The Plasmid R64 Thin Pilus Identified as a Type IV Pilus

SU-RYANG KIM AND TERUYA KOMANO

Department of Biology, Tokyo Metropolitan University, Minamiohsawa, Hachioji, Tokyo 192-03, Japan

Received 14 November 1996/Accepted 2 April 1997

The entire nucleotide sequence of the pil region of the IncI1 plasmid R64 was determined. Analysis of the sequence indicated that 14 genes, designated pil through pilV, are involved in the formation of the R64 thin pilus. Protein products of eight pil genes were identified by the maxicell procedure. The pilN product was shown to be a lipoprotein by an experiment using globomycin. A computer search revealed that several R64 pil genes have amino acid sequence homology with proteins involved in type IV pilus biogenesis, protein secretion, and transformation competence. The pilS and pilV products were suggested to be prepilins for the R64 thin pilus, and the pilN product appears to be a prepilin peptidase. These results suggest that the R64 thin pilus belongs to the type IV family, specifically group IVB, of pili. The requirement of the pilR and pilU genes for R64 liquid mating was demonstrated by constructing their frameshift mutations. Comparison of three type IVB pilus biogenesis systems, the pil system of R64, the toxin-coregulated pilus (tcp) system of Vibrio cholerae, and the bundle-forming pilus (bfp) system of enteropathogenic Escherichia coli, suggests that they have evolved from a common ancestral gene system.

Type IV pili are rod-like surface appendages produced by gram-negative bacteria such as Pseudomonas aeruginosa, Neisseria gonorrhoeae, Moraxella bovis, Myxococcus xanthus, and Vibrio cholerae, as well as enteropathogenic and enterotoxigenic Escherichia coli (for reviews, see references 27, 28, and 35). They are flexible, with a diameter of 6 to 7 nm and a length of up to 20 μm. They are produced at the polar position of the bacterial cell. Many type IV pili play important roles in the attachment of bacterial pathogens to membranes of eukaryotic host cells, as do the other pil (13, 35). Type IV pili are also associated with the twitching motility of various bacteria and with the social motility of myxobacteria (40).

Type IV pili are composed of pilin subunits (35). Pilin molecules from various bacteria have amino acid sequence homology (see Fig. 4E). The type IV pilin family is usually divided into two groups. Group A consists of pilins from P. aeruginosa, N. gonorrhoeae, M. bovis, and so on. They are closely related in amino acid sequence and are produced from prepilin molecules through the cleavage of 6- to 7-amino-acid signal peptides. The N-terminal amino acid of type IVA mature pilins is phenylalanine and is N methylated. Group B pilins, including toxin-coregulated pilus (tcp) in V. cholerae (26) and bundle-forming pilus (bfp) in enteropathogenic E. coli (33, 34), are substantively different from type IVA pilins. Their signal peptides are longer than those of type IVB pilins. The N-terminal amino acid of type IVB mature pilins is methionine or leucine.

The C-terminal amino acid (glutamine) of signal peptides and the 5th amino acid (glutamic acid) of mature pilins are completely conserved among type IVA and type IVB prepilins and related proteins (see Fig. 4E). A long hydrophobic segment is present at the N-terminal region of mature pilins. Cleavage of signal peptides from prepilins is carried out by a signal peptidase, specific for type IV prepilin, such as the PilD protein of P. aeruginosa (36). N methylation of mature pilins is also performed by the same signal peptidase protein.

Various extracellular protein secretion systems of gram-negative bacteria and DNA uptake systems of gram-positive bacteria are known to produce proteins containing extensive sequence homology to proteins required for type IV pilus biogenesis, including prepilins (12, 28).

Many gene products are required for the biogenesis of type IV pili. The gene organizations of the tcp system in V. cholerae (26) and the bfp system in enteropathogenic E. coli (33, 34) have been reported (see Fig. 6). In the tcp and bfp systems, 14 genes are involved in type IV pilus biogenesis. In the case of the bfp system and the longus (11) system in enteropathogenic and enterotoxigenic E. coli, respectively, the type IV pilus formation genes are encoded in large virulent plasmids. The virulent plasmid in enteropathogenic E. coli (EPEC) is called EPEC adherence factor (EAF).

Plasmid R64 is a conjugative plasmid belonging to the incompatibility group I1 (17, 18). The R64 transfer region is located within a 54.0-kb DNA segment, in which the traABCD genes are located at the left end and the oriT sequence is at the right end (Fig. 1A). R64 produces two types of sex pili, thin, flexible pili and thick, rigid pili (3, 4). Thin pili are required only for liquid mating, while thick pili are required for both surface and liquid mating. The leftmost 18.5-kb DNA segment of the R64 transfer region is responsible for R64 thin pilus formation (pil region), since E. coli cells harboring pKK641, containing the 18.5-kb segment together with the R64 rep segment (Fig. 1A), were shown to produce R64 thin pili (18, 19). E. coli cells harboring pKK641 are sensitive to the IncI1-specific phages Io (5) and PR64FS (6), which utilize thin pili as receptors.

The leftmost 3.6-kb segment of pKK641 was sequenced and shown to contain the traABCD genes (15) (Fig. 1B). The traBC genes were inferred to be positive regulators of R64 transfer gene expression, since they were required both for thin-pilus formation and for conjugal transfer in liquid medium as well as on a solid surface.

The rightmost 4.1-kb segment of pKK641 was also sequenced and was shown to contain the pilV gene, the R64 shufflon, and a portion of the rci gene (20, 21). The C-terminal segment of the pilV gene is under the control of DNA rearrangement of the shufflon. The R64 shufflon consists of four DNA segments flanked and separated by seven 19-bp repeat sequences. The rci gene product promotes site-specific recompen...
bination between any two inverted repeat sequences, resulting in the inversion of four DNA segments independently or in groups. The shufflon is thought to function as a biological switch to select one of seven pilV genes, whereby the N-terminal segment remains constant while the C-terminal region is variable. The pilV gene product is a minor component of the R64 thin pilus. The seven variable C-terminal segments of the pilV gene determine the recipient specificity of R64 liquid mating (18, 19).

Recently, a DNA segment carrying the pil region of IncI1 plasmid CoIIb-P9adr-d-1, closely related to R64, was cloned (18, 42). E. coli cells harboring this plasmid were shown to over-produce thin pili which showed a filamentous structure under an electron microscope (42). The purified thin pili consist of 19- and 48-kDa proteins, which were shown to be the products of the CoIIb-P9 pilS and pilV4A genes, respectively.

In the present study, we have determined the nucleotide sequence of the remaining portion of the R64 pil region and have found that 14 genes are present in the same orientation within the R64 pil region. Protein products of eight genes were detected by the maxicell procedure. The deduced amino acid sequences of several pil genes suggest that the R64 thin pilus belongs to the type IV family of pili.

MATERIALS AND METHODS

Bacterial strains, phages, and plasmids. E. coli K-12 strains used were JM83 Δ(lac-proAB) rpsL thr ara80 dacZ8M15, JM109 recA1 Δ(lac-proAB) endA1 gna96 thi hsdR17 ungE44 relA1F tralΔ36 proA lacZ8M15 (38, 41), TN102 (17), and CSR603 uvrA6 recA1 phr thr leuB6 proA2 thi argE3 lacY1 pilK2 ara ytl pilS31 tsu supE33 (31). Phages J5 (3) and PR64Fs (6) were used as IncI1-specific phages. Phage vectors M13mp18 and M13mp19, and plasmid vectors pUC9 and pUC19, were used for sequencing and cloning, respectively (38, 41). The other plasmids used in this study are shown in Fig. 1.

Media. Luria-Bertani and H media were prepared as described by Miller (24). K and Hershey media were used for the maxicell experiment (31). The agar medium consisted of Luria-Bertani or H broth containing 1.5 or 1.2% agar, respectively. Antibiotics were added to liquid or solid medium at the following concentrations: ampicillin, 100 μg/ml; chloramphenicol, 25 μg/ml; kanamycin, 50 μg/ml; nalidixic acid, 20 μg/ml; and tetracycline, 12.5 μg/ml.

DNA manipulation and sequencing. The preparation of plasmid DNA, construction of plasmids, transformation, and other methods of DNA manipulation were performed as previously described (30). pKK686, pKK687, pKK688, pKK689, pKK690, and pKK691 were constructed by inserting 3.0-kb PstI, 2.8-kb HpaI, 6.3-kb PirII, 5.0-kb Hpal, and 4.1-kb Clar-EcoRI fragments of pKK641 into pUC9 in the direction in which pil genes are expressed under the control of the lac promoter of the vector (see Fig. 1B).

Frameshift mutations were introduced into pilR and pilV4A genes of pKK690 as described previously (17). pKK690 DNA was digested partially with Rsal or completely with Hpal, ligated with a 1.45-kb DNA cassette for tetracycline resistance from pUC7Tc, and used to transform E. coli JM83 cells. The DNA fragment for tetracycline resistance was removed from the resultant plasmid DNA by BamH1 digestion of DNA, followed by self-ligation. A 22-bp DNA sequence, AAATCCCCGATTCCGGAATT, remaining at the Rsal site (position 12313 in Fig. 2) or at the Hpal site (position 14378) gave rise to the pilR and pilV4A frameshift mutations, respectively. The pilR1 and pilU1 mutations in pKK690 were transferred into pKK641 by the in vivo gene replacement method (17).

The nucleotide sequence was determined by the dye-deoxy chain termination method (32). The sequences of both strands were determined with overlapping fragments. The resulting sequence was analyzed by using DNA/Sis (Hitachi) and the GenomeNet WWW server (Kyoto University).

Identification of protein products. Plasmids pKK686 through -689, and pKK690 and its insertion derivatives, were introduced into E. coli CSR603 by transformation, and the protein products were labeled with [35S]methionine according to the maxicell procedure (31). For pKK687, globomycin (final concentration, 50 to 200 μg/ml) was added to a 0.5-ml culture of maxicells 20 min prior to the addition of [35S]methionine. The labeled proteins were separated by sodium dodecyl sulfate-17.5% polyacrylamide gel electrophoresis (SDS-17.5% PAGE) and visualized by fluorography. Molecular-weight-standard proteins were stained with Coomassie brilliant blue.

Conjugal transfer and phage sensitivity. Liquid mating was performed as described previously (17). E. coli JM83 and TN102 cells were used as donor and recipient cells, respectively. A culture of log-phase donor cells was mixed with an overnight culture of recipient cells. The mixture was incubated for 90 min at 37°C. Sensitivity to phages J5 and PR64Fs was determined as described previously (17). E. coli JM83 and TN102 cells were used as donor and recipient cells, respectively. A culture of log-phase donor cells was mixed with an overnight culture of recipient cells. The mixture was incubated for 90 min at 37°C. Sensitivity to phages J5 and PR64Fs was determined as described previously (17). A culture of log-phase donor cells was mixed with an overnight culture of recipient cells. The mixture was incubated for 90 min at 37°C. Sensitivity to phages J5 and PR64Fs was determined as described previously (17). Materials. Restriction enzymes, sequencing kits, and other materials for recombinant DNA experiments were commercially obtained and used as recommended by the manufacturers. [α-32P]dCTP (800 Ci/mmol) and [35S]methionine (1,000 Ci/mmol) were from Du Pont, NEN Research Products, and from ICN Biomedicals, respectively. Globomycin was a generous gift from M. Inukai.

Nucleotide sequence accession number. The nucleotide sequence data reported in this paper will appear in the DDBJ, EMBL, and GenBank nucleotide sequence databases as accession no. D85858.
FIG. 2. Nucleotide and deduced amino acid sequences of the R64 pil region. The sequence is numbered from the leftmost BglII site shown in Fig. 1A. Putative promoter sequences and initiation codons for various pil genes are boxed. Predicted Shine-Dalgarno sequences are underlined. The signal peptides and hydrophobic domains of the pil proteins are also underlined.
RESULTS

Nucleotide sequence of the pil region in plasmid R64 Δtrd-11. pKK641 contained all the information necessary for the formation of the R64 thin pilus as described in the Introduction. Hence, to examine the gene organization of the R64 pil region, the nucleotide sequence of a 10.7-kb Smal-HpaI segment (Fig. 1B) of pKK641 was determined by the dideoxy chain termination method (32). Figure 2 shows an 11,160-bp nucleotide sequence of the R64 pil region. A 10,734-bp sequence from the Smal site at position 3644 to the HpaI site at position 14378 was newly determined in this study.

On examination of the sequence, 13 open reading frames (ORFs) were newly identified and tentatively designated as genes pilI through pilU (Fig. 2). The pilI gene, which was previously shown to be required for R64 thin-pilus formation, sequentially follows the pilU gene (20). The gene organization of the R64 pil region, deduced from the nucleotide sequence, is shown schematically in Fig. 1B. These genes are situated side by side in the same orientation as the traABCD genes. All genes start with an ATG initiation codon, preceded by potential Shine-Dalgarno sequences. Most genes stop with a TAA termination codon. Intergenic spaces between neighboring genes vary from 49 to 16 nucleotides (that is, a 49-nucleotide space between genes and a 16-nucleotide overlap). Since no promoter or terminator structure was found in the coding and intergenic regions of the pilI through pilV genes, these 14 genes may form a single long operon. The overlapping nature of the neighboring genes suggests the presence of translational coupling between these genes.

The termination codon of the preceding traD gene is located 267 bp upstream from the pilI gene. In this intergenic region between the traD and pilI genes, sequences CTGTCGA (positions 3657 to 3662) and TTTTTTT (positions 3679 to 3684) may function as −35 and −10 promoter recognition sequences for σ70 of E. coli RNA polymerase. In the shufflon region, two sets of 19-bp repeat sequences, which function as the crossover sites of DNA inversion of the shufflon, are located in the inverted orientation with 8-bp spaces. These 19-bp inverted repeat sequences may also function as transcription terminator signals for the pil operon.

Identification of the pil gene products. To identify the pil gene products, the maxicell procedure was performed (31). Various segments of pKK641 DNA were subcloned into the high-copy-number vector pUC9 to generate pKK686 (containing the pilI to pilM genes), pKK687 (pilIV), pKK688 (pilO to pilS), pKK689 (pil VII to pilT), and pKK690 (pil to pilV) (Fig. 1B). These plasmids were introduced into E. coli CSR603 cells by transformation, and the protein products were analyzed by the maxicell procedure. Maxicells harboring these plasmids produced various pil proteins (Fig. 3). The identification of pil proteins is described below.

Properties of R64 pil gene products. The number of amino acids, molecular mass, pl, and cellular location for each of the 14 pil gene products were predicted from the deduced amino acid sequences and are summarized in Table 1.

The pilI gene encodes an 84-amino-acid basic protein, and the pilJ gene encodes a 150-amino-acid acidic protein. The pilI product was detected as a 18-kDa protein by the maxicell procedure (Fig. 3, lane 2). A computer search of PilI and PilJ amino acid sequences failed to detect any significant sequence similarity with known proteins in the database.

The pilK gene product. The pilK gene encodes a 196-amino-acid basic protein. The pilK product could not be detected by the maxicell procedure. The amino acid sequence of the PilK protein has 27% identity with that of the R64 TraD protein (15) (Fig. 4A). The R64 traD gene is located 1.7 kb upstream from the pilK gene within the R64 transfer region (Fig. 1B). It was shown that the frameshift mutation within the traD coding sequence did not affect transfer frequency in liquid and surface mating of R64 (15). No known protein containing an amino acid sequence with significant similarity to that of the PilK or TraD protein was found.

The pilL and pilM gene products. The pilL gene encodes a 355-amino-acid basic protein, and the pilM gene encodes a 145-amino-acid basic protein. The amino acid sequence of the N-terminal region of the pilL product reveals features indicative of signal peptides of bacterial lipoproteins (Fig. 2). A faint 43-kDa protein band was detected as the pilL product by the maxicell experiment (Fig. 3, lane 2). The observed molecular mass (43 kDa) is significantly higher than the calculated value (39 kDa), suggesting a modification of the pilL product as a lipoprotein. No protein containing sequence similarity to the PilL and PilM proteins could be found in the database.

The pilN gene product is a lipoprotein. The pilN gene encodes a 560-amino-acid neutral protein. Since the amino acid sequence of the N-terminal region of the pilN product was also predicted to encode a lipoprotein, this possibility was tested by using globomycin (14). Globomycin is known to inhibit signal peptidase II, which cleaves the signal peptide of bacterial lipoproteins. Maxicells harboring pKK687 produced a protein with an apparent molecular mass of 67 kDa, which was higher than the calculated value of 60 kDa (Fig. 5, lane 2). Proteins with decreased mobility appeared upon the addition of globomycin at 50 to 200 μg/ml in the maxicell culture (Fig. 5, lanes 3 to 5), suggesting the accumulation of a prePilN protein with a higher molecular mass. Thus, it is likely that the pilN gene encodes a lipoprotein that is processed in the same manner as E. coli lipoproteins. The N-terminal 26-amino-acid signal peptide of the pilN product is believed to be removed by the action of signal peptidase II (Fig. 2). The mature PilN protein may be integrated into the bacterial outer membrane.

The amino acid sequence of the PilN protein has 25% identity to the BfpB sequence (Fig. 4B). The bfpB gene is a member of a gene cluster required for the formation of bundle-forming pilin (bfp) in enteropathogenic E. coli (33, 34). The BfpB protein is also a lipoprotein. Both PilN and BfpB proteins contain serine-rich segments (Fig. 4B). The PilN amino acid sequence
also has a slight similarity to those of the TcpC and PilD proteins, which are related to the formation of the type IV pilus.

The pilO and pilP gene products. The pilO gene encodes a 431-amino-acid basic protein, and the pilP gene encodes a 150-amino-acid basic protein. The products of the pilO and pilP genes were identified as 50- and 16-kDa proteins, respectively (Fig. 3, lane 4). No protein similar to either the PilO or PilP gene product was identified as the 50-kDa protein was identified as the pilQ gene product (Fig. 3, lane 4). The PilP protein may be an integral inner membrane protein, since pKK690 produced the 22-kDa protein and not the 19-kDa protein (compare Fig. 3, lanes 5 and 7). The pilS gene product is a prepilin peptidase, as described above. In fact, the PilS protein was reduced from 22 kDa, suggesting that the PilS protein is a major pilin, since the purified R64 thin pilus was shown to contain a 19-kDa protein as a major component (18). From the similarity of the pilS product to the type IV prepilins (Fig. 4E), the pilS product was assumed to lose the N-terminal 23-amino-acid signal peptide through the putative type IV prepilin peptidase activity of the pilU product. The signal peptide of the PilS protein is unusually long, as is that of the TcpA protein of V. cholerae. The PilS 23rd and 28th amino acid residues are glycine and glutamic acid, respectively, which are completely conserved among all type IV prepilins and related proteins (Fig. 4E).

The pilT gene product. The pilT gene product is a 186-amino-acid basic protein that has sequence similarity to the bfpH gene and gene X (gene 19) of IncF plasmids, including F, R1, and R100 (2) (Fig. 4F). The identity between the amino acid sequences of R64 PilT and the F 19-kDa protein (product of gene X) is as high as 43%. In IncF plasmids, gene X is located in the “leading region” and is the first to enter the recipient cells during conjugation. Recently, gene X was shown to be required for the efficient transfer of R1 (2). The introduction of two stop codons into the coding region of gene X resulted in a 10-fold reduction in R1 transfer frequency. It is interesting that the defective phenotype of the gene X mutation in R1 transfer was complemented in trans by the R64 pilT gene (2).

PilT-homologs are also present in Shigella flexneri (IpgF), Salmonella typhi (IagB), and E. coli (ORF138 and Slt) (Fig. 4F). In S. flexneri and S. typhi, the IpgF and IagB proteins are involved in the invasion of the eukaryotic host cells by the bacterial cells (1, 25). The E. coli slt gene encodes a 70-kDa soluble lytic transglycosylase (7, 8). X-ray crystallography revealed that the C-terminal domain of Slt protein has a three-dimensional structure similar to those of hen egg and T4 lysozymes (37). It is possible that PilT homologs carry peptidoglycan-lytic activity.

The pilU gene product is a prepilin peptidase. The pilU product is a 218-amino-acid basic protein. In the presence of the pilU gene, the size of the PilS protein was reduced from 22 to 19 kDa, suggesting that the pilU gene encodes a prepilin peptidase for the pilS product, as described above. In fact, the amino acid sequence of the PilU protein shows similarity to those of type IV prepilin peptidases, including BfpP, TcpJ,
PilD, and others (Fig. 4G). It is also similar to the R721 PilU protein (16). Since the R721 pilU gene is located just upstream of the R721 pilV gene and the deduced amino acid sequence is 23% identical to that of the R64 pilU gene, the R721 pilU gene may also encode a prepilin peptidase. However, the amino acid sequences of R64 and R721 PilU proteins are greatly divergent from those of the type IV prepilin peptidases of other bacteria.

The pilV gene product. The DNA rearrangement of the shufflon converts the C-terminal segments of the pilV genes, creating seven different pilV genes (18–20). The constant region of the PilV proteins consists of 361 amino acids, while the number of amino acids in the PilV variable region fluctuates between 69 and 113. The N-terminal amino acid sequence of the PilV proteins bears a structure similar to that of type IV prepilin (Fig. 4E). The conserved glycine and glutamic acid appear at the 7th and 12th amino acids of the PilV protein, respectively. It is most likely that the N-terminal 6-amino-acid peptide of the pilV product is cut off by the pilU function. The prediction that the pilV product is a minor pilin is supported by the previous finding that the PilV protein is a minor component of the R64 thin pilus (18, 42).

Effects of pilR and pilU mutations on liquid mating and sensitivity to phages. The effects of the pilR1 and pilU1 mutations on thin-pilus formation were measured by the transfer frequency in liquid medium and sensitivity to IncI1-specific FIG. 4. Alignment of the deduced amino acid sequences of the Pil proteins with those of known proteins. Conserved amino acid residues are printed in white on black. The consensus sequence is given below each alignment. Gaps, marked by dashes, were introduced to reveal the maximal similarity among the sequences. (A) Alignment of amino acid sequences of R64 PilK and R64 TraD (GenBank accession no. D14607). (B) Alignment of sequences of R64 PilN and EAF BfpB (Z68186). Polyserine segments are underlined. (C) Alignment of sequences of R64 PilQ, EAF BfpD (Z68186), V. cholerae (Vch) TcpT (X64098), P. aeruginosa (Pae) PilB (M32066), N. gonorrhoeae (Ngo) PilF (L10291), Klebsiella pneumoniae (Kpn) PilE (M32613), Erwinia chrysanthemi (Ech) OutE (L02214), P. aeruginosa XcpR (X62666), Xanthomonas campestris (Xca) XcpR (X50970), Aeromonas hydrophila (Ahu) XcpP (X56504), Bacillus subtilis (Bsu) ComG1 (M29691), P. aeruginosa PilT (M55524), and pT1A6 (Ti) VirB11 (J03216). Walker boxes A and B, Asp boxes, and CXXC sequences are indicated. (D) Alignment of sequences of R64 PilR, EAF BfpE (Z68186), V. cholerae TcpE (X64098), Pseudomonas putida (Ppu) PilC (X74276), V. campestris XpsF (X59079), P. aeruginosa XcpS (X62666), Erwinia carotovora (Eca) XcpX (X62666), P. aeruginosa XcpT (X62666), Erwinia chrysanthemi XpsG (X59079), and B. subtilis ComG2 (M29691). Hydrophobic domains are underlined. (E) Alignment of sequences of R64 PilS, R64 PilV (X62169), EAF BfpH (Z68186), V. cholerae TcpP (M33514), P. aeruginosa PilA (M14849), N. gonorrhoeae PilE (M32613), Moraxella bovis (Mbo) TfpQ (M59712), M. xanthus (Mxa) PilA (L39904), and B. subtilis ComG3 (M29691). The cleavage site of type IV pilins is shown by a downward arrow. The conserved glycine and glutamic acid residues are indicated by boldface. (F) Alignment of sequences of R64 PilT, EAF BfpH (Z68186), F gene X (M97766), S. flexneri (Sfl) IagB (X80892), S. typhi (Sty) IpgF (U14003), and E. coli ORF138 (U28375). Three motifs conserved in putative lytic transglycosylases are underlined. (G) Alignment of sequences of R64 PilU, R721 PilU (X62169), EAF BfpP (Z68186), V. cholerae TcpJ (M74708), P. aeruginosa PilD (M53066), N. gonorrhoeae PilD (L11715), E. carotovora OutO (X70049), E. coli IpgF (L04309), T. pyov (Sty) IagB (X80892), E. coli (Eco) ORF138 (L28375), and E. coli Slt (U14003). Three motifs conserved in putative lytic transglycosylases are underlined. (H) Alignment of sequences of R64 PilU, R721 PilU (X62169), EAF BfpP (Z68186), V. cholerae TcpJ (M74708), P. aeruginosa PilD (M53066), N. gonorrhoeae PilD (L11715), E. carotovora OutO (X70049), E. coli IpgF (L04309), T. pyov (Sty) IagB (X80892), E. coli (Eco) ORF138 (L28375), and E. coli Slt (U14003). Three motifs conserved in putative lytic transglycosylases are underlined. (I) Alignment of sequences of R64 PilU, R721 PilU (X62169), EAF BfpP (Z68186), V. cholerae TcpJ (M74708), P. aeruginosa PilD (M53066), N. gonorrhoeae PilD (L11715), E. carotovora OutO (X70049), E. coli IpgF (L04309), T. pyov (Sty) IagB (X80892), E. coli (Eco) ORF138 (L28375), and E. coli Slt (U14003). Three motifs conserved in putative lytic transglycosylases are underlined. (J) Alignment of sequences of R64 PilU, R721 PilU (X62169), EAF BfpP (Z68186), V. cholerae TcpJ (M74708), P. aeruginosa PilD (M53066), N. gonorrhoeae PilD (L11715), E. carotovora OutO (X70049), E. coli IpgF (L04309), T. pyov (Sty) IagB (X80892), E. coli (Eco) ORF138 (L28375), and E. coli Slt (U14003). Three motifs conserved in putative lytic transglycosylases are underlined. (K) Alignment of sequences of R64 PilU, R721 PilU (X62169), EAF BfpP (Z68186), V. cholerae TcpJ (M74708), P. aeruginosa PilD (M53066), N. gonorrhoeae PilD (L11715), E. carotovora OutO (X70049), E. coli IpgF (L04309), T. pyov (Sty) IagB (X80892), E. coli (Eco) ORF138 (L28375), and E. coli Slt (U14003). Three motifs conserved in putative lytic transglycosylases are underlined. (L) Alignment of sequences of R64 PilU, R721 PilU (X62169), EAF BfpP (Z68186), V. cholerae TcpJ (M74708), P. aeruginosa PilD (M53066), N. gonorrhoeae PilD (L11715), E. carotovora OutO (X70049), E. coli IpgF (L04309), T. pyov (Sty) IagB (X80892), E. coli (Eco) ORF138 (L28375), and E. coli Slt (U14003). Three motifs conserved in putative lytic transglycosylases are underlined.
phages. For liquid mating, the pKK641-pKK661 system, in which DNA rearrangement of the shufflon does not occur because of lack of the rcl activity, was used (17). E. coli donor cells harboring both pKK641 and pKK661 transmitted pKK661 to recipient cells upon conjugation (Table 2). From donor cells harboring pKK641 pilR1 or pKK641 pilU1 together with pKK661, however, transfer frequency was negligible. In contrast, the pilR1 and pilU1 mutations did not affect the transfer frequency on solid surfaces (data not shown). E. coli cells harboring pKK661 and either pKK641 pilR1 or pKK641 pilU1 were resistant to phages Ia and PR64FS, while cells harboring wild-type pKK641 and pKK661 were sensitive to them (Table 2). These results indicate that pilR and pilU genes are involved in thin-pilus formation and, consequently, in liquid mating.

**DISCUSSION**

**The R64 thin pilus is a type IV pilus.** The present results strongly suggest that the R64 thin pilus belongs to the type IV family of pili, since (i) the PilS and PilV proteins carry structures homologous to those of type IV prepilins, (ii) the PilU protein carries a structure homologous to that of type IV prepilin peptidase, and (iii) the PilN, PilQ, PilR, and PilT proteins carry structures homologous to those of proteins related to type IV pilus biogenesis. The purified thin pili of R64 or ColIb-P9 were shown to consist of 19- and 48-kDa proteins, which might be the products of pilS and pilV genes, respectively (18, 19, 42). The reduction of the size of the PilS protein in the presence of the pilU gene suggests that the pilU product functions as a prepilin peptidase. The N-terminal amino acid sequences of the R64 PilS and PilV proteins have a slight similarity to each other, but they are not so similar to those of type IVA pilins (Fig. 4E). In addition, the N-terminal amino acids of putative mature pilins of R64 pilS and pilV products are tryptophan. These results indicate that the R64 thin pilus belongs to the type IVB group.

**Gene organization of the R64 pil region.** The gene organization of the R64 pil region was compared to those of the bfp system of enteropathogenic E. coli and the tcp system of *V. cholerae* (26, 33, 34) (Fig. 6). The organizations of these type IVB pilus biogenesis systems reveal some similarity to each other. The relative positions of the PilN, PilQ, and PilR homolog genes are conserved among the three systems. Prepilin peptidase genes are located at the downstream positions of these systems. PilT homolog genes are also located at downstream positions, while the relative positions of prepilin peptidase genes and PilT homolog genes are different between the R64 pil and EAF bfp systems. The locations of prepilin genes are different among these systems. The pilin genes are located at upstream positions in the bfp and tcp systems, while pilin genes are located in the downstream region of the R64 pil system. In the R64 pil region, it is reasonable that the pilV gene is located at the 3'-most position of the pil operon, since its C-terminal segments are under the control of the shufflon. Expression of the bfp, tcp, and pil systems is positively controlled by the bfpT, toxRT, and traBC genes, respectively (9, 10, 15). Most of these regulatory genes are located outside the type IV pilus biogenesis gene cluster.

The similarity of amino acid sequences among various genes, as well as the conservation of gene organization, indicates that these type IV pilin biogenesis systems have evolved from a common ancestral gene system.

**Possible function of the R64 thin pilus in conjugation.** To our knowledge, the R64 thin pilus is the first example of a type IV pilus that is involved in bacterial conjugation. The involvement of at least pilR and pilU genes in liquid mating was demonstrated by the present study. Many pathogenic bacteria use type IV pili to attach themselves to eukaryotic host cells. R64 thin pili may be used to attach donor cells to recipient cells at the first step of liquid mating. R64 thin pili contain PilV proteins, the C-terminal segments of which are under the control of the shufflon DNA rearrangement. Thus, seven different kinds of R64 thin pili with different PilV proteins may be produced and used to attach donor cells to different kinds of
recipient cells to determine recipient specificity in liquid mating of R64 (18, 19). For example, thin pili with PilVA′, PilVC, or PilV′C protein are used to recognize E. coli K-12 recipient cells, and those with PilVA′ or PilVB′ protein are used for Salmonella typhimurium LT2. In R64 surface mating, however, this step can be skipped, since thin pili are not required for solid-surface mating (17, 18). R64 thick pili are required for both solid-surface and liquid mating.

In many pili, an adhesin was shown to be located at the tip (13, 22, 23, 29). R64 PilV proteins may also function as an adhesin. The speculation that the PilV protein is localized to the tip of the R64 thin pili awaits further investigation.

ACKNOWLEDGMENTS

We are grateful to M. Inukai for globomycin. We thank Kathy Takayama for critical reading of the manuscript. This work was supported in part by a grant from the Ministry of Education, Science, and Culture of Japan.

REFERENCES