Characterization of the sporulation-related γ-D-glutamyl-(L)*meso*-diaminopimelic-acid-hydrolysing peptidase I of *Bacillus sphaericus* NCTC 9602 as a member of the metallo(zinc) carboxypeptidase A family

Modular design of the protein

Marie-Laure HOURDOU,* Micheline GUINAND,† Marie-Jeanne VACHERON,† Georges MICHEL,† Luc DENOROY,‡ Colette DUEZ,* Serge ENGLEBERT,* Bernard JORIS,* Georges WEBER§ and Jean-Marie GHUYSEN*||

*Centre d'Ingénierie des Protéines, Université de Liège, Institut de Chimie, B6, B-4000 Sart Tilman (Liège 1), Belgium,
†Laboratoire de Biochimie Microbienne, Université Claude Bernard Lyon I, 43 Bd du 11 Novembre 1918, F-69622 Villeurbanne Cedex, France,
‡Service Central d'Analyses du Centre National de la Recherche Scientifique, BP 22, F-69390 Vernaison, France,
and §Institut de Physique Nucléaire Expérimentale (IPNE), Université de Liège, B15, B-4000 Sart Tilman (Liège 1), Belgium

The sporulation-related γ-D-glutamyl-(L)*meso*-diaminopimelic-acid-hydrolysing peptidase I of *Bacillus sphaericus* NCTC 9602 has been analysed by proton-induced X-ray emission. It contains 1 equivalent Zn^{2+} per mol of protein. As derived from gene cloning and sequencing, the *B. sphaericus* Zn peptidase I is a two-module protein. A 100-amino-acid-residue N-terminal domain consisting of two tandem segments of similar sequences, is fused to a 296-amino-acid-residue C-terminal catalytic domain. The catalytic domain belongs to the Zn carboxypeptidase A family, the closest match being observed with the *Streptomyces griseus* carboxypeptidase [Narahashi (1990) J. Biochem. **107**, 879–886] and with the family prototype, bovine carboxypeptidase A. The catalytic domain of the *B. sphaericus* peptidase I possesses, distributed along the amino-acid sequence, peptide segments, a triad His^{162}–Glu^{165}–His^{307} and a dyad Tyr^{347}–Glu^{366} that are equivalent to secondary structures, the zinc-binding triad His^{69}–Glu^{72}–His^{196} and the catalytic dyad Tyr^{248}–Glu^{270} of bovine carboxypeptidase A respectively. The N-terminal repeats of the *B. sphaericus* peptidase I have similarity with the C-terminal repeats of the *Enterococcus hirae* muramidase 2, the *Streptococcus* (now *Enterococcus*) *faecalis* autolysin and the *Bacillus* φPZA and φ29 lysozymes, to which a role in the recognition of a particular moiety of the bacterial cell envelope has been tentatively assigned. Detergents enhance considerably the specific activity of the *B. sphaericus* peptidase I.

INTRODUCTION

Bacillus sphaericus NCTC 9602 produces two sporulation-related, γ-D-glutamyl-(L)diamino-acid-hydrolysing peptidases, known as endopeptidases I and II. They differ with respect to cellular localization (Guinand et al., 1979; Vacheron et al., 1979), molecular mass and catalytic mechanism (Garnier et al., 1985; Bourgogne et al., 1992, Hourdou et al., 1992) and specificity profile (Arminjon et al., 1977; Valentin et al., 1983).

Endopeptidase I, a 45 kDa protein, is produced at stage IV of sporulation (Guinand et al., 1974; Tipper et al., 1977). It has been purified in trace amounts from the sporulation medium (Garnier et al., 1985) and in larger amounts from the integuments of the forespores and spores in the presence of Brij 58 (Baji-Kourda et al., 1989). Endopeptidase I, referred to as peptidase I throughout this paper, is a carboxypeptidase/peptidyl dipeptide hydrolase. It hydrolyses the γ-D-glutamyl-(L)*meso*-diaminopimelic acid (*ms*A_2pm) bond of L-Ala-γ-D-Glu-(L)*ms*A_2pm and L-Ala-γ-D-Glu-(L)*ms*A_2pm(L)-D-Ala peptides. The L-alanine amino group may be free or substituted by an *N*-acetylmuramoyl or *N*-acetylglucosaminyl-β-1,4-*N*-acetylmuramoyl moiety. Conversely, the presence of a *ms*A_2pm residue with free ε-NH_2 and ε-COOH groups is a strict requirement for activity. Peptides containing L-lysine instead of *ms*A_2pm, peptides terminating with the sequence *ms*A_2pm(L)-D-Ala-D-Ala and cross-linked peptide dimers (in which the amino group on the D-centre of *ms*A_2pm of one peptide is linked to the carbonyl group on the L-centre of *ms*A_2pm of another peptide via an intervening D-Ala residue) have no substrate activity. Peptidase I is inactivated by EDTA and reactivated by zinc, cobalt and manganese ions, suggesting that it may be a metallo(zinc)peptidase (Garnier et al., 1985).

In order to gain some understanding of the molecular organization of the *B. sphaericus* peptidase I, the protein has been analysed by proton-induced X-ray emission (p.i.x.e.), the encoding gene has been cloned and sequenced and the derived amino-acid sequence has been submitted to hydrophobic-cluster analysis. The results of these investigations are presented and discussed below.

MATERIALS AND METHODS

Enzyme, substrate and enzyme assay

The *B. sphaericus* peptidase I (purified from the spore integuments) (Baji-Kourda et al., 1989) was at a 0.1% (w/v) concentration in 20 mM Tris/HCl, pH 8.0/10 mM $MgCl_2$ (referred to as buffer) containing 0.2% (w/v) Brij 58. The enzyme (diluted to the extent that the samples contained less than 0.0002% Brij 58 before each experiment) and 0.4 mM *N*-acetylmuramoyl-L-Ala-γ-D-Glu-(L)*ms*A_2pm(L)-D-[^{14}C]Ala (1.1×10^6 c.p.m./μmol) were incubated at 37 °C in 10 μl (final volume) of buffer. Under these conditions, i.e. in the virtual absence of

Abbreviations used: p.i.x.e., proton-induced X-ray emission; *ms*A_2pm, *meso*-diaminopimelic acid; HPC1, 1-hexadecylpyridinium chloride; CHAPS, 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulphonate; ORF, open reading frame.

|| To whom correspondence should be addressed.

detergent, the specific enzyme activity was 14 μmol of *ms*-A_2pm-D-[^{14}C]Ala released/min (i.e. 14 units) per mg of protein (Baji-Kourda et al., 1989).

Detergents

Brij 58 (polyoxyethylene-20 cetyl ether), Sarkosyl (sodium dodecanoylsarcosinate), HPC1 (1-hexadecylpyridinium chloride), Cetavlon (hexadecyltrimethylammonium bromide), Triton X-100 (octylphenoxypolyethoxyethanol), *n*-octylglucoside (*n*-octyl β-D-glucopyranoside), SDS and CHAPS {3-[(3-cholamidopropyl)dimethyl ammonio]-1-propanesulphonate} were tested as modifiers of the hydrolytic activity of the *B. sphaericus* peptidase I. They were used at a final concentration of 0.4 % (w/v) in buffer.

P.l.x.e.

The *B. sphaericus* peptidase I (in buffer containing 0.2 % Brij 58) was dialysed extensively against 2.5 mM Tris/HCl, pH 8.0 and lyophilized. The same volume of buffer containing 0.2 % Brij 58 was treated similarly and served as the control. The pellets from the protein and the control were suspended in water and samples (20 μl) (i.e. 11.25 nmol of peptidase I) were deposited on 4 μm-thick propylene films together with 20 μl of a liposome solution (in order to ensure a good homogeneity of the samples) and 7.7 μg of yttrium nitrate (used as internal standard). After evaporation of the solvent, the films were stretched on commercial 24 mm × 36 mm slide frames and irradiated under vacuum by a 2.5 MeV proton beam (40 nA intensity) using the facilities of the Liège University Cyclotron Research Centre. The beam, of 10 mm diam, covered the entire sample, thus avoiding possible problems of non-uniformity of the target. The measurements were made in duplicate. The protein-bound cations were estimated by subtracting the amounts found in the protein samples from those found in the controls. The values thus obtained depended critically on the amounts of protein used for the experiments. The protein content was estimated on the basis of the A_{230}/A_{260} value (Kalb and Bernlohr, 1977) and by using the method of Bradford (1976) with ovalbumin as standard. Both methods gave identical results.

Papain degradation and amino-acid sequencing

Before proteolysis, the *B. sphaericus* peptidase I was treated with 100 mM EDTA and dialysed against 2 mM Tris/HCl, pH 8.0. Proteolysis was carried out according to Cleveland et al. (1977) in the presence of 1 % (w/v) SDS, at a papain (Sigma)/peptidase I ratio of 1:10 (w/w), in 125 mM Tris/HCl, pH 6.8, for 20 min at 37 °C. The reaction was stopped by heating the solution at 100 °C for 5 min. Samples of untreated and papain-degraded peptidase I (25 μg) were submitted to reverse-phase h.p.l.c. on a 4.6 mm × 220 mm Vydac C4 column. Elution was carried out with a linearly increasing gradient of acetonitrile made in 0.1 % trifluoroacetic acid at a flow rate of 1 ml per min. Peptidase I and the h.p.l.c.-isolated peptides were analysed by SDS/PAGE on 12.5 % (w/v) gels (Laemmli and Favre, 1973) and stained with Coomassie Brilliant Blue. They were submitted to amino-acid microsequencing by using an Applied Biosystems model 470A.

DNA recombinant techniques

The *B. sphaericus* genomic DNA was prepared as described by Hopwood et al. (1985) from cells in the early sporulation phase. Gene cloning was performed using pBR322 (Bolivar et al., 1977) and pSP73 (Promega, Madison, WI, U.S.A.) as vectors and *Escherichia coli* strains HB101 (Boyer et al., 1969) and DH5αMCR (Jessee and Bloom, 1988), grown at 37 °C in Luria–Bertani medium, as hosts. The libraries were constructed using dephosphorylated vectors. *E. coli* DH5αMCR was transformed by electroporation (Dower et al., 1988) using a Bio-Rad Gene Pulser apparatus. Transformants were selected on agar plates containing ampicillin (50 μg/ml) or tetracycline (25 μg/ml). Genomic libraries were screened by hybridization (at t_m −5 °C, for 18 h) with radioactive synthetic oligonucleotides (Eurogentec, Liège, Belgium) using a modified Southern-blot procedure (Wallace et al., 1981; Woods, 1984). Other DNA-manipulation experiments were carried out as described (Maniatis et al., 1982).

DNA segments cloned into plasmid pSL1190 (Brosius, 1989) and double-stranded DNA segments were sequenced by the dideoxynucleotide chain termination method (Sanger et al., 1977) using the T7 sequencing kit (Pharmacia LKB Biotechnology AB, Uppsala, Sweden). Denaturation of double-stranded DNA was performed as described (Zhang et al., 1988) and the sequencing reactions were initiated with synthetic oligonucleotides.

Sequence identity searches

Searches through the nucleic acid (EMBL version 30) and protein (PIR version 32) sequence databases were performed by using the procedure of Pearson and Lipman (1988) (FASTA and TFASTA softwares, GCG package). Alignments of pairs of proteins were made by using the Goad and Kanehisa procedure (1982) and a uniform gap penalty of +8. The significance of the comparison between pairs of aligned sequences was assessed by using the SEQDP program (Kanehisa, 1982). This program gives the score of the best alignment of two sequences. The significance of the score is expressed by the S.D. unit of the scores of 20 random sequences of the same composition. A S.D. unit value of 5 or higher indicates a statistically significant similarity. The theoretical pI value of the protein was computed by using the GCG package (Devereux et al., 1984).

Hydrophobic-cluster analysis

Hydrophobic-cluster analysis (Gaboriaud et al., 1987; Henrissat et al., 1990) is a powerful method for comparing proteins that are weakly related in their primary structures. It rests upon a duplicated representation of the amino-acid sequences on an α-helical two-dimensional pattern (in which the hydrophobic residues tend to form clusters) and compares the distribution of the clusters along the sequences. The shapes of the clusters are usually associated with definite secondary structures. Clusters of similar shapes, sizes and relative positions express similarity in the polypeptide foldings.

Zinc peptidases

Bovine carboxypeptidase A of known primary and three-dimensional structure (Kim and Lipscomb, 1990; Le Huërou et al., 1991) served as a standard of reference. Carboxypeptidase A is a one domain protein the dominant feature of which is an eight-stranded β-sheet that constitutes the core of the molecule. The *Streptomyces griseus* carboxypeptidase of known primary structure (Narahashi, 1990) was also included in the comparison.

Nucleotide sequence accession number

The EMBL accession number for the nucleotide sequence encoding the *B. sphaericus* peptidase I is X69507.

RESULTS

Amino-acid sequence data and nucleotide probes

Figure 1 shows the h.p.l.c. fractionation profile of the papain digest of *B. sphaericus* peptidase I and, as an insert, the SDS/PAGE profile of the untreated enzyme (E) and the isolated peptides 6 (28 kDa), 7 (43 kDa) and 8 (32 kDa). Peptides 2 and 4 (< 14 kDa) ran out of the gels. Figure 2 gives the results of the amino-acid sequencing performed on the *B. sphaericus* peptidase I and the isolated peptides 4, 2, 8 and 6. Peptidase I and peptide 4 gave identical results. The amino-acid numbering was attributed on the basis of the primary structure of the protein as derived from gene cloning and sequencing (see below).

These structural data were used to synthesize the two degenerated nucleotides shown in Figure 2. Nucleotide 1 was complementary to the nucleotide sequence encoding the NYDYSMM amino-acid sequence of peptide 8. Nucleotide 2 was complementary to the nucleotide sequence encoding the PMVNPD(G) amino-acid sequence of peptide 6.

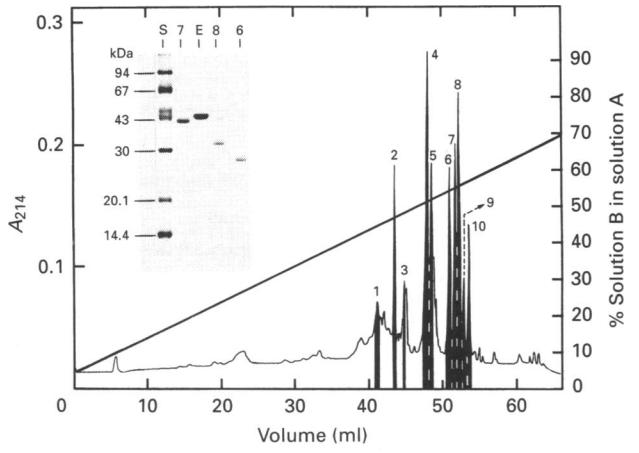

Figure 1 Isolation by reverse-phase h.p.l.c. chromatography on a Vydac C4 column of the peptides produced by papain hydrolysis of the *B. sphaericus* peptidase I

Continuous line: linearly increasing gradient of acetonitrile made by mixing at a constant flow rate (1 ml/min) solution B (acetonitrile/H_2O/trifluoroacetic acid; 90:10:0.1, by vol.) in solution A (0.1% trifluoroacetic acid in water). SDS/PAGE of the untreated peptidase I (E) and the isolated peptides 6, 7 and 8: S: proteins of standard molecular mass (shown in insert).

Peptides		N-terminal amino-acid sequences	
4	1	MDILIRPGDSLWYFSDLFKIP	21
2	51	(T)(S)YTITQGDSL(W)QI	64
8	106	(G)QQNYDYSMM(M)NDI(K)(K)LQTAYP	127
6	205	LVPMVNPDGVNLVINGP	221

Probes		Nucleotide sequences
1	5′	CAT CAT GCT GTA GTC ATA ATT 3′ AGA A A G G T C
2	5′	·CC ATC GGG GTT AAC CAT CGG 3′ G A A G G T T A C C T

Figure 2 N-terminal amino-acid sequences of peptides 4, 2, 8 and 6 produced by papain hydrolysis of the *B. sphaericus* peptidase I and nucleotide sequences of probes 1 and 2

The isolated peptides are those shown in Figure 1. The amino-acid numbering is attributed on the basis of the primary structure of the *B. sphaericus* peptidase I (Figure 4). The amino-acid residues in parentheses were not determined with certainty by chemical sequencing. The two degenerated nucleotides were synthesized on the basis of the sequences of peptide 8 (probe 1) and peptide 6 (probe 2).

Gene cloning

Attempts to clone the *B. sphaericus* peptidase-I-encoding gene were first carried out with *Bam*HI, *Eco*RI, *Hin*dIII, *Pst*I, *Sal*I and *Sph*I libraries made in pBR322 and pSP73, by using $CaCl_2$-treated *E. coli* HB101 as host and the ^{32}P-labelled nucleotides 1 and 2 as probes. Of the 50000 recombinant clones analysed, none gave a hybridization signal.

Given that *B. sphaericus* produces the restriction endonuclease *Bsp*1286, its DNA is probably methylated (Raleigh et al., 1988), making it difficult to clone the desired gene in an *E. coli* host that contains the Mcr/Mrr restriction enzyme (Blumenthal, 1986; Woodcock et al., 1988). Consequently, samples containing 2.5×10^8 cells of *E. coli* DH5αMCR (*recA*$^-$, *mcrA*$^-$, *mcrB*$^-$ and *mrr*$^-$) were electroporated with 5 ng of purified recombinant pBR322 plasmids prepared from *Hin*dIII, *Eco*RI and *Sph*I libraries. The yield was 5×10^5 transformed cells per μg of DNA, with a survival rate of about 1%. Of the 6000 recombinant clones analysed, two originating from the *Hin*dIII library, hybridized with probes 1 and 2. The corresponding plasmids, pDML205 and pDML206, contained an insert of 6.6 kb and 7.8 kb respectively. Upon restriction, each plasmid yielded a small *Spe*I–*Spe*I 461 bp DNA fragment which, after separation by agarose-gel electrophoresis and transfer to a nylon membrane, gave a strong hybridization signal with probe 1. This DNA segment was subcloned in pSL1190. Sequencing, using the universal and reverse primers, showed that it had the information for a 153-amino-acid polypeptide and that this polypeptide contained the N-terminal sequence of peptide 8 at an internal position.

Gene sequencing and primary structure of *B. sphaericus* peptidase I

Nucleotide sequencing of pDML205 was performed in both directions upstream and downstream from the *Spe*I–*Spe*I DNA fragment, by using the strategy shown in Figure 3. The sequenced 1710 bp DNA segment contained an open reading frame (ORF), 1188 nucleotides long, which started with a GTG triplet coding for methionine at position 220 (itself preceded by a putative Shine–Dalgarno sequence) and terminated with a TAA codon at position 1408 (Figure 4). It translated into a 396-amino-acid protein (Figure 4). This protein had no signal peptide. Its theoretical molecular mass, 44724 Da, coincided perfectly with the 45 kDa value attributed to the isolated peptidase I on the basis of its migration by SDS/PAGE (Figure 1). Its theoretical 5.46 pI value also coincided with the experimental 5.40 pI value (Garnier et al., 1985). Peptides 4, 2, 8 and 6 obtained by papain hydrolysis of peptidase I and of established N-terminal sequence (Figure 2) were easily identified along the primary structure of the protein (Figure 4). Downstream from the stop codon of the *B. sphaericus* peptidase-I-encoding gene, a second ORF was identified which started with an ATG codon at position 1594

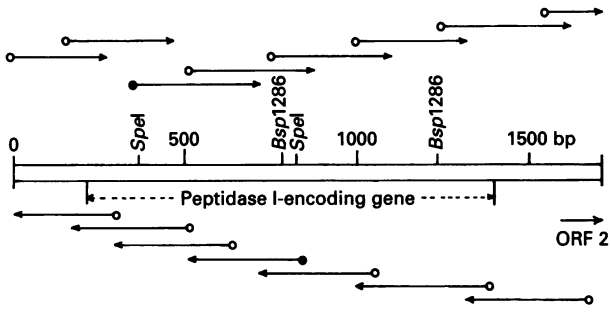

Figure 3 Strategy used for sequencing the *B. sphaericus* peptidase-I-encoding gene. Detection of an additional ORF

Key to symbols: ●, sequences initiated with the M13 universal or reverse primer; ○, sequences initiated with synthetic nucleotides; arrows, orientation and length of the sequenced DNA segments. Abbreviation: ORF, open reading frame.

(Figure 4). This ORF extended downstream from the 3′ end of the sequenced 1710 bp DNA segment.

P.i.x.e.

P.i.x.e. of the *B. sphaericus* peptidase I (see the Materials and methods section) detected the presence of 10.9 equivalent Zn^{2+} and 107 equivalent K^+ for 11.25 nmol of protein, i.e. almost exactly 1 equivalent Zn^{2+} and 10 equivalent K^+ per mol of protein.

Sequence identity searches, hydrophobic-cluster analysis and amino-acid alignments

Provided that the 100-amino-acid-residue N-terminal region of the protein was excluded from the analysis, sequence identity searches (see the Materials and methods section) led to the conclusion that the *B. sphaericus* peptidase I had similarity with the zinc peptidases of the carboxypeptidase A family. Among them, the *S. griseus* carboxypeptidase (Narahashi, 1990) and the prototypic bovine carboxypeptidase A exhibited the closest match.

The amino-acid sequence of the *B. sphaericus* peptidase I and that of the *S. griseus* carboxypeptidase were submitted to hydrophobic-cluster analysis by reference to bovine carboxypeptidase A (Figure 5). In this graphical representation, the hydrophobic residues Phe, Ile, Leu, Met, Val, Trp and Tyr are encircled; the hydrophobic clusters are also delineated; the hydrophobic residues and clusters occurring at equivalent places in the three sequences are in bold; the other amino-acid residues occurring as strict identities are marked by scattered points; and the secondary structures of carboxypeptidase A are identified.

The linear amino-acid alignments derived from the data of Figure 5 are given in Figure 6. The alignments between the pair *B. sphaericus* peptidase I and bovine carboxypeptidase A, the pair *B. sphaericus* peptidase I and *S. griseus* carboxypeptidase, and the pair *S. griseus* carboxypeptidase and bovine carboxypeptidase A had scores which were 16, 18 and 46 S.D. units respectively, above those expected from a run of 20 randomized pairs of proteins having the same amino-acid composition as the pairs under comparison (see the Materials and methods section). The three peptidases bore a similar signature in the form of conserved amino-acid residues (Figure 6).

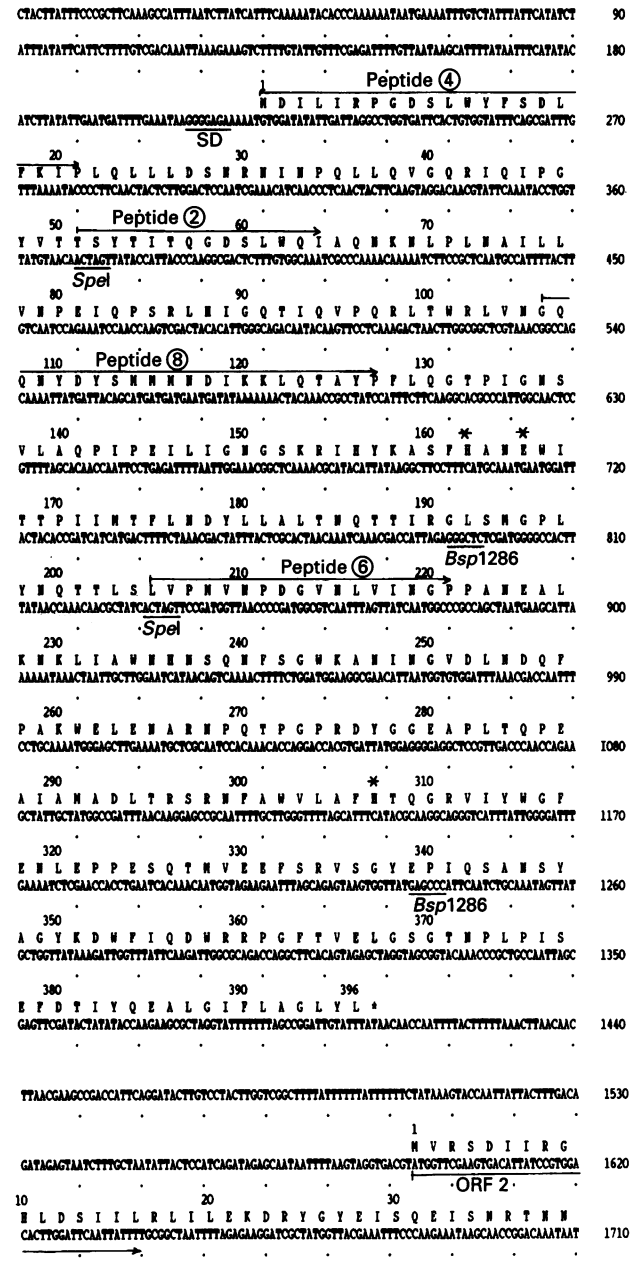

Figure 4 Nucleotide sequence of the *B. sphaericus* peptidase-I-encoding gene and the deduced amino-acid sequence of the protein

The Shine–Dalgarno sequence (SD) and the *SpeI* and *Bsp*1286 restriction sites occurring along the nucleotide sequence of the peptidase-I-encoding gene are indicated. The N-terminal regions of peptides 4, 2, 8 and 6 and the zinc ligands (*) are also indicated. Also shown is the 5′ region of a second ORF located downstream from the peptidase-I-encoding gene.

The 100-amino-acid-residue N-terminal domain of the *B. sphaericus* peptidase I consisted of two peptide segments, from Ile^5 to Thr^{50} and from Ile^{55} to Thr^{100}, having similarity with each other (Figure 7). These segments also had similarity with peptide segments known to occur as repeats in the C-terminal regions of various wall peptidoglycan hydrolases (Joris et al., 1992), namely the *Enterococcus hirae* muramidase 2 (six repeats) (Chu et al., 1992), the *Streptococcus* (*Enterococcus*) *faecalis* autolysin (four repeats) (Beliveau et al., 1991) and the *Bacillus* ϕPZA and ϕ29 lysozymes (two repeats) (Garvey et al., 1986; Paces et al., 1986).

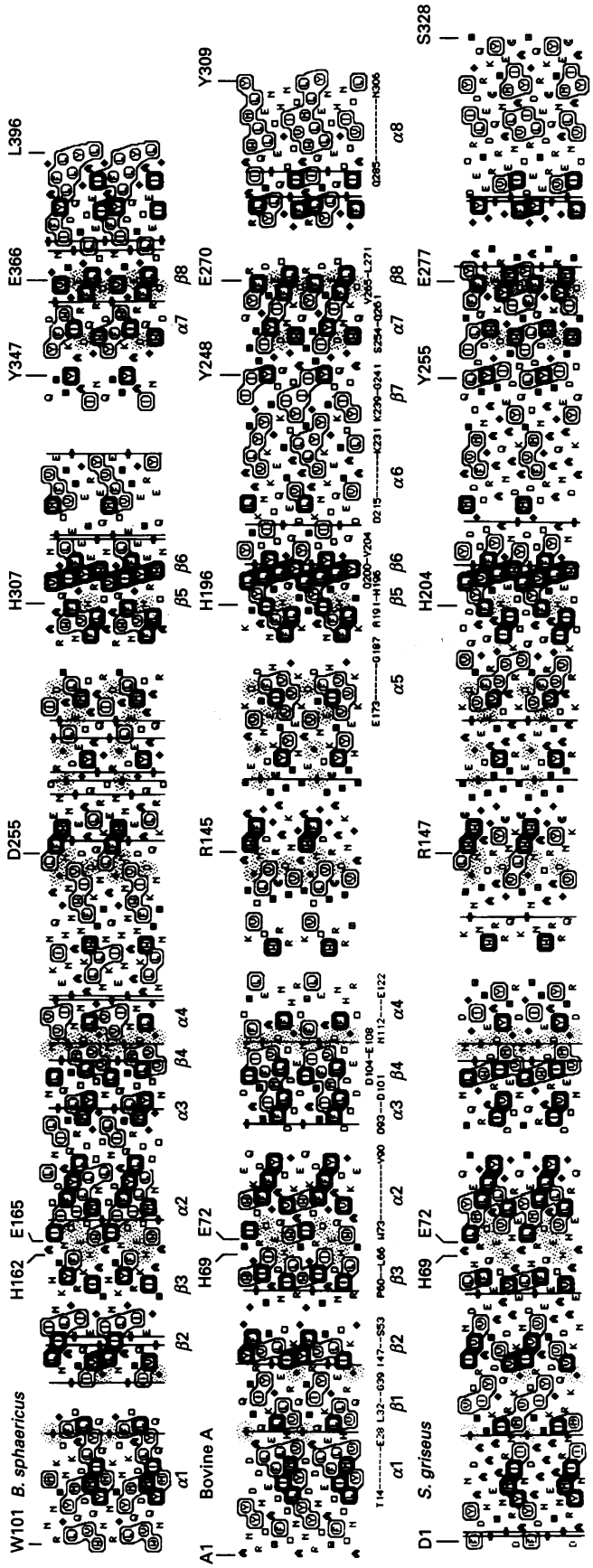

Figure 5 Hydrophobic- and hydrophilic-cluster analysis of the amino-acid sequences of bovine carboxypeptidase A, the *S. griseus* carboxypeptidase and the catalytic domain, from Trp[101] to Leu[306] of the *B. sphaericus* peptidase I

The hydrophobic amino-acid residues and clusters conserved in the three sequences are indicated in bold. The hydrophilic residues occurring as strict identities are marked by scattered points. Structurally/catalytically important amino-acid residues and the known secondary structures of bovine carboxypeptidase A are specified. *, Pro; ◆, Gly; ☇, Cys; ⊡, Ser; □, Thr.

```
B.s.  101   WRLVNGQ QNYDYSMMMN DIKKLQTAYP FLQGT..... .PIGNSVLAQ 141 PIPEILIGNG .SKRIHYKAS FHANEWITTP IIMTFLNDYL LALTNQTTIR 190
Bov.    1   ARSTNTFNYA TYHTLDEIYD FMDLLVAEHP QLVSKLQIGR SYEGRPIYVL  50 KFSTG...GS NRPAIWIDLG IHSREWITQA TGVWFAKKFT EDYGQ.....  92
S.g.    1   ...DFPPADS RYHNYAEMNA AIDARIAANP SIMSKRVIGK TYQGRDVIAV  47 KVSDNVAADE AEPEVLFTAH QHAREHLTVE MALYLLRELG QGYGS.....  92

B.s.        GLSMGPLYNQ TTLSLVPMVN PDGVNLVING PPANEALKNK LIAWNHNSQN 240 FSGWKANING VDLNDQFPAK WELENARNPQ TPGPRDYGGE APLTQPEAIA 290
Bov.        DPSFTAILDS MDIFLEIVTN PDGFAFT..H SQNRL..... ...WRKTRSV 132 ..TSSSLCVG VDANRNWDAG FGKAGA..SS SPCSETYHGK YANSEVEVKS 178
S.g.        DSRITQAVNG RELWIVPDMN PDGGEYDIAS GSYRS..... ...WRKNRQP 134 ..NAGSSAVG TDLNRNWAYK WGCCGG.SSS SPSSETYRGA AAESAPETKV 181

B.s.        MADLTRS... ...RNFAWVL AFHTQGRVIY WGFENLEPPE SQT..MVEEF 332 SRVSGYEP.. .........I QSANSY...A GYKDWFIQDW RRPGFTVELG 368
Bov.        IVDFVKDHG. ....NFKAFL SIHSYSQLLL YPYGYTTQSI PDKTELNQVA 223 KSAVEALKSL YGTSYKYGSI I.TTIYQASG GSIDWSYNQG IKYSFTFELR 272
S.g.        VADFVRSRVV GGKQQITAAI DFHTYSELVL WPFGYTYNDT APG..MTADD 229 RNAFAAVGQK MAASNGYTAE QSSDLYITDG SIDDWLWGSQ KIFGYTFEMY 279

B.s.        SGTNPLPISE FDTIYQEALG IFLAGLYL.. .......... .......... 396 ..........
Bov.        DT.......G RYGFLLPASQ IIPTAQETWL GVLTIMEHTL NNLY...... 309 ..........
S.g.        PRSAS..... GGGFYPPDEV IERETSRNRD AVLQLIENAD CMYRSIGKEA 324 QYCS......

Signature
B.s.        H162(X2)E165(X89)D255(X51)H307(X39)Y347(X18)E366
Bov.        H 69 (X2)E 72 (X72)R145(X50)H196(X51)Y248(X21)E270
S.g.        H 69 (X2)E 72 (X74)R147(X56)H204(X50)Y255(X21)E277
```

Figure 6 Linear amino-acid alignments and signature of the *S. griseus* carboxypeptidase (S.g.), bovine carboxypeptidase A (Bov.) and the catalytic domain of the *B. sphaericus* peptidase I (B.s.)

The proposed alignments derive from the data shown in Figure 5. The catalytic domain of the *B. sphaericus* peptidase I starts at Trp^{101}. Identities: the pair *S. griseus* carboxypeptidase and carboxypeptidase A has 87 identities (28%) (black dots), the pair *S. griseus* carboxypeptidase and *B. sphaericus* peptidase I has 58 identities (19%) (not indicated), the pair *B. sphaericus* peptidase I and carboxypeptidase A has 45 identities (15%) (black dots) and the triad *S. griseus* carboxypeptidase, carboxypeptidase A and *B. sphaericus* peptidase I has 25 identities (8%) (not indicated). For the roles assigned to the amino-acid residues that constitute the signature of bovine carboxypeptidase A, see the Discussion section.

```
B.s.          5 IRPGDSLWYFSDLFKIPLQLLLDSNRNI-NPQLLQVGQRIQIPGYVTTSYT  54
             55 ITQGDSLWQIAQNKNLPLNAILLVNPEI-QPSRLHIGQTIQVPQRLT  100

φPZA          VKSGDNLTKIAKKHNTTVATLLKLNPSIKDPNMIRVGQTINV
Consensus       T S   VDNK   SR MS   E TN  H K     RL

S.f.          VKSGDTLNKIAAQYGVSVANLRSWN-GIS-GDLIFAGQKLIVKKGTS
Consensus     I         S  F  T          K      V  TI     A

E.h.          VKAGESVWGVANKHGISMNQLIEWN-NIKN-NFIYPGQKLIVKKG
Consensus       S D   KISHSFH T A   Q           V    QVVISG
                    S  DDN    D                   E T

General       IK-GDSL--IA---------I---N--I-----I--GQ-I-I
Consensus     VR  ETV  VS         L            L    L L
              (T) (N)  F                       V    V V
```

Figure 7 Amino-acid alignments of the N-terminal tandem segments (Ile^{5}–Thr^{50}; Ile^{55}–Thr^{100}) of the *B. sphaericus* peptidase I (*B.s.*) and the C-terminal repeats of the *Bacillus* sp. φ PZA lysozyme (two repeats), the *Streptococcus* (*Enterococcus*) *faecalis* autolysin (*S.f.*; four repeats) and the *Enterococcus hirae* muramidase 2 (*E.h.*; six repeats)

●: Identities (19 for 46 aligned amino-acid residues; i.e. 41%) present in the two N-terminal repeats of *B. sphaericus* peptidase I. The other aligned sequences are consensus of: (i) the two segments 165–206 and 216–257 of the φ PZA lysozyme; (ii) the four segments 365–409, 433–477, 501–545 and 574–618 of the *S. (E.) faecalis* autolysin; (iii) the six segments 259–301, 340–382, 415–457, 490–532, 565–607 and 624–666 of the *E. hirae* muramidase 2; and (iv) the 14 peptide segments under comparison. For references, see the main text.

Effects of detergents

The activity of the *B. sphaericus* peptidase I was measured in buffer under conditions of initial velocity in the absence and presence of various detergents (0.4%, final concentration; see the Materials and methods section). At 37 °C, the specific enzyme activity, 14 units per mg of protein in the absence of detergent (see the Materials and methods section), increased to 85 and 140 units per mg of protein in the presence of the anionic SDS and Sarkosyl respectively, to 200 units per mg of protein in the presence of the cationic HPC1 and Cetavlon and the non-ionic *n*-octylglucoside and Triton X-100, and to 265 units per mg of protein in the presence of the zwitterionic CHAPS. After a 24 h incubation at 37 °C in the presence of SDS and at 60 °C in the presence of CHAPS, the enzyme retained 20% and 75% of the original activity respectively.

DISCUSSION

The γ-D-Glu-(L)*ms*A_2pm-hydrolysing peptidase I of *B. sphaericus* has similarity in the primary structure with the metallo (Zn) bovine carboxypeptidase A and *S. griseus* carboxypeptidase except that the *B. sphaericus* peptidase I is a two-module protein. The 296-amino-acid-residue C-terminal catalytic domain (Trp^{101}–Leu^{396}) bears, fused at its N-terminal end, a 100-amino-acid-residue polypeptide extension. Hence, similarity with carboxypeptidase A and the *S. griseus* carboxypeptidase is restricted to the C-terminal catalytic domain of the *B. sphaericus* peptidase I. These conclusions rest upon the following observations.

The catalytic domain of the *B. sphaericus* peptidase I, the *S. griseus* carboxypeptidase and carboxypeptidase A have a similar pattern of distribution of hydrophobic clusters and of hydrophilic

residues between conserved hydrophobic clusters (Figure 5). The alignment requires few deletions/insertions to be made in the sequences. Peptide segments equivalent to strands β1 to β8 and to several α-helices of carboxypeptidase A are easily identified in the sequences of the *B. sphaericus* peptidase I and *S. griseus* carboxypeptidase. These two peptidases, however, may lack α5 and the *B. sphaericus* peptidase I also lacks the C-terminal α8.

The family prototypic bovine carboxypeptidase A bears a unique signature in the form of several amino-acid residues (Figure 6). These residues are brought together within the active sites as a result of the polypeptide folding. They play essential roles in zinc binding (His^{69}, Glu^{72}, His^{196}), substrate binding (Arg^{145}) and catalysis (Tyr^{248}, Glu^{270}) (Argos et al., 1978; Kim and Lipscomb, 1990; Vallee and Auld, 1990; Le Huërou et al., 1991). The triad His^{162}–Glu^{165}–His^{307} of the *B. sphaericus* peptidase I and the triad His^{69}–Glu^{72}–His^{204} of the *S. griseus* carboxypeptidase align with the zinc-binding triad His^{69}–Glu^{72}–His^{196} of carboxypeptidase A (Figures 5 and 6). Similarly, the dyad Tyr^{347}–Glu^{366} of the *B. sphaericus* peptidase I and the dyad Tyr^{255}–Glu^{277} of the *S. griseus* carboxypeptidase align with the catalytic dyad Tyr^{248}–Glu^{270} of carboxypeptidase A. However, the substrate-binding Arg^{145} of carboxypeptidase A, which aligns with Arg^{147} of the *S. griseus* carboxypeptidase, aligns with Asp^{255} of the *B. sphaericus* peptidase I. Consistently with the specificity profile of the *B. sphaericus* peptidase I (see the Introduction), substrate binding might be mediated via a salt linkage between the carboxylate of Asp^{255} and the free amino group on the D-centre of *ms*A_2pm in an ε position to the scissile γ-D-Glu–*ms*A_2pm peptide bond. Note also that at variance with carboxypeptidase A and the *S. griseus* carboxypeptidase, the *B. sphaericus* peptidase I has no cysteine residues.

The N-terminal extension of the *B. sphaericus* peptidase I consists of two peptide segments (Ile^5–Thr^{50} and Ile^{55}–Thr^{100}) which have 41% identities (Figure 7). Repeats are found frequently among carbohydrate-binding proteins, wall peptidoglycan autolysins and lytic enzymes. They vary in number, from two to six, per protein molecule; they occur often, but not always, as C-terminal extensions; they fall into several families and, depending on the family to which they belong, they are believed to participate in the recognition of a particular moiety of the ligand. The repeats of *B. sphaericus* peptidase I do not exhibit similarity to the repeats of the carbohydrate-binding proteins reviewed by Wren (1991), the surface proteins from Gram-positive cocci reviewed by Fischetti et al. (1991), and the pneumococcal peptidoglycan LYTA amidase and CPL-1, CPL-7 and CPL-9 muramidases (Sanz et al., 1992). They show similarity with the repeats of other wall peptidoglycan hydrolases such as the *E. hirae* muramidase 2, the *S.* (*E.*) *faecalis* autolysin, the *Bacillus subtilis* φPZA lysozyme and the homologous *Bacillus* gene 15 φ 29 lysozyme (Figure 7) (Joris et al., 1992).

The N-terminal catalytic domains of the pneumococcal LYTA *N*-acetylmuramoyl-L-alanine amidase and CPL-1 and CPL-7 *N*-acetylmuramidases have been expressed in *E. coli* (Sanz et al., 1992). The repeat-free proteins, though able to adopt an active conformation, have nevertheless a lower enzyme activity than that of the complete proteins, suggesting that the presence of a substrate-binding domain is an advantage for enzymes that interact with polymeric substrates. The enhancing effects that detergents exert on the activity of the *B. sphaericus* peptidase I are probably due to specific interactions between the repeats of the protein molecule and the detergent micelles. They suggest a close interplay between the N-terminal repeats and the C-terminal catalytic domain of the protein. Deletion of the N-terminal ligand-binding domain by genetic engineering and expression of the truncated protein should help clarify the point.

This work was supported in part by the Belgian programme on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming (PAI no. 19), the Fonds de la Recherche Scientifique Médicale (contract no. 3.4531.92), the Centre National de la Recherche Scientifique and an Action Intégrée franco–belge (projects 90-20, 91-10 and 90-2-5). M.L.H. was a CEE fellow (Bridge programme S/BIOT-900022). C.D., B.J. and G.W. are chercheurs qualifiés of the Fonds National de la Recherche Scientifique, Brussels.

REFERENCES

Argos, P., Garavito, R. M., Eventoff, N., Rossmann, M. and Brändén, C. I. (1978) J. Mol. Biol. **126**, 141–158
Arminjon, F., Guinand, M., Vacheron, M. J. and Michel, G. (1977) Eur. J. Biochem. **73**, 557–565
Baji-Kourda, F., Guinand, M., Vacheron, M. J. and Michel, G. (1989) Biotechnol. Appl. Biochem. **11**, 169–175
Beliveau, C., Potvin, C., Trudel, J., Asselin, A. and Bellemare, G. (1991) J. Bacteriol. **173**, 5619–5623
Blumenthal, R. M. (1986) Trends Biotechnol. **4**, 302
Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M. C., Keyneker, H. L., Boyer, H. W., Crosa, H. J. and Falkow, S. (1977) Gene **2**, 95–113
Bourgogne, T., Vacheron, M. J., Guinand, M. and Michel, G. (1992) Int. J. Biochem. **24**, 471–476
Boyer, H. W. and Roulland-Dussoix, D. (1969) J. Mol. Biol. **41**, 459–472
Bradford, M. M. (1976) Anal. Biochem. **72**, 248–254
Brosius, J. (1989) DNA **8**, 759
Chu, C. P., Kariyama, R., Daneo-Moore, L. and Shockman, G. D. (1992) J. Bacteriol. **174**, 1619–1625
Cleveland, D. W., Fischer, S. G., Kirschner, M. W. and Laemmli, U. K. (1977) J. Biol. Chem. **252**, 1102–1106
Devereux, J., Haeberli, P. and Smithies, O. (1984) Nucleic Acids Res. **12**, 387–395
Dower, W. J., Miller, F. J. and Ragsdale, C. W. (1988) Nucleic Acids Res. **16**, 6127–6145
Fischetti, U. A., Pancholi, V. and Schneewind, O. (1991) In Genetics and Molecular Biology of Streptococci and Enterococci (Dunny, G. M., Cleary, P. P. and McKey, L. L., eds.), pp. 290–294, American Society for Microbiology, Washington, DC
Gaboriaud, C., Bissery, V., Benchetrit, T. and Mornon, J. P. (1987) FEBS Lett. **224**, 149–155
Garnier, M., Vacheron, M. J., Guinand, M. and Michel, G. (1985) Eur. J. Biochem. **148**, 539–543
Garvey, K. J., Sardi, M. S. and Ito, J. (1986) Nucleic Acids Res. **14**, 10001–10008
Goad, N. B. and Kanehisa, M. T. (1982) Nucleic Acids Res. **10**, 247–263
Guinand, M., Michel, G. and Tipper, D. J. (1974) J. Bacteriol. **120**, 173–184
Guinand, M., Vacheron, M. J., Michel, G. and Tipper, D. J. (1979) J. Bacteriol. **138**, 126–132
Henrissat, B., Salohemo, M., Lavaitte, S. and Knowles, J. K. C. (1990) Proteins: Struct., Funct., Genet. **8**, 251–257
Hopwood, D. A., Bibb, M. J., Chater, K. F., Kieser, T., Bruton, C. J., Kieser, H. M., Lydiate, D. J., Smith, C. P., Ward, J. M. and Schrempf, H. (1985) Genetic Manipulation of Streptomyces: A Laboratory Manual, The John Innes Foundation, Norwich, U.K.
Hourdou, M. L., Duez, C., Joris, B., Vacheron, M. J., Guinand, M., Michel, G. and Ghuysen, J. M. (1992) FEMS Microbiol. Lett. **91**, 165–170
Joris, B., Englebert, S., Chu, C. P., Kariyama, R., Daneo-Moore, L., Shockman, G. D. and Ghuysen, J. M. (1992) FEMS Microbiol. Lett. **91**, 257–264
Jessee, J. and Bloom, F. (1988) BRL Focus **10**, 69–70
Kalb, V. F. and Bernlohr, R. W. (1977) Anal. Biochem. **82**, 362–371
Kanehisa, M. I. (1982) Nucleic Acids Res. **10**, 183–195
Kim, H. and Lipscomb, W. N. (1990) Biochemistry **29**, 5546–5555
Laemmli, U. K. and Favre, M. (1973) J. Mol. Biol. **80**, 575–599
Le Huërou, I., Guilloteau, P., Toullec, K., Puigserver, A. and Wicker, C. (1991) Biochem. Biophys. Res. Commun. **175**, 110–116
Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982) Molecular Cloning. A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
Narahashi, Y. (1990) J. Biochem. **107**, 879–886
Paces, V., Uleek, C. and Urbane, P. (1986) Gene **14** 107–114
Pearson, W. R. and Lipman, D. J. (1988) Proc. Natl. Acad. Sci. U.S.A. **85**, 2444–2448
Raleigh, E. A., Murray, N. E., Revel, H., Blumenthal, R. M., Westaway, D., Reigh, A. D., Rigby, A. D., Rigby, P. W. J., Elhai, J. and Hanakan, D. (1988) Nucleic Acids Res. **16**, 1563–1576
Sanger, F., Nicklen, S. and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 5463–5467
Sanz, J. M., Diaz, E. and Garcia, J. L. (1992) Mol. Microbiol. **6**, 921–931
Tipper, D. J., Pratt, I., Guinand, M., Holt, S. C. and Linnett, P. E. (1977) Microbiology **3**, 50–68
Vacheron, M. J., Guinand, M., Françon, A. and Michel, G. (1979) Eur. J. Biochem. **100**, 189–196

Valentin, C., Vacheron, M. J., Martinez, C., Guinand, M. and Michel, G. (1983) Biochimie **65**, 239–245
Vallee, B. L. and Auld, D. S. (1990) Biochemistry **29**, 5647–5659
Wallace, R. B., Johnson, N. J., Hirose, T., Miyake, M., Kawashima, E. H. and Itakura, K. (1981) Nucleic Acids Res. **9**, 879–894
Woodcock, D. M., Crowther, P. J., Diver, W. P., Graham, M., Bateman, C., Baker, D. J. and Smith, S. S. (1988) Nucleic Acids Res. **16**, 4465
Woods, D. (1984) BRL Focus **6**, 1–3
Wren, B. W. (1991) Mol. Microbiol. **5**, 797–803
Zhang, H., Scholl, R., Browse, J. and Somerville, C. (1988) Nucleic Acids Res. **16**, 1220

Received 5 October 1992/21 December 1992; accepted 4 January 1993