



Gene 198 (1997) 149-157

Molecular characterization of the α-amylase genes of *Lactobacillus plantarum* A6 and *Lactobacillus amylovorus* reveals an unusual 3' end structure with direct tandem repeats and suggests a common evolutionary origin

Eric Giraud ^{a,*}, Gerard Cuny ^b

 ^a Laboratoire de Biotechnologie, ORSTOM, 911 Avenue Agropolis, Campus de Baillarguet, B.P. 5035, 34032 Montpellier Cedex 1, France
^b Laboratoire d'Epidémiologie des Maladies à Vecteurs, ORSTOM, 911 Avenue Agropolis, B.P. 5045, 34032 Montpellier Cedex 1, France

Received 29 January 1997; accepted 6 May 1997; Received by M. Salas

Abstract

The α -amylase gene (amy*A*) of Lactobacillus plantarum A6 was isolated from the genome by polymerase chain reaction with degenerated oligonucleotides, synthesized according to the tryptic peptide amino acid sequences of the purified enzyme. Nucleic acid sequence analysis revealed one open reading frame of 2739 bp encoding a 913 amino acid protein. The amylase appears to be divided into two equal parts. The N-terminal part has the typical characteristics of the well-known α -amylase family (65% identity with the α -amylase of Bacillus subtilis and 97% identity with the partial sequence available for the α -amylase of Lactobacillus amylovorus). The C-terminal part displays a fairly unusual structure. It consists of four direct tandem repeated sequences of 104 amino acids sharing 100% similarity. The complete nucleotide sequence of the α -amylase gene of L. amylovorus was also determined. An open reading frame of 2862 bp encoding a 954 amino acid protein was identified. Perfect homology between the two amyA genes was observed in the N-terminal region. The C-terminal part of L. amylovorus α -amylase also included tandem repeat units but striking differences were observed: (i) the addition of one repeat unit; (ii) a shorter, 91 amino acid repetition unit. These structural homologies suggest that both genes have a common ancestor and may have evolved independently by duplication with subsequent recombination and mutation. © 1997 Elsevier Science B.V.

Keywords: Lactic acid bacteria; Gene organization; Repeated sequences; Evolutionary origin

1. Introduction

Lactobacillus plantarum is recommended as starter culture for fermentation controls of silage, cabbage, cucumber, olive and cassava (Vaughn, 1985; Seale, 1986; Daeschel and Fleming, 1987). Silage sometimes contains too little fermentable sugar to ensure the rapid production of a stabilizing amount of lactic acid. To overcome this difficulty, attempts have been made in recent years to produce a recombined strain of *L. plantarum* releasing a high amylolytic activity by the introduction of the α -amylase gene of *Bacillus amyloliquefaciens* (Jones and Warner, 1990) or *Bacillus stearothermophilus*

* Corresponding author. Tel.: + 33 4 67615800, ext. 4106; Fax: +1 33 4 67593732; e-mail: giraud@melusine.mpl.orstom.fr (Scheirlinck et al., 1989). Nevertheless, the amount of α -amylase excreted by such recombinants was negligible. Recently, Fitzsimons et al. (1994) developed an amylolytic *L. plantarum* strain by the insertion in the chromosome of an active fragment of the gene encoding α -amylase production in *L. amylovorus*. The recombined *L. plantarum* strain was able to achieve amylolytic activity at a level similar to that of *L. amylovorus*. However, the use of such a genetically engineered strain under natural fermentation conditions may raise legal or ecological problems.

A wild strain of *L. plantarum* (strain A6) was previously isolated from retted cassava and selected for its ability to break down soluble starch (Giraud et al., 1991). The characteristics of this strain are particularly attractive. An amylolytic *L. plantarum* strain was reported for the first time. Moreover, it was found to

^{0378-1119/97/\$17.00 © 1997} Elsevier Science B.V. All rights reserved. *PII* S0378-1119(97)00309-0

synthesize large amounts of α -amylase. It was recently shown that *L. plantarum* A6 can also break down cassava raw starch without any preliminary physicochemical treatment (Giraud et al., 1994). New applications of *L. plantarum* A6 might result from these promising results. It could be used as a starter when traditional fermentation processes of starch-containing food fail to produce sufficient lactic acid.

Molecular study of the amylase gene (*amyA*) of *L. plantarum* A6 was used to follow the development and the dispersion of the microorganism during traditional fermentation processes and to understand the origin of its unusual property.

Many microbial and eukaryotic α -amylase genes have been cloned and characterized (Horii et al., 1987; Korman et al., 1990; Sutliff et al., 1991; Rumbak et al., 1991). However, little work has been performed on the α -amylase genes of lactic acid bacteria. Only the α amylase gene of *Streptococcus bovis* has been completely sequenced (Cotta and Whitehead, 1993), and the nucleotide sequence of the 5' end of the *L. amylovorus* amylase gene has been determined (Fitzsimons et al., 1994). No significant homology has been found between these two genes.

The cloning and sequencing of the α -amylase gene of *L. plantarum* A6 are described here and the unusual structure of the 3' end of this gene is shown. Furthermore, as substantial homology was observed with the 5' end of the *L. amylovorus* amylase gene (>97%), the latter was fully characterized to determine the extent of the homology.

2. Materials and methods

2.1. Bacterial strains and growth media

L. plantarum A6 was isolated from retted cassava (Giraud et al., 1991) and L. amylovorus CIP 102989 was purchased from Institut Pasteur (France). For cloning experiments, *Escherichia coli* XL1-blue was used as a host strain. MRS (Difco Laboratories, Detroit, MI, USA) and LB (GIBCO) media were used for routine culturing of lactobacilli and *E. coli*, respectively.

2.2. Purification of the α -amylase and amino acid sequencing

 α -Amylase of *L. plantarum* A6 was purified as described previously (Giraud et al., 1993). An aliquot of the protein (approx. 300 pmoles) was treated with 0.5 µg trypsin in 0.1 M Tris buffer (pH 8.1) overnight at 37°C. The resulting digest was fractionated by reversephase HPLC on a Brownlee C₈, 2×100 mm column (Perkin Elmer, Foster City, CA, USA) and eluted by a linear acetonitrile gradient in 0.1% trifluoracetic acid. The eluate was monitored by UV densitometry (220 nm). The resulting peptides were further purified using a Brownlee C_{18} column with the same experimental conditions. Selected peptides were submitted to Edman N-terminal sequential analysis in a 470-A protein sequencer connected on-line with a 120-A PTH analyser (Perkin Elmer). Both were operated according to the manufacturers' recommendations.

2.3. Southern hybridization analyses

L: plantarum A6 genomic DNA prepared as described by Gasson and Davies (1980) was digested with restriction endonucleases, separated by agarose gel electrophoresis and transferred to nylon membranes (Hybond-N⁺; Amersham, Amersham, UK) using standard methods (Sambrook et al., 1989). DNA probes were labelled with digoxigenin (DIG)-dUTP using the DIG labelling kit supplied by Boehringer-Mannheim (Mannheim, Germany). Hybridization and detection experiments were performed with a DIG detection kit (Boehringer-Mannheim).

2.4. Amylase gene cloning

Total DNA was digested to completion by HinpI and size-fractionated on 1% agarose gel. The 2.7 kb fragments bearing the α -amylase gene were eluted from the gel and ligated into the *AccI* site of plasmid pGEM-3Zf(+) (Promega, Madison, WI, USA). Transformation of *E. coli* electro-competent cells was performed with a Bio-Rad Gene Pulser unit (Bio-Rad, Hercules, CA, USA). After replica plating, the positive clones containing the *amyA* gene were screened from the gene libraries by colony hybridization under the same conditions as those used for southern hybridization analyses.

2.5. DNA amplification and cloning of PCR fragments

PCR reactions were performed in 50 µl final volume containing: 0.1–0.5 µg of DNA and deoxyribonucleoside triphosphate (200 µM each), primers (0.8 µM each), MgCl₂ (1.5 mM), 1.25 unit *Taq* DNA polymerase (Promega) and the buffer supplied with the enzyme. The amplification conditions were: initial denaturation at 94°C for 5 min, 30 cycles with a 30 s denaturation step at 94°C, 30 s annealing step at 50 or 55°C (according to the T_m of primers used) and an elongation step at 72°C lasting 1–3 min (depending on the length of fragment to be generated). After the final elongation step at 72°C for 7 min, the amplified fragments were purified by a Wizard procedure and ligated into pGEM-T vector (Promega).

2.6. Sequencing of the DNA

Both strands of the inserts of the different clones were sequenced by the dideoxy-chain termination method (Sanger et al., 1977) with an 373 DNA Sequencer (Applied Biosystems, Foster City, CA, USA). Protein homology searches were carried out with the SWISS-PROT (release 34) database and genomic sequences with the GenBank (release 98) database and the FASTA program (Pearson and Lipman, 1988).

2.7. Nucleotide sequence accession numbers

The GenBank accession numbers for *L. plantarum* A6 and *L. amylovorus* α -amylase genes are U62095 and U62096, respectively.

3. Results

ł

3.1. Amplification of a specific probe of α -amylase gene of L. plantarum

Purified α -amylase of *L. plantarum* A6 (Giraud et al., 1993) was subjected to trypsin digestion and the resulting peptides were separated by reversed-phase chromatography. Three peptides were selected for sequencing analysis, and the sequences (Table 1) were compared with those of the protein data bank (SWISS-PROT). Strong homology was found for the three peptides with the sequence of the α -amylase of *Bacillus subtilis* (Yang et al., 1983). The degenerate primers A and B were then designed from peptide 1 (amino acids DWTHGNT) and peptide 2 (amino acids WPNITDN). The marked homology found with the α -amylase of *B. subtilis* led us to define primer C (amino acids EFKEMCA) corresponding to the less degenerated amino acid region of the AMYA of *B. subtilis*.

Primers A–B led to the amplification of a 253-bp fragment and primers B–C to the amplification of a 371-bp fragment from the genomic DNA of *L. plantarum* A6. Their sequences revealed a high identity level (65%) with the α -amylase gene of *B. subtilis*.

3.2. Isolation of the L. plantarum A6 amyA gene

Southern hybridization analyses with the 371-bp DIGlabelled fragment as a probe was performed (Fig. 1). Two bands were observed, except for *PstI* digestion. The variation range (600–700 bp) between the two bands was similar for all the digestions. Since the enzyme sites used in this analysis were not present in the probe (except for *Hin*pI, whose site is located at the 5' end of the probe) and since no hybridization signal was observed with the plasmidic fraction of *L. plantarum* A6 DNA (data not shown), these results suggest that *L. plantarum* A6 contains at least two chromosomal *amy*A gene copies.

A first attempt was made at cloning the entire gene. The 5.2-kb *Hin*dIII fragments were eluted from an agarose gel and ligated into the dephosphorylated *Hin*dIII site of vector pGEM-3Zf(+). No hybridization signal was observed after screening 1200 recombinant clones. These results suggest that the entire amylase gene or another region of the chromosome flanking the amylase gene was lethal in *E. coli*, as has already been observed for *B. subtilis amyA* gene cloning (Yang et al., 1983; Emori et al., 1990).

It was therefore decided to isolate the gene in truncated segments. The 2.7-kb *Hin*pI fragment pool was isolated from the gel, and cloned into the vector pGEM-3Zf(+) linearized with *AccI*. Of the 500 recombinant clones, five hybridized with the 371-bp DIGlabelled fragment, one of these was then sequenced.

The 2.7-kb fragment corresponds to approximately 80% of the DNA that encodes the mature part of the α -amylase protein but the 5' end of the entire gene is lacking. This 5' end was then obtained using a PCR screening plasmid approach. The 5.2-kb *Hin*dIII fragment pool was isolated from a gel and ligated into a pGEM-3Zf(+) vector linearized with *Hin*dIII. A primer deduced from the *Hin*pI fragment sequence and the universal oligonucleotides (M13 forward and M13 reverse) of the vector plasmid were used to amplified a 1 kb fragment corresponding to the lacking region of the *amy*A gene.

Table 1

Peptide sequences of the α -amylase of L. plantarum A6 and their homology with the α -amylase of B. subtilis

Peptide	Sequence	Homology with AMYA of <i>B.subtilis</i>		
No.			(annino acid location)	
1	SISDWTHGNTQISNWSDR		SISD WTHGNTQISN WSDR II:: IIIIIII:I IIII	
		151	ISNEVKSIPN WIHGNTQIKN WSDRWDVTQN	181
2	HIELPSQYDGSYGSNFWPNITDNG		HIELPTQYDG SYGSNFWPNI TDNG IIIII: II IIII:IIIII I:::	
		221	HIELPDDG SYGSQFWPNI TNTSAEFQYGEI	251
3	LGWAVVASR		LG WAVVASR	
0		321	WMSDDDIRLG WAVIASRSGS 341	



Fig. 1. Southern hybridization analyses of the α -amylase genes of *L. plantarum* A6 and *L. amylovorus*. Chromosomal DNA of *L. plantarum* A6 and *L. amylovorus* digested with *Hin*dIII, *PsI*, *Hin*pI, *AccI*, *Eco*RV and *DraI* was subjected to electrophoresis on 1% agarose gel and blotted on nylon membranes. The 371-bp DIG-labelled DNA (fragment B-C) was used as a probe.

3.3. Nucleotide sequence of L. plantarum A6 amyA gene

An open reading frame (amyA) of 2739 bp encoding a 913 amino acid protein with a molecular mass of 99 544 Da was identified. Fig. 2 shows that the open reading frame starts with a GTG codon at nucleotide position 363, which is preceded at a canonical distance (9 nucleotides) by a putative ribosome-binding site (AAAGGGGG) complementary to the 3' end of the *L. plantarum* 16S rRNA (UC<u>UUUCCUCCAC</u>; complementary bases underlined) (Woese et al., 1992, unpublished).

The first 108 nucleotides of the structural gene encode a characteristic Gram-positive signal peptide of 36 amino acids (von Heijne, 1985). The entire protein has a predicted theoretical isoelectric point (pI) of 4.63. The amylase gene appears to be divided into two main regions, joined at the *Bam*HI site. The first 5' half, corresponding to the first 480 amino acids, revealed typical features of the well-characterized α -amylase family (Rogers, 1985). The greatest similarity was found with a limited number of α -amylases: (i) 96.8% identity in 155 amino acids overlaps with the partial sequence of α -amylase from *L. amylovorus* (Fitzsimons et al., 1994); (ii) 64.5% identity in 462 amino acids overlaps with the α -amylase of *B. subtilis* (Yang et al., 1983); (iii) 27.2% identity in 430 amino acids overlaps with the α -amylase of *Butyrivibrio fibrisolvens* (Rumbak et al., 1991). Y

5

The 3' half displays an unusual structure consisting of tandem repeated sequences; it consists of four perfectly homologous 104 amino acid repeats, each of them beginning with a microsatellite-like structure (agc agc agc agt agt agt aca aca aca). This microsatellite-like structure has been well documented in eukaryotic genomes (Tautz and Renz, 1984; Bruford and Wayne, 1993), but, to the best of our knowledge, is extremely rare in prokaryotic genomes. Furthermore, a perfectly repeated octanucleotide sequence caagcagc is found at the borders of each repeat, suggesting an insertion event followed by duplication in an ancestral α -amylase gene.

This repeated region was used as a probe in Southern hybridization analyses. The same pattern as for the 371-bp probe (data not shown) was observed, indicating

-360	aagettecacatectggeggate	ag					
-335	335 atcgaattgttgccgacaaatatcgctttggtaatgatgactaaaatgttgctatgatcaccaacttttttgacagtagtttatcaaattacttggtcaaaatcaaagatac						
-223	223 taaacagettetaaaaaagecaataaccacacgeetttggegtgattateagettteagttteagttaetaaaaetaataetgaetataaaacagaageaaaaaatttte						
-111	gatttttatgaaaacggtcgcaaagaagttagcaaaaatatatat	aa					
1	GTG AAA AAA AAG AAA AGT TTC TGG CTT GTT TCT TTT TTA GTT ATA GTA GCT AGT GTT TTC TTT ATA TCT TTT GGA TTG AGC M K K K K S F W L V S F L V I V A S V F F I S F G L S	AAT N					
85 29	TAT TCT AAA CAA GTT GCT CAA GCG GAT AGT GAT AGG ACA TCA ACT GAT CAC TCA AGC AAT GAT ACA GCT GAT TCT GTT AGT Y S K Q V A Q A D S D T T S T D H S S N D T A D S V S	GAC D					
169 57	GET GIT AIT TIG CAI GCA IGG IGC IGG ICG IIC AAC ACG AIT AAA AAC AAC IIG AAA CAG AIT CAI GAC GCC GGC IAC ACA G V I L H A W C W S F N T I K N N L K Q I H D A G Y T	GCG A					
253 85	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	CCA P					
337 113	act aan tat agt att ggt akc tat tat tta gga acg gaa gct gaa tit aag tca atg tgc gct gct gct aaa gaa tat aat t k y s i g n y y l g t e a e f k s m c a a a k e y n	ATC I					
421 141	agg atc att gtc gat gca act ctg aat gat aca aca agt gat tat agt gca att tog gat gaa att aaa agt att tog gat R I I V D A T L N D T T S D Y S A I S D E I K S I S D	TGG W					
505 169	aca cat ggt aac aca caa att tog aat tgg agt gat cgt gaa gat gtt act cat att tog ttg tta ggt ttc tat gat tgg tg agt gat gat gt act act act act tog ttg tta ggt ttc tat gat tgg t R	AAT N					
589 197	act caa aat tet caa get cag acg tat tig aag aat cat tig gaa ege tig att tet gae gea get te gee tet ege tet att te gee tet at tet and tet at t	GAT D					
673 225	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	TCT S					
757 253	GAA TTT CAG TAT GGT GAA GTT TTG CAG GAC TCG ATT TCA AAA GAA TCA GAT TAT GCT AAT TAC ATG AGT GTT ACA GCT TCA E F Q Y G E V L Q D S I S K E S D Y A N Y M S V T A S	AAT N					
841 281	TAC GGC AAT ACG AAT CGC AAT GCG TA AAG AAT CGT GAT TTT ACC GCA AGT ACT TTG CAG AAT TTC AAC ATC AGT GTT CCA Y G N T I R N A L K N R D F T A S T L Q N F N I S V P	GCT A					
925 309	TCT AAA TTA GTA ACT TGG GTC GAA TCG CAT GAT AAT TAT GCT AAC GAT GAT CAA GTT TCG ACT TGG ATG AAT AGT AGT GAT S K L V T W V E S H D N Y A N D D Q V S T W M N S S D	ATT I					
1009 337	and the gec teg get get get teg cet teg cet teg cet teg get age get get cet teg get get get get get get get get g	ACT T					
1093 365	cgg tre cet gge agt tea gan att ggt gat get gge gge gge gge agt trg tat tat gan aaa gea get gte gat aat aaa tre r f p g s s e i g d a g s s l y y d k a v v a v n k f	CAT H					
1177 393	ant gca atg gct ggt can tct gan tat att tct ant cca ant ggc ant acc ang att ttt gan ant gga cgt ggc aga an N A M A G Q S E Y I S N P N G N T K I F E N E R G S K	GGG G					
1261 421	GTT GTT TTT GCA AAC GCT TCC GAC AGT TCA TAT AGT TTG AAT GTT AAA ACT AGT TTA GCT GAT GGG ACT TAT GAA AAC AAG V V F A N A S D S S Y S L N V K T S L A D G T Y E N K	GCT A					
1345 449	GGT TCA GAT GAA TTT ACC GTT AAA AAT GGT TAT TTA ACC GGT ACA ATT CAA GGA CGT GAA GTT GTT GTT CTT TAC GGG GAT G S D E F T V K N G Y L T G T I Q G R E V V V L Y G D	CCA P					
1429 477	RCA AGC AGC AGT AGT AGT ACA ACA ACA GAA ACT ANA AAG GTT TAT TTT GAA AAG CCT TCA AGT TGG GGT AGT ACA GTT T S S S S S S S S S S S T T T T T T T	TAT Y					
1513 505	GCC TAT GTT TAT AAT AAA AAT ACG AAT AAA GCT ATA ACT TCA GCT TGG CCT GGC AAA GAA ATG'ACC GCT TTA GGT CAC GAC A Y V Y N K N T N K A I T S A W P G K E M T A L G H D	GAA E					
1597 533	TAT GAA TTG GAT CTC GAC ACT GAA GAT GAA GAT GAC TCT GAT TTA GCT GTT ATC TTT ACC GAT GGG ACA AAT CAA ACA CCA GCA Y E L D L D T D E D D S D L A V I F T D G T N Q T P A	GCT A					
1681 561	AAT GAG GCT GGT TITT ACC TITT ACG GCT GAT GCC ACT TAT GAT CAA AAT GGT GTC GTA ACA ACT TCT GAT TX AGC AGC AGC N E A G F T F T A D A T Y D Q N G V V T T S D E S S S	AGT S					
1764 589	AGT AGT ACA ACA ACA ACA ACT ANA ANG GIT TAT ITT GAA ANG CCT TCA AGT TGG GGT AGT ACA GIT TAT GCC TAT GIT TAT AGT AGT ACA ACA ACA ACT ANA ANG GIT TAT ITT GAA ANG CCT TCA AGT TGG GGT AGT ACA GIT TAT GCC TAT GIT TAT AGT AGT ACA ACA ACA ACT ANA ANG GIT TAT ITT GAA ANG CCT TCA AGT TGG GGT AGT ACA GIT TAT GCC TAT GIT TAT AGT AGT ACA ACA ACA ACT ANA ANG GIT TAT ITT GAA ANG CCT TCA AGT TGG GGT AGT ACA GIT TAT GCC TAT GIT TAT AGT AGT ACA ACA ACT ANA ANG GIT TAT ITT GAA ANG CCT TCA AGT TGG GGT AGT ACA GIT TAT GCC TAT GIT TAT AGT AGT AGT ACT ACA ACT ANA ANG GIT TAT TIT GAA ANG CCT ACA AGT TGG GGT AGT ACA GIT TAT GCC TAT GIT TAT AGT AGT ACT ACA ACT ANA ANG GIT AT TIT GAA ANG CCT ACA AGT TGG GGT AGT ACA GIT TAT GCC TAT GIT TAT AGT AGT ACT ACA ACT ACA ACT ANA ANG GIT TAT TIT GAA ANG CCT ACA AGT TGG GGT AGT ACA GIT TAT GCC TAT GIT TAT AGT AGT ACA ACA ACT ACA ACT ACA ACT ACA AGT AGT ACA AGT ACA AGT ACA AGT ACA AGT AGT AGT AGT AGT AGT AGT AGT AGT AG	AAT N					
1849 617	ARA ART ACG ART ARA GCT ATA ACT TCA GCT TGG CCT GGC ARA GRA ATG ACC GCT TTA GGT CAC GAC GAA TAT GRA TTG GAT K N T N K A I T S A W P G K E M T A L G H D E Y E L D	CTC L					
1933 645	GAC ACT GAA GAA GAA GAC TCT GAT TTA GCT GTT ATC TTT ACC GAT GGG ACA AAT CAA ACA CCA GCA GCT AAT GAG GCT GGT D T D E D D S D L A V I F T D G T N Q T P A A N E A G	TTT F					
2017 673	ACC TTT ACG GCT GAT GCC ACT TAT GAT CAA AAT GGT GTC GTA ACA ACT TCT GAT GCA AGC AGC AGC AGT AGT AGT ACA ACA	ACA T					
2101 701	GAA ACT AAA AAG GTT TAT TAT ATA AAG CCT TCA AGT TGG GGT AGT ACA GTT TAT GCC TAT GTT TAT AAA AAT ACG AAT \mathbb{R}^{+} T K K V Y F E K P S S W G S T V Y A Y V Y N K N T N	AAA K					
2185 729	GCT ATA ACT TCA GCT TGG CCT GGC AAA GAA ATG ACC GCT TTA GGT CAC GAC GAA TAT GAA TTG GAT CTC GAC ACT GAT GAA A I T S A W P G K E M T A L G H D E Y E L D L D T D E	GAT D					
2269 757	GAC TCT GAT TTA GCT GTT ATC TTT ACC GAT GGG ACA AAT CAA ACA CCA GCA GCT AAT GAG GCT GGT TTT ACC TTT ACG GCT D S D L A V I F T D G T N Q T P A A N E A G F T F T A	GAT D					
2353 785	GCC ACT TAT GAT CAA AAT GGT GTC GTA ACA ACT TCT GAT TCA AGC AGC AGC AGT AGT AGT AGA ACA GAA ACA GAA ACA AAA AAG A T Y D Q N G V V T T S D S S S S S T T T T T E T K K	GTT V					
2437 813	TAT TTT GAA AAG CCT TCA AGT TGG GGT AGT ACA GTT TAT GCC TAT GTT TAT AAA AAT ACG AAT AAA GCT ATA ACT TCA Y F E K P S S W G S T V Y A Y V Y N K N T N K A I T S	GCT A					
2521 841	TGG CCT GGC AAA GAA ATG ACC GCT TTA GGT CAC GAC GAA TAT GAA TTG GAT CTC GAC ACT GAT GAA GAT GAC TCT GAT TTA W P G K E M T A L G H D E Y E L D L D T D E D D S D L	GCT A					
2605 869	GTT ATC TTT ACC GAT GGG ACA ANT CAA ACA CCA GCA GCT AAT GAG GCT GGT TTT ACC TTT ACG GCT GAT GCC ACT TAT GAT V I F T D G T N Q T P A A N E A G F T F T A D A T Y D	CAA Q					
2689 897	AAT GGT GTC GTA ACA ACT TCT GAT TCA AGC AGC ACA TCA AGC AGT TCG TAA gccgataccagcagttcatcaagcagctcggctgtaactg N G V V T T S D S S S T S S S *	icta					
2783 cagaaaccagtggtaacactagttcctctagttctagcgttagccaacctactaaacatcaattggcagttaaattttttgtattacgtgataattatgtcttcaaggcaat							
$2895 \qquad {\tt tgctgtcgaaacgtccactagggtataccctaatagtcatttgcacttttacactatactaggagttatagacgtgcattaggactgttaaaaaaaa$							

3007 attcgtaacgcgactctagaggatc

14

¥

V

Fig. 2. The nucleotide sequence and the deduced amino acid sequence of the *amy*A of *L. plantarum* A6. Putative ribosome-binding sites (underlined), start and stop codon (boldface) are indicated. The presumed signal peptidase cleavage site is indicated with an arrow. The beginning of each repeated sequence is highlighted with a triangle. The 5'-end flanking regions of the 'consensus' repeat unit are shaded and the 3'-end flanking regions of the 'consensus' repeat unit are boxed.

E. Giraud, G. Cuny / Gene 198 (1997) 149-157

153

that these repeats are always associated with α -amylase genes and that other similar repeats are not dispersed in the chromosome.

3.4. Nucleotide sequence of the L. amylovorus amyA gene

The strong homology between the partially sequenced α -amylase gene of *L. amylovorus* and the corresponding region in *L. plantarum* A6 led us to perform full characterization of the former.

Southern hybridization analyses of the *L. amylovorus* chromosomal DNA were carried out using the same restriction enzymes and the same probes as for *L. plantarum*. A main band was observed for each digest (Fig. 1) suggesting that the *amy*A gene is unique in *L. amylovorus*. It is interesting to note that four out of six digests yield very similar fragment sizes in the two different species. Thus, it was decided to use the primers defined for the sequencing of the *amy*A gene of *L. plantarum* A6 to isolate and sequence the *amy*A gene of *L. amylovorus*.

A 2862 bp long open reading frame (amyA) encoding a protein of 954 amino acids with a molecular mass of 104 674 Da was identified (Fig. 3). Perfect homology between the two amyA genes can be observed in the first 5' half (up to the *Bam*HI site) corresponding to the active site, with only seven mismatches in 1600 nucleotides.

As for *L. plantarum* A6 *amy*A, the second part shows the presence of tandem repeat units but striking differences are observed (Fig. 4): (i) Five repeat units are present in *L. amylovorus* as opposed to four in *L. plantarum*. (ii) The repetition unit is shorter (273 nucleotides instead of 321). These repeat units encode 91 amino acids and are homologous with each other and show only five mutations with the corresponding part of the *L. plantarum* A6 repeat unit.

The comparison of repeated sequences of the two *amy*A genes led us to define a 'consensus' repeat unit of 273 nucleotides bordered by two regions: (i) a 5' end flanking region of 35 nucleotides containing the octanucleotide insertion site and the microsatellite-like structure; (ii) a 3' end flanking region of 21 nucleotides containing the octanucleotide insertion site. These flanking regions are found at the extremity of each repeat unit of 273 nucleotides for *L. plantarum* A6, but only once at the borders of the tandem repeated region of *L. amylovorus* (Fig. 2, Fig. 3 and Fig. 4). Note the deletion of the two trinucleotides agt in the microsatellite-like structure of *L. amylovorus*.

The particular structure of the *amy*A genes and the comparison of the tandem repeats suggest the insertion of a 321 nucleotides sequence in an ancestral gene of α -amylase, followed by dispersion in the two *Lactobacillus* species with duplications specific to each strain.

4. Discussion

The molecular characterization of the *amy*A genes of *L. plantarum* A6 and *L. amylovorus* has been described. A very high homology between the two genes is observed, raising the question of the correct identification of the A6 strain. We have sequenced a highly variable region corresponding to the first 400 bp of the 16S rRNA of the A6 strain, and have found a perfect similarity with the *L. plantarum* 16S rRNA sequence (Woese et al., 1992, unpublished). This last result confirms without any ambiguity the identification of the A6 strain to the species *L. plantarum*.

Ų

ķ,

The amino acid sequences deduced indicate that both amyA genes encode a large polypeptide with a molecular weight of about 100 kDa. These polypeptides are approximately twice as large as typical microbial α amylases (50-60 kDa) (Vihinen and Mäntsälä, 1989). However, they are smaller than the estimated molecular weights of the corresponding purified proteins. The amylase protein of L. amylovorus has been estimated to be 126 kDa (Jore and DeParasis, 1993) or 150 kDa (Iman et al., 1991). SDS-PAGE analysis of the purified a-amylase of L. plantarum A6 resulted in a well-defined band (50 kDa) and a diffuse band close to 150 kDa (Giraud et al., 1993). The discrepancies in the estimation of the size can been explained either by migration artefacts of the proteins in SDS-PAGE, as has been reported for some amylases (Robyt and Ackerman, 1973; Arakawa et al., 1992), or by glycosylation of the protein, as observed for bacterial amylases such as those of Alicyclobacillus acidocaldarius (Schwermann et al., 1994), or Bacillus brevis (Stefanova and Emanuilova, 1992).

A remarkable feature of the structure of both amyA genes is the presence, at the 3' end, of direct tandem repeat sequences which share 100% similarity with each other. A striking difference is that the repetition number and the size of repeats vary between the two organisms: four repeated sequences of 321 nucleotides for the amyA of L. plantarum A6, five repeated sequences of 273 nucleotides for L. amylovorus. Southern hybridization analysis showed another noticeable difference between the strains. Whereas only one copy of the amyA gene is observed for L. amylovorus, two copies were revealed in L. plantarum A6. For all the digestion, except for PstI, the size differences between the two copies were always in the same range (600-700 bp). It can be suggested that both gene copies observed in L. plantarum A6 differ by only two 321-bp repeat units and that these duplicated any-genes are located within a single large PstI fragment.

We have investigated by PCR experiments the presence of such repeated sequences in other lactic acid bacteria (*Lactococcus lactis* ATCC 11454, *Leuconostoc mesenteroides* ATCC10832, *Pediococcus pentosaceus* -191

-111 gatttttatgaaaacggtcgcaaagaagttagcaaaaatatataatttcttttgaaattgttcacttggccaagctgcagtttcaatattttaataaagggggcagtaa GTG AAA AAG AAA AAG AAA AGT TTC TGG CTT GTT TCT TTT TTA GTT ATA GTA GCT AGT GTT TTC TTT ATA TCT TTT GGA TTT AGC M K K K K S F W L V S F L V I V A S V F F I S F G F S AAT GCT CAA GCG GCT AGT GAT ACG ACA TCA ACT A Q A A S D T T S T GAT CAC TCA AGC AAT GAT ACA GCT GAT TCT GTT AGC D H S S N D T A D S V S 85 CAT TCT AAA 29 H S K GTT V 169 GGT GTT ATT 57 G V I TTG CAT GCA TGG TGG TGG TGG TTC AAC ACG ATT AAA AAC AAC TTG AAA CAG ATT CAT GAC GCC GGC TAC ACA GCG L H A W C W S F N T I K N N L K O I H D A G Y T A 253 GTT CAA ACT TCA CCT GTT AAT GAA GTT AAA GTT 85 V O T S P V N E V K V GGA AAT AGC GGG G N S G TCT S TCA TTA AAT AAC S L N N TGG W TAT TGG Y W CTA L TAT Y CAG Q CCA P 337 ACT AAA TAT AGT ATT GGT AAC TAT TAT TAT GGA ACG GAA GCT GAA TTT AAG TCA ATG TGC GCT GCT GCT AAA GAA TAT AAT ATC 113 T K Y S I G N Y Y L G T E A E F K S M C A A A K E Y N I 421 agg atc att gtt gat gga act ctg aat gat aca aca agt gat tat agt 141 r i i v d a t l n d t t s d y s GCA ATT TCG GAT GAA ATT AAA AGT ATT TCA GAT TAT GAT 505 ACA CAT GGT AAC ACA CAA ATT TCG AAT TGG AGT GAT CGT GAA GAT 169 T H G N T Q I S N W S D R E D TTG TTA TGG AAT GGT TTC τ. τ. G n w 589 ACT CAA AAT TCT CAA GTT CAG ACG TAT TTG AAG AAT CAT TTG GAA CGC TTG ATT TCT GAC GGA GCT TCA GGC TTC CGT 197 T O N S O V O T Y L K N H L E R L I S D G A S G F R TAT GAT CAT ATT GAA CTT CCA AGT CAA TAT GAT GGC AGC TAT H I E L P S O Y D G S Y GGC AGC AAT TTC TGG CCA AAT ATT ACT GAT AAT GGG TCT 673 GCA GCT ACG 757 GAA TTT CAG TAT GGT GAA GTT TTG CAG GAC TCG ATT 253 E F O Y G E V L Q D S I TCA AAA GAA S K E TCA S GAT TAT GCT AAT TAC ATG AGT GTT D Y A N Y M S V ACA GCT TCA AAT 841 TAC GGC AAT ACG ATT CGC AAT GCG TTA AAG AAT CGT GAT TTT ACC GCA AGT ACT TTG CAG AAT TTC AAC ATC AGT GTT CCA GCT 281 Y G N T I R N A L K N R D F T A S T L Q N F N I S V P A 925 TCT AAA TTA GTA ACT TGG GTC GAA TCG CAT GAT AAT 309 S K L V T W V E S H D N GCT AAC A N TAT Y GAT D GAT D CAA GTT Q V TCG S TGG W AAT N AGT S AGT S GAT ATT D I 1009 AAA TTA GGC TGG GCT GTT GTT TGC TCG CGT TCT GGT AGT GTT CGT CGT CTT GAC CGT CCA GTT GAT GGT GGT AAT GGT ACT 337 K L G W A V V A S R S G S V P L F F D R P V D G G N G T 1093 CGG TTC CCT GGC AGT TCA GAA ATT GGT GAT GCT GGC AGC AGT TTG 365 R F P G S S E I G D A G S S L TAT TAT GAT AAA GCA Y D K A GTT GTA GCT GTT AAT AAA TTC 1177 AAT GCA ATG GCT GGT CAA TCT GAA TAT ATT TCT AAT CCA AAT GGC AAT ACC AAG ATT TTT GAA AAT 393 N A M A G O S E Y I S N F N G N T K I F E N GAA CGT GGC AGC AAA GGG 1261 GTT GTT TTT GCA AAC GCT TCC GAC AGT TCA TAT AGT TTG AAT GTT AAA ACT AGT 421 V V F A N A S D S S Y S L N V K T S TTA GCT GAT GGG ACT TAT GAA AAG GCT 1345 GGT TCA GAT GAA TTT ACC GTT AAA AAT GGT TAT TTA ACC GGT ACA ATT CAA GGA CGT GAA GTT GTT 449 G S D E F T V K N G Y L T G T I O G R F V V GTT V CTT TAC GGG GAT CCA 1429 477 PCA AGC AGC AGC AGT ACA ACA ACA GAA ACT AAA AAG GTT K K V GAA AAG CCT E K P TAT TTT Y F TCA AGT S S TGG GGT W G AGT AGA GTT V TAT GCC TAT 1513 GTT TAT AAT AAA AAT ACG AAT AAA GCT ATA ACT TCA GCT TGG CCT GGC AAA AAA ATG ACC GCT TTA GGT AAC GAC AAA TAT GAA 505 V Y N K N T N K A I T S A W P G K K M T A L G N D K Y E TTT ACC GAT GGG ACA AAG CAA F T D G T K Q GAC D TCT S GAT D TTA GCT L A GTT ATC ACA CCA T P GCA GCT A A AAT N GAG 1681 GCT GGT TTT ACC TTT ACG GCT GAT GCC ACT TAT GAT CAA AAT GGT GTC GTA AAA AAG GTT TAT TTT GAA AAG CCT TCA AGT TGG 561 A G F T F T A D A T Y D Q N G V V K K V Y F E K P S S W ATG 1764 GGT AGT AGA GTT TAT GCC TAT GTT TAT AAT AAA AAT ACG AAT AAA GCT ATA ACT TCA GCT 589 G S R V Y A Y V Y N K N T N K A I T S A TGG CCT GGC AAA AAA ACC 1849 TTA GGT AAC GAC AAA TAT GAA TTG GAT CTC GAC ACT GAT GAA GAT GAC 617 L G N D K Y E L D L D T D E D D TCT GAT TTA GCT GTT ATC TTT ACC GAT GGG ACA AAG 'n τ. GCT A GAT D GCC ACT A T TAT GAT CAA Y D Q 1933 CAA ACA CCA GCA GCT AAT GAG GCT GGT TTT ACC 645 O T P A A N E A G F T TTT ACG F T AAT N GGT G GTC GTA AAA V V K AAG GTT TAT 2017 TTT GAA AAG CCT TCA AGT TGG GGT AGT AGA GTT TAT GCC TAT GTT TAT AAA AAT ACA AAT AAA GCT ATA ACT TCA GCT TGG 673 F E K P S S W G S R V Y A Y V Y N K N T N K A I T S A W 2101 CCT GGC AAA AAA ATG ACC GCT TTA GGT AAC GAC AAA TAT GAA TTG GAT CTC GAC ACT GAT GAA GAT GAC ACT CTT 701 P G K K M T A L G N D K Y E L D L D T D E D D S GAT D TTA L GCT GTT 2185 ATC TTT ACC GAT GGG ACA AAG CAA ACA CCA GCA GCT GAT GAG GCT GGT TTT ACC TTT ACG GCT GAT GCC ACT TAT GAT CAA AAT 729 I F T D G T K O T P A A N E A G F T F T A D A T Y D O N 2269 GGT GTC GTA 757 G V V GAA AAG E K CCT P TCA AGT S S TGG W GGT G AGT S AGA R GTT TAT Y GCC A TAT GTT Y V TAT AAT AAA K AAT N ACG AAT AAA AAG K K GTT TAT 2353 AAA GCT ATA ACT TCA GCT TGG 785 K A I T S A W CCT GGC AAA AAA ATG ACC GCT TTA GGT AAC GAC AAA TAT GAA TTG GAT CTC GAC ACT P G K K M T A L G N D K Y E L D L D T GAT GAA GAG GCT GGT 2437 GAT GAC TCT GAT TTA GCT GTT ATC 813 D D S D L A V I TTT ACC GAT F T D GGG G ACA AAG CAA T K Q ACA CCA GCA GCT AAT T P A A N mm ACC ACG GCT GAA AAG CCT TCA AGT E K P S S TGG GGT AGT AGA GCC 2521 GAT GCC ACT TAT GAT CAA AAT GGT GTC GTA AAA 841 D A T Y D Q N G V AV K AAG GTI V TAT Y TTT F GTT V TAT к 2605 GTT TAT AAT AAA AAT ACG AAT AAA 869 V Y N K N T N K GCT ATA ACT A I T TCA S GCT TGG W CCT P GGC G AAA AAA K K ATG ACC GCT TTA GGT AAC N GAC TAT GAA 2689 TTG GAT CTC GAC ACT GAT GAA GAT GAC TCT GAT GCA GCT A A GAG TTA GCT GTT ATC L A V I AAT 897 τ. D D D D s D 2773 GCT GGT TTT ACC TTT ACG GCT 925 A G F T F T A GAT GCC ACT TAT GAT CAA AAT GGT GTC GTA AGA ACT TCT GAT TCA AGC AGC ACA D A T Y D Q N G V V R T S D S S S T TCA S AGC AAT 2857 TCG **TAN** gccgataccagcagttcatc 953 S *

Fig. 3. The nucleotide sequence and the deduced amino acid sequence of the *amyA* of *L. amylovorus*. Putative ribosome-binding sites (underlined), start and stop codon (boldface) are indicated. The presumed signal peptidase cleavage site is indicated by an arrow. The beginning of each repeated sequence is highlighted with a triangle. The 5'-end flanking region of the 'consensus' repeat unit is shaded and the 3'-end flanking region of the 'consensus' repeat unit is boxed.



Fig. 4. Diagrammatic sketch of the *a*-amylase genes of *L. plantarum* A6 and *L. amylovorus*.

ATCC 43200) and in other *L. plantarum* strains (ATCC 14917, ATCC 10241 and DSM 20174). No amplification products were obtained for all the strains tested; these repeat sequences appear specific to *L. plantarum* A6 and *L. amylovorus*.

Jore and DeParasis (1993), using a series of deletion derivatives in the 3' end of the α -amylase gene of *L. amylovorus*, reported that the N-terminal part (45 kDa) still displays full α -amylase activity. Likewise, Fitzsimons et al. (1994) have demonstrated that the first 590 nucleotides of the *L. amylovorus amyA* gene, including the promoter region, are sufficient to transfer an amylolytic activity to a strain of *L. plantarum* which naturally lacks such activity. The fact that more than 50% of the C-terminal part of the α -amylase is not essential for amylolytic activity raises the question of the function of this region.

The amylases from *B. subtilis*, *L. amylovorus*, and *L.* plantarum A6 have identical physico-chemical properties (same optimal pH and optimal temperature) (Welker and Campbell, 1967; Fogarty, 1983; Giraud et al., 1993; Pompeyo et al., 1993). Nevertheless, only L. plantarum A6 and L. amylovorus possess the special ability to break down raw starch (Giraud et al., 1994; Iman et al., 1991). It is therefore tempting to suggest that the particular structure of the C-terminal half of the α amylases of these two strains is associated with effective substrate binding. This hypothesis is supported by the description of tandem repeat unit located in the C-terminal portion of a family of clostridial and streptococcal ligand-binding proteins (Wren, 1991). This group of proteins includes toxins from *Clostridium difficile* (von Eichel-Streiber et al., 1992), glycosyltransferases from Streptococcus strains (Giffard and Jacques, 1994) and toxin from Streptococcus pneumoniae (Yother and Briles, 1992). This family of ligand-binding proteins appears to display a modular design, with one module providing enzymatic functions and the other module consisting of a repetitive carbohydrate-binding domain located in the C-terminal region (Wren, 1991).

The high degree of homology in the DNA sequences for the two α -amylase genes suggests that these genes are evolutionary related to each other. The structural homologies, particularly in the 3' end part, indicate that both genes probably have a common ancestor and may have evolved independently by duplication and subsequent recombination and mutation, as has been reported for streptococcal glycosyltransferases (Wren, 1991; Giffard and Jacques, 1994) and for Clostridium difficile toxins (von Eichel-Streiber et al., 1992). The significant phylogenetic distance between the two Lactobacillus strains (Collins et al., 1991) led us to consider that the acquisition of this unusual amylolytic property in L. plantarum A6 results from a lateral transfer of this ancestor common gene. Since the repeated sequences are perfectly homologous in one strain and quite well preserved between the two strains, it can be assumed that such an event occurred relatively recently.

The role of these repeated sequences remains to be elucidated. The expression of the entire gene and/or 3'-end truncated forms may confirm their importance in raw starch binding.

Acknowledgement

The authors thank J. Derancourt for expert technical assistance in sequencing the peptides of the α -amylase and G. Giraud for critical reading of the manuscript.

References

Arakawa, T., Hung, L., Narhi, O., 1992. Stability of fungal amylase in sodium dodecylsulfate. J. Protein Chem. 11, 111–117.

- Bruford, M.W., Wayne, R.K., 1993. Microsatellites and their application to population genetical studies. Curr. Opin. Genet. Dev. 3, 939–943.
- Collins, M.D., Rodrigues, U., Ash, C., Aguirre, M., Farrow, J.A.E., Martinez-Murcia, A., Philips, B.A., Williams, A.M., Wallbanks, S., 1991. Phylogenetic analysis of the genus *Lactobacillus* and related lactic acid bacteria as determined by reverse transcriptase sequencing of 16S rRNA. FEMS Microbiol. Lett. 77, 5–12.
- Cotta, M.A., Whitehead, T.R., 1993. Regulation and cloning of the gene encoding amylase activity of the ruminal bacterium *Streptococcus bovis*. Appl. Environ. Microbiol. 59, 189–196.
- Daeschel, M.A., Fleming, H.P., 1987. Achieving pure culture cucumber fermentations: a review, In: Pierce, G. (Ed.), Developments in Industrial Microbiology. Society for Industrial Microbiology, Arlington, VA, pp. 141–148.
- Emori, M., Takagi, M., Maruo, B., Yano, K., Molecular cloning, nucleotide sequencing, and expression of the *Bacillus subtilis* (natto) IAM1212 α-amylase gene, which encodes an α-amylase structurally similar to but enzymatically distinct from that of *B. subtilis* 2633. 1990. J. Bacteriol. 172, 4901–4908.
- Fitzsimons, A., Hols, P., Jore, J., Leer, R.J., O'Connell, M., Delcour, J., 1994. Development of an amylolytic *Lactobacillus plantarum* silage strain expressing the *Lactobacillus amylovorus* α -amylase gene. Appl. Environ. Microbiol. 60, 3529–3535.
- Fogarty, W.M., 1983. Microbial amylases. In: Fogarty, W.M. (Ed.), Microbial Enzymes and Biotechnology. Elsevier, London, pp. 1–92.
- Gasson, M.J., Davies, F.L., 1980. High frequency conjugation associated with *Streptococcus lactis* donor cell aggregation. J. Bacteriol. 143, 1260–1264.
- Giffard, P.M., Jacques, N.A., 1994. Definition of a fundamental repeating unit in streptococcal glucosyltransferase glucan-binding regions and related sequences. J. Dent. Res. 73, 1133-1141.
- Giraud, E., Brauman, A., Keleke, S., Lelong, B., Raimbault, M., 1991. Isolation and physiological study of an amylolytic strain of *Lactobacillus plantarum*. Appl. Microbiol. Biotechnol. 36, 379–383.
- Giraud, E., Champailler, A., Raimbault, M., 1994. Degradation of raw starch by a wild amylolytic strain of *Lactobacillus plantarum*. Appl. Environ. Microbiol. 60, 4319–4323.
- Giraud, E., Gosselin, L., Marin, B., Parada, J.L., Raimbault, M., 1993. Purification and characterization of an extracellular amylase activity from *Lactobacillus plantarum* strain A6. J. Appl. Bacteriol. 75, 276–282.
- Horii, A., Emi, M., Tomita, N., Nishide, T., Ogawa, M., Mori, T., Matsubara, K., 1987. Primary structure of human pancreatic α -amylase gene: its comparison with human salivary α -amylase gene. Gene 60, 57–64.
- Iman, S.H., Burgess-Cassler, A., Cote, G.L., Gordon, S.H., Baker, F.L., 1991. A study of cornstarch granule digestion by an usually high molecular weight α-amylase secreted by *Lactobacillus amylo*vorus. Curr. Microbiol. 22, 365–370.
- Jones, S., Warner, P.J., 1990. Cloning and expression of alpha-amylase from *Bacillus amyloliquefaciens* in a stable plasmid vector in *Lactobacillus plantarum*. Lett. Appl. Microbiol. 11, 214–219.
- Jore, J.P.M., DeParasis, J., 1993. Studies on the α -amylase of *Lactobacillus amylovorus* as a model for heterologous protein secretion by lactobacilli. FEMS Microbiol. Rev. 12, P26
- Korman, D.R., Bayliss, F.T., Barnett, C.C., Carmona, C.L., Kodama, K.H., Royer, T.J., Thompson, S.A., Ward, M., Wilson, L.J., Berka, R.M., 1990. Cloning, characterization, and expression of two αamylase genes from Aspergillus niger var. awamori. Curr. Genet. 17, 203–212.
- Pearson, W.R., Lipman, D.J., 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA 85, 2444–2448.

- Pompeyo, C.C., Gomez, M.S., Gasparian, S., Morlon-Guyot, J., 1993. Comparison of amylolytic properties of *Lactobacillus amylovorus* and of *Lactobacillus amylophilus*. Appl. Microbiol. Biotechnol. 40, 266–269.
- Robyt, J.F., Ackerman, R.J., 1973. Structure and function of amylase. II. Multiple forms of *Bacillus subtilis* α -amylase. Arch. Biochem. Biophys. 155, 445–451.
- Rogers, J.C., 1985. Conserved amino acid sequence domains in a-amylases from plants, mammals, and bacteria. Biochem. Biophys. Res. Commun. 128, 470–476.
- Rumbak, E., Rawlings, D.E., Lindsey, G.G., Woods, D.R., 1991. Cloning, nucleotide sequence, and enzymatic characterization of an α -amylase from the ruminal bacterium *Butyrovibrio fibrisolvens* H17c. J. Bacteriol. 173, 4203–4211.
- Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Sanger, F., Nicklen, S., Coulson, A.R., 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74, 5463–5467.
- Scheirlinck, T., Mahillon, J., Joos, H., Dhaese, P., Michiels, F., 1989. Integration and expression of α -amylase and endoglucanase genes in the *Lactobacillus plantarum* chromosome. Appl. Environ. Microbiol. 55, 2130–2137.
- Schwermann, B., Pfau, K., Liliensiek, B., Schleyer, M., Fischer, T., Bakker, E.P., 1994. Purification, properties and structural aspects of a thermoacidophilic α-amylase from *Alicyclobacillus acidocaldarius* atcc 27009 insight into acidostability of proteins. Eur. J. Biochem. 226, 981–991.
- Seale, D.R., 1986. Bacterial inoculants as silage additives. J. Appl. Bacteriol. Symp. Suppl. 1, 9S-26S.
- Stefanova, M., Emanuilova, E., 1992. Characterisation of a thermostable α-amylase from *Bacillus brevis*. Eur. J. Biochem. 207, 345–349.
- Sutliff, T.D., Huang, N., Litts, J.C., Rodriguez, R.L., 1991. Characterization of an α-amylase multigene cluster in rice. Plant. Mol. Biol. 16, 579-591.
- Tautz, D., Renz, M., 1984. Simple sequences are ubiquitous repetitive components of eukaryotic genomes. Nucleic Acids Res. 12, 4127–4137.
- Vaughn, R.H., 1985. The microbiology of vegetable fermentations, In: Wood, J.B. (Ed.), Microbiology of Fermented Foods. Elsevier, London, Vol. 1, pp. 49–109.
- Vihinen, M., Mäntsälä, P., 1989. Microbial amylolytic enzymes. Crit. Rev. Biochem. Mol. Biol. 24, 329–418.
- von Eichel-Streiber, C., Laufenberg-Feldmann, R., Sartingen, S., Schulze, J., Sauerborn, M., 1992. Comparative sequence analysis of the *Clostridium difficile* toxins A and B. Mol. Gen. Genet. 233, 260-268.
- von Heijne, G., 1985. Signal sequences: the limits of variation. J. Mol. Biol. 184, 99-105.
- Welker, N.E., Campbell, L.L., 1967. Comparison of the α-amylase of Bacillus subtilis and Bacillus amyloliquefaciens. J. Bacteriol. 94, 1131–1135.
- Wren, B.W., 1991. A family of clostridial and streptococcal ligandbinding proteins with conserved C-terminal repeat sequences. Mol. Microbiol. 5, 797–803.
- Yang, M., Galizzi, A., Henner, D., 1983. Nucleotide sequence of the amylase gene from *Bacillus subtilis*. Nucleic Acids Res. 11, 237–249.
- Yother, J., Briles, D.E., 1992. Structural properties and evolutionary relationships of PspA, a surface protein *Streptococcus pneumoniae*, as revealed by sequence analysis. J. Bacteriol. 174, 601-609.

Don Au, (15/12/97)

Reprinted from

AN INTERNATIONAL JOURNAL ON GENES AND GENOMES

GENE

Gene 198 (1997) 149-157

Molecular characterization of the α-amylase genes of *Lactobacillus plantarum* A6 and *Lactobacillus amylovorus* reveals an unusual 3' end structure with direct tandem repeats and suggests a common evolutionary origin

Eric Giraud ^{a,*}, Gerard Cuny ^b

 ^a Laboratoire de Biotechnologie, ÓRSTOM, 911 Avenue Agropolis, Campus de Baillarguet, B.P. 5035, 34032 Montpellier Cedex 1, France
^b Laboratoire d'Epidémiologie des Maladies à Vecteurs, ORSTOM, 911 Avenue Agropolis, B.P. 5045, 34032 Montpellier Cedex 1, France



л



Fonds Documentaire ORSTOM Cote: 合来42484 Ex: 上