Cloning, Nucleotide Sequencing, and Expression in *Escherichia coli* of a *Rhizobium leguminosarum* Gene Encoding a Symbiotically Repressed Outer Membrane Protein

RUUD A. DE MAAGD,†* INE H. M. MULDERS, HAYO C. J. CANTER CREMERS,
AND BEN J. J. LUGTENBERG

*Leiden University, Department of Plant Molecular Biology, Botanical Laboratory,
Nonnensteeg 3, 2311 VJ Leiden, The Netherlands*

We describe the cloning of a gene from *Rhizobium leguminosarum* biovar viciae strain 248 encoding protein IIIa, the 36-kDa outer membrane protein forming a part of the outer membrane protein antigen group III. The expression of this antigen group is repressed in the bacteroid form during symbiosis (R. A. de Maagd, R. de Rijk, I. H. M. Mulders, and B. J. J. Lugtenberg, J. Bacteriol. 171:1136–1142, 1989). A cosmid clone expressing the strain 248-specific MAb38 epitope of this antigen group in a nonrelated strain was selected by a colony blot assay. Sequencing revealed one large open reading frame encoding a 39-kDa protein. N-terminal amino acid sequencing of the purified 36-kDa outer membrane protein IIIa revealed that the isolated gene, now designated *ropA*, is the structural gene for this protein and that the mature protein was formed by processing of the 22-residue N-terminal signal sequence. The gene is preceded by a promoter that was active in *R. leguminosarum* but not in *Escherichia coli*. This promoter, which showed no homology to known promoter sequences, was located approximately by determination of the transcription start site. The region upstream of the putative promoter was shown to contain two potential binding sites for integration host factor protein. Expression of protein IIIa under control of the inducible *lac* promoter in *E. coli* shows that, of its earlier described properties, the peptidoglycan linkage of protein IIIa is specific for *R. leguminosarum* but that outer membrane localization and calcium-stabilized oligomer formation can to a large extent also occur in *E. coli*.

During the establishment of a successful symbiosis with their leguminous host plants, bacteria of the family *Rhizobiaceae* undergo a number of changes that result in the formation of so-called bacteroids in the cytoplasm of the infected plant cells. At the molecular level, a large number of changes occur, of which only a few have been characterized in some detail.

Because of its now well-established role in symbiosis, the bacterial cell envelope and the changes occurring in it during bacteroid development have received increased attention. For *Rhizobium leguminosarum* bv. viciae-pea symbiosis, changes during bacteroid development have been described for both major outer membrane constituents, i.e., lipopolysaccharides and proteins (4, 23, 27). During the establishment of symbiosis, some lipopolysaccharide epitopes may disappear (4, 9), whereas new ones appear (23, 27). We have shown that of the four major outer membrane protein antigen groups that could be defined in free-living bacteria (groups I through IV), the group II and III antigens have almost disappeared in bacteroids isolated from pea nodules (4).

Antigen group III of *R. leguminosarum* bv. viciae strain 248 consists of a group of outer membrane proteins with apparent molecular masses ranging from 36 to 46 kDa; these proteins all react in Western blots with three different monoclonal antibodies (4). In an earlier study we showed that this group probably consists of not more than two different major proteins with apparent molecular masses of 36 and 40 kDa, respectively. The other protein bands, which

appear on protein gels and Western blots only after treatment of the isolated cell envelopes with lysozyme, are probably derivates of the two major proteins that contain increasing numbers of peptidoglycan residues (6). We have also shown that the group III proteins form oligomers, as do the outer membrane porins of other gram-negative bacteria (17). However, in *R. leguminosarum* these oligomers are extremely resistant to denaturation by sodium dodecyl sulfate (SDS) at 100°C in the presence of calcium (6). These oligomers are probably the native form of the protein in the intact bacterium, because MAb38, a monoclonal antibody that preferentially recognizes these oligomers on Western blots, can also bind the epitope on the surface of intact cells (4).

To be able to study the regulation of expression of outer membrane components during symbiosis, we isolated and determined the nucleotide sequence of the gene encoding the smallest of the two major proteins of the group III antigens of *R. leguminosarum* bv. viciae strain 248, i.e., the 36-kDa outer membrane protein here named protein IIIa. Moreover, the protein was expressed in *Escherichia coli* under control of the inducible *lac* promoter, and we studied the localization of the gene product and its interactions with Ca²⁺ ions and peptidoglycan in this system.

MATERIALS AND METHODS

Strains and plasmids. The relevant strains and plasmids used in this study are listed in Table 1.

Construction of a cosmid library of *R. leguminosarum* bv. viciae strain 248 DNA. Total DNA was isolated from strain 248 as described by Meade et al. (19). After digestion with *Eco*RI, the DNA was ligated into broad-host-range cosmid

---

* Corresponding author.
† Present address: Plant Biology Laboratory, Salk Institute for Biological Sciences, P.O. Box 85800, San Diego, CA 92138-9216.

TABLE 1. Strains and plasmids used in this study

Strain or plasmid	Relevant characteristics	Source or reference
Rhizobium leguminosarum		
248	Wild type bv. viciae strain	13
RBL5560	Sym plasmid-cured bv. trifolii strain RCR5, Rif^r^, containing pRL1JI	12, 31
RBL5560-LS47	RBL5560::Tn5spec; LPSI^−	1a
Escherichia coli		
KMBL1164	Δ(*lac-pro*) *thi* F^−	P. van der Putte
JM101	Δ(*lac-pro*), *supE thi* (F' *traD36 proAB lacI^q^lacZdel*M15)	32
DH5α F'	Δ(*lacZYA-arg*) *supE thi recA' lacZdel*M15	Promega
Plasmids		
M13tg130	M13 sequencing vector	14
pLAFR1	Broad-host-range cosmid vector	7
pic19R	Multicopy cloning vector	18
pRK2013	IncColE1, helper plasmid in tripartite mating	7
pMP92	Broad-host-range (IncP) vector	24
pMP220	Promoter-probe vector containing a promoterless *lacZ* gene	24
pMP2201	pLAFR1 clone of genomic DNA of strain 248, insert of 20 kb	This study
pMP2202	Subclone of pMP2201 in pMP92 containing one 2.05-kb *Eco*RI fragment	This study
pMP2206	Subclone of pMP2201 in pMP92 containing a 1.6-kb *Eco*RI-*Bam*HI fragment	This study
pMP2209	Subclone of pMP2201 in pMP220 containing a 1.1-kb *Eco*RI-*Kpn*I fragment	This study
pMP2210	Subclone of pMP2201 in pMP220 containing a 0.9-kb *Eco*RI-*Kpn*I fragment; contains the IIIa promoter	This study
pMP2241	Subclone of pMP2201 in pic19R containing the *ropA* gene under control of the *lac* promoter	This study

pLafRI and then introduced into *E. coli* KMBL1164 by using a phage packaging system (10).

DNA isolation and plasmid constructs. Recombinant DNA techniques were carried out essentially as described by Sambrook et al. (21). Broad-host-range plasmids were mobilized from *E. coli* to *R. leguminosarum* by using pRK2013 as a helper plasmid. Selection of transconjugants was done on YMB medium (26) with 2.0 mg of tetracycline and 20.0 mg of rifampin per liter.

DNA sequencing. DNA sequencing was performed by using the dideoxy-chain termination method (22) with M13tg130 (14) and the Sequenase 2.0 kit. For regions with strong secondary structures, dITP instead of dGTP in the termination reactions and gels supplemented with 50% deionized formamide were used.

Determination of transcriptional start site. mRNA was isolated from bacteria grown on TY medium (1) as described previously (28). The primer was labeled with polynucleotide kinase using standard methods (21). The ^{32}P-labeled primer (1 pmol) was annealed with 30 μg of RNA in 10 mM Tris-HCl (pH 8.3)–250 mM KCl, denatured for 3 min at 80°C, and then annealed for 45 min at 45°C. Primer extension was performed with 27 U of avian myeloblastosis virus reverse transcriptase in a buffer containing 25 mM Tris-HCl (pH 7.6), 10 mM MgCl$_2$, 5 mM dithiothreitol, 50 μg of actinomycin D per ml, and each of the four deoxynucleotides at 0.5 mM for 45 min at 45°C.

Cell envelope isolation. Before cell envelopes were isolated, *Rhizobium* cells were grown in liquid TY medium. After the cells were broken by sonication, the cell envelopes were isolated and treated with lysozyme as described previously (5, 6).

Cell fractionation. *E. coli* cells grown overnight at 37°C in LB medium were harvested and cell envelopes were isolated as described before (5). The soluble protein fraction obtained after cell envelopes were pelleted was precipitated with 5%

trichloroacetic acid for 1 h on ice. Fractionation of *E. coli* cell envelopes into outer and cytoplasmic membranes was performed by the method of Osborn et al. (20).

Electrophoresis and Western blotting. SDS-polyacrylamide gel electrophoresis was performed as described previously with 11% acrylamide gels (16). Suspensions of either lysozyme-treated or nontreated cell envelopes were mixed with concentrated sample buffer and incubated at the appropriate temperature for 10 min before electrophoresis. Separated cell envelope constituents were transferred from gels to nitrocellulose by electroblotting. Immunodetection with monoclonal antibodies and alkaline phosphatase-conjugated rabbit anti-mouse immunoglobulin serum (Sigma) was performed as described elsewhere (4).

Colony blotting. For detection of the MAb38 antigen on intact cells, bacteria were grown on YMB agar. Plates were overlaid with dry nitrocellulose sheets, which were allowed to become completely wet before they were lifted. Excess bacteria and slime were washed off with running tap water. For immunodetection, colony blots were treated as described above for Western blots, except that here horseradish peroxidase-conjugated rabbit anti-mouse immunoglobulin serum (Sigma) was used as the second antibody and 5,5',3,3'-tetramethylbenzidine was used as the substrate.

Amino acid sequencing. Protein was isolated by electroelution from acrylamide gels. The eluted protein was run again on a 11% acrylamide gel and blotted onto a polyvinylidene difluoride membrane (Immobilon-P; Millipore, Bedford, Mass.). After blotting, the membrane was washed overnight in 50% methanol and then stained with 0.1% Coomassie brilliant blue R in 50% methanol. The band corresponding with the purified 36-kDa protein was excised, and sequence analysis was performed with a Applied Biosystems pulsed liquid sequencer type 475A with an on-line PTH analyzer type 120A. The material was applied into the special blot cartridge (Applied Biosystems).

Enzymes and chemicals. A Sequenase version 2.0 kit was obtained from RIH (Capelle a/d IJssel, The Netherlands). Polynucleotide kinase and reverse transcriptase were obtained from Promega Biotech (Leiden, The Netherlands). All other enzymes were purchased from Boehringer (Mannheim, Germany) and Pharmacia LKB (Woerden, The Netherlands). Synthetic primers for sequencing and primer extension were obtained from Pharmacia LKB. α-^{35}S-dATP and [γ-^{32}P]dATP were purchased from Amersham International plc. (Amersham, United Kingdom). All enzymes were used according to the specifications of the manufacturers.

Nucleotide sequence accession number. The sequence shown in Fig. 3 has been submitted to the GenBank/EMBL data base under accession number M69214.

RESULTS

Isolation of a cosmid clone involved in the production of outer membrane antigen group III. For isolation of a clone involved in production of antigen group III of *R. leguminosarum* bv. viciae strain 248, we made use of the strain-specific reactivity of one of the monoclonal antibodies that recognize the antigen group in this strain. MAb38 appears to recognize this group specifically in strain 248 and not in any of the more than 20 other *R. leguminosarum* strains tested (3a). Moreover, MAb38 recognizes its epitope on intact cells, thus allowing detection of correct expression of the epitope in a colony blot assay (4). Therefore we used MAb38 to screen a cosmid library of strain 248 for expression of the MAb38 epitope in transconjugants of an unrelated *R. leguminosarum* strain.

The cosmid library, containing partial *Eco*RI digests of total strain 248 DNA cloned in pLafR1, was crossed into strain RBL5560-LS47, and transconjugants were transferred to fresh YMB-agar plates. The transconjugants were subsequently screened for expression of the MAb38 epitope by using colony blotting and detection with MAb38. Screening of the first 1,200 transconjugants resulted in the selection of one transconjugant that reacted with MAb38. The reaction on a blot of YMB-agar-grown bacteria of this transconjugant containing a cosmid designated as pMP2201 compared with those of the parent strain 248 and the recipient RBL5560-LS47 is shown in Fig. 1A. Comparison of cell envelope constituents by Western blotting with MAb38 is shown in Fig. 1B. In samples of lysozyme-treated cell envelopes incubated at room temperature before electrophoresis, the transconjugant (Fig. 1B, lane 3) contained reactive high-molecular-weight oligomers similar to those of parent strain 248 (lane 1), whereas the recipient strain showed no reaction at all (lane 2). In samples heated for 10 min at 95°C before electrophoresis, the transconjugant (lane 6) contained a 36-kDa protein that reacted with the antibody and also two to three bands of slightly higher apparent molecular mass that reacted with decreasing intensity, together forming the lower half of the group III antigens of strain 248 (lane 4). Again, the recipient showed no reaction (lane 5). We therefore concluded that the isolated transconjugant produces part of the original group III antigens of strain 248, i.e., the smallest of the two major protein constituents (the 36-kDa protein here called protein IIIa) and its derivatives substituted with peptidoglycan residues.

Nucleotide sequencing, identification of an open reading frame, and determination of the direction of transcription. Subcloning of the inserted fragments of pMP2201 into IncP vector pMP92 resulted in the selection of a clone, pMP2202, containing a single *Eco*RI fragment of 2.05 kb; a restriction
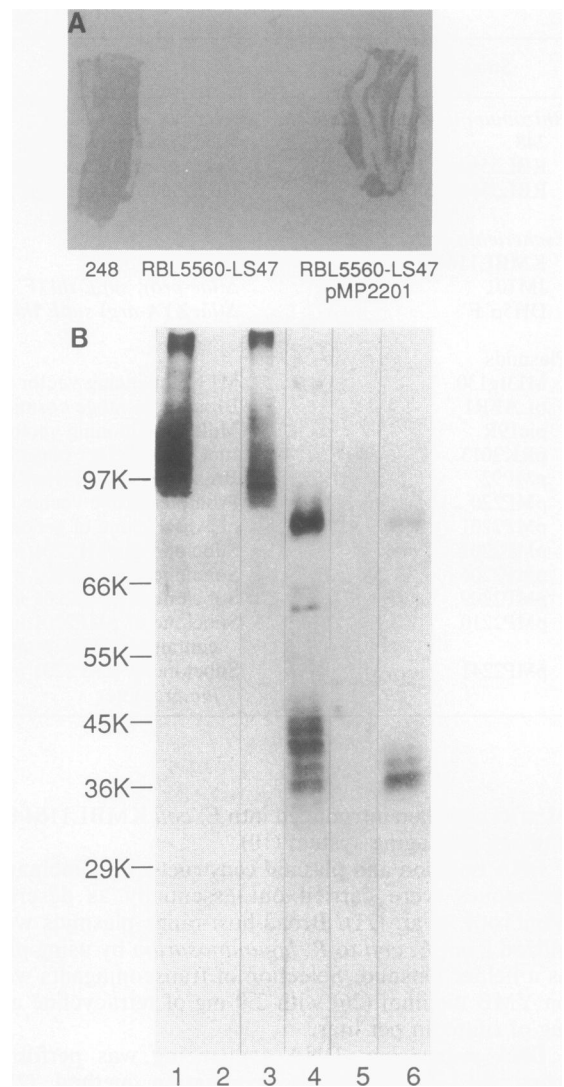

FIG. 1. (A) Blot of YMB-agar-grown wild-type strain 248, recipient strain RBL5560-LS47, and the RBL5560-LS47 transconjugant containing pMP2201. Bacteria were grown in an almost confluent layer rather than individual colonies. The blot was incubated with MAb38. (B) Western blot of cell envelope components of strains 248 (lanes 1 and 4), RBL5560-LS47 (lanes 2 and 5), and RBL5560-LS47 containing pMP2201 (lanes 3 and 6). Before electrophoresis, samples were either kept at 20°C (lanes 1 through 3) or heated for 10 min at 95°C (lanes 4 through 6). The positions of molecular weight markers are indicated at the left.

site map of this fragment is shown in Fig. 2. Still further deletion of the cloned fragment resulted in a 1.6-kb *Eco*RI-*Bam*HI fragment, pMP2206, which was of the minimum length still sufficient for production of the MAb38 epitope. This fragment was sequenced according to the strategy shown in Fig. 2.

The resulting nucleotide sequence of the 1.6-kb *Eco*RI-*Bam*HI fragment is given in Fig. 3. A search of the sequence for possible protein-encoding regions with the codon preference program of the University of Wisconsin GCG Software revealed only one large open reading frame running from nucleotides 407 to 1504 (Fig. 2), which was named ORF38.

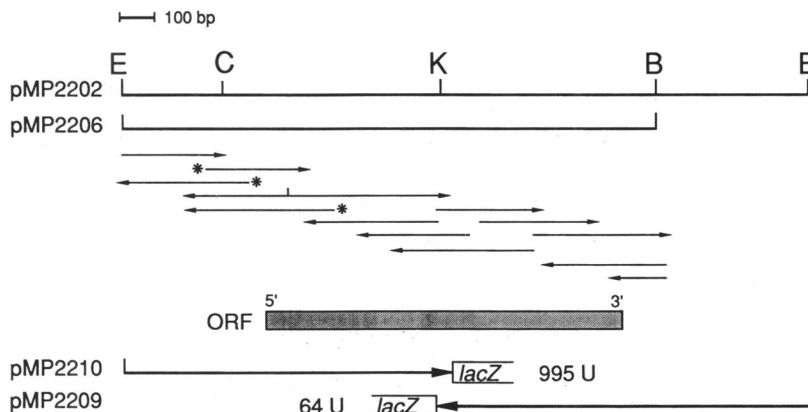To determine the direction of transcription of a putative

FIG. 2. Restriction map of the 2.05-kb *Eco*RI fragment of pMP2202 and strategy of the nucleotide sequence determination. All partial sequences (indicated by arrows) were derived from deletions in M13tg130 by using internal restriction sites, except those marked by an asterisk, for which synthetic primers were used. The shaded box indicates the position and the 5' and 3' ends of the identified open reading frame ORF38. The thick arrows indicate the orientations of the *Eco*RI-*Kpn*I subfragments of pMP2201 in pMP2209 and pMP2210, respectively, relative to the *lacZ* gene of the promoter-probe plasmid pMP220 (see the text). Abbreviations: E, *Eco*RI; B, *Bam*HI, C, *Cla*I; K, *Kpn*I.

gene within the cloned fragment of pMP2202, subfragments were cloned in the promoter-probe plasmid pMP220, which contained a promoterless *lacZ* gene. Of two plasmids containing the complementary halves (*Eco*RI-*Kpn*I fragments) of the *Eco*RI fragment cloned in opposite directions (pMP2209 and pMP2210, respectively, Fig. 2), only pMP2210 gave a significantly elevated β-galactosidase activity in strain 248 (995 U) compared with the background activity of pMP220 in this strain (110 U). Thus the observed direction of transcription, from left to right in Fig. 2, is consistent with the orientation of the open reading frame ORF38. Neither of the two plasmids gave a significantly elevated β-galactosidase activity in *E. coli*, showing that the putative promoter on pMP2210 is not active in this species.

ORF38 encodes protein IIIa. The predicted amino acid sequence of the protein encoded by ORF38 is shown in Fig. 3. It encodes a protein consisting of 366 amino acids with a predicted molecular mass of 39 kDa. The initiation codon is preceded by a putative ribosome binding-site at positions −17 to −12 relative to the initiation codon (Fig. 3). Analysis of the amino-terminal region, with the rules of Von Heijne for prediction of N-terminal signal sequences (29), revealed the very likely presence of such a signal sequence with a predicted processing site between residues 22 and 23. Thus, the predicted molecular mass for the processed protein, 37 kDa, is comparable to the apparent molecular mass (36 kDa) of protein IIIa.

To determine whether the isolated gene was, indeed, the structural gene for the latter protein, we isolated the 36-kDa outer membrane protein for N-terminal amino acid sequence analysis. The first nine amino acids of the protein were determined and were shown to be identical to amino acid residues 23 through 31 in the sequence predicted from ORF38. It was therefore concluded that the identified open reading frame is, indeed, coding for the outer membrane protein IIIa and that export of the protein through the cytoplasmic membrane is accompanied by processing of the signal sequence between residues 22 and 23 of the immature protein product. We propose the name *ropA* (*Rhizobium* outer membrane protein) for this gene encoding outer membrane protein IIIa.

Characterization of the untranslated upstream sequences and determination of the transcription start site. The untrans-

lated part of the *Eco*RI fragment of pMP2202 upstream of the open reading frame consists of 406 bp without any obvious homology to established consensus sequences for promoters. To determine the approximate location of the promoter, we therefore identified the 5' end of the mRNA by primer extension. This experiment, with the middle of the three synthetic primers shown in Fig. 2 (bases 386 through 400 of the sequence shown in Fig. 3), revealed two major extension products differing in length by one base (Fig. 4). We could thereby pinpoint the transcription start site at bases 299 and 300 of the DNA sequence (Fig. 3). The relatively long untranslated 5' leader of the mRNA (107 to 108 bp) appeared to contain an imperfect inverted repeat at positions 341 to 381 of the DNA sequence, which may play a role in the stability of the messenger.

The upstream region contains a very A+T-rich region from positions 166 to 212 that contains two sequences (thick arrows in Fig. 3), in opposite directions, with substantial homology with the recognition site for *E. coli* integration host factor (IHF) protein. The alignment of these sequences with the consensus sequence for IHF binding (2, 15) is shown in Fig. 5. No further homologies to recognition sites for regulatory proteins could be identified.

Expression and localization of protein IIIa in *E. coli*. To establish whether the earlier identified properties of the group III antigens, localization in the outer membrane, peptidoglycan attachment, and calcium-stabilized oligomer formation, are specifically found in *R. leguminosarum* only or are intrinsic properties of the protein IIIa and therefore are also found in other bacteria producing the protein, we expressed the protein in *E. coli* under control of the inducible *lac* promoter. The *Cla*I-*Bam*HI fragment of pMP2202, containing the open reading frame as well as most of the untranslated leader, was cloned behind the *lac* promoter in the multilinker of pic19R, resulting in plasmid pMP2241. In colony blots of this clone, grown without glucose repression (LB without glucose) or under inducing conditions (LB with isopropyl-β-D-thiogalactopyranoside [IPTG]), the MAb38 epitope was detected under these conditions, suggesting that the epitope is expressed at the cell surface of *E. coli* (data not shown). Western blots of cell envelope constituents of repressed, nonrepressed, and IPTG-induced cultures of *E. coli* cells containing pMP2241 (Fig. 6A, lanes 1, 2, and 3,

```
   1  GAATTCAGGACACTATGTTCTAAACCGCCACCTTGTTTGTGCTTGTTAGCTGGCGGTAGA

  61  GTTTATGCGCTTCACCAGTATGTTCGACGGGGTTCGGACGGTGGAGAATTTGACAATGGG

 121  TGTAGCGTGTGCGGCTTGCGATGGAGGATTTCCGTCGCGGGGCCGTTTTGTACGTTTTT

 181  TGTATAAATATTTGAAATTGTTGTATAAAAATCACGTTGCCGGGGAAAGTAGCCCGTGAG

 241  AAGCCATCTGGCGTACATCTACAATTGCTTGATAGAGCCTTGTATGTGATGACGGTTTGG

 301  CGGTTCGATCGATCCGCACGCCGAGATGGCTGTTCAAACTGGGCCGGCTTCAAATCGACT

 361  ACCCGATCTGAAGCCAGCCCCGATCGACAAAGGAATGGATTGGTAAATGAACATCAGAAT
                                                  M  N  I  R  M

 421  GGTTTTGCTTGCATCAGCAGCAGCATTTGCTGCATCGACGCCGGTTCTTGCAGCTGACGC
       V  L  L  A  S  A  A  A  F  A  A  S  T  P  V  L  A  A  D  A

 481  TATCGTTGCTGCCGAGCCGGAACCGGTTGAATATGTTCGCGTCTGCGACGCTTACGGCAC
       I  V  A  A  E  P  E  P  V  E  Y  V  R  V  C  D  A  Y  G  T

 541  TGGCTACTTCTACATCCCGGGCACCGAAACCTGCCTCAAGATCGAAGGCTACATCCGTTT
       G  Y  F  Y  I  P  G  T  E  T  C  L  K  I  E  G  Y  I  R  F

 601  CCAGGTCAACGTTGGCGACAACCCAGGTGGTGACAACGACTCTGATTGGGATGCAGTGAC
       Q  V  N  V  G  D  N  P  G  G  D  N  D  S  D  W  D  A  V  T

 661  CGCGGTCAGGTTCAGTTCACGCAAGAGCGACACCGAGTATGGTCCGCTGACCGGCGTCAT
       A  V  R  F  S  S  R  K  S  D  T  E  Y  G  P  L  T  G  V  I

 721  CGTCATGCAGTTCAATGCTGACAATGCCAGCGATCAGGATGCCATCCTCGACTCCGCTTA
       V  M  Q  F  N  A  D  N  A  S  D  Q  D  A  I  L  D  S  A  Y

 781  CCTCGACGTCGCGGGCTTCCGCGCCGGTCTGTTCTACAGCTGGTGGGACGATGGTCTCTC
       L  D  V  A  G  F  R  A  G  L  F  Y  S  W  W  D  D  G  L  S

 841  TGGCGAAACCGACGACATCGGTTCGGTCGTAACGCTCCACAACTCGATCCGCTATCAGTA
       G  E  T  D  D  I  G  S  V  V  T  L  H  N  S  I  R  Y  Q  Y

 901  TGAAAGCGGCACCTTCTACGCCGGCCTCAGCGTCGATGAACTGGAAGACGGCGTTTACCA
       E  S  G  T  F  Y  A  G  L  S  V  D  E  L  E  D  G  V  Y  Q

 961  GGGTACCTTTACTCCCGGCGTCATTCCCGGCACCACCGACTTTACTGCGGACGATGGTCC
       G  T  F  T  P  G  V  I  P  G  T  T  D  F  T  A  D  D  G  P

1021  GAACAATGTCGGCGTTGCCTTCGGCATCGGCGGCACTGCCGGCGCATTCAGCTACCAGGT
       N  N  V  G  V  A  F  G  I  G  G  T  A  G  A  F  S  Y  Q  V

1081  CACTGGCGGCTGGGACGTCGACAACGAAGACGGCGCTATCCGTGCAATGGGTACGGTTGA
       T  G  G  W  D  V  D  N  E  D  G  A  I  R  A  M  G  T  V  E

1141  AATCGGTCCCGGCACGTTCGGCCTCGCCGGCGTATACTCTTCCGGCCCGAACTCCTACTA
       I  G  P  G  T  F  G  L  A  G  V  Y  S  S  G  P  N  S  Y  Y

1201  CTCTTCAGCTGAATGGGCTGTCGCTGCCGAATACGCTATCAAGGCAACCGACAAGCTCAA
       S  S  A  E  W  A  V  A  A  E  Y  A  I  K  A  T  D  K  L  K

1261  GATCACCCCGGGCCGGTGGCACGGACACGTTCCGGAAGACTTCGACGGTCTCGGCGATGC
       I  T  P  G  R  W  H  G  H  V  P  E  D  F  D  G  L  G  D  A

1321  TTGGAAGGTTGGTCTGACGGTTGATTACCAGATCGTCGAAAACTTCTACGCCAAGGCTTC
       W  K  V  G  L  T  V  D  Y  Q  I  V  E  N  F  Y  A  K  A  S

1381  GGTTCAGTACCTCGATCCGCAAGATGGCGAAGACTCCACTTCGGGCTACTTCGCCTGCAG
       V  Q  Y  L  D  P  Q  D  G  E  D  S  T  S  G  Y  F  A  C  S

1441  CGTGCGTTCTAATCACTTGGTGGATGCTCCCGGTCTTCGGATCGGGAGCACTACAATCAG
       V  R  S  N  H  L  V  D  A  P  G  L  R  I  G  S  T  T  I  S

1501  TTTCTGATCCGAGTGAACCTCCGATCAGCTACGGGGACGGGTCCGGTCTTCCTGCCGGGC
       F

1561  CGTCCTTGCAGTGTTTCCGGCGTACCGTCTTCAAGGCAGTCAATCACGATGCCTTCAGGT

1621  GAGCCGAACCGGTGCGGTGTGTATTGCACTCGGATCCC
```

FIG. 3. Nucleotide sequence of the *Eco*RI-*Bam*HI fragment of pMP2202 and the predicted amino acid sequence of the open reading frame. The cleavage site of the precursor protein is indicated by a vertical arrow. A putative ribosome-binding site preceding the open reading frame is underlined. The transcription start sites are indicated by asterisks. The A+T-rich region containing the IHF recognition sites is indicated by a dotted line. The putative IHF recognition sites, one on each DNA strand, are indicated by two thick arrows each on either site of the nucleotide sequence. The inverted repeats in the untranslated leader RNA are indicated by thin arrows.

respectively) show the IPTG-inducible production of a set of cell envelope proteins reacting with MAb38. In these samples, which were not heated before application on the gel, one can distinguish two proteins of 36 and 38 kDa. The lower of these two comigrated with protein IIIa in a sample of cell envelope constituents of *R. leguminosarum* biovar viciae strain 248 (data not shown) and therefore most likely represents the mature protein IIIa. Moreover, a reaction with a higher-molecular-mass smear, containing one major band with an apparent molecular mass of approximately 80 kDa, was observed in these cell envelope fractions (lanes 2 and 3). The appearance of these high-molecular-weight cross-reacting proteins resembles that of the high-molecular-weight oligomers of the group III proteins observed in *R. legumi-*
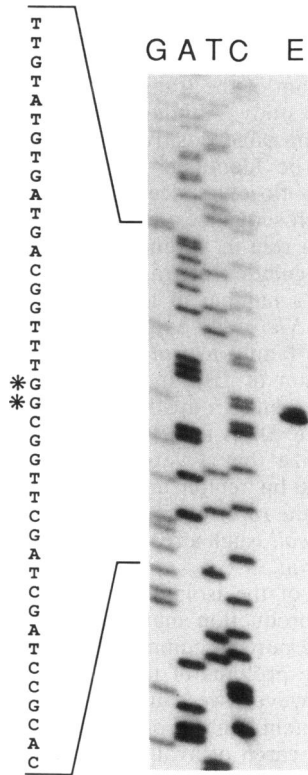
FIG. 4. Determination of the transcription start site by primer extension. The primer used was the middle of the three synthetic primer shown in Fig. 2 (bases 386 through 400 of the sequence shown in Fig. 3). The products of the extended primer (E) in the right lane were run alongside a sequence ladder made with the same primer (G, A, T, and C lanes). The sequence read from this is given on the left, with the transcription start sites marked with asterisks.

FIG. 6. (A) Expression and localization of protein IIIa in *E. coli* DH5α containing pMP2241, detected in a Western blot with MAb 38. Lanes: 1 through 3, nonheated samples of cell envelopes grown in LB medium with 1% glucose (lane 1); LB medium (lane 2), or LB medium with 0.1 mM IPTG (induced culture); 4, soluble fraction proteins of induced culture of lane 3, equivalent to same number of cells; 5, isolated outer membrane fraction of induced culture; 6, isolated cytoplasmic membrane of induced culture; 7, cell envelopes of induced culture without lysozyme treatment; 8, as in lane 7, lysozyme treated; 9, as in lane 8, heated for 10 min at 100°C before electrophoresis. (B) Cell envelopes of induced bacteria grown with various concentrations of added calcium chloride. Samples were heated before electrophoresis. Lanes 1 through 5, samples without EDTA; 6 through 10, samples with 10 mM EDTA. Calcium chloride was added to the growth medium as follows: 0 mM (lanes 1 and 6), 1.0 mM (lanes 2 and 7), 2.0 mM (lanes 3 and 8), 5.0 mM (lanes 4 and 9), and 10.0 mM (lanes 5 and 10).

nosarum (6). These results suggest that protein IIIa is partially capable of forming oligomers in the cell envelope of *E. coli*. To compare the amount of protein IIIa in cell envelopes with that remaining in the soluble fractions of *E. coli*, we compared cell envelopes and precipitated soluble fraction proteins in amounts equivalent to equal numbers of cells (Fig. 6, lanes 3 and 4, respectively). These results show that the major part of the produced protein IIIa is localized in the cell envelope fraction. The soluble cell fractions also do not contain the putative oligomers present in cell envelope fractions, suggesting that membrane localization of the protein is a prerequisite for oligomer formation. Lanes 5 and 6 in Fig. 6 show isolated outer membranes and cytoplasmic membranes, respectively. Most of the protein reacting with MAb38 is localized in the outer membrane fraction. There-

fore it can be concluded that, even at a high level of expression caused by the inducible *lac* promoter, most of the protein IIIa produced in *E. coli* is localized in its proper compartment, i.e., the outer membrane.

Protein IIIa expressed in *E. coli* is not peptidoglycan bound. One of the properties of group III antigens in *R. leguminosarum* demonstrated earlier was the strong, presumably covalent binding of the major part of these proteins to the peptidoglycan, resulting in a large increase in detectability on polyacrylamide gels and Western blots after lysozyme treatment of the cell envelopes (6). To investigate whether this is also the case in *E. coli*, we compared cell envelopes of *E. coli* cells expressing protein IIIa before and after lysozyme treatment by Western blotting (Fig. 6A, lanes 7 and 8, respectively). Only a slight increase in the intensity of the major high-molecular-weight band reacting with MAb38 after lysozyme treatment could be observed, and no multiple bands reminiscent of protein containing murein residues could be detected after lysozyme treatment. From these results we concluded that a very small part, if any, of protein IIIa is covalently bound to the peptidoglycan of the *E. coli* cells expressing it.

Protein IIIa forms calcium-stabilized oligomers in *E. coli* cell envelopes. The detection of protein IIIa in *E. coli* cell envelopes (Fig. 6A, lanes 1 through 8) was initially performed on Western blots of nonheated samples, allowing the detection of oligomers that may be better recognized by MAb38 than the denatured form of the protein (6). Indeed, in *E. coli* part of the protein appeared to be present as oligomers at low temperature, although a substantial amount of the protein was also present in the monomeric form, the 36-kDa protein, which was also recognized in Western blots by MAb37 (data not shown) and does not react with the


```
IHF consensus        Py A A N N N N T T G A T A/T

ropA 1      185    T  A A a t a t T T G A A  A   197

ropA 2      200    T  C A a a t a T T T A T  A   188
```

FIG. 5. Alignment of the two putative IHF recognition sites of the RopA promoter region with the consensus sequence for IHF recognition sites (2, 15). Nucleotides identical to the consensus (in uppercase letters) are underlined.
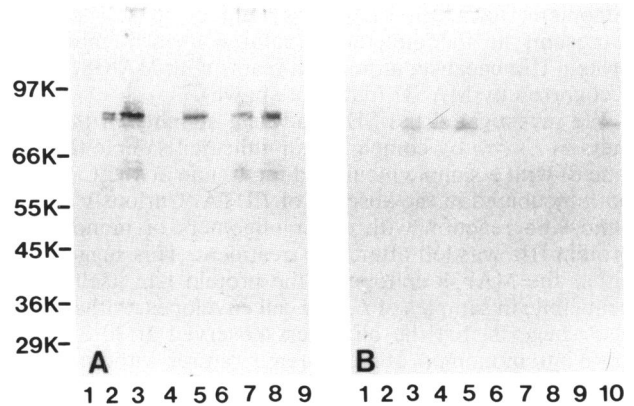
oligomeric form (6). The extra band of 38 kDa probably represents an incompletely denatured form of monomeric protein IIIa because, although it reacts with MAb38, it is not recognized by MAb37 (data not shown).

We investigated the SDS and heat stability of the oligomers in *E. coli* by comparing an unheated sample (Fig. 6A, lane 8) with a sample incubated for 10 min at 100°C (lane 9), both incubated in the absence of EDTA. Curiously enough, almost no reaction with either oligomers or monomers of protein IIIa was left after heat treatment. This suggests that either the MAb38 epitope or the protein IIIa itself is very heat labile in samples of *E. coli* cell envelopes. Although this also suggests that the oligomers observed at 20°C disintegrate into monomers at higher temperatures without the help of EDTA, obviously this could not be proven without either form being detectable in heat-treated samples. In *R. leguminosarum* the oligomers are heat stabilized by calcium, requiring the presence of EDTA in the electrophoresis sample for complete denaturation by heat treatment (6). Because our standard growth medium for *R. leguminosarum*, TY medium, contains relatively high amounts of added calcium (7 mM), we investigated the effects of the addition of calcium to the *E. coli* growth medium LB on the heat stability of the oligomers of protein IIIa. In Fig. 6B, lanes 1 to 5 show the reaction with MAb38 of heat-treated samples (10 min, 100°C, without EDTA) of cell envelopes of *E. coli* grown in LB medium supplemented with calcium chloride added at concentrations of 0, 1, 2, 5, and 10 mM, respectively. At a concentration of 2 mM (lane 3) and higher, calcium apparently stabilizes the oligomeric form, so that it remains detectable after heat treatment of the samples. In Fig. 6B, lanes 6 to 10 show the reaction of the same heat-treated samples containing 10 mM EDTA. The presence of EDTA allows denaturation and thus disappearance of reaction of the protein IIIa oligomers in cell envelopes of cells grown with 0, 1, 2, and 5 mM calcium chloride (Fig. 6B, lanes 6 through 9, respectively) but not completely of cells grown with 10 mM calcium chloride (lane 10). This suggests that in the last case the EDTA concentration is insufficient to remove enough calcium to allow complete denaturation of the oligomers. From these results we conclude that, as in *R. leguminosarum*, protein IIIa oligomers in *E. coli* cell envelopes are stabilized by calcium.

DISCUSSION

In this study we cloned a *R. leguminosarum* biovar viciae gene, designated *ropA*, that encodes the smallest of the two major proteins of the previously defined outer membrane antigen group III, i.e., protein IIIa. The observations in the present study support our hypothesis (6) that antigen group III basically consists of two major proteins that are to a large extent covalently bound to peptidoglycan. Lysozyme treatment of cell envelopes followed by SDS-gel electrophoresis shows the two major proteins and their derivatives that contain increasing numbers of peptidoglycan residues. Thus, *ropA* encodes the 36-kDa major protein IIIa. A gene encoding the other, 40-kDa major protein was not isolated is this study. We expect such a gene to be homologous to the gene isolated here, as its product cross-reacts with all three monoclonal antibodies that recognize protein IIIa (4). Similar genes are expected to be present in other *R. leguminosarum* strains, because these also show cross-reactivity, at least with MAb37 (3).

The expression of group III antigens is severely diminished in cell envelopes of bacteroids as compared with free-living bacteria (4). If regulation of expression takes place at the level of transcription, specific DNA sequences playing a role in this regulation might be found in the 300-bp fragment upstream of the transcriptional start site that was identified in this study. Although neither apparent promoter sequences nor possible positive or negative regulatory sequences could be identified in this DNA fragment, the presence of two closely spaced putative IHF binding sites suggest that such sequences may be present. IHF has been shown to play a role in a number of regulatory processes, including the regulation of transcription initiation by RNA polymerase for a number of genes (8), such as a nitrogen fixation gene in *Klebsiella pneumoniae* and possibly also in members of the family *Rhizobiaceae* (11). Considering the proposed function of IHF in facilitating contact between RNA polymerase and an upstream bound regulator protein by bending of the DNA in between those two (11, 30), one might hypothesize that a yet-unknown regulatory protein binds to the 180-bp region upstream of the putative IHF binding sites in the *ropA* promoter. Since this promoter was not active in *E. coli*, such a regulator may well be unique for *R. leguminosarum*.

Future study of the isolated protein IIIa or of mutants affected in its production may establish a function as an outer membrane pore, a common function for many major outer membrane proteins of the same observed molecular weight (17). However, we could not find significant homology between protein IIIa and other outer membrane proteins by a computer search of available sequences or more specifically by comparison with sequences of porins from various sources.

Expression of protein IIIa in *E. coli* allowed us to determine to what extent the properties of the protein depend on the rhizobial outer membrane environment. We have shown that the majority of the protein IIIa protein produced in *E. coli* is, indeed, exported to the outer membrane. In this context it may be relevant that, although lacking any other homology with other outer membrane proteins, protein IIIa contains a C-terminal phenylalanine residue, a trait common to most outer membrane proteins and probably essential for efficient translocation to the outer membrane (25). We have also shown that in the *E. coli* protein IIIa is to a great extent able to form oligomers that, like *Rhizobium* oligomers, are stabilized by calcium against denaturation by heat and SDS. Apparently both outer membrane localization and calcium-stabilized oligomer formation do not depend on some unique property of *R. leguminosarum*.

A distinctive property of the rhizobial outer membrane is the large proportion of outer membrane proteins covalently linked to the peptidoglycan. Most of the protein IIIa (approximately 80%) in *R. leguminosarum* is only visible on polyacrylamide gels or Western blots after lysozyme treatment of the cell envelopes, which results in the occurrence of protein IIIa derivatives with higher molecular weights. In *E. coli*, there was only a small increase in amount of detectable protein IIIa after lysozyme treatment of cell envelopes. Furthermore, lysozyme treatment did not result in the occurrence of extra higher-molecular-weight derivatives. From these results we conclude that the high extent of peptidoglycan linkage of protein IIIa is unique for *R. leguminosarum* and cannot occur in *E. coli*. This suggests either that *E. coli* peptidoglycan is so different from that of *R. leguminosarum* that linkage of the protein is not possible or that *E. coli* lacks a protein-peptidoglycan linking mechanism that is present in *R. leguminosarum*. Future studies may allow us to determine the function of this peptidoglycan

linkage. Most important, however, is that the study of the regulation of expression of protein IIIa and comparison with other regulatory processes during symbiosis may show the mechanism and function of the surface changes that occur during bacteroid development.

ACKNOWLEDGMENTS

REFERENCES

1. **Beringer, J. E.** 1974. R factor transfer in *Rhizobium leguminosarum*. J. Gen. Microbiol. **84:**188–198.
1a. **Canter Cremers, H. C. C.** Unpublished data.
2. **Craig, N. L., and H. A. Nash.** 1984. *E. coli* integration host factor binds to specific sites in DNA. Cell **39:**707–716.
3. **de Maagd, R.** 1989. Studies on the cell surface of the root-nodulating bacterium *Rhizobium leguminosarum*. PhD thesis, Leiden University, Leiden.
3a. **de Maagd, R.** Unpublished data.
4. **de Maagd, R. A., R. de Rijk, I. H. M. Mulders, and B. J. J. Lugtenberg.** 1989. Immunological characterization of *Rhizobium leguminosarum* outer membrane antigens using polyclonal and monoclonal antibodies. J. Bacteriol. **171:**1136–1142.
5. **de Maagd, R. A., C. Van Rossum, and B. J. J. Lugtenberg.** 1988. Recognition of individual strains of fast-growing rhizobia by using profiles of membrane proteins and lipopolysaccharides. J. Bacteriol. **170:**3782–3785.
6. **de Maagd, R. A., F. B. Wientjes, and B. J. J. Lugtenberg.** 1989. Evidence for divalent cation (Ca²⁺)-stabilized oligomeric proteins and covalently bound protein-peptidoglycan complexes in the outer membrane of *Rhizobium leguminosarum*. J. Bacteriol. **171:**3989–3995.
7. **Ditta, G., S. Stanfield, D. Corbin, and D. R. Helinski.** 1980. Broad host-range DNA cloning system for gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti*. Proc. Natl. Acad. Sci. USA **77:**7347–7351.
8. **Friedman, D. I.** 1988. Integration host factor: a protein for all reasons. Cell **55:**545–554.
9. **Goosen-de Roo, L., and R. A. de Maagd.** 1991. Antigenic changes in lipopolysaccharide (LPS I) of *Rhizobium leguminosarum* bv. viciae strain 248 during the differentiation of bacteria into bacteroids in nodules of *Vicia sativa* subsp. *nigra*. J. Bacteriol. **173:**3177–3183.
10. **Grosveld, F. G., H. M. Dahl, E. De Boer, and R. A. Flavell.** 1981. Isolation of β-globin-related genes from a human cosmid library. Gene **13:**227–237.
11. **Hoover, T. R., E. Santero, S. Porter, and S. Kustu.** 1990. The integration host factor stimulates interaction of RNA polymerase with NIFA, the trancriptional activator for nitrogen fixation operons. Cell **63:**11–22.
12. **Hooykaas, P. J. J., F. G. M. Snijdewint, and R. A. Schilperoort.** 1982. Identification of he Sym plasmid of *Rhizobium leguminosarum* strain 1001 and its transfer to and expression in other rhizobia and *Agrobacterium tumefaciens*. Plasmid **8:**73–82.
13. **Josey, D. P., J. L. Beynon, A. W. B. Johnston, and J. E. Beringer.** 1979. Strain identification in *Rhizobium* using intrinsic antibiotic resistance. J. Appl. Bacteriol. **46:**343–350.

14. **Kieny, M. P., R. Lathe, and J. P. Lecocq.** 1983. New versatile cloning and sequencing vectors based on bacteriophage M13. Gene **26:**91–99.
15. **Leong, J. M., S. Nunes-Duby, C. F. Lesser, P. Youderian, M. M. Susskind, and A. Landy.** 1985. The phi 80 and P22 attachment sites. Primary structure and interaction with *Escherichia coli* integration host factor. J. Biol. Chem. **260:**4468–4477.
16. **Lugtenberg, B., J. Meyers, R. Peters, P. Van der Hoek, and L. Van Alphen.** 1975. Electrophoretic resolution of the major outer membrane protein of *Escherichia coli* K12 into four bands. FEBS Lett. **58:**254–258.
17. **Lugtenberg, B., and L. Van Alphen.** 1983. Molecular architecture and functioning of the outer membrane of *Escherichia coli* and other Gram-negative bacteria. Biochim. Biophys. Acta **737:**51–115.
18. **Marsh, J. L., M. Erfle, and E. J. Wykes.** 1984. The pIC plasmid and phage vectors with versatile cloning sites for recombinant selection by insertional inactivation. Gene **32:**481–485.
19. **Meade, H. M., S. R. Long, G. B. Ruvkun, S. E. Brown, and F. M. Ausubel.** 1982. Physical and genetic characterization of symbiotic and auxotrophic mutants of *Rhizobium meliloti* induced by transposon Tn5 mutagenesis. J. Bacteriol. **149:**114–122.
20. **Osborn, M. J., J. E. Gander, E. Parisi, and J. Carson.** 1972. Mechanism of assembly of the outer membrane of *Salmonella typhimurium*. Isolation and characterization of cytoplasmic and outer membrane. J. Biol. Chem. **247:**3962–392.
21. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
22. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
23. **Sindhu, S. S., N. J. Brewin, and E. L. Kannenberg.** 1990. Immunochemical analysis of lipopolysaccharides from free-living and endosymbiotic forms of *Rhizobium leguminosarum*. J. Bacteriol. **172:**1804–1813.
24. **Spaink, H. P., R. J. H. Okker, C. A. Wijffelman, E. Pees, and B. J. J. Lugtenberg.** 1987. Promoters in the nodulation region of the *Rhizobium leguminosarum* Sym plasmid pRL1JI. Plant Mol. Biol. **9:**27–39.
25. **Struyvé, M., M. Moons, and J. Tommassen.** 1991. Carboxy-terminal phenylalanine is essential for the correct assembly of a bacterial outer membrane protein. J. Mol. Biol. **218:**141–148.
26. **Van Brussel, A. A. N., K. Planqué, and A. Quispel.** 1977. The wall of *Rhizobium leguminosarum* in bacteroid and free-living forms. J. Gen. Microbiol. **101:**52–56.
27. **Vandenbosch, K. A., N. J. Brewin, and E. L. Kannenberg.** 1989. Developmental regulation of a *Rhizobium* cell surface antigen during growth of pea root nodules. J. Bacteriol. **171:**4537–4542.
28. **Van Slogteren, G. M. S., J. H. C. Hoge, P. J. J. Hooykaas, and R. A. Schilperoort.** 1983. Clonal analysis of heterogenous crown gall tumour tissues induced by wild-type and shooter mutant strains of *Agrobacterium tumefaciens*: expression of T-DNA genes. Plant Mol. Biol. **2:**321–333.
29. **Von Heijne, G.** 1986. A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. **14:**4683–4690.
30. **Wedel, A., D. S. Weiss, D. Popham, P. Dröge, and S. Kustu.** 1990. A bacterial enhancer functions to tether a transcriptional activator near a promoter. Science **248:**486–489.
31. **Wijffelman, C. A., E. Pees, A. A. N. van Brussel, R. J. H. Okker, and B. J. J. Lugtenberg.** 1985. Genetic and functional analysis of the nodulation region of the *Rhizobium leguminosarum* Sym plasmid pRL1JI. Arch. Mikrobiol. **143:**225–232.
32. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC vectors. Gene **33:**103–119.