phosphate (78). However, some additional modifications to pilin may occur, since a fraction of the subunits in F pili have been found to migrate at a retarded position on SDS-PAGE. Immunoblot analyses indicated that both the 8-kDa and the more prevalent 7-kDa subunits have the same acetylated N-terminal sequence, while kinetic studies demonstrated that the modification is added to the pilin polypeptide after signal sequence processing occurs (158, 186).

traQ Gene Product

The *traQ* sequence (nt 16585 to 16866) is predicted to encode a 94-aa inner membrane protein. This polypeptide includes two major hydrophobic regions (Ile-14 to Leu-32 and Ala-40 to Leu-60) and has a hydrophilic C-terminal region of the protein containing numerous charged amino acid residues (Table 1) (269). The *traQ* locus was identified as necessary for pilin maturation when various *traA*$^+$ Hfr deletion mutants were found to lack F pilin subunits (191). Subsequent experiments demonstrated that TraQ was required for efficient processing of propilin (119, 137, 271); F derivatives defective only in *traQ* expression are extremely transfer deficient and are unable to synthesize detectable quantities of F pilin or to elaborate F pili (127). The *traQ* product has been characterized, and current data are consistent with a model in which TraQ chaperones and/or translocates propilin residues through the cytoplasmic membrane; in the absence of *traQ*, the majority of the *traA* product is degraded (137, 154, 158, 269).

traX Gene Product

Since the pilin synthesized by lambda *traA* → *traH* transducing phages reacted poorly with F pilus antiserum unless an F$^+$ host was used, Laine et al. (137) suggested that N-terminal acetylation of pilin required an additional gene product. The isolation of two monoclonal antibodies provided a very specific assay for this gene activity: JEL93 recognizes the acetylated F pilin N-terminal sequence, whereas JEL92 reacts with internal pilin residues centered around Met-9 (aa 60 [Fig. 11]) (80). These were used to demonstrate that acetylation of F pilin is dependent on *traX* (186). Expression of cloned *traA*, *traQ*, and *traX* sequences suffices for synthesis of mature F pilin in vivo (158). An F derivative carrying a *traX* mutation was also found to elaborate filaments containing subunits with an unblocked N-terminal sequence that is otherwise identical to that in F pilin (159, 160). Despite previous speculation that *traX* might be essential for conjugation, this derivative is transfer proficient and confers sensitivity to pilus-specific phages. The structure of *traX* pili may differ somewhat from that of wild-type pili, however, since JEL92 antibodies were found to adsorb along the length of pilus filaments elaborated by a *traA* → *traG* clone in which wild-type pili did not adsorb this antibody laterally. Such pili were also defective in R17 phage attachment, and this defect could be complemented by clones containing the distal part of the transfer region (*traD* → *traX*) (99, 100). This suggests that this region, although not essential for transfer, may be important for correct pilus assembly.

The *traX* sequence appears to be highly conserved among F-like plasmids examined (35, 274). F *traX* (nt 31938 to 32681) is predicted to express a highly hydrophobic 248-aa, integral inner membrane protein with four transmembrane domains (35). However, the apparent molecular mass of the inner membrane proteins found to be expressed from *traX* clones is considerably smaller than the 27.5 kDa predicted for this product (155, 157).

Genes Involved in Pilus Assembly

Mutations affecting the expression of genes *traL*, *traE*, *traK*, *traB*, *traV*, *traC*, *traW*, *traU*, *traF*, *traH*, or *trbC* or the N-terminal region of *traG* affect piliation and drastically decrease mating efficiency (11, 12, 70, 105, 162–164, 167, 188, 192, 260, 270). Since mature F pilin subunits have been detected in inner membrane preparations from the mutant strains (186, 189), the products of these genes are assumed to be involved in F pilus assembly. In most cases (*traU* and *trbC* are exceptions), mutations that block the synthesis of one of these gene products also block expression of pilus filaments and result in complete resistance to pilus-specific phages. The nonsense and frameshift mutations in this class that have been sequenced are *traK4* (nt 3731), *traK105* (an additional C in the run at nt 4020 to 4023), *traV569*, *traW546* (nt 11152), and *traF13* (nt 15475) (Fig. 2) (164, 208, 270). The one missense mutation sequenced, *traC1044* (altering R to C at nt 10155), is of particular interest because it differs from *traC* nonsense mutations in that it does not totally block mating-pair formation or filamentous-phage infection (223, 224). Because no pilus filaments could be detected, Schandel et al. (223) suggested that *traC1044* blocks elongation of the pilus but allows the pilus tip to be assembled at the cell surface. Recently, homology between TraC and proteins apparently involved in protein secretion has been detected. TraC shares sequence similarity with the VirB4 protein of the Ti plasmid (232), which, in turn, has sequence similarity to TrbE (IncP), TrwK (IncW), and the PTorfD gene product of the *ptl* operon involved in production of *Bordetella pertussis* toxin (126, 230, 231, 255).

A complete lack of TrbC leads to a phenotype similar to that found for the *traC1044* mutant. Two pOX38 *trbC*::*kan* mutants, containing insertions either at *trbC Pvu*II (nt 12514) or between the *trbC Bsm*I sites (replacing nt 12688 to 12837), both retain some sensitivity to filamentous DNA phages, although they do not express detectable numbers of pili (163). In contrast, inactivation of *traU* reduces but does not severely alter the number of pili that can be expressed; however, a *traU* defect does significantly reduce sensitivity to both RNA and DNA pilus phages (188). Mutants in which a *kan* cassette has been introduced between the *Eco*RV sites in *traU* (replacing nt 11469 to 11840) express about 16% of the number of pili observed on a wild-type control.

Additional gene products that could be involved in pilus expression include those of *traP*, *trbI*, and *traD*, although mutations in these genes do not block pilus production. Kanamycin insertions in *traP* have a discernible effect on the efficiency of both transfer and phage infection (128). Those in *trbI* permit normal levels of plasmid transfer but reduce the efficiency of RNA and DNA phage infection. Interestingly, expression of TrbI in excess has a similar effect, and some *trbI*::*kan* mutants express extremely long pili, suggesting that the kinetics of pilus extension and retraction could be altered in these strains (164). As noted in a previous section, *traD* mutations block the entry of RNA during infection by phages in the f2, R17, and MS2 family and increase the number of pili expressed.

Characteristics of Pilus Assembly Proteins

The predicted size, location, and structure of products expressed by the pilus assembly genes are summarized in Table 1. Presumably, these products interact with one another to remove F pilin subunits from the inner membrane, polymerize them, and extend the F pilus from the cell surface. Sequence-based predictions, as well as other data concerning the locations of the proteins expected to be involved, suggest that such a complex would span the bacterial envelope.

TraC (99.2 kDa) (224, 225) is found in the cytoplasm in F$^-$ cells but fractionates into the inner membrane in F$^+$ cells, suggesting that it is peripherally associated with the inner membrane through interactions with other *tra* proteins. TraB (50.5 kDa) (74) is predicted to be predominantly hydrophilic, with a short hydrophobic segment (residues 12 to 34) which could target it to the inner membrane. Interestingly, TraB has a proline-rich character similar to the TrbI protein of RP4

(143), suggesting an extended conformation, possibly with periplasmic domains. TraL has a region with hydrophobic character (residues 31 to 61), which could also target it to the inner membrane (83). It has sequence similarity Tra2D (TrbD) and VirB3 (230, 232). TraE (21.2 kDa) (83), which may have an N-terminal membrane anchor (residues 12 to 39), has weak similarity to Tra2F (TrbF) of RP4 (74). TraG (102.5 kDa), which has multiple membrane-spanning segments, has been found in the inner membrane (70). TraL, TraE, TraB, TraC, and TraG are all required for piliation, so that interactions among these proteins seem likely. With the exception of TraG, all seem to have analogs in the RP4 Tra2 region involved in pilus synthesis. As mentioned above, a normal piliation phenotype requires that the inner membrane proteins TrbI (10 kDa) (164), TraD (82 kDa) (124, 274), and possibly TraP (22 kDa) (74), also predicted to be located in the inner membrane, be included in this array.

The mature products TraK (23.3 kDa) (208), TraW (21.7 kDa) (164), TrbC (21.2 kDa) (163), TraU (34.3 kDa) (188), TraF (25.9 kDa) (270), and TraH (47.8 kDa) (105) are expected to be periplasmic proteins. These products, which are also required for piliation, are all synthesized with signal sequences (Table 2) and are predicted to be soluble after processing. TraK is similar in this respect to the Tra2G (TrbG) protein of RP4, which is of similar size (297 aa) and has weak homology to F TraK (74). There are hydrophobic regions in TraU and TrbC that might permit peripheral membrane associations, and the possibility that some proteins in this group are outer membrane proteins has not been excluded. However, after synthesis in maxicells, the TraW, TrbC, TraU, and TraF proteins have been shown to concentrate in periplasmic fractions (163, 164, 188, 270).

One gene product required for piliation, TraV, is predicted to be synthesized as an 18.5-kDa product with a signal sequence (Table 2) that could undergo lipoprotein processing (51). Thus, TraV is suggested to be attached to the outer membrane. Although the size of the unmodified, processed TraV polypeptide is predicted to be 16.6 kDa, lipid modification should cause the observed polypeptide to be larger. Whether it then corresponds to the 21-kDa product that has been detected is unclear (192). RP4 also has a predicted lipoprotein, Tra2H (TrbH), of similar size (160 aa), although sequence comparisons have not detected any homology (74).

Since the elaboration of F pili is thought to be energy dependent, it is interesting that, in addition to the motifs in TraD mentioned previously, ATP/GTP-binding motifs were also identified in the sequences of both TraC (GTSGAGKT; nt 9180 to 9203) and TraH (GCTVGGKS; nt 18606 to 18629) (225, 249).

So far, all experimental data available concerning the location and processing of pilus assembly proteins have agreed with sequence-based predictions. The protein products assigned to this group have been identified (see the references in Table 1), and, with the exception of TraV, their observed sizes conform closely with predictions for the M_r of the polypeptide. As yet, no data concerning the interactions of such products with pilin subunits or with each other are available. Presumably, other similarly located *tra* region products that are not known to affect piliation could also be interacting with the assembly proteins.

MATING-AGGREGATE STABILIZATION PROTEINS

There is continuing discussion over the exact mechanism of conjugal DNA transfer from the donor to the recipient cell. Although there is limited evidence that the single strand of DNA can pass to the recipient via an extended pilus (106, 199), there is also considerable evidence that transfer ordinarily (or more usually) occurs between cells in close and specific contact. For discussions of these cellular interactions, the reader is referred to the articles by Achtman and Skurray (10) and Willetts and Skurray (265).

The specific contacts between mating pairs or mating aggregates (since more than two cells are often seen to be involved) (4) are thought to result from the interaction between the F pilus tip on the donor cell and a receptor on the recipient cell surface. The pilin subunits depolymerize into the membrane, drawing the cells together to form a conjugation "bridge" for the passage of DNA. Initial contacts between donor and recipient cells are unstable, but with time they become resistant to disruption by shearing or the addition of SDS (10). The most recent evidence for this model comes from studies with video microscopy and thin-section electron microscopy, which showed the formation of "conjugal junctions" between cells in close wall-to-wall contact (53).

Recent evidence (76) suggests that the pyrophosphorylethanolamine residue on the first heptose of the inner core of the lipopolysaccharide is involved in F mating-pair formation in liquid medium (109). The OmpA outer membrane protein in the recipient has also been implicated in mating-pair formation with F donors in liquid media, with a suggested involvement in either pilus attachment or mating-pair stabilization (9, 168, 169, 193). It is important to recall that neither the lipopolysaccharide nor OmpA is required for F mating on solid surfaces and that OmpA is not required for R100 mating either on solid or in liquid media (238). Thus the roles of OmpA and lipopolysaccharide in conjugation remain in question, and the evidence suggests that mating-pair formation may occur differently on solid and in liquid media.

The products of *traN* and *traG* are essential for DNA transfer and may stabilize mating contacts at an early stage of the interaction between donor and recipient cell surfaces; the Fl *actraN548* amber mutant and some *traG* mutants are able to produce F pili but are defective in stable mating-pair formation (167). It should be noted that the *traG* product, TraG, is bifunctional; it is involved both in the stabilization process and in F pilus assembly (70).

traN Gene Product

DNA sequence analysis combined with polypeptide studies of the *traN* product have suggested that it is a 65.7-kDa protein that undergoes cleavage of a signal sequence (nt 12997 to 14802 [Tables 1 and 2]) and is inserted into the outer membrane (161). The major portion of TraN remained intact after cells were exposed to proteinase K, suggesting that the bulk of the protein was involved in transmembrane segments and periplasmic domains. Internally located segments of TraN might interact with a periplasmic domain of TraG, the second product needed for stabilization. TraN is predicted to have an ATP-binding domain (ATGETGKT; nt 13246 to 13269) (161); it is not known whether mating-pair stabilization is ATP dependent. A *traN* amber mutation, first described by Miki et al. (180), has been sequenced (*traN548*; nt 13387) (161).

traG Gene Product

Predictions from nucleotide sequence analysis, along with polypeptide and protease protection studies, indicate that TraG is a 102.5-kDa protein which spans the inner membrane several times (nt 19403 to 22216 [Tables 1 and 2]), two large periplasmic domains are anchored to the inner membrane by several hydrophobic segments (70). Deletion analysis has

shown that only the N-terminal portion of TraG is essential for F pilus assembly; the C-terminal region encodes sequences essential for aggregate stabilization (8, 70). Moreover, internal cleavage of TraG has been proposed to account for the production of a 50-kDa periplasmic product, TraG*. This product is derived, approximately, from the C-terminal half of TraG and so includes the second large hydrophilic domain and possibly most or all of the sequences necessary for stabilization (70). Either the periplasmically exposed domains of the entire *traG* product or TraG* could interact directly with a periplasmic domain of TraN to stabilize mating-cell interactions.

The failure of a *traG* mutant plasmid to transfer DNA during solid surface matings, conditions which would have been expected to hold mating cells in close contact, and the very limited transfer of a *traN* mutant in these experiments (167) have led to the suggestion that TraG and TraN have a role in more than surface interaction (70, 161). It may well be that TraN and TraG form (part of) the channel or conjugation "bridge" for DNA transfer, since, together, their various domains extend from the cytoplasmic side of the inner membrane to the external surface.

SURFACE EXCLUSION

The surface exclusion mechanism reduces conjugation among cells bearing closely related plasmids of the same exclusion group. Five exclusion groups (Sfx_I to Sfx_V, F, ColB2-K98, R1, R100, and pED208 are the prototypes) have been identified (66, 259). Surface exclusion results from the presence of two *tra* proteins: TraT, an outer membrane protein that blocks the initial steps in mating-pair formation, and TraS, an inner membrane protein which is thought to block DNA transfer after a mating pair has been established (129). Experiments to determine the contribution of these two proteins, which act by very different mechanisms, have shown that TraS exhibits the stronger effect on surface exclusion (100- to 200-fold versus 10- to 20-fold for TraT) (5). TraT also confers serum resistance on bacteria, contributing to the virulence of the organism; the properties of TraT have been reviewed by Sukupolvi and O'Connor (241). The sequence of the F surface exclusion genes, *traS* (nt 22252 to 22698) and *traT* (nt 22796 to 23527), were determined by Jalajakumari et al. (123); the sequence of the R100 *traT* gene was determined by Ogata et al. (197); and the sequence of the ColB-K98 TraT protein was reported by Sukupolvi and O'Connor (241). The sequences of *traT* for F and pED208 were compared by Finlay and Paranchych (66), and the known sequences of TraT were compared by Finlay and Paranchych (66), Jalajakumari et al. (123), and Sukupolvi and O'Connor (241).

TraT Protein

The F *traT* gene product is predicted to be 26 kDa; it undergoes cleavage at the Cys residue at position 22 (Table 2) to give a polypeptide of 23.8 kDa during lipoprotein processing. The precursor protein is modified by the addition of a thioether-linked diglyceride to Cys-22, followed by the addition of two fatty acids to the diglycerides via ester linkages, signal peptidase II cleavage after residue Gly-21, and further acylation at the Cys-22 residue (the first amino acid in the mature protein) via an amide linkage (181, 209).

The difference between the F and R100 TraT sequences is a Gly-to-Ala transition at position 141 in the unprocessed protein, whereas ColB2-K98 has an additional Asn-to-Ser transition at position 136 (107). Recently it has been demonstrated that these small changes in protein sequence define the specificity of TraT in surface exclusion (107). The TraT protein may form a multimeric structure in the membrane (185), which has been visualized by electron microscopy as a five-membered ring of 18.5 nm diameter (182). Thus, the side chains on the amino acids that define TraT specificity may extend into the core of the ring and be recognized by the pilus tip of the homologous plasmid (107).

The TraT protein can be translated from a stable mRNA (145), which is transcribed in a TraJ-independent manner from a weak promoter within the *traS* gene (104, 123, 212, 213). The hairpin structures predicted at the 3′ and 5′ ends of the *traT* transcript (nt 22708 to 22747 and 23569 to 23663) presumably stabilize this mRNA. The former hairpin was initially thought to be a terminator for *traS* transcription (123); *traS* is transcribed from a strong promoter, and *traT* expression may depend on transcription from that promoter with incomplete termination at this hairpin between *traS* and *traT* (104). The F *traT* protein is easily visualized in membranes from F^+ cells which contain about 25,000 copies per cell (5, 165, 183). Overexpression of TraT from a multicopy plasmid increases the level of surface exclusion, suggesting that the effectiveness of TraT is concentration dependent (5).

The TraT protein has two hydrophobic segments (Table 1) which could span the outer membrane; the protein is extremely stable and resists heat denaturation, proteolytic digestion, and solubilization by detergents (165, 185). *traT* mutants with altered permeability properties have been useful in dissecting the permeability barrier in gram-negative bacteria as well as the role of TraT in virulence (211, 241).

The precise role of TraT in surface exclusion is unclear. That an interaction with pili might be involved in surface exclusion was first suggested by Meynell and Ewins (177), who showed that surface exclusion was affected by the production of mixed pili in a cell. Similarly, plasmids in the same exclusion group also usually have the same *traA* alleles, although exceptions do occur (259). Riede and Eschbach (220) have proposed that TraT blocks access to the outer membrane protein OmpA, one of the proposed receptors for the pilus tip, by masking the pilus-binding site. However, purified F pili have been found to bind normally to recipient cells carrying an interrupted *ompA* gene (76). Minkley and Willetts (185) have proposed that the pilus tip binds TraT; thus, there would be competition between TraT and the pilus receptor for the pilus tip, leading to reduced levels of transfer. Purified TraT protein is also capable of specifically blocking transfer; purified F TraT blocks mating by F more efficiently than by R100 (185). Similarly, the R6-5 (Sfx_{IV}) TraT protein blocks transfer by the repressed plasmid R6-5 more efficiently than by the derepressed plasmid R100-1 (Sfx_{IV}), suggesting that the level of piliation may contribute to the efficacy of the surface exclusion system (107); the *traA* sequences of R100-1 and R6-5 are identical (236, 241). Recent evidence suggests that changes in sequence at the N terminus of the pilin subunit do not alter the specificity of the surface exclusion reaction, suggesting that another protein, possibly an adhesin at the pilus tip, is involved in TraT recognition (76).

TraS Protein

The mechanism of action of TraS remains an enigma; expression of TraS in the recipient cell does not prevent mating-aggregate accumulation but does prevent donor DNA synthesis and DNA transfer, suggesting that TraS blocks a signal transduced from the recipient to the donor cell (5, 7, 129). Surface exclusion, mediated by TraS in the recipient cell, also appears to require an intact peptidoglycan in the donor cell, suggesting a role for the cell wall in signal transduction.

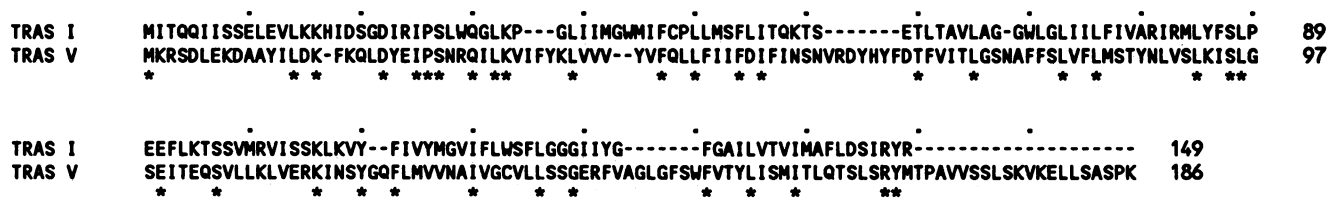


FIG. 12. Comparison of the sequences for TraS from F (I) (123) and pED208 (V) (66). Identical residues are marked with an asterisk.

TraS, which has three or four hydrophobic segments (Table 1), is found in the inner membrane (123). TraS proteins (Fig. 12) may differ in sequence among F-like plasmids, although the only *traS* genes sequenced to date are from the distantly related plasmids F and pED208 (66). Heteroduplex mapping analysis of F-like plasmids (229) suggested a region of nonhomology which approximately coincided with the position of the *traS* gene. Since the *tra* proteins TraM and TraY differ significantly among the F-like plasmids (259), TraS may also differ in sequence among the F-like plasmids and interact with other plasmid-specific *tra* components to interfere with the mating signal that initiates transfer (167). More sequence information on the *traS* gene from plasmids belonging to other exclusion groups may prove informative, and information on its mechanism of action may shed light on the elusive "signal" that initiates transfer.

OTHER *tra* REGION PRODUCTS

The functions of a number of the ORFs identified in Fig. 1 and 2 remain unclear. In addition to *traP* and *trbI*, which were discussed together with pilus assembly products, this group includes *traR*, *trbABDEFGHJ*, and the anti-*tra* ORF, *artA*. As indicated in Table 1, computer analysis of these sequences suggests that their products would include two cytoplasmic proteins (TrbD and TrbG) (51, 74); six inner membrane proteins (TrbA, TrbE, TrbF, TrbH, TrbJ, and ArtA) (104, 156, 161, 269, 271); and, after signal sequence processing, one periplasmic protein, TrbB (Table 2) (271). It seems reasonable to expect that such products would interact with other *tra* proteins in the same subcellular location. However, characterization of this group is still at a preliminary stage. Products have not yet been observed for *trbD*, *trbF*, *trbG*, *trbH*, or *artA*, although an in-frame *artA*::*lacZ* fusion has been shown to be expressed (269). Otherwise, products that correspond in size to the loci have been identified. The small TrbA (12.9-kDa) and TrbE (9.9-kDa) products are known to be inner membrane proteins, and TrbB has been observed to undergo signal peptide processing and to localize in the periplasm (161, 269, 271). The TraR and TrbJ proteins have been identified but not localized (156, 192).

Analysis of *kan* cassette insertion mutations affecting this group of gene products has not yet resolved the question of their functions. Insertion mutations in *traR*, *trbA*, *trbB*, *trbE*, *trbH*, *trbJ*, and *artA* did not significantly alter the efficiency of transfer or pilus-specific phage infection in standard *E. coli* experiments, although these genes could make a more important contribution under other growth or mating conditions (127, 156, 161). Insertions into the *trbD* and *trbG* sequences result in a mutant phenotype, but this may not reflect their individual functions; because the defects cannot be complemented by expression of the genes in *trans*, it is possible that disruption of these sequences affects expression of other *tra* region proteins (187).

Conclusions

The F transfer region could be described as a collage of signaling and export-import systems, designed to allow plasmid transfer to suitable recipient cells under optimal environmental conditions. Pilus assembly and retraction represent two opposing systems for pilin protein export and reabsorption; the equilibrium between these two processes could be disturbed by contact with recipient cells or phage particles or in response to changes in physiological conditions (e.g., changing temperature, entry into stationary phase). This perturbation of the system could trigger pilus retraction, DNA transfer, changes in *tra* gene expression, or a combination of all three responses. The transfer of the DNA is facilitated by a transport system specific for nucleic acids, and similarities in protein sequence and function of the *tra* proteins involved appear to be conserved among the transfer regions studied to date. Overlying these macromolecular transport systems is a network of intercellular signaling mechanisms which do not have obvious similarity to other procaryotic signal transduction systems. Upon identification of a suitable recipient cell, a signal appears to be generated; it is then passed through the pilus into the donor and results in the transfer of the DNA. The TraS protein of the F surface exclusion system either interferes with this signal or generates a new signal that a nonproductive mating pair has formed. A study of the mechanism of action of this protein could be very informative about this remarkably precise and efficient signaling process.

The unique structure of the mating junctions as visualized by electron microscopy (53) suggests ample possibilities for the exchange of signals between the mating cells and could involve many components of the cell envelope. Whether the F transfer operon is solely responsible for construction of the mating bridge or whether it coordinates interactions between components of the envelopes of the partners in the mating pair is an intriguing aspect of this process. Although OmpA and lipopolysaccharide in the recipient have been implicated in mating-pair formation, no mutation in the recipient cell has resulted in a completely conjugation-deficient recipient. This suggests that such a mutation either would be lethal or could be circumvented by increased participation by other components of the junction. It is possible to imagine the construction of a transfer-specific pore spanning the envelope of the donor cell to facilitate DNA transport, but the path the DNA takes once it enters the recipient cell is less obvious. Does the uptake of the DNA require a constitutively expressed import system in the recipient cell, or does the transfer machinery expand into the recipient and inject the DNA into the cytoplasm in a manner reminiscent of the contractile-tailed bacteriophages? The mechanism of mating-pair formation and DNA transfer in intergeneric and interkingdom crosses is even more puzzling. It will be of interest to know whether these unusual conjugative events require the *tra* genes for pilus and mating-pair formation or whether the DNA transfer machinery takes advantage

of cell-to-cell contacts that normally occur at low frequency and have not been previously described.

As information accumulates on the function of the *tra* gene products in the F-like plasmids and other transfer systems in the gram-negative world, similarities in the process of conjugation can be discerned and used to discover the nature of mating-pair formation and precise DNA transfer. On the basis of the two criteria of serological relatedness and shared sensitivity to bacteriophages, the pili of the various transfer systems, which are usually associated with a particular incompatibility group, can be divided into two families: those that are expressed by the group of IncP, IncN, IncU, IncW, IncX, IncM, and IncI plasmids and those that are members of the IncF, IncC, IncD, IncJ, and IncS plasmids (75). Plasmids of the IncH and IncT groups cannot be easily categorized within these two families as yet, whereas IncI plasmids elaborate two types of pili belonging to both families. These similarities are holding true at the genotypic level, where DNA sequence information for plasmids from the IncP family suggests that they are more closely related to each other than to F. Currently, limited DNA sequence information is available for only a few members of the IncF family that are closely related to F. Eventually, it may be possible to distinguish whether there is a continuum of genetic variation throughout the plasmids of gram-negative bacteria whereby F (IncF) and RP4 (IncP) represent distantly related transfer systems or whether F and RP4 have evolved from different ancestors and share only a few processes in common (for pilus assembly and DNA nicking and transfer). This continuum may be extended to include mating systems from gram-positive bacteria as homology with these systems continues to be uncovered (69).

Mating-pair formation may occur by distinct mechanisms in the IncF and IncP families; genes for stabilization have not been detected in the IncP systems, as demonstrated by their preference to mate on solid surfaces. Although a conjugative pilus is expressed by both families, the absolute requirement for a pilus during conjugation has been demonstrated only for F and its relatives. It remains to be seen whether the RP4 pilus is involved in recipient-cell recognition, mating-pair formation, or DNA conduction, functions that have been ascribed to the F pilus.

An examination of conjugative plasmids from gram-negative bacteria reveals a greater degree of similarity in sequence and gene function within the leading region of plasmids of different incompatibility groups; this region may have an important role in establishing the plasmid in the recipient cell. Very few *tra* (or host) proteins appear to be transferred from the donor to the recipient cell during conjugation; only the primase proteins of the IncP and IncI plasmids RP4 and ColIb-P9 (216, 217) have been detected in the recipient, although definitive experiments for other systems have not been done. Thus, the success of the plasmid transfer event could well depend on the expression of the arsenal of genes expressed immediately at the onset of DNA transfer from the leading region as it enters the recipient cell.

Although the number of gene products encoded by the F transfer region, as well as their physical attributes and their location in the cell, is, for the most part, known, the function of the vast majority of these proteins and their role in DNA transfer and/or pilus formation remain to be discovered. The F transfer system shares many of the themes associated with phage infection. The pilus resembles a filamentous phage coat; the DNA transfer event shares mechanistic similarities with ϕX174 replication and packaging. However, the act of bacterial conjugation also represents a simple example of cellular recognition and communication between cells in that a signal generated at the cell surface in the recipient cell is transmitted to the transferosome in the membrane of another cell, initiating DNA transfer. Thus, the study of bacterial conjugation should lead to insight on the complex subject of cell-cell surface interactions and the coordination of events in the cell envelope with the timing of replication initiation.

ACKNOWLEDGMENTS

We thank Neville Firth, Sumit Maneewannakul, Ken Usher, and Tim van Biesen for many valuable discussions and suggestions, computer analyses, and help in compiling the figures. We also thank Fernando de la Cruz, Clarence Kado, Erich Lanka, Steve Matson, and Brian Wilkins for sharing unpublished data and for useful discussions.

REFERENCES

1. **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. USA **80:**4659–4663.
2. **Abo, T., S. Inamoto, and E. Ohtsubo.** 1991. Specific DNA binding of the TraM protein to the *oriT* region of plasmid R100. J. Bacteriol. **173:**6347–6354.
3. **Abo, T., and E. Ohtsubo.** 1993. Repression of the *traM* gene of plasmid R100 by its own product and integration host factor at one of the two promoters. J. Bacteriol. **175:**4466–4474.
4. **Achtman, M.** 1975. Mating aggregates in *Escherichia coli* conjugation. J. Bacteriol. **123:**505–515.
5. **Achtman, M., N. Kennedy, and R. Skurray.** 1977. Cell-cell interactions in conjugating *Escherichia coli*: role of the *traT* protein in surface exclusion. Proc. Natl. Acad. Sci. USA **74:**5104–5108.
6. **Achtman, M., P. A. Manning, C. Edelbluth, and P. Herrlich.** 1979. Export without proteolytic processing of inner and outer membrane proteins encoded by F sex factor *tra* cistrons in *Escherichia coli* minicells. Proc. Natl. Acad. Sci. USA **76:**4837–4841.
7. **Achtman, M., P. A. Manning, B. Kusecek, S. Schwuchow, and N. Willetts.** 1980. A genetic analysis of F sex factor cistrons needed for surface exclusion in *Escherichia coli*. J. Mol. Biol. **138:**779–795.
8. **Achtman, M., G. Morelli, and S. Schwuchow.** 1978. Cell-cell interactions in conjugating *Escherichia coli*: role of F pili and fate of mating aggregates. J. Bacteriol. **135:**1053–1061.
9. **Achtman, M., S. Schwuchow, R. Helmuth, G. Morelli, and P. A. Manning.** 1978. Cell-cell interactions in conjugating *Escherichia coli*: Con$^-$ mutants and stabilization of mating aggregates. Mol. Gen. Genet. **164:**171–183.
10. **Achtman, M., and R. A. Skurray.** 1977. A redefinition of the mating phenomenon in bacteria, p. 233–279. *In* J. L. Reissig (ed.), Microbial interactions: receptors and recognition. Chapman & Hall, Ltd., London.
11. **Achtman, M., N. S. Willetts, and A. J. Clark.** 1971. Beginning a genetic analysis of conjugational transfer determined by the F factor in *Escherichia coli* by isolation and characterization of transfer-deficient mutants. J. Bacteriol. **106:**529–538.
12. **Achtman, M., N. S. Willetts, and A. J. Clark.** 1972. Conjugational complementation analysis of transfer-deficient mutants of *Flac* in *Escherichia coli*. J. Bacteriol. **110:**831–842.
13. **Alifano, P., F. Rivellini, D. Limauro, C. B. Bruni, and M. S. Carlomagno.** 1991. A consensus motif common to all Rho-dependent prokaryotic transcription terminators. Cell **64:**553–563.
14. **Armstrong, G. D., L. S. Frost, P. A. Sastry, and W. Paranchych.** 1980. Comparative biochemical studies on F and EDP208 conjugative pili. J. Bacteriol. **141:**333–341.
15. **Armstrong, G. D., L. S. Frost, H. J. Vogel, and W. Paranchych.** 1981. Nature of the carbohydrate and phosphate associated with ColB2 and EDP208 pilin. J. Bacteriol. **145:**1167–1176.
16. **Bagdasarian, M., A. Bailone, J. Angulo, P. Scholz, and R. Devoret.** 1992. PsiB, an anti-SOS protein, is transiently expressed by the F sex factor during its transmission to an *Escherichia coli*

K-12 recipient. Mol. Microbiol. **6**:885–894.
17. **Bagdasarian, M., A. Bailone, M. M. Bagdasarian, P. A. Manning, R. Lurz, K. N. Timmis, and R. Devoret.** 1986. An inhibitor of SOS induction specified by a plasmid locus in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **83**:5723–5726.
18. **Bailone, A., A. Backman, S. Sommer, J. Celerier, M. M. Bagdasarian, M. Bagdasarian, and R. Devoret.** 1988. PsiB polypeptide prevents activation of RecA protein in *Escherichia coli*. Mol. Gen. Genet. **214**:389–395.
19. **Bairoch, A.** 1992. PROSITE: a dictionary of sites and patterns in proteins. Nucleic Acids Res. **20**:2013–2018.
20. **Bayer, G., D. Aigner, K. Eferl, G. Koraimann, C. Schroller, and G. Hogenauer.** 1993. Gene *19*, a novel transfer gene in resistance plasmid R1, abstr. 96, p. 177. Proc. Seventeenth Int. Congr. Genet.
21. **Benz, I., and H. Müller.** 1990. *Escherichia coli* DNA helicase I. Characterization of the protein and of its DNA-binding properties. Eur. J. Biochem. **189**:267–276.
22. **Beutin, L., P. A. Manning, M. Achtman, and N. Willetts.** 1981. *sfrA* and *sfrB* products of *Escherichia coli* K-12 are transcriptional control factors. J. Bacteriol. **145**:840–844.
23. **Bowie, J. U., and R. T. Sauer.** 1990. TraY proteins of F and related episomes are members of the Arc and Mnt repressor family. J. Mol. Biol. **211**:5–6.
24. **Boyd, A. C., J. A. K. Archer, and D. J. Sherratt.** 1989. Characterization of the ColE1 mobilization region and its protein products. Mol. Gen. Genet. **217**:488–498.
25. **Bradshaw, H. D., Jr., B. A. Traxler, E. G. Minkley, Jr., E. W. Nester, and M. P. Gordon.** 1990. Nucleotide sequence of the *traI* (helicase) gene from the sex factor F. J. Bacteriol. **172**:4127–4131.
26. **Brinton, C. C., Jr.** 1971. The properties of sex pili, the viral nature of "conjugal" genetic transfer systems, and some possible approaches to the control of drug resistance. Crit. Rev. Microbiol. **1**:105–160.
27. **Buxton, R. S., and L. S. Drury.** 1983. Cloning and insertional inactivation of the *dye*(*sfrA*) gene, mutation of which affects sex factor F expression and dye sensitivity of *Escherichia coli*. J. Bacteriol. **154**:1309–1314.
28. **Carter, J. R., and R. D. Porter.** 1991. *traY* and *traI* are required for *oriT*-dependent enhanced recombination between *lac*-containing plasmids and λp*lac5*. J. Bacteriol. **173**:1027–1034.
29. **Chase, J. W., B. M. Merrill, and K. R. Williams.** 1983. F sex factor encodes a single-stranded DNA binding protein (SSB) with extensive sequence homology to *Escherichia coli* SSB. Proc. Natl. Acad. Sci. USA **80**:5480–5484.
30. **Chatfield, L. K., E. Orr, G. J. Boulnois, and B. M. Wilkins.** 1982. DNA primase of plasmid ColIb is involved in conjugal DNA synthesis in donor and recipient bacteria. J. Bacteriol. **152**:1188–1195.
31. **Cheah, K.-C., and R. A. Skurray.** 1986. The F plasmid carries an *IS3* insertion within *finO*. J. Gen. Microbiol. **132**:3269–3275.
32. **Close, S. M., and C. I. Kado.** 1992. A gene near the plasmid pSa origin of replication encodes a nuclease. Mol. Microbiol. **6**:521–527.
33. **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 T-region borders. J. Bacteriol. **174**:6238–6246.
34. **Cotton, F. A., E. E. Hazen, Jr., and M. J. Legg.** 1979. Staphylococcal nuclease: proposed mechanism of action based on structure of enzyme-thymidine 3′,5′-bisphosphate-calcium ion complex at 1.5 A resolution. Proc. Natl. Acad. Sci. USA **76**:2551–2555.
35. **Cram, D. S., S. M. Loh, K.-C. Cheah, and R. A. Skurray.** 1991. Sequence and conservation of genes at the distal end of the transfer region on plasmids F and R6-5. Gene **104**:85–90.
36. **Cram, D. S., A. Ray, L. O'Gorman, and R. Skurray.** 1984. Transcriptional analysis of the leading region in F plasmid DNA transfer. Plasmid **11**:221–233.
37. **Crothers, D. M., T. E. Haran, and J. G. Nadeau.** 1990. Intrinsically bent DNA. J. Biol. Chem. **265**:7093–7096.
38. **Cuozzo, M., and P. M. Silverman.** 1986. Characterization of the F plasmid TraJ protein synthesized in F′ and Hfr strains of *Escherichia coli* K-12. J. Biol. Chem. **261**:5175–5179.
39. **Cuozzo, M., P. M. Silverman, and E. G. Minkley, Jr.** 1984. Overproduction in *Escherichia coli* K-12 and purification of the TraJ protein encoded by the conjugative plasmid F. J. Biol. Chem. **259**:6659–6666.
40. **Dash, P. K., B. A. Traxler, M. M. Panicker, D. D. Hackney, and E. G. Minkley, Jr.** 1992. Biochemical characterization of *Escherichia coli* DNA helicase I. Mol. Microbiol. **6**:1163–1172.
41. **Date, T., M. Inuzuka, and M. Tomoeda.** 1978. Structure of F pili in *Escherichia coli*, p. 301–318. *In* D. E. Bradley, E. Raizen, P. Fives-Taylor, and J. Ou (ed.), Pili. International Conferences on Pili, Washington, D.C.
42. **De La Cruz, F. (University of Cantabria, Spain).** 1992. Personal communication.
43. **Dempsey, W. B.** 1987. Transcript analysis of the plasmid R100 *traJ* and *finP* genes. Mol. Gen. Genet. **209**:533–544.
44. **Dempsey, W. B.** 1987. Integration host factor and conjugative transfer of the antibiotic resistance plasmid R100. J. Bacteriol. **169**:4391–4392.
45. **Dempsey, W. B.** 1989. Sense and antisense transcripts of *traM*, a conjugal transfer gene of the antibiotic resistance plasmid R100. Mol. Microbiol. **3**:561–570.
46. **Dempsey, W. B., and B. E. Fee.** 1990. Integration host factor affects expression of two genes at the conjugal transfer origin of plasmid R100. Mol. Microbiol. **4**:1019–1028.
47. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.
48. **Di Laurenzio, L., B. B. Finlay, L. S. Frost, and W. Paranchych.** 1991. Characterization of the *oriT* region of the IncFV plasmid pED208. Mol. Microbiol. **5**:1779–1790.
49. **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.
50. **Di Laurenzio, L., L. S. Frost, D. S. Scraba, and W. Paranchych.** Unpublished data.
51. **Doran, T., S. Loh, N. Firth, and R. Skurray.** Unpublished data.
52. **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.
53. **Durrenberger, M. B., W. Villiger, and T. Bachi.** 1991. Conjugational junctions: morphology of specific contacts in conjugating *Escherichia coli* bacteria. J. Struct. Biol. **107**:146–156.
54. **Dutreix, M., A. Backman, J. Celerier, M. M. Bagdasarian, S. Sommer, A. Bailone, R. Devoret, and M. Bagdasarian.** 1988. Identification of *psiB* genes of plasmids F and R6-5. Molecular basis for *psiB* enhanced expression in plasmid R6-5. Nucleic Acids Res. **16**:10669–10679.
55. **Ebright, R. H., P. Cossart, B. Gicquel-Sanzey, and J. Beckwith.** 1984. Mutations that alter the DNA sequence specificity of the catabolite gene activator protein of *E. coli*. Nature (London) **311**:232–235.
56. **Eguchi, Y., T. Itoh, and J.-I. Tomizawa.** 1991. Antisense RNA. Annu. Rev. Biochem. **60**:631–652.
57. **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.
58. **Fee, B. E., and W. B. Dempsey.** 1986. Cloning, mapping, and sequencing of plasmid R100 *traM* and *finP* genes. J. Bacteriol. **167**:336–345.
59. **Fee, B. E., and W. B. Dempsey.** 1988. Nucleotide sequence of *geneX* of antibiotic resistance plasmid R100. Nucleic Acids Res. **16**:4726.
60. **Finlay, B. B., L. S. Frost, and W. Paranchych.** 1984. Localization, cloning, and sequence determination of the conjugative plasmid ColB2 pilin gene. J. Bacteriol. **160**:402–407.
61. **Finlay, B. B., L. S. Frost, and W. Paranchych.** 1986. Origin of transfer of IncF plasmids and nucleotide sequences of the type II *oriT*, *traM*, and *traY* alleles from ColB4-K98 and the type IV *traY* allele from R100-1. J. Bacteriol. **168**:132–139.
62. **Finlay, B. B., L. S. Frost, and W. Paranchych.** 1986. Nucleotide

sequences of the R1-19 plasmid transfer genes *traM*, *finP*, *traJ*, and *traY* and the *traYZ* promoter. J. Bacteriol. **166**:368–374.
63. **Finlay, B. B., L. S. Frost, and W. Paranchych.** 1986. Nucleotide sequence of the *traYALE* region from IncFV plasmid pED208. J. Bacteriol. **168**:990–998.
64. **Finlay, B. B., L. S. Frost, W. Paranchych, J. M. Parker, and R. S. Hodges.** 1985. Major antigenic determinants of F and ColB2 pili. J. Bacteriol. **163**:331–335.
65. **Finlay, B. B., L. S. Frost, W. Paranchych, and N. S. Willetts.** 1986. Nucleotide sequences of five IncF plasmid *finP* alleles. J. Bacteriol. **167**:754–757.
66. **Finlay, B. B., and W. Paranchych.** 1986. Nucleotide sequence of the surface exclusion genes *traS* and *traT* from the IncF$_0$*lac* plasmid pED208. J. Bacteriol. **166**:713–721.
67. **Finnegan, D. J., and N. S. Willetts.** 1971. Two classes of F*lac* mutants insensitive to transfer inhibition by an F-like R factor. Mol. Gen. Genet. **111**:256–264.
68. **Finnegan, D. J., and N. S. Willetts.** 1973. The site of action of the F transfer inhibitor. Mol. Gen. Genet. **127**:307–316.
69. **Firth, N., K. P. Ridgway, M. E. Byrne, P. D. Fink, L. Johnson, I. T. Paulsen, and R. A. Skurray.** Analysis of a transfer region from the staphylococcal conjugative plasmid pSK41. Gene, in press.
70. **Firth, N., and R. Skurray.** 1992. Characterization of the F plasmid bifunctional conjugation gene, *traG*. Mol. Gen. Genet. **232**:145–153.
71. **Folkhard, W., K. R. Leonard, S. Malsey, D. A. Marvin, J. Dubochet, A. Engel, M. Achtman, and R. Helmuth.** 1979. X-ray diffraction and electron microscope studies on the structure of bacterial F pili. J. Mol. Biol. **130**:145–160.
72. **Fowler, T., L. Taylor, and R. Thompson.** 1983. The control region of the F plasmid transfer operon: DNA sequence of *traJ* and *traY* genes and characterisation of the *traY-Z* promoter. Gene **26**:79–89.
73. **Fowler, T., and R. Thompson.** 1986. Shadow promoters in the F plasmid transfer operon. Mol. Gen. Genet. **202**:509–511.
74. **Frost, L. S.** Unpublished data.
75. **Frost, L. S.** 1993. Conjugative pili and pilus-specific phages, p. 189–221. *In* D. B. Clewell (ed.), Bacterial conjugation. Plenum Press, New York.
76. **Frost, L. S., K. G. Anthony, C. Sherburne, and R. Sherburne.** Unpublished data.
77. **Frost, L. S., G. D. Armstrong, B. B. Finlay, B. F. Edwards, and W. Paranchych.** 1983. N-terminal amino acid sequencing of EDP208 conjugative pili. J. Bacteriol. **153**:950–954.
78. **Frost, L. S., and D. P. Bazett-Jones.** 1991. Examination of the phosphate in conjugative F-like pili by use of electron spectroscopic imaging. J. Bacteriol. **173**:7728–7731.
79. **Frost, L. S., B. B. Finlay, A. Opgenorth, W. Paranchych, and J. S. Lee.** 1985. Characterization and sequence analysis of pilin from F-like plasmids. J. Bacteriol. **164**:1238–1247.
80. **Frost, L. S., J. S. Lee, D. G. Scraba, and W. Paranchych.** 1986. Two monoclonal antibodies specific for different epitopes within the amino-terminal region of F pilin. J. Bacteriol. **168**:192–198.
81. **Frost, L. S., S. H. Lee, N. Yanchar, and W. Paranchych.** 1989. *finP* and *fisO* mutations in FinP anti-sense RNA suggest a model for FinOP action in the repression of bacterial conjugation by the F*lac* plasmid JCFL0. Mol. Gen. Genet. **218**:152–160.
82. **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.
83. **Frost, L. S., W. Paranchych, and N. S. Willetts.** 1984. DNA sequence of the F *traALE* region that includes the gene for F pilin. J. Bacteriol. **160**:395–401.
84. **Frost, L. S., K. Usher, and W. Paranchych.** 1991. Computer analysis of the F transfer region. Plasmid **25**:226.
85. **Fu, F. Y.-H., M.-M. Tsai, Y. Luo, and R. C. Deonier.** 1991. Deletion analysis of the F plasmid *oriT* locus. J. Bacteriol. **173**:1012–1020.
86. **Furuya, N., and T. Komano.** 1991. Determination of the nick site at *oriT* of IncI1 plasmid R64: global similarity of *oriT* structures of IncI1 and IncP plasmids. J. Bacteriol. **173**:6612–6617.
87. **Gaffney, D., R. Skurray, and N. Willetts.** 1983. Regulation of the F conjugation genes studied by hybridization and *tra-lacZ* fusion. J. Mol. Biol. **168**:103–122.
88. **Gamas, P., L. Caro, D. Galas, and M. Chandler.** 1987. Expression of F transfer functions depends on the *E. coli* integration host factor. Mol. Gen. Genet. **207**:302–305.
89. **Gao, Q., Y. Luo, and R. Deonier.** 1994. Initiation and termination of DNA transfer at F plasmid *oriT*. Mol. Microbiol. **11**:459–470.
90. **Gaudin, H. M., and P. M. Silverman.** 1993. Contributions of promoter context and structure to regulated expression of the F plasmid *traY* promoter in *Escherichia coli* K-12. Mol. Microbiol. **8**:335–342.
91. **Gerdes, K., A. Nielsen, P. Thorsted, and E. G. H. Wagner.** 1992. Mechanism of killer gene activation. Antisense RNA-dependent RNase III cleavage ensures rapid turnover of the stable Hok, SrnB and PndA effector messenger RNAs. J. Mol. Biol. **226**:637–649.
92. **Gerdes, K., T. Thisted, and J. Martinussen.** 1990. Mechanism of post-segregational killing by the *hok/sok* system of plasmid R1: *sok* antisense RNA regulates formation of a *hok* mRNA species correlated with killing of plasmid-free cells. Mol. Microbiol. **4**:1807–1818.
93. **Gerlitz, M., O. Hrabak, and H. Schwab.** 1990. Partitioning of broad-host-range plasmid RP4 is a complex system involving site-specific recombination. J. Bacteriol. **172**:6194–6203.
94. **Golub, E. I., A. Bailone, and R. Devoret.** 1988. A gene encoding an SOS inhibitor is present in different conjugative plasmids. J. Bacteriol. **170**:4392–4394.
95. **Golub, E. I., and K. B. Low.** 1986. Unrelated conjugative plasmids have sequences which are homologous to the leading region of the F factor. J. Bacteriol. **166**:670–672.
96. **Golub, E. I., and H. A. Panzer.** 1988. The F factor of *Escherichia coli* carries a locus of stable plasmid inheritance *stm*, similar to the *parB* locus of plasmid R1. Mol. Gen. Genet. **214**:353–357.
97. **Graus, H., A. Hodl, P. Wallner, and G. Hogenauer.** 1990. The sequence of the leading region of the resistance plasmid R1. Nucleic Acids Res. **18**:1046.
98. **Graus-Goldner, A., H. Graus, T. Schlacher, and G. Hogenauer.** 1990. The sequences of genes bordering *oriT* in the enterotoxin plasmid P307: comparison with the sequences of plasmids F and R1. Plasmid **24**:119–131.
99. **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.
100. **Grossman, T. H., and P. M. Silverman.** 1989. Structure and function of conjugative pili: inducible synthesis of functional F pili by *Escherichia coli* K-12 containing a *lac-tra* operon fusion. J. Bacteriol. **171**:650–656.
101. **Guiney, D. G.** 1982. Host range of conjugation and replication functions of the *Escherichia coli* sex plasmid F′ *lac*: comparison with the broad-host-range plasmid RK2. J. Mol. Biol. **162**:699–703.
102. **Guiney, D. G.** 1993. Broad host range conjugative and mobilizable plasmids in Gram-negative bacteria, p. 75–103. *In* D. B. Clewell (ed.), Bacterial conjugation. Plenum Press, New York.
103. **Guyer, M. S., R. R. Reed, J. A. Steitz, and K. B. Low.** 1980. Identification of a sex-factor affinity site in *E. coli* as gamma-delta. Cold Spring Harbor Symp. Quant. Biol. **45**:135–140.
104. **Ham, L. M., D. Cram, and R. Skurray.** 1989. Transcriptional analysis of the F plasmid surface exclusion region: mapping of *traS*, *traT*, and *traD* transcripts. Plasmid **21**:1–8.
105. **Ham, L. M., N. Firth, and R. Skurray.** 1989. Nucleotide sequence of the F plasmid transfer gene, *traH*: identification of a new gene and a promoter within the transfer operon. Gene **75**:157–165.
106. **Harrington, L. C., and A. C. Rogerson.** 1990. The F pilus of *Escherichia coli* appears to support stable DNA transfer in the absence of wall-to-wall contact between cells. J. Bacteriol. **172**:7263–7264.
107. **Harrison, J. L., I. M. Taylor, K. Platt, and C. D. O'Connor.** 1992. Surface exclusion specificity of the TraT lipoprotein is determined by single alterations in a five amino acid region of the protein. Mol. Microbiol. **6**:2825–2832.

108. **Harwood, C. R., and E. Meynell.** 1975. Cyclic AMP and the production of sex pili by *E. coli* K-12 carrying derepressed sex factors. Nature (London) **254:**628–630.
109. **Havekes, L. M., B. J. J. Lugtenberg, and W. P. M. Hoekstra.** 1976. Conjugation deficient *E. coli* K12 F^- mutants with heptose-less lipopolysaccharide. Mol. Gen. Genet. **146:**43–50.
110. **Heinemann, J. A., and G. F. Sprague, Jr.** 1989. Bacterial conjugative plasmids mobilize DNA transfer between bacteria and yeast. Nature (London) **340:**205–209.
111. **Helmuth, R., and M. Achtman.** 1975. Operon structure of DNA transfer cistrons on the F sex factor. Nature (London) **257:**652–656.
112. **Howland, C. J., C. E. Rees, P. T. Barth, and B. M. Wilkins.** 1989. The *ssb* gene of plasmid ColIb-P9. J. Bacteriol. **171:**2466–2473.
113. **Ihler, G., and W. D. Rupp.** 1969. Strand-specific transfer of donor DNA during conjugation in *E. coli*. Proc. Natl. Acad. Sci. USA **63:**138–143.
114. **Inamoto, S., T. Abo, and E. Ohtsubo.** 1990. Binding sites of integration host factor in *oriT* of plasmid R100. J. Gen. Appl. Microbiol. **36:**287–293.
115. **Inamoto, S., and E. Ohtsubo.** 1990. Specific binding of the TraY protein to *oriT* and the promoter region for the *traY* gene of plasmid R100. J. Biol. Chem. **265:**6461–6466.
116. **Inamoto, S., Y. Yoshioka, and E. Ohtsubo.** 1988. Identification and characterization of the products from the *traJ* and *traY* genes of plasmid R100. J. Bacteriol. **170:**2749–2757.
117. **Inamoto, S., Y. Yoshioka, and E. Ohtsubo.** 1991. Site- and strand-specific nicking *in vitro* at *oriT* by the TraY-TraI endonuclease of plasmid R100. J. Biol. Chem. **266:**10086–10092.
118. **Ippen-Ihler, K., and S. Maneewannakul.** 1991. Conjugation among enteric bacteria: mating systems dependent on expression of pili, p. 35–69. *In* M. Dworkin (ed.), Microbial cell-cell interactions. American Society for Microbiology, Washington, D.C.
119. **Ippen-Ihler, K., D. Moore, S. Laine, D. A. Johnson, and N. S. Willetts.** 1984. Synthesis of F-pilin polypeptide in the absence of F *traJ* product. Plasmid **11:**116–129.
120. **Ippen-Ihler, K., and R. A. Skurray.** 1993. Genetic organization of transfer-related determinants on the sex factor F and related plasmids, p. 23–52. *In* D. Clewell (ed.), Bacterial conjugation. Plenum Press, New York.
121. **Iuchi, S., and E. C. C. Lin.** 1988. *arcA*(*dye*), a global regulatory gene in *Escherichia coli* mediating repression of enzymes in aerobic pathways. Proc. Natl. Acad. Sci. USA **85:**1888–1892.
122. **Jackson, R., D. Cram, A. Ray, D. Di Bernardino, and R. Skurray.** 1984. Cloning and analysis of *pif*, replication and leading regions of the F plasmid. Mol. Gen. Genet. **197:**129–136.
123. **Jalajakumari, M. B., A. Guidolin, H. J. Bukh, P. A. Manning, L. M. Ham, A. L. M. Hodgson, K.-C. Cheah, and R. A. Skurray.** 1987. Surface exclusion genes *traS* and *traT* of the F sex factor of *Escherichia coli* K-12. Determination of the nucleotide sequence and promoter and terminator activities. J. Mol. Biol. **198:**1–11.
124. **Jalajakumari, M. B., and P. A. Manning.** 1989. Nucleotide sequence of the *traD* region in the *E. coli* F sex factor. Gene **81:**195–202.
125. **Jones, A. L., P. T. Barth, and B. M. Wilkins.** 1992. Zygotic induction of plasmid *ssb* and *psiB* genes following conjugative transfer of IncI1 plasmid ColIb-P9. Mol. Microbiol. **6:**605–614.
126. **Kado, C. I.** 1993. *Agrobacterium*-mediated transfer and stable incorporation of foreign genes in plants, p. 243–254. *In* D. B. Clewell (ed.), Bacterial conjugation. Plenum Press, New York.
127. **Kathir, P., and K. Ippen-Ihler.** 1991. Construction and characterization of derivatives carrying insertion mutations in F plasmid transfer region genes, *trbA*, *artA*, *traQ*, and *trbB*. Plasmid **26:**40–54.
128. **Kathir, P., D. Moore, and K. Ippen-Ihler.** Unpublished data.
129. **Kingsman, A., and N. Willetts.** 1978. The requirements for conjugal DNA synthesis in the donor strain during F*lac* transfer. J. Mol. Biol. **122:**287–300.
130. **Kolodkin, A. L., M. A. Capage, E. I. Golub, and K. B. Low.** 1983. F sex factor of *E. coli* K-12 codes for a single-stranded DNA binding protein. Proc. Natl. Acad. Sci. USA **80:**4422–4426.
131. **Koraimann, G., and G. Hogenauer.** 1989. A stable core region of the *tra* operon mRNA of plasmid R1-19. Nucleic Acids Res. **17:**1283–1298.
132. **Koraimann, G., V. Koraimann, S. Schlager, and G. Hogenauer.** 1991. Repression and derepression of conjugation of plasmid R1 by wild-type and mutated *finP* antisense RNA. Mol. Microbiol. **5:**77–87.
133. **Koraimann, G., C. Schroller, H. Graus, D. Angerer, K. Teferle, and G. Hogenauer.** 1991. Expression of gene *19* of the conjugative plasmid R1 is controlled by RNase III. Mol. Microbiol. **9:**717–727.
134. **Koronakis, V., and G. Hogenauer.** 1986. The sequences of the *traJ* gene and the 5′ end of the *traY* gene of the resistance plasmid R1. Mol. Gen. Genet. **203:**137–142.
135. **Koronakis, V. E., E. Bauer, and G. Hogenauer.** 1985. The *traM* gene of the resistance plasmid R1: comparison with the corresponding sequence of the *E. coli* F factor. Gene **36:**79–86.
136. **Lahue, E. E., and S. W. Matson.** 1990. Purified *E. coli* F-factor TraY protein binds *oriT*. J. Bacteriol. **172:**1385–1391.
137. **Laine, S., D. Moore, P. Kathir, and K. Ippen-Ihler.** 1985. Genes and gene products involved in the synthesis of F-pili. Basic Life Sci. **30:**535–553.
138. **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.
139. **Lawn, A. M., and E. Meynell.** 1970. Serotypes of sex pili. J. Hyg. **68:**683–694.
140. **Lawn, A. M., E. Meynell, and M. Cooke.** 1971. Mixed infections with bacterial sex factors: sex pili of pure and mixed phenotype. Ann. Inst. Pasteur (Paris) **120:**3–8.
141. **Lee, S. H., L. S. Frost, and W. Paranchych.** 1992. FinOP repression of the F plasmid involves extension of the half-life of FinP antisense RNA by FinO. Mol. Gen. Genet. **235:**131–139.
142. **Lerner, T., and N. Zinder.** 1979. Chromosomal regulation of sexual expression in *Escherichia coli*. J. Bacteriol. **137:**1063–1065.
143. **Lessl, M., D. Balzer, W. Pansegrau, and E. Lanka.** 1992. Sequence similarities between the RP4 Tra2 and the Ti VirB region strongly support the conjugation model for T-DNA transfer. J. Biol. Chem. **267:**20471–20480.
144. **Lessl, M., W. Pansegrau, and E. Lanka.** 1992. Relationship of DNA-transfer-systems: essential transfer factors of plasmids RP4, Ti and F share common sequences. Nucleic Acids Res. **20:**6099–6100.
145. **Levy, S. B.** 1975. Very stable messenger RNA in chromosomeless *Escherichia coli* minicells. Proc. Natl. Acad. Sci. USA **72:**2900–2904.
146. **Lin, Y. N., J. Tennent, and R. Skurray.** Unpublished data.
147. **Llosa, M., S. Bolland, and F. de la Cruz.** 1991. Structural and functional analysis of the origin of conjugal transfer of the broad-host-range IncW plasmid R388 and comparison with the related IncN plasmid R46. Mol. Gen. Genet. **226:**473–483.
148. **Loh, S., D. Cram, and R. Skurray.** 1989. Nucleotide sequence of the leading region adjacent to the origin of transfer on plasmid F and its conservation among conjugative plasmids. Mol. Gen. Genet. **219:**177–186.
149. **Loh, S., R. Skurray, J. Celerier, M. Bagdasarian, A. Bailone, and R. Devoret.** 1990. Nucleotide sequence of the *psiA* (plasmid SOS inhibition) gene located on the leading region of plasmids F and R6-5. Nucleic Acids Res. **18:**4597.
150. **Loh, S. M., A. Ray, D. S. Cram, L. E. O'Gorman, and R. A. Skurray.** 1986. Location of a second partitioning region (ParL) on the F plasmid. FEMS Microbiol. Lett. **37:**179–182.
151. **Loh, S. M., and R. Skurray.** Unpublished data.
152. **Lopez, J., L. Salazar, I. Andres, J. M. Ortiz, and J. C. Rodriguez.** 1991. Nucleotide sequence of the *oriT-traM-finP* region of the haemolytic plasmid pSU316: comparison to F. Nucleic Acids Res. **19:**3451.
153. **Luo, Y., Q. Gao, and R. Deonier.** 1994. Mutational and physical analysis of F plasmid *traY* protein binding at *oriT*. Mol. Microbiol. **11:**449–458.
154. **Majdalani, N., and K. Ippen-Ihler.** Unpublished data.
155. **Maneewannakul, K., and K. Ippen-Ihler.** Unpublished data.
156. **Maneewannakul, K., and K. Ippen-Ihler.** 1993. Construction and analysis of F plasmid *traR*, *trbJ*, and *trbH* mutants. J. Bacteriol. **175:**1528–1531.
157. **Maneewannakul, K., S. Maneewannakul, and K. Ippen-Ihler.**

1992. Sequence alterations affecting F plasmid transfer gene expression: a conjugation system dependent on transcription by the RNA polymerase of phage T7. Mol. Microbiol. **6**:2961–2973.
158. **Maneewannakul, K., S. Maneewannakul, and K. Ippen-Ihler.** 1993. Synthesis of F pilin. J. Bacteriol. **175**:1384–1391.
159. **Maneewannakul, K., D. Moore, and K. Ippen-Ihler.** Unpublished data.
160. **Maneewannakul, K., D. Moore, and K. Ippen-Ihler.** 1992. The amino terminus of the pilin polypeptide in pilus filaments produced by an F *traX* mutant is not acetylated, abstr. H-67, p. 94. Abstr. 92nd Gen. Meet. Am. Soc. Microbiol. 1992. American Society for Microbiology, Washington, D.C.
161. **Maneewannakul, S., P. Kathir, and K. Ippen-Ihler.** 1992. Characterization of the F plasmid mating aggregation gene *traN* and of the new F transfer region locus *trbE*. J. Mol. Biol. **225**:299–311.
162. **Maneewannakul, S., P. Kathir, D. Moore, L. A. Le, J. H. Wu, and K. Ippen-Ihler.** 1987. Location of F plasmid transfer operon genes *traC* and *traW* and identification of the *traW* product. J. Bacteriol. **169**:5119–5124.
163. **Maneewannakul, S., K. Maneewannakul, and K. Ippen-Ihler.** 1991. Characterization of *trbC*, a new F plasmid *tra* operon gene that is essential to conjugative transfer. J. Bacteriol. **173**:3872–3878.
164. **Maneewannakul, S., K. Maneewannakul, and K. Ippen-Ihler.** 1992. Characterization, localization, and sequence of F transfer region products: the pilus assembly gene product, TraW, and a new product, TrbI. J. Bacteriol. **174**:5567–5574.
165. **Manning, P. A., L. Beutin, and M. Achtman.** 1980. Outer membrane of *Escherichia coli*: properties of the F sex factor *traT* protein which is involved in surface exclusion. J. Bacteriol. **142**:285–294.
166. **Manning, P. A., B. Kusecek, G. Morelli, C. Fisseau, and M. Achtman.** 1982. Analysis of the promoter-distal region of the *tra* operon of the F sex factor of *Escherichia coli* K-12 encoded by *Eco*RI restriction fragments f17, f19, and f2. J. Bacteriol. **150**:76–88.
167. **Manning, P. A., G. Morelli, and M. Achtman.** 1981. TraG protein of the F sex factor of *Escherichia coli* K12 and its role in conjugation. Proc. Natl. Acad. Sci. USA **78**:7487–7491.
168. **Manoil, C.** 1983. A genetic approach to defining the sites of interaction of a membrane protein with different external agents. J. Mol. Biol. **169**:507–519.
169. **Manoil, C., and J. P. Rosenbusch.** 1982. Conjugation-deficient mutants of *Escherichia coli* distinguish classes of functions of the outer membrane OmpA protein. Mol. Gen. Genet. **187**:148–156.
170. **Marvin, D. A., and W. Folkhard.** 1986. Structure of F-pili: reassessment of the symmetry. J. Mol. Biol. **191**:299–300.
171. **Masai, H., N. Nomura, Y. Kubota, and K.-I. Arai.** 1990. Roles of ϕX174 type primosome- and G4 type primase-dependent primings in initiation of lagging and leading strand syntheses of DNA replication. J. Biol. Chem. **265**:15124–15133.
172. **Matson, S. W., and B. S. Morton.** 1991. *Escherichia coli* DNA helicase I catalyzes a site- and strand-specific nicking reaction at the F plasmid *oriT*. J. Biol. Chem. **266**:16232–16237.
173. **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.
174. **Mazodier, P., and J. Davies.** 1991. Gene transfer between distantly related bacteria. Annu. Rev. Genet. **25**:147–171.
175. **McIntire, S. A., and W. B. Dempsey.** 1987. Fertility inhibition gene of plasmid R100. Nucleic Acids Res. **15**:2029–2042.
176. **Meynell, E.** 1978. Experiments with sex pili: an investigation of the characters and function of F-like and I-like sex pili based on their reactions with antibody and phage, p. 207–233. *In* D. E. Bradley, E. Raizen, P. Fives-Taylor, and J. Ou (ed.), Pili. International Conferences on Pili, Washington, D.C.
177. **Meynell, E., and A. Ewins.** 1973. Effect on exclusion of alterations to the F pilus. J. Bacteriol. **113**:71–75.
178. **Meynell, E., R. A. Matthews, and A. M. Lawn.** 1974. Antigenic differences between the ends and shafts of sex pili. J. Gen. Microbiol. **82**:203–205.
179. **Meynell, E., G. G. Meynell, and N. Datta.** 1968. Phylogenetic relationships of drug resistance factors and other transmissible bacterial plasmids. Bacteriol. Rev. **32**:35–83.
180. **Miki, T., T. Horiuchi, and N. S. Willetts.** 1978. Identification and characterization of four new *tra* cistrons on the *E. coli* K12 sex factor F. Plasmid **1**:316–323.
181. **Minkley, E. G., Jr.** 1984. Purification and characterization of proTraTp, the signal sequence-containing precursor of a secreted protein encoded by the F sex factor. J. Bacteriol. **158**:464–473.
182. **Minkley, E. G., Jr. (Carnegie-Mellon University).** Unpublished data, cited in reference 107.
183. **Minkley, E. G., Jr., and K. Ippen-Ihler.** 1977. Identification of a membrane protein associated with expression of the surface exclusion region of the F transfer operon. J. Bacteriol. **129**:1613–1622.
184. **Minkley, E. G., Jr., S. Polen, C. C. Brinton, Jr., and K. Ippen-Ihler.** 1976. Identification of the structural gene for F-pilin. J. Mol. Biol. **108**:111–121.
185. **Minkley, E. G., Jr., and N. S. Willetts.** 1984. Overproduction, purification and characterization of the F *traT* protein. Mol. Gen. Genet. **196**:225–235.
186. **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.
187. **Moore, D., P. Kathir, C. Hamilton, and K. Ippen-Ihler.** Unpublished data.
188. **Moore, D., K. Maneewannakul, S. Maneewannakul, J. H. Wu, K. Ippen-Ihler, and D. E. Bradley.** 1990. Characterization of the F-plasmid conjugative transfer gene *traU*. J. Bacteriol. **172**:4263–4270.
189. **Moore, D., B. A. Sowa, and K. Ippen-Ihler.** 1981. The effect of *tra* mutations on the synthesis of the F-pilin membrane polypeptide. Mol. Gen. Genet. **184**:260–264.
190. **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.
191. **Moore, D., B. A. Sowa, and K. Ippen-Ihler.** 1982. A new activity in the F *tra* operon which is required for F-pilin synthesis. Mol. Gen. Genet. **188**:459–464.
192. **Moore, D., J. H. Wu, P. Kathir, C. M. Hamilton, and K. Ippen-Ihler.** 1987. Analysis of transfer genes and gene products within the *traB-traC* region of the *Escherichia coli* fertility factor, F. J. Bacteriol. **169**:3994–4002.
193. **Morona, R., M. Klose, and U. Henning.** 1984. *Escherichia coli* K-12 outer membrane protein (OmpA) as a bacteriophage receptor: analysis of mutant genes expressing altered proteins. J. Bacteriol. **159**:570–578.
194. **Mullineaux, P., and N. Willetts.** 1985. Promoters in the transfer region of plasmid F. Basic Life Sci. **30**:605–614.
195. **Nelson, W. C., B. S. Morton, E. E. Lahue, and S. W. Matson.** 1993. Characterization of the *Escherichia coli* F factor *traY* gene product and its binding sites. J. Bacteriol. **175**:2221–2228.
196. **Nomura, N., H. Masai, M. Inuzuka, C. Miyazaki, E. Ohtsubo, T. Itoh, S. Sasamoto, M. Matsui, R. Ishizaki, and K.-I. Arai.** 1991. Identification of eleven single-strand initiation sequences (*ssi*) for priming of DNA replication in the F, R6K, R100 and ColE2 plasmids. Gene **108**:15–22.
197. **Ogata, R. T., C. Winters, and R. P. Levine.** 1982. Nucleotide sequence analysis of the complement resistance gene from plasmid R100. J. Bacteriol. **151**:819–827.
198. **Ohnishi, Y., H. Iguma, T. Ono, H. Nagaishi, and A. J. Clark.** 1977. Genetic mapping of the F plasmid gene that promotes degradation of stable ribonucleic acid in *Escherichia coli*. J. Bacteriol. **132**:784–789.
199. **Ou, J. T., and T. F. Anderson.** 1970. Role of pili in bacterial conjugation. J. Bacteriol. **102**:648–654.
200. **Paiva, W. D., T. Grossman, and P. M. Silverman.** 1992. Characterization of F-pilin as an inner membrane component of *Escherichia coli* K-12. J. Biol. Chem. **267**:26191–26197.
201. **Panicker, M. M., and E. G. Minkley, Jr.** 1985. DNA transfer occurs during a cell surface contact stage of F sex factor-mediated bacterial conjugation. J. Bacteriol. **162**:584–590.
202. **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 with affinity for DNA. J. Biol. Chem. **267**:2761–2765.
203. **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.
204. **Paranchych, W.** 1975. Attachment, ejection and penetration stages of the RNA phage infectious process, p. 85–111. *In* N. Zinder (ed.), RNA phages. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
205. **Paranchych, W., B. B. Finlay, and L. S. Frost.** 1986. Studies on the regulation of IncF plasmid transfer operon expression. Banbury Rep. **24**:117–129.
206. **Paranchych, W., and L. S. Frost.** 1988. The physiology and biochemistry of pili. Adv. Microb. Physiol. **29**:53–114.
207. **Penfold, S., and L. S. Frost.** Unpublished data.
208. **Penfold, S. S., K. Usher, and L. S. Frost.** The nature of the *traK4* mutation in the F sex factor of *Escherichia coli*. J. Bacteriol. **176**:1924–1931.
209. **Perumal, N., and E. G. Minkley, Jr.** 1984. The product of the F sex factor *traT* surface exclusion gene is a lipoprotein. J. Biol. Chem. **259**:5357–5360.
210. **Pradel, E., and C. A. Schnaitman.** 1991. Effect of *rfaH* (*sfrB*) and temperature on expression of *rfa* genes of *Escherichia coli* K-12. J. Bacteriol. **173**:6428–6431.
211. **Qi, S.-Y., S. Sukupolvi, and C. D. O'Connor.** 1991. Outer membrane permeability of *Escherichia coli* K12: isolation, cloning and mapping of suppressors of a defined antibiotic-hypersensitive mutant. Mol. Gen. Genet. **229**:421–427.
212. **Rashtchian, A., J. H. Crooks, and S. B. Levy.** 1983. *traJ* independence in expression of *traT* on F. J. Bacteriol. **154**:1009–1012.
213. **Ray, A., K. C. Cheah, and R. Skurray.** 1986. An F-derived conjugative cosmid: analysis of *tra* polypeptides in cosmid-infected cells. Plasmid **16**:90–100.
214. **Ray, A., and R. Skurray.** 1983. Cloning and polypeptide analysis of the leading region in F plasmid DNA transfer. Plasmid **9**:262–272.
215. **Ray, A., and R. Skurray.** 1984. Stabilization of the cloning vector pACYC184 by insertion of F plasmid leading region sequences. Plasmid **11**:272–275.
216. **Rees, C. E. D., and B. M. Wilkins.** 1989. Transfer of *tra* proteins into the recipient cell during bacterial conjugation mediated by plasmid ColIb. J. Bacteriol. **171**:3152–3157.
217. **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.
218. **Reygers, U., R. Wessel, H. Müller, and H. Hoffmann-Berling.** 1991. Endonuclease activity of *Escherichia coli* DNA helicase I directed against the transfer origin of the F factor. EMBO J. **10**:2689–2694.
219. **Richardson, J. P.** 1991. Preventing the synthesis of unused transcripts by Rho factor. Cell **64**:1047–1049.
220. **Riede, I., and M.-L. Eschbach.** 1986. Evidence that TraT interacts with OmpA of *Escherichia coli*. FEBS Lett. **205**:241–245.
221. **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.
222. **Salazar, L., J. Lopez, J. M. Andres, and J. C. Rodriguez.** 1992. Characterization and nucleotide sequence of the *oriT-traM-finP* region of the IncFVII plasmid pSU233. Mol. Gen. Genet. **234**:442–448.
223. **Schandel, K. A., S. Maneewannakul, K. Ippen-Ihler, and R. E. Webster.** 1987. A *traC* mutant that retains sensitivity to f1 bacteriophage but lacks F pili. J. Bacteriol. **169**:3151–3159.
224. **Schandel, K. A., S. Maneewannakul, R. A. Vonder Haar, K. Ippen-Ihler, and R. E. Webster.** 1990. Nucleotide sequence of the F plasmid gene, *traC*, and identification of its product. Gene **96**:137–140.
225. **Schandel, K. A., M. M. Muller, and R. E. Webster.** 1992. Localization of TraC, a protein involved in assembly of the F conjugative pilus. J. Bacteriol. **174**:3800–3806.
226. **Scherzinger, E., R. Lurz, S. Otto, and B. Dobrinski.** 1992. *In vitro* cleavage of double- and single-stranded DNA by plasmid RSF1010-encoded mobilization proteins. Nucleic Acids Res. **20**:41–48.
227. **Schwab, M., H. Gruber, and G. Hogenauer.** 1991. The TraM protein of plasmid R1 is a DNA-binding protein. Mol. Microbiol. **5**:439–446.
228. **Schwab, M., H. Reisenzein, and G. Hogenauer.** 1993. TraM of plasmid R1 regulates its own expression. Mol. Microbiol. **7**:795–803.
229. **Sharp, P. A., S. N. Cohen, and N. Davidson.** 1973. Electron microscope heteroduplex studies of sequence relations among plasmids of *Escherichia coli*. J. Mol. Biol. **75**:235–255.
230. **Shirasu, K., and C. Kado.** 1993. The *virB* operon of *Agrobacterium tumefaciens* virulence regulon has sequence similarities to B, C and D open reading frames downstream of the pertussis toxin-operon and to the DNA transfer-operons of broad-host-range conjugative plasmids. Nucleic Acids Res. **21**:353–354.
231. **Shirasu, K., and C. 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.
232. **Shirasu, K., Z. Koukolikova-Nicola, B. Hohn, and C. Kado.** 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., in press.
233. **Silverman, P. M., K. Nat, J. McEwen, and R. Birchman.** 1980. Selection of *Escherichia coli* K-12 chromosomal mutants that prevent expression of F-plasmid functions. J. Bacteriol. **143**: 1519–1523.
234. **Silverman, P. M., E. Wickersham, and R. Harris.** 1991. Regulation of the F plasmid *traY* promoter in *Escherichia coli* by host and plasmid factors. J. Mol. Biol. **218**:119–128.
235. **Silverman, P. M., E. Wickersham, S. Rainwater, and R. Harris.** 1991. Regulation of the F plasmid *traY* promoter in *Escherichia coli* K12 as a function of sequence context. J. Mol. Biol. **220**:271–279.
236. **Simon, J., and L. Frost.** Unpublished data.
237. **Simonsen, L.** 1990. Dynamics of plasmid transfer on surfaces. J. Gen. Microbiol. **136**:1001–1007.
238. **Skurray, R. A., R. E. W. Hancock, and P. Reeves.** 1974. Con$^-$ mutants: class of mutants of *Escherichia coli* K12 lacking a major cell wall protein and defective in conjugation and adsorption of a bacteriophage. J. Bacteriol. **119**:726–735.
239. **Somers, W. S., and S. E. V. Philips.** 1992. Crystal structure of the *met* repressor-operator complex at 2.8 A resolution reveals DNA recognition by β-strands. Nature (London) **359**:387–393.
240. **Stock, J., A. Ninfa, and A. Stock.** 1989. Protein phosphorylation and regulation of adaptive responses in bacteria. Microbiol. Rev. **53**:450–490.
241. **Sukupolvi, S., and C. D. O'Connor.** 1990. TraT lipoprotein, a plasmid-specified mediator of interactions between gram-negative bacteria and their environment. Microbiol. Rev. **54**:331–341.
242. **Takeda, Y., D. H. Ohlendorf, W. F. Anderson, and B. W. Matthews.** 1983. DNA-binding proteins. Science **221**:1020–1026.
243. **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.
244. **Thompson, R., L. Taylor, K. Kelly, R. Everett, and N. Willetts.** 1984. The F plasmid origin of transfer: DNA sequence of wild-type and mutant origins and location of origin-specific nicks. EMBO J. **3**:1175–1180.
245. **Thompson, T. L., M. B. Centola, and R. C. Deonier.** 1989. Location of the nick at *oriT* of the F plasmid. J. Mol. Biol. **207**:505–512.
246. **Traxler, B. A., and E. G. Minkley, Jr.** 1987. Revised genetic map of the distal end of the F transfer operon: implications for DNA helicase 1, nicking at *oriT*, and conjugal DNA transport. J. Bacteriol. **169**:3251–3259.
247. **Traxler, B. A., and E. G. Minkley, Jr.** 1988. Evidence that DNA helicase I and *oriT* site-specific nicking are both functions of the F TraI protein. J. Mol. Biol. **204**:205–209.
248. **Tsai, M.-M., Y. H. F. Fu, and R. C. Deonier.** 1990. Intrinsic bends and integration host factor binding at F plasmid *oriT*. J. Bacteriol. **172**:4603–4609.

249. **Usher, K., and L. S. Frost.** Unpublished data.
250. **van Biesen, T., and L. S. Frost.** 1992. Differential levels of fertility inhibition among F-like plasmids are related to the cellular concentration of *finO* mRNA. Mol. Microbiol. **6:**771–780.
251. **van Biesen, T., and L. S. Frost.** Unpublished data.
252. **van Biesen, T., F. Soderbom, E. G. H. Wagner, and L. S. Frost.** Structural and functional analysis of the FinP antisense RNA regulatory system of the F conjugative plasmid. Mol. Microbiol., in press.
253. **Van Mansfield, A. D., H. A. Van Teeffelen, P. D. Baas, and H. S. Jansz.** 1986. Two juxtaposed tyrosyl-OH groups participate in ϕX174 gene A protein catalyzed cleavage and ligation of DNA. Nucleic Acids Res. **14:**4229–4238.
254. **Weber, R. F., and P. M. Silverman.** 1988. The Cpx proteins of *Escherichia coli* K-12. Structure of the CpxA polypeptide as an inner membrane protein. J. Mol. Biol. **203:**467–478.
255. **Weiss, A., F. D. Johnson, and D. L. Burnes.** 1993. Molecular characterization of an operon required for pertussis toxin secretion. Proc. Natl. Acad. Sci. USA **90:**2970–2974.
256. **Wilkins, B., and E. Lanka.** 1993. DNA processing and replication during plasmid transfer between Gram-negative bacteria, p. 105–136. *In* D. Clewell (ed.), Bacterial conjugation. Plenum Press, New York.
257. **Willetts, N.** 1977. The transcriptional control of fertility in F-like plasmids. J. Mol. Biol. **112:**141–148.
258. **Willetts, N.** 1980. Interactions between the F conjugal transfer system and CloDF13::TnA plasmids. Mol. Gen. Genet. **180:**213–217.
259. **Willetts, N., and J. Maule.** 1986. Specificities of IncF plasmid conjugation genes. Genet. Res. **47:**1–11.
260. **Willetts, N. S.** 1973. Characterization of the F transfer cistron *traL*. Genet. Res. **21:**205–213.
261. **Willetts, N. S.** 1974. The kinetics of inhibition of F *lac* transfer by R100 in *E. coli*. Mol. Gen. Genet. **129:**123–130.
262. **Willetts, N. S.** 1993. Bacterial conjugation: a historical perspective, p. 1–22. *In* D. Clewell (ed.), Bacterial conjugation. Plenum Press, New York.
263. **Willetts, N. S., J. Maule, and S. McIntire.** 1976. The genetic location of *traO*, *finP* and *tra-4* on the *E. coli* K12 sex factor F. Genet. Res. **26:**255–263.
264. **Willetts, N. S., P. M. Moore, and W. Paranchych.** 1980. Variant pili produced by mutants of the F*lac* plasmid. J. Gen. Microbiol. **117:**455–464.
265. **Willetts, N. S., and R. Skurray.** 1980. The conjugation system of F-like plasmids. Annu. Rev. Genet. **14:**41–76.
266. **Willetts, N. S., and R. Skurray.** 1987. Structure and function of the F factor and mechanism of conjugation, p. 1110–1133. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society For Microbiology, Washington, D.C.
267. **Willetts, N. S., and B. Wilkins.** 1984. Processing of plasmid DNA during bacterial conjugation. Microbiol. Rev. **48:**24–41.
268. **Worobec, E. A., L. S. Frost, P. Pieroni, G. D. Armstrong, R. S. Hodges, J. M. Parker, B. B. Finlay, and W. Paranchych.** 1986. Location of the antigenic determinants of conjugative F-like pili. J. Bacteriol. **167:**660–665.
269. **Wu, J. H., and K. Ippen-Ihler.** 1989. Nucleotide sequence of *traQ* and adjacent loci in the *Escherichia coli* K-12 F-plasmid transfer operon. J. Bacteriol. **171:**213–221.
270. **Wu, J. H., P. Kathir, and K. Ippen-Ihler.** 1988. The product of the F plasmid transfer operon gene, *traF*, is a periplasmic protein. J. Bacteriol. **170:**3633–3639.
271. **Wu, J. H., D. Moore, T. Lee, and K. Ippen-Ihler.** 1987. Analysis of *Escherichia coli* K12 F factor transfer genes: *traQ*, *trbA*, and *trbB*. Plasmid **18:**54–69.
272. **Xiong, X. F., and W. E. Reznikoff.** 1993. Transcriptional slippage during the transcription initiation process at a mutant *lac* promoter *in vivo*. J. Mol. Biol. **231:**569–580.
273. **Yang, C.-C., and H. A. Nash.** 1989. The interaction of *E. coli* IHF protein with its specific binding sites. Cell **57:**869–880.
274. **Yoshioka, Y., Y. Fujita, and E. Ohtsubo.** 1990. Nucleotide sequence of the promoter-distal region of the *tra* operon of plasmid R100, including *traI* (DNA helicase I) and *traD* genes. J. Mol. Biol. **214:**39–53.
275. **Yoshioka, Y., H. Ohtsubo, and E. Ohtsubo.** 1987. Repressor gene *finO* in plasmids R100 and F: constitutive transfer of plasmid F is caused by insertion of IS*3* into F *finO*. J. Bacteriol. **169:**619–623.
276. **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.