Proteus mirabilis MR/P Fimbrial Operon: Genetic Organization, Nucleotide Sequence, and Conditions for Expression

FARAH K. BAHRANI and HARRY L. T. MOBLEY*

Division of Infectious Diseases, Department of Medicine, University of Maryland School of Medicine, Baltimore, Maryland 21201

Received 21 December 1993/Accepted 27 March 1994

Proteus mirabilis, an agent of urinary tract infection, expresses at least four fimbrial types. Among these are the MR/P (mannose-resistant/Proteus-like) fimbriae. MrpA, the structural subunit, is optimally expressed at 37°C in Luria broth cultured statically for 48 h by each of seven strains examined. Genes encoding this fimbria were isolated, and the complete nucleotide sequence was determined. The mrp gene cluster encoded by 7,293 bp predicts eight polypeptides: MrpI (22,133 Da), MrpA (17,909 Da), MrpB (19,632 Da), MrpC (96,823 Da), MrpD (27,686 Da), MrpE (19,470 Da), MrpF (17,363 Da), and MrpG (13,169 Da). mrpl is upstream of the gene encoding the major structural subunit gene mrpA and is transcribed in the direction opposite to that of the rest of the operon. All predicted polypeptides share ≥25% amino acid identity with at least one other enteric fimbrial gene product encoded by the pap, fim, smf, fan, or mrk gene clusters.

Proteus mirabilis, commonly associated with nosocomially acquired urinary tract infection (12, 24, 29, 37, 45), expresses a number of proteins that may contribute to virulence, including urease (15), hemolysin (28, 34, 41), flagella (7, 27, 33), and fimbriae (1, 7, 32, 38, 39, 45, 46). Four fimbria-like structures have been purified from P. mirabilis strains for which the major structural subunits have been identified on sodium dodecyl sulfate (SDS)-polyacrylamide gels and subjected to N-terminal analysis (6–8, 20, 43, 46). The gene clusters for two of the fimbriae, Mrp/R (mannose-resistant/Proteus-like) fimbriae and PMF (P. mirabilis fimbiae), have been cloned, and the nucleotide sequences of the major subunit genes, mrpA (9) and pmfA (6), have been reported.

Among the fimbriae of P. mirabilis, MrpR, MR/P, PMF, and PRF fimbiae are perhaps best understood. The presence of 7- to 8-nm-diameter, channelled fimbiae were first observed on bacterial cells capable of hemagglutinating cytotoxins in a mannosine-resistant fashion (31). The MR/P hemagglutination pattern is expressed by virtually all strains (27, 31) and elicits a serum immunoglobulin response following experimental ascending urinary tract infection in mice (7, 20). Here we report the complete nucleotide sequence for the mrp operon, the distribution of the genes among Proteus and other genera, and the conditions under which MrpA is expressed in the wild-type isolate.

We have previously demonstrated that mice, experimentally infected with P. mirabilis, developed a serum immunoglobulin G response to MrpA, the major subunit of MR/P fimbiae, indicating that the fimbiae are expressed in the urinary tract (7). We sought to mimic these parameters by determining the optimal growth conditions for the in vitro expression of MR/P fimbiae. Expression of MrpA by P. mirabilis DR534, isolated from the urine of a woman with catheter-associated bacteriuria (16, 45), was assessed by Western blotting (immunoblotting) using rabbit polyclonal antiserum raised against purified MR/P fimbiae (Fig. 1A). Optimal expression occurred at 37°C in Luria broth cultured statically for 48 h. Very poor expression occurred on Luria agar or in Luria broth that was aerated (shaken, 200 rpm). No expression was detected in aerated liquid minimal salts medium cultured for 48 h. To determine the effect of incubation temperature, static Luria broth cultures were incubated for 48 h at 23, 30, 37, and 42°C. The strongest expression was seen at 37°C, however, a prominent band was observed at 30°C. Very weak expression of MrpA was observed at 23 and 42°C.

Under these optimal conditions (Luria broth, 37°C, 48 h, static), expression of MrpA by seven strains was compared (Fig. 1B). Five strains produced bands of an apparent molecular size of 16.4 kDa that reacted intensely with the antiserum. Two strains, Pm2 and DR534, showed weak reactivity.

Expression of MrpA by Escherichia coli HB101 containing pMRP101, the cosmid clone carrying the mrp gene cluster (9), was not detected under the conditions found optimal for the wild-type isolates. However, expression of the major fimbrial subunit was detected on a Western blot of a cell lysate when E. coli, containing a clone that included mrpI, mrpA, and mrpB, was cultured on Luria agar (data not shown). Under these conditions, the band was easily detectable but did not approach the intensity of the MrpA band produced by the wild-type strain.

To determine the prevalence of the gene in P. mirabilis strains, an oligonucleotide (5'-GGCCGTAGTAATTCTCGTG3') based on nucleotides 304 to 322 of mrpA (9) was end labeled with [γ-32P]ATP (specific activity, >4,000 Ci/mmol; ICN) and used as a probe for hybridization with a Southern blot (23) of PvuII-digested chromosomal DNA from P. mirabilis HI4320 and nine other P. mirabilis strains. The probe recognized a single band of approximately 8 kb from every strain (data not shown).

To determine the presence of gene sequences homologous to mrpA among gram-negative bacterial species, a 277-bp PstI-PvuII restriction fragment, isolated from within the mrpA open reading frame, was labeled with [α-32P]ATP (specific activity, 3,000 Ci/mmol; Amersham) by random priming (9) and used to probe colony blots of a number of strains representing gram-negative bacterial species. All (38 of 38) P. mirabilis strains and one of two Proteus vulgaris strains hybridized with the probe under stringent conditions. None of 40 other isolates representing 12 species (Providencia stuartii,
Providencia rettgeri, Morganella morganii, Klebsiella pneumoniae, Citrobacter freundii, E. coli, Shigella boydii, Salmonella typhi, Yersinia enterocolitica, Aeromonas hydrophila, Enterobacter aerogenes, and Pseudomonas aeruginosa) reacted with the probe.

The nucleotide sequence of mrpA was previously determined from cosmid clone pMRP101 (9). A restriction map of this plasmid was constructed (data not shown), and nine subclones were prepared that spanned approximately 10 kb, 1.9 kb upstream of the start codon of mrpA and 7.3 kb downstream of the termination codon of mrpA. A nucleotide sequence of 9,692 bp was determined in both directions (Fig. 2) and revealed eight open reading frames within the 7,293 bp predicted to encode polypeptides associated with the synthesis of MR/P fimbriae (Fig. 2). The genetic organization reveals seven genes transcribed from left to right and one gene, upstream of mrpA, predicted to be transcribed in the opposite orientation (Fig. 3). One apparently unrelated open reading frame was located about 280 bp downstream of the mrp operon and was encoded by the complementary strand.

Two open reading frames, mrpA and mrpI, were preceded by putative consensus sequence promoters (Fig. 2). No other putative promoters were identified. The mrpA open reading frame is followed immediately by an inverted repeat that may attenuate transcription by formation of a 10-bp stem (9 of 10 match) and a 6-nucleotide loop. The mrpl open reading frame is not followed by such a stem and loop structure. All open reading frames are preceded by putative ribosomal binding sites separated from the start codon by 6 to 9 nucleotides as identified in Fig. 2. mrpC is followed by a short inverted repeat beginning 23 nucleotides after the stop codon.

The G+C content of the nucleotide sequences within mrp open reading frames is 40.6% (36.9% for mrpl, 38.9% for mrpA, 45.7% for mrpB, 41.2% for mrpC, 40.3% for mrpD, 41.4% for mrpE, 38.1% for mrpF, and 41.0% for mrpG). This is consistent with the overall G+C composition of P. mirabilis DNA, which is 39% (11).

Characteristics of the deduced amino acid sequences of the polypeptides predicted by the nucleotide sequence of the mrp operon are summarized in Table 1. With the exception of MrpI, all polypeptides are predicted to begin with typical hydrophobic signal sequences of 20 to 30 amino acids.

The genes are organized in a fashion similar to that of other fimbrial operons (Figure 3). mrpI, transcribed in the direction opposite to that of the rest of the operon, lies upstream of the major structural subunit gene mrpA. mrpB is downstream of mrpA and is followed by mrpC encoding the characteristic outer membrane “platform” or “usher” (10, 30) polypeptide used as the assembly point for the fimbrial structure. Downstream of mrpC is mrpD, predicted to encode a chaperonin that carries the major and minor fimbrial subunits to the platform for assembly. mrpD is followed by putative minor fimbrial subunit genes mrpE, mrpF, and mrpG.

The putative functions of the predicted gene products are based on significant levels of amino acid identity with proteins found in the SwissProt and PIR data bases. All eight predicted polypeptides share ≥25% amino acid identity with at least one other enteric fimbrial gene product encoded by the fim, pap, smf, fan, or mrk gene clusters (Table 1). No significant homologies (>25% amino acid identity) were identified with any nonfimbrial polypeptides in the data bases. MrpI is homologous with both regulatory proteins encoded by the E. coli type 1 fimbrial gene cluster, namely FimB and FimE (18). Both share about 50% amino acid identity with MrpI. The major structural subunit MrpA is most homologous with the major subunit (SmfA) of another mannose-resistant fimbria expressed by Serratia marcescens (26). The middle three gene products, MrpB, MrpC, and MrpD, are most homologous with corresponding pap gene products, PapH (4), PapC (30), and PapD (SwissProt accession number P15319), of the Gal(1-4)Gal-binding P fimbriae of uropathogenic E. coli. The last three gene products, MrpE, MrpF, and MrpG, are most homologous with the corresponding proteins of the S. marcescens mannose-resistant fimbrial operon (25).
been described for uropathogenic P. mirabilis mrp operon. Double-stranded DNA was used as a template for sequencing by the dideoxy chain termination method (3). Reactions were run using reagents from the Prism Ready Reaction Dye Deoxy Termination Kit (Applied Biosystems) in conjunction with Tag polymerase. Reactions were run on a model 373A DNA Sequencer (Applied Biosystems). The complete nucleotide sequence was determined in both directions. Genepro DNA sequence analysis software (Riverside Scientific Enterprises) was used for analysis of the DNA sequence for base composition and identification of open reading frames, restriction sites, and inverted repeats. The deduced amino acid sequence of the open reading frames was analyzed for signal peptides, amino acid composition, and hydropathy. Putative ribosomal binding sites and promoter-like sequences were identified by visual inspection using the guidelines of von Heijne (44). (A) The mprl open reading frame is the inverse complement of nucleotides 618 to 1184 in the sequence depicted in italics enclosed in a box in panel B. mprl precedes mprA and is predicted to be transcribed in the direction opposite to that of the rest of the genes. (B) Complete nucleotide sequence of the mpr operon cluster. Predicted polypeptides are translated below the nucleotide sequence. Gene product designations are given in the right margin. Numbers in the right margin refer to the nucleotide sequence. –35 and −10 mark the closest match to a putative consensus sequence promoter (44). SD (Shine-Dalgarno) marks the putative ribosomal binding site. The initiating codon is shown (44). The leader peptide of MrpB is marked with a double underline. Inverted repeats are underlined.

FIG. 2. Nucleotide sequence of the P. mirabilis mrp operon. Double-stranded DNA was used as a template for sequencing by the dideoxy chain termination method (3). Reactions were run using reagents from the Prism Ready Reaction Dye Deoxy Termination Kit (Applied Biosystems) in conjunction with Tag polymerase. Reactions were run on a model 373A DNA Sequencer (Applied Biosystems). The complete nucleotide sequence was determined in both directions. Genepro DNA sequence analysis software (Riverside Scientific Enterprises) was used for analysis of the DNA sequence for base composition and identification of open reading frames, restriction sites, and inverted repeats. The deduced amino acid sequence of the open reading frames was analyzed for signal peptides, amino acid composition, and hydropathy. Putative ribosomal binding sites and promoter-like sequences were identified by visual inspection using the guidelines of von Heijne (44). (A) The mprl open reading frame is the inverse complement of nucleotides 618 to 1184 in the sequence depicted in italics enclosed in a box in panel B. mprl precedes mprA and is predicted to be transcribed in the direction opposite to that of the rest of the genes. (B) Complete nucleotide sequence of the mpr operon cluster. Predicted polypeptides are translated below the nucleotide sequence. Gene product designations are given in the right margin. Numbers in the right margin refer to the nucleotide sequence. –35 and −10 mark the closest match to a putative consensus sequence promoter (44). SD (Shine-Dalgarno) marks the putative ribosomal binding site. The initiating codon is shown (44). The leader peptide of MrpB is marked with a double underline. Inverted repeats are underlined.
FIG. 2—Continued.
the *E. coli* type 1 fimbrial genes *fimB* and *fimE* (52 and 48% amino acid identity, respectively). These *E. coli* regulatory genes, which are also homologous to one another and are postulated to have arisen from duplication of an ancestral gene, control the phase variation of type 1 fimbrial expression by reversible inversion of a 314-bp region that includes the promoter of *fimH*, which encodes the primary structural unit of the type 1 fimbria.

The major fimbrial subunit, MrpA, most closely resembles the corresponding subunit of the mannose-resistant hemagglutinating fimbriae of *S. marcescens* (57% amino acid identity) (9, 26). This identity is striking and has been the subject of a previous discussion (8).

The gene products putatively involved in fimbria anchoring and biogenesis, encoded by the *mrp* operon, are most closely related to those polypeptides encoded by the *pap* operon of uropathogenic *E. coli*. MrpB is homologous to PapH, a weakly expressed fimbrial protein that has roles in anchoring fimbriae to the cell and in modulating pilus length (4). One reason for poor expression of *papH* is that the gene follows a transcriptional terminating stem-loop structure found between *papH* and *papG* (5). The identical terminator is found in the *P. mirabilis* *mrp* operon, in which *mrpB* follows a transcriptional terminator that is downstream of *mrpA*. Because of the similar positioning of the genes, the mechanism of attenuation of transcription, and the conservation with its *E. coli* homolog, it is likely that MrpB plays a function similar to that undertaken by PapH.

MrpC is most homologous among data base entries to PapC, an outer membrane protein that may form a transmembrane channel through which fimbrial subunits are localized to the surface (10, 30). Such high-molecular-weight proteins are ubiquitous among enteric fimbrial operons, and thus, this observation leads us to believe that MrpC also carries out this function in *P. mirabilis*.

MrpD shares 51% amino acid identity with PapD, a chaperone protein that forms a complex with newly synthesized and secreted fimbrial subunits as an intermediate step in the biogenesis of new fimbriae (13, 40, 42). The predicted amino acid sequence of PapD, a prototype of a family of periplasmic chaperones produced by gram-negative bacteria, aligns with the sequences of these homologs, including MrpD. Of 23 invariantly conserved residues among these proteins (14, 40),

FIG. 2—Continued.
21 align with corresponding amino acids of MrpD. MrpD follows mrpC in the gene cluster just as papD follows papC in the pap operon. Because of the high level of amino acid identity, conservation of invariant residues, and the similarity of genetic organization, we can speculate that MrpD plays the role of the chaperone protein in the assembly of MR/P fimbriae.

The last genes of the mrp operon are most homologous to the smf gene products encoding the adhesin function of the S. marcescens mannos-resistant hemagglutinating fimbriae. In this region, the mrp genetic organization closely resembles that of both smf and pap operons. The last genes to be transcribed in these sequences, smfG and papG, respectively, both encode the adhesins that mediate adherence to receptors. SmfG mediates binding to human erythrocytes (25); PapG binds to Galα(1-4)Gal on the surface of human uroepithelial cells and erythrocytes (22). That MrpE and MrpF represent minor pilins that are chaperoned by MrpD is also supported by the conservation of carboxy-terminal residues (Tyr is the next to last residue in MrpE and MrpF; Gly is the 14th residue from the end in MrpE and MrpF) demonstrated to interact with conserved residues in the chaperone proteins encoded by gram-negative fimbrial gene clusters (19). Assignment of MrpG as the functional adhesin as well as confirmation of function for mrp-encoded gene products await mutational analysis.

TABLE 1. Characteristics of predicted MR/P fimbrial polypeptides

<table>
<thead>
<tr>
<th>Polypeptide</th>
<th>Predicted mol wt of unprocessed polypeptide (no. of amino acids)</th>
<th>Net charge</th>
<th>Similar polypeptidesa</th>
</tr>
</thead>
<tbody>
<tr>
<td>MrpI</td>
<td>22,133 (188)</td>
<td>+9</td>
<td>E. coli</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FimB</td>
</tr>
<tr>
<td>MrpA</td>
<td>17,909 (175)</td>
<td>−5</td>
<td>E. coli</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FimE</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>S. marcescens</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SmfA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>K. pneumoniae</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FimA</td>
</tr>
<tr>
<td>MrpB</td>
<td>19,632 (182)</td>
<td>−1</td>
<td>E. coli</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PapH</td>
</tr>
<tr>
<td>MrpC</td>
<td>96,823 (871)</td>
<td>−1</td>
<td>E. coli</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PapC</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>K. pneumoniae</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MrkC</td>
</tr>
<tr>
<td>MrpD</td>
<td>27,886 (252)</td>
<td>+6</td>
<td>E. coli</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FanD</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>E. coli</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PapD</td>
</tr>
<tr>
<td>MrpE</td>
<td>19,470 (181)</td>
<td>+2</td>
<td>S. marcescens</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SmfE</td>
</tr>
<tr>
<td>MrpF</td>
<td>17,362 (161)</td>
<td>−5</td>
<td>E. coli</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PapF</td>
</tr>
<tr>
<td>MrpG</td>
<td>13,169 (117)</td>
<td>+3</td>
<td>S. marcescens</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SmfE</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SmfF</td>
</tr>
</tbody>
</table>

* Values for unprocessed polypeptide.

b All polypeptides found to have >25% amino acid identity with predicted Mrp polypeptides are included. Sequences were located in the PIR and SwissProt data bases by using the GCG Wordsearch program.

c GenBank nucleotide sequence accession number.
Nucleotide sequence accession number. The sequence for the \textit{mpf} operon of \textit{P. mirabilis} H14320 has been deposited with GenBank under accession number Z32686.

We thank Rachel McDuff and George Massad for constructing a clone necessary for filling a small, but troublesome, sequencing gap and Michael Island for assistance with homology searches and figure preparation. This work was supported by Public Health Service grants AI23328 and AG4393 from the National Institutes of Health.

REFERENCES


43. Tolson, D. Personal communication.

