

Pectate lyase from *Bacillus subtilis*: molecular characterization of the gene, and properties of the cloned enzyme

W. Nasser*, A.C. Awadé**, S. Reverchon, J. Robert-Baudouy

Laboratoire de Génétique Moléculaire des Microorganismes, URA CNRS 1486, INSA Bâtiment 406, 20 avenue Albert Einstein, 69621 Villeurbanne cedex, France

Received 4 October 1993

Pectate lyases (PL) initiate soft-rot diseases in plants by cleaving pectin which is the major component of the plant cell wall. The present paper reports the first cloning and characterization of a pectate lyase (*pel*) gene from the *Bacillus* genus. This gene was isolated from a *Bacillus subtilis* genomic library constructed in pUC18 as vector and *Escherichia coli* as host. By Southern hybridization this gene was shown to be present in a single copy in the *B. subtilis* genome. The nucleotide sequence of a 1.6 kb-pair *Hind*III restriction fragment, which confers pectate lyase activity to *E. coli*, indicated a 1,260 bp open reading frame encoding a 420 amino acid polypeptide which includes a 21 amino acid signal sequence. The 45,605 Da deduced protein displays homologies to PLs from *Erwinia chrysanthemi*. The *B. subtilis* PL cloned in *E. coli* was located in the periplasm. It was purified to homogeneity in a one-step procedure from the *E. coli* periplasmic fluid after overproduction using the pT7 system. Biochemical properties of the purified enzyme were similar to those found for the PL isolated from *B. subtilis* extracellular media.

pel gene; Gene expression; Purification; *Bacillus subtilis*

1. INTRODUCTION

Pectate lyase (PL) (EC 4.2.2.2) cleaves the α -1,4 glycosidic bond of polygalacturonic acid and generates unsaturated oligogalacturonides [1]. This class of pectinases is widely distributed in bacteria and fungi [2–6], some being phytopathogenic and others, such as members of the genera, *Klebsiella* and *Yersinia* [7–9], being non-phytopathogenic. The most well-known pectinolytic bacteria are the phytopathogenic *Erwiniae* that are the causal agents of soft-rot disease of many plant species [10,11]. These bacteria produce multiple isoenzymes of PL that are responsible for plant tissue maceration [12]. Bacteria from the *Bacillus* genus also produce PL and have been shown to cause soft-rot disease under certain conditions [13–17]. This raises the question of whether *Bacillus* is an opportunistic bacteria or an actual pathogen. The PL enzyme from *B. subtilis* has been purified and characterized by Nasser et al. [4]. This enzyme immunologically cross-reacted with PelB and PelC, the two neutral PL isoenzymes from *E. chrysanthemi*. Moreover crystals of this *B. subtilis* enzyme have been obtained and preliminarily studied by X-rays [18]. Recently, crystallization and thus a model of the three-dimensional structure of PelC from *E. chrysanthemi* has also been achieved [19].

Although PL activity was described in *Bacillus* a long time ago, no genetic characterization of a PL gene had been carried out up to now. In the present paper, we report the first isolation and sequencing of a gene encoding PL activity in *B. subtilis*. Comparison of *B. subtilis* and *E. chrysanthemi* PL reveals the existence of conserved motifs the biological significance of which is discussed. Moreover this work explores the N-terminal sequence of the mature enzyme and some biochemical and enzymatic properties of the enzyme over-produced in *E. coli*.

2. MATERIALS AND METHODS

2.1. Bacterial strains, plasmids, media and growth conditions

Bacterial strains and plasmids used in this study are listed in Table I. *B. subtilis* and *E. coli* were usually grown at 37°C in liquid or solidified agar (15 g/l), LB medium or synthetic M63 minimal medium supplemented with glycerol (0.1%) [20]. When required, antibiotics, ampicillin (Ap), chloramphenicol (Cm), kanamycin (Km) were added at 50 $\mu\text{g}\cdot\text{ml}^{-1}$. For *B. subtilis* genomic DNA preparation, the cells were grown as described by Rodriguez and Tait [21].

2.2. PL activity assay

B. subtilis PL assay, K_m , V_m and optimum pH values were determined as previously described by Nasser et al. [4].

2.3. DNA manipulations

DNA-modifying enzymes and restriction endonucleases were purchased from Boehringer-Mannheim.

Chromosomal and plasmid DNA preparations were carried out as described by Ausubel et al. [22]. DNA digestion and fragment isolation, dephosphorylation, ligation, electrophoresis and bacterial transformation were carried out according to Sambrook et al. [23]. DNA

*Corresponding author. Fax: (33) 72 43 87 14.

**Present address: Laboratoire de Recherches de Technologie Laitière, INRA, 65 rue de St. Briec, 35042 Rennes cedex, France.

fragment isolation was performed after digestion and electrophoresis on agarose gels, with GeneClean II kit (Bio 101 Inc.) or by electroelution, using Biotrap apparatus (Schleicher and Shüell).

For Southern blot hybridization, DNA probes (restriction fragments or oligonucleotides) were labelled by random priming or digoxigenin (DIG) 5' end labelling, using Boehringer-Mannheim kits. Southern blot hybridizations were performed on nylon membrane (Hybond N⁺, Amersham) as proposed by the manufacturer.

2.4. Preparation of a gene library and in situ screening of PL activity

Chromosomal DNA was extracted from *B. subtilis* strain SO113 and partially digested with the restriction endonuclease, *Sau*3A. The obtained random fragments were separated on agarose (0.5%) gels and fragments between 3–6 kb-pair (kbp) were recovered by electroelution, using the Biotrap (Schleicher and Shüell) apparatus. These fragments were ligated with the plasmid vector, pUC18 (Appligene), which had been digested to completion with the restriction endonuclease, *Bam*HI, and dephosphorylated. The library was used to transform *E. coli* NM522 cells and the recombinants were selected on agar plates containing Ap. Recombinant clones were screened for PL activity as described by Keen et al. [24].

2.5. Nucleotide sequence and computer analysis

For nucleotide sequence analysis, a nested series of deletion clones was created using various restriction endonucleases. Sequencing was done using the chain termination method on double-stranded DNA templates. Extension of primers (M13 primer or M13 reverse primer) was carried out with T7 DNA polymerase (T7 sequencing kit from Pharmacia).

The resulting data were analyzed using the Mac Molly programme (Soft Gene, Berlin). Amino acid sequence comparisons were achieved with the CLUSTAL program.

2.6. Over-production and purification of the *B. subtilis* PL in *E. coli*

The 1.6 kbp *Hind*III–*Hind*III DNA fragment encoding the *B. subtilis* PL was cloned into pT7-5 and pT7-6 expression vectors leading to plasmids pNP111 and pNP112 (Table I). Expression of these resulting plasmids was performed in the presence of L-[³⁵S]methionine and cysteine in order to selectively label the *pel* gene product [25].

The periplasmic fluid, which contained PL activity, was released by osmotic shock as described by Nossal and Heppel [26]. The final concentration of this extract was adjusted to 20 mM Tris-HCl, pH 7, 1 mM ethylenediaminetetraacetic (EDTA) and 1 mM dithiothreitol

(DTT) (extraction buffer). This preparation (50 ml) was applied to a Protein-Pack SP 8HR (1 × 10 cm) (Waters) column previously equilibrated with the extraction buffer. The column was washed with the same buffer and the proteins were eluted at 1.4 ml/min with a gradient from 0 to 0.6 M NaCl, using a Waters HPLC system. Fractions of 0.7 ml were collected.

2.7. Analytical methods and other techniques

Protein concentration determination was carried out as proposed by Bradford [27] using the Bio-Rad protein assay, with bovine serum albumin as a standard.

SDS-PAGE was performed according to Laemmli [28] on slab gels (12% resolving gel and 4% stacking gel). Protein bands were detected by Coomassie blue staining. Molecular weight markers were obtained from Bethesda Research Laboratory.

N-Terminal amino acid sequence of the purified PL was determined by automated Edman degradation [29], using a gas-phase protein sequencer.

3. RESULTS AND DISCUSSION

3.1. Cloning and characterization of the *B. subtilis* gene encoding PL activity

The strategy used to isolate the *B. subtilis pel* gene was the selection of recombinant plasmids exhibiting PL activity in *E. coli*. Among 5,000 *E. coli* transformants, two clones showing PL activity were detected. Restriction mapping of the plasmids isolated from these two clones revealed that they contained overlapping inserts of 2.8 and 3.3 kbp. In addition, Southern blot experiments showed that both inserts specifically hybridized to the oligonucleotide probe deduced from the N-terminal amino acid sequence of the purified PL (data not shown). Further studies were performed on the clone with the longest insert, pNP1 (Fig. 1).

To determine the size of the *pel* gene, the plasmid pNP1 was digested with different restriction endonucleases and the resulting fragments were inserted into

Table I
Bacterial strains and plasmids

| Strains and plasmids | Genotypes and characteristics ^a | Origin or reference |
|--------------------------------|--|---------------------|
| Strains | | |
| <i>Bacillus subtilis</i> SO113 | <i>trpC2</i> , <i>amy-3</i> | [40] |
| <i>Escherichia coli</i> NM522 | <i>SupE</i> , <i>thi</i> , Δ (<i>lac-pro AB</i>), <i>dhds5</i> , (<i>r_k-m_k</i>)(<i>F'</i> , <i>pro AB</i> , <i>lacI</i> ^q Δ M15] | Stratagene |
| <i>Escherichia coli</i> K38 | HfrC, λ^+ , <i>phoA4</i> , <i>pit-10</i> , <i>tonA22</i> , <i>ompF627</i> , <i>relA1</i> | [41] |
| Plasmids | | |
| pUC18 | Ap ^R | Appligene |
| pBluescript (pBS) | Cm ^R , <i>lacZ'</i> | Stratagene |
| pT7-5 | Ap ^R , T7 ϕ 10 | [25] |
| pT7-6 | Ap ^R , T7 ϕ 10 | [25] |
| pGP1-2 | Km ^R , P _L -T7 gene 1, <i>Plac-c1857</i> | [25] |
| pNP1 | pUC18 with 3.3 kbp <i>Sau</i> 3A– <i>Sau</i> 3A fragment containing the <i>pel</i> gene of <i>B. subtilis</i> SO113 | This work |
| pNP2 | pUC18 with 2.8 kbp <i>Sau</i> 3A– <i>Sau</i> 3A fragment containing the <i>pel</i> gene of <i>B. subtilis</i> SO113 | This work |
| pNP11 | pBS with 1.6 kbp <i>Hind</i> III– <i>Hind</i> III fragment containing the <i>pel</i> gene of <i>B. subtilis</i> SO113 | This work |
| pNP111 | pT7-5 with the <i>Xba</i> I– <i>Cl</i> aI fragment from pWNPL11 containing the <i>pel</i> gene | This work |
| pNP112 | pT7-6 with the <i>Xba</i> I– <i>Cl</i> aI fragment from pWNPL11 containing the <i>pel</i> gene | This work |

^a Genotype symbols are according to Bachmann [46]. *lacZ'* indicates that the 3' end of this gene is truncated. Ap^R, resistance to ampicillin; Cm^R, resistance to chloramphenicol; Km^R, resistance to kanamycin.

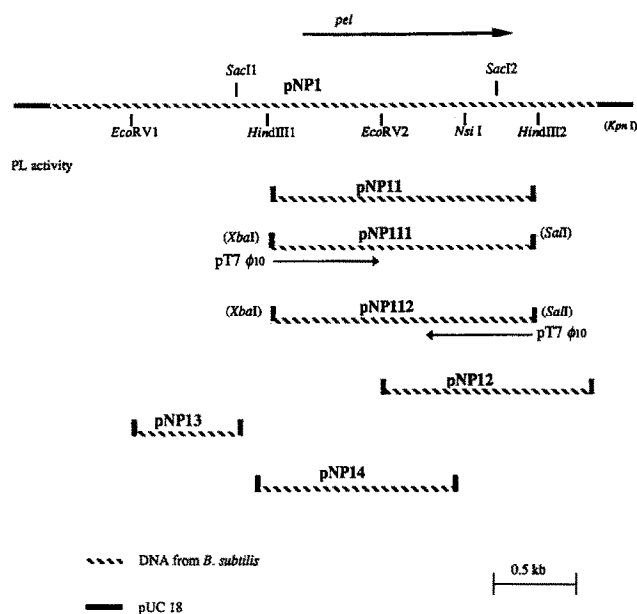


Fig. 1. Physical map of the 3.3 kbp *Sau3A-Sau3A* DNA fragment (pNP1) containing the *pel* gene from *B. subtilis* and subclone derivatives. The PL activity was checked as described by Nasser et al. [4]. pNP11 was constructed by insertion of the *HindIII-HindIII* fragment into the *HindIII* site of pBS; pNP12 by insertion of the *EcoRV-KpnI* fragment into the *EcoRV-SacI* site of pBS; pNP13 by insertion of the *EcoRV-SacI* fragment into the *EcoRV-SacI* site of pBS; pNP14 by insertion of the *HindIII-NsiI* fragment into the *HindIII-PstI* site of pBS. Only *E. coli* harbouring pNP11 displayed PL activity. Restriction sites between brackets are sites from the vector polylinkers. The transcriptional direction of the *pel* gene, demonstrated by phage T7 expression system, is indicated by the arrow.

the appropriate sites of pBS, giving rise to plasmids pNP11, pNP12, pNP13 and pNP14 (Fig. 1). These plasmids were introduced into *E. coli* NM522 and checked for PL activity (Fig. 1). The smallest fragment exhibiting PL activity was the 1.6 kbp *HindIII-HindIII* fragment of pNP11. Cloning of this fragment in the two opposite orientations downstream of the pBS *lac* promoter had no effect on *pel* gene expression. This result indicated that the *pel* gene was probably expressed from its own promoter on the 1.6 kbp DNA fragment. To determine the *pel* gene transcriptional direction, the 1.6 kbp *XbaI-SalI* restriction fragment from pNP11 was cloned into plasmids pT7-5 and pT7-6 digested with the same endonucleases (Fig. 1). These vectors differ by the orientation of the polylinker adjacent to the $\phi 10$ promoter which is specifically recognized by the T7 RNA polymerase. Expression of the genes present on the resulting plasmids, pNP111, pNP112, with specific labelling of the translated proteins, revealed two polypeptides of 42 kDa and 33 kDa produced from pNP111 (Fig. 2). The molecular weight of the first polypeptide is in accordance with the molecular weight of the mature PL protein from *B. subtilis*, previously determined by Nasser et al. [4]. The 33 kDa polypeptide might be a degradation product from the 42 kDa, as indicated by its positive immunological reaction by Western blot

analysis (data not shown). Thus the labelling of the translated products from the T7 promoter-specific mRNAs demonstrated that the *pel* gene is transcribed in the *HindIII-HindIII2* direction (Fig. 1).

Genomic *B. subtilis* DNA digested with *EcoRV*, *EcoRI* or *HindIII* was probed with the 1.6 kbp *HindIII* fragment from plasmid pNP11, containing the *pel* gene. Two different *EcoRV* fragments of 1.3 and 2 kbp hybridized with *pel* DNA. Two *EcoRI* fragments of 1.4 and 9 kbp were also detected and only one *HindIII* band of 1.6 kbp. This hybridization pattern correlates to the restriction map established for the *pel* gene. This suggests that the *pel* gene is present in a single copy on the *B. subtilis* genome.

3.2. Purification and properties of the *B. subtilis* PL expressed in *E. coli*

B. subtilis SO113 PL, expressed from pNP111, was purified with the procedure described in section 2.6. SDS-PAGE analysis of proteins contained in a major peak (not shown) revealed a single homogenous protein band with an apparent molecular weight of 42 kDa (Fig. 3). The purification procedure used allowed the purification to homogeneity in one chromatographic step of 3.2 mg of PL with a yield of 55%. The fact that the enzyme was over-expressed and located in the *E. coli* periplasm facilitated this purification.

In addition to the fact that the enzyme purified from *E. coli* has the same molecular weight as the one isolated

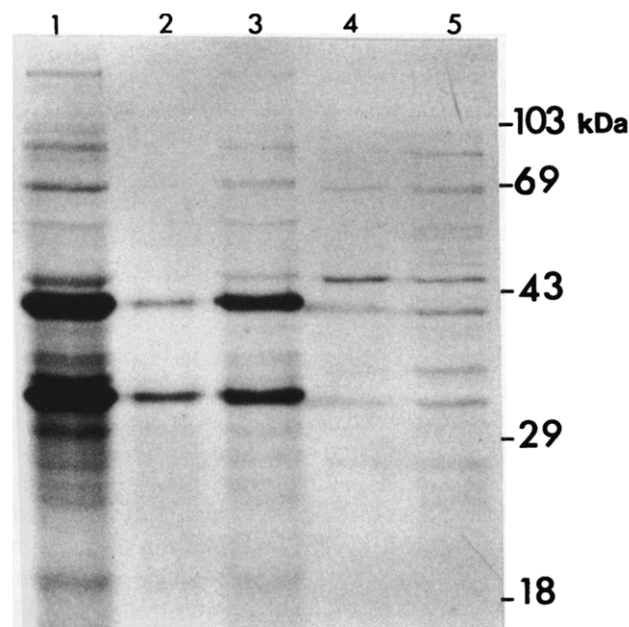


Fig. 2. Analysis of the *pel* gene product using the phage T7 promoter-polymerase expression system. After SDS-PAGE of the proteins synthesized in vivo by K38 (pGP1-2) carrying pNP111 (lane 1 and 3) (500,000 and 200,000 cpm, respectively), pNP112 (lane 2) (500,000 cpm), pT7-5 (lane 4) (500,000 cpm) and pT7-6 (lane 5) (500,000 cpm), the labelled polypeptides were detected by autoradiography. The size markers are indicated at the right.

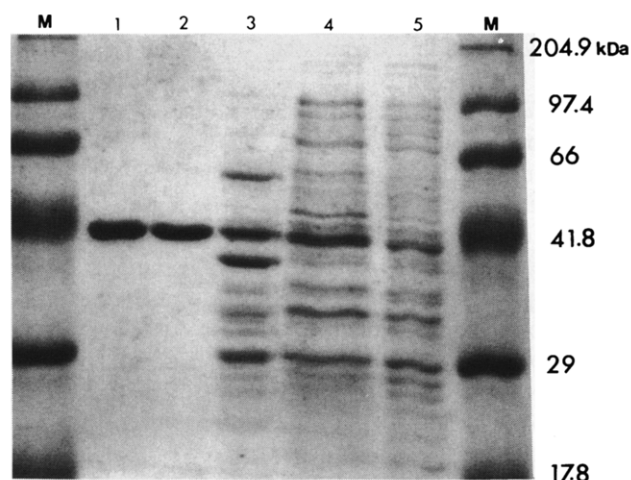


Fig. 3. Purification of the over-expressed *B. subtilis* PL. SDS-PAGE of 5 μ g of purified PL from *B. subtilis* SO113 (lane 1); 5 μ g of purified PL from *E. coli* K38 harbouring pNP111 (lane 2); 25 μ g proteins from the periplasm of *E. coli* K38 harbouring pNP111 (lane 3); 40 μ g of the crude extract from *E. coli* K38 harbouring pNP111 before the temperature shift to 42°C (lane 5) and after the temperature shift (lane 4). M lane denotes molecular weight markers.

from *B. subtilis*, they also have the same K_m and V_m values, exhibit maximum activity at the same pH, and display the same stability (data not shown and [4]). The influence of various enzyme inhibitors was tested on the purified PL. The enzyme was inhibited by hydroxy-mercuribenzoic acid and *N*-ethyl-maleimide which are SH enzyme inhibitors. The inhibition by the second compound was removed by the addition of DTT to the assay medium. These results suggest that a sulphydryl group may form a part of the catalytic site of the *B. subtilis* PL. The enzyme was also inhibited to a lesser degree by diethyl pyrocarbonate (DEPC) (inhibitor of tyrosine and histidine enzyme), whereas phenyl methyl sulfonyl fluoride (PMSF) (inhibitor of serine enzyme) had almost no effect on its activity. This suggests that the *B. subtilis* PL active site also contains a histidine or a tyrosine.

3.3. Nucleotide sequence of the *pel* gene

The complete nucleotide sequence of the 1.6 kbp *Hind*III–*Hind*III insert from pNP11 was determined on both strands, as mentioned in section 2.5. This sequence contains a single open reading frame (ORF) of 1,420 bp (Fig. 4), which starts with an ATG codon at nucleotide (nt) 205 and stops with a TAA termination codon at nt 1,464. The ORF encodes a polypeptide of 420 amino acids with a calculated molecular weight of 45,605 Da and a calculated pI of 7.9. The calculated molecular weight is slightly higher than that of the purified *B. subtilis* PL (42 kDa). This difference can be explained by the presence of a typical 21 amino acid signal peptide at the NH₂ extremity of the protein. The periplasmic localization of the cloned PL in *E. coli* indicates that this 21 amino acid signal peptide is functional in *E. coli* and

is probably processed by the Sec system. The N-terminal amino acid sequence of the mature protein (Ala-Asp-Leu-Gly-His-Gln-Thr-Leu-Gly-Ser-Asn-Asp-Gly-Trp-Gly-Ala-Tyr-Ser-Thr-Gly-Thr-X-Gly-Gly-Ser-Lys-Ala), determined by Edman degradation, corresponds to that predicted from the nucleotide sequence of the *pel* gene and confirmed the signal sequence cleavage site. This signal peptide is sufficient to ensure extracellular localization of the PL in the Gram-positive bacterium *B. subtilis* [4]. In contrast, in Gram-negative bacteria, PLs are secreted into the extracellular medium using a two-step secretion pathway [30–33]. The first step, which is Sec-dependent, allows the passage of the protein through the inner bacterial membrane. This passage is accompanied by the cleavage of the peptide signal. The second step requires the Out proteins and allows the mature periplasmic enzyme to pass through the outer membrane. The fact that the *E. chrysanthemi* PLs cloned in *E. coli* remain localized in the periplasm [2,34] indicates that there is no functional Out machinery in *E. coli*.

Analysis of the nucleotide sequence upstream of the ATG initiation codon of the *B. subtilis pel* gene indicated the presence of a purine-rich sequence (AGAAAATGGGGGTA) that probably contains the ribosome binding site (RBS) [35,36]; it is worth noting that this sequence does not correspond to the well-defined Shine and Dalgarno sequence. Upstream from the putative RBS, there are putative –35 (TGAATG) and –10 (TATATT) promoter signals between nt 121 and nt 126 and between nt 144 and nt 150, respectively. It appears that the putative –10 region has good homology with the –10 promoter sequence recognized by the σ^{43} transcription factor of *B. subtilis* and by the σ^{70} transcription factor of *E. coli*. In contrast, the –35 region is not well conserved, suggesting that transcription of the *pel* gene may require another sigma factor or a specific positive regulatory protein.

Computer-aided searches revealed two inverted sequences (Fig. 4), one located upstream from the putative promoter sequence, between nt 22 and nt 59, and another downstream from the translation stop codon, between nt 1,475 and nt 1,508. The first inverted sequence may be involved in the transcription termination [37] of another gene located upstream from the *pel* gene, or it may also be the target site of a regulatory protein. Indeed a variety of regulators, such as products of *degQ*, *degR*, *degT*, *sacV*, *senN*, *senS*, and *tenA*, have been shown to stimulate the production of many extracellular enzymes in the genus *Bacillus* [38] by binding to sites generally located in the regulatory regions of the degradative enzyme genes. Analysis of the *B. subtilis pel* gene regulation would then be of great interest in order to define the biological significance of this inverted sequence. The second dyad symmetry sequence located 7 bp downstream from the TAA stop codon is able to form a secondary structure that may be involved in transcription termination [37] of the *B. subtilis pel* gene.

| | |
|---|------|
| AAGCTTGGGCATAAAGCAAGGAAAAAACCAGGCAATGTCGGCCTTTTGGTTTTTTTGGCGTCTTTGCG | 72 |
| <div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;"> <div style="width: 100px; border-top: 1px solid black; margin: 0 auto;"></div> <div style="text-align: center;">←</div> </div> <div style="text-align: center;"> <div style="width: 100px; border-top: 1px solid black; margin: 0 auto;"></div> <div style="text-align: center;">→</div> </div> </div> | |
| -35 | |
| GTGGGATTTTGCAGAATGCCGCAATAGGATAGCGGAACATTTTCGGTTCGAATGTCCTCAATTTGCTATT | 144 |
| -10 | |
| Met Lys Lys ATATTTTGTGATAAATTGGAATAAAATCTCACAAAATAGAAAATGGGGGTACATAGTCG ATG AAA AAA | 213 |
| Val Met Leu Ala Thr Ala Leu Phe Leu Gly Leu Thr Pro Ala Gly Ala Asn Ala | |
| GTG ATG TTA GCT ACG GCT TTG TTT TTA GGA TTG ACT CCA GCT GGC GCG AAC GCA | 267 |
| <u>Ala Asp Leu Gly His Gln Thr Leu Gly Ser Asn Asp Gly Trp Gly Ala Tyr Ser</u> | |
| GCT GAT TTA GGC CAC CAG ACG TTG GGA TCC AAT GAT GGC TGG GGC GCG TAC TCG | 321 |
| <u>Thr Gly Thr Thr Gly Gly Ser Lys Ala Ser Ser Ser Asn Val Tyr Thr Val Ser</u> | |
| ACC GGC ACG ACA GGC GGA TCA AAA GCA TCC TCC TCA AAT GTG TAT ACC TCG ACG | 375 |
| Asn Arg Asn Gln Leu Val Ser Ala Leu Gly Lys Glu Thr Asn Thr Thr Pro Lys | |
| AAC AGA AAC CAG CTT GTC TCG GCA TTA GGG AAG GAA ACG AAC ACA ACG CCA AAA | 429 |
| Ile Ile Tyr Ile Lys Gly Thr Ile Asp Met Asn Val Asp Asp Asn Leu Lys Pro | |
| ATC ATT TAT ATC AAG CGA ACG ATT GAC ATG AAC GTG GAT GAC AAT CTG AAG CCG | 483 |
| Leu Gly Leu Asn Asp Tyr Lys Asp Pro Glu Tyr Asp Leu Asp Lys Tyr Leu Lys | |
| CTT GGC CTA AAT GAC TAT AAA GAT CCG GAG TAT GAT TTG GAC AAA TAT TTG AAA | 537 |
| Ala Tyr Asp Pro Ser Thr Trp Gly Lys Lys Glu Pro Ser Gly Thr Gln Glu Glu | |
| GCC TAT GAT CCT AGC ACA TGG GGC AAA AAA GAG CCG TCG GGA ACA CAA GAA GAA | 591 |
| Ala Arg Ala Arg Ser Gln Lys Asn Gln Lys Ala Arg Val Met Val Asp Ile Pro | |
| GCG AGA GCA CGC TCT CAG AAA AAC CAA AAA GCA CGG GTC ATG GTG GAT ATC CCT | 645 |
| Ala Asn Thr Thr Ile Val Gly Ser Gly Thr Asn Ala Lys Val Val Gly Gly Asn | |
| GCA AAC ACG ACG ATC GTC GGT TCA GGG ACT AAC GCT AAA GTC GTG GGA GGA AAC | 699 |
| Phe Gln Ile Lys Ser Asp Asn Val Ile Ile Arg Asn Ile Glu Phe Gln Asp Ala | |
| TTC CAA ATC AAG AGT GAT AAC GTC ATT ATT CGC AAC ATT GAA TTC CAG GAT GCC | 753 |
| Tyr Asp Tyr Phe Pro Gln Trp Asp Pro Thr Asp Gly Ser Ser Gly Asn Trp Asn | |
| TAT GAC TAT TTT CCG CAA TGG GAT CCG ACT GAC GGA AGC TCA GGG AAC TGG AAC | 807 |
| Ser Gln Tyr Asp Asn Ile Thr Ile Asn Gly Gly Thr His Ile Trp Ile Asp His | |
| TCA CAA TAC GAC AAC ATC ACG ATA AAC GGC GGC ACA CAC ATC TGG ATT GAT CAC | 861 |
| Cys Thr Phe Asn Asp Gly Ser Arg Pro Asp Ser Thr Ser Pro Lys Tyr Tyr Gly | |
| TGT ACA TTT AAT GAC GGT TCG CGT CCG GAC AGC ACA TCA CCG AAA TAT TAT GGA | 915 |
| Arg Lys Tyr Gln His His Asp Gly Gln Thr Asp Ala Ser Asn Gly Ala Asn Tyr | |
| AGA AAA TAT CAG CAC CAT GAC GGC CAA ACG GAT GCT TCC AAC GGT GCT AAC TAT | 969 |
| Ile Thr Met Ser Tyr Asn Tyr Tyr His Asp His Asp Lys Ser Ser Ile Phe Gly | |
| ATC ACG ATG TCC TAC AAC TAT TAT CAC GAT CAT GAT AAA AGC TCC ATT TTC GGA | 1023 |
| Ser Ser Asp Ser Lys Thr Ser Asp Asp Gly Lys Leu Lys Ile Thr Leu His His | |
| TCA AGT GAC AGC AAA ACC TCC GAT GAC GGC AAA TTA AAA ATT ACG CTG CAT CAT | 1077 |
| Asn Arg Tyr Lys Asn Ile Val Gln Arg Ala Pro Arg Val Arg Phe Gly Gln Val | |
| AAC CGC TAT AAA AAT ATT GTC CAG CGC GCG CCG AGA GTC CGC TTC GGG CAA GTG | 1131 |
| His Val Tyr Asn Asn Tyr Tyr Glu Gly Ser Thr Ser Ser Ser Tyr Pro Phe | |
| CAC GTA TAC AAC AAC TAT TAT GAA GGA AGC ACA AGC TCT TCA AGT TAT CCT TTT | 1185 |
| Ser Tyr Ala Trp Gly Ile Gly Lys Ser Ser Lys Ile Tyr Ala Gln Asn Asn Val | |
| AGC TAT GCA TGG GGA ATC GGA AAG TCA TCT AAA ATC TAT GCC CAA AAC AAT GTC | 1239 |
| Ile Asp Val Pro Gly Leu Ser Ala Ala Lys Thr Ile Ser Val Phe Ser Gly Gly | |
| ATT GAC GTA CCG GGA CTG TCA GCT GCT AAA ACG ATC AGC GTA TTC AGC GGG GGA | 1293 |
| Thr Ala Leu Tyr Asp Ser Gly Thr Leu Leu Asn Gly Thr Gln Ile Asn Ala Ser | |
| ACG GCT TTA TAT GAC TCC GGC ACG TTG CTG AAC GGC ACA CAG ATC AAC GCA TCG | 1347 |
| Ala Ala Asn Gly Leu Ser Ser Ser Val Gly Trp Thr Pro Ser Leu His Gly Ser | |
| GCT GCA AAC GGG CTG AGC TCT TCT GTC GGC TGG ACG CCG TCT CTG CAT GGA TCG | 1401 |
| Ile Asp Ala Ser Ala Asn Val Lys Ser Asn Val Ile Asn Gln Ala Gly Ala Gly | |
| ATT GAT GCT TCT GCT AAT GTG AAA TCA AAT GTT ATA AAT CAA GCG GGT GCG GGT | 1455 |
| Lys Leu Asn SSS | |
| AAA TTA AAT TAA GAAAGTAAAAACACAAAGGGTGCTAACCTTTGTGTTTTTAATTAATTAATGTT | 1524 |
| <div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;"> <div style="width: 100px; border-top: 1px solid black; margin: 0 auto;"></div> <div style="text-align: center;">←</div> </div> <div style="text-align: center;"> <div style="width: 100px; border-top: 1px solid black; margin: 0 auto;"></div> <div style="text-align: center;">→</div> </div> </div> | |
| TATTAACCTTAGTTAAGGAGTAGAATGGAAGGGGATCGGAAACAGTATATAGGAGGAGACCTATTATG | 1597 |
| CTTCAGAAAAAGACGCAGAAAAACAGTCAGCAGTAAAGCTT | 1638 |

Fig. 4. Nucleotide sequence of the 1.6 kbp *Hind*III fragment including the *pel* gene. The putative promoter region (−35, −10) is in bold print. The peptidic sequence corresponding to the *pel* ORF is shown under the nucleotide sequence. The putative ribosome binding site and the N-terminal amino acid sequence of the mature protein, determined by Edman degradation, are underlined. Inverted repeat sequences are indicated by arrows.

This sequence will appear in Genebank under the accession number: X74880.

| | | |
|-----------------|---|-----|
| PEL B. subtilis | MKKVMLA-----TALFLGLTPAGANAADLGHQTLGS-----NDGWGA | 37 |
| PELB E. chrys. | M-----KSLITPIAAGLLA-----FSQYSLAA-DTGG--- | |
| PELC E. chrys. | M-----KSLITPIAAGLLA-----LSQPLLAATDTGG--- | |
| PELA E. chrys. | MN-----KVSGRSFTRTSTCLLATLIAGVMTSGVSAALVNSKALESAPAAGWAS | |
| PELE E. chrys. | MNNSRMSSVSTQKTTGRS-ALGTSKALAAIATMMVSVVPLVCKPPKRRQRPAGQP | |
| PELD E. chrys. | MNNTRVFSFRST-----KSLLAALIIATSMMTWSVNRATLQTTKATEASTGWAT | |
| | * | |
| PEL B. subtilis | YSTGTTGGSKASSSNVYTVSNRNQLVSALGKETNTTPKIIYIKGTIDMNVDDNLKPLGLN | 97 |
| PELB E. chrys. | -----YTKTDGDDVSGAVKKTASSMQDIVNIEAAKVDAN-----GK | |
| PELC E. chrys. | -----YAATAGGNVTGAVSKTATSMQDIVNIIAARLDAN-----GK | |
| PELA E. chrys. | QNGSTTGGAATSDNIYVVTNISEFTSALS-A-GAVAKIIQITGTVDIS-----GGT | |
| PELE E. chrys. | E-RRHHRGAKASSSKIYAVKSISEFKAALN-GTDSLAKIMPGHGAIDIS-----GGK | |
| PELD E. chrys. | Q-GGTTGGAKAASAKIYAVKNISEFKAALN-GTDDPKIIQVTGAIDIS-----GGK | |
| | * | |
| PEL B. subtilis | DYKDPEYDLDKYLKAYDPSTWGKKEPSGTQEEARARSQKNQKARVMVDIPANTTIVGS-G | 156 |
| PELB E. chrys. | KVKGAYPL---VITYTGN--EDSLINAAANICGQWSKDARGVEIKDFTKGLTIIGANG | |
| PELC E. chrys. | KVKGAYPL---VITYTGN--EDSLINAAANICGQWSKDPRGVEIKEFTKGLTIIGANG | |
| PELA E. chrys. | PYKD--FA-----DQKARSQINIPANTTIVIGI-G | |
| PELE E. chrys. | AYTD--FD-----DQKARSQISIPSNTTIIIGI-G | |
| PELD E. chrys. | AYTS--FD-----DQKARSQISVPSNTTIIIGI-G | |
| | | |
| PEL B. subtilis | TNAKVVGGNFQIKS----DNVIRNIEFQDAYDYFPQWDPDTGSSGNWNSQYDNITINGG | 212 |
| PELB E. chrys. | SSANF--GIWIVNSSD----IVVRNMRIG-----YLPGGAQDG-----DMFRIDNS | |
| PELC E. chrys. | SSANF--GIWIKKSSD---VVVQNMRIK-----YLPGGAQDG-----DMIRVDDS | |
| PELA E. chrys. | TDAKFINGSLIIDGTGNTNVIIRNVIQTPIDVEPHYEKGDGWNAEWDGM-----NITNGA | |
| PELE E. chrys. | NKGKFTNGSLVVKG---VSNVILRNLYIETPVDVAPHYEEGDGWNAEWDVA---VIDST | |
| PELD E. chrys. | SNGKFTNGSLVIK---VSNVILRNLYIETPVDVAPHYEEGDGWNAEWDAA---VIDNS | |
| | | |
| PEL B. subtilis | THIWIDHCTFNDGSRPDSTSPKYYGRKYQHHDGQTDASNGANYITMSYNYIYHDDKSSIF | 272 |
| PELB E. chrys. | PNVWLDHNEFPAAN-HECDGTDGDTTF---ESAIDIKKGATYVTISYNYIHGVKKVGLS | |
| PELC E. chrys. | PNVWLDHNEFPAAN-HECDGTPDNDTTF---ESAVIDIKGASNTVTSYNYIHGVKKVGLD | |
| PELA E. chrys. | HHVWVDHVTISDGSFTDDMYTTKDGETYVQHDGALDIKRGSDYVTISNSLFDQHDKTMIL | |
| PELE E. chrys. | DHVWVDHVTISDGSFTDDKYTTKNGEKYVQHDGSLDIKRGSDYVTISNSRPFELHDKTILI | |
| PELD E. chrys. | TRVWVDHVTISDGSFTDDKYTTKNGEKYVQHDGALDIKKGSDYVTISSSRFELHDKTILI | |
| | | |
| | ----- | |
| | region I | |
| PEL B. subtilis | GSSDSKTSDD-GKLKITLHHNRKYNIVQRAPRVRFQGVHVNYYEGSTSSSSYPFSYAW | 331 |
| PELB E. chrys. | GFSSSDTAERN---ITYHHNIYSDVNARLPLQRGGNVHAYNNLYTGITSS-----GL | |
| PELC E. chrys. | GSSSDTG-RN---ITYHHNIYSDVNARLPLQRGGNVHAYNNLYTNITGS-----GL | |
| PELA E. chrys. | GHSDDTSAQDKGLHVTFLFNNVFNVRTERAPRVYRGSISFNFNVFGDVKDPVRYLYSF | |
| PELE E. chrys. | GHSDDNNGSQDAGKLRTVFNHNLFDVGVGERTPRVRFSGVHAYNNVYVGDVNHKAYRYQSF | |
| PELD E. chrys. | GHSDDNNGSQDAGKLRTVFNHNVFDRVTERTPRVRFSGVHAYNNVYVGLDVKNVYPYLYSF | |
| | * | |
| | ----- | |
| | region II | |
| PEL B. subtilis | GIGKSSKIYAQNNDVIDVPLSAA-----KTIS-VFSGGTA-----LYDSGTLNGTOI | 378 |
| PELB E. chrys. | NVRQNGKALIENNWE-NAVSPVTSRYDGSNFGTQVWLGKGNITKPADFATYNIWTPTDK | |
| PELC E. chrys. | NVRQNGKALIENNWE-KAINPVTSRYDGNFGTQVWLGKGNITKPADFTSYITWTADTK | |
| PELA E. chrys. | GIGTSGSVLSEGSFTIANLSA----SKACKVVK-KFNG-SI-----FSDNGSVLNGSA | |
| PELE E. chrys. | GIGTSGTILSESNFTLSDNKKISGRDKECSVVK-AFNG-KI-----FSDKGSIIINGASY | |
| PELD E. chrys. | GLGTSGTILSESNFTLSNLSIDGKNPECSIVK-QFNS-KV-----FSDNGSVLNGSST | |
| | | |
| PEL B. subtilis | N--ASAANGLSSSVGWTPSLHGSIDASANVKSNNVINOAGAGKLN----- | 420 |
| PELB E. chrys. | EYRNADTWTSTGTYPYVSYSPVS-AQCVDKDLANYAGVGK-NLATLASSACK | |
| PELC E. chrys. | PYVNADSWTSTGTFTYVAYNSPVS-AQCVDKDLPGYAGVGK-NLATLTSTACK | |
| PELA E. chrys. | ADLSGC--GFSAYTSAIPYVYVQPMTELAQSITDHAGSGK--L----- | |
| PELE E. chrys. | -NLNGCGFGFSAYSAPYKYSQITITSLANSISSNAGYGK--L----- | |
| PELD E. chrys. | TKLDTCAV--TAYKPTLPYKYSQITITSLANSISSNAGYGK--L----- | |
| | | |

Fig. 5. Comparison of the *B. subtilis* PL with the five PL isoenzymes from *E. chrysanthemi* (data from PL sequences are from this study for *B. subtilis* PL, from Favey et al. [42] for 3937 PelA, from Van Gijsegem [43] for B374 PelD, from Reverchon et al. [44] for 3937 PelE and from Tamaki et al. [45] for EC16 PelB and PelC). Identical residues are indicated by asterisks; conservative substitutions are indicated by dots. Amino acids mentioned in the text (cysteine 220, aspartic acid 244 and 248, asparagine 210, tyrosine 255, and histidine 310) are in bold print. The two well-conserved regions in *E. chrysanthemi* PL isoenzymes spatially located around the putative Ca^{2+} binding site (Yoder et al. [19]) are underlined; amino acids 246–259 and amino acids 297–313 correspond, respectively, to region I and region II.

3.4. Protein homology and structure

The *B. subtilis* PL has only one cysteine in position 220, indicating that no disulphide bond is present in this enzyme. Moreover, specific inhibition of the *B. subtilis* PL by SH enzyme inhibitors demonstrated that this single cysteine residue may be involved in the catalytic

site. The hydropathy profile, according to Kyte and Doolittle [39], of the deduced amino acid PL sequence showed that besides the region corresponding to the N-terminal signal peptide, which is hydrophobic, the mature enzyme is globally hydrophilic, as expected for a soluble protein.

Comparison of the *B. subtilis* PL amino acid sequence with those of the PLs from *E. chrysanthemi* revealed a globally weak homology. The homology decreases in the following order: PelA the acidic PL (34% identity), PelD and PelE, the two basic isoenzymes (31 and 30% identity, respectively) and PelB and PelC (26% identity) (Fig. 5). Although *B. subtilis* PL appears to be more similar to PelA, PelD, PelE than to PelB and PelC, this protein, purified either from *B. subtilis* (Nasser et al. [4]) or from *E. coli* (data not shown) is immunologically related to the *E. chrysanthemi* PelB and PelC and not to the three other isoenzymes. This result may be explained by a good conservation of sequences corresponding to antibody epitopes in these three proteins. This hypothesis is supported by the fact that the *B. subtilis* PL and *E. chrysanthemi* PelC isoenzymes present a very similar three-dimensional structure (Jenkins, personal communication). After crystallographic studies of the *E. chrysanthemi* PelC isoenzyme, Yoder et al. [19] demonstrated that this enzyme folds into a unique motif of parallel β -strands coiled into a large helix. Within the core, the amino acids form linear stacks and include a new asparagine ladder. It is interesting to note that asparagine residues are also highly conserved in the *B. subtilis* PL (Fig. 5). Moreover, since calcium is essential for PL activity, Yoder et al. [19] suggested the existence of a putative Ca^{2+} binding site involving the residues Asp-131, Glu-166 and Asp-170 in the *E. chrysanthemi* PelC isoenzyme. These acidic amino acids are conserved among all extracellular PLs, including the *B. subtilis* enzyme, where they correspond to Asp-210, Asp-244 and Asp-248, respectively. Further crystallographic analysis on the *B. subtilis* PL will be performed in the presence of Ca^{2+} to confirm this hypothesis. In addition, the two well-conserved regions that are spatially located around the putative Ca^{2+} binding site (Fig. 5), proposed by Yoder et al. [19], contains a conserved tyrosine (region I) and a conserved histidine (region II). These two residues may be involved in the catalytic site of these enzymes, as suggested by the DEPC inhibition. Site-directed mutagenesis will be performed in these regions to evaluate the importance of these two amino acids in the catalytic activity of PL.

Acknowledgements: We are grateful to V. James for correcting the manuscript. We thank N. Hugouvieux-Cotte-Pattat and G. Condemine for valuable discussions and for critical reading of the manuscript. We are grateful to Y. Rahbe for photographic assistance. This work was supported by grants from the Etablissement publique Conseil Général du Rhône, from the Etablissement publique Régional Rhône-Alpes, from UNESCO (Biotechnology programme), and from the CNRS (URA 1486).

REFERENCES

- [1] Rombouts, F.M. and Pilnik, W. (1980) in: *Economic Microbiology: Microbial Enzymes and Bioconversions*, vol. 5 (Rose, A.H., ed.) pp. 228–282, Academic Press, New York.
- [2] Gardner, J.M. and Kado, C.I. (1976) *J. Bacteriol.* 127, 451–460.
- [3] Kamimiyama, S., Itoh, Y. and Izaki, K. (1977) *Agric. Biol. Chem.* 41, 975–981.
- [4] Nasser, W., Chalet, F. and Robert-Baudouy, J. (1990) *Biochimie* 72, 689–695.
- [5] Ayers, W.A., Papavizas, G.C. and Diem, A.F. (1966) *Phytopathology* 56, 1006–1011.
- [6] Sherwood, R.T. (1966) *Phytopathology* 56, 279–286.
- [7] Bagley, S.T. and Starr, M.P. (1979) *Curr. Microbiol.* 2, 381–386.
- [8] Chatterjee, A.K., Buchanan, G.E., Behren, M.K. and Starr, M.P. (1979) *Can. J. Microbiol.* 25, 94–102.
- [9] Walker, M.J. and Pemberton, J.M. (1987) *Arch. Microbiol.* 146, 390–395.
- [10] Chatterjee, A.K. and Vidaver, A.K. (1986) *Adv. Plant Pathol.* 4, 1–213.
- [11] Collmer, A. and Keen, N.T. (1986) *Annu. Rev. Phytopathol.* 24, 383–409.
- [12] Kotoujansky, A. (1987) *Annu. Rev. Phytopathol.* 25, 405–430.
- [13] Davé, B.A. and Vaughn, R.H. (1971) *J. Bacteriol.* 108, 166–174.
- [14] Chesson, A. and Codner, R.C. (1978) *J. Appl. Bacteriol.* 44, 347–364.
- [15] Karbassi, A. and Luh, B.S. (1979) *J. Food Sci.* 44, 1156–1161.
- [16] Ward, O.P. and Fogarty, W. (1974) *Appl. Microbiol.* 27, 346–350.
- [17] Ciampi-Panno, L. (1981) *Proc. Fifth Int. Conf. Plant Path. Bact. Ca.* 374–378.
- [18] Jenkins, J., Nasser, W., Scott, M., Pickersgill, R., Vignon, J.-C. and Robert-Baudouy, J. (1992) *J. Mol. Biol.* 228, 1255–1258.
- [19] Yoder, M.D., Keen, N.T. and Jurnak, F. (1993) *Science* 260, 1503–1507.
- [20] Miller, J.H. (1972) *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- [21] Rodriguez, R.L. and Tait, R.C. (1983) in: *Recombinant DNA Techniques: An Introduction*, Addison-Wesley, Reading.
- [22] Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Smith, J.A., Seidman, J.G. and Struhl, K. (1987) *Current Protocols in Molecular Biology*, Green-Wiley Intersciences, New York.
- [23] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- [24] Keen, N.T., Dahlbeck, D., Staskawicz, B. and Belser, W. (1984) *J. Bacteriol.* 159, 825–831.
- [25] Tabor, S. and Richardson, C. (1985) *Proc. Natl. Acad. Sci. USA* 82, 1074–1078.
- [26] Nossal, N.G. and Heppel, L.A. (1966) *J. Biol. Chem.* 241, 3055–3062.
- [27] Bradford, M.M. (1976) *Anal. Biochem.* 72, 248–254.
- [28] Laemmli, U.K. (1970) *Nature* 227, 680–685.
- [29] Edman, P. (1950) *Acta Chem. Scan.* 4, 283–293.
- [30] He, S.Y., Lindeberg, M. and Collmer, A. (1993) in: *Biotechnology in Plant Disease Control* (I. Chet, ed.) pp. 39–64, Wiley-Liss, New York.
- [31] He, S.Y. et al. (1991) *Proc. Natl. Acad. Sci. USA* 88, 1079.
- [32] Condemine, G. et al. (1992) *Mol. Microbiol.* 6, 3199–3211.
- [33] Lindberg, M. and Collmer, A. (1992) *J. Bacteriol.* 174, 7385.
- [34] Hinton, J.C.D., Sidebotham, J.M., Gill, D.R. and Salmond, G.P.C. (1989) *Mol. Microbiol.* 3, 1785–1795.
- [35] Stormo, G.D., Schneider, T.D. and Gold, L.M. (1982) *Nucleic Acids Res.* 10, 2971–2995.
- [36] Moran, C.P., Lang, N., LeGrice, S.F.J., Lee, G., Stephens, M., Sonenshein, P., Pero, J. and Losick, R. (1982) *Mol. Gen. Genet.* 186, 339–346.
- [37] Rosenberg, M. and Court, D. (1979) *Annu. Rev. Genet.* 13, 319–353.
- [38] Klier, A., Msadek, T. and Rapoport, G. (1992) *Annu. Rev. Microbiol.* 46, 429–459.
- [39] Kyte, J. and Doolittle, R.F. (1982) *J. Mol. Biol.* 157, 105–132.

- [40] Ortlepp, S.A., Ollington, J.F. and McConnell, D.J. (1983) *Gene* 23, 267–276.
- [41] Russel, M. and Model, P. (1985) *J. Bacteriol.* 159, 1034–1039.
- [42] Favey, S., Bourson, C., Bertheau, Y., Kotoujansky, A. and Boccara, M. (1992) *J. Gen. Microbiol.* 138, 499–508.
- [43] Van Gijsegem, F. (1989) *Mol. Microbiol.* 3, 1415–1424.
- [44] Reverchon, S., Huang, Y., Bourson, C. and Robert-Baudouy, J. (1989) *Gene* 85, 125–134.
- [45] Tamaki, S.J., Gold, S., Robeson, M., Manulis, S. and Kenn, N.T. (1988) *J. Bacteriol.* 170, 3468–3478.
- [46] Bachmann, B.S. (1990) 8th edn., *Microbiol. Rev.* 54, 130–197.