Cloning and sequencing of two type 4 (N-methylphenylalanine) pilin genes from *Eikenella corrodens*

VENKATARAMA K. RAO and ANN PROGULSKE-FOX*

Department of Oral Biology and Periodontal Disease Research Center, University of Florida, Gainesville, Florida 32610, USA

(Received 10 July 1992; revised 12 October 1992; accepted 27 October 1992)

*Eikenella corrodens* is a Gram-negative, microaerophilic rod which is gaining recognition as an important human pathogen. We have previously reported the cloning and expression in *Escherichia coli* of a 3-6 kb *Eik. corrodens* genomic DNA fragment which encodes a 31.5 kDa haemagglutinin. Maxicell analysis revealed that this fragment also encodes two proteins of approximately 14 kDa. Nucleotide sequencing of the 2-2 kb fragment upstream of the haemagglutinin gene revealed two open reading frames with strong homology to genes encoding pilin subunit proteins of the type 4 or N-methylphenylalanine class. The two pilin genes, ecpA and ecpB, are complete and are expressed in *E. coli*. Southern analysis of ten additional *Eik. corrodens* strains revealed that all possess fragments homologous to ecpA. These data represent the first molecular evidence for pili in *E. corrodens*.

---

**Introduction**

*Eikenella corrodens* is a Gram-negative, microaerophilic rod which colonizes the oral cavity and upper respiratory tract of humans (Henriksen, 1969; Tanner et al., 1979). In recent years, this unusual organism has gained recognition as an opportunistic pathogen (Brooks et al., 1974; DeMello & Leonard, 1979; Stoloff & Gillies, 1986). It has been isolated from abscesses in a variety of sites (Farrington et al., 1983; Decker et al., 1986; Cheng et al., 1988 a, b; Hofstad & Horn, 1989) and is now regarded as an important cause of infections resulting from human bite injuries (Schmidt & Herkman, 1983; Stoloff & Gillies, 1986). In addition, *Eik. corrodens* is associated with certain forms of periodontal disease and may play an important role in the extensive bone loss which accompanies this complex process (Tanner et al., 1979; Dzink et al., 1985).

In his initial description of *Eik. corrodens*, Eiken (1958) noted the unusual ability of colonies of this organism to form small depressions or pits in the agar. Indeed, it was this characteristic for which the species was named. In addition, the colonies were noted to exhibit evidence of peripheral spread under appropriate growth conditions. Darkfield examination of *Eik. corrodens* revealed that while it lacked flagella, it exhibited an unusual form of random movement which is termed twitching motility and which is presumably responsible for spread on agar plates. The earliest descriptions also noted that at a detectable frequency, strains exhibiting all of these characteristics gave rise to variants which could not pit the agar or spread by twitching motility.

Interestingly, many of these characteristics are shared by a number of unrelated organisms, many of which are pathogens, including *Moraxella* spp. (Marrs et al., 1985), *Dichelobacter (Bacteroides) nodosus* (Elleman, 1988) and *Pseudomonas aeruginosa* (Sastry et al., 1985). In these organisms, these microbiological characteristics have been correlated with the expression of pili or fimbriae (Elleman, 1988). More significantly, the presence of these pili has, in the case of *P. aeruginosa*, been definitively linked with the ability to cause disease (Woods et al., 1980; Paranchych & Frost, 1988) and a similar role in pathogenesis has been suspected for the pili of other organisms. The pili of *D. nodosus* (Elleman, 1988), *Moraxella bovis* (Marrs et al., 1985), *Neisseria gonorrhoeae* (Meyer et al., 1984), *N. meningitidis* (Perry et al., 1988), *P. aeruginosa* (Johnson et al., 1986) and *Vibrio cholerae* (Shaw & Taylor, 1990) belong to a family known as the type 4 or *N*-methylphenylalanine ("N-Met-
Phe') class of pili. These pili are homopolymeric structures composed of pilin subunits, each of molecular mass 15–20 kDa. The pilin subunits are synthesized as single polypeptide chains, which typically include an unusually short six- or seven-amino-acid leader sequence at the N-terminus. This leader is proteolytically removed and the resulting N-terminal amino acid methylated, producing the N-methylphenylanalanine for which these pili are named. In addition, these pilins exhibit a striking conservation of amino acid sequence over the N-terminal 30 amino acids. Genes encoding type 4 pilins have been cloned and sequenced from a variety of these organisms, including N. gonorrhoeae (Meyer et al., 1984), M. bovis (Marrs et al., 1985), D. nodosus (Anderson et al., 1984; Finney et al., 1988) and P. aeruginosa (Pasloske et al., 1985). These genes were each expressed in E. coli, but in no case were the subunit proteins assembled into pili, leading to speculation that the systems responsible for assembly of type 4 pili differ greatly from those responsible for pilus construction in E. coli.

Given the cultural characteristics of Eik. corrodens, it has long been suspected that this organism too possesses pili. While early microscopic observations by Jackson et al. (1971) failed to find pilus-like structures consistently on the facultative Bacteroides corrodens (later renamed Eikenella corrodens), in more recent studies, Progulske & Holt (1980) noted the presence of fibrils, 4 nm thick, arranged in a peritrichous fashion around the cells. While these fibrils were thought to be pili, there remained some uncertainty as to whether they might represent artifacts of processing. To date, pili have not been purified from Eik. corrodens. In the accompanying paper we reported the cloning and sequencing of two haemagglutinin genes from Eik. corrodens ATCC 23834 (Rao et al., 1993). Here we report the identification, cloning and sequencing of two type 4 pilin genes from the chromosome of this bacterium.

Methods

**Bacterial strains, plasmids, media and growth conditions.** Eikenella corrodens ATCC 23834 (the type strain) and ATCC 43278 were obtained from the American Type Culture Collection. Strains FDC373, 470 and 1073 are clinical isolates and were obtained from S. S. Soransky, Foryth Dental Center, Boston, Massachusetts, USA. Strains EC-14, 23, 26, 37, 38 and 50, as described previously (Lacroix & Walker, 1991), were isolated from patients with refractory periodontitis and were obtained from Dr Clay Walker, University of Florida. Identification of strains as Eik. corrodens was based on the cell wall fatty acid profile as detected by gas chromatography using the VPI system (Holdeman et al., 1987). Eik. corrodens strains were cultured on blood agar plates in an atmosphere of 10% (v/v) CO2. Broth-grown cells were cultured in BY broth as previously described (Progulske & Holt, 1987). Escherichia coli JM109 (Yanisch-Perron et al., 1985), which was used for all cloning experiments, was cultured aerobically on Luria–Bertani (LB) medium. Plasmid pUC19, used for subcloning experiments, was purchased from Bethesda Research Laboratories. Plasmid pUC18 lacking the deletion mutation (Lobet et al., 1989) was also used for subcloning and was purchased from Pharmacia LKB Biotechnology. Other plasmids are listed in Table 1.

**DNA preparation and manipulation.** Construction and screening of the genomic library of Eik. corrodens ATCC 23834 in E. coli is described in the accompanying paper (Rao et al., 1993). Chromosomal DNA used in Southern hybridization experiments was isolated from Eik. corrodens strains using standard methods (Sambrook et al., 1989). DNA was further purified by multiple rounds of protease treatment, hexadecyltrimethyl ammonium bromide (CTAB) precipitation phenol/chloroform extraction and ethanol precipitation using conventional techniques (Sambrook et al., 1989). Restriction endonucleases were purchased from Bethesda Research Laboratories or Promega and used according to the manufacturer's instructions.

**DNA sequencing.** Fragments to be sequenced were cloned in either pUC19 or pUC18. Double-stranded sequencing was carried out by the University of Florida DNA Core Sequencing Facility. Details are given in the accompanying paper (Rao et al., 1993).

**Maxicell analysis.** This was done as described by Rao et al. (1993).

**Southern hybridization.** Procedures were as described by Rao et al. (1993), except that after hybridization the membrane was subjected to a wash of low stringency (2 x SSPE, 1%, w/v, SDS) for 60 min at 65 °C followed by a moderate stringency wash (0.5 x SSPE, 1%, w/v, SDS) for 15 min at 65 °C, and positively hybridizing bands were visualized by autoradiography for 12–24 h (rather than 6–12 h) at –70 °C.

### Results

**Cloning of two pilin genes from Eik. corrodens**

In the accompanying paper we reported the construction of a genomic clone bank of Eik. corrodens ATCC 23834 DNA in E. coli JM109, and the identification of two E. coli clones with the ability to agglutinate neuraminidase-treated erythrocytes (Rao et al., 1993). As noted previously, clone 1 harboured plasmid pVKR201, which...
Cloning of two E. corrodens type 4 pilin genes

I-Plasmid pVKR201

Proteins encoded (kDa)

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Proteins encoded (kDa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pVKR201</td>
<td>31.5, 14</td>
</tr>
<tr>
<td>pVKR202</td>
<td>31.5, 14</td>
</tr>
<tr>
<td>pVKR203</td>
<td>14</td>
</tr>
<tr>
<td>pVKR205</td>
<td>14</td>
</tr>
<tr>
<td>pVKR207</td>
<td>14</td>
</tr>
<tr>
<td>pVKR208</td>
<td>14</td>
</tr>
<tr>
<td>pVKR209</td>
<td>14</td>
</tr>
</tbody>
</table>

Fig. 1. Localization of the two pilin genes of pVKR201. A restriction map and the relative positions of various subclones are shown. ORF 1 and 2 represent ecpA and B respectively. ORF 3 represents the gene for the 31.5 kDa haemagglutinin (see Rao et al., 1993). Large arrows indicate the positions of each gene as well as the direction of transcription. Small arrows denote the direction of transcription from the plasmid lac promoter. Sizes of proteins encoded (kDa) are as determined by SDS-PAGE analysis of [35S]methionine-labelled maxicells.

-contained a 3.6 kb fragment of DNA from Eik. corrodens. A restriction map indicating the position of the open reading frame (ORF) encoding the haemagglutinin is shown in Fig. 1 (ORF 3). Maxicells containing this cloned DNA produced, in addition to the 31.5 kDa protein with haemagglutination activity, two closely migrating proteins with molecular masses of approximately 14 kDa each (see Fig. 4, lane 3, of Rao et al., 1993). We previously presented evidence suggesting that these three proteins are expressed in E. coli under the control of a cloned Eikenella promoter. Examination of the nucleotide sequence of a 1670 bp fragment containing the haemagglutinin gene allowed the identification of a potential promoter sequence just upstream of the ORF. Since there were no other identifiable ORFs downstream of the haemagglutinin gene, it seemed likely that the two 14 kDa proteins were encoded by genes contained in the 1400 bp upstream of the haemagglutinin gene and that they are expressed under the control of one or more promoters distinct from that controlling expression of the haemagglutinin. Indeed, transformants harbouring pVKR203, which contains a 1.9 kb EcoRI-HindIII fragment, directed the production of a 14 kDa protein as visualized by the maxicell technique (Fig. 1).

Nucleotide sequence of two pilin genes from Eik. corrodens

The complete nucleotide sequence of the 1.9 kb insert of pVKR205 (see Table 1 for description) was determined. Oligonucleotides were synthesized based on the sequences of the ends of pVKR205 and pVKR204 in order to sequence the joint between these two subclones. In doing so, we identified an additional 82 bp fragment which was deleted during the construction of both pVKR204 and pVKR205 due to the presence of an additional EcoRI site which could not be mapped. Fig. 2 shows a portion of the sequence of the coding strand of the insert of pVKR201 (see Fig. 1) extending from the first KpnI site to the terminal HindIII site. This position includes the complete insert of pVKR205, the 82 bp EcoRI–EcoRI fragment and a 426 bp portion of pVKR204 which has been described previously (Rao et al., 1993). We have identified two complete ORFs on this fragment. The first, ORF 1, begins at position 1098 and
Fig. 2. Nucleotide sequence of a 2386 bp fragment from pVKR201 containing two pilin genes, ecpA (ORF 1) and ecpB (ORF 2), from Eik. corrodens. The deduced amino acid sequence is shown below the nucleotide sequence. The positions of a potential -10 promoter sequence, an ntrA-dependent promoter-like sequence (ntr1), and a potential ribosome-binding site S-D 1 preceding ecpA are indicated above the sequence. The positions of an ntrA-dependent promoter-like sequence and a potential ribosome-binding site preceding ecpB are also indicated above the sequence (ntr2 and S-D 2, respectively). Lines above the sequence indicate the positions of inverted repeats which may function in transcription termination.
extends 477 nucleotides to position 1574. At this position there are two directly consecutive TAA termination codons. ORF 1 encodes a polypeptide of 159 amino acids with a molecular mass of 16.2 kDa as deduced from the nucleotide sequence. ORF 2, which is read in the same direction as ORF 1, but in the −1 frame, begins at position 1652 and extends 477 nucleotides to position 2128. At this point, there is a TGA termination codon. There are two additional stop codons 12 nucleotides downstream, TGA and TAA arranged consecutively, from position 2144 to 2149. This ORF, like the first, encodes a polypeptide of 159 amino acids, with a deduced molecular mass of 17.4 kDa. The third ORF, encoding the previously described haemagglutinin (Rao et al., 1993), begins 122 nucleotides downstream at position 2250 and is read in the same direction but in the +1 frame relative to ORF 1.

Transcription of both ORF 1 and ORF 2 occurs in a direction opposite to that of the vector lac promoter as indicated in Fig. 1 and consistent with our finding that expression was independent of orientation of the insert in the vector. ORF 1 is preceded by a strong Shine–Dalgarno sequence (Gold et al., 1981), AGGA, at positions −11 to −14, while ORF 2 is preceded by a somewhat different potential ribosome-binding site, TGCGG, from positions −9 to −13. Upon examination of sequences upstream of ORF 1, we identified, from positions 914 to 919, a hexanucleotide, TATAAT, which is identical to the consensus sequence for the Pribnow or TATA box portion of prokaryotic promoters (Rosenberg & Court, 1979). However, we could identify no sequences with similarity to the classical −35 sequence. Interestingly, from positions 949 to 960 is the sequence TTGGCATA GCCG TTGCT, which differs in only two nucleotides from the ntrA-dependent promoter consensus, CTGG(C/T)A(C/T)(G/A) NNNN TTGCA (Gussin et al., 1986). A search of the 300 bp immediately upstream of this potential promoter region, however, did not reveal any potential NifA or NtrC binding sites (TGTN,ACA) as were identified upstream of the P. aeruginosa pilin gene and found to be important in the regulation of its expression (Pasloske et al., 1989). Determination of which of these two regions represents the actual promoter will await the completion of primer-extension experiments. We could identify no sequences with strong homology to the E. coli promoter consensus sequences immediately upstream of ORF 2. However, from positions 1525 to 1542 is the sequence CTGGTACC GCCG CTGGC, which is somewhat similar to the ntrA-dependent promoter consensus cited above, but much less so than that upstream of ORF 1. It is of note that this sequence occurs within the coding sequence of ORF 1. Once again, there were no sequences typical of NifA or NtrC binding sites. Beginning 21 nucleotides downstream of the TAA termination codon of ORF 1 at position 1598 and extending to position 1625 is a potential stem–loop structure. Similarly, beginning 21 nucleotides downstream of the TGA termination codon of ORF 2 at position 2152 and extending 30 nucleotides to position 2181 is a second potential stem–loop structure. Each of these may be important in the rho-independent termination of transcription of the respective gene (Rosenberg & Court, 1979).

The deduced amino acid sequences of the proteins encoded by ORFs 1 and 2 were nearly identical over the first 35 amino acids. Moreover, this N-terminal sequence was identical to the N-terminus of the Moraxella bovis Q-pilin subunit (Marrs et al., 1985) and strikingly similar to the N-termini of other type 4 pilin subunit proteins. From these data, it was clear that ORFs 1 and 2 encode the pilin subunits of E. corrodens. We have named these genes ecpA and ecpB respectively.

Subcloning of ecpA and ecpB

As we had the complete nucleotide sequence, it was possible to construct subclones containing each pilin gene. Because of the uncertainty regarding the promoter(s) for these genes, both subclones were constructed such that expression could be under the control of the vector lac promoter. The entire ORF of ecpA is contained in pVKR209, in which a 600 bp AvaI–EcoRV fragment extending from position 1041 to 1630 was ligated to the HincII–AvaI sites of pUC18. This subclone contains little flanking DNA and contains no sequences from the ecpB ORF. As noted above, pUC18 was chosen so as to allow transcription from the lac promoter. Only 117 nucleotides separate the end of ORF 2 (ecpB) and the start of ORF 3, which codes for the haemagglutinin. In this fragment, we could identify no unique restriction sites which would allow convenient subcloning of only ecpB, while allowing its insertion in the proper orientation with respect to the plasmid promoter. We therefore ligated a 750 bp EcoRV–KpnI fragment, extending from position 1630 to 2386 to the KpnI–HincII sites of pUC19 in order to construct pVKR208. While this construct contains the entire ecpB gene with little upstream DNA, it does include 257 nucleotides of downstream DNA, including 140 nucleotides from ORF 3, which encode the N-terminal 46 amino acids of the 31.5 kDa haemagglutinin.

Occurrence of ecpA in other strains of Eik. corrodens

In order to determine if these pilin genes are present in strains of Eik. corrodens other than the type strain, ATCC 23834, we analysed the chromosomal DNA of 10 additional strains of E. corrodens using ecpA as a probe.
we have identified two distinct genes for pilin subunits of the type 4 or N-Met-Phe class. The two genes are located on a 3.6 kb HindIII fragment and are separated by only 75 nucleotides. Both appear to be complete genes and have the same orientation. The nucleotide sequences of the two genes are related, with 70% identity over 162 bp with no gaps. However, this similarity is limited to the first 162 nucleotides of each gene, with no significant relationship observed between the sequences beyond this point or upstream of each gene.

As could be predicted from the nucleotide sequence homology, the amino acid sequences of the two proteins are identical over the N-terminal 29 amino acids. Moreover, over the entire sequence, the two proteins are 42.4% identical and can be aligned with only a single gap. This pattern of identity over the N-terminal 30 amino acids is typical of the type 4 family of pilin subunits, which are all strikingly similar at the N-terminus, with great differences occurring further downstream. These differences occur even between the pilin subunits of a single strain, as is observed here for Eik. corrodens.

As could be inferred from the similarities of the amino acid sequences, the two Eik. corrodens pilin genes ecpA and ecpB exhibit strong homology to type 4 pilin genes from a variety of other bacterial species. ecpA is most similar to the Q-pilin gene from Moraxella bovis, with 61.7% identity over a 632 bp overlap. This homology extends over the entire sequence of each gene, beginning approximately 20 bp upstream of the translation start site and extending to approximately 115 bp downstream of each gene. Interestingly, this region includes the first 35 nucleotides of ecpB, although its homologue in M. bovis is not part of an ORF. ecpB is much less homologous to pilin genes from other bacteria. It is, however, similar to the pilin gene from P. aeruginosa strain P1, with 75% identity over a 111 bp overlap (Pasloske et al., 1988). This portion of the gene includes the coding for only the N-terminal 37 amino acids.

A comparison of the proteins encoded by ecpA and B with other type 4 pilin proteins reveals many interesting features. These two pilins can be aligned with pilins from other bacterial genera as shown in Fig. 4. One common characteristic of this group of proteins is the presence of an unusually short leader sequence, which is cleaved followed by methylation of the terminal phenylalanine. The pilin of Vibrio cholerae also undergoes this methylation process, although in this case, the N-terminal residue is methionine instead of phenylalanine (C. Kaufman, personal communication). Although this must be confirmed by isolation of the pilins from Eikenella and determination of their N-terminal sequences, it can be assumed that these pilins, which exhibit strong homology to other type 4 pilins, also undergo similar post-translational processing. If true, pilin A possesses a seven-amino-acid leader sequence as do the pilins of D. nodosus (Elleman, 1988) and N. gonorrhoeae (Meyer et al., 1984), while pilin B possesses a six-amino-acid leader sequence analogous to pilins of

Fig. 3. Occurrence of ecpA in other strains of Eik. corrodens. Restriction endonuclease digests of genomic DNA from various Eikenella strains were probed with 32P-labelled ecpA DNA. Lanes: 1, E. coli JM109 chromosomal DNA digested with HindIII; 2-12, chromosomal DNA from Eik. corrodens strains digested with EcoRV (2, ATCC 23834; 3, EC-14; 4, EC-23; 5, EC-26; 6, EC-37; 7, EC-38; 8, EC-50; 9, FDC373; 10, FDC470; 11, FDC1073; 12, ATCC 43278). The sizes of λ DNA fragments (digested with HindIII and EcoRI) are shown on the left.

Four of these strains, FDC373, FDC470, FDC1073 and ATCC 43278, are well-characterized laboratory strains. The remaining six, EC-14, 23, 26, 37, 38 and 50, are fresh clinical isolates obtained at the University of Florida. As shown in Figure 3, under conditions of moderate stringency, fragments complementary to ecpA were present in EcoRV digests of genomic DNA from the Eik. corrodens type strain, but not in digests of DNA from E. coli JM109. Moreover, strongly hybridizing bands of varying size were detected in all 10 of the other Eik. corrodens strains tested. In many cases, additional bands were also detected.

Discussion

From a clone bank constructed in E. coli JM109 of the chromosomal DNA from Eik. corrodens ATCC 23834, we have identified two distinct genes for pilin subunits of the type 4 or N-Met-Phe class. The two genes are located on a 3.6 kb HindIII fragment and are separated by only 75 nucleotides. Both appear to be complete genes and have the same orientation. The nucleotide sequences of the two genes are related, with 70% identity over 162 bp with no gaps. However, this similarity is limited to the first 162 nucleotides of each gene, with no significant
Moraxella spp. (Marrs et al., 1985) and the majority of P. aeruginosa pilins (Pasloske et al., 1988). The last three residues of both leaders, Q K G, are identical to those of all of the other pilins as shown in Fig. 4. This intraspecies variability in leader sequence length has not been previously reported in type 4 pilins, although different strains of P. aeruginosa have been found to have leaders of different lengths (Pasloske et al., 1988). Most of the other pilins possess two cysteine residues which are believed to be in the oxidized state (Rothbard et al., 1984; Elleman, 1988). The II class pilins from D. nodosus (serotypes D and H) (Elleman, 1988) as well as one strain of P. aeruginosa (K1224) (Pasloske et al., 1988) possess four cysteine residues. Interestingly, both of the Eikenella pilins possess only three cysteine residues, all of which are located close to the C-terminus. The cysteine residues at positions 132 and 151 of pilin A are at positions comparable to those in the Q-pilin of M. bovis (Marrs et al., 1985). Pilin B, like A, possesses two cysteine residues positioned so as to form a disulfide loop of 18 amino acids, although in this case they are located at positions 131 and 150 of the mature polypeptide. Interestingly, while the third cysteine of pilin A is located outside the potential disulfide loop at position 124, that of pilin B, at position 141, is within the loop. The impact of this difference on pilin and/or pilus structure is unclear.

Experimental data suggested that the pilin genes are expressed in E. coli under the control of an Eikenella promoter. Examination of sequences upstream of ecpA revealed homology to the consensus –10 (TATA box) region of bacterial promoters, but no regions homologous to the –35 sequence (Rosenberg & Court, 1979). Such an arrangement is typical of genes whose expression is regulated and is similar to that observed upstream of the pilE transcriptional initiation site in N. gonorrhoeae (Meyer et al., 1984). It is interesting to note that of the two Eikenella type 4 pilin genes that have been sequenced thus far, only the Eikenella pilin (ecpA) and Neisseria pilin (pilE) possess the TATAAT promoter element. As is the case with the remaining pilin genes, no σ7 promoter-like sequences are present upstream of ecpB. Since both proteins are produced in E. coli harbouring pVKR201, the two genes appear to be transcribed together as a polycistronic message. Evidence for this rests in the fact that transformants harbouring pVKR207, which contains the entire ORF of ecpB as well as the 3' half of ecpA, do not produce the pilin protein (data not shown). Therefore, the region between ecpA and B does not contain all of the sequences necessary for expression of ecpB in E. coli. Johnson et al. (1986) determined the transcriptional initiation site of two P. aeruginosa pilin genes and identified sequences similar to ntrA-dependent promoter sequences. Similar sequences were identified upstream of type 4 pilin genes from other genera as well. In addition, Pasloske et al. (1989) identified enhancer sequences upstream of the P. aeruginosa pilin gene. It is interesting that both pilin genes from Eik. corrodens possess ntrA-dependent promoter-like sequences consistent with that observed with other type 4 pilin genes. However, as we were unable to identify the regulatory sequences which function as the targets of either ntrC or nilA, the functionality of these promoter-like sequences is unclear. While it is likely that in E. coli, the σ7-like promoter
upstream of ecpA is used, the other promoters may be of greater significance in Eikenella. Alternatively, the presence of two different types of promoters may be indicative of a mechanism by which the organism differentially expresses these putative virulence factors under varying environmental conditions.

The two pilin genes from Eik. corrodens described here appear to be complete and are transcribed in the same direction. This type of organization appears to be in contrast to that observed for other type 4 pilin genes. Strains of both P. aeruginosa and D. nodosus have each been found to contain only one complete pilin gene (Elleman, 1988). Moraxella spp. have been found to contain two pilin genes, one of which is non-expressed, as it lacks the N-terminal coding regions and is oriented in a direction opposite to that of the promoter. Switching is believed to occur by an inversion event which brings the non-expressed gene into apposition with the N-terminal coding region and in phase with the promoter (Marrs et al., 1988). N. gonorrhoeae, on the other hand, possesses multiple pilin loci, the majority of which are silent as they lack promoter sequences. Laboratory strain MS11ms, which has been characterized extensively (Meyer et al., 1984), also possesses two pilin genes in expression loci. However, in contrast to the case with Eik. corrodens, in which the two pilin genes are separated by only 75 nucleotides, the two pilE loci of N. gonorrhoeae strain MS1ms are separated by approximately 20 kb. Strain MS1ms is apparently unique in that other strains of N. gonorrhoeae possess only one copy of pilE (Meyer et al., 1990). It is not presently known if both types of pili are expressed by Eik. corrodens simultaneously or if some type of switching event similar to that observed for N. gonorrhoeae or M. bovis occurs here. The fact that colony morphology undergoes spontaneous change at high frequency (corroding to non-corroding transitions) is certainly suggestive of some type of switch mechanism.

Examination of codon frequencies for ecpA reveals several interesting features. There is a definite bias against the use of the codons GGG, GGA, AGG, AGA, ATA, CGG, CGA and CTA, which are recognized by minor tRNA species (Grosjean & Fiers, 1982). Of these codons, only CTA appears in ecpA. Similarly, of the codons UCA, UCG, ACC and CCC, which appear only rarely in E. coli genes which are highly expressed, only CCC is used in ecpA. These data are strikingly similar to the codon usage observed for the Q-pilin gene from M. bovis, in which both CCC and CTA are utilized, but the remaining rare codons are excluded (Marrs et al., 1985). The codon usage in ecpB is, however, very different. Indeed, of the 12 codons listed above, only three, CTA, CGA and TCG, are not used in ecpB. This codon usage is in contrast also to that observed for M. bovis (Marrs et al., 1985), P. aeruginosa (Sastry et al., 1985), and D. nodosus (Elleman & Hoyne, 1984). The significance of these differences in codon usage is unclear, but may once again suggest a means of maintaining expression of pilin genes under different growth conditions. Alternatively, these differences may reflect different origins for the two genes. It is interesting to note that while ecpA is homologous at the nucleotide level to the pilin gene from M. bovis, ecpB is more similar to pilins from Pseudomonas spp. Moreover, while the G+C content of ecpA is 48.3 mol %, that of ecpB is 42.7 mol %. Both are quite different from that determined for the chromosomal DNA of Eik. corrodens (56–58 mol %) (Jackson & Goodman, 1984), which suggests that they may have been derived from different sources. It is tempting to theorize that ecpA was obtained by Eik. corrodens from M. bovis or, more likely, from the human pathogen, Moraxella nonliquefaciens (Tenjum et al., 1991), while ecpB came from Pseudomonas spp. Piliation has often been correlated with competence for genetic transformation. Such a mechanism may have been important in the acquisition of additional pilin genes.

Although these genes appear to have been acquired by Eikenella relatively recently, they seem nevertheless to be a consistent part of the genome of this organism. DNA fragments homologous to ecpA were identified in all 10 of the Eikenella strains tested here, despite their diverse origins. As shown in Fig. 3, the EcoRV fragments containing ecpA are of varying sizes, suggesting a heterogeneity of the DNA surrounding this gene. Although several of the strains possess more than one homologous fragment, only one fragment is observed in the type strain. EcoRV was chosen to digest the chromosomal DNA samples since a site for this enzyme is present between the two pilin genes in the type strain. Since the two pilin genes are homologous only at their 5' ends, conditions of much lower stringency than utilized here must be employed to visualize both bands. Multiple homologous fragments are observed in the chromosomal DNA from several of the strains, including EC-14, EC-37, EC-50, FDC470, FDC1073 and ATCC 43278. In some cases, such as with EC-14, EC-37 and EC-50, in which there is one intense and one fainter band, the additional band may represent a second pilin gene more homologous to ecpA than ecpB. It is interesting that DNAs from both FDC470 and ATCC 43278 each contain two weakly hybridizing bands. These strains may possess two pilin genes neither of which is closely related to ecpA, but nevertheless more so than ecpB. Alternatively, unlike the type strain, the pilin genes of these strains may contain an EcoRV site within the coding region.

ecpA and ecpB are the first pilin genes to be cloned from the genus Eikenella, and they represent the first
molecular evidence for the presence of pili in this organism. The fact that fragments homologous to the cloned pilin gene were identified in all of the *Eikenella* strains tested suggests that these proteins may be important in the survival and/or virulence of this bacterium. Our future studies will be directed toward assessing the roles these pili play in the pathogenic process.

This work was supported in part by Public Health Service grant DE-07117 from the National Institute for Dental Research and by NIDR BSRG grant RR05728. V.R. is the recipient of NIH predoctoral training grant DE-07200.

The University of Florida DNA Sequencing Core Laboratory was supported by Biomedical Research grant 110246212 from the University of Florida Division of Sponsored Research and the Interdisciplinary Center for Biotechnology Research.

The authors express their appreciation to Dr C. B. Walker for providing *E. corrodens* strains, Dr J.-M. Lacroix for confirming their identity, and Ms Lucy Deeter for competent technical assistance in these studies.

References


