Nucleotide Sequence and Molecular Characterization of *pnlA*, the Structural Gene for Damage-Inducible Pectin Lyase of *Erwinia carotovora* subsp. *carotovora* 71†

A. CHATTERJEE,1 J. L. McEVOY,1 J. P. CHAMBOST,2 F. BLASCO,2 AND A. K. CHATTERJEE1†

Department of Plant Pathology, 108 Waters Hall, University of Missouri, Columbia, Missouri 65211,1 and Laboratoire de Chimie Bactérienne, Centre National de la Recherche Scientifique, 13277 Marseille Cedex 9, France2

Received 3 October 1990/Accepted 19 December 1990

In a previous study, *pnlA* (the DNA damage-inducible structural gene for pectin lyase) of *Erwinia carotovora* subsp. *carotovora* 71 was localized to a 1.4-kb DNA segment within a 3.4-kb EcoRI fragment (J. L. McEvoy, H. Murata, and A. K. Chatterjee, J. Bacteriol. 172:3284–3289, 1990). We present here DNA sequence data for a 2.2-kb region revealing an open reading frame of 870 bases, corresponding to a protein (Pnl) of an approximate molecular mass of 32,100 Da and an isoelectric point of 9.92. Although initiation of translation is presumed to occur at the ATG codon, direct protein sequencing revealed alanine as the N-terminal amino acid, probably as a consequence of posttranslational removal of the initiating amino acid. The sequence of the first 20 amino acid residues of Pnl, purified from *E. carotovora* subsp. *carotovora* 71, agreed completely with the predicted amino acid sequence of the N-terminal segment. This finding also indicated that Pnl is not subject to processing by a signal peptidase. The transcriptional start site of *pnlA* was determined to reside 80 bp upstream of the translational start site. Deletion analysis revealed that 218 bp of DNA upstream of the transcriptional start site is sufficient for induction of *pnlA* by mitomycin C. Within 600 bp upstream of the translational start site, no sequences resembling a LexA binding site (SOS box) or a cyclic AMP receptor protein binding site were found. However, palindromic sequences were detected at −187 and −86 bp relative to the translational start site, and these could be potential sites for the binding of a regulatory protein(s). Comparison of the deduced amino acid sequence for PnlA with that of a Pnl from *Aspergillus niger* and with those of various pectinases of *Erwinia* species revealed a low degree of homology dispersed throughout the length of the proteins.

In many soft-rot *Erwinia* species, pectin lyase (Pnl) production is stimulated by DNA-damaging agents such as UV light, mitomycin C, or nalidixic acid (11–13, 19, 27–29, 34). This stimulation is unique in that other pectinases are induced by the catabolic products of pectin or pectate (for reviews, see references 4, 5, and 15) and not by DNA-damaging agents. Most previous studies of Pnl regulation have been done with *Erwinia carotovora* subsp. *carotovora* 71 (19–21, 34). The observations relevant to this report are (i) that RecA function is required in the activation of the transcription of *pnlA* (the structural gene for pectin lyase) and (ii) that *pnlA* is expressed very poorly in *RecA*-*Escherichia coli* and is noninducible by DNA-damaging agents. These findings raised the possibility that Pnl induction requires a positive regulator, the expression of which, and not the expression of *pnlA*, could be controlled by the RecA-LexA pathway. As a prelude to a systematic analysis of cis-acting promoter and regulatory sequences of *pnlA* and of gene structure in relation to enzyme properties, we report here the nucleotide sequence of a 2,172-bp DNA fragment containing *pnlA* and the comparison of the deduced gene product with other pectinases.

**MATERIALS AND METHODS**

**Bacterial strains, plasmids, and bacteriophage.** AC5122 is a *pnlA::Tn5* mutant of the wild-type *E. carotovora* subsp. *carotovora* strain 71 (19, 33). *E. coli* HB101, DH5α, and TG1 have been previously described (2, 19, 35). pAKC272 is a pSF6 (25) derivative harboring the *E. carotovora* subsp. *carotovora* 71 *pnlA* gene on a 4.5-kb insert (19). pAKC278 and pAKC279 are SK− and SK+ pBluescript (Stratagene, La Jolla, Calif.) derivatives carrying the *pnlA* gene on a 3.4-kb EcoRI fragment (19). The deletions used in the sequencing of *pnlA* are exonuclease III-exonuclease VII derivatives of pAKC278 and pAKC279. pAKC289 is a *HindIII-Stul* deletion derivative of pAKC279. The bacteriophage R408 has been previously described (24).

**Sequencing of *pnlA***. Unidirectional 5′ to 3′ deletions within *pnlA* and flanking DNA were made essentially as previously described (9, 19). Derivatives of pAKC278 and pAKC279 containing overlapping deletions differing in size by approximately 200 bp (Fig. 1.) were chosen for sequence analysis. Cultures of *E. coli* TG1 harboring these plasmids were infected with the helper phage R408; single-stranded sequencing templates were isolated and used in sequencing reactions with the Sequenase system of U.S. Biochemicals (Cleveland, Ohio). The nucleotide sequence was analyzed by using the Genetics Computer Group, Inc. (Madison, Wis.) software program (6) at the DNA core facility on the University of Missouri—Columbia campus.

**Determination of *pnlA* transcript size.** RNA was isolated from *E. carotovora* subsp. *carotovora* 71 cultures grown in the presence or absence of mitomycin C. Fifty-milliliter cultures in LB medium (3) were grown to an optical density at 600 nm of 0.6 with shaking at 28°C. At this point, mitomycin C (final concentration, 500 ng/ml) was added to
the culture to be induced. After 3 h of additional incubation, RNA was extracted (1) and treated with DNase.

RNA samples (approximately 20 μg each) and a size marker RNA ladder (GIBCO BRL, Gaithersburg, Md.) were electrophoresed in a formaldehyde-agarose gel, transferred to a Biotrans (ICN, Irvine, Calif.) nylon filter, and probed with a 0.39-kb PvuII-KpnI fragment of pnlA (Fig. 1) radiolabeled with [32P]dCTP (18). Radiolabeled lambda DNA was used to specifically detect the RNA ladder. The blots were washed under stringent conditions as previously described for Southern blots (18).

**Determination of the transcriptional start site of pnlA.** RNA samples used to determine pnlA transcript size (see above) were also utilized in a primer extension assay (21a). The 32P-labeled oligonucleotide primer (5' TCCAGTTACATT TGCCGC, corresponding to base positions 649 to 666 in Fig. 2, was annealed to the RNA. The primer-extended cDNA was run in an 8% acrylamide–urea sequencing gel in parallel with a DNA sequencing ladder obtained by using the same primer and pBluescript SK+ harboring pnlA.

**Construction and characterization of a deletion 5' of pnlA.** pAKC279 was digested with HindIII and Stul (one unique site for each enzyme occurs in the plasmid: for HindIII, it is in the polylinker of the vector, and for Stul, it is 218 bp upstream of the transcriptional start site of pnlA). By using standard procedures (18), the DNA ends were filled in by using Klenow fragment and deoxyxynucleoside triphosphates, the blunt ends were ligated, and the ligation mix was used in transforming DH5α. Plasmids were characterized by restriction analysis and for their ability to complement a pnlA::Tn5 mutation in AC5122 (19).

**Purification of Pnl.** Pnl was purified from a culture of *E. carotovora* subsp. *carotovora* 71 harboring pnlA in pSF6 (pAKC277 [19]). Cells were grown in LB medium containing streptomycin (100 μg/ml) and spectinomycin (50 μg/ml). Mitomycin C at a final concentration of 500 ng/ml was added to the culture grown at 28°C when an optical density at 600 nm of 0.6 was reached. After 18 h of incubation, the cells were removed by centrifugation (20 min at 10,400 × g). The supernatant (250 ml) was dialyzed against Tris-HCl (25 mM, pH 7.0) for 24 h and then passed through a DEAE-Sepharose column (2.6 by 35.5 cm). The active fractions, as determined by using the quantitative Pnl assay (19), were then added to a carboxymethyl cellulose column (2.5 by 37 cm). This column was washed with 75 ml of the Tris-HCl buffer containing 0.125 M NaCl. An NaCl gradient (0.125 to 0.4 M, 300 ml) was then applied to the column. The active fractions were pooled, dialyzed against water, and lyophilized. As an assay for purity, samples of the Pnl preparation (90 and 170 μg) were electrophoresed in a Phastg gel system (Pharmacia LKB Biotechnology, Inc., Piscataway, N.J.) with a 10 to 15% acrylamide gradient gel. The protein and known standards were stained with AgCl.

**Nucleotide sequence accession number.** The sequence of the approximately 2.2-kb fragment containing pnlA has been assigned the GenBank accession number M59909.

**RESULTS AND DISCUSSION**

**Sequence of pnlA.** Previously, we cloned a 3.4-kb EcoRI fragment of the *E. carotovora* subsp. *carotovora* 71 genome harboring the pectin lyase structural gene, pnlA, into the pBluescript vectors SK+ and SK-. By characterizing two series of exonuclease III-exonuclease VII undirectional deletions and MtdI 1734 insertions in this fragment, pnlA was localized to an approximately 1.4-kb segment of DNA (19). Plasmids containing overlapping deletions within pnlA and the flanking DNA (Fig. 1) were used to produce single-stranded DNA templates for sequencing. Upon analysis of the nucleotide sequence, an open reading frame of 870 bases was found (Fig. 2). From this sequence, we deduced that Pnl has a molecular mass of 32,100 Da and an isoelectric point of 9.92.

The amino-terminal end of the deduced Pnl protein does not appear to contain a traditional cleavable signal sequence. Typically, a gram-negative bacterial signal peptide contains one to three positively charged amino acid residues within the amino-terminal region, followed by a long hydrophobic sequence of 14 to 20 amino acid residues (7). The deduced sequence for Pnl contains no positively charged amino acid residues within the first 14 residues of the amino terminus, and the longest stretch of hydrophobic residues within the first 40 residues of the protein constitutes only eight amino acids (residues 5 to 12 of Fig. 2). Several potential signal sequence cleavage sites exist in PnlA; for example, Ala-Ala is found at residues 17 and 18, Gly-Gly at residues 22 and 23 and 25 and 26, and Ser-Ala at residues 39 and 40 and 44 and 45. However, the fact that the amino-terminal sequence of Pnl purified from *E. carotovora* subsp. *carotovora* 71 (see below) corresponds exactly to the amino-terminal sequence of the deduced Pnl supports the notion that PnlA contains no cleavable signal peptide. It should be noted that Pnl produc-
FIG. 3. Primer extension of the pnlA transcript. Total RNA was extracted from a mitomycin C-induced culture of E. carotovora subsp. carotovora 71. A 32P-end-labeled oligonucleotide primer, corresponding to bases 649 to 666 in Fig. 2, was annealed to the RNA and extended by using deoxynucleoside triphosphates and reverse transcriptase. The extended product (lane P) was electrophoresed in parallel with a ladder (GATC) by using the identical primer and the antisense strand of pnlA as the template. The portion of the sequence pertinent to the transcription start site is shown. The arrow indicates the in vivo initiation point corresponding to the extended product.

A potential Shine-Dalgarno sequence (AAGGGA) can be found properly positioned 5' of the putative ATG start codon (Fig. 2). Several E. coli sigma 70-like promoter sequences are found 5' of this putative ribosome binding site. One, TAAGCG(17 bp)IACAAT, that is found at −89 to −117 bp relative to the translational start codon has 7 bases conserved of the 12 bases of the E. coli sigma 70 binding consensus sequence. This may be the promoter for pnlA, since the transcriptional start site was determined by a primer extension assay to be a guanine residue located at −80 bp relative to the translational start site (Fig. 3).

Upstream of the translational start codon, no sequence resembling the consensus sequence for an SOS box (the site for binding of LexA, the repressor of DNA damage-inducible genes of E. coli [31]) was found. This suggests that LexA is not directly involved in regulating pnlA expression. In support of this idea, we now have evidence that the regulation of pnlA expression involves an activator, the gene for which is repressible by E. coli LexA (18a).

Pnl production is not susceptible to catabolite repression (i.e., to glucose effect; our unpublished data). Consistent with this finding is the absence of a sequence similar to the consensus E. coli cyclic AMP receptor protein binding site upstream of the start codon. This region does harbor two palindromic sequences centered at −86 and −187 bp upstream of the start codon (Fig. 2). Furthermore, located at −109 to −94 bp is a sequence which has 10 of 12 bases in...
Characterization of a deletion derivative of pAKC279. pAKC289 contains pnlA along with 298 bp of upstream DNA, including several palindromic sequences (Fig. 2). This plasmid restored the PnI" phenotype to the pnlA::Tn5 mutant AC5122. The induced level of PnI in this construct (specific activity, 71 U/mg of protein) was comparable to that produced by AC5122 carrying pAKC279 (specific activity, 51 U/mg of protein). This finding indicates that pAKC289 harbors all the cis-acting elements required for activation of pnlA transcription.

Comparison of the deduced PnIA sequence with pectin and pectate lyases. Prior to this report, the only pnl sequenced was from Aspergillus niger (8). The deduced product of this fungal gene showed a low degree of homology (23.8% identity dispersed throughout the length of the protein) to the deduced product of E. carotovora subsp. carotovora pnlA. A higher percentage of amino acid identity is seen between pnlA and some pectate lyases (Pels) of Erwinia strains. For example, 28.5% identity in a 281-amino-acid overlap is seen with PElE of E. chrysanthemi 3937 (29). This homology dispersed throughout the length of pnlA and the length of the mature PelE protein, does not extend into the putative signal sequence of this PelE. Similar homologies (about 28.0% identities) were detected with PelE from E. chrysanthemi B374 (30) and EC16 (14), except that some homology was detected between PnlA and the 3' half of the EC16 PelE signal sequence. Homologies between pnlA and other Pels range from 25.1% identity with PelA of E. carotovora subsp. atroseptica EC (17, 22.9% identity with PelB of E. chrysanthemi EC16 (14), and less than 20% identity with PelA of EC16 (26), PelB of E. carotovora subsp. carotovora SCR193 (10), and PelB of E. carotovora subsp. atroseptica EC (16). It is intriguing that 24 of the 76 amino acid residues found to be conserved in at least 10 of 12 Pel sequences examined by Hinton et al. (10) were conserved in pnlA. Additionally, of these 24 residues, 9 were conserved in the fungal Pnl. It may be that these conserved amino acid residues are involved in recognition of the common substrate, pectin.

In summary, the data presented here establish that (i) pnlA is monocistronic, (ii) PnlA does not contain a cleavable signal sequence, (iii) regulatory sequences of pnlA span no more than 298 bp upstream of the translational start site, and (iv) there is only marginal similarity between the deduced amino acid sequence of pnlA and those of Pels of Erwinia species and of a Pnl of A. niger. Our ongoing studies should allow further characterization of the regulatory region upstream of pnlA and indicate whether pnlA sequences are conserved among strains of soft-rot Erwinia.

ACKNOWLEDGMENTS

This investigation was supported by the National Science Foundation (grant DMB-8796262), the Food for the 21st Century program of the University of Missouri, and the Centre National de la Recherche Scientifique (C.N.R.S., Marseille, France).

We thank M. Chippaux and F. Barbas for their interest, J. Forrester and M. McIntosh for their assistance with the sequence analysis and transcript assays, and D. Pinkerton for the photography.

REFERENCES


