

University of Groningen

**Salt-inducible promoter derivable from a lactic acid bacterium, and its use in a lactic acid bacterium for production of a desired protein**

Sanders, Jan Willem; Kok, Jan; Venema, Gerard; Ledebøer, Adrianus Marinus

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*

Publisher's PDF, also known as Version of record

*Publication date:*

1998

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Sanders, J. W., Kok, J., Venema, G., & Ledebøer, A. M. (1998). Salt-inducible promoter derivable from a lactic acid bacterium, and its use in a lactic acid bacterium for production of a desired protein. (Patent No. WO9810080).

**Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

**Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup> :</b> <b>C12N 15/74, 1/21</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 98/10080</b> <b>(43) International Publication Date:</b> 12 March 1998 (12.03.98)
<b>(21) International Application Number:</b> PCT/EP97/04755 <b>(22) International Filing Date:</b> 20 August 1997 (20.08.97) <b>(30) Priority Data:</b> 96202444.4            5 September 1996 (05.09.96)    EP <i>(34) Countries for which the regional or international application was filed:</i> NL et al. 97200744.7            13 March 1997 (13.03.97)    EP <i>(34) Countries for which the regional or international application was filed:</i> NL et al. <b>(71) Applicant (for all designated States except AU BB CA GB GH IE KE LK LS MN MW NZ SD SG SZ TT UG US ZW):</b> UNILEVER N.V. [NL/NL]; Weena 455, NL-3013 AL Rotterdam (NL). <b>(71) Applicant (for AU BB CA GB GH IE KE LK LS MN MW NZ SD SG SZ TT UG ZW only):</b> UNILEVER PLC [GB/GB]; Unilever House, Blackfriars, London EC4P 4BQ (GB). <b>(72) Inventors; and</b> <b>(75) Inventors/Applicants (for US only):</b> SANDERS, Jan, Willem [NL/NL]; Kerklaan 30, NL-9750 AA Haren (NL). KOK, Jan [NL/NL]; Soerabajastraat 57a, NL-9715 LR Groningen (NL). VENEMA, Gerard [NL/NL]; Kerklaan 26, NL-9751		NN Haren (NL). LEDEBOER, Adrianus, Marinus [NL/NL]; B. Lefèvre de Montignyplein 8, NL-3055 ML Rotterdam (NL). <b>(74) Common Representative:</b> UNILEVER N.V.; Patent Division, P.O. Box 137, NL-3130 AC Vlaardingen (NL). <b>(81) Designated States:</b> AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>With international search report.            Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
<b>(54) Title:</b> SALT-INDUCIBLE PROMOTER DERIVABLE FROM A LACTIC ACID BACTERIUM, AND ITS USE IN A LACTIC ACID BACTERIUM FOR PRODUCTION OF A DESIRED PROTEIN		
<b>(57) Abstract</b>  <p>The invention provides a salt-inducible promoter present in SEQ ID NO: 10 and derivable from a lactic acid bacterium in isolation from the coding sequence normally controlled by said promoter in a wild-type lactic acid bacterium, with modifications and important parts thereof. Also provided are a recombinant vector and a transformed lactic acid bacterium comprising such promoter, and the production of a desired protein by such transformed bacterium, whereby the gene encoding said desired protein or a precursor thereof is expressed under control of such promoter. The desired protein can be secreted by the bacterium due to the presence of a signal sequence. The action of the salt-inducible promoter is enhanced at a pH of about 4-4.5 and/or by the presence of glutamic acid. Such process can be used in a fermentation process, in which the desired protein is a lytic protein causing lysis of the cells and release of the cell content. Or the desired protein can be an enzyme involved in flavour formation, e.g. in producing dressings, spreads, sausages and sour dough, or a protein functioning in a cheese production process, such as chymosin or a precursor thereof, or an enzyme involved in cheese flavour formation.</p>		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

<b>AL</b>	Albania	<b>ES</b>	Spain	<b>LS</b>	Lesotho	<b>SI</b>	Slovenia
<b>AM</b>	Armenia	<b>FI</b>	Finland	<b>LT</b>	Lithuania	<b>SK</b>	Slovakia
<b>AT</b>	Austria	<b>FR</b>	France	<b>LU</b>	Luxembourg	<b>SN</b>	Senegal
<b>AU</b>	Australia	<b>GA</b>	Gabon	<b>LV</b>	Latvia	<b>SZ</b>	Swaziland
<b>AZ</b>	Azerbaijan	<b>GB</b>	United Kingdom	<b>MC</b>	Monaco	<b>TD</b>	Chad
<b>BA</b>	Bosnia and Herzegovina	<b>GE</b>	Georgia	<b>MD</b>	Republic of Moldova	<b>TG</b>	Togo
<b>BB</b>	Barbados	<b>GH</b>	Ghana	<b>MG</b>	Madagascar	<b>TJ</b>	Tajikistan
<b>BE</b>	Belgium	<b>GN</b>	Guinea	<b>MK</b>	The former Yugoslav Republic of Macedonia	<b>TM</b>	Turkmenistan
<b>BF</b>	Burkina Faso	<b>GR</b>	Greece	<b>ML</b>	Mali	<b>TR</b>	Turkey
<b>BG</b>	Bulgaria	<b>HU</b>	Hungary	<b>MN</b>	Mongolia	<b>TT</b>	Trinidad and Tobago
<b>BJ</b>	Benin	<b>IE</b>	Ireland	<b>MR</b>	Mauritania	<b>UA</b>	Ukraine
<b>BR</b>	Brazil	<b>IL</b>	Israel	<b>MW</b>	Malawi	<b>UG</b>	Uganda
<b>BY</b>	Belarus	<b>IS</b>	Iceland	<b>MX</b>	Mexico	<b>US</b>	United States of America
<b>CA</b>	Canada	<b>IT</b>	Italy	<b>NE</b>	Niger	<b>UZ</b>	Uzbekistan
<b>CF</b>	Central African Republic	<b>JP</b>	Japan	<b>NL</b>	Netherlands	<b>VN</b>	Viet Nam
<b>CG</b>	Congo	<b>KE</b>	Kenya	<b>NO</b>	Norway	<b>YU</b>	Yugoslavia
<b>CH</b>	Switzerland	<b>KG</b>	Kyrgyzstan	<b>NZ</b>	New Zealand	<b>ZW</b>	Zimbabwe
<b>CI</b>	Côte d'Ivoire	<b>KP</b>	Democratic People's Republic of Korea	<b>PL</b>	Poland		
<b>CM</b>	Cameroon	<b>KR</b>	Republic of Korea	<b>PT</b>	Portugal		
<b>CN</b>	China	<b>KZ</b>	Kazakstan	<b>RO</b>	Romania		
<b>CU</b>	Cuba	<b>LC</b>	Saint Lucia	<b>RU</b>	Russian Federation		
<b>CZ</b>	Czech Republic	<b>LI</b>	Liechtenstein	<b>SD</b>	Sudan		
<b>DE</b>	Germany	<b>LK</b>	Sri Lanka	<b>SE</b>	Sweden		
<b>DK</b>	Denmark	<b>LR</b>	Liberia	<b>SG</b>	Singapore		
<b>EE</b>	Estonia						

**SALT-INDUCIBLE PROMOTER DERIVABLE FROM A LACTIC ACID BACTERIUM, AND ITS USE IN A LACTIC ACID BACTERIUM FOR PRODUCTION OF A DESIRED PROTEIN**

5 The invention relates to a salt-inducible promoter derivable from a lactic acid bacterium.

**Background of the invention and prior art**

It is known that salt-inducible promoters exist in plants and  
10 cyanobacteria; the latter are rather specialised bacteria which can be used for nitrogen-fixation of soil as natural fertilisation and which taxonomically are quite separate from other bacteria.

Although inducible promoter systems are known in Gram-  
15 negative bacteria like *E. coli* and in the Gram-positive bacterium *Bacillus subtilis*, while recently in WO 95/31563 (Quest International B.V. (A. Nauta c.s.); see ref. 36) an inducible promoter system was described for lactic acid bacteria or their phages, no literature was found relating to  
20 salt-inducible promoters active in microorganisms apart from the above mentioned rather specialised cyanobacteria.

Although the expression "salt-initiated induction system" was present in said WO 95/31563 (Quest International B.V. (A. Nauta c.s.); ref. 36), no specific salt-initiated induction  
25 system was disclosed.

The present invention provides for the first time a salt-inducible promoter for lactic acid bacteria and its use in the production of polypeptides by lactic acid bacteria.

**30 Summary of the invention**

The invention provides a salt-inducible promoter derivable from a lactic acid bacterium, in particular a salt-inducible promoter the nucleotide sequence of which is present in SEQ. ID. NO: 10 and in Figure 6. The advantage of a salt-inducible  
35 promoter active in lactic acid bacteria is first of all that salt is a natural food ingredient and therefore can be used as a food-grade inducer in food fermentation processes. For instance, during the salting stage of cheese curd various

processes can be started when the higher salt concentration is used to trigger the formation of various proteins or peptides. Such processes include induced production and **secretion** of compounds which can contribute to the properties of the final cheese, e.g. enzymes and/or peptides that contribute to the formation of flavour compounds, and **lysis** of the lactic acid bacteria whereby peptidolytic enzymes are liberated by which the cheese ripening process is enhanced. Another process in which a salt-inducible promoter can be used advantageously is a process for the production of any protein or secondary metabolite by a food-grade microorganism, especially a lactic acid bacterium, whereby at the end of the culturing at a high cell density the microorganism is induced by a salt to produce the desired protein or secondary metabolite. As will be shown in the Examples, in this specification the term "salt" does only not mean common salt, i.e. sodium chloride, but also includes other halides like alkali metal, earth alkali metal and ammonium halides. In addition to the salt-induction by NaCl, KCl, NH<sub>4</sub>Cl, CaCl<sub>2</sub>, MgCl<sub>2</sub>, NaI, and KI exemplified in Examples 1 and 2 (see also Table 3 below), it is envisaged that also other halides, i.e. bromides and perhaps fluorides, as well as halides with other cations such as substituted ammonium compounds, e.g. tetramethylammonium, or other metallic cations, e.g. Al<sup>3+</sup>, will show the salt-inducing effect.

Still another process which can benefit from the presence of a salt-inducible promoter in a microorganism is the production of secondary metabolites, e.g. flavour or taste ingredients, *in situ* in a fermented product upon the addition of salt, examples of which include dressings and water-containing spreads, as well as sausages and sour dough. The invention also provides a recombinant vector and a transformed lactic acid bacterium each comprising such a salt-inducible promoter.

Brief description of the drawings

Figure 1.

Schematic representation of plasmid pORI13. ori<sup>+</sup>, plus origin of pWV01; Em<sup>r</sup>, erythromycin resistance gene; T, terminator; 5 lacZ, promoterless *E. coli*  $\beta$ -galactosidase gene fused to the ribosome binding site and translational start codon of the lactococcal ORF32 (see Example 1.1 below). Stop codons are indicated by asterisks.

10 Figure 2.

Scheme of the construction of plasmid pLS12 starting from plasmids pBSK+, pKL10, and pMG60 (see Example 1.1.4). The immediate result of the ligation of (1) the multiple cloning site of pBSK+ containing the Km<sup>r</sup> gene in the BamHI site and 15 (2) BamHI-digested pLS11 is a precursor of pLS12 which is not depicted (see 1.1.4 below). After BamHI-digestion of this precursor and self-ligation pLS12 is formed. This is indicated by the dashed arrows.

20 Figure 3.

Scheme of the construction of plasmid pLS13 starting from plasmids pORI28 and pLS12 (see Example 1.1.4).

Figure 4.

25 Scheme of the construction of plasmid pORI13 starting from plasmids pLS13, pMG60 and pORI28 (see Example 1.1.5).

Figure 5.

Deletion analysis NS3 (see Example 1.3). This shows the 30 schematic drawings of the relevant parts of plasmids pNS3 (10 kb insert), pNS3d (2.4 kb insert), pNS3b (1.0 kb insert), pNS3e (0.5 kb insert), and pNS3f (0.4 kb insert).

Figure 6.

35 Nucleotide sequence of the NS3 locus and deduced amino acid sequences of the 3' end of *rnhB*, of *rggL*, and of the 5' end of *orfX* (see Example 1.5). Facing arrows, inverted repeats; rbs, ribosome binding site; -10 and -35 hexanucleotides are

in boldface and underlined; vertical arrow, transcription start point. Stop codons are indicated with asterisks. A number of relevant restriction enzyme sites are shown.

5 Figure 7.

Alignment of the deduced translation products of *rggL* (see SEQ. ID. NO: 11) with *rgg* from *S. gordonii* (see ref. 27; M.C. Sulavik c.s.; 1992 and SEQ. ID. NO: 12) and part of ORF3, found downstream of the *L. lactis* C2 *pip* gene (see ref. 30; 10 B.L. Geller c.s; 1993 and SEQ. ID. NO: 13). Asterisks indicate identical amino acids; conservative amino acids are indicated by dots (see Example 1.4).

Figure 8.

15 Alignment of the translation products of the 3' part of *rnhB* from *L. lactis* (see SEQ. ID. NO: 14) and *Vibrio cholerae* (U30472) (see SEQ. ID. NO: 15) and the complete *rnhB* genes from *E. coli* (see ref. 18; M. Itaya; 1990 and SEQ. ID. NO: 16) and *Haemophilus influenzae* (see ref. 33; R.D. Fleischmann 20 c.s.; 1995 and SEQ. ID. NO: 17). The highest homology was found with *E. coli rnhB* (44% identical plus 13.9% similar amino acids). Asterisks, identical amino acids; periods, similar amino acids (see Example 1.4).

25 Figure 9.

Genomic organization of the NS3 locus. Prgg: *rggL* promoter, IR, inverted repeat, PNaCl: salt-inducible promoter (see Example 1.4).

30 Figure 10.

Sequence analysis NS3. The total *PstI-XbaI* fragment given in Figure 6 is given here schematically with relevant restriction sites and the main features of the DNA. The indicated fragments were subcloned giving plasmids pNS3I - pNS3VI and 35 pNS3i4 (see also Figure 27 and Example 1.4.1). Also primers used for sequencing are given.

Figure 11.

Autoradiogram of a northern blot of total RNA isolated from MG1363 or MGNS3i3 (*rggL*<sup>-</sup>) cells grown in the absence or presence of 0.5 M NaCl. The blot was hybridized with 5 radioactively labeled *XbaI-Sau3A* fragment encoding the 5' end of *orfX*. The sizes of an RNA marker are given in the left margin (see Example 1.6).

Figure 12.

10  $\beta$ -galactosidase activity of NS3 as a function of the NaCl concentration in the culture medium. Samples were taken 7 hours after addition of NaCl to a 1:100 diluted culture (see Example 2).

15 Figure 13.

Nucleotide sequences of the fusion points of truncated *orfX* with either *lytPR* or *acmA* in pNS3PR (Part A) or pNS3AL3 (Part B); see also SEQ. ID. NO: 18 and SEQ. ID. NO: 19, respectively and Examples 3 and 4. The translation of both genes is 20 given below the DNA sequence. Translational stops are indicated with asterisk triplets. The transcription start site is shown by the vertical arrow, while the RBS's are in boldface. The *XbaI* site corresponds to position 1981 in Figure 4. The *EcoRV* and *ScaI* sites were used to make the 25 fusions.

Figure 14.

Renaturing SDS-12.5% PAGE showing cell wall hydrolase activity in cell extracts prepared from samples taken 6 hours 30 after addition of NaCl. Lanes: 1 and 2, LL108(pNS378); 3 and 4, LL108(pNS3PR); 5 and 6, LL108(pNS3AL3); 1, 3 and 5, no NaCl; 2, 4 and 6, plus NaCl. Molecular masses (in kilodaltons) of marker proteins are indicated on the left (see Examples 3 and 4).

35

Figure 15.

Optical density at 600 nm of LL108(pNS378) (squares) and LL108(pNS3PR) (triangles) grown in GM17. NaCl was added to



0.5 M end concentration to part of the culture at A600 = 0.5 (open symbols); see Examples 3 and 5.

N.B. In this specification the indication OD600 (optical density) is used instead of A600 (absorption).

5

Figure 16.

Scheme of the construction of plasmid pNS3PR starting from plasmid pIR1PR and a fragment containing the *rggL* gene and the salt-inducible promoter, which was amplified by PCR using 10 primers NS3-7 and NS3-8 (see Table 2 below) and pNS3 as template (see Example 3.1).

Figure 17.

Optical density at 600 nm of LL108(pNS378) (squares) and 15 LL108(pNS3AL3) (triangles) grown in GM17. NaCl was added to 0.5 M end concentration to part of the culture at A600 = 0.5 (open symbols); see Examples 4 and 5.

Figure 18.

20 Scheme of the construction of plasmid pNS3AL3 starting from plasmids pAL10, pORI28 and the fragment also described in Figure 16 (see Example 4.1).

Figure 19.

25 (A) SDS-12.5% PAA gel stained with Coomassie Brilliant Blue of supernatant samples of *L. lactis* cultures taken 6 hours after addition of NaCl.

(B) Renaturing SDS-12.5% PAGE showing cell wall hydrolase activity. Lanes: 1 and 2, LL108(pNS378); 3 and 4, 30 LL108(pNS3PR); 5 and 6, LL108(pNS3AL3); 1, 3 and 5, no NaCl added; 2, 4 and 6, plus NaCl. Molecular masses (in kilodaltons) of marker proteins are shown on the left. The arrow on the right indicates the position of pre-Acma (see Example 5).

35

Figure 20.

PepXP activity in supernatants of cultures of MG1363acmAA1(pVE6007;pNS3PR), circles, MG1363acmAA1(pVE6007;pNS3AL3), triangles, and 5 MG1363acmAA1(pVE6007;pNS378), squares.

At t=0 (optical density at 600 nm=0.5) NaCl was added to a final concentration of 0.5 M (closed symbols) or 0.1 M (open symbols) to induce the expression of *lytPR*, *acmA*, or *lacZ*, respectively (see Example 5.1).

10

Figure 21.

PepXP activities in the supernatant fractions (black bars) and in the fractions obtained after hypo-osmotic wash of the cells (superimposed grey bars) of MG1363acmAA1(pVE6007) 15 carrying pNS378, pNS3PR, pNS3AL3 or pGKAL1. Cells were induced with 0.025, 0.05, 0.1, 0.25, or 0.5 M of extra NaCl at an optical density of the cultures of 0.5 at 600 nm. As a negative control M17 containing 0.004 M NaCl was used. PepXP levels were determined two days after induction (see Example 20 5.1).

Figure 22.

Gene organization of the NS3 locus. Lollypop, inverted repeat; P, promoter. Relevant restriction enzyme sites are 25 indicated (see Example 6). The black arrows indicate the positions of *lacZ* fusions in the chromosome of the indicated strains (*lacZ* not drawn to scale). The total sequences of the *gadR*, *gadC* and *gadB* genes are given in Figure 29 and SEQ. ID. NO: 21.

30

Figure 23.

Alignment of the deduced translation products of *gadC* from *L. lactis* and *gadC* from *S. flexneri* (Waterman and Small, 1996; ref. 47). Asterisks, identical amino acids; periods, similar 35 amino acids. Putative membrane spanning domains are double underlined. Residues of a putative glutamate binding box are shown in a frame (see Example 6).

Figure 24.

Alignment of GadB from *L. lactis* MG1363, and glutamate decarboxylase from *Synechocystis* sp. strain PCC6803 (Kaneko et al., 1996; ref. 44). Asterisks, identical amino acids; 5 periods, similar amino acids (see Example 6).

Figure 25.

pH and glutamate-dependent expression of *gadC::lacZ*. Strain NS3 was grown in mM17 medium with (A) or without (B) 2%  $\beta$ -glycerophosphate.  $\beta$ -galactosidase activity was followed during growth (fat solid lines) either in the presence of 0.3 M NaCl ( $\blacktriangle$ ) or in the presence of 0.3 M NaCl plus 50 mM glutamic acid ( $\blacksquare$ ). Note that the scale of the left Y-axis in A and B is different. The optical density and pH of the cultures are indicated with dotted lines and thin lines, respectively (see Example 6).

Figure 26.

Schematic representation of the acid resistance mechanism in lactococci, involving the decarboxylation of glutamate to  $\gamma$ -aminobutyrate (see Example 6).

Figure 27.

Schematic representation of plasmids pNS3i4 (see Example 1.4.1), pNS3i5, pNS3i6 and pNS3i7 (for the latter three plasmids see Example 6).

Figure 28.

Schematic representation of plasmid pORI19S (see Example 1.6.1).

Figure 29.

Total nucleotide sequences of the *gadR*, *gadC* and *gadB* genes (see also SEQ. ID. NO: 21).

35

#### Detailed description of the invention

The invention provides a salt-inducible promoter derivable from a lactic acid bacterium, in isolation from the coding

sequence which is normally controlled by said promoter in a wild-type lactic acid bacterium.

Such a salt-inducible promoter was found during the work resulting in the invention and described below in the

5 Examples. This work yielded *inter alia* a DNA fragment originating from chromosomal DNA of *L. lactis* MG1363 in which a salt-inducible promoter appeared to be present.

The nucleotide sequence of this DNA fragment is given in Figure 6 and SEQ. ID. NO: 10.

10

Deletion analysis described in Example 1.3 - 1.3.1 revealed that salt-induction of the *lacZ* gene is still possible, when about 1 kb of genomic DNA isolated from *L. lactis* M1363 is present in front of the *lacZ* gene (which DNA is present in a  
15 *HindIII*-*Sau3A* fragment; see polynucleotide 1476-2426 in Figure 6). With shorter DNA fragments, i.e. containing only 540 bp (see the *EcoRI*-*Sau3A* fragment containing polynucleotide 1882-2422 in Figure 6) or only 440 bp (see the *XbaI*-*Sau3A* fragment containing polynucleotide 1982-2422 in  
20 Figure 6) upstream of the fusion point of the *lacZ* gene with the genomic DNA isolated from *L. lactis* M1363, salt-induction was only weak or not possible at all, respectively.

Therefore, an essential part of this salt-inducible promoter seems to be the polynucleotide 1482-1925 of SEQ. ID. NO: 10,  
25 which is part of the *HindIII* fragment described in Example 1.3.1 below and comprises the 3' part of a gene indicated with *rggL* in Figure 6 up to and including the stop codon. It is believed that this part of the *rggL* gene can be combined with any promoter capable of driving expression of a  
30 structural gene in a lactic acid bacterium, but it is preferred that this part of the *rggL* gene is combined with the real promoter found in the genomic clone, which is the polynucleotide 1926-2000 including an inverted repeat (polynucleotide 1926-1967 of Figure 6) instead of a -35  
35 region and including the -10 region (polynucleotide 1987-1992 of Figure 6), and with the further part of DNA upstream of the ORFX-encoding gene (see polynucleotide 2001-2068 of Figure 6).

Thus a preferred essential part comprises the polynucleotide 1482-2068 of SEQ. ID. NO: 10.

Since the best results were obtained with the about 2.4 kb *Pst*I-*Sau*3A DNA fragment described in Example 1.4, which  
5 contains the full *rgg*L gene preceded by and under control of its own promoter, a salt-inducible promoter comprising the polynucleotide 1-2068 of SEQ. ID. NO: 10 is more preferred, and still more preferred is the use of the full 2.4 kb fragment, thus a salt-inducible promoter, which additionally  
10 comprises part of the ORF X gene together forming polynucleotide 1-2426 of SEQ. ID. NO: 10.

In addition to the various parts of the promoter just mentioned it is envisaged that also modifications thereof can be used according to the invention. Thus the invention also  
15 provides a modification of a salt-inducible promoter according to the invention or an essential part thereof, which comprises a DNA sequence essentially corresponding to a polynucleotide selected from the group consisting of

- (a) polynucleotide 1482-1925 of SEQ. ID. NO: 10 followed  
20 by a promoter functional in a lactic acid bacterium,
- (b) polynucleotide 1482-2068 of SEQ. ID. NO: 10,
- (c) polynucleotide 1-2068 of SEQ. ID. NO: 10, and
- (d) polynucleotide 1-2426 of SEQ. ID. NO: 10.

"Essentially corresponding to a polynucleotide" is understood  
25 as to include genetic variants, such as hybrid sequences containing a salt-inducible promoter or part thereof coupled to other homologous or heterologous DNA sequences including regulatory regions, and sequences containing modifications of the salt-inducible promoter or sequences having mutations,  
30 including mutations which still allow hybridization with the complementary strand of the salt-inducible promoter and genetic variants thereof, while still being capable of exerting the promoter function.

35 It is believed that the *rgg*L gene, or in fact the *rgg*L polypeptide, plays an important role in the effectiveness of the salt-inducible promoter present in the genome of *L. lactis* MG1363, isolated by the present inventors and used to

transform other lactic acid bacteria in order to give them other desirable properties. But it is further believed, that not only the *rggL* gene indicated by polynucleotide 1095-1925 of SEQ. ID. NO: 10 will be functional in this respect, but 5 also other DNA fragments encoding the same polypeptide, and even other DNA fragments encoding a modification of such *rggL* polypeptide still having the same or a similar functionality. Thus an aspect of the present invention is a DNA fragment capable of regulating a salt-inducible promoter active in a 10 lactic acid bacterium, which comprises the polynucleotide 1095-1925 of SEQ. ID. NO: 10, or a modification thereof that (a) encodes the same polypeptide as said polynucleotide 1095-1925, or (b) encodes a modification of such polypeptide still having essentially the same regulating capacity.

15

A further aspect of the invention is a recombinant vector comprising a salt-inducible promoter or an essential part thereof as described above, or a DNA fragment capable of regulating a salt-inducible promoter active in a lactic acid 20 bacterium as described above in combination with a DNA fragment selected from the group consisting of a DNA fragment containing the polynucleotide 1926-2000 of SEQ. ID. NO: 10 and modifications thereof still having essentially the same promoting capacity.

25 The invention also provides a transformed lactic acid bacterium comprising a salt-inducible promoter or an essential part thereof as described above, or a DNA fragment capable of regulating a salt-inducible promoter active in a lactic acid bacterium as described above in combination with 30 a DNA fragment selected from the group consisting of a DNA fragment containing the polynucleotide 1926-2000 of SEQ. ID. NO: 10 and modifications thereof still having essentially the same promoting capacity. The promoter, essential parts thereof and other DNA fragments as described above are 35 preferably present in the chromosome of the lactic acid bacterium, but they can also be present as part of a plasmid that can be maintained during growth of the lactic bacterium.

The lactic acid bacterium containing a salt-inducible promoter according to the invention either can be the natural host from which the salt-inducible promoter is derivable, or it can be a different lactic acid bacterium. If both the  
5 lactic acid bacterium and the salt-inducible promoter applied according to the invention are the same as in the natural situation, the lactic acid bacterium is transformed by incorporating one or more DNA fragments, or the salt-inducible promoter, originating from a lactic acid bacterium,  
10 is used in isolation from the coding sequence which is normally controlled by said promoter in a wild-type lactic acid bacterium.

Further the invention provides a process for the production  
15 of a desired protein by a transformed lactic acid bacterium, whereby the gene encoding said desired protein or a precursor thereof is expressed under control of an inducible promoter, characterised in that the promoter is a salt-inducible promoter or an essential part thereof according to the  
20 invention or a DNA fragment capable of regulating a salt-inducible promoter active in a lactic acid bacterium as described above in combination with a DNA fragment selected from the group consisting of a DNA fragment containing the polynucleotide 1926-2000 of SEQ. ID. NO: 10 and modifications  
25 thereof still having essentially the same promoting capacity. Preferably the transformed lactic acid bacterium is food-grade due to the use of food-grade DNA sequences and/or removal of non-food-grade DNA sequences.

For some embodiments it is desirable that the desired protein  
30 is secreted by the lactic acid bacterium due to the presence of a DNA fragment fused to the gene encoding the desired protein and effecting secretion of the desired protein or a precursor thereof.

A process according to the invention using a salt-inducible  
35 promoter for the expression of a desired gene can be used in a fermentation process, in which the desired protein is a lytic protein causing lysis of the bacterial cells so that the contents of the cells can be released, or in a

TABLE 1.

## Bacterial strains and plasmids

5	Strain or plasmid	Relevant characteristics	Reference or source
<b>Strains</b>			
<i>L. lactis</i>			
	MG1363	Plasmid-free derivative of NCDO712	ref. 5
10	MG1363 <i>acmA</i> Δ1	<i>AcmA</i> <sup>-</sup> derivative of MG1363	ref. 31
	LL108	<i>repA</i> <sup>±</sup> derivative of MG1363, Cm <sup>r</sup>	This work and ref. 40
	LL302	<i>repA</i> <sup>±</sup> derivative of MG1363	This work and ref. 40
15	NS3	NS3:: <i>lacZ</i> , Em <sup>r</sup>	This work
	MGNS3i3	Δ <i>rggL</i> derivative of MG1363, Sp <sup>r</sup>	This work
<i>E. coli</i>			
	MC1000	F <sup>-</sup> , <i>ara</i> D139 ( <i>ara ABC-leu</i> )7679, <i>galU</i> , <i>galK</i> , <i>LacX74</i> , <i>rspL</i> , <i>thi</i>	ref. 7
20	EC1000	<i>repA</i> <sup>±</sup> derivative of MC1000, Km <sup>r</sup> , carrying a single copy of the pWV01 <i>repA</i> gene in the <i>glgB</i> gene	This work and ref. 39
	EC101	<i>repA</i> <sup>±</sup> derivative of JM101, Km <sup>r</sup>	ref. 37
25	NM522	supE, <i>thi</i> , Δ( <i>lac-proAB</i> ), Δ <i>hsd5</i> ( <i>r,m</i> ), [F <sup>'</sup> , <i>proAB</i> , <i>lacI</i> <sup>ΔZ</sup> Δ <i>M15</i> ]	ref. 6
<b>Plasmids</b>			
	pMG60	Em <sup>r</sup> , fusion of promoter 32 with <i>lacZ</i>	ref. 24
30	pORI28	Em <sup>r</sup> , Ori <sup>+</sup> of pWV01, Rep <sup>-</sup>	This work and ref. 40
	pLS28	Em <sup>r</sup> , ' <i>lacZ</i> of pMG60 in pORI28	This work
	pKL10	Em <sup>r</sup> , carrying the pBR322 ori	ref. 17
	pLS10	pKL10, carrying the 5'-end of <i>lacZ</i> from pMG60	This work
35	pLS11	pKL10, carrying the complete <i>E. coli lacZ</i> gene	This work
	pUC7K	Ap <sup>r</sup> , Km <sup>r</sup>	ref. 11
	pBSK+	Amp <sup>r</sup> , α- <i>lacZ</i>	ex Stratagene
40	pLS12Km	Em <sup>r</sup> , Km <sup>r</sup> , pLS11 carrying the multiple cloning site of pBSK+	This work
	pLS12	Em <sup>r</sup> , derivative of pLS12Km	This work
	pLS13	Em <sup>r</sup> , promoterless <i>lacZ</i>	This work
	pORI13	Promoterless <i>lacZ</i> , Em <sup>r</sup> , Ori <sup>+</sup> of pWV01, Rep <sup>-</sup>	This work
45	pVE6007	Cm <sup>r</sup> , pWV01 derivative encoding a temperature sensitive Rep protein	ref. 28



TABLE 1 (continued)

Bacterial strains and plasmids

5	Plasmids (continued)	Relevant characteristics	Reference or source
	pNS3	Em <sup>r</sup> , NS3::lacZ, carries a 10 kb <i>Sau3A</i> chromosomal DNA fragment	This work
10	pNS3b	Em <sup>r</sup> , NS3::lacZ, carries a 1.0 kb <i>HindIII-Sau3A</i> chromosomal DNA fragment	This work
	pNS3d	Em <sup>r</sup> , NS3::lacZ, carries a 2.5 kb <i>PstI-Sau3A</i> chromosomal DNA fragment	This work
	pNS3e	Em <sup>r</sup> , NS3::lacZ, carries a 540 bp <i>EcoRI-Sau3A</i> chromosomal DNA fragment	This work
15	pNS3f	Em <sup>r</sup> , NS3::lacZ, carries a 440 bp <i>XbaI-Sau3A</i> chromosomal DNA fragment	This work
	pNS378	Em <sup>r</sup> , NS3::lacZ, carries a 1280 bp PCR fragment amplified with NS3-7 and NS3-8	This work
20	pUC19	Ap <sup>r</sup> , lacZ'	ref. 8
	pNS3I	Ap <sup>r</sup> , pUC19 with a 401 bp <i>EcoRI-HindIII</i> fragment of pNS3b	This work
	pNS3II	Ap <sup>r</sup> , pUC18 with a 784 bp <i>XbaI</i> fragment of pNS3	This work
25	pNS3III	Ap <sup>r</sup> , pUC18 with a 470 bp <i>XbaI</i> fragment of pNS3	This work
	pNS3IV	Ap <sup>r</sup> , pUC18 with a 604 bp <i>HindIII</i> fragment of pNS3	This work
	pNS3V	Ap <sup>r</sup> , pUC18 with a 743 bp <i>EcoRI-XbaI</i> fragment of pNS3	This work
30	pNS3VI	Ap <sup>r</sup> , pUC19 with a 833 bp <i>HindIII-PstI</i> fragment of pNS3	This work
	pORI19	Em <sup>r</sup> , $\alpha$ -lacZ, Ori <sup>+</sup> of pWV01, Rep <sup>-</sup>	ref. 37
	pNS3i4	Em <sup>r</sup> , pORI19 with a 561 bp <i>Sau3A-XbaI</i> chromosomal DNA fragment	This work
35	pORI19S	Sp <sup>r</sup> , $\alpha$ -lacZ, derivative of pORI19 in which the <i>StuI-XmnI</i> fragment was replaced by an <i>EcoRV-SmaI</i> fragment carrying the Spectinomycin resistance gene	Lab. collection
40	pNS3i3	Sp <sup>r</sup> , internal <i>XbaI-HindIII</i> fragment of <i>rggL</i> in pORI19S	This work
	pIR1PR	Em <sup>r</sup> , phage r1-t <i>lytPR</i> fused to r1-t regulatory cassette	ref. 35
	pNS3PR	Em <sup>r</sup> , NS3::lytPR	This work
	pAL10	Ap <sup>r</sup> , <i>acmA</i> $\Delta$ <i>SacI</i>	ref. 34
45	pAL101	Ap <sup>r</sup> , derivative of pAL10, <i>BglIII</i> sites deleted	This work

TABLE 1 (continued)

Bacterial strains and plasmids

5	Plasmids (continued)	Relevant characteristics	Reference or source
	pAL102	Ap <sup>r</sup> , Em <sup>r</sup> , Ori <sup>+</sup> , derivative of pAL101	This work
10	pNS3AL3S	Em <sup>r</sup> , NS3:: <i>acmA</i> $\Delta$ <i>SacI</i> , derivative of pAL102	This work
	pNS3AL3	Em <sup>r</sup> , NS3:: <i>acmA</i>	This work

fermentation process, in which the desired protein is an  
 15 enzyme involved in the *in situ* production of secondary  
 metabolites as flavour or taste ingredients. Examples of end  
 products include dressings and water-containing spreads, as  
 well as sausages and sour dough.

Still another process which can benefit from the presence of  
 20 a salt-inducible promoter in a microorganism is a fermen-

tation process, in which the desired protein is a protein  
 having a function in a cheese production process, such as  
 chymosin or a precursor thereof, or an enzyme involved in  
 25 cheese flavour formation.

The invention is exemplified by the following Examples 1-6.  
 preceded by a description of the materials and methods that  
 were used.

30

#### MATERIALS AND METHODS

##### **Bacterial strains, plasmids, and growth conditions**

The bacterial strains and plasmids used in the work described  
 are listed in Table 1 above.

35 *L. lactis* was grown at 30°C in M17 medium, with 0.5% glucose;  
 solidified M17 medium contained 1.5% agar. Erythromycin (Em)  
 and chloramphenicol (Cm) were used at final concentrations of  
 5 µg/ml, spectinomycin (Sp) was used at 100 µg/ml. 5-Bromo-4-  
 chloro-3-indolyl- $\beta$ -D-galactopyranoside (X-Gal) was used at a  
 40 final concentration of 0.008%.

*E. coli* was grown in TY broth at 37°C with vigorous agitation or on TY medium supplemented with 1.5% agar. Ampicillin (Ap) and Em were used at 100 µg/ml, Sp at 50 µg/ml.

### 5 Molecular cloning techniques

DNA and RNA techniques were performed essentially as described by Sambrook c.s. (see ref. 16). DNA was introduced by electrotransformation in *E. coli* (see ref. 19; E.R. Zabarovsky & G. Winberg; 1990) and in *L. lactis* (see ref. 15; 10 H. Holo & I.F. Nes; 1989). DNA sequencing was done on double-stranded plasmid DNA by the dideoxy chain-termination method (see ref. 4; F. Sanger c.s.; 1977) and the T7 sequencing kit (Pharmacia LKB Biotechnology AB, Uppsala, Sweden) according to the manufacturer's instructions.

15

Several oligonucleotides prepared and used in the work described are listed in Table 2.

20 TABLE 2. Oligonucleotides used in this work

Name	Nucleotide sequence (5' to 3')	SEQ. ID. NO:
NS3-1	CATTGAGATA ATCAGATAC (19)	1
25 NS3-2	GCAGAGATTG GGAAG (16)	2
NS3-5	GGGGCCCTCT CTTATGTGTT AAATTTTCAG GCGC (34)	3
NS3-6	ATATCGTICA CGTTTTCC (18)	4
NS3-7	GCGATATCCA GACTTCATC	
30 NS3-8	ATACCTCCTT ATATTTATGA TTG (43) GCGAGCTCAG ATCTGAGCGT TGTATAAGCT TTTATGTCTT TC (42)	5
NS3-9	GTTTGACTGA CCAAC (16)	6
NS3-10	CCGCTTCAAT GGTTTG (17)	7
35 NS3-11	CAGTCAAAAC CATTGAAGCG GTTAATGCGA AAAAACCG (38)	8
		9

The oligonucleotides were synthesized with an Applied 40 Biosystems 392A DNA synthesizer (Applied Biosystems Inc. Foster City, CA). DNA sequences were analyzed with the PC/Gene sequence analysis program (IntelliGenetics Inc.,

Geneva, Switzerland). Protein homology searches against the Genbank were carried out using the FASTA program (see ref. 12; W.R. Pearson & D.J. Lipman; 1988). Protein sequence alignments were carried out with the PALIGN program of 5 PC/Gene using the structure genetic matrix or with the CLUSTAL program, both with standard settings.

In this specification the following endonuclease restriction sites are used:

10	<u>giving staggered ends</u>	<u>giving blunt ends</u>
	<i>Apa</i> I            GGGCC↓C	<i>Eco</i> RV        GAT↓ACT
	<i>Bam</i> HI           G↓GATCC	<i>Pvu</i> II        CAG↓CTG
	<i>Bgl</i> III          A↓GATCT	<i>Sca</i> I        AGT↓ACT
	<i>Bss</i> HII         G↓CGCGC	<i>Sma</i> I        CCC↓GGG
15	<i>Cla</i> I            AT↓CGAT	<i>Ssp</i> I        AAT↓ATT
	<i>Eco</i> RI           G↓AATTC	<i>Stu</i> I        AGG↓CCT
	<i>Hind</i> III        A↓AGCTT	<i>Xmn</i> I        GAANN↓NNTTC
	<i>Mlu</i> I            A↓CGCGT	
	<i>Not</i> I            GC↓GGCCGC	
20	<i>Pst</i> I            CTGCA↓G	
	<i>Sac</i> I            GAGCT↓C	
	<i>Sal</i> I            G↓TCGAC	
	<i>Sau</i> 3A          ↓GATC	
	<i>Spe</i> I            A↓CTAGT	
25	<i>Sph</i> I            GCATG↓C	
	<i>Taq</i> I            T↓CGA	
	<i>Xba</i> I            T↓CTAGA	
	<i>Xho</i> I            C↓TCGAG	

### 30 $\beta$ -Galactosidase assays

Cell extracts were prepared from exponentially growing cultures.  $\beta$ -Galactosidase activity was determined as described by Miller (see ref. 2; J.H. Miller; 1972). Protein concentrations in the cell extracts were determined by the 35 method of Bradford (see ref. 3; M.M. Bradford; 1976) with bovine serum albumin as a standard.

### Example 1

#### **Isolation of a salt-inducible promoter**

Several plasmids and transformed strains had to be prepared before a salt-inducible promoter could be isolated. Thus  
5 first some helper strains were constructed containing the *repA* gene of pWV01, i.e. *L. lactis* LL108 and *L. lactis* LL302 (see 1.1.2 below) and *E. coli* EC1000 (see 1.1.1 below). This work is based on the technology described in EP-A1-0 487 159 (see ref. 26; Unilever N.V./PLC (C.J. Leenhouts c.s.); 1992).  
10 Subsequently several plasmids were constructed:

- plasmid pORI28 with plasmids pTC2, pUK24 and pORI24 as intermediates (see 1.1.3 below),
- plasmid pLS13 with plasmids pLS10, pLS11, pLS12 and pORI28 as intermediates (see 1.1.4 below),
- 15 - the promoter screening vector plasmid pORI13 with plasmids pORI28, pLS28 and pLS13 as intermediates using lactococcal helper strains LL108 and LL302 (see 1.1.5 below).

These plasmids carry the ORI<sup>+</sup> of pWV01 but lack the *repA* gene for the replication initiation protein. However, they can be  
20 replicated in the helper strains containing the *repA* gene of pWV01.

#### **1.1 A lactococcal genome bank in pORI13**

To be able to assess expression of chromosomal genes, the  
25 promoter screening vector pORI13 was constructed (Figure 1). To allow only transcriptional fusions one stop codon in each reading frame was present immediately upstream of the *E. coli lacZ* gene. The *lacZ* gene is preceded by lactococcal translation signals derived from ORF32 (see ref. 24; M. van  
30 de Guchte c.s.; 1991 and ref. 10; J.M.B.M. van der Vossen c.s.; 1987). Since plasmid pORI13 carries the ORI<sup>+</sup> of pWV01 but lacks the *repA* gene for the replication initiation protein, it can be used for Campbell-type integrations. Random *Sau3A* fragments obtained from total chromosomal DNA of  
35 *L. lactis* MG1363, ranging in size from 1 to 10 kb, were cloned in the *Bam*HI site of pORI13. With the ligation mixture 48 EC1000 transformants were obtained and checked for their plasmid content. About 73% of these contained pORI13 with a

chromosomal insert while the remainder contained pORI13. The average insert size was 3.3 kb. The restriction enzyme digestion pattern was different for all 48 clones. Plasmid DNA was isolated from 2,460,000 pooled EC1000 transformants 5 (see 1.1.6 below) and used to transform *L. lactis* MG1363(pVE6007). After the temperature-up shift to force integration of the pORI13 derivatives, 9000 Em<sup>r</sup> clones were obtained by plating on selective sucrose GM17 plates containing X-gal and 0.3 M NaCl at 37°C for 16 hours 10 resulting in 195 colonies that stained blue after prolonged incubation at 30°C (see 1.1.7 below).

**1.1.1 Construction of *E. coli* EC1000, being a *repA*<sup>+</sup> derivative of MC1000, Km<sup>r</sup>, and carrying a single copy of the pWV01 *repA* gene in the *glgB* gene** 15

The *repA* gene from pWV01 was introduced onto the chromosome of *E. coli* MC1000 as described for JM101 by Law c.s. (see ref. 37; J. Law c.s.; 1995). pKVB2 (see ref. 37 and ref. 11; J.A.K.W. Kiel c.s.; 1987) is a Tc<sup>r</sup> Km<sup>r</sup> plasmid of 11.7 kb 20 containing the origin of replication of pBR322. It carries the *E. coli* chromosomal *glgB* gene in which an internal 1.2 kb *Bam*HI fragment was replaced by the Km<sup>r</sup> gene from the *Streptococcus faecalis* plasmid pJH1. The *repA* gene from pWV01, driven by the lactococcal consensus promoter P23, was 25 taken as a *Hind*III-*Pvu*II fragment from pUC23rep3 (see ref. 37 and ref. 22; K.J. Leenhouts c.s.; Sept. 1991) and used to replace a 700-bp *Sma*I fragment within *glgB* without interrupting the Km<sup>r</sup> gene. The resultant plasmids pEC1 and pEC2 differ solely in the orientation of *repA* (see ref. 37; 30 J. Law c.s.; 1995).

Plasmids pEC1 and pEC2 were used to transform *E. coli* JM101. Before plating on selective media the transformation mixtures were transferred for 30 generations in the absence of antibiotic and then plated on Km-containing plates. Colonies 35 were tested for glycogen production. Non-glycogen producing colonies were transferred onto plates containing Km and Tc and onto plates containing Km alone. Km<sup>r</sup> Tc<sup>s</sup> colonies were found and were plasmid-free and contained *repA* integrated at

the specific site on the chromosome. Confirmation of the RepA<sup>+</sup> nature of one of the strains (*E. coli* EC1000) was obtained by the successful transformation of this strain with an ORI<sup>+</sup> RepA<sup>-</sup> plasmid. This work will also be described in a publication by Leenhouts c.s. accepted by Mol. Gen. Genet. for publication in 1996 (see ref. 39; K. Leenhouts c.s.; 1996).

### 1.1.2 Construction of RepA<sup>+</sup> *L. lactis* strains LL108 and LL302

Two integration plasmids carrying pWV01 *repA*, pKL15A and pUK30, were constructed.

Plasmid pKL15A is a derivative of the pBR322-based Campbell-type integration plasmid pHV60 (see ref. 14; K.J. Leenhouts; 1989) in which the *repA* gene from plasmid pUC23rep3 (see ref. 22; K.J. Leenhouts; Sept. 1991) was inserted. Integration of pKL15A into the chromosome of *L. lactis* MG1363 by selection for chloramphenicol resistance resulted in *L. lactis* strain LL108 carrying approximately 15 tandem copies of the integration plasmid, as has been described earlier for similar plasmids (see ref. 14; K.J. Leenhouts; 1989).

The replacement-type integration vector pUK30 was obtained by cloning the *repA* fragment of pUC23rep3 in the multiple cloning site of pUK29. The latter plasmid is a derivative of pUK21 (see ref. 23; J. Vieira & J. Messing; 1991) and carries the Em resistance gene of pUC19E (see ref. 17; K.J. Leenhouts; 1990) in the *XhoI* site, the 3'-end of the *L. lactis pepXP* gene (see ref. 20; B. Mayo c.s.; 1991) as a 1.5 kb *XbaI* fragment in the *XbaI* site, and the 5'-end of *pepXP* as

a 1.5 kb *SpeI-MluI* fragment in the *BglIII* and *SphI* sites.

Plasmid pUK30 was used in a two-step gene-replacement strategy (see ref. 21; K.J. Leenhouts c.s.; Aug. 1991) for transforming *L. lactis* MG1363 yielding *L. lactis* strain LL302. This strain contained one copy of pWV01 *repA* inserted in the *pepXP* gene. Both strain LL302 and strain LL108 allow the replication of pWV01-based vectors which lack *repA*. This work will also be described in a publication by Leenhouts c.s. submitted for publication in 1996 (see ref. 40).

### 1.1.3 Construction of pORI28

The Tc<sup>r</sup> of pLS1 (see ref. 9; S.A. Lacks; 1986) was introduced in the *Sma*I site of pMTL25 (see ref. 13; S.P. Chambers c.s.; 1988) to produce pTC2. This Tc resistance gene was isolated 5 from plasmid pTC2 as a 1.6 kb *Bam*HI fragment and plasmid pUK21 (see above) was digested with *Xho*I. Prior to ligation of the two fragments, blunt ends were generated by Klenow enzyme treatment. The ligation resulted in pUK24, in which the Tc<sup>r</sup> gene is flanked by two *Xho*I restriction sites. The 10 1.7 kb *Spe*I fragment of pUK24 carrying the Tc<sup>r</sup> gene was treated with Klenow enzyme to create blunt ends. A 601 bp *Taq*I fragment of pWV01 which carries the plus origin of replication (Ori<sup>+</sup>) but lacks the gene encoding the replication initiation protein (*repA*), was also treated with 15 Klenow enzyme and both fragments were ligated, resulting in pORI24. The *Xho*I fragment of pORI24 carrying the Tc<sup>r</sup> gene was replaced by the Em<sup>r</sup> gene from plasmid pUC19E, located on a 1 kb *Sal*I fragment, which resulted in plasmid pORI28. This work will also be described in a publication by Leenhouts c.s. 20 accepted by Mol. Gen. Genet. for publication in 1996 (see ref. 39; K. Leenhouts c.s.; 1996).

### 1.1.4 Construction of pLS13

The *lacZ* gene of pMG60 (see ref. 24; M. van de Guchte c.s.; 25 1991) was cloned in pKL10 (see ref. 17; K.J. Leenhouts; 1990) in two steps. In the first step the 1250 bp *Ssp*I fragment ex pMG60, containing the 5'-end of *lacZ* gene and the lactococcal RBS of ORF32, was ligated with the Klenow treated *Xba*I site of pKL10. The resulting construct, pLS10, was restricted with 30 *Hind*III, the *Hind*III sticky ends were made blunt and then the fragment was treated with *Cla*I. Subsequently, the resulting fragment was ligated with a *Cla*I-*Xmn*I fragment of pMG60 containing the additional part of the *lacZ* gene, including a transcriptional terminator, resulting in pLS11. 35 In the next step the multiple cloning site (mcs) of pBSK+ had to be inserted upstream of the RBS preceding the *lacZ* gene. To be able to select the mcs of pBSK+ it was marked by cloning a *Bam*HI fragment from pUC7K, coding for a Km<sup>r</sup> gene,



into the *Bam*HI site of the mcs. Subsequently the mcs was excised from this plasmid using *Not*I and *Xho*I. The sticky ends were made blunt-ended and the fragment was ligated into the *Bam*HI site (also made blunt-ended) of pLS11 in front of the *lacZ* gene. Finally the  $Km^r$  gene was deleted using *Bam*HI followed by self ligation. **This resulted in the integration expression vector pLS12** (see Figure 2).

The erythromycin resistance gene of pLS12 was replaced by that from pORI28, including a multiple cloning site by isolating a *Stu*I-*Xba*I fragment from pORI28. This fragment was ligated into pLS12 digested with *Eco*RI, made blunt-ended with Klenow enzyme and subsequently digested with *Xba*I. **The resulting construct was designated pLS13** (see Figure 3).

#### 15 1.1.5 Construction of pORI13

For the construction of a promoter screening vector, the *E. coli lacZ* gene fused to lactococcal translation signals, as present on pMG60 (see ref. 24; M. van de Guchte c.s.; 1991), was used. The *lacZ* gene was placed in the integration vector pORI28 in two steps (see ref. 40; K.J. Leenhouts; submitted for publication 1996).

In the first step, a 2.5 kb *Cla*I-*Xmn*I fragment of pMG60 was ligated in *Xba*I-*Cla*I linearized pORI28 of which the *Xba*I site was made blunt using Klenow. The ligation mixture was digested with *Eco*RI to prevent replication of pMG60 and used to transform the RepA<sup>+</sup> lactococcal helper strain LL108. **The resulting construct, pLS28**, was cut with *Bgl*III and *Bss*HII. The 5' end of the *lacZ* gene was liberated from pLS13 using the same restriction endonucleases. This 1.6 kb fragment was ligated to pLS28 and the ligation mixture was used to transform the RepA<sup>+</sup> *L. lactis* helper strain LL302. **The resulting plasmid was designated pORI13** (see Figure 4).

#### 1.1.6 Construction of a genome bank in pORI13

Total chromosomal DNA of *L. lactis* MG1363 was partially digested with *Sau*3A to obtain fragments ranging in size from 1 to 10 kb which were ligated to *Bam*HI and alkaline phosphatase treated pORI13. The chromosomal fragments were

ligated in the linearized pORI13. This ligation mixture was used to transform the RepA<sup>+</sup> *E. coli* helper strain EC1000 (see 1.1.1 above). Transformants were collected from agar plates by pouring 2 ml of TY broth on each plate, and their plasmid DNA was isolated.

#### 1.1.7 Integration of the bank in the *L. lactis* genome

The plasmid mixture obtained in the previous step was used to transform *L. lactis* MG1363(pVE6007). After electroporation, 10 cells were suspended in recovery medium of 30°C (see ref. 15; H. Holo & I.F. Nes; 1989). After 1.5 hour 5 µg/ml erythromycin was added and incubation at 30°C was prolonged for 0.5 hour. Cells were shifted to 37 °C for 2 hours and then plated on sucrose (0.5 M) GM17 agar containing X-gal, 15 erythromycin and 0.3 M NaCl. Incubation was at 37°C for 24 hours and subsequently at 30°C. The recovery of pORI13 derivatives from the chromosome of selected integrants was done as described by Law c.s. (see ref. 37; J. Law c.s.; 1995).

20

#### 1.2 Identification of a clone expressing *lacZ* in a salt-dependent way

Colonies expressing β-galactosidase in the presence of NaCl were transferred to GM17 agar plates with or without 0.5 M 25 NaCl. Of the 195 selected blue colonies (see 1.1 above) 80 were white on the NaCl-free plates, indicating the absence of β-galactosidase expression. The intensity of the blue colour in the presence of NaCl was similar for all 80 clones. The integrated pORI13 derivatives in 5 of these clones, called 30 NS1-NS5, were rescued and appeared to be identical at the restriction enzyme level. All five plasmids expressed the salt-dependent phenotype. One of these, pNS3 isolated from clone NS3, was selected for further characterization.

#### 35 1.3 Deletion analysis of the genomic region of NS3

Restriction enzyme analysis of pNS3 revealed that about 10 kb of chromosomal DNA had been cloned. A number of restriction enzyme sites was used to map the salt-dependent promoter. A

*Pst*I deletion (pNS3d) showed that NaCl-dependent *lacZ* expression is linked to a 2.4 kb fragment directly upstream of the *Sau*3a site at the original fusion point (see Figure 5 and position 2423 of Figure 6). A larger deletion, leaving 5 1.0 kb upstream of the fusion point (pNS3b), expressed  $\beta$ -galactosidase at a high level in the presence of NaCl but also at a low level in the absence of NaCl, whereas in the original clone  $\beta$ -galactosidase was not expressed in the absence of NaCl. Two other deletion constructs, carrying 540 10 bp (pNS3e) or 440 bp (pNS3f) fragments, showed a low level or no expression with NaCl, respectively. This indicates that the DNA region involved in NaCl-dependent gene expression is located between 1000 and 440 bp upstream of the fusion point (see Figure 5).

15

### 1.3.1 Deletion mapping of NS3

To localize sequences essential for salt-induced gene expression, a number of deletion derivatives of the pORI13 derivative recovered from the chromosome of NS3 (pNS3) were 20 constructed. pNS3 was digested with *Hind*III yielding a fragment carrying *lacZ* and 1.0 kb of upstream lactococcal DNA. This fragment was ligated to a *Hind*III fragment of pORI13 carrying the *Em* resistance marker and the plus origin of replication of pWV01. The resulting construct, pNS3b, was 25 isolated from *E. coli* EC1000. pNS3d was made by self-ligation of *Pst*I digested pNS3 and carries 2.4 kb of chromosomal DNA fused to *lacZ*. The proper construct was obtained in *L. lactis* LL108. Plasmid pNS3e was obtained by using the self-ligation mixture of an *Eco*RI digest of pNS3 to transform *E. coli* 30 EC1000 and carries an about 540-bp chromosomal DNA fragment upstream of *lacZ*. pNS3f was constructed by ligating the *Xba*I fragment from pNS3III (see 1.4.1 below) to the *Xba*I site of pORI13 and transformation of *E. coli* EC1000. pNS3f carries an about 440 bp chromosomal DNA fragment upstream of *lacZ* (see 35 Figure 5).

#### 1.4 Sequence analysis of the NS3 genomic region

The 2.4 kb *Pst*I-*Sau*3A fragment, as present in pNS3d, was sequenced (Figure 6, which also gives the nucleotide sequence of the 561 bp *Sau*3A-*Xba*I fragment described in 1.4.1 below).  
5 This revealed a truncated open reading frame, *orfX*, that had been fused to *lacZ*. *OrfX* started about 360 bp upstream of the *Sau*3A site used for cloning and its putative truncated product shows homology with membrane proteins. The most extensive homology was found with two *E. coli* proteins with  
10 unknown function, YJEM and YGJI (Genbank entries U14003, and U18997, respectively). *OrfX* is preceded by a ribosome binding site (RBS). Upstream of *orfX* an ORF with a weak RBS could encode a protein of 276 amino acid residues (see SEQ. ID. NO: 10 and 11) with a calculated molecular weight of 32990. It  
15 shows homology to the *Streptococcus gordonii rgg* gene product (see SEQ. ID. NO: 12), which regulates expression of the glucosyl transferase gene (see ref. 27; M.C. Sulavik c.s.; 1992) and that of a partially sequenced ORF downstream of the *L. lactis pip* gene (see Figure 7, SEQ.ID. NO: 13 and ref. 30;  
20 B.L. Geller c.s; 1993). In analogy to the *S. gordonii* gene the putative lactococcal regulator gene was named *rggL*. The overall identity between *rggL* and *rgg* is 24.3%, an additional 15.9% of the residues is similar. Directly downstream of *rggL* a 21 bp inverted repeat structure with a  $\Delta G[25^{\circ}\text{C}]$  of -37.8  
25 kcal/mol followed by a stretch of T residues might function as a rho-independent terminator. Upstream of *rggL* a third, incomplete, ORF is present that shows homology to several *rnhB* genes, encoding RNase HII (Figure 8 and SEQ. ID. NO: 14-17). The *rnhB* gene is followed by a rho-independent  
30 terminator-like structure with a  $\Delta G[25^{\circ}\text{C}]$  of -10.6 kcal/mol suggesting that the *rnhB* gene is transcribed independently from *rggL*.

The genomic organization of the NS3 locus is given in Figure 9, which shows that two rare restriction enzyme sites are  
35 present on the sequenced fragment, one for *Not*I (nucleotides 306-313, Figure 6) and one for *Apa*I (nucleotides 2382-2387, Figure 6), separated by only about 2070 bp. From a comparison with the genetic map of *L. lactis* MG1363 (see ref. 32; P. Le

Bourgeois c.s.; 1995) it is clear that the NS3 locus is positioned between the *ldh* gene and the *leu-ilv* gene cluster.

#### 1.4.1 Sequencing of the NS3 locus

5 The 2.5 kb region upstream of the *lacZ*-fusion point in the chromosome of *L. lactis* NS3 was recovered into several subclones and sequenced. A 401 bp *EcoRI-HindIII* fragment from pNS3b was cloned in pUC19 and designated pNS3bI. 784 bp and 470 bp *XbaI* fragments from pNS3 were cloned in the *XbaI* site  
10 of pUC18, resulting in pNS3II and pNS3III, respectively. A 604 bp *HindIII* fragment from pNS3 was cloned in pUC18 and the plasmid was designated pNS3IV and a pUC18 derivative carrying the 743 bp *EcoRI-XbaI* fragment from pNS3 was called pNS3V. A fragment of 745 bp liberated from pNS3 with *HindIII* and *PstI*  
15 and cloned in pUC19 resulted in pNS3VI. *E. coli* NM522 was used as a cloning host for pUC18 or pUC19 constructs. A fragment located at the 3'-end of the chromosomal insert in pNS3 was amplified by inverse-PCR. Self-ligated circularized fragments of a chromosomal *XbaI* digest were used as a  
20 template and NS3-5 and NS3-11 were used as primers. A 561 bp *Sau3A-XbaI* fragment was isolated from the PCR fragment and ligated into *XbaI-BamHI* digested pORI19, resulting in pNS3i4 (see Figure 27). DNA sequencing was done on double-stranded plasmid DNA by the dideoxy chain-termination method (see ref.  
25 4; F. Sanger c.s.; 1977) and the T7 sequencing kit (Pharmacia LKB Biotechnology AB, Uppsala, Sweden) according to the manufacturer's instructions. The double stranded sequence was completed by using synthetic oligonucleotide primers NS3-1, NS3-2, NS3-6, NS3-9, and NS3-10 (see Table 2 above and Figure  
30 10).

#### 1.5 Identification of a salt-inducible promoter structure in NS3

On the basis of the deletion studies and the nucleotide  
35 sequence, an NaCl-dependent start point of transcription was expected in the 100 bp region between the inverted repeat and the start codon of *orfX*. A primer extension was carried out on RNA isolated from LL108(pNS3d) grown in the absence and

presence of 0.5 M NaCl. Using primer NS3-11 transcription was found to start at an A residue 68 bp upstream of the AUG start codon of *orfX* (position 2001, Figure 6). No primer extension product was obtained with RNA of cells grown in the absence of NaCl. 9 bp upstream of the transcription start site a -10 hexanucleotide was identified that differs in only one nucleotide from the consensus sequence (see ref. 25; M. van de Guchte c.s.; 1992). Further upstream no -35 like hexanucleotide could be discerned. Instead, 17 bp upstream of the -10 hexanucleotide the inverted repeat is located, suggesting a role for this structure in gene expression.

#### 1.5.1 RNA analysis

RNA was isolated from exponentially growing *L. lactis* cultures (optical density at 600 nm of 0.5) with or without 0.5 M NaCl as previously described (see ref. 29; M. van Asseldonk; 1993). Northern hybridizations were done at 40°C in a buffer containing 50% formamide, 7% SDS, 2% blocking reagent (Boehringer, Mannheim, Germany), 5 x SSC, 50 mM sodium phosphate pH 7, and 0.1% N-lauryl sarcosine. A 470 bp *Xba*I fragment of *pNS3III* was used as a probe and labeled with [ $\alpha^{32}$ P-dCTP]. A synthetic oligonucleotide (NS3-11) complementary to the mRNA (position +96 to +133 downstream of the transcription start point) was used for primer extension. Twenty five nanogram of primer were added to 5  $\mu$ g of RNA in a reaction mixture containing dCTP, dGTP, dTTP and  $\alpha$ - $^{35}$ S-dATP and cDNA was synthesized using AMV reverse transcriptase (Boehringer, Mannheim, Germany). After 10 minutes incubation at 42°C an excess cold dATP was added and incubation was prolonged for another 10 minutes at 42°C. The product was analyzed on a sequencing gel next to a sequence reaction with the same primer, providing a size marker.

#### 1.6 RggL functions as an activator

The involvement of *rggL* in NaCl-induced gene expression was examined using the *rggL* insertion mutant MGNS3i3 (see 1.6.1 below). RNA was isolated from MG1363 and MGNS3i3 cultured with and without NaCl and transferred to a filter. The 440 bp

*Xba*I-*Sau*3A fragment carrying the 5'-end of *orfX* was used as a probe in Northern hybridization. A strong hybridization signal was seen with RNA from MG1363 grown with NaCl (see Figure 11). Only a very weak signal was detected in the absence of NaCl, confirming the results obtained with the *lacZ* fusions. The same weak signal was seen with RNA from the *rggL* insertion mutant, both from cells grown with and without NaCl. This indicates that *RggL* acts as a positive regulator in NaCl-dependent expression of *orfX*. The size of the NaCl-dependent transcript in MG1363 was estimated to be approximately 3.0 kb.

#### 1.6.1 Construction of an *rggL* insertion mutant

An internal *Xba*I-*Hind*III fragment of the *rggL* gene was cloned in the integration vector pORI19S (see Figure 28), using the RepA<sup>+</sup> *E. coli* helper strain EC101. This plasmid, pNS3i3, was used to disrupt the *rggL* gene in *L. lactis* MG1363, as described before (see ref. 37; J. Law c.s.; 1995). The proper chromosomal location of the integrated plasmid was confirmed by Southern hybridization (not shown) and the strain was named MGNS3i3.

#### Example 2

##### *lacZ* expression from NS3 is halide-ion-dependent

In agar plates a number of compounds were tested to reveal the nature of the inducing condition (see Table 3 below). *lacZ* expression appeared to be independent of the osmolarity of the medium but was strictly linked to the presence of Cl<sup>-</sup> or I<sup>-</sup>. β-galactosidase activity increased with increasing NaCl concentrations in the medium (Figure 12). No induction of β-galactosidase activity was observed by increasing the growth temperature.

#### Example 3

##### Salt-induced expression of phage *ri-t lytPR* genes prevents outgrowth of a lactococcal culture

The transcription start point upstream of *orfX* and *rggL* were amplified as a cassette using primer NS3-7 and NS3-8 and

cloned upstream of the holin and lysin genes (*lytPR*) of the lactococcal temperate bacteriophage r1-t (see ref. 38; D. van Sinderen c.s.; 1996 and ref. 35; Quest International B.V. (A. Nauta c.s.); WO 95/31562). The start codon of *orfX* was fused

5

TABLE 3. *lacZ* expression in strain NS3 in the presence of different compounds

10	Compound <sup>a</sup>	concentration (Mol/l)	$\beta$ -galactosidase activity <sup>b</sup>
	none	-	-
	Maltose	0.5	-
	Mannitol	0.5	-
15	Sorbitol	0.41	-
	Sucrose	0.5	-
	K <sub>2</sub> SO <sub>4</sub>	0.25	-
	MgSO <sub>4</sub>	0.25	-
	NaAcetate	0.25	-
20	Na <sub>2</sub> SO <sub>4</sub>	0.15	-
	Na <sub>2</sub> Succinate	0.25	-
	NH <sub>4</sub> Acetate	0.25	-
	CaCl <sub>2</sub>	0.05	+
	KCl	0.3	+
25	KI	0.3	+
	MgCl <sub>2</sub>	0.125	+
	NaI	0.3	+
	NH <sub>4</sub> Cl	0.25	+
30	<sup>a</sup>	Chemicals were added at the indicated final concentrations to GM17 agar containing Em and X-gal	
	<sup>b</sup>	- : white colonies + : blue colonies	

35

in-frame to ORF5 of r1-t which stops 13 bp upstream of the start codon of *lytP*, resulting in efficient translation of *lytP* and *lytR* (Figure 13A and SEQ. ID. NO: 18). Cultures of the resulting strain, LL108(pNS3PR), were grown in GM17 and

40 induced at an optical density at 600 nm (OD<sub>600</sub>) of 0.5 by the addition of 0.5 M NaCl. Figure 14, lane 4 shows that 6 hours after induction, in addition to wild-type autolysin activity, phage lysin activity was present in the induced cells, and absent in the cells grown without NaCl. The addition of NaCl



led to a stop in the increase of the OD600 of LL108 (pNS3PR) followed by a decrease in optical density 4 hours after induction (Figure 15). The control strain, LL108 (pNS378) expressing *lacZ*, continued to grow after NaCl addition, 5 albeit at a lower rate.

### 3.1 Construction of a transcriptional fusion of the NS3 promoter with the phage *ri-t* holin and lysin genes

A 1280 bp fragment, encoding *rggL* and the salt-inducible 10 promoter, was amplified by PCR using primers NS3-7 and NS3-8 (see Table 2 above) and pNS3 as template. This fragment was digested with *SacI* and *EcoRV* and ligated to pIR1PR (see ref. 35; Quest International B.V. (A. Nauta c.s.); WO 95/31562) linearized with *SacI* and *ScaI*. The ligation mixture was used 15 to transform *L. lactis* LL108 and the resulting plasmid was called pNS3PR (see Figure 16). As a negative control for lytic activity, the same PCR fragment was cloned upstream of *lacZ* in pORI13. The PCR fragment was cut with *BglIII* and *EcoRV* and ligated to *BglIII-SmaI*-digested pORI13. The ligation 20 mixture was used to transform LL108 and the resulting plasmid was labeled pNS378.

#### Example 4

**Salt-induced expression of the lactococcal *acmA* gene results 25 in the stabilization of the OD600 of a culture**

The cassette containing *rggL*, the salt-inducible promoter, and the RBS and start codon of *orfX* was placed upstream of *acmA*, the gene of the major peptidoglycan hydrolase of *Lactococcus lactis* (see ref. 34; Quest International B.V. (G. 30 Buist c.s.); WO 95/31561 and ref. 31; G. Buist c.s.; 1995). Two mutations occurred in the fusion region. An A to G transition in the untranslated leader of the transcript (see nucleotide 64 in Figure 13A + 13B and SEQ. ID. NO: 18 and SEQ. ID. NO: 19) and a deletion of an A residue (see nucleo- 35 tide 99 in Figure 13B and SEQ. ID. NO: 19) resulting in the loss of the *ScaI* site at the fusion point. Translation of the fusion transcript results in the formation of a heptapeptide from the RBS of *orfX* and the translation of *acmA* from its own

RBS (Figure 13B and SEQ. ID. NO: 19). Translation of the small peptide stops 10 bp upstream of the RBS of *acmA*. The growth of the strain carrying the resulting plasmid, pNS3AL3, did not differ from the control strain in GM17 broth (Figure 5 17). However, upon addition of NaCl the cells grew slower and the OD600 stabilized after 4 hours. The OD600 of the control strain increased to a higher level in the presence of salt. More AcmA activity was detected in cells induced with NaCl compared to uninduced cells and to control cells expressing 10  $\beta$ -galactosidase under control of the same induction cassette (Figure 14). Clearly, the expression of *acmA* from pNS3AL3 is induced by NaCl.

#### 4.1 Construction of a transcriptional fusion of the NS3 15 promoter with the lactococcal autolysin gene *acmA*

The *acmA* gene interrupted by a *SacI* fragment on pAL10 was used as a basis (see ref. 34; Quest International B.V. (G. Buist c.s.); WO 95/31561). The *BglIII* sites in pAL10 were deleted by cutting with *BglIII*, filling the overhanging ends 20 with Klenow polymerase, recircularization with T4 ligase and transformation of EC1000. This construct, pAL101, was linearized with *BamHI* and *XbaI* and ligated to pORI28, also linearized with *BamHI* and *XbaI*. The proper construct, pAL102, was isolated from *E. coli* EC1000. A PCR fragment, made by 25 amplification using NS3-7 and NS3-8 on pNS3 as the template and described in Example 3.1 above, was cut with *BglIII* and *ScaI* and ligated to pAL102, linearized with the same restriction enzymes. The ligation mixture was used to transform *L. lactis* LL302 and pNS3AL3S was recovered. This 30 plasmid was digested with *SacI* to delete the insert in *acmA*. After self-ligation the mixture was used to transform LL302. The resulting construct, carrying an intact *acmA* copy fused to the NS3 salt-inducible promoter, was designated pNS3AL3 (see Figure 18).

**Example 5****Intracellular proteins are released upon induction of lysin genes**

The supernatant of the holin-lysin expressing strain (see 5 ref. 35; Quest International B.V. (A. Nauta c.s.); WO 95/31562) contains a significant amount of cytoplasmic proteins 6 hours after NaCl addition, whereas in the supernatant of the same strain grown without induction only secreted proteins are visible (Figure 19.A, lane 3). The 10 addition of NaCl to the control strain already causes the release of a small portion of cytoplasmic proteins (Figure 19.A, lane 2). In the supernatant of cells overexpressing AcmA a much larger quantity of cytoplasmic proteins was observed (Figure 19.A, lane 6) when compared to the amount of 15 intracellular proteins released from LL108(pNS3PR). To detect lytic activity, sodium dodecyl sulphate (12.5%) polyacrylamide gel-electrophoresis was carried out as described (see ref. 1; U.K. Laemmli; 1970) whereby 0.2% autoclaved, lyophilized *Micrococcus lysodeikticus* cells were 20 included in the gel (see ref. 31; G. Buist c.s.; 1995). In Figure 19.B, lane 4 it was shown that LytR activity was detectable in the supernatant of induced LL108(pNS3PR). Using the same method it was also shown that a high level of AcmA activity was present in the supernatant of induced 25 LL108(pNS3AL3). A number of breakdown products of AcmA were visible, as well as the precursor form of AcmA, which is normally present in the cytoplasm (see ref. 34; Quest International B.V. (G. Buist c.s.); WO 95/31561 and ref. 31; G. Buist c.s.; 1995). This latter protein is not processed by 30 the protein export machinery and, therefore, must have been released by cell lysis. AcmA activity in the NaCl-induced samples of LL108(pNS378) and LL108(pNS3PR) was lower than in non-induced samples. This is, most likely, due to the lower OD of the corresponding culture at the moment of sampling 35 (see Figure 15 and Figure 17).

### 5.1 Determination of peptidase activity in the supernatant of *L. lactis* cultures

In order to quantify the release of intracellular proteins PepXP was chosen as an intracellular marker enzyme.

5 Cell lysis was quantified by measuring PepXP release from cells lacking the chromosomal autolysin gene. MG1363acmAA1-(pVE6007) was transformed with either pNS3PR or pNS3AL3. PepXP activity in supernatant samples was determined by following hydrolysis of the synthetic substrate Ala-Pro-p  
10 nitroanilide at 405 nm for 20 minutes at 37°C in a 96-well microtiterplate using a Thermomax microplate reader (Molecular Devices Co., Menlo Park, CA). The presence of protoplasts in induced cultures was tested by washing 1 ml of cells in medium containing a concentration of NaCl which was  
15 identical to that in the original culture. Subsequently, half of the cells were resuspended in 0.5 ml of (hypoosmotic) M17 and the other half in 0.5 ml M17 containing the original NaCl concentration. Samples were incubated for 30 minutes at 37°C and PepXP activity in the supernatants was measured. PepXP  
20 activity in the hypoosmotic extractable cell fraction is given as the activity in the first 0.5 ml corrected for the activity in the latter 0.5 ml.

In order to exclusively study the effect of the salt-induced lysis genes PepXP activity in culture supernatants was fol-  
25 lowed after induction of either lysin gene with 0.5 M NaCl at a culture OD600 of 0.5. During the first hours after induction low levels of PepXP were detected in the culture supernatants (Figure 20). Optimal PepXP levels were obtained 8 hours after induction of cultures of MG1363acmAA1(pVE6007)-  
30 (pNS3AL3) and 30 hours after induction of MG1363acmAA1-(pVE6007)(pNS3PR). The highest level of PepXP activity was present after induction with 0.1 M NaCl (Figure 21). Highest PepXP levels were reached at 30 hours after induction with 0.1 M NaCl of strains carrying the NS3::lytPR fusion and 70  
35 hours after induction of the NS3::acmA strain (Figure 20). The control strain, carrying NS3::lacZ, released low levels of PepXP activity after prolonged incubation. PepXP activity

is stable for at least 40 hours in a cell extract in M17, either with or without NaCl (data not shown). The addition of NaCl to a culture not only induces the activity of the NS3 promoter but also increases the osmolarity of the medium. This probably results in stabilization of osmotically fragile cells formed by lysin activity in the cultures. By incubation in hypoosmotic medium extra PepXP activity could be released from cells carrying the NS3::lytPR fusion induced with 0.1, 0.25, or 0.5 M NaCl, indicative of the formation of fragile cells by the action of LytPR (Figure 21). Induction of *acmA* expression with NaCl resulted in much lower levels of PepXP in the culture medium compared to induction of *lytPR*. However, much more PepXP was extractable with hypoosmotic medium from NS3::*acmA* carrying cells. The sum of PepXP in the supernatant and PepXP extractable from cells is comparable for NS3::*acmA* and NS3::*lytPR* cells induced with 0.25 M NaCl. The amount of PepXP extracted from NS3::*acmA* cells increased with the amount of NaCl used for induction. Only after induction with 0.5 M NaCl a smaller amount of PepXP was extracted. Cells that constitutively overexpress *AcmA* {MG1363(pGKAL1)} released slightly more PepXP in the presence of 0.1 M NaCl as compared to M17 without added NaCl. At higher NaCl concentrations less PepXP was liberated from these cells. Also from this strain, incubated in 0.25 or 0.5 M NaCl, high levels of PepXP were extractable with hypoosmotic M17. Apparently, *AcmA* activity results in a weaker cell wall but not in cell lysis in medium containing NaCl. KCl could replace NaCl as the inducing agent, resulting in only slightly lower levels of released PepXP (data not shown).

#### Example 6

A chloride and low pH-inducible acid resistance mechanism in *Lactococcus lactis*.

In this Example the term *gad* is used to indicate a group of genes involved in glutamate-dependent acid resistance.

Therefore, the genes indicated above with *rggL* and *orfX* have

now been renamed as *gadR* and *gadC*, respectively. The reason for the abbreviation "rggL" was discussed in Example 1.4 above. The reason for the abbreviations "gadB" "gadC" "gadR" is discussed below under the heading RESULTS.

5

#### **Additional MATERIALS AND METHODS**

##### **Bacterial strains, plasmids, and growth conditions.**

The (additional) bacterial strains and plasmids used in this Example are listed in Table 4. *L. lactis* was grown at 30°C in 10 half strength M17 ( $\frac{1}{2}$ M17) broth containing 0.5% glucose and 1.9%  $\beta$ -glycerophosphate; solidified  $\frac{1}{2}$ M17 contained 1.5% agar. Modified M17 (mM17) contained no  $\beta$ -glycerophosphate and no soytone.  $\beta$ -glycerophosphate was added to a final concentration of 2%, where indicated.

15

##### **General DNA techniques.**

Protein homology searches against the Genbank were carried out using the BLAST program (Altschul et al., 1990; ref. 41). Protein sequence alignments were carried out with the PALIGN 20 program of PC/Gene using the structure genetic matrix. Transmembrane segments were predicted using the method of Klein et al. (1985; ref. 45)

##### **Cloning of *gadCB* of *L. lactis***

25 The region downstream of the *lacZ* fusion site in *L. lactis* NS3 was obtained by inverse PCR amplification of a 0.56-kb *Sau3A-XbaI* fragment from *L. lactis* MG1363 chromosomal DNA using primers NS3-5 and NS3-11 (see Table 2 above). The PCR product was cloned in pORI19 using *Escherichia coli* EC101 as 30 a host, resulting in pNS3i4 (see Example 1.4.1 above and Figure 27). In a subsequent inverse PCR the adjoining 0.9-kb *XbaI-Sau3A* chromosomal DNA fragment was amplified using primers NS3-14 (5' GGCAGTCCGTTGCGTCCCACC) and NS3-15 (5' GAGTTATCATTAAGTGCAGGGG) and cloned in pORI19 using *L. lactis* 35 LL108 as the cloning host. This construct was named pNS3i5 (see Figure 27) and used for single cross-over integration, resulting in strain MGNS3i5. Chromosomal DNA of MGNS3i5 was digested with *EcoRI* to clone sequences downstream of the

insertion site of pNS3i5. The chromosomal *EcoRI* fragments were circularized by self-ligation and the ligation mixture was used to transform *L. lactis* LL108. Transformants contained a 4.8-kb plasmid (pNS3i6; see Figure 27) which is pORI19 containing a 2.6-kb *XbaI-EcoRI* fragment from the chromosome (Figure 22).

Table 4. Additional bacterial strains and plasmids for Example 6.

Strain or plasmid	Relevant characteristics	Reference or source
15 Strains		
<i>L. lactis</i>		
LL302	MG1363 derivative, carrying a single <i>repA</i> copy in <i>pepXP</i>	Leenhouts, unpublished, See earlier Examples
MGNS3i3	Sp <sup>r</sup> , derivative of MG1363 with pNS3i3 inserted in <i>gadR</i>	See earlier Examples
20 MGNS3i4	Em <sup>r</sup> , derivative of MG1363 with pNS3i4 inserted in <i>gadC</i>	This work
MGNS3i5	Em <sup>r</sup> , MG1363 derivative carrying a single copy of pNS3i5 on the chromosome	This work
25 MGNS3i7	Em <sup>r</sup> , derivative of MG1363 with pNS3i7 inserted in <i>gadB</i>	This work
MGNS31	Em <sup>r</sup> , MG1363 with <i>lacZ</i> inserted between <i>gadR</i> and P <sup>gad</sup>	This work
30 MGNS32	Em <sup>r</sup> , MG1363 ( <i>gadR::lacZ</i> )	This work
<i>E. coli</i>		
EC101	Km <sup>r</sup> , JM101 derivative, carrying a single copy of the pWV01 <i>repA</i> gene in <i>glgB</i>	Law et al., 1995; ref. 37
35		
Plasmids		
pORI13	Promoterless <i>lacZ</i> , Em <sup>r</sup> , Ori <sup>+</sup> , RepA <sup>+</sup> derivative of pWV01	See earlier Examples
40 pNS3i3	Sp <sup>r</sup> , pORI19S carrying a 240-bp <i>XbaI-HindIII</i> fragment internal to <i>gadR</i>	See earlier Examples
pNS3i5	Em <sup>r</sup> , pORI19 carrying a 0.9-kb <i>XbaI-Sau3A</i> chromosomal DNA fragment encoding the 3'-end of <i>gadC</i> and the 5'-end of <i>gadB</i>	This work
45 pNS3i6	Em <sup>r</sup> , pORI19 carrying a 2.4-kb <i>XbaI-EcoRI</i> chromosomal DNA fragment	This work
50 pNS3i7	Em <sup>r</sup> , pORI19 carrying a 500-bp internal <i>AseI</i> fragment of <i>gadB</i>	This work

#### Construction of insertion mutants.

55 An internal *XbaI-HindIII* fragment of *gadR* was cloned as a blunt fragment in the *SmaI* site of pORI19S (see Figure 28)

using the RepA<sup>+</sup> *E. coli* helper strain EC101. The resulting plasmid, pNS3i3, was used to disrupt *gadR* in *L. lactis* MG1363, as described before (Law et al., 1995; ref. 37). The proper chromosomal location of the integrated plasmid was confirmed by Southern hybridization and the strain was named MGNS3i3. Plasmid pNS3i7 was constructed by cloning a 0.5-kb *AsnI* fragment (internal to *gadB*) as a blunt fragment in the *SmaI* site of pORI19 using *E. coli* EC101 as a host (see Figure 27). Single cross-over integration of pNS3i4 and pNS3i7 in the *L. lactis* MG1363 chromosome resulted in strains MGNS3i4 (*gadC*) and MGNS3i7 (*gadB*), respectively.

#### **β-Galactosidase assays.**

Cell-free extracts were prepared by vigorous shaking of cells in the presence of glass beads (van de Guchte et al., 1991; ref. 24). β-Galactosidase activity was determined as described by Miller (1972; ref. 2). Protein concentrations were determined by the method of Bradford (1976; ref. 3) with bovine serum albumin as the standard.

20

#### **Acid resistance tests.**

Cells of an exponentially growing culture of *L. lactis* in half strength GM17 (½GM17) with or without 0.3 M NaCl was harvested, washed with water and resuspended in an equal volume of MS15 (Cocaign-Bousquet et al., 1995; ref. 42) without glucose and glutamate and containing the same amount of NaCl as the culture. The pH of MS15 was adjusted to 3.5 with either lactic acid or hydrochloric acid prior to cell resuspending. After incubation for 2 hours at 30°C the number of viable cells was determined by plating onto glucose ½M17 plates. The percentage of acid resistant cells was calculated from the number of colony-forming units (cfu) after acid treatment divided by the number of cfu found at the moment of harvest.

35



## RESULTS

### Nucleotide sequence of the genes transcribed by the chloride-dependent promoter.

In the previous Examples the chloride-dependent promoter  
5 ( $P_{gad}$ ) was identified using a random *lacZ* chromosomal  
integration strategy. Here the nucleotide sequence is  
presented of the chromosomal region downstream of the *lacZ*  
integration site in *L. lactis* NS3.

The original *Sau3A* fusion site in the *lacZ* integrant *L.*  
10 *lactis* NS3 is located in an ORF of 503 codons immediately  
downstream of  $P_{gad}$ . This ORF was named *gadC*, as its deduced  
amino acid sequence (Figure 23) is homologous to GadC from  
*Shigella flexneri* (51% identity and 17% similarity, Waterman  
and Small, 1996; ref. 47) and its *E. coli* counterpart XasA  
15 (Hersh et al., 1996; ref. 43). GadC is homologous to a number  
of amino acid antiporters, including the lysine-cadaverine  
antiporter CadB from *E. coli* (Meng and Bennet, 1992; ref.  
46). Lactococcal GadC has a deduced molecular weight of 55369  
and a pI of 9.73 and is highly hydrophobic. The hydrophobic  
20 residues are clustered in 12 domains (Figure 23), whose  
locations coincide with those of the hydrophobic domains in  
*S. flexneri* GadC (as predicted by a number of topology-  
predicting computer programs). This suggests that GadC is an  
integral membrane protein. A conserved domain found in  
25 glutamate transporting proteins is also present in *L. lactis*  
GadC (Figure 23, Waterman and Small, 1996; ref. 47); *gadC* is  
separated by 19-bp from **another ORF of which the deduced  
protein is homologous to glutamate decarboxylases**. Highest  
homology is found with glutamate decarboxylase from  
30 *Synechocystis* sp. (48% identity and 15% similarity, Figure  
24, Kaneko et al., 1996; ref. 44). Both in *E. coli* and *S.*  
*flexneri* *gadB* is linked to the putative glutamate -  $\gamma$ -amino  
butyrate antiporter genes, *xasA* and *gadC*, respectively. **The  
lactococcal gene was, in analogy to *E. coli*, and *S. flexneri***  
35 **named *gadB***. The gene order of *gadB* and *gadC* is inverse in *L.*  
*lactis* compared to these organisms. No possible transcription  
signals could be identified in the 19-bp intergenic region  
between *gadC* and *gadB* which suggests that they form an

operon. Downstream of *gadB* a 16-bp inverted repeat (IR) may function as a rho-independent transcription terminator.

***gadR* is constitutively expressed**

5 Expression of *gadR* was studied in strain MGNS32 carrying a single copy *gadR::lacZ* transcriptional fusion (see Figure 22).  $\beta$ -galactosidase activity in exponentially growing MGNS32 was 3.0 U/mg, independent of the presence of NaCl. In other words, NaCl-dependent expression from  $P_{gad}$  is not regulated by  
10 variations in the level of transcription of *gadR*. Another *lacZ* fusion, located immediately downstream of the 21-bp IR was used to show that transcription of *gadR* is effectively terminated by this IR. No  $\beta$ -galactosidase activity could be detected in strain MGNS31 carrying this fusion (see Figure  
15 22). Transcription of *gadR* starts from a C-residue 116 bp upstream of the AUG start codon (Figure 22). The *gadR* promoter consists of canonical -35 and -10 hexanucleotides separated by 18 bp (Figures 6 and 29).

20 **Expression of *gadCB* is enhanced at low pH and by glutamate**

In *S. flexneri* and *E. coli* *gadC* and *xasA* have been shown to play a role in survival at low pH in the presence of glutamate (Waterman and Small, 1996; ref. 47; Hersh et al., 1996; ref. 43). The expression of lactococcal *gadC* under such  
25 conditions was studied in *L. lactis* NS3. Strain NS3 was previously identified because of its chloride-dependent *lacZ* expression and contains a single copy *gadC::lacZ* fusion on its chromosome (see previous Examples). A modified M17 medium (mM17) was used, lacking the buffer  $\beta$ -glycerophosphate to  
30 modulate the culture pH; mM17 also lacks soytone.  $\beta$ -galactosidase activity in NS3 was still induced in this modified medium in the presence of 0.3 M NaCl and 2%  $\beta$ -glycerophosphate buffer but the induction level was fivefold lower as compared to that in standard  $\frac{1}{2}$ M17 (data not shown). The *lacZ*  
35 expression in mM17 in the presence of NaCl was low in the early stages of exponential growth and increased to an optimum at the onset of the stationary phase (Figure 25). In the absence of buffer, the culture pH decreased to 4.0 to 4.5

in the stationary growth-phase, while cultures containing buffer reached a lowest pH of 5.5. Expression of *lacZ* was increased 10-fold in mM17 containing no buffer. The presence of 50 mM glutamate resulted in an approximately twofold  
 5 additional increase in the expression of *lacZ* both in buffered and in unbuffered mM17 broth (Figure 25). No  $\beta$ -galactosidase activity was detectable in the absence of NaCl whether or not glutamate was present and independent of the culture pH (data not shown). Therefore, expression of *gadCB*  
 10 is optimal at the onset of the stationary growth phase in the presence of NaCl and glutamate, and at low pH.

#### *gadCB* confers acid resistance

15 The ability of *L. lactis* MG1363 to survive acid stress was tested under conditions where *gadCB* is expressed and under conditions where *gadCB* is not expressed. When MG1363 was grown in  $\frac{1}{2}$ GM17 and acid-challenged at pH 3.5 in MS15 for two hours the viability of the cells decreased dramatically  
 20 (Table 5).

Table 5. Acid resistance of *L. lactis* MG1363 under different conditions of challenge

conditions:	% acid resistance at pH 3.5, lactic acid	% acid resistance at pH 3.5, HCl	% acid resistance at pH 6.5
25 -	0.0005	1.12	139
+ 1 mM glutamate	0.0004	0.96	135
+ 0.3 M NaCl	0.0002	1.00	58
+ 0.3 M NaCl + 1 mM glutamate	0.49	16	57

30

Lactic acid was much more deleterious than hydrochloric acid whereas the viability was not affected at pH 6.5. The  
 35 presence of 1 mM glutamate alone did not affect the viability at pH 3.5. In the presence of 0.3 M NaCl during growth,

allowing *gadCB* expression, followed by acid challenge in the presence of 0.3 M NaCl and 1 mM glutamate the viability was reduced only 200-fold when lactic acid was used. Under the latter conditions, the viability in MS15 adjusted to pH 3.5 with hydrochloric acid was also enhanced significantly. Under conditions of *gadCB* expression (by the presence of 0.3 M NaCl) acid resistance was lower in the absence of glutamate than when glutamate was present (Table 5), indicating that glutamate is not only involved in induction of *gadCB* expression but is directly needed to confer acid resistance. To confirm the direct involvement of *gadCB* in acid resistance *L. lactis* MGNS3i3 (which does not express *gadCB*, see earlier Examples) was subjected to acid challenge.

15

Table 6. Acid resistance of *L. lactis gadCB* mutants.

Strain	Genotype;	% acid resistance <sup>a</sup> in MS15	% acid resistance <sup>a</sup> in MS15 + NaCl + glu
MG1363	w.t.	0.0005	0.49
MGNS3i3	<i>gadR</i>	0.0003	0.0013
20 MGNS3i4	<i>gadC</i> (GadB) <sup>b</sup>	0.00007	0.00009
MGNS3i5	(GadB) <sup>b</sup>	0.00003	0.00026
MGNS3i7	<i>gadB</i>	0.00001	0.00008

<sup>a</sup> at pH 3.5, lactic acid

25 <sup>b</sup> Insertions in *gadC* and in between *gadC* and *gadB* will, most likely, have a polar effect on *gadB*.

Survival in the presence of NaCl and glutamate was reduced to levels comparable to that of the wild-type strain in the absence of NaCl and glutamate (Table 6). An insertion mutant in *gadC* (strain MGNS3i4) showed a similar phenotype. This mutation, most likely, has a polar effect on *gadB* which would result in the absence of both GadC and GadB. An insertion mutant in *gadB* (strain MGNS3i7) is similarly acid sensitive, indicating a central role for GadB in the acid resistance

process. Strain MGNS3i5, carrying intact copies of *gadC* and *gadB* separated by an integrated copy of pORI19 is acid sensitive, suggesting that *gadB* expression is abolished. These data show that *gadCB*, which is transcribed from the 5 chloride-dependent *gad* promoter  $P_{\text{gad}}$ , is involved in glutamate-dependent acid resistance, as depicted in Figure 26.

#### SUMMARY

- 10 - *Lactococcus lactis* MG1363 has a glutamate-dependent acid resistance mechanism, that is active in the presence of chloride.
- *L. lactis* has two genes, *gadC* and *gadB*, that encode proteins homologous to a putative glutamate -  $\gamma$ -aminobutyrate  
15 antiporter and a glutamate decarboxylase, respectively, from *Escherichia coli* and *Shigella flexneri*. These genes are involved in glutamate-dependent acid resistance in *E. coli* and *S. flexneri*.
- The expression of *gadCB* in *L. lactis* is induced by chloride  
20 and is optimal at low pH in the presence of glutamate.
- *L. lactis* insertion mutants with a disrupted *gadB* or that are unable to express both *gadB* and *gadC* are more sensitive to low pH than the wild-type when NaCl and glutamate are present, indicating that the lactococcal *gadCB* operon is  
25 involved in glutamate dependent acid resistance, see Figure 26.

#### Conclusions to be drawn from Example 6

The results of further research given in Example 6 have shown  
30 that this salt-inducible promoter is more active at a lower pH. For example, Figure 25A shows that upon induction with 0.3 M NaCl the yield of  $\beta$ -galactosidase is not more than about 8 Units/mg when the medium is buffered such that the pH will not come below 5.5.

35 In contrast therewith, Figure 25B shows that in the absence of a buffer the pH can drop to about 4, while under these conditions the yield of  $\beta$ -galactosidase is about 80 Units/mg.

The salt-inducible promoter is more active at a lower pH and has thus also become pH-inducible in the presence of salt. The activity of the salt- and pH-inducible promoter can be further enhanced in the presence of glutamate/glutamic acid, 5 as is shown in Figure 25A and 25B. They show that production of  $\beta$ -galactosidase is increased from 8 to 15 Units/mg in a buffered medium and from 80 to 225 Units/mg in a non-buffered medium when the pH during fermentation has dropped to about 4.3.

10 Although the scope of the invention is not limited by any theory, it is supposed that these enhancing effects are caused by control of the internal pH by the uptake of glutamate and decarboxylation whereby  $\gamma$ -aminobutyrate is formed and subsequently secreted under influence of the *gadB* 15 and *gadC* genes described in this Example.

Thus Example 6 shows that a desired protein can be produced in a lactic acid bacterium at an improved yield by using a construct in which the gene encoding the desired protein is 20 under control of a salt- and pH-inducible promoter and the medium in which the transformed lactic acid bacterium is cultured is not buffered and preferably contains glutamate-/glutamic acid. For an easy recovery it is generally desirable that the desired protein contains a secretion 25 signal sequence to enable secretion of the protein.

Alternatively, if the transformed lactic acid bacterium also contains at least one gene encoding a lytic protein under control of the salt- and pH-inducible promoter, proteins 30 without a secretion signal sequence can be easily recovered, because the lytic protein will perforate the cell wall so that the contents including the desired protein can be released from the cell.

## REFERENCES

1. U.K. Laemmli; Nature (London) 227 (1970) 680-685; Cleavage of Structural Proteins during the Assembly of the  
5 Head of Bacteriophage T4
2. J.H. Miller; (1972); Experiments in molecular genetics Cold Spring Harbor laboratory, Cold Spring Harbor New York p. 352-359; Assay of  $\beta$ -Galactosidase Induction; Time Course of  $\beta$ -Galactosidase Induction
- 10 3. M.M. Bradford; Anal. Biochem. 72 (1976) 248-254; A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding
4. F. Sanger, S. Nicklen, and A.R. Coulson; Proc. Natl.  
15 Acad. Sci. USA 74 (1977) 5463-5467; DNA sequencing with chain-terminating inhibitors
5. M.J. Gasson; J. Bacteriol. 154 (1983) 1-9; Plasmid - Complements of *Streptococcus lactis* NCD0712 and Other Lactic Streptococci after Protoplast-Induced Curing
- 20 6. J.A. Gough, and N.E. Murray; J. Mol. Biol. 166 (1983) 1-19; Sequence Diversity among Related Genes for Recognition of Specific Targets in DNA Molecules
7. G.M. Weinstock, M.L. Berman, and T.J. Silhavy; In: T.S. Papas, M. Rosenberg, and J.G. Chirikjian (eds), Gene  
25 amplification and analysis III. Elsevier, North Holland, New York, (1983) pp. 27-64; Chimeric genes with  $\beta$ -galactosidase.
8. C. Yanisch-Perron, J. Vieira, and J. Messing; Gene 33 (1985) 103-119; Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19  
30 vectors
9. S.A. Lacks, P. Lopez, B. Greenberg, and M. Espinosa; J. Mol. Biol. 192 (1986) 753-765; Identification and Analysis of Genes for Tetracycline Resistance and Replication Functions in the Broad-host-range Plasmid pLS1
- 35 10. J.M.B.M. van der Vossen, D. van der Lelie, and G. Venema; Appl. Environ. Microbiol. 53 (1987) 2452-2457; Isolation and Characterization of *Streptococcus cremoris* Wg2-Specific Promoters

11. J.A.K.W. Kiel, J.P.M.J. Vossen, and G. Venema; Mol. Gen. Genet. 207 (1987) 294-301; A general method for the construction of *Escherichia coli* mutants by homologous recombination and plasmid segregation
- 5 12. W.R. Pearson and D. J. Lipman; Proc. Natl. Acad. Sci. USA 85 (1988) 2444-2448; Improved tools for biological sequence comparison
13. S.P. Chambers, S.E. Prior, D.A. Barstow, and N.P. Minton; Gene 68 (1988) 139-149; The pMTL *nic*<sup>-</sup>cloning vectors,  
10 I. Improved pUC polylinker regions to facilitate the use of sonicated DNA for nucleotide sequencing
14. K.J. Leenhouts, J. Kok, and G. Venema; Appl. Environ. Microbiol. 55 (1989) 394-400; Campbell-Like Integration of Heterologous Plasmid DNA into the Chromosome of *Lactococcus*  
15 *lactis* subsp. *lactis*.
15. H. Holo, and I.F. Nes; Appl. Environ. Microbiol. 55 (1989) 3119-3123; High-Frequency Transformation, by Electroporation, of *Lactococcus lactis* subsp. *cremoris* Grown with Glycine in Osmotically Stabilized Media
- 20 16. J. Sambrook, E. F. Fritsch, and T. Maniatis; Molecular cloning: a laboratory manual, 2nd ed. (1989); Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
17. K.J. Leenhouts, J. Kok, and G. Venema; Appl. Environ. Microbiol. 56 (1990) 2726-2735; Stability of Integrated  
25 Plasmids in the Chromosome of *Lactococcus lactis*
18. M. Itaya; Proc. Natl. Acad. Sci. USA 87 (1990) 8587-8591; Isolation and characterization of a second RNase H (RNase HII) of *Escherichia coli* K-12 encoded by the *rnhB* gene
19. E.R. Zabarovsky and G. Winberg; Nucleic Acids Res. 18  
30 (1990) 5912; High efficiency electroporation of ligated DNA into bacteria
20. B. Mayo, J. Kok, K. Venema, W. Bockelmann, M. Teuber, H. Reinke, and G. Venema; Appl. Environ. Microbiol. 57 (1991) 38-44; Molecular Cloning and Sequence of an X-Prolyl  
35 Dipeptidyl Aminopeptidase Gene from *Lactococcus lactis* subsp. *cremoris*



21. **K.J. Leenhouts, J. Kok, and G. Venema;** J. Bacteriol. 173 (Aug. 1991) 4794-4798; Replacement Recombination in *Lactococcus lactis*
22. **K.J. Leenhouts, J. Kok, and G. Venema;** Appl. Environ. Microbiol. 57 (Sept. 1991) 2562-2567; Lactococcal Plasmid pWV01 as an Integration Vector for Lactococci
23. **J. Vieira and J. Messing;** Gene 100 (1991) 189-194; New pUC-derived cloning vectors with different selectable markers and DNA replication origins
- 10 24. **M. van de Guchte, J. Kok, and G. Venema;** Mol. Gen. Genet. 227 (1991) 65-71; Distance-dependent translational coupling and interference in *Lactococcus lactis*
25. **M. van de Guchte, J. Kok, and G. Venema;** FEMS Microbiol. Rev. 88 (1992) 73-92; Gene expression in  
15 *Lactococcus lactis*
26. **C.J. Leenhouts, J.D. Marugg and C.T. Verrips** (Unilever N.V./PLC); Patent Application EP-A1-0 487 159 published 27 May 1992; A food-grade vector suitable for transforming a food-grade host cell, use of said vector for transforming  
20 food-grade host cells, and use of said transformed cells in biotransformation processes
27. **M.C. Sulavik, G. Tardif, and D.B. Clewell;** J. Bacteriol. 174 (1992) 3577-3586; Identification of a Gene, *rgg*, Which Regulates Expression of Glucosyltransferase and  
25 Influences the Spp Phenotype of *Streptococcus gordonii*  
Challis
28. **E. Maguin, P. Duwat, T. Hege, D. Ehrlich, and A. Gruss;** J. Bacteriol. 174 (1992) 5633-5638; New Thermosensitive Plasmid for Gram-Positive Bacteria
- 30 29. **M. van Asseldonk, A. Simons, H. Visser, W. M. de Vos, and G. Simons;** J. Bacteriol. 175 (1993) 1637-1644; Cloning, Nucleotide Sequence, and Regulatory Analysis of the *Lactococcus lactis dnaJ* Gene
30. **B.L. Geller, R.G. Ivey, J.E. Trempy, and B. Hettlinger-**  
35 **Smith;** J. Bacteriol. 175 (1993) 5510-5519; Cloning of a Chromosomal Gene Required for Phage Infection of *Lactococcus lactis* subsp. *lactis* C2

31. G. Buist, J. Kok, K.J. Leenhouts, M. Dabrowska, G. Venema, and A. Haandrikman; J. Bacteriol. 177 (1995) 1554-1563; Molecular Cloning and Nucleotide Sequence of the Gene Encoding the Major Peptidoglycan Hydrolase of *Lactococcus*  
5 *lactis*, a Muramidase Needed for Cell Separation
32. P. Le Bourgeois, M. Lautier, L. van den Berghe, M.J. Gasson, and P. Ritzenthaler; J. Bacteriol. 177 (1995) 2840-2850; Physical and Genetic Map of the *Lactococcus lactis* subsp. *cremoris* MG1363 Chromosome: Comparison with That of  
10 *Lactococcus lactis* subsp. *lactis* IL1403 Reveals a Large Genome Inversion
33. R.D. Fleischmann c.s.; Science 269 (1995) 496-512; Whole-Genome Random Sequencing and Assembly of *Haemophilus influenzae* Rd
- 15 34. G. Buist, G. Venema, J. Kok, and A.M. Ledeboer (Quest International B.V.); International Patent Application WO-A1-95/31561 published 23 Nov. 1995 (PCT); PROCESS FOR THE LYSIS OF A CULTURE OF LACTIC ACID BACTERIA BY MEANS OF A LYSIN, AND USES OF THE RESULTING LYSED CULTURE
- 20 35. A. Nauta, G. Venema, J. Kok, and A.M. Ledeboer (Quest International B.V.); International Patent Application WO-A1-95/31562 published 23 Nov. 1995 (PCT); PROCESS FOR INHIBITING THE GROWTH OF A CULTURE OF LACTIC ACID BACTERIA, AND OPTIONALLY LYSING THE BACTERIAL CELLS, AND USES OF THE  
25 RESULTING LYSED CULTURE
36. A. Nauta, G. Venema, J. Kok, and A.M. Ledeboer (Quest International B.V.); International Patent Application WO-A1-95/31563 published 23 Nov. 1995 (PCT); COMPLEX INDUCIBLE PROMOTER SYSTEM DERIVABLE FROM A PHAGE OF A LACTIC ACID  
30 BACTERIUM (LAB), AND ITS USE IN A LAB FOR PRODUCTION OF A DESIRED PROTEIN
37. J. Law, G. Buist, A. Haandrikman, J. Kok, G. Venema, and K. Leenhouts; J. Bacteriol. 177 (1995) 7011-7018; A System to Generate Chromosomal Mutations in *Lactococcus*  
35 *lactis* Which Allows Fast Analysis of Targeted Genes.
38. D.van Sinderen, H. Karsens, J. Kok, P. Terpstra, M.H.J. Ruiters, G. Venema, and A. Nauta; Mol. Microbiol. 19 (No.6) (1996) 1343-1355; Sequence analysis and molecular

- characterization of the temperate lactococcal bacteriophage  
rlt
39. K. Leenhouts, G. Buist, A. Bolhuis, A. ten Berge, J.  
Kiel, I. Mierau, M. Dabrowska, G. Venema, and J. Kok; Mol.  
5 Gen. Genet. (1996, Accepted for publication); A general  
system for generating unlabelled gene replacements in  
bacterial chromosomes  
= K.J. Leenhouts, G. Buist, A. Bolhuis, A. ten Berge, J.  
Kiel, I. Mirau, M. Dabrowska, G. Venema, and J. Kok; Mol.  
10 Gen. Genet. 253 (1996) 217-224; A general system for  
generating unlabeled gene replacements in bacterial  
chromosomes (Published after 5 Sep 1996)
40. K.J. Leenhouts, A. Bolhuis, G. Venema, and J. Kok;  
(Submitted for publication 1996); Construction of a food-  
15 grade multiple copy integration system for *Lactococcus*  
*lactis*.
41. S.F. Altschul, W. Gish, W. Miller, E. W. Myers, and D.  
J. Lipman; J. Mol. Biol. 215 (1990) 403-410; Basic local  
alignment search tool
- 20 42. M. Coccagn-Bousquet, C. Garrigues, L. Novak, N. D.  
Lindley, and P. Loubiere; J. Appl. Bacteriol. 79 (1995) 108-  
116; Rational development of a simple synthetic medium for  
the sustained growth of *Lactococcus lactis*
43. B.M. Hersh, F. T. Farooq, D. N. Barstad, D. L.  
25 Blankenhorn, and J. L. Slonczewski; J. Bacteriol. 178; (1996)  
3978-3981; A glutamate-dependent acid resistance gene in  
*Escherichia coli*
44. T. Kaneko, S. Sato, H. Kotani, A. Tanaka, E. Asamizu,  
Y. Nakamura, N. Miyajima, M. Hirose, M. Sugiura, S.  
30 Sasamoto, T. Kimura, T. Hosouchi, A. Matsuno, A. Muraki, N.  
Nakazaki, K. Naruo, S. Okumura, S. Shimpo, C. Takeuchi, T.  
Wada, A. Watanabe, M. Yamada, M. Yasuda, and S. Tabata; J.  
DNA Res. 3 (1996) 109-136; Sequence analysis of the genome of  
the unicellular cyanobacterium *Synechocystis* sp. strain  
35 PCC6803. II. Sequence determination of the entire genome and  
assignment of potential protein-coding regions

45. P. Klein, M. Kanehisa, and C. DeLisi; *Biochim. Biophys. Acta* 815 (1985) 468-476; The detection and classification of membrane spanning proteins
46. S. Meng, and G.N. Bennett; *J. Bacteriol.* 174 (1992) 5 2659-2669; Nucleotide sequence of the *Escherichia coli cad* operon: a system for neutralization of low extracellular pH
47. S.R. Waterman, and P.L.C. Small; *Mol. Microbiol.* 21 (1996) 925-940; Identification of  $\sigma^S$ -dependent genes associated with the stationary-phase acid-resistance  
10 phenotype of *Shigella flexneri*

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT (except for U.S.A.):
  - (A) NAME: Unilever N.V.
  - (B) STREET: Weena 455
  - (C) CITY: Rotterdam
  - (E) COUNTRY: The Netherlands
  - (F) POSTAL CODE (ZIP): NL-3013 AL
  
- (i) APPLICANT (except for U.S.A.):
  - (A) NAME: Unilever PLC
  - (B) STREET: Unilever House, Blackfriars
  - (C) CITY: London
  - (E) COUNTRY: United Kingdom
  - (F) POSTAL CODE (ZIP): EC4P 4BQ (GB)
  
- (i) APPLICANT (for U.S.A. only):
  - (A) NAME: Jan Willem SANDERS  
c/o Rijksuniversiteit Groningen, Afdeling Biologie
  - (B) STREET: Kerklaan 30 (P.O. Box 14)
  - (C) CITY: Haren
  - (E) COUNTRY: The Netherlands
  - (F) POSTAL CODE (ZIP): NL-9750 AA
  
- (i) APPLICANT (for U.S.A. only):
  - (A) NAME: Adrianus Marinus LEDEBOER  
c/o Unilever Research Vlaardingen
  - (B) STREET: Olivier van Noortlaan 120
  - (C) CITY: Vlaardingen
  - (E) COUNTRY: The Netherlands
  - (F) POSTAL CODE (ZIP): NL-3133 AT
  
- (i) APPLICANT (for U.S.A. only):
  - (A) NAME: Gerard VENEMA  
c/o Rijksuniversiteit Groningen, Afdeling Biologie
  - (B) STREET: Kerklaan 30 (P.O. Box 14)
  - (C) CITY: Haren
  - (E) COUNTRY: The Netherlands
  - (F) POSTAL CODE (ZIP): NL-9750 AA
  
- (i) APPLICANT (for U.S.A. only):
  - (A) NAME: Jan KOK  
c/o Rijksuniversiteit Groningen, Afdeling Biologie
  - (B) STREET: Kerklaan 30 (P.O. Box 14)
  - (C) CITY: Haren
  - (E) COUNTRY: The Netherlands
  - (F) POSTAL CODE (ZIP): NL-9750 AA
  
- (ii) TITLE OF INVENTION: Salt-inducible promoter derivable  
from a lactic acid bacterium
  
- (iii) NUMBER OF SEQUENCES: 25
  
- (iv) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
  
- (v) CURRENT APPLICATION DATA:
  - APPLICATION NUMBER: .....

## (2) INFORMATION FOR SEQ ID NO: 1:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "synthetic DNA"

- (vii) IMMEDIATE SOURCE:
  - (B) CLONE: primer NS3-1

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

CATTGAGATA ATCAGATAC

19

- (2) INFORMATION FOR SEQ ID NO: 2:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 16 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "synthetic DNA"

- (vii) IMMEDIATE SOURCE:
  - (B) CLONE: primer NS3-2

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

GCAGAGATTG GGAAG

16

- (2) INFORMATION FOR SEQ ID NO: 3:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 34 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "synthetic DNA"

- (vii) IMMEDIATE SOURCE:
  - (B) CLONE: primer NS3-5

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GGGGCCCTCT CTTATGTGTT AAATTTTCAG GCGC

34

- (2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "synthetic DNA"

- (vii) IMMEDIATE SOURCE:
  - (B) CLONE: primer NS3-6

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

ATATCGTTCA CGTTTTCC

18

## (2) INFORMATION FOR SEQ ID NO: 5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 43 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

## (vii) IMMEDIATE SOURCE:

- (B) CLONE: primer NS3-7

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

GCGATATCCA GTACTTCATC ATACCTCCTT ATATTTATGA TTG

43

## (2) INFORMATION FOR SEQ ID NO: 6:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

## (vii) IMMEDIATE SOURCE:

- (B) CLONE: primer NS3-8

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

GCGAGCTCAG ATCTGAGCGT TGTATAAGCT TTTATGTCTT TC

42

## (2) INFORMATION FOR SEQ ID NO: 7:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

## (vii) IMMEDIATE SOURCE:

- (B) CLONE: primer NS3-9

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GTTTGACTGA CCCAAC

16

## (2) INFORMATION FOR SEQ ID NO: 8:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

## (vii) IMMEDIATE SOURCE:

- (B) CLONE: primer NS3-10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

CCGCTTCAAT GGTTTTG

17

(2) INFORMATION FOR SEQ ID NO: 9:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 38 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "synthetic DNA"

- (vii) IMMEDIATE SOURCE:
  - (B) CLONE: primer NS3-11

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

CAGTCAAAC CATTGAAGCG GTTAATGCGA AAAAACCG

38

(2) INFORMATION FOR SEQ ID NO: 10:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 2989 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "genomic DNA of NS3 locus"

- (vii) IMMEDIATE SOURCE:
  - (B) CLONE: PstI-XbaI fragment

- (ix) FEATURE:
  - (A) NAME/KEY: mat\_peptide
  - (B) LOCATION:3..500
  - (D) OTHER INFORMATION:/product= "C-terminus of rnhB3"

- (ix) FEATURE:
  - (A) NAME/KEY: mat\_peptide
  - (B) LOCATION:1095..1922
  - (D) OTHER INFORMATION:/product= "rggL"

- (ix) FEATURE:
  - (A) NAME/KEY: mat\_peptide
  - (B) LOCATION:2069..2989
  - (D) OTHER INFORMATION:/product= "N-terminus of lacZ"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

CT GCA GTA ATT TTG CCA AAG AAT TGT AAA ATT CGT GGT TTG AAT GAT 47  
 Ala Val Ile Leu Pro Lys Asn Cys Lys Ile Arg Gly Leu Asn Asp  
   5  10  15

AGT AAA AAA GTG CCA AAA TCA AAG CAT CAT GCT ATT CTA TCT GAA ATT 95  
 Ser Lys Lys Val Pro Lys Ser Lys His His Ala Ile Leu Ser Glu Ile  
   20  25  30

CAA GAA AAA GCG CTA GCG ATT GGA GTT GGA ATT GTT GAT GCC GAA AAA 143  
 Gln Glu Lys Ala Leu Ala Ile Gly Val Gly Ile Val Asp Ala Glu Lys  
   35  40  45

ATT GAT GAA GTA AAT ATT TAT GAG GCG ACA AAA ATT GCA ATG ATT CAG 191  
 Ile Asp Glu Val Asn Ile Tyr Glu Ala Thr Lys Ile Ala Met Ile Gln  
   50  55  60



GCA GTA TCA AAA TTA TCT CTT AAA CCT GAA CAT CTC TTA ATA GAT GCG	239
Ala Val Ser Lys Leu Ser Leu Lys Pro Glu His Leu Leu Ile Asp Ala	
65 70 75	
ATG GTT TTA GAT TTG CCC ATT GCT CAG ACG AAA ATC ATT CAT GGA GAT	287
Met Val Leu Asp Leu Pro Ile Ala Gln Thr Lys Ile Ile His Gly Asp	
80 85 90 95	
GCT CGT TCA GCT TCA ATT GCG GCC GCA TCA ATT GTA GCT AAA GTG ACT	335
Ala Arg Ser Ala Ser Ile Ala Ala Ala Ser Ile Val Ala Lys Val Thr	
100 105 110	
CGT GAT GAA ATG ATG AAG GAT TTC GCT TTA GAA TTT CCA GAA TAT GAT	383
Arg Asp Glu Met Lys Asp Phe Ala Leu Glu Phe Pro Glu Tyr Asp	
115 120 125	
TTT GAA CAT AAT GCA GGC TAT GGA ACA GCA AAA CAT CTT GCA GCT CTG	431
Phe Glu His Asn Ala Gly Tyr Gly Thr Ala Lys His Leu Ala Ala Leu	
130 135 140	
ACA AAA TAT GGT ATC ACA AGA ATT CAT CGG AAA TCC TAT GAA CCA ATT	479
Thr Lys Tyr Gly Ile Thr Arg Ile His Arg Lys Ser Tyr Glu Pro Ile	
145 150 155	
AAA TCG ATG GTC AAT TTC AAA TAG TAGATTATGT AAGTAAAAAA AGGAAAACGT	533
Lys Ser Met Val Asn Phe Lys ***	
160 165	
GAACGATATT TGGTCACGTT TTTTTGCTGA CAAGTCTGTC AGTAATTATT TTCAAAGGTT	593
TCAAAAAATAT AGTCTAGTAA TTTGCTAGAC TAATCTTCTG TTTTTTAATA ATAACTAAT	653
TTTTTGTTAA TCTAAATGAC AAAATTAATA AGCAGAGTTT TTTATAAAAT TAGCTACTTA	713
TAAAAAATTT GAAATTGGTA TAGTTAAATC TGTTATAAAT TCCAATATTT TTTAATAATA	773
ATTATTTTAA CAAAATACTT ATATCAAAAC TCTTTCAAAG TATATAATGA GCGTTGTATA	833
AGCTTTTATG TCTTTCTATA TCAACTTTTA ATAGAAATAT AAAGTAATAT AAATGTTTTT	893
ATAATAAATTT ATGTGAGATA TATTTTTTTG TCCGTACTGG TATAGATTTG ACGATTAAGT	953
CTTAAATAAG TTATAATCTC AATTGCGTAA TTTCTTAAAT ACAGAAATAA CACTACATT	1013
GGTAGACTGA TTAAAAAGTG TACTTGATGA ACTGTTATAA ACCTTAAAAA AATAAAAAATA	1073
ATAGTTTGGG GGATGTTAAA G ATG TAT AAA AAA TAT GGA GAT TGT TTT AAA	1124
Met Tyr Lys Lys Tyr Gly Asp Cys Phe Lys	
5 10	
AAG TTG CGA AAC CAA AAG AAT TTA GGG TTA TCA TAC TTT AGT AAA CTT	1172
Lys Leu Arg Asn Gln Lys Asn Leu Gly Leu Ser Tyr Phe Ser Lys Leu	
15 20 25	
GGA ATA GAC CGT TCA AAT ATA TCT AGA TTT GAA CAT GGA AAA TGT ATG	1220
Gly Ile Asp Arg Ser Asn Ile Ser Arg Phe Glu His Gly Lys Cys Met	
30 35 40	
ATG AGT TTT GAG CGT ATA GAT TTG ATG TTA GAA GAA ATG CAA GTT CCG	1268
Met Ser Phe Glu Arg Ile Asp Leu Met Leu Glu Glu Met Gln Val Pro	
45 50 55	
TTA TCT GAG TAC GAA TTG ATT GTA AAT AAT TTT ATG CCG AAT TTC CAA	1316
Leu Ser Glu Tyr Glu Leu Ile Val Asn Asn Phe Met Pro Asn Phe Gln	
60 65 70	

GAA TTT TTT ATA TTA GAA TTG GAA AAA GCT GAA TTT AGC CAA AAT CGA 1364  
 Glu Phe Phe Ile Leu Glu Leu Glu Lys Ala Glu Phe Ser Gln Asn Arg  
 75 80 85 90

GAT AAA ATA AAA GAG TTG TAT TCT GAG GTC AAA GAA ACG GGG AAT CAT 1412  
 Asp Lys Ile Lys Glu Leu Tyr Ser Glu Val Lys Glu Thr Gly Asn His  
 95 100 105

TTA CTG ACG GTT ACC GTG AAA ACG AAG CTT GGG AAT ATA AGT CAG ACA 1460  
 Leu Leu Thr Val Thr Val Lys Thr Lys Leu Gly Asn Ile Ser Gln Thr  
 110 115 120

GAA GTT AAA GAA ATT GAA GCT TAT CTT TGC AAT ATT GAA GAG TGG GGA 1508  
 Glu Val Lys Glu Ile Glu Ala Tyr Leu Cys Asn Ile Glu Glu Trp Gly  
 125 130 135

TAT TTT GAA CTT ACT TTA TTT TAT TTT GTA TCT GAT TAT CTC AAT GTC 1556  
 Tyr Phe Glu Leu Thr Leu Phe Tyr Phe Val Ser Asp Tyr Leu Asn Val  
 140 145 150

AAT CAA TTA GAA TTG CTG CTT TTT AAT TTT GAT AAA AGA TGT GAA AAT 1604  
 Asn Gln Leu Glu Leu Leu Leu Phe Asn Phe Asp Lys Arg Cys Glu Asn  
 155 160 165 170

TAC TGT AGA GTC TTA AAA TAT AGA AGG AGA CTA TTG CAA ATA GCC TAT 1652  
 Tyr Cys Arg Val Leu Lys Tyr Arg Arg Arg Leu Leu Gln Ile Ala Tyr  
 175 180 185

AAA AGT GTT GCG ATA TAC GCG GCT AAA GGA GAA AGA AAA AAA GCC GAA 1700  
 Lys Ser Val Ala Ile Tyr Ala Ala Lys Gly Glu Arg Lys Lys Ala Glu  
 190 195 200

AAT ATT TTA GAA ATG ACT AAA AAA TAT CGA ACT GTG GGA GTC GAT TTA 1748  
 Asn Ile Leu Glu Met Thr Lys Lys Tyr Arg Thr Val Gly Val Asp Leu  
 205 210 215

TAT TCA GAA GTA TTA AGA CAT CTT GCT AGA GCT ATC ATT ATT TTT AAT 1796  
 Tyr Ser Glu Val Leu Arg His Leu Ala Arg Ala Ile Ile Ile Phe Asn  
 220 225 230

TTT GAA AAT GCA GAG ATT GGG GAA GAA AAA ATA AAT TAT GCT CTT GAG 1844  
 Phe Glu Asn Ala Glu Ile Gly Glu Glu Lys Ile Asn Tyr Ala Leu Glu  
 235 240 245 250

ATT TTG GAA GAA TTT GGA GGA AAG AAG ATA AAA GAA TTC TAT CAG AAT 1892  
 Ile Leu Glu Glu Phe Gly Gly Lys Lys Ile Lys Glu Phe Tyr Gln Asn  
 255 260 265

AAA ATG GAA AAG TAT TTG AAA AGG TCA ATT TAG TCTCTTTTGA GCTGTTGCTT 1945  
 Lys Met Glu Lys Tyr Leu Lys Arg Ser Ile \*\*\*  
 270 275

TAAAGCAACA GCTCAAAGA GATTTTCTTT ATTCTAGAGC ATATACTAGA GGGTGAAGAT 2005

AGGTTGTCTG AAGCATTATA ACTTGTCTTT TAAAAATTC AATCATAAAT ATAAGGAGGT 2065

ATG ATG AAT CAA AAA AAA TTA TCA TTA TTC GGT TTT TTC GCA TTA ACC 2113  
 Met Asn Gln Lys Lys Leu Ser Leu Phe Gly Phe Phe Ala Leu Thr  
 5 10 15

GCT TCA ATG GTT TTG ACT GTC TAT GAG TAT CCG ACT TTT GCC ACG TCA 2161  
 Ala Ser Met Val Leu Thr Val Tyr Glu Tyr Pro Thr Phe Ala Thr Ser  
 20 25 30

AAA TTA CAT TTG GTG TTC TTT TTA CTT CTC GGA GGA CTA CTA TGG TTT 2209  
 Lys Leu His Leu Val Phe Phe Leu Leu Gly Gly Leu Leu Trp Phe  
 35 40 45



## (2) INFORMATION FOR SEQ ID NO: 11:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 276 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein rggL ex *L. lactis* MG1363

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

MYKKYGDCFK	KLRNQKNLGL	SYFSKLGIDR	SNISRFEHGK	CMMSFERIDL	50
MLEEMQVPLS	EYELIVNNFM	PNFQEFFILE	LEKAEFSONR	DKIKELYSEV	100
KETGNHLLTV	TVKTKLGNIS	QTEVKEIEAY	LCNIEEWGYF	ELTLFYFVSD	150
YLNVNQLELL	LFNFDKRCEN	YCRVLKYRRR	LLQIAYKVA	IYAAKGERKK	200
AENILEMTKK	YRTVGVDLYS	EVLRLARAI	IIFNFENAEI	GEEKINYALE	250
ILEEFGGKKI	KEFYQNKMEK	YLKRSI			276

## (2) INFORMATION FOR SEQ ID NO: 12:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 297 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein rgg1 ex *S. gordonii*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

MLIVKSSGKI	LKI IRESK NM	SLKEVAAGDI	SVAQLSRYER	GISSLTVDSF	50
YSCLRNMVS	LAEFQYVYHN	YREADDVVLS	QKLSEAQREN	NIVKLESILA	100
GSEAMAQEF	EKKNYKLNTI	VIRATLTSCN	PDYQVSKGDI	EFLTDYLFSV	150
EEWGRYELWL	FTNSVNLTL	ETLETFASEM	INRTQFYNNL	PENRRRIK M	200
LLNVVSACIE	NNHLQVAMKF	LNYIDNTKIP	ETDLYDRVLI	KYHKALYSYK	250
VGNPHARHDI	EQCLSTFEYL	DSFGVARKLK	EQFERIQLTV	VADLQIE	297

## (2) INFORMATION FOR SEQ ID NO: 13:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 97 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein ORF3 ex *L. lactis* C2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

MPYKRYGEIF	KKLREQKNFS	LSHFSEIGIS	KASLSRFELG	QTMISFERLD	50
SALQEMNVTL	AEYEHFINNF	SMDYKEEFLE	DIILADIADD	VDKLNK L	97

## (2) INFORMATION FOR SEQ ID NO: 14:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 166 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein C-terminus of rnhB ex *L. lactis*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

AVILPKNCKI	RGLNDSKKVP	KSKHHAILSE	IQEKALAIGV	GIVDAEKIDE	50
VNIYEATKIA	MIQAVSKLSL	KPEHLLIDAM	VLDLPIAQTK	IIHGDARSAS	100
IAAASIVAKV	TRDEMMK DFA	LEFPEYDFEH	NAGYG TAKHL	AALTKYGITR	150
IHRKSYEPIK	SMVNFK				166

## (2) INFORMATION FOR SEQ ID NO: 15:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 161 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein C-terminus of rnhB ex  
*Vibrio cholera* (U30472)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

PNRPIMGLND SKKLSEKKRL ALFPEIQVKA LAWAVGRCSP QEIDELNIFQ	50
ATMVAMQRAV AGLRIQPDV LIDGNKIPKL PMEAVVKG DLRVAQISAA	100
SIIAKVIRDQ EMEALDKQYP QFGFANHKGY PTAHFAAIE QHGVIEQHRK	150
SFGPVKRALG E	161

## (2) INFORMATION FOR SEQ ID NO: 16:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 213 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein rnhB ex *E. coli*

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

MIEFVYPHTQ LVAGVDEVGR GPLVGAVVTA AVILDPARPI AGLNDSKCLS	50
EKRRLALYEE IKEKALSWSL GRAEPHEIDE LNILHATMLA MQRVAVGLHI	100
APEYVLIDGN RCPKLPMPAM AVVKGDSRVP EISAASILAK VTRDAEMAAL	150
DIVFPQYGFA QHKGYPYTA FH LEKLAHEGAT EHHRRSFGPV KRALGTCVLI	200
LVSRLSKPES EDV	213

## (2) INFORMATION FOR SEQ ID NO: 17:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 197 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein rnhB ex *Haemophilus influenza*

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

MFYEYQGYKL IAGVDEVGRG PLVGAVVTAA VILDPHNPIE GLADSKLSE	50
KKRLALAEI KEKARAWALG RAEADEIDEI NILQASLLAM TRAVKSLKIQ	100
PHFVLIDGNK IPKDLAIPAQ AVVKGDSLVA EISAASILAK VARDQEMEEL	150
DKQYPEYafa QHKGYPYTKLH LEKLAELGAL PQHRRSFAPV KKALEQF	197

## (2) INFORMATION FOR SEQ ID NO: 18:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 360 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "fusion point of truncated orfX  
with lytPR in pNS3PR"

(vii) IMMEDIATE SOURCE:

- (B) CLONE: pNS3PR

(ix) FEATURE:

- (A) NAME/KEY: mat\_peptide
- (B) LOCATION: 92..283

(D) OTHER INFORMATION:/product= "C-terminus of ORFX"

(ix) FEATURE:

- (A) NAME/KEY: mat\_peptide
- (B) LOCATION:297..359
- (D) OTHER INFORMATION:/product= "lytP"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

```

TCTAGAGCAT ATACTAGAGG GTGAAGATAG GTTGTCTGAA GCATTATAAC TTGTCTTTTA      60
AAAAATTCAA TCATAAATAT AAGGAGGTAT G ATG AAG TAC TGG ATA CTT ATC      112
                               Met Lys Tyr Trp Ile Leu Ile
                               5

AAT GAT GAA CCT TGG TTT GTC GGA AAA GAT GTA GCA ATT GCT ATT GGT      160
Asn Asp Glu Pro Trp Phe Val Gly Lys Asp Val Ala Ile Ala Ile Glu
   10                15                20

TAC AAG AAT TTC AGG GAT GCT TTG AAA TCT CAT GTA AAA GAC AAA TAT      208
Tyr Lys Asn Phe Arg Asp Ala Leu Lys Ser His Val Lys Asp Lys Tyr
   25                30                35

AAG AGG GAG TCG GAC AGC AGT GAT TGG TTC AAC GAC AAT ATA TTA TTG      256
Lys Arg Glu Ser Asp Ser Ser Asp Trp Phe Asn Asp Asn Ile Leu Leu
   40                45                50                55

GAA ACG AAC TGC ATA AAA AAT AAA AAA TAG GAGAAAGAAC ATG AAA ACA      305
Glu Thr Asn Cys Ile Lys Asn Lys Lys *** Met Lys Thr
   60

TTT TTT AAA GAT ATG GCA GAA CGT GCC ATT AAA ACA TTT GCA CAA GCA      353
Phe Phe Lys Asp Met Ala Glu Arg Ala Ile Lys Thr Phe Ala Gln Ala
   5                10                15

ATG ATT G      360
Met Ile
   20
    
```

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 180 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "fusion point of truncated orfX with acmA in pNS3AL3"

(vii) IMMEDIATE SOURCE:

- (B) CLONE: pNS3AL3

(ix) FEATURE:

- (A) NAME/KEY:mat\_peptide
- (B) LOCATION:92..154
- (D) OTHER INFORMATION:/product= "truncated ORFX"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

```

TCTAGAGCAT ATACTAGAGG GTGAAGATAG GTTGTCTGAA GCATTATAAC TTGTCTTTTA      60
AAAGATTCAA TCATAAATAT AAGGAGGTAT G ATG AAG TAC TTA TTA TAT TTT      112
                               Met Lys Tyr Leu Leu Tyr Phe
                               5
    
```

GTA ATC TTT AGA AAG GTA ATT ATT TAT GCC AGT ATC ACG TGT TAA 157  
 Val Ile Phe Arg Lys Val Ile Ile Tyr Ala Ser Ile Thr Cys \*\*\*  
 10 15 20

AGTTAAAAAT AGACATTTAA AAA 180

(2) INFORMATION FOR SEQ ID NO: 20:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 179 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "fusion point of truncated orfx with acmA in pNS3AL3"
- (vii) IMMEDIATE SOURCE:
  - (B) CLONE: pNS3AL3
- (ix) FEATURE:
  - (A) NAME/KEY: mat\_peptide
  - (B) LOCATION: 92..112
  - (D) OTHER INFORMATION: /product= "truncated ORFX"
- (ix) FEATURE:
  - (A) NAME/KEY: mat\_peptide
  - (B) LOCATION: 137..178
  - (D) OTHER INFORMATION: /product= "N-terminus of acmA"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

TCTAGAGCAT ATACTAGAGG GTGAAGATAG GTTGTCTGAA GCATTATAAC TTGTCITTTA 60  
 AAAGATTCAA TCATAAATAT AAGGAGGTAT G ATG AAG TCT TAT TAT ATT TTG TAA 115  
 Met Lys Ser Tyr Tyr Ile Leu \*\*\*  
 5  
 TCTTTAGAAA GGTAATTATT T ATG CCA GTA TCA CGT GTT AAA GTT AAA AAT AGA 169  
 Met Pro Val Ser Arg Val Lys Val Lys Asn Arg  
 5 10  
 CAT TTA AAA A 179  
 His Leu Lys

(2) INFORMATION FOR SEQ ID NO: 21:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 5565 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "genomic DNA of NS3 locus containing gadRCB"
- (vii) IMMEDIATE SOURCE:
  - (B) CLONE: PstI-EcoRI fragment
- (ix) FEATURE:
  - (A) NAME/KEY: mat\_peptide
  - (B) LOCATION: 3..500
  - (D) OTHER INFORMATION: /product= "C-terminus of rnhB"
- (ix) FEATURE:
  - (A) NAME/KEY: mat\_peptide
  - (B) LOCATION: 1095..1922
  - (D) OTHER INFORMATION: /product= "rggL = gadR"

(ix) FEATURE:

- (A) NAME/KEY: mat\_peptide
- (B) LOCATION: 2069..3577
- (D) OTHER INFORMATION:/product= "orfX = gadC"

(ix) FEATURE:

- (A) NAME/KEY: mat\_peptide
- (B) LOCATION: 3600..4997
- (D) OTHER INFORMATION:/product= "gadB"

(ix) FEATURE:

- (A) NAME/KEY: mat\_peptide
- (B) LOCATION: 5563..5078
- (D) OTHER INFORMATION:/product= "C-terminus of unknown ORF in opposite direction"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

```

CTG GCA GTA ATT TTG CCA AAG AAT TGT AAA ATT CGT GGT TTG AAT GAT      47
Ala Val Ile Leu Pro Lys Asn Cys Lys Ile Arg Gly Leu Asn Asp
                    5                      10

AGT AAA AAA GTG CCA AAA TCA AAG CAT CAT GCT ATT CTA TCT GAA ATT      95
Ser Lys Lys Val Pro Lys Ser Lys His His Ala Ile Leu Ser Glu Ile
                    20                      25                      30

CAA GAA AAA GCG CTA GCG ATT GGA GTT GGA ATT GTT GAT GCC GAA AAA      143
Gln Glu Lys Ala Leu Ala Ile Gly Val Gly Ile Val Asp Ala Glu Lys
                    35                      40                      45

ATT GAT GAA GTA AAT ATT TAT GAG GCG ACA AAA ATT GCA ATG ATT CAG      191
Ile Asp Glu Val Asn Ile Tyr Glu Ala Thr Lys Ile Ala Met Ile Gln
                    50                      55                      60

GCA GTA TCA AAA TTA TCT CTT AAA CCT GAA CAT CTC TTA ATA GAT GCG      239
Ala Val Ser Lys Leu Ser Leu Lys Pro Glu His Leu Leu Ile Asp Ala
                    65                      70                      75

ATG GTT TTA GAT TTG CCC ATT GCT CAG ACG AAA ATC ATT CAT GGA GAT      287
Met Val Leu Asp Leu Pro Ile Ala Gln Thr Lys Ile Ile His Gly Asp
                    80                      85                      90                      95

GCT CGT TCA GCT TCA ATT GCG GCC GCA TCA ATT GTA GCT AAA GTG ACT      335
Ala Arg Ser Ala Ser Ile Ala Ala Ala Ser Ile Val Ala Lys Val Thr
                    100                      105                      110

CGT GAT GAA ATG ATG AAG GAT TTC GCT TTA GAA TTT CCA GAA TAT GAT      383
Arg Asp Glu Met Met Lys Asp Phe Ala Leu Glu Phe Pro Glu Tyr Asp
                    115                      120                      125

TTT GAA CAT AAT GCA GGC TAT GGA ACA GCA AAA CAT CTT GCA GCT CTG      431
Phe Glu His Asn Ala Gly Tyr Gly Thr Ala Lys His Leu Ala Ala Leu
                    130                      135                      140

ACA AAA TAT GGT ATC ACA AGA ATT CAT CGG AAA TCC TAT GAA CCA ATT      479
Thr Lys Tyr Gly Ile Thr Arg Ile His Arg Lys Ser Tyr Glu Pro Ile
                    145                      150                      155

AAA TCG ATG GTC AAT TTC AAA TAG TAGATTATGT AAGTAAAAAA AGGAAAACGT      533
Lys Ser Met Val Asn Phe Lys ***
160                      165

GAACGATATT TGGTCACGTT TTTTGTCTGA CAAGTCTGTC AGTAATTATT TTCAAAGGTT      593

TCAAAAATAT AGTCTAGTAA TTTGCTAGAC TAATCTTCTG TTTTTTAATA ATAACTAAT      653

TTTTTGTTAA TCTAAATGAC AAAATTAATA AGCAGAGTTT TTTATAAAAT TAGCTACTTA      713

TAAAAAATTT GAAATGGTA TAGTTAAATC TGTTATAATT TCCAATATTT TTTAATAATA      773
    
```



ATTATTTTAA CAAAATACTT ATATCAAAAC TCTTTCAAAG TATATAATGA GCGTTGTATA	833
AGCTTTTATG TCTTTCTATA TCAACTTTTA ATAGAAATAT AAAGTAATAT AAATGTTTTT	893
ATAATAAATT ATGTGAGATA TATTTTTTTG TCCGTACTGG TATAGATTTG ACGATTAAGT	953
CTTAAATAAG TTATAATCTC AATTGCGTAA TTTCTTAAAT ACAGAAATAA CAACTACATT	1013
GGTAGACTGA TTAAAAAGTG TACTTGATGA ACTGTTATAA ACCTTAAAAA AATAAAAAATA	1073
ATAGTTTGGG GGATGTTAAA G ATG TAT AAA AAA TAT GGA GAT TGT TTT AAA	1124
Met Tyr Lys Lys Tyr Gly Asp Cys Phe Lys	5 10
AAG TTG CGA AAC CAA AAG AAT TTA GGG TTA TCA TAC TTT AGT AAA CTT	1172
Lys Leu Arg Asn Gln Lys Asn Leu Gly Leu Ser Tyr Phe Ser Lys Leu	15 20 25
GGA ATA GAC CGT TCA AAT ATA TCT AGA TTT GAA CAT GGA AAA TGT ATG	1220
Gly Ile Asp Arg Ser Asn Ile Ser Arg Phe Glu His Gly Lys Cys Met	30 35 40
ATG AGT TTT GAG CGT ATA GAT TTG ATG TTA GAA GAA ATG CAA GTT CCG	1268
Met Ser Phe Glu Arg Ile Asp Leu Met Leu Glu Glu Met Gln Val Pro	45 50 55
TTA TCT GAG TAC GAA TTG ATT GTA AAT AAT TTT ATG CCG AAT TTC CAA	1316
Leu Ser Glu Tyr Glu Leu Ile Val Asn Asn Phe Met Pro Asn Phe Gln	60 65 70
GAA TTT TTT ATA TTA GAA TTG GAA AAA GCT GAA TTT AGC CAA AAT CGA	1364
Glu Phe Phe Ile Leu Glu Leu Glu Lys Ala Glu Phe Ser Gln Asn Arg	75 80 85 90
GAT AAA ATA AAA GAG TTG TAT TCT GAG GTC AAA GAA ACG GGG AAT CAT	1412
Asp Lys Ile Lys Glu Leu Tyr Ser Glu Val Lys Glu Thr Gly Asn His	95 100 105
TTA CTG ACG GTT ACC GTG AAA ACG AAG CTT GGG AAT ATA AGT CAG ACA	1460
Leu Leu Thr Val Thr Val Lys Thr Lys Leu Gly Asn Ile Ser Gln Thr	110 115 120
GAA GTT AAA GAA ATT GAA GCT TAT CTT TGC AAT ATT GAA GAG TGG GGA	1508
Glu Val Lys Glu Ile Glu Ala Tyr Leu Cys Asn Ile Glu Glu Trp Gly	125 130 135
TAT TTT GAA CTT ACT TTA TTT TAT TTT GTA TCT GAT TAT CTC AAT GTC	1556
Tyr Phe Glu Leu Thr Leu Phe Tyr Phe Val Ser Asp Tyr Leu Asn Val	140 145 150
AAT CAA TTA GAA TTG CTG CTT TTT AAT TTT GAT AAA AGA TGT GAA AAT	1604
Asn Gln Leu Glu Leu Leu Phe Asn Phe Asp Lys Arg Cys Glu Asn	155 160 165 170
TAC TGT AGA GTC TTA AAA TAT AGA AGG AGA CTA TTG CAA ATA GCC TAT	1652
Tyr Cys Arg Val Leu Lys Tyr Arg Arg Arg Leu Leu Gln Ile Ala Tyr	175 180 185
AAA AGT GTT GCG ATA TAC GCG GCT AAA GGA GAA AGA AAA AAA GCC GAA	1700
Lys Ser Val Ala Ile Tyr Ala Ala Lys Gly Glu Arg Lys Lys Ala Glu	190 195 200
AAT ATT TTA GAA ATG ACT AAA AAA TAT CGA ACT GTG GGA GTC GAT TTA	1748
Asn Ile Leu Glu Met Thr Lys Lys Tyr Arg Thr Val Gly Val Asp Leu	205 210 215

TAT	TCA	GAA	GTA	TTA	AGA	CAT	CTT	GCT	AGA	GCT	ATC	ATT	ATT	TTT	AAT	1796
Tyr	Ser	Glu	Val	Leu	Arg	His	Leu	Ala	Arg	Ala	Ile	Ile	Ile	Phe	Asn	
	220					225					230					
TTT	GAA	AAT	GCA	GAG	ATT	GGG	GAA	GAA	AAA	ATA	AAT	TAT	GCT	CTT	GAG	1844
Phe	Glu	Asn	Ala	Glu	Ile	Gly	Glu	Glu	Lys	Ile	Asn	Tyr	Ala	Leu	Glu	
235					240					245					250	
ATT	TTG	GAA	GAA	TTT	GGA	GGA	AAG	AAG	ATA	AAA	GAA	TTC	TAT	CAG	AAT	1892
Ile	Leu	Glu	Glu	Phe	Gly	Gly	Lys	Lys	Ile	Lys	Glu	Phe	Tyr	Gln	Asn	
				255					260						265	
AAA	ATG	GAA	AAG	TAT	TTG	AAA	AGG	TCA	ATT	TAG	TCTCTTTTGA	GCTGTTGCTT				1945
Lys	Met	Glu	Tyr	Leu	Lys	Arg	Ser	Ile	***							
			270					275								
TAAAGCAACA	GCTCAAAGA	GATTTTCTTT	ATTCTAGAGC	ATATACTAGA	GGTGAAGAT											2005
AGGTTGTCTG	AAGCATTATA	ACTTGTCTTT	TAAAAAATTC	AATCATAAAT	ATAAGGAGGT											2065
ATG	ATG	AAT	CAA	AAA	AAA	TTA	TCA	TTA	TTC	GGT	TTT	TTC	GCA	TTA	ACC	2113
	Met	Asn	Gln	Lys	Lys	Leu	Ser	Leu	Phe	Gly	Phe	Phe	Ala	Leu	Thr	
					5					10					15	
GCT	TCA	ATG	GTT	TTG	ACT	GTC	TAT	GAG	TAT	CCG	ACT	TTT	GCC	ACG	TCA	2161
Ala	Ser	Met	Val	Leu	Thr	Val	Tyr	Glu	Tyr	Pro	Thr	Phe	Ala	Thr	Ser	
				20					25					30		
AAA	TTA	CAT	TTG	GTG	TTC	TTT	TTA	CTT	CTC	GGA	GGA	CTA	CTA	TGG	TTT	2209
Lys	Leu	His	Leu	Val	Phe	Phe	Leu	Leu	Leu	Gly	Gly	Leu	Leu	Trp	Phe	
			35					40					45			
TTG	CCT	GTA	GCG	CTC	TGC	GCA	GCA	GAA	ATG	GCG	ACG	GTT	GAA	GGC	TGG	2257
Leu	Pro	Val	Ala	Leu	Cys	Ala	Ala	Glu	Met	Ala	Thr	Val	Glu	Gly	Trp	
		50				55						60				
AAA	AAT	GGT	GGA	ATC	TTT	AGT	TGG	GTC	AGT	CAA	ACT	TTA	GGT	GAG	CGC	2305
Lys	Asn	Gly	Gly	Ile	Phe	Ser	Trp	Val	Ser	Gln	Thr	Leu	Gly	Glu	Arg	
	65					70					75					
TTT	GGT	TTT	GCA	GCC	ATA	TTT	TTT	CAG	TGG	TTC	CAA	ATT	ACA	GTA	GGT	2353
Phe	Gly	Phe	Ala	Ala	Ile	Phe	Phe	Gln	Trp	Phe	Gln	Ile	Thr	Val	Gly	
80					85					90					95	
TTT	GTC	ACT	ATG	ATC	TAT	TTC	ATT	TTA	GGG	GCC	CTC	TCT	TAT	GTG	TTA	2401
Phe	Val	Thr	Met	Ile	Tyr	Phe	Ile	Leu	Gly	Ala	Leu	Ser	Tyr	Val	Leu	
				100					105					110		
AAT	TTT	CAG	GCG	CTC	AAT	ACA	GAT	CCA	TTG	ATA	AAA	TTT	ATT	GGT	TTA	2449
Asn	Phe	Gln	Ala	Leu	Asn	Thr	Asp	Pro	Leu	Ile	Lys	Phe	Ile	Gly	Leu	
			115					120					125			
CTA	ATC	ATT	TTT	TGG	GGA	TTG	ACT	TTT	TCT	CAA	TTA	GGT	GGG	ACG	CAA	2497
Leu	Ile	Ile	Phe	Trp	Gly	Leu	Thr	Phe	Ser	Gln	Leu	Gly	Gly	Thr	Gln	
			130					135					140			
CGG	ACT	GCC	AAA	TTA	GTA	AAA	GCT	GGC	TTT	GTA	GTT	GGA	ATA	GTG	ATT	2545
Arg	Thr	Ala	Lys	Leu	Val	Lys	Ala	Gly	Phe	Val	Val	Gly	Ile	Val	Ile	
	145					150						155				
CCA	TCG	GTT	ATC	TTG	TTT	GGA	TTA	GCA	GCG	GCA	TAC	TTT	ATC	GGA	GGC	2593
Pro	Ser	Val	Ile	Leu	Phe	Gly	Leu	Ala	Ala	Ala	Tyr	Phe	Ile	Gly	Gly	
160					165					170					175	
AAT	CCT	ATA	GAA	ATA	CCA	ATT	AAC	AGC	CAT	GCT	TTT	GTA	CCA	GAT	TTT	2641
Asn	Pro	Ile	Glu	Ile	Pro	Ile	Asn	Ser	His	Ala	Phe	Val	Pro	Asp	Phe	
				180					185						190	

TCA Ser	CAG Gln	GTA Val	TCA Ser	ACT Thr	TTA Leu	GTA Val	GTT Val	TTT Phe	GTT Val	TCT Ser	TTT Phe	ATT Ile	CTG Leu	GCT Ala	TAT Tyr	2689
			195					200					205			
ATG Met	GGG Gly	GTA Val	GAA Glu	GCC Ala	TCA Ser	GCT Ala	TCA Ser	CAT His	ATT Ile	AAT Asn	GAA Glu	CTT Leu	GAA Glu	AAT Asn	CCA Pro	2737
		210					215					220				
AAA Lys	CGA Arg	AAT Asn	TAT Tyr	CCC Pro	TTA Leu	GCA Ala	ATG Met	ATT Ile	TTA Leu	TTA Leu	GTA Val	ATT Ile	TTG Leu	GCT Ala	ATT Ile	2785
	225					230					235					
TCT Ser	TTA Leu	GAT Asp	GCC Ala	ATA Ile	GGT Gly	GGA Gly	TTT Phe	TCT Ser	GTA Val	GCA Ala	GCA Ala	GTT Val	ATT Ile	CCT Pro	CAA Gln	2833
240					245					250					255	
AAA Lys	GAG Glu	TTA Leu	TCA Ser	TTA Leu	AGT Ser	GCA Ala	GGG Gly	GTA Val	ATC Ile	CAA Gln	ACT Thr	TTT Phe	CAA Gln	ACG Thr	TTA Leu	2881
				260					265					270		
ATC Ile	TTA Leu	CAT His	TTT Phe	AAT Asn	CAT His	CAT His	TTG Leu	GGA Gly	TGG Trp	TTA Leu	GTT Val	AAA Lys	GTG Val	ATT Ile	GCA Ala	2929
			275					280					285			
CTA Leu	ATG Met	ATT Ile	GCC Ala	TTT Phe	GGG Gly	GTT Val	ATG Met	GGA Gly	GAA Glu	GTG Val	AGT Ser	TCA Ser	TGG Trp	GTT Val	GTT Val	2977
		290				295						300				
GGT Gly	CCT Pro	TCT Ser	AGA Arg	GGG Gly	ATG Met	TTT Phe	GCA Ala	GCA Ala	GCA Ala	CAA Gln	AGA Arg	GGT Gly	TTA Leu	TTA Leu	CCA Pro	3025
	305					310					315					
AAA Lys	TTT Phe	TTA Leu	CGT Arg	AAA Lys	ACG Thr	AAT Asn	ACA Thr	CAT His	GAA Glu	GTC Val	CCT Pro	GTT Val	CCT Pro	TTA Leu	GTT Val	3073
				320		325				330					335	
ATG Met	ATT Ile	CAA Gln	GGA Gly	ATC Ile	ATT Ile	GTT Val	ACA Thr	CTT Leu	TGG Trp	GGC Gly	GCT Ala	GTA Val	TTA Leu	ACT Thr	TTT Phe	3121
				340					345					350		
GGA Gly	GGA Gly	GGA Gly	GGA Gly	AAT Asn	AAT Asn	TTA Leu	TCT Ser	TTC Phe	TTA Leu	GTT Val	GCC Ala	ATT Ile	TCA Ser	CTG Leu	ACT Thr	3169
			355					360					365			
GTA Val	GTG Val	ATT Ile	TAT Tyr	TTG Leu	GTG Val	GGT Gly	TAC Tyr	CTC Leu	TTG Leu	TTC Phe	TTT Phe	ATT Ile	GTT Val	TAC Tyr	TTT Phe	3217
		370					375					380				
GTT Val	TTA Leu	ATC Ile	TAT Tyr	AAA Lys	AAA Lys	CAA Gln	AAT Asn	TTA Leu	AAG Lys	CGT Arg	ACT Thr	TAT Tyr	AAT Asn	GTT Val	CCA Pro	3265
		385				390					395					
GGT Gly	AAA Lys	ATA Ile	ATA Ile	GGA Gly	AAA Lys	ACA Thr	ATC Ile	ATT Ile	GCA Ala	GGA Gly	ATT Ile	GGA Gly	TTC Phe	TTA Leu	TTA Leu	3313
400				405						410					415	
TCA Ser	ATT Ile	TTT Phe	GCT Ala	CTA Leu	TTT Phe	ATT Ile	TCC Ser	TTT Phe	GTT Val	CCT Pro	CCA Pro	GCA Ala	TCA Ser	ATT Ile	GCG Ala	3361
				420					425					430		
AAA Lys	AAT Asn	GAA Glu	ACT Thr	CAC His	ACC Thr	TAT Tyr	CAA Gln	ATG Met	ATA Ile	CTT Leu	CTT Leu	ATA Ile	AGT Ser	TTT Phe	GTT Val	3409
			435					440					445			
GTG Val	ACC Thr	GCT Ala	ATC Ile	TTG Leu	CCA Pro	TTT Phe	ATT Ile	ATT Ile	TAT Tyr	GAA Glu	TTG Leu	CAT His	GAT Asp	AAA Lys	AAG Lys	3457
		450					455					460				

GGA	CAT	GAT	ACT	ATT	GAA	GAA	CCA	ACA	CAC	TTT	AAA	GCA	GGA	GAT	GTG	3505
Gly	His	Asp	Thr	Ile	Glu	Glu	Pro	Thr	His	Phe	Lys	Ala	Gly	Asp	Val	
	465					470					475					
AAC	CCT	GCG	ATT	TAT	CCA	GCA	GCT	CGT	GGA	GAG	CAT	CAT	ATT	ATT	AAA	3553
Asn	Pro	Ala	Ile	Tyr	Pro	Ala	Ala	Arg	Gly	Glu	His	His	Ile	Ile	Lys	
480					485				490						495	
AAA	GAA	GAA	CAT	ATC	TTA	AAA	CAT	TGA	AAAATTggag	gATGTACAT	ATG					3602
Lys	Glu	Glu	His	Ile	Leu	Lys	His	***			Met					
				500			503									
TTA	TAC	GGA	AAA	GAA	AAT	CGA	GAT	GAA	GCG	GAG	TTC	TTG	GAA	CCA	ATT	3650
Leu	Tyr	Gly	Lys	Glu	Asn	Arg	Asp	Glu	Ala	Glu	Phe	Leu	Glu	Pro	Ile	
			5					10					15			
TTT	GGT	TCA	GAA	AGT	GAA	CAA	GTG	GAT	TTA	CCT	AAA	TAT	AAA	TTA	GCT	3698
Phe	Gly	Ser	Glu	Ser	Glu	Gln	Val	Asp	Leu	Pro	Lys	Tyr	Lys	Leu	Ala	
		20					25					30				
CAA	CAA	TCA	ATT	GAG	CCT	CGA	GTG	GCC	TAT	CAG	TTA	GTT	CAA	GAT	GAA	3746
Gln	Gln	Ser	Ile	Glu	Pro	Arg	Val	Ala	Tyr	Gln	Leu	Val	Gln	Asp	Glu	
		35				40					45					
ATG	CTA	GAT	GAA	GGG	AAC	GCT	CGT	TTA	AAT	TTG	GCC	ACA	TTC	TGT	CAA	3794
Met	Leu	Asp	Glu	Gly	Asn	Ala	Arg	Leu	Asn	Leu	Ala	Thr	Phe	Cys	Gln	
	50				55					60					65	
ACT	TAT	ATG	GAA	CCT	GAA	GCA	GTC	AAG	CTG	ATG	AGT	CAG	ACC	TTG	GAA	3842
Thr	Tyr	Met	Glu	Pro	Glu	Ala	Val	Lys	Leu	Met	Ser	Gln	Thr	Leu	Glu	
				70					75					80		
AAA	AAT	GCG	ATT	GAC	AAA	TCA	GAA	TAT	CCA	AGA	ACA	ACT	GAA	ATT	GAA	3890
Lys	Asn	Ala	Ile	Asp	Lys	Ser	Glu	Tyr	Pro	Arg	Thr	Thr	Glu	Ile	Glu	
			85					90					95			
AAC	CGT	TGC	GTC	AAC	ATG	ATC	GCT	GAC	CTT	TGG	AAT	GCG	AGT	GAA	AAA	3938
Asn	Arg	Cys	Val	Asn	Met	Ile	Ala	Asp	Leu	Trp	Asn	Ala	Ser	Glu	Lys	
		100				105						110				
GGA	AAA	ATT	TAT	GGG	ACT	TCG	ACA	ATT	GGT	TCT	TCA	GAA	GCT	TGT	ATG	3986
Gly	Lys	Ile	Tyr	Gly	Thr	Ser	Thr	Ile	Gly	Ser	Ser	Glu	Ala	Cys	Met	
	115					120					125					
CTT	GGG	GGA	ATG	GCT	ATG	AAG	TTT	TCT	TGG	CGT	AAG	CGA	GCA	GAA	AAA	4034
Leu	Gly	Gly	Met	Ala	Met	Lys	Phe	Ser	Trp	Arg	Lys	Arg	Ala	Glu	Lys	
	130				135					140					145	
TTA	GGC	CTA	GAT	ATT	AAT	GCG	AAA	AAG	CCA	AAC	TTA	GTC	ATT	TCC	TCT	4082
Leu	Gly	Leu	Asp	Ile	Asn	Ala	Lys	Lys	Pro	Asn	Leu	Val	Ile	Ser	Ser	
				150					155					160		
GGT	TAT	CAA	GTT	TGC	TGG	GAA	AAA	TTC	TGT	GTT	TAT	TGG	GAT	ATT	GAA	4130
Gly	Tyr	Gln	Val	Cys	Trp	Glu	Lys	Phe	Cys	Val	Tyr	Trp	Asp	Ile	Glu	
			165					170				175				
ATG	AGA	GAA	GTG	CCA	ATG	GAT	AGA	GAA	CAT	ATG	TCA	ATC	AAT	TTG	GAA	4178
Met	Arg	Glu	Val	Pro	Met	Asp	Arg	Glu	His	Met	Ser	Ile	Asn	Leu	Glu	
		180				185						190				
AAA	GTG	ATG	GAT	TAT	GTT	GAT	GAA	TAT	ACG	ATT	GGA	GTA	GTT	GGA	ATT	4226
Lys	Val	Met	Asp	Tyr	Val	Asp	Glu	Tyr	Thr	Ile	Gly	Val	Val	Gly	Ile	
		195				200					205					
ATG	GGG	ATT	ACT	TAT	ACT	GGT	CGT	TAT	GAT	GAT	ATC	AAA	GCT	TTG	GAT	4274
Met	Gly	Ile	Thr	Tyr	Thr	Gly	Arg	Tyr	Asp	Asp	Ile	Lys	Ala	Leu	Asp	
	210				215					220					225	

AAT	TTG	ATT	GAA	GAA	TAT	AAT	AAA	CAG	ACA	GAC	TAC	AAA	GTT	TAT	ATT	4322
Asn	Leu	Ile	Glu	Glu	Tyr	Asn	Lys	Gln	Thr	Asp	Tyr	Lys	Val	Tyr	Ile	241
			230						235				240			
CAC	GTA	GAT	GCT	GCT	TCA	GGA	GGA	CTT	TAT	GCT	CCT	TTT	GTT	GAG	CCA	4370
His	Val	Asp	Ala	Ala	Ser	Gly	Gly	Leu	Tyr	Ala	Pro	Phe	Val	Glu	Pro	
			245					250					255			
GAA	CTT	GAG	TGG	GAT	TTC	CGT	TTG	AAA	AAT	GTC	ATT	TCA	ATC	AAT	ACT	4418
Glu	Leu	Glu	Trp	Asp	Phe	Arg	Leu	Lys	Asn	Val	Ile	Ser	Ile	Asn	Thr	
		260					265					270				
TCA	GGA	CAT	AAA	TAT	GGT	TTA	GTA	TAT	CCT	GGT	GTA	GGT	TGG	GTC	TTG	4466
Ser	Gly	His	Lys	Tyr	Gly	Leu	Val	Tyr	Pro	Gly	Val	Gly	Trp	Val	Leu	
	275					280					285					
TGG	CGT	GAC	AAA	AAA	TAT	TTA	CCT	GAA	GAG	TTA	ATT	TTT	AAA	GTA	AGT	4514
Trp	Arg	Asp	Lys	Lys	Tyr	Leu	Pro	Glu	Glu	Leu	Ile	Phe	Lys	Val	Ser	
290					295					300					305	
TAT	CTT	GGA	GGA	GAA	TTA	CCA	ACA	ATG	GCG	ATT	AAT	TTT	TCT	CAC	AGT	4562
Tyr	Leu	Gly	Gly	Glu	Leu	Pro	Thr	Met	Ala	Ile	Asn	Phe	Ser	His	Ser	
				310					315					320		
GCT	TCT	CAA	TTA	ATC	GGT	CAA	TAC	TAT	AAT	TTT	GTA	CGT	TAT	GGA	TTT	4610
Ala	Ser	Gln	Leu	Ile	Gly	Gln	Tyr	Tyr	Asn	Phe	Val	Arg	Tyr	Gly	Phe	
			325					330					335			
GAT	GGA	TAT	AAA	GCT	ATT	CAT	GAG	AGA	ACG	CAT	AAA	GTA	GCC	ATG	TAT	4658
Asp	Gly	Tyr	Lys	Ala	Ile	His	Glu	Arg	Thr	His	Lys	Val	Ala	Met	Tyr	
		340					345					350				
TTA	GCA	GAA	GAA	ATT	GAA	AAA	ACA	GGA	ATG	TTT	GAG	ATT	ATG	AAC	GAT	4706
Leu	Ala	Glu	Glu	Ile	Glu	Lys	Thr	Gly	Met	Phe	Glu	Ile	Met	Asn	Asp	
	355					360					365					
GGG	GCA	CAA	TTA	CCA	ATT	GTC	TGC	TAC	AAA	TTA	AAA	GAA	AAT	TCA	AAC	4754
Gly	Ala	Gln	Leu	Pro	Ile	Val	Cys	Tyr	Lys	Leu	Lys	Glu	Asn	Ser	Asn	
370				375						380					385	
CGT	GGT	TGG	AAT	CTT	TAT	GAT	TTG	GCA	GAT	CGT	TTA	TTA	ATG	AAG	GGA	4802
Arg	Gly	Trp	Asn	Leu	Tyr	Asp	Leu	Ala	Asp	Arg	Leu	Leu	Met	Lys	Gly	
				390					395					400		
TGG	CAA	GTG	CCT	GCT	TAT	CCA	CTT	CCT	AAA	AAT	TTG	GAA	AAT	GAA	ATC	4850
Trp	Gln	Val	Pro	Ala	Tyr	Pro	Leu	Pro	Lys	Asn	Leu	Glu	Asn	Glu	Ile	
			405					410					415			
ATT	CAA	CGT	TTA	GTA	ATT	CGA	GCA	GAT	TTC	GGG	ATG	AAT	ATG	GCA	TTT	4898
Ile	Gln	Arg	Leu	Val	Ile	Arg	Ala	Asp	Phe	Gly	Met	Asn	Met	Ala	Phe	
		420					425					430				
AAC	TAT	GTT	CAA	GAT	ATG	CAA	GAA	GCA	ATT	GAT	GCA	CTA	AAT	AAG	GCT	4946
Asn	Tyr	Val	Gln	Asp	Met	Gln	Glu	Ala	Ile	Asp	Ala	Leu	Asn	Lys	Ala	
	435					440					445					
CAT	ATT	CTA	TTT	CAT	CAG	GAA	CCT	GAA	AAT	AAA	ACA	TAT	GGC	TTT	ACT	4994
His	Ile	Leu	Phe	His	Gln	Glu	Pro	Glu	Asn	Lys	Thr	Tyr	Gly	Phe	Thr	
450				455						460					465	
CAC	TAA	AGATAAAAGC	GATATATCTA	AGATATATCG	CTTTTATTTT	GTTTTAGGCT										5050
His																
ATTTACTAAT	TAGCTTGTCG	CTTA	TTA	TTT	TTC	ATA	GTA	TTT	ATC	CAA	AAT					5101
			***	Lys	Glu	Tyr	Tyr	Lys	Asp	Leu	Ile					

TTC CAT TTT TAA AGG AGT AAT TTT AGA TAG GGG GGC AGT TAG ACT TGT	5149
Glu Met Lys Leu Pro Thr Ile Lys Ser Leu Pro Ala Thr Leu Ser Thr	
-10 -15 -20	
TCT TAG GAA GAG CTT ATC TTC AAT GTT GAT GAT ACC CAG ATA TTT AAC	5197
Arg Leu Phe Leu Lys Asp Glu Ile Asn Ile Ile Gly Leu Tyr Lys Val	
-25 -30 -35 -40	
TTG AGG GTA GTT AGC ATT GAC TTC TAT AAT TTG GGC TTT TTT CTC ACT	5245
Gln Pro Tyr Asn Ala Asn Val Glu Ile Ile Gln Ala Lys Lys Glu Ser	
-45 -50 -55	
AAT ATT TTC GTC TGT CAC GGG CAC ATC TAG GTT GAC CGT TCT TTC TTT	5293
Ile Asn Glu Asp Thr Val Pro Val Asp Leu Asn Val Thr Arg Glu Lys	
-60 -65 -70	
ATA AGA GTA ATT TTT GAG AGC AGC AAT ATT TCG GTT TGG AAT AAA AGT	5341
Tyr Ser Tyr Asn Lys Leu Ala Ala Ile Asn Arg Asn Pro Ile Phe Thr	
-75 -80 -85	
AGT CGC ACC GTC GGC TCC GAT AAC AGT AAT GGA ACG AAT TCC TAC GGT	5389
Thr Ala Gly Asp Ala Gly Ile Val Thr Ile Ser Arg Ile Gly Val Thr	
-90 -95 -100	
TTT CAC CAC TCC CTC AAT ATC AAG ACC GGC AAA GGC AAC CGT ATC CGC	5437
Lys Val Val Gly Glu Ile Asp Leu Gly Ala Phe Ala Val Thr Asp Ala	
-105 -110 -115 -120	
AAC ATT GAT TTG ATG TTC AAC GAT AAT AAA GAA ACC ATT AAT GAT ATC	5485
Val Asn Ile Gln His Glu Val Ile Ile Phe Phe Gly Asn Ile Ile Asp	
-125 -130 -135	
AGC GAC TAA ATC TCG TCC AGC GAA ACC AAG AGC GAC TCC AAG GAT ACC	5533
Ala Val Leu Asp Arg Gly Ala Phe Gly Leu Ala Val Gly Leu Ile Gly	
-140 -145 -150	
CGC GCC AGC TAG GAC ATT TGC AAC AGG AATTC	5565
Ala Gly Ala Leu Val Asn Ala Val Pro Leu	
-155 -160	

(2) INFORMATION FOR SEQ ID NO: 22:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 166 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein, C-terminus of rnhB ex *L. lactis* MG1363

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

Ala Val Ile Leu Pro Lys Asn Cys Lys Ile Arg Gly Leu Asn Asp Ser  
 5 10 15  
 Lys Lys Val Pro Lys Ser Lys His His Ala Ile Leu Ser Glu Ile Gln  
 20 25 30  
 Glu Lys Ala Leu Ala Ile Gly Val Gly Ile Val Asp Ala Glu Lys Ile  
 35 40 45  
 Asp Glu Val Asn Ile Tyr Glu Ala Thr Lys Ile Ala Met Ile Gln Ala  
 50 55 60  
 Val Ser Lys Leu Ser Leu Lys Pro Glu His Leu Leu Ile Asp Ala Met  
 65 70 75 80  
 Val Leu Asp Leu Pro Ile Ala Gln Thr Lys Ile Ile His Gly Asp Ala  
 85 90 95

Arg Ser Ala Ser Ile Ala Ala Ala Ser Ile Val Ala Lys Val Thr Arg  
 100 105 110  
 Asp Glu Met Met Lys Asp Phe Ala Leu Glu Phe Pro Glu Tyr Asp Phe  
 115 120 125  
 Glu His Asn Ala Gly Tyr Gly Thr Ala Lys His Leu Ala Ala Leu Thr  
 130 135 140  
 Lys Tyr Gly Ile Thr Arg Ile His Arg Lys Ser Tyr Glu Pro Ile Lys  
 145 150 155 160  
 Ser Met Val Asn Phe Lys  
 165

## (2) INFORMATION FOR SEQ ID NO: 23:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 276 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein, rggL = gadR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

Met Tyr Lys Lys Tyr Gly Asp Cys Phe Lys Lys Leu Arg Asn Gln Lys  
 5 10 15  
 Asn Leu Gly Leu Ser Tyr Phe Ser Lys Leu Gly Ile Asp Arg Ser Asn  
 20 25 30  
 Ile Ser Arg Phe Glu His Gly Lys Cys Met Met Ser Phe Glu Arg Ile  
 35 40 45  
 Asp Leu Met Leu Glu Glu Met Gln Val Pro Leu Ser Glu Tyr Glu Leu  
 50 55 60  
 Ile Val Asn Asn Phe Met Pro Asn Phe Gln Glu Phe Phe Ile Leu Glu  
 65 70 75 80  
 Leu Glu Lys Ala Glu Phe Ser Gln Asn Arg Asp Lys Ile Lys Glu Leu  
 85 90 95  
 Tyr Ser Glu Val Lys Glu Thr Gly Asn His Leu Leu Thr Val Thr Val  
 100 105 110  
 Lys Thr Lys Leu Gly Asn Ile Ser Gln Thr Glu Val Lys Glu Ile Glu  
 115 120 125  
 Ala Tyr Leu Cys Asn Ile Glu Glu Trp Gly Tyr Phe Glu Leu Thr Leu  
 130 135 140  
 Phe Tyr Phe Val Ser Asp Tyr Leu Asn Val Asn Gln Leu Glu Leu Leu  
 145 150 155 160  
 Leu Phe Asn Phe Asp Lys Arg Cys Glu Asn Tyr Cys Arg Val Leu Lys  
 165 170 175  
 Tyr Arg Arg Arg Leu Leu Gln Ile Ala Tyr Lys Ser Val Ala Ile Tyr  
 180 185 190  
 Ala Ala Lys Gly Glu Arg Lys Lys Ala Glu Asn Ile Leu Glu Met Thr  
 195 200 205  
 Lys Lys Tyr Arg Thr Val Gly Val Asp Leu Tyr Ser Glu Val Leu Arg  
 210 215 220

His Leu Ala Arg Ala Ile Ile Ile Phe Asn Phe Glu Asn Ala Glu Ile  
 225 230 235 240

Gly Glu Glu Lys Ile Asn Tyr Ala Leu Glu Ile Leu Glu Glu Phe Gly  
 245 250 255

Gly Lys Lys Ile Lys Glu Phe Tyr Gln Asn Lys Met Glu Lys Tyr Leu  
 260 265 270

Lys Arg Ser Ile \*\*\*  
 275

## (2) INFORMATION FOR SEQ ID NO: 24:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 503 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein, orfX = gadC

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

Met Asn Gln Lys Lys Leu Ser Leu Phe Gly Phe Phe Ala Leu Thr Ala  
 5 10 15

Ser Met Val Leu Thr Val Tyr Glu Tyr Pro Thr Phe Ala Thr Ser Lys  
 20 25 30

Leu His Leu Val Phe Phe Leu Leu Leu Gly Gly Leu Leu Trp Phe Leu  
 35 40 45

Pro Val Ala Leu Cys Ala Ala Glu Met Ala Thr Val Glu Gly Trp Lys  
 50 55 60

Asn Gly Gly Ile Phe Ser Trp Val Ser Gln Thr Leu Gly Glu Arg Phe  
 65 70 75 80

Gly Phe Ala Ala Ile Phe Phe Gln Trp Phe Gln Ile Thr Val Gly Phe  
 85 90 95

Val Thr Met Ile Tyr Phe Ile Leu Gly Ala Leu Ser Tyr Val Leu Asn  
 100 105 110

Phe Gln Ala Leu Asn Thr Asp Pro Leu Ile Lys Phe Ile Gly Leu Leu  
 115 120 125

Ile Ile Phe Trp Gly Leu Thr Phe Ser Gln Leu Gly Gly Thr Gln Arg  
 130 135 140

Thr Ala Lys Leu Val Lys Ala Gly Phe Val Val Gly Ile Val Ile Pro  
 145 150 155 160

Ser Val Ile Leu Phe Gly Leu Ala Ala Ala Tyr Phe Ile Gly Gly Asn  
 165 170 175

Pro Ile Glu Ile Pro Ile Asn Ser His Ala Phe Val Pro Asp Phe Ser  
 180 185 190

Gln Val Ser Thr Leu Val Val Phe Val Ser Phe Ile Leu Ala Tyr Met  
 195 200 205

Gly Val Glu Ala Ser Ala Ser His Ile Asn Glu Leu Glu Asn Pro Lys  
 210 215 220

Arg Asn Tyr Pro Leu Ala Met Ile Leu Leu Val Ile Leu Ala Ile Ser  
 225 230 235 240



```

Leu Asp Ala Ile Gly Gly Phe Ser Val Ala Ala Val Ile Pro Gln Lys
      245                250                255
Glu Leu Ser Leu Ser Ala Gly Val Ile Gln Thr Phe Gln Thr Leu Ile
      260                265                270
Leu His Phe Asn His His Leu Gly Trp Leu Val Lys Val Ile Ala Leu
      275                280                285
Met Ile Ala Phe Gly Val Met Gly Glu Val Ser Ser Trp Val Val Gly
      290                295                300
Pro Ser Arg Gly Met Phe Ala Ala Ala Gln Arg Gly Leu Leu Pro Lys
      305                310                315                320
Phe Leu Arg Lys Thr Asn Thr His Glu Val Pro Val Pro Leu Val Met
      325                330                335
Ile Gln Gly Ile Ile Val Thr Leu Trp Gly Ala Val Leu Thr Phe Gly
      340                345                350
Gly Gly Gly Asn Asn Leu Ser Phe Leu Val Ala Ile Ser Leu Thr Val
      355                360                365
Val Ile Tyr Leu Val Gly Tyr Leu Leu Phe Phe Ile Val Tyr Phe Val
      370                375                380
Leu Ile Tyr Lys Lys Gln Asn Leu Lys Arg Thr Tyr Asn Val Pro Gly
      385                390                395                400
Lys Ile Ile Gly Lys Thr Ile Ile Ala Gly Ile Gly Phe Leu Leu Ser
      405                410                415
Ile Phe Ala Leu Phe Ile Ser Phe Val Pro Pro Ala Ser Ile Ala Lys
      420                425                430
Asn Glu Thr His Thr Tyr Gln Met Ile Leu Leu Ile Ser Phe Val Val
      435                440                445
Thr Ala Ile Leu Pro Phe Ile Ile Tyr Glu Leu His Asp Lys Lys Gly
      450                455                460
His Asp Thr Ile Glu Glu Pro Thr His Phe Lys Ala Gly Asp Val Asn
      465                470                475                480
Pro Ala Ile Tyr Pro Ala Ala Arg Gly Glu His His Ile Ile Lys Lys
      485                490                495
Glu Glu His Ile Leu Lys His
      500                503

```

## (2) INFORMATION FOR SEQ ID NO: 25:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 466 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein, gadB

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

```

Met Leu Tyr Gly Lys Glu Asn Arg Asp Glu Ala Glu Phe Leu Glu Pro
      5                10                15
Ile Phe Gly Ser Glu Ser Glu Gln Val Asp Leu Pro Lys Tyr Lys Leu
      20                25                30

```

Ala Gln Gln Ser Ile Glu Pro Arg Val Ala Tyr Gln Leu Val Gln Asp  
35 40 45

Glu Met Leu Asp Glu Gly Asn Ala Arg Leu Asn Leu Ala Thr Phe Cys  
50 55 60

Gln Thr Tyr Met Glu Pro Glu Ala Val Lys Leu Met Ser Gln Thr Leu  
65 70 75 80

Glu Lys Asn Ala Ile Asp Lys Ser Glu Tyr Pro Arg Thr Thr Glu Ile  
85 90 95

Glu Asn Arg Cys Val Asn Met Ile Ala Asp Leu Trp Asn Ala Ser Glu  
100 105 110

Lys Gly Lys Ile Tyr Gly Thr Ser Thr Ile Gly Ser Ser Glu Ala Cys  
115 120 125

Met Leu Gly Gly Met Ala Met Lys Phe Ser Trp Arg Lys Arg Ala Glu  
130 135 140

Lys Leu Gly Leu Asp Ile Asn Ala Lys Lys Pro Asn Leu Val Ile Ser  
145 150 155 160

Ser Gly Tyr Gln Val Cys Trp Glu Lys Phe Cys Val Tyr Trp Asp Ile  
165 170 175

Glu Met Arg Glu Val Pro Met Asp Arg Glu His Met Ser Ile Asn Leu  
180 185 190

Glu Lys Val Met Asp Tyr Val Asp Glu Tyr Thr Ile Gly Val Val Gly  
195 200 205

Ile Met Gly Ile Thr Tyr Thr Gly Arg Tyr Asp Asp Ile Lys Ala Leu  
210 215 220

Asp Asn Leu Ile Glu Glu Tyr Asn Lys Gln Thr Asp Tyr Lys Val Tyr  
225 230 235 240

Ile His Val Asp Ala Ala Ser Gly Gly Leu Tyr Ala Pro Phe Val Glu  
245 250 255

Pro Glu Leu Glu Trp Asp Phe Arg Leu Lys Asn Val Ile Ser Ile Asn  
260 265 270

Thr Ser Gly His Lys Tyr Gly Leu Val Tyr Pro Gly Val Gly Trp Val  
275 280 285

Leu Trp Arg Asp Lys Lys Tyr Leu Pro Glu Glu Leu Ile Phe Lys Val  
290 295 300

Ser Tyr Leu Gly Gly Glu Leu Pro Thr Met Ala Ile Asn Phe Ser His  
305 310 315 320

Ser Ala Ser Gln Leu Ile Gly Gln Tyr Tyr Asn Phe Val Arg Tyr Gly  
325 330 335

Phe Asp Gly Tyr Lys Ala Ile His Glu Arg Thr His Lys Val Ala Met  
340 345 350

Tyr Leu Ala Glu Glu Ile Glu Lys Thr Gly Met Phe Glu Ile Met Asn  
355 360 365

Asp Gly Ala Gln Leu Pro Ile Val Cys Tyr Lys Leu Lys Glu Asn Ser  
370 375 380

Asn Arg Gly Trp Asn Leu Tyr Asp Leu Ala Asp Arg Leu Leu Met Lys  
385 390 395 400

Gly Trp Gln Val Pro Ala Tyr Pro Leu Pro Lys Asn Leu Glu Asn Glu  
405 410 415

Ile Ile Gln Arg Leu Val Ile Arg Ala Asp Phe Gly Met Asn Met Ala  
420 425 430

Phe Asn Tyr Val Gln Asp Met Gln Glu Ala Ile Asp Ala Leu Asn Lys  
435 440 445

Ala His Ile Leu Phe His Gln Glu Pro Glu Asn Lys Thr Tyr Gly Phe  
450 455 460

Thr His  
465

## C L A I M S

1. A salt-inducible promoter **derivable from a lactic acid bacterium**, in isolation from the coding sequence which is normally controlled by said promoter in a wild-type lactic acid bacterium.
2. An essential part of a salt-inducible promoter according to claim 1, which comprises the polynucleotide 1482-1925 of SEQ. ID. NO: 10.
3. An essential part according to claim 2, which comprises the polynucleotide 1482-2068 of SEQ. ID. NO: 10.
4. A salt-inducible promoter according to claim 1, which comprises the polynucleotide 1-2068 of SEQ. ID. NO: 10.
5. A salt-inducible promoter according to claim 4, which additionally comprises part of the ORF X gene together forming polynucleotide 1-2426 of SEQ. ID. NO: 10.
6. A modification of a salt-inducible promoter according to claim 1 or an essential part thereof, which comprises a DNA sequence essentially corresponding to a polynucleotide selected from the group consisting of
  - (a) polynucleotide 1482-1925 of SEQ. ID. NO: 10 followed by a promoter functional in a lactic acid bacterium,
  - (b) polynucleotide 1482-2068 of SEQ. ID. NO: 10,
  - (c) polynucleotide 1-2068 of SEQ. ID. NO: 10, and
  - (d) polynucleotide 1-2426 of SEQ. ID. NO: 10.
7. A DNA fragment capable of regulating a salt-inducible promoter active in a lactic acid bacterium, which comprises the polynucleotide 1095-1925 of SEQ. ID. NO: 10, or a modification thereof that
  - (a) encodes the same polypeptide as said polynucleotide 1095-1925, or

(b) encodes a modification of such polypeptide still having essentially the same regulating capacity.

8. A recombinant vector comprising a salt-inducible promoter or an essential part thereof as claimed in any one of claims 1-6, or a DNA fragment as claimed in claim 7 in combination with a DNA fragment selected from the group consisting of a DNA fragment containing the polynucleotide 1926-2000 of SEQ. ID. NO: 10 and modifications thereof still having essentially the same promoting capacity.

9. A transformed lactic acid bacterium comprising a salt-inducible promoter or an essential part thereof as claimed in any one of claims 1-6, or a DNA fragment as claimed in claim 7 in combination with a DNA fragment selected from the group consisting of a DNA fragment containing the polynucleotide 1926-2000 of SEQ. ID. NO: 10 and modifications thereof still having essentially the same promoting capacity.

10. A process for the production of a desired protein by a transformed lactic acid bacterium, whereby the gene encoding said desired protein or a precursor thereof is expressed under control of an inducible promoter, **characterised in that** the promoter is a salt-inducible promoter or an essential part thereof as claimed in any one of claims 1-6, or a DNA fragment as claimed in claim 7 in combination with a DNA fragment selected from the group consisting of a DNA fragment containing the polynucleotide 1926-2000 of SEQ. ID. NO: 10 and modifications thereof still having essentially the same promoting capacity.

11. A process according to claim 10, in which the desired protein is secreted by the lactic acid bacterium due to the presence of a DNA fragment fused to the gene encoding the desired protein and effecting secretion of the desired protein or a precursor thereof.

12. A process according to claim 10, in which the action of the salt-inducible promoter is enhanced by allowing the pH of the medium in which the lactic acid bacterium is cultured to decrease to a value of about 4 to about 4.5.

5

13. A process according to anyone of claims 10-12, in which the action of the salt-inducible promoter is enhanced by incorporating a sufficient amount of glutamate or glutamic acid, preferably about 50 mM, in the medium in which the  
10 lactic acid bacterium is cultured.

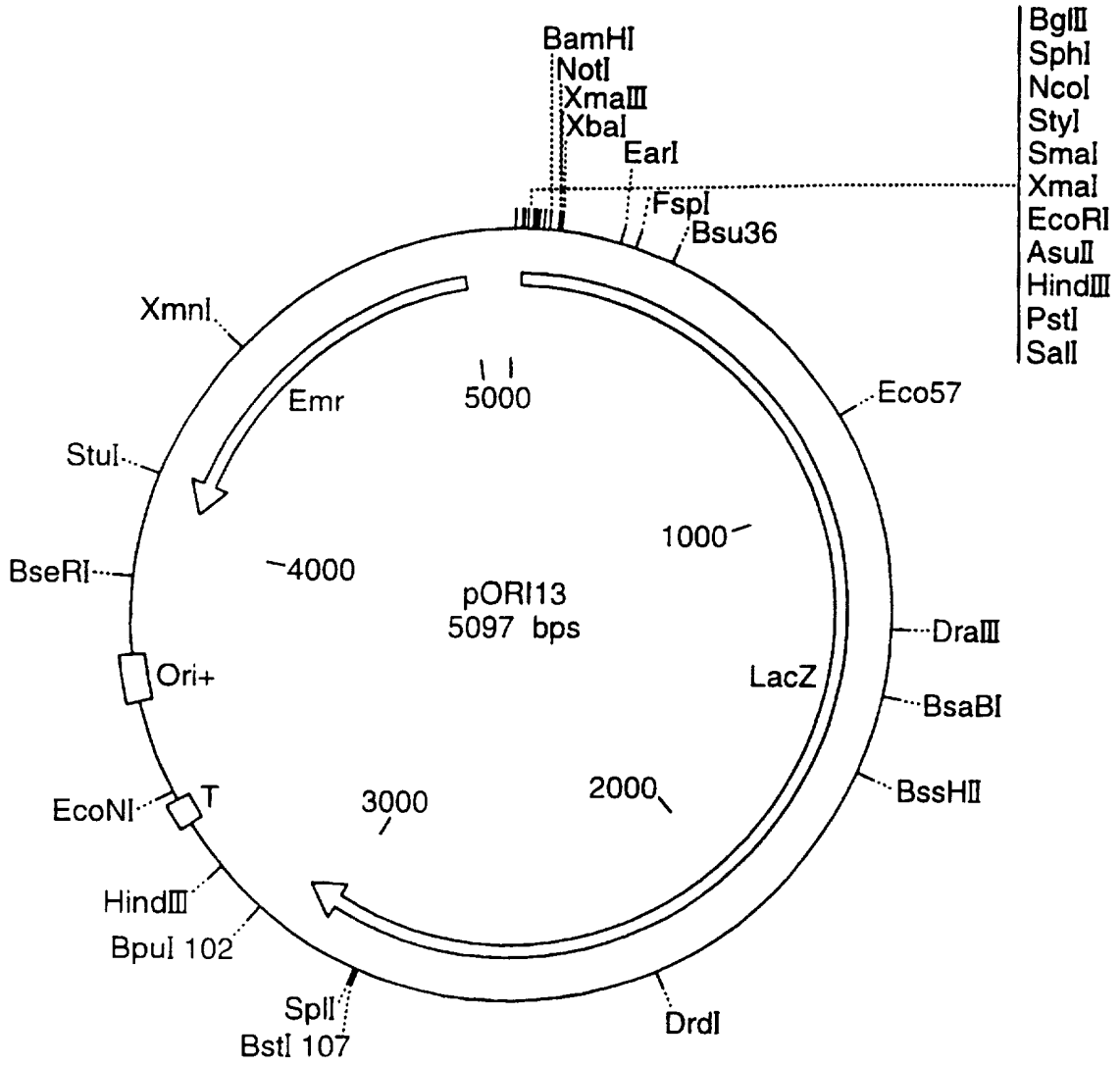
14. Use of a process as claimed in any one of claims 10-13 in a fermentation process, in which the desired protein is a lytic protein causing lysis of the bacterial cells so that  
15 the contents of the cells can be released.

15. Use of a process as claimed in any one of claims 10-13 in a fermentation process, in which the desired protein is an enzyme involved in flavour formation.

20

16. Use of a process as claimed in any one of claims 10-13 in a fermentation process, in which the desired protein is a protein having a function in a cheese production process, such as chymosin or a precursor thereof, or an enzyme  
25 involved in cheese flavour formation.

Fig. 1.



...MCS...XbaI RBS Met Lys Gly  
 ...TCTAGATTCGGAGGAATTTTGAA ATG AAA GGG.. lacZ  
 \*\*\* stop \*\*\* stop \*\*\*

Fig.2.

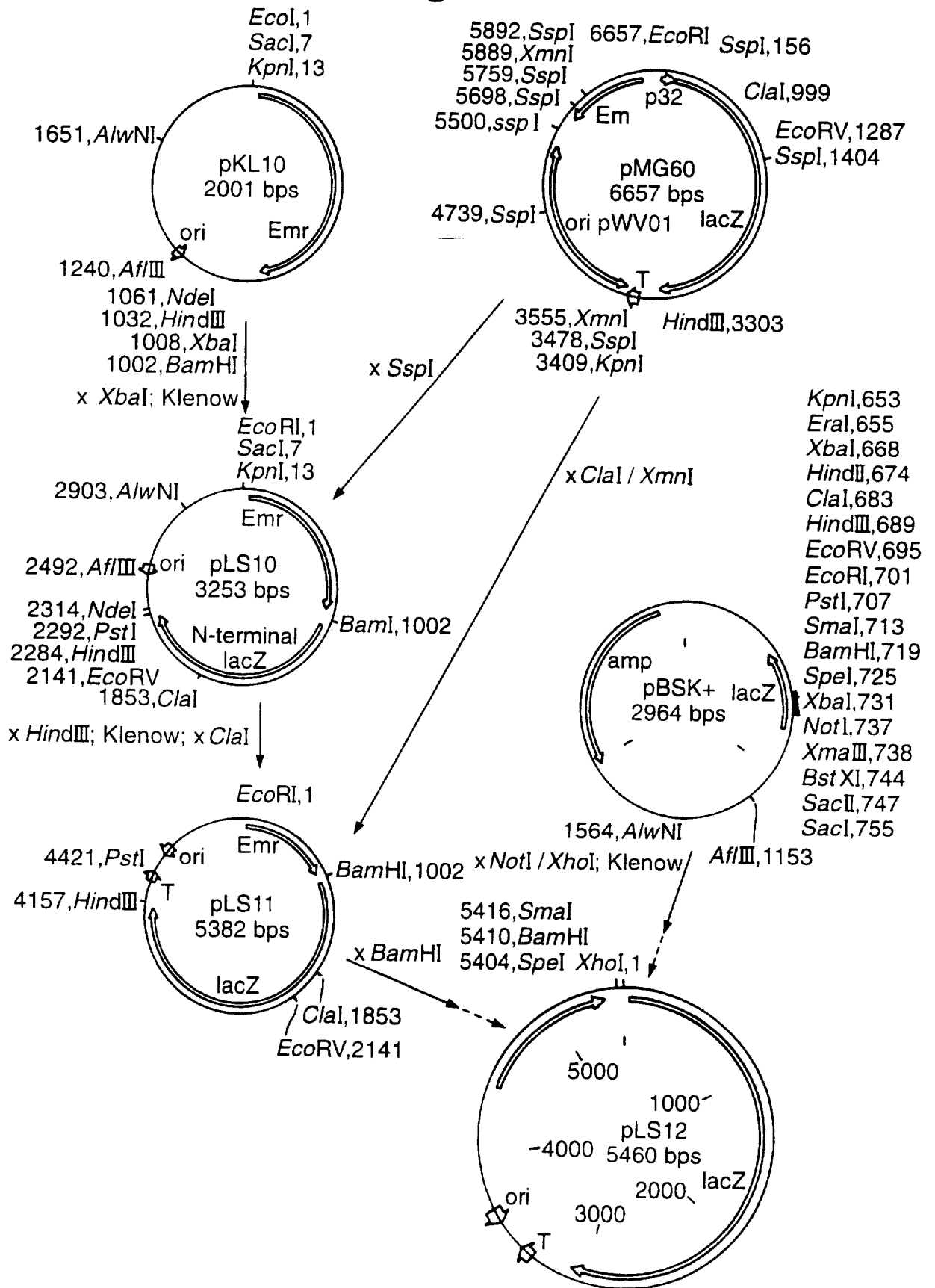




Fig.3.

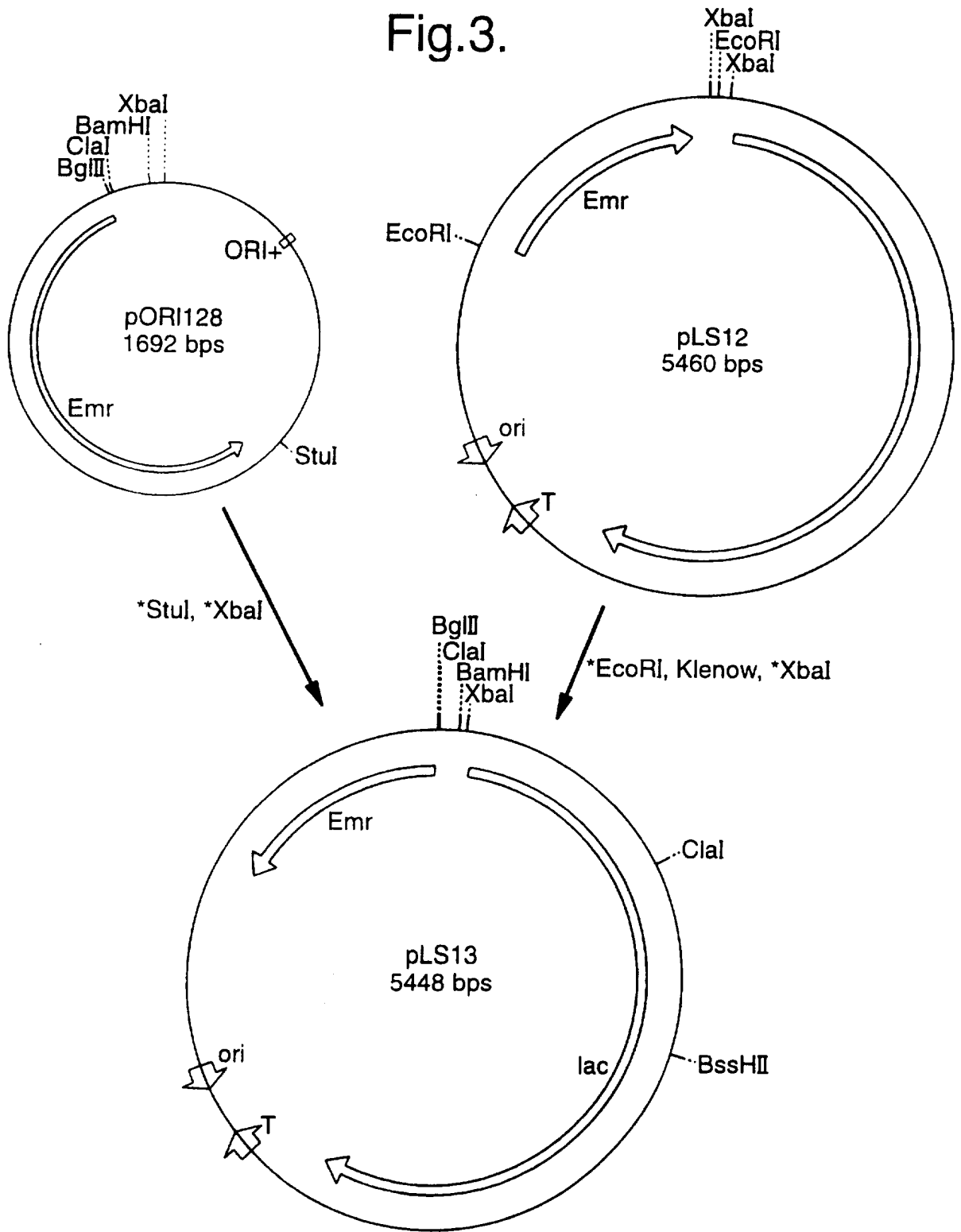
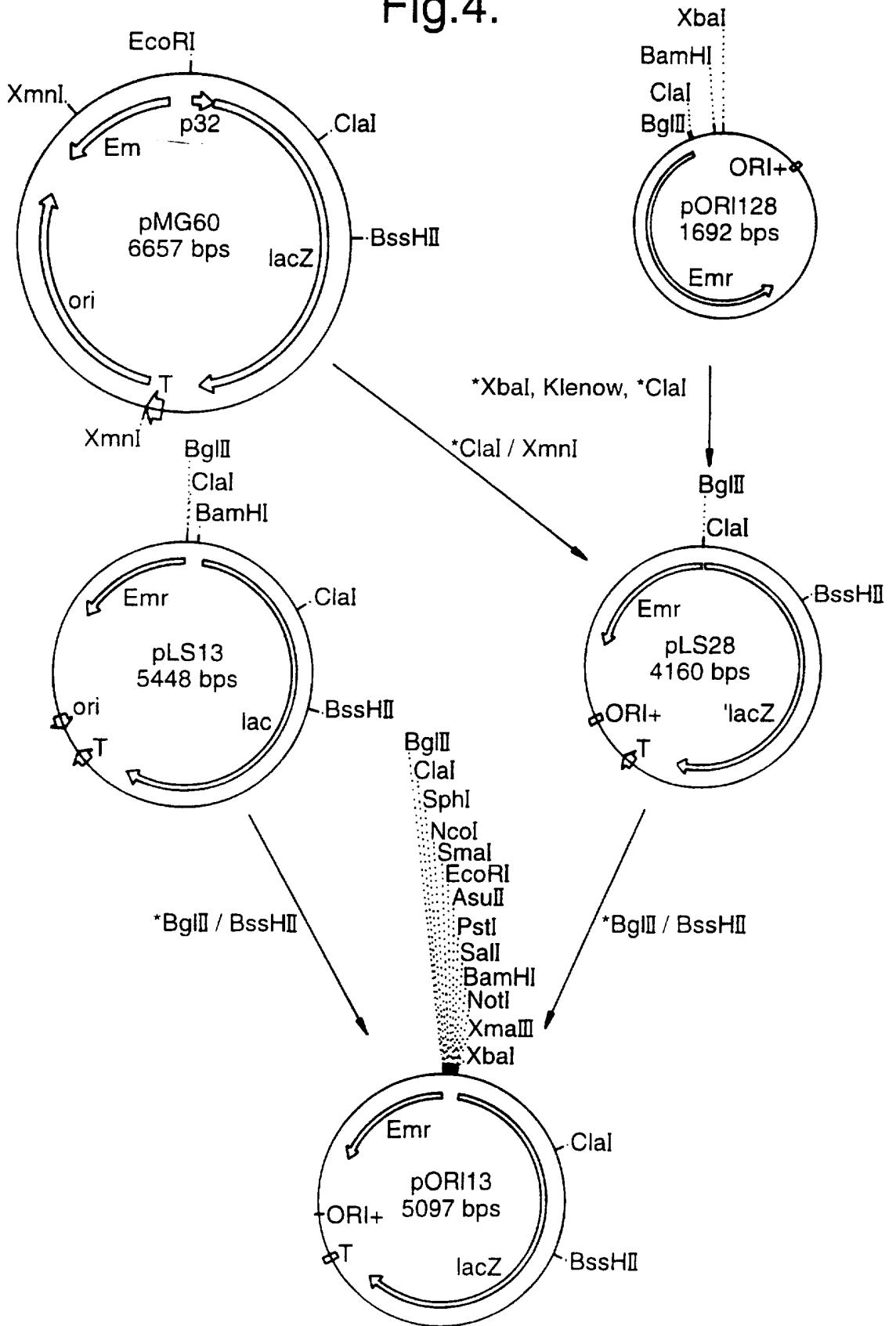
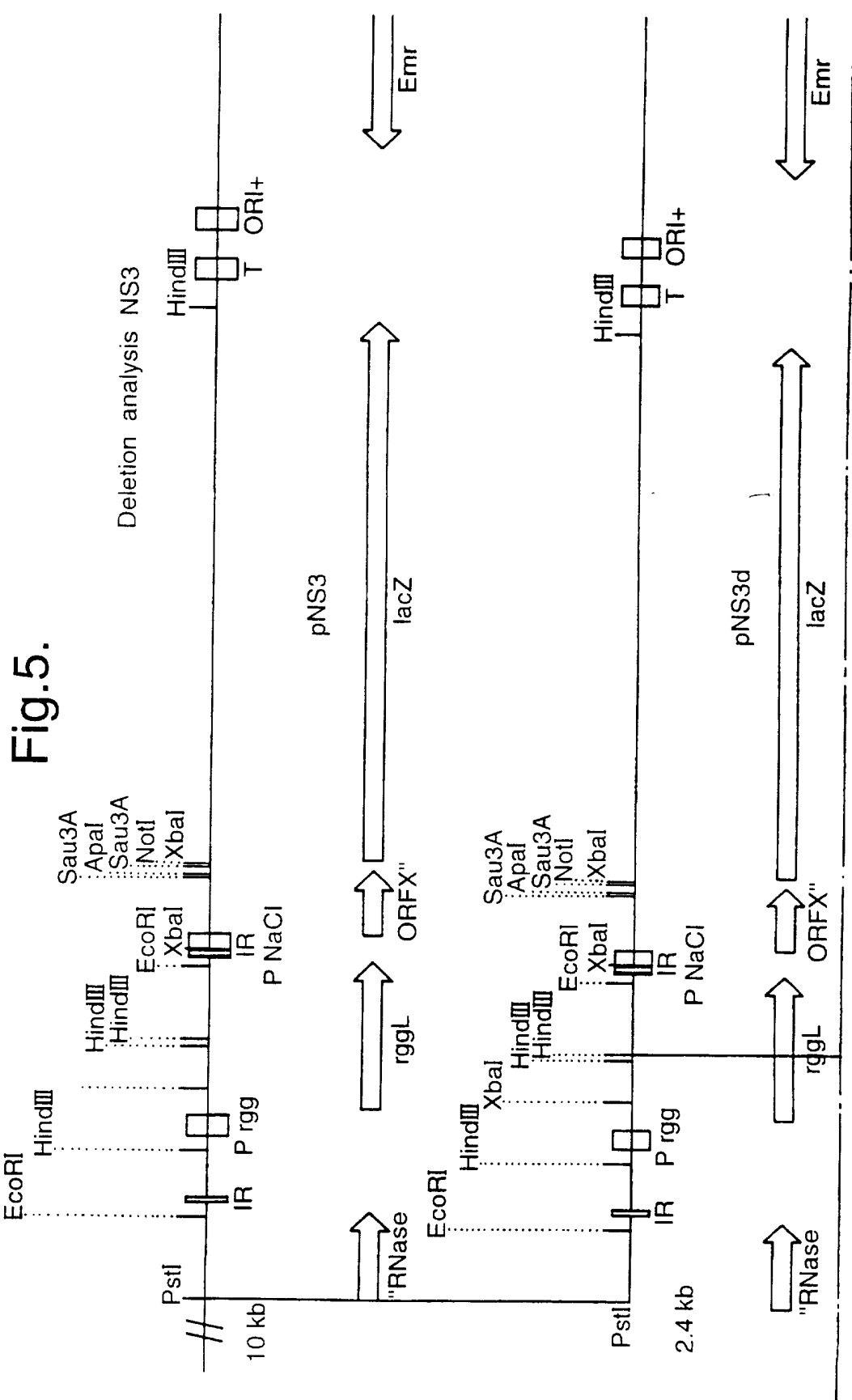


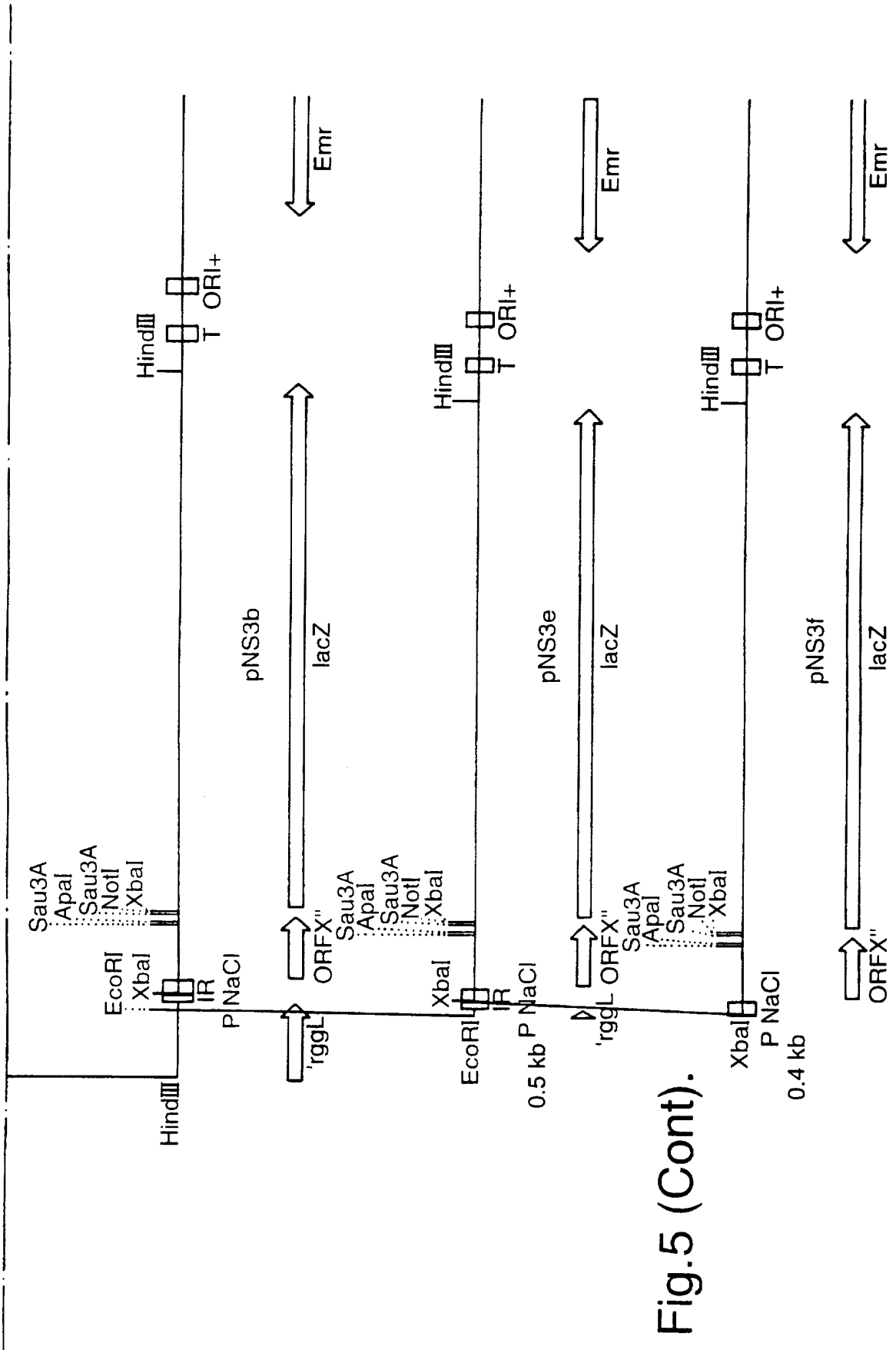
Fig.4.



5/34



6/3A



7/34

Fig.6. (1/3)

```

PstI   'rnhB
1  CTGCAGTAAT TTTGCCAAAG AATTGTAAAA TTCGTGGTTT GAATGATAGT AAAAAAGTGC
1  AlaVal IleLeuProLys AsnCysLys IleArgGly LeuAsnAspSer LysLysVal

61  CAAAATCAAA GCATCATGCT ATTCTATCTG AAATTCAAGA AAAAGCGCTA GCGATTGGAG
20  ProLysSer LysHisHisAla IleLeuSer GluIleGln GluLysAlaLeu AlaIleGly

121  TTGGAATTGT TGATGCCGAA AAAATTGATG AAGTAAATAT TTATGAGGCG ACAAAAATTG
40  ValGlyIle ValAspAlaGlu LysIleAsp GluValAsn IleTyrGluAla ThrLysIle

181  CAATGATTCA GGCAGTATCA AAATTATCTC TTAACCTGA ACATCTCTTA ATAGATGCCA
60  AlaMetIle GlnAlaValSer LysLeuSer LeuLysPro GluHisLeuLeu IleAspAla

241  TGGTTTTAGA TTTGCCCATF GCTCAGACGA AAATCATTCA TGGAGATGCT CGTTCAGCTT
80  MetValLeu AspLeuProIle AlaGlnThr LysIleIle HisGlyAspAla ArgSerAla

NotI
301  CAATTGCGGC CGCATCAATT GTAGCTAAAG TGAATGATG TGAAATGATG AAGGATTTTCG
100 SerIleAla AlaAlaSerIle ValAlaLys ValThrArg AspGluMetMet LysAspPhe

361  CTTTAGAATT TCCAGAATAT GATTTTGAAC ATAATGCAGG CTATGGAACA GCAAAACATC
120 AlaLeuGlu PheProGluTyr AspPheGlu HisAsnAla GlyTyrGlyThr AlaLysHis

EcoRI
421  TTGCAGCTCT GACAAAATAT GGTATCACAA GAATTCATCG GAAATCCTAT GAACCAATTA
140 LeuAlaAla LeuThrLysTyr GlyIleThr ArgIleHis ArgLysSerTyr GluProIle

481  AATCGATGGT CAATTTCAAA TAGTAGATTA TGTAAGTAAA AAAAGGAAAA CGTGAACGAT
160 LysSerMet ValAsnPheLys ***          ----- >

541  ATTTGGTCAC GTTTTTTTGC TGACAAGTCT GTCAGTAATT ATTTTCAAAG GTTTCAAAAA
< -----

601  TATAGTCTAG TAATTTGCTA GACTAATCTT CTGTTTTTTA ATAATAAACT AATTTTTTGT

661  TAATCTAAAT GACAAAATTA ATAAGCAGAG TTTTTTATAA AATTAGCTAC TTATAAAAAA

721  TTTGAAATTG GTATAGTTAA ATCTGTTATA ATTTCCAATA TTTTTTAATA ATAATTATTT

HindIII
781  TAACAAAATA CTTATATCAA AACTCTTTCA AAGTATATAA TGAGCGTTGT ATAAGCTTTT

841  ATGCTTTTCT ATATCAACTT TTAATAGAAA TATAAAGTAA TATAAATGTT TTTATAATAA

-35
901  ATTATGTGAG ATATATTTTT TTGTCCGTAC TGGTATAGAT TTGACGATTA AGTCITTAAT

-10
961  AAGTTATAAT CTCAATTGCG TAATTTCTTA AATACAGAAA TAACAACTAC ATTGGTAGAC
      ↑

1021  TGATTAAAAA GTGTACTTGA TGAAGTGTTA TAAACCTTAA AAAAATAAAA ATAATAGTTT

start rggL
1081  GGGggaTgtT AAAGATGTAT AAAAAATATG GAGATTGTTT TAAAAAGTTG CGAAACCAAA
1     rbs      MetTyr LysLysTyr GlyAspCys PheLysLysLeu ArgAsnGln

XbaI
1141  AGAATTTAGG GTTATCATACT TTTAGTAAAC TTGGAATAGA CCGTTCAAAT ATATCTAGAT
16  LysAsnLeu GlyLeuSerTyr PheSerLys LeuGlyIle AspArgSerAsn IleSerArg

1201  TTGAACATGG AAAATGTATG ATGAGTTTTG AGCGTATAGA TTTGATGTTA GAAGAAATGC
36  PheGluHis GlyLysCysMet MetSerPhe GluArgIle AspLeuMetLeu GluGluMet

```

8/34

Fig.6. (2/3)

```

1261 AAGTTCGGTT ATCTGAGTAC GAATTGATTG TAAATAATTT TATGCCGAAT TTCCAAGAAT
 56 GlnValPro LeuSerGluTyr GluLeuIle ValAsnAsn PheMetProAsn PheGlnGlu

1321 TTTTATATATT AGAATTGGAA AAAGCTGAAT TTAGCCAAAA TCGAGATAAA ATAAAAGAGT
 76 PhePheIle LeuGluLeuGlu LysAlaGlu PheSerGln AsnArgAspLys IleLysGlu

1381 TGTATTCTGA GGTCAAAGAA ACGGGGAATC ATTTACTGAC GGTTACCGTG AAAACGAAGC
 96 LeuTyrSer GluValLysGlu ThrGlyAsn HisLeuLeu ThrValThrVal LysThrLys
                                     HindIII

1441 TTGGGAATAT AAGTCAGACA GAAGTTAAAG AAATTGAAGC TTATCTTTGC AATATTGAAG
 116 LeuGlyAsn IleSerGlnThr GluValLys GluIleGlu AlaTyrLeuCys AsnIleGlu

1501 AGTGGGGATA TTTTGAACCT ACTTTATTTT ATTTTGTATC TGATTATCTC AATGTCAATC
 136 GluTrpGly TyrPheGluLeu ThrLeuPhe TyrPheVal SerAspTyrLeu AsnValAsn

1561 AATTAGAATT GCTGCTTTTT AATTTTGATA AAAGATGTGA AAATTACTGT AGAGTCTTAA
 156 GlnLeuGlu LeuLeuLeuPhe AsnPheAsp LysArgCys GluAsnTyrCys ArgValLeu

1621 AATATAGAAG GAGACTATTG CAAATAGCCT ATAAAAGTGT TGCGATATAC GCGGCTAAAG
 176 LysTyrArg ArgArgLeuLeu GlnIleAla TyrLysSer ValAlaIleTyr AlaAlaLys

1681 GAGAAAGAAA AAAAGCCGAA AATATTTTAG AAATGACTAA AAAATATCGA ACTGTGGGAG
 196 GlyGluArg LysLysAlaGlu AsnIleLeu GluMetThr LysLysTyrArg ThrValGly

1741 TCGATTTATA TTCAGAAGTA TTAAGACATC TTGCTAGAGC TATCATTATT TTTAATTTTG
 216 ValAspLeu TyrSerGluVal LeuArgHis LeuAlaArg AlaIleIleIle PheAsnPhe

1801 AAAATGCAGA GATTGGGGAA GAAAAAATAA ATTATGCTCT TGAGATTTTG GAAGAATTTG
 236 GluAsnAla GluIleGlyGlu GluLysIle AsnTyrAla LeuGluIleLeu GluGluPhe

                                     EcoRI

1861 GAGGAAAGAA GATAAAAGAA TTCTATCAGA ATAAAATGGA AAAGTATTTG AAAAGGTCAA
 256 GlyGlyLys LysIleLysGlu PheTyrGln AsnLysMet GluLysTyrLeu LysArgSer

1921 TTTAGTCTCT TTTGAGCTGT TGCTTTAAAG CAACAGCTCA AAAGAGATTT TCTTTATCT
 276 Ile***-----><-----

XbaI      -10
1981 AGAGCATATA CTAGAGGGTG AAGATAGGTT GTCTGAAGCA TTATAACTTG TCTTTTAAAA
      ↑
                                     Start orfX
2041 AATTCAATCA TAAATATAag gaggtATGAT GAATCAAAAA AAATTATCAT TATTCGGTTT
 1      rbs      Met AsnGlnLys LysLeuSer LeuPheGlyPhe

2101 TTTGCGCATT ACCGCTTCAA TGTTTTTGAC TGTCTATGAG TATCCGACTT TTGCCACGTC
 12 PheAlaLeu ThrAlaSer MetValLeu ThrValTyrGlu TyrProThrP heAlaThrSer

2161 AAAATTACAT TTGGTGTCT TTTTACTTCT CGGAGGACTA CTATGGTTTT TGCCTGTAGC
 32 LysLeuHis LeuValPhe PheLeuLeu LeuGlyGlyLeu LeuTrpPhe LeuProValAla

2221 GCTCTGCGCA GCAGAAATGG CGACGGTTGA AGGCTGGAAA AATGGTGGAA TCTTTAGTTG
 52 LeuCysAla AlaGluMet AlaThrVal GluGlyTrpLys AsnGlyGly IlePheSerTrp

2281 GGTCAGTCAA ACTTTAGGTG AGCGCTTTGG TTTTGCAGCC ATATTTTTTC AGTGGTTCCA
 72 ValSerGln ThrLeuGly GluArgPhe GlyPheAlaAla IlePhePhe GlnTrpPheGln

                                     ApaI
2341 AATTACAGTA GGTTTGTGCA CTATGATCTA TTTCATTTTA GGGGCCCTCT CTTATGTGTT
 92 IleThrVal GlyPheVal ThrMetIle TyrPheIleLeu GlyAlaLeu SerTyrValLeu
    
```

9/34

## Fig.6. (3/3)

Sau3A

2401 AAATTTTCAG GCGCTCAATA CAGATCCATT GATAAAATTT ATTGGTTTAC TAATCATTTT  
 112 AsnPheGln AlaLeuAsn ThrAspPro LeuIleLysPhe IleGlyLeu LeuIleIlePhe

2461 TTGGGGATTG ACTTTTTCTC AATTAGGTGG GACGCAACGG ACTGCCAAAT TAGTAAAAGC  
 132 TrpGlyLeu ThrPheSer GlnLeuGly GlyThrGlnArg ThrAlaLys LeuValLysAla

2521 TGGCTTTGTA GTTGAATAG TGATTCCATC GGTATCTTG TTTGGATTAG CAGCGGCATA  
 152 GlyPheVal ValGlyIle ValIlePro SerValIleLeu PheGlyLeu AlaAlaAlaTyr

2581 CTTTATCGGA GGCAATCCTA TAGAAATACC AATTAACAGC CATGCTTTTG TACCAGATTT  
 172 PheIleGly GlyAsnPro IleGluIle ProIleAsnSer HisAlaPhe ValProAspPhe

2641 TTCACAGGTA TCAACTTTAG TAGTTTTTGT TTCTTTTATT CTGGCTTATA TGGGGGTAGA  
 192 SerGlnVal SerThrLeu ValValPhe ValSerPheIle LeuAlaTyr MetGlyValGlu

2701 AGCCTCAGCT TCACATATTA ATGAACCTGA AAATCCAAAA CGAAATTATC CCTTAGCAAT  
 212 AlaSerAla SerHisIle AsnGluLeu GluAsnProLys ArgAsnTyr ProLeuAlaMet

2761 GATTTTATTA GTAATTTTGG CTATTTCTTT AGATGCCATA GGTGGATTTT CTGTAGCAGC  
 232 IleLeuLeu ValIleLeu AlaIleSer LeuAspAlaIle GlyGlyPhe SerValAlaAla

2821 AGTTATTCCT CAAAAGAGT TATCATTAG TGCAGGGGTA ATCCAAACTT TTCAAACGTT  
 252 ValIlePro GlnLysGlu LeuSerLeu SerAlaGlyVal IleGlnThr PheGlnThrLeu

2881 AATCTTACAT TTTAATCATC ATTTGGGATG GTTAGTTAAA GTGATTGCAC TAATGATTGC  
 272 IleLeuHis PheAsnHis HisLeuGly TrpLeuValLys ValIleAla LeuMetIleAla

XbaI

2941 CTTTGGGGTT ATGGGAGAAG TGAGTTCATG GGTGTTGGT CCTTCTAGA  
 292 PheGlyVal MetGlyGlu ValSerSer TrpValValGly ProSerArg

10/34

Fig.7.

RGGL	MY - - KKYGDCFKKLRNQKNLGLSYFSKLGIDRSNISRFEHGKCMMSFERI	48
RGG1	MLIVKSSGKILKIIRESKNMSLKEVAAGDISVAQLSRYERGISSLTVDSF	50
ORF3	M- PYKRYGEIFKKLREQKNFSLSHFSEIGISKASLSRFELGQTMISFERL	49
	* * . * . * . * . * . * . * . . . * . . . * . * * * . . . . . .	
RGGL	DLMLEEMQVPLSEYELIVNNFMPNFQEFFILELEKAEFSQNRDKIKELYS	98
RGG1	YSLRNMSVSLAEFQYVYHNYREADDVVLSQKLSEAQRENNIVKLESILA	100
ORF3	DSALQEMNVTLAEYEHFINNFSDYKEEFLEDIILADIADDVDKHLHL - -	97
	* * * * * . * . . . . * . . . . * . . . . * . . . .	
RGGL	- - - - - EVKETGNHLL - TVTVKTKLGN - - - - - ISQTEVKEIEAYLCNI	134
RGG1	GSEAMAQEFPEKKNYKLNTIVIRATLTSCNPDYQVSKGDIEFLTLYLFSV	150
	* * * * * . * . . . . * . . . . * * . . .	
RGGL	EEWGYFELTLFYFVSDYLVNVQLELLLNFNFKRCENYCRVLKYRRRLQLI	184
RGG1	EEWGRYELWLFTNSVNLLTLETLETFASEMINRTQFYNNLPENRRRIIKM	200
	**** * . * . * * . * . * * * . * * . . . **** . . .	
RGGL	AYKSVAIYAAKGERKKAENILEMTKKYRTVGVDLYSEVLRHLARAIIFN	234
RGG1	LLNVVSACIENNLQVAMKFLNYIDNTKIPETDLYDRVLIKYHKALYSYK	250
	* . * . * * * . * * * * * * . * . .	
RGGL	FENAEIGE - - EKINYALEILEEFG - GKKIKEFYQNKMEKYLKR - SI	276
RGG1	VGNPHARHDIEQCLSTFEYLDSEFGVARKLKEQFERIQLTVVADLQIE	297
	* * . * * * * * * * * * . * . . . . *	



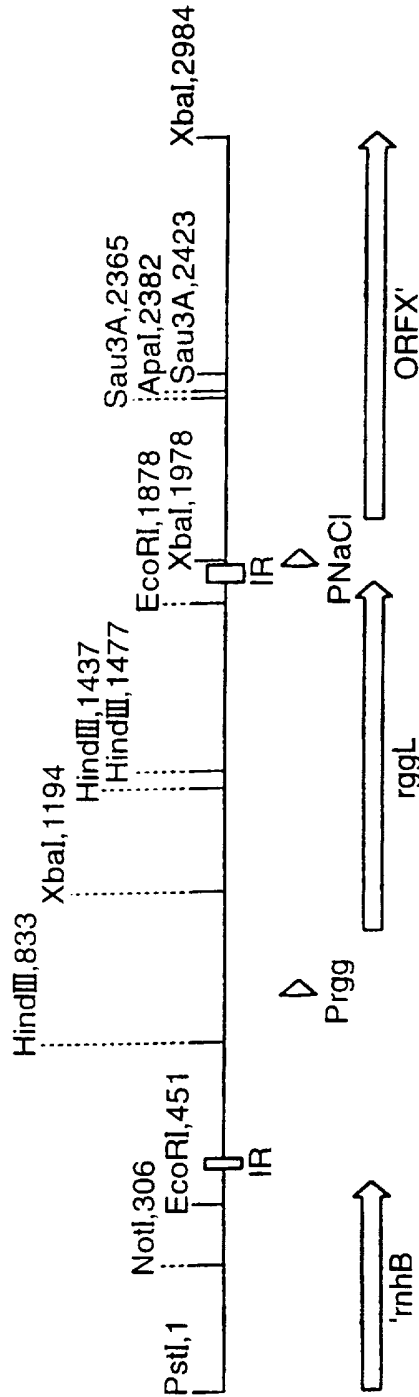
11/34

Fig.8.

L. lactis	-----AVILPKNCKIRGLNDSKKV	19
V. cholerae	-----PNRPIMGLNDSKKL	14
E. coli	MIEFVYPHT-QLVAGVDEVGRGPLVGAVVTAAVILDPARP IAGLNDSSKKL	49
H. influenzae	M--FEYPQGYKLLIAGVDEVGRGPLVGAVVTAAVILDPHPNPIEGLADSSKKL	48
	**** * ** . **** .	
L. lactis	PKSKHAILSEIQEKALAIIGVIGVDAEKIDEVNIYEATKIAMIQAVSKLS	69
V. cholerae	SEKKRLALFPEIQVKALAWAVGRCSPOEIDELNIFQATMVAMQRAVAGLR	64
E. coli	SEKRRRLALYEEIKEKALSWSLGRAEPHEIDELNILHATMLAMQRAVAGLH	99
H. influenzae	SEKKRLALAAEEIKEKARAWALGRAEADEIDEINILQASLLAMTRAVKSLK	98
	..... * . ** . ** . . . . . **** . ** . ** . ** . * .	
L. lactis	LKPEHLLIDAMVL- -DLPIAQTKI IHGDARSASIAAASIVAKVTRDEMVK	117
V. cholerae	IQPDVLIDGNKIPK- LPMEAQAVVKGDLRVAQISAASIIIAKVIRDQEME	113
E. coli	IAPFYVLIDGNRCPK- LPMPAMAVVKGDSRVPEISAASIIIAKVTRDAEMA	148
H. influenzae	IQPHFVLIDGNKIPKDLAIPAQAVVKGDSLVAEISAASIIIAKVARDQEME	148
	. * . . . . **** . * . . . . * . . . . * . . . . * . . . . * .	
L. lactis	DFALEFPEYDFEHNAGYGTAKHLAALTKYGITRIHRKSYEPIKSMVNFK-	166
V. cholerae	ALDKQYPQFGFANHKGYPTAAHFAAIEQHGVIHQHRKSFQVVKRALGE--	162
E. coli	ALDIVFPQYGFQAQHKGYPTAFHLEKLAEHGATEHRRRSFGPVRALGTCV	198
H. influenzae	ELDKQYPEYAFQAQHKGYPTKLHLEKLAELGALPQHRRSFAPVKKALEQF-	197
	... * . . . . * . . . . * . . . . * . . . . * . . . . * . . . .	
L. lactis	-----	166
V. cholerae	-----	162
E. coli	LILVSRLSKPESEDV	213
H. influenzae	-----	197

12/3A

Fig.9.



3/34

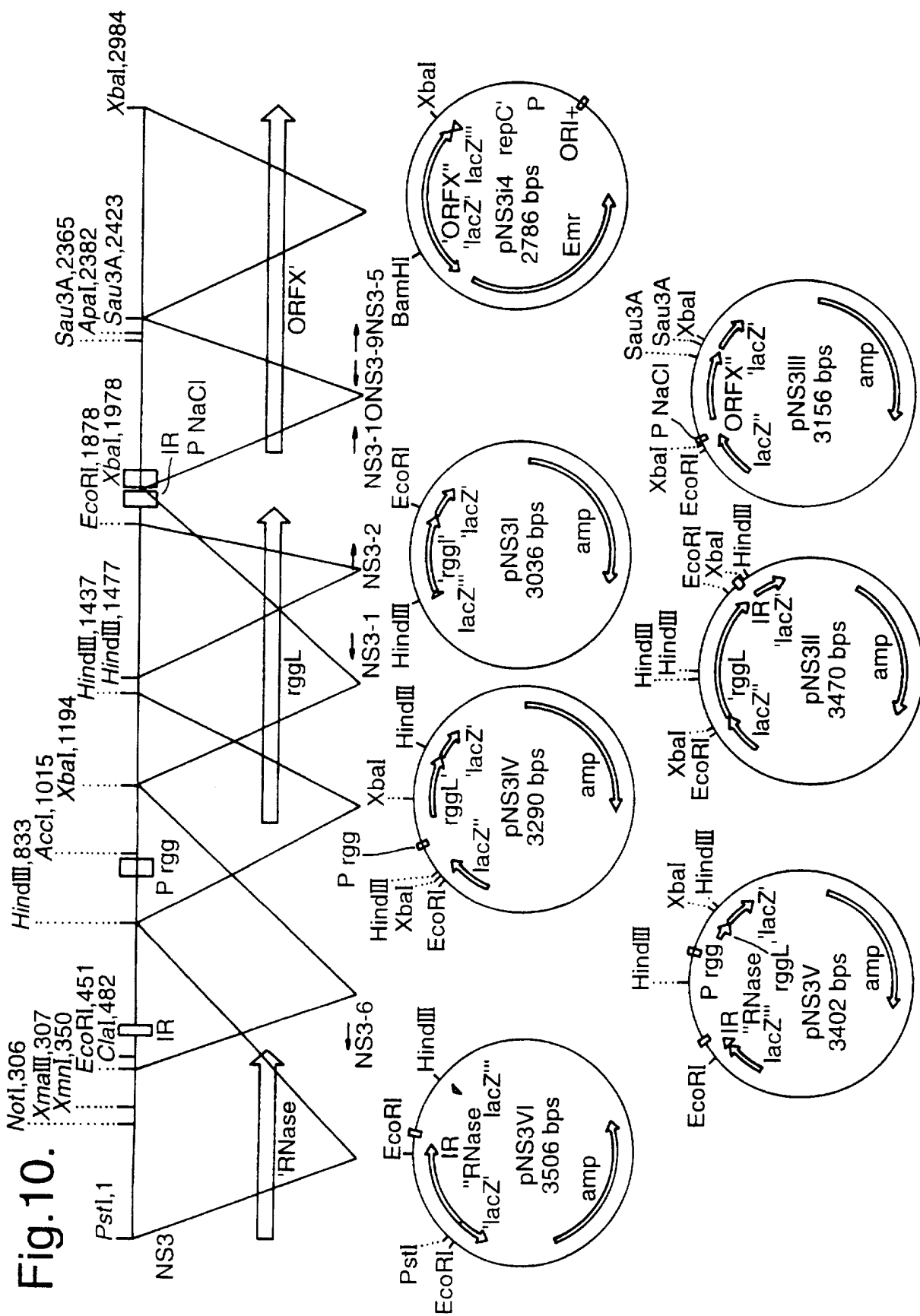


Fig.10.

14/34

Fig.11.

MG1363 MGNS3i3  
- + - + NaCl

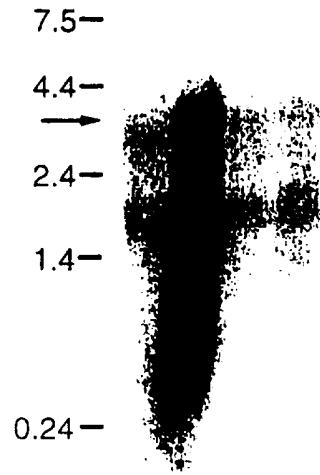
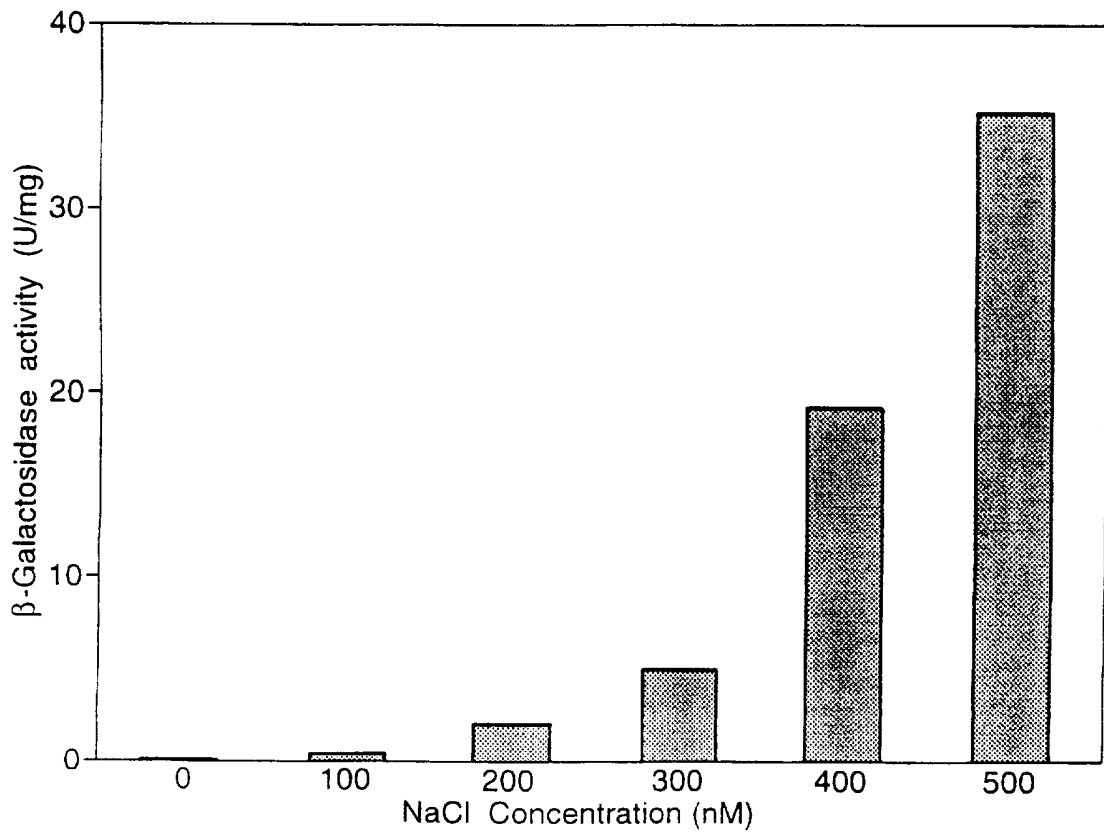


Fig.12.



15/34

## Fig.13.

## Part A

XbaI -10 ↓ Transcription start  
 1 TCTAGAGCAT ATACTAGAGG GTGAAGATAG GTTGTCTGAA GCATTATAAC TTGTCTTTTA

61 AAAAAATTCAA TCATAAATAT AAGGAGGTAT GatgAAGTAC TGGAT|ACTTA TCAATGATGA  
 RBS ScaI EcoRV/ScaI  
 ORFX': MetLysTyr TrpIle Leu IleAsnAspGlu

121 ACCTTGGTTT GTCGGAAAAG ATGTAGCAAT TGCTATTGGT TACAAGAATT TCAGGGATGC  
 ProTrpPhe ValGlyLys AspValAlaIle AlaIleGlu TyrLysAsn PheArgAspAla

181 TTTGAAATCT CATGTAAAAG ACAAATATAA GAGGGAGTCG GACAGCAGTG ATTGGTTCAA  
 LeuLysSer HisValLys AspLysTyrLys ArgGluSer AspSerSer AspTrpPheAsn

241 CGACAATATA TTATTGGAAA CGAACTGCAT AAAAAATAAA AAATAGGAGA AAGAACatgA  
 RBS  
 AspAsnIle LeuLeuGlu ThrAsnCysIle LysAsnLys Lys\*\*\*  
 lytP: Met

301 AAACATTTTT TAAAGATATG GCAGAACGTG CCATTAAAAC ATTTGCACAA GCAATGATTG  
 LysThrPhePhe LysAspMet AlaGluArg AlaIleLysThr PheAlaGln AlaMetIle

## Part B

XbaI -10 ↓ Transcription start  
 1 TCTAGAGCAT ATACTAGAGG GTGAAGATAG GTTGTCTGAA GCATTATAAC TTGTCTTTTA

61 AAAGATTCAA TCATAAATAT AAGGAGGTAT GatgAAGTAC TTATTATATT TTGTAATCTT  
 RBS ScaI  
 ORFX': MetLysTyr LeuLeuTyr PheValIlePhe  
 MetLys Ser TyrTyrIle Leu\*\*\*

121 TAGAAAGGTA ATTATTTatg CCAGTATCAC GTGTTAAAGT TAAAAATAGA CATTTAAAAA  
 RBS  
 ArgLysVal IleIleTyr AlaSerIleThr Cys\*\*\*  
 acmA: Met ProValSer ArgValLys ValLysAsnArg HisLeuLys

16/3A

Fig.14.

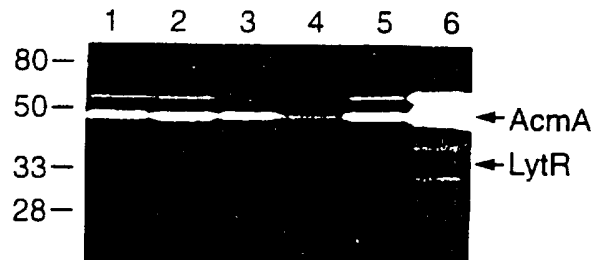
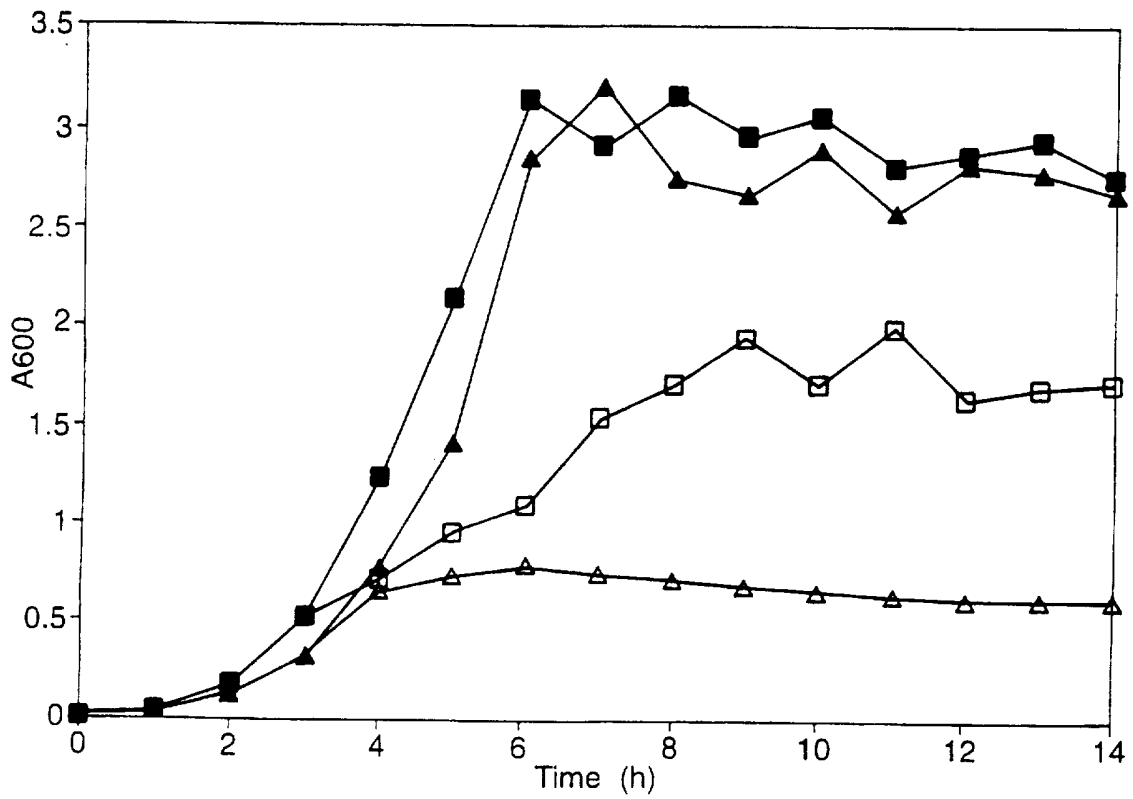
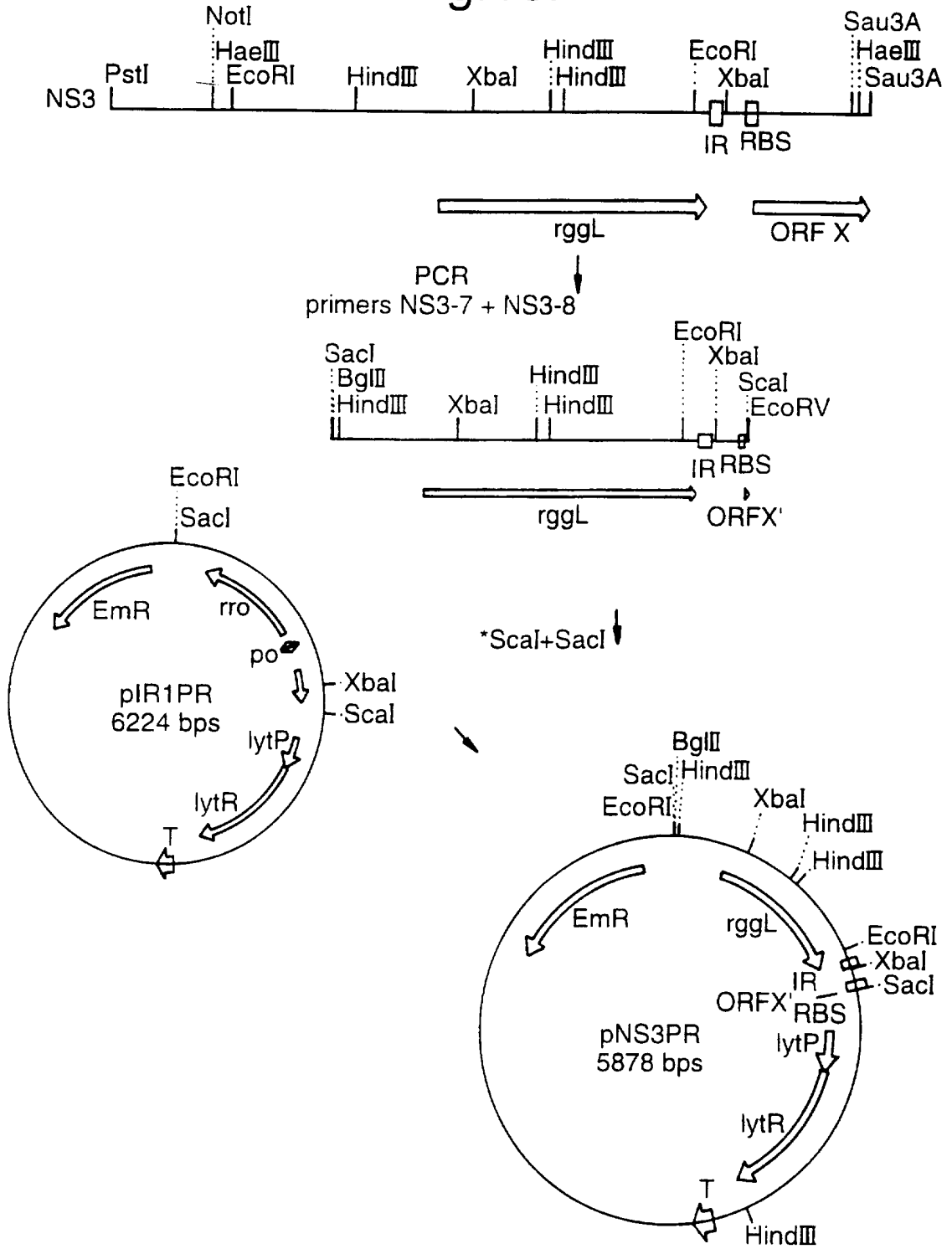


Fig.15.



17/34

Fig.16.



18/34

Fig.17.

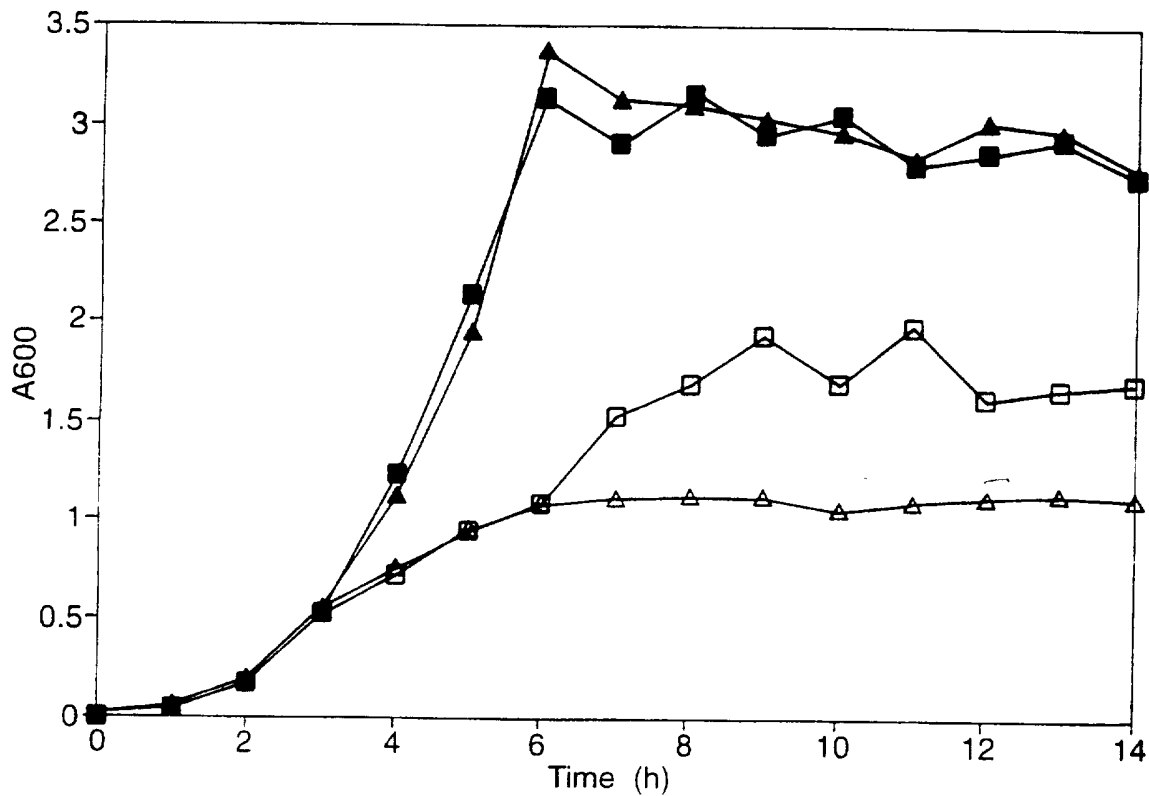
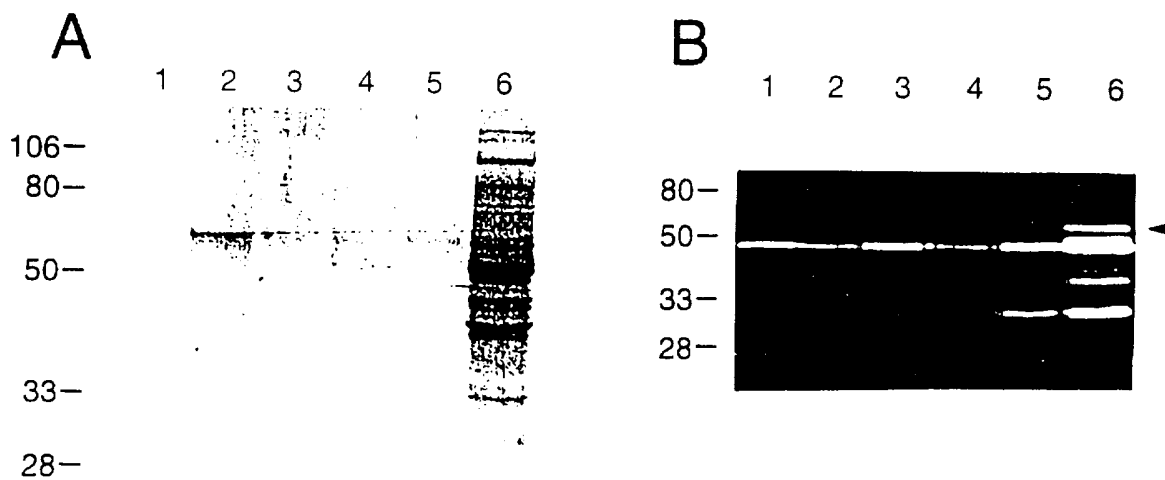


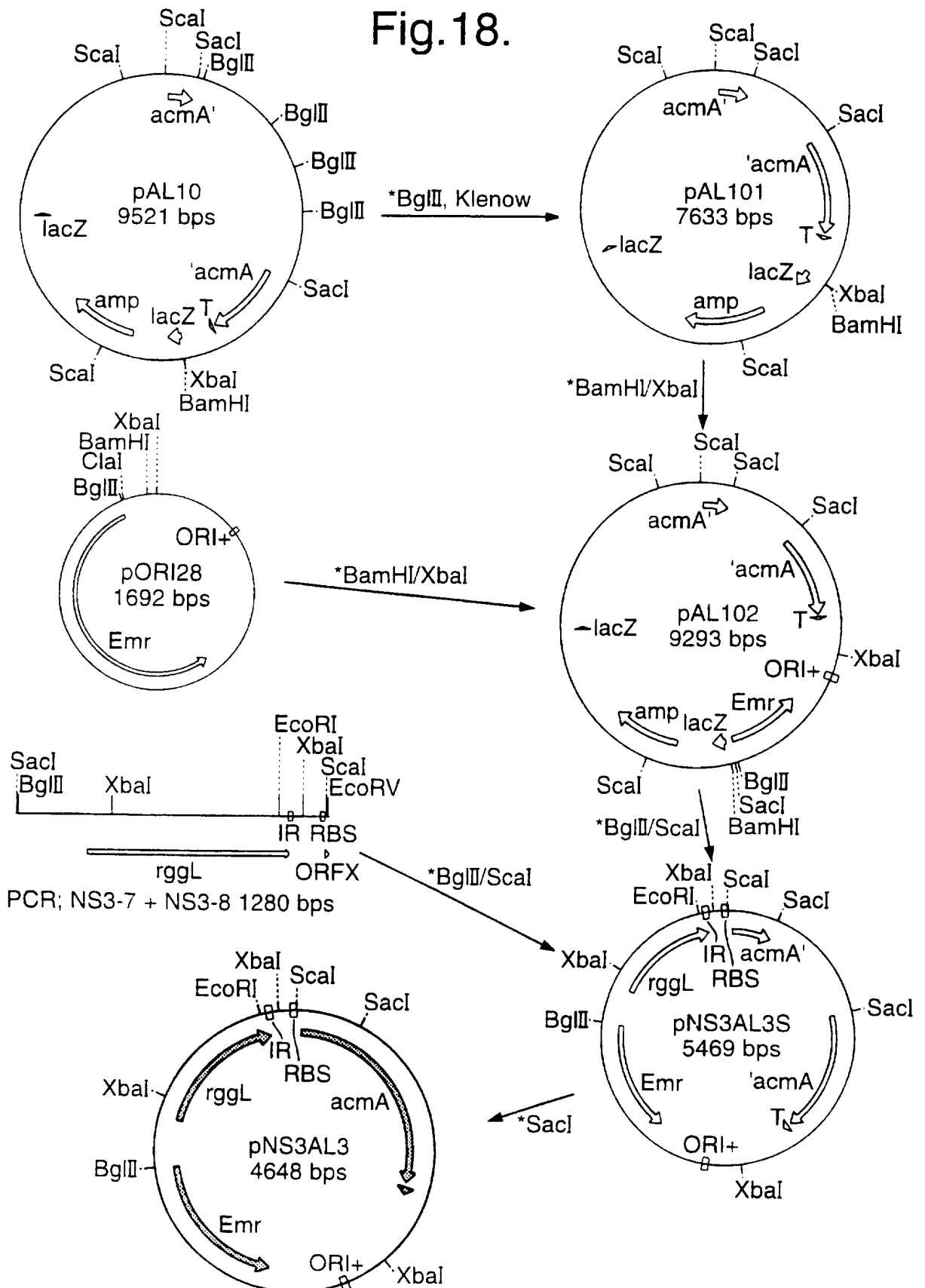
Fig.19.



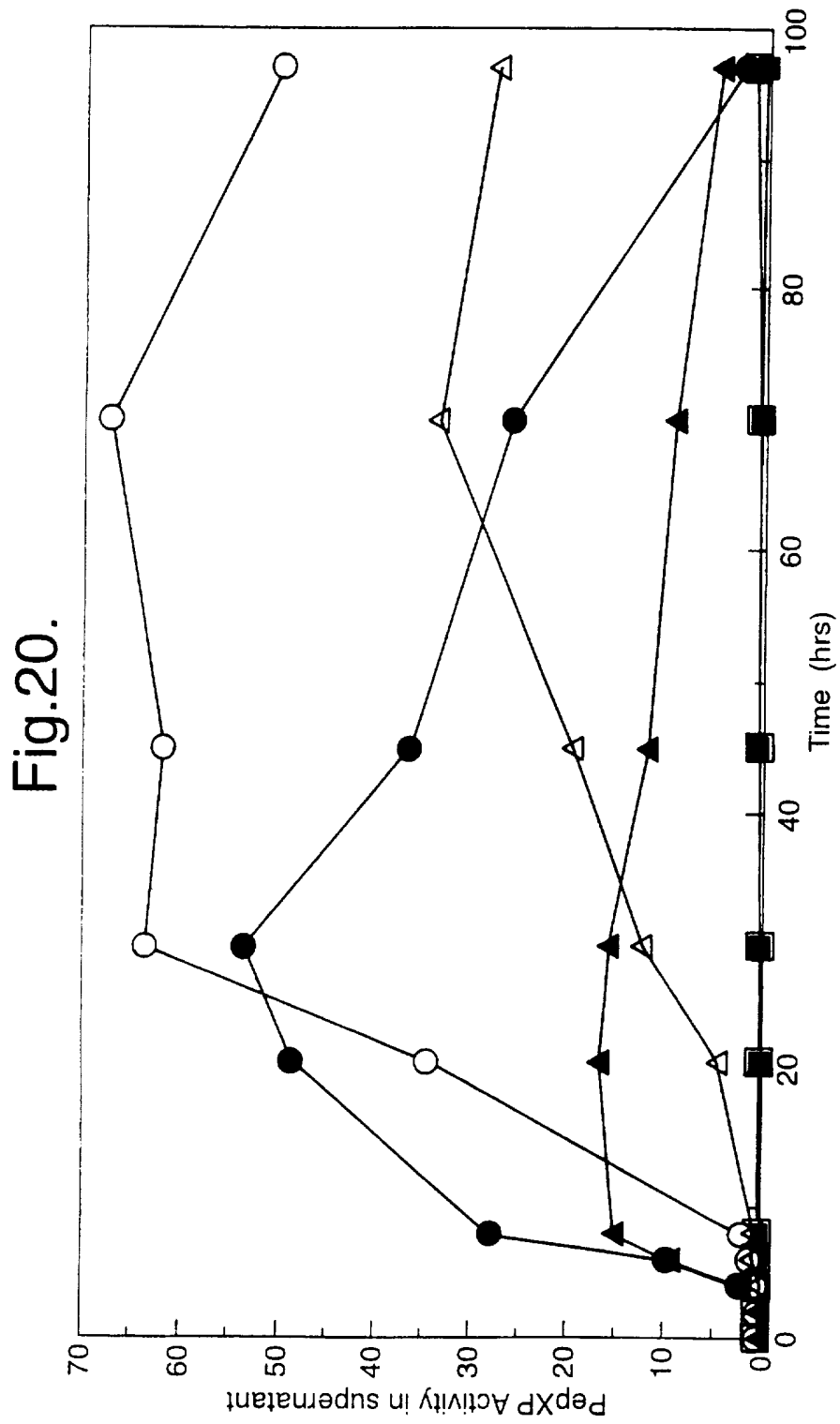


19/34

Fig. 18.

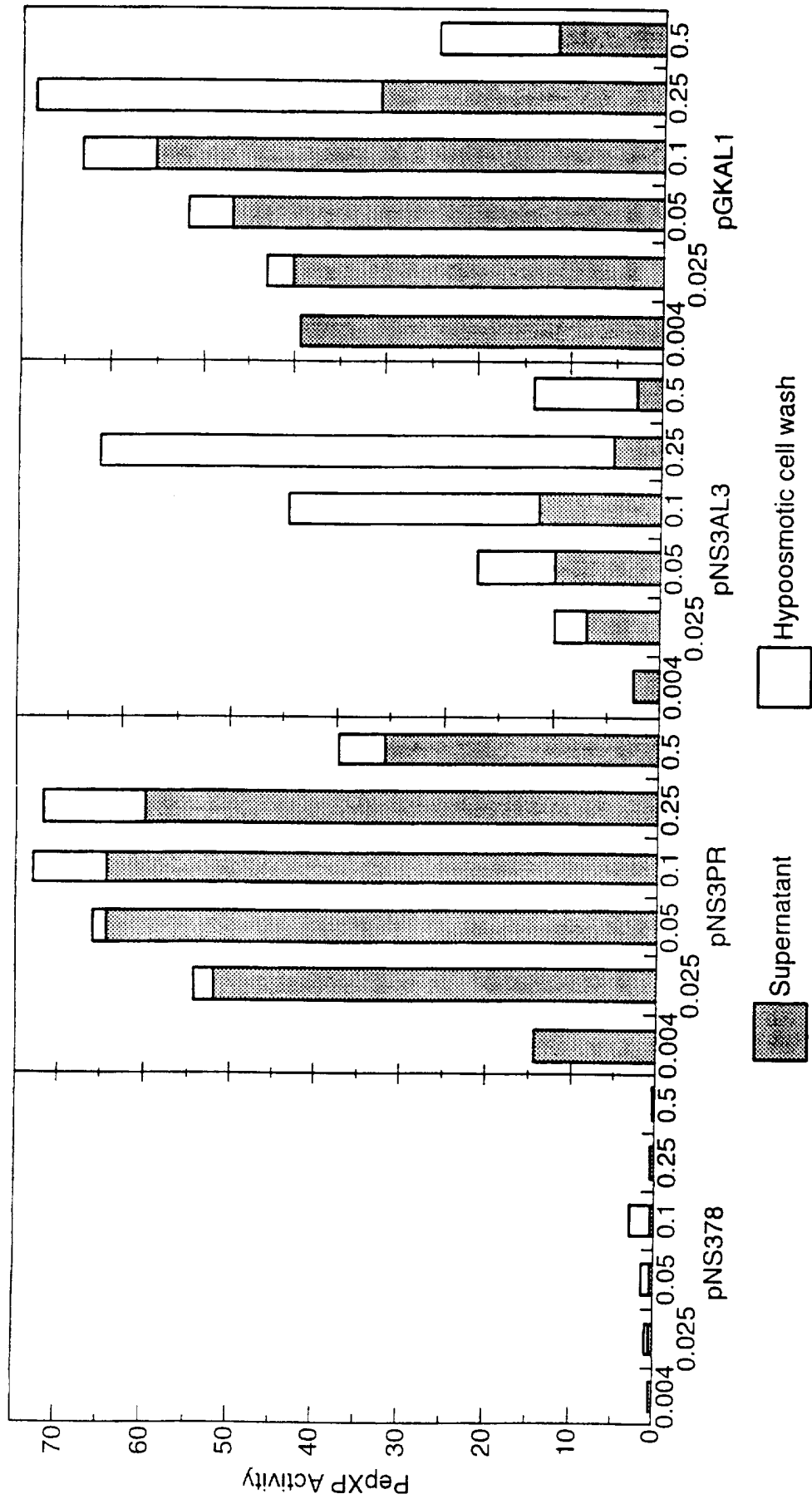


20/34



2/34

Fig.21.



22/3A

Fig.22.

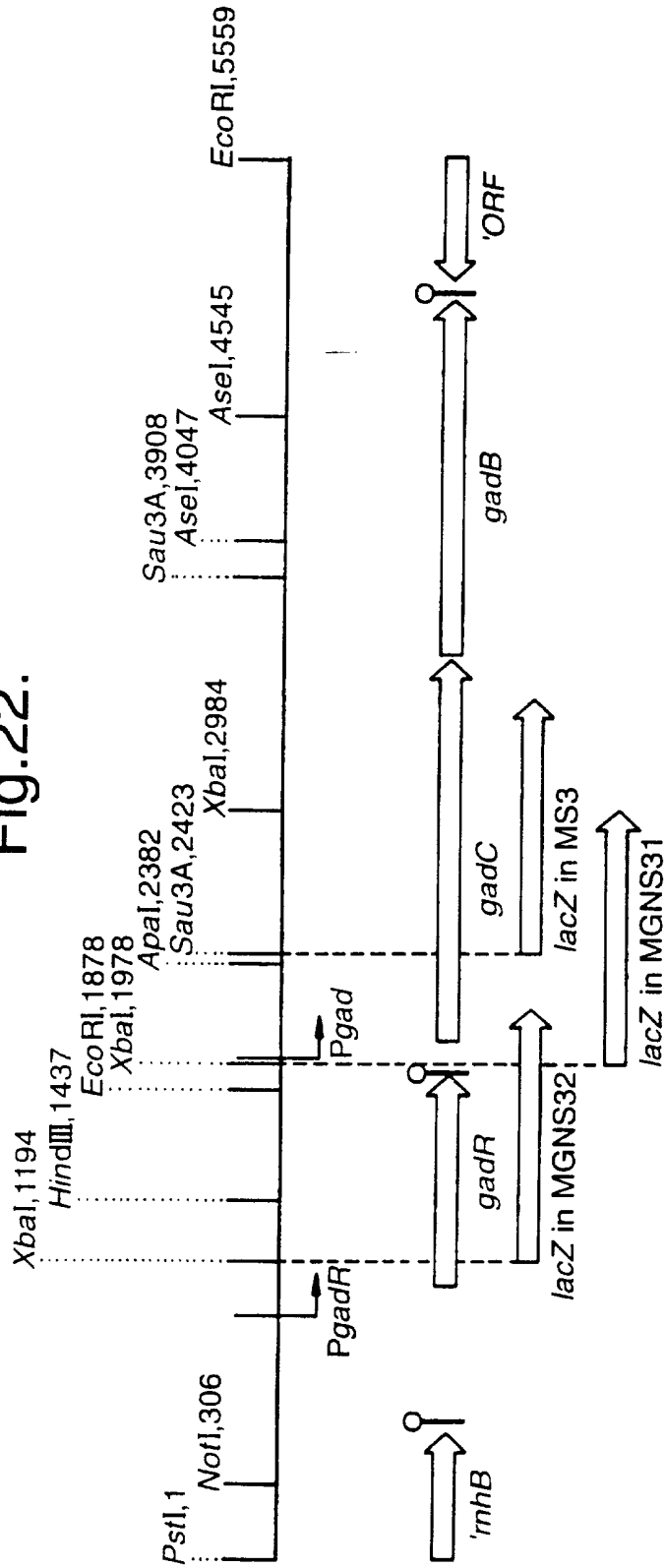


Fig.23.

<i>L.lactis</i>	<u>M-----NQKKLSLFGFFALTASMLTVYVYPTTFATSKLHLVFFLLGG</u>	43
	* * * * *	
<i>Shigella</i>	<u>MATSVQGTGKAKQLTLGFFAITASMVMVAVYVYPTTFATSGFSLVFFLLGG</u>	50
<i>L.lactis</i>	<u>LLWFLLPVALCAEMATVEGWKNGGIFSWVSQTLGERFGFAAIFQWFQIT</u>	93
	.***.*** ***** ** **.*.***.*** * .***** * .**	
<i>Shigella</i>	<u>ILWFI PVGLCAAEMATVDGWEEGGVFAWVSNTLGPRWGFAAISFGYLQIA</u>	100
<i>L.lactis</i>	<u>VGFTMIYFILGALSYVLNFQALNTDPLIKFIGLLIFWGLTFSQLGGTO</u>	143
	.***.*** ***** * .*** ** * * * * * * . * * * *	
<i>Shigella</i>	<u>IGFIPMLYFVLGALSYILKWPALNEDPITKTIAALIILWALALTQFGGTK</u>	150
<i>L.lactis</i>	<u>RTAKLVKAGFVVGVIVIPSVILFGLAAAYFIGGNPIEIPINSHAFVPDFSQ</u>	193
	** . *	
<i>Shigella</i>	<u>YTARIAKVGFFAGILLPAFILIALAIYLHSGAPVAIEMDSKTFPPDFSK</u>	200
<i>L.lactis</i>	<u>VSTLVVFSFILAYMGVEASASHINELENPKRNYPLAMILLVILAISLDA</u>	243
	* *	
<i>Shigella</i>	<u>VGTLVVFVAFIGSYMGVEASATHVNEMSNPGRDYP LAMLLLMVAACLSS</u>	250
<i>L.lactis</i>	<u>IGGFSVAAVIPOKELSLSAGVIQTFOTLILHFNHHLGWLKVIAMIAFG</u>	293
	. *	
<i>Shigella</i>	<u>VGGLSIAMVIPGNEINLSAGVMQFTVLM SHVAPEIEWTVRVISALLLG</u>	300
<i>L.lactis</i>	<u>VMGEVSSWVVGPSRGMFAAAQRGLLPKFLRKTNTHEVPVPLVMIQGIIVT</u>	343
	* . *	
<i>Shigella</i>	<u>VLAELASWIVGPSRGMVYTAQKNLLPAFAKMNKNGVPTLVISOLVITS</u>	350
<i>L.lactis</i>	<u>LWGAULTFGGGNNLSFLVAISLTVVIYLVGYLLFFIVYFVLIYKKQNLK</u>	393
	. *	
<i>Shigella</i>	<u>IALIILTNTGGGNNMSFLIALALT VVIYLCAYFMLFIGYIVLVLKHPDLK</u>	400
<i>L.lactis</i>	<u>RTYNVPGKIIIGKTI IAGIGFLLSIFALFISFVPPASIAKNETHYOMILL</u>	443
	* *	
<i>Shigella</i>	<u>RTFNIPGGKGVKLVVAIVGLLTSIMAFIVSFLPPDNIQGDSTDMYVELLV</u>	450
<i>L.lactis</i>	<u>ISFVVTAAILPFIYELHDKKGHD---TIEEPHFAGDVNPAIYPAARG</u>	489
	.***.*** ***** * .*** ** * * * * * * * * * * * * *	
<i>Shigella</i>	<u>VSFLVVLALPFILYAVHDRK GKANTGVTL EPI NSQNAPKGHFFLHPRARS</u>	500
<i>L.lactis</i>	<u>EHHIKKEEHILKH</u>	503
	* * * * *	
<i>Shigella</i>	<u>PHYIVMNDK---KH</u>	511

24/34

Fig.24.

```

L.lactis GadB MLYGKENRD---EAE-FLEPIFGSESEQVDLPKYKLAQQSIEPRVAYQLV 46
                *   *           *** * *           **   *   *   *   *
Synecocystis MVHKKIDLNLQSEAESLLTPTYAARGLANSVSKYEMPETEMLP AIAYNLI 50

L.lactis      QDEMLDEGNARLNLATFCQTYMEPEAVKLSQTLKNAIDKSEYPRTEI 96
                **   ** ***** * ***** ** * ** ** * ** * **
Synecocystis HDELGLDGN SRLNLATFVTTWMEPEARQLMADTFDKNMIDKDEYPQTAEI 100

L.lactis      ENRCVNMIADLWNASEKGIYGTSTIGSSEACMLGGMAMKFSWRKRAEKL 146
                * ***** ** * ***** ***** ** *
Synecocystis ELRCVNILSRLWNAPASAEATGCSTIGSSEAAMLGGMAMKWKWRQR--RQ 148

L.lactis      GLDINAKKPNLVISSGYQVCWEKFCVYWDIEMREVPMDREHMSINLEKVM 196
                ***** ***** ** * * ** * * *
Synecocystis AAGKPGDRPNLVMGINVQVCWEKFCRYWEVEPRFVPMEGDRYHISPEEAV 198

L.lactis      DYVDEYTIQVVGIMGITYTGRYDDIKALDNLIEEYNKQTDYKVIYHVDA 246
                ** ***** ** * * * * * * * * * * * * * * *
Synecocystis KLIDENTIGVIGILGSTFDGSEYEP IALNDALETNLQRTGWQVPLHIDAA 248

L.lactis      SGGLYAPFVEPELEWDFRLKNVISINTSGHKGVLVYPGVGWLWRDKKYL 296
                *** ** * * * * * * * * * * * * * * * * * * *
Synecocystis SGGFIAPFLDPDLRWFRLPWVKSINTSGHKGVLVYPGVGWIIWRDKKEEL 298

L.lactis      PEELIFKVSYLGGELPTMAINFSHSASQLIGQYVYVRYGFDGYKAIHER 346
                ***** ***** ** * * * * * * * * * * *
Synecocystis PEELIFHCNYLGGDLPNFALNFSRPGNQVVAQYVYVFLRLGKEGYRKIQQT 348

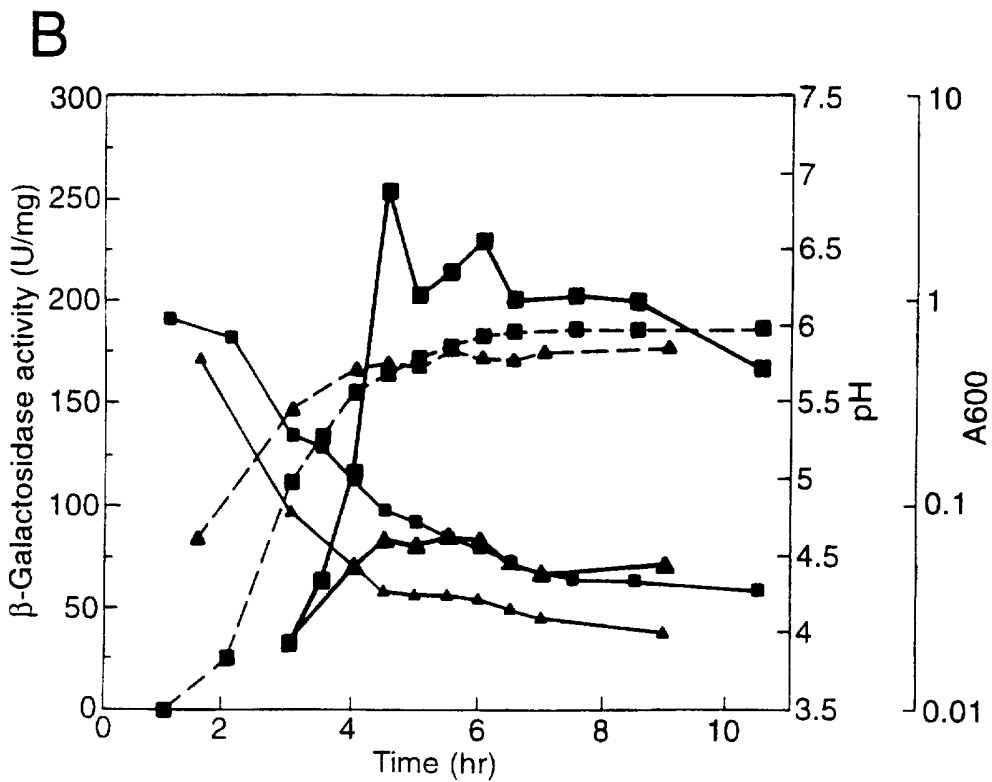
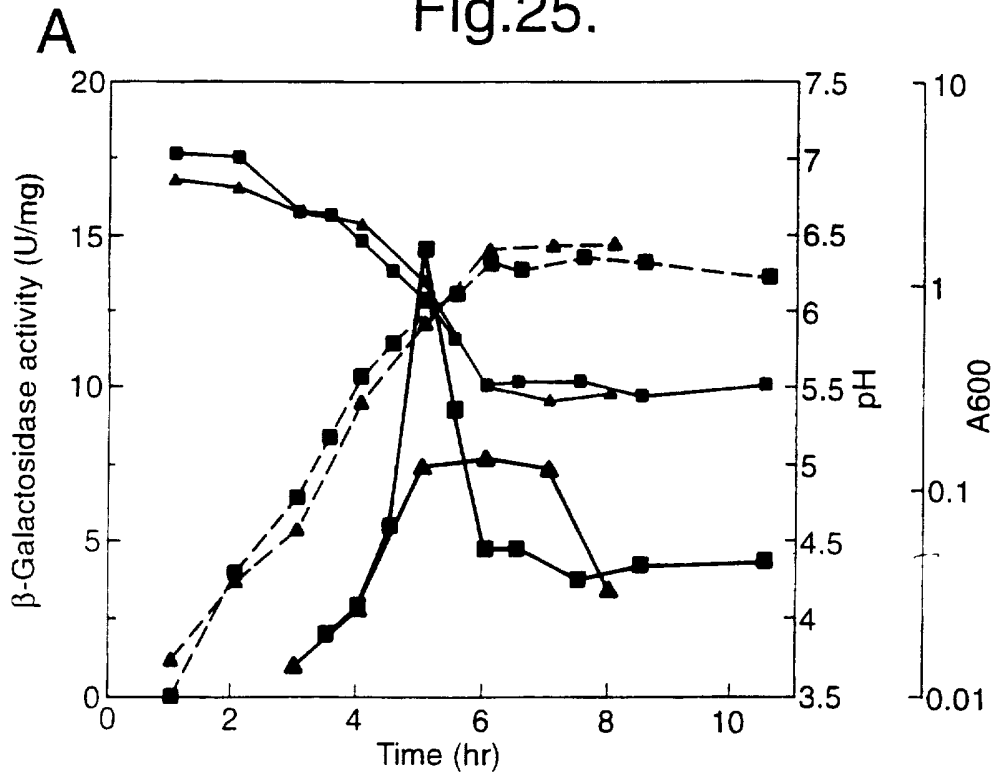
L.lactis      THKVAMYLAEEIEKTGMFEIMNDGAQLPIVCYKLNEN--SNRGNLYDLA 394
                * ** * * * * * * * * * * * * * * * * * * *
Synecocystis CRDTALYLSGKIAQLGPFELLDGGDIPVFAWRLKDEVLANCYTLYDMA 398

L.lactis      DRLLMKGWQVPAYPLPKNLENI IQRXXRADFGMNMFAFNYVQDMQEAI 444
                * * ** * * * * * * * * * * * * * * * * * *
Synecocystis DKLRERGWLVPAYRMPKNREDLVVQRIVVKEGFSRDMADLLLADMERAI 448

L.lactis      ALNKAHILFHQEPENKTYGFTH 466
                * * * *
Synecocystis YFASQP---DHKPKQEGSHFSH 467
    
```

25/34

Fig.25.



26/3A

Fig.26.

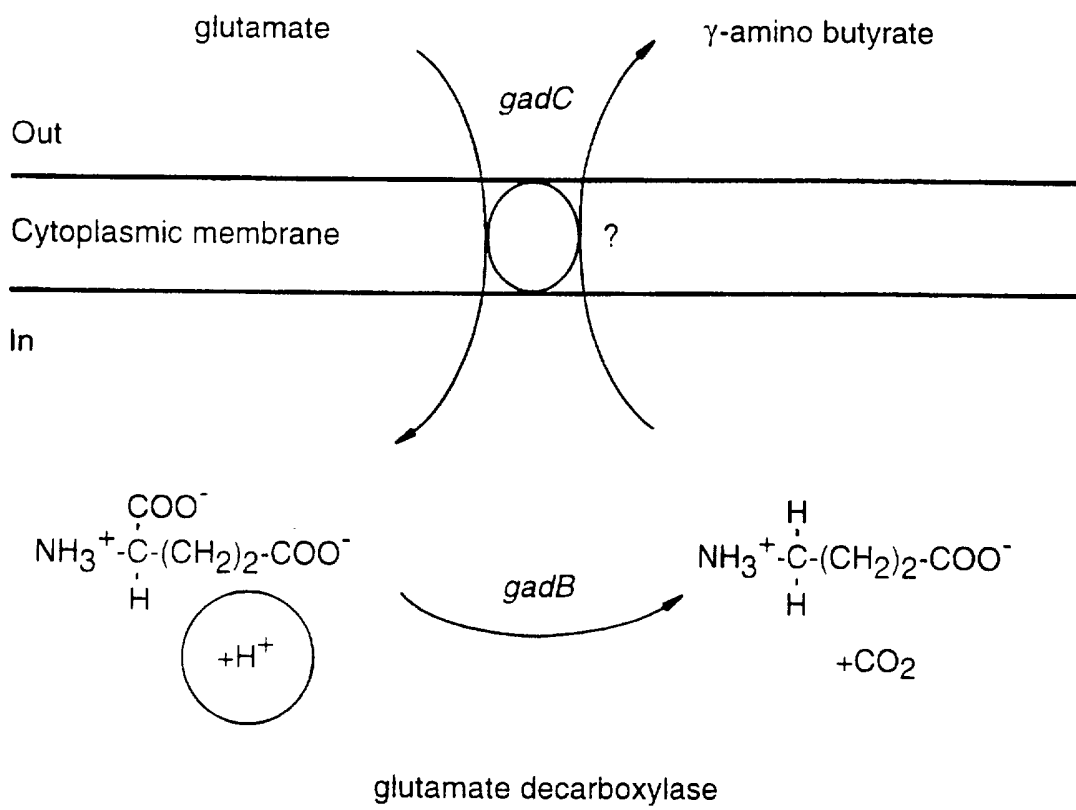




Fig.27.

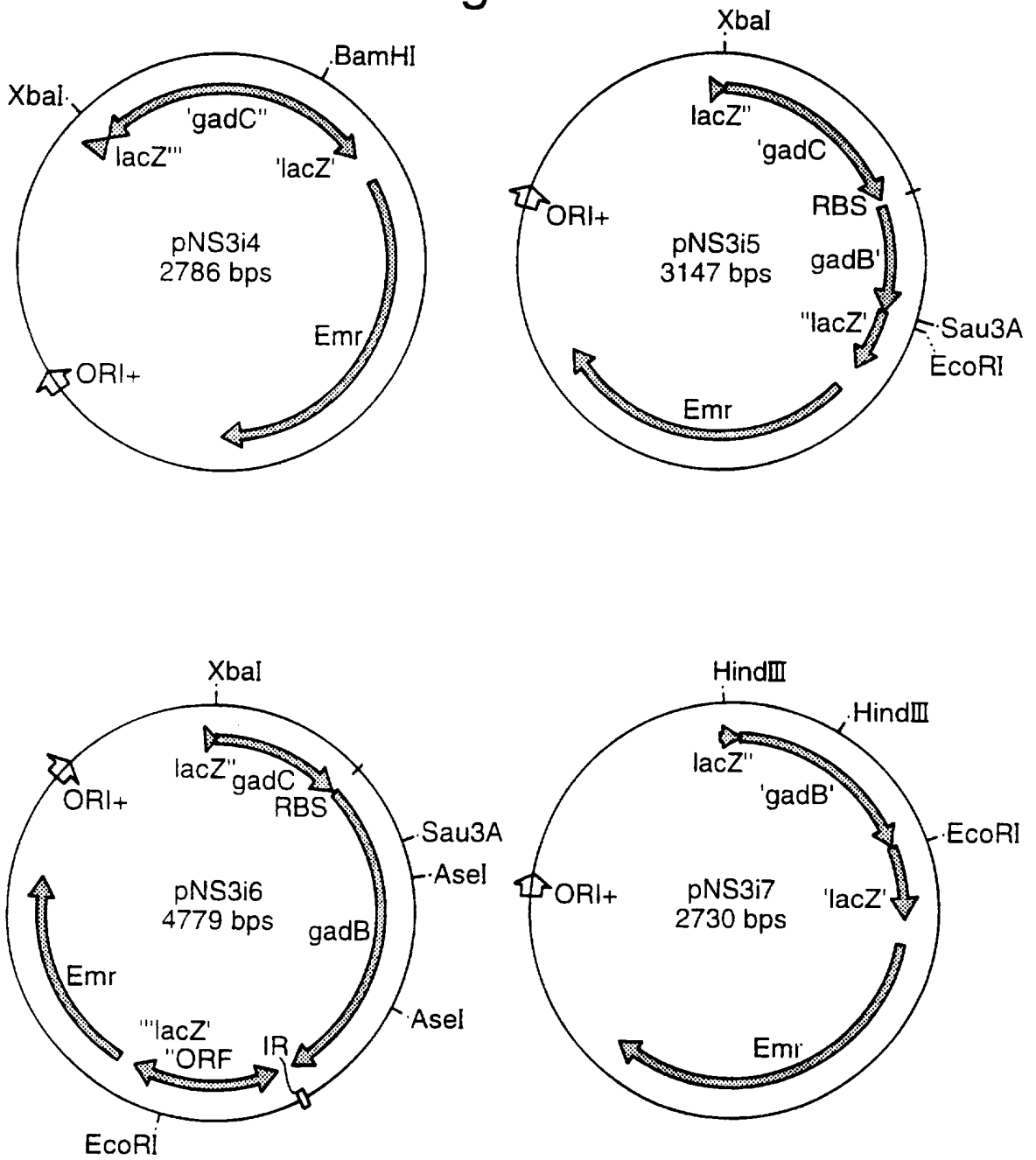
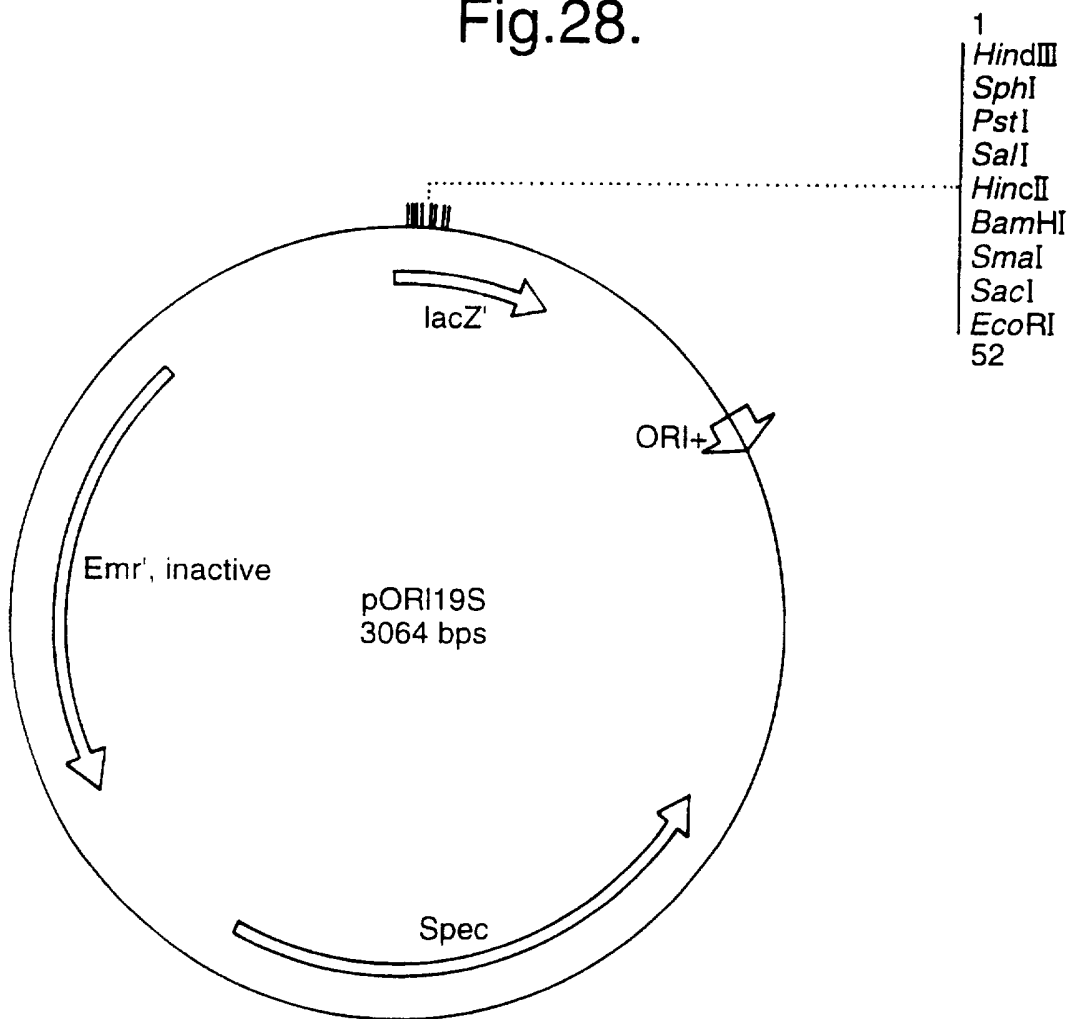


Fig.28.



29/34

## Fig.29. (1/6)

<i>Pst</i> I     ' <i>rnhB</i>																
CT	GCA	GTA	ATT	TTG	CCA	AAG	AAT	TGT	AAA	ATT	CGT	GGT	TTG	AAT	GAT	47
	Ala	Val	Ile	Leu	Pro	Lys	Asn	Cys	Lys	Ile	Arg	Gly	Leu	Asn	Asp	15
AGT	AAA	AAA	GTG	CCA	AAA	TCA	AAG	CAT	CAT	GCT	ATT	CTA	TCT	GAA	ATT	95
Ser	Lys	Lys	Val	Pro	Lys	Ser	Lys	His	His	Ala	Ile	Leu	Ser	Glu	Ile	31
CAA	GAA	AAA	GCG	CTA	GCG	ATT	GGA	GTT	GGA	ATT	GTT	GAT	GCC	GAA	AAA	143
Gln	Glu	Lys	Ala	Leu	Ala	Ile	Gly	Val	Gly	Ile	Val	Asp	Ala	Glu	Lys	47
ATT	GAT	GAA	GTA	AAT	ATT	TAT	GAG	GCG	ACA	AAA	ATT	GCA	ATG	ATT	CAG	191
Ile	Asp	Glu	Val	Asn	Ile	Tyr	Glu	Ala	Thr	Lys	Ile	Ala	Met	Ile	Gln	63
GCA	GTA	TCA	AAA	TTA	TCT	CTT	AAA	CCT	GAA	CAT	CTC	TTA	ATA	GAT	GCG	239
Ala	Val	Ser	Lys	Leu	Ser	Leu	Lys	Pro	Glu	His	Leu	Leu	Ile	Asp	Ala	79
ATG	GTT	TTA	GAT	TTG	CCC	ATT	GCT	CAG	ACG	AAA	ATC	ATT	CAT	GGA	GAT	287
Met	Val	Leu	Asp	Leu	Pro	Ile	Ala	Gln	Thr	Lys	Ile	Ile	His	Gly	Asp	95
<i>Not</i> I																
GCT	CGT	TCA	GCT	TCA	ATT	GCG	GCC	GCA	TCA	ATT	GTA	GCT	AAA	GTG	ACT	335
Ala	Arg	Ser	Ala	Ser	Ile	Ala	Ala	Ala	Ser	Ile	Val	Ala	Lys	Val	Thr	111
CGT	GAT	GAA	ATG	ATG	AAG	GAT	TTC	GCT	TTA	GAA	TTT	CCA	GAA	TAT	GAT	383
Arg	Asp	Glu	Met	Met	Lys	Asp	Phe	Ala	Leu	Glu	Phe	Pro	Glu	Tyr	Asp	127
TTT	GAA	CAT	AAT	GCA	GGC	TAT	GGA	ACA	GCA	AAA	CAT	CTT	GCA	GCT	CTG	431
Phe	Glu	His	Asn	Ala	Gly	Tyr	Gly	Thr	Ala	Lys	His	Leu	Ala	Ala	Leu	143
ACA	AAA	TAT	GGT	ATC	ACA	AGA	ATT	CAT	CGG	AAA	TCC	TAT	GAA	CCA	ATT	479
Thr	Lys	Tyr	Gly	Ile	Thr	Arg	Ile	His	Arg	Lys	Ser	Tyr	Glu	Pro	Ile	159
AAA	TCG	ATG	GTC	AAT	TTC	AAA	TAG	TAGATTA	TGTAAGTAAA	AAAAGGAAAA						530
Lys	Ser	Met	Val	Asn	Phe	Lys	***									166
CGTGAACGAT	ATTTGGTCAC	GTTTTTTTGC	TGACAAGTCT	GTCAGTAATT	ATTTTCAAG											590
----- >	< -----	-----														
GTTTTCAAAA	TATAGTCTAG	TAATTTGCTA	GACTAATCTT	CTGTTTTTTA	ATAATAAACT											650
AATTTTTTGT	TAATCTAAAT	GACAAAATTA	ATAAGCAGAG	TTTTTTATAA	AATTAGCTAC											710
TTATAAAAA	TTTGAAATTG	GTATAGTTAA	ATCTGTTATA	ATTTCCAATA	TTTTTTAATA											770
ATAATTATTT	TAACAAAATA	CTTATATCAA	AACTCTTTCA	AAGTATATAA	TGAGCGTTGT											830
ATAAGCTTTT	ATGTCTTTCT	ATATCAACTT	TTAATAGAAA	TATAAAGTAA	TATAAATGTT											890
TTTATAATAA	ATTATGTGAG	ATATATTTTT	TTGTCCGTAC	TGGTATAGAT	<u>TTGACGATTA</u>											950
		-10														
AGTCTTAAAT	AAGT <u>TATAAT</u>	CTCAATTGCG	TAATTTCTTA	AATACAGAAA	TAACAACTAC											1010
ATTGGTAGAC	TGATTA AAAA	GTGTACTTGA	TGAACTGTTA	TAAACCTTAA	AAAAATAAAA											1070
		rbs														
ATAATAGTTT	GGGggaTgtT	AAAGATG	TAT	AAA	AAA	TAT	GGA	GAT	TGT	TTT	AAA					1124
																10
AAG	TTG	CGA	AAC	CAA	AAG	AAT	TTA	GGG	TTA	TCA	TAC	TTT	AGT	AAA	CTT	1172
Lys	Leu	Arg	Asn	Gln	Lys	Asn	Leu	Gly	Leu	Ser	Tyr	Phe	Ser	Lys	Leu	26

30/34

Fig.29.(2/6)

XbaI																
GGA	ATA	GAC	CGT	TCA	AAT	ATA	TCT	AGA	TTT	GAA	CAT	GGA	AAA	TGT	ATG	1220
Gly	Ile	Asp	Arg	Ser	Asn	Ile	Ser	Arg	Phe	Glu	His	Gly	Lys	Cys	Met	42
HindIII																
ATG	AGT	TTT	GAG	CGT	ATA	GAT	TTG	ATG	TTA	GAA	GAA	ATG	CAA	GTT	CCG	1268
Met	Ser	Phe	Glu	Arg	Ile	Asp	Leu	Met	Leu	Glu	Glu	Met	Gln	Val	Pro	58
HindIII																
TTA	TCT	GAG	TAC	GAA	TTG	ATT	GTA	AAT	AAT	TTT	ATG	CCG	AAT	TTC	CAA	1316
Leu	Ser	Glu	Tyr	Glu	Leu	Ile	Val	Asn	Asn	Phe	Met	Pro	Asn	Phe	Gln	74
HindIII																
GAA	TTT	TTT	ATA	TTA	GAA	TTG	GAA	AAA	GCT	GAA	TTT	AGC	CAA	AAT	CGA	1364
Glu	Phe	Phe	Ile	Leu	Glu	Leu	Glu	Lys	Ala	Glu	Phe	Ser	Gln	Asn	Arg	90
HindIII																
GAT	AAA	ATA	AAA	GAG	TTG	TAT	TCT	GAG	GTC	AAA	GAA	ACG	GGG	AAT	CAT	1412
Asp	Lys	Ile	Lys	Glu	Leu	Tyr	Ser	Glu	Val	Lys	Glu	Thr	Gly	Asn	His	106
HindIII																
TTA	CTG	ACG	GTT	ACC	GTG	AAA	ACG	AAG	CTT	GGG	AAT	ATA	AGT	CAG	ACA	1460
Leu	Leu	Thr	Val	Thr	Val	Lys	Thr	Lys	Leu	Gly	Asn	Ile	Ser	Gln	Thr	122
HindIII																
GAA	GTT	AAA	GAA	ATT	GAA	GCT	TAT	CTT	TGC	AAT	ATT	GAA	GAG	TGG	GGA	1508
Glu	Val	Lys	Glu	Ile	Glu	Ala	Tyr	Leu	Cys	Asn	Ile	Glu	Glu	Trp	Gly	138
HindIII																
TAT	TTT	GAA	CTT	ACT	TTA	TTT	TAT	TTT	GTA	TCT	GAT	TAT	CTC	AAT	GTC	1556
Tyr	Phe	Glu	Leu	Thr	Leu	Phe	Tyr	Phe	Val	Ser	Asp	Tyr	Leu	Asn	Val	154
HindIII																
AAT	CAA	TTA	GAA	TTG	CTG	CTT	TTT	AAT	TTT	GAT	AAA	AGA	TGT	GAA	AAT	1604
Asn	Gln	Leu	Glu	Leu	Leu	Leu	Phe	Asn	Phe	Asp	Lys	Arg	Cys	Glu	Asn	170
HindIII																
TAC	TGT	AGA	GTC	TTA	AAA	TAT	AGA	AGG	AGA	CTA	TTG	CAA	ATA	GCC	TAT	1652
Tyr	Cys	Arg	Val	Leu	Lys	Tyr	Arg	Arg	Arg	Leu	Leu	Gln	Ile	Ala	Tyr	186
HindIII																
AAA	AGT	GTT	GCG	ATA	TAC	GCG	GCT	AAA	GGA	GAA	AGA	AAA	AAA	GCC	GAA	1700
Lys	Ser	Val	Ala	Ile	Tyr	Ala	Ala	Lys	Gly	Glu	Arg	Lys	Lys	Ala	Glu	202
HindIII																
AAT	ATT	TTA	GAA	ATG	ACT	AAA	AAA	TAT	CGA	ACT	GTG	GGA	GTC	GAT	TTA	1748
Asn	Ile	Leu	Glu	Met	Thr	Lys	Lys	Tyr	Arg	Thr	Val	Gly	Val	Asp	Leu	218
HindIII																
TAT	TCA	GAA	GTA	TTA	AGA	CAT	CTT	GCT	AGA	GCT	ATC	ATT	ATT	TTT	AAT	1796
Tyr	Ser	Glu	Val	Leu	Arg	His	Leu	Ala	Arg	Ala	Ile	Ile	Ile	Phe	Asn	234
HindIII																
TTT	GAA	AAT	GCA	GAG	ATT	GGG	GAA	GAA	AAA	ATA	AAT	TAT	GCT	CTT	GAG	1844
Phe	Glu	Asn	Ala	Glu	Ile	Gly	Glu	Glu	Lys	Ile	Asn	Tyr	Ala	Leu	Glu	250
EcoRI																
ATT	TTG	GAA	GAA	TTT	GGA	GGA	AAG	AAG	ATA	AAA	GAA	TTC	TAT	CAG	AAT	1892
Ile	Leu	Glu	Glu	Phe	Gly	Gly	Lys	Lys	Ile	Lys	Glu	Phe	Tyr	Gln	Asn	266
EcoRI																
AAA	ATG	GAA	AAG	TAT	TTG	AAA	AGG	TCA	ATT	TAG	TCTCTTTTGA	GCTGTTGCTT				1945
Lys	Met	Glu	Lys	Tyr	Leu	Lys	Arg	Ser	Ile	***	-----	-----				276
XbaI																
TAAAGCAACA	GCTCAAAGA	GATTTTCTTT	ATTCTAGAGC	ATATACTAGA	GGGTGAAGAT	2005										
><-----	-----	-----	-----	-----	-----											
rbs																
GTCTGAAGCA	TTATAACTTG	TCTTTTAAAA	AGGTTAATTC	AATCATAAAT	Ataaggaggt	2065										
Start gadC																
ATG	ATG	AAT	CAA	AAA	AAA	TTA	TCA	TTA	TTC	GGT	TTT	TTC	GCA	TTA	ACC	2113
Met	Asn	Gln	Lys	Lys	Leu	Ser	Leu	Phe	Gly	Phe	Phe	Ala	Leu	Thr	15	

3/34

## Fig.29. (3/6)

GCT TCA ATG GTT TTG ACT GTC TAT GAG TAT CCG ACT TTT GCC ACG TCA	2161
Ala Ser Met Val Leu Thr Val Tyr Glu Tyr Pro Thr Phe Ala Thr Ser	31
AAA TTA CAT TTG GTG TTC TTT TTA CTT CTC GGA GGA CTA CTA TGG TTT	2209
Lys Leu His Leu Val Phe Phe Leu Leu Leu Gly Gly Leu Leu Trp Phe	47
TTG CCT GTA GCG CTC TGC GCA GCA GAA ATG GCG ACG GTT GAA GGC TGG	2257
Leu Pro Val Ala Leu Cys Ala Ala Glu Met Ala Thr Val Glu Gly Trp	63
AAA AAT GGT GGA ATC TTT AGT TGG GTC AGT CAA ACT TTA GGT GAG CGC	2305
Lys Asn Gly Gly Ile Phe Ser Trp Val Ser Gln Thr Leu Gly Glu Arg	79
TTT GGT TTT GCA GCC ATA TTT TTT CAG TGG TTC CAA ATT ACA GTA GGT	2353
Phe Gly Phe Ala Ala Ile Phe Phe Gln Trp Phe Gln Ile Thr Val Gly	95
ApaI	
TTT GTC ACT ATG ATC TAT TTC ATT TTA GGG GCC CTC TCT TAT GTG TTA	2401
Phe Val Thr Met Ile Tyr Phe Ile Leu Gly Ala Leu Ser Tyr Val Leu	111
Sau3A	
AAT TTT CAG GCG CTC AAT ACA GAT CCA TTG ATA AAA TTT ATT GGT TTA	2449
Asn Phe Gln Ala Leu Asn Thr Asp Pro Leu Ile Lys Phe Ile Gly Leu	127
CTA ATC ATT TTT TGG GGA TTG ACT TTT TCT CAA TTA GGT GGG ACG CAA	2497
Leu Ile Ile Phe Trp Gly Leu Thr Phe Ser Gln Leu Gly Gly Thr Gln	143
CGG ACT GCC AAA TTA GTA AAA GCT GGC TTT GTA GTT GGA ATA GTG ATT	2545
Arg Thr Ala Lys Leu Val Lys Ala Gly Phe Val Val Gly Ile Val Ile	159
CCA TCG GTT ATC TTG TTT GGA TTA GCA GCG GCA TAC TTT ATC GGA GGC	2593
Pro Ser Val Ile Leu Phe Gly Leu Ala Ala Ala Tyr Phe Ile Gly Gly	175
AAT CCT ATA GAA ATA CCA ATT AAC AGC CAT GCT TTT GTA CCA GAT TTT	2641
Asn Pro Ile Glu Ile Pro Ile Asn Ser His Ala Phe Val Pro Asp Phe	191
TCA CAG GTA TCA ACT TTA GTA GTT TTT GTT TCT TTT ATT CTG GCT TAT	2689
Ser Gln Val Ser Thr Leu Val Val Phe Val Ser Phe Ile Leu Ala Tyr	207
ATG GGG GTA GAA GCC TCA GCT TCA CAT ATT AAT GAA CTT GAA AAT CCA	2737
Met Gly Val Glu Ala Ser Ala Ser His Ile Asn Glu Leu Glu Asn Pro	223
AAA CGA AAT TAT CCC TTA GCA ATG ATT TTA TTA GTA ATT TTG GCT ATT	2785
Lys Arg Asn Tyr Pro Leu Ala Met Ile Leu Leu Val Ile Leu Ala Ile	239
TCT TTA GAT GCC ATA GGT GGA TTT TCT GTA GCA GCA GTT ATT CCT CAA	2833
Ser Leu Asp Ala Ile Gly Gly Phe Ser Val Ala Ala Val Ile Pro Gln	255
AAA GAG TTA TCA TTA AGT GCA GGG GTA ATC CAA ACT TTT CAA ACG TTA	2881
Lys Glu Leu Ser Leu Ser Ala Gly Val Ile Gln Thr Phe Gln Thr Leu	271
ATC TTA CAT TTT AAT CAT CAT TTG GGA TGG TTA GTT AAA GTG ATT GCA	2929
Ile Leu His Phe Asn His His Leu Gly Trp Leu Val Lys Val Ile Ala	287
CTA ATG ATT GCC TTT GGG GTT ATG GGA GAA GTG AGT TCA TGG GTT GTT	2977
Leu Met Ile Ala Phe Gly Val Met Gly Glu Val Ser Ser Trp Val Val	303
XbaI	
GGT CCT TCT AGA GGG ATG TTT GCA GCA GCA CAA AGA GGT TTA TTA CCA	3025
Gly Pro Ser Arg Gly Met Phe Ala Ala Ala Gln Arg Gly Leu Leu Pro	319
AAA TTT TTA CGT AAA ACG AAT ACA CAT GAA GTC CCT GTT CCT TTA GTT	3073
Lys Phe Leu Arg Lys Thr Asn Thr His Glu Val Pro Val Pro Leu Val	335

32/34

Fig.29. (4/6)

ATG ATT CAA GGA ATC ATT GTT ACA CTT TGG GGC GCT GTA TTA ACT TTT	3121
Met Ile Gln Gly Ile Ile Val Thr Leu Trp Gly Ala Val Leu Thr Phe	351
GGA GGA GGA GGA AAT AAT TTA TCT TTC TTA GTT GCC ATT TCA CTG ACT	3169
Gly Gly Gly Gly Asn Asn Leu Ser Phe Leu Val Ala Ile Ser Leu Thr	367
GTA GTG ATT TAT TTG GTG GGT TAC CTC TTG TTC TTT ATT GTT TAC TTT	3217
Val Val Ile Tyr Leu Val Gly Tyr Leu Leu Phe Phe Ile Val Tyr Phe	383
GTT TTA ATC TAT AAA AAA CAA AAT TTA AAG CGT ACT TAT AAT GTT CCA	3265
Val Leu Ile Tyr Lys Lys Gln Asn Leu Lys Arg Thr Tyr Asn Val Pro	399
GGT AAA ATA ATA GGA AAA ACA ATC ATT GCA GGA ATT GGA TTC TTA TTA	3313
Gly Lys Ile Ile Gly Lys Thr Ile Ile Ala Gly Ile Gly Phe Leu Leu	415
TCA ATT TTT GCT CTA TTT ATT TCC TTT GTT CCT CCA GCA TCA ATT GCG	3361
Ser Ile Phe Ala Leu Phe Ile Ser Phe Val Pro Pro Ala Ser Ile Ala	431
AAA AAT GAA ACT CAC ACC TAT CAA ATG ATA CTT CTT ATA AGT TTT GTT	3409
Lys Asn Glu Thr His Thr Tyr Gln Met Ile Leu Leu Ile Ser Phe Val	447
GTG ACC GCT ATC TTG CCA TTT ATT ATT TAT GAA TTG CAT GAT AAA AAG	3457
Val Thr Ala Ile Leu Pro Phe Ile Ile Tyr Glu Leu His Asp Lys Lys	463
GGA CAT GAT ACT ATT GAA GAA CCA ACA CAC TTT AAA GCA GGA GAT GTG	3505
Gly His Asp Thr Ile Glu Glu Pro Thr His Phe Lys Ala Gly Asp Val	479
AAC CCT GCG ATT TAT CCA GCA GCT CGT GGA GAG CAT CAT ATT ATT AAA	3553
Asn Pro Ala Ile Tyr Pro Ala Ala Arg Gly Glu His His Ile Ile Lys	495
AAA GAA GAA CAT ATC TTA AAA CAT TGA AAAATTggag gATGTACAT ATG	3602
Lys Glu Glu His Ile Leu Lys His *** Met	1
TTA TAC GGA AAA GAA AAT CGA GAT GAA GCG GAG TTC TTG GAA CCA ATT	3650
Leu Tyr Gly Lys Glu Asn Arg Asp Glu Ala Glu Phe Leu Glu Pro Ile	17
TTT GGT TCA GAA AGT GAA CAA GTG GAT TTA CCT AAA TAT AAA TTA GCT	3698
Phe Gly Ser Glu Ser Glu Gln Val Asp Leu Pro Lys Tyr Lys Leu Ala	33
CAA CAA TCA ATT GAG CCT CGA GTG GCC TAT CAG TTA GTT CAA GAT GAA	3746
Gln Gln Ser Ile Glu Pro Arg Val Ala Tyr Gln Leu Val Gln Asp Glu	49
ATG CTA GAT GAA GGG AAC GCT CGT TTA AAT TTG GCC ACA TTC TGT CAA	3794
Met Leu Asp Glu Gly Asn Ala Arg Leu Asn Leu Ala Thr Phe Cys Gln	65
ACT TAT ATG GAA CCT GAA GCA GTC AAG CTG ATG AGT CAG ACC TTG GAA	3842
Thr Tyr Met Glu Pro Glu Ala Val Lys Leu Met Ser Gln Thr Leu Glu	81
AAA AAT GCG ATT GAC AAA TCA GAA TAT CCA AGA ACA ACT GAA ATT GAA	3890
Lys Asn Ala Ile Asp Lys Ser Glu Tyr Pro Arg Thr Thr Glu Ile Glu	97
AAC CGT TGC GTC AAC ATG ATC GCT GAC CTT TGG AAT GCG AGT GAA AAA	3938
Asn Arg Cys Val Asn Met Ile Ala Asp Leu Trp Asn Ala Ser Glu Lys	113
GGA AAA ATT TAT GGG ACT TCG ACA ATT GGT TCT TCA GAA GCT TGT ATG	3986
Gly Lys Ile Tyr Gly Thr Ser Thr Ile Gly Ser Ser Glu Ala Cys Met	129
CTT GGG GGA ATG GCT ATG AAG TTT TCT TGG CGT AAG CGA GCA GAA AAA	4034
Leu Gly Gly Met Ala Met Lys Phe Ser Trp Arg Lys Arg Ala Glu Lys	145
TTA GGC CTA GAT ATT AAT GCG AAA AAG CCA AAC TTA GTC ATT TCC TCT	4082
Leu Gly Leu Asp Ile Asn Ala Lys Lys Pro Asn Leu Val Ile Ser Ser	161

33/34  
Fig.29.(5/6)

GGT TAT CAA GTT TGC TGG GAA AAA TTC TGT GTT TAT TGG GAT ATT GAA Gly Tyr Gln Val Cys Trp Glu Lys Phe Cys Val Tyr Trp Asp Ile Glu	4130 177
ATG AGA GAA GTG CCA ATG GAT AGA GAA CAT ATG TCA ATC AAT TTG GAA Met Arg Glu Val Pro Met Asp Arg Glu His Met Ser Ile Asn Leu Glu	4178 193
AAA GTG ATG GAT TAT GTT GAT GAA TAT ACG ATT GGA GTA GTT GGA ATT Lys Val Met Asp Tyr Val Asp Glu Tyr Thr Ile Gly Val Val Gly Ile	4226 209
ATG GGG ATT ACT TAT ACT GGT CGT TAT GAT GAT ATC AAA GCT TTG GAT Met Gly Ile Thr Tyr Thr Gly Arg Tyr Asp Asp Ile Lys Ala Leu Asp	4274 225
AAT TTG ATT GAA GAA TAT AAT AAA CAG ACA GAC TAC AAA GTT TAT ATT Asn Leu Ile Glu Glu Tyr Asn Lys Gln Thr Asp Tyr Lys Val Tyr Ile	4322 241
CAC GTA GAT GCT GCT TCA GGA GGA CTT TAT GCT CCT TTT GTT GAG CCA His Val Asp Ala Ala Ser Gly Gly Leu Tyr Ala Pro Phe Val Glu Pro	4370 257
GAA CTT GAG TGG GAT TTC CGT TTG AAA AAT GTC ATT TCA ATC AAT ACT Glu Leu Glu Trp Asp Phe Arg Leu Lys Asn Val Ile Ser Ile Asn Thr	4418 273
TCA GGA CAT AAA TAT GGT TTA GTA TAT CCT GGT GTA GGT TGG GTC TTG Ser Gly His Lys Tyr Gly Leu Val Tyr Pro Gly Val Gly Trp Val Leu	4466 289
TGG CGT GAC AAA AAA TAT TTA CCT GAA GAG TTA ATT TTT AAA GTA AGT Trp Arg Asp Lys Lys Tyr Leu Pro Glu Glu Leu Ile Phe Lys Val Ser	4514 305
TAT CTT GGA GGA GAA TTA CCA ACA ATG GCG ATT AAT TTT TCT CAC AGT Tyr Leu Gly Gly Glu Leu Pro Thr Met Ala Ile Asn Phe Ser His Ser	4562 321
GCT TCT CAA TTA ATC GGT CAA TAC TAT AAT TTT GTA CGT TAT GGA TTT Ala Ser Gln Leu Ile Gly Gln Tyr Tyr Asn Phe Val Arg Tyr Gly Phe	4610 337
GAT GGA TAT AAA GCT ATT CAT GAG AGA ACG CAT AAA GTA GCC ATG TAT Asp Gly Tyr Lys Ala Ile His Glu Arg Thr His Lys Val Ala Met Tyr	4658 353
TTA GCA GAA GAA ATT GAA AAA ACA GGA ATG TTT GAG ATT ATG AAC GAT Leu Ala Glu Glu Ile Glu Lys Thr Gly Met Phe Glu Ile Met Asn Asp	4706 369
GGG GCA CAA TTA CCA ATT GTC TGC TAC AAA TTA AAA GAA AAT TCA AAC Gly Ala Gln Leu Pro Ile Val Cys Tyr Lys Leu Lys Glu Asn Ser Asn	4754 385
CGT GGT TGG AAT CTT TAT GAT TTG GCA GAT CGT TTA TTA ATG AAG GGA Arg Gly Trp Asn Leu Tyr Asp Leu Ala Asp Arg Leu Leu Met Lys Gly	4802 401
TGG CAA GTG CCT GCT TAT CCA CTT CCT AAA AAT TTG GAA AAT GAA ATC Trp Gln Val Pro Ala Tyr Pro Leu Pro Lys Asn Leu Glu Asn Glu Ile	4850 417
ATT CAA CGT TTA GTA ATT CGA GCA GAT TTC GGG ATG AAT ATG GCA TTT Ile Gln Arg Leu Val Ile Arg Ala Asp Phe Gly Met Asn Met Ala Phe	4898 433
AAC TAT GTT CAA GAT ATG CAA GAA GCA ATT GAT GCA CTA AAT AAG GCT Asn Tyr Val Gln Asp Met Gln Glu Ala Ile Asp Ala Leu Asn Lys Ala	4946 449
CAT ATT CTA TTT CAT CAG GAA CCT GAA AAT AAA ACA TAT GGC TTT ACT His Ile Leu Phe His Gln Glu Pro Glu Asn Lys Thr Tyr Gly Phe Thr	4994 465
CAC TAA AGATAAAAGC GATATATCTA AGATATATCG CTTTTATTTT GTTTTAGGCT His *** > <	5050 466
ATTACTAAT TAGCTTGTCG CTTA TTA TTT TTC ATA GTA TTT ATC CAA AAT *** Lys Glu Tyr Tyr Lys Asp Leu Ile	5101 -8
TTC CAT TTT TAA AGG AGT AAT TTT AGA TAG GGG GGC AGT TAG ACT TGT Glu Met Lys Leu Pro Thr Ile Lys Ser Leu Pro Ala Thr Leu Ser Thr	5149 -24

3A/34

## Fig.29. (6/6)

TCT TAG GAA GAG CTT ATC TTC AAT GTT GAT GAT ACC CAG ATA TTT AAC	5197
Arg Leu Phe Leu Lys Asp Glu Ile Asn Ile Ile Gly Leu Tyr Lys Val	-40
TTG AGG GTA GTT AGC ATT GAC TTC TAT AAT TTG GGC TTT TTT CTC ACT	5245
Gln Pro Tyr Asn Ala Asn Val Glu Ile Ile Gln Ala Lys Lys Glu Ser	-56
AAT ATT TTC GTC TGT CAC GGG CAC ATC TAG GTT GAC CGT TCT TTC TTT	5293
Ile Asn Glu Asp Thr Val Pro Val Asp Leu Asn Val Thr Arg Glu Lys	-72
ATA AGA GTA ATT TTT GAG AGC AGC AAT ATT TCG GTT TGG AAT AAA AGT	5341
Tyr Ser Tyr Asn Lys Leu Ala Ala Ile Asn Arg Asn Pro Ile Phe Thr	-88
AGT CGC ACC GTC GGC TCC GAT AAC AGT AAT GGA ACG AAT TCC TAC GGT	5389
Thr Ala Gly Asp Ala Gly Ile Val Thr Ile Ser Arg Ile Gly Val Thr	-104
TTT CAC CAC TCC CTC AAT ATC AAG ACC GGC AAA GGC AAC CGT ATC CGC	5437
Lys Val Val Gly Glu Ile Asp Leu Gly Ala Phe Ala Val Thr Asp Ala	-120
AAC ATT GAT TTG ATG TTC AAC GAT AAT AAA GAA ACC ATT AAT GAT ATC	5485
Val Asn Ile Gln His Glu Val Ile Ile Phe Phe Gly Asn Ile Ile Asp	-136
AGC GAC TAA ATC TCG TCC AGC GAA ACC AAG AGC GAC TCC AAG GAT ACC	5533
Ala Val Leu Asp Arg Gly Ala Phe Gly Leu Ala Val Gly Leu Ile Gly	-152
'ORF EcoRI	
CGC GCC AGC TAG GAC ATT TGC AAC AGG AATTC	5565
Ala Gly Ala Leu Val Asn Ala Val Pro Leu	-162



INTERNATIONAL SEARCH REPORT

Intern: .nal Application No  
PCT/EP 97/04755

A. CLASSIFICATION OF SUBJECT MATTER  
IPC 6 C12N15/74 C12N1/21

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
IPC 6 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 95 31563 A (QUEST INT ;NAUTA ARJEN (NL); VENEMA GERARD (NL); KOK JAN (NL); LED) 23 November 1995 cited in the application see the whole document ---	1,8-10
A	WO 95 31562 A (QUEST INT ;NAUTA ARJEN (NL); VENEMA GERARD (NL); KOK JAN (NL); LED) 23 November 1995 see the whole document -----	1

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

° Special categories of cited documents :

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

- \*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- \*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- \*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- \*Z\* document member of the same patent family

Date of the actual completion of the international search

16 January 1998

Date of mailing of the international search report

30. 01. 98

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Hillenbrand, G

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/EP 97/04755

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9531563 A	23-11-95	AU 2374395 A EP 0759084 A	05-12-95 26-02-97
-----			
WO 9531562 A	23-11-95	AU 2354095 A EP 0759999 A	05-12-95 05-03-97
-----			