FEMS Microbiology Letters 71 (1990) 215–220
Published by Elsevier

FEMSLE 04135

Cloning, nucleotide sequence and amplified expression of the gene encoding the extracellular metallo (Zn) DD-peptidase of *Streptomyces albus* G

Colette Duez, Bernard Lakaye, Simone Houba, Jean Dusart and Jean-Marie Ghuysen

Service de Microbiologie, Université de Liège, Institut de Chimie, B6, Sart Tilman, Belgium

Received 2 April 1990
Revision received 7 May 1990
Accepted 14 May 1990

Key words: Cloning of methylated DNA; *Streptomyces* plasmids; *Streptomyces* signal peptides

1. SUMMARY

The gene encoding the extracellular metallo (Zn) DD-peptidase of *Streptomyces albus* G has been cloned in *Escherichia coli* DH5αMCR via pBR322 or 325, and then transferred into *Streptomyces lividans* TK24 via pIJ486, with substantial amplification of the expressed DD-peptidase. The gene has the information for the synthesis of a 255 amino acid precursor, the amino terminal region of which has the characteristic features of a signal peptide. The primary structure as deduced from nucleotide sequencing confirms that previously determined by chemical methods except for the occurrence of an Asp instead of Asn at position 1 and an additional Ala immediately downstream of Pro67.

Correspondence to: J. Dusart, Service de Microbiologie, Université de Liège, Institut de Chimie, B6, B-4000 Sart Tilman (Liège 1), Belgium.

2. INTRODUCTION

The metallo (Zn) DD-peptidase of *Streptomyces albus* G is a powerful bacteriolytic agent that hydrolyses D-alanyl-D-αα peptide bonds in α-position to a free carboxylate, which bonds serve to cross-link the peptide-substituted glycan chains in wall peptidoglycans of chemotypes I and IV [1–4]. The primary [5] and tertiary [6] structures of the DD-peptidase have been established by chemical methods and X-ray crystallography, respectively. On the basis of these data, the DD-peptidase is a 212 amino acid two-domain protein. A small amino terminal domain consisting of three α-helices is connected to a large carboxy terminal domain consisting of three α-helices and a five stranded β-sheet. The active site, located in the carboxy terminal domain, is topologically defined by the β-sheet at the back and two loops on each side of the cavity. While establishment of the atomic structure of the protein at high resolution is in progress, cloning of the relevant gene has

0378-1097/90/$03.50 © 1990 Federation of European Microbiological Societies

been undertaken. Indeed, site-directed mutagenesis and other DNA recombinant techniques should allow to map out the catalytically important amino acid side-chains and to unravel the possible functions(s) played by the small amino terminal domain.

3. MATERIALS AND METHODS

Streptomyces albus G was from this laboratory. *Streptomyces lividans* TK24 (str-6; a strain cured of its natural plasmids) [7] and the non-conjugative, high-copy number plasmids pIJ702 [8], pIJ385 [9], pIJ699 [10] and pIJ486 [11] were from the John Innes Institute, Norwich, U.K. *Escherichia coli* strains HB101 [12] and DH5αMRC [13] and plasmids pBR322 [14] and pBR325 [15] were also used.

The recombinant DNA techniques were applied and *E. coli* HB101 was transformed as described in Maniatis et al. [16]. *Streptomyces* chromosomal and plasmid DNAs were prepared and *S. lividans* TK24 was transformed as described in [9]. All the libraries were constructed by shotgun cloning, using dephosphorylated vectors. *E. coli* DH5αMRC was transformed by electroporation with a Bio-Rad Gene pulser apparatus [17]. Synthetic nucleotide probes were purchased from CRI (Centre Recherche Industrielle, Nivelles, Belgium) or Eurogentec (Liège, Belgium). Restriction DNA fragments were separated by agarose gel electrophoresis, transferred to nitrocellulose membranes and hybridized with the ^{32}P-labelled probes. Preparation of *E. coli* colony filters was as in [18]. Hybridizations were performed at tm–5°C for 18 h using a modified Southern procedure [19].

DNA sequencing was made using the dideoxynucleotide chain-termination method [20]. Zones of base compression, due to high G + C contents, were resolved using the formamide procedure [9] or dITP instead of dGTP (sequenase kit, USB, Cleveland, U.S.A.). Codon usage was assessed with Staden's program [21] and Fickett's test [22] using, as reference, the gene that encodes the *S. albus* G β-lactamase [23].

Rabbit antiserum directed against the *S. albus* G DD-peptidase was prepared by Gamma S.A. (Tavier, Belgium) and served to screen DD-peptidase secretory clones according to Bio-Rad immuno Blot Alkaline Phosphatase Assay System.

DD-peptidase-producing *Streptomyces* strains were grown at 28°C with vigorous orbital shaking in 250-ml flasks containing 30 ml of a modified buffered YEME medium [24] and 5 μg thiostrepton ml^{-1}. The pH was adjusted at time intervals. The level of excreted DD-peptidase was estimated (reaction catalysed: Ac_2-L-Lys-D-Ala-D-Ala + $H_2O \rightarrow$ D-Ala + Ac_2-L-Lys-D-Ala) as described in [25].

4. RESULTS AND DISCUSSION

Attempts to clone the gene encoding the *S. albus* G Zn^{2+} DD-peptidase using pIJ702, pIJ385 and pIJ699 as vectors, *Bcl*I, *Bam*HI, *Sac*I, *Sph*I, *Cla*I and *Bgl*II as restricted enzymes and *S. lividans* as host, failed. None of the 140 000 recombinant clones analysed produced the desired protein as evidenced by the immunological test.

Attempts were then made to clone the gene in *E. coli* via PBR322 and pBR325, using for this purpose, an expression-independent screening procedure. The 17-nucleotide probe 3′GTG-TAC-ATG-CCI-GTG-CG5′ complementary to the sequence encoding the His154-Met-Tyr-Gly-His-(Ala)159 hexapeptide of the *S. albus* G protein (the numbering is that used in Fig. 2), and the 23-nucleotide probe 3′TAC-ACC-TTC-$^{A}_{G}$A$^{G}_{C}$-GT$^{T}_{C}$-GGG-TAC-GC5′ complementary to the sequence encoding the Met116-Trp-Lys-Leu-Gln-Ala-Met-(Arg)123 octapeptide were synthesized and served to analyse the recombinant clones. Note that the probes show little degeneracy due to the bias codon usage in *Streptomyces* genus. One single clone originating from the *Bgl*II library, gave a hybrization signal with the two probes. The relevant plasmid pDML451 contained a 8 kb insert.

A 512-bp *Hin*dII DNA fragment was the smallest one that gave a hybridization signal. Subcloning into M13mp10 (digested with *Sma*I) and nucleotide sequencing showed that this fragment had the information for the carboxy terminal region, from Asn134 to Ile213, of the *S. albus* G

DD-peptidase. In turn, the 182-bp *Bam*HI/*Bgl*II-*Hin*dII nucleotide segment located immediately upstream of the 512-bp *Hin*dII stretch was sequenced. It had the information for the peptide Lys72-Val133 of the *S. albus* G DD-peptidase. That pDML451 lacked the information for the amino terminal region of the protein located upstream of Lys72 (i.e. for the small amino terminal domain of the protein; see the Introduction) was not surprising. Indeed, as deduced from the known primary structure, the gene encoding the *S. albus* G DD-peptidase was expected to contain a unique *Bgl*II site located close to this critical position.

To clone the entire *S. albus* G DD-peptidase-encoding gene, *Bam*HI, *Bcl*I, *Sph*I, *Pst*I, *Nco*I and *Not*I genomic libraries were introduced in *E. coli* via pBR322 or pBR325. Recombinant clones were screened with the help of the 28-nucleotide probe 3′C-TAG-ATG-GTC-AAC-GTC-CTG-CTG-CTG-ACG5′ complementary to the sequence encoding the (Lys72)-Ile-Tyr-Gln-Leu-Gln-Asp-Asp-Cys 81 decapeptide. The assays made with *E. coli* HB101 as host failed. A possible interpretation was that since *S. albus* G produces the restriction nuclease *Sal*I, its DNA is very probably methylated [26] making it difficult to achieve gene cloning in *E. coli* hosts that contain the Mcr/Mrr restriction enzymes (as reported in other similar cases [27,28]).

Consequently, *E. coli* DH5αMCR (*rec*A$^-$, *mcr*A$^-$, *mcr*B$^-$ and *mrr*$^-$) [13] was used as an alternative host. Samples of 2.4×10^8 cells were electroporated with 2 ng of the ligation mixtures. The yields varied from $3 \cdot 10^5$ to $7 \cdot 10^5$ transformed cells per μg of DNA and the survival rate was about 1%. About 7300 recombinant clones were examined. One of them, from the *Not*I library, hybridized with the 28-nucleotide probe. The corresponding plasmid pDML453 (Fig. 1A) had a 5-kb insert which contained an internal *Bgl*II site.

Nucleotide sequencing of pDML453 (in both directions and starting from this *Bgl*II site) was achieved using the strategy shown in Fig. 1B. The 5′A-A-GAGGT3′ segment at position −17 to −9 upstream of the ATG translation start codon presents the characteristics of a Shine-Dalgarno sequence [29], and, in turn, the 28-nucleotide sequence that starts at position 23 downstream of the TGA opal stop codon has a perfect inverted repeat of 7 bases and may function as a terminator [29,30]. The open reading frame had the information for the synthesis of a 255 amino acid precursor (Fig. 2), the 42 amino acid amino terminal region of which has the characteristic features

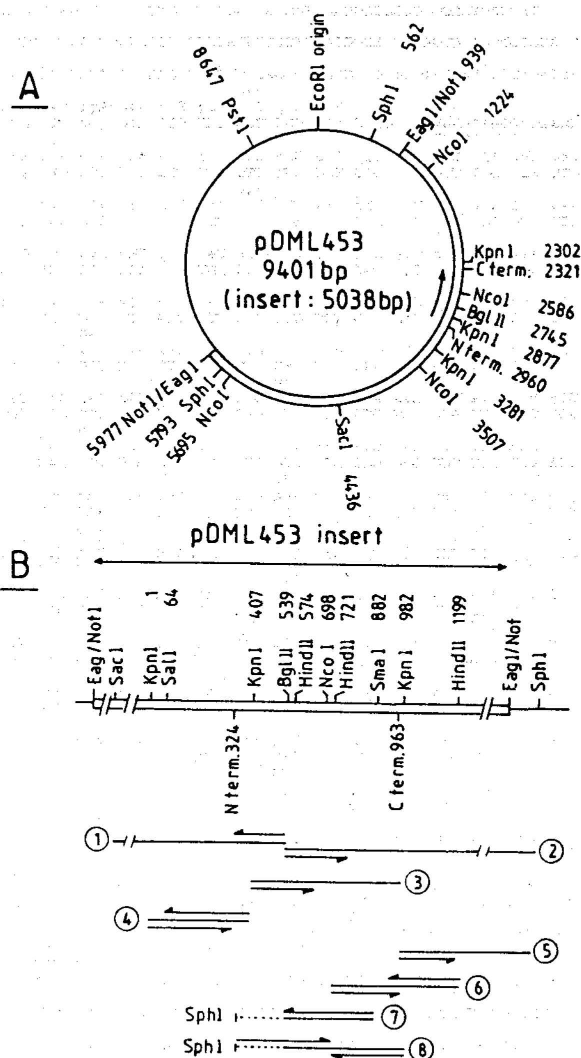

Fig. 1. Restriction map of pDML453 (A) and strategy used for nucleotide sequencing (B). The arrow (A) indicates the orientation of the *S. albus* G DD-peptidase gene. Subfragments (1–4) and (5–8) (B) derived from pDML453 and pDML451, respectively. The positions corresponding to the amino and carboxy terminal residues of the mature DD-peptidase are indicated. The arrows show the length of DNA sequenced in each fragment. ——— pBR322; ▭ *S. albus* G DNA.

```
GGTACCACTGGCGGGCGCGCGGGCCTGGAACCGGCGCTCCTCGGGGGTGGAGGTGGA
CGAGGCGTCGACCTGATGGGCCCGGACGTGGGCGTACGCGACCCGGGCCCAAGGATTTG
ACGAGTTCTTGACAATCGGGCCGGTGCGCGGACAGTCTCGGGGGCCCGTCCGGCACCAC
                                       -40
                           Met Arg Pro Arg Pro Ile Arg Leu Leu
GGATCAAGAAGAGGTGGCCCCCC ATG CGC CCA CGC CCC ATC AGA CTT CTG
               -30                                         -20
Leu Thr Ala Leu Val Gly Ala Gly Leu Ala Phe Ala Pro Val Ser
CTC ACC GCC CTC GTC GGG GCC GGT CTC GCG TTC GCC CCC GTC TCC
                                -10
Ala Val Ala Ala Pro Thr Ala Thr Ala Ser Ala Ser Ala Asp Val
GCC GTC GCC GCC CCC ACC GCG ACG GCC TCC GCC TCC GCC GAC GTC
                +1                                  10
Gly Ala Leu Asp Gly Cys Tyr Thr Trp Ser Gly Thr Leu Ser Glu
GGC GCC CTG GAC GGC TGC TAC ACC TGG AGT GGC ACC CTC AGT GAG
                                 20
Gly Ser Ser Gly Glu Ala Val Arg Gln Leu Gln Ile Arg Val Ala
GGC TCC TCC GGT GAA GCG GTG CGC CAG CTC CAG ATC CGC GTC GCC
            30                                      40
Gly Tyr Pro Gly Thr Gly Ala Gln Leu Ala Ile Asp Gly Gln Phe
GGG TAC CCG GGG ACC GGC GCC CAG CTC GCC ATC GAC GGG CAG TTC
                                 50
Gly Pro Ala Thr Lys Ala Ala Val Gln Arg Phe Gln Ser Ala Tyr
GGC CCC GCC ACC AAG GCG GCC GTG CAG CGC TTC CAG TCG GCC TAC
            60                                      70
Gly Leu Ala Ala Asp Gly Ile Ala Gly Pro [Ala] Thr Phe Asn Lys
GGA CTG GCC GCC GAC GGC ATC GCC GGG CCC [GCC] ACC TTC AAC AAG
                                 80
Ile Tyr Gln Leu Gln Asp Asp Asp Cys Thr Pro Val Asn Phe Thr
ATC TAC CAG TTG CAG GAC GAC GAC TGC ACG CCC GTC AAC TTC ACC
            90                                      100
Tyr Ala Glu Leu Asn Arg Cys Asn Ser Asp Trp Ser Gly Gly Lys
TAC GCC GAG CTG AAC CGC TGC AAC TCG GAC TGG TCC GGC GGC AAG
                                 110
Val Ser Ala Ala Thr Ala Arg Ala Asn Ala Leu Val Thr Met Trp
GTC TCC GCC GCC ACC GCC CGC GCC AAC GCC CTG GTC ACC ATG TGG
            120                                     130
Lys Leu Gln Ala Met Arg His Ala Met Gly Asp Lys Pro Ile Thr
AAG CTC CAG GCC ATG CGG CAC GCC ATG GGT GAC AAG CCG ATC ACC
                                 140
Val Asn Gly Gly Phe Arg Ser Val Thr Cys Asn Ser Asn Val Gly
GTC AAC GGC GGC TTC CGC AGC GTC ACC TGC AAC AGC AAC GTC GGC
            150                                     160
Gly Ala Ser Asn Ser Arg His Met Tyr Gly His Ala Ala Asp Leu
GGC GCC TCC AAC AGC CGC CAC ATG TAC GGC CAC GCG GCC GAC CTC
                                 170
Gly Ala Gly Ser Gln Gly Phe Cys Ala Leu Ala Gln Ala Ala Arg
GGC GCC GGC TCG CAG GGC TTC TGC GCG CTG GCC CAG GCC GCC CGC
            180                                     190
Asn His Gly Phe Thr Glu Ile Leu Gly Pro Gly Tyr Pro Gly His
AAC CAC GGC TTC ACC GAG ATC CTC GGC CCG GGC TAC CCG GGC CAC
                                 200
Asn Asp His Thr His Val Ala Gly Gly Asp Gly Arg Phe Trp Ser
AAC GAC CAC ACG CAC GTC GCC GGC GGC GAC GGC CGC TTC TGG TCG
            210             213
Ala Pro Ser Cys Gly Ile ***
GCG CCC AGC TGC GGC ATC TGAGCCCCGGCTCGCCGCGGTACCCCCGCCCTCGC
CACCCGGCGGCGGAGGGCGCCGTGCCGTGTCCGGGGGCGGCCGGTGCACGTCACACAGA
TGTCACACCGGAAAGCTACCGCTCAGTAGCCTCCCCGTGTAGCGTGGCCGGCATGGCTC
AGGCAGCACAGGCAACCATCGGTGACAGCGAGTTCGACCCGGACACGGCCGTCACCCTC
CGGGCCCCCGGCGTCTACGACACCGAACTCTCCGCGGGCTGGACGATCATCTCCGCCGT
```

Fig. 2. Nucleotide sequence of the *Streptomyces albus* G gene encoding the DD-peptidase precursor. Deduced amino acid sequence of the protein precursor. The additional alanine, missed by chemical sequencing of the mature protein, is boxed. The putative Shine-Dalgarno sequence is underlined. The inverted repeat of the putative transcription termination signal is shown by horizontal arrows. └───┘: *Bgl*II site.

of a signal peptide; three Arg occur at positions −41, −39 and −36 upstream of a Leu-18 Leu-8 hydrophobic segment, itself prolonged by several possible processing sites. The exact processing site remains uncertain. This deduced amino acid sequence of the precursor confirmed that previously established by chemical degradation of the secreted, mature protein except for the occurrence of an additional Ala immediately downstream of Pro67 and for the occurrence of Asp, instead of Asn, at position 1. That the secreted DD-peptidase consists of 213 amino acid residues, and not 212 residues as previously proposed [5], has been fully confirmed by refined X-ray crystallographic data [31; and unpublished]. No leader peptidase is known that would cleave a Leu(-1)-Asp(+1) peptide bond but immediately upstream there occurs a classical Val(-4)-Gly(-3)-Ala(-2) recognition site for such an enzyme. On this basis, cleavage of the Ala(-2)-Leu(-1) peptide bond would be followed by aminopeptidase elimination of Leu.

E. coli HB101 harbouring pDML453 did not express any detectable DD-peptidase activity as tested in supernatant of liquid cultures. The same observation has been made with the *S. albus* G β-lactamase gene cloned in *E. coli* by Dehottay et al. [32]. The 4.4 kb *Eco*RI-*Sac*I DNA fragment containing the desired gene was excised from pDML453 and transferred into *S. lividans* via pIJ486 (which plasmid contained a polylinker possessing the unique *eco*RI and *Sac*I sites). The DD-peptidase was expressed (as shown by immunoblotting of the colonies) and exported (as shown by measuring the DD-peptidase activity in the culture filtrates). Maximal yield was 10 mg enzyme × litre culture^{-1} after 72 h of growth. This level was 80-fold higher than that observed with the original *S. albus* G strain (0.13 mg enzyme × litre culture^{-1}) [33]. Cloning of the gene encoding the *S. albus* G DD-peptidase in *S. lividans* was thus achieved only via an intermediate subcloning step in *E. coli* DH5αMCR. A likely explanation of this observation is that replication in this particular *E. coli* strain gives rise to a non-methylated DNA that is table in the *Streptomyces* host.

ACKNOWLEDGEMENTS

This work was supported in part by the Fonds de la Recherche Scientifique Médicale (contract N° 3.4507.83), an Action concertée with the Belgian Government (convention 86/91-90), the Fonds de Recherche de la Faculté de Médecine

ULg and a contract with the EEC (BAP-0197-B). C.D. and J.D. are research associates of the Fonds National de la Recherche Scientifique (Brussels, Belgium). We are grateful to Bio-Rad Laboratories for lending a Gene Pulser apparatus.

REFERENCES

[1] Ghuysen, J.M., Leyh-Bouille, M., Bonaly, R., Nieto, M., Perkins, H.R., Schleifer, K.H. and Kandler, O. (1970) Biochemistry 9, 2955–2961.

[2] Leyh-Bouille, M.,Ghuysen, J.M., Bonaly, R., Nieto, M., Perkins, H.R., Schleifer, K.H. and Kandler, O. (1970) Biochemistry 9, 2961–2971.

[3] Leyh-Bouille, M., Ghuysen, J.M., Nieto, M., Perkins, H.R., Schleifer, K.H. and Kandler, O. (1970) Biochemistry 9, 2971–2975.

[4] Dideberg, O., Joris, B., Frère, J.M., Ghuysen, J.M., Weber, G., Robaye, R., Delbrouck, J.M. and Roelandts, I. (1980) FEBS Lett. 117, 215–218.

[5] Joris, B., Van Beeumen, J., Casagrande, F., Gerday, Ch., Frère, J.M. and Ghuysen, J.M. (1983) Eur. J. Biochem. 130, 53–69.

[6] Dideberg, O., Charlier, P., Dive, G., Joris, B., Frère, J.M. and Ghuysen, J.M. (1982) Nature 299, 469–470.

[7] Hopwood, D.A., Kieser, T., Wright, H.M. and Bibb, M.J. (1983) J. Gen. Microbiol. 129, 2257–2269.

[8] Katz, E., Thompson, C.J. and Hopwood, D.A. (1983) J. Gen. Microbiol. 129, 2703–2714.

[9] 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, UK.

[10] Kieser, T. and Melton, R.E. (1988) Gene 65, 83–91.

[11] Ward, J.M., Janssen, G.R., Kieser, T., Bibb, M.J., Buttner, M.J. and Bibb, M.J. (1986) Mol. Gen. Genet. 203, 468–478.

[12] Boyer, H.W. and Roulland-Dussoix, D. (1969) J. Mol. Biol. 41, 459–472.

[13] Jessee, J. and Bloom, F. (1988) BRL Focus 10, N°4, 69–70.

[13] 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.

[15] Bolivar, F. (1978) Gene 4, 121–136.

[16] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning. A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

[17] Dower, W.J., Miller, F.J. and Ragsdale, C.W. (1988) Nucl. Acid Res. 16, 6127–6145.

[18] Woods, D. (1984) Focus 6, 1–3.

[19] Wallace, R.B., Johnson, N.J., Hirose, T., Miyake, M., Kawashima, E.H. and Itakura, K. (1981) Nucleic Acids Res. 9, 879–894.

[20] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci.U.S.A. 74, 5463–5467.

[21] Staden, R. (1984) Nucl. Acids Res. 12, 551–567.

[22] Fickett, J.W. (1982) Nucl. Acids Res. 10, 5303–5318.

[23] Dehottay, P., Dusart, J., De Meester, F., Joris, B., Van Beeumen, J., Erpicum, T., Frère, J.M. and Ghuysen, J.M. (1987) Eur. J. Biochem. 166, 345–350.

[24] Erpicum, T., Granier, B., Delcour, M., Lenzini, V., Nguyen-Distèche, M., Dusart, J. and Frère, J.M. (1990) Biotech. Bioeng. 35, in press.

[25] Frère, J.M., Leyh-Bouille, M., Ghuysen, J.M., Nieto, M. and Perkins, H.R. (1976) Methods Enzymol. 45B, 610–636.

[26] 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) Nucl. Acids Res. 16, 1563–1576.

[27] Blumenthal, R.M. (1986) Trends Biotechnol. 4, 302.

[28] Woodcock, D.M., Crowther, P.J., Diver, W.P., Graham, M., Bateman, C., Baker, D.J. and Smith, S.S. (1988) Nucl. Acids Res. 16, 4465.

[29] Hopwood, D.A., Bibb, M.J., Chater, K.F., Janssen, G.R., Malpartida, F. and Smith, C.P. (1986) in Regulation of Gene Expression. 25 Years on (Booth, E.R. and Higgins, C.F., eds.), pp. 251–276, Cambridge University Press, Cambridge, UK.

[30] Bibb, Me.J., Bibb, Ma.J., Ward, J.M. and Cohen, S.N. (1985) Mol. Gen. Genet. 199, 26–36.

[31] Wéry, J.P. (1987). Ph.D. Thesis, University of Liège Belgium.

[32] Dehottay, P., Dusart, J., Duez, C., Lenzini, M.V., Martial, J.A., Frère, J.M., Ghuysen, J.M. and Kieser, T. (1986). Gene 42, 31–36.

[33] Duez, C., Frère, J.M., Geurts, F., Ghuysen, J.M., Dierickx, L. and Delcambe, L. (1978) Biochem. J. 175, 793–800.